suggest that ISDR mutations will remain a significant predictor of good response to IFN therapies, including 72-week extension.

In conclusion, ISDR mutations are the most effective predictors of treatment outcomes in multivariate analysis. The number of mutations in the ISDR sequence of HCV-1b ($\geq$2) is the most effective parameter which will facilitate further the selection of patients with a high likelihood of response to PEG-IFN plus RBV treatment.

Acknowledgments This study was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology-Japan; the Japan Society for the Promotion of Science; the Ministry of Health, Labour and Welfare-Japan; the Japan Health Sciences Foundation; the Miyakawa Memorial Research Foundation; and the National Institute of Biomedical Innovation. The following hospitals participated in the Ochanomizu-Liver Conference Study Group: Oume City General Hospital, Kashiwa City Hospital, Kudanzaka Hospital, Showa General Hospital, Shuwa General Hospital, Soka Municipal Hospital, Tama-Nambu Chiiki Hospital, Tuchiura Kyodo General Hospital, Tokyo Kyosai Hospital, Tokyo Metropolitan Ohtsuka Hospital, Tokyo Metropolitan Fuchu Hospital, Tokyo Metropolitan Bokutoh Hospital, Toride Kyodo General Hospital, Nakano General Hospital, Hokushin General Hospital, Mishima Social Insurance Hospital, Musashino Red Cross Hospital, Yokosuka Kyosai Hospital, Yokohama City Minato Red Cross Hospital.

References

1. Major ME, Feinstone SM. The molecular virology of hepatitis C. Hepatology. 1997;25:1527–38.
2. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science. 1989;244:359–62.
3. Reichard O, Andersson J, Schvarcz R, Weiland O. Ribavirin treatment for chronic hepatitis C. Lancet. 1991;337:1058–61.
4. Dibisceglie AM, Shindo M, Fong TL, Fried MW, Swain MG, Bergasa NV, et al. A pilot-study of ribavirin therapy for chronic hepatitis-C. Hepatology. 1992;16:649–54.
5. Dusheiko G, Main J, Thomas H, Reichard O, Lee C, Dhillon A, et al. Ribavirin treatment for patients with chronic hepatitis C: results of a placebo-controlled study. J Hepatol. 1996;25:591–8.
6. Reichard O, Norkrans G, Fryden A, Braconier JH, Sonnerborg A, Weiland O, et al. Randomised, double-blind, placebo-controlled trial of interferon alpha-2b with and without ribavirin for chronic hepatitis C. Lancet. 1998;351:83–7.
7. Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, et al. Randomised trial of interferon alpha 2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha 2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. Lancet. 1998;352:1426–32.
8. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. N Engl J Med. 1998;339:1485–92.
9. Glue P, Fang JWS, Rouzier-Panis R, Raffanel C, Sabo R, Gupta SK, et al. Pegylated interferon-alpha 2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. Clin Pharmacol Ther. 2000;68:556–67.
10. Fried MW, Shiffman ML, Reddy KR, Smith C, Marionos G, Goncales FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med. 2002;347:975–82.
11. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J Clin Invest. 1995;96:224–30.
12. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med. 1996;334:77–81.
13. Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. Hepatology. 1997;25:750–3.
14. Enomoto N, Kurosaki M, Tanaka Y, Marumo F, Sato C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. J Gen Virol. 1994;75:1361–9.
15. Kato N, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, et al. Molecular-cloning of the human hepatitis-C virus genome from Japanese patients with non-A, non-B hepatitis. Proc Natl Acad Sci USA. 1990;87:9524–8.
16. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology. 2005;48:372–80.
17. Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. Hepatology. 2007;46:1357–64.
18. Zeuzem S, Lee JH, Roth WK. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon Alfa. Hepatology. 1997;25:740–4.
19. Khorsi H, Castelain S, Wyseur A, Izopet J, Canva V, Rombout A, et al. Mutations of hepatitis C virus 1b NS5A 2209–2248 amino acid sequence do not predict the response to recombinant interferon-alfa therapy in French patients. J Hepatol. 1997;27:72–7.
20. Squadrito G, Leone F, Sartori M, Nalpas B, Berthelot P, Raimondo G, et al. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alfa. Gastroenterology. 1997;113:567–72.
21. Hofgartner WT, Polyak SJ, Sullivan DG, Carithers RL, Gretch DR. Mutations in the NS5A gene of hepatitis C virus in North American patients infected with HCV genotype 1a or 1b. J Med Virol. 1997;53:118–26.
22. Squadrito G, Orlando ME, Cacciola I, Rumi MG, Artini M, Picciotto A, et al. Long-term response to interferon alpha is unrelated to "interferon sensitivity determining region" variability in patients with chronic hepatitis C virus-1b infection. J Hepatol. 1999;30:1023–7.
23. Chung RT, Monto A, Dienstag JL, Kaplan LM. Mutations in the NS5A region do not predict interferon-responsiveness in American patients infected with genotype 1b hepatitis C virus. J Med Virol. 1999;58:353–8.
24. Sarrazin C, Berg T, Lee JH, Teuber G, Dietrich CF, Roth WK, et al. Improved correlation between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. J Hepatol. 1999;30:1004–13.
25. Honda T, Katano Y, Urano F, Murayama M, Hayashi K, Ishigami M, et al. Efficacy of ribavirin plus interferon-alpha in patients aged $\geq$60 years with chronic hepatitis C. J Gastroenterol Hepatol. 2007;22:989–95.

26. Hung CH, Chen CH, Lee CM, Wu CM, Hu TH, Wang JH, et al. Association of amino acid variations in the NS5A and E2-PePHD region of hepatitis C virus 1b with hepatocellular carcinoma. J Viral Hepatitis. 2008;15:58–65.
27. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Predictive factors of virological non-response to interferon-ribavirin combination therapy for patients infected with hepatitis C virus of genotype1b and high viral load. J Med Virol. 2006;78:83–90.
28. Mori N, Imamura M, Kawakami Y, Saneto H, Kawaoka T, Takaki S, et al. Randomized trial of high-dose interferon-alpha-2b combined with ribavirin in patients with chronic hepatitis C: correlation between amino acid substitutions in the core/NS5A region and virological response to interferon therapy. J Med Virol. 2009;81:640–9.
29. Yen YH, Hun CH, Hu TH, Chen CH, Wu CM, Vvang JH, et al. Mutations in the interferon sensitivity-determining region (nonstructural 5A amino acid 2209–2248) in patients with hepatitis C-1b infection and correlating response to combined therapy of pegylated interferon and ribavirin. Aliment Pharmacol Ther. 2008;27:72–9.
30. Hung CH, Lee CM, Lu SN, Lee JF, Wang JH, Tung HD, et al. Mutations in the NS5A and E2-PePHD region of hepatitis C virus type 1b and correlation with the response to combination therapy with interferon and ribavirin. J Viral Hepatitis. 2003;10: 87–94.
31. Murayama M, Katano Y, Nakano I, Ishigami M, Hayashi K, Honda T, et al. A mutation in the interferon sensitivity-determining region is associated with responsiveness to interferon-ribavirin combination therapy in chronic hepatitis patients infected with a Japan-specific subtype of hepatitis C virus genotype 1b. J Med Virol. 2007;79:35–40.
32. Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. Hepatology. 2008;48:1753–60.
33. Murphy MD, Rosen HR, Marousek GI, Chou SW. Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. Dig Dis Sci. 2002;47:1195–205.
34. Yang SS, Lai MY, Chen DS, Chen GH, Kao JH. Mutations in the NS5A and E2-PePHD regions of hepatitis C virus genotype 1b and response to combination therapy of interferon plus ribavirin. Liver Int. 2003;23:426–33.
35. Chayama K, Tsubota A, Kobayashi M, Okamoto K, Hashimoto M, Miyano Y, et al. Pretreatment virus load and multiple amino acid substitutions in the interferon sensitivity-determining region predict the outcome of interferon treatment in patients with chronic genotype 1b hepatitis C virus infection. Hepatology. 1997;25:745–9.
36. Watanabe H, Nagayama K, Enomoto N, Itakura J, Tanabe Y, Hamano K, et al. Sequence elements correlating with circulating viral load in genotype 1b hepatitis C virus infection. Virology. 2003;311:376–83.
37. Gale MJ, Korth MJ, Tang NM, Tan SL, Hopkins DA, Dever TE, et al. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. Virology. 1997;230:217–27.
38. Gale M, Blakely CM, Kwieciszewski B, Tan SL, Dossett M, Tang NM, et al. Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. Mol Cell Biol. 1998;18:5208–18.
39. Abe T, Kaname Y, Hamamoto I, Tsuda Y, Wen XY, Taguwa S, et al. Hepatitis C virus nonstructural protein 5A modulates the toll-like receptor–MyD88-dependent signaling pathway in macrophage cell lines. J Virol. 2007;81:8953–66.
40. Miyanari Y, Atsuzawa K, Usuda N, Watashi K, Hishiki T, Zayas M, et al. The lipid droplet is an important organelle for hepatitis C virus production. Nat Cell Biol. 2007;9:1089–97.
41. Watanabe H, Enomoto N, Nagayama K, Izumi N, Marumo F, Sato C, et al. Number and position of mutations in the interferon (IFN) sensitivity-determining region of the gene for nonstructural protein 5A correlate with IFN efficacy in hepatitis C virus genotype 1b infection. J Infect Dis. 2001;183:1195–203.
42. Witherell GW, Beineke P. Statistical analysis of combined substitutions in nonstructural 5A region of hepatitis C virus and interferon response. J Med Virol. 2001;63:8–16.
43. Blight KJ, Kolykhalov AA, Rice CM. Efficient initiation of HCV RNA replication in cell culture. Science. 2000;290:1972–4.
44. Maekawa S, Enomoto N, Sakamoto N, Kurosaki M, Ueda E, Kohashi T, et al. Introduction of NS5A mutations enables subgenomic HCV replicon derived from chimpanzee-infectious HC-J4 isolate to replicate efficiently in Huh-7 cells. J Viral Hepatitis. 2004;11:394–403.
45. Kohashi T, Maekawa S, Sakamoto N, Kurosaki M, Watanabe H, Tanabe Y, et al. Site-specific mutation of the interferon sensitivity-determining region (ISDR) modulates hepatitis C virus replication. J Viral Hepatitis. 2006;13:582–90.
46. Berg T, von Wagner M, Nasser S, Sarrazin C, Heintges T, Gerlach T, et al. Extended treatment duration for hepatitis C virus type 1: comparing 48 versus 72 weeks of peginterferon-alfa-2a plus ribavirin. Gastroenterology. 2006;130:1086–97.
47. Ferenci P, Laferl H, Scherzer TM, Maieron A, Hofer H, Stauber R, et al. Peginterferon alfa-2a/ribavirin for 48 or 72 weeks in hepatitis C types 1 and 4 patients with slow virologic response. Gastroenterology 2009 [Epub ahead of print].
48. Watanabe S, Enomoto N, Koike K, Izumi N, Takikawa H, Hashimoto E, et al. Prolonged treatment with pegylated interferon a 2b plus ribavirin improves sustained virological response in chronic hepatitis C genotype 1 patients with late response in a clinical real-life setting in Japan. Hepatol Res 2009 [Epub ahead of print].

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors

Masayuki Kurosaki[1], Yasuhito Tanaka[2], Nao Nishida[3], Naoya Sakamoto[4], Nobuyuki Enomoto[5], Masao Honda[6], Masaya Sugiyama[2], Kentaro Matsuura[2], Fuminaka Sugauchi[2], Yasuhiro Asahina[1], Mina Nakagawa[4], Mamoru Watanabe[4], Minoru Sakamoto[5], Shinya Maekawa[5], Akito Sakai[6], Shuichi Kaneko[6], Kiyoaki Ito[7], Naohiko Masaki[7], Katsushi Tokunaga[3], Namiki Izumi[1,*], Masashi Mizokami[2,7]

[1]*Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan;* [2]*Department of Virology, Liver Unit, Nagoya City University, Graduate School of Medical Sciences, Nagoya, Japan;* [3]*Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan;* [4]*Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan;* [5]*First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan;* [6]*Department of Gastroenterology, Kanazawa University, Graduate School of Medicine, Kanazawa, Japan;* [7]*Research Center for Hepatitis and Immunology, International Medical Center of Japan, Konodai Hospital, Ichikawa, Japan*

Background & Aims: Pegylated interferon and ribavirin (PEG-IFN/RBV) therapy for chronic hepatitis C virus (HCV) genotype 1 infection is effective in 50% of patients. Recent studies revealed an association between the *IL28B* genotype and treatment response. We aimed to develop a model for the pre-treatment prediction of response using host and viral factors.

