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

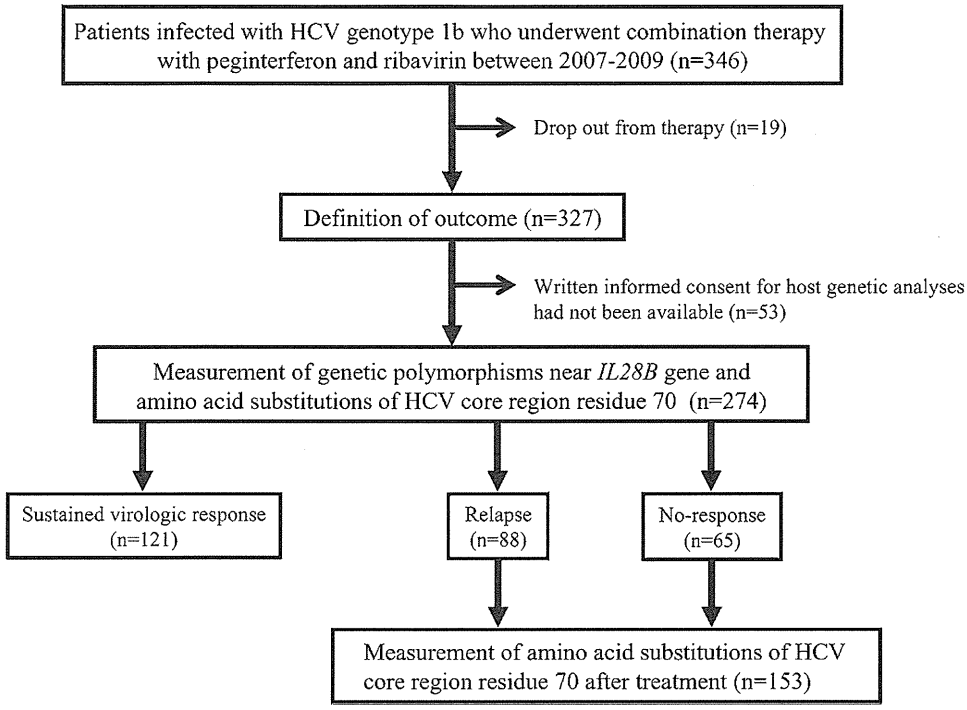

Fig. 1. Schematic representation of the study design.

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.

Journal of Human Genetics (2011), 1–5

www.nature.com/jhg

ORIGINAL ARTICLE

A polymorphism of the *POLG2* gene is genetically associated with the invasiveness of urinary bladder cancer in Japanese males

Chanavee Ratanajaraya[1], Hiroyuki Nishiyama[2], Meiko Takahashi[1,3], Takahisa Kawaguchi[1], Ryoichi Saito[2], Yoshiki Mikami[4], Mikita Suyama[1], Mark Lathrop[5,6], Ryo Yamada[1], Osamu Ogawa[2] and Fumihiko Matsuda[1,3,5,7]

Urinary bladder cancer (UBC) is a common cancer with male predominance. Pathologically it is classified into two distinct tumor entities related to the risk of patients. The low-grade tumors with relatively well-differentiated tumor histology (G1 and G2) at stage Ta are non-invasive and pose a minimal risk, whereas high-grade tumors (G2 and G3) with stages T1 to T4 are aggressive with invasion, and therefore, pose a serious risk for the patients. DNA repair and metabolic process genes may have major roles in cancer progression and development. To identify genes associated with invasiveness of UBC, we have extensively genotyped 802 single nucleotide polymorphisms in 114 genes related to DNA repair mechanisms and metabolic processes. A genetic association study was performed between non-invasive (G1 and G2 with Ta) and invasive (G2 and G3 with T1 to T4) groups of Japanese UBC patients. We found that rs17650301 in *POLG2* showed marked difference in genotype distribution between the two groups in males ($P=6.93\times10^{-4}$), which was further confirmed in an independent sample set (overall $P=1.67\times10^{-4}$). We also found by an *in silico* analysis that the risk allele of rs17650301 increased the transcription of *POLG2*. In conclusion, rs17650301 is a good candidate marker for UBC invasiveness in Japanese males.
Journal of Human Genetics advance online publication, 7 July 2011; doi:10.1038/jhg.2011.60

Keywords: genotyping; invasiveness; *POLG2*; single nucleotide polymorphism; urinary bladder cancer

INTRODUCTION

Urinary bladder cancer (UBC) is one of the most common cancers of the urinary system and is much more common in males worldwide, including in Japan. UBC is classified by the stage and grade of the tumor, which are highly correlated with recurrence, progression and patient survival rates. Tumor stage is determined by the degree of invasiveness and metastasis, whereas tumor grade classification is based on the degree of differentiation. Non-invasive UBCs are designated as stage Ta, whereas stages T1, T2, T3 and T4 refer to invasion into the subepithelial connective tissue, muscle, perivesical tissue and adjacent organs, respectively. With regard to tumor grades, the degrees of differentiation are based on tumor histology: well differentiated (G1), moderately differentiated (G2) and poorly differentiated (G3) cancer.[1,2] Most of the UBC cases clinically fall into one of the two distinct tumor entities at diagnosis related to the risk of patients.[3,4] The first comprises non-invasive low-grade carcinoma of relatively well-differentiated histology (G1 or G2) with stage Ta, which rarely progresses to a higher stage and pose a minimal risk for the patients. On the other hand, the second refers to aggressive tumors corresponding to stages T1 to T4 with higher grades (G2 to G3), showing high recurrence and progression rates, and therefore pose a serious risk for the patients.

Carcinogen exposure is one of the major risk factors contributing to UBC incidence.[5] Genes involved in DNA repair mechanisms, DNA replication, transcription, DNA damage signaling, cell cycle and metabolic processes influence the development and progression of UBC. A number of genetic polymorphisms in these types of genes were reported to be associated with genetic susceptibility to UBC.[6–8] However, genetic association analyses focusing on the invasiveness of UBC are yet to be performed.

In this study, we focused on 121 genes acting on the four pathways of DNA repair (base excision repair, nucleotide excision repair, double strand break repair and mismatch repair), as well as those related to DNA synthesis, cell cycle and metabolism. We designed a panel of 1536 single nucleotide polymorphism (SNP) markers from these genes for extensive genotyping study, of which 802 SNPs corresponding to 114 genes that passed quality control process were statistically analyzed for their association with UBC invasiveness in Japanese patients.

[1]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; [2]Department of Urology, Kyoto University Graduate School of Medicine, Kyoto, Japan; [3]Institut National de la Santé et de la Recherche Médicale (INSERM) Unit U852, Kyoto University Graduate School of Medicine, Kyoto, Japan; [4]Department of Diagnostic Pathology, Kyoto University Hospital, Kyoto, Japan; [5]Centre National de Génotypage, Institut Génomique, Commissariat à l'Énergie Atomique (CEA), Evry, France; [6]Foundation Jean Dausset Centre d'Étude du Polymorphisme Humain (CEPH), Paris, France and [7]CREST program, Japan Science and Technology Agency, Kawaguchi, Japan
Correspondence: Professor F Matsuda, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Yoshida, Kyoto 606-8501, Japan.
E-mail: fumi@genome.med.kyoto-u.ac.jp
Received 4 February 2011; revised 20 April 2011; accepted 7 May 2011

MATERIALS AND METHODS

Study population

DNA samples were collected at the Department of Urology, Kyoto University, Kyoto, Japan. The stage and the grade of UBC were judged by the histological examinations at diagnosis. Non-invasive group (group-N) corresponds to stage Ta patients with grades G1 or G2, whereas invasive group (group-I) refers to those of stages T1 to T4 with grades G2 or G3. The characteristics and the detailed clinical diagnosis of the study populations were summarized in Table 1 and Supplementary Table 1, respectively. All the patients provided their written informed consents according to the protocols approved by the ethical review board of Kyoto University Graduate School of Medicine.

Selection of candidate genes and construction of SNP genotyping panel

Identification of SNPs was performed by resequencing exons and flanking regions of 121 genes (summarized in Supplementary Table 2), using 32 each of DNA samples of Japanese, French, Congolese and Thai populations as reported.[9] Alignment and genotyping were performed using the Genalys software (Centre National de Génotypage, Institut Génomique, Commissariat à l'Énergie Atomique (CEA), Evry, France, http://software.cng.fr/).[10] As for the selection of SNP markers, SNPs located in linkage disequilibrium (LD) blocks encompassing the 121 genes and essentially covering all the major haplotypes with frequencies greater than 5% were initially chosen from the International HapMap Project data of Japanese, Caucasians and Africans.[11] Additional SNPs identified in at least one of the four populations by the SNP discovery were included in the selection. Finally, we generated a panel of 1536 SNPs, which were chosen as tag SNP markers for genetic analyses of multiple populations using GoldenGate technology (Illumina, San Diego, CA, USA).

SNP genotyping

Genomic DNAs were prepared from peripheral blood leukocytes using Qiagen DNA Extraction kits (Qiagen, Hilden, Germany). The first screening was performed using GoldenGate assay (Illumina). The validation analysis of marker rs17650301 was performed with additional samples by sequencing. The forward primer (5′-AGGCTGGTAGGGTCCAAAGT-3′) and the reverse primer (5′-AGGGTTAGGTTGAGCATCCC-3′) were used for the PCR and an internal primer (5′-GAAGTTTTCACCGTGTTGCC-3′) was used for the sequencing with BigDye Terminator v3.1 Cycle Sequencing kit (Life Technologies, Foster City, CA, USA).

