Clinical Infectious Diseases Advance Access published September 24, 2012

Single Nucleotide Polymorphisms in *ABCC2* Associate With Tenofovir-Induced Kidney Tubular Dysfunction in Japanese Patients With HIV-1 Infection: A Pharmacogenetic Study

Takeshi Nishijima,[1,2] Hirokazu Komatsu,[3] Koichiro Higasa,[4] Misao Takano,[1] Kiyoto Tsuchiya,[1] Tsunefusa Hayashida,[1] Shinichi Oka,[1,2] and Hiroyuki Gatanaga[1,2]

[1]AIDS Clinical Center, National Center for Global Health and Medicine, Tokyo; [2]Center for AIDS Research, Kumamoto University; [3]Department of Community Care, Saku Central Hospital, Nagano; and [4]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Japan

Background. Tenofovir is a widely used antiretroviral drug although it can cause kidney tubular dysfunction (KTD). The aim of this study was to determine the association between polymorphisms in genes encoding drug transporters and KTD in Japanese patients treated with tenofovir.

Methods. The association between tenofovir-induced KTD and 14 single nucleotide polymorphisms (SNPs) in the *ABCC2*, *ABCC4*, *ABCC10*, *SCL22A6*, and *ABCB1* genes was investigated in 190 Japanese patients. KTD was diagnosed by the presence of at least 3 abnormalities in the following parameters: fractional tubular resorption of phosphate, fractional excretion of uric acid, urinary β2-microglobulin, urinary α1-microglobulin, and urinary N-acetyl-β-D-glucosaminidase. Genotyping was performed by allelic discrimination using TaqMan 5′-nuclease assays with standard protocols. Associations between genotypes and KTD were tested by univariate and multivariate logistic regression analyses.

Results. KTD was diagnosed in 19 of the 190 (10%) patients. Univariate and multivariate analyses showed a significant association between KTD and genotype CC at position −24 CC (adjusted odds ratio [OR], 20.08; 95% confidence interval [CI], 1.711–235.7; $P = .017$) and genotype AA at position 1249 (adjusted OR, 16.21; 95% CI, 1.630–161.1; $P = .017$) of *ABCC2*. Multivariate analysis showed higher adjusted OR for patients with both homozygotes (adjusted OR, 38.44; 95% CI, 2.051–720.4; $P = .015$). *ABCC2* haplotype −24T and 1249G was a protective haplotype for KTD (OR, 0.098; 95% CI, .002–.603; $P = .003$

Conclusions. This is the first study of our knowledge to identify the association between SNPs in *ABCC2* and tenofovir-induced KTD in an Asian population. Close monitoring of renal function is warranted in tenofovir-treated patients with these SNPs.

Tenofovir disoproxil fumarate (TDF), a prodrug of tenofovir, is a nucleotide reverse transcriptase inhibitor widely used for the treatment of human immunodeficiency virus type 1 (HIV-1) infection and hepatitis B infection [1–4]. Tenofovir is excreted by a combination of glomerular filtration and active tubular secretion. Although the nephrotoxicity of tenofovir is regarded mild and tolerable [5–7], several cases of tenofovir-induced nephrogenic diabetes insipidus, Fanconi syndrome, and acute renal failure have been reported, and prognosis of renal function with long-term tenofovir use remains unknown [8–10].

The mechanism of tenofovir-induced kidney damage is not fully understood. However, mitochondrial damage in the proximal renal tubular cells was observed in patients with prominent tenofovir-induced kidney tubular dysfunction (KTD) [11, 12].

Received 20 May 2012; accepted 24 August 2012.

Correspondence: Hiroyuki Gatanaga, MD, PhD, AIDS Clinical Center, National Center for Global Health and Medicine, 1-21-1, Toyama, Shinjuku, Tokyo 162-0052, Japan (higatana@acc.ncgm.go.jp).

Clinical Infectious Diseases

DOI: 10.1093/cid/cis772

Because the characteristics and severity of tenofovir-induced KTD vary widely among individuals, the role of host genetics has drawn a particular attention. Single nucleotide polymorphisms (SNPs) in transporter proteins of renal tubular cells have been investigated to elucidate their roles in tenofovir-induced KTD [13–15].

Tenofovir enters kidney tubular cells through the basolateral membrane and is transported mainly by organic anion transporter (OAT) 1 and, to a lesser extent, OAT 3, encoded by genes *SLC22A6* and *SLC22A8*, respectively [16]. Tenofovir is excreted into the urine at the apical membrane by 2 transporters on the luminal membrane; multidrug resistance protein (MRP) 4 and MRP 2, encoded by the adenosine triphosphate–binding cassette (ABC) genes *ABCC4* and *ABCC2*, respectively [17, 18]. Although the role of MRP4 in transporting tenofovir has been well established, that of MRP 2 remains controversial [19, 20]. Recently, MRP 7, encoded by *ABCC10* gene, was also reported to take part in the excretion of tenofovir [21]. P-glycoprotein is a membrane protein expressed on the cells of renal proximal tubule, intestine, and hepatocytes. Encoded by *ABCB1* gene, P-glycoprotein transports TDF, the prodrug of tenofovir. SNPs on *ABCB1* might alter the expression of P-glycoprotein and thus affect exposure of tenofovir [22–24].

Previous studies reported inconsistent findings on the association of the SNPs of the transporter protein on tenofovir-induced KTD [13–15]. Several pathological processes could induce KTD, such as active infection, inflammation, diabetic nephropathy, concurrent use of nephrotoxic drugs, and preexisting renal impairment, and thus it is difficult to evaluate KTD induced exclusively by tenofovir [25]. Moreover, drug interaction with other antiretrovirals, especially ritonavir-boosted protease inhibitors, modifies tenofovir clearance and thus the severity of tenofovir-induced KTD [26, 27]. Previous studies examined patients treated with various antiretroviral combinations, which might also contribute to the inconsistent findings. Thus, the effect of SNPs on tenofovir-induced KTD remains to be clarified and isolated from other abovementioned conventional risk factors for KTD [15, 28]. Of note, the population investigated in previous studies on the role of SNPs in tenofovir-induced KTD was mostly whites, and patients of other genetic background have hardly been examined.

Based on the above background, the present study was designed to elucidate the association between polymorphisms in genes encoding drug transporters in renal tubular cells and tenofovir-induced KTD, in a setting designed to exclude other predisposing or intervening factors: the inclusion of Japanese patients with HIV infection on the same antiretroviral combination with suppressed HIV-1 viral load, and free of preexisting renal impairment, major comorbidities, and active infections.

METHODS

Ethics Statement

This study was approved by the Human Genetics Research Ethics Committee of the National Center for Global Health and Medicine, Tokyo, Japan. Each patient included in this study provided a written informed consent for genetic testing and publication of clinical data for research purposes. The study was conducted according to the principles expressed in the Declaration of Helsinki.

Study Design

We performed a single-center cohort study to cross-sectionally elucidate the association between SNPs in genes encoding renal tubular transporters in Japanese patients with HIV infection and tenofovir-induced KTD.

Study Subjects

The study included consecutive Japanese patients with HIV infection, aged >17 years, with HIV-1 viral load <200 copies/mL, and on at least 4-week treatment with once-daily ritonavir (100 mg)–boosted darunavir (800 mg) plus fixed dose tenofovir (300 mg)/emtricitabine (200 mg), seen at our clinic between 1 October 2011 and 31 March 2012. The exclusion criteria were (1) active infection, (2) malignancy, (3) diabetes mellitus, defined by the use of anti-diabetic agents or fasting plasma glucose >126 mg/dL or plasma glucose >200 mg/dL on two different days, (4) alanine aminotransferase 2.5 times more than the upper limit of normal, (5) estimated glomerular filtration rate (eGFR) calculated by Cockcroft-Gault equation of <50 mL/minutes [creatinine clearance = [(140 − age) × weight (kg)]/(serum creatinine × 72)(×0.85 for females)] [29], and (6) patients without consent to the study.

Measurements

Blood and spot urine samples were collected either on the day of enrollment or on the next visit, together with body weight measurement. The blood samples were used to measure serum creatinine, serum uric acid, serum phosphate, CD4 count, and C-reactive protein, whereas urine samples were used to measure phosphate, uric acid, creatinine, β2-microglobulin (β2M), α1-microglobulin (α1M), and N-acetyl-β-D-glucosaminidase (NAG). The values of β2M, α1M, and NAG measured in the urine samples were expressed relative to urinary creatinine of 1 g/L (/g Cr).

Urinary concentrations of β2M and α1M were measured with latex aggregation assay kits (β2M: BMG-Latex X1 "Seiken"; Denka Seiken Co, Niigata, Japan; α1M: Eiken α1M-III; Eiken Chemical Co, Tokyo, Japan), and those of NAG by colorimetric assay of enzyme activity with 6-methyl-2-pyridyl-N-acetyl-1-thio-β-D-glucosaminide as substrate (Nittobo Medical Co, Tokyo).

Definition of Renal Proximal Tubular Dysfunction

KTD was defined as the presence of at least 3 abnormalities in the following 5 parameters: fractional tubular resorption of phosphate $\{1 - [(\text{urine phosphate} \times \text{serum creatinine})/(\text{urine creatinine} \times \text{serum phosphate})]\} \times 100$ of <82%, fractional excretion of uric acid $\{[(\text{urine uric acid} \times \text{serum creatinine})/(\text{urine creatinine} \times \text{serum uric acid})] \times 100)\}$ of >15%, β2-microglobulinuria (β2M > 1000 µg/g Cr), α1-microglobulinuria (α1M > 16.6 mg/g Cr), and high-NAG level in urine (NAG > 5.93 U/g Cr). The above cutoff levels were selected on the basis of data reported previously by various investigators [15, 30, 31].

The potential risk factors for KTD were determined according to previous studies and collected together with the basic demographics from the medical records [6, 27, 32, 33]. They included age, sex, body weight, and presence or absence of other medical conditions (concurrent use of nephrotoxic drugs such as ganciclovir, sulfamethoxazole/trimethoprim, and nonsteroidal antiinflammatory agents, coinfection with hepatitis B, defined by positive hepatitis B surface antigen, coinfection with hepatitis C, defined by positive HCV viral load, hypertension, defined by current treatment with antihypertensive agents or 2 successive measurements of systolic blood pressure >140 mmHg or diastolic blood pressure >90 mmHg at the clinic, dyslipidemia, defined by current treatment with lipid-lowering agents or 2 successive measurements of either low-density lipoprotein cholesterol >140 mg/dL, high-density lipoprotein cholesterol <40 mg/dL, total cholesterol >240 mg/dL, triglyceride >500 mg/dL. At our clinic, blood pressure and body weight are measured every visit. We used the data on or closest to and preceding the day of blood/urine sample collection by no more than 180 days.

Genetic Polymorphisms

SNPs in genes encoding tubular transporters were selected on the basis of their functional significance, findings of previously published reports, and/or reported minor-allele frequencies >5% in the Japanese [13–15, 21, 28]. The allele frequency data for the Japanese were obtained from the Japanese Single Nucleotide Polymorphisms (JSNP) database [34]. The 14 SNPs selected were (1) *ABCC2* (encodes MRP2) −24C → T (in the promoter; rs717620); 1249G → A (Val417Ile; rs2273697); 2366C → T (Ser789Phe; rs56220353); 2934G → A (Ser978Ser; rs3740070), (2) *ABCC4* (encodes MRP4) 559G → T (Gly187Trp; rs11568658); 912G → T (Lys304Asn; rs2274407); 2269G → A (Glu757Lys; rs3765534); 3348A → G (Lys1116Lys; rs1751034); 4135T → G [in the 3′ untranslated region (UTR); (rs3742106)]; 4976T → C (3′ UTR; rs1059751), (3) *ABCC10* (encodes MRP10) 526G → A (intron; rs9349256); 2759T → C (Ile920Thr; rs2125739), (4) *SLC22A6* (encodes OAT1) 180C → T (Asn60Asn; rs11568630), and (5) *ABCB1* (encodes P-glycoprotein) 2677T → A/G (A:Ser893Thr, G:Ser893Ala; rs2032582).

