表1　*KCNJ 15* の多型は非肥満2型糖尿病と関連する[7]

		試料数	頻度(%)		オッズ比	
			アレル	遺伝子型	アレル	遺伝子型
3つの試料セット総計	健常者	1,700	3.1	6.1	–	–
	全患者	1,568	5.4	10.2	1.76	1.75
	非肥満患者(Ⅰ)[a]	875	5.8	11.1	1.93	1.92
	非肥満患者(Ⅱ)[b]	285	7.6	14.1	2.54	2.51

a)非肥満患者(Ⅰ)群：診断時に肥満でなかった患者(BMI<24)．
b)非肥満患者(Ⅱ)群：これまで肥満になったことがない患者(BMI<24)．
3つの独立な患者群・健常者群試料セットを合計した値を示す．

GWASの成果が顕著な疾患の1つが2型糖尿病である．2007年に複数の欧米のグループから発表された結果は，いずれも数千人規模のGWASに基づくものであり，合わせて11個以上の感受性遺伝子が同定され，その半数以上が新規に見いだされた遺伝子であった．[1~4] 翌2008年にはその3大グループが共同でメタ解析を行い，更に6個の新規感受性遺伝子を同定している．日本からも2つの独立なグループが，新規遺伝子 *KCNQ1* を報告した．[5,6] 我々もその一方のグループにおいて，大規模SNP解析の一部を担当した．

また我々は最近，GWASではなくゲノムワイド連鎖解析によって検出された候補領域から段階的に絞り込むことによって，新たな感受性遺伝子 *KCNJ15* を同定した．[7] 興味深いことに，見いだされたリスクアレルは2型糖尿病患者全体よりも肥満でない患者とより強い関連を示した．すなわち，BMI(body mass index)が24以上になったことのない患者では，オッズ比が2.5に達した(表1)．肥満でない2型糖尿病患者は，日本をはじめアジア諸国には多いが欧米諸国にはほとんどいない．実際，デンマークの共同研究者によれば，見いだされたリスクアレルの頻度は1%以下であった．興味深いことに，この遺伝子はインスリン産生の制御に関わる遺伝子と考えられ，特に高血糖時のインスリン産生誘導に抑制的な機能を持つと推定された(図1)．

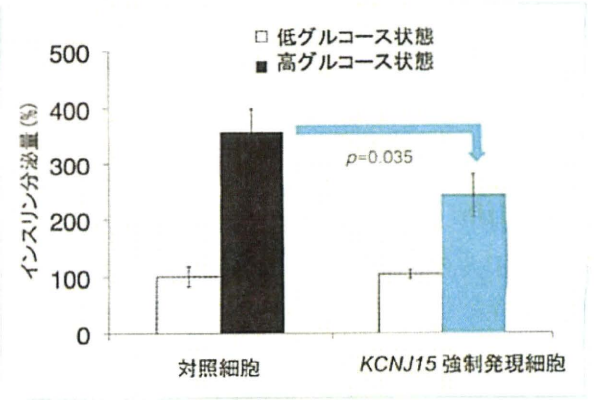

図1　インスリン分泌細胞に *KCNJ 15* を強制発現すると応答性インスリン分泌量が減少する[7]

代表的な過眠症であるナルコレプシーの発症にも，ストレスなどの環境要因に加えて複数の遺伝要因が関わると考えられる．従来，遺伝要因として確定しているのは *HLA-DQB1* 遺伝子のみであったため，我々がGWASを用いて探索した結果，22番染色体上の *CPT1B* 及び *CHKB* 遺伝子近傍の多型がナルコレプシーに関連することを見いだした(図2)．[8] このリスクアレルは両遺伝子の低発現レベルと関連した(図3)．さらに，スタンフォード大を中心とする国際共同研究グループのGWASによって，新たな感受性遺伝子 *TCRA*(T細胞リセプターα)を見いだすことができた(表2)．[9] *TCRA* と上述した *HLA* とは，多様な抗原に対する免疫応答性を制御する主役であることが知られている．これらの成果から，我々はナルコレプシー発症には少なくとも2つの機序，すなわちオレキシン(ヒポクレチン)産生細胞に対する自己免疫，及び脂肪酸β酸化あるいはコリン代謝系の異常が関与していると推定している．

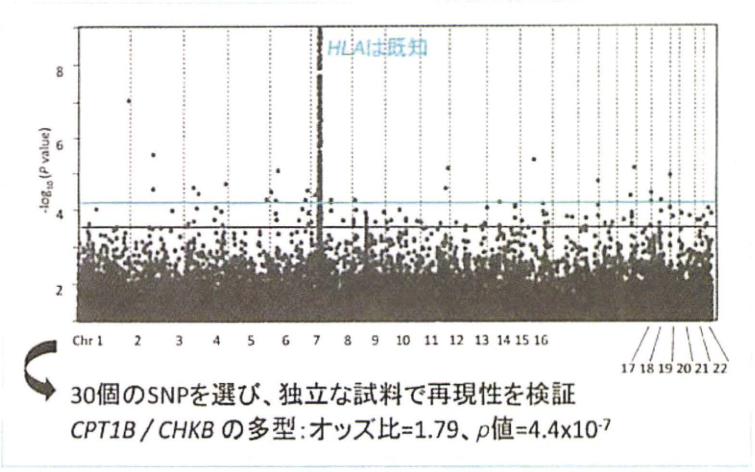

図2　GWASで検出されたナルコレプシーの感受性座位[8]