Methods: Data were collected from 496 patients with HCV genotype 1 treated with PEG-IFN/RBV at five hospitals and universities in Japan. *IL28B* genotype and mutations in the core and IFN sensitivity determining region (ISDR) of HCV were analyzed to predict response to therapy. The decision model was generated by data mining analysis.

Results: The *IL28B* polymorphism correlated with early virological response and predicted null virological response (NVR) (odds ratio = 20.83, p <0.0001) and sustained virological response (SVR) (odds ratio = 7.41, p <0.0001) independent of other covariates. Mutations in the ISDR predicted relapse and SVR independent of *IL28B*. The decision model revealed that patients with the minor *IL28B* allele and low platelet counts had the highest NVR (84%) and lowest SVR (7%), whereas those with the major *IL28B* allele and mutations in the ISDR or high platelet counts had the lowest NVR (0–17%) and highest SVR (61–90%). The model had high reproducibility and predicted SVR with 78% specificity and 70% sensitivity.

Conclusions: The *IL28B* polymorphism and mutations in the ISDR of HCV were significant pre-treatment predictors of response to PEG-IFN/RBV. The decision model, including these host and viral factors may support selection of optimum treatment strategy for individual patients.

Introduction

Hepatitis C virus (HCV) infection is the leading cause of cirrhosis and hepatocellular carcinoma worldwide [1]. The successful eradication of HCV, defined as a sustained virological response (SVR), is associated with a reduced risk of developing hepatocellular carcinoma. Currently, pegylated interferon (PEG-IFN) plus ribavirin (RBV) is the most effective standard of care for chronic hepatitis C but the rate of SVR is around 50% in patients with HCV genotype 1 [2,3], the most common genotype in Japan, Europe, the United States, and many other countries. Moreover, 20–30% of patients with HCV genotype 1 have a null virological response (NVR) to PEG-IFN/RBV therapy [4]. The most reliable method for predicting the response is to monitor the early decline of serum HCV-RNA levels during treatment [5] but there is no established method for prediction before treatment. Because PEG-IFN/RBV therapy is costly and often accompanied by adverse effects such as flu-like symptoms, depression and hematological abnormalities, pre-treatment predictions of those patients who are unlikely to benefit from this regimen enables ineffective treatment to be avoided.

Recently, it has been reported through a genome-wide association study (GWAS) of patients with genotype 1 HCV that single nucleotide polymorphisms (SNPs) located near the *IL28B* gene are strongly associated with a response to PEG-IFN/RBV therapy in

Keywords: *IL28B*; ISDR; Peg-interferon; Ribavirin; Data mining; Decision tree.
Received 14 March 2010; received in revised form 22 June 2010; accepted 7 July 2010
* Corresponding author. Address: Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. Tel.: +81 422 32 3111; fax: +81 422 32 9551.
E-mail address: nizumi@musashino.jrc.or.jp (N. Izumi).

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

Research Article

Table 1. Baseline characteristics of all patients, and patients assigned to the model building or validation groups.

	All patients n = 496	Model group n = 331	Validation group n = 165
Gender: male	250 (50%)	170 (51%)	80 (48%)
Age (years)	57.1 ± 9.9	56.8 ± 9.7	57.5 ± 10.2
ALT (IU/L)	78.6 ± 60.8	78.1 ± 61.4	79.7 ± 59.6
GGT (IU/L)	59.3 ± 63.6	58.9 ± 62.0	60.2 ± 66.9
Platelets (10^9/L)	154 ± 53	153 ± 52	154 ± 56
Fibrosis: F3-4	121 (24%)	80 (24%)	41 (25%)
HCV-RNA: >600,000 IU/ml	409 (82%)	273 (82%)	136 (82%)
ISDR mutation: ≤1	220 (88%)	290 (88%)	145 (88%)
Core 70 (Arg/Gln or His)	293 (59%)/203 (41%)	197 (60%)/134 (40%)	96 (58%)/69 (42%)
Core 91 (Leu/Met)	299 (60%)/197 (40%)	200 (60%)/131 (40%)	99 (60%)/66 (40%)
IL28B: Minor allele	151 (30%)	101 (31%)	50 (30%)
SVR	194 (39%)	129 (39%)	65 (39%)
Relapse	152 (31%)	103 (31%)	49 (30%)
NVR	150 (30%)	99 (30%)	51 (31%)

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Gln, glutamine; His, histidine; Leu, leucine; Met, methionine; Minor, heterozygote or homozygote of minor allele; SVR, sustained virological response; NVR, null virological response.

Japanese [6], European [7], and a multi-ethnic population [8,9]. The last three studies focused on the association of SNPs in the *IL28B* region with SVR [7–9] but we found a stronger association with NVR [6]. In addition to these host genetic factors, we have reported that mutations within a stretch of 40 amino acids in the NS5A region of HCV, designated as the IFN sensitivity determining region (ISDR), are closely associated with the virological response to IFN therapy: a lower number of mutations is associated with treatment failure [10–13]. Amino acid substitutions at positions 70 and 91 of the HCV core region (Core70, Core91) also have been reported to be associated with response to PEG-IFN/RBV therapy: glutamine (Gln) or histidine (His) at Core70 and methionine (Met) at Core91 are associated with treatment resistance [4,14]. The importance of substitutions in the HCV core and ISDR was confirmed recently by a Japanese multicenter study [15]. How these viral factors contribute to response to therapy is yet to be determined. For general application in clinical practice, host genetic factors and viral factors should be considered together.

Data mining analysis is a family of non-parametric regression methods for predictive modeling. Software is used to automatically explore the data to search for optimal split variables and to build a decision tree structure [16]. The major advantage of decision tree analysis over logistic regression analysis is that the results of the analysis are presented in the form of flow chart, which can be interpreted intuitively and readily made available for use in clinical practice [17]. The decision tree analysis has been utilized to define prognostic factors in various diseases [18–25]. We have reported recently its usefulness for the prediction of an early virological response (undetectable HCV-RNA within 12 weeks of therapy) to PEG-IFN/RBV therapy in chronic hepatitis C [26].

This study aimed to define the pre-treatment prediction of response to PEG-IFN/RBV therapy through the integrated analysis of host factors, such as the *IL28B* genetic polymorphism and various clinical covariates, as well as viral factors, such as mutations in the HCV core and ISDR and serum HCV-RNA load. In addition, for the general application of these results in clinical practice, decision models for the pre-treatment prediction of response were determined by data mining analysis.

Materials and methods

Patients

This was a multicentre retrospective study supported by the Japanese Ministry of Health, Labor and Welfare. Data were collected from a total of 496 chronic hepatitis C patients who were treated with PEG-IFN alpha and RBV at five hospitals and universities throughout Japan. Of these, 98 patients also were included in the original GWAS analysis [6]. The inclusion criteria in this study were as follows (1) infection by genotype 1b, (2) lack of co-infection with hepatitis B virus or human immunodeficiency virus, (3) lack of other causes of liver disease, such as autoimmune hepatitis, and primary biliary cirrhosis, (4) completion of at least 24 weeks of therapy, (5) adherence of more than 80% to the planned dose of PEG-IFN and RBV for the NVR patients, (6) availability of DNA for the analysis of the genetic polymorphism of *IL28B*, and (7) availability of serum for the determination of mutations in the ISDR and substitutions of Core70 and Core91 of HCV. Patients received PEG-IFN alpha-2a (180 μg) or 2b (1.5 μg/kg) subcutaneously every week and were administered a weight adjusted dose of RBV (600 mg for <60 kg, 800 mg for 60–80 kg, and 1000 mg for >80 kg daily) which is the recommended dosage in Japan. Written informed consent was obtained from each patient and the study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the institutional ethics review committee. The baseline characteristics are listed in Table 1. For the data mining analysis, 67% of the patients (331 patients) were assigned randomly to the model building group and 33% (165 patients) to the validation group. There were no significant differences in the clinical backgrounds between these two groups.

Laboratory and histological tests

Blood samples were obtained before therapy and were analyzed for hematologic tests and for blood chemistry and HCV-RNA. Sequences of ISDR and the core region of HCV were determined by direct sequencing after amplification by reverse-transcription and polymerase chain reaction as reported previously [4,11]. Genetic polymorphism in one tagging SNP located near the *IL28B* gene (rs8099917) was determined by the GWAS or DigiTag2 assay [27]. Homozygosity (GG) or heterozygosity (TG) of the minor sequence was defined as having the *IL28B* minor allele, whereas homozygosity for the major sequence (TT) was

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

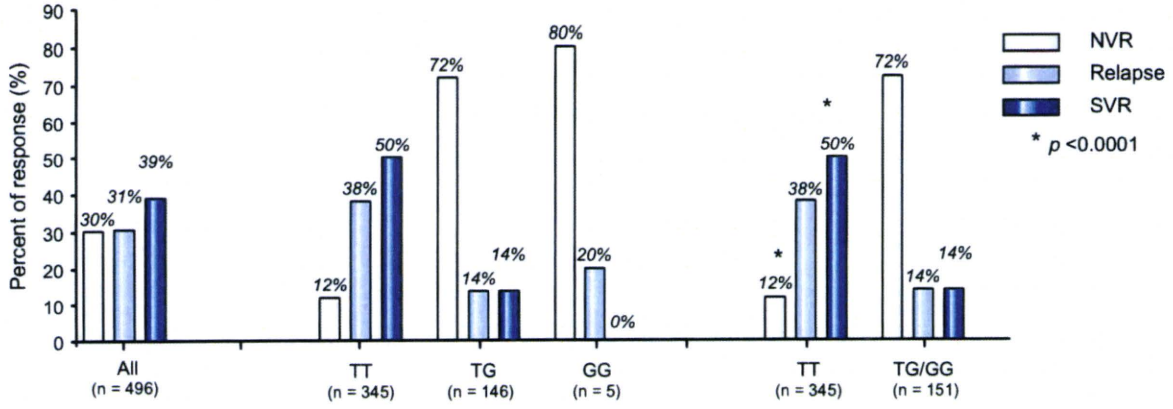

Fig. 1. Association between the IL28B genotype (rs8099917) and treatment response. The rates of response to treatment are shown for each rs8099917 genotype. The rate of null virological response (NVR), relapse, and sustained virological response (SVR) is shown. The *p* values are from Fisher's exact test. The rate of NVR was significantly higher ($p < 0.0001$) and the rate of SVR was significantly lower ($p < 0.0001$) in patients with the *IL28B* minor allele compared to those with the major allele. [This figure appears in colour on the web.]

defined as having the *IL28B* major allele. In this study, NVR was defined as a less than 2 log reduction of HCV-RNA at week 12 and detectable HCV-RNA by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor, Roche Diagnostic systems, CA) at week 24 during therapy. RVR (rapid virological response) and complete early virological response (cEVR) were defined as undetectable HCV-RNA at 4 weeks and 12 weeks during therapy and SVR was defined as undetectable HCV-RNA 24 weeks after the completion of therapy. Relapse was defined as reappearance of HCV-RNA after the completion of therapy. The stage of liver fibrosis was scored according to the METAVIR scoring system: F0 (no fibrosis), F1 (mild fibrosis: portal fibrosis without septa), F2 (moderate fibrosis: few septa), F3 (severe fibrosis: numerous septa without cirrhosis) and F4 (cirrhosis). Percentage of steatosis was quantified in 111 patients by determining the average proportion of hepatocytes affected by steatosis.

Statistical analysis

Associations between pre-treatment variables and treatment response were analyzed by univariate and multivariate logistic regression analysis. Associations between the *IL28B* polymorphism and sequences of HCV were analyzed by Fisher's exact test. SPSS software v.15.0 (SPSS Inc., Chicago, IL) was used for these analyses. For the data mining analysis, IBM-SPSS Modeler version 13.0 (IBM-SPSS Inc., Chicago, IL) software was utilized as reported previously [26]. The patients used for model building were divided into two groups at each step of the analysis based on split variables. Each value of each variable was considered as a potential split. The optimum variables and cut-off values were determined by a statistical search algorithm to generate the most significant division into two prognostic subgroups that were as homogeneous as possible for the probability of SVR. Thereafter, each subgroup was evaluated again and divided further into subgroups. This procedure was repeated until no additional significant variable was detected or the sample size was below 15. To avoid over-fitting, 10-fold cross validation was used in the tree building process. The reproducibility of the resulting model was tested with the data from the validation patients.

