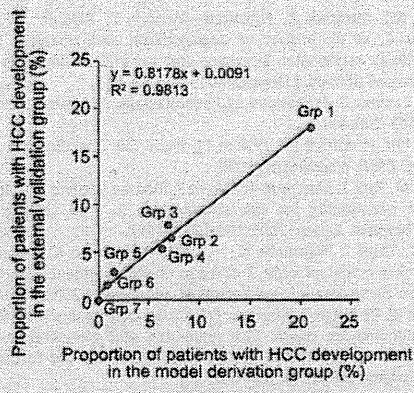

Fig. 2. External validation of the decision tree model with an independent cohort. Each patient in the external validation group was allocated to groups 1–7 following the flowchart of the decision tree. The HCC development rates were then calculated for each group and the graph plotted. The *x*-axis represents the HCC development rate in the model derivation group, and the *y*-axis represents the HCC development rate in the external validation group. The HCC development rates in each subgroup of patients are closely correlated between the model derivation group and the external validation group (correlation coefficient: $R^2 = 0.981$).

simple test values that are readily obtained in routine care and can therefore be easily used at the patient bedside. The model can be used to identify patients with a high risk of HCC development and therefore requiring surveillance, thereby allowing the formulation of surveillance plans personalized for individual patients.

Advanced fibrosis has been reported as independent risk factors for HCC development [7,8]. Platelet counts and albumin levels, which were factors selected for discrimination of the risk of HCC development, are closely related to the stage of fibrosis. Their correlation with the HCC risk has been repeatedly demonstrated [9–11,29–31]. The present study confirmed the impact of old age and advanced fibrosis, as reflected by low platelet counts and albumin levels. These results are consistent with our previous report [32]. What is unique to the present study was the study design to build a simple and reliable model for

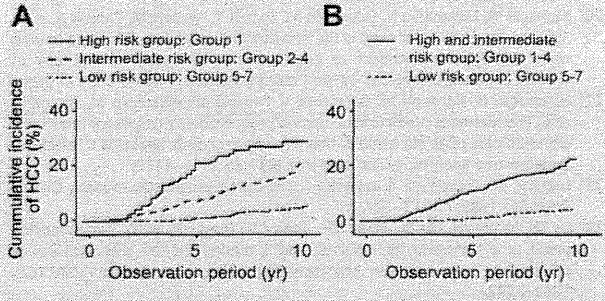

Fig. 3. Cumulative incidence of HCC development beyond 5 years in subgroups of patients defined by the decision tree model. Cumulative incidences of HCC in the groups classified by the decision tree model are compared. (A) The cumulative HCC development rate beyond 5 years is higher in the high- (group 1) and intermediate-risk (groups 2–4) groups compared to the low-risk group (groups 5–7). (B) The high and intermediate-risk group created by pooling data from the high- and intermediate-risk groups has a significantly higher cumulative HCC development rate than the low-risk group (5-year rate, 11.6% *vs.* 1.0%; 10-year rate, 24.5% *vs.* 4.8%; $p < 0.0001$).

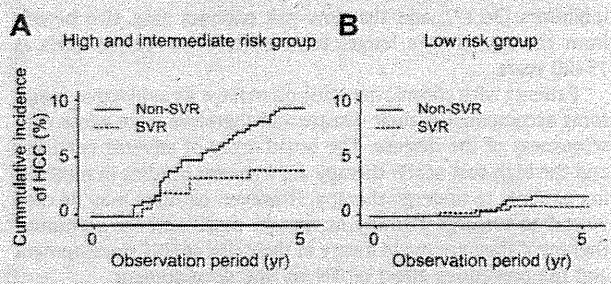

Fig. 4. Sustained virological response to PEG-IFN plus RBV therapy reduces the incidence of HCC development after stratification by the HCC risk. The 600 nonSVR patients and the 472 SVR patients in the external cohort were fitted into the HCC risk prediction model and classified into the high and intermediate-risk group or the low-risk group. The HCC development rate is significantly lower in SVR patients than in nonSVR patients in the high and intermediate-risk group (groups 1–4) (5-year HCC rate, 9.5% *vs.* 4.5%; $p = 0.040$). In the low-risk group (groups 5–7), the 5-year rate is 1.8% in nonSVR patients and 0.9% in SVR patients. Both rates are low and not significantly different ($p = 0.331$).

the prediction of HCC development that could be easily used in the clinic. For this purpose, a novel statistical method was used, histological factors were excluded in the analysis, the model derivation cohort was restricted to those who had nonSVR and had a long follow-up period duration (5 years), and the reproducibility of the model was independently validated by an external cohort. These are the major differences of the present study compared to our previous report. Many researchers have put a lot of efforts to formulate regression models for HCC prediction [9,10,33]. These prediction models are useful for identifying high-risk patients but are somewhat complicated to use at the bedside because they require calculations to be performed. Our prediction model is used simply by incorporating patients' data obtained through simple tests into the decision tree and following the flowchart. These prediction models based on factors easily accessible in routine clinical settings help physicians identify high-risk patients out of chronic hepatitis.

Viral eradication is the short-term goal of IFN therapy, but the ultimate goal is the prevention of HCC occurrence. Previous reports have shown that SVR to IFN therapy suppresses HCC occurrence in patients with type C liver cirrhosis and chronic hepatitis [7,12,30,34,35]. However, there is a marked heterogeneity in the magnitude of the treatment effect on the risk of HCC among studies, probably due to differences in the baseline risk of HCC among different trials [12]. Thus, the question remains whether the preventive effect of IFN therapy on HCC development could apply to all patients with chronic hepatitis C, especially those without liver cirrhosis. The result of the present study indicated that among high- and intermediate-risk patients, as assessed with our HCC risk prediction model, the cumulative HCC development rate was significantly reduced in SVR patients compared with nonSVR patients. This finding suggests that patients with chronic hepatitis, in whom disease has not yet progressed to hepatic cirrhosis but who are at a high risk of HCC development, benefit from antiviral treatment. The preventive effect of IFN on HCC development was not evident in low-risk patients within 5 years of observation. A longer observation term may be required to analyze the possible effect of antiviral therapy in these patients. Application of the present model on treatment decision may have limitations in that effect to prevent HCC development may differ in newer therapeutic agents such as protease

Please cite this article in press as: Kurosaki M et al. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. J Hepatol (2011), doi:10.1016/j.jhep.2011.09.011

inhibitors [36,37], and that low-risk patients may also benefit from therapy after a longer term observation period such as 15–20 years.

