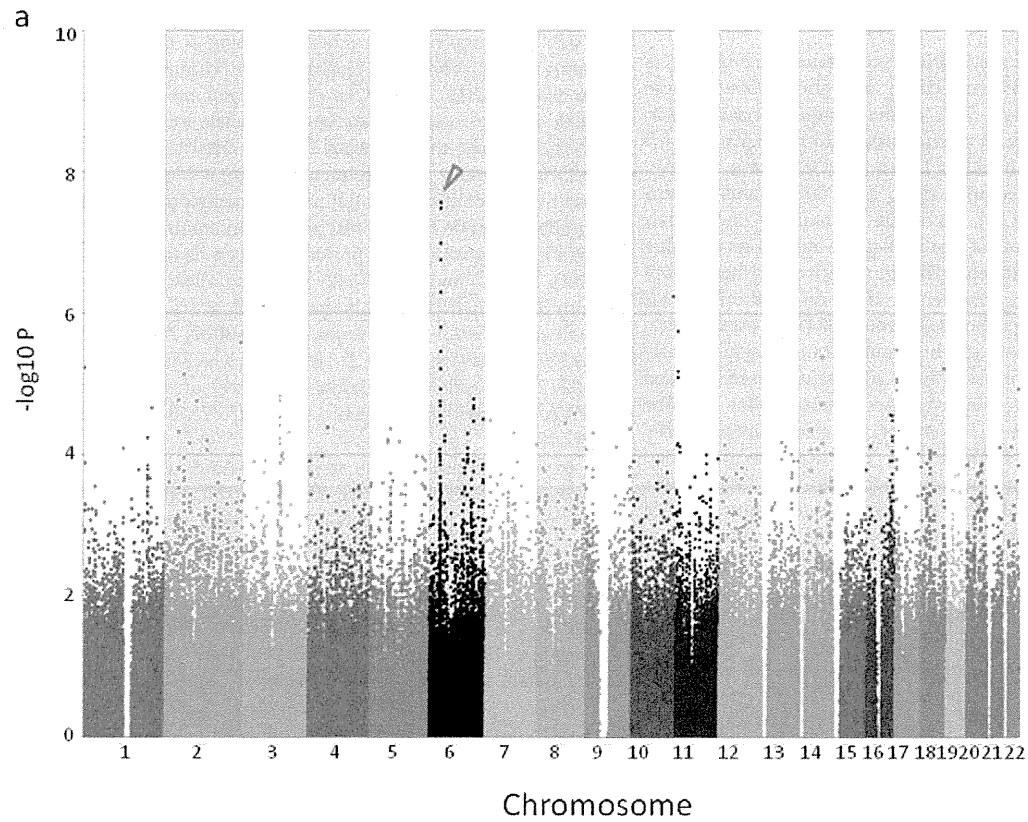
a
-log10 P
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Chromosome

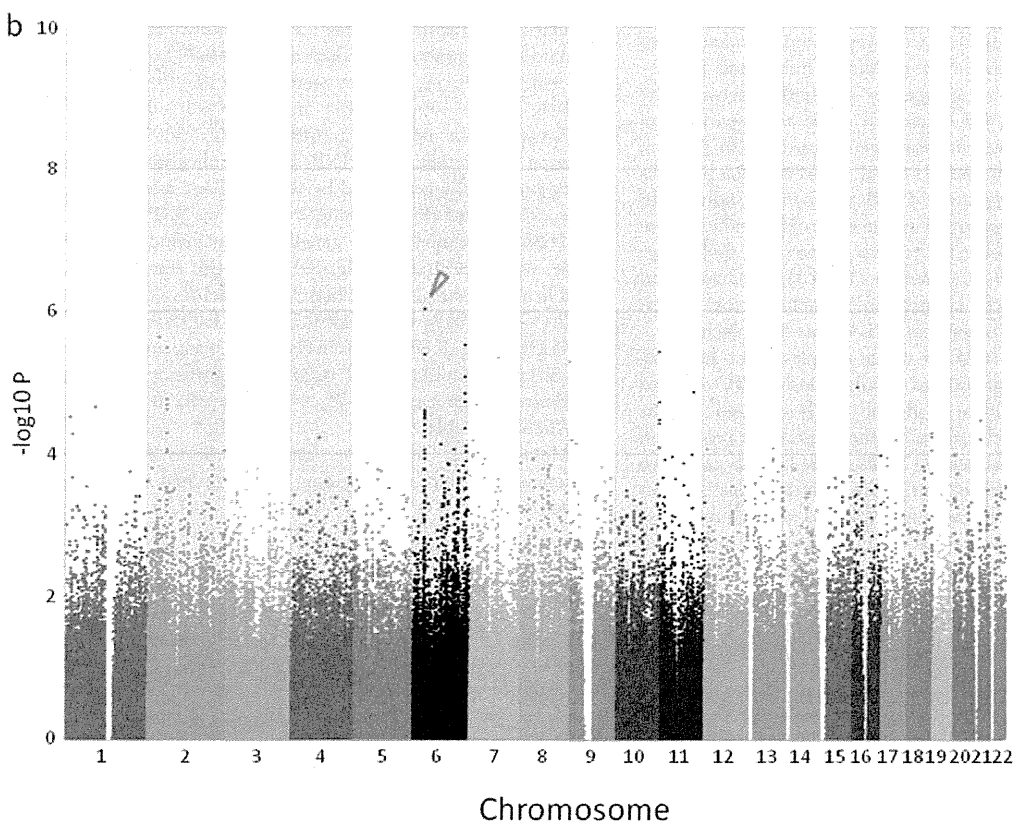
b
-log10 P
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Chromosome

Figure 1. Results of genome-wide association studies. a) HBV carriers and healthy controls, and b) HBV carriers and HBV-resolved individuals were compared. *P* values were calculated by chi-squared test for allele frequencies. Dots with arrows on chromosome 6 show strong associations with protective effects against persistent HB infection and with HBV clearance.
doi:10.1371/journal.pone.0039175.g001

Clearance of Hepatitis B virus in Japanese and Korean Individuals

We also conducted a GWAS to identify the host genetic factors related to clearance of HBV in the above 181 Japanese HBV carriers and 185 Japanese HBV-resolved individuals using a genome-wide SNP typing array (Affymetrix Genome-Wide Human SNP Array 6.0 for 900 K SNPs). The same two SNPs (rs3077 and rs9277542) showed strong associations in the allele frequency model ($P=9.24\times10^{-7}$ and $P=3.15\times10^{-5}$) with clearance of HBV (Figure 1b).

The above 32 SNPs, including the two associated SNPs (rs3077 and rs9277542), were selected for a replication study in two independent sets of HBV carriers and HBV resolved individuals (replication-1:256 Japanese HBV carriers and 150 Japanese HBV resolved individuals; and replication-2:344 Korean HBV carriers and 106 Korean HBV resolved individuals; Table 1). All 32 SNPs were genotyped using the DigiTag2 assay and 29 of 32 SNPs were successfully genotyped (Table S3). The associations of the original SNPs were replicated in both replication sets [replication-1 (Japanese): rs3077, $P=3.32\times10^{-2}$, OR $=0.72$ and rs9277542, $P=1.25\times10^{-2}$, OR $=0.68$; replication-2 (Korean): rs3077, $P=2.35\times10^{-7}$, OR $=0.41$ and rs9277542, $P=4.97\times10^{-6}$, OR $=0.46$; Table 3]. Meta-analysis using random effects model showed $P_{meta}=1.56\times10^{-4}$ for rs3077 (OR $=0.51$, 95% CI $=0.36$–0.72), and 5.91×10^{-7} for rs9277542 (OR $=0.55$, 95% CI $=0.43$–0.69). While there was evidence of heterogeneity between these studies for rs3077 ($P_{het}=0.03$) and no evidence for rs9277542 ($P_{het}=0.19$), significant associations with HBV clearance were observed with Mantel-Haenszel $P_{meta}=3.28\times10^{-12}$ for rs3077 and 1.42×10^{-10} for rs9277542, when using CMH fixed-effects model. Among the remaining 27 SNPs in the replication study, two SNPs (rs9276431 and rs7768538), located in a genetic region including *HLA-DQ* gene, were marginally replicated in the two sets of HBV carriers and HBV resolved individuals with Mantel-Haenszel *P* values of 2.10×10^{-5} (OR $=0.59$) and 1.10×10^{-5} (OR $=0.56$), respectively (Table S3), when using CMH fixed-effect model. Due to the existing heterogeneity among three groups (GWAS, Replication-1 and Replication-2) ($P_{het}=0.03$ for rs9276431 and 0.04 for rs7768538), weak associations were observed with $P_{meta}=0.03$ for rs9276431 and 0.02 for rs7768538 by the random effects model meta-analysis.

Meta-analysis across 6 independent studies, including 5 additional published data, showed $P_{meta}=1.48\times10^{-9}$, OR $=0.60$ for rs3077, $P_{meta}=1.08\times10^{-17}$, OR $=0.66$ for rs9277535 and $P_{meta}=5.14\times10^{-5}$, OR $=0.55$ for rs9277542 (Table S4). As shown in Table S4, the OR for the rs9277535 and rs9277542 were similar among the 6 independent studies, and heterogeneity was negligible ($P_{het}=0.03$ for rs9277535 and 0.14 for rs9277542). However, significant level of heterogeneity for rs3077 was observed with $P_{het}=9.57\times10^{-6}$ across 5 independent studies, including our study.

Table 1. Number of study samples.

		GWAS	Replication-1	Replication-2
population		**Japanese**	**Japanese**	**Korean**
HBV carriers	Total	**181**	**256**	**344**
	IC	20	94	–
	CH	67	101	177
	LC	3	10	–
	HCC	91	51	167
Healthy controls		**184**	**236**	**151**
Resolved individuals		**185**	**150**	**106**

Abbreviation: IC, Inactive Carrier; CH, Chronic Hepatitis; LC, Liver Cirrhosis; HCC, Hepatocellular Carcinoma.
doi:10.1371/journal.pone.0039175.t001

URLs

The results of the present GWAS are registered at a public database: https://gwas.lifesciencedb.jp/cgi-bin/gwasdb/gwas_top.cgi.

Discussion

The recent genome-wide association study showed that the SNPs located in a genetic region including *HLA-DPA1* and *HLA-DPB1* genes were associated with chronic HBV infection in the Japanese and Thai population [10,11]. In this study, we confirmed a significant association between SNPs (rs3077 and rs9277542) located in the same genetic region as *HLA-DPA1* and *HLA-DPB1* and protective effects against CHB in Korean and Japanese individuals. Mata-analysis using the random effects model across 6 independent studies including our study suggested that, widely in East Asian populations, variants in antigen binding sites of *HLA-DP* contribute to protective effects against persistent HBV infection (Table S2).

On GWAS and replication analysis with Japanese and Korean individuals, we identified associations between the same SNPs (rs3077 and rs9277542) in the *HLA-DPA1* and *HLA-DPB1* genes and HBV clearance; however, no new candidate SNPs from the GWAS were detected on replication analysis (Table S3). When the data of reference#18 was excluded from the meta-analysis across 6 independent studies, heterogeneity among 4 studies was estimated to be $P_{het}=0.15$ and significant association of rs3077 with HBV clearance was observed with $P_{meta}=5.88\times10^{-24}$, OR $=0.56$ (Table S4). In our study, a negligible level of heterogeneity for rs3077 was also observed ($P_{het}=0.03$) on meta-analysis by adding replication-1 (Table 3). Despite the heterogeneity in replication-1, a marginal association was observed for rs3077 with the same downward trend in the odds ratio ($P=3.32\times10^{-2}$, OR $=0.72$). Moreover, meta-analysis using GWAS and replication-2 showed significant association of $P_{meta}=1.89\times10^{-12}$, OR $=0.43$ for rs3077 with no evidence of heterogeneity ($P_{het}=0.75$). Although the reason why heterogeneity was observed in replication-1 is unclear, one possible reason is the clinical heterogeneity due to different kits being used for antibody testing. The associations of *HLA-DPA1/−DPB1* with CHB and HBV clearance showed the same level of significance in the comparison of HBV patients with HBV resolved individuals (OR $=0.43$ for rs3077 and 0.49 for rs9277542) as the one with healthy controls (OR $=0.46$ for rs3077 and 0.50 for rs9277542), when the replication-1 was excluded in the analysis (Table 2 and Table 3). The results of meta-analysis across 6 independent studies including our study also showed the same or slightly weaker associations in the

Table 2. Results of replication study for protective effects against CHB.

