田中靖人, 徳永勝士, 溝上雅史.	IL28B:C型肝炎治療効果を規定する遺伝子多型C型肝炎に対するテーラーメイド治療の確立を目指したゲノムワイド関連解析	肝胆膵 2009	59 (6)	1187-1193	2009
Kurbanov F, Tanaka Y, Mizokami M.	Geographical and genetic diversity of the human hepatitis B virus.	Hepatol Res.	40(1)	14-30	2010
Mukaide M, Tanaka Y, Shin-I T, Yuen MF, Kurbanov F, Yokosuka O, Sata M, Karino Y, Yamada G, Sakaguchi K, Orito E, Inoue M, Baqai S, Lai CL, Mizokami M.	Mechanism of entecavir resistance of hepatitis B virus with viral breakthrough as determined by long-term clinical assessment and molecular docking simulation.	Antimicrob Agents Chemother.	54(2)	882-889	2010
Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M.	A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J.	J Virol.	83(20)	10538-10547	2009
Koike A, Nishida N, Inoue I, Tsuji S, Tokunaga K.	Genome-wide association database developed in the Japanese Integrated Database Project.	Journal of Human Genetics	54(9)	543-546	2009
西田奈央、徳永勝士	疾患感受性遺伝子とゲノムワイド関連解析	治療学	43 (3)	13-18	2009
西田奈央、徳永勝士	テーラーメイド医療をめざした疾患感受性遺伝子のゲノムワイド探索	遺伝子医学MOOK	14	191-195	2009
西田奈央、徳永勝士	疾患関連遺伝子を探し出すためのSNP解析	肝胆膵	59 (6)	1131-1138	2009

Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, Tamaki N, Umeda N, Hosokawa T, Ueda K, Tsuchiya K, Nakanishi H, Ikeda H, Itakura J, Takahashi Y, Asahina Y, Enomoto N, Izumi N.	A Predictive Model of Response to Peginterferon Ribavirin in Chronic Hepatitis C using Classification and Regression Tree Analysis.	Hepatology Research	in press		2010
Jun Itakura Masayuki Kurosaki, Yoshie Itakura Shinya Maekawa Yasuhiro Asahina, Namiki Izumi Nobuyuki Enomoto.	Reproducibility and usability of chronic virus infection model using agent-based simulation; comparing with a mathematical model	Biosystems	99(1)	70-78	2010
Hodo Y, Hashimoto SI, Honda M, Yamashita T, Suzuki Y, Sugano S, Kaneko S, Matsushima K.	Comprehensive gene expression analysis of 5'-end of mRNA identified novel intronic transcripts associated with hepatocellular carcinoma.	Genomics	2010 Jan 21. [Epub ahead of print]		
Takatori H, Yamashita T, Honda M, Nishino R, Arai K, Yamashita T, Takamura H, Ohta T, Zen Y, Kaneko S.	dUTP pyrophosphatase expression correlates with a poor prognosis in hepatocellular carcinoma.	Liver Int.	2009 Nov 30. [Epub ahead of print]		
Komura T, Sakai Y, Honda M, Takamura T, Matsushima K, Kaneko S.	CD14+monocytes are vulnerable and functionally impaired under ER stress in patients with type 2 diabetes.	Diabetes.	2009 Dec 3. [Epub ahead of print]		

Ura S, Honda M, Yamashita T, Ueda T, Takatori H, Nishino R, Sunakozaka H, Sakai Y, Horimoto K, Kaneko S.	Differential microRNA expression between hepatitis B and hepatitis C leading disease progression to hepatocellular carcinoma.	Hepatology	49(4)	1098-1112	2009
Ootsuji H, Honda H, Kaneko S, Usui S, Okajima M, Okada H, Sakai Y, Takamura T, Horimoto K, Takamura M.	Altered Hepatic Gene Expression Profiles Associated with MyocardialIschemia.	Circ Cardiovasc Genet.	3(1)	68-77	2010
Shirasaki T, Honda M, Mizuno H, Shimakami T, Okada H, Sakai Y, Murakami S, Wakita Tand Kaneko S.	La protein required for IRES-directed translation is a potential therapeutic target for hepatitis C virus replication.	Journal of Infectious Diseases	in press		2010
Nagao Y, Matsuoka H, Kawaguchi T, Sata M.	Aminofeel® improves the sensitivity to taste in patients with HCV-infected liver disease.	Med Sci Monit.	in press		2010
Nagao Y, Matsuoka H, Seike M, Yamasaki K, Kato J, Nakajima T, Miyazaki Y, Ohno T, Inuzuka S, Ohira H, Yokosuka O, Yatsuhashi H, Mori T, Honda K, Kawaguchi T, Ide T, Sata M.	Knowledge about *Vibrio vulnificus* infection in Japanese patients with liver diseases: A prospective multicenter study.	Med Sci Monit.	15(10)	115-120	2009
Nagao Y, Sata M.	High incidence of multiple primary carcinomas in HCV-infected patients with oral squamous cell carcinoma.	Med Sci Monit.	15(9)	453-459	2009

佐田通夫, 長尾由実子, 大坪維範, 岡村 孝.	C型肝炎. HCV感染とB cell clonality、口腔癌、インスリン抵抗性についての検討.	犬山シンポジウム記録刊行会	27	137-142	2010
Kobayashi Y, Suzuki, Y, Itou T, Carvalho AA, Cunha EM, Ito FH, Gojobori T, Sakai T.	Low genetic diversities of rabies virus populations within different hosts in Brazil.	Infect Genet Evol.	in press		
Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, Ikuta K, Jern P, Gojobori T.	Endogenous non-retroviral RNA virus elements in mammalian genomesIdentification of endogenous non-retroviral RNA virus elements in mammalian genomes.	Nature	463(7277)	84-87.	2010
Kaminuma E, Mashima J, Kodama Y, Gojobori T, Ogasawara O, Okubo K, Takagi T, Nakamura Y.	DDBJ launches a new archive database with analytical tools for next-generation sequence data.	Nucleic Acids Res.	38(Database issue)	D33-38.	2010
Takeda J, Suzuki Y, Sakate R, Sato Y, Gojobori T, Imanishi T, Sugano, S.	H-DBAS: Human-transcriptome DataBase for Alternative Splicing, update 2010	Nucleic Acids Res.	38(Database issue):	D86-90.	2010
Yamasaki C, Murakami K, Takeda J, Sato Y, Noda A, Sakate R, Habara T, Nakaoka H, Todokoro F, Matsuya A, Imanishi T, Gojobori T.	H-InvDB in 2009, extended database and data-mining resources for human genes and transcripts.	Nucleic Acids Res.	38(Database issue)	D626-632.	2010

The HUGO Pan-Asian SNP Consortium, Abdu M, Ahmed I, Brahmachari S, Gojobori T, Liu E, Sugano S, Suzuki Y, Tokunaga K, Zilfalil B A, Indian Genome Variation Consortium.	Mapping Human Genetic Diversity in Asia.	Science	326	1541-1545.	2009
Suzuki Y, Gojobori T, Kumar S.	Methods for incorporating the hypermutability of CpG dinucleotides in detecting natural selection operating at the amino acid sequence level.	Mol. Biol. Evol.	26 (10)	2275-84.	2009
Clemente JC, Ikeo K, Valiente G, Gojobori T.	Optimized ancestral state reconstruction using Sankoff parsimony.	BMC Bioinformatics	10 (51)	1-27.	2009
Sugawara H, Ikeo K, Fukuchi S, Gojobori T, Tateno Y.	DDBJ Dealing with Mass Data Produced by the Second Generation Sequencer.	Nucleic Acids Res.	37(Database issue)	D16-8.	2009
Shimada MK, Matsumoto R, Hayakawa Y, Sanbonmatsu R, Gough C, Yamaguchi-Kabata Y, Yamasaki C, Imanishi T, Gojobori T.	VarySysDB: a human genetic polymorphism database based on all H-InvDB transcripts.	Nucleic Acids Res.	37(Database issue)	D810-5	2009
Fukuchi S, Homma K, Sakamoto S, Sugawara H, Tateno Y, Gojobori T, Nishikawa K.	The GTOP database in 2009: updated content and novel features to expand and deepen insights into protein structures and functions.	Nucleic Acids Res.	37(Database issue)	D333-7	2009

IV. 研究成果の刊行物・別冊

nature genetics

Genome-wide association of *IL28B* with response to pegylated interferon-α and ribavirin therapy for chronic hepatitis C

Yasuhito Tanaka[1,18], Nao Nishida[2,18], Masaya Sugiyama[1], Masayuki Kurosaki[3], Kentaro Matsuura[1], Naoya Sakamoto[4], Mina Nakagawa[4], Masaaki Korenaga[5], Keisuke Hino[5], Shuhei Hige[6], Yoshito Ito[7], Eiji Mita[8], Eiji Tanaka[9], Satoshi Mochida[10], Yoshikazu Murawaki[11], Masao Honda[12], Akito Sakai[12], Yoichi Hiasa[13], Shuhei Nishiguchi[14], Asako Koike[15], Isao Sakaida[16], Masatoshi Imamura[17], Kiyoaki Ito[17], Koji Yano[17], Naohiko Masaki[17], Fuminaka Sugauchi[1], Namiki Izumi[3], Katsushi Tokunaga[2] & Masashi Mizokami[1,17]

