with the lack of virologic response to IFN-based antiviral therapy. As a host-related factor, recent studies reported that genetic polymorphisms near the *IL28B* gene (rs8099917, rs12979860) on chromosome 19 are strongly associated with a resistance to the combination therapy in patients infected with HCV genotype 1 [Ge et al., 2009; Suppiah et al., 2009; Tanaka et al., 2009; McCarthy et al., 2010; Rauch et al., 2010]. Patients having the TT genotype at a polymorphic locus (rs8099917) near the *IL28B* gene show a favorable response to the combination therapy with PEG-IFN and ribavirin, whereas patients having the GG genotype or those who are TG heterozygote show a resistance to the therapy. As for virus-related factors, amino acid (AA) mutations at residue 70 in the HCV core region have been reported to be associated strongly with a resistance to PEG-IFN/ribavirin combination therapy in patients infected with HCV genotype 1b [Akuta et al., 2005, 2007a; Donlin et al., 2007]. Patients with the mutant-type AA (glutamine or histidine) at residue 70 in the HCV core region show a resistance to the combination therapy in comparison to those with the wild-type AA (arginine) at this residue. These host- and virus-related factors are both associated with the outcome of the combination therapy with PEG-IFN and ribavirin independently in a previous report [Hayes et al., 2011].

A previous study reported that the percentage of patients with the mutant-type AA at residue 70 of the HCV core region increases with the progression of chronic hepatitis, suggesting that the mutation of AA at residue 70 (from arginine to glutamine or histidine) occurs in the natural course of chronic HCV infection [Kobayashi et al., 2010a]. Several recent studies have reported a higher prevalence of the mutant-type AA at residue 70 in patients who have the TG/GG genotype of genetic polymorphism of rs8099917 near the *IL28B* gene, which is associated with an unfavorable response to the combination therapy with PEG-IFN and ribavirin, than in patients who have the TT genotype [Abe et al., 2010; Kobayashi et al., 2010b]. These reports suggest that the mutation of AA residue 70 of the HCV core region may occur more frequently in patients with the TG/GG genotype. Especially, the induction of this mutation may occur easily in patients who underwent PEG-IFN/ribavirin combination therapy and failed to clear HCV (non-sustained virologic response), wherein HCV obtained a resistance to combination therapy.

Mutation at HCV core region residue 70 has reportedly been associated with a hepatocarcinogenesis and an insulin resistance [Akuta et al., 2007b, 2009; Nakamoto et al., 2010]. In addition, a recent study reported that patients who have both the TG/GG genotype of rs8099917 near the *IL28B* gene and the mutant-type AA at residue 70 of the HCV core region have shown further resistance even to the triple therapy with telaprevir, PEG-IFN, and ribavirin [Akuta et al., 2010]. It is, therefore, important to clarify whether PEG-IFN/ribavirin combination therapy induces the mutation of the HCV core region residue 70 in patients who failed to eradicate HCV, and whether genetic polymorphism near the *IL28B* gene are correlated with this mutation. If so, some patients should not undergo the current standard combination therapy in order to prevent the acquisition of the resistance (i.e., mutation at residue 70).

The present study investigated the effects of the combination therapy with PEG-IFN and ribavirin and genetic polymorphisms near the *IL28B* gene on the mutation of HCV core region residue 70 in patients who failed to achieve a sustained virologic reponse.

PATIENTS AND METHODS

Patients and Treatment

Three hundred and forty six patients with chronic hepatitis C who had been infected with HCV genotype 1b (as assessed by amplification of core-gene sequences with polymerase chain reaction (PCR) using genotype-specific primers [Ohno et al., 1997]) and pretreatment HCV-RNA level of $>100 \times 10^3$ IU/ml [as assessed by a quantitative PCR assay (Amplicor GT-HCV Monitor, Version 2.0; Roche Molecular Systems, Pleasanton, CA)] underwent antiviral combination therapy with PEG-IFN and ribavirin between January, 2007 and December, 2009 at the Ogaki Municipal Hospital or the Nagoya University Hospital. Of these patients, 19 patients dropped out and their outcome could not be defined. Among the remaining 327 patients, 274 patients who gave written informed consent for genetic analyses were enrolled to the study (Fig. 1). No patients were coinfected with hepatitis B virus or human immunodeficiency virus.

All patients were given PEG-IFN alpha-2b (Pegintron, Schering-Plough, Tokyo, Japan) weekly and ribavirin (Rebetol, Schering-Plough) daily. The initial doses of PEG-IFN and ribavirin and the dose reductions were according to the manufacturer's recommendations. All patients were scheduled to undergo 48 weeks of the treatment. Some patients had an extended treatment duration of up to 72 weeks. In some patients, the treatment was discontinued before 48 weeks because they had a low likelihood of achieving a sustained virologic response, when serum HCV-RNA was positive 24 weeks after starting the therapy. The outcomes of the combination therapy were classified as a sustained virologic response when serum HCV-RNA became undetectable during the treatment and remained undetectable for 6 months after the treatment ended (i.e., eradication of HCV), a relapse when the serum HCV-RNA became undetectable during the treatment period but returned detectable after the treatment, and no-response when the serum HCV-RNA remained detectable during and after the treatment period.

The study protocol was in compliance with the Helsinki Declaration and was approved by the ethics committee of the Ogaki Municipal Hospital and the Nagoya University School of Medicine. Written

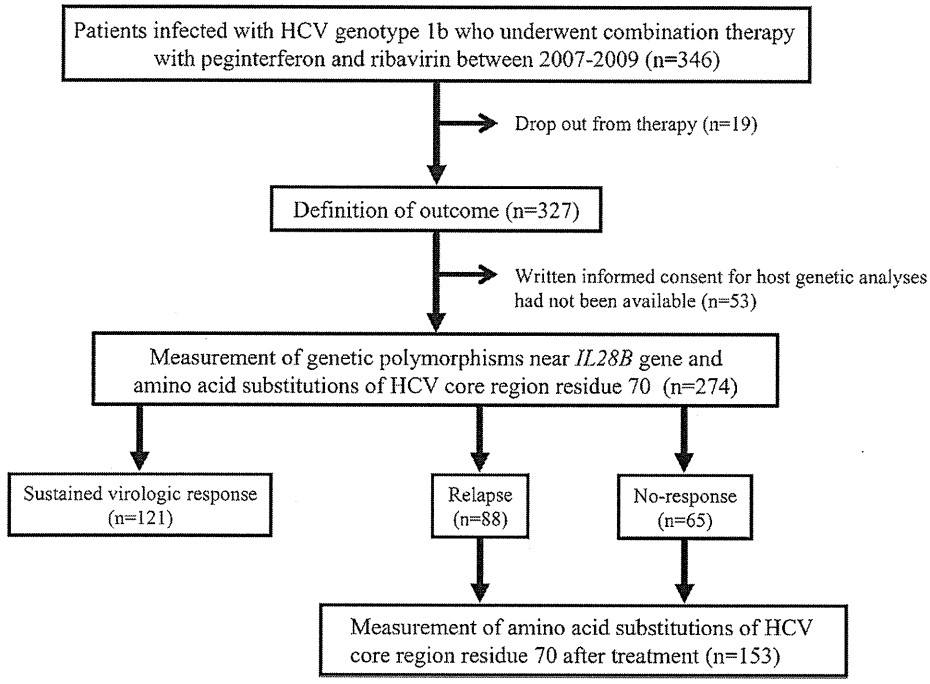

Fig. 1. Schematic representation of the study design.

informed consent was obtained from all patients prior to the study for the measurement of genetic polymorphism of rs8099917 near *IL28B* gene and AA substitution of HCV core region residue 70, and for the use of the laboratory data.

Measurements of Genetic Polymorphism Near the *IL28B* Gene and Amino Acid Substitution of the HCV Core Region Residue 70

Genotyping of polymorphisms of the rs8099917 locus near the *IL28B* gene was carried out in all 274 patients using a Taqman SNP assay (Applied Biosystems, Foster City, CA) according to the manufacturer's guidelines. A pre-designed and functionally tested probe was used for rs8099917 (C__11710096_10, Applied Biosystems).

The AA at residue 70 of the core region of HCV was measured before the treatment in all patients. In patients who failed to achieve a sustained virologic response, that is, patients who showed a relapse or no-response, the AA identity was measured at residue 70 after the treatment and compared pre- to post-treatment AA identity at this residue (Fig. 1). The AA at residue 70 after the treatment was measured in serum samples obtained at the end of treatment in patients who showed no-response. In patients with a relapse, it was measured in serum samples obtained upon the reappearance of HCV-RNA after the completion of the therapy. The AA identity was analyzed by direct nucleotide sequencing according to a previous report [Akuta et al., 2007c]. The primer pairs used for PCR for direct sequencing the HCV core region were 5′-GCCATAGTGGTCTGCGGAAC-3′ (outer, sense primer), 5′-GGAGCAGTCCTTCGTGACATG-3′ (outer, antisense primer), 5′-GCTAGCCGAGTAGTGTT-3′ (inner, sense primer), and 5′-GGAGCAGTCCTTCGTGACATG-3′ (inner, antisense primer).

Statistical Analysis

The chi-square test was used to analyze the differences in percentages between groups.

RESULTS

Patient Characteristics and the Outcome of the Combination Therapy

The characteristics of study patients are shown in Table I. The study patients comprised 139 males (50.7%) and 135 females (49.3%), with a mean age of 58.0 ± 10.4 years. The grade of liver fibrosis according to the METAVIR score [The French METAVIR Cooperative Study Group, 1994] was F0 in 31 patients (11.6%), F1 in 122 patients (45.9%), F2 in 75 patients (28.2%), and F3 in 38 patients (14.3%). Analysis of the genetic polymorphism of the rs8099917 near the *IL28B* gene indicated 202 patients (73.7%) had the TT genotype, three patients (1.1%) had the GG genotype, and the remaining 69 patients (25.2%) were TG heterozygous. Before the treatment, 204 patients (74.4%)

TABLE I. Baseline Characteristics of the Study Patients (n = 274)

Age (years)	55.9 ± 11.2
Sex (female/male)	135 (49.3)/139 (50.7)
Body weight (kg)	58.0 ± 10.4
Alanine aminotransferase (IU/L)	64.5 ± 56.3
Aspartate aminotransferase (IU/L)	53.7 ± 42.2
Gamma-glutamyl transpeptidase (IU)	49.7 ± 48.5
Alkaline phosphatase (IU/L)	267.9 ± 100.6
Albumin (g/dl)	4.07 ± 0.38
Total bilirubin (mg/dl)	0.79 ± 0.30
White blood cell count (/μl)	4933 ± 1331
Hemoglobin (g/dl)	14.0 ± 1.4
Platelet count ($\times 10^3$/μl)	164 ± 50
Liver histology-activity (A0/A1/A2/A3)[a]	2 (0.7)/147 (55.3)/99 (37.2)/18 (6.8)
Liver histology-fibrosis (F0/F1/F2/F3)[a]	31 (11.6)/122 (45.9)/75 (28.2)/38 (14.3)
HCV-RNA concentration ($\log_{10}$ IU/ml)[b]	6.34 ± 0.54
Genetic polymorphisms near the *IL28B* gene (TT/TG/GG)[b]	202 (73.7)/69 (25.2)/6 (2.2)
Amino acid at HCV core 70 (wild type/mutant type/both)[c]	204 (74.4)/64 (23.4)/6 (2.2)
Response (SVR/relapse/NR)	121 (44.2)/88 (32.1)/65 (23.7)

HCV, hepatitis C virus; SVR, sustained virologic response; NR, no-response.
Percentages are shown in parentheses.
[a]Liver biopsy was not performed in eight patients.
[b]rs8099917 genetic polymorphism
[c]Before the treatment.

carried HCV with the wild-type AA at residue 70 of the HCV core region, 64 patients (23.4%) carried the mutant-type AA at residue 70, and both the wild-type AA and the mutant-type AA were identified at residue 70 in the remaining six patients (3.5%).

As a final outcome, 121 patients (44.2%) achieved a sustained virologic response, 88 patients (32.1%) relapsed, and the remaining 65 patients (23.7%) showed no-response (Fig. 1). Treatment was discontinued before 48 weeks in 11 of 65 patients who showed no-response because HCV-RNA remained detectable in serum 24 weeks after starting the therapy. The identity of the AA 70 of the core region of HCV was determined after the treatment in serum obtained at the discontinuation of the therapy in these 11 patients. Table II shows the association between the genetic polymorphisms of the rs8099917 near the *IL28B* gene, the AA substitutions of the HCV core region residue 70, and the outcome of the combination therapy. The wild-type AA was more frequently identified at residue 70 in patients with the TT genotype in comparison to those with the TG/GG genotype (82.2% vs. 52.8%, $P < 0.0001$). The rate of a sustained virologic response was significantly higher in patients with the TT genotype than those with the TG/GG genotype (107 of 202 patients, 53.0% vs. 14 of 72 patients, 19.4%, $P < 0.0001$), as well as being higher in patients carrying HCV with the wild-type AA at residue 70 of the core region than those with the mutant-type AA at this residue (101 of 204 patients, 49.5% vs. 19 of 64 patients, 29.7%, $P = 0.0083$, one patient had both the wild-type and the mutant-type AAs).

