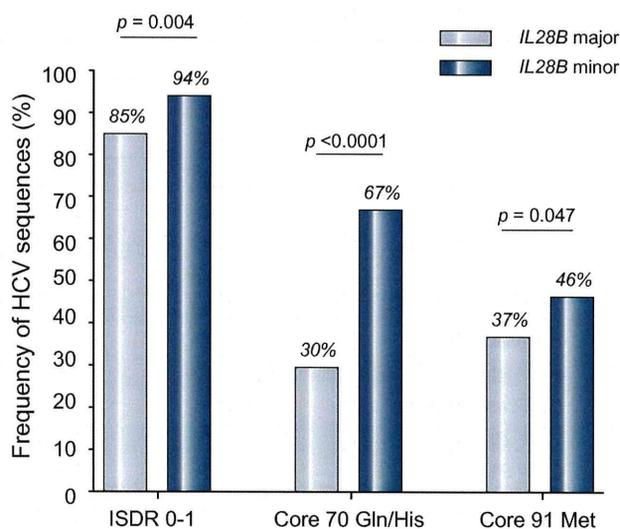Fig. 3. **Association between relapse and the *IL28B* allele or mutations in the ISDR.** The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele (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.]

portion of patients with advanced fibrosis (F3-4) was 39% (84/217) in patients with low platelet counts ($<140 \times 10^9$/L) compared to 13% (37/279) in those with high platelet counts ($\geqslant 140 \times 10^9$/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 nitazoxianide [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.

12%). After stratification by the *IL28B* allele, patients with low platelet counts ($<140 \times 10^9$/L) had a lower probability of SVR and higher probability of NVR than those with high platelet counts ($\geqslant 140 \times 10^9$/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 pro-

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.

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

Research Article
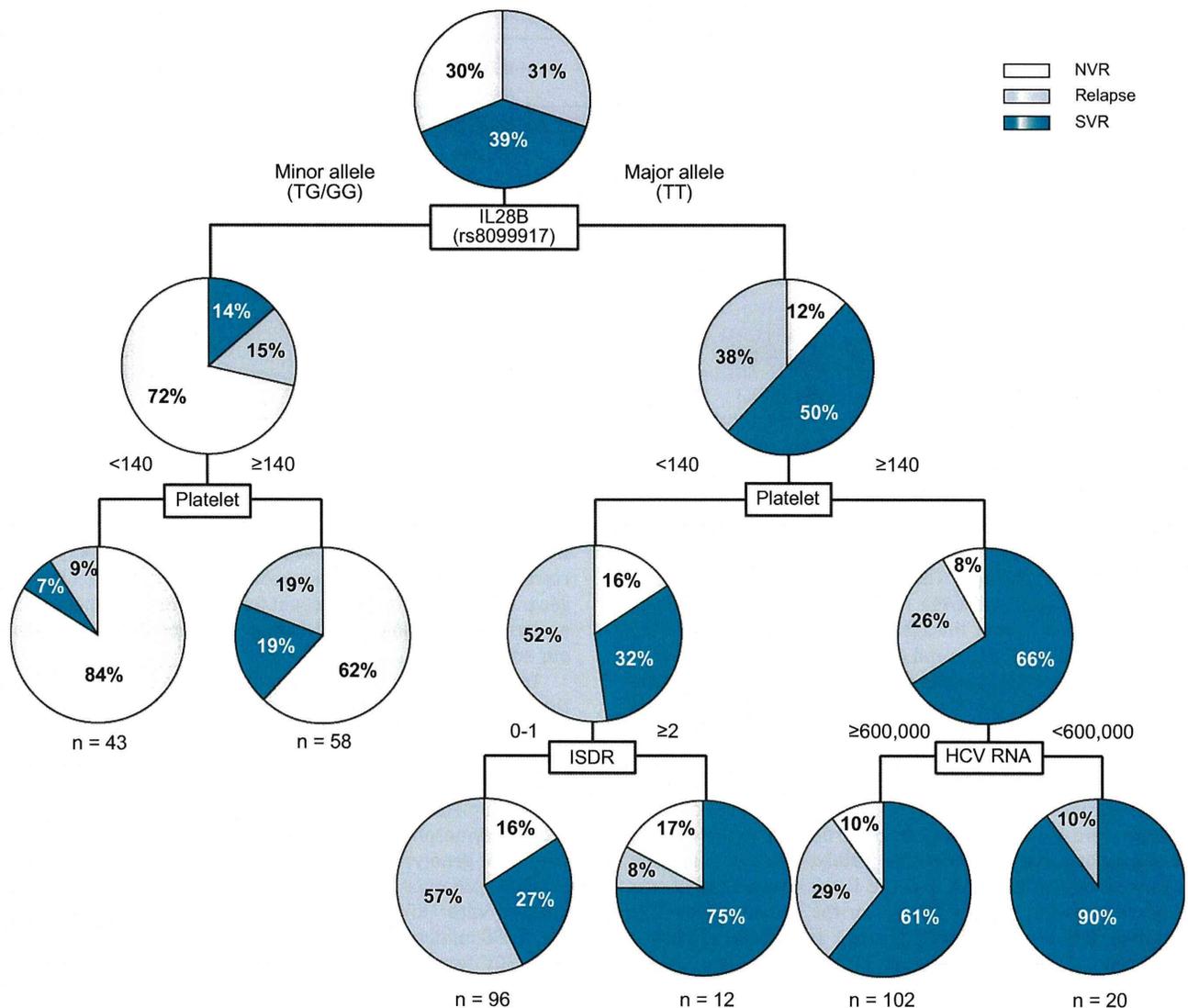
Fig. 5. **Decision tree for the prediction of response to therapy**. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown. [This 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

A

B
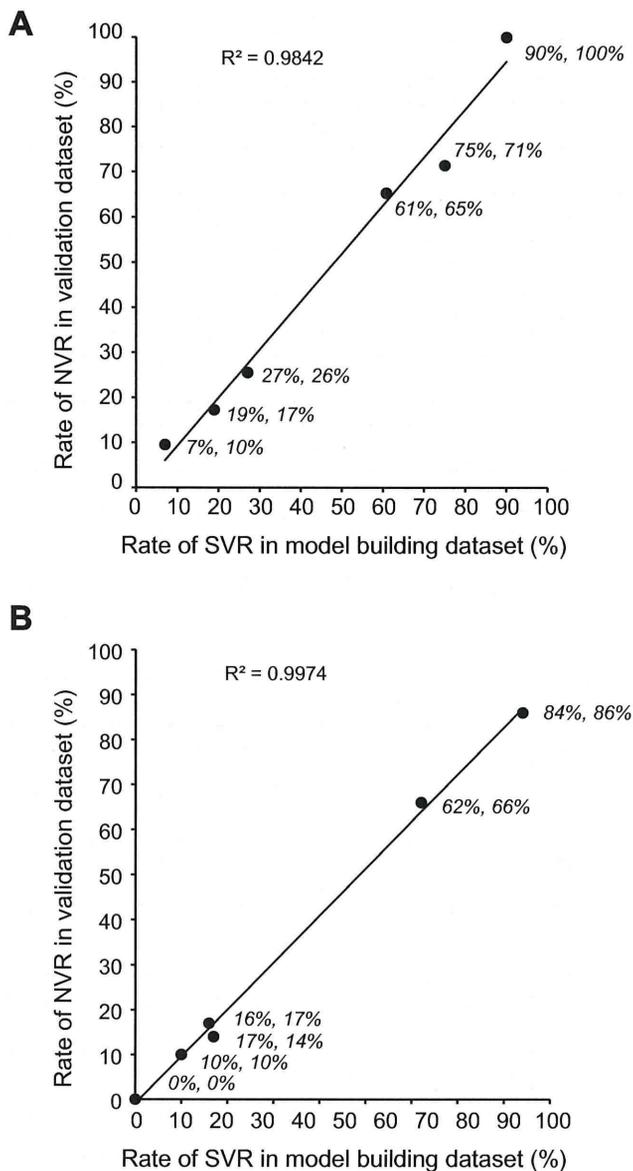
Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and 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

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.

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.

Genome-wide association study identified *ITPA/DDRGK1* variants reflecting thrombocytopenia in pegylated interferon and ribavirin therapy for chronic hepatitis C

Yasuhito Tanaka[1,2,†], Masayuki Kurosaki[3,†], Nao Nishida[4], Masaya Sugiyama[1,2,5], Kentaro Matsuura[1,6], Naoya Sakamoto[7], Nobuyuki Enomoto[8], Hiroshi Yatsuhashi[9], Shuhei Nishiguchi[10], Keisuke Hino[11], Shuhei Hige[12], Yoshito Itoh[13], Eiji Tanaka[14], Satoshi Mochida[15], Masao Honda[16], Yoichi Hiasa[17], Asako Koike[18], Fuminaka Sugauchi[1,6], Shuichi Kaneko[16], Namiki Izumi[3], Katsushi Tokunaga[4] and Masashi Mizokami[5,*]

[1]Department of Virology and [2]Liver Unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan, [3]Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan, [4]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [5]The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Chiba, Japan, [6]Department of Gastroenterology and Metabolism, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan, [7]Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan, [8]First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan, [9]Clinical Research Center, National Nagasaki Medical Center, Nagasaki, Japan, [10]Department of Internal Medicine, Hyogo College of Medicine, Nishinomiya, Japan, [11]Division of Hepatology and Pancreatology, Kawasaki Medical College, Kurashiki, Japan, [12]Department of Internal Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan, [13]Molecular Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kyoto, Japan, [14]Department of Medicine, Shinshu University School of Medicine, Matsumoto, Japan, [15]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Saitama Medical University, Saitama, Japan, [16]Department of Gastroenterology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan, [17]Department of Gastroenterology and Metabology, Ehime University Graduate School of Medicine, Ehime, Japan, and [18]Central Research Laboratory, Hitachi Ltd, Kokubunji, Japan

---

Hematologic abnormalities during current therapy with pegylated interferon and ribavirin (PEG-IFN/RBV) for chronic hepatitis C (CHC) often necessitate dose reduction and premature withdrawal from therapy. The aim of this study was to identify host factors associated with IFN-induced thrombocytopenia by genome-wide association study (GWAS). In the GWAS stage using 900K single-nucleotide polymorphism (SNP) microarrays, 303 Japanese CHC patients treated with PEG-IFN/RBV therapy were genotyped. One SNP (rs11697186) located on *DDRGK1* gene on chromosome 20 showed strong associations in the minor-allele-dominant model with the decrease of platelet counts in response to PEG-IFN/RBV therapy [$P = 8.17 \times 10^{-9}$; odds ratio (OR) = 4.6]. These associations were replicated in another sample set ($n = 391$) and the combined P-values reached 5.29×10^{-17} (OR = 4.5). Fine mapping with 22 SNPs around *DDRGK1* and *ITPA* genes showed that rs11697186 at the GWAS stage had a strong linkage disequilibrium with rs1127354, known as a functional variant in the *ITPA* gene. The

---

*To whom correspondence should be addressed at: The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, 1-7-1 Kohnodai, Ichikawa, Chiba 272-8516, Japan. Tel: +81 473723501; Fax: +81 473754766; Email: mmizokami@hospk.ncgm.go.jp
†These authors contributed equally.