Results

Association between the IL28B (rs8099917) genotype and the PEG-IFN/RBV response

The rs8099917 allele frequency was 70% for TT ($n = 345$), 29% for TG ($n = 146$), and 1% for GG ($n = 5$). We defined the *IL28B* major allele as homozygous for the major sequence (TT) and the *IL28B* minor allele as homozygous (GG) or heterozygous (TG) for the minor sequence. The rate of NVR was significantly higher (72% vs. 12%, $p < 0.0001$) and the rate of SVR was significantly lower (14% vs. 50%, $p < 0.0001$) in patients with the *IL28B* minor allele compared to those with the major allele (Fig. 1).

Effect of the IL28B polymorphism, substitutions in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV

Patients were stratified according to their IL28B allele type, the number of mutations in the ISDR, the amino acid substitutions in Core70 and Core91, and the rate of undetectable HCV-RNA at 4, 8, 12, 24, and 48 weeks after the start of therapy was analyzed (Fig. 2A–D). The rate of undetectable HCV-RNA was significantly higher in patients with the IL28B major allele than the minor allele, in patients with two or more mutations in the ISDR compared to none or only one mutation, in patients with arginine (Arg) at Core70 rather than Gln/His, and in patients with leucine (Leu) at Core91 rather than Met. The difference was most significant when stratified by the *IL28B* allele type. The rate of RVR and cEVR was significantly more frequent in patients with the *IL28B* major allele compared with those with the *IL28B* minor allele: 9% vs. 3% for RVR ($p < 0.005$) and 57% vs. 11% for cEVR ($p < 0.0001$). These findings suggest that IL28B has the greatest impact on early virological response to therapy.

Association between substitutions in the ISDR and relapse after the completion of therapy

Patients were stratified according to the *IL28B* allele, number of mutations in the ISDR, and amino acid substitutions of Core70 and Core91, and the rate of relapse was analyzed (Fig. 3A and B). Among patients who achieved cEVR, the rate of relapse was significantly lower in patients with two or more mutations in the ISDR compared to those with only one or no mutations (15% vs. 31%, $p < 0.005$) (Fig. 3 B). On the other hand, the relapse rate was not different between the *IL28B* major and minor alleles within patients who achieved RVR (3% vs. 0%) or cEVR (28% vs. 29%) (Fig. 3A). Amino acid substitutions of Core70 and Core91 were not associated with the rate of relapse (data not shown).

Factors associated with response by multivariate logistic regression analysis

By univariate analysis, the minor allele of *IL28B* ($p < 0.0001$), one or no mutations in the ISDR ($p = 0.03$), high serum level of

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

Research Article

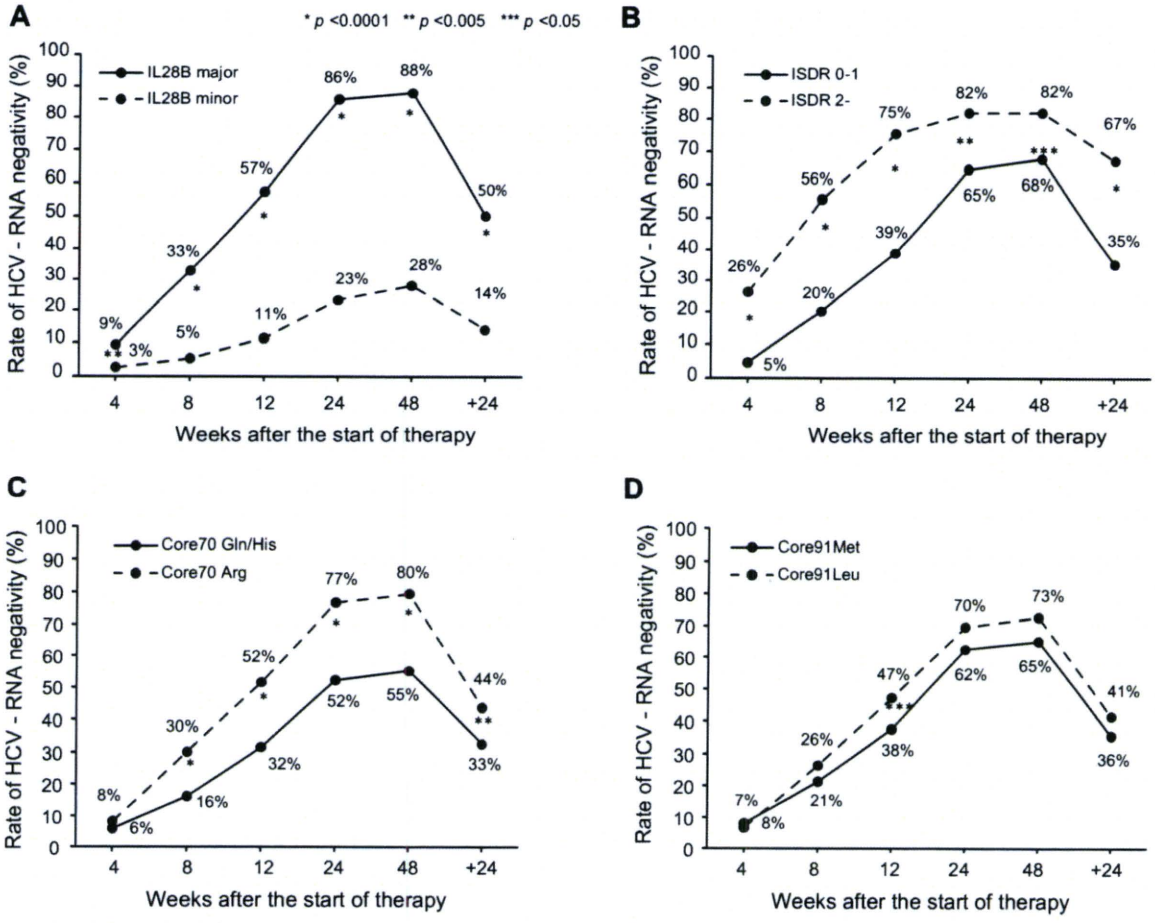

Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70 and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele), (B) the number of mutations in the ISDR (0–1 mutation vs. 2 or more mutations), amino acid substitutions of (C) Core70 (Gln/His vs. Arg), and (D) Core91 (Met vs. Leu). The *p* values are from Fisher's exact test.

HCV-RNA ($p = 0.035$), Gln or His at Core70 ($p < 0.0001$), low platelet counts ($p = 0.009$), and advanced fibrosis ($p = 0.0002$) were associated with NVR. By multivariate analysis, the minor allele of *IL28B* (OR = 20.83, 95%CI = 11.63–37.04, $p < 0.0001$) was associated with NVR independent of other covariates (Table 2). Notably, mutations in the ISDR ($p = 0.707$) and at amino acid Core70 ($p = 0.207$) were not significant in multivariate analysis due to the positive correlation with the *IL28B* polymorphism ($p = 0.004$ for ISDR and $p < 0.0001$ for Core70, Fig. 4).

Genetic polymorphism of *IL28B* also was associated with SVR (OR = 7.41, 95% CI = 4.05–13.57, $p < 0.0001$) independent of other covariates, such as platelet counts, fibrosis, and serum levels of HCV-RNA. Mutation in the ISDR was an independent predictor of SVR (OR = 2.11, 95% CI = 1.06–4.18, $p = 0.033$) but the amino acid at Core70 was not (Table 3).

Factors associated with the IL28B polymorphism

Patients with the *IL28B* minor allele had significantly higher serum level of gamma-glutamyltransferase (GGT) and a higher frequency of hepatic steatosis (Table 4). When the association between the *IL28B* polymorphism and HCV sequences was analyzed, Gln or His at Core70, that is linked to resistance to PEG-IFN and RBV therapy [4,14,15], was significantly more frequent in patients with the minor *IL28B* allele than in those with the major allele (67% vs. 30%, $p < 0.0001$) (Fig. 4). Other HCV sequences with an IFN resistant phenotype also were more prevalent in patients with the minor *IL28B* allele than those with the major allele: Met at Core91 (46% vs. 37%, $p = 0.047$) and one or no mutations in the ISDR (94% vs. 85%, $p = 0.004$) (Fig. 4).

Data mining analysis

Data mining analysis was performed to build a model for the prediction of SVR and the result is shown in Fig. 5. The analysis selected four predictive variables, resulting in six subgroups of patients. Genetic polymorphism of *IL28B* was selected as the best predictor of SVR. Patients with the minor *IL28B* allele had a lower probability of SVR and a higher probability of NVR than those with the major *IL28B* allele (SVR: 14% vs. 50%, NVR: 72% vs.

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

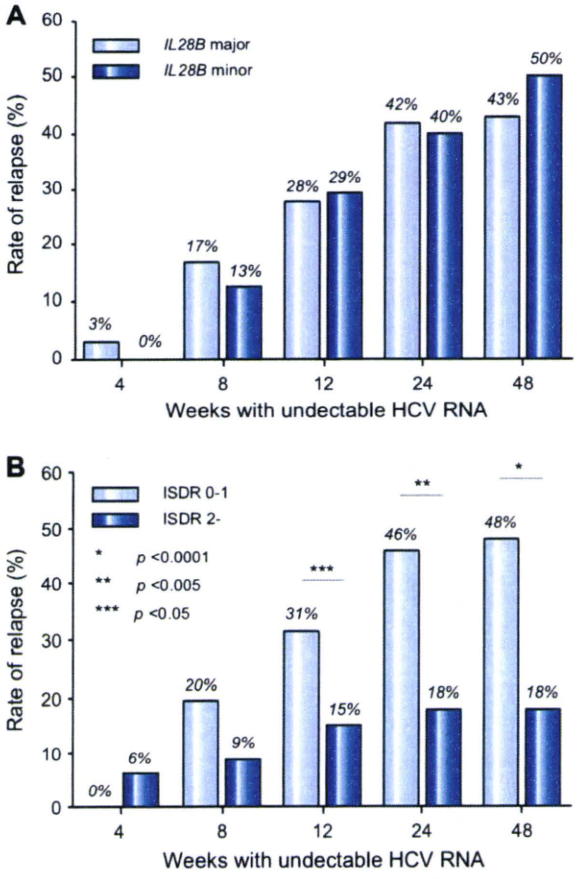

Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele) and (B) the number of mutations in the ISDR (0–1 mutation vs. 2 or more mutations). The *p* values are from Fisher's exact test. [This figure appears in colour on the web.]

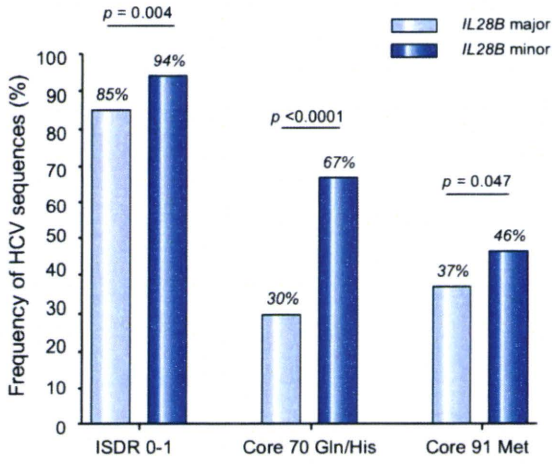

Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met at Core91. *p* values are from Fisher's exact test. [This figure appears in colour on the web.]

12%). After stratification by the *IL28B* allele, patients with low platelet counts (<140 × 10^9/L) had a lower probability of SVR and higher probability of NVR than those with high platelet counts (≥140 × 10^9/L): for the minor *IL28B* allele, SVR was 7% vs. 19%, and NVR was 84% vs. 62%, and for the major *IL28B* allele, SVR was 32% vs. 66% and NVR was 16% vs. 8%. Among patients with the major *IL28B* allele and low platelet counts, those with two or more mutations in the ISDR had a higher probability of SVR and lower probability of relapse than those with one or no mutations in the ISDR (SVR: 75% vs. 27%, and relapse: 8% vs. 57%). Among patients with the major *IL28B* allele and high platelet counts, those with a low HCV-RNA titer (<600,000 IU/ml) had a higher probability of SVR and lower probability of NVR and relapse than those with a high HCV-RNA titer (SVR: 90% vs. 61%, NVR: 0% vs. 10%, and relapse: 10% vs. 29%). The sensitivity and specificity of the decision tree were 78% and 70%, respectively. The area under the receiver operating characteristic (ROC) curve of the model was 0.782 (data not shown). The proportion of patients with advanced fibrosis (F3-4) was 39% (84/217) in patients with low platelet counts (<140 × 10^9/L) compared to 13% (37/279) in those with high platelet counts (≥140 × 10^9/L).

