15. Gaynor JJ, Feuer EJ, Tan CC, *et al.* On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *J Am Stat Assoc.* 1993; **88**: 400–9.
16. Sacks D, McClenny TE, Cardella JF, Lewis CA. Society of interventional radiology clinical practice guidelines. *J Vasc Interv Radiol.* 2003; **14**: S199–202.
17. Lencioni R, Caramella D, Bartolozzi C. Response of hepatocellular carcinoma to percutaneous ethanol injection: CT and MR evaluation. *J Comput Assist Tomogr.* 1993; **17**: 723–9.
18. Lin SM, Lin CJ, Lin CC, Hsu CW, Chen YC. Radiofrequency ablation improves prognosis compared with ethanol injection for hepatocellular carcinoma < or = 4 cm. *Gastroenterology.* 2004; **127**: 1714–23.
19. Sala M, Llovet JM, Vilana R, *et al.* Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. *Hepatology.* 2004; **40**: 1352–60.
20. Llovet JM, Fuster J, Bruix J. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: resection versus transplantation. *Hepatology.* 1999; **30**: 1434–40.
21. Minagawa M, Makuuchi M, Takayama T, Kokudo N. Selection criteria for repeat hepatectomy in patients with recurrent hepatocellular carcinoma. *Ann Surg.* 2003; **238**: 703–10.
22. Poon RT, Fan ST, Lo CM, Liu CL, Wong J. Long-term survival and pattern of recurrence after resection of small hepatocellular carcinoma in patients with preserved liver function: implications for a strategy of salvage transplantation. *Ann Surg.* 2002; **235**: 373–82.
23. Johnson P, Bruix J. Hepatocellular carcinoma and the art of prognostication. *J Hepatol.* 2000; **33**: 1006–8.
24. Di Stasi M, Buscarini L, Livraghi T, *et al.* Percutaneous ethanol injection in the treatment of hepatocellular carcinoma. A multicenter survey of evaluation practices and complication rates. *Scand J Gastroenterol.* 1997; **32**: 1168–73.
25. Capussotti L, Muratore A, Amisano M, *et al.* Liver resection for hepatocellular carcinoma on cirrhosis: analysis of mortality, morbidity and survival–a European single center experience. *Eur J Surg Oncol.* 2005; **31**: 986–93.
26. Taketomi A, Kitagawa D, Itoh S, *et al.* Trends in morbidity and mortality after hepatic resection for hepatocellular carcinoma: an institute's experience with 625 patients. *J Am Coll Surg.* 2007; **204**: 580–7.
27. Imamura H, Seyama Y, Kokudo N, *et al.* One thousand fifty-six hepatectomies without mortality in 8 years. *Arch Surg.* 2003; **138**: 1198–206; discussion 206.
28. Shiina S, Tateishi R, Arano T, *et al.* Radiofrequency ablation for hepatocellular carcinoma: 10-year outcome and prognostic factors. *Am J Gastroenterol.* 2012; **107**: 569–77.
29. Lin SM, Lin CJ, Lin CC, Hsu CW, Chen YC. Randomised controlled trial comparing percutaneous radiofrequency thermal ablation, percutaneous ethanol injection, and percutaneous acetic acid injection to treat hepatocellular carcinoma of 3 cm or less. *Gut.* 2005; **54**: 1151–6.
30. Lencioni RA, Allgaier HP, Cioni D, *et al.* Small hepatocellular carcinoma in cirrhosis: randomized comparison of radio-frequency thermal ablation versus percutaneous ethanol injection. *Radiology.* 2003; **228**: 235–40.
31. Germani G, Pleguezuelo M, Gurusamy K, *et al.* Clinical outcomes of radiofrequency ablation, percutaneous alcohol and acetic acid injection for hepatocelullar carcinoma: a meta-analysis. *J Hepatol.* 2010; **52**: 380–8.
32. Giorgio A, Di Sarno A, De Stefano G, *et al.* Percutaneous radiofrequency ablation of hepatocellular carcinoma compared to percutaneous ethanol injection in treatment of cirrhotic patients: an Italian randomized controlled trial. *Anticancer Res.* 2011; **31**: 2291–5.
33. Park YK, Kim BW, Wang HJ, Kim MW. Hepatic resection for hepatocellular carcinoma meeting Milan criteria in Child-Turcotte-Pugh class a patients with cirrhosis. *Transplant Proc.* 2009; **41**: 1691–7.
34. Wang CC, Iyer SG, Low JK, *et al.* Perioperative factors affecting long-term outcomes of 473 consecutive patients undergoing hepatectomy for hepatocellular carcinoma. *Ann Surg Oncol.* 2009; **16**: 1832–42.
35. Kamiyama T, Nakanishi K, Yokoo H, *et al.* Recurrence patterns after hepatectomy of hepatocellular carcinoma: implication of Milan criteria utilization. *Ann Surg Oncol.* 2009; **16**: 1560–71.
36. Yamamoto J, Kosuge T, Saiura A, *et al.* Effectiveness of hepatic resection for early-stage hepatocellular carcinoma in cirrhotic patients: subgroup analysis according to Milan criteria. *Jpn J Clin Oncol.* 2007; **37**: 287–95.
37. Nuzzo G, Giuliante F, Gauzolino R, *et al.* Liver resections for hepatocellular carcinoma in chronic liver disease: experience in an Italian centre. *Eur J Surg Oncol.* 2007; **33**: 1014–8.
38. Hanazaki K, Kajikawa S, Shimozawa N, *et al.* Survival and recurrence after hepatic resection of 386 consecutive patients with hepatocellular carcinoma. *J Am Coll Surg.* 2000; **191**: 381–8.
39. Shimada K, Sano T, Sakamoto Y, Kosuge T. A long-term follow-up and management study of hepatocellular carcinoma patients surviving for 10 years or longer after curative hepatectomy. *Cancer.* 2005; **104**: 1939–47.
40. Huang GT, Lee PH, Tsang YM, *et al.* Percutaneous ethanol injection versus surgical resection for the treatment of small hepatocellular carcinoma: a prospective study. *Ann Surg.* 2005; **242**: 36–42.
41. Castells A, Bruix J, Bru C, *et al.* Treatment of small hepatocellular carcinoma in cirrhotic patients: a cohort study comparing surgical resection and percutaneous ethanol injection. *Hepatology.* 1993; **18**: 1121–6.
42. Livraghi T, Bolondi L, Buscarini L, *et al.* No treatment, resection and ethanol injection in hepatocellular carcinoma: a retrospective analysis of survival in 391 patients with cirrhosis. Italian Cooperative HCC Study Group. *J Hepatol.* 1995; **22**: 522–6.
43. Ryu M, Shimamura Y, Kinoshita T, *et al.* Therapeutic results of resection, transcatheter arterial embolization and percutaneous transhepatic ethanol injection in 3225 patients with hepatocellular carcinoma: a retrospective multicenter study. *Jpn J Clin Oncol.* 1997; **27**: 251–7.
44. Arii S, Yamaoka Y, Futagawa S, *et al.* Results of surgical and nonsurgical treatment for small-sized hepatocellular carcinomas: a retrospective and nationwide survey in Japan. The liver cancer study group of Japan. *Hepatology.* 2000; **32**: 1224–9.