Quality control and association analysis

Among the 1536 SNPs genotyped, 20 SNPs on X-chromosome and one triallelic SNP were initially removed before quality control process. During the quality control, we first examined the presence of DNA samples showing high degrees of kinship by plink (Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA, http://pngu.mgh.harvard.edu/~purcell/plink/)[12] and found there were no such samples. Three DNA samples with call rates smaller than 0.90 were removed. Regarding the SNP markers, 104 SNPs with calling rates smaller than 0.95 and 609 SNPs with minor allele frequencies smaller than 0.01 were also excluded. After these steps, the remaining 802 SNP markers corresponding to 114 genes were used for statistical analyses. The association was examined by trend exact test for genotype distribution. Odds ratios and 95% confidence intervals were estimated with additive model. The heterogeneity of the effect size between the first screening and the validation analysis was evaluated by Breslow–Day test.[13] The LD structure of the *POLG2* locus was established by using Japanese data from the International HapMap Project.[11]

Table 1 Characteristics of UBC patients used for statistical analysis

Study	*Group-N*	*Group-I*
First screening		
Number	106	93
Number of males (%)	75 (70.8)	70 (75.3)
Age at diagnosis (mean ± s.d.)	68.5 ± 11.9	68.8 ± 10.5
Validation analysis		
Number	65	114
Number of males (%)	50 (76.9)	98 (86.0)
Age at diagnosis (mean ± s.d.)	67.9 ± 13.3	71.1 ± 10.0
Combined analysis		
Number	171	207
Number of males (%)	125 (73.1)	168 (81.2)
Age at diagnosis (mean ± s.d.)	68.3 ± 12.4	70.1 ± 10.3

Abbreviations: Group-I, invasive group; Group-N, non-invasive group; UBC, urinary bladder cancer.

Expression quantitative trait loci analysis

Allele-specific effect on the expression of rs17650301 was performed using Japanese data set GSE6536 of expression profile Gene Expression Omnibus (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, USA, http://www.ncbi.nlm.nih.gov/geo/).[14] The association *P*-values were obtained by Jonckheere-Terpstra method using R software (Department of Statistics, The University of Auckland, Auckland, New Zealand, http://www.r-project.org/) or SPSS (version 18, IBM, Armonk, NY, USA, http://www-01.ibm.com/software/data/spss/products/statistics/statistics-desktop/).

RESULTS

First screening

Among 1536 SNPs corresponding to 121 candidate genes in the genotyping panel, a total of 802 SNPs derived from 114 candidate genes were examined for their association with UBC invasiveness, using 106 group-N samples and 93 group-I samples (for details, see Materials and Methods). Although no SNPs showed significant association ($P=6.23\times10^{-5}$) after Bonferroni's correction for multiple testing, we identified 55 SNPs with nominal *P*-values smaller than 0.05, of which nine SNPs showed *P*-values smaller than 0.01 (Table 2). The strongest *P*-value was obtained for rs17650301 ($P=3.40\times10^{-3}$) located in intron 5 of DNA polymerase subunit γ2 gene (*POLG2*) at 17q24.1. Five SNPs in the *POLL* gene, rs3730476, rs1055364, rs3730465, rs3730472 and rs3095795, were in complete LD with each other.

We then performed a subgroup analysis by dividing the UBC patients according to their gender. In the male study, six SNPs showed *P*-values smaller than 0.01, of which the strongest association was again obtained for rs17650301 ($P=6.93\times10^{-4}$) (Table 2). However, rs17650301 did not show association in the female study. In addition, rs2518968, which showed the second lowest *P*-value in the joint study ($P=4.50\times10^{-3}$), showed association in the male study ($P=4.81\times10^{-3}$), although it was much weaker than that of rs17650301.

Validation analysis

We focused on rs17650301 and performed a validation analysis by direct sequencing, using DNA samples of an independent set of UBC patients consisting of 65 group-N samples and 114 group-I samples. Although the association was not observed in the male–female joint study ($P=0.432$), we obtained a marginal *P*-value ($P=0.0396$) in the male study (Table 3). The combined analysis by pooling the genotypes of the two studies returned a *P*-value of 9.45×10^{-3} and 1.67×10^{-4} in the joint study and in the male study, respectively (Table 3). The *P*-value of heterogeneity did not show evidence of differentiation in effect size ($P=0.17$ for the joint study and $P=0.47$ for the male study). A logistic regression analysis adjusted for age and sex returned a *P*-value of 5.28×10^{-3} for the joint study.

Table 2 Results of association analysis in the first screening

				Non-invasive (N)				Invasive (I)				
Study	*SNP ID*	*A1/A2*[a]	*Gene*	*A1/A1*	*A1/A2*	*A2/A2*	*Freq. A1*	*A1/A1*	*A1/A2*	*A2/A2*	*Freq. A1*	P-value[b]
Male and female												
	rs17650301	A/C	*POLG2*	55	41	10	0.712	31	43	19	0.565	3.40×10^{-3}
	rs2518968	C/G	*BLM*	30	55	21	0.542	46	36	11	0.688	4.50×10^{-3}
	rs3730476	A/G	*POLL*	81	23	0	0.889	85	5	1	0.962	7.29×10^{-3}
	rs1055364	C/A	*POLL*	83	23	0	0.892	87	5	1	0.962	7.31×10^{-3}
	rs3730465	A/G	*POLL*	83	23	0	0.892	87	5	1	0.962	7.31×10^{-3}
	rs3730472	T/G	*POLL*	83	23	0	0.892	87	5	1	0.962	7.31×10^{-3}
	rs3095795	T/C	*POLL*	83	23	0	0.892	87	5	1	0.962	7.31×10^{-3}
	rs1801127	G/A	*OGG1*	98	7	1	0.958	93	0	0	1.000	7.63×10^{-3}
	rs2276332	A/C	*ADH1A*	84	22	0	0.896	86	7	0	0.962	8.95×10^{-3}
Male												
	rs17650301	A/C	*POLG2*	43	26	6	0.747	22	32	16	0.543	6.93×10^{-4}
	rs8192772	T/C	*CYP2E1*	38	30	5	0.726	53	15	1	0.877	1.79×10^{-3}
	rs2518968	C/G	*BLM*	19	41	15	0.527	36	25	9	0.693	4.81×10^{-3}
	rs2070676	G/C	*CYP2E1*	4	23	45	0.215	1	10	57	0.088	6.34×10^{-3}
	rs9634161	T/C	*RAD52*	65	10	0	0.933	46	21	1	0.831	6.84×10^{-3}
	rs4937	T/C	*POLR2C*	25	38	11	0.595	13	34	22	0.435	8.68×10^{-3}
Female												
	rs12611088	G/A	*XRCC1*	12	18	1	0.677	18	5	0	0.891	4.40×10^{-3}
	rs2023614	C/G	*XRCC1*	26	5	0	0.919	12	9	2	0.717	9.92×10^{-3}
	rs3213266	G/A	*XRCC1*	26	5	0	0.919	12	9	2	0.717	9.92×10^{-3}

Abbreviation: SNP, single nucleotide polymorphism.
SNP markers that showed statistical *P*-values smaller than 0.01 are shown.
[a]The alleles A1 and A2 represent the reference and non-reference allele, respectively, in NCBI build 36.
[b]Statistical *P*-values were calculated by trend exact test.

Table 3 Summary of the association analysis of rs17650301 with UBC invasiveness

	Non-invasive (N)				Invasive (I)					
Study	*AA*	*AC*	*CC*	*Freq. A*	*AA*	*AC*	*CC*	*Freq. A*	*OR (95% CI)*	P-value[a]
First screening										
Male and female	55	41	10	0.712	31	43	19	0.565	1.91 (1.26–2.89)	**3.40×10^{-3}**
Male	43	26	6	0.747	22	32	16	0.543	2.48 (1.51–4.08)	**6.93×10^{-4}**
Validation analysis										
Male and female	32	26	7	0.692	49	49	16	0.645	1.24 (0.78–1.97)	0.432
Male	28	19	3	0.750	39	45	14	0.628	1.78 (1.04–3.05)	**0.0396**
Combined analysis										
Male and female	87	67	17	0.705	80	92	35	0.609	1.53 (1.13–2.08)	**9.45×10^{-3}**
Male	71	45	9	0.748	61	77	30	0.592	2.04 (1.43–2.93)	**1.67×10^{-4}**

Abbreviations: CI, confidence intervals; OR, odds ratio; UBC, urinary bladder cancer.
Reference (A) and variant (C) alleles are based on NCBI build36.
P-values that are less than 0.05 are shown in bold characters.
[a]Statistical *P*-values were calculated by trend exact test.

Also, the *P*-value did not change ($P=1.67\times10^{-4}$) after adjusting for age in the male study.

Linkage disequilibrium block

The LD block of the DNA region encompassing rs17650301 in the Japanese population was established using genotype data of the International Hapmap Project.[11] In the LD block encompassing a 34-kb DNA between rs8068311 and rs6504238, there were no SNP markers showing strong LD ($r^2\geqslant0.8$) with rs17650301, except rs1991401 ($r^2=1$). Rs1991401 is located in the 5′ untranslated region of the *DDX5* (DEAD box polypeptide 5) gene adjacent to *POLG2* (Figure 1).

Rs17650301 and expression of *POLG2*

Although rs17650301 is located in an intronic region, it may have functional impact in the transcription of *POLG2*. We performed an expression quantitative trait loci (eQTL) analysis by examining the expression profiles of 44 lymphoblastoid cells of Japanese origin in Gene Expression Omnibus database[14] and found an increase of

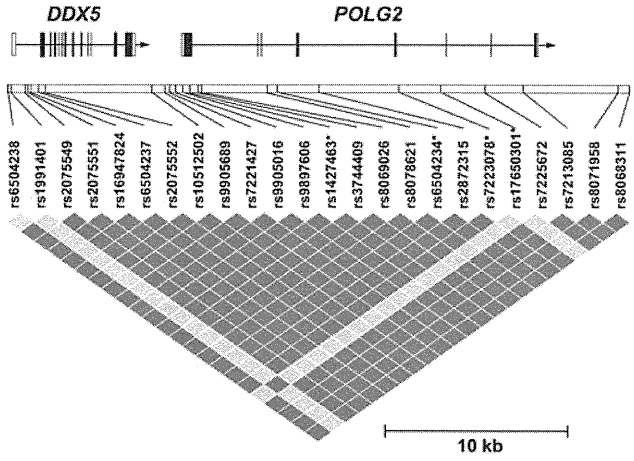

Figure 1 LD structure of the 34-kb region spanning the *POLG2* and *DDX5* genes. The structure and transcriptional polarity of *POLG2* and *DDX5* are shown according to NCBI Reference Sequence Build 36. Exons are shown by filled and open rectangles representing the coding and untranslated regions respectively. Pairwise LD estimation scores between SNPs within the region were converted into colors according to the color scheme of Haploview. The SNPs included in the genotyping panel for the first screening are indicated by an asterisk.

