HLA-DQB1*03 Confers Susceptibility to Chronic Hepatitis C in Japanese: A Genome-Wide Association Study

Daiki Miki1,2,3,*, Hidenori Ochi1,2,3,*, Atsushi Takahashi4, C. Nelson Hayes1,2,3, Yuji Urabe2, Hiromi Abe1,2, Tomokazu Kawaoka1,2,3, Masatake Tsuge8,5, Nobuhiko Hiraga1,2,3, Michio Imamura1,2,3, Yoshiiku Kawakami1,2,3, Hiroshi Aikata1,2,3, Shoichi Takahashi2,3, Norio Akuta6, Fumitaka Suzuki6, Kenji Ikeda6, Hiromitsu Kumada6, Yoshiyasu Karino7, Joji Toyota7, Tatsuhiro Tsunoda7, Michiaki Kubo7, Naoyuki Kamatani4, Yusuke Nakamura10, Kazuaki Chayama1,2,3*

1 Laboratory for Digestive Diseases, Center for Integrative Medical Sciences, RIKEN, Hiroshima, Japan, 2 Department of Gastroenterology and Metabolism, Applied Life Sciences, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan, 3 Liver Research Project Center, Hiroshima University, Hiroshima, Japan, 4 Laboratory for Statistical Analysis, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan, 5 Natural Science Center for Basic Research and Development, Hiroshima University, Hiroshima, Japan, 6 Department of Hepatology, Toranomon Hospital, Tokyo, Japan, 7 Department of Gastroenterology, Sapporo Kosei General Hospital, Hokkaido, Japan, 8 Laboratory for Medical Science Mathematics, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan, 9 Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan, 10 Laboratory of Molecular Medicine, Human Genome Center, The Institute of Medical Science, University of Tokyo, Tokyo, Japan

Abstract

Hepatitis C virus (HCV) establishes a chronic infection in 70-80% of infected individuals. Many researchers have examined the effect of human leukocyte antigen (HLA) on viral persistence because of its critical role in the immune response against exposure to HCV, but almost all studies have proven to be inconclusive. To identify genetic risk factors for chronic HCV infection, we analyzed 458,207 single nucleotide polymorphisms (SNPs) in 481 chronic HCV patients and 2,963 controls in a Japanese cohort. Next, we performed a replication study with an independent panel of 4,358 cases and 1,114 controls. We further confirmed the association in 1,379 cases and 25,817 controls. In the GWAS phase, we found 17 SNPs that showed suggestive association (P < 1 × 10^{-3}). After the first replication study, we found one intronic SNP in the HLA-DQ locus associated with chronic HCV infection, and when we combined the two studies, the association reached the level of genome-wide significance. In the second replication study, we again confirmed the association (P_{combined} = 3.59 × 10^{-10}, odds ratio [OR] = 0.79). Subsequent analysis revealed another SNP, rs1130380, with a stronger association (OR=0.72). This nucleotide substitution causes an amino acid substitution (R55P) in the HLA-DQB1 protein specific to the DQB1*03 allele, which is common worldwide. In addition, we confirmed an association with the previously reported IFNL3-IFNL4 locus and propose that the effect of DQB1*03 on HCV persistence might be affected by the IFNL4 polymorphism. Our findings suggest that a common amino acid substitution in HLA-DQB1 affects susceptibility to chronic infection with HCV in the Japanese population and may not be independent of the IFNL4 genotype.


Editor: Ludmila Prokunina-Olsson, National Cancer Institute, National Institutes of Health, United States of America

Received September 19, 2013; Accepted November 20, 2013; Published December 20, 2013

Copyright: © 2013 Miki et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partially supported by Grants-in-Aid for scientific research and development from the Ministry of Health, Labor and Welfare (http://www.mhlw.go.jp/seisakunitsuite/bunya/hokabunya/kenkyyjigyou/) and Ministry of Education Culture Sports Science and Technology (http://www.mext.go.jp/a_menu/shinkou/hoyo/main5_a5.htm), Government of Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: chayama@hiroshima-u.ac.jp

These authors contributed equally to this work.