Liver International (2012)

OPEN ACCESS Freely available online

PLOS ONE

Evolutionary Analysis of Classical *HLA* Class I and II Genes Suggests That Recent Positive Selection Acted on *DPB1*04:01* in Japanese Population

Minae Kawashima[1]*, Jun Ohashi[2]*, Nao Nishida[1,3], Katsushi Tokunaga[1]

1 Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, **2** Molecular and Genetic Epidemiology, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan, **3** The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Chiba, Japan

Abstract

The human leukocyte antigen (*HLA*) genes exhibit the highest degree of polymorphism in the human genome. This high degree of variation at classical *HLA* class I and class II loci has been maintained by balancing selection for a long evolutionary time. However, little is known about recent positive selection acting on specific *HLA* alleles in a local population. To detect the signature of recent positive selection, we genotyped six *HLA* loci, *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQB1*, and *HLA-DPB1* in 418 Japanese subjects, and then assessed the haplotype homozygosity (*HH*) of each *HLA* allele. There were 120 *HLA* alleles across the six loci. Among the 80 *HLA* alleles with frequencies of more than 1%, *DPB1*04:01*, which had a frequency of 6.1%, showed exceptionally high *HH* (0.53). This finding raises the possibility that recent positive selection has acted on *DPB1*04:01*. The *DPB1*04:01* allele, which was present in the most common 6-locus *HLA* haplotype (4.4%), *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01*, seems to have flowed from the Korean peninsula to the Japanese archipelago in the Yayoi period. A stochastic simulation approach indicated that the strong linkage disequilibrium between *DQB1*06:04* and *DPB1*04:01* observed in Japanese cannot be explained without positive selection favoring *DPB1*04:01*. The selection coefficient of *DPB1*04:01* was estimated as 0.041 (95% credible interval 0.021–0.077). Our results suggest that *DPB1*04:01* has recently undergone strong positive selection in Japanese population.

Citation: Kawashima M, Ohashi J, Nishida N, Tokunaga K (2012) Evolutionary Analysis of Classical *HLA* Class I and II Genes Suggests That Recent Positive Selection Acted on *DPB1*04:01* in Japanese Population. PLoS ONE 7(10): e46806. doi:10.1371/journal.pone.0046806

Editor: Henry Harpending, University of Utah, United States of America

Received May 2, 2012; **Accepted** September 10, 2012; **Published** October 3, 2012

Funding: This work was supported by National Bioscience Database Center (NBDC) of Japan Science and Technology Agency (JST), by KAKENHI (22133008) Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by KAKENHI (23133502) Grant-in-Aid for Scientific Research on Innovative Areas. 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: mina-e@umin.ac.jp (MK); juno-tky@umin.ac.jp(JO)

Introduction

The crucial immunological function of human leukocyte antigen (HLA) molecules is to present pathogen-derived antigenic peptides to T lymphocytes [1]. The HLA proteins are encoded by genes in the major histocompatibility complex region, which spans approximately 4 megabases (Mb) on the short arm of chromosome 6 (6p21.3) and includes the most polymorphic loci in the human genome [2]. A remarkable feature of the classical *HLA* class I and class II genes is the high degree of polymorphism. More than 1,750 *HLA-A*, 2,330 *HLA-B*, 1,300 *HLA-C*, 1,060 *HLA-DRB1*, 160 *HLA-DQB1*, and 150 *HLA-DPB1* alleles have been reported (IMGT/HLA database; http://www.ebi.ac.uk/imgt/hla/).

Positive selection has been shown as a driving force for the high degree of polymorphism at *HLA* loci [3,4]. The *HLA* genes show three remarkable signatures of positive selection: (1) the rate of nonsynonymous (amino acid altering) nucleotide substitution is substantially higher than that of synonymous substitution at antigen-recognition sites [5,6], (2) there are trans-species polymorphisms (i.e., similar alleles are present in multiple species) [7], and (3) there is a significant excess of heterozygosity [8,9]. Balancing selection, including overdominant selection and frequency-dependent selection, can easily account for these observations [3,4].

A number of studies have reported common long-range *HLA* haplotypes [10–16]. The extended length of common haplotype is a key feature of recent positive selection [17,18]. The *HLA* alleles on long-range haplotypes may have been subject to recent positive selection. In this study, to identify the signature of recent positive selection that has acted on specific *HLA* alleles in a local (i.e., geographically restricted) population, we investigated the allele frequencies and haplotype frequencies at *HLA-A*, *HLA-C*, *HLA-B*, *HLA-DRB1*, *HLA-DQB1*, and *HLA-DPB1* in 418 Japanese individuals. Our theoretical and computer simulation analyses suggested that *DPB1*04:01* has recently undergone strong positive selection in Japanese population.

Results

HLA Class I and Class II Alleles in Japanese

The genotypes of six *HLA* genes (three class I and three class II genes) were determined for each of 418 Japanese individuals. The frequencies of the 67 alleles found at the three *HLA* class I genes are listed in Table 1. Of the 17 *HLA-A* alleles, two–*A*02:01* and *A*24:02*–had frequencies higher than 10% (10.2 and 37.7 percent,

respectively). Of the 17 *HLA-C* alleles, four–*C*01:02*, *C*03:03*, *C*03:04*, and *C*07:02*–had frequencies higher than 10%: 16.5, 13.5, 12.6, and 14.5 percent, respectively. There were 33 *HLA-B* alleles, and not one had an allele frequency greater than 10%. The allele with the highest frequency (9.6%) was *B*52:01*; this allele was followed by *B*15:01* (8.5%), *B*51:01* (8.5%), *B*4403* (8.1%), and *B*35:01* (8.0%).

The frequencies of 53 alleles at three *HLA* class II genes are listed in Table 2. Of the 27 alleles at the *HLA-DRB1* locus, two–*DRB1*09:01* and *DRB1*04:05*–had frequencies of more than 10% (15.2% and 14.6%, respectively), and five–*DRB1*15:02* (8.4%), *DRB1*15:01* (8.0%), *DRB1*13:02* (7.8%), *DRB1*08:03* (7.5%), and *DRB1*01:01* (6.8%)–were also common. Of the 14 alleles at *HLA-DQB1*, four–*DQB1*03:03*, *DQB1*06:01*, *DQB1*04:01*, and *DQB1* 03:01*–were observed at frequencies of greater than 10% (15.9%, 15.9%, 14.6%, and 11.8%, respectively). There were four other common alleles at *HLA-DQB1–DQB1*03:02* (9.2%), *DQB1*06:02* (7.8%), *DQB1*05:01*, and *DQB1*06:04* (7.5%). Of the six *HLA* loci genotyped, *HLA-DPB1* had the fewest alleles with just 12. The *DPB1*05:01* (38.5%) and *DPB1*02:01* (25.1%) alleles were the most frequent alleles at this locus.

Of the six *HLA* loci examined, the *HLA-B* locus showed the highest heterozygosity (0.937), and *HLA-DPB1* showed the lowest (0.765) (Tables 1 and 2). None of the *HLA* class I or II loci

Table 1. Frequencies of *HLA* class I alleles.