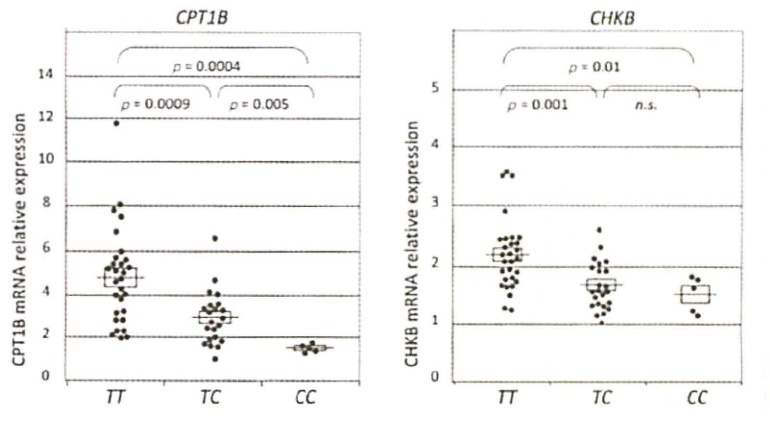

図3　感受性アレルは *CPT1B* 及び *CHKB* の低発現レベルと関連する[8]
末梢血白血球中のmRNA発現レベルをリアルタイムPCR法で定量した．

表2　*TCRA* の多型とナルコレプシーの関連[9]

集団	rs 12587781 C	rs 1154155 C	rs 1263646 G
ヨーロッパ系集団			
患者中頻度	0.22	0.22	0.24
健常者中頻度	0.14	0.14	0.16
p 値	3.58×10^{-7}	3.67×10^{-7}	2.19×10^{-4}
オッズ比	1.79	1.80	1.65
日本人及び韓国人			
患者中頻度	0.68	0.57	0.51
健常者中頻度	0.61	0.47	0.45
p 値	8.70×10^{-4}	2.30×10^{-7}	1.73×10^{-5}
オッズ比	1.34	1.54	1.30

4. GWASによる治療・薬剤応答性遺伝子の探索

ウイルス性肝炎に関する多施設共同研究においても，我々はGWASを担当している．最近，C型肝炎のペグインターフェロンα・リバビリン併用療法の無効例と極めて強く関連する遺伝子として，全く予想されていなかった19番染色体上の *IL28B* 遺伝子を見いだした．[10] すなわちGWAS段階で既に 10^{-12} レベルのp値を示し，ゲノムワイド有意水準をクリアした(図4)．再現性検討段階を合わせるとp値は 10^{-27}～10^{-32} となり，本治療法の無効性に関

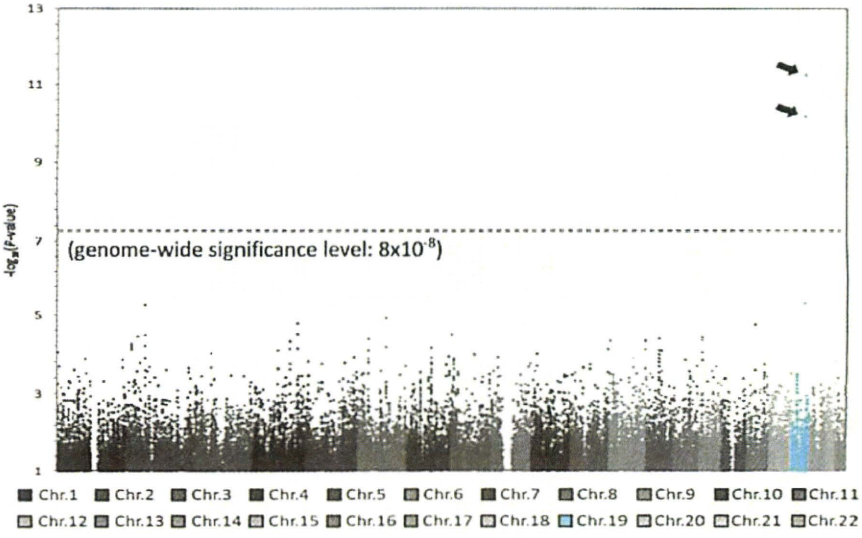

図4　19番染色体上の2個のSNPがGWAS段階でゲノムワイド有意水準の関連を示した[10]

してマイナーアレル陽性患者群のオッズ比は17～30に達した(図5)．またリスクアレルを持つ患者群では，*IL-28* の発現レベルが有意に低いことも分かった．従来，ウイルスの型や量が主に関わると考えられてきたインターフェロン療法への応答性が，実はほぼヒトの遺伝要因で決定されていることの発見は大きなインパクトを持つ．さらに興味深いことに，*IL28B* はインターフェロンλファミリーの一員であり，治療に用いられるインターフェロンαと類似するレセプター及び細胞内シグナル伝達系を介して，ウイルス感染に対する防御作用を発揮する．したがって，治療効果の見込めない患者の予測のみならず，新たな治療薬の開発への貢献が期待される．