	Position			MAF[a]	Allele	Stage	HBV carriers			Healthy controls				OR[b]		
dbSNP rsID	Chr	Buld 36.3	Nearest Gene	(allele)	(1/2)	(population)	11	12	22	11	12	22	HWEp	95% CI	*P*-value[c]	P_{het}[d]
rs3077	6	33141000	HLA-DPA1	0.44	T/C	GWAS	13	51	117	28	88	67	0.919	0.42	1.14×10^{-7}	
				(T)		(Japanese)	(7.2)	(28.2)	(64.6)	(15.3)	(48.1)	(36.6)		(0.30–0.58)		
						Replication-1	26	95	134	46	125	65	0.309	0.48	2.70×10^{-8}	
						(Japanese)	(10.2)	(37.3)	(52.5)	(19.5)	(53.0)	(27.5)		(0.37–0.62)		
						Replication-2	23	81	111	31	74	40	0.767	0.47	2.08×10^{-6}	
						(Korean)	(10.7)	(37.7)	(51.6)	(21.4)	(51.0)	(27.6)		(0.35–0.65)		
						Meta-analysis[e]								0.46	4.40×10^{-19}	0.80
														(0.39–0.54)		
rs9277542	6	33163225	HLA-DPB1	0.45	T/C	GWAS	18	53	110	29	102	52	0.073	0.42	5.32×10^{-8}	
				(T)		(Japanese)	(9.9)	(29.3)	(60.8)	(15.8)	(55.7)	(28.4)		(0.31–0.58)		
						Replication-1	30	106	118	54	114	67	0.681	0.54	3.33×10^{-6}	
						(Japanese)	(11.8)	(41.7)	(46.5)	(23.0)	(48.5)	(28.5)		(0.42–0.70)		
						Replication-2	30	87	94	35	72	36	0.933	0.54	8.29×10^{-5}	
						(Korean)	(14.2)	(41.2)	(44.5)	(24.5)	(50.3)	(25.2)		(0.40–0.74)		
						Meta-analysis[e]								0.50	1.28×10^{-15}	0.40
														(0.43–0.60)		

[a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19).
[b]Odds ratio of minor allele from two-by-two allele frequency table.
[c]P value of Pearson's chi-square test for allelic model.
[d]Heterogeneity was tested using general variance-based method.
[e]Meta-analysis was tested using the random effects model.
doi:10.1371/journal.pone.0039175.t002

comparison of HBV patients with HBV resolved individuals (OR = 0.56 for rs3077, 0.66 for rs9277535 and 0.55 for rs9277542) than in the one with healthy controls (OR = 0.55 for rs3077, 0.61 for rs9277535 and 0.51 for rs9277542), which was the opposite result as we expected (Table S2 and Table S4). These results may suggest that other unknown immune system(s) exist to eliminate the HBV in the HBV resolved individuals.

Among the HLA class II loci (*HLA-DPA1*, *HLA-DPB1* and *HLA-DQB2*), which were associated with CHB and HBV clearance, a weak linkage disequilibrium ($r^2<0.1$) was observed between *HLA-DQB2* locus and *HLA-DPA1/−DPB1* loci in Japanese and Korean populations (Figure S2). We also found that similar linkage disequilibrium blocks (r^2) were observed among three subgroups (HBV carriers, HBV resolved individuals and Healthy controls). Moreover, logistic regression analysis of *HLA-DP* (rs3077 and rs92775542) with use of *HLA-DQ* (rs9276431 and rs768538) as covariates showed that the same level of significant associations of *HLA-DP* with CHB and HBV clearance as shown in the single-point association analysis, while no associations of *HLA-DQ* with $P_{log}>0.05$ were detected both in Japanese and in Korean (Table S5). These results show that *HLA-DP* is the main genetic factor for susceptibility to CHB and HBV clearance, and the associations of *HLA-DQB2* would result from linkage disequilibrium of *HLA-DP*A1/−DPB1.

In this study, we confirmed the significant associations between *HLA-DPA1* and *HLA-DPB1*, and protective effects against CHB and HBV clearance in Japanese and Korean individuals. These results suggest that the associations between the *HLA-DP* locus, CHB and HBV clearance are widely replicated in East Asian populations, including Chinese, Thai, Japanese and Korean individuals; however, there have been no similar GWAS performed in Caucasian and African populations. Moreover, there were no significant SNPs associated with HCC development in this study, thus suggesting that it is necessary to increase the sample size. To clarify the pathogenesis of CHB or the mechanisms of HBV clearance, further studies are necessary, including a functional study of the *HLA-DP* molecule, identification of novel host genetic factors other than *HLA-DP*, and variation analysis of HBV.

Materials and Methods

Ethics Statement

All study protocols conform to the relevant ethical guidelines, as reflected in the *a priori* approval by the ethics committees of all participating universities and hospitals. The written informed consent was obtained from each patient who participated in this study and all samples were anonymized.

Genomic DNA Samples and Clinical Data

All of the 1,793 Japanese and Korean samples, including individuals with CHB, healthy controls and HBV-resolved individuals (HBsAg-negative and anti-HBc-positive), were collected at 20 multi-center hospitals (liver units with hepatologists) throughout Japan and Korea. The 19 hospitals in Japan were grouped into the following 8 areas: Hokkaido area (Hokkaido University Hospital, Teine Keijinkai Hospital), Tohoku area (Iwate Medical University Hospital), Kanto area (Musashino Red Cross Hospital, Saitama Medical University, Kitasato University Hospital, University of Tokyo), Koshin area (Shinshu University Hospital, Kanazawa University Hospital), Tokai area (Nagoya City University Hospital, Nagoya Daini Red Cross Hospital), Kinki area (Kyoto Prefectural University of Medicine Hospital, National Hospital Organization Osaka National Hospital, Osaka

Table 3. Results of replication study for clearance of hepatitis B virus.

		Position	MAF[a]	Allele	Stage	HBV carriers			Resolved individuals			OR[b]		
dbSNP rsID	Chr	Buld 36.3 Nearest Gene	(allele)	(1/2)	(population)	11	12	22	11	12	22	95% CI	*P*-value[c]	P_{het}[d]
rs3077	6	33141000 HLA-DPA1	0.44	T/C	GWAS	13	51	117	29	82	74	0.44	9.24×10^{-7}	
			(T)		(Japanese)	(7.2)	(28.2)	(64.6)	(15.7)	(44.3)	(40.0)	(0.32–0.61)		
					Replication-1	26	95	134	20	64	60	0.72	3.32×10^{-2}	
					(Japanese)	(10.2)	(37.3)	(52.5)	(13.9)	(44.4)	(41.7)	(0.53–0.97)		
					Replication-2	23	81	111	29	48	28	0.41	2.35×10^{-7}	
					(Korean)	(10.7)	(37.7)	(51.6)	(27.6)	(45.7)	(26.7)	(0.29–0.58)		
					Meta-analysis[e]							0.51	1.56×10^{-4}	0.03
												(0.36–0.72)		
					Meta-analysis[e]							0.43	1.89×10^{-12}	0.75
					(GWAS+replication-2)							(0.34–0.54)		
rs9277542	6	33163225 HLA-DPB1	0.45	T/C	GWAS	18	53	110	28	88	69	0.51	3.15×10^{-5}	
			(T)		(Japanese)	(9.9)	(29.3)	(60.8)	(15.1)	(47.6)	(37.3)	(0.37–0.70)		
					Replication-1	30	106	118	28	62	52	0.68	1.25×10^{-2}	
					(Japanese)	(11.8)	(41.7)	(46.5)	(19.7)	(43.7)	(36.6)	(0.51–0.92)		
					Replication-2	30	87	94	30	53	22	0.46	4.97×10^{-6}	
					(Korean)	(14.2)	(41.2)	(44.5)	(28.6)	(50.5)	(21.0)	(0.33–0.64)		
					Meta-analysis[e]							0.55	5.91×10^{-7}	0.19
												(0.43–0.69)		
					Meta-analysis[e]							0.49	9.69×10^{-10}	0.65
					(GWAS+replication-2)							(0.39–0.61)		

[a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19).
[b]Odds ratio of minor allele from two-by-two allele frequency table.
[c]P value of Pearson's chi-square test for allelic model.
[d]Heterogeneity was tested using general variance-based method.
[e]Meta-analysis was tested using the random effects model.
doi:10.1371/journal.pone.0039175.t003

City University), Chugoku/Shikoku area (Tottori University Hospital, Ehime University Hospital, Yamaguchi University Hospital, Kawasaki Medical College Hospital) and Kyushu area (Kurume University Hospital). Korean samples were collected at Yonsei University College of Medicine.

HBV status was measured based on serological results for HBsAg and anti-HBc with a fully automated chemiluminescent enzyme immunoassay system (Abbott ARCHITECT; Abbott Japan, Tokyo, Japan, or LUMIPULSE f or G1200; Fujirebio, Inc., Tokyo, Japan). For clinical staging, inactive carrier (IC) state was defined by the presence of HBsAg with normal ALT levels over 1 year (examined at least four times at 3-month intervals) and without evidence of portal hypertension. Chronic hepatitis (CH) was defined by elevated ALT levels (>1.5 times the upper limit of normal [35 IU/L]) persisting over 6 months (at least by 3 bimonthly tests). Liver cirrhosis (LC) was diagnosed principally by ultrasonography (coarse liver architecture, nodular liver surface, blunt liver edges and hypersplenism), platelet counts $<100{,}000/cm^3$, or a combination thereof. Histological confirmation by fine-needle biopsy of the liver was performed as required. Hepatocellular carcinoma (HCC) was diagnosed by ultrasonography, computerized tomography, magnetic resonance imaging, angiography, tumor biopsy or a combination thereof.

The Japanese control samples from HBV-resolved subjects (HBsAg-negative and anti-HBc-positive) at Nagoya City University-affiliated healthcare center were used by comprehensive agreement (anonymization in an unlinkable manner) in this study. Some of the unrelated Japanese healthy controls were obtained from the Japan Health Science Research Resources Bank (Osaka, Japan). One microgram of purified genomic DNA was dissolved in 100 µl of TE buffer (pH 8.0) (Wako, Osaka, Japan), followed by storage at −20°C until use.

SNP Genotyping and Data Cleaning

For GWAS, we genotyped a total of 550 individuals, including 181 Japanese HBV carriers, 184 Japanese healthy controls and 185 spontaneously HBV-resolved Japanese individuals (HBsAg-negative and anti-HBc-positive), using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Inc., Santa Clara, CA), in accordance with the manufacturer's instructions. The average QC call rate for 550 samples reached 98.47% (95.00–99.92%), which had an average sample call rate of 98.91% (93.55–99.74%) by determining the genotype calls of over 900 K SNPs using the Genotyping Console v4.1 software (with Birdseed v1 algorithm) provided by the manufacturer [19]. We then applied the following thresholds for SNP quality control in data cleaning: SNP call rate ≥95% and MAF ≥1% for three groups (HBV carriers, healthy controls and HBV-resolved individuals), and HWE *P*-value ≥0.001 for healthy controls [20]. Here, SNP call rate is defined for each SNP as the number of successfully genotyped samples divided by the number of total samples genotyped. A total of 597,789 SNPs and 590,278 SNPs on autosomal chromosomes

passed the quality control filters in the genome-wide association analysis using HBV carriers and healthy controls, and using HBV carriers and HBV-resolved individuals, respectively (Figure 1). All cluster plots for the SNPs showing $P<0.0001$ on association analyses in the allele frequency model were confirmed by visual inspection, and SNPs with ambiguous cluster plots were excluded.

In the following replication stage, we selected a set of 32 SNPs with $P<0.0001$ in the GWAS using HBV carriers and HBV-resolved individuals. SNP genotyping in two independent sets of 256 Japanese HBV carriers, 236 Japanese healthy controls and 150 Japanese HBV-resolved individuals (Table 1, replication-1), and 344 Korean HBV carriers, 151 Korean healthy controls and 106 Korean HBV-resolved individuals (Table 1, replication-2) was completed for the selected 32 SNPs using the DigiTag2 assay [21,22] and custom TaqMan SNP Genotyping Assays (Applied Biosystems, Foster City, CA) on the LightCycler 480 Real-Time PCR System (Roche, Mannheim, Germany).

Statistical Analysis

The observed associations between SNPs and the protective effects on chronic hepatitis B or clearance of hepatitis virus B were assessed by chi-squared test with a two-by-two contingency table in allele frequency model. SNPs on chromosome X were removed because gender was not matched among HBV carriers, healthy controls and HBV-resolved individuals. A total of 597,789 SNPs and 590,278 SNPs passed the quality control filters in the GWAS stage; therefore, significance levels after Bonferroni correction for multiple testing were $P=8.36\times10^{-8}$ (0.05/597,789) and $P=8.47\times10^{-8}$ (0.05/590,278), respectively. For the replication study, 29 of 32 SNPs were successfully genotyped; therefore, we applied $P=0.0017$ (0.05/29) as a significance level, and none of the 29 markers genotyped in the replication stage showed deviations from the Hardy-Weinberg equilibrium in healthy controls ($P>0.01$).

The genetic inflation factor λ was estimated by applying the Cochrane-Armitage test on all SNPs and was found to be 1.056 and 1.030 in the GWAS using HBV carriers and healthy controls, and using HBV carriers and HBV-resolved individuals, respectively (Figure S3). These results suggest that the population substructure should not have any substantial effect on statistical analysis. In addition, the principal component analysis in a total of 550 individuals in the GWAS stage together with the HapMap samples also revealed that the effect of population stratification was negligible (Figure S4).

Based on the genotype data of a total of 1,793 samples including 1,192 Japanese samples and 601 Korean samples in both GWAS and replication stages, haplotype blocks were estimated using the Gabriel's algorithm using the Haploview software (v4.2) (Figure S2). In the logistic regression analysis, two SNPs (rs9276431 and rs7768538) within the HLA-DQ locus were individually involved as a covariate (Table S5). Statistical analyses were performed using the SNP & Variation Suite 7 software (Golden Helix, MT, USA).