HLA-A						HLA-C						HLA-B					
Allele	Count	Freq.	H[a]	HWE[b] P-val	EW[c] P-val	Allele	Count	Freq.	H[a]	HWE[b] P-val	EW[c] P-val	Allele	Count	Freq.	H[a]	HWE[b] P-val	EW[c] P-val
*A*01:01*	10	0.012	0.810	0.667	0.294	*C*01:02*	138	0.165	0.891	0.919	0.003	*B*07:02*	57	0.068	0.937	0.286	0.002
*A*02:01*	85	0.102				*C*01:03*	4	0.005				*B*13:01*	13	0.016			
*A*02:06*	61	0.073				*C*03:02*	3	0.004				*B*15:01*	71	0.085			
*A*02:07*	23	0.028				*C*03:03*	113	0.135				*B*15:07*	5	0.006			
*A*02:10*	2	0.002				*C*03:04*	105	0.126				*B*15:11*	5	0.006			
*A*03:01*	4	0.005				*C*04:01*	42	0.050				*B*15:18*	13	0.016			
*A*03:02*	1	0.001				*C*05:01*	5	0.006				*B*15:27*	1	0.001			
*A*11:01*	80	0.096				*C*06:02*	7	0.008				*B*15:28*	1	0.001			
*A*24:02*	315	0.377				*C*07:02*	121	0.145				*B*27:04*	2	0.002			
*A*24:08*	1	0.001				*C*07:04*	7	0.008				*B*35:01*	67	0.080			
*A*24:20*	10	0.012				*C*08:01*	47	0.056				*B*37:01*	7	0.008			
*A*26:01*	67	0.080				*C*08:03*	12	0.014				*B*39:01*	34	0.041			
*A*26:02*	12	0.014				*C*12:02*	81	0.097				*B*39:04*	5	0.006			
*A*26:03*	22	0.026				*C*12:03*	1	0.001				*B*40:01*	46	0.055			
*A*26:05*	1	0.001				*C*14:02*	50	0.060				*B*40:02*	57	0.068			
*A*31:01*	66	0.079				*C*14:03*	69	0.083				*B*40:03*	7	0.008			
*A*33:03*	76	0.091				*C*15:02*	31	0.037				*B*40:06*	34	0.041			
												*B*40:52*	1	0.001			
												*B*44:02*	5	0.006			
												*B*44:03*	68	0.081			
												*B*46:01*	38	0.045			
												*B*48:01*	22	0.026			
												*B*51:01*	71	0.085			
												*B*51:02*	4	0.005			
												*B*52:01*	80	0.096			
												*B*54:01*	64	0.077			
												*B*55:02*	20	0.024			
												*B*55:04*	1	0.001			
												*B*56:01*	5	0.006			
												*B*56:03*	2	0.002			
												*B*58:01*	3	0.004			
												*B*59:01*	16	0.019			
												*B*67:01*	11	0.013			

[a]Heterozygosity.
[b]Hardy-Weinberg equilibrium test.
[c]Ewens-Watterson test.
doi:10.1371/journal.pone.0046806.t001

Table 2. Frequencies of *HLA* class II alleles.

HLA-DRB1						*HLA-DQB1*						*HLA-DPB1*					
Allele	**Count**	**Freq.**	**H[a]**	**HWE[b]**	**EW[c]**	**Allele**	**Count**	**Freq.**	**H[a]**	**HWE[b]**	**EW[c]**	**Allele**	**Count**	**Freq.**	**H[a]**	**HWE[b]**	**EW[c]**
				P-val	P-val					P-val	P-val					P-val	P-val
*DRB1*01:01*	57	0.068	0.918	0.247	0.013	*DQB1*02:01*	1	0.001	0.885	0.222	0.001	*DPB1*02:01*	210	0.251	0.765	0.398	0.225
*DRB1*03:01*	1	0.001				*DQB1*03:01*	99	0.118				*DPB1*02:02*	35	0.042			
*DRB1*04:01*	10	0.012				*DQB1*03:02*	77	0.092				*DPB1*03:01*	36	0.043			
*DRB1*04:03*	24	0.029				*DQB1*03:03*	133	0.159				*DPB1*04:01*	51	0.061			
*DRB1*04:04*	2	0.002				*DQB1*04:01*	122	0.146				*DPB1*04:02*	83	0.099			
*DRB1*04:05*	122	0.146				*DQB1*04:02*	26	0.031				*DPB1*05:01*	322	0.385			
*DRB1*04:06*	28	0.033				*DQB1*05:01*	63	0.075				*DPB1*06:01*	5	0.006			
*DRB1*04:07*	1	0.001				*DQB1*05:02*	17	0.020				*DPB1*09:01*	65	0.078			
*DRB1*04:10*	12	0.014				*DQB1*05:03*	30	0.036				*DPB1*13:01*	12	0.014			
*DRB1*08:02*	32	0.038				*DQB1*06:01*	133	0.159				*DPB1*14:01*	10	0.012			
*DRB1*08:03*	63	0.075				*DQB1*06:02*	65	0.078				*DPB1*19:01*	5	0.006			
*DRB1*09:01*	127	0.152				*DQB1*06:03*	5	0.006				*DPB1*41:01*	2	0.002			
*DRB1*10:01*	6	0.007				*DQB1*06:04*	63	0.075									
*DRB1*11:01*	23	0.028				*DQB1*06:09*	2	0.002									
*DRB1*12:01*	30	0.036															
*DRB1*12:02*	18	0.022															
*DRB1*13:01*	5	0.006															
*DRB1*13:02*	65	0.078															
*DRB1*14:02*	1	0.001															
*DRB1*14:03*	11	0.013															
*DRB1*14:05*	17	0.020															
*DRB1*14:06*	13	0.016															
*DRB1*14:07*	3	0.004															
*DRB1*14:54*	26	0.031															
*DRB1*15:01*	67	0.080															
*DRB1*15:02*	70	0.084															
*DRB1*16:02*	2	0.002															

[a]Heterozygosity.
[b]Hardy-Weinberg equilibrium test.
[c]Ewens-Watterson test.
doi:10.1371/journal.pone.0046806.t002

exhibited significant deviation from HWE. Results of a Ewens-Watterson neutrality test [19,20] of *HLA* allele frequencies in this study population revealed that the observed distributions of allele frequencies at *HLA-C* ($P = 0.003$), *HLA-B* ($P = 0.002$), *HLA-DRB1* ($P = 0.013$), and *HLA-DQB1* ($P = 0.001$) differed significantly (i.e., there was excess heterozygosity) from the distributions expected based on the assumption of neutrality, whereas there was no significant difference between the expected and observed distributions of allele frequencies at *HLA-A* or *HLA-DPB1* (Tables 1 and 2).