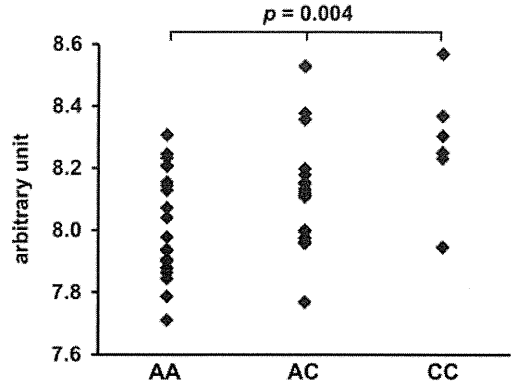

Figure 2 Allelic expression of *POLG2* with the genotypes of rs17650301. The levels of POLG2 transcription in cell lines carrying the three genotypes are indicated by a square. 'C' and 'A' represent the risk and the alternative alleles, respectively.

POLG2 transcription by the risk allele (C) of rs17650301 (P=0.004, Figure 2). However, a search for *cis*-regulatory elements using Jaspar database (The Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen, Denmark, http://jaspar.genereg.net/)[15] failed to identify any known motifs covering rs17650301.

DISCUSSION

In this study, we conducted a case-control association analysis of UBC invasiveness by taking a candidate gene approach of 1536 SNPs in 121 genes related to DNA repair mechanisms, DNA synthesis and metabolic processes. To our knowledge, this is the first genetic study focusing on the invasiveness of UBC. Statistical analysis was performed for 802 SNPs in 114 genes that passed quality control. The majority of the SNP markers (627 out of 734) excluded from the statistical test showed minor allele frequencies smaller than 0.01 in the Japanese population. This is due to the design of the genotyping panel for trans-ethnic study by covering SNPs identified in at least one of the four populations (Japanese, French, Thai and Congolese). Indeed, there were as many as 934 SNPs whose minor allele frequencies in the SNP discovery were smaller than 0.01 in Japanese and greater than 0.01 in at least one of the three other populations.

There was no SNP showing a significant association after Bonferroni's correction for multiple testing ($P=6.23\times10^{-5}$). The Bonferroni method is a very stringent test; although the possibilities of false positives are decreased, it comes together with the risk of losing potential candidates with true associations. In order to overcome this drawback, we took a strategy to screen SNP markers for potential associations in two stages. The most important aspect of our results is that the association of rs17650301 with UBC invasiveness was reproducibly obtained in two independent sample collections, without evidence of differentiation in effect size. However, the detection of significant association will require replication analyses using other sample collections.

The genotyping of 802 SNP markers followed by the validation analysis identified rs17650301 located in intron 5 of the *POLG2* gene as the strongest candidate for the invasiveness of UBC in Japanese male patients. *POLG2* encodes the 55-kDa accessory subunit of mitochondrial DNA polymerase. This subunit, together with the 140-kDa catalytic subunit (POLG), stimulates the polymerase and exonuclease activities in the replication process of mitochondrial DNA.[16,17] Although the association between *POLG2* polymorphisms and bladder cancer has not yet been established, its association with the risk of head and neck cancer was demonstrated in the French population.[18] There is no strong biological evidence to support the functional importance of *POLG2* to the invasiveness of UBC. However, the increase of transcription level with the risk allele of rs17650301 may be a reason for the association. Because there are no other known SNPs in strong LD ($r^2\geqslant0.8$) with rs17650301 in the *POLG2* locus, rs17650301 is a good candidate marker for the invasiveness of UBC in Japanese males.

The strong association of rs17650301 with the UBC invasiveness was observed only in males. As such, the reasons for the observed higher UBC incidence in males than females remain uncertain. Different non-genetic risk factors including sex hormones, life style and environment may contribute to the disease onset, resulting in the sex-specific association. For instance, involvement of androgen and its receptor in bladder cancer is demonstrated in a mouse study.[19] Epidemiologically, postmenopausal women have a higher risk of development and progression of UBC than premenopausal women.[20] Also, smoking is one of the major risks of urinary tract cancer and considered to increase the cancer risk by approximately threefold,[21,22] and the percentage of smokers is much higher in males than in females in Japan (36.8 versus 9.1%, as of 2008 in the National Survey by the Ministry of Health, Labour and Welfare of Japan). If such environmental and life style-related factors contribute to the predisposition of bladder cancer together with *POLG2*, this may explain the observed effect of *POLG2* with UBC invasiveness in males only. However, further functional characterization of *POLG2* for its involvement in carcinogenesis of indolent and aggressive tumors are required for the elucidation of the molecular mechanism underlying the prognosis of UBC.

ACKNOWLEDGEMENTS

We would like to express our gratitude to all the study participants who made this research possible. We also thank M Kokubo and M Mizutani for technical assistance, and A Yoshizumi and H Uneme for informatics management. This study was supported, in part, by Core Research of Evolutional Science & Technology program (CREST) from Japan Science and Technology Agency (JST) and by Institut National de la Santé et de la Racherche Médicale.

1 Greene, F. L., Compton, C. C., Fritz, A. G., Shah, J. & Winchester, D. P. *AJCC Cancer Staging Atlas* (Springer, New York, 2006).
2 Meiden, A.v.d. in *Invasive Bladder Cancer* (eds Bassi, P. & Pagano, F.) 67–76 (Springer, London, 2007).
3 Gofrit, O. N., Pode, D., Lazar, A., Katz, R. & Shapiro, A. Watchful waiting policy in recurrent Ta G1 bladder tumors. *Eur. Urol.* **49,** 303–306 discussion 306-307 (2006).
4 McNally, M. L., Rashid, H. H. & Messing, E. M. in *Invasive Bladder Cancer* (eds Bassi, P. & Pagano, F.) 15–38 (Springer, London, 2007).
5 Taylor, J. A., Umbach, D. M., Stephens, E., Castranio, T., Paulson, D., Robertson, C. *et al.* The role of N-acetylation polymorphisms in smoking-associated bladder cancer: evidence of a gene-gene-exposure three-way interaction. *Cancer Res.* **58,** 3603–3610 (1998).
6 Sanyal, S., Festa, F., Sakano, S., Zhang, Z., Steineck, G., Norming, U. *et al.* Polymorphisms in DNA repair and metabolic genes in bladder cancer. *Carcinogenesis* **25,** 729–734 (2004).
7 Garcia-Closas, M., Malats, N., Silverman, D., Dosemeci, M., Kogevinas, M., Hein, D. W. *et al.* NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* **366,** 649–659 (2005).
8 Figueroa, J. D., Malats, N., Rothman, N., Real, F. X., Silverman, D., Kogevinas, M. *et al.* Evaluation of genetic variation in the double-strand break repair pathway and bladder cancer risk. *Carcinogenesis* **28,** 1788–1793 (2007).
9 Danoy, P., Sonoda, E., Lathrop, M., Takeda, S. & Matsuda, F. A naturally occurring genetic variant of human XRCC2 (R188H) confers increased resistance to cisplatin-induced DNA damage. *Biochem. Biophys. Res. Commun.* **352,** 763–768 (2007).
10 Takahashi, M., Matsuda, F., Margetic, N. & Lathrop, M. Automated identification of single nucleotide polymorphisms from sequencing data. *J. Bioinform. Comput. Biol.* **1,** 253–265 (2003).
11 The International HapMap Consortium. A haplotype map of the human genome. *Nature* **437,** 1299–1320 (2005).
12 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
13 Breslowk, N. E. & Day, N. E. Statistical methods in cancer research. Volume I—The analysis of case-control studies. *IARC Sci. Publ.* 5–338 (1980).
14 Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315,** 848–853 (2007).
15 Portales-Casamar, E., Thongjuea, S., Kwon, A. T., Arenillas, D., Zhao, X., Valen, E. *et al.* JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **38,** D105–D110 (2010).
16 Lim, S. E., Longley, M. J. & Copeland, W. C. The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. *J. Biol. Chem.* **274,** 38197–38203 (1999).
17 Yakubovskaya, E., Chen, Z., Carrodeguas, J. A., Kisker, C. & Bogenhagen, D. F. Functional human mitochondrial DNA polymerase gamma forms a heterotrimer. *J. Biol. Chem.* **281,** 374–382 (2006).
18 Michiels, S., Danoy, P., Dessen, P., Bera, A., Boulet, T., Bouchardy, C. *et al.* Polymorphism discovery in 62 DNA repair genes and haplotype associations with risks for lung and head and neck cancers. *Carcinogenesis* **28,** 1731–1739 (2007).
19 Miyamoto, H., Yang, Z., Chen, Y. T., Ishiguro, H., Uemura, H., Kubota, Y. *et al.* Promotion of bladder cancer development and progression by androgen receptor signals. *J. Natl. Cancer Inst.* **99,** 558–568 (2007).
20 McGrath, M., Michaud, D. S. & De Vivo, I. Hormonal and reproductive factors and the risk of bladder cancer in women. *Am. J. Epidemiol.* **163,** 236–244 (2006).
21 Zeegers, M. P., Tan, F. E., Dorant, E. & van Den Brandt, P. A. The impact of characteristics of cigarette smoking on urinary tract cancer risk: a meta-analysis of epidemiologic studies. *Cancer* **89,** 630–639 (2000).
22 Puente, D., Hartge, P., Greiser, E., Cantor, K. P., King, W. D., Gonzalez, C. A. *et al.* A pooled analysis of bladder cancer case-control studies evaluating smoking in men and women. *Cancer Causes Control* **17,** 71–79 (2006).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