ITPA-AA/CA genotype was independently associated with a higher degree of reduction in platelet counts at week 4 ($P < 0.0001$), as well as protection against the reduction in hemoglobin, whereas the CC genotype had significantly less reduction in the mean platelet counts compared with the AA/CA genotype ($P < 0.0001$ for weeks 2, 4, 8, 12), due to a reactive increase of the platelet count through weeks 1–4. Our present results may provide a valuable pharmacogenetic diagnostic tool for tailoring PEG-IFN/RBV dosing to minimize drug-induced adverse events.

INTRODUCTION

Chronic infection with hepatitis C virus (HCV) presents a significant health problem worldwide, with ∼2.3% of the world population, i.e. more than 120–130 million people, being infected (1). Only 20–30% of HCV-infected individuals recover spontaneously. The remaining 70–80% go on to develop chronic infection, being at significant risk for progressive liver fibrosis and subsequent liver cirrhosis (LC) and hepatocellular carcinomas (HCC). Successful treatment of chronic hepatitis C (CHC) leads to a reduction of liver fibrosis stage of patients, and also prevents HCC development (2).

Antiviral treatment has been shown to improve liver histology and decrease incidence of hepatocellular carcinoma in CHC (3,4). Current therapy for CHC consists of treatment with pegylated interferon (IFN), which acts both as an antiviral and as an immunoregulatory cytokine, and ribavirin (RBV), an antiviral pro-drug that interferes with RNA metabolism (5,6). However, <50% of patients infected with HCV genotype 1 treated in this way achieve a sustained viral response (SVR) or cure of the infection (5,7). Older patients with liver fibrosis showed a significantly lower SVR rate due to poor adherence resulting from adverse events and laboratory abnormalities (8–10). In particular, hematologic abnormalities often necessitate dose reduction, and premature withdrawal from therapy in 10–14% of patients (5,11–14). New drugs and therapeutic approaches for CHC are actively developed and several candidates are in early trial phase (15,16). Given this background, effective pre-treatment screening for predictive biomarkers with the aim of evaluating possible risks over benefits of currently available treatment will avoid these side effects in patients who will not be helped by treatment, as well as reduce the substantial cost of treatment.

The completion of the Human Genome Project has led to the advent of a new era of scientific research, including a revolutionary approach: the genome-wide association study (GWAS). Several recent studies, including our study, have demonstrated marked associations between single-nucleotide polymorphisms (SNPs) within and around *IL28B* gene, which codes for IFN−λ3 (16–21). Another recent study indicated that genetic variants of *ITPA* gene leading to inosine triphosphatase (ITPA) deficiency could protect against hemolytic anemia (HA) in CHC patients receiving RBV (22).

In Japan, HCV-infected patients are relatively old and some of them have had severe fibrosis (9). Thrombocytopenia is one of the critical adverse events by IFN-based therapy among liver cirrhotic patients (23), because low platelet count (PLT), i.e. <30.0 (10^9/l), would be a risk factor for any bleeding, as well as it would lead to poor treatment efficiency due to the initial or early dose reduction of PEG-IFN. Based on its pathogenesis, drug-induced thrombocytopenia is usually due to bone marrow

suppression, immune-mediated destruction and platelet aggregation (24). In this study, we firstly found that genetic variants in the *ITPA/DDRGK1* genes were associated with IFN-induced thrombocytopenia, and then examined the correlation between IFN-induced thrombocytopenia and RBV-induced HA in Japanese CHC patients under PEG-IFN/RBV treatment.

RESULTS

Genetic variants associated with IFN-induced thrombocytopenia

In this study, we conducted a GWAS to identify host genes associated with the decrease of platelets in response to PEG-IFN/RBV treatment in 303 Japanese HCV patients (107 patients with the decrease of PLT versus 196 patients without the decrease of PLT based on the criteria described in Materials and Methods), using a genome-wide SNP typing array (Affymetrix SNP 6.0 for 900K SNPs). The characteristics of patients for each GWAS stage and replication stage are summarized in Table 1. Figure 1 shows a genome-wide view of the single-point association data based on allele frequencies. One SNP (rs11697186) located on *DDRGK1* gene on chromosome 20 showed strong associations in the allele frequency model ($P = 8.17 \times 10^{-9}$) with the decrease of PLT in response to PEG-IFN plus RBV treatment. The association reached genome-wide level of significance [Bonferroni criterion $P < 8.40 \times 10^{-8}$ (0.05/595052)], and another SNP (rs6139030) near *ITPA* gene had a marginal significance ($P = 4.30 \times 10^{-7}$, in Table 2).

To validate the results of the GWAS stage, 22 SNPs were selected for the replication in a set of 391 Japanese HCV patients with and without platelet reduction (Supplementary Material, Table S1). The associations of the original significant SNP (rs11697186) and the marginal SNP (rs6139030) at the GWAS stage were replicated in the second set of 391 patients in the minor-allele-dominant model [$P = 5.88 \times 10^{-10}$, odds ratio (OR) = 4.6 for rs11697186; $P = 3.83 \times 10^{-10}$, OR = 4.3 for rs6139030, Table 2]. The combined *P*-values for both stages reached 5.29×10^{-17} (OR = 4.5; 95% CI = 3.1–6.5) and 1.33×10^{-15} (OR = 3.9; 95% CI = 2.8–5.5), respectively (Table 2).

Genetic variants associated with RBV-induced anemia

We also conducted a GWAS to identify host genes associated with a quantitative change in hemoglobin (Hb) levels from baseline to week 4 of PEG-IFN/RBV treatment in the above 303 Japanese HCV patients (94 patients with an Hb reduction of ≥3 g/dl at week 4 and 209 patients without Hb reduction), using a genome-wide SNP typing array (Affymetrix SNP 6.0 for 900K SNPs). Two SNPs (rs11697186 and rs6139030)

Table 1. Clinical characteristics of patients in this study

	GWAS ($n = 303$)	Replication ($n = 391$)
Age	57.4 (9.7)	56.8 (9.9)
Sex (M/F)	151/152	209/182
Weight (kg)	60.6 (10.4)	61.3 (10.7)
Body mass index	23.5 (3.1)	23.7 (4.1)
Baseline Hb (g/dl)	14.1 (1.4)	14.1 (1.4)
Baseline platelet count (10^9/l)	151.3 (54.3)	159.7 (55.0)
Baseline ALT (IU/l)	83.5 (79.4)	86.8 (71.9)
Baseline creatinine (mg/dl)	0.70 (0.15)	0.72 (0.16)
Baseline liver fibrosis (F0–2/F3–4/ND)	153/77/73	175/59/43
rs8099917: TT/non-TT	165/138	296/95
rs1127354: AA/CA/CC	4/79/220	6/101/284
Week 4 Hb (g/dl)	11.8 (1.7)	11.9 (1.5)
Week 4 platelet count (10^9/l)	127.6 (48.2)	132.4 (51.0)
Hb reduction at week 4	−2.3 (1.4)	−2.2 (1.4)
Platelet reduction at week 4	−22.2 (38.4)	−24.7 (30.4)

located on *DDRGK1* gene and *ITPA* gene on chromosome 20 showed strong associations in the allele frequency model ($P = 3.29 \times 10^{-10}$ and $P = 2.56 \times 10^{-9}$) with Hb reduction in response to PEG-IFN plus RBV treatment (Table 3).

The above 22 SNPs were selected for the replication study and fine mapping, including rs1127354, which was reported by the US group (22) to be strongly associated with Hb reduction (Supplementary Material, Table S2). All SNPs were genotyped using the DigiTag2 assay in an independent set of 391 Japanese HCV patients with quantitative change in Hb in response to PEG-IFN/RBV treatment [137 patients with Hb reduction versus 254 patients without Hb reduction (Table 3)]. The associations of the original SNPs were replicated in the second set of 391 patients in the minor-allele-dominant model ($P = 3.86 \times 10^{-16}$, OR = 0.02 for rs11697186; $P = 6.90 \times 10^{-18}$, OR = 0.03 for rs6139030, Table 3). The combined *P*-values for both stages reached 9.43×10^{-25} (OR = 0.03; 95% CI = 0.01–0.08) and 2.12×10^{-25} (OR = 0.04; 95% CI = 0.02–0.09), respectively (Table 3). The rs1127354 was also strongly associated with a quantitative change in Hb in response to PEG-IFN/RBV treatment in a set of 694 Japanese HCV patients (303 patients from the GWAS stage plus the second set of 391 patients) with and without Hb reduction ($P = 4.58 \times 10^{-26}$, OR = 0.03; 95% CI = 0.01–0.08).

Fine mapping with 22 SNPs around *DDRGK1* and *ITPA* genes showed that four significant SNPs (rs11697186, rs6139030, rs1127354 and rs13830) at the GWAS stage had a strong linkage disequilibrium (LD) ($r^2 > 0.86$) within the 22.7 kb region (Fig. 2). As the rs1127354 is known as a functional variant in the *ITPA* gene that caused ITPase deficiency and protected against RBV-induced HA (22,25), the representative SNP was applied for the following detailed studies.

### *ITPA/DDRGK1* variants reflect anemia and reactive increase of the platelet count

The mean quantitative reduction of blood cells from the baseline according to the *ITPA* rs1127354 genotypes is shown in Figure 3. Patients with the rs1127354 genotypes AA and CA showed lower degree of Hb reduction at weeks 2, 4, 8 and

12 during therapy compared with those with the CC genotype ($P < 0.0001$ for weeks 2, 4, 8 and 12 in Fig. 3A). The most difference of mean Hb reduction was found at week 4 (AA/CA −1.14 versus CC −2.72). These results show that the AA and CA genotypes are significantly associated with less absolute reduction in Hb levels, especially during the early weeks of therapy, and protect against the development of severe anemia. Interestingly, the CC genotype had significantly less reduction in the mean platelet count compared with the AA/CA genotype ($P < 0.0001$ for weeks 2, 4, 8; $P = 0.019$ for week 12 in Fig. 3B), due to a reactive increase of platelet count through weeks 1–4. The most difference of mean platelet reduction was found at week 4 [AA/CA −41.2 versus CC −18.0 (10^9/l)]. There was no difference in the neutrophil leukocyte count between genotypes (Fig. 3C). We then compared the percentage of patients with platelet count reduction in the *ITPA* rs1127354 genotypes at week 4 of PEG-IFN/RBV therapy (Fig. 4). The percentage of patients with a platelet count reduction of <30 (10^9/l) at week 4 was significantly higher in the rs1127354 genotypes CC ($P < 0.0001$), indicating that the degree of platelet count reduction was less in patients with the rs1127354 genotype CC. A multivariate analysis for factors associated with a platelet reduction >30 (10^9/l) at week 4 showed that lower platelet count at the baseline and the rs1127354 genotypes AA/CA were independently associated with platelet reduction (OR = 1.15; 95% CI = 1.11–1.20; $P < 0.0001$, OR = 5.92; 95% CI = 3.82–9.17; $P < 0.0001$, respectively).

