

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

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

Statistical analysis

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

Results

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

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

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

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

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

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

Factors associated with response by multivariate logistic regression analysis

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

Research Article

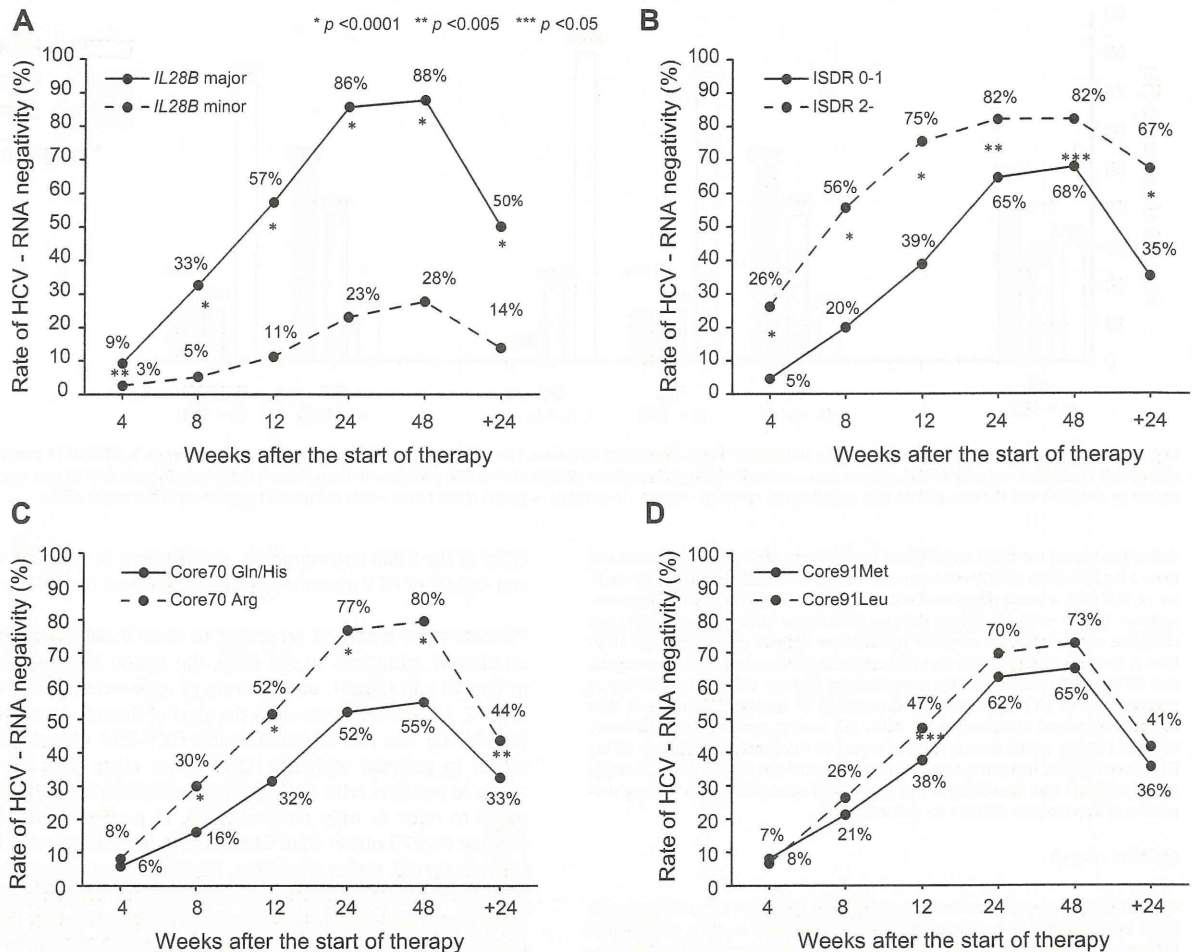


Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified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

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

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

Factors associated with the *IL28B* polymorphism

Patients with the *IL28B* minor allele had significantly higher serum level of gamma-glutamyltransferase (GGT) and a higher

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

Data mining analysis

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

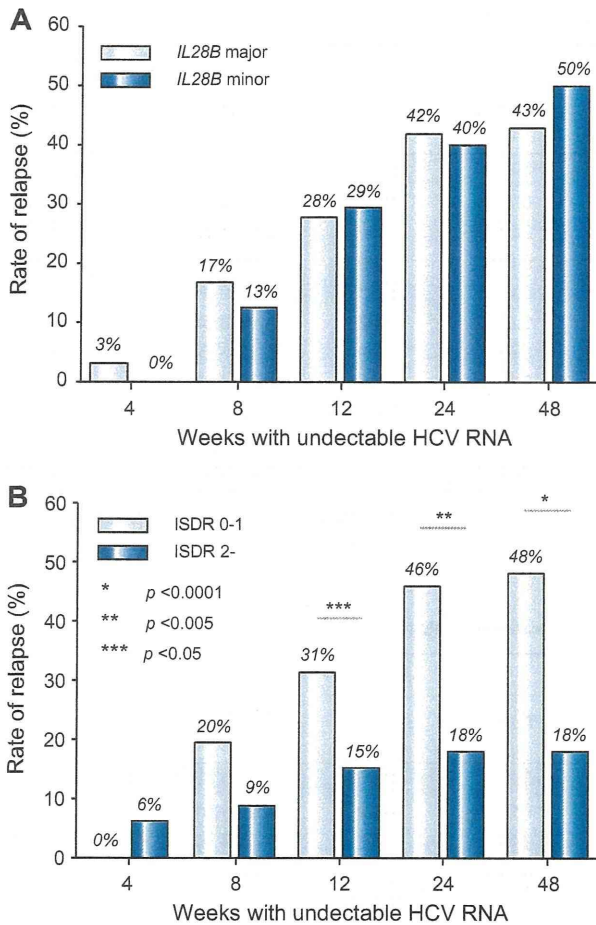


Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele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

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 ($\geq 140 \times 10^9/L$): 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-

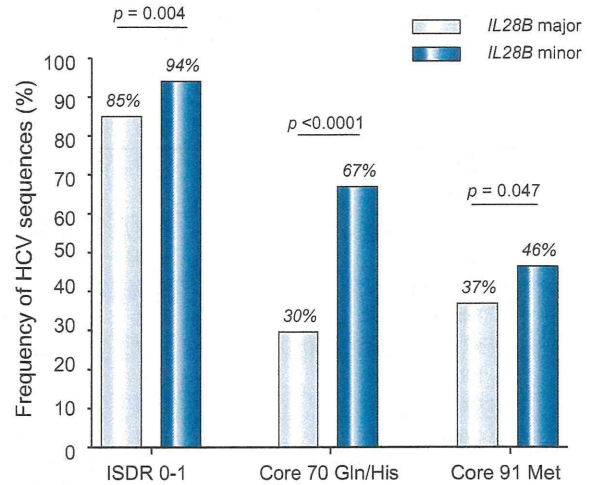


Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met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

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 ($\geq 140 \times 10^9/L$).

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 nitazoxanide [30], are in clinical trials and may become available in the near future, a pre-treatment prediction of the likelihood of response may be helpful for patients and physicians, to support clinical decisions about whether to begin the current standard of care or whether to wait for emerging therapies. This study revealed that the *IL28B* polymorphism was the overwhelming predictor of NVR and is independent of host factors and viral sequences reported previously. The *IL28B* encodes a protein also known as IFN-lambda 3, which is thought to suppress the replication of various viruses including HCV [31,32]. The results of the current study and the findings of the GWAS studies [6-9] may provide the rationale for developing diagnostic testing or an IFN-lambda based therapy for chronic hepatitis C in the future.

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: $\geq 600,000$ 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
<i>IL28B</i> : 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	1.01-1.16	<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
<i>IL28B</i> : 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.

	<i>IL28B</i> major allele n = 345	<i>IL28B</i> 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 RNA 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 group of patients. Using this model, we can rapidly develop an