Pairwise LD between *HLA* Alleles

The pairwise linkage disequilibrium (LD) parameters, r^2 and $|D'|$ [21], for each possible pair of two *HLA* alleles were estimated (Figure 1 and Data S1). Most alleles at *HLA-A* were not in strong LD with any of the alleles at the other loci because the physical distance from *HLA-A* to each of the other loci is large. To evaluate the relative strength of LD between two *HLA* loci, 2-locus r^2 and 2-locus $|D'|$ (see Materials and Methods for details), were calculated based on the pairwise LD parameters for all the allelic pairs (Table S1). The values of 2-locus $|D'|$ for *HLA-C* and *HLA-B* ($|D'| = 0.91$) and for *HLA-DRB1* and *HLA-DQB1* ($|D'| = 0.80$) were high, whereas the lowest 2-locus $|D'|$ value was observed for *HLA-A* and *HLA-DPB1* ($|D'| = 0.25$). These values reflected the physical distances between the respective loci. The values of 2-locus $|D'|$ for *HLA-DRB1* and *HLA-DPB1* and for *HLA-DQB1* and *HLA-DPB1* were relatively low compared to the values for the other pairs (Figure 2). These low values probably result from the recombination hotspot in the *HLA* class II region [22–24].

Major 6-locus *HLA* Haplotypes in Japanese

Frequencies of multi-locus haplotypes were estimated using the PHASE program [25,26] (Table 3 and Tables S2, S3, S4, S5). In 418 Japanese subjects (i.e., 839 chromosomes), 489 different 6-locus *HLA* haplotypes were inferred. Based on the frequencies of 6-locus *HLA* haplotypes, the probability of selecting two identical 6-locus *HLA* haplotypes at random from the Japanese population was estimated as 0.0075. Six 6-locus *HLA* haplotypes had

frequencies higher than 1% (Table 3). Of these, *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* was the most common (4.4%).

The intensity of recombination in the *HLA* region has been estimated at 0.67 cM/Mb [27], which corresponds to a recombination fraction of approximately 2% between *HLA-A* and *HLA-DPB1*. Thus, association between the six *HLA* alleles in any 6-locus *HLA* haplotype is not generally strong due to the frequent recombination in the *HLA* region. The expected frequency of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype is 2.5×10^{-7} under the assumption of linkage equilibrium, which is much smaller than the observed frequency of 0.044. The strong LD among *HLA* alleles on the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype may result from recent positive selection acting on one of *HLA* alleles on the haplotype, although other mechanisms such as neutral random genetic drift, recent admixture, recent migration, recent bottlenecks, and suppression of recombination can also cause the strong LD [10,12,13,15,16].

Haplotype Omozygosity

Strong positive selection leads to a rapid increase in the frequency of a selected (target) allele in a population. The number of recombination events between the target allele and the surrounding polymorphic sites is limited while the advantageous allele increases in frequency; therefore, the diversity of haplotypes carrying the advantageous allele becomes low. Accordingly, strong LD is expected in the genomic region bearing the selected allele. In this study, the degree of LD for each HLA allele was measured by haplotype homozygosity (*HH*); this term is defined as the probability that any two randomly chosen samples of haplotype bearing a focal *HLA* allele have the same 6-locus *HLA* haplotype. Like *EHH* [17], a high *HH* value can be regarded as a signature of recent positive selection acting on a focal *HLA* allele.

To detect *HLA* alleles that have been subject to recent positive selection, *HH* was calculated for each allele based on the estimated number of 6-locus haplotypes in 418 Japanese subjects. Of the 80

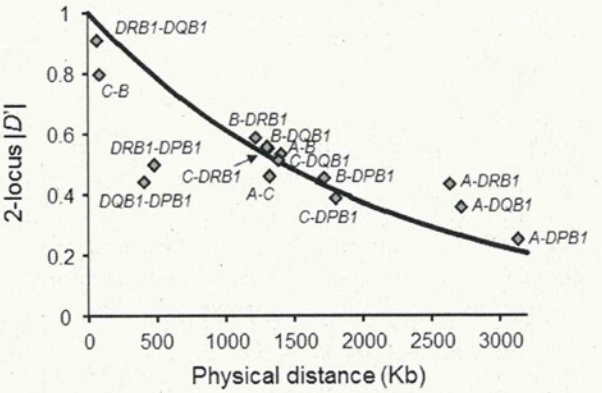


Figure 2. Relationship between two-locus $|D'|$ and physical distance (Kb). A solid-line curve, $2-locus|D'|=(1-0.67\times10^{-5}\times x)$[75,13], was obtained using the least-squares method, where x represents the physical distance (Kb). The recombination rate in the *HLA* region was assumed to be 0.67 cM/Mb [27]. Spearman's rank correlation coefficient between 2-locus $|D'|$ and the physical distance was −0.8607 ($P<0.0001$).
doi:10.1371/journal.pone.0046806.g002

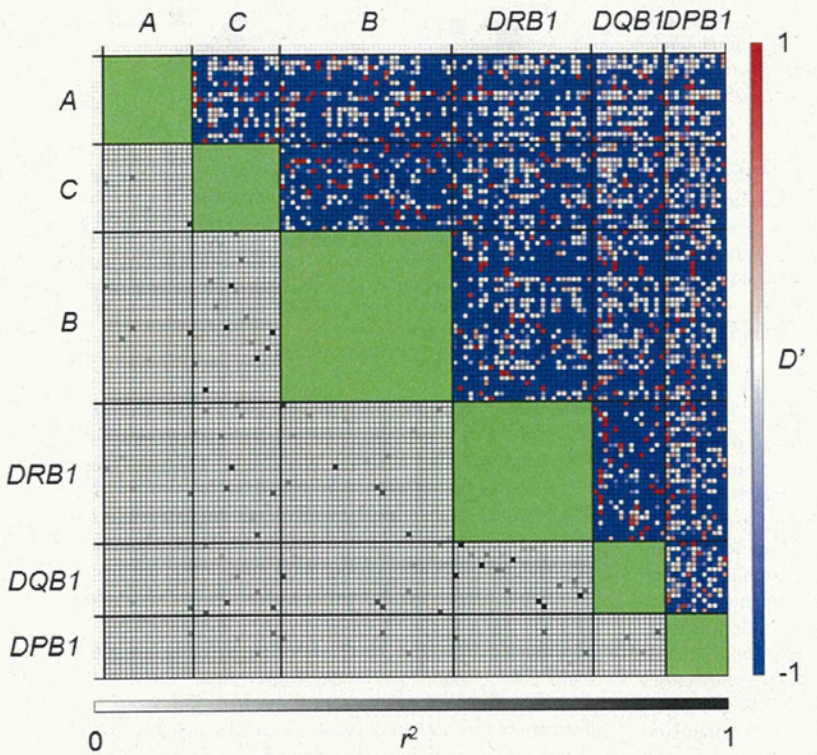


Figure 1. Pairwise estimates of LD parameters, $|D'|$ (upper diagonal) and r^2 (lower diagonal) for every pair of *HLA* alleles. The name of each allele is presented in Data S1.
doi:10.1371/journal.pone.0046806.g001

Table 3. Estimated frequencies of 6-locus *HLA* haplotypes.