Figure 5 showed reactive increase of the platelet count through weeks 1–4 of PEG-IFN/RBV therapy. Patients with anemia (Hb reduction ≥3.0 g/dl) at week 4 had a significantly higher degree of the reactive increase of the platelet count than those without anemia ($P < 0.0001$ in Fig. 5A). Within a subgroup of patients with the rs1127354 genotypes CC, patients with anemia still had a significantly higher degree of reactive increase of the platelet count than those without anemia ($P = 0.004$ in Fig. 5B). On the other hand, patients with the rs1127354 genotypes CC had a significantly higher degree of the reactive increase of the platelet count than those with genotypes AA/CA ($P < 0.0001$ in Fig. 5C), and a similar result was obtained in a subgroup of patients without anemia (Fig. 5D). To elucidate the significant factors associated with the rs1127354 genotypes by multivariate analysis, the rs1127354 genotypes AA/CA were independently associated with protection against the reduction in Hb and more reduction in platelet counts at week 4 due to a lower degree of the reactive increase of the platelet count (OR = 0.029; 95% CI = 0.009–0.092; $P < 0.0001$, OR = 4.73; 95% CI = 3.04–7.37; $P < 0.0001$, respectively). Indeed, the reactive increase of the platelet count through weeks 1–4 was positively correlated with a high platelet count at the baseline and anemia (Hb reduction ≥3.0 g/dl) at week 4, but was negatively correlated with rs1127354 genotypes AA/CA and a platelet count reduction of ≥30 (10^9/l) at week 4 (Table 4).

Relationship between *ITPA rs1127354* genotypes and treatment outcome due to dose reduction of PEG-IFN or RBV

In this population, a multivariate analysis showed that SVR was significantly associated with *IL28B* TT-genotype [OR
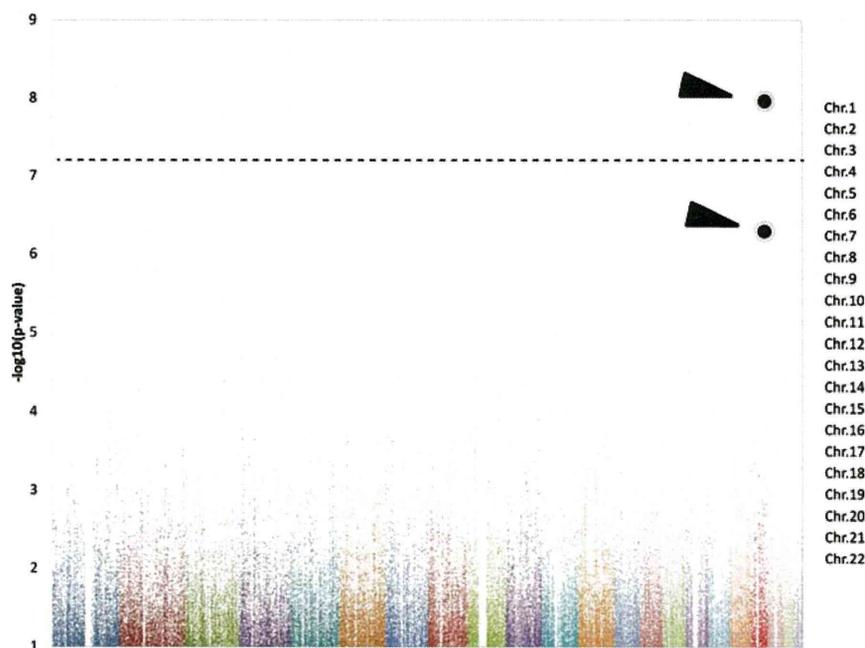
Figure 1. Genome-wide association results in 303 Japanese HCV patients with the decrease of platelets in response to PEG-IFN plus RBV treatment (107 patients with the decrease of PLT and 196 patients without the decrease of PLT). *P*-values were calculated using the χ^2 test for allele frequencies. Dots with arrow on chromosome 20 showed a significant SNP ($P = 8.17 \times 10^{-9}$ for rs11697186) and a candidate SNP with a marginal significance ($P = 4.30 \times 10^{-7}$ for rs6139030) associated with the decrease of PLT with response to PEG-IFN/RBV treatment. The dotted line indicates a genome-wide significance ($P < 8.40 \times 10^{-8}$).

Table 2. Two SNPs (rs11697186 and rs6139030) significantly associated with the decrease of PLT in response to PEG-IFN/RBV treatment

dbSNP rsID	Nearest gene	MAF[a] (allele)	Allele (1/2)	Stage	Patients with the decrease of PLT			Patients without the decrease of PLT			OR (95% CI)[b]	*P*-value[c]
					11	12	22	11	12	22		
rs11697186	*DDRGK1*	0.15 (T)	T/A	GWAS	3 (2.8)	48 (44.9)	56 (52.3)	0 (0.0)	32 (16.6)	161 (83.4)	4.6 (2.7–7.8)	8.17×10^{-9}
				Replication	3 (1.8)	65 (39.9)	95 (58.3)	3 (1.4)	25 (12.0)	181 (86.6)	4.6 (2.8–7.7)	5.88×10^{-10}
				Combined	6 (2.2)	113 (41.9)	151 (55.9)	3 (0.7)	57 (14.2)	342 (85.1)	4.5 (3.1–6.5)	5.29×10^{-17}
rs6139030	*ITPA*	0.17 (C)	T/C	GWAS	56 (52.3)	48 (44.9)	3 (2.8)	157 (80.1)	38 (19.4)	1 (0.5)	3.7 (2.2–6.1)	4.30×10^{-7}
				Replication	96 (54.9)	74 (42.3)	5 (2.9)	181 (83.8)	32 (14.8)	3 (1.4)	4.3 (2.7–6.8)	3.83×10^{-10}
				Combined	152 (53.9)	122 (43.3)	8 (2.8)	338 (82.0)	70 (17.0)	4 (1.0)	3.9 (2.8–5.5)	1.33×10^{-15}

[a]Minor allele frequency and minor allele in 184 healthy Japanese individuals.
[b]OR for the minor allele in a dominant model.
[c]*P*-value by χ^2 test for the minor allele dominant model.

6.12 (2.78–13.46), $P < 0.0001$] as well as platelet counts [OR 1.18 (1.11–1.26), $P < 0.00001$]. We analyzed whether the rs1127354 genotype could influence the treatment outcome by PEG-IFN/RBV therapy. When analyzed in the patients available for treatment outcome (172 with *ITPA*-AA/CA and 450 with *ITPA*-CC), the percentage of patients receiving >80% of the expected PEG-IFN and RBV dose at baseline and week 4 was not significantly different among the rs1127354 genotypes. However, the rate of SVR tended to be higher in patients with *ITPA*-AA/CA genotype than those with *ITPA*-CC (48.8 versus 37.3%), because the relapse rate was lower in patients with *ITPA*-AA/CA. To investigate the influence on treatment outcome by dose reduction of PEG-IFN, in a subgroup of patients with low platelet counts (<10) at baseline (19 with *ITPA*-AA/CA and 53 with *ITPA*-CC) we analyzed the treatment outcome according to

rs1127354 genotypes. The SVR rate was very low in each group (21.1% in *ITPA*-AA/CA and 17.0% in *ITPA*-CC), because many patients had the initial dose reduction of PEG-IFN (<80% of standard dose)—36.8% of patients with *ITPA*-AA/CA and 44.6% of patients with *ITPA*-CC genotype. Further prospective studies are required among the pre-cirrhotic or cirrhotic patients with low platelet counts.

DISCUSSION

Recent genome-wide association studies, including our study on HCV infection, have identified two important host genetic variants: the SNP in *IL28B* gene, which is strongly associated with response to therapy for chronic genotype 1 HCV infection (16–21), and the SNP in *ITPA* gene, which precisely predicts RBV-induced anemia in

Table 3. Two SNPs (rs11697186 and rs6139030) significantly associated with quantitative change in Hb levels from baseline to week 4 of PEG-IFN/RBV treatment

dbSNP rsID	Nearest gene	MAF[a] (allele)	Allele (1/2)	Stage	Patients with quantitative change in Hb			Patients without quantitative change in Hb			OR (95% CI)[b]	*P*-value[c]
					11	12	22	11	12	22		
rs11697186	*DDRGK1*	0.15 (T)	T/A	GWAS	0 (0.0)	3 (3.3)	89 (96.7)	3 (1.5)	77 (37.0)	128 (61.5)	0.06 (0.02–0.16)	3.29×10^{-10}
				Replication	0 (0.0)	2 (1.5)	134 (98.5)	6 (2.5)	88 (37.3)	142 (60.2)	0.02 (0.01–0.09)	3.86×10^{-16}
				Combined	0 (0.0)	5 (2.2)	223 (97.8)	9 (2.0)	165 (37.2)	270 (60.8)	0.03 (0.01–0.08)	9.43×10^{-25}
rs6139030	ITPA	0.17 (C)	T/C	GWAS	88 (93.6)	6 (6.4)	0 (0.0)	125 (59.8)	80 (38.3)	4 (1.9)	0.08 (0.03–0.22)	2.56×10^{-9}
				Replication	134 (97.8)	3 (2.2)	0 (0.0)	143 (56.3)	103 (40.6)	8 (3.1)	0.03 (0.01–0.08)	6.90×10^{-18}
				Combined	222 (96.1)	9 (3.9)	0 (0.0)	268 (57.9)	183 (39.5)	12 (2.6)	0.04 (0.02–0.09)	2.12×10^{-25}