Patients with chronic hepatitis often have no subjective symptoms accompanying their disease and therefore have a low consciousness of the disease. The broad array of adverse reactions and the high cost of IFN therapy are frequent hurdles in motivating patients to undergo therapy. However, patients may be convinced to undergo therapy or remain motivated for continued therapy if they are made aware of their risk of HCC development and the preventive effect of IFN on HCC development.

In conclusion, a reproducible HCC risk prediction model, which includes the factors such as age, platelet count, albumin levels, and AST levels, was constructed to predict the 5-year HCC development rate in patients with chronic hepatitis C. The model requires only a combination of readily available test values and can therefore be easily used at the bedside. The information provided by the model allows the physician to identify patients requiring IFN therapy for the prevention of HCC and formulate plans for imaging HCC surveillance.

Conflict of interest

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

Financial support

This study was supported by a Grant-in-Aid from the Ministry of Health, Labor and Welfare, Japan (H20-kanen-006).

References

[1] Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
[2] Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. Gastroenterology 2004;127:S5–S16.
[3] Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology 2005;42:1208–1236.
[4] Fattovich G, Giustina G, Degos F, Tremolada F, Diodati G, Almasio P, et al. Morbidity and mortality in compensated cirrhosis type C: a retrospective follow-up study of 384 patients. Gastroenterology 1997;112:463–472.
[5] Niederau C, Lange S, Heintges T, Erhardt A, Buschkamp M, Hurter D, et al. Prognosis of chronic hepatitis C: results of a large, prospective cohort study. Hepatology 1998;28:1687–1695.
[6] Degos F, Christidis C, Ganne-Carrie N, Farmachidi JP, Degott C, Guettier C, et al. Hepatitis C virus related cirrhosis: time to occurrence of hepatocellular carcinoma and death. Gut 2000;47:131–136.
[7] Yoshida H, Tateishi R, Arakawa Y, Sata M, Fujiyama S, Nishiguchi S, et al. Benefit of interferon therapy in hepatocellular carcinoma prevention for individual patients with chronic hepatitis C. Gut 2004;53:425–430.
[8] Kurosaki M, Hosokawa T, Matsunaga K, Hirayama I, Tanaka T, Sato M, et al. Hepatic steatosis in chronic hepatitis C is a significant risk factor for developing hepatocellular carcinoma independent of age, sex, obesity, fibrosis stage and response to interferon therapy. Hepatol Res 2010;40: 870–877.
[9] Velazquez RF, Rodriguez M, Navascues CA, Linares A, Perez R, Sotorrios NG, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. Hepatology 2003;37:520–527.
[10] Lok AS, Seeff LB, Morgan TR, di Bisceglie AM, Sterling RK, Curto TM, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. Gastroenterology 2009;136:138–148.
[11] Cardoso AC, Moucari R, Figueiredo-Mendes C, Ripault MP, Giuily N, Castelnau C, et al. Impact of peginterferon and ribavirin therapy on hepatocellular carcinoma: incidence and survival in hepatitis C patients with advanced fibrosis. J Hepatol 2010;52:652–657.
[12] Craxi A, Camma C. Prevention of hepatocellular carcinoma. Clin Liver Dis 2005;9:329–346, viii.
[13] Breiman LJH, Friedman RA, Olshen CJ, Stone CM. Classification and regression trees. Calif: Wadsworth; 1980.
[14] Garzotto M, Park Y, Mongoue-Tchokote S, Bledsoe J, Peters L, Blank BH, et al. Recursive partitioning for risk stratification in men undergoing repeat prostate biopsies. Cancer 2005;104:1911–1917.
[15] Miyaki K, Takei I, Watanabe K, Nakashima H, Omae K. Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. J Epidemiol 2002;12:243–248.
[16] Averbook BJ, Fu P, Rao JS, Mansour EG. A long-term analysis of 1018 patients with melanoma by classic Cox regression and tree-structured survival analysis at a major referral center: implications on the future of cancer staging. Surgery 2002;132:589–602.
[17] Leiter U, Buettner PG, Eigentler TK, Garbe C. Prognostic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of the german dermatological society. J Clin Oncol 2004;22: 3660–3667.
[18] Valera VA, Walter BA, Yokoyama N, Koyama Y, Iiai T, Okamoto H, et al. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. Ann Surg Oncol 2007;14:34–40.
[19] Zlobec I, Steele R, Nigam N, Compton CC. A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. Clin Cancer Res 2005;11:5440–5443.
[20] Baquerizo A, Anselmo D, Shackleton C, Chen TW, Cao C, Weaver M, et al. Phosphorus as an early predictive factor in patients with acute liver failure. Transplantation 2003;75:2007–2014.
[21] Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, et al. A predictive model of response to peginterferon ribavirin in chronic hepatitis c using classification and regression tree analysis. Hepatol Res 2010;40: 251–260.
[22] Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, et al. Sequences in the interferon sensitivity-determining region and core region of hepatitis C virus impact pretreatment prediction of response to PEG-interferon plus ribavirin: data mining analysis. J Med Virol 2011;83: 445–452.
[23] Izumi N, Asahina Y, Kurosaki M. Predictors of virological response to a combination therapy with pegylated interferon plus ribavirin including virus and host factors. Hepat Res Treat 2010;2010:703602.
[24] Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, et al. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. J Gastroenterol 2011;46:401–409.
[25] Kurosaki M, Tanaka Y, Nishida N, Sakamoto N, Enomoto N, Honda M, et al. Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in IL28B and viral factors. J Hepatol 2011;54:439–448.
[26] Kurosaki M, Hiramatsu N, Sakamoto M, Suzuki Y, Iwasaki M, Tamori A, et al. Age and total ribavirin dose is an independent predictor of relapse among early virological responders to peg-interferon plus ribavirin therapy in chronic hepatitis C revealed by data mining analysis. Antivir Ther, in press.
[27] Hiramatsu N, Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, et al. Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: decision-tree analysis. J Gastroenterol 2011;46:1111–1119.
[28] LeBlanc M, Crowley J. A review of tree-based prognostic models. Cancer Treat Res 1995;75:113–124.
[29] Lu SN, Wang JH, Liu SL, Hung CH, Chen CH, Tung HD, et al. Thrombocytopenia as a surrogate for cirrhosis and a marker for the identification of patients at high-risk for hepatocellular carcinoma. Cancer 2006;107: 2212–2222.
[30] Bruno S, Stroffolini T, Colombo M, Bollani S, Benvegnu L, Mazzella G, et al. Sustained virological response to interferon-alpha is associated with improved outcome in HCV-related cirrhosis: a retrospective study. Hepatology 2007;45:579–587.
[31] Bonis PA, Tong MJ, Blatt LM, Conrad A, Griffith JL. A predictive model for the development of hepatocellular carcinoma, liver failure, or liver transplantation for patients presenting to clinic with chronic hepatitis C. Am J Gastroenterol 1999;94:1605–1612.