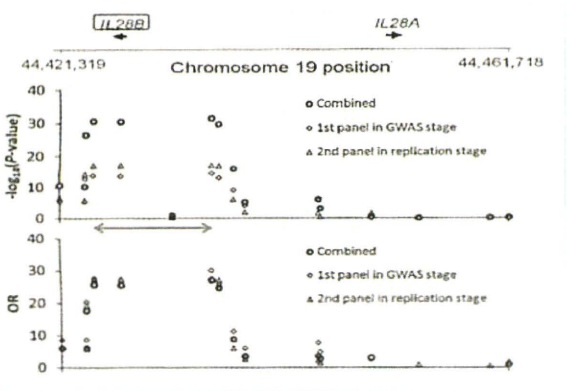

図5　C型肝炎のPEG-IFN-α/RBV治療応答性が *IL 28 B* 遺伝子領域にマップされた[10]

このほか，各種の薬剤に対する反応性に関わる遺伝子が報告されており，GWASによって初めて同定された遺伝子も増えている．また，薬剤過敏症に関わる遺伝子の探索においても，GWASが用いられる研究が多くなっている．一般に薬剤応答性遺伝子は，疾患感受性遺伝子に比較してオッズ比が大きい傾向にあるため，比較的少数の試料でも遺伝子を特定できる可能性が高いことから，今後ますますの成果が期待される．

5　疾患関連遺伝子に見られる集団差

次に，先に述べた2型糖尿病の感受性遺伝子に見られた集団差を取り上げたい．表3は，ヨーロッパ系集団で見いだされた代表的な2型糖尿病感受性遺伝子 *TCF7L2* と，日本人で見いだされた *KCNQ1* について，両集団におけるアレル頻度，オッズ比，p 値を比較したものである．*TCF7L2* は，ヨーロッパ系集団では 10^{-55} と明確な関連を示すことが分かる．一方，日本人では p 値が 10^{-4} レベルにとどまるが，その最大の理由はマイナーアレルの頻度の違

表3　2型糖尿病感受性遺伝子にみられる集団差[1~5,11]

遺伝子	集団	オッズ比	p 値	マイナーアレル頻度
TCF7L2	ヨーロッパ人	1.37	1.0×10^{-18}	0.31/0.25
	日本人	1.70	7.0×10^{-4}	0.05/0.02
KCNQ1	ヨーロッパ人	1.29	7.8×10^{-4}	0.03/0.05
	日本人	1.43	2.8×10^{-29}	0.31/0.40

いによる．日本人ではこれが1桁低いために，ヨーロッパ系集団と同様なオッズ比を示すものの，2,000人の患者と2,000人の健常者を解析しても明瞭な関連は観察されない．[11]

これとは対照的な状況が*KCNQ1*について見られる．すなわち，日本人ではp値が10^{-29}と明確な関連が認められ，また韓国人，中国人試料についても同様に明確な関連が認められた一方，ヨーロッパ系では同様なオッズ比が認められるものの，p値が10^{-4}レベルにとどまる．[5] すなわち，ヨーロッパ系集団とアジア系集団の各々において代表的な2型糖尿病感受性遺伝子は，いずれも集団差を越えて共通する遺伝要因であるけれども，その頻度が大きく異なるために，それぞれの集団における重要性は異なるといえる．いうまでもなく，前述の非肥満型糖尿病に特徴的な感受性遺伝子*KCNJ15*もまた，明瞭な集団差を示す典型例といえる．

同様な集団差は，上述したナルコレプシーの新規感受性領域*CPT1B/CHKB*についても認められた(表4)．[8] すなわち，日本人及び韓国人ではアレル頻度が近似して共に有意な関連が認められたが，ヨーロッパ系アメリカ人及びアフリカ系アメリカ人ではオッズ比で同じ傾向を示すものの，アレル頻度が低いために有意差には至らなかった．これもまた，それぞれの集団における寄与度が異なる例と考えられる．

表4　ナルコレプシーの感受性遺伝子に見られる集団差[8]

集団	患者数/健常者数	mAF 患者	mAF 健常者	オッズ比	p 値
日本人	381/579	0.251	0.158	1.79	4.4×10^{-7}
韓国人	115/309	0.248	0.191	1.40	0.03
USA ヨーロッパ系	388/397	0.053	0.040	1.33	0.12
USA アフリカ系	86/98	0.047	0.026	1.86	0.14

mAF：マイナーアレル頻度．

このように，2型糖尿病及びナルコレプシーの成果からヨーロッパ系集団だけ大規模に解析していれば，すべての遺伝要因が見いだされるわけではないことが分かる．すなわち，日本人/アジア系集団において重要な遺伝要因の全容を知るためには，日本人/アジア系集団自体の研究が必須であることを教えてくれる．