[a]Minor allele frequency and minor allele in 184 healthy Japanese individuals.
[b]OR for the minor allele in a dominant model.
[c]*P*-value by χ^2 square test for the minor allele dominant model.
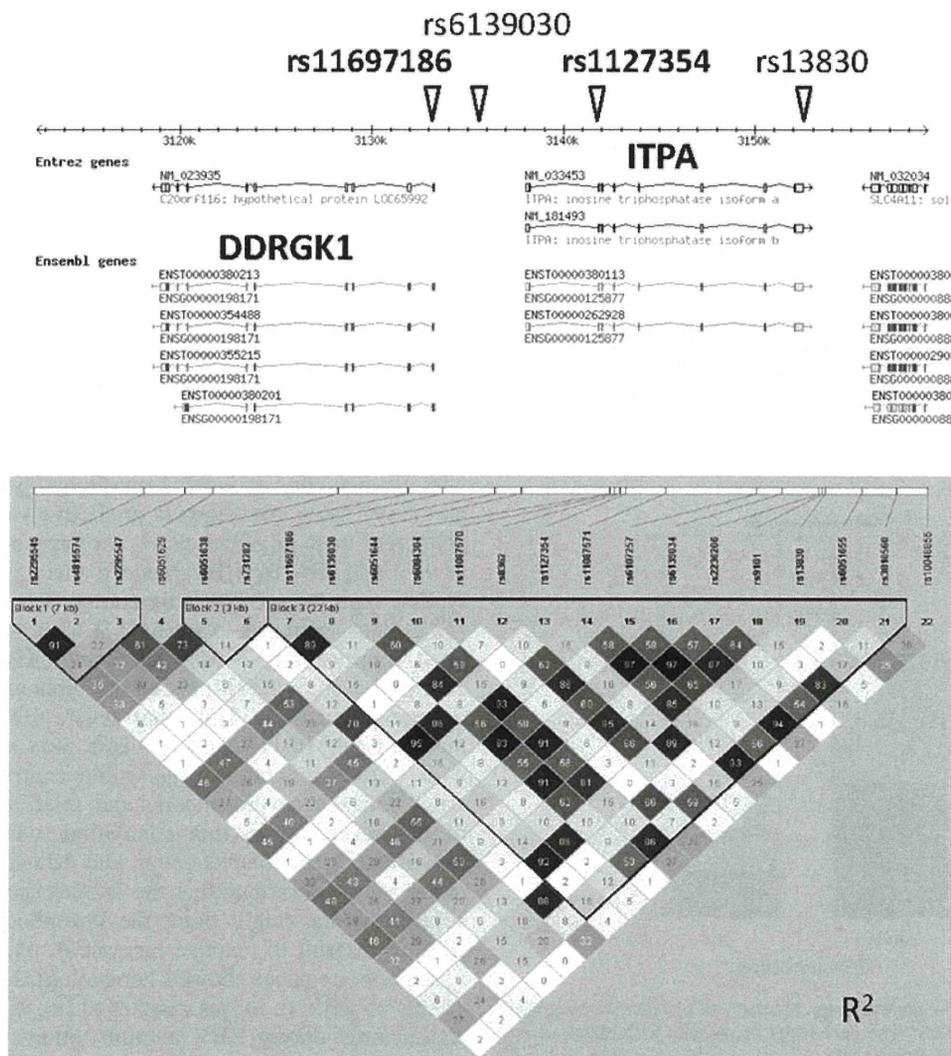

Figure 2. Pairwise LD (r^2) diagrams for *DDRGK1* and *ITPA*. Lower panel shows estimates of pairwise r^2 for 22 SNPs selected in the replication study using the second set of 391 Japanese HCV patients with and without quantitative change in PLT levels from baseline to week 4 of PEG-IFN/RBV treatment.

European-American population (22) and Japanese population (26). The genetic variation of *ITPA* causing an accumulation of inosine triphosphate (ITP) has been shown to protect patients against RBV-induced anemia during treatment for CHC infection. A recent report showed the biologic mechanism that ITP confers protection against RBV-induced ATP reduction by substituting for erythrocyte GTP, which is depleted by RBV, in the biosynthesis of ATP (25).
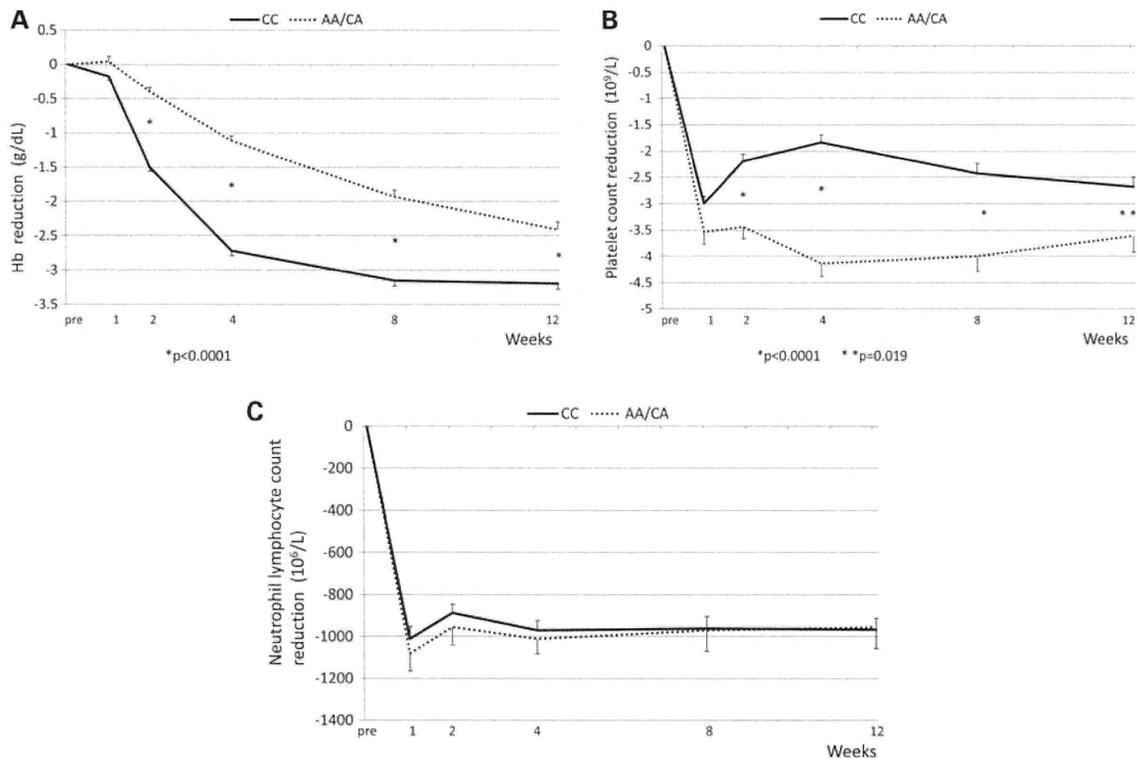
Figure 3. *ITPA* rs1127354 genotypes and the quantitative reduction of blood cells from baseline. Mean reduction of (**A**) Hb levels, (**B**) platelet counts and (**C**) neutrophil leukocyte counts during treatment according to rs1127354 genotype is shown. Solid and dotted lines indicate patients with CC and AA/CA genotypes, respectively. Error bars indicate standard error. CC genotype had more reduction in mean Hb levels during therapy compared with the AA/CA genotype (*$P < 0.0001$ for weeks 2, 4, 8, 12). CC genotype had less of a reduction in mean platelet counts (*$P < 0.0001$ for weeks 2, 4, 8, and **$P = 0.019$ for week 12), and showed a reactive increase of platelet counts through weeks 1–4.
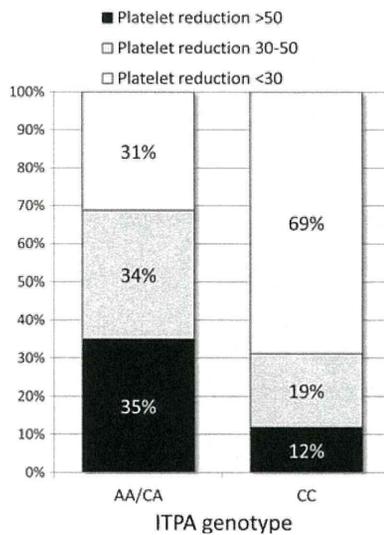

Figure 4. *ITPA* rs1127354 genotypes and reduction of platelet counts at week 4 of PEG-IFN/RBV therapy. The percentage of patients with platelet count reduction of >50 (10^9/l) (black bar), 30–50 (10^9/l) (gray bar) and <30 (10^9/l) (white bar) at week 4 is shown for rs1127354 genotypes. The incidence of platelet count reduction of >50 and <30 was significantly lower in patients with the rs1127354 genotypes CC compared with AA/CA genotypes: 12 versus 35%, $P < 0.0001$, and 69 versus 31%, $P < 0.0001$, respectively.

In this study, two SNPs, rs11697186 and rs6139030, which were within and around *DDRGK1* gene on chromosome 20, were strongly associated with thrombocytopenia as well as

with Hb reduction at week 4. In clinical practice, the positive predictive value and negative predictive value by rs11697186 genotypes were 66.5 and 69.4% for thrombocytopenia, as well as 97.2 and 45% for RBV-induced anemia at week 4. As previously reported (22,26), a functional SNP (rs1127354) in the *ITPA* locus, which is in strong LD with rs11697186, was the most significant SNP associated with RBV-induced anemia and, in this study, IFN-induced thrombocytopenia in Japanese genetic populations. Note that severe Hb decline, which is mainly found in ITPA-CC patients, was inversely correlated with platelet reduction. This would contribute to an association between severe anemia and relative reactive increase of platelet count in this population, which attenuated the IFN effect on the platelet count. Our data supported a previous report which described that the current use of RBV, inducing severe anemia, might blunt the thrombocytopenic effect of IFNs as a result of reactive increase of platelet counts (27).

A previous paper showed hematological and bone marrow effects of RBV in rhesus monkeys (28). Hb values decreased significantly during RBV administration due to dose-related erythroid hypoplasia in bone marrow and returned to normal following withdrawal. On the other hand, increase of the platelet count occurred in both low- and high-dose treatment groups during RBV administration, with a fall of the platelet count to normal after drug withdrawal. The effect on platelet count was clearly dose related, with maximum counts rising to twice and three times above baseline levels in the low- and high-dose groups, respectively. This caused a significant increase of
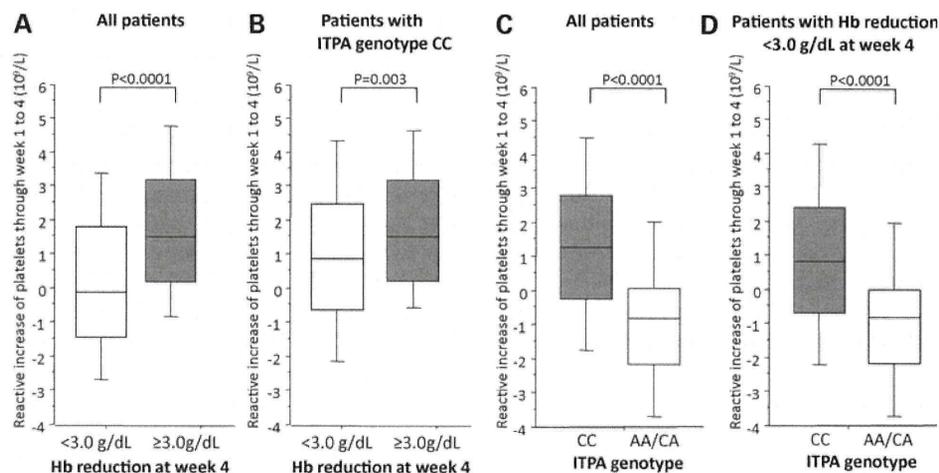
Figure 5. Reactive increase of platelet counts through weeks 1–4. Box plots of reactive increase of platelet count through weeks 1–4 according to the degree of anemia at week 4 are shown for all patients (**A**) and a subgroup of patients with the rs1127354 genotypes CC (**B**). Patients with anemia (Hb reduction $\geq$3.0 g/dl) at week 4 had a significantly higher degree of reactive increase of platelet count than those without anemia ($P < 0.0001$). Box plots of reactive increase of platelet counts according to the rs1127354 genotype CC are shown for all patients (**C**) and a subgroup of patients without anemia (**D**) (Hb reduction <3.0 g/dl) at week 4. Patients with the rs1127354 genotypes CC had a significantly high degree of reactive increase of platelet counts compared with those with genotypes AA/CA ($P < 0.0001$).