Comparison of the Amino Acid at Residue 70 of the HCV Core Region Before and After the Combination Therapy in Patients Who Showed a Relapse or No-Response

Table III shows the comparison of the AA at residue 70 of the HCV core region before and after the combination therapy in patients who showed a relapse or no-response, according to the genetic polymorphisms of the rs8099917 near the *IL28B* gene. In three of five

TABLE II. Association Between the Genetic Polymorphisms Near the *IL28B* Gene, the Amino Acid at the HCV Core Region Residue 70, and the Final Outcome of Peginterferon/Ribavirin Combination Therapy

Genetic polymorphism of rs8099917 near *IL28B* gene	Amino acid at residue 70 of the HCV core region		
	Wild type (n = 204)	Mutant type (n = 64)	Wild type + mutant type (n = 6)
TT (n = 202)	166 (92/60/14)	31 (14/9/8)	5 (1/2/2)
TG/GG (n = 72)	38 (9/9/20)	33 (5/7/21)	1 (0/1/0)

Outcomes of the combination therapy with peginterferon and ribavirin are shown in parentheses as sustained virologic response/relapse/no-response.

TABLE III. Amino Acid Substitutions of HCV Core Region Residue 70 Before and After the Combination Therapy With Peginterferon and Ribavirin in No-Responders or Relapsers

Amino acid at HCV core region residue 70	After treatment		
Before treatment	Wild type	Wild + Mutant	Mutant type
(A) Genetic polymorphisms near the *IL28B* gene (rs8099917): TT (n = 91)			
No-responders (n = 24)			
Wild type (n = 14)	13	1	0
Wild + mutant (n = 2)	0	0	2
Mutant type (n = 8)	0	0	8
Relapsers (n = 71)			
Wild type (n = 60)	60	0	0
Wild + mutant (n = 2)	0	1	1
Mutant type (n = 9)	0	0	9
(B) Genetic polymorphisms near the *IL28B* gene (rs8099917): TG/GG (n = 57)			
No-responders (n = 41)			
Wild type (n = 20)	19	1	0
Wild + mutant (n = 0)	0	0	0
Mutant type (n = 21)	0	0	21
Relapsers (n = 17)			
Wild type (n = 9)	9	0	0
Wild + mutant (n = 1)	0	1	0
Mutant type (n = 7)	0	0	7

HCV, hepatitis C virus.

patients in whom both the wild-type and mutant-type AAs had been identified at residue 70 of the HCV core region before treatment, only the mutant-type AA was identified at this residue after the treatment. All three of these patients (two no-responders and one relapser) had the TT genotype of the rs8099917. Both the wild-type and mutant-type AAs were identified at residue 70 after the treatment in two no-responders in whom only the wild-type AA had been identified before the treatment. One of them had the TT genotype at the rs 8099917 and the other patient was TG heterozygous. No change in the HCV core region residue 70 was found after the treatment in patients with the mutant-type AA at this residue before the treatment.

DISCUSSION

The present study investigated whether the combination therapy with PEG-IFN and ribavirin causes the mutation of residue 70 of the HCV core region, and whether the genetic polymorphisms of the rs8099917 locus near the *IL28B* gene influence this mutation. It is thought to be important to verify this issue, because it may be advisable to avoid the treatment of patients who have the TG/GG genotypes by the combination therapy with PEG-IFN and ribavirin so as to avoid an acquisition of the further resistance to emerging new therapies against HCV, as well as to avoid a potential enhancement of hepatocarcinogenesis.

The mutation of the AA at residue 70 was not observed before and after the treatment in all patients who had failed to achieve a sustained virologic response. The mutant-type AA was identified solely at residue 70 after the treatment in three patients who had both the wild-type and the mutant-type AAs at residue 70 before the treatment. This could be due to the selection of HCV strains with the mutant-type AA at residue 70 by the combination therapy with PEG-IFN and ribavirin, as reported previously [Kurbanov et al., 2010]. In two patients who carried only the wild-type AA before the treatment, the HCV with the mutant-type AA at residue 70 was also detected with the persistence of the wild-type AA at this residue after the treatment. The very minor HCV strain with the mutant-type AA at residue 70 which were not detected before the treatment may have been detected after the treatment due to the reduction of HCV with the wild-type AA at residue 70 by the combination therapy. Indeed, HCV with the mutant-type AA at core region residue 70 was not detectable in serum 6 months after the end of the combination therapy, suggesting that it returned to being a very minor population (data not shown). These two phenomena were observed in patients with both the TT genotype of the rs8099917, that is associated with a favorable response to the combination therapy and those with the TG/GG genotypes that is associated with an unfavorable response, without difference in the prevalence according to the genetic polymorphisms at the rs8099917 near the *IL28B* gene.

In conclusion, PEG-IFN/ribavirin combination therapy does not appear to induce the mutation of the AA at the HCV core region residue 70 regardless of the genetic polymorphism near the *IL28B* gene in Japanese patients infected with HCV genotype 1b. The combination therapy can be attempted regardless of the genetic polymorphisms near the *IL28B* gene in

treatment-naïve patients without the anxiety for the acquisition of the further resistance to the antiviral therapy. However, future studies should be undertaken to confirm the absence of the mutation at residue 70 of the HCV core region induced by the combination therapy with PEG-IFN and ribavirin. In addition, the effect of the genetic polymorphisms near the *IL28B* gene on the mutation of the AA at the HCV core region residue 70 should be investigated in the long-term observation of the natural course of chronic hepatitis C.

REFERENCES

Abe H, Ochi H, Maekawa T, Hayes CN, Tsuge M, Miki D, Mitsui F, Hiraga N, Imamura M, Takahashi S, Ohishi W, Arihiro K, Kubo M, Nakamura Y, Chayama K. 2010. Common variation of IL28 affects gamma-GTP levels and inflammation of the liver in chronically infected hepatitis C virus patients. J Hepatol 53:439–443.

Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, Kobayashi M, Saitoh S, Watahiki S, Sato J, Matsuda M, Kobayashi M, Arase Y, Ikeda K, Kumada H. 2005. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 48:372–380.

Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Kobayashi M, Arase Y, Ikeda K, Kumada H. 2007a. 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.

Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Kobayashi M, Arase Y, Ikeda K, Kumada H. 2007b. Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. Hepatology 46:1357–1364.

Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Kobayashi M, Arase Y, Ikeda K, Miyakawa Y, Kumada H. 2007c. Prediction of response to pegylated interferon and ribavirin in hepatitis C by polymorphisms in the viral core protein and very early dynamics of viremia. Intervirology 50:361–368.

Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Kumada H. 2009. Amino acid substitutions in the hepatitis C virus core region of genotype 1b are the important predictor of severe insulin resistance in patients without cirrhosis and diabetes mellitus. J Med Virol 81:1032–1039.

Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Chayama K, Nakamura Y, Kumada H. 2010. Amino acid substitution in hepatitis C virus core region and genetic variation near the interleukin 28B gene predict viral response to telaprevir with peginterferon and ribavirin. Hepatology 52:421–429.

Donlin MJ, Cannon NA, Yao E, Li J, Wahed A, Taylor MW, Belle SH, Di Bisceglie AM, Aurora R, Tavis JE. 2007. Pretreatment sequence diversity differences in the full-length hepatitis C virus open reading frame correlate with early response to therapy. J Virol 81:8211–8224.

Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461:399–401.

Ghany MG, Strader DB, Thomas DL, Seeff LB. 2009. Diagnosis, management, and treatment of hepatitis C: An update. Hepatology 49:1335–1374.

Hayes NC, Kobayashi M, Akuta N, Suzuki F, Kumada H, Abe H, Miki D, Imamura M, Ochi H, Kamatani N, Nakamura Y, Chayama K. 2011. HCV substitutions and IL28B polymorphisms on outcome of peg-interferon plus ribavirin combination therapy. Gut 60:261–267.

Kenny-Walsh E. 1999. Clinical outcomes after hepatitis C infection from contaminated anti-D immune globulin. Irish Hepatology Research Group. N Engl J Med 340:1228–1233.

Kobayashi M, Akuta N, Suzuki F, Hosaka T, Sezaki H, Kobayashi M, Suzuki Y, Arase Y, Ikeda K, Watahiki S, Mineta R, Iwasaki S, Miyakawa Y, Kumada H. 2010a. Influence of amino-acid polymorphism in the core protein on progression of liver disease in patients infected with hepatitis C virus genotype 1b. J Med Virol 82:41–48.

Kobayashi M, Suzuki F, Akuta N, Suzuki Y, Sezaki H, Yatsuji H, Hosaka T, Kobayashi M, Kawamura Y, Hirakawa M, Arase Y, Ikeda K, Mineta R, Iwasaki S, Watahiki S, Nakamura Y, Chayama K, Kumada H. 2010b. Relationship between SNPs in the *IL28B* region and amino acid substitutions in HCV core region in Japanese patients with chronic hepatitis C. Kanzo 51:322–323 [in Japanese with English abstract].

Kurbanov F, Tanaka Y, Matsuura K, Sugauchi F, Elkady A, Khan A, Hasegawa I, Ohno T, Tokuda H, Mizokami M. 2010. Positive selection of core 70Q variant genotype 1b hepatitis C virus strains induced by pegylated interferon and ribavirin. J Infect Dis 201:1663–1671.

McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, Patel K, Tillmann HL, Muir AJ, McHutchison JG. 2010. Replicated association between an IL28B gene variant and a sustained response to pegylated interferon and ribavirin. Gastroenterology 138: 2307–2314.

Nakamoto S, Imazeki F, Fukai K, Fujiwara K, Arai M, Kanda T, Yonemitsu Y, Yokosuka O. 2010. Association between mutations in the core region of hepatitis C virus genotype 1 and hepatocellular carcinoma development. J Hepatol 52:72–78.

Niederau C, Lange S, Heintges T, Erhardt A, Buschkamp M, Hürter D, Nawrocki M, Kruska L, Hensel F, Petry W, Häussinger D. 1998. Progress of chronic hepatitis C: Results of a large, prospective cohort study. Hepatology 28:1687–1695.

Ohno T, Mizokami M, Wu R-R, Saleh MG, Ohba K-I, Orito E, Mukaide M, Williams R, Lau JY. 1997. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J Clin Microbiol 35:201–207.

Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, Bochud M, Battegay M, Bernasconi E, Borovicka J, Colombo S, Cerny A, Dufour JF, Furrer H, Günthard HF, Heim M, Hirschel B, Malinverni R, Moradpour D, Müllhaupt B, Witteck A, Beckmann JS, Berg T, Bergmann S, Negro F, Telenti A, Bochud PY. Swiss Hepatitis C Cohort Study; Swiss HIV Cohort Study. 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: A genome-wide association study. Gastroenterology 138:1338–1345.

Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E, Riordan S, Sheridan D, Smedile A, Fragomeli V, Müller T, Bahlo M, Stewart GJ, Booth DR, George J. 2009. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 41:1100–1104.

Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M. 2009. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41:1105–1109.

The French METAVIR Cooperative Study Group. 1994. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. Hepatology 20:15–20.

OPEN ACCESS Freely available online

PLoS one

Myelin Basic Protein as a Novel Genetic Risk Factor in Rheumatoid Arthritis—A Genome-Wide Study Combined with Immunological Analyses

Chikashi Terao[1,2,3], Koichiro Ohmura[2], Masaki Katayama[2], Meiko Takahashi[1], Miki Kokubo[1,4], Gora Diop[1], Yoshinobu Toda[5], Natsuki Yamamoto[2], Human Disease Genomics Working Group[¶a], Rheumatoid Arthritis (RA) Clinical and Genetic Study Consortium[¶b], Reiko Shinkura[6], Masakazu Shimizu[2], Ivo Gut[7], Simon Heath[7], Inga Melchers[8], Toshiaki Manabe[9], Mark Lathrop[7,10], Tsuneyo Mimori[2], Ryo Yamada[1,11], Fumihiko Matsuda[1,4,12]*

1 Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **2** Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **3** Global Centers of Excellence (COE) program, Kyoto University Graduate School of Medicine, Kyoto, Japan, **4** Core Research of Evolutional Science and Technology (CREST) program, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan, **5** Center for Anatomical Studies, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **6** Department of Immunology and Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **7** Commisariat a l'energie Atomique (CEA), Institut Genomique, Centre National de Genotypage, Evry, France, **8** Clinical Research Unit for Rheumatology, University Medical Center, Freiburg, Germany, **9** Laboratory of Diagnostic Pathology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **10** Fondation Jean Dausset, Centre d'Etude du Polymorphisme Humain, Paris, France, **11** Unit of Statistical Genetics, Center for Genomic Medicine Graduate School of Medicine Kyoto University, Kyoto, Japan, **12** Institut National de la Sante et de la Recherche Medicale (INSERM) Unite U852, Kyoto University Graduate School of Medicine, Kyoto, Japan

Abstract

Rheumatoid arthritis (RA) is a major cause of adult chronic inflammatory arthritis and a typical complex trait. Although several genetic determinants have been identified, they account for only a part of the genetic susceptibility. We conducted a genome-wide association study of RA in Japanese using 225,079 SNPs genotyped in 990 cases and 1,236 controls from two independent collections (658 cases and 934 controls in collection1; 332 cases and 302 controls in collection2), followed by replication studies in two additional collections (874 cases and 855 controls in collection3; 1,264 cases and 948 controls in collection4). SNPs showing $p<0.005$ in the first two collections and $p<10^{-4}$ by meta-analysis were further genotyped in the latter two collections. A novel risk variant, rs2000811, in intron2 of the myelin basic protein (*MBP*) at chromosome 18q23 showed strong association with RA ($p=2.7\times10^{-8}$, OR 1.23, 95% CI: 1.14–1.32). The transcription of *MBP* was significantly elevated with the risk allele compared to the alternative allele ($p<0.001$). We also established by immunohistochemistry that MBP was expressed in the synovial lining layer of RA patients, the main target of inflammation in the disease. Circulating autoantibody against MBP derived from human brain was quantified by ELISA between patients with RA, other connective tissue diseases and healthy controls. As a result, the titer of anti-MBP antibody was markedly higher in plasma of RA patients compared to healthy controls ($p<0.001$) and patients with other connective tissue disorders ($p<0.001$). ELISA experiment using citrullinated recombinant MBP revealed that a large fraction of anti-MBP antibody in RA patients recognized citrullinated MBP. This is the first report of a genetic study in RA implicating MBP as a potential autoantigen and its involvement in pathogenesis of the disease.