Supporting Information

Figure S1 GWAS using samples from HBV carriers with LC or HCC, and HBV carriers without LC and HCC. *P* values were calculated using chi-squared test for allele frequencies.
(PPTX)

Figure S2 Estimation of linkage disequilibrium blocks in HBV patients, HBV resolved individuals and healthy controls in Japanese and Korean. The LD blocks (r^2) were analyzed using the Gabriel's algorithm.
(PPTX)

Figure S3 Quantile-quantile plot for test statistics (allele-based chi-squared tests) for GWAS results. Dots represent P values of each SNP that passed the quality control filters. Inflation factor λ was estimated to be: a) 1.056 in the analysis with HBV carriers and healthy controls; and b) 1.030 with HBV carriers and HBV-resolved individuals.
(PPTX)

Figure S4 Principal component analysis on a total of 550 individuals in GWAS, together with HapMap samples (CEU, YRI and JPT).
(PPTX)

Table S1 Results for 29 SNPs selected in replication study using samples of HBV carriers and healthy controls. [a]*P* values by chi-squared test for allelic model. [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]Meta-analysis was tested using additive, two-tailed CMH fixed-effects model.
(XLSX)

Table S2 Results of meta-analysis for protective effects against persistent HB infection across 6 independent studies, including this study. [a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19). [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]*P* value of Pearson's chi-squared test for allele model. [d]Heterogeneity was tested using general variance-based method. [e]Meta-analysis was tested using the random effects model.
(XLSX)

Table S3 Results for 29 SNPs selected in replication study using samples from HBV carriers and HBV-resolved individuals. [a]*P* values by chi-squared test for allelic model. [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]Meta-analysis was tested using additive, two-tailed CMH fixed-effects model.
(XLSX)

Table S4 Results of meta-analysis for clearance of HBV across 6 independent studies, including this study. [a]Minor allele frequency and minor allele in 198 healthy Japanese (ref#19). [b]Odds ratio of minor allele from two-by-two allele frequency table. [c]*P* value of Pearson's chi-squared test for allele model. [d]Heterogeneity was tested using general variance-based method. [e]Meta-analysis was tested using the random effects model.
(XLSX)

Table S5 Logistic regression analysis of *HLA-DP* (rs3077 and rs9277542) and *HLA-DQ* (rs9276431 and rs7768538) with susceptibility to CHB and HBV clearance using the *HLA-DQ* genotypes individually as a covariate.
(XLSX)

Acknowledgments

We thank all the patients and families who contributed to the study and Ms. Yasuka Uehara-Shibata and Ms. Yoshimi Ishibashi for technical assistance.

Author Contributions

Conceived and designed the experiments: NN HS YT. Performed the experiments: HS Y. Mawatari M. Sageshima YO. Analyzed the data: NN MK AK. Contributed reagents/materials/analysis tools: KM M. Sugiyama SHA JYP SH JHK KS M. Kurosaki YA SM MW ET MH SK EO YI EM AT Y. Murawaki YH IS M. Korenaga KH TI NI KHH YT MM. Wrote the paper: NN M. Kawashima YT KT MM.

References

1. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO (2002) Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 83: 2059–2073.
2. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok ASF (2007) Management of hepatitis B: Summary of a clinical research workshop. Hepatology 45: 1056–1075.
3. Yokosuka O, Kurosaki M, Imazeki F, Arase Y, Tanaka Y, et al. (2011) Management of hepatitis B: Consensus of the Japan society of Hepatology 2009. Hepatol Res 41: 1–21.
4. Tada H, Uga N, Fuse Y, Shimizu M, Nemoto Y, et al. (1992) Prevention of perinatal transmission of hepatitis B virus carrier state. Acta Paediatr Jpn 34: 656–659.
5. Stevens CE, Toy PT, Taylor PE, Lee T, Yip HY (1992) Prospects for control of hepatitis B virus infection: implications of childhood vaccination and long-term protection. Pediatrics 90: 170–173.
6. Szmuness W (1979) Large-scale efficacy trials of hepatitis B vaccines in the USA: baseline data and protocols. J Med Virol 4: 327–340.
7. Kwon H, Lok AS (2011) Hepatitis B therapy. Nat Rev Gastroenterol Hepatol 8: 275–284.
8. Kobayashi M, Ikeda K, Arase Y, Suzuki F, Akuta N, et al. (2008) Change of Hepatitis B virus genotypes in acute and chronic infections in Japan. J Med Virol 80: 1880–1884.
9. Yano K, Tamada Y, Yatsuhashi H, Komori A, Abiru S, et al. (2010) Dynamic epidemiology of acute viral hepatitis in Japan. Intervirology 53: 70–75.
10. Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, et al. (2009) A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat Genet 41: 591–595.
11. Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, et al. (2011) A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. Hum Mol Genet 20: 3884–3892.
12. Li J, Yang D, He Y, Wang M, Wen Z, et al. (2011) Associations of HLA-DP variants with hepatitis B virus infection in southern and northern Han Chinese populations: a multicenter case-control study. PLoS ONE 6: e24221.
13. Guo X, Zhang Y, Li J, Ma J, Wei Z, et al. (2011) Strong influence of human leukocyte antigen (HLA)-DP gene variants on development of persistent chronic hepatitis B virus carriers in the Han Chinese population. Hepatol 53: 422–428.
14. Thursz MR, Kwiatkowski D, Allsopp CEM, Greenwood BM, Thomas HC, et al. (1995) Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N Engl J Med 332: 1065–1069.
15. Godkin A, Davenport M, Hill AVS (2005) Molecular analysis of HLA class II associations with hepatitis B virus clearance and vaccine nonresponsiveness. Hepatology 41: 1383–1390.
16. An P, Winkler C, Guan L, O'Brien SJ, Zeng Z, et al. (2011) A common HLA-DPA1 variant is a major determinant of hepatitis B virus clearance in Han Chinese. J Infect Dis 203: 943–947.
17. Wang L, Wu X-P, Zhang W, Zhu D-H, Wang Y, et al. (2011) Evaluation of genetic susceptibility loci for chronic hepatitis B in Chinese: two independent case-control study.
18. Hu L, Zhai X, Liu J, Chu M, Pan S, et al. (2011) Genetic variants in HLA-DP/DQ influence both hepatitis B virus clearance and Hepatocellular carcinoma development. Hepatology (in press).
19. Nishida N, Koike A, Tajima A, Ogasawara Y, Ishibashi Y, et al. (2008) Evaluating the performance of Affymetrix SNP Array 6.0 platform. BMC Genomics 9: 431.
20. Miyagawa T, Nishida N, Ohashi J, Kimura R, Fujimoto A, et al. (2008) Appropriate data cleaning methods for genome-wide association study. J Hum Genet 53: 886–893.
21. Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 364: 78–85.
22. Nishida N, Mawatari Y, Sageshima M, Tokunaga K (2012) Highly parallel and short-acting amplification with locus-specific primers to detect single nucleotide polymorphisms by the DigiTag2 assay. PLoS ONE 7: e29967.

Human Molecular Genetics, 2011, Vol. 20, No. 19 **3884–3892**
doi:10.1093/hmg/ddr301
Advance Access published on July 12, 2011

A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population

Hamdi Mbarek[1], Hidenori Ochi[2,3], Yuji Urabe[1], Vinod Kumar[1], Michiaki Kubo[2], Naoya Hosono[2], Atsushi Takahashi[2], Yoichiro Kamatani[1], Daiki Miki[2,3], Hiromi Abe[3], Tatsuhiko Tsunoda[2], Naoyuki Kamatani[2], Kazuaki Chayama[2,3], Yusuke Nakamura[1,2] and Koichi Matsuda[1,*]

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, [2]Center for Genomic Medicine, RIKEN, Kanagawa, Japan and [3]Division of Frontier Medical Science, Department of Medicine and Molecular Science, Programs for Biomedical Research, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima, Japan

Received February 23, 2011; Revised June 23, 2011; Accepted July 7, 2011

Hepatitis B virus (HBV) infection is a major health issue worldwide which may lead to hepatic dysfunction, liver cirrhosis and hepatocellular carcinoma. To identify host genetic factors that are associated with chronic hepatitis B (CHB) susceptibility, we previously conducted a two-stage genome-wide association study (GWAS) and identified the association of *HLA-DP* variants with CHB in Asians; however, only 179 cases and 934 controls were genotyped using genome-wide single nucleotide polymorphism (SNP) arrays. Here, we performed a second GWAS of 519 747 SNPs in 458 Japanese CHB cases and 2056 controls. After adjustment with the previously identified variants in the *HLA-DP* locus (rs9277535), we detected strong associations at 16 loci with *P*-value of $<5 \times 10^{-5}$. We analyzed these loci in three independent Japanese cohorts (2209 CHB cases and 4440 controls) and found significant association of two SNPs (rs2856718 and rs7453920) within the *HLA-DQ* locus (overall *P*-value of 5.98×10^{-28} and 3.99×10^{-37}). Association of CHB with SNPs rs2856718 and rs7453920 remains significant even after stratification with rs3077 and rs9277535, indicating independent effect of *HLA-DQ* variants on CHB susceptibility (*P*-value of 1.52×10^{-21} – 2.38×10^{-30}). Subsequent analyses revealed *DQA1*0102-DQB1*0604* and *DQA1*0101-DQB1*0501* [odds ratios (OR) =0.16, and 0.39, respectively] as protective haplotypes and *DQA1*0102-DQB1*0303* and *DQA1*0301-DQB1*0601* (OR = 19.03 and 5.02, respectively) as risk haplotypes. These findings indicated that variants in antigen-binding regions of *HLA-DP* and *HLA-DQ* contribute to the risk of persistent HBV infection.

INTRODUCTION

Hepatitis B virus (HBV) is the most common cause of infectious liver diseases, and about 400 million people are suffering from chronic viral infection worldwide. Routes of infection include vertical transmission during neonatal period and horizontal transmission in childhood (bites, lesions and sanitary habits) or adulthood (sexual contact, drug use and medical exposure). In Japan, most of the chronic hepatitis B (CHB) patients were infected through vertical transmission and become HBV carrier (1). Nearly 90% of the HBV carrier will clear HBV (negative for HBsAg and positive for HBc ab) during adolescence, and only 10% of the HBV carrier indicate persistent liver dysfunction and develop chronic hepatitis (2). CHB dramatically increases the risk to progress to liver cirrhosis and hepatocellular carcinoma over a period of several decades (3,4). Currently, CHB is a serious public health problem worldwide, however pathogenesis of HBV-related diseases still remains elusive.

*To whom correspondence should be addressed at: Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, 4-6-1, Shirokanedai, Minato, Tokyo 108-8639, Japan. Tel: +81 354495376; Fax: +81 354495123; Email: koichima@ims.u-tokyo.ac.jp

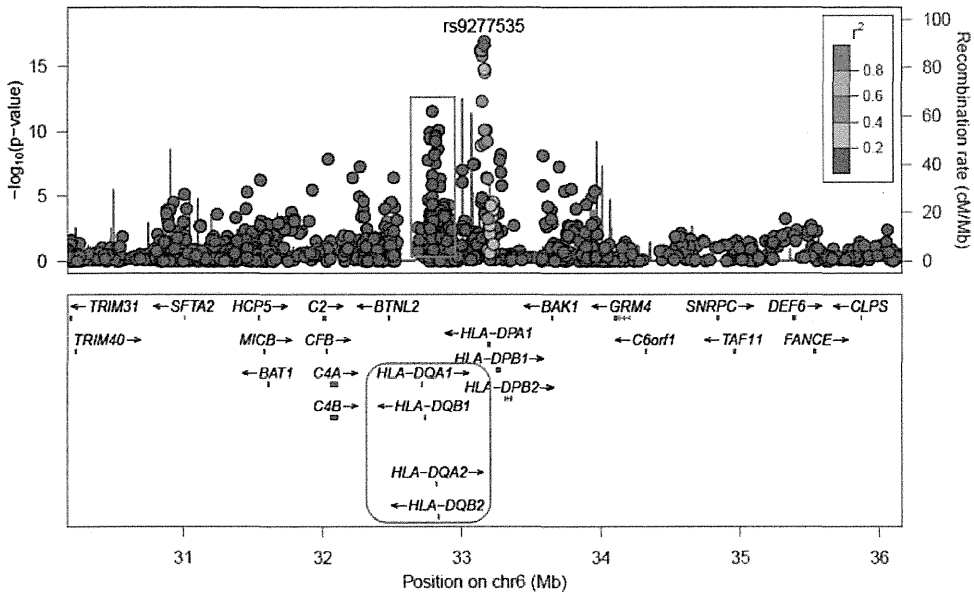

Figure 1. Signal of association with CHB in the *HLA* region of the GWAS stage. This figure shows the regional plots of the negative decadic logarithm trend *P*-values in a ~ 3000 kb window centered on the association peak, located at rs9277535 in *HLA-DPB1*. The top panel shows all SNPs in this region plotted according to the significance of their association with CHB and color coded according to their LD (r^2) with the most significant SNP, rs9277535 (see right corner of the plot). Vertical blue lines indicate local recombination rate. The bottom panel shows the genes in the region. The strongest signal on 6p21.32 localizes to *HLA-DP* genes and the second strongest signal localizes to *HLA-DQ* genes.

In addition to the viral and environmental factors, host genetic factors are considered to govern the pathology of disease development, progression or regression. Genetic epidemiological studies provide robust evidence that genetic variations contribute to progression from acute to chronic hepatitis (5). In 2009, our group conducted a genome-wide association study (GWAS) in the Asian population and identified a strong association of CHB with variants in the *HLA-DP* genes (6). In addition to our report, several association studies have suggested that genetic factors such as *HLA* (7–9), cytokines (10–12) and immune response-related genes (13–15) could influence the outcomes of HBV infection. However, these susceptibility loci were not identified in our previous study probably due to smaller sample size or smaller phenotypic effects of these loci. Here we conducted a second GWAS in the Japanese population to identify new susceptibility loci for CHB by increasing the number of samples in the screening stage from 179 cases and 934 controls to 458 case and 2056 controls.