Table 4. Multivariate analysis of factors associated with reactive increase of platelets $\geq$20 (10^9/l) through weeks 1–4

	OR	95% CI	*P*-value
Baseline platelet counts	1.168	1.101–1.239	<0.0001
ITPA AA/CA	0.379	0.168–0.856	0.0196
Platelet reduction $\geq$30 (10^9/l) at week 4	0.051	0.021–0.120	<0.0001
Hb reduction $\geq$3.0 g/dl at week 4	1.602	0.914–2.809	0.0996

the platelet count associated with increased numbers of mega-karyocytes. Additionally, the sequence homology of thrombo-poietin (TPO) and erythropoietin (EPO) may explain the synergy of the physiologic role of TPO and EPO in platelet production. When EPO is elevated, as in iron deficiency anemia, an amino acid sequence similar to TPO may increase the platelet count (29).

Another possibility is a direct association between *ITPA* SNPs or the related SNPs with a strong LD and IFN-induced thrombocytopenia. *DDRGK1* (DDRGK domain-containing protein 1) is a novel C53/LZAP-interacting protein. C53/LZAP (also named as Cdk5rap3) is a putative tumor suppressor that plays important roles in multiple cell signaling pathways, including DNA damage response and NF-kappaB signaling (30); however, it remains largely unknown how the function of *DDRGK1* variants is regulated. Further studies are required to elucidate the possible association between *DDRGK1* variants and thrombocytopenia.

Multivariate analysis demonstrated that rs1127354 in the *ITPA* gene was independently associated with RBV-induced severe anemia and IFN-induced thrombocytopenia. This finding suggests that rs1127354 would be a useful marker to predict these hematological side effects by PEG-IFN/RBV therapy, indicating that genetic testing of *ITPA* variant might be applied to establish personalized dosages of PEG-IFN/RBV therapy. The rate of SVR tended to be higher in patients with *ITPA*-AA/CA genotype than those

with *ITPA*-CC in this population. This might reflect decreased treatment efficacy (higher relapse rate) due to dose reduction of RBV in patients with *ITPA*-CC genotype. Our recent paper also demonstrated that the incidence of early dose reduction was significantly higher in *ITPA*-major (CC) patients as expected and, more importantly, that a significantly higher SVR rate was achieved in *ITPA*-hetero/minor (CA/AA) patients with HCV non-1b or low viral load strains (31) and in a subset of Japanese patients with the favorable TT genotype at rs8099917 of *IL28B* (32). Taken together, our results indicate that the *ITPA* minor variant A is not only a protective allele against PEG-IFN and RBV treatment-associated anemia in Japanese population, but also a significant predictor of SVR in certain HCV strains that show good response to IFN. The possible mechanism of protection against RBV-induced hemolysis is that ITP deficiency or low-activity variants (*ITPA* minor variant A) in turn lead to the accumulation of ITP in red blood cells (33,34), and the ITP confers protection against RBV-induced ATP reduction by substituting for erythrocyte GTP (25). On the other hand, half of the *ITPA*-major (CC) patients did not develop a significant Hb decline. This finding suggests other low-frequency *ITPA* variants or SNPs in other enzymes that are involved in erythrocyte purine nucleoside metabolism.

In Japan, the older HCV-infected patients developing liver fibrosis have been prevalent (mean age 62 years) (9). Thrombocytopenia by PEG-IFN/RBV therapy could lead to poor treatment efficiency among such Japanese patients with LC due to the initial or early dose reduction of PEG-IFN. In fact, ~40% of such population in this study had the initial dose reduction of PEG-IFN, resulting in a low SVR rate. Splenectomy or embolization of the splenic artery might be one of the options to increase the SVR rate, but a sufficient treatment outcome had not been obtained at present (35). Based on the recently accumulated SNP data, if patients had favorable *IL28B* genotype and *ITPA*-CC (lower reduction of platelet counts), a standard dose of PEG-IFN might be available for

the patients with lower platelet counts and the SVR rate might be increased due to sufficient dose of PEG-IFN.

Several STAT-C agents (specifically targeted antiviral therapies for hepatitis C) are being tested for clinical efficacy against hepatitis C (12,13,15,16). Most experts believe that when new drugs are approved to treat hepatitis C, they will be used in combination with PEG-IFN and RBV. Moreover, recent clinical trials, including NS3 protease inhibitors, have shown that PEG-IFN plus RBV would be necessary to achieve optimal treatment responses (12,13). Our present results may provide a valuable pharmacogenetic diagnostic tool for tailoring PEG-IFN and RBV dosing to minimize drug-induced adverse events and for further optimization of clinical anti-HCV chemotherapeutics.

MATERIALS AND METHODS

Patients

From April 2007 to April 2010, samples were obtained from 303 patients with chronic HCV (genotype 1) infection who were treated at 14 multi-center hospitals (liver units with hepatologists) throughout Japan. Each patient was treated with PEG-IFN-α2b (1.5 μg/kg body weight, subcutaneously once a week) or PEG-IFN-α2a (180 μg once a week) plus RBV (600–1000 mg daily according to body weight) for 48 weeks. Treatment duration was extended in some patients up to 72 weeks, according to the physicians' preferences. The dose of PEG-IFN or RBV was reduced according to the recommendations on the package inserts or the clinical conditions of the individual patients. EPO or other growth factors were not given. Written informed consent was obtained from each patient and the study protocol conformed to the ethics guidelines of the Declaration of Helsinki and was approved by the institutional ethics review committees. HBsAg-positive and/or anti-HIV-positive patients were excluded from this study.

In the following stage of replication study, SNP genotyping in an independent set of 391 Japanese HCV patients treated with PEG-IFN plus RBV treatment was completed using the DigiTag2 or TaqMan assay (ABI) following the manufacturer's protocol. The characteristics of patients for each GWAS stage and replication stage are summarized in Table 1.

SNP genotyping and data cleaning

In the GWAS stage, we genotyped 303 Japanese HCV patients with and without the decrease of platelet counts from baseline to week 4 of PEG-IFN/RBV treatment [107 patients with a decrease of >30 (10^9/l) in platelet counts and 196 patients without a decrease of >30 (10^9/l) in platelet counts], using the Affymetrix Genome-Wide Human SNP Array 6.0 according to the manufacturer's instructions. The cut-off value was calculated to maximize the difference, which was also close to the median change. The average overall call rate of patients with and without the decrease of PLT reached 98.69 and 98.72%, respectively. We then applied the following thresholds for SNP QC in data cleaning: SNP call rate ≥95% for all samples, MAF ≥1% for all samples. A total of 595 052 SNPs on autosomal chromosomes passed the QC filters and were used for association analysis. All cluster

plots of SNPs showing $P < 0.0001$ in association analyses by comparing allele frequencies in both groups with and without the decrease of PLT were checked by visual inspection, and SNPs with ambiguous genotype calls were excluded.

In the following stage of the replication study and high-density association mapping, we selected 23 tag SNPs from the 44.7 kb region, including *DDRGK1* gene and *ITPA* gene by analyzing LD and haplotype structure based on the HapMap data of Japanese, using the Haploview software. Of these tag SNPs, rs1127354 within the *ITPA* gene, which was associated with RBV-induced anemia (22), was included; however, rs7270101 was excluded because recent papers studying Japanese patients showed no variants in rs7270101 (26,31,32). The SNP genotyping in an independent set of 391 Japanese HCV patients with and without quantitative change in PLT levels from baseline to week 4 of PEG-IFN/RBV treatment (175 patients with quantitative change in PLT and 216 patients without quantitative change in PLT) was completed using the DigiTag2 assay (36). Twenty-two of the 23 SNPs were successfully analyzed and were used for SNP genotyping and data cleaning. All 22 SNPs in the replication study cleared HWE P-value >0.001.

Based on the above SNPs data obtained from 303 Japanese HCV patients, using the Affymetrix Genome-Wide Human SNP Array 6.0, we also performed GWAS between 94 patients with a quantitative change of >3 g of reduction in Hb and 209 patients without quantitative change in Hb levels from baseline to week 4 of PEG-IFN/RBV treatment. SNP genotyping in an independent set of 391 Japanese HCV patients with and without quantitative change in Hb levels from baseline to week 4 of PEG-IFN/RBV treatment (137 patients with quantitative change in Hb and 254 patients without quantitative change in Hb) was also completed using the DigiTag2 assay (36). Twenty-two of the 23 SNPs were successfully analyzed and were used for SNP genotyping and data cleaning.

An application of the Cochrane–Armitage test on all the SNPs showed the genetic inflation factor $\lambda = 1.000$ for thrombocytopenia and $\lambda = 1.006$ for anemia in the GWAS stage (Supplementary Material, Figs S1 and S2). In addition, principal component analysis was performed in 303 samples for the GWAS stage together with the HapMap samples (CEU, YRI, CHB and JPT) (Supplementary Material, Fig. S3). These results implied that the effect of population stratification was negligible, except one sample, which was excluded from further analysis.

Laboratory and histological tests

Blood samples were obtained at baseline, 1, 2, 4, 8 and 12 weeks after the start of therapy and for hematologic tests after the start of therapy and for hematologic tests, blood chemistry and HCV-RNA. Genetic polymorphism in the *IL28B* gene (rs8099917) was determined using the ABI TaqMan assay (Applied Biosystems, Carlsbad, CA, USA). Fibrosis was evaluated on a scale of 0–4 according to the METAVIR scoring system. The SVR was defined as an undetectable HCV-RNA level by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor, Roche Diagnostic Systems, CA, USA) or by Cobas Ampliprep/Cobas TaqMan assay (CAP/CTM) with a lower detection limit of

15 IU/ml (Roche Diagnostic Systems) 24 weeks after the completion of therapy.

Statistical analysis

The observed association between an SNP and the decrease of platelets/quantitative change in Hb levels with response to PEG-IFN plus RBV treatment was assessed by χ^2 test with a two-by-two contingency table in three genetic models: allele frequency model, dominant-effect model and recessive-effect model. SNPs on chromosome X were removed because gender was not matched between groups with and without the decrease of PLT and quantitative change in Hb levels. A total of 595 052 SNPs passed the quality control filters in the GWAS stage; therefore, significance levels after Bonferroni correction for multiple testing were $P = 8.40 \times 10^{-8}$ (0.05/595052) in the GWAS stage and $P = 2.27 \times 10^{-3}$ (0.05/22) in the replication stage.