Research Article

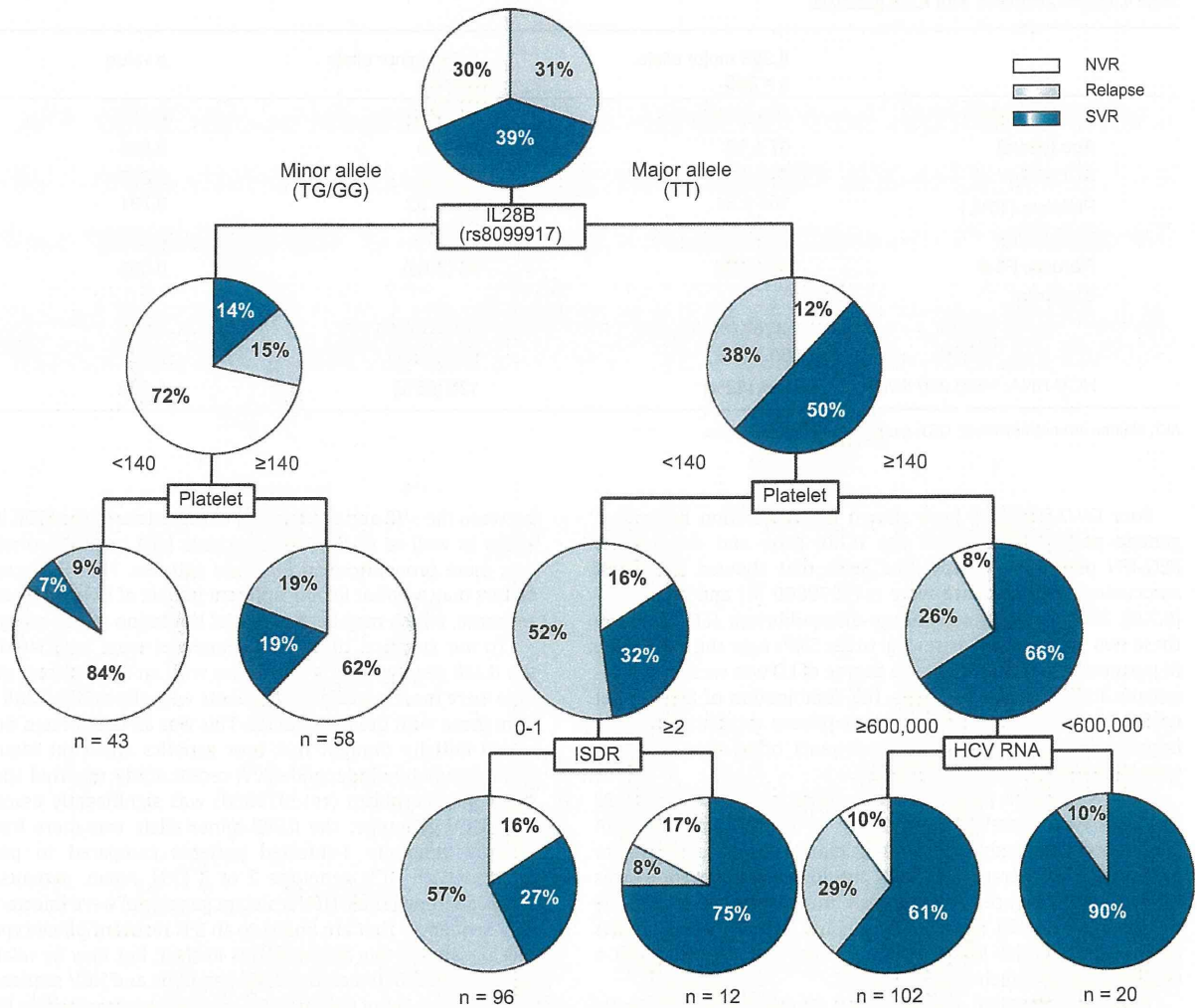


Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown.

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 signif-

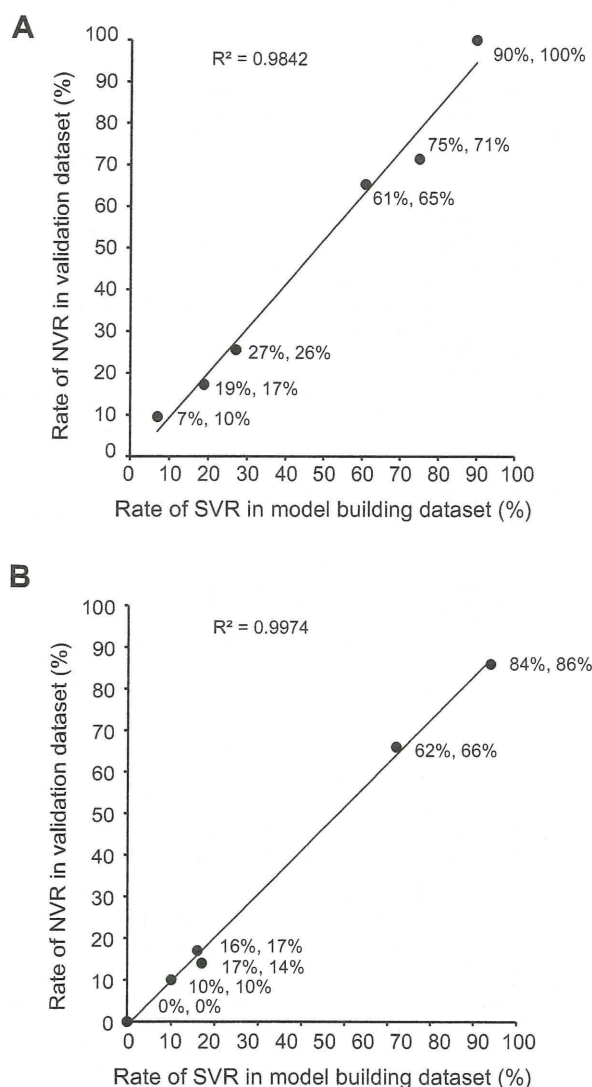


Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and 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: $r^2 = 0.98-0.99$).

icantly associated with *IL28B* genotype in the present study (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.

Conflict of interest

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

Financial support

This study was supported by a grant-in-aid from the Ministry of Health, Labor and Welfare, Japan, (H19-kannen-013), (H20-kannen-006).

References

- [1] Ray Kim W. Global epidemiology and burden of hepatitis C. *Microbes Infect* 2002;4 (12):1219–1225.
- [2] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncalves 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, 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;41:1105–1109.
- [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;41:1100–1104.
- [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, Iai 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.

The rs8099917 Polymorphism, When Determined by a Suitable Genotyping Method, Is a Better Predictor for Response to Pegylated Alpha Interferon/Ribavirin Therapy in Japanese Patients than Other Single Nucleotide Polymorphisms Associated with Interleukin-28B[†]

Kiyoaki Ito,^{1,‡} Katsuya Higami,^{1,‡} Naohiko Masaki,¹ Masaya Sugiyama,¹ Motokazu Mukaide,¹ Hiroaki Saito,¹ Yoshihiko Aoki,¹ Yo Sato,¹ Masatoshi Imamura,¹ Kazumoto Murata,¹ Hideyuki Nomura,² Shuhei Hige,³ Hiroshi Adachi,⁴ Keisuke Hino,⁵ Hiroshi Yatsushashi,⁶ Etsuro Orito,⁷ Satomi Kani,⁸ Yasuhito Tanaka,⁸ and Masashi Mizokami^{1*}

The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Japan¹; The Center for Liver Diseases, Shin-Kokura Hospital, Kitakyushu, Japan²; Department of Internal Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan³; Department of Virology and Liver Unit, Tonami General Hospital, Tonami, Japan⁴; Division of Gastroenterology, Department of Medicine, Kawasaki Medical School, Okayama, Japan⁵; Clinical Research Center, NHO Nagasaki Medical Center, Nagasaki, Japan⁶; Department of Gastroenterology and Hepatology, Nagoya Daini Red Cross Hospital, Nagoya, Japan⁷; and Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan⁸

Received 22 October 2010/Returned for modification 4 January 2011/Accepted 28 February 2011

We focused on determining the most accurate and convenient genotyping methods and most appropriate single nucleotide polymorphism (SNP) among four such polymorphisms associated with interleukin-28B (IL-28B) in order to design tailor-made therapy for patients with chronic hepatitis C virus (HCV) patients. First, five different methods (direct sequencing, high-resolution melting analysis [HRM], hybridization probe [HP], the InvaderPlus assay [Invader], and the TaqMan SNP genotyping assay [TaqMan]) were developed for genotyping four SNPs (rs11881222, rs8103142, rs8099917, and rs12979860) associated with IL-28B, and their accuracies were compared for 292 Japanese patients. Next, the four SNPs associated with IL-28B were genotyped by Invader for 416 additional Japanese patients, and the response to pegylated interferon/ribavirin (PEG-IFN/RBV) treatment was evaluated when the four SNPs were not in linkage disequilibrium (LD). HRM failed to genotype one of the four SNPs in five patients. In 2 of 287 patients, the results of genotyping rs8099917 by direct sequencing differed from the results of the other three methods. The HP, TaqMan, and Invader methods were accurate for determination of the SNPs associated with IL-28B. In 10 of the 708 (1.4%) patients, the four SNPs were not in LD. Eight of nine (88.9%) patients whose rs8099917 was homozygous for the major allele were virological responders, even though one or more of the other SNPs were heterozygous. The HP, TaqMan, and Invader methods were suitable to determine the SNPs associated with IL-28B. The rs8099917 polymorphism should be the best predictor for the response to the PEG-IFN/RBV treatment among Japanese chronic hepatitis C patients.

Hepatitis C virus (HCV) infection is a global health problem, with worldwide estimates of 120 to 130 million carriers (7). Chronic HCV infection can lead to progressive liver disease, resulting in cirrhosis and complications, including decompensated liver disease and hepatocellular carcinoma (25). The current standard of care treatment for suitable patients with chronic HCV infection consists of pegylated alpha 2a or 2b interferon (PEG-IFN) given by injection in combination with

oral ribavirin (RBV), for 24 or 48 weeks, dependent on HCV genotype. Large-scale treatment programs in the United States and Europe showed that 42 to 52% of patients with HCV genotype 1 achieved a sustained virological response (SVR) (3, 8, 13), and similar results were found in Japan. This treatment is associated with well-described side effects (such as a flu-like syndrome, hematologic abnormalities, and neuropsychiatric events) resulting in reduced compliance and fewer patients completing treatment (2). It is valuable to predict an individual's response before treatment with PEG-IFN/RBV to avoid these side effects, as well as to reduce the treatment cost. The HCV genotype, in particular, is used to predict the response: patients with HCV genotype 2 or 3 have a relatively high rate of SVR (70 to 80%) with 24 weeks of treatment, whereas those infected with genotype 1 have a much lower rate of SVR despite 48 weeks of treatment (8).