Introduction

At least 1.5 million people in Japan and more than 200 million people worldwide are chronically infected with the hepatitis C virus (HCV) [1,2]. In many western countries and Japan, HCV infection is the most common risk factor for hepatocellular carcinoma (HCC) [3,4]. HCV establishes a chronic infection in approximately 70-80% of infected individuals [5,6], and host genetic factors in addition to viral factors are assumed to partially explain the heterogeneity in HCV persistence or clearance. Many researchers have examined the effect of human leukocyte antigen (HLA) on viral
Table 1. Summary of GWAS and replication studies.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Case</th>
<th>Control</th>
<th>MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>locus</td>
<td>Case</td>
<td>Control</td>
<td>MAF</td>
</tr>
<tr>
<td>rs9275572</td>
<td>96</td>
<td>217</td>
<td>0.245</td>
</tr>
<tr>
<td>HLA-DQB1</td>
<td>2049</td>
<td>1773</td>
<td>0.381</td>
</tr>
<tr>
<td>Combined GWAS and 1st replication study</td>
<td>206</td>
<td>695</td>
<td>0.402</td>
</tr>
<tr>
<td>Combined all studies</td>
<td>7.63E-06</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MAF: minor allele frequency; OR; odds ratio; CI; confidence interval.
Odds ratios and P values for independence test were calculated by the Mantel-Haenszel method.

Results

GWAS and Replication Studies

In the GWAS phase, single nucleotide polymorphism (SNP) genotyping was performed using the Illumina HumanHap610-Quad BeadChip for cases and the Illumina HumanHap550v3 BeadChip for controls. Genotype concordance between these two BeadChips was 99.99% among 182 duplicated samples, indicating the low possibility of genotyping error. 458,207 SNPs passed quality control filters and were analyzed using an additive model for genotype-phenotype association in 481 patients with chronic HCV infection and 2,963 controls. Principal component analysis revealed no population substructure in our population. In addition, a quantile-quantile plot using the results of the Cochran-Armitage trend test showed that the inflation factor, λ, was 1.007, indicating a low probability of false-positive associations resulting from population stratification (Figure S1A). Using the additive model, one SNP reached the genome-wide significance level for association after Bonferroni correction (calculated as \( P < 0.05/458,207 = P < 0.05 \times 10^{-5} \)), and another 24 SNPs showed suggestive association (\( P < 1 \times 10^{-5} \)) in our GWAS (Figure S1B and Table S1). Next, we conducted the first replication study to validate the results of the GWAS phase using 4,358 cases and 1,114 controls. We performed genotyping of 18 SNPs with \( P \) values < 1x10^{-5} in the GWAS phase, after excluding 7 SNPs with \( r^2 = 1 \) in the same locus based on the HapMap-JPT database, and successfully genotyped all SNPs using multiplex PCR-based Invader assay (Table S2), except for rs9275563, which is moderately linked with rs9275572 \( (r^2 = 0.56) \). Only rs9275572 in the HLA-DQ locus on chromosome 6 showed significant replication of the association \( (P < 0.05/17), \) and the association reached a genome-wide significance level when we combined the two studies using the Mantel-Haenszel method \( (P = 2.04 \times 10^{-12}, \text{ odds ratio } [OR] = 0.79; \text{ Table 1}) \). We further confirmed the association in 1,379 cases and 25,817 controls and found a highly significant association with chronic HCV after meta-analysis of all three studies \( (P_{\text{combined}} = 3.59 \times 10^{-16}, \text{ OR} = 0.79, 95\% \text{ confidence interval } [CI] 0.75-0.84) \). As shown in Table 1, we observed no heterogeneity among the three studies (heterogeneity test \( P = 0.113 \)).

Multiple Logistic Regression Analysis and Stratified Analysis

We next evaluated the association in more detail using the first replication set. Both HCV case and healthy control samples were collected at Hiroshima University. After adjusting for gender and age using multiple logistic regression analysis, the rs9275572 C allele remained highly significant with an OR = 0.79 (95% CI 0.70-0.89) (Table S3). Our study also confirmed a stronger protective effect of female gender compared to male gender [12,16]. Subsequently, we analyzed...
the effect of rs9275572 according to gender and age. As shown in Table S4A, we observed no significant difference between subgroups. In addition to host genetic factors, such as HLA alleles, the distribution of viral genotypes is also highly variable around the world, and it has been suggested that different HCV genotypes result in different clinical outcomes [17,18]. For example, HCV genotype 1, the most common genotype in Japan, shows stronger resistance against IFN-based anti-viral therapy than HCV genotype 2, the second most frequently encountered genotype in Japan [19,20].

