



OPEN

Polymorphism at rs9264942 is associated with HLA-C expression and inflammatory bowel disease in the Japanese

Hiroshi Suzuki¹, Satoru Joshita¹✉, Atsuhiko Hirayama^{1,5}, Akihiro Shinji³, Kenji Mukawa², Minako Sako⁴, Naoki Yoshimura⁴, Tomoaki Suga¹, Takeji Umemura^{1,6}, Norihiro Ashihara¹, Tomoo Yamazaki¹ & Masao Ota¹

An expression quantitative trait locus (eQTL) single-nucleotide polymorphism (SNP) at rs9264942 was earlier associated with human leukocyte antigen (HLA)-C expression in Europeans. HLA-C has also been related to inflammatory bowel disease (IBD) risk in the Japanese. This study examined whether an eQTL SNP at rs9264942 could regulate HLA-C expression and whether four SNP haplotypes, including the eQTL SNP at rs9264942 and three SNPs at rs2270191, rs3132550, and rs6915986 of IBD risk carried in the HLA-C*12:02~B*52:01~DRB1*15:02 allele, were associated with IBD in the Japanese. HLA-C expression on CD3e⁺CD8a⁺ lymphocytes was significantly higher for the CC or CT genotype than for the TT genotype of rs9264942. The TACC haplotype of the four SNPs was associated with a strong susceptibility to ulcerative colitis (UC) but protection against Crohn's disease (CD) as well as with disease clinical outcome. While UC protectivity was significant but CD susceptibility was not for the CGTT haplotype, the significance of UC protectivity disappeared but CD susceptibility reached significance for the CGCT haplotype. In conclusion, our findings support that the eQTL SNP at rs9264942 regulates HLA-C expression in the Japanese and suggest that the four SNPs, which are in strong linkage disequilibrium, may be surrogate marker candidates of a particular HLA haplotype, HLA-C*12:02~B*52:01~DRB1*15:02, related to IBD susceptibility and disease outcome.

Inflammatory bowel disease (IBD) is a chronic or remission-relapse inflammatory disease of the gastrointestinal tract. IBD includes two types of chronic gut disorders: ulcerative colitis (UC) and Crohn's disease (CD)¹. UC primarily affects the mucous membranes and often forms erosions and ulcers in the colon, which can lead to inflammation of the rectum and extend proximally. There are three UC phenotypes in the clinical setting: proctitis, left-sided colitis, and pancolitis². CD produces granulomatous inflammation with ulceration and fibrosis mainly in the small intestine, but also anywhere in the digestive tract. Long-term severe intestinal inflammation can result in ulcers, constriction, and perforation of the intestinal tract. Indeed, IBD chronically injures the digestive tract, leading to tissue damage, function loss, disability, and systemic inflammation¹.

The exact etiology of IBD remains unknown, although a multifactorial pathogenesis that includes genetic, immunological, environmental, and microbial factors are likely involved³⁻⁷. Specifically, the influence of genetic factors on IBD pathogenesis has been supported by the high concordance rate among monozygotic twins as well as the high relative risk in affected siblings^{8,9}.

To date, genome-wide association studies and the candidate-gene approach have identified more than 240 IBD susceptibility genes or loci outside of the human leukocyte antigen (HLA) region, including *nucleotide*

¹Department of Medicine, Division of Gastroenterology and Hepatology, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan. ²Department of Gastroenterology, Japanese Red Cross Society Suwa Red Cross Hospital, Suwa, Japan. ³Department of Medical Oncology, Japanese Red Cross Society Suwa Red Cross Hospital, Suwa, Japan. ⁴Center for Inflammatory Bowel Disease, Tokyo Yamate Medical Center, Tokyo, Japan. ⁵Department of Inflammatory Bowel Disease, Yokohama City University Medical Center, Yokohama, Japan. ⁶Department of Life Innovation, Institute for Biomedical Sciences, Shinshu University, Matsumoto, Japan. ✉email: joshita@shinshu-u.ac.jp

Characteristic	Genotype of SNP at rs9264942				p value
	All (n = 32)	CC (n = 7)	CT (n = 12)	TT (n = 13)	
Age, median years (first-third quartile)	35 (26–67)	40 (27–58)	40 (29–67)	31 (26–45)	0.094
Male / female	16/16	2/5	4/8	10/3	0.012

Table 1. Age and gender proportions of healthy subjects for HLA-C expression analysis. *HLA* human leukocyte antigen, *SNP* single-nucleotide polymorphism.

oligomerization domain 2, *autophagy related 16-like 1*, *interleukin-23*, *PR domain-containing 1*, and *caspase recruitment domain 9*^{3,10–17}. Among the many candidate genes proposed, however, the HLA region has been consistently and strongly associated with IBD onset. HLA class I proteins play a crucial role in human immune responses^{18–20}. High expression levels of HLA-C molecules, which contribute to viral control in HIV-infected individuals, reportedly have a deleterious effect on CD in patients of European descent²¹. Moreover, among several single-nucleotide polymorphisms (SNPs) associated with HLA-C expression^{22,23}, a SNP located 35 kb upstream of the coding region of the *HLA-C* gene in chromosome 6 (-35C/T; rs9264942) was suggested to play an important role in controlling HLA-C cell surface molecule expression as an expression quantitative trait locus (eQTL) in subjects of European ancestry²⁴. However, no such association studies on Japanese IBD patients have been published²⁴.

Including HLA-DQB1, HLA-DRB1, and HLA-DQA1, the HLA class II region has also been associated with IBD onset across human groups^{7,25,26}. A particular haplotype, HLA-C*12:02~B*52:01~DRB1*15:02, was related to an increased risk for UC but a reduced risk for CD in a Japanese population²⁵. Specifically, both *HLA-B*52:01* and *HLA-DRB1*15:02* alleles were directly associated with UC susceptibility and CD protection in the Japanese, whereas HLA-C genetic involvement in IBD remained unknown²⁵. Since HLA genes are the most highly polymorphic in the human genome, a high-resolution HLA and SNP haplotype map analysis was developed for disease association studies^{27–29}. Based on that dataset and Okada's report²⁵, three SNPs at rs2270191, rs3132550, and rs6915986 have strong linkage disequilibrium (LD) with the *HLA-C*12:02* ($r^2 = 1$), *HLA-B*52:01* ($r^2 = 0.94$), and *HLA-DRB1*15:02* ($r^2 = 0.89$) alleles, respectively, in Japanese populations^{22,25} (Supplementary Fig. 1). However, no association studies of these tag SNPs have been reported in terms of Japanese IBD.