SHORT REPORT

The *FOXE1* and *NKX2-1* loci are associated with susceptibility to papillary thyroid carcinoma in the Japanese population

Michiko Matsuse,[1] Meiko Takahashi,[2] Norisato Mitsutake,[1,3] Eijun Nishihara,[4] Mitsuyoshi Hirokawa,[4] Takahisa Kawaguchi,[2] Tatiana Rogounovitch,[1] Vladimir Saenko,[5] Andrey Bychkov,[1] Keiji Suzuki,[1] Keitaro Matsuo,[6] Kazuo Tajima,[6] Akira Miyauchi,[4] Ryo Yamada,[2] Fumihiko Matsuda,[2,7] Shunichi Yamashita[1,5]

[1]Department of Molecular Medicine, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[2]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan
[3]Nagasaki University Research Centre for Genomic Instability and Carcinogenesis (NRGIC), Nagasaki, Japan
[4]Kuma Hospital, Kobe, Japan
[5]Department of International Health and Radiation Research, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[6]Aichi Cancer Center Hospital and Research Institute, Nagoya, Japan
[7]Institut National de la Santé et de la Recherche Médicale (INSERM) Unit U852, Paris, France

Correspondence to
Dr Norisato Mitsutake, Department of Molecular Medicine, Atomic Bomb Disease Institute, Nagasaki University Graduate School of Biomedical Sciences, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan; mitsu@nagasaki-u.ac.jp

MM, MT and NM are the authors who contributed equally to this work.

Received 16 March 2011
Revised 17 May 2011
Accepted 16 June 2011
Published Online First 5 July 2011

ABSTRACT

Background *FOXE1* and *NKX2-1* are two known genetic risk factors for the predisposition to sporadic papillary thyroid carcinoma (PTC) in Europeans, but their association in other ethnicities is still unknown.

Objective We aim to examine the association of the two genes with Japanese sporadic PTC, which exhibits high $BRAF^{V600E}$ mutation rate.

Methods 507 Japanese sporadic PTC cases and 2766 controls were genotyped for rs965513 (*FOXE1*) and rs944289 (*NKX2-1*). PTC cases were also examined for their $BRAF^{V600E}$ mutational status.

Results The association of both rs965513 ($p=1.27\times10^{-4}$, OR=1.69, 95% CI 1.29 to 2.21) and rs944289 (p=0.0121, OR=1.21, 95% CI 1.04 to 1.39) with the risk of sporadic PTC was confirmed. Subgroup analysis based on the *BRAF* mutational status showed strong association of rs965513 with $BRAF^{V600E}$-positive cases ($p=2.26\times10^{-4}$, OR=1.72, 95% CI 1.29 to 2.29), but not with $BRAF^{V600E}$-negative cases (p=0.143, OR=1.52, 95% CI 0.87 to 2.65). However, there was no difference in the observed effect size between both subgroups. For rs944289, both subgroups showed marginal association (p=0.0585, OR=1.17, 95% CI 0.99 to 1.37 for $BRAF^{V600E}$-positive cases; p=0.0492, OR=1.35, 95% CI 1.00 to 1.81 for $BRAF^{V600E}$-negative cases).

Conclusions Both *FOXE1* and *NKX2-1* were associated with the increased risk of sporadic Japanese PTC. No clear associations were observed for either SNP with $BRAF^{V600E}$ status.

INTRODUCTION

Papillary thyroid carcinoma (PTC) is the most common malignant tumour in endocrine organs, and its only well-known environmental risk factor is exposure to ionising radiation. Genetic polymorphisms have been shown to contribute to individual susceptibility to PTC.[1] The identification and further assessment of the relevant genetic variations are important for understanding the potential mechanisms involved in thyroid carcinogenesis.

Recently, two genome-wide association (GWA) studies on thyroid cancer have been achieved. The first study dealt with sporadic thyroid cancer in Icelandic population using 192 and 37 196 cases and controls, respectively, followed by a replication study in individuals of European descent.[2] The *FOXE1* (*TTF2*) gene on 9q22.33 ($p=1.7\times10^{-27}$, odds ratio (OR)=1.75, 95% CI 1.59 to 1.94 for rs965513) and *NKX2-1* (*TTF1*) on 14q13.3 ($p=2.0\times10^{-9}$, OR=1.37, 95% CI 1.24 to 1.52 for rs944289) showed the strongest association signals. Both genes encode thyroid-specific transcription factors and appear to contribute to an increased risk of both PTC and follicular thyroid carcinoma. The second GWA study focused on radiation-related PTC using 667 young patients exposed to radioiodine fallouts during childhood and 1275 age-matched control subjects residing in the radio-contaminated regions of Belarus at the time of the Chernobyl accident.[3] In this study, the *FOXE1* gene showed strong association with radiation-related PTC ($p=4.8\times10^{-12}$, OR=1.65, 95% CI 1.43 to 1.91 for rs965513), whereas no association was found with *NKX2-1* (p=0.17, OR=1.13, 95% CI 0.95 to 1.36 for rs944289). These results demonstrate that *FOXE1* is a major genetic determinant of predisposition to thyroid carcinoma regardless of aetiology and age. In contrast, *NKX2-1* may be associated only with adult sporadic PTC. Furthermore, the involvement of these genes in thyroid carcinogenesis is yet to be examined in non-European populations.

Thyroid carcinoma in the Japanese population shows distinct characteristics: higher incidence of PTC than in European populations, most PTCs are low-risk tumours with classic papillary morphology, and higher $BRAF^{V600E}$-positive rate (~80%) than in European populations (~50%).[4–8] Similar findings have also been observed in the Korean population.[9] These characteristics are presumably due to high iodine intake, since distribution of thyroid carcinoma type seems to be related to the intake of iodine: more aggressive follicular and anaplastic carcinomas in iodine-deficient areas and more papillary carcinomas in iodine-rich areas.[10 11] A recent study demonstrated significant association between the prevalence of $BRAF^{V600E}$ mutation and high iodine intake.[12] The $BRAF^{V600E}$ mutation is the most prevalent genetic alteration in adult sporadic PTC and is related to aggressive clinicopathological characteristics including extrathyroidal invasion, lymph node

metastasis, advanced tumour stages and poorer prognosis.[8] This mutation is also observed in micropapillary carcinomas and may thus be an early event in thyroid carcinogenesis.

In this study, we aimed to examine the association of the formerly identified genetic loci, namely, *FOXE1* and *NKX2-1*, with PTC in a Japanese case-control series. In addition, we have evaluated if the two genes are associated with the presence of the $BRAF^{V600E}$ mutation in PTC.

METHODS

Subjects

A total of 509 patients with sporadic PTC (mean (SD) age 51.3 (16.0) years, range 13–87 years; 84.4% women) were recruited from Kuma Hospital (Kobe, Japan). Histological diagnosis was performed by a thyroid pathologist (MH). Two thousand seven hundred and sixty-six Japanese individuals were collected as population controls at Kyoto University. All patients and controls have no history of radiation exposure. The protocol was approved by the ethics committees of Nagasaki University, Kuma Hospital and Kyoto University.

DNA extraction and *BRAF* status screening

DNA from PTC subjects was extracted from formalin-fixed paraffin-embedded primary tumour tissues using QIAamp DNA mini kit (QIAGEN, Tokyo, Japan) according to the manufacturer's protocol. DNAs with sufficient quality for genotyping and sequencing were obtained from 507 of 509 PTC specimens. *BRAF* status was screened by direct DNA sequencing. Primer sequences used for PCR and sequencing were BRAFi14F, 5′-ACATACTTATTGACTCTAAGAGGAAAGATGAA-3′, and BRAFi15R, 5′-GATTTTTGTGAATACTGGGAACTATGA-3′. PCR products were treated with ExoSAP-IT PCR clean-up reagent (GE Healthcare Japan, Tokyo, Japan), and sequencing was performed with Big Dye Terminator sequencing kit version 3.1 (Life Technologies, Foster City, California, USA) on an ABI3100 automated sequencer (Life Technologies). We prepared five negative controls (without tissue section) per 96 samples to ensure contamination-free amplifications.

Genotyping

Genotyping was performed for PTC cases using the ABI TaqMan SNP assays (Life Technologies) in accordance with the manufacturer's guidelines. A predesigned and functionally tested probe for rs965513 (C_1593670_20) and rs944289 (C_1444137_10) were used. The conditions were denaturation at 95°C for 10 min followed by 50 cycles of 92°C for 15 s and 60°C for 1 min for both probes.

Table 1 Specification of the DNA collections recruited for the study

Sample	Number	Mean age (SD)	Age range	% Male
Cases	507	51.3 (16.0)	13–87	15.6
$BRAF^{V600E}$ (+)	388	52.4 (15.5)	17–87	16.1
$BRAF^{V600E}$ (−)	104	48.1 (17.2)	13–81	13.6
Not available	15	–	–	–
Controls	2766	50.1 (15.4)	20–79	36.9

As control subjects, genotypes of rs965513 and rs944289 were extracted from the genome scan results using Illumina Human610-Quad BeadChip of 2766 healthy Japanese individuals (mean (SD) age 50.1 (15.4) years, range 20–79 years; 63.1% women).