The recommended treatment for patients with chronic hepatitis C, pegylated interferon-α (PEG-IFN-α) plus ribavirin (RBV), does not provide sustained virologic response (SVR) in all patients. We report a genome-wide association study (GWAS) to null virological response (NVR) in the treatment of patients with hepatitis C virus (HCV) genotype 1 within a Japanese population. We found two SNPs near the gene *IL28B* on chromosome 19 to be strongly associated with NVR (rs12980275, $P = 1.93 \times 10^{-13}$, and rs8099917, 3.11×10^{-15}). We replicated these associations in an independent cohort (combined P values, 2.84×10^{-27} (OR = 17.7; 95% CI = 10.0–31.3) and 2.68×10^{-32} (OR = 27.1; 95% CI = 14.6–50.3), respectively). Compared to NVR, these SNPs were also associated with SVR (rs12980275, $P = 3.99 \times 10^{-24}$, and rs8099917, $P = 1.11 \times 10^{-27}$). In further fine mapping of the region, seven SNPs (rs8105790, rs11881222, rs8103142, rs28416813, rs4803219, rs8099917 and rs7248668) located in the *IL28B* region showed the most significant associations ($P = 5.52 \times 10^{-28}$–2.68×10^{-32}; OR = 22.3–27.1). Real-time quantitative PCR assays in peripheral blood mononuclear cells showed lower *IL28B* expression levels in individuals carrying the minor alleles ($P = 0.015$).

Hepatitis C is a global health problem that affects a significant proportion of the world's population. The World Health Organization estimated that in 1999, there were 170 million HCV carriers worldwide, with 3–4 million new cases appearing each year. HCV infection affects more than 4 million people in the United States, where it represents the leading cause of cirrhosis and hepatocellular carcinoma as well as the leading cause of liver transplantation[1].The American Gastroenterological Association estimated that drugs are the largest direct costs of hepatitis C[1].

The most effective current standard of care in patients with chronic hepatitis C, a combination of PEG-IFN-α with ribavirin, does not produce SVR in all patients treated. Large-scale studies on 48-week-long PEG-IFN-α/RBV treatment in the United States and Europe showed that 42–52% of patients with HCV genotype 1 achieved SVR[2–4], and similar results were found in Japan. However, older patients (greater than 50 years of age) had a significantly lower rate of SVR due to poor adherence resulting from adverse events and laboratory-detectable abnormalities such as neutropenia and thrombocytopenia[5,6]. Specifically, various well-described side effects (such as a flu-like syndrome, hematologic abnormalities and adverse neuropsychiatric events) often necessitate dose reduction, and 10–14% of patients require premature withdrawal from interferon-based therapy[7]. To avoid these side effects in patients who will not be helped by the treatment, as well as to reduce the substantial cost of PEG-IFN-α/RBV treatment, it would be useful to be able to predict an individual's response before or early in treatment. Several viral factors, such as genotype 1, high baseline viral load, viral

[1]Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan. [2]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [3]Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan. [4]Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan. [5]Division of Hepatology and Pancreatology, Kawasaki Medical College, 577 Matsushima, Kurashiki, Japan. [6]Department of Internal Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan. [7]Molecular Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kyoto, Japan. [8]National Hospital Organization Osaka National Hospital, Osaka, Japan. [9]Department of Medicine, Shinshu University School of Medicine, Matsumoto, Japan. [10]Division of Gastroenterology and Hepatology, Internal Medicine, Saitama Medical University, Saitama, Japan. [11]Second department of Internal Medicine, Faculty of Medicine, Tottori University, Yonago, Japan. [12]Department of Gastroenterology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan. [13]Department of Gastroenterology and Metabology, Ehime University Graduate School of Medicine, Ehime, Japan. [14]Department of Internal Medicine, Hyogo College of Medicine, Nishinomiya, Japan. [15]Central Research Laboratory, Hitachi Ltd., Kokubunji, Japan. [16]Gastroenterology and Hepatology, Yamaguchi University Graduate School of Medicine, Yamaguchi, Japan. [17]Research Center for Hepatitis and Immunology, International Medical Center of Japan Konodai Hospital, Ichikawa, Japan. [18]These authors contributed equally to this work. Correspondence should be addressed to M.M. (mmizokami@imcjk2.hosp.go.jp).

Received 29 June; accepted 21 August; published online 13 September 2009; doi:10.1038/ng.449

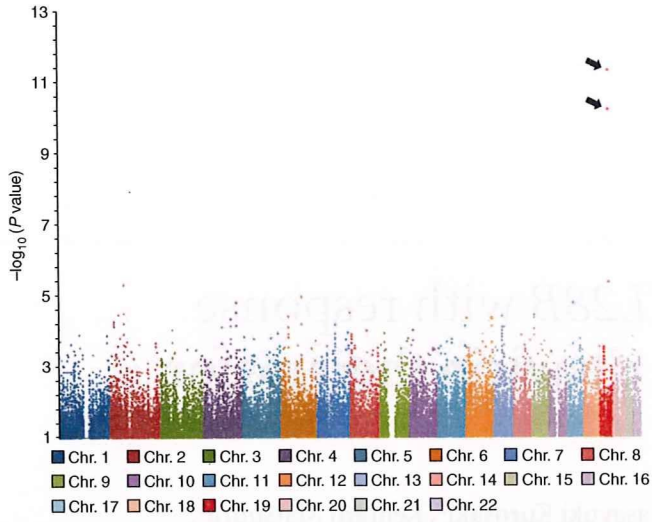


Figure 1 Genome-wide association results with PEG-IFN-α/RBV treatment in 142 Japanese patients with HCV (78 NVR and 64 VR samples). *P* values were calculated by using a χ^2 test for allele frequencies. The dots with arrows for chromosome 19 denote SNPs that showed significant genome-wide associations ($P < 8.05 \times 10^{-8}$) with response to PEG-IFN-α/RBV treatment.

kinetics during treatment, and amino acid pattern in the interferon sensitivity–determining region, have been reported to be significantly associated with the treatment outcome in a number of independent studies[8–10]. Studies have also provided strong evidence that ~20% of patients with HCV genotype 1 and 5% of patients with genotype 2 or 3 have a null response to PEG-IFN-α/RBV. No definite predictor of this resistance is currently available that make it possible to bypass the initial 12–24 weeks' treatment before deciding whether treatment should be continued. If a reliable predictor of non-response were identified for use in patients before treatment initiation, then an estimated 20%, including those who have little or no chance to achieve SVR, could be spared the side effects and cost of treatment.

Host factors, including age, sex, race, liver fibrosis and obesity, have also been reported to be associated with PEG-IFN-α/RBV therapy outcome[11,12]. However, little is known about the host genetic factors that might be associated with the response to therapy: thus far only a few candidate genes, including those encoding type I interferon receptor-1 (*IFNAR1*) and mitogen-activated protein kinase–activated protein kinase 3 (*MAPKAPK3*), have been reported to be associated with treatment response[13,14]. We describe here a GWAS for response to PEG-IFN-α/RBV treatment.

We conducted this GWAS to identify host genes associated with response to PEG-IFN-α/RBV treatment in 154 Japanese patients with HCV genotype 1 (82 with NVR and 72 with virologic response (VR), based on the selection criteria as described in Online Methods). We used the Affymetrix SNP 6.0 genome-wide SNP typing array for 900,000 SNPs. A total of 621,220 SNPs met the following criteria: (i) SNP call rate ≥95%, (ii) minor allele frequency (MAF) ≥1% and (iii) deviation from Hardy-Weinberg equilibrium (HWE) $P \geq 0.001$ in VR samples. After excluding 4 NVR and 8 VR samples that showed quality control (QC) call rates of <95%, 78 NVR and 64 VR samples were included in the association analysis. **Figure 1** shows a genome-wide view of the single-point association data based on allele frequencies. Two SNPs located close to *IL28B* on chromosome 19 showed strong associations, with a minor allele dominant model (rs12980275, $P = 1.93 \times 10^{-13}$, and rs8099917, $P = 3.11 \times 10^{-15}$, respectively), with NVR to PEG-IFN-α/RBV treatment (**Table 1**). The rs8099917 lies between *IL28B* and *IL28A*, ~8 kb downstream from *IL28B* and ~16 kb upstream from *IL28A*. These associations reached genome-wide levels of significance for both SNPs in this initial GWAS cohort (Bonferroni criterion $P < 8.05 \times 10^{-8}$ (0.05/621,220)). The frequencies of minor allele–positive patients were much higher in the NVR group than in the VR group for both SNPs (74.3% in NVR, 12.5% in VR for rs12980275; 75.6% in NVR, 9.4% in VR for rs8099917). Notably, individuals homozygous for the minor allele were observed only in the NVR group. The VR group, as compared to the NVR group, showed genotype frequencies closer to those in the healthy Japanese population[15], yet the minor allele frequencies were slightly higher in the transient virologic response (TVR) group (23.1%, 15.4%) than in the SVR group (9.8%, 7.8%) (**Table 1**). We applied the Cochrane-Armitage test on all the SNPs and found a genetic inflation factor, λ, of 1.029 for the GWAS stage (**Supplementary Fig. 1**). We also carried out principal component analysis in 142 samples for the GWAS stage together with the HapMap samples (CEU, YRI, CHB and JPT) (**Supplementary Fig. 2**); this suggested that the effect of population stratification was negligible.