The aim of this study was to determine whether the eQTL SNP at rs9264942 could regulate HLA-C expression in the Japanese and analyze the relationships of SNPs in strong LD of a particular HLA haplotype with UC and CD susceptibility. It sought to uncover important findings on IBD from the in vivo, genetic, and functional aspects in well-defined patient groups and controls.

Results

Comparisons of HLA-C expression on peripheral blood mononuclear cells by the eQTL rs9264942 SNP genotype in healthy Japanese subjects.

A total of 32 healthy control subjects were included for the analysis of HLA-C expression on peripheral blood mononuclear cells (PBMC) (Table 1). The gating strategy for PBMC is shown in Fig. 1a. Although the cell surface expression of HLA-C on CD3e⁺CD8a⁺ T lymphocytes (Fig. 1b) as detected by flow cytometry in healthy subjects was comparable between males and females (Fig. 1e), it was significantly higher for the CC or CT genotype than for the TT genotype at rs9264942. HLA-C expression on CD3e⁺CD8a⁺ T lymphocytes (Fig. 1f), macrophages (Fig. 1c,g), and neutrophils (Fig. 1d,h) were significantly higher for the CC or CT genotype than for the TT genotype. Since another SNP at rs2395471 was also reported to be an eQTL SNP of HLA-C²³, we compared HLA-C expression for the AA or AG genotype and the GG genotype at rs2395471 on PBMC. HLA-C expression on CD3e⁺CD8a⁺ T lymphocytes (Fig. 1i) was significantly higher for the AA or AG genotype than for the GG genotype at rs2395471, while HLA-C expression on macrophages (Fig. 1j) and neutrophils (Fig. 1k) were comparable between the groups.

Association of the eQTL rs9264942 SNP of HLA-C expression with IBD susceptibility.

A total of 160 patients with UC and 275 patients with CD along with 325 healthy subjects were enrolled for this association study (Table 2). The C allele frequency of the eQTL SNP at rs9264942, which has been associated with HLA-C expression, was significantly higher in UC patients than in the healthy control group (UC vs. controls: 49.7% vs. 35.5%, odds ratio [OR] 1.79; $pc = 9.52 \times 10^{-4}$) (Table 3). In terms of clinical findings, there were no significant differences for UC between the CC or CT genotype and the TT genotype (Supplementary Table 1). There was also no remarkable difference for the C allele frequency of the rs9264942 SNP between the CD and healthy control groups (OR 1.00; $pc = 1.000$) (Table 3).

Association of three SNPs at rs2270191, rs3132550, and rs6915986 with IBD susceptibility.

The allele frequencies of three SNPs at rs2270191, rs3132550, and rs6915986 in patients with UC and CD and in healthy and primary biliary cholangitis (PBC) disease controls are shown in Table 3. The frequency of the T allele at rs2270191 in strong LD with the *HLA-C*12:02* allele ($r^2 = 1$), the A allele at rs3132550 in strong LD with the *HLA-B*52:01* allele ($r^2 = 0.94$), and the C allele at rs6915986 in strong LD with the *HLA-DRB1*15:02* allele ($r^2 = 0.89$) were significantly higher in UC patients than in the healthy group but significantly lower in CD patients than in healthy controls (Table 3). These allele frequencies were comparable between healthy and PBC disease controls (Table 3).