The association between an SNP of the *ITPA* gene (rs1127354) and the incidence of platelet reduction at week 4 was analyzed by Fisher's exact test. The association between *ITPA* polymorphisms and the degree of reduction in platelet counts and Hb levels at each time point during therapy were analyzed by Mann–Whitney U test. Multivariable regression analysis was used to analyze the factors associated with *ITPA*, the rs1127354 genotype, factors associated with platelet count reductions and factors associated with the reactive increase in platelet counts. IBM-SPSS software v.15.0 (SPSS, Inc., Chicago, IL, USA) was used for these analyses.

Possible heterogeneity in allele frequencies at rs1127354 was assessed by Tarone's test. The association between the SNP and thrombocytopenia/anemia were analyzed by the Cochran–Mantel–Haenszel test. Both analyses were performed using the R (version 2.9.0) software (Supplementary Material, Table S3).

AUTHORS' CONTRIBUTIONS

Drafting of the paper, statistical analysis and approval of the final draft submitted: M.M.; drafting of the paper, statistical analysis, collecting samples and clinical data and approval of the final draft submitted: Y.T. and M.K.; statistical analysis and approval of the final draft submitted: N.N., M.S. and K.T.; collecting samples and clinical data and approval of the final draft submitted: K.M., N.S., N.E., H.Y., S.N., K.H., S.H., Y.I., E.T., S.M., M.H., Y.H., F.S., S.K. and N.I.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

This study is based on 14 multi-center hospitals throughout Japan: Hokkaido area (Hokkaido University Hospital), Kanto area (Saitama University Hospital; Konodai Hospital; Musashino Red Cross Hospital; Tokyo Medical and Dental University Hospital; Yamanashi University Hospital), Koshin area (Shinshu University Hospital; Kanazawa University Hospital), Tokai area (Nagoya City University Hospital), Kinki area (Kyoto Prefectural University of Medicine Hospital; Hyogo College of Medicine Hospital), Chugoku/Shikoku area (Ehime University Hospital; Kawasaki Medical College Hospital) and Kyushu area (National Nagasaki Medical Center). We thank Ms Yasuka Uehara-Shibata, Yuko Ogasawara-Hirano, Yoshimi Ishibashi, Natsumi Baba and Megumi Yamaoka-Sageshima (Tokyo University) for technical assistance. We also thank Dr Masaaki Korenaga (Kawasaki), Dr Akihiro Matsumoto (Shinshu), Dr Kayoko Naiki (Saitama), Dr Takeshi Nishimura (Kyoto), Dr Hirayuki Enomoto (Hyogo), Dr Minako Nakagawa (Tokyo Medical and Dental University) and Ochanomizu Liver Conference Study Group for collecting samples, and Dr Mamoru Watanabe (Tokyo Medical and Dental University) and Dr Moriichi Onji (Ehime University) for their advice throughout the study.

REFERENCES

1. Global Burden of Hepatitis C Working Group (2004) Global burden of disease (GBD) for hepatitis C. *J. Clin. Phamacol.*, **44**, 20–29.
2. Shiratori, Y., Shiina, S., Imamura, M., Kato, N., Kanai, F., Okudaira, T., Teratani, T., Tohgo, G., Toda, N., Ohashi, M. *et al.* (1995) Characteristic difference of hepatocellular carcinoma between hepatitis B- and C-viral infection in Japan. *Hepatology*, **22**, 1027–1033.
3. Yoshida, H., Tateishi, R., Arakawa, Y., Sata, M., Fujiyama, S., Nishiguchi, S., Ishibashi, H., Yamada, G., Yokosuka, O., Shiratori, Y. *et al.* (2004) Benefit of interferon therapy in hepatocellular carcinoma prevention for individual patients with chronic hepatitis C. *Gut*, **53**, 425–430.
4. George, S.L., Bacon, B.R., Brunt, E.M., Mihindukulasuriya, K.L., Hoffmann, J. and Di Bisceglie, A.M. (2009) Clinical, virologic, histologic, and biochemical outcomes after successful HCV therapy: a 5-year follow-up of 150 patients. *Hepatology*, **49**, 729–738.
5. Fried, M.W., Shiffman, M.L., Reddy, K.R., Smith, C., Marinos, G., Goncales, F.L. Jr, Haussinger, D., Diago, M., Carosi, G., Dhumeaux, D. *et al.* (2002) Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.*, **347**, 975–982.
6. Manns, M.P., McHutchison, J.G., Gordon, S.C., Rustgi, V.K., Shiffman, M., Reindollar, R., Goodman, Z.D., Koury, K., Ling, M. and Albrecht, J.K. (2001) Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet*, **358**, 958–965.
7. Hadziyannis, S.J., Sette, H. Jr, Morgan, T.R., Balan, V., Diago, M., Marcellin, P., Ramadori, G., Bodenheimer, H. Jr, Bernstein, D., Rizzetto, M. *et al.* (2004) Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann. Intern. Med.*, **140**, 346–355.
8. Hiramatsu, N., Oze, T., Tsuda, N., Kurashige, N., Koga, K., Toyama, T., Yasumaru, M., Kanto, T., Takehara, T., Kasahara, A. *et al.* (2006) Should aged patients with chronic hepatitis C be treated with interferon and ribavirin combination therapy? *Hepatol. Res.*, **35**, 185–189.
9. Iwasaki, Y., Ikeda, H., Araki, Y., Osawa, T., Kita, K., Ando, M., Shimoe, T., Takaguchi, K., Hashimoto, N., Kobatake, T. *et al.* (2006) Limitation of

combination therapy of interferon and ribavirin for older patients with chronic hepatitis C. *Hepatology*, **43**, 54–63.

10. Sezaki, H., Suzuki, F., Akuta, N., Yatsuji, H., Hosaka, T., Kobayashi, M., Suzuki, Y., Arase, Y., Ikeda, K., Miyakawa, Y. *et al.* (2009) An open pilot study exploring the efficacy of fluvastatin, pegylated interferon and ribavirin in patients with hepatitis C virus genotype 1b in high viral loads. *Intervirology*, **52**, 43–48.

11. Bruno, R., Sacchi, P., Maiocchi, L., Patruno, S. and Filice, G. (2006) Hepatotoxicity and antiretroviral therapy with protease inhibitors: a review. *Dig. Liver Dis.*, **38**, 363–373.

12. Hezode, C., Forestier, N., Dusheiko, G., Ferenci, P., Pol, S., Goeser, T., Bronowicki, J.P., Bourliere, M., Gharakhanian, S., Bengtsson, L. *et al.* (2009) Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N. Engl. J. Med.*, **360**, 1839–1850.

13. McHutchison, J.G., Everson, G.T., Gordon, S.C., Jacobson, I.M., Sulkowski, M., Kauffman, R., McNair, L., Alam, J. and Muir, A.J. (2009) Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N. Engl. J. Med.*, **360**, 1827–1838.

14. Suzuki, F., Akuta, N., Suzuki, Y., Sezaki, H., Yatsuji, H., Kawamura, Y., Hosaka, T., Kobayashi, M., Arase, Y., Ikeda, K. *et al.* (2009) Rapid loss of hepatitis C virus genotype 1b from serum in patients receiving a triple treatment with telaprevir (MP-424), pegylated interferon and ribavirin for 12 weeks. *Hepatol. Res.*, **39**, 1056–1063.

15. Sakamoto, N. and Watanabe, M. (2009) New therapeutic approaches to hepatitis C virus. *J. Gastroenterol.*, **44**, 643–649.

16. Afdhal, N.H., McHutchison, J.G., Zeuzem, S., Mangia, A., Pawlotsky, J.M., Murray, J.S., Shianna, K.V., Tanaka, Y., Thomas, D.L., Booth, D.R. *et al.* (2010) Hepatitis C pharmacogenetics: state of the art in 2010. *Hepatology*, **53**, 336–345.

17. Tanaka, Y., Nishida, N., Sugiyama, M., Kurosaki, M., Matsuura, K., Sakamoto, N., Nakagawa, M., Korenaga, M., Hino, K., Hige, S. *et al.* (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat. Genet.*, **41**, 1105–1109.

18. Ge, D., Fellay, J., Thompson, A.J., Simon, J.S., Shianna, K.V., Urban, T.J., Heinzen, E.L., Qiu, P., Bertelsen, A.H., Muir, A.J. *et al.* (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature*, **461**, 399–401.

19. Suppiah, V., Moldovan, M., Ahlenstiel, G., Berg, T., Weltman, M., Abate, M.L., Bassendine, M., Spengler, U., Dore, G.J., Powell, E. *et al.* (2009) IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat. Genet.*, **41**, 1100–1104.

20. Thomas, D.L., Thio, C.L., Martin, M.P., Qi, Y., Ge, D., O'Huigin, C., Kidd, J., Kidd, K., Khakoo, S.I., Alexander, G. *et al.* (2009) Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature*, **461**, 798–801.

21. Tanaka, Y., Nishida, N., Sugiyama, M., Tokunaga, K. and Mizokami, M. (2010) lambda-Interferons and the single nucleotide polymorphisms: a milestone to tailor-made therapy for chronic hepatitis C. *Hepatol. Res.*, **40**, 449–460.

22. Fellay, J., Thompson, A.J., Ge, D., Gumbs, C.E., Urban, T.J., Shianna, K.V., Little, L.D., Qiu, P., Bertelsen, A.H., Watson, M. *et al.* (2010) ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. *Nature*, **464**, 405–408.

23. Afdhal, N., McHutchison, J., Brown, R., Jacobson, I., Manns, M., Poordad, F., Weksler, B. and Esteban, R. (2008) Thrombocytopenia associated with chronic liver disease. *J. Hepatol.*, **48**, 1000–1007.

24. Wazny, L.D. and Ariano, R.E. (2000) Evaluation and management of drug-induced thrombocytopenia in the acutely ill patient. *Pharmacotherapy*, **20**, 292–307.

25. Hitomi, Y., Cirulli, E.T., Fellay, J., McHutchison, J.G., Thompson, A.J., Gumbs, C.E., Shianna, K.V., Urban, T.J. and Goldstein, D.B. (2011) Inosine triphosphate protects against ribavirin-induced adenosine triphosphate loss by adenylosuccinate synthase function. *Gastroenterology*, **140**, 1314–1321.