Hence, we stratified HCV cases by HCV genotype and compared each with healthy controls. We observed that the effect tended to be slightly greater in genotype 1 (OR = 0.74 [0.66-0.82]) than genotype 2 (OR = 0.83 [0.74-0.94]), but there was no significant difference (heterogeneity test \( P = 0.127 \)) (Table S4B).

**Detailed Analysis of HLA-DQ alleles**

Although no allele showed significant association (\( P < 0.05/7 \)) among HLA-DQA1 alleles, the HLA-DQB1*0303 allele showed significant association among HLA-DQB1 alleles (\( P < 0.05/12, \) OR = 0.69). We also performed haplotype analysis using 1,025 cases and 1,038 controls that had been successfully genotyped for HLA-DQA1, HLA-DQB1 and rs9275572 to investigate the effects of haplotype combinations; however, no haplotype showed stronger association than the single association of HLA-DQB1*0303 (Table S6). In our Japanese cohort, DQB1*0303 was the most frequent of the DQB1*03 alleles and showed the strongest effect against HCV persistence, but we also found that all DQB1*03 alleles tended to have a protective effect. Hence, we speculate that not only DQB1*0303, but all DQB1*03 alleles share protective residues. We searched for a variant that is common to all DQB1*03 alleles but that differs from non-DQB1*03 alleles and found that SNP rs1130380 is a variant specific to DQB1*03 (Table S5A and S5B). Then, we evaluated the single marker effect of several nucleotide variants, which are useful to distinguish among common DQB1 alleles using the same individuals with the haplotype analysis, and finally determined that rs1130380 (\( P = 6.08 \times 10^{-7}, \) OR = 0.72) had a stronger effect than the landmark SNP, rs9275572 (\( P = 2.07 \times 10^{-5}, \) OR = 0.76) (Table S7). After mutual adjustment between rs1130380 and rs9275572 using multiple logistic regression analysis, the association for both variants was remarkably attenuated (\( P = 2.38 \times 10^{-3} \) and \( P = 0.0534, \) respectively), suggesting their mutual linkage (\( D^2 = 0.99, \) \( r^2 = 0.28 \)) (Figure S4). Therefore, these variants probably represent the same genetic signal and rs1130380 could be more compatible as a representative variant of the signal. This nucleotide polymorphism causes an amino acid substitution at position 55 on the HLA-DQB1 molecule. We examined the position of this amino acid substitution using on a three-dimensional structure of HLA-DQ molecule that was determined by X-ray diffraction method [21].

As shown in Figure S5, the amino acid at B55 forms part of a peptide-binding pocket in the HLA-DQ molecule, suggesting a critical role in antigen presentation. The change from Arg to Pro implies a physicochemical change from basic to hydrophobic (Figure S6), but we have no relevant data with which to show a change of affinity to certain peptides at this time. Further functional analysis is therefore needed.

**Evaluation of IFNL3-IFNL4 locus in GWAS**

In spite of our failure to detect it in our GWAS screen, variants close to IFNL3 have been reported to be associated with spontaneous clearance of HCV, so we tried to confirm the association using our GWAS data. Because genotype data for rs12979860 was not available in our platform, we checked the genotype data using another frequently reported SNP, rs8099917, which is highly linked with rs12979860 in the Japanese population (\( r^2 = 0.94 \)) [22,23]. As shown in Table S8, the association was confirmed in our GWAS set (\( P = 3.22 \times 10^{-5}, \) OR = 0.73) with the same allele conferring a protective effect.