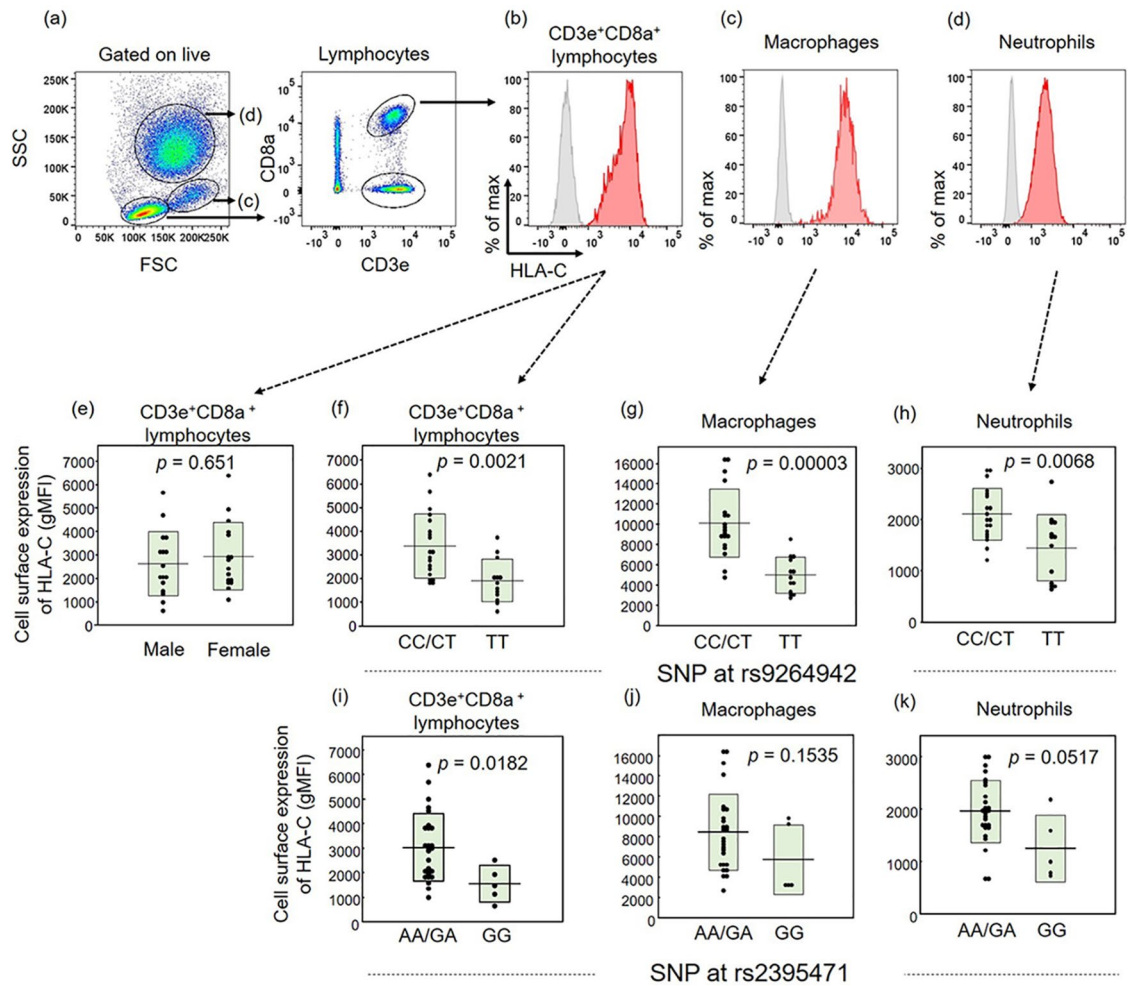


Figure 1. HLA-C expression on PBMC by flow cytometry analysis. Representative gating strategy (a). Representative HLA-C expression on CD3e⁺CD8a⁺ lymphocytes (b), macrophages (c), and neutrophils (d) (red) in relation to the isotype control (gray). Comparisons of geometric mean fluorescence intensity quantification of the cell surface expression of HLA-C on CD3e⁺CD8a⁺ lymphocytes between males and females (e) and between the SNP of the CC or CT genotype and of the TT genotype at rs9264942 on CD3e⁺CD8a⁺ lymphocytes (f), macrophages (g), and neutrophils (h) as well as between the SNP of the AA or AG genotype and the GG genotype at rs2395471 on CD3e⁺CD8a⁺ T lymphocytes (i), macrophages (j), and neutrophils (k).

Haplotype analysis of the SNPs at rs2270191, rs3132550, and rs6915986. Haplotype frequencies were calculated using an expectation–maximization algorithm for the three SNPs (rs2270191, rs3132550, and rs6915986) and compared among UC, CD, PBC, and healthy control groups. The three-SNP haplotype of TAC, which was predicted as associated with the particular HLA haplotype of HLA-C*12:02~B*52:01~DRB1*15:02, showed a strong susceptibility correlation between UC and controls (OR 2.53; $p_c = 3.92 \times 10^{-7}$) but a protective association between CD and controls (OR 0.50; $p_c = 0.002$) (Table 4). In contrast, the three-SNP haplotype of CGT showed a strong protectivity correlation between UC and controls (OR 0.39; $p_c = 1.58 \times 10^{-8}$) but a significant susceptibility association between CD and controls (OR 1.90; $p_c = 1.19 \times 10^{-3}$) (Table 4). The three-SNP haplotypes of CGT and TAC accounted for the vast majority of the cohort, which indicated that the SNPs were in strong LD with each other, as shown in Fig. 2. Moreover, these three-SNP haplotype frequencies were comparable between healthy and PBC disease controls (Table 4).

Logistic regression analysis of SNPs associated with IBD susceptibility. The strong LD of the three SNPs at rs2270191, rs3132550, and rs6915986 indicated that simultaneous logistic regression analysis to determine which SNP was independently associated with IBD susceptibility was inappropriate due to multi-collinearity. However, since rs9264942 SNP distribution was different from that of the other three SNPs regardless of being within a haplotype block (Fig. 2 and Supplementary Fig. 1), the SNP at rs9264942 was compared with the representative SNP at rs3132550 by logistic regression analysis. The SNP at rs3132550 was independently associated with UC susceptibility [OR 3.507, 95% confidence interval (95% CI) 2.176–5.654; $p < 0.00001$], whereas the SNP at rs9264942 did not reach statistical significance (OR 1.236, 95% CI 0.744–2.054; $p = 0.413$).

Characteristic	UC	CD	PBC	Control	UC versus control	CD versus control	PBC versus control
	(n = 160)	(n = 275)	(n = 328)	(n = 325)	p value	p value	p value
Age, years	48 (19–84)	41 (13–87)	71 (18–95)	47 (29–87)	0.880	0.157	3.83×10^{-72}
Age of onset, median years (first-third quartile)	34 (9–76)	23 (7–64)					
Male/female	83/77	190/85	53/275	90/235	1.72×10^{-7}	4.16×10^{-24}	3.70×10^{-4}
Disease phenotype							
Location*	31/55/73/1	45/48/182/0					
PSL administration	84 (53.2%)	108 (39.3%)					
Anti-TNF- α administration	47 (29.7%)	150 (54.5%)					
Surgical history	12 (7.5%)	176 (64%)					
Presence of anal lesion	-	177 (64.4%)					
Family history of IBD	5 (3.1%)	26 (9.5%)					
Liver cirrhosis			48 (14.6%)				
ANA ($\times 40 <$)			209 (63.7%)				
AMA			265 (80.8%)				

Table 2. Demographic and clinical data of UC, CD, PBC and healthy subjects. *Defined as colonic/ileal/ileocolonic/not determined in CD and proctitis/left-sided colitis/pancolitis/not determined in UC. UC ulcerative colitis, CD Crohn's disease, PBC primary biliary cholangitis, PSL prednisolone, ANA anti-nuclear antibody, AMA anti-mitochondrial antibody.

SNP	Position	Allele	UC (n = 160)	CD (n = 275)	Control (n = 325)	PBC (n = 328)	UC versus control	CD versus control	PBC versus control
			Minor allele frequencies				pc-value OR (95% CI)	pc-value OR (95% CI)	pc-value OR (95% CI)
rs2270191	31112543	T<C	0.256	0.067	0.117	0.111	1.16×10^{-7} 2.61 (1.85–3.69)	0.014 0.53 (0.36–0.82)	1.000 0.95 (0.68–1.34)
rs3132550	31118271	A<G	0.269	0.069	0.120	0.117	2.17×10^{-8} 2.70 (1.92–3.81)	0.012 0.55 (0.36–0.82)	1.000 0.98 (0.7–1.37)
rs9264942	31306603	C<T	0.497	0.355	0.355	0.352	9.52×10^{-4} 1.79 (1.37–2.35)	1.000 1.00 (0.79–1.26)	1.000 0.99 (0.79–1.24)
rs6915986	32716121	C<T	0.231	0.058	0.112	0.099	4.28×10^{-6} 2.61 (1.85–3.69)	0.004 0.49 (0.32–0.75)	1.000 0.87 (0.61–1.24)

Table 3. Association of four SNPs with IBD susceptibility. SNP single-nucleotide polymorphism, IBD inflammatory bowel disease, UC ulcerative colitis, CD Crohn's disease, PBC primary biliary cholangitis, pc corrected p value (Bonferroni correction), OR odds ratio, CI confidence interval. p values were calculated by chi-squared test 2×2 contingency tables.