RESULTS

We performed a two-stage GWAS followed by two independent replications as described in the Supplementary Material, Figure S1. In the GWAS stage, we genotyped 458 Japanese patients with CHB and 2056 control individuals using Illumina gene chip and obtained the genotyping results of 423 627 single nucleotide polymorphisms (SNPs) after quality control (QC). Examination of the quantile–quantile plots of the GWAS stage indicated no evidence for inflation of the test statistics, which could occur in the presence of population substructure ($\lambda = 1.028$) and also revealed an enrichment of significant *P*-values, suggesting the possible existence of candidates (Supplementary Material, Fig. S2A). The results of genome-wide association analysis are represented in Supplementary Material, Table S2, where a total of 34 SNPs in the major histocompatibility complex (MHC) region satisfied the genome-wide significance level ($P < 5.0 \times 10^{-8}$). We also found 54 SNPs (40 in the MHC region and 14 in the non-MHC region) with suggestive associations ($P < 5.0 \times 10^{-5}$) (Supplementary Material, Fig. S2B and Tables S2 and S3). We confirmed the most significant association at the *HLA-DP* locus as described in our previous report (rs9277535 and rs3077, $P = 3.72 \times 10^{-17}$ and 1.28×10^{-16}, respectively) (6) and found another significant peak around the *HLA-DQ* locus which is located ~300 kb telomeric to the *HLA-DP* locus (Fig. 1). To identify SNPs that are associated with CHB independently from *HLA-DP* SNPs, we conducted the association analysis after adjustment for a top SNP in the *HLA-DP* locus (rs9277535) using a logistic regression model (Fig. 2). As a result, five SNPs in the MHC region indicated suggestive associations ($P < 5.0 \times 10^{-5}$) even after stratification with rs9277535. Finally, 5 SNPs in the MHC region and 11 SNPs in the non-MHC region were selected for further analysis (Supplementary Material, Table S4).

Subsequently, we analyzed these 16 SNPs in the first replication set consisting of 606 cases and 2022 controls and found 2 SNPs within the MHC region [rs2856718, $P = 1.6 \times 10^{-5}$, odds ratios (OR) =1.33; rs7453920, $P = 5.72 \times 10^{-4}$, OR = 1.43] to be significantly associated with CHB after stratification for rs9277535 ($P_{corrected} < 3.0 \times 10^{-3}$, Supplementary Material, Table S5). The SNP rs2856718 is located in the intergenic region between *HLA-DQA2* and *HLA-DQB1*, while rs7453920 is located in intron 1 of *HLA-DQB2*

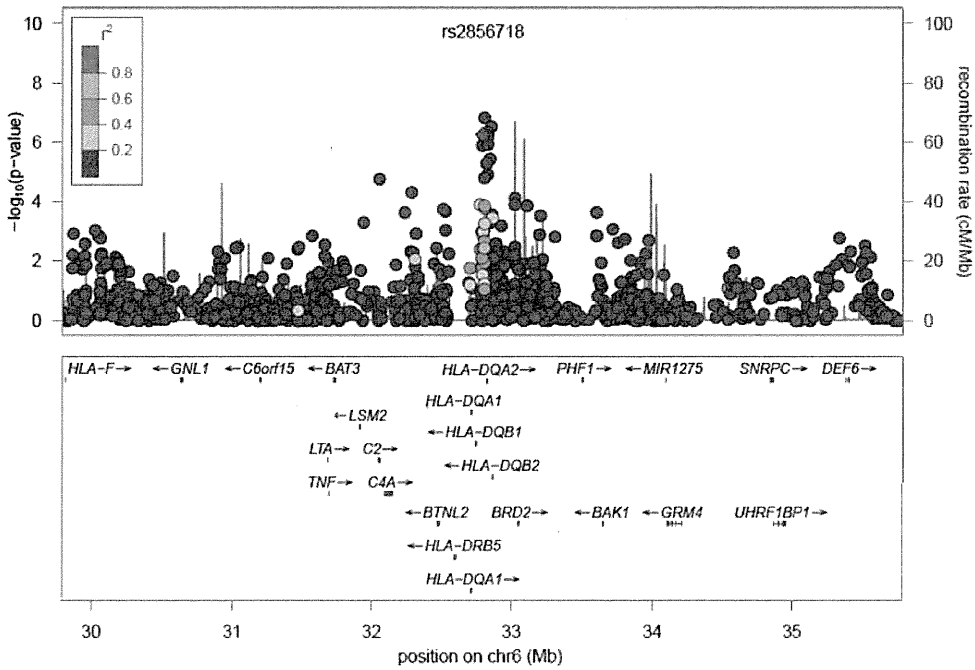

Figure 2. Regional association plot of the 6p21.32 locus after adjustment for the top SNP (rs9277535) in the *HLA-DP* locus in the GWAS stage. This figure shows the evidence of independent association with CHB based on logistic regression analysis. Only one strong peak remained after adjustment for rs9277535. This peak, represented by three top SNPs: rs3892710, rs7453920 and rs2856718, is located in the *HLA-DQ* locus (6p21.32).

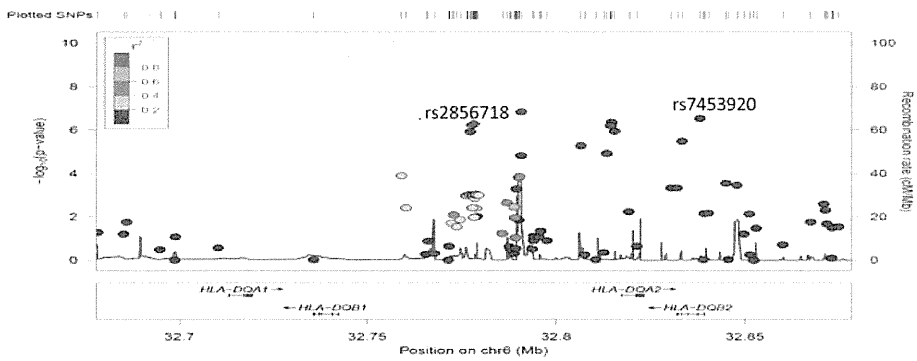

Figure 3. Regional association plot of the *HLA-DQ* locus. This figure indicates a ~200 kb region centered on the association peak, located between rs2856718 and rs7453920. The middle panel shows the genes in this region including the *HLA-DQ* locus.

(Fig. 3). To further validate these results, we analyzed these SNPs in two additional Japanese cohorts consisting of 381 cases and 1539 controls from Biobank Japan as well as 1222 cases and 879 controls from Hiroshima University. Association for these SNPs loci was confirmed in both replication sets (P-value $= 3.14 \times 10^{-5}$–3.59×10^{-12}; Table 1). To combine these studies, we conducted a meta-analysis with a fixed-effects model using the Mantel–Haenszel method. As shown in Table 1 and Supplementary Material, Figure S3, the OR were quite similar among the four studies and no heterogeneity was observed. Mantel–Haenszel P-values for independence were 3.99×10^{-37} for rs2856718 [OR = 1.77, 95% confidence interval (CI) = 1.65–1.91], and 5.98×10^{-28} for rs7453920 (OR = 1.81, 95% CI = 1.62–2.01). Two previously reported SNPs on the *HLA-DP* locus (rs9277535 on *HLA-DPB1* and rs3077 on *HLA-DPA1*) were also associated with CHB ($P_{\text{meta-analysis}} = 2.55 \times 10^{-54}$ and 1.57×10^{-61}) (Table 1).

To test whether the strong association observed in these regions is due to the effect of one of them, we performed logistic regression analysis based on the effect of each top SNP in both *HLA-DP* and *HLA-DQ* loci. Notably, rs2856718 and rs7453920 did show strong association with CHB after adjusting for the effect of rs3077 ($P = 8.12 \times 10^{-27}$ and $P = 1.52 \times 10^{-21}$, respectively) and rs9277535 ($P = 2.38 \times 10^{-30}$ and $P = 2.21 \times 10^{-22}$, respectively), indicating variants at the *HLA-DQ* locus are associated with CHB independent of the effect of *HLA-DP* polymorphisms (Table 2).

Table 1. Summary of results for GWAS and replication study

Chr (position)	SNP	Nearest gene	Allele (1/2)	Stage	Case 11	12	22	MAF[a]	Control 11	12	22	MAF[a]	*P*-value[b]	OR[c] (95% CI)	P_{het}^{d}
6 (33141000)	rs3077	*HLA-DPA1*	A/G	GWAS	38	156	264	0.25	330	991	735	0.40	1.28×10^{-16}	1.98 (1.68–2.32)	
				First replication	42	240	324	0.27	313	947	762	0.39	1.93×10^{-14}	1.74 (1.51–2.01)	
				Second replication	36	139	204	0.28	268	742	529	0.42	9.52×10^{-12}	1.84 (1.55–2.19)	
				Third replication	115	430	681	0.27	155	420	304	0.42	1.53×10^{-21}	1.93 (1.69–2.2)	
				Meta-analysis[e]									1.57×10^{-61}	1.87 (1.73–2.01)	0.62
6 (33162839)	rs9277535	*HLA-DPB1*	A/G	GWAS	40	179	239	0.28	384	1020	652	0.43	3.72×10^{-17}	1.95 (1.67–2.28)	
				First replication	58	254	294	0.31	364	963	696	0.42	3.70×10^{-12}	1.63 (1.42–1.87)	
				Second replication	42	145	192	0.30	301	758	480	0.44	5.43×10^{-12}	1.83 (1.54–2.17)	
				Third replication	133	464	628	0.30	160	429	290	0.43	1.02×10^{-16}	1.75 (1.54–1.99)	
				Meta-analysis[e]									2.55×10^{-54}	1.77 (1.65–1.91)	0.40
6 (32778233)	rs2856718	*HLA-DQB1*	A/G	GWAS	158	226	73	0.41	477	1001	568	0.48	4.41×10^{-10}	1.59 (1.37–1.85)	
				First replication	209	266	127	0.43	484	966	572	0.48	1.07×10^{-7}	1.43 (1.27–1.64)	
				Second replication	128	191	62	0.41	325	746	468	0.45	7.49×10^{-11}	1.72 (1.45–2)	
				Third replication	465	530	227	0.40	216	420	243	0.48	3.59×10^{-12}	1.59 (1.39–1.79)	
				Meta-analysis[e]									3.99×10^{-37}	1.56 (1.45–1.67)	0.24
6 (32837990)	rs7453920	*HLA-DQB2*	A/G	GWAS	4	72	382	0.09	67	582	1407	0.17	1.27×10^{-10}	2.20 (1.73–2.81)	
				First replication	5	127	471	0.11	50	575	1397	0.17	5.47×10^{-6}	1.56 (1.28–1.9)	
				Second replication	4	75	302	0.11	53	422	1064	0.17	3.14×10^{-5}	1.69 (1.32–2.17)	
				Third replication	14	198	1011	0.09	19	245	615	0.16	2.21×10^{-11}	1.88 (1.56–2.27)	
				Meta-analysis[e]									5.98×10^{-28}	1.81 (1.62–2.01)	0.16

[a]MAF, minor allele frequency.
[b]*P*-value of the Cochrane–Armitage trend test for each stage.
[c]OR and CI are calculated using the non-susceptible allele as reference.
[d]*P*-value of the Breslow–Day test.
[e]Results of meta-analysis were calculated by the Mantel–Haenzel method.

Table 2. Logistic regression results for the top SNPs in HLA-DP and HLA-DQ loci associated with CHB in all stages

SNP	P-value[a]	$P_{adjusted\ for\ rs3077}$	OR (95% CI)	$P_{adjusted\ for\ rs9277535}$	OR (95% CI)	$P_{adjusted\ for\ rs2856718}$	OR (95% CI)	$P_{adjusted\ for\ rs7453920}$	OR (95% CI)
rs3077	1.57×10^{-61}	NA	–	2.05×10^{-10}	1.43 (1.3–1.67)	7.45×10^{-48}	1.7 (1.58–1.83)	9.42×10^{-51}	1.73 (1.61–1.85)
rs9277535	2.55×10^{-54}	1.67E−05	1.25 (1.15–1.45)	NA	–	6.80×10^{-47}	1.67 (1.55–1.79)	9.03×10^{-48}	1.67 (1.56–1.8)
rs2856718	3.99×10^{-37}	8.12E−27	1.43 (1.33–1.54)	2.38×10^{-30}	1.43 (1.37–1.56)	NA	–	6.34×10^{-26}	1.43 (1.34–1.53)
rs7453920	5.98×10^{-28}	1.52E−21	1.66 (1.49–1.85)	2.21×10^{-22}	1.67 (1.51–1.85)	4.96×10^{-18}	1.60 (1.44–1.77)	NA	–

Trend P-values are shown with or without adjusting the analysis for the most associated SNPs in *HLA-DP* and *HLA-DQ* loci.
[a]Meta-analysis P-value was calculated by the Mantel–Haenzel method.

Subsequently, we examined the interaction of four SNPs in *HLA-DP* and *HLA-DQ* genes on CHB susceptibility. We only found evidence for interactive effects between HLA-DP SNPs and also between HLA-DQ SNPs (Supplementary Material, Table S6). For all other pairwise combinations, each locus had an independent role in CHB ($P_{interaction} > 0.10$). CHB risk increases with increasing number of risk alleles for four SNPs (Fig. 4 and Supplementary Material, Table S7). Individuals with seven or eight risk alleles have more than 5-fold higher CHB risk than those with two or less risk alleles. Taken together, our findings clearly indicated the additive effects of variants in *HLA-DP* and *HLA-DQ* loci on CHB susceptibility.