Please cite this article in press as: Kurosaki M et al. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. J Hepatol (2011), doi:10.1016/j.jhep.2011.09.011

[32] Asahina Y, Tsuchiya K, Tamaki N, Hirayama I, Tanaka T, Sato M, et al. Effect of aging on risk for hepatocellular carcinoma in chronic hepatitis C virus infection. Hepatology 2010;52:518–527.
[33] Ganne-Carrie N, Chastang C, Chapel F, Munz C, Pateron D, Sibony M, et al. Predictive score for the development of hepatocellular carcinoma and additional value of liver large cell dysplasia in Western patients with cirrhosis. Hepatology 1996;23:1112–1118.
[34] Nishiguchi S, Kuroki T, Nakatani S, Morimoto H, Takeda T, Nakajima S, et al. Randomised trial of effects of interferon-alpha on incidence of hepatocellular carcinoma in chronic active hepatitis C with cirrhosis. Lancet 1995;346:1051–1055.
[35] Shiratori Y, Ito Y, Yokosuka O, Imazeki F, Nakata R, Tanaka N, et al. Antiviral therapy for cirrhotic hepatitis C: association with reduced hepatocellular carcinoma development and improved survival. Ann Intern Med 2005;142:105–114.
[36] Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goeser T, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. N Engl J Med 2009;360:1839–1850.
[37] McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N Engl J Med 2009;360:1827–1838.

Please cite this article in press as: Kurosaki M et al. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. J Hepatol (2011), doi:10.1016/j.jhep.2011.09.011

HEPATOLOGY
Official Journal of the American Association for the Study of Liver Diseases

Comprehensive Analysis for Viral Elements and Interleukin-28B Polymorphisms in Response to Pegylated Interferon Plus Ribavirin Therapy in Hepatitis C Virus 1B Infection

Shinya Maekawa,[1,2] Minoru Sakamoto,[2] Mika Miura,[2] Makoto Kadokura,[2] Ryota Sueki,[2] Kazuki Komase,[2] Hiroko Shindo,[2] Nobutoshi Komatsu,[2] Kuniaki Shindo,[2] Asuka Kanayama,[2] Takako Ohmori,[2] Fumitake Amemiya,[2] Shinichi Takano,[2] Tatsuya Yamaguchi,[2] Yasuhiro Nakayama,[2] Takatoshi Kitamura,[2] Taisuke Inoue,[1,2] Shunichi Okada,[2] and Nobuyuki Enomoto[2]

To comprehensively characterize the contribution of virological factors as well as interleukin-28B (IL28B) single-nucleotide polymorphisms (SNPs) in determining treatment responses in pegylated-interferon plus ribavirin (Peg-IFN/RBV) therapy for chronic hepatitis C virus (HCV)-1b infection, we undertook a retrospective cohort analysis for the pretreatment dominant complete HCV open reading frame (ORF) amino-acid (aa) sequence study in 103 consecutive HCV-1b Japanese patients. The dominant HCV sequences classified by the response were subjected to systematic sliding-window comparison analysis to characterize response-specific viral sequences, along with IL28B SNP analyses (rs8099917). In each comparison of the patients between with and without rapid viral response (RVR), nonearly viral response (nEVR), sustained virological response (SVR), or relapse, the following regions were extracted as most significantly associated with the different responses respectively: nonstructural protein 5A (NS5A) aa.2224-2248 (P = 1.2E-07); core aa.70 (P = 4E-04); NS5A aa.2340-2382 (P = 7.0E-08); and NS5A aa.2360-2377 (P = 1.1E-05). Those NS5A regions nearly coincided with the interferon (IFN) sensitivity-determining region (NS5A aa.2209-2248) and the IFN/RBV resistance-determining region (NS5A aa.2339-2379). In a multivariate analysis, the IL28B SNP (odds ratio [OR] = 16.8; P = 0.009) and NS5A aa.2340-2382 (OR = 13.8; P = 0.0003) were extracted as the two most-significant independent variables contributing to the final outcome. *Conclusion:* In Peg-IFN/RBV therapy, polymorphisms in IL28B, NS5A aa.2224-2248, core aa.70, and, most important, NS5A aa.2340-2382 have a tremendous influence on treatment response in association with viral kinetics, resulting in significantly different outcomes in chronic HCV-1b infection. (HEPATOLOGY 2012;56:1611-1621)

Hepatitis C vrus (HCV) is a major cause of chronic liver disease (CLD) worldwide, causing CLD that may progress to hepatocellular carcinoma (HCC).[1] Treatment response of the conventional pegylated interferon (Peg-IFN) plus ribavirin (RBV) therapy is highly variable, and half of the patients cannot eradicate the virus (i.e., sustained virological response; SVR).[2] Recently, direct-acting