6　今後の課題

2型糖尿病については，既に世界から20以上の感受性遺伝子が報告されており，そのうちの10個余りは日本人患者においてもリスク要因となっている．[12] これらのリスクアレル間に相加的効果が認められ，リスクアレルを多く持つほどオッズ比が上昇することも分かった．しかしながら，それらは遺伝要因全体の一部しか説明できないことから，GWASを用いる戦略に懐疑的な議論もある．筆者は，この問題提起に答えるにはまだ時期尚早だと考えている．その最大の理由は，ほとんどの研究がまだGWASの可能性を活かし切っていないことにある．例えば，同じ病名のついた患者群の試料をただ集めただけで行ったGWASでは見いだせない感受性遺伝子を，患者毎の臨床情報も注意深く収集し臨床的亜型に着目した解析を行うことで，見いだせる可能性が大きいと考えている．前述した非肥満型糖尿病の感受性遺伝子*KCNJ15*は，その典型例である．むしろ，ありふれた病気をその遺伝素因から見れば，互いに類似するが異質性もある多くの病気の集合体と見るのが自然ではないかと思う．また，欧米のGWASによく見られるように，統計学的な検出力だけを考慮して圧倒的な数の試料を各地からかき集めた結果，地域集団間の異質性(階層化)がいわばノイズとなって，真の感受性遺伝子を発見し難くしている状況もあると考える．

さらには感受性遺伝子多型の多くが，いわゆるゲノムワイド有意水準には到達せず中間的なp値を

示していると考えられることから，それらの“gray zone”から真の感受性遺伝子多型を同定する方法の確立が，今後解決すべき大きな課題といえる．GWASをスクリーニング段階ととらえ，遺伝子アノテーション，パスウェイ情報や他分野からの知見を組み合わせることも必要であろう．

見いだされた複数の感受性遺伝子が，特定のパスウェイあるいはネットワークに帰属することが分かれば，疾患の発症・病態形成の機序を解明し，新たな治療法を開発するための極めて有用なヒントになる．このことは，遺伝要因の全容が明らかになる以前に，期待できる大きな成果であるといえる．

7 おわりに

疾患関連遺伝子の探索について，ゲノム全域に分布するSNPを利用する戦略の現状と課題を解説した．一方，現在急速な発展を見せている次世代シークエンサーを用いたリシークエンス技術は，初めは家系内伝達が明瞭な，あるいは低頻度ながらオッズ比が高い変異を同定するのに用いられ，徐々に頻度の高い疾患関連多型の検出にも用いられていくであろう．この新技術から生み出される膨大なデータから，良質の望ましい情報を取り出すシステムの確立も今後の大きな課題である．

なお我々は，疾患関連遺伝子の同定を促進するために「統合データベースプロジェクト」のなかでGWASデータベースを構築している[13]（https://gwas.lifesciencedb.jp/index.Japanese.html）．なるべく多くのGWAS結果を受け入れ，個人特定につながらない情報を公開し，より詳細な情報を研究者の申請に応じて提供することによって，更に多数の疾患関連遺伝子が特定され，発症・病態機序の理解が進むことを期待している．

文　献

1) Sladek R. *et al*., *Nature*, 445, 881–885 (2007).
2) Saxena R. *et al*., *Science*, 316, 1331–1336 (2007).
3) Scott L. J. *et al*., *Science*, 316, 1341–1345 (2007).
4) Wellcome Trust Case Control Consortium, *Nature*, 447, 661–678 (2007).
5) Yasuda K. *et al*., *Nat. Genet*., 40, 1092–1097 (2008).
6) Unoki H. *et al*., *Nat. Genet*., 40, 1098–1102 (2008).
7) Okamoto K. *et al*., *Am. J. Hum. Genet*., 86, 54–64 (2010).
8) Miyagawa T. *et al*., *Nat. Genet*., 40, 1324–1328 (2008).
9) Hallmayer J. *et al*., *Nat. Genet*., 41, 708–711 (2009).
10) Tanaka Y. *et al*., *Nat. Genet*., 41, 1105–1109 (2009).
11) Miyake K. *et al*., *J. Hum. Genet*., 53, 174–180 (2008).
12) Miyake K. *et al*., *J. Hum. Genet*., 54, 236–241 (2009).
13) Koike A. *et al*., *J. Hum. Genet*., 54, 543–546 (2009).

新刊紹介

珈琲一杯の薬理学

岡　希太郎　著

医薬経済社／A5・120頁・1,575円

私の1日は，1杯の珈琲から始まる．外出時以外は研究室で毎朝必ず飲む．ただし，これまで珈琲の効能を気にした覚えはない．単純に私にはおいしく思え，飲まないと1日が始まった気分にならない．効能といえば，緑茶の情報はよく耳にするが，珈琲については関心がないせいか，これといった覚えがないのが本音である．しかし，この本を読んで目から鱗が落ちた．これほどまでに学術論文に発表された珈琲の効能があるのかと，ましてや，豆の煎れ方や種類で効能が変わるとは思いもよらなかった．

本書は，珈琲の奥深さを実に詳細に教えてくれる．内容は，タイトル通りの「珈琲一杯の薬理学」と，その前に珈琲が薬として用いられてきた「珈琲一杯の薬史学」の2部構成になっている．

薬学史の章では，珈琲発見にまつわる伝説から始まり，世界最古の薬用珈琲文献，宗教との関わり，どのように世界的に広がったのか，日本への上陸など，歴史が苦手な私にも面白く読める．

薬理学の章では，カフェインばかりが薬用成分ではなく，他の成分で薬になっているもの，あるいは今後期待できる成分の解説から始まり，珈琲の味を決定する焙煎が，実は薬用成分の含有量まで変えること．

各論として肝臓がんや肝炎，2型糖尿病，パーキンソン病の予防効果などが取り上げられている．

珈琲にうるさいヒトには必読書である．珈琲の味と共に焙煎の仕方で各種疾患に対する効果も変わるらしく，読者各位もご自身のテーラーメード珈琲を創られてみたらいかがであろうか．

山田久陽　Hisaharu YAMADA

※本書は，日本薬学会「薬学情報コーナー」で閲覧できます．

Book Review

Research Article

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

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

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

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

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

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

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

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

Keywords: *IL28B*; ISDR; Peg-interferon; Ribavirin; Data mining; Decision tree.
Received 14 March 2010; received in revised form 22 June 2010; accepted 7 July 2010; available online 19 September 2010

* Corresponding author. Address: Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. Tel.: +81 422 32 3111; fax: +81 422 32 9551.
E-mail address: nizumi@musashino.jrc.or.jp (N. Izumi).