Pharmacogenetic Analyses

Genomic DNA was extracted from peripheral-blood leukocytes using the protocol described in the sheet enclosed with the QIAamp DNA MiniKit (Qiagen, Valencia, California). All genotyping was performed by allelic discrimination using TaqMan 5′-nuclease assays with standard protocols (TaqMan SNP Genotyping Assays; Applied Biosystems, Foster City, California). The primer and probe sequences are available on request.

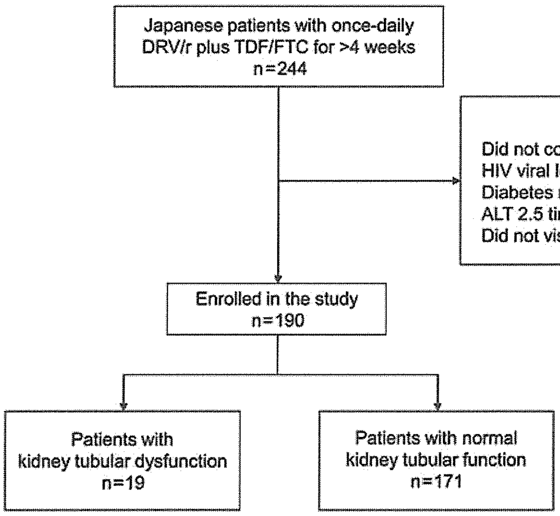


Figure 1. Patient enrollment. Abbreviations: ALT, alanine transaminase; DRV/r, ritonavir-boosted darunavir; HIV, human immunodeficiency virus; TDF/FTC, tenofovir/emtricitabine; ULN, upper limit of normal.

Statistical Analysis

Baseline characteristics were compared between patients with KTD and without tubular dysfunction by the Student t test for continuous variables and by either the χ^2 test or Fisher exact test for categorical variables. Statistical comparisons for genotype frequencies between 2 groups were made by use of 2 × 3 table Fisher exact test (2 × 6 table for rs2032582). Associations between genotypes and KTD were tested by univariate and multivariate logistic regression analyses. The impact of other variables was estimated with univariate analysis, and those with $P < .20$ were incorporated into multivariate analysis, in addition to the basic demographics such as age and sex. Statistical significance was defined at 2-sided P value < .05. We used odds ratios (ORs) and 95% confidence intervals (95% CIs) to estimate the impact of each variable on KTD. The Haploview software was used to test Hardy-Weinberg equilibrium and *ABCC2* and *ABCC4* haplotype analysis. All other statistical analyses were performed with the Statistical Package for Social Sciences ver. 17.0 (SPSS, Chicago, Illinois).

RESULTS

A total of 190 patients who provided blood and urine samples and satisfied the inclusion and exclusion criteria were enrolled in the study (Figure 1). KTD was diagnosed in 19 of the 190 patients (10%). The baseline characteristics and laboratory data for patients with and without KTD are listed in Table 1. Patients with KTD were older ($P < .001$), had smaller body weight ($P = .006$) and lower eGFR ($P = .003$), and were more likely to be hypertensive than patients with normal tubular function ($P = .088$). The median duration of tenofovir therapy was 71.5 weeks (interquartile range [IQR]: 36.8–109.2 weeks) for the entire study population, which was not different between the 2 groups ($P = .888$).

Table 1. Characteristics of Patients With and Without Kidney Tubular Dysfunction

	Patients With KTD (n = 19)	Patients With Normal Tubular Function (n = 171)	*P* Value
Variables for kidney tubular markers			
Urinary β2M (μg/g Cr)[a]	3066 (2247–10068)	209.2 (114.2–536.2)	<.001
Urinary α1M (mg/g Cr)[a]	26.5 (19.8–37.4)	7.95 (5.02–11.9)	<.001
Urinary NAG (U/g Cr)[a]	9 (6.2–14.3)	3.74 (2.84–4.95)	<.001
Fractional tubular resorption of phosphate[a]	83.9 (81.7–92)	91.9 (88.8–94.4)	<.001
Fractional excretion of uric acid[a]	9.7 (8.1–12.4)	6.4 (5.0–9.0)	<.001
Contribution of each parameter to KTD			
Urinary β2M > 1000 μg/g Cr, No. (%)	19 (100)	21 (12.3)	<.001
Urinary α1M > 16.6 mg/g Cr, No. (%)	18 (94.7)	17 (9.9)	<.001
Urinary NAG >5.93 U/g Cr, No. (%)	17 (89.5)	23 (13.5)	<.001
Fractional tubular resorption of phosphate <82%, No. (%)	5 (26.3)	2 (1.2)	<.001
Fractional excretion of uric acid >15%, No. (%)	2 (10.5)	4 (2.3)	.112
Characteristics			
Sex (male), No. (%)	18 (94.7)	166 (97.1)	.473
Age[a]	60 (41–62)	38 (32–42)	<.001
Route of transmission (homosexual contact), No. (%)	16 (84.2)	153 (89.5)	.528
Weight (kg)[a]	56 (53.5–66.5)	67.2 (58.1–75)	.006
Estimated glomerular filtration rate (mL/minutes/1.73 m^2)[a]	75.5 (62.8–93.5)	87.7 (77.5–98)	.003
Serum creatinine (mg/dL)[a]	0.85 (0.68–0.96)	0.80 (0.73–0.88)	.168
CD4 cell count (μL)[a]	380 (194–501)	379 (275–533)	.261
Serum phosphate (mg/dL)[a]	3.4 (2.7–3.7)	3.2 (2.9–3.6)	.815
Serum uric acid (mg/dL)[a]	4.7 (4.2–5.7)	5.6 (4.8–6.4)	.080
Nephrotoxic drug, No. (%)	2 (10.5)	12 (7.0)	.420
Hepatitis C, No. (%)	0 (0)	3 (1.8)	.728
Hepatitis B, No. (%)	2 (10.5)	24 (14)	.501
Dyslipidemia, No. (%)	4 (21.1)	54 (31.6)	.253
Hypertension, No. (%)	8 (42.1)	42 (24.6)	.088
C-reactive protein (mg/dL)[a]	0.07 (0.03–0.28)	0.07 (0.03–0.16)	.277
Duration of treatment with TDF (weeks)[a]	60.3 (17.7–115.4)	73.3 (37.7–109.1)	.888

Abbreviations: KTD, kidney tubular dysfunction; NAG, N-acetyl-β-D-glucosaminidase; TDF, tenofovir disoproxil fumarate.

[a] Median (interquartile range).

Table 2. Genotype Frequencies at *ABCC2*, *ABCC4*, *ABCC10*, *SLC22A6*, and *ABCB1* in Patients With and Without Kidney Tubular Dysfunction

Genotype	Amino Acid	Patients With KTD (n = 19)	Patients With Normal Tubular Function (n = 171)	*P* Value[a]
ABCC2 (MRP2)				
−24 C → T, rs717620				
C/C		18 (94.7)	108 (63.2)	
C/T		1 (5.3)	52 (30.4)	.018
T/T		0 (0)	11 (6.4)	
1249 G → A, rs2273697	Val417Ile			
G/G		11 (57.9)	133 (77.8)	
A/G		5 (26.3)	34 (19.9)	.017
A/A		3 (15.8)	4 (2.3)	
2366 C → T, rs56220353	Ser789Phe			
C/C		19 (100)	167 (97.7)	
C/T		0 (0)	3 (1.8)	1.000
T/T		0 (0)	1 (0.6)	
2934 G → A, rs3740070	Ser978Ser			
G/G		18 (94.7)	159 (93.0)	
G/A		1 (5.3)	11 (6.4)	1.000
A/A		0 (0)	1 (0.6)	
ABCC4 (MRP4)				
559 G → T, rs11568658	Gly187Trp			
G/G		13 (68.4)	133 (77.8)	
G/T		4 (21.1)	34 (19.9)	.126
T/T		2 (10.5)	4 (2.3)	
912G → T, rs2274407				
G/G		13 (68.4)	102 (59.6)	
T/G		6 (31.6)	52 (30.4)	.461
T/T		0 (0)	17 (9.9)	
2269 G → A, rs3765534	Glu757Lys			
G/G		15 (78.9)	129 (75.4)	
G/A		2 (10.5)	35 (20.5)	.241
A/A		2 (10.5)	7 (4.1)	
3348 A → G, rs1751034	Lys1116Lys			
A/A		13 (68.4)	98 (57.3)	
A/G		3 (15.8)	58 (33.9)	.185
G/G		3 (15.8)	15 (8.8)	
4135 T → G, rs3742106				
T/T		6 (31.6)	46 (26.9)	
T/G		7 (36.8)	79 (46.2)	.707
G/G		6 (31.6)	46 (26.9)	
4976T → C, rs1059751				
T/T		6 (31.6)	46 (26.9)	
T/C		5 (26.3)	86 (50.3)	.090
C/C		8 (42.1)	39 (22.8)	
ABCC10 (MRP7)				
526G → A, rs9349256				
G/G		4 (21.1)	32 (18.7)	
A/G		9 (47.4)	65 (38)	.569
A/A		6 (31.6)	74 (43.3)	

Table 2 continued.

Genotype	Amino Acid	Patients With KTD (n = 19)	Patients With Normal Tubular Function (n = 171)	*P* Value[a]
2759T → C, rs2125739	Ile920Thr			
T/T		15 (71.4)	131 (77.5)	
T/C		6 (28.6)	31 (18.3)	.488
C/C		0 (0)	7 (4.1)	
SLC22A6 (OAT1)				
180C → T, rs11568630				
C/C		18 (94.7)	164 (95.9)	
C/T		1 (5.3)	7 (4.1)	.577
T/T		0 (0)	0 (0)	
ABCB1 (P-glycoprotein)				
2677T → A/G, rs2032582	A:Ser893Thr G:Ser893Ala			
T/T		0 (0)	47 (27.5)	
T/A		3 (15.8)	14 (8.2)	
G/G		4 (21.1)	36 (21.1)	.002
G/T		8 (42.1)	46 (26.9)	
G/A		1 (5.3)	24 (14)	
A/A		3 (15.8)	4 (2.3)	

Abbreviation: KTD, kidney tubular dysfunction.

[a] By Fisher exact test.

Table 2 summarizes the distribution of genotypes at the *ABCC2*, *ABCC4*, *ABCC10*, *SLC22A11*, and *ABCB1* genes in the 2 groups. All polymorphisms were in Hardy-Weinberg equilibrium with a cutoff *P* value of .001. In single SNP analysis, a higher percentage of patients with KTD were found among genotype CC at position −24 and genotype AA at position 1249 of *ABCC2*, compared to patients with other genotypes (−24 CC; 14.3% [in 18 of 126 patients] vs 1.6% [in 1 of 64 patients]; $P = .004$) (1249 AA; 42.9% [in 3 of 7 patients] vs 8.7% [in 16 of 183 patients]; $P = .023$), respectively. The percentage of patients with KTD was also higher among genotype AA at position 2677 of *ABCB1*, compared to patients with other genotypes (2677 AA; 42.9% [in 3 of 7 patients] vs. 8.7% [in 16 of 183 patients]; $P = .023$). KTD was marginally associated with genotype AA at position 559 and genotype GG at position 4976 of *ABCC4* ($P = .112$, and .090, respectively).

Association of Genotypes with KTD

Univariate analysis showed a significant association between KTD and patients with genotype CC at position −24 (OR, = 10.50; 95% CI, 1.369–80.55; $P = .024$) and patients with genotype AA at position 1249 (OR, 7.828; 95% CI, 1.609–38.10; $P = .011$) of *ABCC2* (Table 3). The risk for KTD was higher in patients with both genotype CC at position −24 and genotype AA at position 1249 (OR, 31.88; 95% CI, 3.131–324.5; $P = .003$). Genotype AA at position 2677 of *ABCB1* was also significantly associated with KTD (OR, 7.828; 95% CI, 1.609–38.10; $P = .011$). Furthermore, old age (per 1 year, OR, 1.165; 95% CI, 1.100–1.233; $P < .001$), low body weight (per 1 kg decrement, OR, 1.076; 95% CI, 1.021–1.135; $P = .007$), and low eGFR (per 1 mL/minutes/1.73 m^2 decrement, OR, 1.052; 95% CI, 1.016–1.090; $P = .004$) were also associated with KTD.