Abbreviations: aa, amino acid; AFP, alpha-fetoprotein; ALB, albumin; ALT, alanine aminotransferase; BMI, body mass index; cEVR, complete early viral response; cEVR-8w, HCV RNA <50 IU/mL at between weeks 5 and 8; cEVR-12w, HCV RNA <50 IU/mL at between weeks 9 and 12; CI, confidence interval; CLD, chronic liver disease; DAAs, direct-acting antiviral agents; ETR, end-of-treatment response; EVR, early viral response; Hb, hemoglobin; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; IFN, interferon; IL28B, interleukin-28B; IRRDR, IFN/RBV resistance-determining region; ISDR, IFN sensitivity-determining region; nEVR, nonearly viral response; NS5A, nonstructural protein 5A; OR, odds ratio; ORF, open reading frame; PCR, polymerase chain reaction; Peg-IFN, pegylated IFN; PePHD, PKR-eIF2 phosphorylation homology domain; pEVR, partial early viral response; PKR-BD, PKR-binding domain; PLT, platelet count; RBV, ribavirin; RVR, rapid viral response; SNPs, single-nucleotide polymorphisms; SVR, sustained viral response; T-Cho, total cholesterol.

From the [1]Department of Advanced Medicine for Liver Diseases, Faculty of Medicine, University of Yamanashi, Yamanashi, Japan; and [2]First Department of Medicine, Faculty of Medicine, University of Yamanashi, Yamanashi, Japan.

Received February 20, 2012; accepted April 30, 2012.

This study was supported, in part, by a grant-in-aid scientific research fund of the Ministry of Education, Science, Sports, and Culture (grant nos.: 21590836, 21590837, and 23390195) and, in part, by a grant-in-aid from the Ministry of Health, Labor, and Welfare of Japan (grant nos.: H22-kanen-006, H22-kanen-003, and H23-kanen-001).

antiviral agents (DAAs) have been under development, and telaprevir and boceprevir have now been included in HCV treatment regimens in the Unites States. However, it has gradually became learned that HCVs showing resistance to Peg-IFN/RBV therapy might demonstrate higher resistance to these new regimens of Peg-IFN/RBV plus DAAs.[3] In this background, it is urgent to clarify a comprehensive characterization of viral and host determinants for Peg-IFN/RBV therapy and to determine the most appropriate candidates for the new therapies.

In interferon (IFN)-based therapy, treatment response is influenced by multiple host and viral factors. Among the host factors, younger age, milder fibrosis stage, being nonobese,[4] being Asian or Caucasian rather than African,[5] and, recently, the interleukin-28B (IL28B) major allele type[6-8] are associated with favorable responses. Among the viral factors, low baseline viral load and genotype 2/3, rather than genotype 1/4, show favorable responses.[9] On the other hand, the contribution of other viral factors, such as polymorphisms in several restricted viral genetic regions, has long been debated in terms of their association with treatment responses. HCV genetic elements, including the IFN sensitivity-determining region (ISDR) in nonstructural protein 5A (NS5A),[10,11] PKR-binding domain (PKR-BD) in NS5A,[12,13] the V3 region in NS5A,[14] the IFN/RBV resistance-determining region (IRRDR) in NS5A,[15] the PKR-eIF2 phosphorylation homology domain (PePHD) of E2,[16] the C-terminal region of NS5A (G404S and E442G),[17] F415Y in NS5B,[18] polymerase motif in NS5B,[19] and amino acid (aa).70 and 91 in core,[20] have been investigated for their correlation with the clinical outcome of IFN-based therapy or RBV in genotype 1 infection. Complete open reading frame (ORF) analyses in Peg-IFN/RBV therapy also revealed the link between treatment response at day 28 or treatment outcome with viral diversities in several viral genomic regions in genotype 1 infection.[21,22] Importantly, most recent studies reported the strong contribution of core aa.70, ISDR, and IL28B polymorphisms in the response of Peg-IFN/RBV therapy in genotype 1b infection.[11,23]

Nevertheless, a comprehensive analysis of how these viral elements affect treatment response has not been presented clearly yet, especially along with IL28B single-nucleotide polymorphisms (SNPs). Moreover, inconsistent results that have been reported on for some of those regions made the association with the response obscure. Under these circumstances, the previous studies had limitations regarding the following points: (1) Viral regions selected for analysis were partial; (2) associations among different viral regions were not evaluated; (3) most studies investigated the associations only with the final SVR rate, although this is influenced by multiple factors, other than a simple virological response; (4) some studies have included patients with different racial backgrounds; and (5) most studies lacked analysis with IL28B polymorphisms.

To overcome these limitations, we have recently determined complete HCV ORF sequences of 88 patients receiving Peg-IFN/RBV, and confirmed that the NS5A-ISDR and core 70 were specifically extracted as regions most significantly correlated to rapid viral response (RVR) and nonearly viral response (nEVR), respectively.[24] In the present study, we undertook more comprehensive, detailed analysis to disclose the effect of HCV ORF on determining early viral response (EVR), final outcome, and relapse by extending the previous result through adding the information of IL28B polymorphisms in Japanese patients given Peg-IFN/RBV therapy for genotype 1b HCV.

Patients and Methods

Study Patients. We retrospectively analyzed consecutive patients with chronic HCV-1b infection treated with combination therapy of Peg-IFN/RBV at the Yamanashi University Hospital (Yamanashi, Japan) between December 2004 and July 2008. Eligible patients were 18-75 years of age, seronegative for hepatitis B surface antigen and antibodies against human immunodeficiency virus, and had an absolute neutrophil count $\geq$1,500/mm^3, a normal hemoglobin (Hb) level, and available pretreatment serum sample conserved for HCV-sequence analysis. Patients were excluded if they had decompensated liver cirrhosis or HCC. Consequently, 103 patients were eligible for this study. In addition to those 103 patients, 30

Address reprint requests to: Shinya Maekawa, M.D., Ph.D., First Department of Medicine, Faculty of Medicine, University of Yamanashi, Faculty of Medicine, 1110 Shimokato, Chuo 409-3898, Yamanashi, Japan. E-mail: maekawa@yamanashi.ac.jp; fax: +81-55-273-6748.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.25826

Potential conflict of interest: Shinya Maekawa and Taisuke Inoue belong to a donation-funded department that is funded by MSD Co., Ltd., Tokyo, Japan. Nobuyuki Enomoto received research funded by MSD Co., Ltd., Tokyo, Japan and Chugai Pharmaceutical Co., Ltd., Tokyo, Japan.