Table 1 Significant association of two SNPs (rs12980275 and rs8099917) with response to PEG-IFN-α/RBV treatment

					Null responder (NVR[a], *n* = 128)			Responder (VR[a], *n* = 186)			Responder (SVR[a], *n* = 140)			NVR vs. VR		NVR vs. SVR	
dbSNP rsID	Nearest gene	MAF[b] (allele)	Allele (1/2)	Stage	11	12	22	11	12	22	11	12	22	OR (95% CI)[c]	*P* value[d]	OR (95% CI)[c]	*P* value[d]
rs12980275	*IL28B*	0.15 (G)	A/G	GWAS	20 (25.6)	54 (69.2)	4 (5.1)	56 (87.5)	8 (12.5)	0 (0.0)	46 (90.2)	5 (9.8)	0 (0.0)	20.3 (8.3–49.9)	1.93×10^{-13}	26.7 (9.3–76.5)	7.41×10^{-13}
				Replication	10 (20.0)	37 (74.0)	3 (6.0)	101 (82.8)	21 (17.2)	0 (0.0)	73 (82.0)	16 (18.0)	0 (0.0)	19.2 (8.3–44.4)	5.46×10^{-15}	18.3 (7.6–44.0)	8.37×10^{-13}
				Combined	30 (23.4)	91 (71.1)	7 (5.5)	157 (84.4)	29 (15.6)	0 (0.0)	119 (85.0)	21 (15.0)	0 (0.0)	17.7 (10.0–31.3)	2.84×10^{-27}	18.5 (10.0–34.4)	3.99×10^{-24}
rs8099917	*IL28B*	0.12 (G)	T/G	GWAS	19 (24.4)	56 (71.8)	3 (3.8)	58 (90.6)	6 (9.4)	0 (0.0)	47 (92.2)	4 (7.8)	0 (0.0)	30.0 (11.2–80.5)	3.11×10^{-15}	36.5 (11.6–114.6)	5.00×10^{-14}
				Replication	11 (22.0)	37 (74.0)	2 (4.0)	108 (88.5)	14 (11.5)	0 (0.0)	78 (87.6)	11 (12.4)	0 (0.0)	27.4 (11.5–65.3)	9.47×10^{-18}	25.1 (10.0–63.1)	1.00×10^{-14}
				Combined	30 (23.4)	93 (72.7)	5 (3.9)	166 (89.2)	20 (10.8)	0 (0.0)	125 (89.3)	15 (10.7)	0 (0.0)	27.1 (14.6–50.3)	2.68×10^{-32}	27.2 (13.9–53.4)	1.11×10^{-27}

[a]NVR, null virologic response; VR, virologic response; SVR, sustained virologic response. The 186 VRs consisted of 46 transient virologic response (TVRs) and 140 SVRs. [b]Minor allele frequency and minor allele in 184 healthy Japanese individuals[15]. The MAF of the SNPs in SVR is similar to that of TVR group, whereas that of NVR is much higher (76.6%). [c]Odds ratio for the minor allele in a dominant model. [d]*P* value by χ^2 test for the minor allele dominant model.

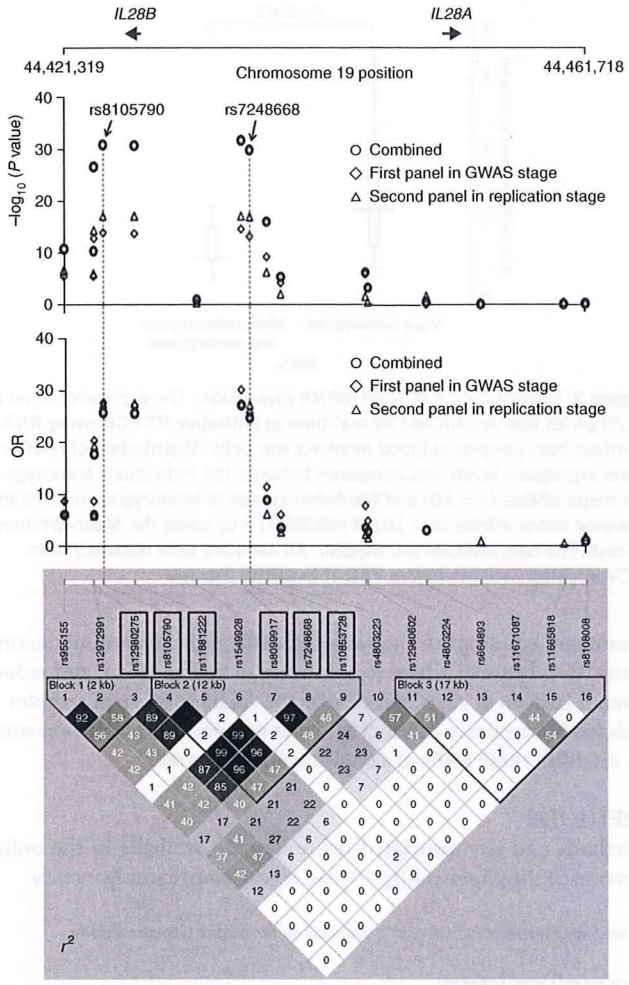


Figure 2 Genomic structure, *P* value and OR plots in association analysis and LD map around *IL28B* and *IL28A* (chr.19, nucleotide positions 44421319–44461718; build 35). *P* values by the χ^2 test for minor allele dominant effect model are shown for the first panel of 142 samples in the GWAS stage, the second panel of 172 samples in the replication stage, and the combined analysis. Below are estimates of pairwise r^2 for 16 SNPs selected in the replication study using a total of 314 Japanese patients with HCV treated with PEG-IFN-α/RBV. Boxes indicate the significantly associated SNPs with response to PEG-IFN-α/RBV treatment both in the GWAS stage and in the replication stage. Dotted lines indicate the region with the strongest associations from the positions of rs8105790 to rs7248668.

We analyzed the region of ~40 kb (chr. 19, nucleotide positions 44421319–44461718; build 35) containing the significantly associated SNPs (rs12980275 and rs8099917) using Haploview software for linkage disequilibrium (LD) and haplotype structure based on the HapMap data for individuals of Japanese ancestry. The LD blocks were analyzed using the four-gamete rule, and four blocks were observed (**Supplementary Fig. 3**). We selected 16 SNPs for both replication study and high-density association mapping, including tagging SNPs estimated on the basis of the haplotype blocks, one SNP located within *IL28B* (rs11881222) and the significantly associated SNPs from the GWAS stage (rs12980275 and rs8099917) (**Supplementary Table 1**).

To validate the results of the GWAS stage, 16 SNPs selected for the replication stage, including the original SNPs, were genotyped using the DigiTag2 assay in an independent set of 172 Japanese patients with HCV treated with PEG-IFN-α/RBV treatment (50 NVR and 122 VR samples), together with the first panel of 142 samples analyzed in the GWAS stage (**Supplementary Table 1**). The associations of the original SNPs were replicated in the replication cohort of 172 patients ($P = 5.46 \times 10^{-15}$, OR = 19.2 for rs12980275; $P = 9.47 \times 10^{-18}$, OR = 27.4 for rs8099917; **Table 1**). The combined *P* values for both stages reached 2.84×10^{-27} (OR = 17.7; 95% CI = 10.0–31.3) and 2.68×10^{-32} (OR = 27.1; 95% CI = 14.6–50.3), respectively (**Table 1**). Notably, when we compared the SVR ($n = 140$) with the NVR group ($n = 128$), the original two SNPs (rs12980275 and rs8099917) again showed strong associations: both *P* values and ORs were similar to those observed in the comparison between VR and NVR, and the combined *P* values for both stages reached 3.99×10^{-24} (OR = 18.5; 95% CI = 10.0–34.4) and 1.11×10^{-27} (OR = 27.2; 95% CI = 13.9–53.4), respectively (**Table 1**). Comparing SVR ($n = 140$) versus NVR plus TVR ($n = 174$), we again found that these SNPs were significantly associated ($P = 1.71 \times 10^{-16}$, OR = 8.8; 95% CI 5.1–15.4 for rs12980275; $P = 1.18 \times 10^{-18}$, OR = 12.1; 95% CI 6.5–22.4 for rs8099917, **Supplementary Table 2**), suggesting that these SNPs would predict NVR as well as SVR before PEG-IFN-α/RBV therapy.