Moreover, the SNP at rs3132550 was independently associated with CD protectivity (OR 0.478, 95% CI 0.298–0.767; $p = 0.00221$), unlike the SNP at rs9264942 (OR 1.376, 95% CI 0.966–1.961; $p = 0.077$).

Haplotype analysis of the SNPs at rs2270191, rs3132550, rs9264942, and rs6915986. Haplotype frequencies were calculated using an expectation–maximization algorithm for the four SNPs (rs2270191, rs3132550, rs9264942, and rs6915986) and compared among UC, CD, PBC, and healthy control groups. The four-SNP haplotype of TACC, which was predicted as associated with the particular HLA haplotype of HLA-C*12:02~B*52:01~DRB1*15:02, showed a strong susceptibility correlation between UC and controls (OR 2.53; $pc = 3.96 \times 10^{-7}$), but a protective association between CD and controls (OR 0.50; $pc = 0.002$) (Table 4). The four-SNP haplotype of TATC could not be calculated, which indicated that all subjects carrying the particular HLA haplotype of HLA-C*12:02~B*52:01~DRB1*15:02 had the TACC haplotype, i.e., the SNP at rs9264942 of subjects carrying the HLA haplotype of HLA-C*12:02~B*52:01~DRB1*15:02 was the C allele (CC or CT genotype).

Haplotype	UC (n = 160)	CD (n = 275)	Control (n = 325)	PBC (n = 328)	UC versus control	CD versus control	PBC versus control
	Haplotype frequencies				<i>pc</i> -value OR (95% CI)	<i>pc</i> -value OR (95% CI)	<i>pc</i> -value OR (95% CI)
3 SNPs*							
CGT	0.721	0.927	0.869	0.873	1.58×10^{-8} 0.39 (0.28–0.54)	1.19×10^{-3} 1.90 (1.28–2.82)	0.835 1.04 (0.75–1.43)
TAC	0.221	0.053	0.101	0.090	3.92×10^{-7} 2.53 (1.76–3.65)	0.002 0.50 (0.32–0.78)	0.487 0.88 (0.61–1.27)
TAT	0.035	0.013	0.015	0.022	0.055 2.29 (0.96–5.44)	0.708 0.83 (0.31–2.20)	0.421 1.40 (0.62–3.18)
CGC	0.010	0.004	0.011	0.009	0.843 0.87 (0.22–3.40)	0.156 0.34 (0.07–1.63)	0.773 0.85 (0.28–2.55)
CAT	0.012	0.004	0.003	0.006	0.078 4.11 (0.75–22.6)	0.862 1.19 (0.17–8.48)	0.417 1.99 (0.36–10.9)
4 SNPs**							
CGTT	0.498	0.645	0.639	0.645	2.66×10^{-5} 0.56 (0.42–0.73)	0.828 1.03 (0.81–1.30)	0.828 1.03 (0.82–1.29)
CGCT	0.224	0.284	0.231	0.228	0.803 0.96 (0.70–1.32)	0.042 1.31 (1.01–1.70)	0.933 0.98 (0.76–1.27)
TACC	0.221	0.053	0.101	0.090	3.96×10^{-7} 2.53 (1.77–3.62)	0.002 0.50 (0.32–0.78)	0.486 0.88 (0.60–1.27)
TACT	0.035	0.013	0.015	0.022	0.053 2.28 (0.98–5.30)	0.705 0.82 (0.31–2.17)	0.418 1.40 (0.62–3.18)

Table 4. Association of the three-SNP* haplotype or four-SNP* haplotype with IBD susceptibility. Haplotypes were created by an accelerated EM algorithm in Haploview software. SNP single-nucleotide polymorphism, IBD inflammatory bowel disease, UC ulcerative colitis, CD Crohn's disease, PBC primary biliary cholangitis, OR odds ratio, CI confidence interval. *SNPs at rs2270191, rs3132550, and rs6915986. **SNPs at rs2270191, rs3132550, rs9264942, and rs6915986.

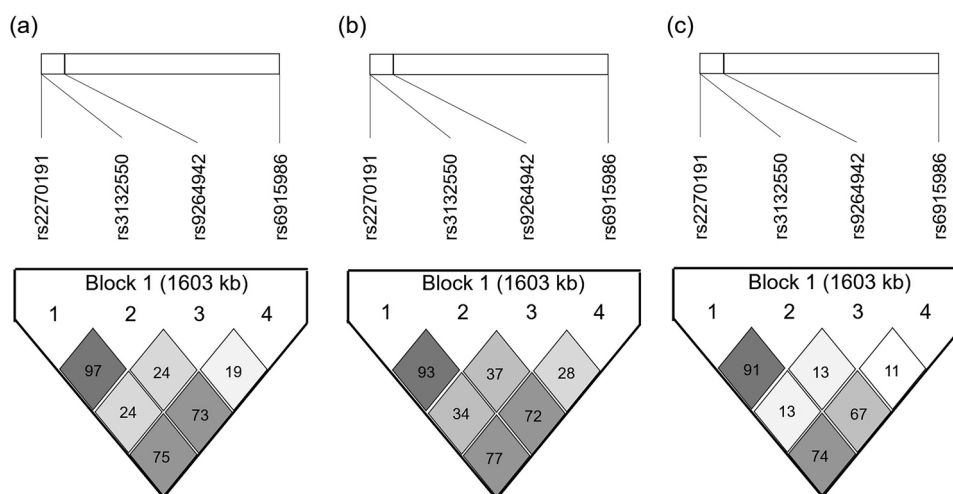


Figure 2. LD plots of four SNPs at rs2270171, rs3132550, rs9264942, and rs6915986 of healthy controls (a), ulcerative colitis patients (b) and Crohn's disease patients (c). Values of r^2 corresponding to each pair are expressed as a percentage and shown within the respective square.

The four-SNP haplotype of CGTT in individuals with the SNP T allele at rs9264942 showed a protective correlation between UC and controls (OR 0.56; $pc = 2.66 \times 10^{-5}$), whereas the four-SNP haplotype of CGCT in individuals with the SNP C allele at rs9264942 displayed no significance ($pc = 0.803$). On the other hand, the four-SNP haplotype of CGTT showed no significance between CD and controls ($pc = 0.828$), while the four-SNP haplotype of CGCT exhibited significant susceptibility (OR 1.31; $pc = 0.042$). The above four-SNP haplotype frequencies were comparable between healthy and PBC disease controls (Table 4).