Additional Supporting Information may be found in the online version of this article.

consecutive patients who received the standard length of Peg-IFN/RBV at the Yamanashi University Hospital from August 2008 to April 2011 and were meeting the above-mentioned criteria were also included in the study to perform uni- and multivariate analysis for SVR and relapse. The study was approved by the ethics committees of the University of Yamanashi, and the study protocol conformed to the ethical guidelines of the 2000 Declaration of Helsinki.

Doses and treatment periods were determined according to a standard treatment protocol for Japanese patients, established by a hepatitis study group of the Ministry of Health, Labor, and Welfare, Japan. Patients were treated with Peg-IFN-α-2b (1.5 μg/kg, once-weekly, subcutaneously) and RBV (600-800 mg daily, per os) for 48 weeks. When patients failed to achieve a 2-log reduction of HCV RNA at week 12 (nEVR), or failed to achieve HCV RNA clearance (HCV RNA, <50 IU/mL) at week 24 (null viral response), the therapy was discontinued if they did not desire to continue. For patients without viral clearance by week 13, the therapy period was extended up to 72 weeks if they agreed. For patients having achieved viral clearance (HCV RNA, <50 IU/mL) within 4 weeks (RVR), the therapy could be reduced to 24 weeks if they agreed.

Analytic Methods. The following patient characteristics were analyzed: age; sex; stage of fibrosis on liver biopsy; body mass index (BMI); alanine aminotransferase (ALT); Hb; gamma-glutamyl transpeptidase (γ-GTP); total cholesterol (T-Cho); albumin (ALB); platelet counts (PLTs); alpha-fetoprotein (AFP); serum HCV RNA; Peg-IFN dose; and RBV dose. Liver-biopsy specimens were evaluated blindly by an independent interpreter. HCV RNA was determined by polymerase chain reaction (PCR) (Amplicor HCV RNA kit, version 2.0; Roche Diagnostics Corp., Indianapolis, IN).

Viral Response. Patients were subdivided into four groups according to the initial response at week 12. Each group was defined as follows: RVR (<50 IU/mL at week 4); complete early viral response (cEVR; HCV RNA <50 IU/mL at between weeks 5 and 12); partial EVR (pEVR; HCV RNA ≥2-log reduction, but still detectable [≥50 IU/mL] at week 12); and nEVR (HCV RNA <2-log drop at week 12). SVR was defined as undetectable HCV RNA 24 weeks after completion of therapy. Viral relapse after the achievement of end-of-treatment response (ETR) were also evaluated. In some analysis, cEVR was further divided into two groups of cEVR-8w (HCV RNA <50 IU/mL at between weeks 5 and 8) and cEVR-12w (HCV RNA <50 IU/mL at between weeks 9 and 12).

Complete HCV ORF Sequencing. Extraction of RNA, complementary DNA synthesis, and nested PCR were performed using patient serum collected before starting therapy, as described previously.[25] The full-length HCV genome was amplified by nested PCR with 20 partially overlapping primer sets. Both strands of PCR products were cycle-sequenced with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Tokyo, Japan), according to the manufacturer's instructions, using an M13 forward as well as reverse primers. Products were sequenced by an automated DNA sequencer (3130 series; Applied Biosystems). Nucleotide and predicted aa sequences of 20 HCV genomic fragments were determined and assembled using vector NTI software (Invitrogen, Tokyo, Japan).

Sliding-Window Analysis. A sliding-window analysis was introduced to search for HCV polypeptide regions related to treatment response. Briefly, the total number of aa substitutions, compared to the consensus sequence, within a given number of consecutive aas (window) was counted at each aa position in each HCV sequence. The distribution of aa substitutions in the HCV ORF was scanned, applying these windows from aa.1 to aa.3010. The substitution numbers in each window and the treatment response was compared statistically between the two groups, showing different treatment response by Mann-Whitney's U test for each aa window. In each comparison, the length of peptide window was changed from 1 to 100 aas to search for those regions. Consequently, approximately 300,000 windows (100 width × 3,010 aas) were analyzed for each HCV aa sequence. To visualize the result, windows showing significantly low P values were colored in red and nonsignificant P values were colored in green to generate a "heat map" appearance using Microsoft Excel (Microsoft Corp., Redmond, WA), whereas the window with the lowest P value was colored in white to be distinguished clearly.

IL28B SNP Analysis. Human genomic DNA was extracted from peripheral blood using a blood DNA extraction kit (QIAGEN, Tokyo, Japan), according to the manufacturer's protocol. The allele typing of each DNA sample was performed by real-time PCR (model 7500; Applied Biosystems) using fluorescein-amidite–labeled SNP primer for the locus rs8099917 (purchased from Applied Biosystems).

Statistical Analysis. Statistical differences in parameters, including all available patient demographic, biochemical, hematological, and virological data, was determined between patients in various groups by the Student t test or Mann-Whitney's U test for numerical variables and Fisher's exact probability test for categorical variables.

Table 1. Baseline Characteristics of 103 Patients and SVR Rate

Variables	Initial 103 Patients
Age, years	56 (31-70)
Gender, male (%)	64 (62)
Fibrosis, F2-F4 (%)	46 (44)
HCV RNA, kIU/mL	1,500 (28-8,392)
BMI	22.7 (17.5-31.7)
ALB, g/dL	4.1 (3.0-4.9)
γ-GTP, IU/mL	43 (11-289)
ALT, IU/mL	68 (20-413)
T-Cho, mg/dL	165 (104-240)
WBCs, per μL	4,450 (2,520-7,850)
Hb, g/dL	14.2 (11.2-17.9)
PLT, $\times 10^4/\mu L$	14.5 (6.5-27.3)
AFP, ng/mL	5.8 (0.7-468.4)
IL28B TT (%)	65 (73)*
Peg-IFN dose (%)	89 (43-147)
RBV dose (%)	98 (49-133)
SVR rate (n, %)	
All (n = 103)	55 (53)
Standard therapy (n = 76)	
RVR (n = 10)	10 (100)
cEVR (n = 35)	28 (80)
pEVR (n = 15)	3 (20)
nEVR (n = 16)	0 (0)
Extended therapy (n = 27)	
RVR (n = 0)	–
cEVR (n = 5)	3 (60)
pEVR (n = 18)	11 (61)
nEVR (n = 4)	0 (0)

Abbreviation: WBCs, white blood cells.
*n = 89.