Citation: Terao C, Ohmura K, Katayama M, Takahashi M, Kokubo M, et al. (2011) Myelin Basic Protein as a Novel Genetic Risk Factor in Rheumatoid Arthritis—A Genome-Wide Study Combined with Immunological Analyses. PLoS ONE 6(6): e20457. doi:10.1371/journal.pone.0020457

Editor: Amanda Ewart Toland, Ohio State University Medical Center, United States of America

Received March 10, 2011; **Accepted** April 21, 2011; **Published** June 3, 2011

Funding: C.T. is an associate fellow of Global Centers of Excellence program supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan (http://www.mext.go.jp/english/). G.D. is a postdoctoral fellow of Japan Society for the Promotion of Science (JSPS, http://www.jsps.go.jp/english/) and M.A. is a research member of Core University Program supported by JSPS (http://www.jsps.go.jp/english/). The study was supported in part by Core Research of Evolutional Science & Technology (http://www.jst.go.jp/kisoken/crest/en/index.html), Solution-Oriented Research for Science and Technology (http://www.jst.go.jp/kisoken/sorst/en/index.html), Japan Science and Technology Agency (http://www.jst.go.jp/EN/index.html), and by grants-in-aid for scientific research from MEXT (http://www.mext.go.jp/english/) and from the Ministry of Health, Labour and Welfare in Japan (http://www.mhlw.go.jp/english/index.html), by the Institut National de la Sante et de la Recherche Medicale in France (http://english.inserm.fr/), and by Okawa Foundation for Information and Telecommunications (http://www.okawa-foundation.or.jp/e/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fumi@genome.med.kyoto-u.ac.jp

¶a Membership of the Human Disease Genomics Working Group is provided in the Acknowledgments.
¶b Membership of the RA Clinical and Genetic Study Consortium is provided in the Acknowledgments.

Introduction

Rheumatoid arthritis (RA) is the most common cause of adult inflammatory arthritis, affecting 0.5–1% of the adult population worldwide, and is associated with joint pain, dysfunction and deformity. Both genetic and environmental risk factors have been implicated in RA [1–2]. *HLA-DRB1* is a major genetic component of RA across ethnicities and is estimated to contribute to 30 to 50% of the total genetic risk [3]. However, the other risk loci identified to date show ethnic-specific patterns of disease association. Large-scale genetic analyses including genome-wide association (GWA) studies have shown that more than 20 genes such as *PTPN22*, *TRAF1/C5*, *CD40*, and *TNFAIP3* are associated with RA in populations of European descent [4–12]. A different set of non-*HLA* genes, namely, *PADI4*, *SLC22A4*, *FCRL3*, *CD244* and *CCR6* were first reported for their association with RA using Japanese DNA collections [13–17]. Among them, several genes including *CCR6*, *STAT4* and *TNFAIP3* were later proven their association beyond ethnicity [12,18–19]. On the other hand, some other genes showed strong specificity to a certain ethnic group. The association of the *PTPN22* has been repeatedly reproduced by subsequent genetic studies in Europeans [5,20–21]. However, no evidence of strong disease risk in *PTPN22* was shown in Japanese in part due to a much lower frequency of the risk allele [22]. Similarly the association of *PADI4* with RA, which has been confirmed by multiple genetic studies in Japanese and Koreans [23–24], is found to be much weaker in Europeans [25–26]. Moreover, the size of DNA collections used for GWA studies is much larger in European populations than in Japanese, suggesting the existence of unknown genetic factors in Japanese [12,17]. For these reasons, we decided to conduct a new large-scale GWA study of RA in Japanese. Independent collections of RA patients and controls were enrolled from four clinical centers in our study. The collections from two centers, totaling 990 cases and 1,241 controls, were characterized with genome-wide SNP arrays, and the data were analyzed to identify potential disease-associated loci. For replication, SNPs at these loci were examined in the two remaining collections, totaling 2,138 cases and 1,803 controls.

Results

Genome scan and validation studies

We accumulated data on 3,128 cases and 3,039 controls of four independent RA collections (termed as collections 1,2,3 and 4, Table S1). Collections 1 and 2 (totaling 990 cases and 1,236 controls) were used for GWA analysis and collections 3 and 4 (totaling 2,138 cases and 1,803 controls) were used as replication samples. Quality control of the GWA genotyping results was undertaken separately in cohorts 1 and 2 because of differences in the SNP arrays used (see Methods, Table S2). For 225,079 markers that were common between the arrays and fulfilled our inclusion criteria, we found no evidence of population stratification between cases and controls (genomic control inflation factor $\lambda = 1.03$, Figure S1). We undertook analysis of each collection individually, and a meta-analyses to pool the results in the two collections in the association analysis (see Methods for further details). We report p-values from the meta-analysis unless otherwise stated.

We found a strong association of disease risk with markers in the *HLA* complex [27] ($p = 5.0\times10^{-31}$, Table S3). Although no other chromosomal loci showed genome-wide significance, we detected evidence of association signal in *PADI4* ($p = 2.3\times10^{-5}$), as previously reported in Japanese [13,23,27] (Table S3). The association results of five reported genes in Japanese were shown in Table S4. These results support the quality of our study populations for genetic analysis. In addition, we identified 10 SNPs in five additional chromosomal regions that met our statistical criteria for testing in the replication collections ($p<0.005$ in both collections 1 and 2, and $p<10^{-4}$ in the meta-analysis). None of them showed potential association with p-value being smaller than 10^{-5} in the other Japanese GWA study [17]. From each of these regions, the SNP with the smallest p-value was selected for examination in collections 3 and 4. One of these SNPs, rs2000811 ($p = 0.0036$ in collection1; $p = 5.7\times10^{-4}$ in collection2; $p = 1.2\times10^{-5}$ in the meta-analysis) located on chromosome 18q23, was significantly associated with RA in both replication collections ($p = 0.023$ in collection3; $p = 0.0041$ in collection4, and

Table 1. Association of *MBP* locus with rheumatoid arthritis in the Japanese population.

Chr	dbSNPID	Gene	Allele	DNA Collection		Genotype counts			Success rate	HWEp	RAF**	p-value	OR	mhp***
			Ref.(A1)/Var.(A2)			A1A1	A1A2	A2A2					(95%CI)	
18q23	rs2000811	*MBP*	C/T*	1	case	203	303	136	99.8	0.25	0.45	0.0036	1.25	
					control	344	442	148	100	0.76	0.4		(1.08–1.44)	
				2	case	95	152	79	99.7	0.24	0.48	5.7×10^{-4}	1.49	
					control	120	131	46	100	0.31	0.38		(1.19–1.87)	
				3	case	283	392	182	98.1	0.034	0.44	0.023	1.17	
					control	298	404	134	97.8	0.88	0.4		(1.02–1.34)	
				4	case	393	622	233	98.7	0.63	0.44	0.0041	1.19	
					control	341	451	141	98.4	0.68	0.39		(1.06–1.35)	
				3+4	case	676	1014	415	98.5	0.32	0.44	3.0×10^{-4}	1.18	
					control	639	855	275	98.1	0.69	0.4		(1.08–1.30)	
				pooled	case	974	1469	630	98.9	0.078	0.44	4.0×10^{-8}	1.23	2.7×10^{-8}
					control	1103	1428	469	98.9	0.85	0.39		(1.14–1.32)	

*risk allele for the disease,
**risk allele frequency, and
***p-value in meta-analysis using Cochran-Mantel-Haenszel test.
doi:10.1371/journal.pone.0020457.t001

$p = 3.0 \times 10^{-4}$ in collections 3 and 4). When the four collections were combined, the evidence of association at rs2000811 exceeded genome-wide significance when evaluated either by meta-analysis ($p = 2.7 \times 10^{-8}$; OR = 1.23; 95%CI 1.14–1.32) or by pooling of the genotype counts ($p = 4.0 \times 10^{-8}$; OR = 1.23; 95%CI 1.14–1.32; see Table 1). There was no difference in the effect size among the four collections ($p = 0.28$). To be more conservative, however, we have calculated corrected *p*-values with Principle Component Analysis (PCA), using subsets of case and control collections for which individual genotypes were available (970 cases and 297 controls, for details, see Materials and Methods). There was no difference in *p*-values with and without the correction ($p = 6.5 \times 10^{-4}$ and corrected $p = 6.5 \times 10^{-4}$). The four SNPs from the other regions that were tested showed no evidence of association in collection3 for replication study (Table S5).

The disease associated marker rs2000811 is located in the second intron of the *MBP* (myelin basic protein) gene at chromosome 18q23 within a 156-kb region that contains the *MBP* gene (NCBI MapViewer, build 36.3). Linkage disequilibrium (LD) was evaluated using genotyping results obtained in collections 1 and 2; rs2000811 did not show significant LD with other markers from the region ($r^2 < 0.14$; Figure 1), or elsewhere in the genome. An imputation analysis using the Japanese HapMap data identified a SNP, rs9958028, which was 358-bp apart and in strong LD with rs2000811 ($r^2 = 0.96$), as the second strongest association. However, no other marker was in strong LD with these two markers ($r^2 = 0.35$ or smaller) (Figure S2). To determine if unidentified polymorphisms within *MBP* were in LD with rs2000811, we performed a sequencing of the exons and the promoter region of the *MBP* gene in 84 Japanese population control DNAs (Method S1). We identified 66 SNPs, 37 of which were not registered in dbSNP, and three of which were deleterious polymorphisms (Tables S6 and S7). Again, none of these polymorphisms was in strong LD with rs2000811 ($r^2 = 0.35$ or smaller) (Figure S3). An imputation analysis using the genotyping results obtained by sequencing did not discover any other polymorphisms showing stronger association signals ($p > 0.0070$) than that of rs2000811. Taken together, these data suggest that rs2000811 and/or one or more other as yet unidentified non-coding polymorphisms within or near *MBP* are responsible for the genetic association.

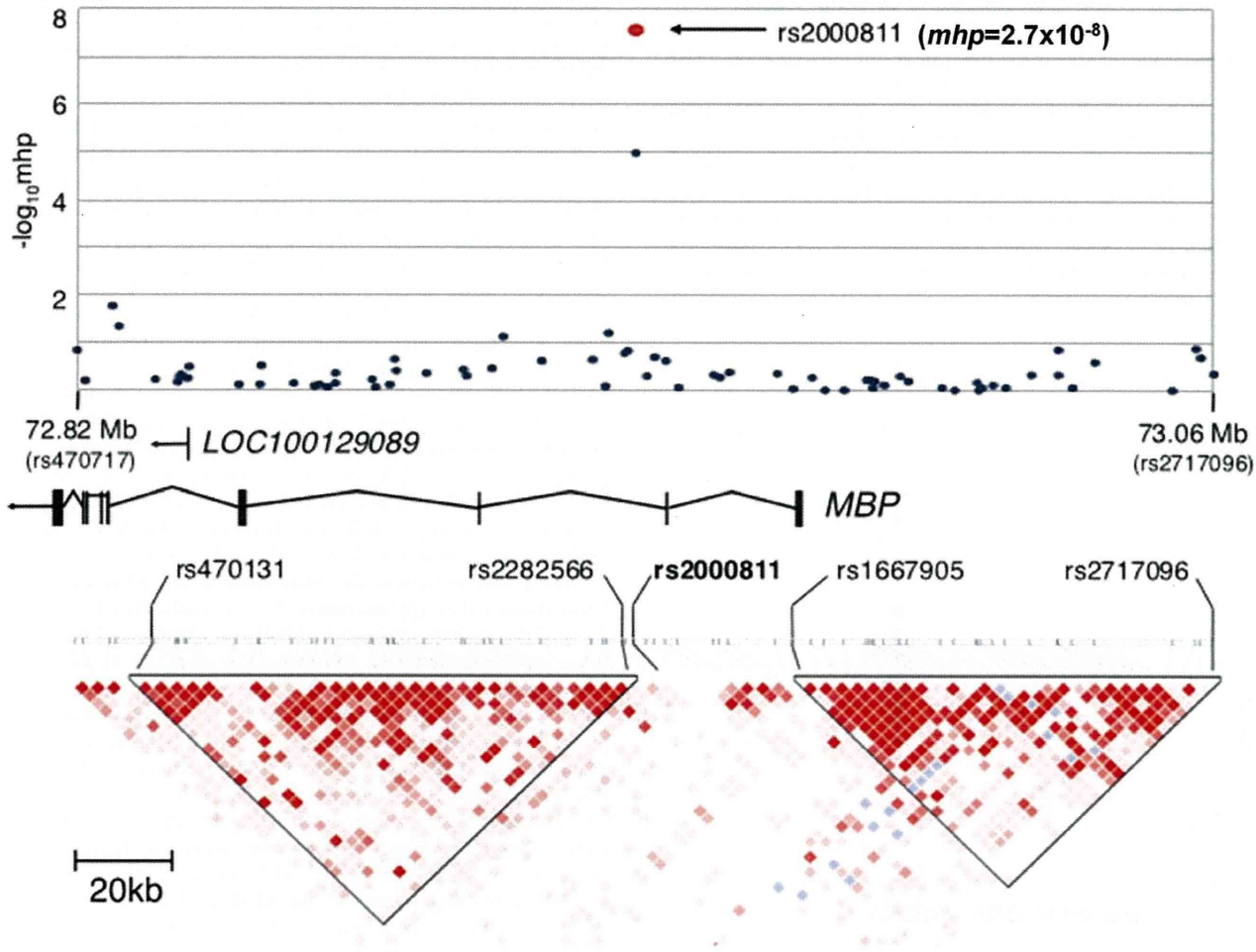

Figure 1. A schematic view of the association results and LD structure of the human *MBP* gene locus at chromosome 18q23. SNPs located between rs470131 and rs2717096 are plotted in $-\log_{10}$ scale according to their chromosomal positions and *p*-values calculated with Cochran-Mantel-Haenszel test. Red circle indicates *mhp*-value of rs2000811 by meta-analysis using the combined results of collections 1 to 4. Relative locations of the genes in the region are shown with their transcriptional orientations by arrows. LD blocks were generated using the genome scan results.