Recently, we reported from genome-wide association stud-

* Corresponding author. Mailing address: The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, 1-7-1, Konodai, Ichikawa 272-8516, Japan. Phone: 81-47-372-3501. Fax: 81-47-375-4766. E-mail: mmizokami@hospk.ncgm.go.jp.

‡ These authors contributed equally to the manuscript.

† Supplemental material for this article may be found at <http://jcm.asm.org/>.

‡ Published ahead of print on 9 March 2011.

TABLE 1. Characteristics of the patients examined

Parameter	Result for:	
	1st stage (n = 292)	2nd stage (n = 416)
Age (yr)	57.2 ± 10.2	56.6 ± 10.9
No. of patients male/female	145/147	194/222
No. (%) of patients in institution ^a :		
1	18 (6.2)	0 (0)
2	178 (61.0)	0 (0)
3	57 (19.5)	0 (0)
4	39 (13.3)	0 (0)
5	0 (0)	249 (59.9)
6	0 (0)	94 (22.6)
7	0 (0)	52 (12.5)
8	0 (0)	21 (5.0)

^a Institutions: 1, The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine; 2, The Center for Liver Diseases, Shin-Kokura Hospital, Kitakyushu; 3, Tonami General Hospital, Tonami; 4, Department of Internal Medicine, Virology and Liver Unit, Hokkaido University Graduate School of Medicine, Sapporo; 5, Clinical Research Center, NHO Nagasaki Medical Center, Nagasaki; 6, Nagoya City University Graduate School of Medical Sciences, Nagoya; 7, Department of Gastroenterology and Hepatology, Nagoya Daini Red Cross Hospital; and 8, Division of Gastroenterology, Department of Medicine, Kawasaki Medical School, Okayama.

ies (GWAS) that several highly correlated common single nucleotide polymorphisms (SNPs), located in the vicinity of the lambda 3 interferon (IFN- λ 3), coded for by the interleukin-28B (IL-28B) gene on chromosome 19, are implicated in non-virological response (NVR) to PEG-IFN/RBV among patients with HCV genotype 1 (21). At almost exactly the same time as our report, the association between response to PEG-IFN/

RBV and SNPs associated with IL-28B was reported from the results of GWAS by two other groups (6, 19). Determination of these SNPs associated with IL-28B before PEG-IFN/RBV treatment will provide extremely valuable information, because the patients predicted as showing NVR to PEG-IFN/RBV treatment could avoid the treatment. There are two questions to be asked before using these SNPs in clinical practice: (i) which methods for genotyping these SNPs are efficient, and (ii) which SNP is most informative in cases where the SNPs are not in linkage disequilibrium (LD)? We have developed five different methods for detecting the SNPs associated with IL-28B and compared their accuracies to establish the most efficient genotyping method. The response to PEG-IFN/RBV treatment was evaluated, when the SNPs associated with IL-28B were not in LD, to determine the best SNP to predict the response to PEG-IFN/RBV treatment.

MATERIALS AND METHODS

Study population. Samples were obtained from 708 Japanese chronic hepatitis C patients and divided into groups of 292 patients (145 males and 147 females; mean age, 57.2 years) and 416 patients (194 males and 222 females; mean age, 56.6 years) for the first and second stages (Table 1). In the first stage, we focused on analyzing the effective methods for determining the genotypes of four SNPs (rs11881222, rs8103142, rs12979860, and rs8099917) associated with IL-28B (Fig. 1A). Figure 2 shows the locations of these four SNPs in chromosome 19; rs11881222 and rs8103142 are located in the IL-28B gene, and rs12979860 and rs8099917 are located downstream from the IL-28B gene. The results of genotyping the four SNPs by five different methods, described below, were compared and evaluated for consistency. For this first stage, the 292 chronic hepatitis C patients were recruited from the National Center for Global Health and Medicine, Hokkaido University Hospital, Tonami General Hospital, and Shin-Kokura Hospital in Japan (Table 1). From the results of the first stage, the InvaderPlus assay was chosen as one of the best methods to determine the genotypes of the four SNPs associated with IL-28B and was used for genotyping 416 patients (Fig.

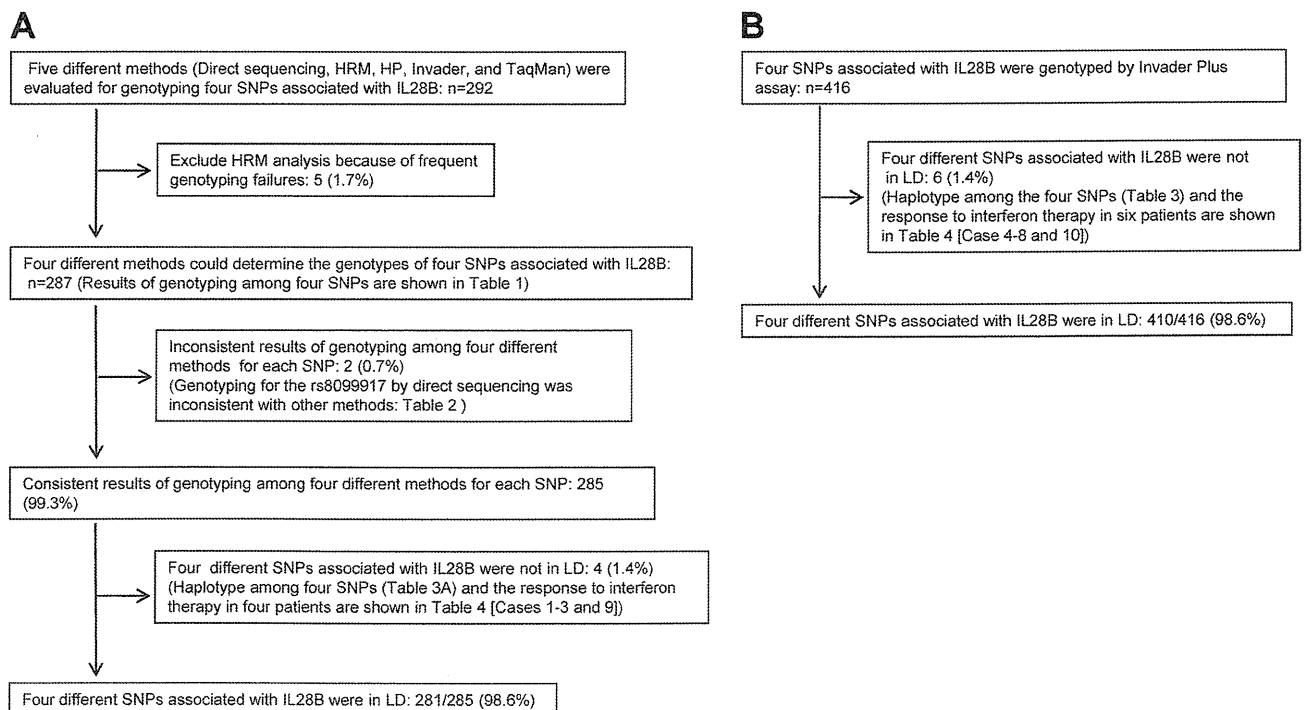


FIG. 1. Schema for the flowchart of the examinations.

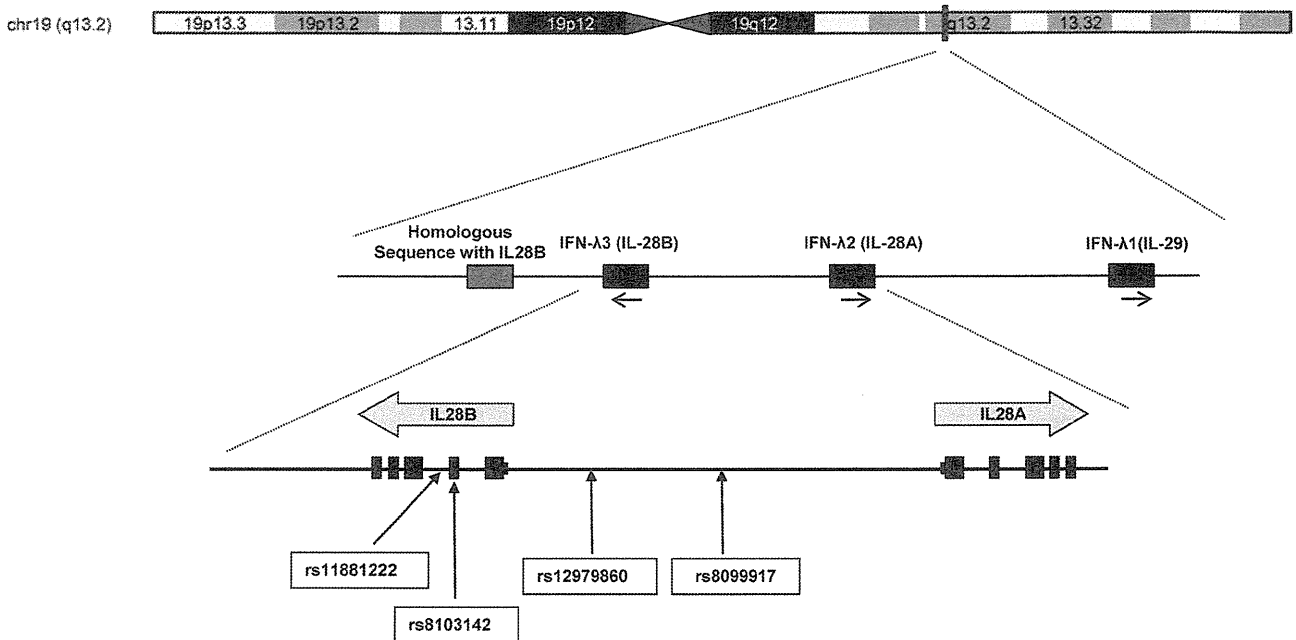


FIG. 2. Location of interferon lambda genes and the four SNPs (rs11881222, rs8103142, rs12979860, and rs8099917) associated with IL-28B. chr19, chromosome 19.