26. Ochi, H., Maekawa, T., Abe, H., Hayashida, Y., Nakano, R., Kubo, M., Tsunoda, T., Hayes, C.N., Kumada, H., Nakamura, Y. *et al.* (2010) ITPA polymorphism affects ribavirin-induced anemia and outcomes of therapy—a genome-wide study of Japanese HCV virus patients. *Gastroenterology*, **139**, 1190–1197.

27. Ong, J.P. and Younossi, Z.M. (2004) Managing the hematologic side effects of antiviral therapy for chronic hepatitis C: anemia, neutropenia, and thrombocytopenia. *Cleve. Clin. J. Med.*, **71** (Suppl. 3), S17–S21.

28. Canonico, P.G., Kastello, M.D., Cosgriff, T.M., Donovan, J.C., Ross, P.E., Spears, C.T. and Stephen, E.L. (1984) Hematological and bone marrow effects of ribavirin in rhesus monkeys. *Toxicol. Appl. Pharmacol.*, **74**, 163–172.

29. Akan, H., Guven, N., Aydogdu, I., Arat, M., Beksac, M. and Dalva, K. (2000) Thrombopoietic cytokines in patients with iron deficiency anemia with or without thrombocytosis. *Acta Haematol.*, **103**, 152–156.

30. Wu, J., Lei, G., Mei, M., Tang, Y. and Li, H. (2010) A novel C53/LZAP-interacting protein regulates stability of C53/LZAP and DDRGK domain-containing protein 1 (DDRGK1) and modulates NF-kappaB signaling. *J. Biol. Chem.*, **285**, 15126–15136.

31. Sakamoto, N., Tanaka, Y., Nakagawa, M., Yatsuhashi, H., Nishiguchi, S., Enomoto, N., Azuma, S., Nishimura-Sakurai, Y., Kakinuma, S., Nishida, N. *et al.* (2010) ITPA gene variant protects against anemia induced by pegylated interferon-alpha and ribavirin therapy for Japanese patients with chronic hepatitis C. *Hepatol. Res.*, **40**, 1063–1071.

32. Kurosaki, M., Tanaka, Y., Tanaka, K., Suzuki, Y., Hoshioka, Y., Tamaki, N., Kato, T. and Yasui, Y. (2011) Analysis of the correlations between genetic polymorphisms of the ITPA gene and hemolytic anemia or outcome after treatment with pegylated-interferon and ribavirin in genotype 1b chronic hepatitis C. *Antivir. Ther.*, in press.

33. Shipkova, M., Lorenz, K., Oellerich, M., Wieland, E. and von Ahsen, N. (2006) Measurement of erythrocyte inosine triphosphate pyrophosphohydrolase (ITPA) activity by HPLC and correlation of ITPA genotype-phenotype in a Caucasian population. *Clin. Chem.*, **52**, 240–247.

34. Fraser, J.H., Meyers, H., Henderson, J.F., Brox, L.W. and McCoy, E.E. (1975) Individual variation in inosine triphosphate accumulation in human erythrocytes. *Clin. Biochem.*, **8**, 353–364.

35. Kumada, H., Okanoue, T., Onji, M., Moriwaki, H., Izumi, N., Tanaka, E., Chayama, K., Sakisaka, S., Takehara, T., Oketani, M. *et al.* (2010) Guidelines for the treatment of chronic hepatitis and cirrhosis due to hepatitis C virus infection for the fiscal year 2008 in Japan. *Hepatol. Res.*, **40**, 8–13.

36. Nishida, N., Tanabe, T., Takasu, M., Suyama, A. and Tokunaga, K. (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal. Biochem.*, **364**, 78–85.

ORIGINAL ARTICLE

Analysis of viral amino acids sequences and the IL28B SNP influencing the development of hepatocellular carcinoma in chronic hepatitis C

Mika Miura · Shinya Maekawa · Makoto Kadokura · Ryota Sueki · Kazuki Komase ·
Hiroko Shindo · Takako Ohmori · Asuka Kanayama · Kuniaki Shindo · Fumitake Amemiya ·
Yasuhiro Nakayama · Takatoshi Kitamura · Tomoyoshi Uetake · Taisuke Inoue ·
Minoru Sakamoto · Shunichi Okada · Nobuyuki Enomoto

Abstract

Background and aims The association between hepatitis C virus (HCV) sequences with interleukin 28B (IL28B) single-nucleotide polymorphism (SNP) in the development of hepatocellular carcinoma (HCC) has not been well clarified.

Methods Complete HCV open-reading frame sequences were determined in 20 patients developing HCC and 23 non-HCC patients with HCV-1b infection in two distant time points. An additional 230 patients were studied cross-sectionally for core and NS5A sequences with HCC development. Among them, 98 patients with available samples were investigated for changes in viral core sequences over time. Finally, IL28B SNPs and HCC development were investigated in 228 patients.

Results During observation period (HCC for 10.8 years, and non-HCC for 11.1 years), changes in core a.a. 70 and three amino acid positions in NS5A were characteristics of the patients developing HCC. In 230 patients, Q (glutamine) or H (histidine) to R (arginine) ratio at core a.a. 70 was significantly higher in the HCC group (HCC group 43:22 vs. non-HCC group 66:99, $p = 0.001$). A change in core R70Q was observed over time in 11 patients associated with a decrease in platelets ($p = 0.005$) and albumin ($p = 0.005$), while a Q70R change was observed in 4 patients without associated changes in platelets (nonsignificant) and albumin (nonsignificant). IL28B SNP showed significant correlation with the core a.a. 70 residue. There was no evident link between IL28B SNPs and the occurrence of HCC.

Conclusions Hepatitis C virus core a.a. 70 residue is associated with liver disease progression and is independent factor for HCC development in genotype-1b infection. IL28B SNPs are related to core a.a. 70 residue, but not to HCC. The functional relevance of core a.a. 70 residue in hepatitis C pathogenesis should be further investigated.

Keywords HCV · HCC · Core · IL28B

Introduction

Hepatitis C virus (HCV) infection is a major risk factor for hepatocellular carcinoma (HCC). Chronic HCV infection can result in liver cirrhosis (LC) and HCC over the course of 20–30 years [1]. However, the rate of progression is variable; some patients remain for a long time with persistently normal ALT values, while others progress rapidly to LC and HCC.

Viral factors, host factors, and their interplay appear to play an important role in determining the progression of chronic hepatitis C to LC and HCC. In terms of viral factors, most previous clinical studies have focused on searching for HCV regions correlated with the response to interferon (IFN)-based therapy. In those analyses, correlation between amino acid substitutions and treatment response have been reported for the IFN sensitivity

M. Miura · S. Maekawa (✉) · M. Kadokura · R. Sueki ·
K. Komase · H. Shindo · T. Ohmori · A. Kanayama ·
K. Shindo · F. Amemiya · Y. Nakayama · T. Kitamura ·
T. Uetake · T. Inoue · M. Sakamoto · S. Okada · N. Enomoto
First Department of Internal Medicine, Faculty of Medicine,
University of Yamanashi, 1110, Shimokato, Chuo,
Yamanashi 409-3898, Japan
e-mail: maekawa@yamanashi.ac.jp

🖄 Springer

determining region in nonstructural (NS)5A [2], a.a. 70 and 91 in core [3], the PKR/eIF-2alpha phosphorylation homology domain (PePHD) in envelope (E)2 [4], and the IFN–ribavirin resistance determining region (IRRDR) in NS5A [5].

Regarding the viral factors related to disease progression, among the various HCV proteins, the core protein has been thought widely to contribute because it has been shown experimentally to affect multiple cellular functions, in addition to the evidence from clinical studies [6–10]. Core protein modifies cellular apoptosis, oncogenic signaling, reactive oxygen species formation, lipid metabolism, transcriptional activation, transformation, and immune reactivity. Core protein has oncogenic potential in transgenic mice [11]. In contrast, fewer clinical studies to date have systematically investigated the correlation between the variability of HCV regions and disease progression. However, some of those limited clinical studies reported a correlation between amino acid substitutions in core or NS5A with disease progression [12–15]. Despite those reports, few studies to date support the correlation. Moreover, it is unclear whether those viral sequences change during disease progression or how the disease activity is modified by those viral sequences in the long course of chronic hepatitis.

On the other hand, regarding host factors, recent reports disclosed a significant correlation between polymorphisms in the IL28B gene and responses to pegylated-IFN plus ribavirin therapy for HCV patients [16–19]. This single-nucleotide polymorphism (SNP) also showed significant correlation with natural HCV clearance [20]. However, it remains unknown whether the IL28B SNP is related to disease progression or the development of HCC.

In this study, we first undertook the analysis to identify the viral regions related to disease progression and HCC development through the analysis of complete HCV open-reading frame (ORF) sequences. Because some regions in HCV core and NS5A showed characteristic changes over time in patients developing HCC during the observation period, we proceeded further to analyze the contribution of those regions to disease progression, in association with time and with the IL28B SNP.

Patients and methods

Patients

This study is based on the analysis of two groups of patients, 43 in Group 1 and 230 patients in Group 2.