**Analysis of IFNL3-IFNL4 locus and HLA-DQ*03**

We genotyped a IFNL4 variant, ss469415590, and two representative variants close to IFNL3, rs12979860 and rs8099917, by Invader assay using the same individuals used in the detailed analysis of the HLA-DQ alleles above (1,114 cases and 1,114 controls). Each of these three polymorphisms were successfully genotyped in 1,074 cases and 1,073 controls. rs12979860 and ss469415590 were almost complete LD, and ss469415590 showed the strongest association with HCV persistence under the additive model (\( P = 2.96 \times 10^{-5}, \) OR = 0.75) (Table 3). After adjusting for the effect of the HLA-DQ*03 allele using multiple logistic regression analysis, the significance level of ss469415590 increased slightly (\( P = 1.67 \times 10^{-3}, \) OR=0.74). We additionally analyzed the association under different genetic models other than additive. Among several genetic models, the IFNL4 variant, ss469415590, showed the strongest association under the recessive model for the protective allele (the dominant model for the risk allele), whereas rs1130380 showed the strongest association under the allelic model (Table S9). We further evaluated the effect of the DQB1*03 allele stratified by ss469415590 genotype (Table 4). Interestingly, DQB1*03 showed a stronger effect in individuals with the ss469415590-TT/TT genotype, and this tendency became more pronounced for rs9275572, the landmark SNP in our GWAS (heterogeneity test \( P = 2.45 \times 10^{-3} \)).
Discussion

The MHC region containing the HLA locus is widely believed to play an important role in viral infection, and a recent GWAS revealed a strong association between the HLA-DP and DQ locus and hepatitis B virus (HBV) persistence [24,25] as well as HBV-induced HCC [26] in East Asia. In our GWAS with two replication studies, we aimed to clarify the genetic factors involved in chronic HCV infection, and we successfully identified an association with the HLA-DQB1 locus based on the landmark SNP rs9275572. An association between the MICA locus and the risk of developing HCV-induced HCC was previously reported in a Japanese GWAS [27], and this study identified rs9275572 as a second associated SNP. HCV-positive cases in the previous study, which are independent of ours, showed a similar minor allele frequency (MAF) of rs9275572 to our cases (0.407 in GWAS stage and 0.429 in replication stage), which strongly supports our results. Furthermore, according to the 1000 Genomes database, the MAF of this SNP is 0.348 in the Japanese population, which is consistent with the MAF of our controls.

On the other hand, while we observed no significant heterogeneity across the three studies, there were some differences in MAF among studies in both case and control populations. For both cases and controls, MAFs were lower in the first replication study than in the other two studies. Among the three studies, only the first replication set was collected from the same geographical area, therefore, geographical differences might affect the differences in MAF among studies, but this is only speculation. There are also differences in the

Figure 1. Linkage disequilibrium structure around the HLA-DQ region based on D’ using HapMap-JPT Data. P-value plot and genomic structure of the GWAS stage. Blue lines represent the LD block containing the HLA-DQA2 and HLA-DQB2 loci. Black lines represent two LD blocks containing the HLA-DQA1 and HLA-DQB1 loci, and the black dotted line represents the boundary between them. The red line represents the most strongly associated SNP, rs9275572, which is located within the HLA-DQB1 locus. The LD maps were created using HaploView software.

doi: 10.1371/journal.pone.0084226.g001
In the current study, we failed to detect the previously most replicated IFNL3-IFNL4 locus during GWAS screening even though the association indeed existed in our GWAS set ($P = 3.22 \times 10^{-3}$, OR = 0.73; Table S8). The odds ratio for rs8099917 is nearly the same as for rs9275572, but the $P$ values were quite different, with rs9275572 ranked in the top 9 candidate SNPs, whereas the other is far below at number 1,640. This crucial difference may be explained by differences in allele frequencies. The minor allele frequency of rs9275572 is over three times higher than that of rs8099917, and the statistical power to detect each SNP in our GWAS was calculated to be only 65% and 8%, respectively, compared to the 80% recommended to detect an association of the expected effect size. We are concerned that there are several remaining undetected SNPs that might affect HCV clearance or persistence, in addition to the IFNL3-IFNL4 locus. To overcome this problem, it will be necessary to use a larger sample size, as well as more appropriate control samples. In fact, the major limitation of our study involves the selection of controls. Undoubtedly, the most appropriate control set would include patients who have been exposed to HCV (anti-HCV positive) but who spontaneously cleared the infection (HCV-negative). Although the risk of exposure to HCV in the Japanese general population is very low, making it difficult to locate suitable control subjects, we admit concede that our choice of study design has reduced the statistical power of the study and might have undermined the conclusions.