Variables with $P < 0.05$ in univariate analysis were entered into multiple logistic regression analysis to identify significant independent factors with the odds ratios (ORs) as well as 95% confidence intervals (CIs). All P values of <0.05 by the two-tailed test were considered significant.

Results

Patient Characteristics. Clinical background factors of the 103 patients are shown in Table 1. Responses at 12 weeks were closely related to the final outcome of therapy. In the standard therapy up to 48 weeks, the SVR rate was 100%, 80%, 20%, and 0% for the RVR, the cEVR, the pEVR, and the nEVR, respectively. Among 103 patients, 27 patients from three groups received extended therapy (5 from cEVR-12w, 18 from pEVR, and 4 from nEVR). Although improvement of SVR was observed in the pEVR (from 20% to 61%), there was no improvement in cEVR or nEVR.

Clinical background factors of the 30 patients who were additionally included for uni- and multivariate analysis for SVR and relapse receiving the standard period of Peg-IFN/RBV therapy are also shown (Supporting Table 1).

IL28B SNPs and Their Relationship to Viral Diversity. To evaluate the contribution of the IL28B polymorphism in the 103-patient study group, we investigated the rs8099917 SNPs in 89 patients available for analysis. The polymorphism was closely related to the viral response at week 12 (Table 2). To clarify the relationship between viral diversity and IL28B SNPs, we compared viral sequences between the major allele groups showing favorable initial response (TT) and the minor allele groups showing poor initial responses (TG or GG). IL28B SNP was significantly correlated with the aa residue at core aa.70 in full HCV ORF analysis ($P = 3.4E-06$); non-arginine at core aa.70 was closely related to minor IL28B alleles and vice versa (Supporting Fig. 1).

HCV Sequences Related to RVR and nEVR. To characterize the HCV sequences related to RVR and nEVR, we determined the full dominant HCV ORF sequences by direct sequencing and searched for polymorphic aa positions specifically related to the different responses. Though aa.2240 was extracted as the most-different single position between the RVR and the remainder (data not shown), successive sliding-window analysis revealed that aa.2224 to aa.2248 of the NS5A region, being completely included in the ISDR (aa.2209 to aa.2248), was the region most significantly related to the RVR ($P = 0.00037$; Fig. 1A). On the other hand, when the nEVR and the remainder were compared, core aa.70 was extracted as the most-significant single aa position discriminating the two groups ($P = 7.0E-8$; Fig. 1B). In this comparison of the nEVR versus the remainder, a sliding-window analysis also extracted regions around aa.70 to be the most significantly different (data not shown).

HCV Sequences Related to Final Outcome. We also compared the viral sequence between SVR and non-SVR patients. In comparing complete HCV ORFs, we confined this analysis to HCV sequences obtained from the standard therapy (n = 76) to exclude the influence of therapy duration. In the analysis of each single aa, various differences were observed

Table 2. IL28B SNPs at rs8099917 and the Initial Viral Responses*

	RVR (%) (n = 8)	cEVR-8w (%) (n = 17)	cEVR-12w (%) (n = 15)	pEVR (%) (n = 31)	nEVR (%) (n = 18)
TT	8 (100)	16 (94)	13 (87)	24 (77)	4 (22)
TG	0 (0)	1 (6)	1 (7)	7 (23)	12 (67)
GG	0 (0)	0 (0)	1 (7)	0 (0)	2 (11)

*n = 89.