1B), recruited from NHO Nagasaki Medical Center, Nagoya City University Hospital, Nagoya Daini Red Cross Hospital, and Kawasaki Medical University Hospital in Japan, in the second stage (Table 1). We then focused on 10 patients whose four SNPs were found in the first and second stages not to be in LD and investigated the response to PEG-IFN/RBV treatment in detail for these patients. Informed consent was obtained from each patient who participated in the study. This study was conducted in accordance with provisions of the Declaration of Helsinki.

Definition of treatment responses. Nonvirological response (NVR) was defined as less than a 2-log-unit decline in the serum level of HCV RNA from the pretreatment baseline value within the first 12 weeks or detectable viremia 24 weeks after treatment. Virological response (VR) was defined in this study as the achievement of sustained VR (SVR) or transient VR (TVR); SVR was defined as undetectable HCV RNA in serum 6 months after the end of treatment, whereas TVR was defined as a reappearance of HCV RNA in serum after treatment was discontinued in a patient who had undetectable HCV RNA during

the therapy or had achieved a more than 2-log-unit decline within the first 12 weeks after treatment.

DNA extraction. Whole blood was collected from all participants and centrifuged to separate the buffy coat. Genomic DNA was extracted from the buffy coat with Genomix (Talent SRL, Italy).

Five different genotyping methods. Four SNPs (rs11881222, rs8103142, rs12979860, and rs8099917) (Fig. 2) were determined in 292 patients by five different genotyping methods. We developed the five methods (direct sequencing, high-resolution melting analysis [HRM], hybridization probe (HP), Invader-Plus assay (Invader), and the TaqMan SNP genotyping assay (TaqMan) to determine the genotypes of the rs11881222 and rs8103142 polymorphisms. We also developed four different methods (direct sequencing, HRM, HP, and Invader) to determine the genotypes of the rs12979860 and rs8099917 polymorphisms. The genotype of rs12979860 was also determined by the TaqMan genotyping method developed by Duke University, and the genotype of rs8099917 was also determined with the TaqMan predesigned SNP genotyping assay. Figures 3,

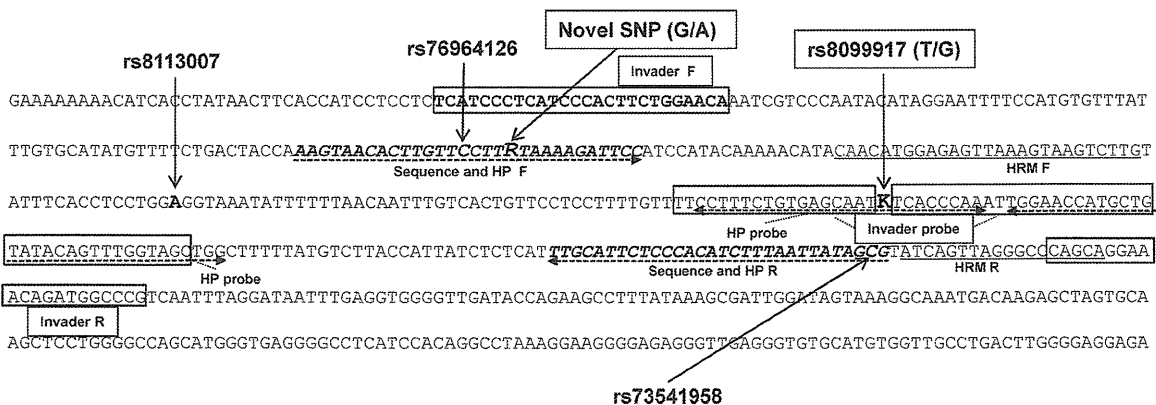


FIG. 3. The nucleotide sequence around rs8099917 is shown. Primers and probes for four different methods (Sequence, direct sequencing; HRM, high-resolution melting analysis; HP, hybridization probe; Invader, InvaderPlus assay) to determine rs8099917 polymorphism are shown. F, forward primer; R, reverse primer.

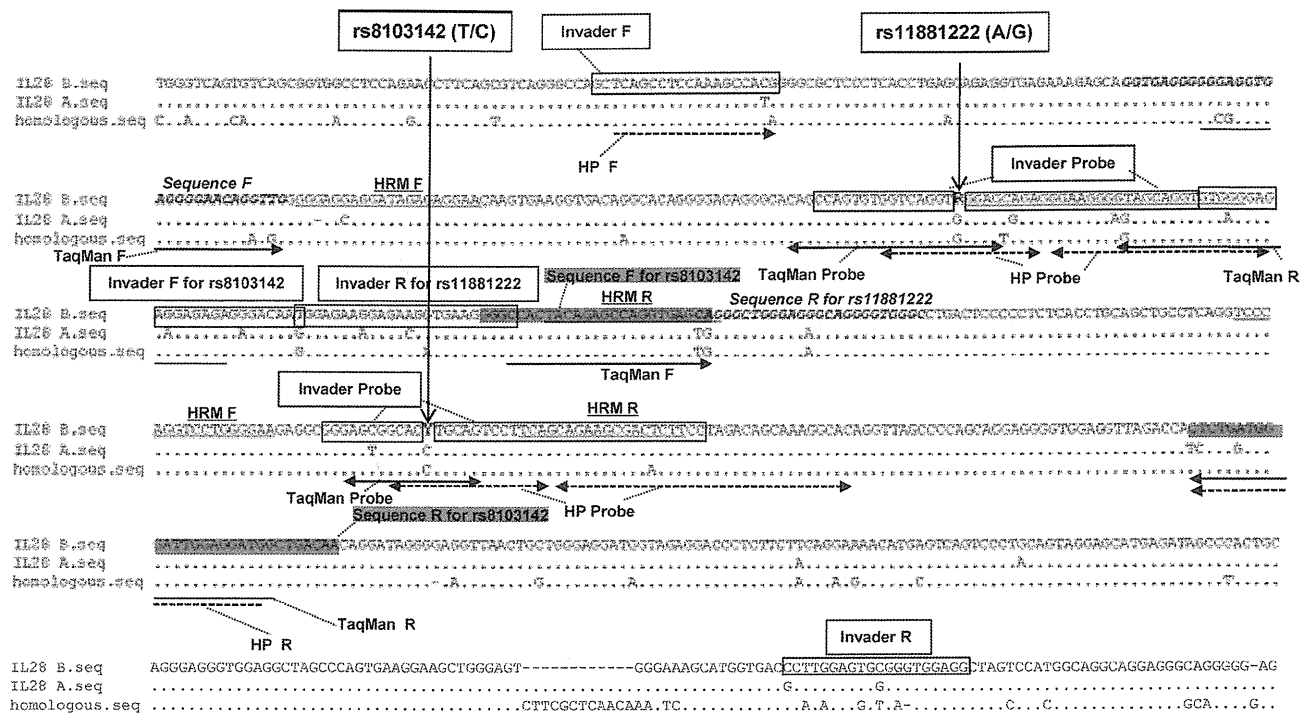


FIG. 4. The nucleotide sequence around rs11881222 and rs103142 is shown. Primers and probes for five different methods (Sequence, direct sequencing; HRM, high-resolution melting analysis; HP, hybridization probe; Invader, InvaderPlus assay; TaqMan, TaqMan assay) to determine rs11881222 and rs103142 polymorphisms are shown. F, forward primer; R, reverse primer.

4, and 5 show the primers and probes for each genotyping method. Because the sequence of IL-28B is very similar to those of IL-28A, IL-29, and a homologous sequence upstream of IL-28B, we had to design the primers and probe for each method to distinguish IL-28B from the other sequences. First, primers were designed with Visual OMP Nucleic Acid software, and then we confirmed that the candidate primers should not amplify sequences other than the target region by using UCSC Genome Browser. Next, we confirmed that the amplicon was resolved as a single band, when the PCR products amplified by the primers under evaluation were electrophoresed. Finally, we had to optimize each set of primers and probe for each method (Fig. 3 to 5; see the table in the supplemental material).

Direct sequencing. PCR was carried out with 12.5 μ l AmpliTaq Gold 360 master mix (Applied Biosystems), 10 pmol of each primer, and 10 ng of genomic DNA under the following thermal cycler conditions: stage 1, 94°C for 5 min; stage 2, 94°C for 30 s, 65°C for 30 s, 72°C for 45 s, for a total of 35 cycles; and stage 3, 72°C for 7 min. For sequencing, 1.0 μ l of the PCR products was incubated with the use of a BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). After ethanol purification, the reaction products were applied to the Applied Biosystems 3130xl DNA analyzer.