Statistical analysis

A case-control association in each study was examined using trend exact test to compare genotypic distributions between cases and controls.[13] A subtype analysis was performed based on the *BRAF* mutational status (*BRAF* mutant versus control, and *BRAF* wild-type versus control), and the homogeneity of the ORs between the two studies was examined with the Breslow–Day test.[14] For each case-control study, p values and ORs adjusted for age and sex were calculated using multiple logistic regression analyses. Haplotype frequency was calculated using haplo.stats R package, and linkage disequilibrium (LD) between SNPs was estimated with gap package.[15]

RESULTS

DNA samples extracted from 507 PTC tissues were genotyped by TaqMan assay for rs965513 and rs944289 (table 1). Genotyping success rates for cases were 479/507 (94.5%) and 467/507 (92.1%) for rs965513 and rs944289, respectively. As for controls, the genotyping results that could be extracted from the genome scan data were 2764/2766 (99.9%) and 2766/2766 (100%) for rs965513 and rs944289, respectively. The genotype distributions of the two SNPs conformed to Hardy–Weinberg equilibrium both in cases and in controls. A case-control association was examined using trend exact test to compare the genotypic distributions. Significant associations were obtained for both SNPs (rs965513: $p=1.27\times10^{-4}$, OR=1.69, 95% CI 1.29 to 2.21; rs944289: p=0.0121, OR=1.21, 95% CI 1.04 to 1.39) (table 2). Our results confirmed the previously reported risk alleles, namely, allele A for rs965513 and allele T for rs944289.

We next screened for *BRAF* mutation in PTC tissues by DNA sequencing. Of 507 samples, 492 (97.0%) were successfully

Table 2 Association results for rs965513 and rs944289 in Japanese PTC

SNP (ref/var)*	Study group	Genotyped samples		Allele frequency		HWE exact p value‡		p Value§	OR (95% CI)¶	Power**
		Case†	Cont	Case	Cont	Case	Cont			
rs965513 (A*/G)	All cases	479	2764	0.090	0.057	0.255	0.721	1.27×10^{-4}	1.69 (1.29 to 2.21)	0.76
	$BRAF^{V600E}$ (+)	381	2764	0.092	0.057	0.113	0.721	2.26×10^{-4}	1.72 (1.29 to 2.29)	0.69
	$BRAF^{V600E}$ (−)	95	2764	0.079	0.057	1.000	0.721	0.143	1.52 (0.87 to 2.65)	0.31
rs944289 (C/T*)	All cases	467	2766	0.466	0.411	0.306	0.695	0.0121	1.21 (1.04 to 1.39)	0.60
	$BRAF^{V600E}$ (+)	373	2766	0.458	0.411	0.118	0.695	0.0585	1.17 (0.99 to 1.37)	0.52
	$BRAF^{V600E}$ (−)	93	2766	0.489	0.411	0.411	0.695	0.0492	1.35 (1.00 to 1.81)	0.19

*The reference (ref) and variant (var) alleles refer to NCBI Build 36.3, and the risk allele is indicated by an asterisk.
†*BRAF* mutational statuses were not available in three samples of 479 cases successfully typed for rs965513 and in one sample of 467 cases successfully typed for rs944289.
‡The exact p values for Hardy–Weinberg equilibrium (HWE) are shown.
§The p values using trend exact test adjusted for age and sex are shown.
¶ORs are calculated for the risk allele with a 95% CI.
**Statistical power is calculated using power function in Hmisc package (http://cran.r-project.org/web/packages/Hmisc/index.html) of R.
PTC, papillary thyroid carcinoma.

genotyped, of which 388 were found to carry a heterozygous $BRAF^{V600E}$ mutation and 104 were negative for the mutation (table 1). Subgroup analysis based on the *BRAF* mutational status showed a strong association for rs965513 between 381 $BRAF^{V600E}$-positive cases and 2766 controls ($p=2.26\times10^{-4}$, OR=1.72, 95% CI 1.29 to 2.29) (table 2). On the other hand, no statistically significant association was found between 95 $BRAF^{V600E}$-negative cases and controls (p=0.143, OR=1.52, 95% CI 0.87 to 2.65), although there was no reversal of the risk allele, and a similar trend was observed (table 2). There was no difference in the observed effect size between the $BRAF^{V600E}$-positive and $BRAF^{V600E}$-negative groups (p=0.615).

For rs944289, both analyses showed marginal association (p=0.0585, OR=1.17, 95% CI 0.99 to 1.37 for $BRAF^{V600E}$-positive cases; p=0.0492, OR=1.35, 95% CI 1.00 to 1.81 for $BRAF^{V600E}$-negative cases) (table 2). Again, there was no reversal of the risk allele, and there was no difference in the observed effect size between the $BRAF^{V600E}$-positive and $BRAF^{V600E}$-negative groups (p=0.455).

DISCUSSION

Here, we report for the first time an association between the *FOXE1* gene and PTC in the Japanese population by genotyping rs965513 located 57 kb upstream to the *FOXE1* gene. The association of the *FOXE1* gene has been previously demonstrated by a GWA study for sporadic PTC[2] as well as for radiation-related PTC[3] in European populations. Furthermore, a recent study using an SNP panel of 97 genes related to thyroid cell differentiation and proliferation identified rs1867277, a causal SNP within the *FOXE1* 5′ UTR, functioning as a genetic risk factor associated with susceptibility to PTC.[16] The sequence containing the risk allele was demonstrated to recruit the USF1/USF2 transcription factors that in turn increased *FOXE1* transcriptional activity. Indeed, animal model experiments have shown that mice lacking the *FOXE1* locus exhibit neonatal hypothyroidism that shows similarity to thyroid dysgenesis in humans.[17] We additionally genotyped rs1867277 in 64 randomly selected cases in our series to estimate LD with rs965513. There was no strong evidence for LD between these two SNPs (D′=0.23), suggesting that the functional significance of rs965513 may be different from that of rs1867277.

The association between *NKX2-1* at chromosome 14q13.3 and sporadic PTC was also successfully reproduced in the Japanese population. However, the association was weaker for *NKX2-1* compared to *FOXE1* in our study in concordance with the results of the Icelandic study. Although rs944289 lies in a 249 kb LD region with no known genes, transcription units or predicted exons, *NKX2-1* is one of the closest genes residing in this region. NKX2-1 is another thyroid-specific transcription factor, which together with *FOXE1*, is expressed from early stages of thyroid morphogenesis and plays a major role in the development of the thyroid gland. Knockout mice lacking the *NKX2-1* gene die at birth because they lack normal thyroid and lungs, demonstrating the essential role of the gene in embryonic differentiation of these organs.[18] Interestingly, rs944289 was strongly associated with sporadic PTC in the Icelandic population[2] as well as in our Japanese series, but not in the Belarusian radiation-related PTC, suggesting that this variant may be associated only with sporadic PTC.

The strong association of rs965513 with Japanese PTC was also found between $BRAF^{V600E}$-positive cases and controls ($p=2.26\times10^{-4}$, OR=1.72, 95% CI 1.29 to 2.29) but not between $BRAF^{V600E}$-negative cases and controls (p=0.143, OR=1.52, 95% CI 0.87 to 2.65). However, the effect size is similar between the two groups (p=0.615), and the statistical power is relatively low (0.31) in the latter analysis, suggesting that the lack of significance is due to the lower minor allele frequency of the SNP in the Japanese population (0.090 in cases, 0.057 in controls) than in Europeans (0.462–0.490 in cases, 0.352–0.367 in controls)[2 3] and/or the much smaller number of $BRAF^{V600E}$-negative cases ($n=95$). On the other hand, a marginal association was observed for rs944289 in both subgroups (p=0.0585, OR=1.17, 95% CI 0.99 to 1.37 for $BRAF^{V600E}$-positive subgroup; p=0.0492, OR=1.35, 95% CI 1.00 to 1.81 for $BRAF^{V600E}$-negative subgroup). Again, the effect size is similar between the two groups (p=0.455). However, for both SNPs, the number of $BRAF^{V600E}$-negative cases needs to be increased to draw significant conclusions in the subtype analyses, especially for rs965513 for which the minor allele frequency in the Japanese population is so much lower compared to Europeans.

In conclusion, our study successfully confirms the association of both rs965513 and rs944289 with sporadic PTC in the Japanese population. Conceivably, *FOXE1* is likely to be the most important genetic determinant of susceptibility to PTC regardless of ethnicity. There was no clear difference in genetic impact for either of the SNPs with $BRAF^{V600E}$ status.

Funding This work was supported in part by Grant-in-Aid for Scientific Research (#22256004 and #22390189), Grant-in-Aid for Young Scientists (#22791204 and #21790337) and Global COE Program from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. FM was in part supported by INSERM. NM was in part supported by The Yasuda Medical Foundation.

Competing interests None.

Ethics approval This study was conducted with the approval of the Nagasaki University, Kuma Hospital and Kyoto University.