Multivariate analysis identified genotype CC at position −24 and genotype AA at position 1249 of *ABCC2* as independent risks for KTD after adjustment for sex, age, weight, eGFR, and hypertension (adjusted OR, = 20.08; 95% CI, 1.711–235.7; $P = .017$) (adjusted OR, 16.21; 95% CI, 1.630–161.1; $P = .017$), respectively (Table 4). Patients with both of the abovementioned two homozygotes showed higher adjusted OR in multivariate analysis (adjusted OR, 38.44; 95% CI, 2.051–720.4; $P = .015$) (Table 4). On the other hand, genotype AA at position 2677 of *ABCB1* was not significantly associated with KTD in multivariate analysis adjusted for the abovementioned variables (adjusted OR, 1.686; 95%CI, .163–17.43; $P = .661$).

Association of Haplotypes at ABCC2 and ABCC4 with KTD

Haplotype construction was performed with the 4 identified SNPs with $P < .10$ in univariate analysis: *ABCC2*, −24 C → T, 1249 G → A; *ABCC4*, 559 G → T, 4976T → C (Table 4). Haplotypes with frequency of >1% were analyzed. *ABCC2* haplotype CA was significantly associated with TDF-induced KTD (OR, 2.910; 95% CI, 1.295–6.221; $P = .011$), whereas *ABCC2* haplotype TG was a protective haplotype (OR, 0.098; 95% CI, .002–.603; $P = .003$). *ABCC4* haplotype TT was marginally

Table 3. Univariate Analysis of Risks for Kidney Tubular Dysfunction in Patients With HIV Infection Treated With Tenofovir

Characteristic	OR	95% CI	*P* Value
Female sex	1.844	.204–16.67	.586
Age per 1 year	1.165	1.100–1.233	<.001
Weight per 1 kg decrement	1.076	1.021–1.135	.007
CD4 count per 1/μL decrement	1.002	.999–1.004	.261
Baseline eGFR per 1 mL/minutes/1.73 m^2 decrement	1.052	1.016–1.090	.004
Concurrent use of nephrotoxic drugs	1.559	.322–7.555	.581
Hepatitis B	0.721	.156–3.319	.674
C-reactive protein per 1 mg/dL	1.551	.689–3.494	.289
Hypertension	2.234	.843–5.922	.106
Dyslipidemia	0.578	.183–1.823	.349
Duration of treatment with tenofovir disoproxil fumarate (weeks)	0.999	.992–1.007	.888
ABCC2			
−24 CC	10.50	1.369–80.55	.024
1249 AA	7.828	1.609–38.10	.011
−24 CC plus 1249 AA	31.88	3.131–324.5	.003
2934 GG	1.358	.167–11.07	.775
ABCC4			
559 TT	4.912	.837–28.81	.078
912 TT	1.466	.531–4.042	.460
2269 AA	2.756	.530–14.34	.228
3348 GG	1.950	.510–7.463	.329
4135 GG	1.254	.450–3.494	.665
4976 CC	2.462	.925–6.547	.071
ABCC10			
526 GG	1.158	.360–3.725	.805
2759 TT	0.619	.220–1.738	.363
ABCB1			
2677 AA	7.828	1.609–38.10	.011

Abbreviations: CI, confidence interval; eGFR: estimated glomerular filtration rate; HIV, human immunodeficiency virus; OR, odds ratio.

[a] Due to low prevalence of minor alleles, rs56220353, rs11568630, and rs2274407 were not included in this analysis.

associated with tenofovir-induced KTD (OR, 2.497; 95% CI, .902–6.949; *P* = .077).

DISCUSSION

The present study demonstrated that genotype CC at position −24 and genotype AA at position 1249 of *ABCC2* gene are associated with tenofovir-induced KTD in Japanese patients with HIV-1 infection. The effect of SNPs was more evident in patients with both −24 CC and 1249 AA homozygotes than in those with either homozygote only. The findings of this study resolve long-term controversy over the role of genetic polymorphisms in tenofovir-induced KTD and confirm the effect of the SNPs in *ABCC2* gene in tenofovir-induced KTD.

Table 4. Multivariate Analysis for the Risk of Tenofovir-Induced Kidney Tubular Dysfunction With Homozygotes at −24 and 1249 of *ABCC2* in Patients With HIV Infection

ABCC2	Adjusted OR	95% CI	*P* Value
Homozygote at −24 CC	20.08	1.711–235.7	.017
Homozygote at 1249 AA	16.21	1.630–161.1	.017
Homozygotes at −24 CC plus 1249 AA	38.44	2.051–720.4	.015

Each variable was adjusted for sex, age, weight, estimated glomerular filtration rate, and hypertension.

Abbreviations: CI, confidence interval; OR, odds ratio.

CA haplotype (−24C, 1249A) of *ABCC2* was associated with tenofovir-induced KTD, whereas TG was a protective haplotype (Table 5). Izzedine et al [13] reported the role of CATC haplotype (−24C, 1249A, 3563T, 3972C) of *ABCC2* in KTD. However, 3563T did not play such role in this haplotype analysis, because the prevalence of 3563T is 0% in the Japanese, according to the HapMap data, and haplotype with only −24C plus 1249A still exhibited its effect on tenofovir-induced KTD (Table 5; www.hapmap.org). The reported association between tenofovir-induced KTD and 526G and 2759C of *ABCC10* described by Pushpakom et al [21] was also not reproduced in this study. Furthermore, SNPs in *ABCC4*, *SLC22A6*, and *ABCB1* investigated in the present study did not show a significant association with tenofovir-induced KTD (Table 3).

Three main aspects of our study are important. First, this is the first study to our knowledge that elucidated the effect of SNPs on tenofovir-induced KTD conducted in a country other than European countries or the United States. Our study examined Japanese patients of genetic background different from patients of previous studies, which consisted mostly of whites. While SNPs −24C and 1249A of *ABCC2* have been speculated to correlate with tenofovir-induced KTD in previous studies, the present study confirmed that these SNPs are risk factors for tenofovir-induced KTD in nonwhites.

The result that the SNPs in *ABCC2* are a risk for tenofovir-induced KTD can also be applied to patients with other genetic backgrounds who host SNPs −24C and 1249A. Notably, the impact of SNPs on tenofovir-induced KTD might be more significant in Africans and Indians than in Japanese or whites, considering that the allele frequencies of −24C and 1249A are higher in these population according to the HapMap data (−24C; Africans 96.9%, Indians 92.6%, Japanese 80.8%, whites 81.9%, 1249A; Africans 21.7%, Indians 30.7%, Japanese 8.9%, whites 23.7%; www.hapmap.org).

Second, the study was designed to evaluate the exclusive effect of SNPs on tenofovir-induced KTD by excluding

Table 5. Association Between Haplotype in *ABCC2* and *ABCC4* and Kidney Tubular Dysfunction

		Allele/Haplotype Frequency, %			
SNP Marker/Haplotype	Allele	KTD Group (n = 19)	Control Group (n = 171)	OR (95% CI)[a]	*P* Value
ABCC2					
−24 C → T	C	97.4	78.4	10.22 (1.658–419.8)	.003
1249 G → A	A	28.9	12.3	2.91 (1.345–6.296)	.011
ABCC2 haplotype	CA	28.9	12.3	2.91 (1.295–6.221)	.011
	TG	2.6	21.6	0.098 (.002–.603)	.003
ABCC4					
559 G → T	T	21.1	12.3	1.905 (.705–4.614)	.213
4976 T → C	T	48	55.3	0.746 (.375–1.470)	.399
ABCC4 haplotype					
TT	TT	17.6	7.9	2.497 (.902–6.949)	.077

Abbreviations: CI, confidence interval; KTD, kidney tubular dysfunction; OR, odds ratio; SNP, single-nucleotide polymorphism.

[a] ORs and *P* values are for comparisons of allele/haplotype frequencies between the kidney tubular dysfunction and control groups.

possible predisposing factors for KTD, for example, active infection, malignancies, diabetes mellitus, and preexisting renal impairment, which are known risks for KTD [35]. Patients who showed no HIV-1 viral suppression were also excluded. Furthermore, the enrolled patients were Japanese only, and this helped to examine a study population with comparatively similar genetic background. The study population was also on the same antiretroviral regimen (ritonavir-boosted darunavir plus tenofovir/emtricitabine), and this also helped to evaluate more precisely the effect of SNPs, because plasma concentration of tenofovir is affected by concomitant antiretrovirals and the delta change in plasma tenofovir concentration likely differs in the presence of each concomitant drug [26].

Third, SNPs were examined in 190 patients in this study. To our knowledge, the number of enrolled patients is the largest among the studies that have so far examined the effect of SNPs on tenofovir-induced KTD. Thus, this feature provided the study a higher statistical power than previous studies.

Why are polymorphisms in *ABCC2* a risk for tenofovir-induced KTD, even though it is controversial whether MRP2 plays a role in the excretion of tenofovir via the luminal membrane? [18, 20] The exact mechanism has not been determined yet, but we speculate 2 hypotheses. First, there might be unknown endogenous substances that influence tenofovir nephrotoxicity in renal tubular cells, and SNPs in *ABCC2* modulate the function or transportation of such substances [15]. Second, MRP2 may indeed take part in transporting tenofovir, because various substances including methotrexate are reported to be a substrate of MRP2, and *ABCC2* mutation alters excretion of those substances [36, 37]. Further studies are warranted to elucidate the exact mechanism of these SNPs on tenofovir-induced KTD. Furthermore, the impact of these SNPs on KTD with long-term TDF use needs to be evaluated in prospective studies.

Several limitations need to be acknowledged. First, not all polymorphisms in genes of the targeted transporter proteins were examined. Thus, we might have missed other important SNPs on the function of tenofovir transportation. There might be other unknown transporter proteins for tenofovir excretion in the kidney that contribute to susceptibility to tenofovir-induced KTD as well. Second, the diagnostic criteria for TDF-induced KTD are not uniformly established in the field and are different in the published studies. The criteria applied in this study are not entirely similar to the ones used in previous studies that examined the role of SNPs in tenofovir-induced KTD. However, by excluding other predisposing factors for KTD and enrolling a large number of patients, this study succeeded in providing a clear-cut association between SNPs and tenofovir-induced KTD.

In conclusion, the present study demonstrated that SNPs in *ABCC2* associate with tenofovir-induced KTD in Japanese patients, in a setting that excluded other predisposing factors. Assessment of renal tubular function is more cumbersome and costly to monitor than serum creatinine. However, monitoring tubular function is clinically important, because undetected long-term tubular dysfunction might lead to premature osteopenia due to phosphate wasting and accelerated progression of renal dysfunction. Close monitoring of tubular function is warranted in patients with *ABCC2* −24C and 1249A under TDF treatment.

Notes

Acknowledgments. The authors thank Ryo Yamada, Takuro Shimbo, Fumihiko Hinoshita, Yoshimi Kikuchi, Katsuji Teruya, Kunihisa Tsukada, Junko Tanuma, Hirohisa Yazaki, Haruhito Honda, Ei Kinai, Koji

Watanabe, Takahiro Aoki, Daisuke Mizushima, Yohei Hamada, Michiyo Ishisaka, Mikiko Ogata, Mai Nakamura, Akiko Nakano, Fumihide Kanaya, and all other staff at the AIDS Clinical Center for their help in completion of this study.

Financial support. This work was supported by a Grant-in-Aid for AIDS research from the Japanese Ministry of Health, Labour, and Welfare (H23-AIDS-001), and the Global Center of Excellence Program (Global Education and Research Center Aiming at the Control of AIDS) from the Japanese Ministry of Education, Science, Sports, and Culture.