Association						# of haplotypes[a]	HF[b]
*A*33:03*	*C*14:03*	*B*44:03*	*DRB1*13:02*	*DQB1*06:04*	*DPB1*04:01*	37	0.044
*A*24:02*	*C*12:02*	*B*52:01*	*DRB1*15:02*	*DQB1*06:01*	*DPB1*09:01*	33	0.039
*A*24:02*	*C*07:02*	*B*07:02*	*DRB1*01:01*	*DQB1*05:01*	*DPB1*04:02*	29	0.035
*A*24:02*	*C*01:02*	*B*54:01*	*DRB1*04:05*	*DQB1*04:01*	*DPB1*05:01*	13	0.016
*A*24:02*	*C*12:02*	*B*52:01*	*DRB1*15:02*	*DQB1*06:01*	*DPB1*02:01*	12	0.014
*A*11:01*	*C*04:01*	*B*15:01*	*DRB1*04:06*	*DQB1*03:02*	*DPB1*02:01*	11	0.013

[a]Estimated by the PHASE program version 2.1.
[b]Haplotype frequency.
doi:10.1371/journal.pone.0046806.t003

HLA alleles that had frequencies of more than 1%, one allele at each class I locus (*A*33:03*, *C*14:03*, and *B*44:03*) had the highest *HH* for that locus; similarly, one allele at each class II locus (*DRB1*13:02*, *DQB1*06:04*, and *DPB1*04:01*) had the highest *HH* for that locus (Figure 3). These six *HLA* alleles made up the 6-locus haplotype, *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01*, with the highest frequency in this Japanese population (Table 3).

The *HH* values are generally reduced by loci with high heterozygosity. Therefore, it was relatively difficult for an allele at *HLA-DPB1* to show high *HH*, because heterozygosities at the other loci are high. Nevertheless, the *DPB1*04:01* allele, which had a population frequency of 6.1%, showed the highest *HH* value (0.53) of the 80 *HLA* alleles with frequencies higher than 1% (Figure 3). The values of *HH* of the remaining 79 *HLA* alleles were less than 0.33. This finding suggests that *DPB1*04:01* had undergone recent positive selection in Japan. The large *HH* values of the five other alleles (*A*33:03*, *C*14:03*, *B*44:03*, *DRB1*13:02*, and *DQB1*06:04*) in this 6-locus *HLA* haplotype (i.e., *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01*) appear to be due to the hitchhiking effect of *DPB1*04:01*.

To investigate the effect of recombination on the decay of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype, the value of extended haplotype homozygosity (*EHH*) was calculated for *DPB1*04:01* (Figure 4). Although the *EHH* of *DPB1*04:01* was reduced at *HLA-DQB1*, the decrease in *EHH* was almost negligible at *HLA-DRB1*, *HLA-B*, and *HLA-C* loci; these findings indicate that, in this haplotype, recombination mainly has occurred between *DQB1*06:04* and *DPB1*04:01*.

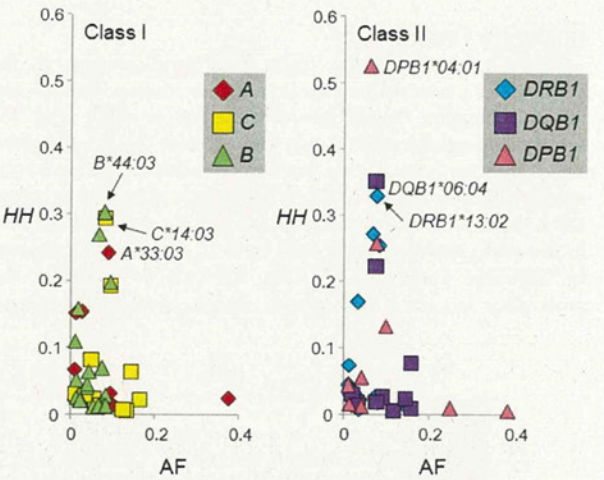


Figure 3. Haplotype homozygosity (*HH*) × allele frequency (AF) of each *HLA* allele. The left and right panels show *HH* values of *HLA* class I alleles and *HLA* class II alleles, respectively. The class I alleles were designated as follows: *HLA-A* (red diamond), *HLA-C* (yellow square), and *HLA-B* (green triangle); the class II alleles were designated as follows: *HLA-DRB1* (blue diamond), *HLA-DQB1* (purple square), and *HLA-DPB1* (pink triangle). In both panels, only *HH* values of alleles with frequencies of more than 0.01 are shown.
doi:10.1371/journal.pone.0046806.g003

Origin of *DPB1*04:01* in Japanese

*DPB1*04:01* is common (>30%) in European populations [9,28], whereas the frequency of *DPB1*04:01* is 6.1% in Japanese (Table 2). Given the worldwide distribution of *DPB1*04:01*, it is unlikely that *DPB1*04:01* originated in Japan. *DPB1*04:01* seems to have entered Japan. Archaeological studies of Japanese history have suggested that the Yayoi people came from the Korean peninsula circa 300 B.C., and mixed with the indigenous Jomon people. A recent large-scale survey of single nucleotide polymorphisms (SNPs) on autosomal chromosomes [29] revealed that most people presently inhabiting mainland Japan are genetically closer to Koreans than to Ryukuans. Ryukuans are considered to be more pure descendants of the Jomon people than are mainland Japanese. These observations indicate that a large population of Yayoi people migrated from the Korean peninsula. Although the frequency of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype in Koreans has not been reported, *DPB1*04:01*, which was carried by *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01*, appears to have derived from the Korean population because the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04* and *DRB1*13:02-DQA1*01:02-DQB1*06:04-DPB1*04:01* haplotypes are observed at the frequencies of 4.2% and 4.7% in Korean populations [28,30,31]. These and similar haplotypes have not been reported in other Asian populations (http://www.allelefrequencies.net) [28].

If the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype has a single origin, the current genetic diversity of this haplotype must be low. To assess the genetic diversity of this haplotype, we performed a sliding window analysis of individual heterozygosity, defined as a proportion of heterozygous SNPs to all SNPs in the window (Figure 5). Reduced individual heterozygosity was only found in the *HLA* region on the short arm of chromosome 6 in all the three subjects that were

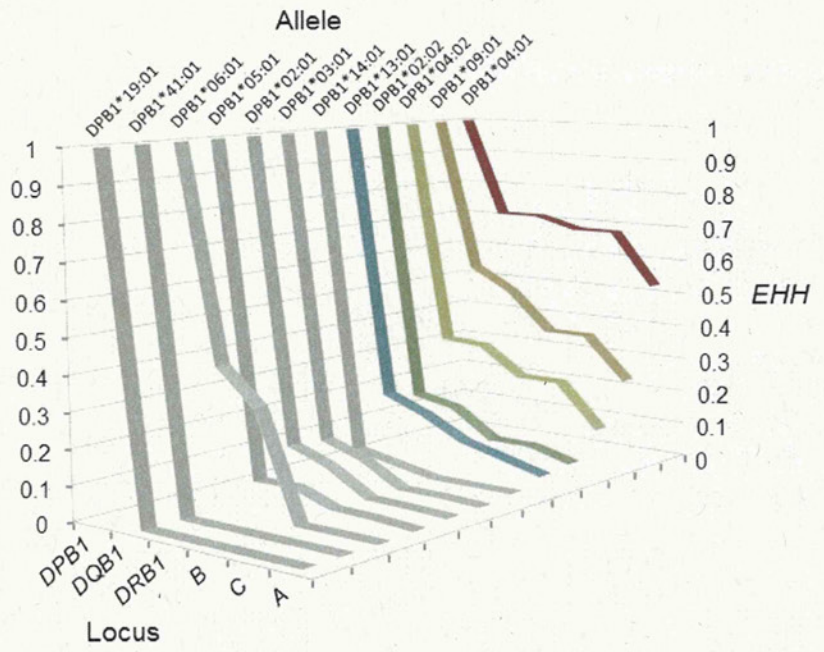


Figure 4. Extended *HH* (*EHH*) × relative locus position for12 *HLA-DPB1* alleles.
doi:10.1371/journal.pone.0046806.g004