HLA-DQ molecules function as a heterodimer of α and β subunits, those are encoded by the *HLA-DQA1* and the *HLA-DQB1* genes, respectively. The SNP rs2856718 is located in a linkage disequilibrium (LD) block including *HLA-DQB1* and *HLA-DQA1* genes, and rs7453920 and rs2856718 are in LD with r^2 of 0.1 and D' of 0.73 (Fig. 3 and Supplementary Material, Fig. S4). Similar to *HLA-DPs*, *HLA-DQs* are highly polymorphic especially in exon 2 which encode antigen-binding sites. We therefore considered that the association of these SNPs with CHB might reflect variations in antigen-binding sites of *HLA-DQA1* and *DQB1* that would affect the immune response to HBV. Hence, we genotyped *HLA-DQA1* and *DQB1* alleles by direct sequencing of exon 2 (cases and controls from the GWAS and first replication sets) and found *HLA-DQB1*0303* and *DQB1*0602* were significantly associated with CHB susceptibility ($P = 1.49 \times 10^{-6}$ and 1.87×10^{-5}, OR = 1.64 and 2.51, respectively), while *DQB1*0501* and *DQB1*0604* were significantly associated with protection from persistent HBV infection ($P = 3.61 \times 10^{-4}$ and 5.38×10^{-16}, OR = 0.50 and 0.22, respectively) (Supplementary Material, Table S8). To further investigate the relationship between *HLA-DQ* alleles and CHB susceptibility, we performed logistic regression analysis using SNPs rs2856718 and rs7453920 as covariates. Interestingly, *HLA-DQB1*0303* and**0604* showed strong association with CHB after adjustment for rs2856718 and rs7453920 ($P = 6.3 \times 10^{-4}$ and $P = 2.59 \times 10^{-8}$, respectively). In addition, we performed logistic regression analysis using the top *HLA-DQ* alleles that show the strongest association (*DQB1*0303*, **0602*, **0501*, **0604*) as covariate. As expected, HLA-DQ SNPs rs2856718 and rs7453920 failed to find the association between CHB and those SNPs ($P = 0.36$, and $P = 0.08$, respectively). Finally, we performed conditional analysis of the *DQB1, DPA1* and *DPB1* alleles together. As a result, HLA-DP SNPs rs3077 and rs9277535 as well as HLA-DQ SNPs rs2856718 and rs7453920 did not show any further association beyond these *HLA-DQ* and *DP* alleles (rs9277535, $P = 0.55$, OR = 0.88; rs3077, P = NA; rs2856718, $P = 0.63$, OR = 0.95 and rs7453920, $P = 0.30$, OR = 0.85). We also performed conditional analysis of the *DPA1* and *DPB1* and we found that *HLA-DQ* alleles **0303*, **0602* and **0604* still showed strong association ($P = 0.0006$, OR = 1.5; $P = 0.00047$, OR = 2.28 and $P = 6.66 \times 10^{-7}$, OR = 0.31) except for *DQB1*0501* ($P = 0.35$, OR = 0.81) which already showed weak association before adjustment as shown in Supplementary Material, Table S8. Collectively, these results together confirmed our findings for the

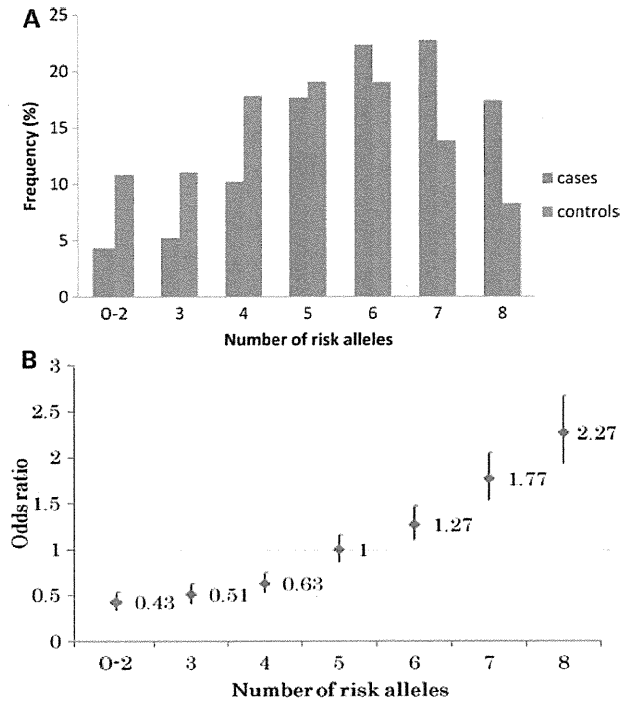

Figure 4. Cumulative effects of CHB risk alleles. (**A**) Distribution of risk alleles in CHB cases (red bars) and controls (blue bars). (**B**) Plot of the increasing OR for CHB with increasing number of risk alleles. The OR are relative to the median number of four risk alleles (rs3077, rs9277535, rs2856718 and rs7453920). Vertical bars correspond to 95% CIs. Horizontal line marks the null value (OR = 1).

causality of *HLA-DQ* and *HLA-DP* alleles and their independent effects on the CHB susceptibility. We further performed haplotype analysis and found four haplotypes showing the highest association ($8.39 \times 10^{-5} - 3.42 \times 10^{-13}$); *DQA1*0102-DQB1*0604* and *DQA1*0101-DQB1*0501* were considered to have protective effects ($P = 3.42 \times 10^{-13}$, OR = 0.16 and $P = 1.06 \times 10^{-5}$, OR = 0.39, respectively), whereas *DQA1*0102-DQB1*0303* and *DQA1*0301-DQB1*0601* increased a risk of CHB ($P = 8.39 \times 10^{-5}$, OR = 19.03, and $P = 7.34 \times 10^{-5}$, OR = 5.02, respectively, Table 3). Furthermore, we performed integrated analysis to test the haplotypic relationship between HLA-DP and DQ. We found seven associated haplotypes: *DQA1*0501-DQB1*0301-DPA1*0202-DPB1*0501*, *DQA1*0301-DQB1*0401 DPA1*0103DPB1*0201*, *DQA1*0301-DQB1*0302-DPA1*0202-DPB1*0501* and *DQA1*0102-DQB1*0604-DPA1*0103-DPB1*0401* showed protective effects ($P = 1.90 \times 10^{-4}$, OR = 0.18; $P = 5.30 \times 10^{-3}$, OR = 0.27; $P = 5.90 \times 10^{-3}$, OR = 0.43 and $P = 9.70 \times 10^{-3}$, OR = 0.41, respectively), whereas *DQA1*0301-DQB1*0301-DPA1*0103-DPB1*0201*, *DQA1*0102-DQB1*0602-DPA1*0202-DPB1*0501* and *DQA1*0301-DQB1*0601-DPA1*0202-DPB1*0501* were associated with susceptibility to CHB ($P = 2.30 \times 10^{-3}$, OR = 4.9; $P = 9.30 \times 10^{-4}$, OR = 4.8 and $P = 3.30 \times 10^{-5}$, OR = 11, respectively, Supplementary Material, Table S9). Taken together, our findings strongly implicated the significant association of *HLA-DQ-DP* haplotypes with CHB.

Recent GWASs have identified several SNPs that are associated with viral and non-viral liver diseases as well as response to HBV vaccination and liver function test (16–18). More recently, Zhang *et al.* (19) performed a GWAS of hepatocellular carcinoma in chronic HBV carriers of Chinese ancestry. They successfully identified one intronic SNP rs17401966 in *KIF1B* on chromosome 1p36.22 that was highly associated with HBV-related hepatocellular carcinoma. We analyzed those loci in our GWAS data, but failed to find the association between CHB and those SNPs (Supplementary Material, Table S10).

DISCUSSION

Here, we present the results of the two-stage GWAS followed by two independent replications on a total of 2667 cases with CHB and 6496 controls in Japanese population. In this study, we genotyped additional 279 cases and 1122 controls by using Illumina Human610-Quad BeadChip. As a result, we increased the number of samples in the first screening from 179 cases and 934 controls in the previous study to 458 cases and 2056 controls in current study. As a result, the statistic power to detect SNPs with moderate effects (i.e. OR of 1.4 and risk allele frequency of 0.2) increased from 23 to 85% at a significance threshold of 5×10^{-5}. Indeed, two SNPs in *HLA-DQ* locus did not indicate significant association in the GWAS stage of our previous GWAS ($P = 5.62 \times 10^{-2}$ for rs2856718 and $P = 4.88 \times 10^{-2}$ for rs7453920), confirming the importance of sample size in GWAS (20).

Most of significant SNPs with *P*-value of smaller than 5×10^{-5} (74 among 88 SNPs) are located in the MHC region which encompasses a large number of genes involved in our immunological response.

Three groups of *HLA* class II genes produce cell-surface Ag, designated *HLA-DR*, *HLA-DQ* and *HLA-DP*. It is suggested that the host immune response to HBV is under T lymphocyte control, and this response has been shown to be HLA-restricted (21). The *HLA-DQ* locus is located ~300 kb telomeric of the *HLA-DP* locus in a different LD block. Indeed, the analysis of the HLA complex revealed several recombination hot spots distributing across the HLA complex, including two hot spots near *DP* and *DQ* genes (22,23). The result of conditional analyses also demonstrated that the association of the *HLA-DQ* locus with CHB is independent from that of the *HLA-DP* locus.

Previous reports showed an association of *HLA* class II alleles with susceptibility of persistent HBV infection (24–27), but the results were inconsistent even within the same population except for *HLA-DR13*. *HLA-DR13* (corresponding to *HLA-DRB1*1301* and **1302* alleles) was consistently associated with HBV clearance across the population, and we found that rs11752643 which is strongly linked with *HLA-DR13* (28) showed a strong association in the GWAS stage ($P = 1.26 \times 10^{-10}$). The SNP rs3892710 which is in strong LD with rs11752643 ($r^2 = 0.8$, $D' = 1$) and showed higher association in the GWAS stage ($P = 4.49 \times 10^{-12}$) was selected for replication in the first independent replication set. However, rs3892710 failed to clear Bonferroni correction

Table 3. Haplotype analysis

No.	Haplotype		Haplotype frequencies		*P*-value[a]	OR[a] (95% CI)
	HLA.DQA1	*HLA.DQB1*	Case (%)	Control (%)		
1	*0102	*0604	1.22	6.59	3.42×10^{-13}	0.16 (0.09–0.29)
2	*0101	*0501	1.68	4.77	1.06×10^{-5}	0.39 (0.24–0.65)
3	*0501	*0301	3.06	5.79	1.52×10^{-3}	0.53 (0.35–0.79)
4	*0301	*0401	9.73	13.40	2.98×10^{-3}	0.76 (0.57–1.02)
5	*0301	*0302	5.08	7.56	1.67×10^{-2}	0.72 (0.50–1.02)
6	*0301	*0402	2.55	3.49	1.73×10^{-1}	0.74 (0.45–1.22)
7	*0401	*0402	1.31	1.62	4.91×10^{-1}	0.72 (0.36–1.44)
8	*0101	*0503	4.23	4.34	8.69×10^{-1}	0.94 (0.62–1.42)
9	*0103	*0601	18.70	18.90	9.11×10^{-1}	Reference
10	*0601	*0301	1.38	0.89	2.53×10^{-1}	1.46 (0.68–3.11)
11	*0301	*0503	1.48	0.95	2.06×10^{-1}	1.65 (0.74–3.68)
12	*0301	*0301	2.46	1.79	1.97×10^{-1}	1.33 (0.76–2.33)
13	*0101	*0502	2.09	1.39	1.89×10^{-1}	1.67 (0.90–3.11)
14	*0301	*0303	16.90	13.10	7.50×10^{-3}	1.32 (1–1.74)
15	*0102	*0602	3.39	1.55	3.47×10^{-3}	2.24 (1.28–3.92)
16	*0102	*0303	1.91	0.25	8.39×10^{-5}	19.03 (2.53–143.39)
17	*0301	*0601	2.45	0.42	7.34×10^{-5}	5.02 (1.87–13.45)

[a]*P*-values, OR and its 95% CIs of each haplotype were calculated as described in Materials and Methods.

for multiple testing after adjustment for rs9277535 ($P = 4.73 \times 10^{-2}$). In addition, the association of hepatitis B with *HLA-DQ* SNPs rs2856718 and rs7453920 remarkably attenuated after adjustment for rs11752643 using the logistic regression model ($P = 2.53 \times 10^{-6}$ and $P = 5.84 \times 10^{-4}$, respectively). Unlike HLA-DP SNPs, rs3077 and rs9277535 remained highly significant ($P = 7.74 \times 10^{-13}$ and 2.52×10^{-12}, respectively). Therefore, our findings clearly indicated that hepatitis B is associated with the variants on *HLA-DP* loci independent of the association with SNP rs11752643 that is closely linked with *HLA-DR13* and also reinforce the previous report of *HLA-DQ-DR* linkage. Thus, our study demonstrated that the association of CHB with the variants in the *HLA-DQ* locus was more prominent and consistent than those with *HLA-DR13* in the Japanese population. However, the 19 major haplotypes shown in Supplementary Material, Table S9 accounted for only 51.80% of cases and 57.92% of controls, and other 314 haplotypes were missed due to low haplotype frequency (<1% in both cases and controls). Therefore, the result of *DP-DQ* haplotype analyses should be carefully interpreted. Subsequently, further functional analysis including *HLA-DR*, *DQ* and *DP* is essential to fully elucidate the molecular mechanism whereby these variations confer CHB susceptibility.