Potential conflicts of interest. S. O. has received honorariums and research grants from MSD K.K., Abbott Japan, Janssen Pharmaceutical K.K., Pfizer, and Roche Diagnostics K.K.; has received honorariums from Astellas Pharmaceutical K.K., Bristol-Myers K.K., Daiichisankyo, Dainippon Sumitomo Pharma, GlaxoSmithKline, K.K., Taisho Toyama Pharmaceutical, Torii Pharmaceutical, and ViiV Healthcare. H. G. has received honorariums from MSD K.K., Abbott Japan, Janssen Pharmaceutical K.K., Torii Pharmaceutical, Roche Diagnostics K.K., and ViiV Healthcare. The remaining authors declare no conflict of interest.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Sax PE, Tierney C, Collier AC, et al. Abacavir/lamivudine versus tenofovir DF/emtricitabine as part of combination regimens for initial treatment of HIV: final results. J Infect Dis **2011**; 204:1191–201.
2. Post FA, Moyle GJ, Stellbrink HJ, et al. Randomized comparison of renal effects, efficacy, and safety with once-daily abacavir/lamivudine versus tenofovir/emtricitabine, administered with efavirenz, in antiretroviral-naive, HIV-1-infected adults: 48-week results from the ASSERT study. J Acquir Immune Defic Syndr **2010**; 55:49–57.
3. Arribas JR, Pozniak AL, Gallant JE, et al. Tenofovir disoproxil fumarate, emtricitabine, and efavirenz compared with zidovudine/lamivudine and efavirenz in treatment-naive patients: 144-week analysis. J Acquir Immune Defic Syndr **2008**; 47:74–8.
4. de Vries-Sluijs TE, Reijnders JG, Hansen BE, et al. Long-term therapy with tenofovir is effective for patients co-infected with human immunodeficiency virus and hepatitis B virus. Gastroenterology **2010**; 139:1934–41.
5. Izzedine H, Hulot JS, Vittecoq D, et al. Long-term renal safety of tenofovir disoproxil fumarate in antiretroviral-naive HIV-1-infected patients. Data from a double-blind randomized active-controlled multicentre study. Nephrol Dial Transplant **2005**; 20:743–6.
6. Nelson MR, Katlama C, Montaner JS, et al. The safety of tenofovir disoproxil fumarate for the treatment of HIV infection in adults: the first 4 years. AIDS **2007**; 21:1273–81.
7. Cooper RD, Wiebe N, Smith N, Keiser P, Naicker S, Tonelli M. Systematic review and meta-analysis: renal safety of tenofovir disoproxil fumarate in HIV-infected patients. Clin Infect Dis **2010**; 51:496–505.
8. Verhelst D, Monge M, Meynard JL, et al. Fanconi syndrome and renal failure induced by tenofovir: a first case report. Am J Kidney Dis **2002**; 40:1331–3.
9. Schaaf B, Aries SP, Kramme E, Steinhoff J, Dalhoff K. Acute renal failure associated with tenofovir treatment in a patient with acquired immunodeficiency syndrome. Clin Infect Dis **2003**; 37:e41–3.
10. Peyriere H, Reynes J, Rouanet I, et al. Renal tubular dysfunction associated with tenofovir therapy: report of 7 cases. J Acquir Immune Defic Syndr **2004**; 35:269–73.
11. Kohler JJ, Hosseini SH, Hoying-Brandt A, et al. Tenofovir renal toxicity targets mitochondria of renal proximal tubules. Lab Invest **2009**; 89:513–9.
12. Herlitz LC, Mohan S, Stokes MB, Radhakrishnan J, D'Agati VD, Markowitz GS. Tenofovir nephrotoxicity: acute tubular necrosis with distinctive clinical, pathological, and mitochondrial abnormalities. Kidney Int **2010**; 78:1171–7.
13. Izzedine H, Hulot JS, Villard E, et al. Association between ABCC2 gene haplotypes and tenofovir-induced proximal tubulopathy. J Infect Dis **2006**; 194:1481–91.
14. Kiser JJ, Aquilante CL, Anderson PL, King TM, Carten ML, Fletcher CV. Clinical and genetic determinants of intracellular tenofovir diphosphate concentrations in HIV-infected patients. J Acquir Immune Defic Syndr **2008**; 47:298–303.
15. Rodriguez-Novoa S, Labarga P, Soriano V, et al. Predictors of kidney tubular dysfunction in HIV-infected patients treated with tenofovir: a pharmacogenetic study. Clin Infect Dis **2009**; 48:e108–16.
16. Uwai Y, Ida H, Tsuji Y, Katsura T, Inui K. Renal transport of adefovir, cidofovir, and tenofovir by SLC22A family members (hOAT1, hOAT3, and hOCT2). Pharm Res **2007**; 24:811–5.
17. Mallants R, Van Oosterwyck K, Van Vaeck L, Mols R, De Clercq E, Augustijns P. Multidrug resistance-associated protein 2 (MRP2) affects hepatobiliary elimination but not the intestinal disposition of tenofovir disoproxil fumarate and its metabolites. Xenobiotica **2005**; 35:1055–66.
18. Imaoka T, Kusuhara H, Adachi M, Schuetz JD, Takeuchi K, Sugiyama Y. Functional involvement of multidrug resistance-associated protein 4 (MRP4/ABCC4) in the renal elimination of the antiviral drugs adefovir and tenofovir. Mol Pharmacol **2007**; 71:619–27.
19. Kohler JJ, Hosseini SH, Green E, et al. Tenofovir renal proximal tubular toxicity is regulated by OAT1 and MRP4 transporters. Lab Invest **2011**; 91:852–8.
20. Ray AS, Cihlar T, Robinson KL, et al. Mechanism of active renal tubular efflux of tenofovir. Antimicrob Agents Chemother **2006**; 50:3297–304.
21. Pushpakom SP, Liptrott NJ, Rodriguez-Novoa S, et al. Genetic variants of ABCC10, a novel tenofovir transporter, are associated with kidney tubular dysfunction. J Infect Dis **2011**; 204:145–53.
22. Hoffmeyer S, Burk O, von Richter O, et al. Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. Proc Natl Acad Sci USA **2000**; 97:3473–8.
23. Horinouchi M, Sakaeda T, Nakamura T, et al. Significant genetic linkage of MDR1 polymorphisms at positions 3435 and 2677: functional relevance to pharmacokinetics of digoxin. Pharm Res **2002**; 19:1581–5.
24. Kurata Y, Ieiri I, Kimura M, et al. Role of human MDR1 gene polymorphism in bioavailability and interaction of digoxin, a substrate of P-glycoprotein. Clin Pharmacol Ther **2002**; 72:209–19.
25. Han WK, Waikar SS, Johnson A, et al. Urinary biomarkers in the early diagnosis of acute kidney injury. Kidney Int **2008**; 73:863–9.
26. Kiser JJ, Carten ML, Aquilante CL, et al. The effect of lopinavir/ritonavir on the renal clearance of tenofovir in HIV-infected patients. Clin Pharmacol Ther **2008**; 83:265–72.
27. Goicoechea M, Liu S, Best B, et al. Greater tenofovir-associated renal function decline with protease inhibitor-based versus nonnucleoside reverse-transcriptase inhibitor-based therapy. J Infect Dis **2008**; 197:102–8.
28. Rodriguez-Novoa S, Labarga P, Soriano V. Pharmacogenetics of tenofovir treatment. Pharmacogenomics **2009**; 10:1675–85.
29. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron **1976**; 16:31–41.
30. Salem MA, el-Habashy SA, Saeid OM, el-Tawil MM, Tawfik PH. Urinary excretion of n-acetyl-beta-D-glucosaminidase and retinol binding protein as alternative indicators of nephropathy in patients with type 1 diabetes mellitus. Pediatr Diabetes **2002**; 3:37–41.
31. Ezinga M, Wetzels J, van der Ven A, Burger D. Kidney tubular dysfunction is related to tenofovir plasma concentration, abstract 603. In: Program and abstracts of the 19th Conference on Retroviruses and Opportunistic Infections. 5–8 March 2012, Seattle, Washington.
32. Gupta SK, Eustace JA, Winston JA, et al. Guidelines for the management of chronic kidney disease in HIV-infected patients: recommendations of the HIV medicine association of the infectious diseases society of America. Clin Infect Dis **2005**; 40:1559–85.

33. Gatanaga H, Tachikawa N, Kikuchi Y, et al. Urinary beta2-microglobulin as a possible sensitive marker for renal injury caused by tenofovir disoproxil fumarate. AIDS Res Hum Retroviruses **2006**; 22:744–8.
34. Hirakawa M, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, Nakamura Y. JSNP: a database of common gene variations in the Japanese population. Nucleic Acids Res **2002**; 30:158–62.
35. Ando M, Yanagisawa N, Ajisawa A, Tsuchiya K, Nitta K. Kidney tubular damage in the absence of glomerular defects in HIV-infected patients on highly active antiretroviral therapy. Nephrol Dial Transplant **2011**; 26:3224–9.
36. Hulot JS, Villard E, Maguy A, et al. A mutation in the drug transporter gene ABCC2 associated with impaired methotrexate elimination. Pharmacogenet Genomics **2005**; 15:277–85.
37. Suzuki H, Sugiyama Y. Single nucleotide polymorphisms in multidrug resistance associated protein 2 (MRP2/ABCC2): its impact on drug disposition. Adv Drug Deliv Rev **2002**; 54:1311–31.

RHEUMATOLOGY

Rheumatology 2013;52:1769–1774
doi:10.1093/rheumatology/ket241
Advance Access publication 19 July 2013

Concise report

Association of Takayasu arteritis with HLA-B*67:01 and two amino acids in HLA-B protein

Chikashi Terao[1,2], Hajime Yoshifuji[2], Koichiro Ohmura[2], Kosaku Murakami[3], Daisuke Kawabata[2], Kimiko Yurugi[4], Junichi Tazaki[5], Hideyuki Kinoshita[6], Akinori Kimura[7], Masashi Akizuki[8], Yasushi Kawaguchi[9], Hisashi Yamanaka[9], Yasuo Miura[4], Taira Maekawa[4], Hiroo Saji[10], Tsuneyo Mimori[2] and Fumihiko Matsuda[1]

Abstract

Objective. Takayasu arteritis (TAK) is a rare autoimmune arteritis that affects large arteries. Although the association between TAK and HLA-B*52:01 is established, the other susceptibility *HLA-B* alleles are not fully known. We performed genetic association studies to determine independent *HLA-B* susceptibility alleles other than HLA-B*52:01 and to identify important amino acids of HLA-B protein in TAK susceptibility.

Methods. One hundred patients with TAK and 1000 unrelated healthy controls were genotyped for *HLA-B* alleles in the first set, followed by a replication set containing 73 patients with TAK and 1000 controls to compare the frequencies of *HLA-B* alleles. Step-up logistic regression analysis was performed to identify susceptibility amino acids of HLA-B protein.

Results. Strong associations of susceptibility to TAK with HLA-B*52:01 and HLA-B*67:01 were observed ($P = 1.0 \times 10^{-16}$ and 9.5×10^{-6}, respectively). An independent susceptibility effect of HLA-B*67:01 from HLA-B*52:01 was also detected ($P = 1.8 \times 10^{-7}$). Amino acid residues of histidine at position 171 and phenylalanine at position 67, both of which are located in antigen binding grooves of the HLA-B protein, were associated with TAK susceptibility ($P \leqslant 3.8 \times 10^{-5}$) with a significant difference from other amino acid variations ($\Delta\text{AIC} \geqslant 9.65$).

Conclusion. HLA-B*67:01 is associated with TAK independently from HLA-B*52:01. Two amino acids in HLA-B protein are strongly associated with TAK susceptibility.

Key words: Takayasu arteritis, genetic association study, HLA-B, aortitis, vasculitis.

Introduction

Takayasu arteritis (TAK) is a relatively rare systemic arteritis that affects mainly large branches of the arteries, including the aorta, carotid artery, subclavian artery and coronary artery. TAK affects mainly young females. Both environmental and genetic components have been shown to be involved with the onset of TAK [1]. Among genetic components, HLA-B52, mainly B*52:01, has shown a strong association with TAK with odds ratios (ORs) of ~3 [2, 3]. The association with HLA-B52 has been repeatedly shown in various studies and the association is established beyond ethnicity [4, 5]. Other associations between TAK and HLA alleles have not been confirmed due to the low prevalence of the disease and the lack of

[1]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, [2]Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, [3]Department of Rheumatology and Clinical Immunology, Osaka Red Cross Hospital, Osaka, [4]Department of Transfusion Medicine and Cell Therapy, Kyoto University Hospital, Kyoto, [5]Department of Cardiovascular Medicine, Kyoto University Graduate School of Medicine, Kyoto, [6]Department of Medicine and Clinical Science, Kyoto University Graduate School of Medicine, Kyoto, [7]Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, [8]Akizuki Clinic, Yokohama, [9]Institute of Rheumatology, Tokyo Women's Medical University, Tokyo and [10]HLA Laboratory, Kyoto, Japan.

Submitted 29 January 2013; revised version accepted 21 May 2013.