Validation of the data mining analysis

The results of the data mining analysis were validated with 165 patients who differed from those used for model building. Each patient was allocated to one of the six subgroups for the validation using the flow-chart form of the decision tree. The rate of SVR and NVR in each subgroup was calculated. The rates of SVR and NVR for each subgroup of patients were closely correlated between the model building and the validation patients (r^2 = 0.99 and 0.98) (Fig. 6).

Discussion

The rate of NVR after 48 weeks of PEG-IFN/RBV therapy among patients infected with HCV of genotype 1 is around 20–30%. Previously, there have been no reliable baseline predictors of NVR or SVR. Because more potent therapies, such as protease and polymerase inhibitor of HCV [28,29] and nitazoxanide [30], are in clinical trials and may become available in the near future, a pre-treatment prediction of the likelihood of response may be helpful for patients and physicians, to support clinical decisions about whether to begin the current standard of care or whether to wait for emerging therapies. This study revealed that the *IL28B* polymorphism was the overwhelming predictor of NVR and is independent of host factors and viral sequences reported previously. The *IL28B* encodes a protein also known as IFN-lambda 3, which is thought to suppress the replication of various viruses including HCV [31,32]. The results of the current study and the findings of the GWAS studies [6–9] may provide the rationale for developing diagnostic testing or an IFN-lambda based therapy for chronic hepatitis C in the future.

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

Research Article

Table 2. Factors associated with NVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	*p* value	Odds ratio	95%CI	*p* value
Gender: female	0.98	0.67-1.45	0.938	1.29	0.75-2.23	0.363
Age	1.01	0.97-1.01	0.223	0.99	0.97-1.02	0.679
ALT	1.00	1.00-1.00	0.867	1.00	0.99-1.00	0.580
GGT	1.004	1.00-1.01	0.029	1.00	1.00-1.00	0.715
Platelets	0.95	0.91-0.99	0.009	0.92	0.87-0.98	0.006
Fibrosis: F3-4	2.23	1.46-3.42	0.0002	1.97	1.09-3.57	0.025
HCV-RNA: ≥600,000 IU/ml	1.83	1.05-3.19	0.035	2.49	1.17-5.29	0.018
ISDR mutation: ≤1	2.14	1.08-4.22	0.030	0.96	0.78-1.18	0.707
Core 70 (Gln/His)	3.23	2.16-4.78	<0.0001	1.41	0.83-2.42	0.207
Core 91 (Met)	1.39	0.95-2.06	0.093	1.21	0.72-2.04	0.462
IL28B: Minor allele	19.24	11.87-31.18	<0.0001	20.83	11.63-37.04	<0.0001

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Gln, glutamine; His, histidine; Met, methionine; Minor allele, heterozygote or homozygote of minor allele.

Table 3. Factors associated with SVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	*p* value	Odds ratio	95%CI	*p* value
Gender: female	0.81	0.56-1.16	0.253	0.86	0.55-1.35	0.508
Age	0.97	0.95-0.99	0.0003	0.99	0.96-1.01	0.199
ALT	1.00	1.00-1.00	0.337	1.00	1.00-1.01	0.108
GGT	1.00	1.00-1.00	0.273	1.00	1.00-1.00	0.797
Platelets	1.12	1.01-116	<0.0001	1.13	1.08-1.19	<0.0001
Fibrosis: F0-2	2.64	1.65-4.22	<0.0001	1.87	1.07-3.28	0.029
HCV-RNA: <600,000 IU/ml	2.49	1.55-3.98	0.0001	2.75	1.55-4.90	0.001
ISDR mutation: 2≤	3.78	2.14-6.68	<0.0001	2.11	1.06-4.18	0.033
Core 70 (Arg)	1.61	1.11-2.28	0.012	0.84	0.52-1.35	0.470
Core 91 (Leu)	1.28	0.88-1.85	0.185	1.26	0.81-1.96	0.300
IL28B: Major allele	6.21	3.75-10.31	<0.0001	7.41	4.05-13.57	<0.0001

ALT, alanine aminotransferase; GGT, Gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Leu, leucine; Major allele, homozygote of major allele.

Among baseline factors, IL28B was the most significant predictor of NVR and SVR. Moreover, the IL28B allele type was also correlated with early virological response: the rate of RVR and cEVR was significantly high for the IL28B major allele compared to the IL28B minor allele: 9% vs. 3% for RVR and 57% vs. 11% for cEVR (Fig. 2). On the other hand, the relapse rate was not different between the IL28B genotypes within patients who achieved RVR or cEVR (Fig. 3). We believe that optimal therapy should be based on baseline features and a response-guided approach. Our findings suggest that the IL28B genotype is a useful baseline predictor of virological response which should be used for selecting the treatment regimen: whether to treat patients with PEG-IFN and RBV or to wait for more effective future therapy including direct acting antiviral drugs. On the other hand, baseline IL28B genotype might not be suitable for determining the treatment duration in patients who started PEG-IFN/RBV therapy and whose virological response is determined because the IL28B genotype is not useful for the prediction of relapse. The duration of therapy should be personalized based on the virological response. Future studies need to explore whether the combination of baseline IL28B genotype and response-guided approach further improves the optimization of treatment duration.

The SVR rate in patients having the IL28B minor allele was 14% in the present study while it was 23% in Caucasians and 9% in African Americans in a study by McCarthy et al. [33]. On the other hand, the SVR rate in patients having the IL28B minor allele was 28% in genotypes 1/4 compared to 80% in genotypes 2/3 in a study by Rauch et al. [9]. These data imply that the impact of the IL28B polymorphism on response to therapy may be different in terms of race, geographical areas, or HCV genotypes, and that our data need to be validated in future studies including different populations and geographical areas before generalization.

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

Table 4. Factors associated with IL28B genotype.

	IL28B major allele n = 345	*IL28B* minor allele n = 151	*p* value
Gender: male	166 (48%)	84 (56%)	0.143
Age (years)	57 ± 10	57 ± 10	0.585
ALT (IU/L)	79 ± 60	78 ± 62	0.842
Platelets (10^9/L)	153 ± 54	155 ± 52	0.761
GGT (IU/L)	51 ± 45	78 ± 91	0.001
Fibrosis: F3-4	76 (22%)	45 (30%)	0.063
Steatosis:			
>10%	16/88 (18%)	13/23 (57%)	0.024
>30%	6/88 (7%)	6/23 (26%)	0.017
HCV-RNA: >600,000 IU/ml	284 (82%)	125 (83%)	1.000

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase.

Four GWAS studies have shown the association between a genetic polymorphism near the IL28B gene and response to PEG-IFN plus RBV therapy. The SNPs that showed significant association with response were rs12979860 [8] and rs8099917 [6,7,9]. There is a strong linkage-disequilibrium (LD) between these two SNPs as well as several other SNPs near the IL28B gene in Japanese patients [34] but the degree of LD was weaker in Caucasians and Hispanics [8]. Thus, the combination of SNPs is not useful for predicting response in Japanese patients but may improve the predictive value in patients other than Japanese who have weaker LD between SNPs.

Other significant predictors of response independent of IL28B genotype were platelet counts, stage of fibrosis, and HCV RVA load. A previous study reported that platelet count is a predictor of response to therapy [35], and the lower platelet count was related with advanced liver fibrosis in the present study. The association between response to therapy and advanced fibrosis independent of the IL28B polymorphism is consistent with a recent study by Rauch et al. [9].

There is agreement that the viral genotype is significantly associated with the treatment outcome. Moreover, viral factors such as substitutions in the ISDR of the NS5A region [10] or in the amino acid sequence of the HCV core [4] have been studied in relation to the response to IFN treatment. The amino acid Gln or His at Core70 and Met at Core91 are repeatedly reported to be associated with resistance to therapy [4,14,15] in Japanese patients but these data wait to be validated in different populations or other geographical areas. In this study, we confirmed that patients with two or more mutations in the ISDR had a higher rate of undetectable HCV-RNA at each time point during therapy. In addition, the rate of relapse among patients who achieved cEVR was significantly lower in patients with two or more mutations in ISDR compared to those with only one or no mutations (15% vs. 31%, $p < 0.05$). Thus, the ISDR sequence may be used to predict a relapse among patients who achieved virological response during therapy, while the IL28B polymorphism may be used to predict the virological response before therapy. A higher number of mutations in the ISDR are reported to have close association with SVR in Japanese [11–13,15,36] or Asian [37,38] populations but data from Western countries have been controversial [39–42]. A meta-analysis of 1230 patients including 525 patients from Europe has shown that there was a positive correlation between the SVR and the number of mutations in the ISDR in Japanese as well as in European patients [43] but this correlation was more pronounced in Japanese patients. Thus, geographical factors may account for the different impact of ISDR on treatment response, which may be a potential limitation of our study.

To our surprise, these HCV sequences were associated with the IL28B genotype: HCV sequences with an IFN resistant phenotype were more prevalent in patients with the minor *IL28B* allele than those with the major allele. This was an unexpected finding, as we initially thought that host genetics and viral sequences were completely independent. A recent study reported that the IL28B polymorphism (rs12979860) was significantly associated with HCV genotype: the IL28B minor allele was more frequent in HCV genotype 1-infected patients compared to patients infected with HCV genotype 2 or 3 [33]. Again, patients with the IL28B minor allele (IFN resistant genotype) were infected with HCV sequences that are linked to an IFN resistant phenotype. The mechanism for this association is unclear, but may be related to an interaction between the IL28B genotype and HCV sequences in the development of chronic HCV infection as discussed by McCarthy et al., since the *IL28B* polymorphism was associated with the natural clearance of HCV [44]. Alternatively, the HCV sequence within the patient may be selected during the course of chronic infection [45,46]. These hypotheses should be explored through prospective studies of spontaneous HCV clearance or by testing the time-dependent changes in the HCV sequence during the course of chronic infection.

How these host and viral factors can be integrated to predict the response to therapy in future clinical practice is an important question. Because various host and viral factors interact in the same patient, predictive analysis should consider these factors in combination. Using the data mining analysis, we constructed a simple decision tree model for the pre-treatment prediction of SVR and NVR to PEG-IFN/RBV therapy. The classification of patients based on the genetic polymorphism of *IL28B*, mutation in the ISDR, serum levels of HCV-RNA, and platelet counts, identified subgroups of patients who have the lowest probabilities of NVR (0%) with the highest probabilities of SVR (90%) as well as those who have the highest probabilities of NVR (84%) with the lowest probability of SVR (7%). The reproducibility of the model was confirmed by the independent validation based on a second

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

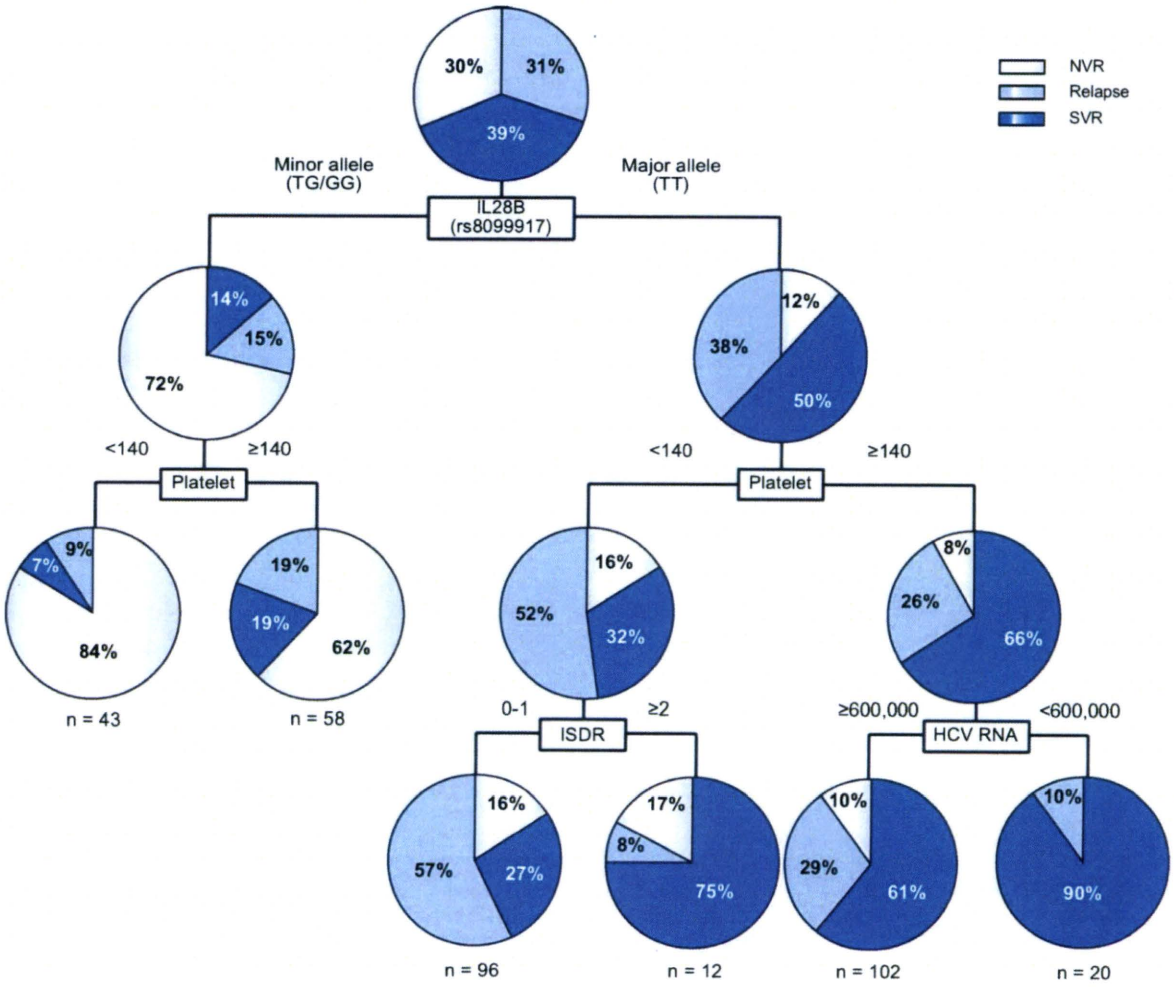

Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown. [This figure appears in colour on the web.]