HRM analysis. HRM analysis was performed on a LightCycler 480 (LC480; Roche Diagnostics) as described previously (5, 15, 24). We designed pairs of primers flanking each SNP (Fig. 3 to 5) to amplify DNA fragments shorter than 200 bp. PCR was performed in a 20- μ l volume containing 10 μ l LightCycler 480 high-resolution melting master mix (Roche Applied Science), 4 pmol of each primer, and 10 ng genomic DNA. The cycling conditions were as follows: SYBR green I detection format, 1 cycle of 95°C for 10 min and 50 cycles of 95°C for 5 s, 60°C for 10 s, and 72°C for 20 s, followed by an HRM step of 95°C for 1 min, 40°C for 1 min, and 74°C for 5 s and continuous acquisition to 90°C at 25 acquisitions per 1°C. HRM data were analyzed with Gene Scanning software (Roche Diagnostics).

Hybridization probe. We designed oligonucleotide primers and hybridization probes for the four SNPs (Fig. 3 to 5). All assays were performed with the LC480 as described previously (4, 18). The amplification mixture consisted of 4 μ l of 5 \times reaction mixture (LightCycler 480 genotyping master; Roche Diagnostics), 5 pmol of each oligonucleotide primer, 3.2 pmol of each oligonucleotide probe, and 10 ng of template DNA in a final volume of 20 μ l. Samples were amplified

as follows: 45 cycles of denaturation at 95°C for 10 s, annealing at 60°C for 10 s, and extension at 72°C for 20 s. The generation of target amplicons for each sample was monitored between the annealing and elongation steps at 610 and 640 nm. Samples positive for target genes were identified by the instrument at the cycle number where the fluorescence attributable to the target sequences exceeded that measured as background. Those scored as positive by the instrument were confirmed by visual inspection of the graphical plot (cycle number versus fluorescence value) generated by the instrument.

InvaderPlus assay. The InvaderPlus assay, which combines PCR and the Invader reaction (11, 12), was performed with the LC480. The enzymes used in InvaderPlus are native *Taq* polymerase (Promega Corporation, Madison, WI) and Cleavase enzyme (Third Wave Technologies, Madison, WI). The reaction is configured to use PCR primers with a melting temperature (T_m) of 72°C and Invader detection probe with a target-specific T_m of 63°C. The Invader oligonucleotide overlaps the probe by one nucleotide, forming at 63°C an overlap flap substrate for the Cleavase enzyme. The first step of InvaderPlus is PCR target amplification, in which the reaction is subjected to 18 cycles of a denaturation step (95°C for 15 s) and hybridization and extension steps (70°C for 1 min). At the end of PCR cycling, the reaction mixture is incubated at 99°C for 10 min to inactivate the *Taq* polymerase. Next, the reaction temperature is lowered to 63°C for 15 to 30 min to permit the hybridization of the probe oligonucleotide and the formation of the overlap flap structure. Data were analyzed by endpoint genotyping software (Roche Diagnostics).

TaqMan assay. The rs8099917 polymorphism was determined by using TaqMan predesigned SNP genotyping assays, as recommended by the manufacturer. The TaqMan assay for determination of the genotype of rs12979860 was kindly provided by David B. Goldstein at Duke University. We designed primers and probes for TaqMan genotyping assays for the other two SNPs. Each genomic DNA sample (20 ng) was amplified with TaqMan universal PCR master mix reagent (Applied Biosystems, Foster City, CA) combined with the specific TaqMan SNP genotyping assay mixture, corresponding to the SNP to be genotyped. The assays were carried out using the LC480 (Roche Applied Science) and the following conditions: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Data were analyzed by endpoint genotyping software (Roche Diagnostics).

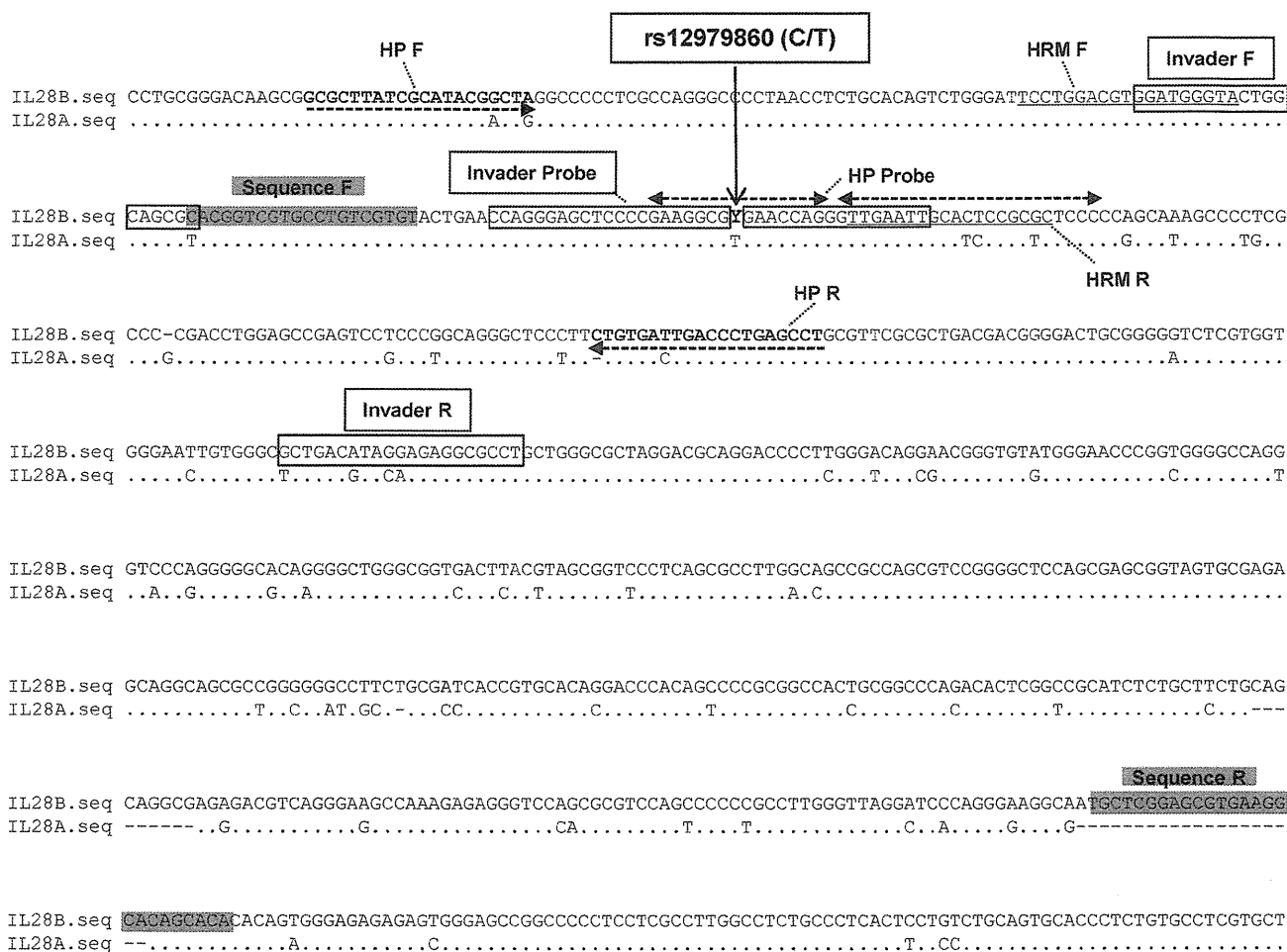


FIG. 5. The nucleotide sequence around rs12979860 is shown. Primers and probes for four different methods (Sequence, direct sequencing; HRM, high-resolution melting analysis; HP, hybridization probe; Invader, InvaderPlus assay) to determine rs12979860 are shown. F, forward primer; R, reverse primer.

RESULTS

Genotyping for four SNPs associated with IL-28B was unsuccessful by HRM in five cases. Figure 1A shows the patients' flowchart of the first stage. Genotyping of four SNPs (rs11881222, rs8103142, rs12979860, and rs8099917) was attempted by five different methods (direct sequencing, HRM, HP, Invader, and TaqMan) for 292 patients. In five cases, one of the four SNPs could not be genotyped by HRM. Therefore, we excluded the HRM method from further study. The genotyping failures by HRM involved two cases for rs11881222, two cases for rs8103142, and one case for rs8099917.