Correspondence to: Chikashi Terao, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan. E-mail: a0001101@kuhp.kyoto-u.ac.jp

large-scale comprehensive genetic analysis using TAK patients. While a previous study showed that HLA-B39 is also associated with TAK [6], another recent study failed to replicate the association [7]. The recent study also suggested that a relatively rare *HLA-B* allele, HLA-B*67:01, was associated with TAK susceptibility (corrected $P = 0.023$) [7], but there are no replication studies. No studies have addressed independent associations of *HLA-B* alleles from HLA-B*52:01. Moreover, no large studies have ever been performed to analyse which amino acid residues are important for TAK susceptibility. As some TAK patients develop autoantibodies against aortic endocardium [8], the detection of susceptibility amino acids to TAK would lead to identification of a possible antigen that provokes an autoimmune response in TAK patients.

Materials and methods

Study subjects

DNA samples of 100 Japanese patients with TAK were collected at Kyoto University Hospital and Tokyo Women's Medical University. DNA samples of 73 patients with TAK were also collected at Kyoto University Hospital. Two patients in the first set are a parent and a child. Each of 1000 control DNA samples in the first and replication sets were collected at the HLA laboratory from unrelated healthy individuals. All the patients were diagnosed with TAK based on ACR criteria [9] or guidelines of the Japanese Circulation Society [10] or were registered to the Japanese national registry for rare and intractable diseases (http://www.nanbyou.or.jp/english/index.htm). Information on classifications of TAK based on criteria by Hata *et al.* [11] and of complications of aortic regurgitation (AR) were obtained from 75 and 85 patients, respectively, by reviewing clinical charts from the Kyoto University Hospital. A summary of the subjects is shown in supplementary Table S1, available at *Rheumatology* Online. This study was approved by the local ethics committees at each institution (Kyoto University Graduate School and Faculty of Medicine Ethics Committee, Tokyo Women's Medical University Genome Ethics Committee and Tokyo Women's Medical University Ethics Committee) and written informed consent was obtained from all subjects.

HLA-B genotyping

The cases and controls were genotyped for *HLA-B* alleles in four-digit resolution using the WAKFlow system at the Kyoto University Hospital and HLA laboratory, respectively.

Amino acid sequences of *HLA-B* alleles

Amino acid sequences were obtained for each *HLA-B* allele for four-digit resolution from the IMGT database (http://www.ebi.ac.uk/ipd/imgt/hla/). Amino acid variations were searched across all the *HLA-B* alleles contained in the study population. In total, 99 variants of amino acids over 53 positions were determined in the *HLA-B* alleles. A three-dimensional structure analysis of the HLA-B protein was performed using UCSF Chimera software (University of California, San Francisco, San Francisco, CA, USA) [12].

Statistical analysis

The association between TAK and *HLA-B* alleles was investigated using 2×2 contingency tables and tested for statistical significance using the χ^2 test or Fisher's exact test for alleles with a frequency >1% in either the cases or the controls. Heterogeneity between the two studies was analysed using the Breslow-Day test. The ORs and 95% CIs were also calculated. The relative predispositional effect (RPE) method [13] was applied to identify the associations of more than one *HLA-B* allele sequentially according to their strength. The associations between clinical phenotypes and HLA-B*67:01 were assessed by the Wilcoxon rank sum test for age at onset or Fisher's exact test for other clinical phenotypes. Crude *P*-values were indicated unless we mentioned the use of corrected *P*-values.

The 99 amino acids over 53 positions were used as independent variables in logistic regression analysis (supplementary Table S2, available at *Rheumatology* Online). When we found only two amino acids in the same position, we set one independent variable in this position. When we found more than three amino acids in the same position, we distinguished all amino acids as independent variables. Step-up multiple logistic regression analyses for TAK susceptibility were performed to identify susceptibility amino acids in *HLA-B* in the combined study until an amino acid whose *P*-value did not reach significance was chosen. The Akaike information criterion (AIC) was also calculated.

Significant levels were set according to Bonferroni's correction, namely, 0.0028 for *HLA-B* alleles, 0.05 for clinical phenotypes and 0.00051 for amino acids. Amino acid variations were considered significant over other variations when the logistic regression model containing the variation showed the smallest AIC with the difference in AIC >7 in comparison with those containing other alleles ($\Delta \text{AIC} > 7$) [14]. Permutation tests were performed 10 000 times to assess the probability that the observed improvements of deviance in multiple logistic regression analysis were obtained by chance. Deviances of the best-fitting model using one, two and three amino acids were calculated in each permutation test and sequential improvements of deviance by the best three amino acids were compared with the observed improvements. Deviance of logistic regression analysis is defined as $-2 \times \log$ likelihood of logistic models with degrees of freedom of k, where $k-1$ is the number of alleles in the models.

Results

The strong association of TAK with HLA-B*52:01 was confirmed in the first set ($P = 2.6 \times 10^{-13}$). A significant association with HLA-B*67:01 was also observed ($P = 0.00081$; Table 1), and no other *HLA-B* alleles showed significant associations with TAK. The replication

608

TABLE 1 Associations of HLA-B alleles with TAK susceptibility

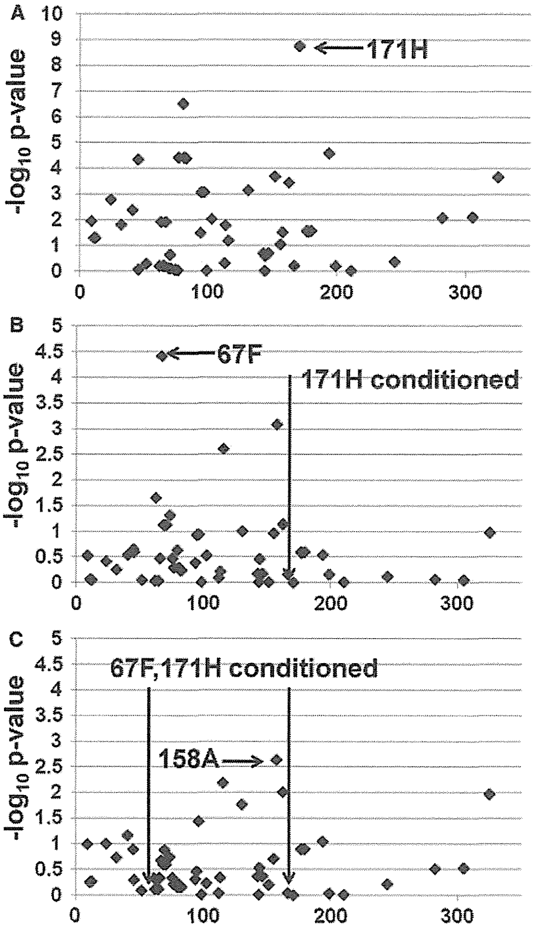
	First set				Replication set				Combined study		
HLA-B allele[a]	Case	Control	*P*	OR (95% CI)	Case	Control	*P*	OR (95% CI)	*P*	OR (95% CI)	RPE *P*
B*07:02	7	122	0.14	0.56 (0.26, 1.21)	6	119	0.36	0.68 (0.29, 1.57)	0.085	0.61 (0.34, 1.08)	
B*13:01	1	28	0.51	0.35 (0.0086, 2.17)	1	27	1.00	0.50 (0.012, 3.10)	0.32	0.42 (0.049, 1.59)	
B*15:01	16	153	0.86	1.05 (0.61, 1.80)	11	152	0.98	0.99 (0.52, 1.87)	0.90	1.03 (0.68, 1.54)	
B*15:18	1	27	0.31	0.51 (0.0089, 2.25)	0	35	0.17	0 (0, 1.52)	0.059	0.18 (0.0046, 1.08)	
B*35:01	10	160	0.13	0.61 (0.31, 1.17)	9	164	0.38	0.74 (0.37, 1.47)	0.084	0.66 (0.41, 1.06)	
B*39:01	3	68	0.20	0.43 (0.086, 1.34)	8	61	0.11	1.84 (0.86, 3.93)	0.96	0.99 (0.53, 1.84)	
B*40:01	6	108	0.14	0.54 (0.24, 1.25)	9	115	0.84	1.08 (0.53, 2.17)	0.33	0.77 (0.45, 1.31)	
B*40:02	14	177	0.38	0.78 (0.44, 1.36)	6	153	0.11	0.52 (0.22, 1.19)	0.11	0.68 (0.43, 1.09)	
B*40:06	8	92	0.70	0.86 (0.41, 1.81)	1	102	0.0085	0.13 (0.0032, 0.74)	0.057	0.52 (0.27, 1.03)	
B*44:03	13	125	0.89	1.04 (0.58, 1.88)	6	115	0.41	0.7 (0.30, 1.62)	0.70	0.91 (0.56, 1.47)	
B*46:01	11	87	0.45	1.28 (0.67, 2.44)	6	102	0.6	0.8 (0.34, 1.85)	0.87	1.04 (0.63, 1.73)	
B*48:01	5	60	0.69	0.83 (0.33, 2.09)	3	61	0.80	0.67 (0.13, 2.08)	0.45	0.76 (0.37, 1.57)	
B*51:01	10	161	0.12	0.60 (0.31, 1.16)	14	169	0.63	1.15 (0.65, 2.04)	0.39	0.83 (0.54, 1.27)	
B*52:01	60	232	$2.6x10^{-13}$	3.27 (2.34, 4.55)	34	235	4.8×10^{-5}	2.28 (1.52, 3.43)	1.6×10^{-16}	2.82 (2.19, 3.64)	1.6×10^{-16}
B*54:01	12	133	0.72	0.90 (0.49, 1.65)	9	149	0.57	0.82 (0.41, 1.63)	0.57	0.85 (0.54, 1.35)	
B*55:02	2	47	0.31	0.42 (0.049, 1.62)	5	58	0.72	1.19 (0.47, 3.01)	0.72	0.77 (0.35, 1.66)	
B*59:01	2	41	0.43	0.48 (0.056, 1.88)	1	39	0.52	0.35 (0.0085, 2.08)	0.21	0.43 (0.086, 1.31)	
B*67:01	9	27	0.00081	3.44 (1.60, 7.43)	5	19	0.0061	3.70 (1.36, 10.05)	9.5×10^{-6}	3.62 (1.97, 6.66)	1.8×10^{-7}

[a] Alleles with a frequency of >1% in either the case or control are indicated.

study was performed to confirm the findings in the first set. As a result, the association of HLA-B*67:01 with TAK was observed ($P = 0.0061$; Table 1), as well as that of HLA-B*52:01 ($P = 4.8 \times 10^{-5}$). HLA-B*40:06 showed a protective association against TAK ($P = 0.0085$). In the combined analysis, solid evidence of associations of HLA-B*67:01 and HLA-B*52:01 with TAK was obtained [$P = 9.5 \times 10^{-6}$ and 1.0×10^{-16}, OR = 3.62 (95% CI 1.97, 6.66) and 2.82 (95% CI 2.19, 3.64), respectively; Table 1]. Any signs of heterogeneity between the two studies were observed for these two alleles ($P \geqslant 0.18$). No other alleles, including HLA-B*40:06, showed significant associations. To assess the independence of HLA-B*67:01's association from HLA-B*52:01 and to detect further candidates of independent susceptibility *HLA-B* alleles, RPE analysis was performed in the combined study. As a result, HLA-B*67:01 showed a significant association with TAK susceptibility that was independent of HLA-B*52:01 ($P = 1.8 \times 10^{-7}$; Table 1). The analysis did not show further candidates with independent associations with TAK ($P \geqslant 0.17$). Since previous studies suggested associations between *HLA-B* alleles and disease phenotypes [15], we analysed the associations between HLA-B*67:01 and age at onset, female ratio, classifications of TAK or complications of AR. As a result, we did not observe significant associations ($P \geqslant 0.61$; supplementary Tables S3 and S4, available at *Rheumatology* Online).