group of patients. Using this model, we can rapidly develop an estimate of the response before treatment, by simply allocating patients to subgroups by following the flow-chart form, which may facilitate clinical decision making. This is in contrast to the calculating formula, which was constructed by the traditional logistic regression model. This was not widely used in clinical practice as it is abstruse and inconvenient. These results support the evidence based approach of selecting the optimum treatment strategy for individual patients, such as treating patients with a low probability of NVR with current PEG-IFN/RBV combination therapy or advising those with a high probability of NVR to wait for more effective future therapies. Patients with a high probability of relapse may be treated for a longer duration to avoid a relapse. Decisions may be based on the possibility of a response against a potential risk of adverse events and the cost of the therapy, or disease progression while waiting for future therapy.

We have previously reported the predictive model of early virological response to PEG-IFN and RBV in chronic hepatitis C [26]. The top factor selected as significant was the grade of steatosis, followed by serum level of LDL cholesterol, age, GGT, and blood sugar. The mechanism of association between these factors and treatment response was not clear at that time. To our interest, a recent study by Li et al. [47] has shown that high serum level of LDL cholesterol was linked to the IL28B major allele (CC in rs12979860). High serum level of LDL cholesterol was associated with SVR but it was no longer significant when analyzed together with the IL28B genotype in multivariate analysis. Thus, the association between treatment response and LDL cholesterol levels may reflect the underlining link of LDL cholesterol levels to IL28B genotype. Steatosis is reported to be correlated with low lipid levels [48] which suggest that IL28B genotypes may be also associated with steatosis. In fact, there were significant correlations between the IL28B genotype and the presence of steatosis in the present study (Table 4). In addition, the serum level of GGT, another predictive factor in our previous study, was significantly associated with IL28B genotype in the present study

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

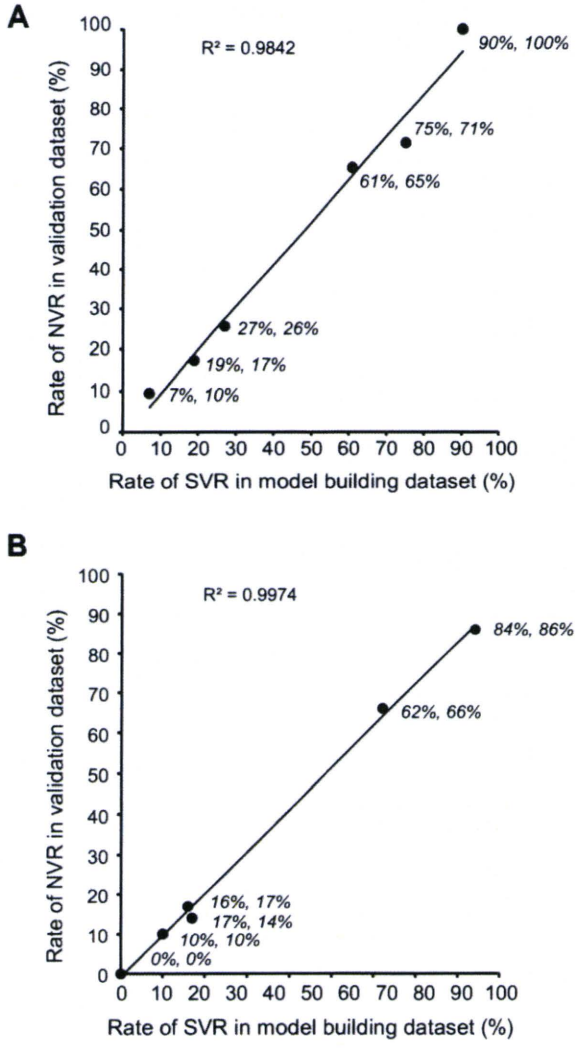

Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and plotted. The *X*-axis represents the rate of SVR or NVR in the model building patients and the *Y*-axis represents those in the validation patients. The rate of SVR and NVR in each subgroup of patients is closely correlated between the model building and the validation patients (correlation coefficient: $r^2 = 0.98–0.99$).

(Table 4). The serum level of GGT was significantly associated with NVR when examined independently but was no longer significant when analyzed together with the IL28B genotype. These observations indicate that some of the factors that we have previously identified may be associated with virological response to therapy through the underlining link to the IL28B genotype.

In conclusion, the present study highlighted the impact of the *IL28B* polymorphism and mutation in the ISDR on the pre-treatment prediction of response to PEG-IFN/RBV therapy. A decision model including these host and viral factors has the potential to support selection of the optimum treatment strategy for individual patients, which may enable personalized treatment.

Conflicts of interest

The authors who have taken part in this study declare 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, (H19-kannen-013), (H20-kannen-006).

References

[1] Ray Kim W. Global epidemiology and burden of hepatitis C. Microbes Infect 2002;4 (12):1219–1225.
[2] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales Jr FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 2002;347 (13):975–982.
[3] Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 2001;358 (9286):958–965.
[4] Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 2005;48 (6):372–380.
[5] Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. Hepatology 2003;38 (3):645–652.
[6] Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009;10:1105–1109.
[7] Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 2009;10:1100–1104.
[8] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461 (7262):399–401.
[9] Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 2010;138 (4):1338–1345.
[10] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J Clin Invest 1995;96 (1):224–230.
[11] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 1996;334 (2):77–81.
[12] Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. Hepatology 1997;25 (3):750–753.
[13] Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. Hepatology 2008;48 (6):1753–1760.
[14] Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Predictive factors of early and sustained responses to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b: amino acid substitutions in the core region and low-density lipoprotein cholesterol levels. J Hepatol 2007;46 (3):403–410.

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

Research Article

[15] Okanoue T, Itoh Y, Hashimoto H, Yasui K, Minami M, Takehara T, et al. Predictive values of amino acid sequences of the core and NS5A regions in antiviral therapy for hepatitis C: a Japanese multi-center study. J Gastroenterol 2009;44 (9):952–963.
[16] Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. Stat Med 1989;8 (5):539–550.
[17] LeBlanc M, Crowley J. A review of tree-based prognostic models. Cancer Treat Res 1995;75:113–124.
[18] Garzotto M, Beer TM, Hudson RG, Peters L, Hsieh YC, Barrera E, et al. Improved detection of prostate cancer using classification and regression tree analysis. J Clin Oncol 2005;23 (19):4322–4329.
[19] 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 (4):589–602.
[20] 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 (18):3660–3667.
[21] 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 (1):34–40.
[22] 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 (15):5440–5443.
[23] Thabane M, Simunovic M, Akhtar-Danesh N, Marshall JK. Development and validation of a risk score for post-infectious irritable bowel syndrome. Am J Gastroenterol 2009;104 (9):2267–2274.
[24] Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. Gut 2008;57 (12):1698–1703.
[25] Fonarow GC, Adams Jr KF, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. Jama 2005;293 (5):572–580.
[26] 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 (3):251–260.
[27] Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 2007;364 (1):78–85.
[28] 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 (18):1839–1850.
[29] 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 (18):1827–1838.
[30] Rossignol JF, Elfert A, El-Gohary Y, Keeffe EB. Improved virologic response in chronic hepatitis C genotype 4 treated with nitazoxanide, peginterferon, and ribavirin. Gastroenterology 2009;136 (3):856–862.
[31] Marcello T, Grakoui A, Barba-Spaeth G, Machlin ES, Kotenko SV, MacDonald MR, et al. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. Gastroenterology 2006;131 (6):1887–1898.
[32] Robek MD, Boyd BS, Chisari FV. Lambda interferon inhibits hepatitis B and C virus replication. J Virol 2005;79 (6):3851–3854.
[33] McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, Patel K, et al. Replicated association between an IL28B Gene Variant and a Sustained Response to Pegylated Interferon and Ribavirin. Gastroenterology 2010;138:2307–2314.
[34] Tanaka Y, Nishida N, Sugiyama M, Tokunaga K, Mizokami M. Λ-interferons and the single nucleotide polymorphisms: a milestone to tailor-made therapy for chronic hepatitis C. Hepatol Res 2010;40:449–460.
[35] Backus LI, Boothroyd DB, Phillips BR, Mole LA. Predictors of response of US veterans to treatment for the hepatitis C virus. Hepatology 2007;46 (1):37–47.
[36] Mori N, Imamura M, Kawakami Y, Saneto H, Kawaoka T, Takaki S, et al. Randomized trial of high-dose interferon-alpha-2b combined with ribavirin in patients with chronic hepatitis C: correlation between amino acid substitutions in the core/NS5A region and virological response to interferon therapy. J Med Virol 2009;81 (4):640–649.
[37] Hung CH, Lee CM, Lu SN, Lee JF, Wang JH, Tung HD, et al. Mutations in the NS5A and E2-PePHD region of hepatitis C virus type 1b and correlation with the response to combination therapy with interferon and ribavirin. J Viral Hepat 2003;10 (2):87–94.
[38] Yen YH, Hung CH, Hu TH, Chen CH, Wu CM, Wang JH, et al. Mutations in the interferon sensitivity-determining region (nonstructural 5A amino acid 2209–2248) in patients with hepatitis C-1b infection and correlating response to combined therapy of pegylated interferon and ribavirin. Aliment Pharmacol Ther 2008;27 (1):72–79.
[39] Zeuzem S, Lee JH, Roth WK. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. Hepatology 1997;25 (3):740–744.
[40] Squadrito G, Leone F, Sartori M, Nalpas B, Berthelot P, Raimondo G, et al. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alfa. Gastroenterology 1997;113 (2):567–572.
[41] Sarrazin C, Berg T, Lee JH, Teuber G, Dietrich CF, Roth WK, et al. Improved correlation between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. J Hepatol 1999;30 (6):1004–1013.
[42] Murphy MD, Rosen HR, Marousek GI, Chou S. Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. Dig Dis Sci 2002;47 (6):1195–1205.
[43] Pascu M, Martus P, Hohne M, Wiedenmann B, Hopf U, Schreier E, et al. Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. Gut 2004;53 (9):1345–1351.
[44] Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009;461 (7265):798–801.
[45] Kurosaki M, Enomoto N, Marumo F, Sato C. Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. Virology 1994;205 (1):161–169.
[46] Enomoto N, Kurosaki M, Tanaka Y, Marumo F, Sato C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. J Gen Virol 1994;75 (Pt 6):1361–1369.
[47] Li JH, Lao XQ, Tillmann HL, Rowell J, Patel K, Thompson A, et al. Interferon-lambda genotype and low serum low-density lipoprotein cholesterol levels in patients with chronic hepatitis C infection. Hepatology 1904;51 (6):1904–1911.
[48] Serfaty L, Andreani T, Giral P, Carbonell N, Chazouilleres O, Poupon R. Hepatitis C virus induced hypobetalipoproteinemia: a possible mechanism for steatosis in chronic hepatitis C. J Hepatol 2001;34 (3):428–434.