Introduction

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

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

Research Article

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

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

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

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

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

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

Materials and methods

Patients

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

Laboratory and histological tests

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

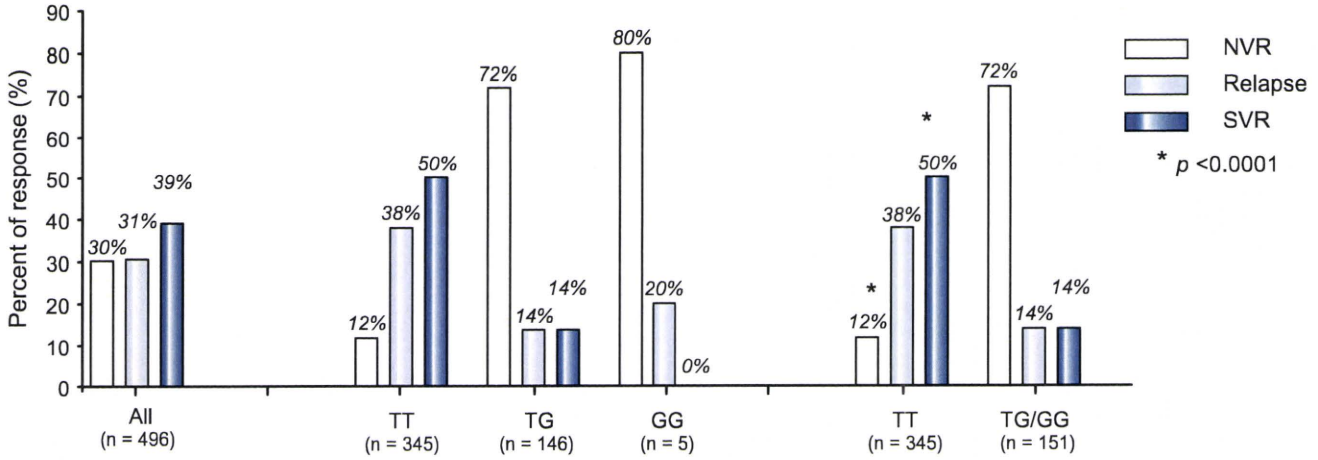

Fig. 1. Association between the IL28B genotype (rs8099917) and treatment response. The rates of response to treatment are shown for each rs8099917 genotype. The rate of null virological response (NVR), relapse, and sustained virological response (SVR) is shown. The *p* values are from Fisher's exact 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. 3 B). On the other hand, the relapse rate was not different between the *IL28B* major and minor alleles within patients who achieved RVR (3% vs. 0%) or cEVR (28% vs. 29%) (Fig. 3A). Amino acid substitutions of Core70 and Core91 were not associated with the rate of relapse (data not shown).