HLA-DQAs belong to the HLA class II molecules that form α-β heterodimers on the cell surface of antigen-presenting cells, such as macrophages, dendritic cells and B lymphocytes. HLA-DQ genes play a central role in immune-mediated diseases by presenting peptides derived from extracellular proteins to CD4-positive T lymphocytes. Both HLA-DQB1 and HLA-DQA1 are highly polymorphic, especially in exon 2, which encodes antigen-binding sites and that somewhat different effects of rs9275572 between HCV genotypes might reflect differences in binding specificities between the HLA-DQ pocket and HCV genotype-specific epitopes (Table S4B). Finally, we found that a single amino acid polymorphism (R55P) corresponding to DQB1*03,
a very common allele worldwide [28], has the strongest protective effect against HCV persistence. Among DQB1*03 alleles, DQB1*0301 has the highest frequency various populations worldwide and has been reported in association with viral clearance in several European and American studies [7,13]. A suggestive paper by Cangussu et al. found an association between protection against HCV persistence and not only DQB1*0301 but also DQB1*0302 and DQB1*0303 in a white Brazilian cohort [29]. They suggested that HLA-DQ beta molecules responsible for protection are present not only in molecules encoded by DQB1*0301 but also by DQB1*0302 and DQB1*0303, and the common polymorphic residues shared by DQB1*03 alleles are responsible for the selection of particular HCV epitopes. In the current study, we identified that not only the allele with the strongest effect, DQB1*0303, but all DQB1*03 alleles share protective residues at position 55 in at least the Japanese and white Brazilian populations. Although the human MHC region encompasses a complex and extended LD structure, and LD among SNPs in the HLA-class II locus is known to vary among ethnic populations, various immunological functions of MHC molecules have gradually become clear. Several issues had to be overcome in order improve our understanding of whether or not the HLA-DQB1 molecule has common functions across multiple populations.

In this study, we also made an intriguing observation that HLA-DQB1*03 showed a stronger effect in individuals with the ss469415590-TT/TT genotype (Table 4), suggesting that the effect of HLA-DQB1*03 on HCV persistence might be affected by the genotype of the IFNL4 polymorphism. The fact that in the Japanese population the frequency of the ss469415590-TT/TT genotype is nearly equal to that of the rs12979860-CC genotype (JPT 78.7%, CEU 57.6%, AFR 15.4% based on the 1000 Genomes database), with a higher frequency than most other populations, might partially explain the differences among the top loci in European versus Japanese GWAS results. Finally, we determined the relationship between HLA-DQB1*03 (rs1130380) and IFNL4 variants (ss469415590) in cases and controls (Figure S7A and S7B). In cases, the proportion of individuals with the protective ss469415590-TT/TT genotype significantly increased while the proportion of protective DQB1*03 alleles decreased, but this tendency was not observed in controls. During any phase of HCV infection, these two polymorphisms might affect each other in complementary or exclusive fashion.

A recent GWAS reported that variants close to IFNL3 and DQB1*0301 are independently associated with spontaneous resolution of HCV infection [13]. On the other hand, our study newly suggests that a single amino acid polymorphism, common throughout the world, at position 55 on the HLA-DQB1 molecule has the strongest association with chronic HCV infection and may not be independent of the IFNL4 polymorphism.

In summary, we conducted a GWAS followed by two independent replication studies and detected genetic variants in the HLA-DQ genes strongly associated with chronic hepatitis C in the Japanese population. Our findings suggest that variation in antigen-binding sites involved in antigen presentation on HLA-DQ molecules might play an important role in HCV persistence or clearance. Further research is required to determine the distinct roles of HLA-DQB1 and IFNL4 in chronic HCV infection.

Materials and Methods

Study Population

Characteristics of each case-control group are shown in Table S10. Case samples used in this study were obtained between 2002 and 2012 from Toranomon Hospital Department of Hepatology, Sapporo Kosei General Hospital and Hiroshima University-affiliated hospitals. We used 481 samples from Toranomon Hospital as cases for the GWAS and the remaining 1,004 samples, along with 375 samples from Sapporo Kosei General Hospital, as cases for the second replication study. We used all 4,358 samples from Hiroshima University-affiliated hospitals as cases for the first replication study. All cases had abnormal levels of serum alanine aminotransferase for more than 6 months and were positive for both anti-HCV and serum HCV RNA. All patients were negative for hepatitis B surface antigen (HBsAg), had no evidence of other liver diseases, and had not received immunosuppressive therapy before enrollment. Control samples used for the GWAS and second replication studies were obtained from BioBank Japan at the Institute of Medical Science, University of Tokyo [30]. The control groups included 2,057 individuals for the GWAS and
25,817 independent samples for the second replication study and consisted of GWAS samples for unrelated diseases in the BioBank Japan project. We excluded samples that were positive for HBsAg or anti-HCV or that were registered as positive for any liver disease. We also obtained DNA from 906 Japanese control subjects from volunteers without liver disease in the Osaka-Midosuji Rotary Club, Osaka, Japan. In addition, we obtained 1,114 samples as healthy controls for the second replication study from volunteers without any liver diseases at Hiroshima University, Hiroshima, Japan.