In summary, we have demonstrated that genetic variations in the *HLA-DQ* genes were strongly associated with CHB in the Japanese population, and this association was independent from the *HLA-DP* genes which we reported previously. Considering the importance of the MHC region in the clearance after the infection of HBV, our findings should provide a novel insight that the antigen presentation on the HLA-DP and HLA-DQ molecules might be critical for virus elimination and play an important role in the development of CHB. We are confident that our findings would serve to allow better understanding of the pathogenesis of hepatitis B and contribute to better clinical outcome of the disease.

MATERIALS AND METHODS

Study population

A total of 2667 cases and 6496 control subjects were analyzed in this study. Characteristics of each cohort are shown in Supplementary Material, Table S1. DNA samples from both CHB patients and non-HBV controls used in this study were obtained from the BioBank Japan at the Institute of Medical Science, the University of Tokyo (29) except for samples for the third replication. Among the BioBank Japan samples, we selected HBsAg-seropositive CHB patients with elevated serum aminotransferase levels for more than six months, according to the guideline for diagnosis and treatment of chronic hepatitis from The Japan Society of Hepatology (http://www.jsh.or.jp/medical/gudelines/index.html). The control groups for the GWAS and first replication as well as for the second replication consisted of subjects with diseases other than CHB (uterine cancer, esophageal cancer, hematological cancer, pulmonary tuberculosis, ovarian cancer, keloid, peripheral artery disease and ischemic stroke) that were also negative for HBsAg. Case and control samples for the third replication cohort were collected from hospitals participating to the Hiroshima Liver Study Group (listing of participating doctors in this study group can be obtained at http: //home.hiroshima-u.ac.jp/naika1/research_profile/pdf/liver_ study_ group_e.pdf) and Toranomon Hospital. All the participants provided written informed consent. This project was approved by the ethical committees at each institute.

SNP genotyping and QC

In the GWAS stage, 458 patients with CHB and 2056 non-HBV controls were genotyped using Illumina Infinium HumanHap550v3 or Illumina Infinium Human610-Quad DNA Analysis Genotyping BeadChip. SNP QC for all sets of samples was applied as follows: SNP call rate of

≥ 0.99 in both cases and controls and *P*-value of the Hardy–Weinberg equilibrium test of $\geq 1.0 \times 10^{-6}$ in controls. SNPs with minor allele frequency of ≤ 0.01 in both case and control samples were excluded from the further analysis. In the first replication, we genotyped an additional panel of 616 cases by multiplex polymerase chain reaction (PCR)-based Invader assay (Third Wave Technologies, Madison, WI, USA) (30). After excluding 10 cases with the call rate of <0.95, all cluster plots were visually analyzed by trained staffs and SNPs with ambiguous calls were excluded. Randomly selected 94 case samples in the GWAS stage were re-genotyped in the first replication and SNPs with concordance rates of <98% between two assays (Illumina and Invader) were excluded. In the subsequent replication analyses, we used the TaqMan genotyping system (Applied Biosystems, Foster City, CA, USA) or the multiplex PCR-based Invader assay.

HLA-DQA1 and HLA-DQB1 genotyping

We analyzed *HLA-DQ* genotypes using 748 cases and 614 controls (from GWAS and first replication sets). The second exons of the *HLA-DQA1* and *HLA-DQB1* genes were amplified and directly sequenced according to the protocol reported previously (31–33). *HLA-DQA1* and *DQB1* alleles were determined based on the alignment database of dbMHC.

Statistical analysis

In the GWAS stage and replication analyses, statistical significance of the association with each SNP was assessed using 1-df Cochrane–Armitage trend test and logistic regression analysis adjusted with top SNP (rs9277535) in the *HLA-DP* locus. Significance levels after Bonferroni correction for multiple testing were $P = 3.0 \times 10^{-3}$ (0.05/16) in the first replication and $P = 0.025$ (0.05/2) in second and third replication. OR and CIs were calculated using the non-susceptible allele as a reference. The meta-analysis was conducted using the Mantel–Haenszel method. Heterogeneity among studies was examined by the Breslow–Day test. To assess the association of each *HLA* allele, we used Fisher's exact test on two-by-two contingency tables with or without each *HLA* allele. To analyze the association of haplotypes, we used R package haplo.stats. *P*-values for each haplotype were given by the results of a score test, and OR and 95% CIs were calculated from coefficients of the generalized linear model. OR of each haplotype were calculated relative to the major haplotype. All of these statistical values were calculated by function haplo.cc.

Software

For general statistical analysis, we used R statistical environment version 2.11.1 (http://cran.r-project.org) or plink-1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/). Estimation of haplotype frequencies and analysis of haplotype association were performed by R package haplo.stats (34). Sequence variants in the second exons of *HLA-DQA1* and *HLA-DQB1* were analyzed by Sequencher 4.8. Haploview software was employed to analyze LD values and draw LD map.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We would like to acknowledge members of Hiroshima Liver Study Group and the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. We also thank the technical staff of the Laboratory for Genotyping Development at RIKEN.

Conflict of Interest statement. None declared.

FUNDING

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

REFERENCES

1. Pungpapong, S., Kim, W.R. and Poterucha, J.J. (2007) Natural history of hepatitis B virus infection: an update for clinicians. Mayo Clin. Proc., **82**, 967–975.
2. Okada, K., Kamiyama, I., Inomata, M., Imai, M. and Miyakawa, Y. (1976) e Antigen and anti-e in the serum of asymptomatic carrier mothers as indicators of positive and negative transmission of hepatitis B virus to their infants. N. Engl. J. Med., **294**, 746–749.
3. Chemin, I. and Zoulim, F. (2009) Hepatitis B virus induced hepatocellular carcinoma. Cancer Lett., **286**, 52–59.
4. McMahon, B. (2009) The natural history of chronic hepatitis B virus infection. *Hepatology*, **49**, S45–S55.
5. Lin, T., Chen, C., Wu, M., Yang, C., Chen, J., Lin, C., Kwang, T., Hsu, S., Lin, S. and Hsu, L. (1989) Hepatitis B virus markers in Chinese twins. Anticancer Res., **9**, 737–741.
6. Kamatani, Y., Wattanapokayakit, S., Ochi, H., Kawaguchi, T., Takahashi, A., Hosono, N., Kubo, M., Tsunoda, T., Kamatani, N., Kumada, H. *et al.* (2009) A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat. Genet., **41**, 591–595.
7. Thio, C., Carrington, M., Marti, D., O'Brien, S., Vlahov, D., Nelson, K., Astemborski, J. and Thomas, D. (1999) Class II HLA alleles and hepatitis B virus persistence in African Americans. J. Infect. Dis., **179**, 1004–1006.
8. Kummee, P., Tangkijvanich, P., Poovorawan, Y. and Hirankarn, N. (2007) Association of HLA-DRB1*13 and TNF-alpha gene polymorphisms with clearance of chronic hepatitis B infection and risk of hepatocellular carcinoma in Thai population. J. Viral. Hepat., **14**, 841–848.
9. Hwang, S., Sohn, Y., Oh, H., Hwang, C., Lee, S., Shin, E. and Lee, K. (2007) Human leukocyte antigen alleles and haplotypes associated with chronicity of hepatitis B virus infection in Koreans. Arch. Pathol. Lab. Med., **131**, 117–121.
10. Ben-Ari, Z., Mor, E., Papo, O., Kfir, B., Sulkes, J., Tambur, A., Tur-Kaspa, R. and Klein, T. (2003) Cytokine gene polymorphisms in patients infected with hepatitis B virus. Am. J. Gastroenterol., **98**, 144–150.
11. Höhler, T., Kruger, A., Gerken, G., Schneider, P., Meyer zum Büschenefelde, K. and Rittner, C. (1998) A tumor necrosis factor-alpha (TNF-alpha) promoter polymorphism is associated with chronic hepatitis B infection. Clin. Exp. Immunol., **111**, 579–582.
12. Migita, K., Maeda, Y., Abiru, S., Nakamura, M., Komori, A., Miyazoe, S., Nakao, K., Yatsuhashi, H., Eguchi, K. and Ishibashi, H. (2007) Polymorphisms of interleukin-1beta in Japanese patients with hepatitis B virus infection. J. Hepatol., **46**, 381–386.
13. Chong, W., To, Y., Ip, W., Yuen, M., Poon, T., Wong, W., Lai, C. and Lau, Y. (2005) Mannose-binding lectin in chronic hepatitis B virus infection. *Hepatology*, **42**, 1037–1045.

14. Thio, C., Mosbruger, T., Kaslow, R., Karp, C., Strathdee, S., Vlahov, D., O'Brien, S., Astemborski, J. and Thomas, D. (2004) Cytotoxic T-lymphocyte antigen 4 gene and recovery from hepatitis B virus infection. J. Virol., **78**, 11258–11262.
15. Zhou, J., Lu, L., Yuen, M., Lam, T., Chung, C., Lam, C., Zhang, B., Wang, S., Chen, Y., Wu, S. *et al.* (2007) Polymorphisms of type I interferon receptor 1 promoter and their effects on chronic hepatitis B virus infection. J. Hepatol., **46**, 198–205.
16. Davila, S., Froeling, F., Tan, A., Bonnard, C., Boland, G., Snippe, H., Hibberd, M. and Seielstad, M. (2010) New genetic associations detected in a host response study to hepatitis B vaccine. Genes Immun., **11**, 232–238.
17. Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Nakamura, Y. and Kamatani, N. (2010) Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat. Genet., **42**, 210–215.
18. Romeo, S., Kozlitina, J., Xing, C., Pertsemlidis, A., Cox, D., Pennacchio, L., Boerwinkle, E., Cohen, J. and Hobbs, H. (2008) Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat. Genet., **40**, 1461–1465.
19. Zhang, H., Zhai, Y., Hu, Z., Wu, C., Qian, J., Jia, W., Ma, F., Huang, W., Yu, L., Yue, W. *et al.* (2010) Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. Nat. Genet., **42**, 755–758.
20. Wellcome Trust Case Control Consortium. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
21. Milich, D. (1988) T- and B-cell recognition of hepatitis B viral antigens. *Immunol. Today*, **9**, 380–386.
22. Jeffreys, A., Kauppi, L. and Neumann, R. (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat. Genet., **29**, 217–222.
23. Cullen, M., Perfetto, S., Klitz, W., Nelson, G. and Carrington, M. (2002) High-resolution patterns of meiotic recombination across the human major histocompatibility complex. Am. J. Hum. Genet., **71**, 759–776.
24. Liu, C. and Cheng, B. (2007) Association of polymorphisms of human leucocyte antigen-DQA1 and DQB1 alleles with chronic hepatitis B virus infection, liver cirrhosis and hepatocellular carcinoma in Chinese. Int. J. Immunogenet., **34**, 373–378.
25. Xun, Y., Guo, J., Shi, W., Shi, J. and Liu, C. (2009) Association between HLA-DQA1 gene polymorphism and the outcomes of hepatitis B virus infection. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi*, **23**, 430–433.
26. Thio, C., Thomas, D., Karacki, P., Gao, X., Marti, D., Kaslow, R., Goedert, J., Hilgartner, M., Strathdee, S., Duggal, P. *et al.* (2003) Comprehensive analysis of class I and class II HLA antigens and chronic hepatitis B virus infection. J. Virol., **77**, 12083–12087.
27. Singh, R., Kaul, R., Kaul, A. and Khan, K. (2007) A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J. Gastroenterol., **13**, 1770–1787.
28. Thursz, M., Kwiatkowski, D., Allsopp, C., Greenwood, B., Thomas, H. and Hill, A. (1995) Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N. Engl. J. Med., **332**, 1065–1069.
29. Nakamura, Y. (2007) The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.*, **5**, 696–697.
30. Ohnishi, Y., Tanaka, T., Ozaki, K., Yamada, R., Suzuki, H. and Nakamura, Y. (2001) A high-throughput SNP typing system for genome-wide association studies. J. Hum. Genet., **46**, 471–477.
31. van der Zwan, A., Griffith, B., Rozemuller, E., Williams, T. and Tilanus, M.G.J. (2002) Sequence-based typing for HLA-DQB1 strategy for ABI sequencing equipment. In Tilanus, M.G.J. (ed.), International Histocompatibility Working Group, Seattle, Washington, 2002.
32. van Dijk, A., Melchers, R., Tilanus, M. and Rozemuller, E. (2007) HLA-DQB1 sequencing-based typing updated. *Tissue Antigens*, **69**(Suppl. 1), 64–65.
33. Witter, K., Zahn, R., Volgger, A. and Reininger, A.J. (2010) HLA-DQB1*0404, a novel DQB1-allele detected in a volunteer blood platelet donor typed for HLA. *Tissue Antigens*, **76**, 256–258.
34. Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M. and Poland, G.A. (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am. J. Hum. Genet., **70**, 425–434.

Associations of *HLA-DP* Variants with Hepatitis B Virus Infection in Southern and Northern Han Chinese Populations: A Multicenter Case-Control Study

Jin Li[1], Daguo Yang[2], Yongwen He[3], Mengyi Wang[4], Zirong Wen[5], Lifeng Liu[1], Jinjian Yao[1], Koichi Matsuda[6], Yusuke Nakamura[6], Jinling Yu[1], Xiaorui Jiang[1], Shuzhen Sun[1], Qing Liu[1], Xiang Jiang[1], Qilong Song[1], Man Chen[1], Hong Yang[1], Feng Tang[1], Xiaowen Hu[1], Jing Wang[1], Ying Chang[1], Xingxing He[1], Yuan Chen[4]*, Jusheng Lin[1]*

1 Institute of Liver Disease, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, China, **2** The Third People's Hospital of Shenzhen, Shenzhen, Guangdong Province, China, **3** Department of Infection Disease, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, China, **4** Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, China, **5** Qingdao Infectious Disease Hospital, Qingdao, Shandong Province, China, **6** Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Abstract

Background: Human leukocyte antigen DP (*HLA-DP*) locus has been reported to be associated with hepatitis B virus (HBV) infection in populations of Japan and Thailand. We aimed to examine whether the association can be replicated in Han Chinese populations.