Please cite this article in press as: Kurosaki 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 (2010), doi:10.1016/j.jhep.2010.07.037

J Gastroenterol (2010) 45:523–536
DOI 10.1007/s00535-009-0162-3

ORIGINAL ARTICLE—LIVER, PANCREAS, AND BILIARY TRACT

Comparison of HCV-associated gene expression and cell signaling pathways in cells with or without HCV replicon and in replicon-cured cells

Yuki Nishimura-Sakurai · Naoya Sakamoto · Kaoru Mogushi · Satoshi Nagaie · Mina Nakagawa · Yasuhiro Itsui · Megumi Tasaka-Fujita · Yuko Onuki-Karakama · Goki Suda · Kako Mishima · Machi Yamamoto · Mayumi Ueyama · Yusuke Funaoka · Takako Watanabe · Seishin Azuma · Yuko Sekine-Osajima · Sei Kakinuma · Kiichiro Tsuchiya · Nobuyuki Enomoto · Hiroshi Tanaka · Mamoru Watanabe

Received: 2 September 2009 / Accepted: 2 November 2009 / Published online: 12 December 2009

Abstract

Background Hepatitis C virus (HCV) replication is affected by several host factors. Here, we screened host genes and molecular pathways that are involved in HCV replication by comprehensive analyses using two genotypes of HCV replicon-expressing cells, their *cured* cells and naïve Huh7 cells.

Methods Huh7 cell lines that stably expressed HCV genotype 1b or 2a replicon were used. The *cured* cells were established by treating HCV replicon cells with interferon-alpha. Expression of 54,675 cellular genes was analyzed by GeneChip DNA microarray. The data were analyzed by using the KEGG Pathway database.

Results Hierarchical clustering analysis showed that the gene-expression profiles of each cell group constituted clear clusters of naïve, HCV replicon-expressed, and cured cell lines. The pathway process analysis between the replicon-expressing and the *cured* cell lines identified significantly altered pathways, including MAPK, steroid biosynthesis and TGF-beta signaling pathways, suggesting that these pathways were affected directly by HCV replication. Comparison of *cured* and naïve Huh7 cells identified pathways, including steroid biosynthesis and sphingolipid metabolism, suggesting that these pathways were required for efficient HCV replication. Cytoplasmic lipid droplets were obviously increased in replicon-expressing and *cured* cells as compared to naïve cells. HCV replication was significantly suppressed by peroxisome proliferator-activated receptor (PPAR)-alpha agonists but augmented by PPAR-gamma agonists.

Conclusion Comprehensive gene expression and pathway analyses show that lipid biosynthesis pathways are crucial to support proficient virus replication. These metabolic pathways could constitute novel antiviral targets against HCV.

Keywords DNA microarray · KEGG database · HCV replicon · Lipid metabolism

Y. Nishimura-Sakurai and N. Sakamoto contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00535-009-0162-3) contains supplementary material, which is available to authorized users.

Y. Nishimura-Sakurai · N. Sakamoto (✉) · M. Nakagawa · Y. Itsui · M. Tasaka-Fujita · Y. Onuki-Karakama · G. Suda · K. Mishima · M. Yamamoto · M. Ueyama · Y. Funaoka · T. Watanabe · S. Azuma · Y. Sekine-Osajima · S. Kakinuma · K. Tsuchiya · M. Watanabe
Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan
e-mail: nsakamoto.gast@tmd.ac.jp

N. Sakamoto · M. Nakagawa · S. Kakinuma
Department for Hepatitis Control, Tokyo Medical and Dental University, Tokyo, Japan

K. Mogushi · S. Nagaie · H. Tanaka
Information Center for Medical Science, Tokyo Medical and Dental University, Tokyo, Japan

Y. Itsui
Department of Internal Medicine, Soka Municipal Hospital, Saitama, Japan

N. Enomoto
First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan

Abbreviations

HCV	Hepatitis C virus
TLR	Toll-like receptor
BMP	Bone morphogenetic protein

TGF	Transforming growth factor
FKBP	FK-binding protein
HSP	Heat shock proteins
FBS	Fetal bovine serum
YFP	Yellow fluorescence protein
FACS	Fluorescent activated cell sorting
RIN	RNA integrity number
SAM	Significance analysis of microarray
KEGG	Kyoto Encyclopedia of Genes and Genomes
EGID	NCBI Entrez Gene ID
RT-PCR	Reverse transcription-polymerase chain reaction
MTS	Dimethylthiazol carboxymethoxyphenyl sulfophenyl tetrazolium
PPAR	Peroxisome proliferator-activated receptor

Introduction

Hepatitis C virus (HCV) infection is one of the most important causative agents of acute and chronic hepatitis, liver cirrhosis and hepatocellular malignancies [1]. Currently, the most efficient combination treatment of ribavirin plus peginterferon can eliminate the virus in almost half of the patients treated [2, 3]. Thus, it is our high priority goal to understand the HCV life cycle precisely, to identify cellular cofactors for HCV replication and to develop new class antiviral therapeutics.

Molecular analyses of the HCV life cycle, virus–host interactions, and mechanisms of liver cell damage by the virus are not understood completely, mainly because of the lack of cell culture systems. These problems have been partly overcome by the development of the HCV subgenomic replicon [4] and HCV cell culture systems [5, 6]. These systems have allowed us to study the complete HCV life cycle: virus-cell entry, translation, protein processing, RNA replication, virion assembly and virus release.

Several host proteins and drugs have been reported to have a direct affect on HCV replication in vitro [7]. These include factors that affect immune responses (interferons and their related genes [8, 9], RIG-I, TLRs [10]), cell proliferation (BMP7 [11], TGF-beta [12], nucleolin [13]), molecular chaperone function (cyclophilin [14], ER-stress proteins [15], FKBP [16], HSP27 [17], HSP90 [18]) and lipid metabolism (cholesterol, sphingolipid [19]). However, it is often difficult to determine whether these genes are changed by HCV replication or the changes are essential for HCV replication in the host cells.

In this study, we investigated the effects of host cellular gene expression using our HCV replicon system [20, 21]. We performed DNA microarray analyses using cells expressing the replicons, the corresponding *cured* cells, from which the replicon had been eliminated by prolonged treatment with interferon-alpha, and naïve Huh7 cells. Furthermore, we investigated the signaling pathways using DNA microarrays to study molecular pathways that are involved in the HCV life cycle and its pathogenesis.

Materials and methods

Cells and cell culture

Huh7 cells were maintained in Dulbecco's modified minimal essential medium (Sigma, St. Louis, MO) supplemented with 10% fetal calf serum at 37°C under 5% CO_2. To maintain cell lines carrying the HCV replicon (Huh7/Rep cells), G418 (Nakalai Tesque, Kyoto, Japan) was added to the culture medium to a final concentration of 500 μg/ml.

HCV replicon and cell culture

The HCV-1b replicon plasmid, pHCV1bneo-delS, was provided by Dr. Christoph Seeger (Fox Chase Cancer Center, Philadelphia, PA) [22]. HCV-2a replicon plasmid, pSGR-JFH1, was provided by Dr. Takaji Wakita (National Institute of Infectious Diseases, Tokyo, Japan). The neomycin phosphotransferase (Neo) gene of pHCV1bneo-delS and pSGR-JFH1 was replaced by a chimeric gene coding for yellow fluorescent protein fused in-frame with the foot-and-mouth disease virus peptide 2A (P2A) autocleavage motif followed by neomycin phosphotransferase, which we designated *Yeo* (Fig. 1) [23]. An HCV *Feo*-replicon that expresses chimeric firefly luciferase and the neomycin resistance gene has been described [20, 21]. In vitro replicon RNA synthesis, RNA transfection and selection of G418-resistant cell lines were carried out as described previously [21, 24]. Briefly, replicon RNAs were transfected into Huh7 cells. By cell culture in the presence of G418, we established Huh7 cell lines that stably express the *Yeo*-replicons: Huh7/Rep-1b-*Yeo* and Huh7/Rep-2a-*Yeo*.

Fluorescence microscopy and FACS analysis

The cells were plated onto eight-well chamber slides (Lab-Tek® Chamber Slide™ System, Nalgen Nunc International, Rochester, NY), and the YFP expression was detected by fluorescence microscopy (BZ-8000,

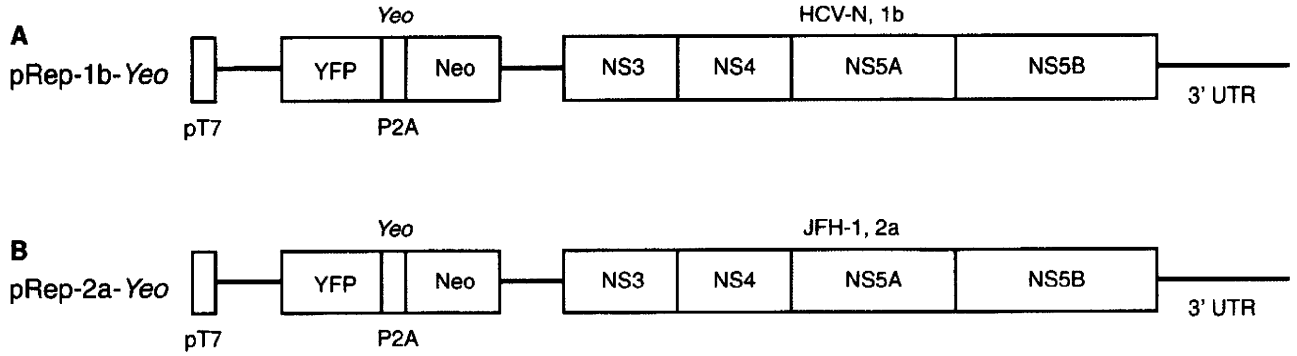

Fig. 1 Structure of replicon plasmid constructs. A hepatitis C virus (HCV) replicon plasmid, pRep-1b-Yeo (**a**) and pRep-2a-Yeo (**b**), was reconstructed from pHCV1bneo-delS [22] and pSGR-JFH1 [47] by replacing the neomycin phosphotransferase (Neo) gene with a fusion gene of yellow fluorecence protein (YFP) and Neo, which we designate "*Yeo*." *NS* Nonstructural region, *pT7* T7 promoter, *3'UTR* 3'untranslated region, *P2A* foot-and-mouth disease peptide 2A (see also "Materials and methods") [23]

KEYENCE, Osaka, Japan) and FACS Caliber using CellQuest software (BD Biosciences, Flanklin Lakes, NJ).

Cell sorting

Cells were treated for 5 min with trypsin/EDTA at 37°C and then resuspended in 10% FBS/DMEM. A single cell suspension was prepared by passage through a 35-μm nylon filter. The cell populations that support a high level of *Yeo*-replicon expression (Huh7/Rep-1b-*Yeo*^{high} and Huh7/Rep-2a-*Yeo*^{high}) were separated using a FACS Vantage SE cell-sorting system (BD Biosciences). The YFP-directed fluorescence of sorted cells was confirmed by fluorescence microscopy and FACS.

Establishment of the cured Huh7 cells

Cured Huh7 cells (cHuh7) were established by eliminating the HCV replicon from the *Yeo*-1b^{high} and -2a^{high} replicon expressing Huh7 cells by treatment with 100 U/ml of interferon-alpha for 14 days [6, 25]. Clearance of replicon RNA was confirmed by FACS analysis and by the loss of resistance to G418.

RNA preparation and microarray hybridization

Total cellular RNA was extracted from the 1b^{high} and 2a^{high} *Yeo*-replicon cells, *cured*-1b and -2a cells and naïve Huh7 cells using ISOGEN (Wako). Integrity of obtained RNA was assessed using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). All samples had an RNA Integrity Number (RIN) greater than 9.4 [26]. Complementary RNA was prepared from 1 μg total RNA, using one-cycle target labeling and a control reagents kit (Affymetrix, Santa Clara, CA). Hybridization and signal detection of the Human Genome U133 Plus 2.0 array (Affymetrix) were performed in accordance with the manufacturer's instructions. Assays were performed in duplicate.

Analysis of gene expression data

A total of ten microarray datasets was normalized using the robust multi-array average (RMA) method under R 2.8.1 statistical software (http://www.R-project.org). Estimated gene expression levels were log_2-transformed, data from 62 control probe sets were removed, and we selected 18,613 probe sets that were categorized as "present" or "marginal" among all samples. We performed two sets of gene comparisons to examine effects of HCV replicons on host cellular gene expression: one was the *high Yeo*-replicon-expressing cells versus *cured* cells and the other was parental Huh7 cells versus *cured* cells. We selected differently expressed genes using the significance analysis of microarray (SAM) as described by Tusher et al. [27], and the fold changes and the *Q*-values were calculated for each probe sets. We used delta = 0.1 as a cutoff parameter for SAM. A hierarchical clustering with selected genes was performed with R software. Euclidean distance was used to calculate the similarity matrix among genes or cell conditions, respectively. The complete linkage method was used for agglomeration.