Contributors Conceived and designed the experiments: MM, MT, NM, FM, SY; performed experiments: MM, MT, AB; analyzed the data: MT, NM, TK, VS, RY, FM; contributed materials/analysis tools: EN, MH, TR, KS, KM, KT, AM; wrote the paper: MM, MT, NM, VS, FM.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Adjadj E,** Schlumberger M, de Vathaire F. Germ-line DNA polymorphisms and susceptibility to differentiated thyroid cancer. *Lancet Oncol* 2009;**10**:181–90.
2. **Gudmundsson J,** Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT, He H, Blondal T, Geller F, Jakobsdottir M, Magnusdottir DN, Matthiasdottir S, Stacey SN, Skarphedinsson OB, Helgadottir H, Li W, Nagy R, Aguillo E, Faure E, Prats E, Saez B, Martinez M, Eyjolfsson GI, Bjornsdottir US, Holm H, Kristjansson K, Frigge ML, Kristvinsson H, Gulcher JR, Jonsson T, Rafnar T, Hjartarsson H, Mayordomo JI, de la Chapelle A, Hrafnkelsson J, Thorsteinsdottir U, Kong A, Stefansson K. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet* 2009;**41**:460–4.
3. **Takahashi M,** Saenko VA, Rogounovitch TI, Kawaguchi T, Drozd VM, Takigawa-Imamura H, Akulevich NM, Ratanajaraya C, Mitsutake N, Takamura N, Danilova LI, Lushchik ML, Demidchik YE, Heath S, Yamada R, Lathrop M, Matsuda F, Yamashita S. The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. *Hum Mol Genet* 2010;**19**:2516–23.
4. **Kumagai A,** Namba H, Akanov Z, Saenko VA, Meirmanov S, Ohtsuru A, Yano H, Maeda S, Anami M, Hayashi T, Ito M, Sagandikova S, Eleubaeva Z, Mussinov D, Espenbetova M, Yamashita S. Clinical implications of pre-operative rapid BRAF analysis for papillary thyroid cancer. *Endocr J* 2007;**54**:399–405.
5. **Nakayama H,** Yoshida A, Nakamura Y, Hayashi H, Miyagi Y, Wada N, Rino Y, Masuda M, Imada T. Clinical significance of BRAF (V600E) mutation and Ki-67 labeling index in papillary thyroid carcinomas. *Anticancer Res* 2007;**27**:3645–9.
6. **Watanabe R,** Hayashi Y, Sassa M, Kikumori T, Imai T, Kiuchi T, Murata Y. Possible involvement of BRAFV600E in altered gene expression in papillary thyroid cancer. *Endocr J* 2009;**56**:407–14.
7. **Zuo H,** Nakamura Y, Yasuoka H, Zhang P, Nakamura M, Mori I, Miyauchi A, Kakudo K. Lack of association between BRAF V600E mutation and mitogen-activated protein kinase activation in papillary thyroid carcinoma. *Pathol Int* 2007;**57**:12–20.
8. **Xing M.** BRAF mutation in papillary thyroid cancer: pathogenic role, molecular bases, and clinical implications. *Endocr Rev* 2007;**28**:742–62.
9. **Kim SW,** Lee JI, Kim JW, Ki CS, Oh YL, Choi YL, Shin JH, Kim HK, Jang HW, Chung JH. BRAFV600E mutation analysis in fine-needle aspiration cytology specimens for evaluation of thyroid nodule: a large series in a BRAFV600E-prevalent population. *J Clin Endocrinol Metab* 2010;**95**:3693–700.

10. **Feldt-Rasmussen U.** Iodine and cancer. *Thyroid* 2001;**11**:483–6.
11. **Williams ED,** Abrosimov A, Bogdanova T, Demidchik EP, Ito M, LiVolsi V, Lushnikov E, Rosai J, Tronko MD, Tsyb AF, Vowler SL, Thomas GA. Morphologic characteristics of Chernobyl-related childhood papillary thyroid carcinomas are independent of radiation exposure but vary with iodine intake. *Thyroid* 2008;**18**:847–52.
12. **Guan H,** Ji M, Bao R, Yu H, Wang Y, Hou P, Zhang Y, Shan Z, Teng W, Xing M. Association of high iodine intake with the T1799A BRAF mutation in papillary thyroid cancer. *J Clin Endocrinol Metab* 2009;**94**:1612–17.
13. **Yamada R,** Okada Y. An optimal dose-effect mode trend test for SNP genotype tables. *Genet Epidemiol* 2009;**33**:114–27.
14. **Breslow NE,** Day NE. Statistical methods in cancer research. Volume I—The analysis of case-control studies. *IARC Sci Publ* 1980;**32**:5–338.
15. **Zhao H,** Nettleton D, Dekkers JC. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. *Genet Res* 2007;**89**:1–6.
16. **Landa I,** Ruiz-Llorente S, Montero-Conde C, Inglada-Perez L, Schiavi F, Leskela S, Pita G, Milne R, Maravall J, Ramos I, Andia V, Rodriguez-Poyo P, Jara-Albarran A, Meoro A, del Peso C, Arribas L, Iglesias P, Caballero J, Serrano J, Pico A, Pomares F, Gimenez G, Lopez-Mondejar P, Castello R, Merante-Boschin I, Pelizzo MR, Mauricio D, Opocher G, Rodriguez-Antona C, Gonzalez-Neira A, Matias-Guiu X, Santisteban P, Robledo M. The variant rs1867277 in FOXE1 gene confers thyroid cancer susceptibility through the recruitment of USF1/USF2 transcription factors. *PLoS Genet* 2009;**5**:e1000637.
17. **De Felice M,** Ovitt C, Biffali E, Rodriguez-Mallon A, Arra C, Anastassiadis K, Macchia PE, Mattei MG, Mariano A, Scholer H, Macchia V, Di Lauro R. A mouse model for hereditary thyroid dysgenesis and cleft palate. *Nat Genet* 1998;**19**:395–8.
18. **Kimura S,** Hara Y, Pineau T, Fernandez-Salguero P, Fox CH, Ward JM, Gonzalez FJ. The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev* 1996;**10**:60–9.

The *FOXE1* and *NKX2-1* loci are associated with susceptibility to papillary thyroid carcinoma in the Japanese population

Michiko Matsuse, Meiko Takahashi, Norisato Mitsutake, et al.

J Med Genet 2011 48: 645-648 originally published online July 5, 2011
doi: 10.1136/jmedgenet-2011-100063

ARD Online First, published on January 10, 2012 as 10.1136/annrheumdis-2011-200907

Quantitative effect of HLA-DRB1 alleles to ACPA levels in Japanese rheumatoid arthritis: no strong genetic impact of shared epitope to ACPA levels after stratification of HLA-DRB1*09:01

Anti-citrullinated peptide antibody (ACPA) is a highly specific serological marker for rheumatoid arthritis (RA).[1–3] Different HLA-DRB1 alleles have been shown to be associated with the susceptibility to ACPA-positive RA.[4 5] Former studies demonstrated that HLA-DRB alleles carrying a shared epitope (SE),[6] consisting of a conserved amino acid motif at positions 70–74 of the HLA-DRβ chain, were strongly associated with ACPA-positive RA and with higher ACPA levels in European and Japanese populations.[7–9] On the other hand, HLA-DRB1*09:01 was recently found to be negatively associated with ACPA levels in the Japanese.[9] These observations imply that combinations of HLA-DRB1 alleles differentially influence ACPA levels in ACPA-positive RA.

To address this question, we conducted a genetic association study employing 2457 ACPA-positive Japanese RA patients. ACPA was quantified by MESACUP CCP ELISA kit (MBL Co Ltd, Nagoya, Japan) with a cut-off level of 4.5 U/ml. The patients were then divided into three groups based on their ACPA titres: low (~4.5–13.5 U/ml), intermediate (~13.5–100 U/ml) and high (≥100 U/ml) groups. These groups were defined according to the 2010 ACR/EULAR classification criteria for RA and a measurement limit of the kit. HLA-DRB1 genotyping was carried out using either the Wakflow system (Wakunaga Pharmaceutical Co Ltd, Osaka, Japan) or the sequencing-based AlleleSEQR HLA-DRB1 typing kit (Abbott Japan, Nagoya, Japan). Frequencies of HLA-DRB1 alleles were compared among the three groups using the Cochran-Armitage Trend test. The relative predispositional effect (RPE) method was applied to identify the associations of more than one HLA-DBR1 allele sequentially according to their strength.[10] Briefly, associations of HLA-DRB1 alleles with ACPA categories were estimated for each allele using the Cochran-Armitage Trend test. When we detected the strongest association with a significant p value, the allele was excluded from the whole data and the same steps were repeated until no further significant alleles were found.

As expected from the previous studies,[9] HLA-DRB1*09:01 showed the strongest association with ACPA levels in a decreasing manner ($p=1.0\times10^{-21}$) and the SE alleles were significantly associated with an increasing effect ($p=3.2\times10^{-7}$) (table 1). In addition, HLA-DRB1*04:07 showed negative association with ACPA levels (p=0.0013), and HLA-DRB1*15:01 and HLA-DRB1*15:02 were positively associated with ACPA levels ($p=2.3\times10^{-5}$ and 0.0011, respectively) (table 1). Of note, the association between the SE and ACPA levels lost significance after

Table 1 Association of HLA-DRB1 alleles with ACPA levels

HLA-DRB1	Low n=594	Intermediate n=1510	High n=2810	p Value	RPE p Value	RPE (OR)	Effect on ACPA levels
SE							
SEall	216 (36.4%)	616 (40.8%)	1303 (46.4%)	3.2×10^{-7}	0.16†	1.08 (0.98–1.20)†	
DRB1*01:01	32 (5.4%)	96 (6.4%)	223 (7.9%)	0.0096			
DRB1*04:01	18 (3.0%)	47 (3.1%)	82 (2.9%)	0.78			
DRB1*04:04	2 (0.3%)	1 (0.1%)	14 (0.5%)	0.13			
DRB1*04:05	138 (23.2%)	409 (27.1%)	840 (29.9%)	0.00053			
DRB1*04:10	17 (2.9%)	33 (2.2%)	67 (2.4%)	0.71			
DRB1*10:01	6 (1.0%)	13 (0.9%)	28 (1.0%)	0.87			
DRB1*14:06	3 (0.5%)	14 (0.9%)	44 (1.6%)	0.013			
Non-SE							
DRB1*04:03	12 (2.0%)	30 (2.0%)	31 (1.1%)	0.019			
DRB1*04:06	17 (2.9%)	14 (0.9%)	57 (2.0%)	0.96			
DRB1*04:07	5 (0.8%)	11 (0.7%)	4 (0.1%)	0.0013	0.00034	0.30 (0.16–0.57)	(−)
DRB1*08:02	15 (2.5%)	30 (2.0%)	60 (2.1%)	0.74			
DRB1*08:03	36 (6.1%)	66 (4.4%)	119 (4.2%)	0.10			
DRB1*09:01	158 (26.6%)	334 (22.1%)	367 (13.1%)	1.0×10^{-21}	1.0×10^{-21}	0.56 (0.50–0.62)	(−)
DRB1*11:01	8 (1.3%)	27 (1.8%)	50 (1.8%)	0.57			
DRB1*12:01	14 (2.4%)	30 (2.0%)	68 (2.4%)	0.63			
DRB1*12:02	8 (1.3%)	26 (1.7%)	50 (1.8%)	0.52			
DRB1*13:02	22 (3.7%)	53 (3.5%)	102 (3.6%)	0.98			
DRB1*14:01	4 (0.7%)	32 (2.1%)	32 (1.1%)	0.64			
DRB1*14:03	6 (1.0%)	17 (1.1%)	37 (1.3%)	0.46			
DRB1*14:05	5 (0.8%)	19 (1.3%)	21 (0.7%)	0.36			
DRB1*15:01	20 (3.4%)	53 (3.5%)	180 (6.4%)	2.3×10^{-5}	0.0011	1.53 (1.21–1.92)	(+)
DRB1*15:02	36 (6.1%)	120 (7.9%)	276 (9.8%)	0.0011			
DRB1*16:02	4 (0.7%)	20 (1.3%)	29 (1.0%)	0.83			

HLA-DRB1 alleles with frequencies greater than 0.5% are shown. Significant levels were set as 0.0022 for HLA-DRB1 alleles after Bonferroni's correction for multiple testing.
†p Value and OR after removal of HLA-DRB1*09:01.
ACPA, anti-citrullinated peptide antibody; RPE, relative predispositional effect; SE, shared epitope.