Among the newly analyzed SNPs in the replication study, six (rs12980275, rs8105790, rs11881222, rs8099917, rs7248668 and rs10853728) showed significant associations both in the GWAS stage ($P < 8.05 \times 10^{-8}$) and in the replication stage ($P < 0.0031$ (0.05/16)) after Bonferroni correction. These SNPs are located within a 15.7-kb region that includes *IL28B* (**Fig. 2** and **Supplementary Table 1**). In particular, the strongest associations with NVR were observed for four SNPs, rs8105790, rs11881222, rs8099917 and rs7248668, that are located in the downstream flanking region, the third intron and the upstream flanking region of *IL28B*. The combined *P* values for these polymorphisms were 1.98×10^{-31} (OR = 25.7; 95% CI = 13.9–47.6), 2.84×10^{-31} (OR = 25.6; 95% CI = 13.8–47.3), 2.68×10^{-32} (OR = 27.1; 95% CI = 14.6–50.3) and 1.84×10^{-30} (OR = 24.7; 95% CI = 13.3–45.8), respectively (**Supplementary Table 1**). We then sequenced this region to identify further variants and found three SNPs (rs8103142, rs28416813 and rs4803219) located in the third exon, the first intron and the upstream flanking region of *IL28B*, and a few infrequent variations. These SNPs also showed strong associations in the combined dataset of 128 NVR and 186 VR samples ($P = 1.40 \times 10^{-29}$, OR = 26.6 for rs8103142; $P = 5.52 \times 10^{-28}$, OR = 22.3 for rs28416813; $P = 2.45 \times 10^{-29}$, OR = 23.3 for rs4803219; **Supplementary Table 3**). We also performed LD and haplotype analyses with seven SNPs. These SNPs were in strong LD, and the risk haplotype showed a level of association similar to those of individual SNPs ($P = 1.35 \times 10^{-25}$, OR = 11.1; 95% CI = 6.6–18.6) (**Table 2**). These results suggest that the association with NVR was primarily driven by one of these SNPs.

Table 2 Association analysis of response to treatment by *IL28B* haplotype

SNP							Frequencies			
rs8105790	rs11881222	rs8103142	rs28416813	rs4803219	rs8099917	rs7248668	NVR group	VR group	*P* value	OR (95% CI)
T	**A**	**T**	C	C	T	G	0.543	0.942	1.81×10^{-32}	0.1 (0.04–0.12)
C	**G**	**C**	G	T	G	A	0.387	0.054	1.35×10^{-25}	11.1 (6.6–18.6)

Association analysis of haplotypes consisting of seven SNPs with response to PEG-IFN-α/RBV treatment in 314 Japanese patients with HCV. Boldface letters: rs11881222 (third intron); rs8103142 (third exon).

Table 3 Factors associated with NVR by logistic regression model

Factors	Odds ratio	95% CI	*P* value
rs8099917 (G allele)	37.68	16.71–83.85	<0.0001
Age	1.02	0.98–1.07	0.292
Gender (Female)	3.32	1.49–7.39	0.003
Re-treatment[a]	1.12	0.55–2.33	0.750
Platelet count	0.93	0.87–1.01	0.080
Aminotransferase level	1.00	0.99–1.00	0.735
Fibrosis stage[20]	1.10	0.73–1.66	0.658
HCV-RNA level	1.01	0.99–1.02	0.139

[a]Re-treatment, non-response to previous treatment with interferon-α (plus RBV).

To examine the relative contribution of factors associated with NVR, we used a logistic regression model. One tagging SNP located within *IL28B* (minor allele of rs8099917) was the most significant factor for predicting NVR, followed by gender (**Table 3**). Clinically, viral factors such as HCV genotype and HCV RNA level are important for the outcome of PEG-IFN-α/RBV therapy. Indeed, mean HCV-RNA level was significantly lower in SVR (SVR versus TVR, $P = 0.002$; SVR versus NVR, $P = 0.016$; **Supplementary Table 4**). Mean platelet count and the proportion of mild fibrosis (F1–F2) were significantly higher in SVR than in NVR.

Real-time quantitative PCR assays in peripheral blood mononuclear cells revealed a significantly lower level of *IL28* mRNA expression in individuals with the minor alleles (**Fig. 3**), suggesting that variant(s) regulating *IL28* expression is associated with a response to PEG-IFN-α/RBV treatment. *IL28B* encodes a cytokine distantly related to type I (α and β) interferons and the interleukin (IL)-10 family. This gene and *IL28A* and *IL29* (encoding IL-28A and IL-29, respectively) are three closely related cytokine genes that encode proteins known as type III IFNs (IFN-λs) and that form a cytokine gene cluster at chromosomal region 19q13 (ref. 16). The three cytokines are induced by viral infection and have antiviral activity[16,17]. All three interact with a heterodimeric class II cytokine receptor that consists of IL-10 receptor beta (IL10Rβ) and IL-28 receptor alpha (IL28Rα, encoded by *IL28RA*)[16,17], and they may serve as an alternative to type I IFNs in providing immunity to viral infection.

Notably, a recent report showed that the strong antiviral activity evoked by treating mice with TLR3 or TLR9 agonists was significantly reduced in both *IL28RA*$^{-/-}$ and *IFNAR*$^{-/-}$ mice, indicating that IFN-λ is important in mediating antiviral protection by ligands for TLR3 and TRL9 (ref. 18). IFN-λ induced a steady increase in the expression of a subset of IFN-stimulated genes, whereas IFN-α induced the same genes with more rapid and transient kinetics[19]. Therefore, it is possible that IFN-λ induces a slower but more sustained response that is important for TLR-mediated antiviral protection. This might be one of the ways that a genetic variant regulating *IL28* expression influences the response to PEG-IFN-α/RBV treatment. Further research will be required to fully understand the specific mechanism by which a genotype might affect the response to treatment.

In conclusion, the strongest associations with NVR were observed for seven SNPs, rs8105790, rs11881222, rs8103142, rs28416813, rs4803219, rs8099917 and rs7248668, that are located in the downstream flanking region, the third intron, the third exon, the first intron and the upstream flanking region of *IL28B*. Further studies following our report of this robust genetic association to NVR may make it possible to develop a pre-treatment predictor of which individuals are likely to respond to PEG-IFN-α/RBV treatment. This would remove the need for the initial 12–24 weeks of treatment that is currently used as a basis for a clinical decision about whether treatment should be continued. That would allow better targeting of PEG-IFN-α/RBV treatment, avoiding the unpleasant side effects that commonly accompany the treatment where it is unlikely to be beneficial, and reduce overall treatment costs. Because of the small number of samples in this study, we plan to conduct a further prospective multicenter study to establish these SNPs as a clinically useful marker.

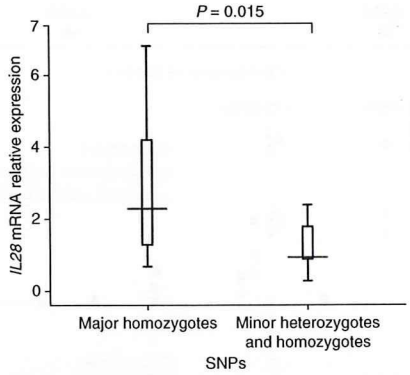


Figure 3 Quantification of *IL28* mRNA expression. The expression level of *IL28* genes was determined by real-time quantitative RT-PCR using RNA purified from peripheral blood mononuclear cells. Distribution of relative gene expression levels was compared between the individuals homozygous for major alleles ($n = 10$) and the heterozygous or homozygous individuals carrying minor alleles ($n = 10$) of rs8099917 by using the Mann-Whitney *U*-test. The bars indicate the median. All samples were obtained from HCV-infected patients before PEG-IFN-α/RBV therapy.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This study was supported by a grant-in-aid from the Ministry of Health, Labour, and Welfare of Japan (H19-kannen-013). This study is based on 15 multicenter hospitals throughout Japan, in the Hokkaido area (Hokkaido University Hospital), Kanto area (Saitama University Hospital; Konodai Hospital; Musashino Red Cross Hospital; Tokyo Medical and Dental University Hospital), Koshin area (Shinshu University Hospital; Kanazawa University Hospital), Tokai area (Nagoya City University Hospital), Kinki area (Kyoto Prefectural University of Medicine Hospital; National Hospital Organization Osaka National Hospital; Hyogo College of Medicine Hospital) and Chugoku/Shikoku area (Tottori University Hospital; Ehime University Hospital; Yamaguchi University Hospital; Kawasaki Medical College Hospital). We thank Y. Uehara-Shibata, Y. Ogasawara, Y. Ishibashi and M. Yamaoka-Sageshima (Tokyo University) for technical assistance; A. Matsumoto (Shinshu), K. Naiki (Saitama), K. Nishimura (Kyoto), H. Enomoto (Hyogo), K. Oyama (Tottori) and the Ochanomizu Liver Conference Study Group for collecting samples; M. Watanabe (Tokyo Medical and Dental University), S. Kaneko (Kanazawa University) and M. Onji (Ehime University) for their advice throughout the study; and H. Ito (Aichi Cancer Center) for conducting statistical analyses.