Factors associated with response by multivariate logistic regression analysis

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

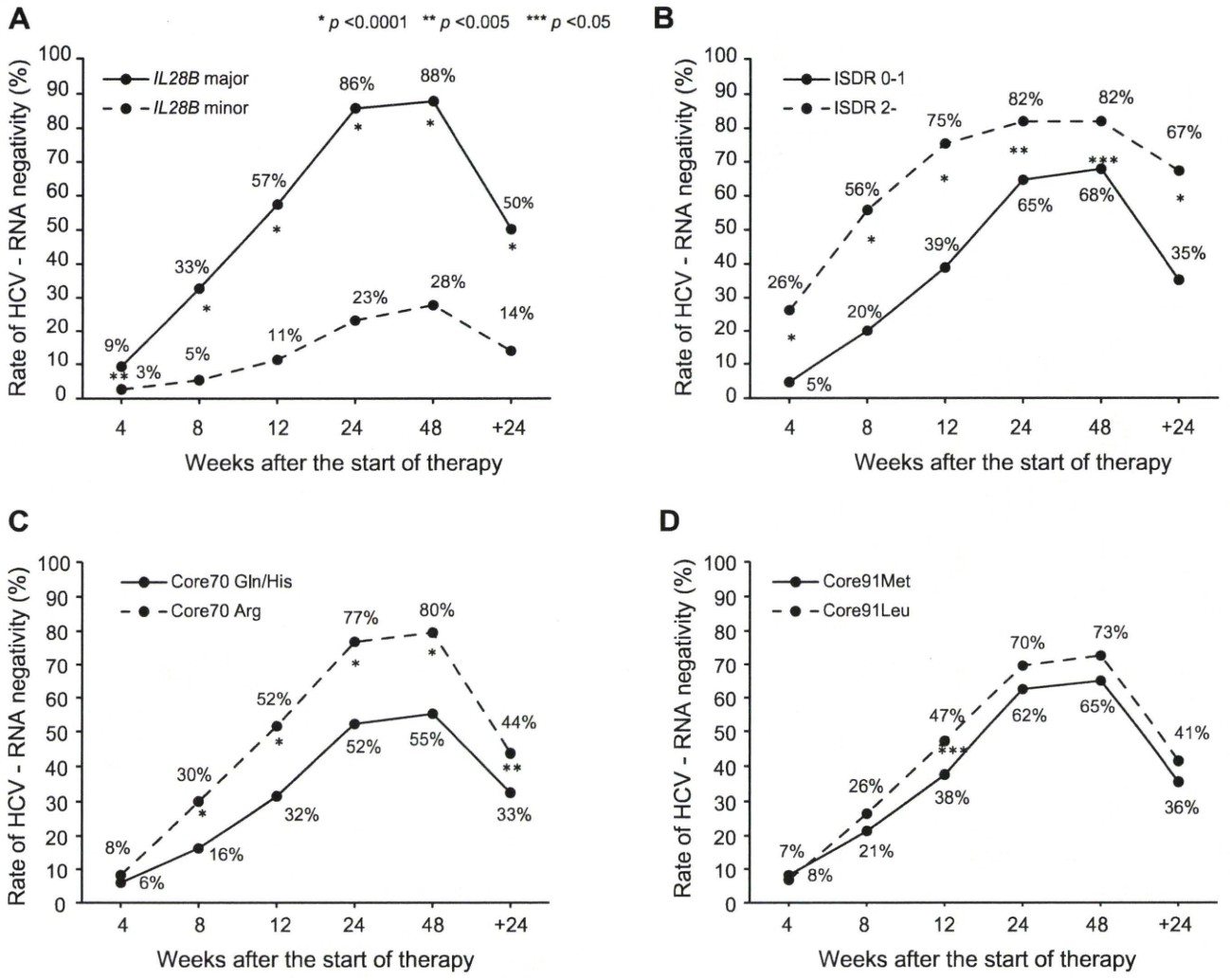

Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified 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.

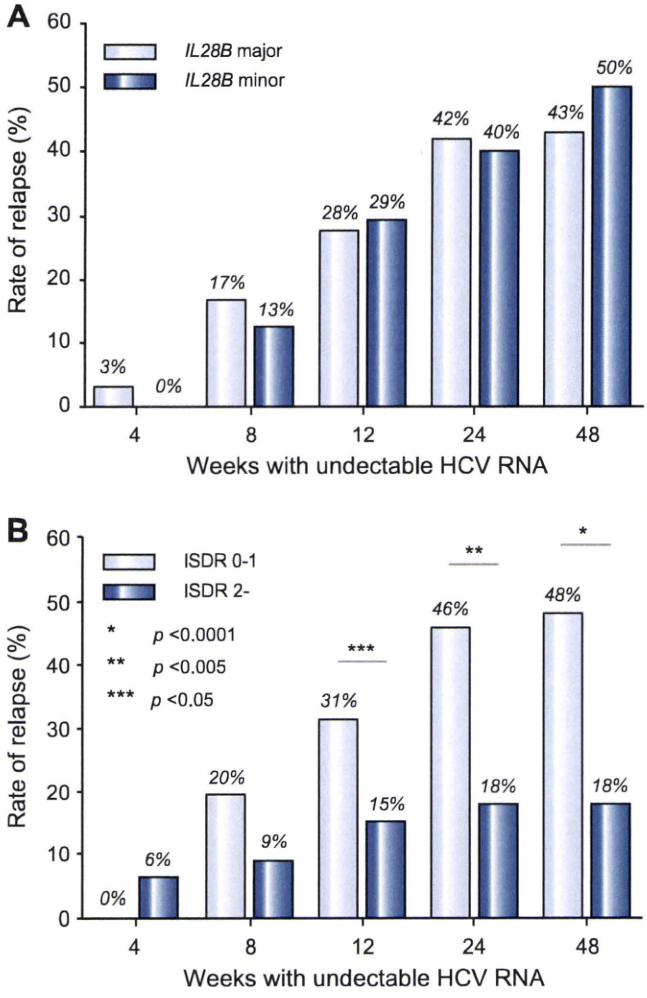

Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele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

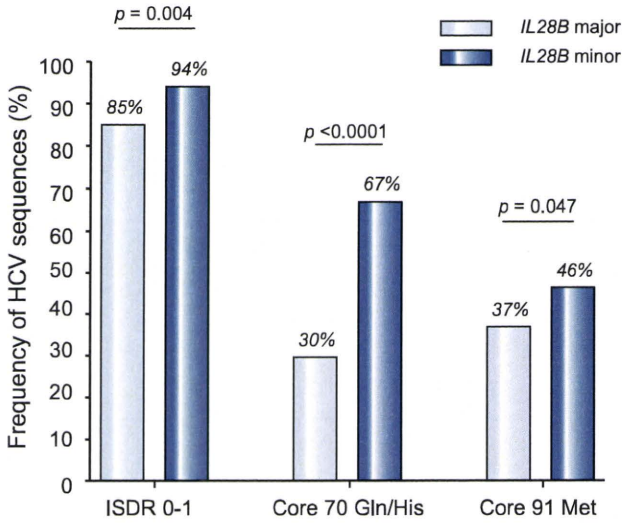

Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met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