doi:10.1371/journal.pone.0020457.g001

Evaluation of MBP transcription

Quantitative RT-PCR experiments showed only very low levels of *MBP* expression in RNA from Epstein-Barr virus (EBV)-transformed human B-lymphoblastoid cell lines, and we could detect no discernable correlation of *MBP* transcript levels with different risk genotypes ($p = 0.36$, Figure S4). By a similar reason, *in-silico* expression analysis using GEO database did not return clear association [28]. However, when we performed allele-specific quantitative RT-PCR [16] using genomic DNA and cDNA of these cell lines, we observed elevated allele-specific transcription associated with the risk allele ($p<0.001$, Figure 2, for detailed procedure, see Materials and Methods). This suggests that rs2000811, and/or other variants in linkage disequilibrium with this marker, impact the quantitative pattern of *MBP* transcription. However, bioinformatics analysis identified no known *cis*-acting elements covering rs2000811 that could be inferred to have functional effects (Method S2). In addition, the alignment of the 4-kb region ranging between 2-kb centromeric and 2-kb telomeric to rs2000811 revealed that this segment has very low interspecies conservation among placental mammals.

Expression of the MBP protein in synovial tissues

Next we investigated the expression of the MBP protein in synovial tissue, as this is the main target of inflammation in RA. Microscopic observation revealed that MBP was highly expressed along the lining layer of synovial tissues in 20 out of 23 RA patients tested, while the expression of MBP was observed in only one out of five controls ($p = 0.0017$), and then generally at a weaker level (Figure 3A, B). In synovial tissue from RA patients, the detected MBP expression was weaker in synovial lining layer adjacent to the follicules of infiltrated lymphocytes (Figure 3C). In synoviocytes, the expression of MBP was mainly observed in the plasma membrane (Figure 3D).

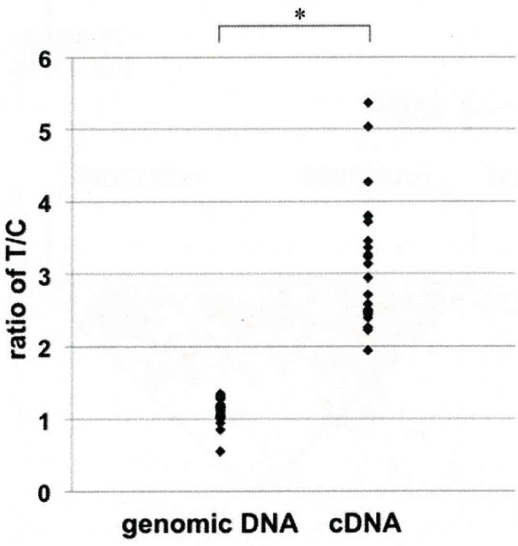

Figure 2. Allelic difference in *MBP* transcription using allele specific quantitative RT-PCR. The amount of *MBP* primary transcripts transcribed from chromosomes carrying rs2000811 risk (T) and alternative (C) alleles was compared in each cell line, and the ratio (T/C) was plotted. Genomic DNA was used as a control for equimoler biallelic representation. Experiments were done twice independently.
doi:10.1371/journal.pone.0020457.g002

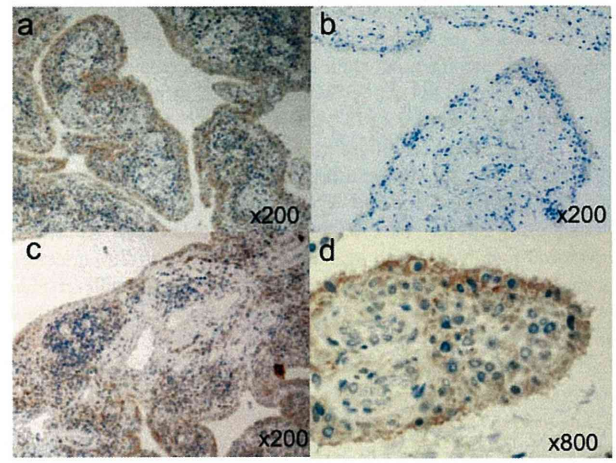

Figure 3. Immunohistochemistry of the MBP protein in human synovial tissues stained by monoclonal anti-MBP antibody. Synovial tissue of RA patients, in particular, along the synovial lining layer strongly expressed MBP (A), whereas that of non-inflammatory osteoarthritis patients was much weaker (B). The expression of MBP in the synovial lining layer was weaker near follicules of infiltrated lymphocytes (C). Localized expression of MBP was observed at the plasma membrane of synoviocytes (D).
doi:10.1371/journal.pone.0020457.g003

Quantification of antibodies to MBP in RA patients

Antibodies to MBP are the major component of autoantibodies in multiple sclerosis, a human autoimmunity with a neurodegenerative phenotype [29]. To assess a possible association of circulating antibodies to MBP with RA, we quantified these in plasma from 323 RA cases, 131 healthy controls and 162 patients with other connective tissue diseases (disease controls) by Enzyme-linked immunosorbent assay (ELISA) with MBP purified from human brain as antigen. The average levels of anti-MBP antibody in plasma of RA patients were much higher than those of healthy controls and patients with seven other connective tissue diseases ($p<0.001$; Figure 4A). Specificity in detection of anti-MBP antibody in ELISA experiments was confirmed by immunoblotting using plasma of a subset of RA patients and controls (Figure S5, Method S3). We also confirmed that the enhancement of ELISA signals by non-specific binding of IgG- and IgM-RF in patients' sera was negligible (for details, see Method S4 and Figure S6).

Amino acid analysis of the MBP protein derived from human brain showed that approximately 21% (citrulline/arginine = 6.0/23.1) of the arginine residues in MBP was citrullinated in human brain under physiological conditions (Method S5). It is possible that anti-MBP antibody recognizes and binds to citrullinated MBP protein. We then performed ELISA using recombinant MBP protein and compared the antibody titers with those of human brain-derived MBP. We could not observe correlation between the two results ($r = -0.19$, Figure 4B). However, when we used recombinant MBP protein artificially citrullinated *in-vitro*, the ELISA results showed strong correlation in titers of the autoantibody ($r = 0.88$, Figure 4C). From these results, we concluded that higher levels of anti-MBP antibody in RA patients than in healthy controls and in patients with other connective tissue diseases was attributed to autoantibodies binding to citrullinated MBP.

Discussion

We conducted a large-scale GWA-based genetic study of RA in the Japanese population. A genome scan of 225,079 SNPs in two

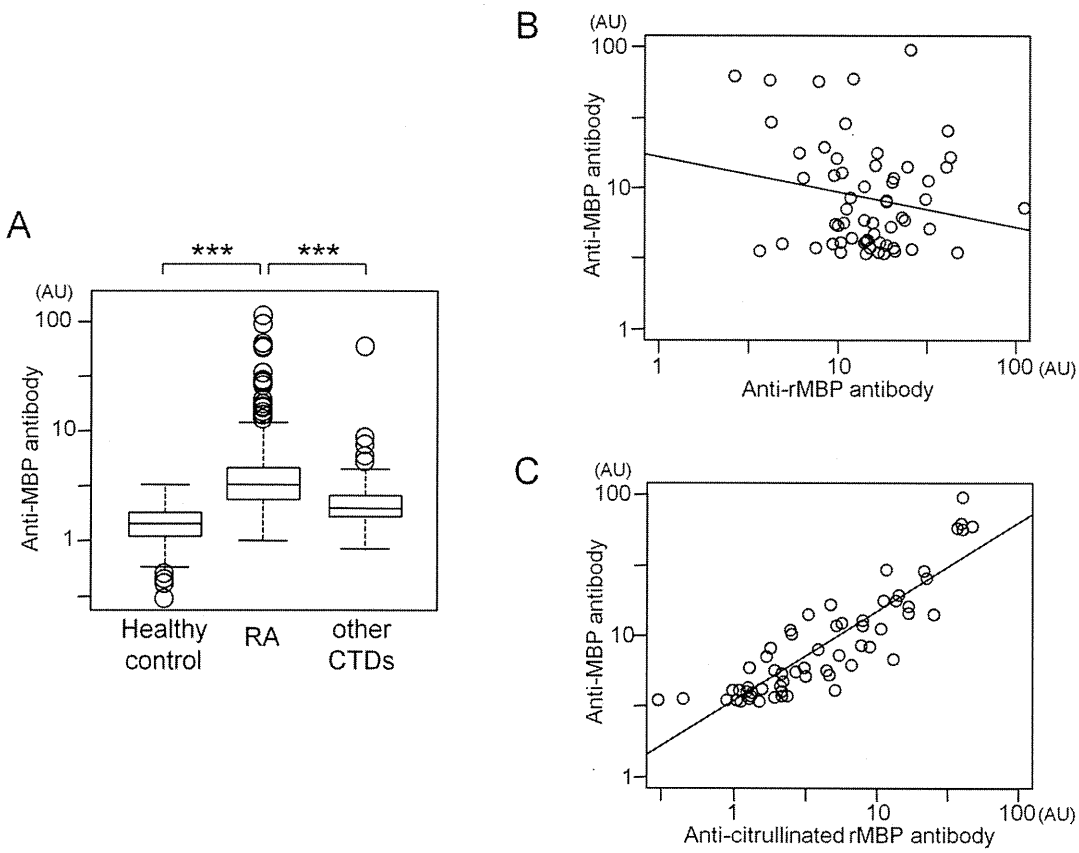

Figure 4. Quantification of circulating antibodies against MBP protein. A. Each boxplot indicates distribution of anti-MBP antibody titers in healthy controls, in RA patients, and in other connective tissue diseases (CTDs) (for detailed composition of the disease patients, see materials and methods). Results were representative of the two independent experiments. *** indicates statistical *p*-value smaller than 0.001. B. Correlation of autoantibody levels between human brain-derived MBP and recombinant MBP (rMBP). C. Correlation of autoantibody levels between human brain-derived MBP and citrullinated rMBP. Sixty RA patients who were positive for anti-MBP antibody were used for the analysis. In each figure, antibody titers were shown as arbitrary unit (AU).
doi:10.1371/journal.pone.0020457.g004

DNA collections of RA patients followed by replication in two additional collections led to identification of a novel risk variant, rs2000811 ($p = 2.7 \times 10^{-8}$, OR 1.23, 95% CI: 1.14–1.32), in the second intron of the *MBP* gene at chromosome 18q23. This is the second largest genome-wide association study to date in the Japanese population, and the first to have identified chromosome 18q23 as a potential susceptibility locus for RA. 18q23 was not reported as a candidate genetic locus in recent GWA study of RA in Japanese [17]. As summarized in Table S4, we confirmed in the current study the association of three out of seven genes reported to date whereas there was no evidence of association for the other four genes. A comprehensive genetic study by a meta-analysis using the results of different genome scans followed by the validation study with a large number of patients and controls will clarify relative contribution of each genetic locus to RA in the Japanese population. There are no GWA studies to date that indicate chromosome 18q23 as a potential genetic locus related to the susceptibility to RA in European populations [5–12]. By taking into account the detection power of the meta-analysis by Stahl *et al.*, it is natural to consider that the association of *MBP* with RA is limited to Japanese (and possibly other Asian populations).

The disease-associated marker rs2000811 is isolated in an 18-kb segment of the intron, and does not show evidence of significant LD with known markers in the region or neighboring genomic regions (Figure 1). The LD structure of the human *MBP* locus was very similar between Europeans and Japanese, and there was no clear LD block in the region encompassing rs2000811 in both populations (Figure S2). Sequencing of the DNAs of 84 healthy controls failed to identify other polymorphisms that were in strong LD with rs2000811 in the exons or the promoter region of the *MBP* gene. While bioinformatics analysis failed to identify any *cis*-acting elements encompassing rs2000811, allele-specific expression analysis demonstrated the elevated transcription level of *MBP* with the risk allele. This raises the possibility that *cis*-acting regulatory elements that encompass unknown genetic variations in strong LD with rs2000811.

A strong expression of the MBP protein was observed in synoviocytes of RA patients while expression was weaker in those of non-inflammatory controls. The fact that the expression of MBP was strongest in the synovial lining layer suggests the role of anti-MBP antibody as a trigger of inflammatory reactions through attacking synoviocytes. Although a study investigating citrullinated proteins in synoviums did not clearly identify MBP [30], MBP is possibly one of the unidentified subset of the citrullinated proteins. MBP expression was weaker in synoviocytes adjacent to the follicules of infiltrated lymphocyte. It would be of interest to examine whether MBP in synovium is

citrullinated and such phenomena are correlated with the disease activity, in particular, at different stages of proliferative synovitis but data for such an investigation are not available in our study. In the absence of a quantitative assay of expression in synoviocytes, it was not possible to discern a correlation between the levels of *MBP* expression with the risk genotype in our experiments.