AUTHOR CONTRIBUTIONS

Study design and discussion: Y.T., N.N., N.M., K.T., M.M.; sample collection: Y.T., M.K., K.M., N.S., M.N., M.K., K.H., S.H., Y.I., E.M., E.T., S.M., Y.M., M.H., A.S., Y.H., S.N., I.S., M.I., K.I., K.Y., F.S., N.I.; genotyping: N.N.; statistical analysis: N.N., A.K., K.I.; quantitative RT-PCR: M.S.; manuscript writing: Y.T., N.N., K.T., M.M.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Ray Kim, W. Global epidemiology and burden of hepatitis C. *Microbes Infect.* **4**, 1219–1225 (2002).
2. Manns, M.P. *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* **358**, 958–965 (2001).

3. Fried, M.W. *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* **347**, 975–982 (2002).
4. Hadziyannis, S.J. *et al.* 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 (2004).
5. Bruno, S. *et al.* Peginterferon alfa-2b plus ribavirin for naive patients with genotype 1 chronic hepatitis C: a randomized controlled trial. *J. Hepatol.* **41**, 474–481 (2004).
6. Sezaki, H. *et al.* Poor response to pegylated interferon and ribavirin in older women infected with hepatitis C virus of genotype 1b in high viral loads. *Dig. Dis. Sci.* **54**, 1317–1324 (2009).
7. Fried, M.W. Side effects of therapy of hepatitis C and their management. *Hepatology* **36**, S237–S244 (2002).
8. Pascu, M. *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* **53**, 1345–1351 (2004).
9. Shirakawa, H. *et al.* Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. *Hepatology* **48**, 1753–1760 (2008).
10. Akuta, N. *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.* **46**, 403–410 (2007).
11. Walsh, M.J. *et al.* Non-response to antiviral therapy is associated with obesity and increased hepatic expression of suppressor of cytokine signalling 3 (SOCS-3) in patients with chronic hepatitis C, viral genotype 1. *Gut* **55**, 529–535 (2006).
12. Gao, B., Hong, F. & Radaeva, S. Host factors and failure of interferon-alpha treatment in hepatitis C virus. *Hepatology* **39**, 880–890 (2004).
13. Matsuyama, N. *et al.* The dinucleotide microsatellite polymorphism of the IFNAR1 gene promoter correlates with responsiveness of hepatitis C patients to interferon. *Hepatol. Res.* **25**, 221–225 (2003).
14. Tsukada, H. *et al.* A polymorphism in MAPKAPK3 affects response to interferon therapy for chronic hepatitis C. *Gastroenterology* **136**, 1796–1805 (2009).
15. Nishida, N. *et al.* Evaluating the performance of Affymetrix SNP Array 6.0 platform with 400 Japanese individuals. *BMC Genomics* **9**, 431 (2008).
16. Kotenko, S.V. *et al.* IFN-λs mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat. Immunol.* **4**, 69–77 (2003).
17. Sheppard, P. *et al.* IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat. Immunol.* **4**, 63–68 (2003).
18. Ank, N. *et al.* An important role for type III interferon (IFN-lambda/IL-28) in TLR-induced antiviral activity. *J. Immunol.* **180**, 2474–2485 (2008).
19. Marcello, T. *et al.* Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology* **131**, 1887–1898 (2006).
20. Desmet, V.J., Gerber, M., Hoofnagle, J.H., Manns, M. & Scheuer, P.J. Classification of chronic hepatitis: diagnosis, grading and staging. *Hepatology* **19**, 1513–1520 (1994).

IL28B：C型肝炎治療効果を規定する遺伝子多型 C型肝炎に対するテーラーメイド治療の確立を目指したゲノムワイド関連解析

田中靖人* 徳永勝士** 溝上雅史***

索引用語：IL28B，SNPs，IFNラムダ，GWAS，HCV

1 背景

わが国のC型肝炎ウイルス(HCV)感染者は約200万人存在するとされ，わが国における最大の感染症である．HCVは一旦感染すると6～8割が慢性肝炎に移行し，自然に治ることはほとんどなく，多くは肝硬変・肝癌へと進展し，本邦では年間約2万5千人が肝がんで死亡しているのが現状である．

そのHCVの根治治療で，現時点で最強治療であるペグインターフェロン＋リバビリン併用療法で根治させることができるようになったが，日本人に最も多いGenotype 1型高ウイルス量の症例では50％程度の根治しか得られず，約20％はペグインターフェロン＋リバビリン併用療法が全く効かないのが現状である．しかも，全治療期間が長期にわたるため，高齢者ではさまざまな副作用により減量・中断を余儀なくされる治療法である．したがって，治療前の効果予測が極めて重要と考えられる．その治療効果予測因子として，HCVゲノタイプ，ウイルス量，コア領域やNS5A領域のアミノ酸変異などのウイルス側因子に加えて，ペグインターフェロン，リバビリンのアドヒアランス(薬剤因子)，年齢，性差，肝線維化進展度，インシュリン抵抗性などの宿主側因子の重要性が多数報告されているが，それらの因子を総動員して解析しても治療前効果予測は約50％程度に留まる．

一方，ヒトゲノム計画の成功により，ヒト遺伝子は個人差として約300個に1個の遺伝子変異(SNP)が存在し，このSNPが個々の薬剤反応性や副作用に大きく関与することが続々と明らかとなってきている．近年，ゲノムワイドに均一に配置された約90万箇所(日

Yasuhito TANAKA et al: IL28B/genetic variation associated with response to pegylated interferon-alfa plus ribavirin therapy for chronic hepatitis C

*名古屋市立大学大学院医学研究科病態医科学〔〒467-8601 名古屋市瑞穂区瑞穂町字川澄1〕

東京大学大学院人類遺伝学，*国立国際医療センター国府台病院 肝炎・免疫研究センター

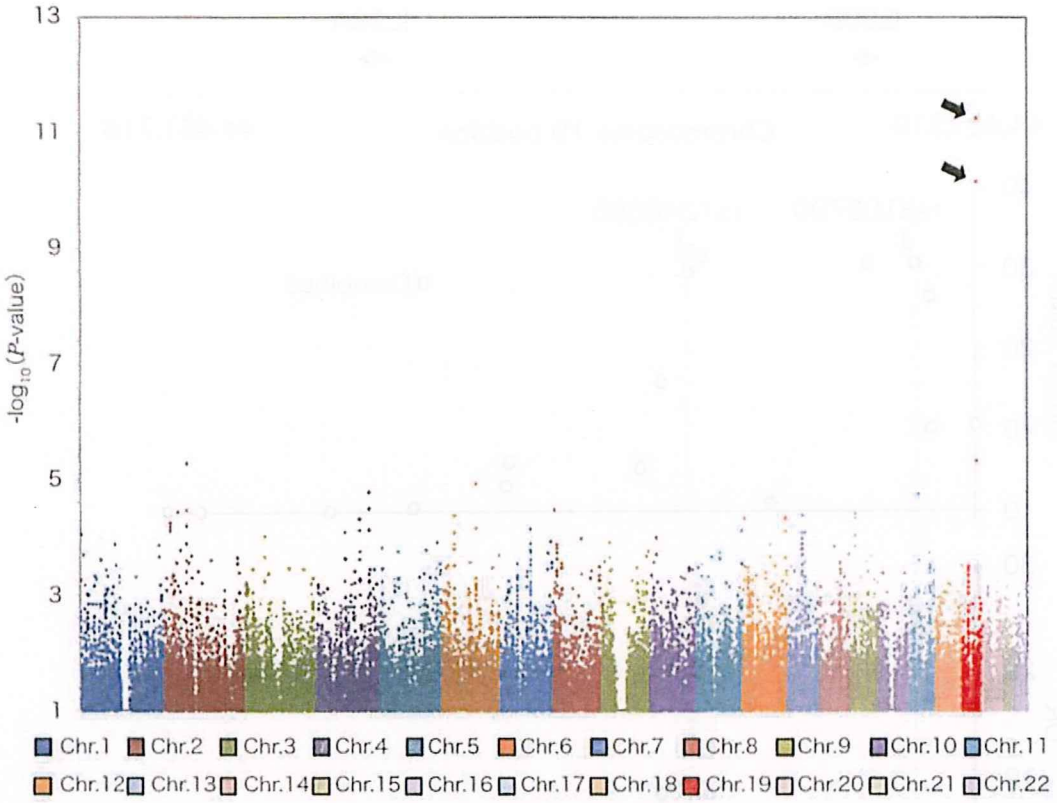


図1　ゲノムワイド関連解析：90万SNPs
Genome-wide association study（GWAS）
⇒19番染色体に有意な遺伝子多型（SNPs）

表1　治療無効に関連する因子

因子	Odds Ratio	95% Conf. Interval	P-values
rs8099917（マイナーアリル）	37.68	16.71-83.85	<0.0001
年齢	1.02	0.98-1.07	0.292
性別（女性）	3.32	1.49-7.39	0.003
再治療例	1.12	0.55-2.33	0.750
血小板数	0.93	0.87-1.01	0.080
Aminotransferase	1.00	0.99-1.00	0.735
Fibrosis stage	1.10	0.73-1.66	0.658
HCV-RNA量	1.01	0.99-1.02	0.139