Comparisons of clinical findings of the 4-SNP haplotype for IBD. There were no significant differences for any clinical findings between UC patients with and without the TACC haplotype (Supplementary Table 2). In contrast, significant differences were observed for male frequency (OR 2.30, 95% CI 1.06–5.01; $p = 0.032$), disease phenotype location of ileocolonic pattern (OR 3.17, 95% CI 1.44–6.96; $p = 0.003$), intestinal

	TACC-negative (n = 246)	TACC-positive (n = 29)	p value	OR (95% CI)
Male, (%)	175 (71.1%)	15 (51.7%)	0.032	2.30 (1.06–5.01)
Age, years	41 (13–77)	40 (22–87)	0.116	–
Disease duration	15 (1–51)	12 (3–34)	0.116	–
Disease phenotype location: colonic, ileal, ileocolonic, and not determined	36 (14.6%), 40 (16.3%), 170 (69.1%), 0 (0%)	9 (31%), 8 (27.2%), 12 (41.4%), 0 (0%)	0.003	–
Ileocolonic	170 (69.1%)	12 (41.4%)	0.003	3.17 (1.44–6.96)
Age of onset	23 (7–62)	23 (15–64)	0.412	–
Family history of IBD	23 (9.3%)	3 (10.3%)	0.862	0.89 (0.25–3.18)
Medical therapy				
5-ASA	244 (99.2%)	29 (100%)	0.626	–
PSL	97 (39.4%)	11 (37.9%)	0.876	1.07 (0.48–2.35)
Anti-TNF- α administration	134 (54.5%)	16 (55.2%)	0.943	0.97 (0.45–2.11)
Anti-IL-12/23p40	7 (2.8%)	2 (6.9%)	0.246	0.40 (0.08–2.0)
AZA or 6-MP	124 (50.4%)	18 (62.1%)	0.235	0.62 (0.28–1.37)
Anal surgery history	116 (47.2%)	11 (37.9%)	0.346	1.46 (0.66–3.22)
History of intestinal surgery	165 (67.1%)	11 (37.9%)	0.002	3.33 (1.50–7.39)
Balloon expansion	22 (8.9%)	1 (3.4%)	0.312	2.75 (0.36–21.2)
Disease complication				
Presence of anal lesion	161 (65.4%)	16 (55.2%)	0.275	1.54 (0.71–3.35)
Intestinal complication	187 (76.0%)	14 (48.3%)	0.001	3.40 (1.55–7.45)
Extra-intestinal complication	46 (18.7%)	4 (13.8%)	0.517	1.44 (0.48–4.33)
Intestinal carcinoma	2 (0.8%)	0 (0%)	0.626	–

Table 5. Relationship between haplotypes from four SNPs (rs2270171, rs3132550, rs9264942, and rs6915986) and clinical findings in CD. Data are expressed as the number (%) except for age, which is expressed as the median (first-third quartile). SNP single-nucleotide polymorphism, CD Crohn's disease, OR odds ratio, CI confidence interval, IBD inflammatory bowel disease, 5-ASA 5-aminosalicylic acid, PSL prednisolone, CAP cytapheresis, AZA azathioprine, 6-MP 6-mercaptopurine, Tac tacrolimus, CyA cyclosporin.

complication frequency (OR 3.40, 95% CI 1.55–7.45; $p=0.001$), and frequency of history of intestinal surgery (OR 3.33, 95% CI 1.50–7.39; $p=0.002$) between CD patients with and without the TACC haplotype (Table 5).

Discussion

This study clearly demonstrated a significant association of an eQTL SNP of HLA-C with the regulation of HLA-C expression on PBMC in Japanese subjects. Moreover, four SNP haplotypes consisting of the eQTL SNP of HLA-C and three SNPs of HLA-C*12:02~B*52:01~DRB1*15:02 were significantly related to susceptibility to UC but protection against CD.

The cell surface expression of HLA-C is regulated through multiple mechanisms³⁰. Located 35 kb upstream of the *HLA-C* gene (Supplementary Fig. 1), the SNP variant at rs9264942 has been associated with HLA-C surface expression in subjects of European-descent^{24,31,32}, but not in an African American population²¹, suggesting that the function of the eQTL SNP at rs9264942 in HLA-C expression varies among population groups. Therefore, we examined for and identified a significant association of the eQTL SNP at rs9264942 with HLA-C expression on PBMC, specifically on CD3e⁺CD8a⁺ lymphocytes, macrophages, and neutrophils, in the Japanese. HLA-C cell surface expression differed among cell populations, which supported previous findings of relatively low HLA-C expression on lymphocytes^{33–35} and high expression on antigen-presenting cells, such as macrophages³⁶. The increased expression of HLA-C molecules promotes antigen presentation and recognition by CD8a⁺ cytotoxic T cells and natural killer cells for host protection, but in some instances leads to multiple diseases. HLA-C surface expression is also regulated by microRNA miR-148a binding on the 3' untranslated region (3'UTR) of the *HLA-C* gene^{37,38}. Therefore, interactions between the eQTL SNP of HLA-C at rs9264942 and miR-148a binding on the 3'UTR region of the *HLA-C* gene merit future consideration to uncover the precise genetic and molecular mechanisms of HLA-C expression. We also compared the HLA-C expression on PBMC between the AA or AG genotype and the GG genotype at rs2395471, which was reported as another eQTL SNP of HLA-C²³. HLA-C expression on CD3e⁺CD8a⁺ T lymphocytes was significantly higher for the AA or AG genotype than for the GG genotype at rs2395471, but no differences were observed on macrophages and neutrophils. Additional molecular mechanisms of HLA-C expression regulation on PBMC subsets by eQTL SNP differences require assessment in future studies.