homozygous for the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype (Figure 5A); in contrast, such a reduction was not observed in two subjects that were heterozygous for this haplotype (Figure 5B). Furthermore, three subjects that were homozygous for the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype shared the same SNP haplotype that spanned more than 4 Mb in the *HLA* region (Figure 5A). These observations suggest that the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype in Japanese has a single origin, and has not been generated repeatedly by recombination.

Computer Simulation

The analysis of *EHH* revealed that the reduction in *EHH* for *DPB1*04:01* resulted from recombination between *DQB1*06:04* and *DPB1*04:01* that inhabited the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype (Figure 4). Therefore, the relationship between *DQB1*06:04* and *DPB1*04:01* was focused in the following analyses. The high *HH* and *EHH* values of *DPB1*04:01* (Figures 3 and 4) may merely reflect that a neutral random genetic drift, rather than a recent positive selection, occurred after the Yayoi people reached the Japanese archipelago (300 B.C. or 2300 years ago). To assess this possibility, we conducted a computer simulation assuming a two-locus two-allele model in which changes in the frequency of four haplotypes carrying *DPB1*04:01* or non-*DPB1*04:01* alleles at the *HLA-DPB1* locus and *DQB1*06:04* or non-*DQB1:06:04* alleles at the *HLA-DQB1* locus were evaluated. In the simulation, the values of three parameters: selection intensity, s, recombination rate, c, and frequency of *DQB1*06:04-DPB1*04:01* haplotype, $f_1(0)$, in the beginning of the Yayoi period were drawn by a random number generator in every run. Haplotype frequencies were subject to change based on a stochastic model of positive selection, recombination, and random genetic drift. Dominant selection was assumed for *DPB1*04:01*, and, for the sake of simplicity, no selection (i.e., selectively neutral) was assumed for all alleles at the *DQB1* locus. The rejection method [18,32,33] was applied to accept only simulation runs that gave results similar to the observed values (see Materials and Methods for details). The uniform distribution was used for each parameter as a prior distribution (see Materials and Methods for detail). Figure 6A shows 2,500 parameter sets (i.e., posterior distributions) that were accepted in these simulations. The posterior distribution of the initial frequency of *DQB1*06:04-DPB1*04:01* haplotype was similar to the prior one, whereas the posterior distributions of selection intensity and recombination rate were different from the prior ones. In the posterior distribution, s ranged from 0.009 to 0.098, and the mean and 95% credible interval of s were 0.041 and 0.021−0.077, respectively (Figure 6B). It should be noted that neutral random genetic drift (i.e., $s \approx 0$) did not yield the results similar to the observed values. The findings from the simulations indicated that *DPB1*04:01* has been subject to relatively strong positive selection in Japanese since the Yayoi period.

Discussion

A number of *HLA* alleles have been shown to be associated with variations in immune responses to infectious diseases (e.g., human immunodeficiency virus [HIV]/AIDS, malaria, tuberculosis, hepatitis, leprosy, leishmaniasis, and schistosomiasis) caused by pathogenic microorganisms (see review by Blackwell et al. [34]). The most plausible explanation for positive selection favoring *DPB1*04:01* would be its function in resistance to infections. A recent genome-wide association study showed that the *DPA1*01:03-DPB1*04:01* haplotype confers protection against hepatitis B virus (HBV) infection (OR = 0.57, 95% CI = 0.33–0.96) [35]. Hepatitis B is a deadly infectious disease. Acute hepatitis B, which can cause fatal complications such as fulminant hepatitis, occurs in a percentage of the people infected with HBV. Although the estimated selection coefficient of s (0.0254–0.0550) for

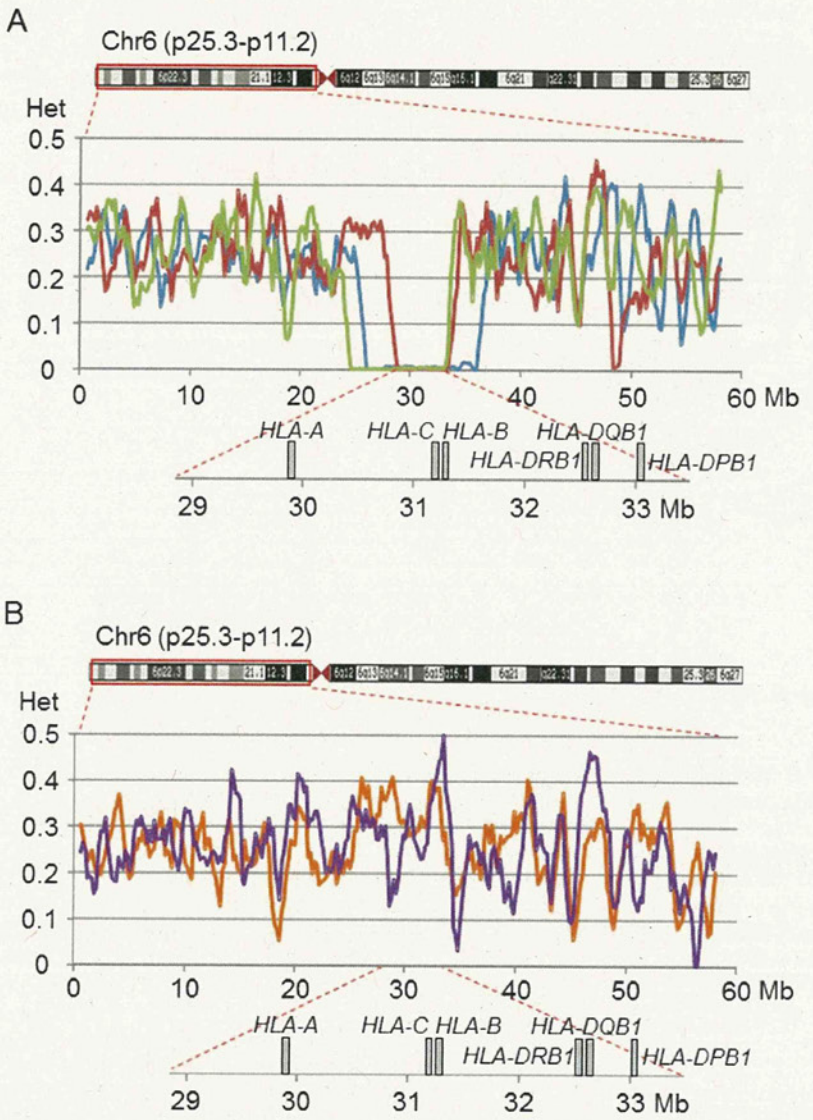
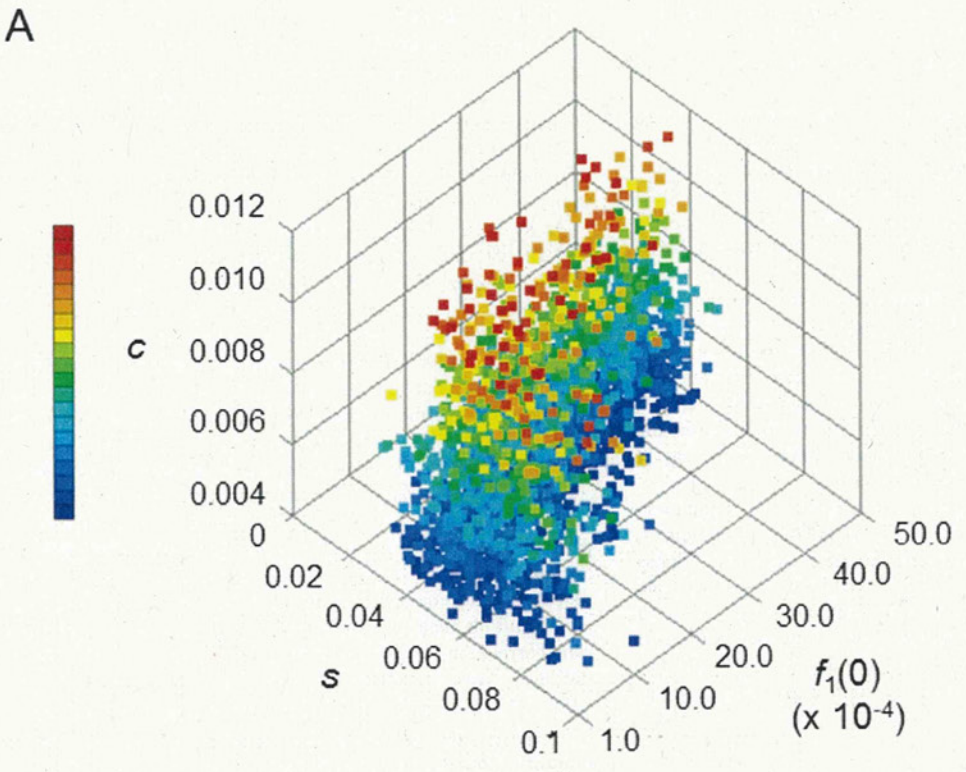