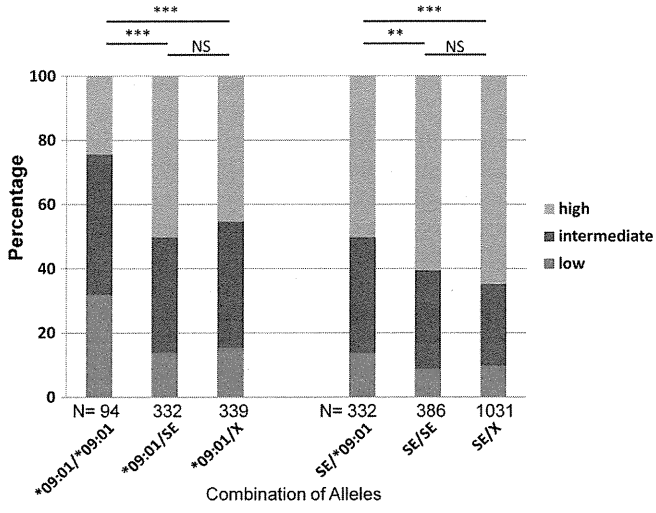

Figure 1 Comparisons of blood anti-citrullinated peptide antibody (ACPA) levels among HLA-DRB1*09:01, shared epitope (SE) and other alleles in combination. Frequencies of three rheumatoid arthritis subgroups based on ACPA levels were compared among different HLA-DRB1 combinations containing HLA-DRB1*09:01 and/or SE. X indicates HLA-DRB1 alleles other than HLA-DRB1*09:01 and SE. 'Low', 'intermediate' and 'high' categories correspond to patients with ACPA titres of ~4.5–13.5, ~13.5–100 and ≥100 U/ml, respectively. **p<0.005 and ***p<0.00005. NS, not significant.

stratification of HLA-DRB1*09:01 using RPE (p=0.16) whereas HLA-DRB1*04:07 and HLA-DRB1*15:01 remained significant after RPE (p=0.00034 and p=0.0011, respectively) (table 1). To confirm the dominant effect of HLA-DRB1*09:01 on ACPA levels over SE, we compared ACPA levels in two sets: first between HLA-DRB1*09:01/*09:01 and HLA-DRB1*09:01/SE or HLA-DRB1*09:01/X, and second between SE/HLA-DRB1*09:01 and SE/SE or SE/X. We found that HLA-DRB1*09:01 showed a significant association with low ACPA category compared with the other two groups in both sets of analyses (p<0.005, figure 1). On the other hand, we could not observe any difference between SE and the other alleles.

In this study, we aimed to identify HLA-DRB1 alleles showing quantitative effects on ACPA levels using a large collection of Japanese ACPA-positive RA patients. RPE was applied to avoid misleading frequency deviation by the allele with the strongest association to other associated alleles. We demonstrated that HLA-DRB1*09:01 was the strongest genetic determinant for lower ACPA levels, and the quantitative effects of HLA-DRB1 alleles carrying the SE were not a primary effect but merely an expected consequence of the decreased frequency of HLA-DRB1*09:01. We also identified two novel HLA-DRB1 alleles, HLA-DRB1*04:07 and HLA-DRB1*15:01, being associated with ACPA levels. It is interesting and feasible to perform similar studies in other populations and investigate whether or not the same set of HLA-DRB1 alleles are related to the quantitative effects beyond ethnicities and to examine if such alleles share conserved amino acid motifs.

Chikashi Terao,[1,2] Katsunori Ikari,[3] Koichiro Ohmura,[2] Taku Suzuki,[3] Takuji Iwamoto,[3] Kiyoshi Takasugi,[4] Hiroo Saji,[5] Atsuo Taniguchi,[3] Shigeki Momohara,[3] Hisashi Yamanaka,[3] Fumihiko Matsuda,[1,6,7] Tsuneyo Mimori[2]

[1]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan
[2]Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan
[3]Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan
[4]Dohgo Spa Hospital, Matsuyama, Japan
[5]HLA Laboratory, Kyoto, Japan
[6]CREST Program, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan
[7]Institut National de la Sante et de la Recherche Medicale (INSERM) Unite U852, Kyoto University Graduate School of Medicine, Kyoto, Japan

Correspondence to Chikashi Terao, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan; a0001101@kuhp.kyoto-u.ac.jp

Acknowledgements We would like to thank all the doctors and staff who collected DNA samples and helped us with ACPA quantification and HLA genotyping. This study was performed with the support of Genetics and Allied research in Rheumatic diseases Networking (GARNET) consortium.

Funding This work was supported by Grants-in-aid from the Ministry of Health, Labour and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science and Technology of Japan as well as by research grants from the Japan Rheumatism Foundation, the Waksman Foundation and the Mitsubishi Pharma Research Foundation.

Competing interest None.

Provenance and peer review Not commissioned; externally peer reviewed.

Accepted 11 December 2011
Received 12 October 2011

REFERENCES

1. **van Venrooij WJ**, Hazes JM, Visser H. Anticitrullinated protein/peptide antibody and its role in the diagnosis and prognosis of early rheumatoid arthritis. *Neth J Med* 2002;**60**:383–8.
2. **Schellekens GA**, Visser H, de Jong BA, *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum* 2000;**43**:155–63.
3. **van Jaarsveld CH**, ter Borg EJ, Jacobs JW, *et al.* The prognostic value of the antiperinuclear factor, anti-citrullinated peptide antibodies and rheumatoid factor in early rheumatoid arthritis. *Clin Exp Rheumatol* 1999;**17**:689–97.
4. **van der Woude D**, Lie BA, Lundström E, *et al.* Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations. *Arthritis Rheum* 2010;**62**:1236–45.
5. **Gorman JD**, Lum RF, Chen JJ, *et al.* Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. *Arthritis Rheum* 2004;**50**:400–12.
6. **Gregersen PK**, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* 1987;**30**:1205–13.
7. **Balsa A**, Cabezón A, Orozco G, *et al.* Influence of HLA-DRB1 alleles in the susceptibility of rheumatoid arthritis and the regulation of antibodies against citrullinated proteins and rheumatoid factor. *Arthritis Res Ther* 2010;**12**:R62.
8. **Ohmura K**, Terao C, Maruya E, *et al.* Anti-citrullinated peptide antibody-negative RA is a genetically distinct subset: a definitive study using only bone-erosive ACPA-negative rheumatoid arthritis. *Rheumatology (Oxford)* 2010;**49**:2298–304.
9. **Okada Y**, Suzuki A, Yamada R, *et al.* HLA-DRB1*0901 lowers anti-cyclic citrullinated peptide antibody levels in Japanese patients with rheumatoid arthritis. *Ann Rheum Dis* 2010;**69**:1569–70.
10. **Payami H**, Joe S, Farid NR, *et al.* Relative predispositional effects (RPEs) of marker alleles with disease: HLA-DR alleles and Graves disease. *Am J Hum Genet* 1989;**45**:541–6.

Quantitative effect of HLA-DRB1 alleles to ACPA levels in Japanese rheumatoid arthritis: no strong genetic impact of shared epitope to ACPA levels after stratification of HLA-DRB1*09:01

Chikashi Terao, Katsunori Ikari, Koichiro Ohmura, et al.

Ann Rheum Dis published online January 10, 2012
doi: 10.1136/annrheumdis-2011-200907

A Genome-Wide Association Study Identified *AFF1* as a Susceptibility Locus for Systemic Lupus Eyrthematosus in Japanese

Yukinori Okada[1,2,3☯], Kenichi Shimane[1,2☯], Yuta Kochi[1,2☯*], Tomoko Tahira[4], Akari Suzuki[1], Koichiro Higasa[3], Atsushi Takahashi[3], Tetsuya Horita[5], Tatsuya Atsumi[5], Tomonori Ishii[6], Akiko Okamoto[2], Keishi Fujio[2], Michito Hirakata[7], Hirofumi Amano[8], Yuya Kondo[9], Satoshi Ito[9], Kazuki Takada[10], Akio Mimori[11], Kazuyoshi Saito[12], Makoto Kamachi[13], Yasushi Kawaguchi[14], Katsunori Ikari[14], Osman Wael Mohammed[15], Koichi Matsuda[15], Chikashi Terao[16,17], Koichiro Ohmura[16], Keiko Myouzen[1], Naoya Hosono[18], Tatsuhiko Tsunoda[19], Norihiro Nishimoto[20], Tsuneyo Mimori[16], Fumihiko Matsuda[17], Yoshiya Tanaka[12], Takayuki Sumida[9], Hisashi Yamanaka[14], Yoshinari Takasaki[8], Takao Koike[5], Takahiko Horiuchi[21], Kenshi Hayashi[4], Michiaki Kubo[18], Naoyuki Kamatani[3], Ryo Yamada[1,17], Yusuke Nakamura[15], Kazuhiko Yamamoto[1,2]