In the first part, we tried to characterize and extract viral sequences specific to disease progression through the analysis of complete HCV ORFs (Fig. 1a). In particular, we focused our investigation on the changes in viral sequences over time in association with disease progression by comparing HCV sequences of two sufficiently distant time points. With this aim, we determined to investigate patients with a history of IFN therapy, because those patients often were followed long-term with preservation of old and recent sera. However, we excluded sustained
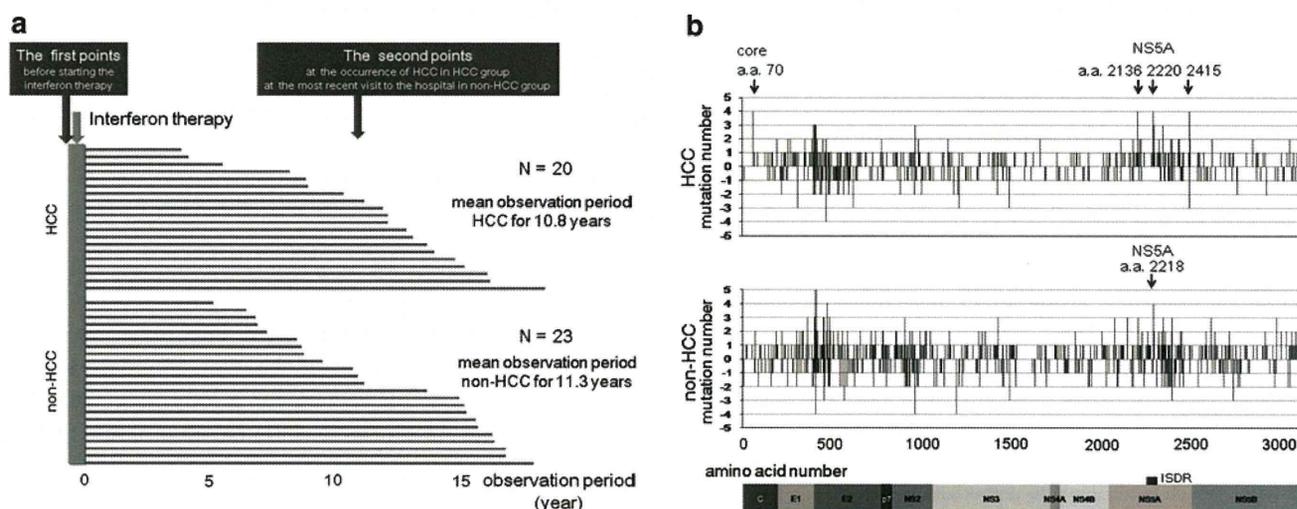
Fig. 1 **a** A total of 43 patients were analyzed for complete HCV ORF sequences. They were all non-responders to the previous IFN therapy. Twenty patients developed HCC during the observation period, while the 23 patients did not. HCV ORF sequences were determined for the paired samples and the predicted amino acid changes were compared in each patient. **b** Specific HCV amino acid changes associated with disease progression was evaluated by the analysis of the full-length viral ORF during the observation period for each patient according to the following rules: *1* +1 point for consensus to non-consensus, *2* −1 point for non-consensus to consensus, *3* 0 point for non-consensus to non-consensus. These points were added together and are shown for HCC and non-HCC patients. Patient with later HCC development (*upper panel*). Non-HCC patients (*lower panel*)

virologic response (SVR) patients because we thought that the viral clearance leads to improvement of the liver disease and, therefore, viral regions influencing the IFN response would be extracted as affecting the course of disease. Between March 1992 and April 2004, 273 consecutive patients with HCV-1b infection were given IFN monotherapy at Yamanashi University Hospital, and 133 were followed long-term. A total of 65 patients showed SVR, while 68 showed non-SVR. Among these 68 non-SVRs, 43 patients were included in the study because laboratory data and sera were available from the two distant time points (Group 1). Twenty patients developed HCC during the observation period, while the remaining 23 did not. Regarding the sera, the first time point for both groups was before starting IFN therapy, while the second points were at the occurrence of HCC for the HCC group and at the most recent visit to the hospital for the non-HCC group (Fig. 1a).

An additional 230 HCV-1b patients were recruited (Group 2) as the second study group. They were mainly outpatients at Yamanashi University Hospital and were selected randomly from those with stored sera at the time of disease diagnosis. Sixty-five had HCC, while the remaining 165 did not. They all were positive for HCV RNA at the time of study, although 74 patients had a history of IFN therapy. Parts of the core and NS5A sequences were determined at HCC onset in 65 HCC patients and at the most recent visit to the hospital in 165 non-HCC patients. Because historical sera around 10 years before were also available for 55 of these 230 patients, HCV sequence analysis was also performed for core and NS5A at those previous time points in those patients.

From these two study groups, 228 patients (68 HCCs and 160 non-HCCs) with available genomic DNAs were examined to determine the IL28B SNP.

All the patients studied all fulfilled following criteria: (1) Negative for hepatitis B surface antigen. (2) No other forms of hepatitis, such as primary biliary cirrhosis, autoimmune liver disease, or alcoholic liver disease. (3) Free of coinfection with human immunodeficiency virus. (4) A signed consent was obtained for the study protocol that had been approved by Human Ethics Review Committee of Yamanashi University Hospital.

Complete and partial HCV ORF sequence determination by direct sequencing from sera

HCV RNA extraction, complementary DNA synthesis and amplification by two-step nested PCR from serum samples were done using the specific primers for full HCV ORF or partial viral regions as described previously [15]. PCR amplicons were sequenced directly by Big Dye Terminator Version 3.1 (ABI, Tokyo, Japan) with universal M13

forward and reverse primers using an ABI prism 3130 sequencer (ABI). Generated sequence files were assembled using Vector NTI software (Invitrogen, Tokyo, Japan) and base-calling errors were corrected following inspection of the chromatogram.

IL28B SNP analysis

Human genomic DNA was extracted from peripheral blood using a blood DNA extraction kit (QIAGEN, Tokyo, Japan) according to the manufacturer's protocol. The allele typing of each DNA sample was performed by real-time PCR with a model 7500 (ABI) using FAM-labeled SNP primer for the locus rs8099917 (ABI).

Statistical analysis

Statistical differences in the parameters, including all available patients' demographic, biochemic, hematologic, and virologic data, were determined between different groups of patients by Student's t test for numerical variables and Fisher's exact probability test for categorical variables. Odds ratios and their 95% confidence intervals were used to quantify the level of association. All p values of <0.05 by the two-tailed test were considered significant throughout. Multiple logistic regression analyses were used to identify the independent variables influencing core a.a. 70 residue and HCC development. Because most variables used for the analyses were generally considered to correlate with the disease progression, we entered all the variables into the multiple logistic regression analysis even if some of them did not reach significant differences in individual univariate analysis.

Results

Comparing complete HCV amino acid sequences between patients with and without HCC

The clinical characteristics of the 43 patients (Group 1) analyzed for HCV ORF changes over time are shown in Table 1. At the start of observation, clinical characteristics did not differ significantly between the HCC group and the non-HCC group. The mean observation period was comparable between the two groups and was 10.8 years for the HCC group and 11.3 years for non-HCC group ($p = 0.745$). On the other hand, platelets ($p < 0.001$), albumin ($p < 0.001$), and AFP ($p = 0.001$) became significantly lower or higher in the HCC group at the end of observation (Table 1).

We proceeded to investigate viral amino acid changes during the course of disease in each patient to determine

Table 1 Patient characteristics in Group 1

	At the start of observation			At the end of observation		
	HCC ($N = 20$)	Non-HCC ($N = 23$)	p value	HCC ($N = 20$)	Non-HCC ($N = 23$)	p value
Observation period (years)				10.8 ± 3.6	11.3 ± 3.8	0.745
Sex (male/female)	11/9	12/11	0.999	11/9	12/11	0.999
Age (years)	51.5 ± 8.0	50.0 ± 9.9	0.604	$61.7 \pm 10.0*$	$61.0 \pm 10.9*$	0.818
Stage of fibrosis (F1/2/3/4)	1/7/6/6	5/11/4/3	0.190	N/A	N/A	–
AST (IU/L)	102 ± 114	74 ± 40	0.695	$71 \pm 36*$	51 ± 30	0.048
ALT (IU/L)	124 ± 86	104 ± 71	0.411	$69 \pm 47*$	$52 \pm 31*$	0.159
Platelets (10^{-4}/mm^3)	16.2 ± 4.8	18.3 ± 6.2	0.217	$9.7 \pm 3.9*$	15.3 ± 5.1	<0.001
Albumin (g/dL)	4.1 ± 0.4	4.1 ± 0.2	0.639	3.6 ± 0.4	4.1 ± 0.5	<0.001
γGTP (IU/L)	90 ± 60	71 ± 46	0.275	69 ± 59	$45 \pm 38*$	0.114
T.Chol (mg/dL)	169 ± 28	156 ± 22	0.110	146 ± 21	$164 \pm 5,108$	0.086
Alpha-fetoprotein (ng/mL)	10.5 ± 6.8	9.3 ± 10.8	0.695	42.4 ± 41.1	4.7 ± 2.7	0.001
HCV RNA concentration (kIU/mL)	706 ± 696	$614 \pm 1,181$	0.760	$3,325 \pm 415*$	$4,508 \pm 5,108*$	0.426

* Factors with significant changes over time (<0.05)

whether specific amino acid changes related to disease progression could be identified. First, the consensus amino acid was determined at each amino acid position in the HCV ORF after determination of all sequences in these 43 patients. Amino acid changes were determined according to the following rules to highlight directional changes according to disease progression: When an amino acid changed from the consensus to the non-consensus during the observation period, we scored +1 point. Conversely, a change from the non-consensus to the consensus scored −1 point. We scored 0 point for a change from one non-consensus amino acid to another. As shown in Fig. 1b, directional amino acid changes were observed throughout the HCV genome to some degree both in patients with and without HCC development during the clinical course of almost 10 years, and frequent substitutions in E2 hypervariable region were common in both groups. On the other hand, in patients with HCC development, as many as four directional changes were observed at core a.a. 70 and at three amino acid positions of NS5A (Fig. 1b, upper panel). In contrast, in patients without HCC, the significant change ($n = 4$) was observed at a.a. 2,218 of NS5A when E2 hypervariable region was excluded (Fig. 1b, lower panel).

Core and NS5A sequences in patients with and without HCC

Because the first analysis suggested that the patients with later HCC development might accumulate specific mutations in core and NS5A at the time of HCC occurrence, additional sequences were analyzed from 230 HCV-1b patients to confirm the result. The clinical backgrounds of the additional 230 patients are shown in Table 2 (Group 2).

Table 2 Patient characteristics in Group 2

	HCC ($N = 65$)	Non-HCC ($N = 165$)	p value
Observation period (years)			
Sex (male/female)	42/23	76/89	0.018
Age (years)	68.2 ± 9.2	62.4 ± 11.7	<0.001
AST (IU/L)	66 ± 35	41 ± 21	<0.001
ALT (IU/L)	67 ± 47	44 ± 47	<0.001
Platelets (10^{-4}/mm^3)	11.3 ± 5.8	15.3 ± 6.2	<0.001
Albumin (g/dL)	3.6 ± 0.5	4.4 ± 2.9	0.025
γGTP (IU/L)	59 ± 53	38 ± 40	0.001
T.Chol (mg/dL)	153 ± 30	165 ± 31	0.004
Alpha-fetoprotein (ng/mL)	$302 \pm 1,670$	10 ± 25	0.025
HCV RNA concentration (kIU/mL)	$5,400 \pm 13,574$	$7,990 \pm 8,512$	0.104

All patients were positive for HCV RNA. Between the HCC (65 patients) and non-HCC (165 patients) groups, HCC patients were older ($p < 0.001$) and more frequently tended to be males ($p = 0.018$). Moreover, AST, ALT, γGTP, and AFP were significantly higher, and platelets, albumin, and cholesterol were significantly lower in the HCC group. Different predicted amino acids in the core and NS5A regions, between the two groups, are demonstrated in Fig. 2a. The ratio of the core a.a. 70Q (glutamine) or H (histidine) to R (arginine) was significantly higher with the existence of HCC as demonstrated in Fig. 2a (left panel). On the other hand, evident correlations were not confirmed between mutations in NS5A and disease progression (Fig. 2a, right panel). The ratio of Q or H