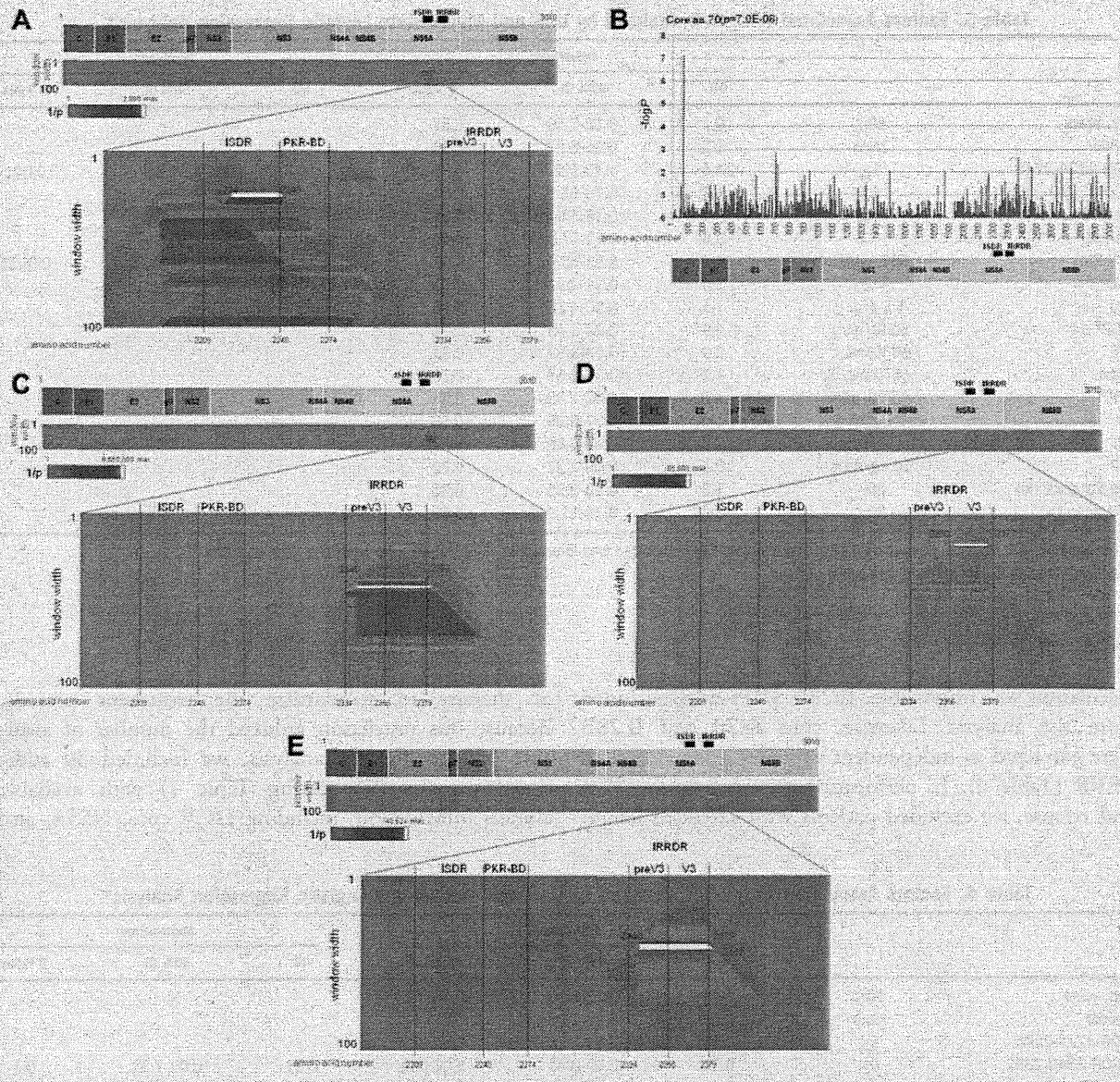

Fig. 1. The contribution of viral sequences and IL28B SNPs in the treatment response to Peg-IFN/RBV was studied. (A) Sliding-window analysis for RVR versus the remainder (n = 103). (B) Single aa analysis for nEVR versus the remainder (n = 103). (C) Sliding-window analysis for SVR versus non-SVR (n = 76). (D) Sliding-window analysis for relapsers versus nonrelapsers among ETR (n = 57). (E) Sliding-window analysis for SVR versus non-SVR in IL28B TT patients with standard therapy (n = 47).

in the HCV ORF, including core aa.70 and NS5B (data not shown). However, a sliding-window analysis disclosed that NS5A region aa.2340 to aa.2382, the region almost coinciding with IRRDR, was extracted as the most clearly related to the final outcome (P = 1.2E-07; Fig. 1C).

HCV Sequences Related to Relapse. To identify the viral regions related to relapse, we compared SVR patients and non-SVR patients among 57 patients with standard therapy achieving ETR (40 nonrelapsers and 17 relapsers). A sliding-window analysis disclosed that the NS5A region aa.2360 to aa.2377, the region almost coinciding with the V3 region in the IRRDR, could be extracted as the most strongly related to relapse (P = 1.1E-05; Fig. 1D).

Uni- and Multivariate Analyses. We performed further analyses to extract the factors associated with RVR, nEVR, SVR, and relapse by univariate, as well as multivariate, analyses. For achieving RVR, ISDR aa.2224-2248 and HCV-RNA were extracted as independent variables (Table 3). Because all the RVR patients possessed IL28B TT alleles and OR

Table 3. Factors Associated With RVR Analyzed by Uni- and Multivariate Logistic Regression Analysis*

		Univariate			Multivariate		
		OR	95% CI	*P* Value	OR	95% CI	*P* Value
Age, years	60≤	0.7	0.18-2.59	0.57			
Gender	Male	1.5	0.36-6.07	0.59			
ISDR 2224-2248	1≤	24.6	4.70-129	8.5E-07†	14.7	1.10-198	0.04‡
IRRDR 2340-2382	4≤	6.2	0.76-51.1	0.06			
Core 70	Arg	0.7	0.18-3.07	0.68			
Fibrosis	<2	3.6	0.72-17.8	0.10			
HCV RNA	<600 k/Ul/mL	74.7	8.55-653	8.3E-10†	51.2	3.97-662	0.003†
BMI	<23	1.3	0.34-4.87	0.71			
ALB	4.1 g/dL≤	1.1	0.30-4.28	0.85			
γ-GTP	50 IU/mL≤	0.9	0.24-3.49	0.91			
ALT	60 IU/mL<	0.9	0.25-3.59	0.94			
T-Cho	<170 mg/dL	1.2	0.33-4.67	0.76			
WBC	4,700/μL≤	1.9	0.47-7.89	0.36			
Hb	14 g/dL≤	1.5	0.37-6.35	0.55			
PLT	150,000/μL≤	1.8	0.48-6.88	0.37			
AFP	10 ng/mL≤	0.3	0.03-2.37	0.22			
Peg-IFN dose (%)	80≤	1.3	0.33-5.55	0.68			
RBV dose (%)	80≤	3.0	0.79-11.4	0.09			

Because all RVR patients possessed IL28B TT alleles and OR calculation was impossible, IL28B SNPs were secluded from analysis.
Abbreviation: WBC, white blood cell count.
*n = 103.
†$P < 0.01$.
‡$P < 0.05$.

calculation was impossible, IL28B SNPs were excluded from the analysis. Likewise, core aa.70 and IL28B were extracted as independent variables associated with nEVR (Table 4). In performing the analysis for SVR and relapse, we excluded patients with extended length of therapy to standardize the treatment periods. Because this restriction reduced the number of available patients for the analysis, we included 30 additional patients (Supporting Table 1) with available clinical information, including HCV core, NS5A, and

Table 4. Factors Associated with nEVR Analyzed by Uni- and Multivariate Logistic Regression Analysis*