Figure 5. Individual heterozygosity of each subject with the most common 6-locus *HLA* haplotype. The individual heterozygosity in the genomic region on the short arm of chromosome 6 was assessed using the sliding window analysis; in this analysis, the window and step sizes were set to be 1 Mb and 200 kb, respectively. The individual heterozygosity was defined as a proportion of heterozygous SNPs to SNPs genotyped in a single subject. This analysis was performed for five Japanese subjects with the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype: (A) three of these five subjects were homozygous for this haplotype (blue, red, and green) and (B) two subjects had the heterozygous genotypes of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*07:02-B*07:02-DRB1*01:01–DQB1*05:01-DPB1*04:02* haplotype (orange) and of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*12:02-B*52:01-DRB1*15:02-DQB1*06:01-DPB1*09:01* haplotype (purple).
doi:10.1371/journal.pone.0046806.g005

*DPB1*04:01* does not seem to result solely from protection against infection with HBV, HBV infection may have been one of the key driving forces for the rapid increase in frequency of *DPB1*04:01* in the Japanese population.

Here, the analysis of *HH* was used to detect a signature of recent positive selection. The advantage of using *HH* in the analysis of *HLA* genes is that alleles with similar frequencies not only at the same *HLA* locus, but also at different loci, can be compared. This feature of analyses based on *HH* allows us to compare *HLA* alleles even within the same long-range haplotype. Since the same polymorphic markers are used for all *HLA* alleles in the calculation of *HH*, the effect of recombination on the value of *HH* can be well controlled. However, the *HH* analysis has a disadvantage in that the empirical distribution of *HH* value has to be obtained from only those alleles that are in the targeted region. Therefore, unlike conventional long-range haplotype tests based on *EHH* values [17,36], the statistical test based on *HH* values cannot be performed using genome-wide data. Nevertheless, *HH*-based test is thought to be suitable for analysis of *HLA* genes because each locus has a number of alleles to be examined and strong LD exists between alleles even at distant loci. The use of *HH* in the analysis of various human populations would help us to detect other *HLA*



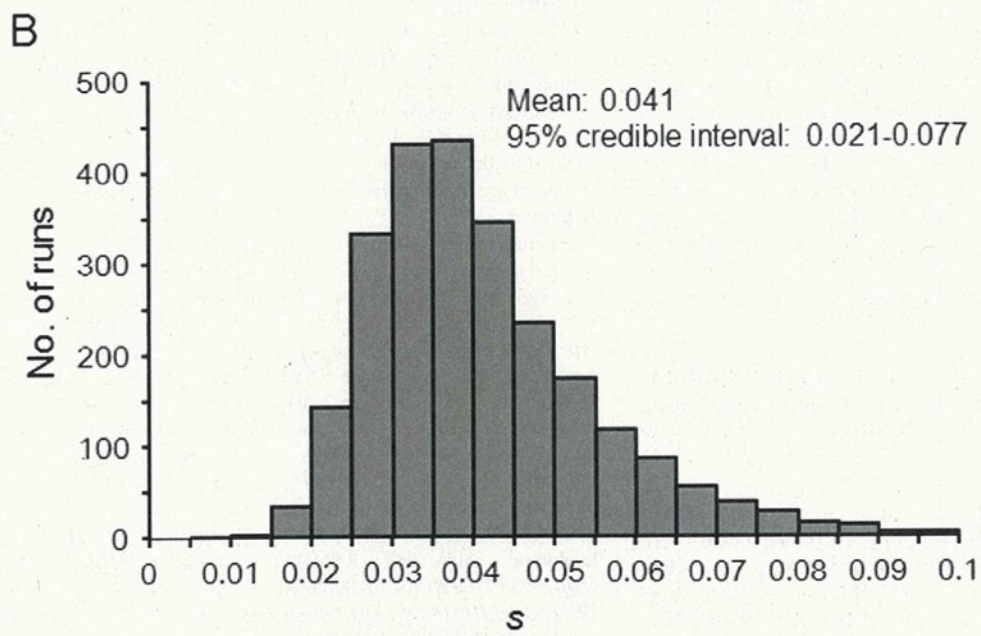


Figure 6. Estimation of model parameters for positive selection acting on *DPB1*04:01*. The recombination rate (*c*), initial haplotype frequency ($f_1(0)$), and selection coefficient (*s*), were estimated by comparing the four haplotype frequencies observed in our study population with the respective values predicted via simulation. (A) Posterior distributions of the three parameters that produced simulated data that resemble the observed data. (B) Frequency distribution of *s* accepted in simulation runs. The mean and 95% credible interval of *s* are 0.041 and 0.021−0.077.
doi:10.1371/journal.pone.0046806.g006

alleles that have been subject to geographically-restricted positive selection and to understand the role of *HLA* genes in the adaptation of human population to local environments over evolutionary time.