Consistencies of four different methods to determine genotypes for four SNPs associated with IL-28B. Consistencies among the results of genotyping by the remaining four methods were 100%, except for the results for rs8099917 (Table 2). For rs8099917, the results determined by direct sequencing were inconsistent with the other three methods in two cases (Tables 2 and 3). The HP, TaqMan, and Invader methods were accurate and reliable for genotyping the four SNPs associated with IL-28B. Invader was chosen for genotyping in the second stage, because the analysis time was the shortest and the sen-

TABLE 2. Determination of four SNPs associated with IL-28B by four different methods^a

SNP	Genotype	No. (%) of cases with genotype by:			
		Direct sequencing	HP	Invader	TaqMan
rs11881222	AA	199 (69.3)	199 (69.3)	199 (69.3)	199 (69.3)
	AG	84 (29.3)	84 (29.3)	84 (29.3)	84 (29.3)
	GG	4 (1.4)	4 (1.4)	4 (1.4)	4 (1.4)
rs8103142	TT	199 (69.3)	199 (69.3)	199 (69.3)	199 (69.3)
	TC	84 (29.3)	84 (29.3)	84 (29.3)	84 (29.3)
	CC	4 (1.4)	4 (1.4)	4 (1.4)	4 (1.4)
rs12979860	CC	198 (69.0)	198 (69.0)	198 (69.0)	198 (69.0)
	CT	85 (29.6)	85 (29.6)	85 (29.6)	85 (29.6)
	TT	4 (1.4)	4 (1.4)	4 (1.4)	4 (1.4)
rs8099917	TT	204 (71.1)	202 (70.4)	202 (70.4)	202 (70.4)
	TG	79 (27.5)	81 (28.2)	81 (28.2)	81 (28.2)
	GG	4 (1.4)	4 (1.4)	4 (1.4)	4 (1.4)

^a There was 100% consistency for rs11881222, rs8103142, and rs12979860, and there was 99.3% consistency for rs8099917.

TABLE 3. Inconsistency in two cases between rs8099917 genotyping by direct sequencing and three other methods

Case no.	rs8099917 genotype by ^a :			
	Direct sequencing	HP	Invader	TaqMan
1	T/T	T/G	T/G	T/G
2	T/T	T/G	T/G	T/G

^a Homozygous genotypes are highlighted in boldface.

sitivity was the greatest of the three methods (HP, TaqMan, and Invader), as reported previously (20).

Genotyping error for rs8099917 by direct sequencing due to novel SNP. In two cases, the results of genotyping for rs8099917 by direct sequencing were inconsistent with the results by the other methods (Table 3). Direct sequencing determined the genotype for rs8099917 as T/T in cases 1 and 2; however, the other three genotyping methods (HP, Invader, and TaqMan) determined the genotypes for rs8099917 as T/G in both cases. Further study using alternative primers for direct sequencing revealed that the correct genotypes were T/G and revealed a novel minor SNP present in the forward primer binding site in these two cases (data on file) and which interfered with the PCR amplification step (Fig. 3).

Distribution of haplotypes among four SNPs associated with IL-28B. In the first stage, the four SNPs were in LD in 281 (98.6%) of 285 cases and not in LD in the remaining 4 (1.4%). The first stage revealed five different haplotypes (no. 1 to 5 in Table 4). In haplotypes 1 to 3, the four SNPs were in LD (haplotype 1, homozygous of the major allele among 4 SNPs; $n = 198$ [69.5%]; haplotype 2, heterozygous among 4 SNPs; $n = 79$ [27.7%]; and haplotype 3, homozygous of the minor allele among 4 SNPs; $n = 4$ [1.4%]). In haplotype 4 (3 cases) rs11881222, rs8103142, rs12979860, and rs8099917 were AG, TC, CT, and TT, respectively. In haplotype 5 (one case), rs11881222, rs8103142, rs12979860, and rs8099917 were AA, TT, CT, and TT, respectively. Genotyping by the Invader method of the four SNPs associated with IL-28B in 416 patients in the second stage revealed that the four SNPs were not in LD in 6 cases (1.4%) (Table 4). A total of 410 (98.6%) of 416 cases were in LD for the four different SNPs. The second stage showed six different haplotypes (haplotypes 1 to 4, 6, and 7). Haplotypes 1 to 4 were detected in the first stage, but haplotypes 6 and 7 were not. The distribution of haplotypes was such that haplotypes 1, 2, 3, and 4 were found in 294 (70.7%), 110 (26.5%), 6 (1.4%), and 4 (1.0%) cases, respectively. In haplotype 6 (one case), rs11881222, rs8103142, rs12979860, and rs8099917 were AG, TT, CC, and TT, respectively. In haplotype 7 (one case), rs11881222, rs8103142, rs12979860, and rs8099917 were AA, TT, CT, and TG, respectively.

Response to PEG-IFN/RBV treatment in 10 cases in which the four SNPs associated with IL-28B were not in LD. In 7 (cases 1 to 7 [70%]) of the 10 cases where the four SNPs were not in LD, the haplotype was such that rs11881222, rs8103142, rs12979860, and rs8099917 were AG, TC, CT, and TT, respectively (Table 5). In nine cases (cases 1 to 9), rs8099917 was homozygous for the major allele, while one or more of the other SNPs were heterozygous. Eight (cases 1 to 8) of these

TABLE 4. Distribution of haplotypes among four SNPs associated with IL-28B in stages 1 and 2

Stage	Haplotype no.	Genotype for SNP:				No. (%) of cases with haplotype shown
		rs11881222	rs8103142	rs12979860	rs8099917	
1	1	AA	TT	CC	TT	198 (69.5)
	2	AG	TC	CT	TG	79 (27.7)
	3	GG	CC	TT	GG	4 (1.4)
	4	AG	TC	CT	TT	3 (1.0)
	5	AA	TT	CT	TT	1 (0.4)
2	1	AA	TT	CC	TT	294 (70.7)
	2	AG	TC	CT	TG	110 (26.5)
	3	GG	CC	TT	GG	6 (1.4)
	4	AG	TC	CT	TT	4 (1.0)
	6	AG	TT	CC	TT	1 (0.2)
	7	AA	TT	CT	TG	1 (0.2)

nine cases were viral responders who met the following criteria: HCV had disappeared during therapy, or HCV RNA had decreased more than 2 log copies/ml before 12 weeks after beginning of therapy, although some cases were under treatment or before determination of the final response to PEG-IFN/RBV. Case 9 was NVR due to poor adherence of PEG-IFN (<50% dose), even though rs8099917 was homozygous of the major allele. The haplotype of case 9 showed that rs11881222, rs8103142, rs12979860, and rs8099917 were AA, TT, CT, and TG, respectively. NVR in case 10 was reasonable from the genotypes of rs8099917 and rs12979860, because they were heterozygous, although rs11881222 and rs8103142 were homozygous for the major allele.

DISCUSSION

The relationship between SNPs associated with IL-28B and the response to PEG-IFN/RBV therapy for chronic hepatitis C was found by SNP array, using GWAS technology, by three different groups throughout the world, including our own, in 2009 (6, 19, 21). Following these reports, many studies have confirmed the association between the response to PEG-IFN/RBV and SNPs associated with IL-28B (14, 16). Therefore, it is obvious that these SNPs may be valuable for predicting the response to PEG-IFN/RBV therapy. Recently, it was reported that various SNPs were associated with development of disease and response to therapy and correlated with adverse effects. Several SNPs, such as the UGT1A1 polymorphism for the treatment with irinotecan (1, 17), have already been exploited in clinical practice to avoid severe adverse effects. These tailor-made therapies are expected to become more common in clinical practice in the near future (9). The next step toward tailor-made therapy for PEG-IFN/RBV therapy against chronic hepatitis C involved the development of simple, accurate, and inexpensive methods to determine the genotype of SNPs and determination of the best SNP where the four SNPs associated with IL-28B were not in LD, so that they may be applied in clinical practice.

Genotyping of IL-28B SNPs is quite different from other SNPs, because the sequence of IL-28B is very similar to those of IL-28A, IL-29, and an additional homologous sequence upstream of IL-28B (Fig. 2). We had to design primers and probes for each method to distinguish IL-28B specifically. We

TABLE 5. Clinical characteristics of 10 cases in which the SNPs associated with IL-28B were not in LD

Case no. ^a	SNP of IL-28B ^b				Age (yr)	Gender	Genotype	Viral titer	Final response to PEG-IFN/RBV	VR or NVR	Period of disappearance of HCV
	rs11881222	rs8103142	rs12979860	rs8099917							
1	A/G	T/C	C/T	T/T	64	Female	1b	6.5	TR	VR	4 wk
2	A/G	T/C	C/T	T/T	72	Male	1b	2.9	SVR	VR	4 wk
3	A/G	T/C	C/T	T/T	64	Male	1b	7	ND ^c	VR	8 wk
4	A/G	T/C	C/T	T/T	51	Female	1b	7.2	Under treatment	VR	3.6 log units down after 12 wk
5	A/G	T/C	C/T	T/T	60	Female	2	5.8	Under treatment	VR	12 wk
6	A/G	T/C	C/T	T/T	56	Female	1b	5.9	Under treatment	VR	2.0 log units down after 2 wk
7	A/G	T/C	C/T	T/T	62	Male	1b	5.4	SVR	VR	4 wk
8	A/G	T/T	C/C	T/T	58	Male	1b	6.2	TR	VR	12 wk
9	A/A	T/T	C/T	T/T	68	Male	1b	7	NVR	NVR	— ^d
10	A/A	T/T	C/T	T/G	48	Female	1b	6	NVR	NVR	—

^a All cases shown were treated with PEG-IFN/RBV.

^b Homozygous genotypes are highlighted in boldface.

^c ND, not determined. The final response to PEG-IFN/RBV was not determined in this patient because 6 months had not passed after the end of treatment.