Next, which amino acid residues contribute to TAK susceptibility was analysed by a step-up multiple logistic regression analysis in the combined study, using 99 amino acid variations over the HLA-B protein (see the Materials and methods section). Logistic regression analysis revealed the strongest association of histidine residue at the position of the 171st amino acid residue ($P = 1.8 \times 10^{-9}$; Fig. 1 and supplementary Table S5, available at *Rheumatology* Online), with a ΔAIC of 11.17 in comparisons with other amino acids. In the case of conditioning with the 171st histidine, phenylalanine at the 67th amino acid residue showed the strongest association ($P = 3.8 \times 10^{-5}$; Fig. 1), with an ΔAIC of 9.65. An alanine residue at the 158th position showed a suggestive association in conditioning with these two amino acids ($P = 0.0024$; Fig. 1) without significant ΔAIC (0.13), indicating uncertainty of the selection of this amino acid as the third susceptibility amino acid. We did not find any further amino acid variations with significant associations in the condition with these three amino acids ($P \geqslant 0.065$). A total of 10 000 permutation tests revealed that improvements in deviance from these three amino acids were less likely to be obtained by chance (permutation *P*-value 0.0001, 0.0001 and 0.0024, respectively). The 171st and 67th amino acid residues are located in peptide binding grooves, implying that an antigen binding capacity conferred by these amino acids might influence TAK susceptibility (supplementary Fig. S1, available at *Rheumatology* Online). When HLA-B*52:01 was excluded from the association studies, the two amino acid variations did not show significant associations (data not shown), indicating that the associations of the two amino acids were brought by HLA-B*52:01.

Fig. 1 Associations of amino acids of the HLA-B protein with TAK susceptibility.



P-values are plotted according to the amino acid positions in (**A**) single logistic regression analysis, (**B**) in the condition with histidine at position 171 and (**C**) with 171 histidine and 67 phenylalanine.

Discussion

This is the first study to provide solid evidence of an *HLA-B* TAK susceptibility allele independent of HLA-B*52:01 and to report on TAK susceptibility amino acids in HLA-B protein. As the top two residues of 171 and 67 were located at peptide binding grooves and have shown their significance in peptide binding [16, 17], the susceptibility effects on TAK appear to be reasonable.

HLA-B39 did not show an association in the current study, or in another recent study [7]. Thus association between HLA-B39 and TAK is not likely. Although the previous study suggested the association of TAK with HLA-B*67:01, the association was not conclusive [7]. The current study has clearly revealed that HLA-B*67:01 is associated with TAK independently of HLA-B*52:01. The association and independence of

HLA-B*67:01 were observed in the two independent sets. The two relatives in our study did not have HLA-B*67:01, thus having no effect on the conclusion. When we analysed the previous data in which only serological *HLA-B* type was available [6], HLA-B67 showed a suggestive association with TAK after excluding HLA-B52 (95 cases and 232 controls) [6]. As HLA-B*67:01 is present specifically in East Asians, this association should be specific to East Asians. The RPE analysis did not reveal an association of HLA-B*40:06. This indicates that the protective association of HLA-B*40:06 in the replication set was conferred by positive associations of HLA-B*52:01 and HLA-B*67:01. While we detected the association of HLA-B*67:01 with TAK, HLA-B*67:01 did not show significant associations with age at onset, female ratio, TAK classifications and AR. Based on the fact that the *P*-values are far from significant, it is not very likely that HLA-B*67:01 has a strong effect on clinical phenotype. Further detailed clinical information, including affected branches of the aorta, and disease activity were available for only a limited number of patients in our study, so we could not analyse the association between the susceptibility allele and detailed clinical phenotypes.

The 171st histidine was found to be a susceptibility amino acid and the 67th phenylalanine was found to be a protective amino acid. While the permutation *P*-value supported the existence of the third susceptibility amino acid, alanine at position 158 did not show a significant ΔAIC in comparison with other amino acids. Thus we cannot specify the third susceptibility amino acid to TAK. The significant associations of the top two amino acids were mainly attributed to HLA-B*52:01. No other alleles share the two amino acids with HLA-B*52:01. Although HLA-B*51:01 has similar amino acid sequences to HLA-B*52:01 with two different amino acid residues, it is not associated with TAK [4, 7]. This lack of association can be explained by our findings because HLA-B*51:01 includes the 67th phenylalanine, the protective allele against TAK. HLA-B*67:01 did not have histidine in position 171. The effects of amino acids in HLA-B*67:01 were not very apparent due to its low frequency. These two amino acids did not explain TAK susceptibility more efficiently than HLA-B*52:01 (ΔAIC over the null model = 53.94 and 66.01, respectively), indicating that these amino acids cannot explain all the susceptibility effects of HLA-B*52:01. This suggests the existence of a haplotypic effect of amino acids in HLA-B*52:01 or further susceptibility amino acids.

The 171st tyrosine forming an A pocket in the $\alpha 2$ domain is one of the highly conserved amino acid residues among *HLA-B* alleles. The conversion into histidine has been shown to modulate the binding groove, coordinating the N-terminus of the binding peptides in HLA-B14 and HLA-B*51:01 [17, 18]. The 67th serine in the B pocket of the peptide binding groove in HLA-B27 has been shown to influence peptide presentation [16]. These reports suggest important roles for the two amino acids in TAK susceptibility by modulating peptide binding. A previous Mexican study, which included 19 patients with TAK, suggested that serine at 67 and glutamic acid at 63 was associated with TAK [19]. The previous report seemed to support our results, as the current study showed non-phenylalanine amino acid residues at position 67, including serine, were positively associated with TAK. As the current study did not specify the third susceptibility amino acid, the specific association of alanine at position 158 with TAK is inconclusive. Thus these combinations of amino acids should be regarded as incomplete and further investigation is required.

As previous studies surveying susceptibility HLA alleles beyond *HLA-B* did not find significant alleles beyond the *HLA-B* alleles [6, 7], it is not very likely that amino acids in other HLA alleles linked with HLA-B*52:01 and HLA-B*67:01 can explain the TAK susceptibility of the HLA locus. Nevertheless, it is undeniable that other genes in the HLA locus, without belonging to HLA alleles, play a critical role in TAK susceptibility. Amino acid residues in the HLA locus should be analysed to grasp the whole picture of the relationship between the HLA locus and TAK susceptibility. It would be interesting to determine what kinds of protein bind to the alleles containing the susceptibility amino acid residues.

Rheumatology key messages

- HLA-B*67:01 is an independent susceptibility allele to TAK from HLA-B*52:01.
- Amino acids at positions 171 and 67 of the HLA-B protein show TAK susceptibility effects .

Acknowledgements

We would like to thank Ms Kayo Umemoto, who coordinated the acquisition of blood samples. We would like to thank all the patients who gave blood samples and the medical staff members who helped us with this study. C.T. and H. Yoshifuji participated in conception and design of this study. C.T., H. Yoshifuji, K.O., K.M., D.K., J.T., H.K., K.A., K.Y., Y.M., T. Maekawa, H.S., M.A., Y.K., H. Yamanaka, F.M. and T. Mimori participated in acquisition of data. C.T. performed analysis and interpretation of the data. C.T. wrote the initial draft and all the authors reviewed and approved the final manuscript.

Funding: This study was supported by a Kyoto University Step-up grant as well as grants-in-aid from the Ministry of Health, Labor, and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Disclosure statement: The authors declare no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Arnaud L, Haroche J, Mathian A *et al.* Pathogenesis of Takayasu's arteritis: a 2011 update. Autoimmun Rev 2011; 11:61-7.

2 Yoshida M, Kimura A, Katsuragi K *et al.* DNA typing of HLA-B gene in Takayasu's arteritis. Tissue Antigens 1993; 42:87-90.

3 Yajima M, Numano F, Park YB *et al.* Comparative studies of patients with Takayasu arteritis in Japan, Korea and India—comparison of clinical manifestations, angiography and HLA-B antigen. Jpn Circ J 1994;58:9-14.

4 Sahin Z, Bicakcigil M, Aksu K *et al.* Takayasu's arteritis is associated with HLA-B*52, but not with HLA-B*51, in Turkey. Arthritis Res Ther 2012;14:R27.

5 Vargas-Alarcon G, Flores-Dominguez C, Hernandez-Pacheco G *et al.* Immunogenetics and clinical aspects of Takayasu's arteritis patients in a Mexican Mestizo population. Clin Exp Rheumatol 2001;19:439-43.

6 Kimura A, Kitamura H, Date Y *et al.* Comprehensive analysis of HLA genes in Takayasu arteritis in Japan. Int J Cardiol 1996;54(Suppl):S61-9.

7 Takamura C, Ohhigashi H, Ebana Y *et al.* New human leukocyte antigen risk allele in Japanese patients with Takayasu arteritis. Circ J 2012;76:1697-702.

8 Eichhorn J, Sima D, Thiele B *et al.* Anti-endothelial cell antibodies in Takayasu arteritis. Circulation 1996;94: 2396-401.

9 Arend WP, Michel BA, Bloch DA *et al.* The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum 1990; 33:1129-34.

10 Japanese Circulation Society Joint Working Group. Guideline for management of vasculitis syndrome (JCS 2008). Circ J 2011;75:474-503.

11 Hata A, Noda M, Moriwaki R *et al.* Angiographic findings of Takayasu arteritis: new classification. Int J Cardiol 1996; 54(Suppl):S155-63.

12 Pettersen EF, Goddard TD, Huang CC *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004;25: 1605-12.

13 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.

14 Burnham KP, Anderson, David R. Model selection and multi-model inference: a practical information-theoretic approach. 2nd edn. New York: Springer-Verlag, 2002.

15 Kitamura H, Kobayashi Y, Kimura A *et al.* Association of clinical manifestations with HLA-B alleles in Takayasu arteritis. Int J Cardiol 1998;66(Suppl 1): S121-6.

16 Appel H, Kuon W, Kuhne M *et al.* The solvent-inaccessible Cys67 residue of HLA-B27 contributes to T cell recognition of HLA-B27/peptide complexes. J Immunol 2004;173:6564-73.

17 Maenaka K, Maenaka T, Tomiyama H *et al.* Nonstandard peptide binding revealed by crystal structures of HLA-B*5101 complexed with HIV immunodominant epitopes. J Immunol 2000;165:3260-7.

18 Kumar P, Vahedi-Faridi A, Saenger W *et al.* Structural basis for T cell alloreactivity among three HLA-B14 and HLA-B27 antigens. J Biol Chem 2009;284: 29784-97.

19 Vargas-Alarcon G, Hernandez-Pacheco G, Soto ME *et al.* Comparative study of the residues 63 and 67 on the HLA-B molecule in patients with Takayasu's arteritis. Immunol Lett 2005;96:225-9.

Functional Variants in *NFKBIE* and *RTKN2* Involved in Activation of the NF-κB Pathway Are Associated with Rheumatoid Arthritis in Japanese

Keiko Myouzen[1], Yuta Kochi[1,2]*, Yukinori Okada[1,2,3], Chikashi Terao[4,5], Akari Suzuki[1], Katsunori Ikari[6], Tatsuhiko Tsunoda[7], Atsushi Takahashi[3], Michiaki Kubo[8], Atsuo Taniguchi[6], Fumihiko Matsuda[4,9,10], Koichiro Ohmura[5], Shigeki Momohara[6], Tsuneyo Mimori[5], Hisashi Yamanaka[6], Naoyuki Kamatani[11], Ryo Yamada[12], Yusuke Nakamura[13], 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, the University of Tokyo, Tokyo, Japan, **3** Laboratory for Statistical Analysis, CGM, RIKEN, Yokohama, Japan, **4** Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **5** Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **6** Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan, **7** Laboratory for Medical Informatics, CGM, RIKEN, Yokohama, Japan, **8** Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan, **9** CREST Program, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan, **10** Institut National de la Santé et de la Recherche Médicale (INSERM), Unité U852, Kyoto University Graduate School of Medicine, Kyoto, Japan, **11** Laboratory for International Alliance, CGM, RIKEN, Yokohama, Japan, **12** Unit of Statistical Genetics, Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **13** Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan

Abstract

Rheumatoid arthritis is an autoimmune disease with a complex etiology, leading to inflammation of synovial tissue and joint destruction. Through a genome-wide association study (GWAS) and two replication studies in the Japanese population (7,907 cases and 35,362 controls), we identified two gene loci associated with rheumatoid arthritis susceptibility (*NFKBIE* at 6p21.1, rs2233434, odds ratio (OR) = 1.20, $P = 1.3 \times 10^{-15}$; *RTKN2* at 10q21.2, rs3125734, OR = 1.20, $P = 4.6 \times 10^{-9}$). In addition to two functional non-synonymous SNPs in *NFKBIE*, we identified candidate causal SNPs with regulatory potential in *NFKBIE* and *RTKN2* gene regions by integrating *in silico* analysis using public genome databases and subsequent *in vitro* analysis. Both of these genes are known to regulate the NF-κB pathway, and the risk alleles of the genes were implicated in the enhancement of NF-κB activity in our analyses. These results suggest that the NF-κB pathway plays a role in pathogenesis and would be a rational target for treatment of rheumatoid arthritis.