Ethics statement
All subjects received a detailed explanation, and all signed a written informed consent form. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki and was approved a priori by the ethical committees at University of Tokyo, Hiroshima University, Toranomon Hospital and Sapporo Kosei General Hospital and by the Ethical Committee at the SNP Research Center, the Institute of Physical and Chemical Research (RIKEN), Yokohama.

SNP Genotyping and Quality Control
Genomic DNA was extracted from peripheral blood leukocytes using a standard method. For the GWAS stage, we genotyped 496 cases using the Illumina HumanHap610-Quad BeadChip (San Diego, CA, USA). We excluded 2 samples with call rates < 0.98. 13 other samples with apparent kinship or sample duplication were excluded from the analysis based on PI_HAT values (> 0.4). We assessed population stratification using the smartpcpa program in the EIGENSOFT package using SNPs informative for the Japanese population according to a previously described method [31]. Analysis was performed based on the GWAS data and the Japanese (JPT), Han-Chinese (CHB), European (CEU), and African (YRI) individuals from the International HapMap Project. Principal component analysis identified no outliers from the JPT/CHB clusters. 458,207 autosomal SNPs passed the quality control filters (call rate ≥ 0.99 in both cases and controls, a MAF > 0 and a Hardy-Weinberg equilibrium P value ≥ 1.0 × 10−6 in controls). We used multiplex-PCR-based Invader assays (Third Wave Technologies, Madison, WI, USA) for the replication studies [32]. Samples for both cases and controls were distributed randomly on genotyping plates in both phases of the study, and all persons performing genotyping and interpretation of results were blind to case / control status.

HLA-DQA1 and HLA-DQB1 Genotyping
We analyzed HLA-DQ genotypes using 1,114 cases and 1,114 controls in the first replication set. Exon 2 of the HLA-DQA1 and HLA-DQB1 genes were amplified and directly sequenced following the protocol of the International Histocompatibility Workshop Group [33]. HLA-DQA1 and DQB1 alleles were determined based on the alignment database of dbMHC (http://www.ncbi.nlm.nih.gov/gv/mhc/align.cgi?cmd=aligndisplay&locus_name=HLA-DQB11&banner=1#).

Protein Analysis
To represent the position of amino acid substitution caused by the SNP, we edited the image of a three-dimensional structure of HLA-DQ molecule that was previously determined by X-ray diffraction method [21] using Protein Data Bank Japan (PDBJ) Viewer (http://www.pdbj.org/index_j.html). The amino acid numbering excludes the 32 amino acid signal peptide.

Statistical Analysis
Genotype-based associations were tested using the Cochran-Armitage trend test. ORs and CIs were calculated from a two-by-two allele frequency table. Combined analysis was performed following the Mantel-Haenszel method. Heterogeneity among studies was examined using the Breslow-Day test. We used HaploView software to analyze the association of haplotypes and LD values between HLA-DQs and SNPs [34]. Chi-squared or Fisher exact tests were used to analyze categorical data, as appropriate. For general statistical analysis, we used the R statistical environment version 2.12.0 (http://www.r-project.org/) or PLINK 1.06 (http://pngu.mgh.harvard.edu/~purcell/plink/) [35]. Power analysis was performed using Power for Genetic Association Analyses software [36]. SNP allele frequencies were collected from the 1000 Genomes database (http://browser.1000genomes.org/).

Supporting Information

Figure S1. Results of the GWAS. (A) Quantile-quantile plot for the Cochran-Armitage trend tests for the GWAS phase. The horizontal axis represents P values expected under a null distribution, and the vertical axis shows the observed P values. Under the null hypothesis of no association at any locus, the points would be expected to fall along the line (y=x). (B) Manhattan plot showing the -log10 P value of each SNP calculated using the 1-d.f. Cochran-Armitage trend test. The red line shows the Bonferroni cutoff for genome-wide significance (P = 1.09×10−7) given the number of SNPs analyzed in this study (0.05/458207).