本人では62万箇所）のSNPsを一括タイピングすることが可能になり，病態進展に多因子が関与すると想定されてきたⅡ型糖尿病，クローン病，B型慢性肝炎などにおいて疾患感受性遺伝子の同定が矢継ぎ早に報告されている[1~8]．

2　IL28B遺伝子多型と治療効果

われわれは，ペグインターフェロン＋リバビリン併用療法の有効性に関連するSNPsを

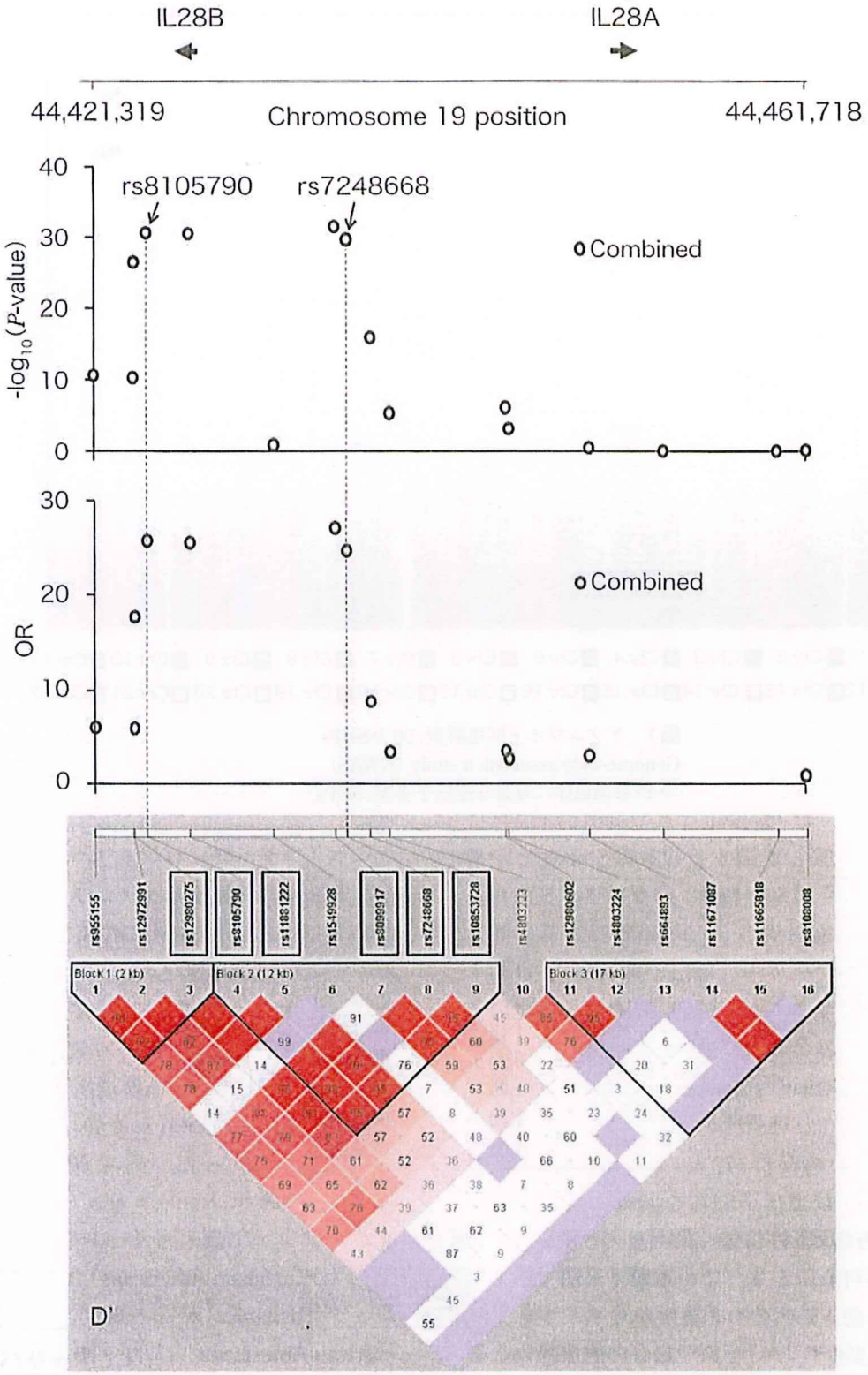


図2　IL28B，IL28A周辺のLD map．D'は連鎖不平衡係数のことで，連鎖不平衡の強さを示している．

同定するために，ゲノムワイド関連解析を実施した．すなわち，ペグインターフェロン＋リバビリン併用療法が有効(再燃例も含む)であった日本人患者と無効であった患者142人に関して，ヒト遺伝子の中で個人差があるとされる約90万箇所をAffymetrix Genome-Wide Human SNP Array 6.0（以下，SNP Array 6.0)を用い分析した結果，19番染色体のIL28B遺伝子周辺に治療無効に関連する有意なSNPsを発見した(図1)[9]．この代表的なSNPであるrs8099917（マイナーアリルG)を持つHCV患者群は，危険率約30倍の確率($P = 2.68 \times 10^{-32}$)でペグインターフェロン＋リバビリン併用療法が無効となることがわかった．さらに，年齢，性別，線維化，治療歴，ウイルス量，IL28B SNP, ALT，血小板にて多変量解析を行った結果，IL28B SNP（37.68 (16.71-83.85)，p<0.001）と女性(3.32（1.49-7.39)．p=0.003)が治療無効に関与する因子であった(表1)．

SNPなどの遺伝子多型は連鎖することが多いため，1つのSNPなどの遺伝子多型だけでなく近接し連鎖する複数のSNPなどの遺伝子多型を組み合わせ，ハプロタイプを作った上で，検討することが重要である．例えば，このマーカー（遺伝子)内のハプロタイプを表現するタグ(代表）SNPなどの遺伝子多型を組み合わせて検査することで，C型肝炎の治療効果の予測が，より一層，効率的かつ確実となる可能性が高いからである．別のコホート(検証群172人)を用いて，IL28B遺伝子および遺伝子周辺を詳細に検討した結果，治療反応性に強く関与するSNPsは複数存在し，しかもすべてが連鎖不平衡であった(図2)．ハプロタイプ解析からマイナーアリル(リスクアリル)を持つ場合の治療無効となるオッズ比は11.1 ($P=1.35 \times 10^{-25}$)であった．

IL28BはIFNλ3（ラムダ）と呼ばれ，類似の構造を持つIL28A (λ2)およびIL29 (λ1)の存在も知られており，特にペグ化IFNλ1は，すでに臨床試験(phase Ⅰb)がヨーロッパで開始され，副作用が少なくその有効性が期待されている．IL28B遺伝子は19番染色体長腕に位置し，約1.5 kと非常に小さいが，その詳細な機能は不明である．唯一，IFNλは共通のクラスⅡサイトカインレセプター（IL-28R)に結合し，ISGの発現レベルを向上させ，抗ウイルス活性を発揮することが報告されているが(図3)，実際の臨床に使用されているIFNαやβに比べてISGの誘導は弱い．しかし，今回同定されたIL28Bマイナーアリルを持つグループではIL28遺伝子発現レベルが低く(図4)，治療効果別に検討してみても無効群でその発現レベルが低いことから，IL28の補填あるいは誘導する治療が期待される．

3 IL28B遺伝子多型に関する世界の状況

ペグインターフェロン＋リバビリン併用療法の有効性に関連するゲノムワイド関連解析の結果は，ほぼ同時期に欧米でも報告された[10,11]．いずれの報告もキーワードは"IL28B"であった．Geらの報告では，白人(871人)，黒人(191人)，ヒスパニック(75人)で検討した結果，白人おいてIL28B遺伝子から3 kb上流のSNP (rs12979860)が著効に強く関連することがわかった．興味深いことに，rs12979860のメジャーアリル(C-allele)の頻度は，アジアで最も多く(80～90％)，続いて白人(European-Americans)およびヒスパニック(Hispanics)が70～80％，そして黒人(African-Americans）は30～50％と低値であり，この頻度が著効率と正に相関しているこ