^d —, HCV did not disappear.

think that the results in this paper are especially applicable to IL-28B genotyping. In this study, only HRM failed to determine the genotype of SNPs associated with IL-28B. The reason HRM failed more frequently than the other genotyping methods is attributable to the characteristics of this specific method. Because HRM determines the genotype of each SNP by distinguishing the melting curve of an amplicon of around 200 bp, it may tend to be influenced by another SNP. As a matter of fact, minor SNPs around rs8099917 were found in cases of genotyping failure by HRM (data not shown). Although this specific characteristic of the HRM method is useful for detecting novel mutations or SNPs, it is not suitable for determination of the genotype of SNPs associated with IL-28B.

Direct sequencing erroneously reported the T/G genotype as T/T for the rs8099917 polymorphism. We found that the cause of this genotyping error was a novel rare SNP in the forward primer binding site used for amplification and direct sequencing (data on file). Because this novel SNP was not registered as an SNP in the NCBI database, the primer was designed at this site. Since the novel SNP correlated with the rs8099917 polymorphism in LD, adenine for the novel SNP is present on the same allele as guanine in the rs8099917 polymorphism. Therefore, the forward PCR primer (AAGTAACACTTGTCCTT GTAAAAGATTCC) could not anneal to the binding site, which was changed from guanine (G) to adenine (A) at the underlined nucleotide position: only the allele which has T at the rs8099917 was amplified, the genotype was determined as T/T. Rare sequence variations not registered in the database, might be present in the primer binding sites for amplification and might be the cause of erroneous direct sequencing. Ikegawa et al. reported that annealing efficiency in direct sequencing led to the mistyping of an SNP (10). Although our results in this paper are especially applicable to IL-28B genotyping, it should be recognized that allele-dependent PCR amplification and erroneous typing can occur when SNPs are genotyped by a PCR-based approach. Should SNPs associated with IL-28B be found not to be in LD, it would be preferable to confirm the genotype by another method.

In 10 cases, four SNPs associated with IL-28B were not in LD. In seven (70%) of the 10 cases, the haplotype showed that

rs11881222, rs8103142, rs12979860, and rs8099917 were AG, TC, CT, and TT, respectively. Only the rs8099917 polymorphism differed frequently from the other three SNPs. The reason for the high frequency of this haplotype is thought to be attributable to the location of these SNPs. The location of rs8099917 is downstream and quite far from the two SNPs (rs11881222 and rs8103142) in the IL-28B gene (Fig. 2). The SNPs rs11881222 and rs8103142 were almost perfectly in LD, because they are located close to each other.

It is well described that homozygosity for the major allele of SNPs associated with IL-28B is correlated with a better response to PEG-IFN/RBV treatment, and minor allele-positive patients are poor responders. However, the response to PEG-IFN/RBV remains unknown when several SNPs associated with IL-28B are not in LD. Because cases in which the SNPs are not in LD are quite rare, it was thought to be difficult to study such cases. In this study, 10 (1.4%) of 708 patients showed haplotypes in which the four SNPs were not in LD. We focused on the response to PEG-IFN/RBV therapy in these 10 cases (Table 5). We evaluated the response to PEG-IFN/RBV treatment from the viewpoint of virological response, because some patients had not completed their PEG-IFN/RBV treatment. (Case 3 was before determination for the final response after finishing the treatment, and cases 4 to 6 were under treatment.)

Thomas et al. reported that allele frequencies for rs12979860 varied among racial and ethnic groups (23). Indeed, the observation that the major allele is less frequent among individuals of African descent than those of European descent might explain the observed discrepancy in the frequencies of viral clearance in these two ethnic groups, where clearance occurs in 36.4% of HCV infections in individuals of non-African ancestry, but in only 9.3% of infections in individuals of African ancestry (22). We have recruited only Japanese chronic hepatitis C patients for this study. Since the distribution of haplotype and response to PEG-IFN/RBV treatment should vary among populations, further study will be necessary for any other populations except Japanese.

We have shown that the rs8099917 polymorphism determined by Invader assay should be the best predictor of the

response to PEG-IFN/RBV in Japanese chronic hepatitis C patients.

ACKNOWLEDGMENT

This study was supported by grants (22-302 and 21-113) from the National Center for Global Health and Medicine in Japan.

REFERENCES

1. Ando, Y., et al. 2000. Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. *Cancer Res.* **60**:6921–6926.
2. Bruno, S., et al. 2004. Peginterferon alfa-2b plus ribavirin for naive patients with genotype 1 chronic hepatitis C: a randomized controlled trial. *J. Hepatol.* **41**:474–481.
3. Fried, M. W., et al. 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* **347**:975–982.
4. Froehlich, T., and O. Geulen. 2008. Hybridization probe pairs and single-labeled probes: an alternative approach for genotyping and quantification. *Methods Mol. Biol.* **429**:117–133.
5. Garritano, S., et al. 2009. Determining the effectiveness of high resolution melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *BMC Genet.* **10**:5.
6. Ge, D., et al. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **461**:399–401.
7. Global Burden of Hepatitis C Working Group. 2004. Global burden of disease (GBD) for hepatitis C. *J. Clin. Pharmacol.* **44**:20–29.
8. Hadziyannis, S. J., 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.
9. Iadonato, S. P., and M. G. Katze. 2009. Genomics: hepatitis C virus gets personal. *Nature* **461**:357–358.
10. Ikegawa, S., A. Mabuchi, M. Ogawa, and T. Ikeda. 2002. Allele-specific PCR amplification due to sequence identity between a PCR primer and an amplicon: is direct sequencing so reliable? *Hum. Genet.* **110**:606–608.
11. Lyamichev, V., et al. 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* **17**:292–296.
12. Lyamichev, V. I., et al. 2000. Experimental and theoretical analysis of the invasive signal amplification reaction. *Biochemistry* **39**:9523–9532.
13. Manns, M. P., et al. 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* **358**:958–965.
14. Montes-Cano, M. A., et al. 2010. Interleukin-28B genetic variants and hepatitis virus infection by different viral genotypes. *Hepatology* **52**:33–37.
15. Nguyen-Dumont, T., et al. 2009. Description and validation of high-throughput simultaneous genotyping and mutation scanning by high-resolution melting curve analysis. *Hum. Mutat.* **30**:884–890.
16. Rauch, A., et al. 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* **138**:1338–1345.e7.
17. Rouits, E., et al. 2004. Relevance of different UGT1A1 polymorphisms in irinotecan-induced toxicity: a molecular and clinical study of 75 patients. *Clin. Cancer Res.* **10**:5151–5159.
18. Stamer, U. M., B. Bayerer, S. Wolf, A. Hoeffl, and F. Stuber. 2002. Rapid and reliable method for cytochrome P450 2D6 genotyping. *Clin. Chem.* **48**:1412–1417.
19. Suppiah, V., et al. 2009. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat. Genet.* **41**:1100–1104.
20. Tadokoro, K., et al. 2010. Rapid quantification of periodontitis-related bacteria using a novel modification of Invader PLUS technologies. *Microbiol. Res.* **165**:43–49.
21. Tanaka, Y., 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.
22. Thomas, D. L., et al. 2000. The natural history of hepatitis C virus infection: host, viral, and environmental factors. *JAMA* **284**:450–456.
23. Thomas, D. L., et al. 2009. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* **461**:798–801.
24. Whitehall, V., et al. 2009. A multicenter blinded study to evaluate KRAS mutation testing methodologies in the clinical setting. *J. Mol. Diagn.* **11**:543–552.
25. Younossi, Z., J. Kallman, and J. Kincaid. 2007. The effects of HCV infection and management on health-related quality of life. *Hepatology* **45**:806–816.

Original Article

Easy-to-use phylogenetic analysis system for hepatitis B virus infection

Masaya Sugiyama,^{1,2*} Ayano Inui,³ Tadasu Shin-I,¹ Haruki Komatsu,³ Motokazu Mukaide,^{1,4} Naohiko Masaki,¹ Kazumoto Murata,¹ Kiyooki Ito,¹ Makoto Nakanishi,² Tomoo Fujisawa³ and Masashi Mizokami¹

¹The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa,

²Department of Biochemistry and Cell Biology, Nagoya City University Graduate School of Medical Sciences,

Nagoya, ³Department of Pediatrics, Eastern Yokohama Hospital, Yokohama, and ⁴SRL, Inc. Tokyo, Japan

Aim: The molecular phylogenetic analysis has been broadly applied to clinical and virological study. However, the appropriate settings and application of calculation parameters are difficult for non-specialists of molecular genetics. In the present study, the phylogenetic analysis tool was developed for the easy determination of genotypes and transmission route.

Methods: A total of 23 patients of 10 families infected with hepatitis B virus (HBV) were enrolled and expected to undergo intrafamilial transmission. The extracted HBV DNA were amplified and sequenced in a region of the S gene.

Results: The software to automatically classify query sequence was constructed and installed on the Hepatitis Virus Database (HVDB). Reference sequences were retrieved from HVDB, which contained major genotypes from A to H. Multiple-alignments using CLUSTAL W were performed before the genetic distance matrix was calculated with the six-parameter method. The phylogenetic tree was output by the

neighbor-joining method. User interface using WWW-browser was also developed for intuitive control. This system was named as the easy-to-use phylogenetic analysis system (E-PAS). Twenty-three sera of 10 families were analyzed to evaluate E-PAS. The queries obtained from nine families were genotype C and were located in one cluster per family. However, one patient of a family was classified into the cluster different from her family, suggesting that E-PAS detected the sample distinct from that of her family on the transmission route.