Methodology/Principal Findings: Two *HLA-DP* variants rs2395309 and rs9277535 (the most strongly associated SNPs from each *HLA-DP* locus) were genotyped in three independent Han cohorts consisting of 2 805 cases and 1 796 controls. By using logistic regression analysis, these two SNPs in the *HLA-DPA1* and *HLA-DPB1* genes were significantly associated with HBV infection in Han Chinese populations ($P=0.021\sim3.36\times10^{-8}$ at rs2395309; $P=8.37\times10^{-3}\sim2.68\times10^{-10}$ at rs9277535). In addition, the genotype distributions of both sites (rs2395309 and rs9277535) were clearly different between southern and northern Chinese population ($P=8.95\times10^{-5}$ at rs2395309; $P=1.64\times10^{-9}$ at rs9277535). By using asymptomatic HBV carrier as control group, our study showed that there were no associations of two *HLA-DP* variants with HBV progression ($P=0.305\sim0.822$ and $0.163\sim0.881$ in southern Chinese population, respectively; $P=0.097\sim0.697$ and $0.198\sim0.615$ in northern Chinese population, respectively).

Conclusions: Our results confirmed that two SNPs (rs2395309 and rs9277535) in the *HLA-DP* loci were strongly associated with HBV infection in southern and northern Han Chinese populations, but not with HBV progression.

Citation: Li J, Yang D, He Y, Wang M, Wen Z, et al. (2011) Associations of *HLA-DP* Variants with Hepatitis B Virus Infection in Southern and Northern Han Chinese Populations: A Multicenter Case-Control Study. PLoS ONE 6(8): e24221. doi:10.1371/journal.pone.0024221

Editor: Sunil K. Ahuja, South Texas Veterans Health Care System, United States of America

Received January 27, 2011; **Accepted** August 5, 2011; **Published** August 31, 2011

Funding: This study was financially supported by the National Basic Research Program of China (973 Program, No. 2007CB512903) and the National Natural Science Foundation of China (No. 30872237). 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: jslin2010@Hotmail.com (J.Lin); chenyuan008@163.com (Y.Chen)

Introduction

More than 2 billion people have been infected with the hepatitis B virus (HBV) worldwide, of which 350 million are chronic carriers and about 600 000 die annually of HBV-related acute or chronic liver disease [1]. Although many individuals eventually achieve a state of nonreplicative infection, the prolonged immunologic response to infection leads to the development of cirrhosis, liver failure, or hepatocellular carcinoma (HCC) in up to 40% of patients [2]. In China, where HBV infection is endemic, there are estimated 93 million HBV carriers, and among them 30 million are patients with chronic hepatitis B [3]. Multiple causes influence the risk of chronic HBV infection in china, for example, age, gender, viral genotype, ethnicity, variation in genes of the immune system and so on [4].

Several polymorphisms of the *HLA* loci have been reported for hepatitis B virus infection [5,6]. A study in Gambian found that the allele DRB1*1302 was associated with the clearance of the virus [7]. Hepatitis B virus persistence and disease chronicity were associated with *HLA-DQA*1*0501 and *HLA-DQB*1*0301 in Chinese [8] and with *HLA-DR*9 in Koreans [9]. Although the association between common diseases and these *HLA* (or non-*HLA*) genes has become increasingly evident [10], their results are conflicting among the studies, and have not been confirmed by other investigators [11].

A recent study found that the *HLA-DP* locus was associated with chronic hepatitis B in Japanese and Thais [12]. As the frequencies of these *HLA-DP* alleles in Chinese populations were similar to those in Japanese populations, it would be necessary to confirm whether there was the association between the *HLA-DP* genetic

variation and HBV infection in Chinese populations. To this end, we selected the most strongly associated SNPs (the previous GWAS results) from each *HLA-DP* locus (rs9277535 at the *HLA-DPB1* and rs2395309 at the *HLA-DPA1*, respectively) and genotyped these two polymorphisms in a population-based case-control study of Chinese Hans, including 2 805 cases and 1 796 controls from Hubei province (Central China), Shandong province (North China) and Guangdong province (South China).

Materials and Methods

Ethic statement

The study was approved by the local research ethics committee (REC) at the Tongji Hospital of Huazhong University of Science and Technology in accordance with the principle of the Helsinki Declaration II. All written informed consent documents from each participant were obtained during the enrollment phase.

Study subjects

A total of 4 601 unrelated Han Chinese were recruited in this study between September 2007 and June 2011. All subjects were divided into six groups: a) HBV clearance group(Clear); b) Healthy control group(Health); c) Persistent asymptomatic HBV carriers group(AsC); d) Chronic active hepatitis B group(CHB); e) HBV-related liver cirrhosis group(LC); and f) HBV-related heptocellular carcinoma group(HCC). The diagnostic criteria for study inclusion were listed in Table S1, which had been described in the previous publication [13,14]. All individuals were gathered from three Han Chinese cohorts. First, we recruited 2 280 subjects from Tongji hostital and Union hospital in Wuhan, Hubei province. Second, we gathered additional 1 304 subjects from The Affiliated Hospital of Binzhou Medical College and Qingdao Infectious Disease Hospital in Shandong province, 1 017 subjects from Shenzhen Third People's Hospital, Shenzhen Fourth People's Hospital and Shenzhen Sixth People's Hospital in Guangdong province.

A uniform questionnaire was used at three enrollment sites and recorded self-report of risk factors for HBV transmission, family history of HBV infection, past and current smoking, alcohol ingestion, etc. The demographic information included gender, birth-date, birthplace, and past and current residency.

DNA Isolation and Genotyping

Genomic DNA was isolated from peripheral whole blood using TIANamp blood DNA kit (Tiangen Biotech [Beijing] Co., Ltd., China). The concentration and purity of the DNA were determined with a NanoDrop spectrophotometer and diluted to a final concentration of 8 ng/μL. The genotyping of genetic polymorphisms was performed via the TaqMan method according to the protocol of TaqMan® SNP Genotyping Assays (Applied Biosystems, California, USA). Allelic category was measured automatically using the Sequence Detection System 2.3 software (Applied Biosystems) according to the intensity of VIC and FAM dye. To detect these SNPs (rs2395309 and rs9277535), we customized the TaqMan® MGB Probe as well as the primers for PCR amplification (Table S2.).

Statistical analysis

Statistical analysis was conducted by using haploview 4.2, Arlequin 3.5, Stata10.0 and SPSS 17.0 softwares. Linkage disequilibrium was assessed by the haploview 4.2 softwares using frequencies obtained from the Health group. The (Bayesian) ELB algorithm was used to infer haplotypes by using Arlequin 3.5. The Hard-Weinberg equilibrium of alleles and population pairwise comparisons were also evaluated by using Arlequin 3.5 [15]. A meta-analysis of all studies was performed for each SNP associated with chronic hepatitis B by using Stata10.0 softwares. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated on the basis of the binary logistic regression analysis (adjustment for gender and age). The strength of association between the genotypes or alleles and HBV infection was estimated by using SPSS 17.0 softwares. A best-fit model was constructed by means of comparisons with other models. Values of $P<0.05$ were considered statistically significant.

Results

Hardy-Weinberg equilibrium test

Hardy-Weinberg equilibrium was estimated by Fisher's exact test using Arlequin 3.5 software. There was no significant difference between observed and expected frequencies of each genotype in these involved populations ($P>0.05$). This result indicated that these populations had a relatively stable genetic background and were suitable for further genetic statistical analysis.

The clinic and demographic characteristics

The clinical and demographic characteristics of the case-control study were summarized in Table 1, including gender, age, drinkers, serum total bilirubin level (T-Bil), HBV-DNA load, alanine transaminase (ALT) and serum markers of hepatitis B virus. There was no significant difference in the percentage of hepatitis B e antigen (HBeAg) positive ($P=0.10$) between asymptomatic HBV carriers (17.1%) and the patients of chronic hepatitis B group (20.2%). In addition, there was more alcohol consumption in patients ($P<0.05$) with HBV-related liver cirrhosis group (24.3%) and HBV-related heptocellular carcinoma group (30.5%) than those in HBV clearance group (10.2%) and healthy control group (8.9%). The difference in the alcohol consumption status was due to few drinkers in Chinese female population. Although an effort was made to obtain a good match on age and sex, there were more men in four case groups (averaged 73.5%) than those in HBV clearance group (51.6%, $P<0.05$) and healthy control group (47.5%, $P<0.05$).

Population pairwise comparisons and grouping of subjects

To explore whether differences in susceptibility loci were caused by the disease or by genetic background between populations, we first needed to determine which populations should be compared with each other, and whether there were populations that could be lumped together to simplify statistical analysis. To this end, we performed population pairwise comparisons $\mathbf{F_{ST}}$ testing between each population using Arlequin 3.5 software. The principle of population pairwise comparisons states that: if there is no difference in heredity between two populations, the data permuting of genotypes or haplotypes between two populations should not cause a significant difference, which can be evaluated by $\mathbf{F_{ST}}$ P value ($P>0.05$). According to the results shown in Table 2, we could infer that Shandong population had greater difference than Hubei population or Guangdong population in genetic background ($P<0.0001$). As the pairwise comparisons for Hubei population and Guangdong population were not significantly different ($P=0.191$), and both of them were the same geographic position (southern of china) [16], we had determined to merge Hubei population and Guangdong population into southern Chinese population. Meanwhile, Shandong population was taken as northern Chinese population. Furthermore, in order

Table 1. Clinical characteristics of study subjects.

Characteristics	Health(n = 962)	Clear(n = 834)	AsC(n = 910)	CHB(n = 964)	LC(n = 544)	HCC(n = 387)
Gender, no. (%)						
Male	454(47.5)	421(51.6)	519(58.4)	765(81.0)	435(80.3)	342(88.8)
Female	501(52.5)	395(48.4)	370(41.6)	180(19.0)	107(19.7)	43(11.2)
Age (years), mean(sd)	51.70±10.96	51.48±11.04	45.64±10.93	38.45±10.73	48.30±11.02	49.82±10.01
Drinkers, no. (%)	86(8.9)	85(10.2)	102(11.2)	179(18.6)	132(24.3)	118(30.5)
HBsAg	All−	All−	All+	All+	All+	All+
Anti-HBs IgG	All−	All+	All−	All−	All−	All−
HBeAg-positive,no. (%)	All−	All−	156(17.1)	195(20.2)	64(11.8)	39(10.1)
Anti-HBc IgG, no. (%)	All−	All+	All+	All+	All+	All+
Family history, no. (%)	No	No	36(4.0)	142(14.7)	69(12.7)	77(19.9)
ALT(U/L)	No	No	19.74±15.09	447.53±402.71	101.57±119.90	83.64±86.17
TBil (μmol/L)	No	No	11.71±9.01	161.55±171.06	124.19±125.86	69.24±65.74
HBV-DNA (copy/ml)	No	No	No	3.23E7±4.76E7	1.38E6±5.16E6	3.72E6±1.13E7

'No' means non-detected.
'Drinkers' was defined as alcohol consumption of >40 g/week, which included occasional drinkers and daily drinkers.
Abbreviations: Clear, HBV clearance group; Health, Healthy control group; AsC, Asymptomatic HBV carriers group; CHB, Chronic active hepatitis B group; LC,HBV-related liver cirrhosis group; HCC, HBV-related heptocellular carcinoma group;
doi:10.1371/journal.pone.0024221.t001

to identify whether the two polymorphisms were associated with HBV infection or clearance, we combined all the types of HBV infection populations into one group by using the healthy group or clearance group as the reference.

Logistic regression analysis of the *HLA-DP* loci polymorphisms

Then, to investigate which genotypic models were significantly associated with the various outcomes, we conducted comparisons of four models (Multiplicative model, Additive model, Dominant model and Recessive model) in southern and northern Chinese populations respectively (data not show). For the four models, the best-fit genotypic effect of these two SNPs (rs2395309 and rs9277535) was observed in the dominant model which was the protective genotype AA and AG (see Table 3). After compared with the Healthy control group, both single nucleotide polymorphism (SNP) sites (rs2395309 and rs9277535) showed associations with HBV infection in southern Chinese population (Odds ratio [OR] = 0.57; 95% Confidence intervals [CI] :0.47, 0.70; $P = 3.36\times10^{-8}$ at rs2395309; OR = 0.52; 95% CI :0.43, 0.64; $P = 2.68\times10^{-10}$ at rs9277535), as well as in northern Chinese population (OR = 0.50; 95% CI :0.35, 0.71; $P = 1.23\times10^{-4}$ at rs2395309; OR = 0.50; 95% CI :0.36, 0.68; $P = 1.74\times10^{-5}$ at rs9277535). And, interestingly, *HLA-DP* rs2395309 and rs9277535 sites also showed a strong protective effect for HBV clearance not only in southern Chinese population (OR = 1.31; 95% CI :1.17, 1.45; $P = 9.63\times10^{-7}$ at rs2395309; OR = 1.33; 95% CI :1.20, 1.49; $P = 1.67\times10^{-7}$ at rs9277535) but also in northern Chinese population (OR = 1.20; 95% CI :1.03, 1.40; $P = 0.021$ at rs2395309; OR = 1.26; 95% CI :1.06, 1.49; $P = 8.37\times10^{-3}$ at rs9277535). As shown in Table 3, notably, the genotype distributions of both sites (rs2395309 and rs9277535) were clearly different between southern and northern Healthy populations (P values = 8.95×10^{-5} and 1.64×10^{-9}, respectively. P values of Pearson's x^2 test for allele model). The two minor-allele frequencies (MAF) in both Healthy populations (southern and northern Han Chinese) were 30.1% vs 38.8% at rs2395309, 38.1% vs 52.2% at rs9277535. In addition, to decrease the bias of sex and age in population sampling, we further conducted the stratified analysis for sex and age. As presented in Table S3, male and female patients showed different associations with HBV diseases in these two SNPs (rs2395309 and rs9277535). Specially, in the northern Chinese population, this difference was notable between male patients and female patients. Furthermore, in the stratified analysis of age, most cases were no significant differences in genotype distributions of two SNPs sites between patients with age≤45 years and patients with age>45 years (Table S4.).