Molecular pathway analysis and visualization of gene expression data

We used the KEGG Pathway database to investigate the molecular reactions and pathways that showed significant gene expression changes [28]. The KEGG Pathway is a database of biological systems, consisting of over 4,252 genes and 204 molecular pathway-wiring diagrams of interaction and reaction networks (http://www.genome.jp/kegg/pathway.html). Prior to the pathway analysis, we selected probe sets that were differentially expressed

between Huh7 and *cured* cells and between *cured* cells and replicon cells. For Huh7 versus *cured* cells analyses, we selected probe sets that showed 20% upregulation or downregulation (i.e., fold change of grater than 1.2) in both Huh7 versus *cured*-1b and Huh7 versus *cured*-2a cells. For replicon cell versus *cured* cell analyses, we selected probe sets that showed 20% upregulation or downregulation in both *cured*-1b versus replicon-1b cells or *cured*-2a versus replicon-2a cells. Association between the obtained gene list and each pathway was evaluated by Fisher's exact test. The significance level for KEGG analysis was set to a false discovery rate (FDR) of lower than 0.3 using the Benjamini and Hochberg method [29].

We next visualized functional associations between the differentially expressed genes and biological pathway processes. The KEGG Pathway provides a reference knowledge base for linking genomes to biological systems and also to environments by the processes of Pathway mapping and BRITE mapping. NCBI Entrez Gene IDs (EGIDs) for each gene in the pathways were extracted from the database. The relationship between probe sets on the microarray and EGIDs was obtained from a gene annotation file provided by Affymetrix. Thereafter, gene expression changes were mapped on the pathway by combining the results of fold-change analyses with the data sets above.

Real-time PCR analysis of mRNA expression

To confirm the results of the microarray analysis, we examined the expression levels of several mRNA by real-time RT-PCR (7500 Real Time PCR Systems, Applied Biosystems, Foster City, CA). Single-stranded cDNA was synthesized from total RNA using SuperScript II reverse transcriptase (Invitrogen) and random hexamers (Takara Bio Inc., Shiga, Japan) as primers. Expression of mRNA was quantified using QuantiTect SYBR Green PCR master Mix (QIAGEN, Valencia, CA). The primers used were as follows: HMGCR, SQLE, CYP51A1, TM7SF2, NSDHL, EBP and beta-actin. The nucleotide sequences of primers and corresponding product sizes are as indicated (see Supplementary Table 1).

Oil red O staining

Huh7 cells, replicon cells and *cured* cells were cultured on 18-mm-round micro cover glasses (Matsunami, Tokyo, Japan). These cells were fixed with 4% paraformaldehyde for 5 min at room temperature. After washing with PBS, the cells were permeabilized with 0.05% Triton X-100 in PBS for 5 min at room temperature. Staining of intracellular neutral lipids was performed with Oil red O, and nuclei were stained with Mayer's hematoxylin using Oil red O stain kit procedure (Diagnostic Biosystems Inc., Pleasanton, CA).

Immunofluorescence analysis

Huh7 cells, replicon cells and *cured* cells were cultured on 18-mm-round micro cover glasses. For immunostaining, the cells were fixed in 4% paraformaldehyde for 5 min at room temperature. For detection of HCV-NS5A, cells were incubated with the primary antibody (Biodesign International, Saco, ME) for 1 h at 37°C. The fluorescent secondary antibodies were Alexa Fluor 594 goat anti-mouse IgG antibody (Invitrogen, Carlsbad, CA). Nuclei were labeled with 4′,6-diamidino-2-phenylindole (DAPI). Lipid droplets were visualized with BODIPY 493/503 (Invitrogen). Analysis was performed on a Delta-Vision microscope system (Applied Precision, Seattle, WA).

Luciferase-based expression analysis of HCV replicon and analysis of cell viability

Huh7/Rep-Feo cells [20, 21] were cultured with various concentrations of peroxisome proliferator-activated receptor (PPAR)-alpha and -gamma agonists. After 48 h of culture, levels of HCV replication were quantified by internal luciferase assay using a Bright-Glo Luciferase Assay System (Promega). Assays were performed in triplicate, and the results were expressed as mean ± SD as percentages of the controls. To evaluate cell viability, dimethylthiazol carboxymethoxyphenyl sulfophenyl tetrazolium (MTS) assay was performed using a Cell Titer 96 Aqueous One Solution Cell Proliferation Assay (Promega) according to manufacturer's directions.

Statistical analyses

Statistical analyses were performed using the Student's *t*-test, and *P*-values of less than 0.05 were considered as statistically significant.

Results

Fluorescence detection of Yeo replicon

Genotypes 1b and 2a *Yeo*-replicon RNAs were stably transfected into Huh7 cells (Huh7/Rep-1b-Yeo and Huh7/Rep-2a-Yeo, respectively, Fig. 1). In these transfected cells, expression of the HCV replicon was visualized by HCV-IRES-driven, YFP-mediated fluorescence (Fig. 2, left panels). The expression levels of individual cells could be measured by fluorescence intensity and cytogram analysis using flow cytometry (Fig. 2, right panels).

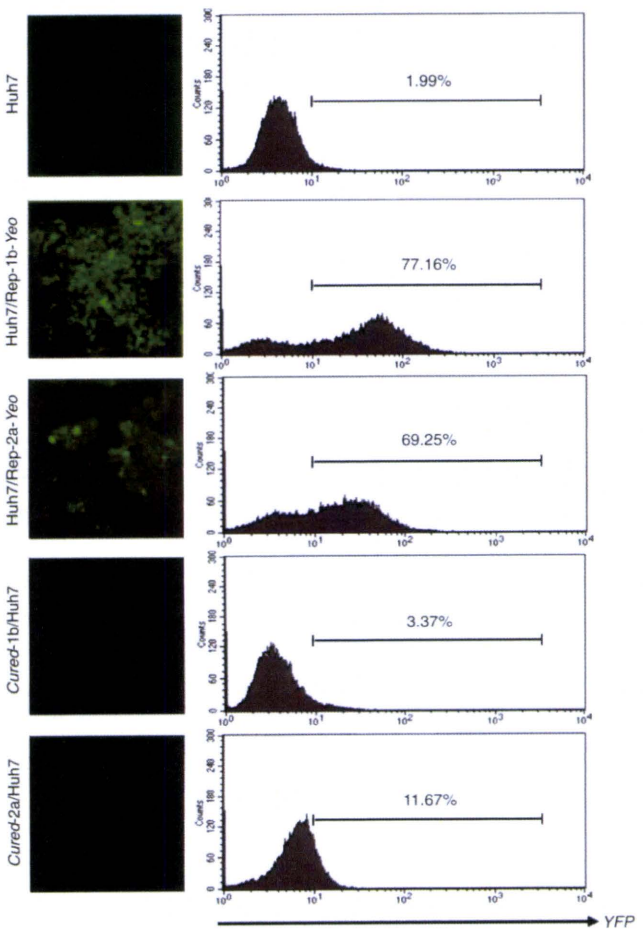

Fig. 2 Visualization of YFP replicon expression. Detection of YFP expression by fluorescence microscopy analysis and of intracellular YFP expression by FACS analysis

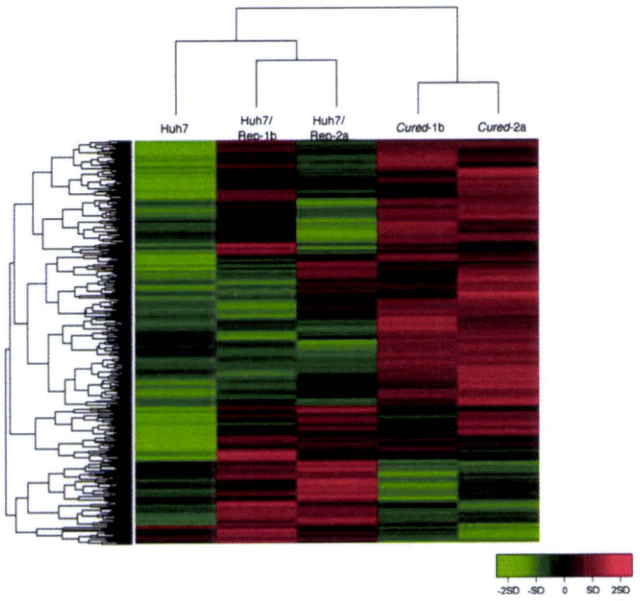

Fig. 3 Hierarchical clustering of gene expression profiles obtained from the 1b and 2a Yeo-replicon expressiong cells, cured cells and Huh7. The 1,870 probe sets that changed expression more than 1.2-fold in either Huh7 versus *cured* or *cured* versus replicon were selected. Dendrograms show the classification determined by hierarchical clustering analysis. *Red* and *green colors* indicate relative overexpression and underexpression, respectively

Our initial trial was to compare gene expression profiles between replicon-expressing cells and parental Huh7 cells. However, these comparisons identified gene expressional changes induced by HCV infection and by adaptation of the host cell to support efficient HCV replication, because transfection of replicon RNA and G418-treatment of cells resulted in selection of a cell population that can support a high level of HCV subgenomic replication. Therefore, we used *cured* cell lines, which were established from Huh7/Rep-1b-Yeo and -2a-Yeo by interferon-alpha treatment. These cured cell lines are highly permissive for HCV replication on re-introduction of virus or replicon RNA (Fig. 2). With these backgrounds, we performed two sets of gene comparison using microarray analyses: comparison of replicon-expressing cell lines (Huh7/Rep-1b-Yeo and Huh7/Rep-2a-Yeo) and *cured* cells (Cured-1b/Huh7 and Cured-2a/Huh7) was intended to identify genes that are affected by HCV replication, and comparison of parental Huh7 cells and *cured* cell lines was intended to identify genes that are essential for a high level of HCV replication in cultured cells.

Hierarchical clustering gene expression profiles in naïve, replicon-expressing and cured cells

Datasets from the microarrays were normalized using the robust multi-array average (RMA) method, and differentially expressed genes were extracted in replicon cells, *cured* cells and naïve Huh7 cells. Gene expression profiles were well correlated each other between duplicate microarray data from the same cell line with the Pearson's correlation coefficients (R^2) of greater than 0.975 (see Supplementary Fig. 1). In this analysis, 1,870 probe sets showed differences in expression levels of more than 1.2-fold under delta <0.1 in either Huh7 versus *cured* (1,516 probe sets) or *cured* versus replicon (372 probe sets). A hierarchical clustering analysis showed that the gene-expression profiles of each cell group constituted clear clusters, Huh7/Rep-2a and Huh7/Rep-1b, Cured-2a and Cured-1b, and Huh7 cells (Fig. 3). Among genes whose expression differed significantly between replicon-expression cells and *cured* cells, 15 showed changes of more than two-fold (Table 1). These included cell cycle- or cell growth-related genes (nuclear protein 1, growth duf-ferentiation factor 15, urothelial cancer associated 1, inhibin and tubulin), oncogene (Ras-related GTP binding D) and interferon-related gene (IFIM3). On the other hand, 37

Table 1 Microarray analysis: genes for which the expression changed more than two-fold in Rep-1b/Huh7 and Rep-2a/Huh7 cells compared to their *cured* cells

Probe set	Title	Rep1b/cured1b		Rep2a/cured2a	
		Fold change	*Q*-value	Fold change	*Q*-value
209230_s_at	Nuclear protein 1	5.41	0.00	13.78	7.44
221523_s_at	Ras-related GTP binding D	3.82	0.00	2.32	24.60
221524_s_at	Ras-related GTP binding D	3.17	0.00	2.40	17.08
205923_at	Reelin	3.14	0.00	2.70	12.42
200924_s_at	Hypothetical protein LOC442497/solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	2.84	0.00	2.63	0.69
221577_x_at	Growth differentiation factor 15	2.62	0.00	3.16	0.00
233030_at	Patatin-like phospholipase domain containing 3	2.07	0.00	2.31	3.16
201471_s_at	Sequestosome 1	2.52	0.81	2.10	2.45
227919_at	Urothelial cancer associated 1	5.09	0.90	2.90	54.18
207076_s_at	Argininosuccinate synthetase 1	2.99	0.90	2.22	61.28
205749_at	Cytochrome P450, family 1, subfamily A, polypeptide 1	2.69	0.90	2.63	4.55
217127_at	Cystathionase (cystathionine gamma-lyase)	2.32	3.05	2.38	6.54
210587_at	Inhibin, beta E	2.51	4.25	3.89	2.45
214023_x_at	Tubulin, beta 2B	2.41	4.25	4.31	54.18
212203_x_at	Interferon induced transmembrane protein 3 (1-8U)	2.34	5.75	2.09	54.18

genes were up-regulated by more than two-fold between *cured* and naïve cells (Table 2), which included genes such as chemokine (CCL14), solute carrier family and metallothionein family.