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

Validation of the data mining analysis

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

Discussion

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

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 group of patients. Using this model, we can rapidly develop an

Research Article

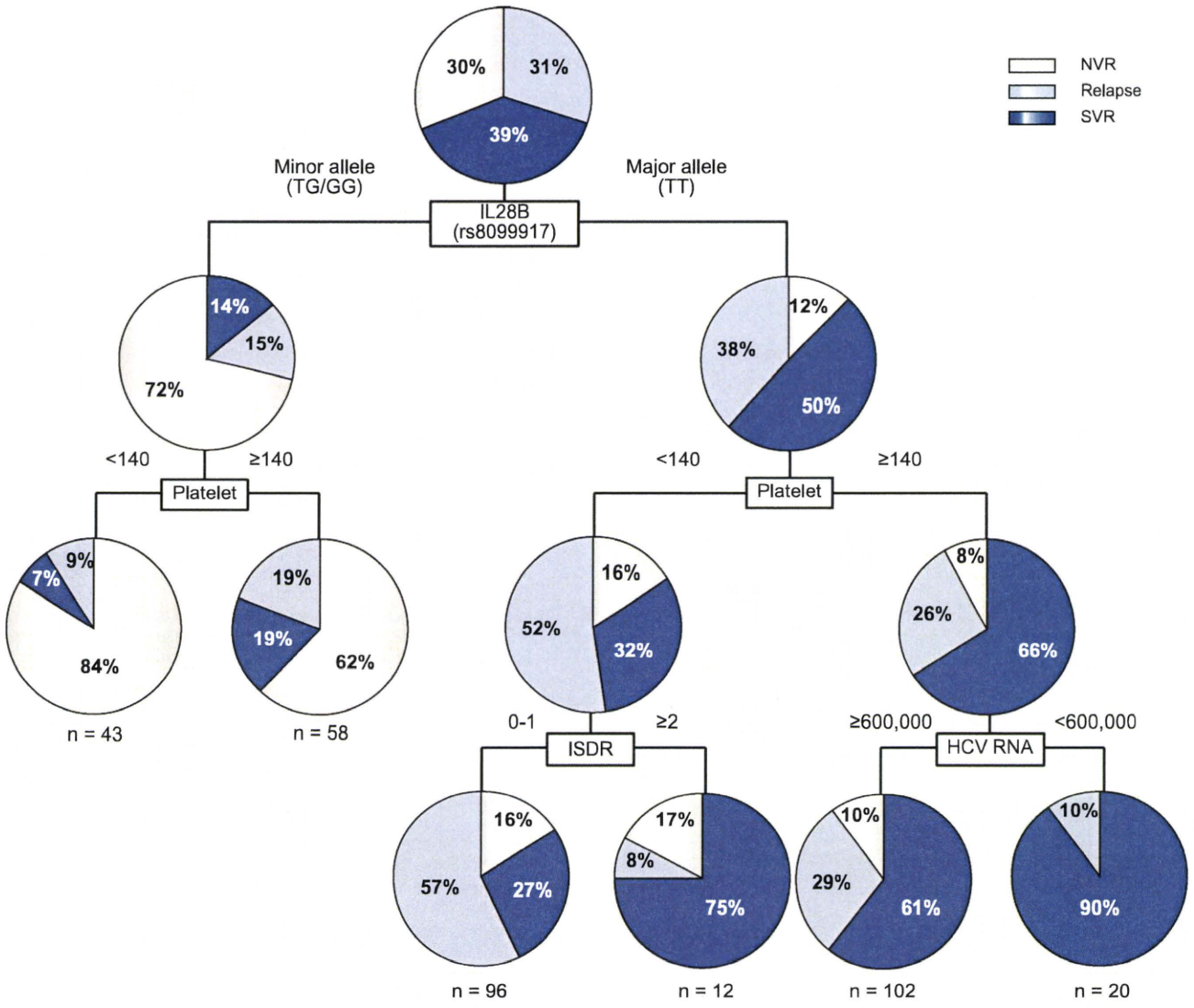

Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown.