Okada et al. reported that the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype was associated with susceptibility to UC as well as protection against CD in the Japanese²⁵. If patients had the particular three-SNP haplotype of TAC at rs2270191, rs3132550, and rs6915986, this haplotype showed very similar sensitivity to UC as did the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype²⁵. Moreover, the TAC haplotype frequency was low in CD patients, which supported a similar findings for the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype frequency

in a previous report²⁵. Our result indicated that three-SNP haplotype assignments could serve as surrogate markers for disease susceptibility or protectivity to IBD instead of HLA ones. Although the HLA-C eQTL SNP at rs9264942 was located on this haplotype, the r^2 values for each SNP were relatively high apart from those between the SNP at rs9264942 and the others of less than 30. This indicated that LD was broken only by rs9264942 and therefore the possibility that the eQTL SNP at rs9264942 had an independent function in terms of IBD susceptibility. Thus, we analyzed the relationship of four-SNP haplotypes with IBD risk. The SNP at rs9264942 was related to CD susceptibility in European patients by the regulation of HLA-C expression^{21,24}. We detected an association of the TACC haplotype at rs9264942, but not the TATC haplotype, which suggested that IBD susceptibility might not be related to the eQTL SNP at rs9264942 if patients carried the three SNPs of the TAC haplotype, thereby depending on a particular HLA-C*12:02~B*52:01~DRB1*15:02 haplotype²⁵. Reciprocally, the CGTT haplotype in patients with the T allele at rs9264942 (low HLA-C expression group) was associated with UC protection, whereas the CGCT haplotype in patients with the C allele at rs9264942 (high HLA-C expression group) exhibited no remarkable association (Table 4). On the other hand, the CGTT haplotype was not associated with CD, although the CGCT haplotype had a significant association with CD susceptibility. This supported a previous paper describing that high HLA-C expression was related to CD susceptibility²¹. Attempts to confirm the independent involvement of the SNP at rs9264942 in IBD susceptibility or protection by stratified analysis using the Mantel-Haenszel test were inconclusive (data not shown), which might have been due to a relatively small number of subjects or a stronger association of a particular haplotype. Larger studies are needed to validate our findings.

Regarding disease phenotype, it was interesting that CD patients without the TACC haplotype had significantly more intestinal complications (intestinal stenosis, perforation, and fistula) and history of intestinal surgery, suggesting that the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype might impart a stronger effect on clinical results. Based on the above findings, these four SNPs may be simple but useful surrogate markers for predicting HLA-C*12:02~B*52:01~DRB1*15:02 haplotypes as well as disease outcome. Other HLA class II alleles have been associated with IBD onset among populations, such as *HLA-DRB1*01:03* with disease predisposition to CD and UC in Europeans, but not in East-Asian Backgrounds^{39,40}. Further studies are required to identify the direct mechanisms of involvement of these genetic variants on disease susceptibility and outcome.

The present study had several limitations. First, it evaluated HLA-C expression on the PBMC of healthy controls. As the number of normal controls was too small for a definitive conclusion, a larger validation analysis with more subjects is needed to confirm our results. However, earlier studies have used as few as 30 healthy controls³¹. HLA-C expression in tissue samples should also be analyzed to confirm the molecular mechanisms of HLA-C involvement in IBD. Second, this investigation was preliminary in nature because the numbers of cases and controls were limited. However, we adopted a PBC disease control cohort for the association analysis, whereby the IBD cohort was significantly different from healthy controls as well as from PBC patient controls regarding genetic frequency. Additional investigation is required to validate these newly discovered associations in a larger number of individuals; however, power calculations based on the study subjects of 160 UC patients and 275 CD patients and an effect size of 0.28 at rs9264942 showed sufficient detection power (0.99) at the 0.05 level of significance⁴¹. Moreover, the calculated type I error and type II were 0.001 and 0.004, respectively, which suggested that statistical error could be avoided. Third, the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype should be examined in more detail to strengthen the results of this study. Fourth, this investigation was retrospective in nature and did not include a segregation analysis; prospective longitudinal studies are needed to clarify the associations of genetic polymorphisms with IBD outcomes.

In conclusion, our findings supported that the eQTL SNP at rs9264942 regulated HLA-C expression in the Japanese. Our association study also implicated four SNPs in strong LD of a particular HLA haplotype, HLA-C*12:02~B*52:01~DRB1*15:02, with IBD susceptibility and outcome, which might be potential surrogate markers for the disease.

Patients and methods

Research ethics considerations. This study was conducted in accordance with the principles of the 1975 Declaration of Helsinki and approved by the ethics committees of each participating institution [Shinshu University School of Medicine, Matsumoto, Japan (approval numbers: 639 and 4533), Japanese Red Cross Society Suwa Red Cross Hospital, Suwa, Japan (approval number: 30-18), and Tokyo Yamate Medical Center, Tokyo, Japan (approval number: 188)]. Informed written consent was obtained from all patients and healthy subjects.

Study subjects. A total of 32 healthy control subjects were included for the analysis of HLA-C expression on PBMC (Table 1). The controls were volunteers from hospital staff who had indicated the absence of any major illnesses and no direct familial relations in a standard questionnaire.

A total of 160 patients with UC and 275 patients with CD who were seen between April 2014 and May 2019 at Shinshu University School of Medicine, Matsumoto, Japan, Japanese Red Cross Society Suwa Red Cross Hospital, Suwa, Japan, or Tokyo Yamate Medical Center, Tokyo, Japan, along with 325 healthy subjects including the above-described 32 healthy control subjects were enrolled for this association study (Table 2). A total of 328 PBC patients who were diagnosed as having PBC at Shinshu University Hospital between 1986 and 2015 were also included as a disease control group in this analysis. No participants were direct relatives of each other, and all were of an East-Asian genetic background.

The diagnoses of UC and CD were based on established Japanese Society of Gastroenterology guidelines using endoscopic, histologic, radiographic, and serologic findings⁴². We also sub-classified the IBD patients into disease parameters based on the phenotype of disease location (colonic, ileal, ileocolonic, and not determined in CD and proctitis, left-sided colitis, pancolitis, and not determined in UC) (Table 2)^{2,43,44}. The introduction of

prednisolone (PSL) and anti-TNF- α was recommended by the guidelines for moderate or severe UC or CD⁴². Oral or intravenous PSL was advised as a remission induction therapy for left-sided colitis or pancolitis in UC, while anti-TNF- α was advocated for induction or maintenance therapy. A PSL-positive status was designated for patients with a history of PSL administration. A PSL- and anti-TNF- α -positive status was designated for patients who had received both PSL and anti-TNF- α .

The diagnosis of PBC was based on the criteria from the Japan Society of Hepatology⁴⁵. This sub-cohort has been well characterized in prior genetic and clinical studies^{46–54}.