Table 2. Matrix of significant F_{ST} P values among populations.

Populations	Hubei	Guangdong	Shandong
Hubei	*		
Guangdong	0.191	*	
Shandong	**<0.0001**	**<0.0001**	*

Population pairwise comparisons F_{ST} tests were performed between pairs of groups using Arlequin 3.5software. Statistically significant values are in shown in bold.
Abbreviations: Hubei, Hubei populations; Guangdong, Guangdong populations; Shandong, Shandong populations.
doi:10.1371/journal.pone.0024221.t002

Associations of the *HLA-DP* loci polymorphisms with HBV progression

Considering the function of *HLA-DP* molecules, we were interested in the possible association between the polymorphisms in *HLA-DP* gene and the disease progression of chronic hepatitis B. To test our prediction, we further analysed the difference in two SNPs genotype distributions by using asymptomatic HBV carrier as control group. Unfortunately, there were not associations in chronic active hepatitis B group (OR = 1.03; 95% CI : 0.79, 1.34; $P = 0.822$ at rs2395309; OR = 0.92; 95% CI : 0.71, 1.18; $P = 0.501$ at rs9277535, in southern Chinese population; OR = 0.92; 95% CI : 0.62, 1.38; $P = 0.697$ at rs2395309; OR = 1.33; 95% CI : 0.86, 2.06; $P = 0.198$ at rs9277535, in northern Chinese population),

Table 3. Associations of two SNPs (rs2395309, rs9277535) with HBV infection and clearance in Han Chinese populations.

	South of china		North of china	
	Control group	**Case group**	**Control group**	**Case group**
HLA-DPB1 (rs2395309)- dominant model (AA+AGvsGG)				
AA/AG/GG	57/234/288[†]	112/709/1367[‡]	52/193/138[†]	63/249/302[‡]
P value OR (95%CI)	Reference	3.36×10^{-8} 0.57 (0.47,0.70)	Reference	1.23×10^{-4} 0.50 (0.35,0.71)
AA/AG/GG	112/709/1367[‡]	35/235/257[‖]	63/249/302[‡]	56/130/121[‖]
P value OR (95%CI)	Reference	9.63×10^{-7} 1.31 (1.17,1.45)	Reference	0.021 1.20 (1.03,1.40)
HLA-DPA1 (rs9277535)- dominant model (AA+AGvsGG)				
AA/AG/GG	80/277/216[†]	177/830/1195[‡]	97/203/80[†]	118/287/206[‡]
P value OR (95%CI)	Reference	2.68×10^{-10} 0.52 (0.43,0.64)	Reference	1.74×10^{-5} 0.50 (0.36,0.68)
AA/AG/GG	177/830/1195[‡]	67/251/208[‖]	118/287/206[‡]	67/165/75[‖]
P value OR (95%CI)	Reference	1.67×10^{-7} 1.33 (1.20,1.49)	Reference	8.37×10^{-3} 1.26 (1.06,1.49

[†]Healthy control group.
[‖]HBV clearance group.
[‡]HBV infection groups, including Asymptomatic HBV carriers, Chronic active hepatitis B group, HBV-related liver cirrhosis group, HBV-related heptocellular carcinoma group.
The *P* values, odds ratios (OR), and 95% confidence intervals (CI) were calculated on the basis of the binary logistic regression analysis, adjusted for sex and age.
doi:10.1371/journal.pone.0024221.t003

HBV-related liver cirrhosis group (OR = 1.11; 95% CI : 0.82, 1.52; $P=0.499$ at rs2395309; OR = 1.24; 95% CI : 0.92, 1.67; $P=0.163$ at rs9277535, in southern Chinese population; OR = 0.74; 95% CI : 0.48, 1.16; $P=0.189$ at rs2395309; OR = 1.29; 95% CI : 0.81, 2.06; $P=0.286$ at rs9277535, in northern Chinese population) and HBV-related heptocellular carcinoma group(OR = 0.85; 95% CI : 0.63, 1.16; $P=0.305$ at rs2395309; OR = 0.98; 95% CI : 0.73, 1.31; $P=0.881$ at

Table 4. Results of the association test for two SNPs(rs2395309,rs9277535) haplotypes in Han Chinese populations.

South of china						
Haplotype	**Health(2n = 1106)**	**Clear(2n = 1048)**	**AsC(2n = 1342)**	**CHB(2n = 1486)**	**LC(2n = 754)**	**HCC(2n = 632)**
A-A	247(22.3)	205(17.0)	227(16.8)	238(16.0)	122(16.2)	92(14.5)
A-G	88(8.0)	97(11.8)	62(4.6)	95(6.4)	36(4.8)	23(3.7)
G-A	173(15.6)	177(19.5)	140(10.4)	143(9.6)	95(12.6)	78(12.4)
G-G	598(54.1)	569(51.7)	913(68.1)	1010(68.0)	501(66.4)	439(69.4)
P value[II*]	Reference		1.47×10^{-6}	6.47×10^{-8}	2.53×10^{-5}	6.07×10^{-7}
OR (95%CI)			0.60 (0.49,0.74)	0.57 (0.47,0.70)	0.59 (0.46,0.76)	0.51 (0.39,0.66)
P value[#I]		Reference	7.35×10^{-4}	8.92×10^{-5}	2.45×10^{-3}	1.06×10^{-4}
OR (95%CI)			1.45 (1.17,1.80)	1.53 (1.24,1.89)	1.48 (1.15,1.91)	1.72 (1.31,2.27)
North of china						
Haplotype	**Health(2n = 734)**	**Clear(2n = 608)**	**AsC (2n = 422)**	**CHB (2n = 378)**	**LC(2n = 300)**	**HCC(2n = 100)**
A-A	226(30.8)	200(31.5)	118(28.0)	103(27.3)	71(23.6)	21(20.5)
A-G	55(7.5)	41(8.1)	27(6.4)	18(4.8)	16(5.4)	4(3.5)
G-A	157(21.4)	98(17.5)	66(15.6)	60(15.9)	62(20.7)	15(15.4)
G-G	296(40.3)	269(42.9)	211(50.0)	197(52.1)	151(50.3)	60(60.5)
P value[II*]	Reference		0.032	0.012	0.004	0.003
OR (95%CI)			0.73 (0.55,0.97)	0.68 (0.51,0.92)	0.62 (0.44,0.86)	0.46 (0.27,0.78)
P value[#I]		Reference	0.054	0.021	0.007	0.005
OR (95%CI)			1.33 (0.99,1.78)	1.42 (1.05,1.92)	1.58 (1.13,2.21)	2.12 (1.25,3.61)

[II*]Two SNPs haplotypes G-G, A-A in Health group compared with those in HBV infection groups.
[#I]Two SNPs haplotypes G-G, A-A in HBV infection groups compared with those in Clearance group.
The *P* values, odds ratios (OR), and 95% confidence intervals (CI) were calculated by Pearson Chi-Square test.
Abbreviations: Clear, HBV clearance group; Health, Healthy control group; AsC, Asymptomatic HBV carriers group; CHB, Chronic active hepatitis B group; LC,HBV-related liver cirrhosis group; HCC, HBV-related heptocellular carcinoma group; OR, odds ratio; CI, confidence interval.
doi:10.1371/journal.pone.0024221.t004

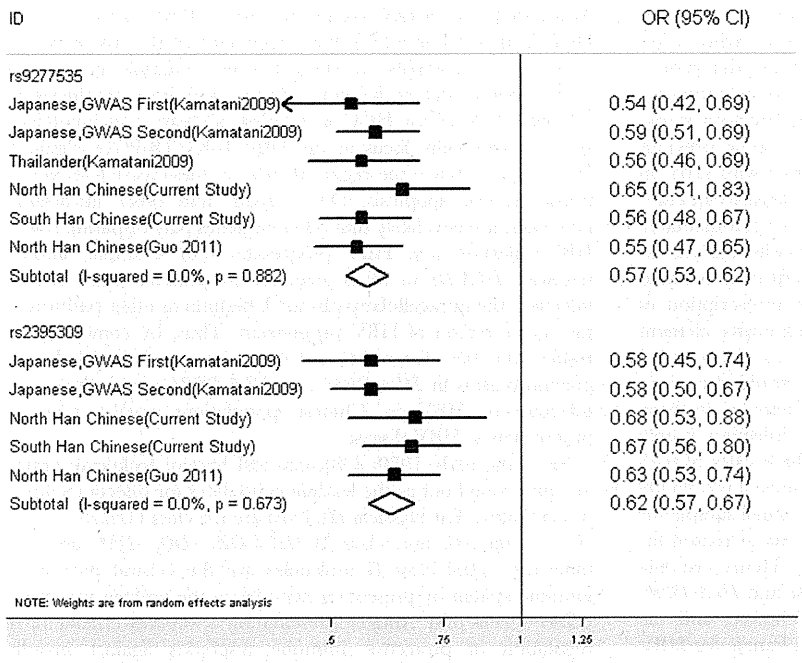

Figure 1. Meta-analasis of the rs9277535 and rs2395309. The meta-analysis combined with the results of previous studies, including more than 2,243 cases and 4,137 controls. Each effect size is shown with its confidence interval. Abbreviations: p, *P* heterogeneity value; OR, odds ratios; 95%CI, 95% confidence interval.
doi:10.1371/journal.pone.0024221.g001

rs9277535, in southern Chinese population; OR = 0.56; 95% CI : 0.28, 1.11; $P = 0.097$ at rs2395309; OR = 0.84; 95% CI : 0.42, 1.68; $P = 0.615$ at rs9277535, in northern Chinese population), compared with asymptomatic HBV group(Table S5.).

Associations of the *HLA-DP* loci polymorphisms with clinical factors

In order to analyze the associations between two SNPs and clinical factors (HBV-DNA load, ALT and TB), we used the independent-sample Kolmogorov-Smirnov t test in CHB group, LC group and HCC group. Although the GG patients have a higher mean on the HBV-DNA load, no significant difference was found between patients of different genotypes (see Fig. S1). In the analysis of ALT, the associations between two SNPs and the ALT level only be found in HBV-related liver cirrhosis group ($P = 0.002$ at rs2395309; $P = 0.009$ at rs9277535), rather than in other groups. Meanwhile, for the associations of the TB level, there was no difference between GG patients and AG+AA patients ($P > 0.05$ in each group).

Results of the Haplotype analysis and Meta-analysis

To further understand the contributions of these loci to HBV susceptibility, two-locus haplotypes were constructed for two SNPs rs2395309 and rs9277535 (Table 4.). Pairwise linkage disequilibrium (LD) analyses performed using all individuals from the health group showed that rs2395309 and rs9277535 SNPs were in LD with each other ($D' = 0.57$, $r^2 = 0.23$ in southern Chinese population; $D' = 0.58$, $r^2 = 0.20$ in northern Chinese population). In trying to derive HBV infection-specific haplotypes, the haplotype frequencies of two SNPs (rs2395309 and rs9277535) were evaluated in both Chinese populations. Four haplotypes were observed, and among them three haplotypes had frequencies more than 5% (Table 4.). Compared with protective A-A haplotype homozygotes, only G-G haplotype homozygotes had a significant increased risk for HBV infection (*P* value and odds ratios were shown in Table 4). Then, we summarized a meta-analysis combined with the results of related studies [12,17], including more than 2,243 cases and 4,137 controls. As shown in Figure 1 and Table S6, these odds ratios were quite similar among the three ethnic groups (Japanese, Thais and Chinese) and no heterogeneity was observed (P het = 0.673 at rs2395309; P het = 0.882 at rs 9577535).

Discussion

In this analysis, we confirmed that two SNPs sites (rs2395309 and rs9277535) in the *HLA-DPA1* and *HLA-DPB1* genes were significantly associated with HBV infection in southern and northern Han Chinese populations. Again, our haplotype analysis showed the frequency of G-G haplotype had a significant increase in the HBV infected populations, as compared with the healthy control group or HBV clearance group. As a result, we inferred that these persons with G-G haplotype have a higher risk of HBV infection than those persons with A-A haplotype. Meanwhile, the A-A haplotype could be strongly predictive for HBV clearance in HBV infection populations. Although our manuscript suggested that the genotype distributions of both sites (rs2395309 and rs9277535) were different between southern and northern Chinese population, the frequencies of two protective alleles A in Chinese populations were also similar to those in Asian populations, compared with European and Central American populations (data from public databases, HapMap). The results of the genetic association in our study were consistent with the previous study [12]. Hence, we could confirm that the polymorphisms of *HLA-DPA1* and *HLA-DPB1* gene play a very important role in chronic hepatitis B virus infection in southern and northern Han Chinese populations.