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-

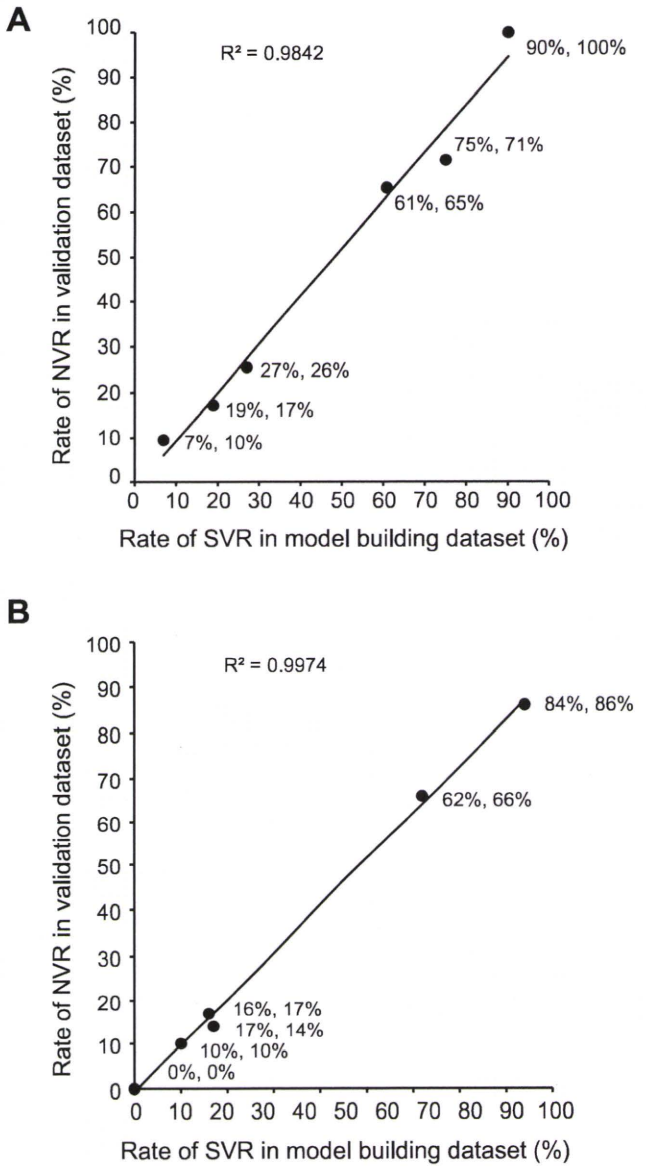

Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and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

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

Journal:	*Antiviral Therapy*
Manuscript ID:	AVT-10-OA-1861.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a

Complete List of Authors:	Kurosaki, Masayuki; Musashino Red Cross hospital, Division of Gastroenterology and Hepatology Tanaka, Yasukito Tanaka, Keisuke; Musashino Red Cross Hospital, Gastroenterology and Hepatology Suzuki, Yuichiro; Musashino Red Cross Hospital, Gastroenterology and Hepatology Hoshioka, Yasuhide; Musashino Red Cross Hospital, Gastroenterology and Hepatology Tamaki, Nobuharu; Musashino Red Cross Hospital, Gastroenterology and Hepatology Kato, Tomoji; Musashino Red Cross Hospital, Gastroenterology and Hepatology Yasui, Yutaka; Musashino Red Cross Hospital, Gastroenterology and Hepatology Hosokawa, Takanori; Musashino Red Cross Hospital, Gastroenterology and Hepatology Ueda, Ken; Musashino Red Cross Hospital, Gastroenterology and Hepatology Tsuchiya, Kaoru; Musashino Red Cross Hospital, Gastroenterology and Hepatology Kuzuya, Teiji; Musashino Red Cross Hospital, Gastroenterology and Hepatology Nakanishi, Hiroyuki; Musashino Red Cross Hospital, Gastroenterology and Hepatology Itakura, Jun; Musashino Red Cross Hospital, Gastroenterology and Hepatology Takahashi, Yuka; Musashino Red Cross Hospital, Gastroenterology and Hepatology Asahina, Yasuhiro; Musashino Red Cross Hospital, Gastroenterology and Hepatology Matsuura, Kentaro Sugauchi, Fuminaka Enomoto, Nobuyuki Nishida, Nao Tokunaga, Katsushi

	Mizokami, Masashi Izumi, Namiki
Keywords:	Antiviral therapy, Hepatitis C virus, ITPA, Peg-interferon, Ribavirin, IL28B, Anemia

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

Masayuki Kurosaki, M.D.[1], Yasuhito Tanaka, M.D.[2], Keisuke Tanaka, M.D.[1],Yuichiro Suzuki, M.D.[1, 3], Yoshihide Hoshioka, M.D.[1], Nobuharu Tamaki, M.D.[1], Tomoji Kato, M.D.[1], Yutaka Yasui, M.D.[1], Takanori Hosokawa, M.D.[1], Ken Ueda, M.D.[1], Kaoru Tsuchiya, M.D.[1], Teiji Kuzuya, M.D.[1], Hiroyuki Nakanishi, M.D.[1], Jun Itakura, M.D.[1], Yuka Takahashi, M.D.[1], Yasuhiro Asahina, M.D.[1], Kentaro Matsuura, M.D.[2], Fuminaka Sugauchi, M.D.[2], Nobuyuki Enomoto, M.D.[3], Nao Nishida, PhD[4], Katsushi Tokunaga, PhD[4], Masashi Mizokami, M.D.[5], Namiki Izumi, M.D.[1]*

Short title: Association of *ITPA* with anemia and treatment outcome

[1]Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan, [2]Department of Virology, Liver unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan, [3]First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan, [4]Department of Human Genetics, Graduate School of

Medicine, University of Tokyo, Tokyo, Japan, [5]Research Center for Hepatitis and Immunology, International Medical Center of Japan Konodai Hospital, Ichikawa, Japan.

Word count: 4012 words including figure legends

5 Tables, 4 Figures

Address correspondence to

Namiki Izumi, M.D.

Division of Gastroenterology and Hepatology,

Musashino Red Cross Hospital

1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan

Tel: +81-422-32-3111

Fax: +81-422-32-9551

E-mail: nizumi@musashino.jrc.or.jp

Grant Support: This study was supported by a grant-in-aid from Ministry of Health, Labor and Welfare, Japan.

Disclosures: No conflicts of interest exist for all authors.