Conclusions: The E-PAS to output phylogenetic tree was developed since requisite material was sequence data only. E-PAS could expand to determine HBV genotypes as well as transmission routes.

Key words: database, genotype, hepatitis B virus, intrafamilial transmission, phylogenetic analysis

INTRODUCTION

HEPATITIS B VIRUS (HBV) infects approximately 350 million people worldwide. Chronic HBV infection causes liver cirrhosis and liver cancer. In Japan, chronic hepatitis B patients are estimated to be approximately one million.¹ As HBV has high infectivity, almost all advanced countries have launched the

universal infant immunization program against HBV. The patients with HB antigen seropositive, fulminant hepatitis and HCC have substantially declined in these countries.^{2,3} The current immunization strategy represents favorable effects to prevent the HBV transmission. In countries without the universal vaccination program, however, substantial intrafamilial transmission and horizontal transmission have been reported.⁴ In Japan, only high-risk infants born to chronic HBV-infected mothers have been given the HBV vaccine according to a selective vaccination policy of Japanese governments since 1986. This strategy has led to successful reduction of HBV carrier infants⁵ since the major route of transmission has been perinatal transmission,^{6,7} and horizontal transmission in early childhood has occurred as a result of close family contact.^{8–10} However, the main

Correspondence: Dr Masashi Mizokami, The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, 1-7-1, Kohmodai, Ichikawa 272-8516, Japan. Email: mmizokami@hospk.ncgm.go.jp

*JSPS Research Fellow.

Received 14 April 2011; revision 13 June 2011; accepted 15 June 2011.

route of acute hepatitis B in adult is sexual transmission in countries without a universal vaccination program.

Hepatitis B virus genotypes are identified worldwide and classified into at least eight genotypes (A–H) on the basis of a divergence of 8% or more of the entire nucleotide sequences.^{1,11–13} In a Japanese population of chronic hepatitis B, the distribution of major HBV genotypes (HBV/A, B, C, and D) was reported to be 1.7%, 12.2%, 84.7%, and 0.4% respectively.¹⁴ However, the prevalence of HBV/A increased to approximately 40% in acute HBV infection, and 4.3% in chronic HBV infection in Japan.¹⁵ The main transmission routes of HBV/A are a contact among men who have sex with men as well as among members of a heterosexual population.^{16–19} The expansion of HBV infection has been led by the sexual transmission of HBV/A since little infection by contaminated medical materials has recently been reported in Japan.

As previously reported, phylogenetic analyses based on virus genome revealed the transmission route.²⁰ The

identification of the transmission route provides beneficial information for epidemiologic study and healthcare reform. Although phylogenetic analysis is gradually known in this field, the handlings of the data are difficult to achieve for researchers who are not familiar with genetic analysis. In this study, we have developed a novel computing system to output phylogenetic trees automatically when users simply input their sequence data. The result from this system also shows the type of transmission route such as intrafamilial (mother-to-child, father-to-child, or child-to-child) or horizontal in the general population. For this purpose, sera collected from 10 families with possible intrafamilial transmission were used to reveal their accurate transmission route.

METHODS

Patients

TEN FAMILIES CONSISTING of 23 patients being inactive carriers or having chronic hepatitis B were enrolled in this study after 32 individuals of 10 families

Table 1 Characteristics of intrafamilial transmission cases

Family	Feature	Age	Sex	HBV DNA (Log copies/mL)	Status 1	Status 2	ID
FM1	C	11	F	>9	eAg +	IC	1
	Mo	40	F	8.7	eAg +	IC	2
FM2	C	7	M	>9	eAg +	IC	3
	Mo/S1	41	F	8.6	eAg +	CH	4
	S2	27	F	<2.6	eAb +	IC	5
	S3	30	F	3	eAb +	IC	6
FM3	C	3	F	>9	eAg +	IC	7
	Mo	37	F	>9	eAg +	IC	8
FM4	C1	15	M	>9	eAg +	IC	9
	C2	20	F	>9	eAg +	IC	10
	Mo	47	F	>9	eAg +	IC	11
FM5	C	15	M	5.9	eAb +	IC	12
	Fa	46	M	4.2	eAb +	IC	13
FM6	C	12	M	>9	eAg +	CH	14
	Mo	36	F	>9	eAg +	CH	15
FM7	C	2	F	2.8	eAb +	IC	16
	Mo	33	F	>9	eAg +	IC	17
FM8	C	2	M	>9	eAg +	IC	18
	Mo	41	F	>9	eAg +	IC	19
FM9	C	7	M	3.5	eAb +	IC	20
	Mo	39	F	>9	eAg +	CH	21
FM10	C1	4	M	>9	eAg +	CH	22
	C2	2	F	<2.6	eAb +	IC	23

All patients were hepatitis B surface antigen (HBsAg) positive. Feature: C, children; Mo, mother; Fa, father; S, sister. Sex: M, male; F, female. Status1: Status of HB antigen or antibody except for HbsAg. Status2: IC, inactive carrier; CH, chronic hepatitis. ID 4, 5, 6 are sister and ID 4 is mother of ID 3. ID 22 and 23 are brother and sister.

were tested on HB markers (Table 1). All patients were hepatitis B surface antigen (HBsAg) positive, and the family cases had a possibility of intrafamilial transmission. The histories of familial clustering of HBV infection and hepatitis B vaccination were recorded. The serum samples obtained from all patients were tested for hepatitis B surface antigen (HBsAg), hepatitis B surface antibody (anti-HBs), and hepatitis B core antibody (anti-HBc). Four sera of chronic hepatitis B patients were collected for operation-checks of the developed system before a trial of 23 samples. Serum samples were divided into aliquots and kept at -80°C until testing. The study protocol conformed to the 1975 declaration of Helsinki and was approved by the ethics committees of the respective institutions. Every patient or his/her next of kin gave informed consent to the purpose of the study. Consent of children for participating in the study was filled by their parents.

Serological testing

Their sera were tested for alanine aminotransferase (ALT), and hepatitis B e antigen (HBeAg) and hepatitis B s antigen (HBsAg), as well as antibodies to HBeAg (anti-HBe) and HBsAg (anti-HBs) (Dinabot, Tokyo, Japan). Antibodies to HBcAg (anti-HBc) were tested by ARCHITECT (Abbott Japan, Tokyo, Japan). The inactive carrier state was defined by the presence of HBV surface antigen (HBsAg) with normal ALT levels over 1 year (examined at least four times at 3-month intervals). Chronic hepatitis was defined by elevated ALT levels (>1.5 times the upper limit of normal [35 IU/L]) persisting over 6 months (with at least three bimonthly tests). HBV DNA levels in sera were quantitated with a commercial kit (Taqman Real-time polymerase chain reaction [PCR] or Amplicor HBV Monitor; Roche Diagnostics, Basel, Switzerland) with a detection range from 2.6 to 9 log copies/mL.

Viral DNA extraction

HBV DNA was extracted from 200 μL of serum using QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instruction. The extracted DNA was used for amplification and direct sequencing of S gene as described below.

HBV DNA sequencing

The target of S gene (255 bp, nucleotide positions 458–712) was amplified by nested PCR and sequenced to

detect in high sensitivity. The forward primers of S gene were HBS/F2, 5'- AGGTATGTTGCCCGTTTGC -3' for the outer set and HBS/F1, 5'- GTATGTTGCCCGTTTGCCT -3' for the inner set. The reverse primers of the gene were HBS/R2, 5'- AAAGCCCTACGAACCACTGA -3' for the outer set and HBS/R1, 5'- AAGCCCTACGAACCACTGAA-3' for the inner set. Nested PCRs were performed with these primers for 35 cycles (95°C , 15 s; 58°C , 30 s; 72°C , 30 s) in 1st and 2nd PCR using Veriti (Applied Biosystems, Foster City, CA, USA). The PCR products were sequenced on both strands with the BigDye Terminator V3.1 cycle sequencing kit (Applied Biosystems) with the same primers used for the 2nd PCR. The sequencing products were analyzed with an ABI 3130xl DNA analyzer (Applied Biosystems). The obtained sequences were aligned with GenBank sequences corresponding to HBV genotypes.

Database and the system for phylogenetic analysis

The determination of transmission routes using phylogenetic analyses is considered based on age difference, mutation rate, clinical background, and the amplicon size between queries. The determination of non-intrafamilial transmission between queries was performed using a significance level of $P < 10^{-99}$. The criterion was calculated by the simulation of random sampling from Hepatitis Virus Database (HVDB) and the genetic factors described above. The presently developed system, the easy-to-use phylogenetic analysis system (E-PAS) for the determination of genotypes and/or transmission route is implemented in the account mode of the HVDB (<http://s2as02.genes.nig.ac.jp>). We recommend that researchers contact the web master before use.

Table 2 Reference data for the analysis of transmission root

Genotype	Country	Number of references
A	Foreign	2
B	Japan	24
C	Japan	64
D	Japan	1
E	Foreign	1
F	Foreign	1
G	Foreign	1
H	Japan	3

a

Figure 1 WWW-based user interface. To avoid bias such as age and selective pressure of antiviral treatment, adjustment parameters are prepared to obtain appropriate results. (a) The input field of query sequences and patient's age when serum was collected is shown. (b) The upper field configures masked sites for the elimination of drug-resistant sites. The lower field is a mutation rate of hepatitis B virus (HBV).