		Univariate			Multivariate		
		OR	95% CI	*P* Value	OR	95% CI	*P* Value
Age, years	60≤	1.18	0.42-3.30	0.75			
Gender	Male	0.86	0.31-2.38	0.77			
ISDR 2224-2248	1≤	0.97	0.29-3.28	0.96			
IRRDR 2340-2382	4≤	0.25	0.09-0.69	5.0E-03‡	0.21	0.03-1.33	0.1
Core 70	Arg	0.03	0.01-0.16	2.0E-08‡	0.04	0.00-0.04	0.008‡
IL28B†	Major allele	0.05	0.01-0.17	5.4E-08‡	0.1	0.01-0.57	0.011§
Fibrosis	<2	0.28	0.08-1.0	0.04§	0.5	0.03-0.57	0.55
HCV RNA	<600 k/Ul/mL	0.19	0.02-1.5	0.08			
BMI	<23	0.97	0.36-2.58	0.95			
ALB	4.1 g/dL≤	0.69	0.26-1.85	0.46			
γ-GTP	50 IU/mL≤	1.95	0.73-5.22	0.18			
ALT	60 IU/mL<	0.38	0.14-1.03	0.05			
T-Cho	<170 mg/dL	0.34	0.11-1.03	0.06			
WBC	4,700/μL≤	0.64	0.23-1.76	0.38			
Hb	14 g/dL≤	0.82	0.29-2.26	0.70			
PLT	150,000/μL≤	0.42	0.15-1.19	0.10			
AFP	10 ng/mL≤	5.12	1.82-14.4	0.001‡	3.5	0.52-23.2	0.20
Peg-IFN dose (%)	80≤	0.37	0.14-1.01	0.048§	0.9	0.13-5.93	0.89
RBV dose (%)	80≤	0.38	0.12-1.23	0.10			

Abbreviation: WBC, white blood cell count.
*n = 103.
†n = 89.
‡$P < 0.01$.
§$P < 0.05$.

Table 5. Factors Associated With SVR Analyzed by Uni- and Multivariate Logistic Regression Analysis*

		Univariate			Multivariate		
		OR	95% CI	*P* Value	OR	95% CI	*P* Value
Age, years	60≤	0.8	0.34-1.78	0.55			
Gender	Male	1.4	0.61-3.22	0.43			
ISDR 2224-2248	1≤	6.3	1.98-20.26	0.001†	13.4	1.86-96.5	0.010†
IRRDR 2340-2382	4≤	11.1	4.07-30.54	4.08E-07‡	13.8	3.31-57.4	0.0003‡
Core 70	Arg	3.2	1.37-7.59	0.007‡	2.2	0.43-11.7	0.34
IL28B	Major allele	9.6	2.92-31.34	0.00003‡	16.8	2.04-139	0.009‡
Fibrosis	<2	3.1	1.33-7.23	0.008‡	1.4	0.31-6.64	0.65
HCV RNA	<600 k/UI/mL	3.5	1.39-9.02	0.007‡	3.5	0.72-17.3	0.12
BMI	<23	1.0	0.44-2.20	0.97			
ALB	4.1 g/dL≤	0.9	0.39-1.96	0.75			
γ-GTP	<50 IU/mL	2.6	1.13-5.88	0.02†	3.5	0.90-13.47	0.07
ALT	≤60 IU/mL	0.8	0.35-1.77	0.57			
T-Cho	<170 mg/dL	1.7	0.71-3.94	0.24			
WBC	<4,700/μL	0.8	0.36-1.87	0.64			
Hb	<14 g/dL	0.9	0.35-2.13	0.75			
PLT	150,000/μL≤	2.6	1.06-6.56	0.03†	3.5	0.71-16.8	0.20
AFP	<10 ng/mL	3.7	1.49-9.29	0.004‡	3.4	0.54-21.2	0.20
Peg-IFN dose (%)	80≤	2.2	0.96-5.13	0.06			
RBV dose (%)	80≤	0.8	0.37-1.92	0.68			

Abbreviation: WBC, white blood cell count.
*n = 97.
†$P < 0.05$.
‡$P < 0.01$.

IL28B SNPs. Those 30 patients were consecutively introduced the Peg-IFN/RBV therapy at Yamanashi University Hospital in succession to the initial 103 patients. As a result, 97 patients were available for SVR analysis, and 78 patients were available for relapse analysis. ISDR aa.2224-2248, IRRDR aa.2340-2382, and IL28B SNPs were extracted as the independent variables affecting SVR (Table 5). On the other hand, IRRDR-V3 aa.2360-2377 was extracted as an independent factor for relapse (Supporting Table 2).

Contribution of IL28B SNPs and NS5A aa.2340-2382 in Determining Treatment Response. Because multivariate analysis finally extracted IL28B SNPs and IRRDR aa.2340-2382 as the two most-significant variables determining final outcome, the correlation of IL28B SNPs and IRRDR aa.2340-2382 in association with final outcome was further investigated. Alignment of IRRDR aa.2340-2382 in association with SVR was demonstrated (Fig. 2). By this analysis, it was evident that three or more mutations in IRRDR aa.2340-2382 were significantly associated with SVR. Last, to disclose the viral sequence contribution in the determination of final outcome in IL28B TT haplotype patients with the standard therapy (n = 47), sliding-window analysis was performed (Fig. 1E). As demonstrated here, NS5A IRRDR aa.2340-2379 (~2382) was finally extracted as the most-significant viral region contributing to final outcome ($P = 2.47E-05$).

The contribution of these three viral regions in the phase-specific treatment responses is schematically illustrated (Fig. 3).

Discussion

In this study, we determined 103 complete HCV ORF sequences in consecutive Japanese patients, infected with genotype 1b HCV and given PEG-IFN/RBV therapy, and systematically searched and investigated the contribution of viral regions associated with the phase-specific treatment responses with IL28B SNP haplotypes. To our knowledge, this study is most comprehensive in the following aspects: (1) complete HCV ORF studied with the largest analyzed number of patients; (2) analyzed according to viral kinetics closely related to outcome; (3) unified to a single genotype (1b); (4) unified background of patients; (5) introduction of a sliding-window method to screen the responsible viral regions systematically; and (6) analysis of IL28B SNPs.

In a recent randomized, controlled study of Peg-IFN/RBV combination therapy, the status of patients according to response to Peg-IFN/RBV therapy at 12 weeks showed a marked correlation with final outcome, and viral response at week 12 has been considered as a useful predictor in early-response–guided therapy.[26] In agreement with the previous study, virological responses to Peg-IFN/RBV at week 12 had a