Citation: Myouzen K, Kochi Y, Okada Y, Terao C, Suzuki A, et al. (2012) Functional Variants in *NFKBIE* and *RTKN2* Involved in Activation of the NF-κB Pathway Are Associated with Rheumatoid Arthritis in Japanese. PLoS Genet 8(9): e1002949. doi:10.1371/journal.pgen.1002949

Editor: Panos Deloukas, The Wellcome Trust Sanger Institute, United Kingdom

Received March 31, 2012; **Accepted** July 12, 2012; **Published** September 13, 2012

Funding: This work was conducted as a part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Sciences, and Technology of the Japanese government. 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

Introduction

Rheumatoid arthritis (RA [MIM 180300]) is an autoimmune disease [1] with a complex etiology involving several genetic factors as well as environmental factors. Previous genome-wide association studies (GWAS) for RA have discovered many genetic loci [2–6], although the causal mechanisms linking the variants in these loci and disease etiology are largely unknown, except for in a few cases [6–8]. In contrast to mutations in Mendelian, monogenic diseases, most disease-associated variants in complex diseases, including autoimmune diseases, have moderate effects on disease susceptibility. This is because the disease causal variants in complex diseases are thought to have moderate effects on gene function, while amino acid changes introduced by the mutations of monogenic diseases have critical impacts on protein function [9]. Moreover, it has been demonstrated that the majority of autoimmune disease loci are expression quantitative trait loci (eQTLs) [10,11], indicating that accumulation of quantitative, but not qualitative, changes in gene function likely predisposes individuals to the disease. This renders it difficult to pinpoint the causal variants in the GWAS loci, especially in eQTLs, because all the variations in strong linkage disequilibrium (LD) with the marker SNP in a GWAS, the majority of which are not covered by the SNP array, are possible candidates for causal variants.

In recent years, with the emergence of next-generation sequencing technologies, the way we approach disease-causing variants has dramatically changed. First, a comprehensive map of human genetic variations is now available owing to the 1000 Genome Project [12], which allows us to grasp most of the potential common variants. This also enables us to perform genotype imputation of SNPs that are not directly genotyped in the GWAS, and consequently, to test them for association. Second, genomic studies using new technologies, such as chromatin immunoprecipitation-sequencing (ChIP-seq) and DNase I hypersensitive sites sequencing (DNase-seq), have advanced our understanding of how each genomic cluster regulates gene

Author Summary

Rheumatoid arthritis (RA) is a chronic autoimmune disease affecting approximately 1% of the general adult population. More than 30 susceptibility loci for RA have been identified through genome-wide association studies (GWAS), but the disease-causal variants at most loci remain unknown. Here, we performed replication studies of the candidate loci of our previous GWAS using Japanese cohorts and identified variants in *NFKBIE* and *RTKN2* gene loci that were associated with RA. To search for causal variants in both gene regions, we first examined non-synonymous (ns)SNPs that alter amino-acid sequences. As *NFKBIE* and *RTKN2* are known to be involved in the NF-κB pathway, we evaluated the effects of nsSNPs on NF-κB activity. Next, we screened *in silico* variants that may regulate gene transcription using publicly available epigenetic databases and subsequently evaluated their regulatory potential using *in vitro* assays. As a result, we identified multiple candidate causal variants in *NFKBIE* (2 nsSNPs and 1 regulatory SNP) and *RTKN2* (2 regulatory SNPs), indicating that our integrated *in silico* and *in vitro* approach is useful for the identification of causal variants in the post-GWAS era.

transcription. If disease-associated variants are present in a critical site for gene regulation suggested by the ChIP-seq and DNase-seq studies, the disease-associated variants might possibly influence gene transcription levels such as through altered transcription factor-DNA binding avidity.

In the present study, we first performed replication studies of candidate loci in our previous GWAS and identified two association signals with genome-wide significance ($P{<}5\times10^{-8}$) in nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon (*NFKBIE* [MIM 604548]) and rhotekin 2 (*RTKN2*) loci. By utilizing publicly available datasets yielded by the above-mentioned genomic studies, we then performed integrated *in silico* and *in vitro* analysis to identify plausible causal variants in *NFKBIE* and *RTKN2* loci.

Results

Identification of rheumatoid arthritis susceptibility genes

We previously performed a GWAS of RA using a Japanese case-control cohort (2,303 cases and 3,380 controls) and identified significant associations in major histocompatibility complex, class II, DR beta 1 (*HLA-DRB1* [MIM 142857]), and chemokine (C-C motif) receptor 6 (*CCR6* [MIM 601835]) loci ($P_{\mathrm{GWAS}}{<}5\times10^{-8}$) [6]. To reveal additional risk loci from those showing moderate associations in the GWAS (31 loci, $5\times10^{-8}{<}P_{\mathrm{GWAS}}{<}5\times10^{-5}$), we selected a landmark SNP from each locus and genotyped it for an additional cohort (replication-1: 2,187 cases and 28,219 controls) (Table S1, S2). Among the 31 SNPs genotyped, seven SNPs were nominally associated with RA ($P{<}0.05$), which included SNPs in the tumor necrosis factor, alpha-induced protein 3 (*TNFAIP3* [MIM 191163]), and signal transducer and the activator of transcription 4 (*STAT4* [MIM 600558]) gene loci that were previously reported to be associated with RA [13,14] (Table S2). In a combined analysis of the GWAS and the 1st replication study, we identified two associations with genome-wide significance ($P{<}5\times10^{-8}$) in *NFKBIE* (6p21.1, rs2233434, $P = 4.1\times10^{-11}$, odds ratio (OR) = 1.21, 95% confidence interval (CI) = 1.14–1.28) and in *RTKN2* (10q21.2, rs3125734, $P = 3.7\times10^{-8}$, OR = 1.23, 95% CI = 1.14–1.32) (Table 1 and Figure 1). *NFKBIE* was previously reported as a novel RA susceptibility gene locus in a meta-analysis of three GWASs for RA in the Japanese population, which included the GWAS set that the present study used [15]. *RTKN2* is located in the same region (10q21) as *ARID5B*, in which a significant association signal was also reported in the meta-analysis [15]. In our GWAS set, however, two significant signals were observed at rs3125734 (*RTKN2*: $P = 4.8\times10^{-5}$) and rs10821944 (*ARID5B*: $P = 7.4\times10^{-4}$), the former of which was tested as a landmark in the replication study. These two SNPs were in weak LD ($r^2 = 0.11$) and the independent effect of each SNP was observed after conditioning on each SNP (*RTKN2*: $P = 1.5\times10^{-3}$, *ARID5B*: $P = 0.024$, respectively). This indicated that two independent associations existed in this region, and the association of *RTKN2* is novel. We also confirmed the association in the *STAT4* locus [14] with genome-wide significance (2q32.2, rs10168266, $P = 3.2\times10^{-8}$, OR = 1.16, 95% CI = 1.10–1.22) (Table S2). The associations in *NFKBIE* and *RTKN2* were further replicated in the 2nd replication cohort (3,417 cases and 3,763 controls; rs2233434, $P = 1.1\times10^{-5}$, OR = 1.19, 95% CI = 1.10–1.30 and rs3125734, $P = 0.016$, OR = 1.14, 95% CI = 1.02–1.26, respectively), confirming the associations in these loci (a combined analysis of three sets; rs2233434, $P = 1.3\times10^{-15}$, OR = 1.20, 95% CI = 1.15–1.26 and rs3125734, $P = 4.6\times10^{-9}$, OR = 1.20, 95% CI = 1.13–1.27, respectively) (Table 1 and Figure 1). We also genotyped these SNPs for individuals with systemic lupus erythematosus (SLE [MIM 152700]) ($n = 657$) and Graves' disease (GD [MIM 275000]) ($n = 1{,}783$). We identified a significant association of *RTKN2* (rs3125734) with GD ($P = 3.4\times10^{-5}$, OR = 1.24, 95% CI = 1.12–1.37), whereas no significant associations were detected in *NFKBIE* (rs2233434) with either disease or in *RTKN2* (rs3125734) with SLE (Table S3).

Functional analysis of non-synonymous SNPs

NFKBIE and *RTKN2* genes are both involved in the NF-κB pathway: *NFKBIE* encodes IκB epsilon (IκBε), a member of the IκB family [16], and its binding to NF-κB inhibits the nuclear translocation of NF-κB [17]; *RTKN2* encodes a member of Rho-GTPase effector proteins highly expressed in CD4$^+$ T cells [18] and is implicated in the activation of the NF-κB pathway [19]. Considering that the NF-κB pathway is critical for the pathogenesis of RA [20], these two genes could be strong candidates in these regions. To identify disease-causing variants, we first sequenced the coding regions of the genes using DNA from patients ($n = 48$) to find variants that alter amino acid sequences. We identified four non-synonymous (ns)SNPs, which were all registered in the dbSNP database: two nsSNPs in *NFKBIE* (rs2233434 (Val194Ala) and rs2233433 (Pro175Leu)) and two in *RTKN2* (rs3125734 (Arg462His) and rs61850830 (Ala288Thr)), where rs2233434 and rs3125734 were the same as the landmark SNPs in the GWAS (Figure 1 and Figure 2A). The two nsSNPs of each locus were in strong LD (Figure 2B) and were both associated with disease (Table S4). In the haplotype analysis, a single common risk haplotype with a frequency >0.05 was observed in each locus, and significant associations with disease risk were detected (*NFKBIE*, $P = 5.3\times10^{-8}$, Table S5; *RTKN2*, $P = 5.7\times10^{-5}$, Table S6).

To investigate the effect of these nsSNPs on protein function, we evaluated them by *in silico* analysis using PolyPhen and SIFT software, which predicts possible impacts of amino acid substitutions on the structure and function of proteins, but all four nsSNPs were predicted to have little effect (Table S7), contrasting with the effect of Mendelian disease mutations [9]. We next examined their influence on the NF-κB activity in cells by performing NF-κB

Table 1. Association analysis of *NFKBIE* and *RTKN2* with rheumatoid arthritis.

		Allele		**Number of subjects**		**Frequency of allele 1**			
Gene	**dbSNP ID**	**(1/2)**	**Study set**	**Case**	**Control**	**Case**	**Control**	**Odds ratio (95% CI)**	**P-value**[a]
NFKBIE	rs2233434	G/A	GWAS	2,303	3,380	0.254	0.216	1.24 (1.13–1.35)	2.2×10^{-6}
			Replication study-1	2,186	28,204	0.245	0.215	1.19 (1.10–1.27)	4.2×10^{-6}
			Replication study-2	3,396	3,756	0.239	0.209	1.19 (1.10–1.30)	1.1×10^{-5}
			Combined analysis	7,885	35,340	0.245	0.215	1.20 (1.15–1.26)	1.3×10^{-15}
RTKN2	rs3125734	T/C	GWAS	2,303	3,380	0.125	0.101	1.27 (1.13–1.43)	4.8×10^{-5}
			Replication study-1	2,185	28,218	0.129	0.110	1.20 (1.09–1.31)	1.4×10^{-4}
			Replication study-2	3,402	3,751	0.115	0.103	1.14 (1.02–1.26)	0.016
			Combined analysis	7,890	35,349	0.122	0.108	1.20 (1.13–1.27)	4.6×10^{-9}

[a]: Cochran-Armitage trend test was used for the GWAS and replication studies. Mantel-Haenszel method was used for the combined analysis.

doi:10.1371/journal.pgen.1002949.t001

reporter assays with haplotype-specific expression vectors. In *NFKBIE*, the non-risk haplotype (A-C: rs2233434 (non-risk allele (NR))-rs2233433 (NR)) displayed an inhibitory effect on NF-κB activity compared with the mock construct, which reflected compulsorily binding of exogenous IκBε to the endogenous NF-κB, as shown in a previous study [16]. Of note, the risk haplotype (G-T: risk allele (R)-R) showed higher NF-κB activity than A-C (NR-NR) (Figure 3A), suggesting impaired inhibitory potential of G-T (R-R) products. No haplotypic difference was detected in the protein expression levels of these constructs (Figure 3C). We also examined two additional constructs of G-C (R-NR) and A-T (NR-R) haplotypes to evaluate the effect of each nsSNP (Figure S1A, S1B). Because NF-κB activity increased in the order of A-C<G-C<A-T<G-T (rs2233434-rs2233433: NR-NR<R-NR<NR-R<R-R) when cells were stimulated with TNF-α, the C>T substitution (Pro175Leu) in rs2233433 may have more impact on the protein function of IκBε compared with the A>G substitution (Val194Ala) in rs2233434. In contrast to the observations in *NFKBIE*, no clear difference was detected between the two common haplotype products of *RTKN2* in either their effect on NF-κB activity or protein expression levels, although both products enhanced NF-κB activity as reported previously (Figure 3B, 3D) [19]. These functional analyses of nsSNPs suggest that two nsSNPs (rs2233434 and rs2233433) in the *NFKBIE* region are candidates for causal SNPs.