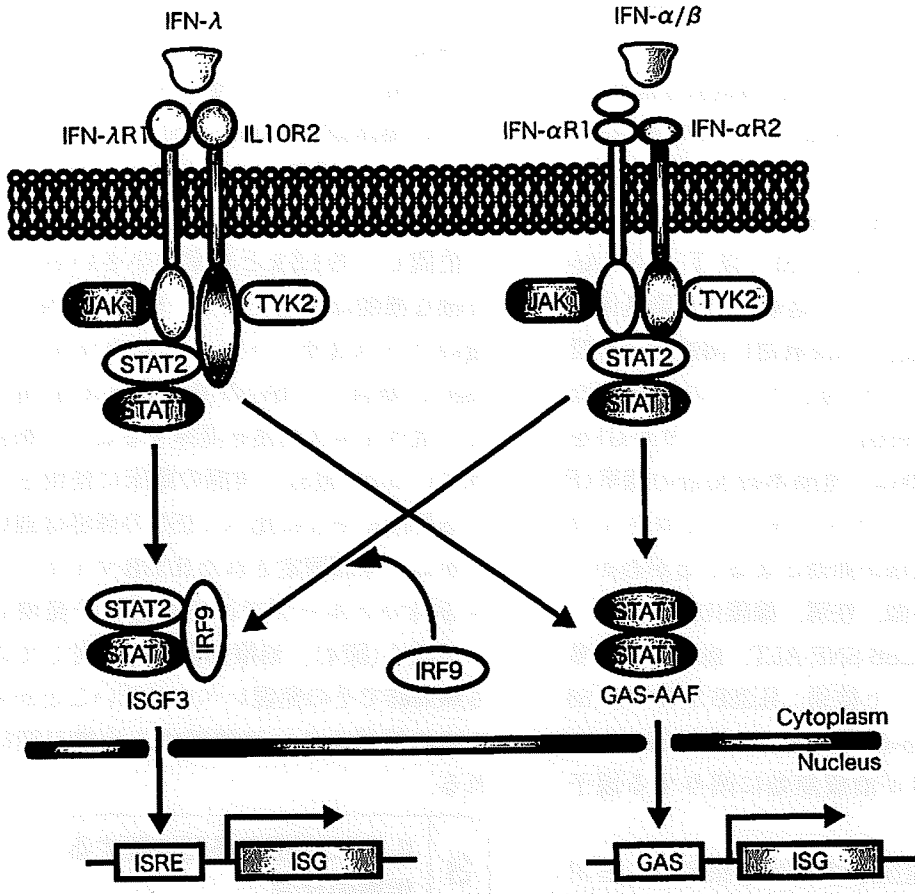


図3　IFNシグナル伝達系

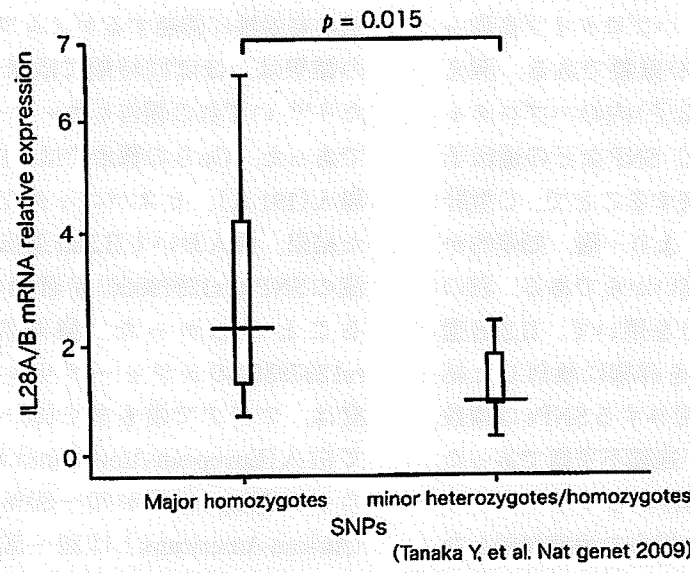
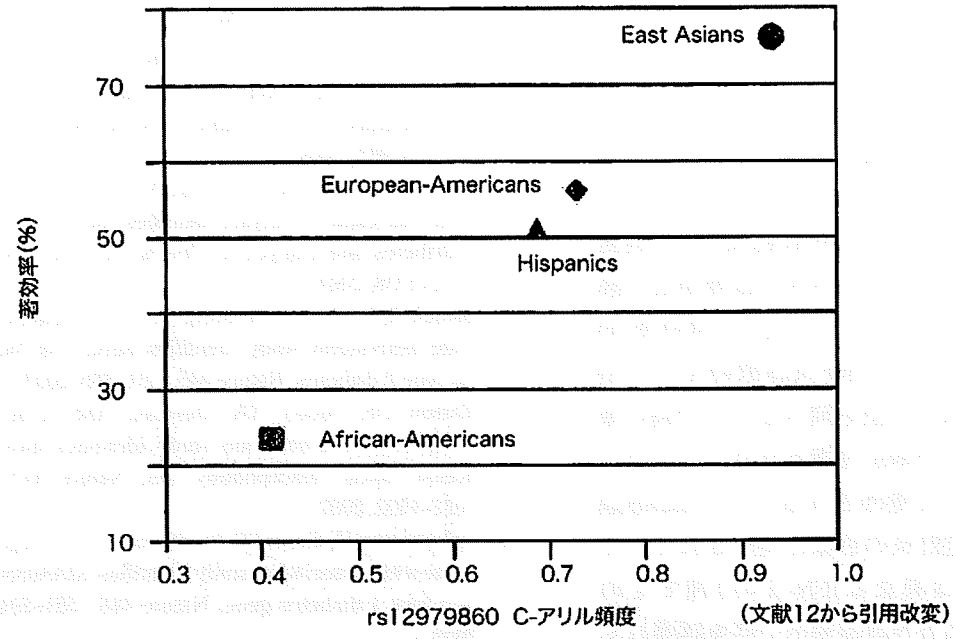


図4　アリル別のIL28A/B遺伝子発現レベル



(文献12から引用改変)

図5 人種別のrs12979860 C-アリル頻度と著効率

とが述べられている(図5). さらに, rs12979860はHCV自然治癒にも関連していることが報告され[12], IL28Bがウイルス排除に重要な役割を果たしている可能性が示唆された. Suppiah Vらは, オーストラリア(北ヨーロッパ起源)およびヨーロッパ諸国(イギリス, ドイツ, イタリア, オーストリア)で検討し, われわれと全く同じSNPであるrs8099917がペグインターフェロン+リバビリン併用療法の治療効果に有意に関連していることを報告している[11].

これらの論文の特徴をまとめてみると, 1) 人種によるマイナーアリル頻度(MAF)に差があり, これにより著効率に差がでている. 2) Caseの取り方が異なるため, オッズ比に差がでている. すなわち, われわれは無効群(NVR) vs. 無効群以外(TVR + SVR)で比較しているが, 他の2報は著効群(SVR) vs. 著効群以外(TVR + NVR)で比較している. われわれのデータからはSVR群とTVR群(再燃群)のMAFが類似しているので, TVR群をどちらのグループに含めて検討するかによってオッズ比は変わる可能性がある. 3) NVRの定義を12週までペグインターフェロン, リバビリンの両剤が80%以上の投与量にも関わらず, 12週で2 log減少がなく24週でHCV-RNAが消失しない例と厳密に設定したため, 有意差がより明確になったものと思われる.

全く独立した3つのグループから, しかも人種を超えた遺伝要因であることが一気に証明されたことで, 今後IL28B遺伝子多型に基づいたテーラーメイド治療が展開されることが予測される. また, ペグインターフェロン+リバビリン併用無効例のほとんどは, IL28A/B (IFNλ2/3)の産生不十分な遺伝子型であることを考慮すると, IFNλ2あるいはIFNλ3を補填あるいは誘導する方法の確立が必要と考えられ, 新規治療薬の開発が期待される.

4 おわりに

実際の臨床において，ペグインターフェロン＋リバビリン併用療法の前にこの遺伝子多型(SNPs)を測定することで，根治の見込める患者群を高い確率(的中率約80%)で選別できるし，効かない人たちからは無用な苦痛や出費から免れることができる(的中率約80%)．すなわち，C型肝炎診療の中で，治療前にこの遺伝子多型を調べることで高い確率で治療効果の予測が可能となり，テーラーメード医療として期待される．このIL28B遺伝子は通常C型肝炎の治療に使用されているIFN-αやβとは異なるIFN-λの1種でその下流に存在するIFN誘導遺伝子群を誘導して抗ウイルス効果をもたらすので，今後このIL28Bを増強する新規薬剤を開発することで，現在ペグインターフェロン＋リバビリン併用療法で効かない人達や効果の不十分な人達も根治が望める可能性がある．現在，肝炎治療の効果的促進(経済的負担軽減)をはかるため2008年4月1日より「B型・C型肝炎患者医療給付事業」がスタートしているが，これらの公費助成も効率的運用が図れることを意味する．