Flow cytometry analysis of HLA-C expression on CD3e⁺CD8a⁺ lymphocytes. PBMC were obtained one day after the hemolysis of red blood cells in ACK lysing buffer (Thermo Fisher Scientific K.K., Tokyo, Japan) according to the manufacturer's instructions. PBMC were incubated with antibodies that included 7-AAD (BioLegend, San Diego, USA), APC-anti-CD3e (BioLegend), FITC-anti-CD8a (BioLegend), and PE-anti-HLA-C (BD Biosciences) or PE-HLA-C-isotype control by IgG1 k (BioLegend) for 30 min on ice, washed twice with staining buffer, and then analyzed using a BD FACSCanto™ II flow cytometer (BD Biosciences, New Jersey, USA). HLA-C expression was determined by median fluorescence intensity on CD3e⁺CD8a⁺ lymphocytes, CD3e⁺CD8a⁻ lymphocytes, neutrophils, and macrophages in relation to the isotype control. The gating strategy is shown in Fig. 1a. Obtained data were analyzed by FlowJo software (BD Biosciences). All samples were simultaneously stained and analyzed in a one-day experiment.

SNP determination. Genomic DNA from all participants was isolated from whole blood samples using Quick Gene-610L assays (Fujifilm, Tokyo, Japan). The DNA concentration was adjusted to 10–15 ng/ μ L for SNP genotyping experiments. Two eQTL SNPs at rs9264942 and rs2395471 were selected based on a previous study^{22,23}. Three SNPs at rs2270191, rs3132550, and rs6915986, which were in strong LD with the HLA-C*12:02~B*52:01~DRB1*15:02 haplotype, were also adopted based on an earlier report²². The genotyping of five SNPs located at rs9264942, rs2395471, rs2270191, rs3132550, and rs6915986 was performed with a TaqMan 5 exonuclease assay using primers supplied by Applied Biosystems (Foster City, CA, USA). The probe's fluorescence signal was detected with a StepOne Plus Real-Time PCR System (Thermo Fisher Scientific) according to the manufacturer's instructions.

SNP haplotype analysis. Haploview version 4.2^{55,56} was used to evaluate the haplotype structure of the four SNPs at rs2270191, rs3132550, rs9264942, and rs6915986 (Fig. 2). Pairwise LD patterns and haplotype frequency analysis for all SNPs in patients and controls were analyzed by the block definition established by Gabriel et al.⁵⁷.

Statistical analysis. The significance of associations was evaluated using chi-squared analysis or Fisher's exact test. *p* values were subjected to Bonferroni's correction by multiplication by the number of different alleles observed in each locus (*pc*). The Mann–Whitney U-test was used to analyze continuous variables. Stepwise logistic regression analysis with a forward approach was performed to identify independent SNPs associated with IBD susceptibility. A two-sided *p* value of < 0.05 was considered statistically significant. Association strength was estimated by calculating the OR and 95% CI with StatFlex software (version 7.0.3, Artech, Osaka, Japan) as well as IBM SPSS Statistics version 23.0 (IBM, Chicago, IL, USA).

Received: 30 April 2020; Accepted: 8 July 2020

Published online: 24 July 2020

References

1. Abraham, C. & Cho, J. H. Inflammatory bowel disease. *N. Engl. J. Med.* **361**, 2066–2078 (2009).
2. Moran, C., Sheehan, D. & Shanahan, F. The changing phenotype of inflammatory bowel disease. *Gastroenterol. Res. Pract.* **2016**, 1619053 (2016).
3. Hugot, J. P. et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
4. Ivanov, I. I. et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).
5. Lee, J. C. & Lennard-Jones, J. E. Inflammatory bowel disease in 67 families each with three or more affected first-degree relatives. *Gastroenterology* **111**, 587–596 (1996).
6. Bridger, S., Lee, J. C., Bjarnason, I., Jones, J. E. & Macpherson, A. J. In siblings with similar genetic susceptibility for inflammatory bowel disease, smokers tend to develop Crohn's disease and non-smokers develop ulcerative colitis. *Gut* **51**, 21–25 (2002).
7. Jostins, L. et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
8. Halme, L. et al. Family and twin studies in inflammatory bowel disease. *World J. Gastroenterol.* **12**, 3668–3672 (2006).
9. Halfvarson, J., Bodin, L., Tysk, C., Lindberg, E. & Jarnerot, G. Inflammatory bowel disease in a Swedish twin cohort: A long-term follow-up of concordance and clinical characteristics. *Gastroenterology* **124**, 1767–1773 (2003).
10. Duerr, R. H. et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
11. Ellinghaus, D. et al. Association between variants of PRDM1 and NDP52 and Crohn's disease, based on exome sequencing and functional studies. *Gastroenterology* **145**, 339–347 (2013).
12. Rioux, J. D. et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat. Genet.* **39**, 596–604 (2007).
13. Lamas, B. et al. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* **22**, 598–605 (2016).