ASTQ analysis suggested the existence of regulatory variants

As the majority of autoimmune disease loci have been implicated as eQTL [11], we speculated that variants in the *NFKBIE* and *RTKN2* loci would influence gene function by regulating gene expression, in addition to changing the amino acid sequences. To address this possibility, we performed allele-specific transcript quantification (ASTQ) analysis by using allele-specific probes targeting the nsSNPs in exons (rs2233434 for *NFKBIE* and rs3125734 for *RTKN2*, both of which were the GWAS landmarks). The genomic DNAs and cDNAs were extracted from peripheral blood mononuclear cells (PBMCs) in individuals with heterozygous genotype ($n = 14$ for *NFKBIE* and $n = 6$ for *RTKN2*) and from lymphoblastoid B-cell lines ($n = 9$) for *NFKBIE*. As the expression levels of *RTKN2* were low in lymphoblastoid B cells, only PBMCs were used. When quantified by allele-specific probes, transcripts from the risk allele of *NFKBIE* showed 1.1-fold and 1.2-fold lower amounts (in PBMCs and lymphoblastoid B cells, respectively) than those from non-risk alleles ($P = 0.012$ and 5.3×10^{-4}, respectively; Figure 3E and Figure S2). In contrast, 1.5-fold higher amounts of transcripts were observed in the risk allele of *RTKN2* ($P = 0.016$; Figure 3F). These allelic imbalances suggested that both gene loci were eQTL and that there existed variants with *cis*-regulatory effects. Moreover, considering the inhibitory effects of *NFKBIE* and the activating potential of *RTKN2* on NF-κB activity, which might both be dose dependent (Figure 3G, 3H), these regulatory variants in the risk alleles should enhance NF-κB activity *in vivo*.

Integrated *in silico* and *in vitro* analysis to search for regulatory variants

To comprehensively search the two genomic regions for causal regulatory variants, we performed an integrated *in silico* and *in vitro* analysis with multiple steps (Figure 4 and Figures S3, S4). We first determined the target genomic region by selecting LD blocks containing disease-associated SNPs ($P_{\mathrm{GWAS}} < 1.0\times10^{-3}$) (Step 1). We then extracted SNPs with frequencies of >0.05 from HapMap and 1000 Genome Project databases in the region (Step 2). We excluded uncommon variants (MAF<0.05) from the analysis because of their low imputation accuracy in the GWAS (93% of uncommon variants in *NFKBIE* and 76% in *RTKN2* exhibited *Rsq* <0.6). There is neither structural variation (>1 kbps) nor indels (100 bps to 1 kbs) that are common in the population (frequency >0.01) in these loci. To evaluate the *cis*-regulatory potential of sequences around the SNPs *in silico*, we used the regulatory potential (RP) score [21,22]. This score was calculated based on the extent of sequence conservation among species or similarity with known regulatory motifs. We selected SNPs from the genomic elements with an RP score >0.1 (Step 3a). Subsequently, we selected SNPs from sites of transcriptional regulation as demonstrated by previous ChIP-seq studies (transcription factor binding sites [23,24] and histone modification sites [25,26]) or a DNase-seq study (DNase I hypersensitivity sites) [27] (Step 3b). Finally, these SNPs with regulatory potential were further screened out by the disease-association status ($P < 0.05$) using an imputed GWAS dataset (Step 4). As a consequence, we selected 14 SNPs in *NFKBIE* and 10 SNPs in *RTKN2* that had regulatory potential predicted *in silico*.

To further investigate the regulatory potential of the SNPs, we evaluated 31-bp sequences around the SNPs by *in vitro* assays. First, we examined their ability to bind nuclear proteins by EMSAs (Step 5a) using nuclear extracts from lymphoblastoid B cells (PSC cells) and Jurkat cells. Of the 24 SNPs examined, nine

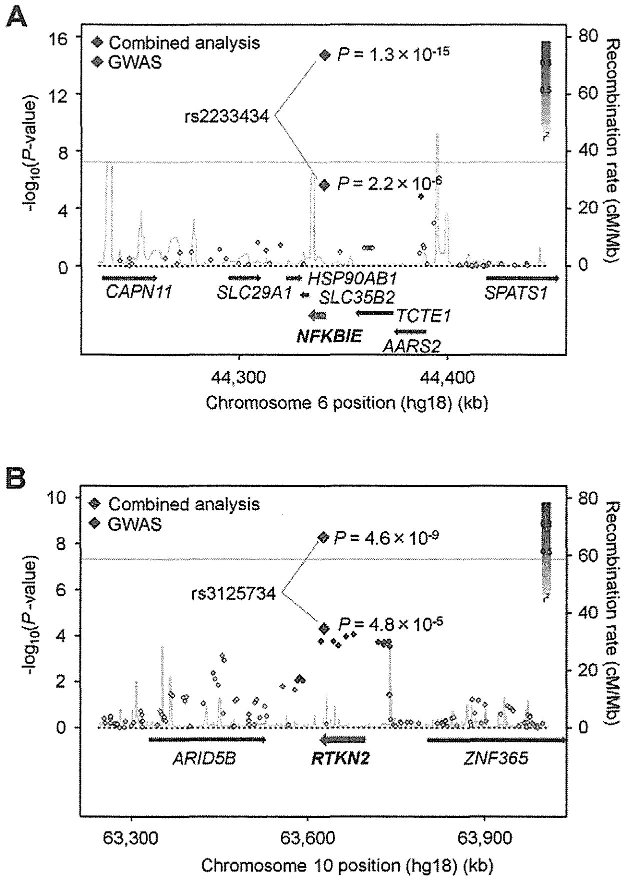


Figure 1. Association plots of *NFKBIE* and *RTKN2* regions. The diamonds represent the -$\log_{10}$ of the Cochran-Armitage trend *P*-values. Large diamonds show landmark SNPs in *NFKBIE* (rs2233434: A) and *RTKN2* (rs3125734: B). Red: GWAS, Blue: combined analysis. Red colors of each SNP indicate its r^2 with landmark SNP. Gray lines indicate the genome-wide significance threshold ($P<5\times10^{-8}$). For each plot, the -$\log_{10}$ *P*-values (y-axis) of the SNPs are presented according to their chromosomal positions (x-axis). Physical positions are based on NCBI build 36.3 of the human genome. Genetic recombination rates, estimated using the 1000 Genome Project (JPT and CHB), are represented by the blue line.
doi:10.1371/journal.pgen.1002949.g001

SNPs displayed allelic differences, implying differential potential of transcriptional activity between these alleles (Figure 5A and Figure S5). We then evaluated the enhancing or repressing activity of the sequences by luciferase reporter assays (Step 5b). We cloned them into the pGL4.24 vector, which has minimal promoter activity, and transfected these constructs into HEK293A cells (for *NFKBIE* and *RTKN2*), lymphoblastoid B cells (for *NFKBIE*), and Jurkat cells (for *RTKN2*). Among the three SNPs examined in *NFKBIE*, the risk allele of rs2233424 (located −396 bps from the 5′ end) displayed stronger repression activity (Figure 2A and Figure 5B) than that of the non-risk allele. Among the six SNPs in *RTKN2*, the risk alleles of rs12248974 (approximately 10 kb from the 3′ end) and rs61852964 (−215 bps from the 5′ end) showed higher enhancing activity compared with the non-risk alleles (Figure 2A and Figure 5B). These results corresponded to the results of ASTQ analyses (Figure 3E, 3F). Other SNPs showed no allelic differences or had the opposite trend of transcriptional activity in the risk allele compared to the results of ASTQ analysis (Figure S6).

To confirm the regulatory potential of these SNPs, we investigated the correlation between genotypes and gene expression levels in lymphocytes utilizing the data from the previous eQTL studies. We evaluated the expression of *RTKN2* in primary T cells from Western European individuals by using Genevar software [28,29]. Though *NFKBIE* is also expressed in primary T cells, the genotypes of rs2233424 are not available. We thus evaluated gene expression data of lymphoblastoid B-cell lines obtained from HapMap individuals (Japanese (JPT) + Han Chinese in Beijing (CHB), European (CEU), and African (YRI)) [30,31] instead. The *NFKBIE* expression level decreased with the number of risk alleles of rs2233424 ($R = -0.18$, $P = 0.020$), and the *RTKN2* expression levels increased with that of rs1432411 (a proxy for rs12248974, $r^2 = 0.97$) ($R = 0.27$, $P = 0.018$) (Figure 5C), corresponding to the results of the *in vitro* assays. The data for rs61852964 in *RTKN2* was not available. Among the SNPs that displayed opposite transcriptional activities in the reporter assays compared to the results of ASTQ, the data for rs2233434, rs77986492, and rs3852694 (a proxy for rs1864836, $r^2 = 1.0$) were available (Figure S7 and S8). These SNPs displayed the opposite direction of the correlation trend as compared to the results of reporter assays, but parallel to ASTQ, implying that the regulatory effects observed in the *in vitro* assays were cancelled out by the effects of other regulatory variants on the same haplotype *in vivo*.

Finally, we validated the associations of these regulatory (r)SNPs observed in the imputed GWAS dataset. We directly genotyped them by TaqMan assay and confirmed significant associations (Table S8). As the candidate causal variants (nsSNPs and rSNPs) and the landmark SNPs of GWAS were in strong LD at each locus (Figure 2A, 2B), we evaluated the independent effect of each SNP by haplotype analysis in both loci (Table S9 and S10) and the conditional logistic regression analysis in *RTKN2* (Table S11). The conditional analysis was not performed in *NFKBIE* because three candidate causal variants were in strong LD ($r^2>0.9$). However, the analyses for these two loci did not demonstrate any evidence of primary or independent effects across the candidate causal variants, and it remains a possibility that all of the functional variants were involved in the pathogenesis. In addition, although the landmark nsSNP (rs3125734) in *RTKN2* did not display any influence on NF-κB activity in our *in vitro* assays, rs3125734 might influence functions of *RTKN2* other than those in the NF-κB pathway; alternatively, it is still possible that rs3125734 tags the effects of other unknown variants, such as rare variants, in addition to the other two rSNPs (rs12248974 and rs61852964).

Discussion

In the present study, we performed a replication study of our previously reported GWAS and identified variants in *NFKBIE* and *RTKN2* loci that were associated with RA susceptibility. The associations of *NFKBIE* and *RTKN2* loci have not been reported in other populations with genome-wide significance. However, rs2233434 in *NFKBIE* showed a suggestive association (589 cases vs. 1,472 controls, $P = 0.0099$, OR = 1.57, 95% CI = 1.11–2.21) in a previous meta-analysis in European populations [32]. The weak association signal in Europeans may be partially due to the lower frequency of the risk allele (0.04 in Europeans compared to 0.22 in Japanese). On the other hand, the association of rs3125734 in *RTKN2* was not observed in a GWAS meta-analysis of European populations (cases 5,539 vs. controls 20,169, $P = 0.11$, OR = 1.04, 95% CI = 0.99–1.09). As the association of *RTKN2* locus was also implicated in Graves' disease in a Han Chinese population [33], the association in *RTKN2* locus may be unique to Asian populations.

To find the disease causal variants in disease-associated loci, target re-sequencing and variant genotyping with a large sample set followed by conditional association analysis examining the