Pathway process analyses and hierarchical clustering of genes in each functional category

Using the KEGG Pathway database, we analyzed pathway processes that were altered between replicon-expressing cells and *cured* cells as well as between *cured* cells and Huh7 cells (Supplementary Tables 2, 3). Comparison of the pathway processes between replicon-expressing and *cured* cells identified six pathways that showed differences of FDR < 0.3, including pathways related to MAPK ($P = 4.0 \times 10^{-4}$, FDR 0.08), biosynthesis of steroids ($P = 4.21 \times 10^{-3}$, FDR 0.21) and TGF-beta ($P = 8.4 \times 10^{-3}$, FDR 0.29) (KEGG Pathway maps for each significant pathways are shown in Supplementary Fig. 2A–F). Comparison of the pathway processes between *cured* and naïve Huh7 cells identified 11 significant pathways (KEGG Pathway maps for each significant pathway are shown in Fig. 5 and Supplementary Fig. 3A–J). These included pathways that were related to TGF-bata ($P = 8.42 \times 10^{-3}$), cell cycle ($P = 9.0 \times 10^{-3}$) and sphingolipid metabolism ($P = 1.32 \times 10^{-2}$). Interestingly, there were significant changes in the biosynthesis of steroids ($P = 1.75 \times 10^{-4}$) between cured and naïve Huh7 cells. These results suggested that several lipid metabolism processes were substantially associated with efficient HCV replication in host cells.

Hierarchical clustering analyses of representative genes included in functional pathway categories

Based on pathway process analyses using the KEGG database, we performed hierarchical clustering analyses of each functional subset of genes (fold change >1.2, Fig. 4a–c). The cell cycle, cholesterol biosynthesis and sphingolipid metabolism-related genes demonstrated clear clusters in replicon cells, *cured* cell and parental Huh7, respectively. In particular, cholesterol biosynthesis-related genes were activated in replicon cells and *cured* cells.

Mapping between pathway information and gene expression data

Knowing that cholesterol metabolism pathway was changed substantially in *cured* cells, we performed graphical mapping of the related genes to the KEGG Pathway map database (Fig. 5). Similar to the pathway analyses, cholesterol biosynthesis related genes, which are involved in the mevalonate pathway or sterol biosynthesis, were clearly activated in *cured* cells compared to naïve Huh7 (Fig. 5).

To verify the microarray results, we performed real-time RT-PCR of cholesterol biosynthesis-related genes including HMGCR, SQLE, NSDHL, CYP51A1, TM7SF2 and EBP. All the genes were upregulated in replicon-expressing and *cured* cells compared to the naïve Huh7 cells (Fig. 6). These results were consistent with the microarray data.

Table 2 Microarray analysis: genes for which the expression changed more than two-fold in *cured*-1b and *cured*-2a cells compared to Huh7

Probe set	Title	Cured1b/Huh7		Cured2a/Huh7	
		Fold change	*Q*-value	Fold change	*Q*-value
210390_s_at	Chemokine (C–C motif) ligand 14/15	4.49	0.00	2.56	0.00
221168_at	PR domain containing 13	2.38	0.00	2.15	0.00
1553995_a_at	5′-nucleotidase, ecto (CD73)	2.15	0.00	3.64	0.36
204897_at	Prostaglandin E receptor 4 (subtype EP4)	2.07	2.40	2.19	0.36
214522_x_at	Histone cluster 1, H2ad/H3d	3.01	4.12	2.98	0.46
214472_at	Histone cluster 1, H2ad/H3a-j	3.36	5.08	3.53	0.46
218280_x_at	Histone cluster 2, H2aa3/H2aa4	4.65	5.20	5.17	0.61
232035_at	Histone cluster 1, H4a-f, H4 h-l/histone cluster 2, H4a-b/histone cluster 4, H4	5.56	6.51	5.37	0.95
214455_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	4.72	6.51	2.68	0.61
202708_s_at	Histone cluster 2, H2be	4.48	6.51	5.31	0.95
214290_s_at	Histone cluster 2, H2aa3/H2aa4	4.06	6.51	2.98	2.08
209398_at	Histone cluster 1, H1c	3.40	6.51	3.60	2.08
230795_at	–	2.91	6.51	4.51	0.95
1553994_at	5′-nucleotidase, ecto (CD73)	2.29	6.51	2.31	3.34
208180_s_at	Histone cluster 1, H4a-f, H4 h-l/histone cluster 2, H4a, H4b/histone cluster 4, H4	6.27	7.65	3.40	1.46
215779_s_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	3.53	7.65	5.86	1.46
206110_at	–	3.82	9.12	10.68	0.00
206535_at	Solute carrier family 2 (facilitated glucose transporter), member 2	3.58	9.12	2.84	5.02
213880_at	Leucine-rich repeat-containing G protein-coupled receptor 5	2.81	9.12	3.02	5.02
210387_at	Histone cluster 1, H2bc, H2be, H2bf, H2bg, H2bi	3.24	10.68	2.28	11.66
203044_at	Chondroitin sulfate synthase 1	2.36	11.85	2.67	2.08
207102_at	Aldo-keto reductase family 1, member D1 (delta 4-3-ketosteroid-5-beta-reductase)	2.14	13.35	5.58	0.36
219596_at	THAP domain containing 10	2.20	16.73	3.32	2.08
217997_at	Pleckstrin homology-like domain, family A, member 1	2.15	23.48	3.01	3.34
217996_at	Pleckstrin homology-like domain, family A, member 1	2.12	23.48	2.21	11.66
217165_x_at	Metallothionein 1F	3.36	29.61	3.25	11.66
213629_x_at	Metallothionein 1F	3.13	29.61	3.70	11.66
210524_x_at	–	2.39	29.61	2.46	17.81
206143_at	Solute carrier family 26, member 3	2.35	29.61	2.49	17.81
212859_x_at	Metallothionein 1E	3.46	35.93	4.07	11.66
208581_x_at	Metallothionein 1X	3.31	35.93	3.52	17.81
204326_x_at	Metallothionein 1X	3.24	35.93	3.49	17.81
211456_x_at	Metallothionein 1 pseudogene 2	3.19	35.93	3.71	17.81
206461_x_at	Metallothionein 1H	3.09	35.93	3.49	17.81
216336_x_at	Metallothionein 1E, 1H, 1 M/metallothionein 1 pseudogene 2	3.02	35.93	3.27	17.81
212185_x_at	Metallothionein 2A	2.92	35.93	2.85	17.81
204745_x_at	Metallothionein 1G	2.73	35.93	3.31	17.81

Detection of intracellular lipid droplets in naïve, replicon-expressing and cured cells

Because several lipid-related pathways were extracted (Supplementary Tables 2 and 3), we examined phenotypes of the cell lines featuring different lipid metabolism gene expression profiles by carrying out detection of cellular lipid droplets (Fig. 8a, b). The cells were stained by Oil red O or BODIPY493/503, dye solutions specific for neutral lipids. We found a large number of lipid droplets in the cytoplasm of each Huh7 cell line. The number of lipid droplets obviously was increased more in

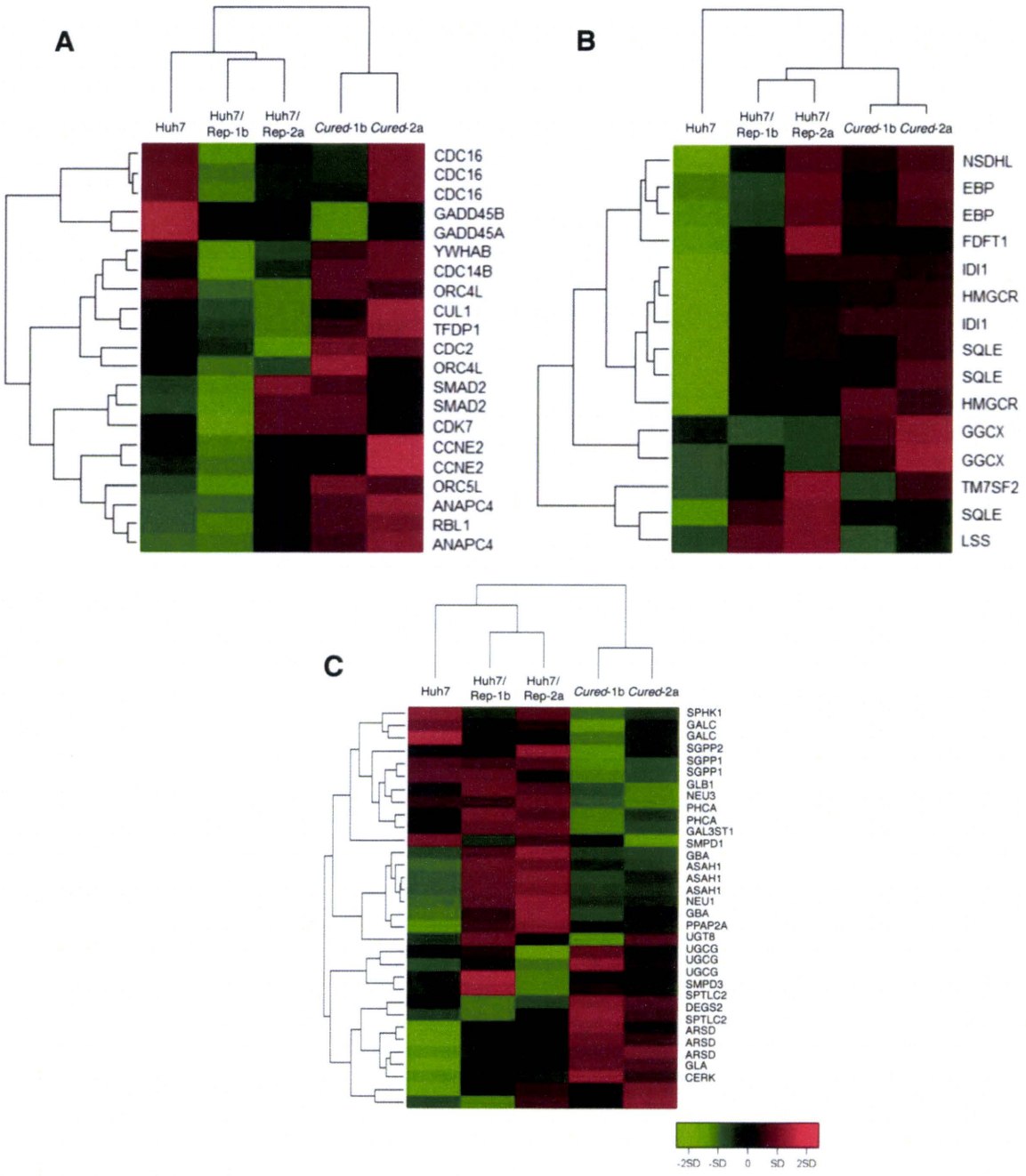

Fig. 4 Hierarchical clustering of representative genes included in each KEGG Pathway map. **a** Cell cycle, **b** cholesterol biosynthesis, **c** sphingolipid metabolism. Dendrograms shows the classification determined by hierarchical clustering analysis. *Red* and *green colors* indicate relative overexpression and underexpression, respectively

the two replicon-expressing cells and the *cured* cells than in the parental Huh7 cells. Lipid and HCV-NS5A double staining showed an increase in lipid droplets in cells that expressed HCV proteins (Fig. 8b). Analyses of the KEGG fatty acid metabolism pathway showed that a substantial number of the genes of these pathways were up-regulated in the *cured* cells compared to the naïve cells, although these could not reach statistical significance (Fig. 7).

Effects of hepatitis C virus replication by PPAR-alpha and gamma agonists

To assess the effects of lipid metabolic status on the intracellular replication of the HCV genome, Huh7/Rep-Feo cells were cultured with various concentrations of several PPAR-alpha agonists (clofibrate, fenofibrate and bezafibrate) and gamma agonists (pioglitazone and troglitazone) (Fig. 9). The luciferase activities of the Huh7/