MBP is a well-known target autoantigen in multiple sclerosis (MS), a human neurodegenerative disease with an active destruction of myelin sheath [29]. The MBP protein has six different isoforms of which isoforms 1, 2, 3 and 4 with shorter open reading frames are expressed preferentially in the central nervous system (CNS). Isoforms 7 and 8 with higher molecular weights (also called Golli-MBP) are known to be expressed relatively ubiquitously including cells of the hematopoietic lineage [31]. The fact that RA patients rarely present CNS symptoms may suggest the involvement of Golli-MBP protein in the generation of anti-MBP antibody and expression in the synovium. An immunological study using relatively small numbers of plasma samples showed that anti-MBP antibody is present in 60% (or six out of ten) of RA patients [32]. To our knowledge, however, there is no study which quantitatively characterized anti-MBP antibody in RA patients as well as in patients of other connective tissue diseases and in healthy controls. On the other hand, ACPA which recognizes citrullinated proteins by peptidylarginine deiminase is considered as a specific and predictive marker for RA [13,33–34]. Indeed, several studies showed the existence of antibodies to citrullinated MBP in RA patients [35–36]. However, the role of MBP and anti-MBP antibody in the pathogenesis of RA is yet to be elucidated. In the current study, we found that a part of brain-derived MBP was citrullinated in physiological conditions and that higher levels of anti-MBP antibody in RA patients can be attributed to MBP citrullination, although anti-MBP antibody is not in complete cross-reactivity with ACPA. We examined whether or not rs2000811 was predominantly associated with patients who were positive for ACPA. However, there was no statistical difference in allele frequency of rs2000811 between patients with and without ACPA ($p = 0.40$, data not shown).

We observed no correlation between the levels of anti-MBP antibody and genotypes at the risk locus. The repertoire of autoantibody is likely to depend on numerous factors such as the immunogenicity and extent of citrullination of the MBP protein. Also, a series of immunological reactions including antigen presentation in thymus and peripheral organs, activation of dendritic cells, T- and B-lymphocytes, and balance of Th1/Th2 and effector/regulatory T-lymphocytes affect antibody production. A possible explanation for the lack of correlation is that the disease-associated genotypes affect these factors in a way that impacts risk without directly influencing the quantitative values of anti-MBP antibody, which we find to be highly variable between RA patients that we have studied. Therefore, it will be of interest to compare the expression levels of *MBP* transcripts and MBP epitopes in synovial tissue, and between different genotypes and disease activity.

Taken together, this is the first genetic study which identified *MBP*, an autoantigen gene, to be associated with RA in Japanese. Transcription of *MBP* was increased with the risk allele of the associated SNP, rs2000811. Strong expression of the MBP protein was observed in the synovial tissues of the patients. Furthermore, significant increase of circulating autoantibodies against MBP protein was demonstrated in RA patients as compared to those with other connective tissue diseases, implicating its role as a disease-associated biomaker.

Materials and Methods

Ethics Statement

Written informed consent was obtained from all the participants at the institute of sample collection after being approved for genetic studies by the local ethical committee, namely, Kyoto University Graduate School and Faculty of Medicine, Ethics Committee, The Ethic Committee, Sagamihara National Hospital, NHO, Dohgo Spa Hospital Ethical Committee, University of Tokyo Medical Research Center Ethics Committee, Tokyo Women's Medical University Genome Ethics Committee, Ethics Committee of Tokyo Women's Medical University and Aichi Cancer Center Ethical Committee for human genome research.

Study subjects

RA collections 1 to 4 consisted of 658 affected individuals and 934 controls, 332 and 302, 874 and 855, and 1,264 and 948, respectively (summarized in Table S1). The case subjects of collections 1, 2 and 3 were recruited at the rheumatology departments of Kyoto University Hospital, Dohgo Spa Hospital, Sagamihara National Hospital, and Tokyo University Hospital, and those for collection4 were from Tokyo Women's Medical University. All cases fulfilled the revised criteria (1987) of the American College of Rheumatology (ACR) for rheumatoid arthritis [37]. Genotype count data of the Japanese Single Nucleotide Polymorphism (JSNP) database [38] were used as controls for collection1. The control subjects for collection2 were from the Department of Ophthalmology and Visual Science at Kyoto University Hospital [39]. DNA samples of healthy Japanese volunteers in Pharma SNP Consortium [40] and in Aichi Cancer Center Hospital and Research Institute [41] were used for collections 3 and 4, respectively. Plasma of 323 RA patients and 162 patients of other connective tissue diseases (38 of SLE, 25 each of Sjögren's syndrome and systemic sclerosis, 20 each of Behçet's disease and mixed connective tissue disease, 19 of polymyositis/dermatomyositis and 15 of vasculitis) were obtained at Kyoto University Hospital, and those of 131 healthy controls were from Dohgo Spa Hospital.

Genome-wide association analysis

Genome scan for collections 1 and 2 was performed using Infinium Technology (Illumina Inc., San Diego, CA). Case subjects of collection1 were genotyped with Human-Hap300 (version 1.0, 302,627 SNPs) or Human CNV370-Duo (version 1.0, 332,270 SNPs). For collection2, case and control subjects were genotyped on Human610-Quad (version 1.0, 577,348 SNPs), and HumanHap550 (version 3.0, 547,163 SNPs), respectively. Validation studies using collections 3 and 4 were performed using Taqman technology (Applied Biosystems Inc., Foster City, CA) according to the manufacturer's specifications.

Quality control and statistical tests for case-control association

277,420 SNPs that were common among the four arrays described above were chosen for association study. Publicly available genotype counts from the JSNP project were used as control collection1. For this collection, detailed information such as individual genotypes and cluster plots are not disclosed. DNA samples with a call rate smaller than 0.90 (three in collection1 cases), showing high degree of kinship (PI_HAT>0.10 by PLINK [42], eleven in collection1 cases, one in collection2 cases, and four in collection2 controls) and with evidence of possible contamination (one in collection2 controls) were removed from statistical analyses. PCA was performed using the genome scan results of the remaining 644 cases of collection1, in addition to 331 cases and 297 controls of collection2. At this stage, five DNA samples (one in

collection1 cases, and four in collection2 cases) that did not fall into the Japanese cluster were removed. Regarding the SNP markers, a total of 225,079 SNPs with call rate greater than 0.95 for both cases and controls and minor allele frequency greater than 0.05 either in case or in control of each collection were used for analysis. The *p*-value for rs2000811 corrected by PCA was calculated using the remaining 970 cases and 297 controls and compared with the *p*-value without correction.

The case-control association was examined independently for collections 1 and 2 with Cochran-Armitage trend test, followed by meta-analysis with Cochran-Mantel-Haenszel (CMH) test by combining the two collections. Population stratification in collections 1 and 2 was examined and corrected with Genomic Control [43]. The SNPs that showed $p<0.005$ in both collections and meta-analysis *p*-value smaller than 10^{-4} were selected as candidates for further evaluation. Among multiple SNPs in the same region that fulfilled the above criteria, the SNP with the smallest *p*-value in the meta-analysis was chosen for validation with collections 3 and 4. To be more specific, among the four SNPs in the *PLEKHK1* region on chromosome 10q21, namely rs3910172 ($p=4.4\times10^{-5}$), rs6479805 ($p=5.4\times10^{-5}$), rs10733769 ($p=5.8\times10^{-5}$), and rs4147233 ($p=6.6\times10^{-5}$), rs3910172 was selected for the replication study. Likewise, in the region of chromosome 10p14, rs2026628 ($p=1.6\times10^{-5}$) was chosen over rs11253857 ($p=5.3\times10^{-5}$), and rs687848 ($p=5.1\times10^{-5}$) was chosen over rs587404 ($p=7.1\times10^{-5}$) in the *MACF1* region on chromosome 1p31–32. SNPs in the *HLA* and *PADI4* loci were not included in validation studies. Haploview version 4.1 software [44] was used for LD evaluation, and MapViewer (build 36.3) [45] was used to identify the location and structure of the genes in the region.

Quantification of allelic difference in gene expression in MBP transcription

Allele specific gene expression analysis was performed as previously described [16]. Briefly, human B-lymphoblastoid cell lines transformed by EBV were obtained from the Health Science Research Resources Bank of Japan (Osaka). Genomic DNA and total RNA were extracted by standard procedures from 22 cell lines heterozygous for rs2000811 alleles. The ratio of MBP primary transcripts (hnRNA) was quantified between the risk and wild-type alleles by TaqMan assay with SNP genotyping probes. Genomic DNA of a cell line homozygous for the wild-type allele (C) were mixed with DNA homozygous for the risk allele (T) at eight different molar ratios (2:1, 3:2, 1:1, 2:3, 1:2, 1:3, 1:4, 1:6) to draw a standard curve for the evaluation of RNA quantity.

Immunohistochemistry

Synovial tissue specimens of 23 RA patients and five non-inflammatory controls were obtained from Department of Diagnostic Pathology and Department of Orthopaedic Surgery in Kyoto University Hospital. Paraffin-embedded tissues were prepared in a standard method and sectioned at a thickness of 3 μm. The section was mounted on a glass slide coated with 2% 3-aminopropyl triethoxy silane (Tokyo Kasei, Tokyo, Japan). Immunohistochemical staining of MBP was performed by using the standard avidin-biotin-peroxidase complex (ABC) method, as previously described [46]. The sections were incubated overnight at 4°C with an affinity-purified murine anti-human MBP monoclonal antibody (Leica Microsystems, Wetzlar, Germany) diluted at 1:100 in PBS. The sections were then incubated with biotinylated horse anti-mouse IgG antibody (Vector Lab, Burlingame, CA) for 40 minutes, followed by incubation with peroxidase-conjugated streptavidin (Vector Lab, Burlingame, CA) at room temperature for 50 minutes. The coloring reaction was performed with 0.3 mg/ml diaminobenzidine and 0.003% H_2O_2 in 50 mM Tris-HCl (pH 7.6). Each section was counterstained with haematoxylin. Evaluation of MBP expression was performed by a blind test by two rheumatologists as well as a pathologist, and statistical significance in positivity was calculated with Fisher's exact test.

Enzyme-linked immunosorbent assay (ELISA)

Microtiter plates (Nalge Nunc International K.K., Tokyo, Japan) were coated with 50 μl of MBP protein at 5 μg/ml extracted from human brain (Sigma, St. Louis, MO) or with recombinant human MBP (Genscript, Piscataway, NJ.) in 50 mM bicarbonate buffer (pH 9.6) and incubated at 4°C overnight. Citrullination of recombinant human MBP was done *in-vitro* by rabbit skeletal PAD (Sigma, St. Louis, MO) for 3 hours as previously described [47]. After wells were washed and coated with phosphate-buffered-saline (PBS) containing 2% bovine serum albumin (BSA), 50 μl of plasma samples diluted to 1:150 with 2% BSA in PBS containing 5 U/ml heparin (Mochida Pharmaceutical Co., Ltd. Tokyo, Japan) were added and incubated at room temperature for 2 hours. 50 μl each of purified goat anti-human IgG polyclonal antibody conjugated to alkaline phosphate (Millipore. Billerica, MA) diluted to 1:2000 was added and incubated at room temperature for 1.5 hours. 50 μl of BCIP/NBT substrate (Sigma, St. Louis, MO) in 2 μM $MgCl_2$ was then added and incubated for 45 minutes in the dark. The optical density (OD) value at 405 nm was measured by a SpectraMax Plus^{384} Microplate Reader (Molecular Devices, Sunnyvale, CA). A standard curve was generated by serial dilution (1:50, 1:150, 1:500, 1:1500 and 1:5000) of a plasma sample with a high titer of anti-MBP antibody. The titer of diluted standard sample in 1:50 was set as 100 U for human brain-derived MBP. All samples were examined in duplicate except for negative control in which plasma was replaced by PBS with 2% BSA and was measured in quadruplicate. The specificity of ELISA results was confirmed by standard immunoblotting analysis using sera of 10 each of RA patients and controls.

Statistical analysis of ELISA results

The titer of anti-MBP antibody between RA or its subgroups and controls was compared with Wilcoxon rank-sum test. The correlation between the titers of autoantibody was estimated with Pearson's correlation coefficient in logarithm scale. These statistics were performed in the R statistical system (http://www.R-project.org) and SPSS(ver18).