Figure S2. Case-control association results of the MHC region. P-value plot and genomic structure of the GWAS stage within the extended MHC region of chromosome 6. The black dotted line represents the SNP with the strongest association, rs9275572, which is located within the HLA-DQ locus.

Figure S3. Linkage disequilibrium structure around the HLA-DQ region based on D’ and r² using GWAS data. The red line represents the most strongly associated SNP, rs9275572, which is located within the HLA-DQB1 locus. The LD maps were created using HaploView software.

Figure S4. Linkage disequilibrium structure between rs1130380 and rs9275572 based on D’ and r² using data.
from 2,063 individuals (1,025 cases and 1,038 controls). The LD maps were created using HaploView software.

Figure S5. The position of the amino acid at β55 on the HLA-DQ molecule. An edited representation of the three-dimensional structure of HLA-DQ molecule that was previously determined by X-ray diffraction method [21] is shown. Protein Data Bank Japan (PDBj) Viewer (http://www.pdbj.org/index_j.html) was used for editing. The alpha and beta chains are represented by green and yellow, respectively. The amino acid at β55 is indicated as red. Amino acid numbering excludes the 32 amino acid signal peptide.

Figure S6. Chromatograms of different alleles of rs1130380 and their predicted effects on protein translation. Two different PCR amplifications were performed for genotyping of HLA-DQB1 alleles: amplicon 1 for the 02/03/04 alleles and amplicon 2 for the 05/06 alleles [33]. The G to C nucleotide change of codon 55 leads to amino acid substitution from Arg to Pro. The properties of Arg and Pro are basic and hydrophobic, respectively. We also observed the G to T nucleotide change, but this variant was not considered for further analysis because of its low frequency (< 1%).

Figure S7. Relationship between HLA-DQB1*03 (rs1130380) and IFNL4 variant (ss469415590) in chronic HCV patients (A) and healthy controls (B). Proportion of individuals with ss469415590 TT/TT in cases and controls according to HLA-DQB1*03 status. P values were calculated using the chi-squared test.

Table S1. Results of GWAS.

Table S2. Results of 1st replication study.

Table S3. Multiple logistic regression analysis for the risk of chronic hepatitis C using 1st replication samples (4,347 cases and 1,097 controls).

Table S4. (A) Effect of rs9275572 on the susceptibility of chronic hepatitis C stratified by gender and age using 1st replication samples. (B) Effect of rs9275572 on the susceptibility of chronic hepatitis C stratified by HCV genotype using 1st replication samples.

Table S5. (A) Useful variants for genotyping common HLA-DQA1 haplotypes in the Japanese population. (B) Useful variants for genotyping common HLA-DQB1 haplotypes in the Japanese population.

Table S6. Haplotype analysis was performed using the landmark SNP (rs9275572) and 12 variants shown in Table S5A and S5B by HaploView software.

Table S7. Single marker effect on chronic hepatitis C.

Table S8. Results of GWAS of two SNPs.

Table S9. Case-control association analysis of 5 variants under different genetic models.

Table S10. Basic characteristics of study population.

Acknowledgements

The authors thank the patients who agreed to participate in this study. We also thank team members at Toranomon Hospital, Sapporo Kosei General Hospital, Hiroshima University Hospital, and Hiroshima Liver Study Group for clinical sample collection. We thank Yasufumi Hayashida, Kana Izumoto and Tomoko Imai for technical assistance; Junko Sakamiya and Mika Tsuzuno for clerical assistance; and other members of the RIKEN Center for Integrative Medical Sciences and Hiroshima University for assistance with various aspects of this study.

Author Contributions

Conceived and designed the experiments: DM HO MK NK YN KC. Performed the experiments: DM HO YU H. Abe MK. Analyzed the data: DM HO YU H. Abe MK. Contributed reagents/materials/analysis tools: TK MT NH MI Y. Karino H. Aikata ST NA FS KI HK Y. Kawakami JT MK NK YN KC. Wrote the manuscript: DM HO CNH KC.

References