1 Laboratory for Autoimmune Diseases, Center for Genomic Medicine (CGM), RIKEN, Yokohama, Japan, **2** Department of Allergy and Rheumatology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, **3** Laboratory for Statistical Analysis, CGM, RIKEN, Yokohama, Japan, **4** Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan, **5** Department of Medicine II, Hokkaido University Graduate School of Medicine, Sapporo, Japan, **6** Department of Hematology and Rheumatology, Tohoku University Graduate School of Medicine, Sendai, Japan, **7** Division of Rheumatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan, **8** Department of Internal Medicine and Rheumatology, Juntendo University School of Medicine, Tokyo, Japan, **9** Division of Clinical Immunology, Doctoral Program in Clinical Sciences, Graduate School of Comprehensive Human Science, University of Tsukuba, Tsukuba, Japan, **10** Departments of Medicine and Rheumatology, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan, **11** Division of Rheumatic Diseases, National Center for Global Health and Medicine, Tokyo, Japan, **12** First Department of Internal Medicine, University of Occupational and Environmental Health, Kitakyushu, Japan, **13** Department of Immunology and Rheumatology, Unit of Translational Medicine, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan, **14** Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan, **15** Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan, **16** Department of Rheumatology and Clinical immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **17** Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **18** Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan, **19** Laboratory for Medical Informatics, CGM, RIKEN, Yokohama, Japan, **20** Laboratory of Immune Regulation, Wakayama Medical University, Wakayama, Japan, **21** Department of Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan

Abstract

Systemic lupus erythematosus (SLE) is an autoimmune disease that causes multiple organ damage. Although recent genome-wide association studies (GWAS) have contributed to discovery of SLE susceptibility genes, few studies has been performed in Asian populations. Here, we report a GWAS for SLE examining 891 SLE cases and 3,384 controls and multi-stage replication studies examining 1,387 SLE cases and 28,564 controls in Japanese subjects. Considering that expression quantitative trait loci (eQTLs) have been implicated in genetic risks for autoimmune diseases, we integrated an eQTL study into the results of the GWAS. We observed enrichments of cis-eQTL positive loci among the known SLE susceptibility loci (30.8%) compared to the genome-wide SNPs (6.9%). In addition, we identified a novel association of a variant in the AF4/FMR2 family, member 1 (*AFF1*) gene at 4q21 with SLE susceptibility (rs340630; $P = 8.3 \times 10^{-9}$, odds ratio = 1.21). The risk A allele of rs340630 demonstrated a cis-eQTL effect on the *AFF1* transcript with enhanced expression levels ($P < 0.05$). As *AFF1* transcripts were prominently expressed in $CD4^+$ and $CD19^+$ peripheral blood lymphocytes, up-regulation of *AFF1* may cause the abnormality in these lymphocytes, leading to disease onset.

Citation: Okada Y, Shimane K, Kochi Y, Tahira T, Suzuki A, et al. (2012) A Genome-Wide Association Study Identified *AFF1* as a Susceptibility Locus for Systemic Lupus Eyrthematosus in Japanese. PLoS Genet 8(1): e1002455. doi:10.1371/journal.pgen.1002455

Editor: Mark I. McCarthy, University of Oxford, United Kingdom

Received May 16, 2011; **Accepted** November 18, 2011; **Published** January 26, 2012

Funding: This study was supported by a grant from the CGM, RIKEN, and a grant from the autoimmune disease study group of Research in Intractable Diseases, Japanese Ministry of Health, Labor, and Welfare, Japan. 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: ykochi@src.riken.jp

☯ These authors contributed equally to this work.

Author Summary

Although recent genome-wide association study (GWAS) approaches have successfully contributed to disease gene discovery, many susceptibility loci are known to be still uncaptured due to strict significance threshold for multiple hypothesis testing. Therefore, prioritization of GWAS results by incorporating additional information is recommended. Systemic lupus erythematosus (SLE) is an autoimmune disease that causes multiple organ damage. Considering that abnormalities in B cell activity play essential roles in SLE, prioritization based on an expression quantitative trait loci (eQTLs) study for B cells would be a promising approach. In this study, we report a GWAS and multi-stage replication studies for SLE examining 2,278 SLE cases and 31,948 controls in Japanese subjects. We integrated eQTL study into the results of the GWAS and identified *AFF1* as a novel SLE susceptibility loci. We also confirmed cis-regulatory effect of the locus on the *AFF1* transcript. Our study would be one of the initial successes for detecting novel genetic locus using the eQTL study, and it should contribute to our understanding of the genetic loci being uncaptured by standard GWAS approaches.

Introduction

Systemic lupus erythematosus (SLE) is an autoimmune disease characterized by autoantibody production, complement activation, and multi-organ damage [1]. Familial aggregation demonstrates that both genetic and environmental factors play a role in pathogenesis of SLE [2]. Genetic studies using candidate gene-approaches, and recently, genome-wide association studies (GWAS), have uncovered more than 25 SLE susceptibility genes, including *HLA-DRB1*, *IRF5*, *STAT4*, *ITGAM*, *BLK*, *TNFAIP3*, and others [3–18]. However, most of these studies were conducted in European populations [3–13,15,17], and few studies have been conducted in Asian populations [14,16,18]. Since the epidemiology of SLE has demonstrated that the prevalence of disease substantially differs among populations, genetic backgrounds of SLE should be also heterogeneous across populations [19,20]. Therefore, additional studies in Asians might provide novel insights. It is of note that GWAS for SLE in Chinese populations identified novel loci that had not been detected in Europeans, such as *ETS1*, *IKZF1*, and *WDFY4* [14,16].

Another issue raised by the previous GWASs for complex diseases is that many susceptibility loci still remained uncaptured, owing to its strict significance threshold for multiple hypothesis testing [21]. In SLE, for example, the 26 risk loci identified by the previous GWAS explained only an estimated 8% of the total genetic susceptibility to the disease [15]. Therefore, it is still important to examine the sub-loci of GWAS, in order to reveal the entire picture of genetic etiology. To effectively explore these uncaptured loci, prioritization of GWAS results by incorporating additional information implicated in the disease pathophysiology is recommended [22,23]. Considering that abnormalities in B cell activity play essential roles in SLE [1] and that expression quantitative trait loci (eQTL) have been implicated to comprise approximately a half of genetic risks for autoimmune diseases [24], prioritization based on an eQTL study for B cells would be a promising approach for SLE [25]. Moreover, an eQTL itself assures the presence of functional variant(s) that regulate gene expression. Thus, eQTL increases the prior probability of the presence of disease-causal variant(s) in the locus more effectively and unbiasedly, compared to other knowledge-based prioritizations such as gene pathway analysis [24].

Here, we report a GWAS and multi-stage replication studies for SLE examining 2,278 SLE cases and 31,948 controls in Japanese subjects. We integrated eQTL study into the results of the GWAS, which effectively enabled to detect a novel SLE susceptibility locus.

Results

GWAS for SLE

In the GWAS, 891 SLE cases and 3,384 controls in Japanese subjects were genotyped over 550,000 single nucleotide polymorphism (SNP) markers (Table S1, S2 and Figure 1). We applied stringent quality control (QC) criteria and evaluated associations of 430,797 autosomal SNPs, as previously described [26]. No substantial population stratification was demonstrated through principal component analysis (Figure S1) or a Quantile–Quantile plot of *P*-values (inflation factor, $\lambda_{GC,} = 1.088$, Figure S2), suggesting homogenous ancestries of our study population [27].

We identified significant associations in six chromosomal loci that satisfied the genome-wide significance threshold of $P<5.0\times10^{-8}$ (Table 1 and Figure 2A). These loci have been reported to be associated with SLE susceptibility (*STAT4*, *TNFAIP3*, *HIP1*, *BLK*, *ETS1*, and the HLA region) [3–18]. We also observed significant replications in 17 of the previously reported SLE susceptibility loci [3–18] ($\alpha = 0.01$; Table 2). Of these, significant replications were enriched in the loci identified through the studies in Asian populations (80%; 8 of the 10 loci), including *RASGRP3*, *IKZF1*, *HIP1*, *WDFY4*, intergenic region at 11q23, *ETS1*, *SLC15A4*, *ELF1*, and *HIC2-UBE2L3* [14,16,18], compared to those in European populations (56.3%; 9 of the 16 loci) [3–13,15,17].

Incorporation of eQTL study into GWAS results

For the selection of SNPs incorporated in the replication studies of the potential association signals, we evaluated cis-eQTL effects of the SNPs using publically available gene expression data [28], and prioritized the results of the GWAS. After applying QC criteria, we evaluated the expression levels of 19,047 probes assayed in lymphoblastoid B cell lines from Phase II HapMap East-Asian individuals [29] using Illumina's human whole-genome expression array (WG-6 version 1) [28]. For each of the SNPs included in our GWAS, probes located within ±300 kbp regions were focused on as cis-eQTLs (average 4.93 probes per SNP). We denoted the SNPs which exhibited significant associations with expression levels of any of the corresponding cis-eQTLs as eQTL positive (false discovery rate (FDR) *Q*-values<0.2). We observed enrichments of eQTL positive loci among the SLE susceptibility loci (30.8%; 8 of the 26 evaluated loci) including a well-known eQTL gene of *BLK* [11,25] (Table 2), compared to the genome-wide SNPs (6.9%) and compared even to the SNPs in the vicinity of expressed loci (among the SNPs located within ±10 kbp of probes used for the expression analysis, 13.1% were eQTL positive; Table S3).

By prioritizing the results of the GWAS using the eQTL study, we selected 57 SNPs from 1,207 SNPs that satisfied $P<1.0\times10^{-3}$ in the GWAS. We subsequently referred the associations of the selected SNPs using the results of the concurrent genome-wide scan for SLE in an independent Japanese population (Tahira T et al. Presented at the 59th Annual Meeting of the American Society of Human Genetics, October 21, 2009). In the scan, 447 SLE cases and 680 controls of Japanese origin were evaluated using a pooled DNA approach [30]. We selected SNPs if any association signals were observed in the neighboring SNPs of the