Supporting Information

Figure S1 QQ plot to show the observed and expected *p*-values of the combined genome scan results. Vertical and horizontal axes indicate observed and expected *p*-values, respectively (A) and in logarithmic scale (B). The analysis using genomic control method showed no significant effect of population stratification ($\lambda_{GC}=1.03$) between the case and control groups.
(TIF)

Figure S2 Imputation analysis and LD structure of the human *MBP* locus. Imputation was performed using the case genotypes of collections 1 and 2 and control genotypes of collection2. Individual genotypes of control population of collection1 (JSNP) were not available. Determination of LD structure was performed by using the GWAS results in this study and the HapMap results of Japanese and Caucasians (Hap JPT and Hap CEU, respectively).
(TIF)

Figure S3 LD structure of the 156-kb region spanning the *MBP* gene. LD plot was generated with Haploview using polymorphisms with reference allele frequencies between 0.05 and 0.95.
(TIF)

Figure S4 Quantification of allelic difference in *MBP* transcription. Human B-lymphoblastoid cell lines transformed by EBV were obtained from the Health Science Research Resources Bank of Japan (Osaka, Japan). Total RNA was extracted by standard procedures from the cell lines that were either homozygous for the wild-type allele (50 cell lines), heterozygous (50 cell lines) or homozygous for the risk allele (49 cell lines) of rs2000811. The amount of *MBP* cDNA in each cell line was measured and normalized to that of β-glucronidase using Taqman Gene Expression Assay (for MBP; Hs00921943-m1, for β-glucronidase; Hs99999908_m1, Applied Biosystems Inc., Foster City, CA) in GeneAmp 7500 Sequence Detection System. The comparative ΔΔCT method and Jonckheere-Terpstra test were used for the analysis.
(TIF)

Figure S5 Immunoblotting of anti-MBP antibody. Immunoblotting analysis was performed to confirm specific binding of circulating anti-MBP antibody. Lane 1, 2 to 6 and 7 were incubated with control plasma, plasma of patients, and rabbit polyclonal anti-human MBP antibody, respectively. The intensity was variable between RA patients whereas no signal was obtained in controls. Similar results were obtained using plasma of the other five RA patients and nine controls.
(TIF)

Figure S6 Comparison of circulating anti-MBP antibody and IgG-RF. We investigated whether the presence of RF in sera enhances the signal in ELISA through binding of IgG-RF to the constant region of anti-MBP antibody or through non-specific binding of IgM-RF to secondary antibody used in the experiments. Initially, correlation between IgG-RF and anti-MBP antibody titers was assessed to examine whether IgG-RF had any influence on anti-MBP antibody titers. 114 patients with RA and 13 other connective tissue disease patients for whom IgG-RF measurement was available were enrolled in this evaluation. As a result, no correlation was observed between the titers of anti-MBP antibody and IgG-RF (Spearman's rank-sum coefficient being 0.145 with *p*-value of 0.103). Subsequently, ELISA experiments using human IgM or IgG as target antigens were undertaken. Non-specific binding of secondary antibodies to human IgM compared to IgG was less than 1%.
(TIF)

Table S1 Summary of the study populations used for the association analysis. Abbreviations were as follows; ACPA: antibodies to citrullinated peptide antigens, RF: rheumatoid factor, SD: standard deviation, N/A, not available.
(DOC)

Table S2 Summary of quality control for genome scan results.
(DOC)

Table S3 Association of *HLA* and *PADI4* loci with rheumatoid arthritis in the Japanese population. *risk allele for the disease, **risk allele frequency, and ***p-value in meta-analysis using Cochran-Mantel-Haenszel test.
(DOC)

Table S4 Association of previously reported non-*HLA* genes in the current study. *p-value in meta-analysis using Cochran-Mantel-Haenszel test.
(DOC)

Table S5 Replication results of the four regions. The order of SNPs is in accordance with *mhp*-value. Chromosome and dbSNPID refer to NCBI build 36.3. *P*-values are calculated using the Cochran-Armitage trend test. *risk allele for the disease, **risk allele frequency, and ***OR, odds ratio with 95% confidence interval. ***p-value in meta-analysis using Cochran-Mantel-Haenszel test.
(DOC)

Table S6 Oligonucleotide primers used for sequencing of the exons and the promoter region of the *MBP* gene.
(DOC)

Table S7 The list of genetic polymorphisms discovered by sequencing the exons and the promoter region of the *MBP* gene. *Positions of polymorphisms are according to NCBI Refseq Build 36.3. Polymorphisms are located between position 1 and position 2.
(DOC)

Method S1 Sequencing of the exons and the promoter region of the MBP gene.
(DOC)

Method S2 Bioinformatics analysis.
(DOC)

Method S3 Immunoblotting of anti-MBP antibody.
(DOC)

Method S4 Evaluation of non-specific binding of secondary antibodies.
(DOC)

Method S5 Amino acid analysis.
(DOC)

Acknowledgments

We are grateful to all patients and medical staffs who were concerned with the establishment of the RA cohorts.

The Human Disease Genomics Working Group includes Hisako Imamura, Mikita Suyama, Marie-Ange Palomares, Monthikan Aksornworanart, Akiko Yoshizumi, Hiroyuki Uneme, Katsura Hirosawa, Takahisa Kawaguchi, Chanavee Ratanajaraya, Mika Mizutani at The Unit of Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, Hideo Nakanishi, Nagahisa Yoshimura at Department of Ophthalmology and Visual Science, Kyoto University Graduate School of Medicine, Kyoto, Japan, Keitaro Matsuo, Hideo Tanaka at Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Aichi, Japan, Kazuo Tajima, Director, at Aichi Cancer Center Hospital and Research Institute, Nagoya, Aichi, Japan, Roubila Meziani at Commisariat a l'energie Atomique (CEA), Institut Genomique, Centre National de Genotypage, Evry, France, CREST program, Japan Science and Technology Agency, Kawaguchi, Japan and Fondation Jean Dausset, Centre d'Etude du Polymorphisme Humain, Paris, France.

The RA Clinical and Genetic Study Consortium includes Naoichiro Yukawa, Daisuke Kawabata, Takaki Nojima, Takashi Usui, Takao Fujii at Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan, Kota Shimada, Hiroshi Furukawa, Shigeto Tohma at Department of Rheumatology, National Hospital Organization Sagamihara Hospital, Sagamihara, Kanagawa, Japan, Kenichiro Matoba, Kiyoshi Takasugi at Department of Internal Medicine, Center for Rheumatic Diseases, Dohgo Spa Hospital, Matsuyama, Ehime, Japan, Kazuyuki Nakagome at Department of Allergy and Rheumatology, The University of Tokyo Hospital, Tokyo, Japan, Atsuo Taniguchi, Hisashi Yamanaka, Naoyuki Kamatani at Institute of

Rheumatology, Tokyo Women's Medical University, Tokyo, Japan, Hiromu Ito, Hiroyuki Yoshitomi, Takashi Nakamura at Department of Orthopaedic Surgery, Kyoto University Graduate School of Medicine, Kyoto, Japan, Ryoko Ito at Laboratory of Diagnostic Pathology, Graduate School of Medicine, Kyoto University, Kyoto, Japan.

Author Contributions

Conceived and designed the experiments: CT KO T. Mimori IM ML RS RY FM. Performed the experiments: CT M. Kokubo NY GD MS M. Katayama MT. Analyzed the data: CT RY SH IG ML YT T. Manabe HDGWG. Contributed reagents/materials/analysis tools: HDGWG RCGSC. Wrote the paper: CT KO MT T. Mimori ML FM.

References

1. Firestein GS (2003) Evolving concepts of rheumatoid arthritis. Nature 423: 356–361.
2. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, et al. (2000) Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 43: 30–37.
3. Deighton CM, Walker DJ, Griffiths ID, Roberts DF (1989) The contribution of HLA to rheumatoid arthritis. Clin Genet 36: 178–182.
4. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, et al. (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 75: 330–337.
5. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, et al. (2007) TRAF1-C5 as a risk locus for rheumatoid arthritis–a genomewide study. N Engl J Med 357: 1199–1209.
6. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, et al. (2007) STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 357: 977–986.
7. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, et al. (2008) Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 40: 1216–1223.
8. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, et al. (2007) Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 39: 1477–1482.
9. Thomson W, Barton A, Ke X, Eyre S, Hinks A, et al. (2007) Rheumatoid arthritis association at 6q23. Nat Genet 39: 1431–1433.
10. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, et al. (2009) Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nat Genet 41: 1313–1318.
11. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, et al. (2009) REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat Genet 41: 820–823.
12. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, et al. (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42: 508–514.
13. Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, et al. (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34: 395–402.
14. Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, et al. (2003) An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet 35: 341–348.
15. Kochi Y, Yamada R, Suzuki A, Harley JB, Shirasawa S, et al. (2005) A functional variant in FCRL3, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. Nat Genet 37: 478–485.
16. Suzuki A, Yamada R, Kochi Y, Sawada T, Okada Y, et al. (2008) Functional SNPs in CD244 increase the risk of rheumatoid arthritis in a Japanese population. Nat Genet 40: 1224–1229.
17. Kochi Y, Okada Y, Suzuki A, Ikari K, Terao C, et al. (2010) A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility. Nat Genet 42: 515–519.
18. Kobayashi S, Ikari K, Kaneko H, Kochi Y, Yamamoto K, et al. (2008) Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. Arthritis Rheum 58: 1940–1946.
19. Shimane K, Kochi Y, Yamada R, Okada Y, Suzuki A, et al. (2009) A single nucleotide polymorphism in the IRF5 promoter region is associated with susceptibility to rheumatoid arthritis in the Japanese population. Ann Rheum Dis 68: 377–383.
20. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.
21. Orozco G, Sanchez E, Gonzalez-Gay MA, Lopez-Nevot MA, Torres B, et al. (2005) Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum 52: 219–224.
22. Mori M, Yamada R, Kobayashi K, Kawaida R, Yamamoto K (2005) Ethnic differences in allele frequency of autoimmune-disease-associated SNPs. J Hum Genet 50: 264–266.
23. Ikari K, Kuwahara M, Nakamura T, Momohara S, Hara M, et al. (2005) Association between PADI4 and rheumatoid arthritis: a replication study. Arthritis Rheum 52: 3054–3057.
24. Kang CP, Lee HS, Ju H, Cho H, Kang C, et al. (2006) A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans. Arthritis Rheum 54: 90–96.
25. Barton A, Bowes J, Eyre S, Spreckley K, Hinks A, et al. (2004) A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. Arthritis Rheum 50: 1117–1121.
26. Caponi L, Petit-Teixeira E, Sebbag M, Bongiorni F, Moscato S, et al. (2005) A family based study shows no association between rheumatoid arthritis and the PADI4 gene in a white French population. Ann Rheum Dis 64: 587–593.
27. Kochi Y, Yamada R, Kobayashi K, Takahashi A, Suzuki A, et al. (2004) Analysis of single-nucleotide polymorphisms in Japanese rheumatoid arthritis patients shows additional susceptibility markers besides the classic shared epitope susceptibility sequences. Arthritis Rheum 50: 63–71.
28. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848–853.
29. Lutton JD, Winston R, Rodman TC (2004) Multiple sclerosis: etiological mechanisms and future directions. Exp Biol Med (Maywood) 229: 12–20.
30. Skriner K, Adolph K, Jungblut PR, Burmester GR (2006) Association of citrullinated proteins with synovial exosomes. Arthritis Rheum 54: 3809–3814.
31. Campagnoni AT, Pribyl TM, Campagnoni CW, Kampf K, Amur-Umarjee S, et al. (1993) Structure and developmental regulation of Golli-mbp, a 105-kilobase gene that encompasses the myelin basic protein gene and is expressed in cells in the oligodendrocyte lineage in the brain. J Biol Chem 268: 4930–4938.
32. Reindl M, Linington C, Brehm U, Egg R, Dilitz E, et al. (1999) Antibodies against the myelin oligodendrocyte glycoprotein and the myelin basic protein in multiple sclerosis and other neurological diseases: a comparative study. Brain 122(Pt 11): 2047–2056.
33. Simon M, Girbal E, Sebbag M, Gomes-Daudrix V, Vincent C, et al. (1993) The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis. J Clin Invest 92: 1387–1393.
34. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ (1998) Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 101: 273–281.
35. Boire G, Cossette P, de Brum-Fernandes AJ, Liang P, Niyonsenga T, et al. (2005) Anti-Sa antibodies and antibodies against cyclic citrullinated peptide are not equivalent as predictors of severe outcomes in patients with recent-onset polyarthritis. Arthritis Res Ther 7: R592–603.
36. Ioan-Facsinay A, el-Bannoudi H, Scherer HU, van der Woude D, Menard HA, et al. (2011) Anti-cyclic citrullinated peptide antibodies are a collection of anti-citrullinated protein antibodies and contain overlapping and non-overlapping reactivities. Ann Rheum Dis 70: 188–193.
37. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, et al. (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31: 315–324.
38. Hirakawa M, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, et al. (2002) JSNP: a database of common gene variations in the Japanese population. Nucleic Acids Res 30: 158–162.
39. Nakanishi H, Yamada R, Gotoh N, Hayashi H, Yamashiro K, et al. (2009) A genome-wide association analysis identified a novel susceptible locus for pathological myopia at 11q24.1. PLoS Genet 5: e1000660.
40. Kamatani N, Sekine A, Kitamoto T, Iida A, Saito S, et al. (2004) Large-scale single-nucleotide polymorphism (SNP) and haplotype analyses, using dense SNP Maps, of 199 drug-related genes in 752 subjects: the analysis of the association between uncommon SNPs within haplotype blocks and the haplotypes constructed with haplotype-tagging SNPs. Am J Hum Genet 75: 190–203.
41. Suzuki T, Matsuo K, Sawaki A, Mizuno N, Hiraki A, et al. (2008) Alcohol drinking and one-carbon metabolism-related gene polymorphisms on pancreatic cancer risk. Cancer Epidemiol Biomarkers Prev 17: 2742–2747.
42. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.
43. Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997–1004.
44. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.
45. Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, et al. (2001) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 29: 11–16.

46. Toda Y, Kono K, Abiru H, Kokuryo K, Endo M, et al. (1999) Application of tyramide signal amplification system to immunohistochemistry: a potent method to localize antigens that are not detectable by ordinary method. Pathol Int 49: 479–483.

47. Lundberg K, Kinloch A, Fisher BA, Wegner N, Wait R, et al. (2008) Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. Arthritis Rheum 58: 3009–3019.