14. Park, S. C. & Jeen, Y. T. Genetic studies of inflammatory bowel disease-focusing on Asian patients. *Cells* **8**, 404 (2019).
15. de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* **49**, 256–261 (2017).
16. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
17. Yamazaki, K. *et al.* Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum. Mol. Genet.* **14**, 3499–3506 (2005).
18. Bashirova, A. A., Martin, M. P., McVicar, D. W. & Carrington, M. The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. *Annu. Rev. Genomics Hum. Genet.* **7**, 277–300 (2006).
19. Dendrou, C. A., Petersen, J., Rosjohn, J. & Fugger, L. HLA variation and disease. *Nat. Rev. Immunol.* **18**, 325–339 (2018).
20. Parham, P. & Moffett, A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat. Rev. Immunol.* **13**, 133–144 (2013).
21. Apps, R. *et al.* Influence of HLA-C expression level on HIV control. *Science* **340**, 87–91 (2013).
22. Kitajima, H., Sonoda, M. & Yamamoto, K. HLA and SNP haplotype mapping in the Japanese population. *Genes Immun.* **13**, 543–548 (2012).
23. Vince, N. *et al.* HLA-C level is regulated by a polymorphic Oct1 binding site in the HLA-C promoter region. *Am. J. Hum. Genet.* **99**, 1353–1358 (2016).
24. Thomas, R. *et al.* HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. *Nat. Genet.* **41**, 1290–1294 (2009).
25. Okada, Y. *et al.* HLA-Cw*1202-B*5201-DRB1*1502 haplotype increases risk for ulcerative colitis but reduces risk for Crohn's disease. *Gastroenterology* **141**, 864–71.e1–5 (2011).
26. Arimura, Y. *et al.* Characteristics of Japanese inflammatory bowel disease susceptibility loci. *J. Gastroenterol.* **49**, 1217–1230 (2014).
27. Miretti, M. M. *et al.* A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am. J. Hum. Genet.* **76**, 634–646 (2005).
28. de Bakker, P. I. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* **38**, 1166–1172 (2006).
29. Pillai, N. E. *et al.* Predicting HLA alleles from high-resolution SNP data in three Southeast Asian populations. *Hum. Mol. Genet.* **23**, 4443–4451 (2014).
30. Kulpa, D. A. & Collins, K. L. The emerging role of HLA-C in HIV-1 infection. *Immunology* **134**, 116–122 (2011).
31. Corrah, T. W. *et al.* Reappraisal of the relationship between the HIV-1-protective single-nucleotide polymorphism 35 kilobases upstream of the HLA-C gene and surface HLA-C expression. *J. Virol.* **85**, 3367–3374 (2011).
32. Charoudeh, H. N. *et al.* Quantity of HLA-C surface expression and licensing of KIR2DL+ natural killer cells. *Immunogenetics* **64**, 739–745 (2012).
33. Snary, D., Barnstable, C. J., Bodmer, W. F. & Crumpton, M. J. Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur. J. Immunol.* **7**, 580–585 (1977).
34. McCutcheon, J. A., Gumperz, J., Smith, K. D., Lutz, C. T. & Parham, P. Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. *J. Exp. Med.* **181**, 2085–2095 (1995).
35. Neisig, A., Melief, C. J. & Neefjes, J. Reduced cell surface expression of HLA-C molecules correlates with restricted peptide binding and stable TAP interaction. *J. Immunol.* **160**, 171–179 (1998).
36. Schaefer, M. R. *et al.* A novel trafficking signal within the HLA-C cytoplasmic tail allows regulated expression upon differentiation of macrophages. *J. Immunol.* **180**, 7804–7817 (2008).
37. Kulkarni, S. *et al.* Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature* **472**, 495–498 (2011).
38. Kulkarni, S. *et al.* Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc. Natl. Acad. Sci. USA* **110**, 20705–20710 (2013).
39. Han, B. *et al.* Amino acid position 37 of HLA-DRβ1 affects susceptibility to Crohn's disease in Asians. *Hum. Mol. Genet.* **27**, 3901–3910 (2018).
40. Goyette, P. *et al.* High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nat. Genet.* **47**, 172–179 (2015).
41. Faul, F., Erdfelder, E., Lang, A.-G. & Ag, B. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods Instrum. Comput.* **39**, 175–191 (2007).
42. Matsuoka, K. *et al.* Evidence-based clinical practice guidelines for inflammatory bowel disease. *J. Gastroenterol.* **53**, 305–353 (2018).
43. Bhat, M. *et al.* Phenotypic and genotypic characteristics of inflammatory bowel disease in French Canadians: comparison with a large North American repository. *Am. J. Gastroenterol.* **104**, 2233–2240 (2009).
44. Hirayama, A. *et al.* Lymphocyte antigen 75 polymorphisms are associated with disease susceptibility and phenotype in Japanese patients with inflammatory bowel disease. *Dis. Markers* **2016**, 6485343 (2016).
45. Guidelines for the management of primary biliary cirrhosis: The Intractable Hepatobiliary Disease Study Group supported by the Ministry of Health, Labour and Welfare of Japan. *Hepatol Res* **44** (Suppl S1), 71–90 (2014).
46. Joshita, S. *et al.* STAT4 gene polymorphisms are associated with susceptibility and ANA status in primary biliary cirrhosis. *Dis. Markers* **2014**, 727393 (2014).
47. Joshita, S. *et al.* Serum autotaxin is a useful disease progression marker in patients with primary biliary cholangitis. *Sci. Rep.* **8**, 8159 (2018).
48. Joshita, S. *et al.* Biochemical and plasma lipid responses to pemaifibrate in patients with primary biliary cholangitis. *Hepatol. Res.* **49**, 1236–1243 (2019).
49. Joshita, S. *et al.* Association analysis of cytotoxic T-lymphocyte antigen 4 gene polymorphisms with primary biliary cirrhosis in Japanese patients. *J. Hepatol.* **53**, 537–541 (2010).
50. Joshita, S. *et al.* Clinical utility of FibroScan as a non-invasive diagnostic test for primary biliary cholangitis. *J. Gastroenterol. Hepatol.* **35**, 1208–1214 (2019).
51. Morita, S. *et al.* Association analysis of toll-like receptor 4 polymorphisms in Japanese primary biliary cirrhosis. *Hum. Immunol.* **74**, 219–222 (2013).
52. Sekiguchi, T. *et al.* Serum cell death biomarkers for prediction of liver fibrosis and poor prognosis in primary biliary cirrhosis. *PLoS ONE* **10**, e0131658 (2015).
53. Umemura, T. *et al.* Human leukocyte antigen class II molecules confer both susceptibility and progression in Japanese patients with primary biliary cirrhosis. *Hepatology* **55**, 506–511 (2012).
54. Umemura, T. *et al.* Serum *Wisteria floribunda* agglutinin-positive Mac-2-binding protein level predicts liver fibrosis and prognosis in primary biliary cirrhosis. *Am. J. Gastroenterol.* **110**, 857–864 (2015).
55. Sole, X., Guino, E., Valls, J., Iñiesta, R. & Moreno, V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics* **22**, 1928–1929 (2006).
56. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
57. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).

Acknowledgments

The authors thank Yuki Akahane, Asami Yamazaki, and Naoya Sakuraba for their technical assistance, and Trevor Ralph for his English editorial assistance. We sincerely appreciate the research support provided in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant-in-Aid for Young Scientists (20K17019) and Shinshu Public Utility Foundation for Promotion of Medical Sciences.

Author contributions

Concept and design of study: H.S., S.J., and M.O. Collection of samples: H.S., A.H., A.S., K.M., M.S., N.Y., and T.S. Analysis and interpretation of data: H.S., S.J., T.U., N.A., T.Y., and M.O. Preparation of the manuscript: H.S., S.J., T.U., and M.O. Critical review: S.J., T.U., and M.O. Approval of final manuscript: all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69370-8>.

Correspondence and requests for materials should be addressed to S.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020