謝　辞

「テーラーメイド治療を目指した肝炎ウイルスデータベース構築に関する研究」の一環として行われる．H19～21：厚労省科研費肝炎等克服緊急対策研究事業

文　献

1) Barrett JC, Hansoul S, Nicolae DL et al : Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40 : 955–962, 2008
2) Saxena R, Voight BF, Lyssenko V et al : Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316 : 1331–1336, 2007.
3) Sladek R, Rocheleau G, Rung J et al : A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445 : 881–885, 2007
4) Easton DF, Pooley KA, Dunning AM et al : Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447 : 1087–1093, 2007
5) Hakonarson H, Grant SF, Bradfield JP et al : A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 448 : 591–594, 2007
6) Shi J, Levinson DF, Duan J et al : Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 460 : 753–757, 2009
7) Wang K, Zhang H, Ma D et al : Common genetic variants on 5p14.1 associate with autism spectrum disorders. Nature 459 : 528–533, 2009
8) Weiss LA, Arking DE, Daly MJ et al : A genome-wide linkage and association scan reveals novel loci for autism. Nature 461 : 802–808, 2009
9) Tanaka Y, Nishida N, Sugiyama M et al : Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41 : 1105–1109, 2009
10) Ge D, Fellay J, Thompson AJ et al : Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461 : 399–401, 2009
11) Suppiah V, Moldovan M, Ahlenstiel G et al : IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 41 : 1100–1104, 2009
12) Thomas DL, Thio CL, Martin MP et al : Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 461 : 798–801, 2009

*　　*　　*

Hepatology Research 2010; 40: 14–30 doi: 10.1111/j.1872-034X.2009.00601.x

Review Article

Geographical and genetic diversity of the human hepatitis B virus

Fuat Kurbanov,[1] Yasuhito Tanaka[1] and Masashi Mizokami[1,2]

[1]*Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, and* [2]*Research Center for Hepatitis and Immunology, International Medical Center of Japan, Konodai Hospital, Konodai, Japan*

Hepatitis B virus (HBV) is one of the most widely distributed viruses that infect humankind. Distinct clinical and virological characteristics of the HBV-infection have been reported in different geographical parts of the world and are increasingly associated with genetic diversity of the infecting virus. HBV is classified into genotypes and subgenotypes that are associated with ethnicity and geography. The genetic diversity of HBV in its various aspects has been the subject of extensive investigations during the last few decades. Since molecular epidemiology research tools have become widely available, the number of new publications in this field has grown exponentially. This review summarises the recent publications on the geographical distribution of genetic variants of HBV, and proposes updated criteria for the identification of new genotypes and subgenotypes of the virus.

Key words: genotypes, hepatitis B virus, molecular epidemiology, recombination

INTRODUCTION

FOUR DECADES AGO, in 1965, the "Australian antigen" (now reffered to as the hepatitis B surface antigen, HBsAg) was detected in hemophilia patients[1] and was soon specifically associated with hepatits B virus (HBV).[2,3] Three decades ago, the HBV strains were divided into nine major serotypes based on antigenic determinants of HBsAg.[4] Two decades ago, the classification of the HBV by genome nucleotide sequence divergence was proposed.[5] A decade ago, a "unique phylogenetic cluster within genotype A strains was described, triggering consecutive investigations on HBV subgenotypes.[6]

Outlined are the most important key-events in a chain of findings that accumulated in the current image of the HBV diversity. The chain was tortuous before powerful tools such as PCR and nucleotide sequencing became available to researchers. These tools enabled the annual number of publications in this field to grow exponentially. The current review will discuss the most recently published observations on HBV diversity, particularly their geographical distribution, and taxonomical aspects.

Correspondence: Dr Masashi Mizokami, The Research Center for Hepatitis and Immunology, International Medical Center of Japan, Konodai Hospital, 1-7-1, Kounodai, Ichikawa 272-8516, Japan.
Email: mmizokami@imcjk2.hosp.go.jp
Received 30 March 2009; revision 10 August 2009; accepted 25 August 2009.

CURRENTLY KNOWN HBV GENOTYPES

A TOTAL OF 18 complete genome sequences were available for comparison when the first four genotypes of HBV (designated A to D, consecutively) were originally proposed and divergence exceeding 8% of the complete genome was indicated as a criterion for genotype identification.[5] Almost at the same time, genotyping based on the phylogenetic clustering was proposed.[7] Consecutively, by sequencing the HBsAg coding region, four new strains were designated as novel genotypes E and F based on both, percent evaluation of nucleotide divergence and phylogenetic analysis. This added new criteria for genotyping; 4% of nucleotide divergence in HBsAg coding sequence.[8] Shortly after, the genotype F was confirmed by an analysis of the full genome sequence.[9] Relatively recent reports identified the last two of the 8 currently known genotypes, genotype G[10] and H.[11]

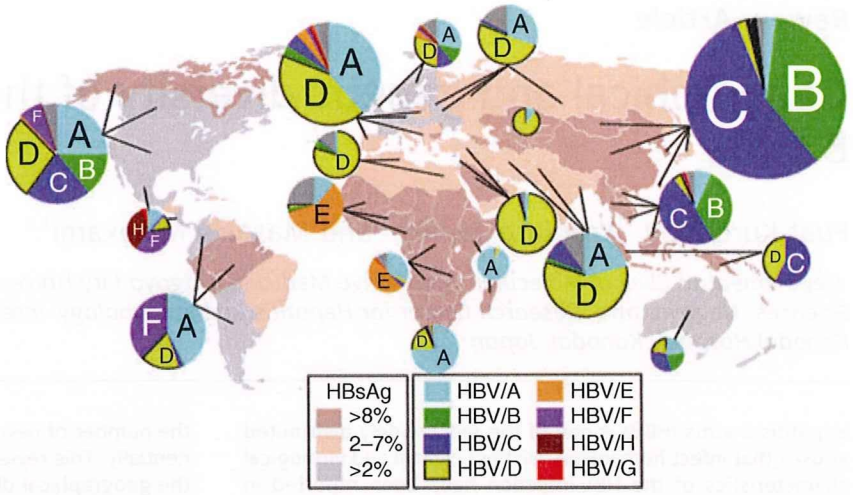


Figure 1 Hepatitis B virus (HBV) infection endemicity is based on the 2005 estimation of hepatitis B surface antigen (HBsAg) seroprevalence (Centers for Disease Control and Prevention Travelers' Health: Yellow Book Chapter 4 – Prevention of Specific Infectious Diseases: Hepatitis, Viral, Type B URL: http://wwwn.cdc.gov/travel/yellowbook/ch4/hep-b.aspx). Percentile distribution of genotypes is indicated for each geographic region.

DISTRIBUTION OF GENOTYPES IN THE WORLD

EPIDEMIOLOGICAL GEOGRAPHICAL DISTRIBUTION of HBV genotypes is being continuously investigated in different parts of the world. For this review we summarised epidemiological studies published within the last decade. A total of 256 papers were analysed. The results of the geographical distribution of genotypes are graphically summarised in Figure 1A. The detailed summary presented in Table 1 contains the number of examined HBV carriers in each particular subregion, and the corresponding references.

The total number of HBV-infected individuals genotyped during the last 10 years consist of approximately 45 000, with more than half of that number in Eastern Asia. From the data accumulated, it can be seen that a single genotype can only be predominately found in

Table 1 Prevalence of hepatitis B virus genotypes in different geographical regions

Geographical subregion	*n*	A	B	C	D	E	F	H	G	Mixed	UT	References
Eastern Africa	43	93.0			2.3						4.7	12–14
South Africa	404	74.3	0.7	1.5	19.3	1.2					3.0	15–17
Central Africa	126	31.0		3.2	3.2	49.2			1.6	11.9		18–20
Western Africa	759	11.3	0.3		1.6	59.2				2.9	24.8	21–29
Northern Africa	331	0.3	5.7	0.9	79.2					9.4	4.5	30–34
Western Asia	1652	0.9	0.2	0.5	94.8		0.1				3.5	35–58
Central Asia	118	11.0		0.8	88.1							59–61
Southern Asia	3023	21.5	0.9	8.9	58.7					3.9	6.1	25,62–79
East Europe	1674	30.5	0.9	0.7	50.4					6.0	11.5	25,80–86
European Union	4968	38.5	3.3	4.3	42.6	3.4	1.4	0.2	0.7	2.0	3.7	10,20,25,81,87–116
North Europe	442	28.3	10.9	10.6	30.8	5.0	1.4	0.2	0.2	2.0	10.6	117–122
North America	3412	25.1	14.3	20.8	27.7	0.2	7.3	0.1	0.9		3.6	10,25,108,123–131
Central America	225	11.6		0.4	11.6		36.0	35.1	3.6	1.3	0.4	132–135
South America	1393	42.6	0.5	1.9	17.4	0.1	35.9		0.1	0.6	0.9	25,136–157
Atlantic Island	84	54.8		1.2	23.8		2.4			17.9		110
Southeastern Asia	2024	6.7	35.2	47.3	4.1	0.7	0.4		0.9	3.4	1.2	20,25,108,158–170
Eastern Asia	23577	2.0	36.9	55.0	2.2					1.9	1.9	25,108,171–255
Pacific Islands	274			57.7	42.3							225,239,256,257
Australia	132	22.7	22.7	31.8	21.2				0.8		0.8	258–260
TOTAL	44661	13.1	22.9	34.5	19.9	1.6	2.1	0.2	0.2	2.1	3.3	

UT, untypeable