OPEN ACCESS Freely available online

PLoS one

Dysregulation of IFN System Can Lead to Poor Response to Pegylated Interferon and Ribavirin Therapy in Chronic Hepatitis C

Koji Onomoto[1,7¤], Shiho Morimoto[1], Takahisa Kawaguchi[2], Hidenori Toyoda[3], Masami Tanaka[4], Masahiko Kuroda[4], Kazuko Uno[5], Takashi Kumada[3], Fumihiko Matsuda[2], Kunitada Shimotohno[6], Takashi Fujita[1], Yoshiki Murakami[2]*

1 Institute for Viral Research and Graduate School of Bioscience, Kyoto University, Kyoto, Japan, **2** Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **3** Department of Gastroenterology, Ogaki Municipal Hospital, Ogaki, Japan, **4** Department of Molecular Pathology, Tokyo Medical University, Tokyo, Japan, **5** Louis Pasteur Center for Medical Research, Kyoto, Kyoto, Japan, **6** Research Institute, Chiba Institute for Technology, Narashino, Japan, **7** Research Institute for Science and Engineering, Waseda University, Tokyo, Japan

Abstract

Background: Despite being expensive, the standard combination of pegylated interferon (Peg-IFN)- α and ribavirin used to treat chronic hepatitis C (CH) results in a moderate clearance rate and a plethora of side effects. This makes it necessary to predict patient outcome so as to improve the accuracy of treatment. Although the antiviral mechanism of genetically altered IL28B is unknown, IL28B polymorphism is considered a good predictor of IFN combination treatment outcome.

Methodology: Using microarray, we quantified the expression profile of 237 IFN related genes in 87 CH liver biopsy specimens to clarify the relationship between IFN pathway and viral elimination, and to predict patients' clinical outcome. In 72 out of 87 patients we also analyzed IL28B polymorphism (rs8099917).

Principal Findings: Five IFN related-genes (IFI27, IFI 44, ISG15, MX1, and OAS1) had expression levels significantly higher in nonresponders (NR) than in normal liver (NL) and sustained virological responders (SVR); this high expression was also frequently seen in cases with the minor (TG or GG) IL28B genotype. The expression pattern of 31 IFN related-genes also differed significantly between NR and NL. We predicted drug response in NR with 86.1% accuracy by diagonal linear discriminant analysis (DLDA).

Conclusion: IFN system dysregulation before treatment was associated with poor IFN therapy response. Determining IFN related-gene expression pattern based on patients' response to combination therapy, allowed us to predict drug response with high accuracy. This method can be applied to establishing novel antiviral therapies and strategies for patients using a more individual approach.

Citation: Onomoto K, Morimoto S, Kawaguchi T, Toyoda H, Tanaka M, et al. (2011) Dysregulation of IFN System Can Lead to Poor Response to Pegylated Interferon and Ribavirin Therapy in Chronic Hepatitis C. PLoS ONE 6(5): e19799. doi:10.1371/journal.pone.0019799

Editor: Mathias Lichterfeld, Massachusetts General Hospital, United States of America

Received December 21, 2010; **Accepted** April 11, 2011; **Published** May 13, 2011

Funding: Y.M. was financially supported by the Japanese Ministry of Health, Labour and Welfare. Grants-in-Aid for scientific research were received from the Ministry of Education, Culture, Sports, Science and Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ymurakami@genome.med.kyoto-u.ac.jp

¤ Current address: Division of Molecular Immunology, Medical Mycology Research Center, Chiba University, Chiba, Japan

Introduction

Hepatitis C virus (HCV) infection affects more than 3% of the world population. Without suitable treatment, chronic hepatitis C (CH) frequently leads to the development of chronic liver diseases such as liver cirrhosis (LC) and hepatocellular carcinoma (HCC) [1]. The current standard treatment for CH is a combination of pegylated-IFN (Peg-IFN)- α and ribavirin (hereafter CH combination therapy). Over a 15-year observation period, the rate of hepatocarcinogenesis was found to be significantly lower in sustained viral responders (SVR) and relapse (R) patients than in non responders (NR) and interferon (IFN) untreated patients [2]. However, CH combination therapy achieves a sustained virological response in 50–55% of patients with HCV genotype 1b infection [3]. Consequently, this creates a pressing need to develop alternative strategies for treating CH.

IFN Type-I and III play various important immunomodulatory roles in both innate immune and acquired immune responses. Four main effector pathways of the IFN-mediated antiviral response have been recognized by gene targeting studies: the Mx GTPase pathway, the 2′, 5′-oligoadenylate-synthetase-directed ribonuclease L (OASL) pathway, the protein kinase R (PKR) pathway and the interferon stimulated gene (ISG) 15 ubiquitin-like pathway. These effector-pathways individually block viral

transcription, degrade viral RNA, inhibit translation and modify protein function to control all steps of viral replication [4–5].

IFN treatment for CH usually results in a high incidence of side effects; therefore, it is important to adjust IFN treatment accurately using a prediction method. Viral factors (HCV genotype, pretreatment viral load, and sequence of HCV gene core and NS5A), [6–7] host factors (obesity, cirrhosis, ethnic background, serum cytokine levels, liver fibrosis grades) [8], and treatment factors (adequate course of treatment, adherence to the treatment, management of side effects) [9] has been utilized in prior research to predict the outcome of combination therapy. Hepatic microRNA expression pattern before anti-viral treatment has also been utilized as a prediction biomarker of drug response in CH [10], while other studies have shown that there is a possible association between two SNPs near the gene interleukin 28B (IL28B) on chromosome 19 and lack of response to combination therapy [11–13].

In this study, we evaluated the IFN related gene expression profiles in CH patients before administering CH combination treatment. After the anti-viral therapy, patients were classified according to their clinical outcome: sustained viral response (SVR), relapse (R), and non responder (NR). It was observed that in the NR group, the expression level of some IFN related genes was significantly higher than that in normal liver (NL) groups, and that the expression level of the other IFN related genes was significantly lower than in NL. Moreover, the significantly high expression of IFN related genes was associated with low response to combination therapy. This suggests that dysregulation of the IFN system can be related to cases of CH combination therapy failure.

Results

In order to provide specific information with less data analysis, we developed a custom-made focused DNA microarray called Genopal (Mitsubishi Rayon, Tokyo, Japan) using genes that target human innate-immunity. Based on the results from the expression profiles, we carefully selected 237 gene probes (materials and methods) by activating RIG-I with Agilent DNA microarray. A microarray platform was used to establish IFN-related gene expression profiles in the specimens collected from the 87 CH and 5 NL samples (Table 1). The results of the analysis of these genes using the DNA chip strongly correlated with those obtained by real-time PCR (Pearson's correlation coefficient $R2 = 0.996$, $P<0.0001$; data not shown).

IFN related genes associated with the final response to combination therapy

We determined unique IFN gene expression patterns for liver specimens with or without HCV based on the final virological response to the combination therapy. The expression level of 66 genes significantly differed among NR, R, SVR, and normal liver (NL) groups (Figure 1). To clearly identify the IFN-related genes associated with the clinical outcome, we extracted genes that showed significant differences ($p<0.05$). It was observed that the expression level of 5 genes (myxovirus (influenza virus) resistance 1 (MX1), 2″,5″-oligoadenylate synthetase 1 (OAS1), ISG15 ubiquitin-like modifier (ISG15), interferon, alpha-inducible protein 27 (IFI27), and interferon, alpha-inducible protein 44 (IFI44)) were significantly higher in NR than in SVR samples (Table 2). The expression levels of 3 genes (MX1, IFI27, and ISG15) were significantly higher in NR than in R samples (Table 2). We also analyzed the IFN-related genes expression pattern according to the grade of inflammation or stage of fibrosis, however, no significant differences was observed between the two (data not shown).

Table 1. Clinical characteristics of patients.

Characteristics	SVR (n = 38)	R (n = 26)	NR (n = 23)	NL (n = 5)
Age	56.7±10.3	61.3±8.6	60.8±7.8	57.2±9.5
Male (%)	28 (61%)	11 (39%)	9 (36%)	3(60%)
Weight (kg)	59.5±8.9	57.2±10.3	55.7±7.2	ND
HCV RNA ($\times10^6$ copies/ml)	2.00±2.07	1.79±1.02	1.55±0.95	ND
Fibrosis stage				
F 0	1	1	1	
F 1	29	13	10	
F 2	9	7	5	
F 3	6	4	6	
F 4	0	0	1	
WBC($\times10^3$/mm^3)	5.42±1.63	5.23±1.25	4.69±1.13	ND
Hemoglobin (g/dl)	14.3±1.14	13.5±1.35	13.6±1.09	ND
Platelet ($\times10^4$/mm^3)	16.7±5.3	16.6±4.0	15.0±5.7	ND
AST (IU/L)	59.2±51.0	48.7±30.1	57.4±29.7	ND
ALT (IU/L)	80.8±93.7	49.3±29.6	69.1±44.4	ND
γGTP (IU/L)	60.3±74.2	41.2±29.7	76.2±60.2	ND
ALP (IU/L)	255±74.0	246±71.3	314±144	ND
Total bilirubin (mg/dl)	0.66±0.22	0.73±0.31	0.69±0.19	ND
Albumin (g/dl)	4.20±0.34	4.14±0.25	4.02±0.48	ND

Abbreviations; NR, non-virological responder; R, relapse; SVR, sustained virological responder; AST, aspartate aminotransferase;ALT, alanine aminotransferase; WBC, white blood cell; ALP, alkaline phosphatase; γGTP, gamma-glutamyl transpeptidase; ND, not detected.
doi:10.1371/journal.pone.0019799.t001

Comparison of IFN related genes between CH and NL

We also compared the gene expression pattern in NR and NL. After extracting genes with a fold change $<1/3$, $3<$ and p-value<0.05, we found that the expression level of 6 genes (growth arrest and DNA-damage-inducible, beta (GADD45B), hairy and enhancer of split 1 (HES1), B-cell CLL/lymphoma 3 (BCL3), signal transducer and activator of transcription 3 (STAT3), suppressor of cytokine signaling 3 (SOCS3), and DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (DDX11)) was significantly lower in NR than in NL. The expression level of SOCS3 and DDX11 in NR was significantly lower than in SVR. The expression level of 25 genes were significantly higher in NR than in NL. The expression levels of most of these genes were significantly higher in NR than in SVR, but the expression level of tumor necrosis factor (ligand) superfamily, member 10 (TRAIL), major histocompatibility complex, class I, C (HLA–C), major histocompatibility complex, class I, B (HLA–B), and chemokine (C-X-C motif) ligand 10 (CXCL10 (IP10)) were similar in NR and SVR samples (Table 3).

Validation of the microarray result by real-time qPCR

The five genes (ISG15, MX1, OAS1, IFI27 and IFI44) with the largest difference in fold change between NR and SVR groups were chosen to confirm the microarray results using real-time

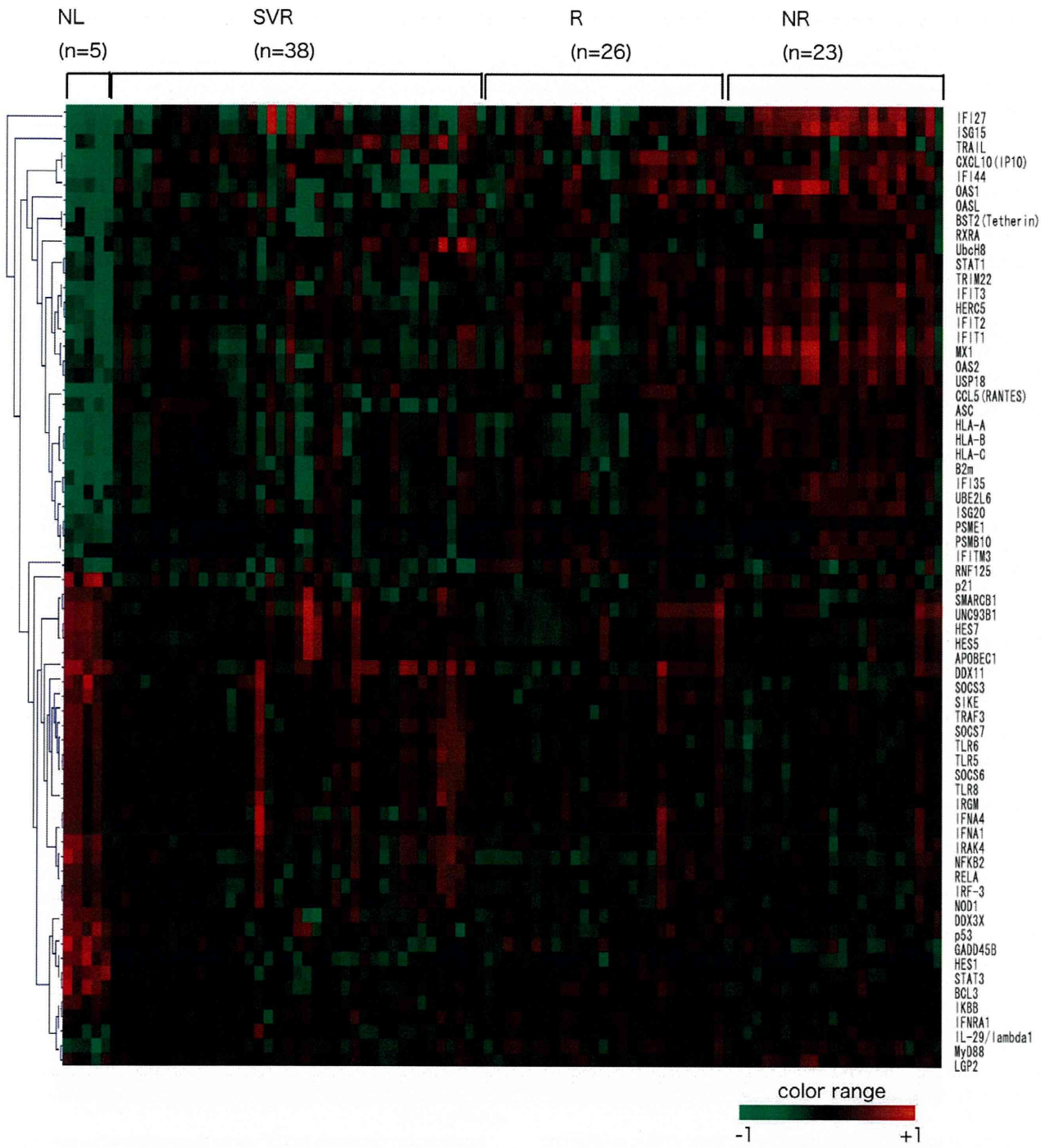

Figure 1. Clustering of IFN related gene expression. Clustering of CH patients according to the expression profiles of the 66 genes that showed significant differences among SVR, R, NR, and NL. Vertical bars represent the IFN related genes and the horizontal bars represent the samples. Green bars reflect down-regulated genes and red bars up-regulated genes.
doi:10.1371/journal.pone.0019799.g001

qPCR. The result from real-time qPCR supported the results from the microarray analysis (Figure S1).

Prediction of the clinical outcome by DLDA

We attempted to simulate the clinical outcome of the CH combination therapy using diagonal linear discriminant analysis (DLDA). Patients were randomly divided into TS (training set) and VS (validation set) (Table 4) in the order in which their samples were obtained. Samples within each group were then classified as NR or non-NR (SVR+R). DLDA showed that the accuracy, sensitivity, specificity, positive and negative predictive value of these two classifications were 86.1%, 87.5%, 81.8%, 93.3%, and

Table 2. Extracted genes related to the clinical outcome with a fold change greater than or equal to 1.5 between two groups (NR/SVR, NR/R) ($p<0.05$).

Accession No.	gene	symbol	fold change (NR/SVR)	p-value
NM_006417.4	interferon, alpha-inducible protein 44	IFI44	2.13	2.01E-03
NM_005532.3	interferon, alpha-inducible protein 27	IFI27*	2.37	2.01E-03
NM_016816.2	2″,5″-oligoadenylate synthetase 1, 40/46kDa, transcript variant 1	OAS1	2.51	1.36E-02
NM_005101.2	ISG15 ubiquitin-like modifier	ISG15*	2.68	1.18E-03
NM_002462.2	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	MX1*	2.71	1.57E-03
Accession No.	gene	symbol	fold change (NR/R)	p-value
NM_002462.2	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	MX1*	2.27	1.11E-03
NM_005532.3	interferon, alpha-inducible protein 27	IFI27*	2.33	1.69E-03
NM_005101.2	ISG15 ubiquitin-like modifier	ISG15*	2.5	1.11E-03

Asterisk deposits extracted genes that are common to both SVR and NR and to NR and R.
doi:10.1371/journal.pone.0019799.t002

69.2% respectively (Table 5). Additionally, we attempted to predict (1) SVR and nonSVR (R+NR), and (2) SVR, R, and NR by DLDA. The accuracy with which patients were classified as SVR and nonSVR, was 56.8% and as SVR, R, and NR was 56.9%.

Genetic variation of IL28B is correlated with the expression of IFN related genes

To examine the relationship between the genetic variation of IL28B and IFN related gene expression, we determined the IL28B polymorphism in 72 patients (Table 6). Patients with the minor genotype of IL28B displayed higher levels of hepatic ISGs expression, whereas patients with the major genotype showed significantly lower expression levels (Figure 2A). In order to further widen our understanding of the above relationship, we significantly identified individual genetic variations in IL28B at the clinical outcome (Figure 2B). We then individually compared the expression level of several IFN-lambda related genes at the clinical outcome with the genetic variation of IL28B. The expression level of interleukin 28A (IL28A), IL28B, interleukin 29 (IL29), interleukin 10 receptor, beta (IL10RB), signal transducer and activator of transcription 1 (STAT1), STAT5A, and tyrosine kinase 2 (TYK2) in IL28B genotype minor allele and major allele did not differ; however, the expression level of STAT5A and IRF9 was significantly higher in IL28B minor allele cases than in major allele (Figure 3A). The expression levels of these nine genes did not significantly differ among the clinical outcomes (NR, R, and SVR) (Figure 3B).

Finally, in regards to genes which contribute to IFN production (interferon regulatory factor 7 (IRF7), interleukin-1 receptor-associated kinase 1 (IRAK1), myeloid differentiation primary response gene (MyD88), and toll-like receptor 7 (TLR7)) there was not much difference in their expression level prior to CH combination treatment and their expression level at the clinical outcome (Figure 4A) [14]. Unlike IRF7 and MyD88, there was no significant difference in the expression level of IRAK1 and TLR7 according to the IL28B genetic variation (Figure 4B). When we attempted to predict NR and nonNR by using ISG genes with and without IL28B polymorphism using DLDA by using 72 patients (36 patients for training set, 36 patients for validation set). DLDA with IFN related gene and IL28B polymorphism showed that the accuracy, sensitivity, specificity, positive and negative predictive value of these two classifications were 83.3%, 85.1%, 77.8%, 92.0%, 63.6%, respectively (Table 7). DLDA with IFN related gene only showed that the accuracy, sensitivity, specificity, positive and negative predictive value were 83.3%, 81.5%, 88.9%, 95.7%, 61.5%, respectively (Table 8).

Discussion

Our comprehensive analysis identified 66 genes with expression levels that consistently differed depending on the drug response of 87 CH patients and 5 normal liver specimens (Figure 1). Comparing the gene expression pattern in NR and NL showed the expression levels of 31 genes were significantly different (Table 3). In addition, most genes with expression levels in NR that were higher or lower than in NL, also differed between NR and SVR. Therefore, it is possible that innate immunity in the early period of HCV infection strongly influences IFN reaction.

HCV infection induces the impairment of cell subset number and the function of plasmacytoid dendritic cells (PDC) and natural killer cells [15]. The amount of PDC, which are the most potent producers of antiviral Type-I and III IFN [16], decreased in patients' peripheral blood [17], however, PDC was trapped in the HCV infected liver tissue. Therapeutic non-responders had increased PDC migration to inflammatory chemokines before therapy, compared with therapeutic responders [18]. This situation resulted in elevated expressions of IFN-related genes in the CH samples and was associated with their inability to eliminate the virus [19].

Inadequate expression of IFN related genes has been associated with several diseases. High expression of ISG can induce a refractory state in IFN therapy [20] and impaired IFN production leads to high risk of HCV-related hepatocarcinogenesis [21]. Lymphocyte IFN signaling was less responsive in patients with breast cancer, melanoma, and gastrointestinal cancer and these defects may represent a common cancer-associated mechanism of immune dysfunction. Alternately, since immunotherapeutic strategies require functional immune activation, such impaired IFN signaling may hinder therapeutic approaches designed to stimulate anti-tumor immunity [22]. In this way, the dysregulation of the IFN system can influence the progression of diseases and decrease curative effects.

Table 3. List of genes that had significantly different expression levels in NR and NL (fold change <1/3, 3<, and $p<0.05$).

symbol	NR/NL (fold change)	NR/NL (t-test)	NR/SVR (fold change)	NR/SVR (t-test)
GADD45B	0.20	1.14E-02	1.01	NS
HES1	0.26	1.26E-03	0.97	NS
BCL3	0.26	1.84E-02	1.02	NS
STAT3	0.26	5.81E-04	0.97	NS
SOCS3	0.27	7.96E-03	0.68	2.15E-02
DDX11	0.28	4.33E-05	0.59	9.52E-03
TRIM22	3.06	2.91E-03	1.37	7.97E-03
ASC	3.19	1.35E-03	1.33	4.07E-03
UBE2L6	3.32	1.06E-02	1.41	1.01E-03
STAT1	3.38	6.04E-04	1.33	1.86E-02
ISG20	3.64	2.42E-04	1.42	2.37E-03
TRAIL	3.81	2.08E-02	0.78	NS
OAS2	4.02	2.91E-03	1.89	1.07E-04
IFIT2	4.60	1.48E-03	1.56	8.34E-05
BST2(Tetherin)	5.14	8.17E-03	1.49	5.67E-04
IFI35	5.29	1.35E-03	1.63	2.37E-05
HERC5	5.32	1.16E-03	1.68	4.07E-05
MX1	6.21	1.33E-03	2.94	8.46E-07
HLA-C	6.49	6.34E-04	1.21	NS
CCL5(RANTES)	6.73	5.48E-04	1.25	3.77E-02
HLA-B	6.84	4.91E-04	1.22	NS
OAS1	7.80	5.52E-04	2.75	1.92E-04
HLA-A	8.49	5.92E-05	1.41	9.08E-04
B2m	9.09	7.78E-04	1.25	1.89E-02
IFIT1	9.42	1.86E-03	2.11	1.41E-05
OASL	10.38	3.97E-06	1.48	1.24E-02
IFIT3	10.45	4.33E-05	2.11	5.63E-06
CXCL10(IP10)	15.67	8.89E-07	1.28	NS
IFI44	17.00	9.40E-05	2.22	4.83E-06
ISG15	21.12	1.05E-04	2.85	3.99E-05
IFI27	43.74	1.80E-05	2.56	5.62E-05

doi:10.1371/journal.pone.0019799.t003

Genes which participate in IFN production (TLR7, MyD88, IRAK1, and IRF7) did not show any significant difference in their expression level prior to CH combination therapy, and their level at the clinical outcome (Figure 4A and 4B). However, the gene expression pattern of down-stream IFN pathway genes (IFI27, IFI44, ISG15, MX1, and OAS1) was significantly different among SVR, R, and NR (Table 2). IFN is usually up-regulated in HCV infected cells; however in some cases, the mechanism that controls IFN becomes abnormal, and the expression levels of IFN and ISG remain high without any curative effect [23]. The ISG family was generally up-regulated in NR compared to SVR [24–27] and this high expression of ISG related genes was associated with poor response to IFN therapy in previous, as well as in this present study. ISG15 has been linked to innate immune response to viruses and to cellular response to IFN. Although over-expression of ISG15 enhances the antiviral activity of IFN in vitro in acute infection [28], in chronic infection, extended pre-activation of IFN induced genes leads to dysregulation of the IFN system.

CH therapy is still imperfect at present and therefore suitable prediction methods are necessary to avoid adverse effects. Treatment failure using CH combination therapy is associated with up-regulation of a specific set of IFN-responsive genes thereby making it possible to predict non-response to exogenous therapy [29]. Early gene expression during anti-HCV therapy may elucidate important molecular pathways that might be influencing the probability of achieving a virological response [30]. Our study supports this fact by demonstrating that CH and NL differ fundamentally in their innate response to CH combination therapy.

IFN related gene expression suggests novel aspects of HCV pathogenesis, and form the basis for a subset of genes that can predict treatment response before initiation of combination therapy. After proper external validation, these gene sets may provide the basis for a diagnostic biomarker that can determine early on whether a patient treated with combination therapy is likely to be NR or not. In this respect, what sets our analysis apart is the effect of using DLDA to predict final response with high accuracy in NR and non-NR groups. This prediction showed that the expectation in NR (proportion of actual non-NR versus the predicted number of non-NR) was 93.3% and overall accuracy was 86.1%. In prior report, Dill et al. successfully predicted SVR, but were unable to predict R and NR with high accuracy [31]. In our experiments on the other hand, we predicted NR with high accuracy but were unable to do so for SVR and R. Possible causes for differences between our results and those received by Dill et al. may be (1) the differences in the races of subjects; European patients vs. Japanese patients in our study, (2) the composition of genotype; genotype 1 and 4 vs. genotype 1b in our study, and (3) the difference of the ISG genes extracted.

Genome-wide association studies have described alleic variants near the IL28B gene that are associated with treatment response and with spontaneous clearance of HCV [11–13]. In order to clarify the relationship between IL28B polymorphism and drug response, we compared the expression level of IFN-lambda related gene at the clinical outcome with any genetic variation in IL28B. The expression of hepatic ISG and related genes was strongly associated with treatment response and genetic variation of IL28B [32]. Classification of the patients into SVR and NR revealed that ISG expression was conditionally independent of the IL28B genotype. In CH patients in Europe, the expression pattern of genes induced by IFN more accurately predicts CH combination treatment clinical outcome than polymorphism of IL28B [31]. We observed that curative effect prediction using IFN gene expression pattern resulted in high level of accuracy, however, IFN with IL28B or IFN alone resulted in approximately similar levels of accuracy, therefore, the polymorphism of IL28B did not contribute significantly to our prediction. These findings are accordance with Dill et al. results (Table 7). There was an increased expression in NR compared to SVR irrespective of the IL28B genotype. However, there was no significant difference in their expression at the clinical outcome or in the genetic variation of IL28B (Figure 3A and 3B). Genetic variation of IL28B polymorphism is effective in predicting curative effect; however, the reason for this is not fully understood.

In conclusion, comprehensive analysis of IFN related gene showed that dysregulation of the IFN system might be related to treatment failure and that IFN related gene expression before treatment can enable accurate prediction of CH combination therapy clinical outcome. By focusing the full course of treatment on only those patients who have the highest likelihood of achieving