To estimate the selection coefficient of *DPB1*04:01*, we used a simple two-locus two-allele genetic model that was based on two assumptions, directional selection at *DPB1* and selective neutrality at *HLA-DQB1*. The problem associated with the use of this model was that the Ewens-Watterson test revealed that the allele frequency distribution at *HLA-DQB1* in this study population deviated significantly from that expected under neutrality (Table 2); therefore, the assumption of selective neutrality at *HLA-DQB1* may not be valid. If balancing selection is operating at *HLA-DQB1*, the allele frequency of *DQB1*06:04* is maintained at a certain frequency, and the change in the allele frequency of *DPB1*04:01* must be influenced by this selection at *HLA-DQB1*, although the effect of balancing selection at *HLA-DQB1* on the estimation of s is considered to be much smaller than that of directional selection favoring *DPB1*04:01*.

In this study, six *HLA* loci were investigated in 418 Japanese subjects. Of *HLA* alleles with high population frequencies, *DPB1*04:01*, which was present in the most common 6-locus *HLA* haplotype spanning more than 4 Mb, showed exceptionally high *HH*. A computer simulation estimated the selection coefficient of *DPB1*04:01* as 0.041. Taken together with high *HH* value of *DPB1*04:01*, we conclude that *DPB1*04:01* has recently undergone strong positive selection in Japanese population.

Materials and Methods

Subjects

All 418 individuals investigated in this study were unrelated Japanese adults living in Tokyo or neighboring areas. The genomic DNAs were extracted from peripheral blood samples using a commercial kit (QIAamp Blood Kit [Qiagen, Hilden, Germany]). All blood and DNA samples were de-identified. Verbal informed consent was obtained from all the participants before 1990. In this study, written informed consent was not obtained because the blood sampling was conducted before the "Ethical Guidelines for Human Genome and Genetic Sequencing Research" were established in Japan. Under the condition that DNA sample is permanently de-linked from the individual, this study was approved by the Research Ethics Committee of the Faculty of Medicine, University of Tokyo.

HLA Typing

DNA typing of *HLA* alleles was performed by HLA LABORATORY (Kyoto, Japan) using a Luminex Multi-Analyte profiling system (xMAP; Luminex, Austin, TX, USA) [37].

SNP Typing

Five Japanese subjects who had at least one *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype were genotyped using the Axiom™ Genome-Wide ASI 1 Array Plate (Affymetrix Inc., Santa Clara, CA, USA). Of five subjects, three subjects were homozygous for the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and two subjects had the heterozygous genotypes of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*07:02-B*07:02-DRB1*01:01-DQB1*05:01-DPB1*04:02* haplotype and of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*12:02-B*52:01-DRB1*15:02-DQB1*06:01-DPB1*09:01* haplotype.

Statistical Analysis

Deviation from HWE for each *HLA* locus was tested using an exact test available in a web-based software, Genepop 4.0.10 [38]. Using Arlequin version 3.5 [39], the Ewens-Watterson test [40], which is based on Ewens sampling theory of neutral alleles [19], was performed to assess whether the observed distribution of allele frequencies at each *HLA* locus was different from an expectation that was based on neutrality.

To evaluate the degree of LD between *HLA* alleles, values of r^2 and D' [21] for all pairwise combinations of *HLA* alleles were calculated based on the haplotype frequencies estimated using the expectation maximization algorithm [20]. Here, each *HLA* allele was regarded as a single nucleotide polymorphism (SNP). For example, the *A*01:01* allele and the other alleles at the *HLA-A* locus were designated as "A" and "G", respectively. Accordingly, the algorithm for the estimation of haplotype frequencies for two loci, each with two alleles, could be applied to the *HLA* loci with multiple alleles for the purposes of these pairwise comparisons.

The LD parameter, 2-locus $|D'|$, between any two *HLA* loci (locus 1 and locus 2) was calculated based on the pairwise LD parameter, D'_{ij}, between ith allele at locus 1 and jth allele at locus 2 as follows: 2-locus $|D'| = \sum_{i=1}^{m} \sum_{j=1}^{n} p_i q_j |D'_{ij}|$, where p_i and q_j represent the frequencies of ith allele at locus 1 with m different alleles and jth allele at locus 2 with n different alleles. Spearman's rank correlation coefficient between 2-locus $|D'|$ and the physical distance was calculated. Assuming a model: 2-locus $|D'| = \left(1 - 0.67 \times 10^{-5} \times x\right)^a$, the curve fitting model parameter, a, was estimated using the least squares method; this method minimizes the sum-of-squared residual between an observed value and a fitted value that was determine by a model. In the above equation, the physical distance (Kb) between two loci is denoted by x and the recombination intensity in the *HLA* region was set at 0.65 cM/Mb [27,41].

The phased haplotypes consisting of two or more *HLA* loci were estimated using the PHASE program version 2.1 [25,26]. The estimated 6-locus haplotypes were further used for the calculation of extended haplotype homozygosity (*EHH*) [17] and of haplotype homozygosity (*HH*). In this study, *HH* of each *HLA* allele was defined as the probability that any two randomly chosen samples of haplotype bearing the *HLA* allele have the same 6-locus *HLA* haplotype.

A sliding window analysis of individual heterozygosity, which was defined as the proportion of heterozygous SNPs to SNPs genotyped in a single subject, was conducted to examine whether the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype had a single origin in Japan. 19,949 SNPs located on 6p were genotyped, and the average SNP density was 0.34 SNP/kb. The window and step sizes were 1 Mb and 200 kb, respectively. This analysis was performed using the SNP data from the five subject included in the SNP typing: three subjects were homozygous for the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and two subjects had the heterozygous genotypes of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*07:02-B*07:02-DRB1*01:01-DQB1*05:01-DPB1*04:02* haplotype and of the *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01* haplotype and the *A*24:02-C*12:02-B*52:01-DRB1*15:02-DQB1*06:01-DPB1*09:01* haplotype.

Computer Simulation

To estimate the intensity of recent positive selection acting on *DPB1*04:01*, a stochastic population genetic model (two-locus two-allele model) assuming both positive selection and random

genetic drift was built and assessed. The diploid population size, N, was set to be 10,000 (i.e., 20,000 chromosomes). Four haplotypes carrying *DPB1*04:01* or non-*DPB1*04:01* alleles (designated by *DPB1*X*) at the *HLA-DPB1* locus and *DQB1*06:04* or non-*DQB1:06:04* alleles (designated by *DQB1*X*) at the *HLA-DQB1* locus were used in this model. The frequencies of the *DQB1*06:04-DPB1*04:01*, *DQB1*X-DPB1*04:01*, *DQB1*06:04-DPB1*X*, and *DQB1*X-DPB1*X* haplotypes at generation t were denoted by $f_1(t)$, $f_2(t)$, $f_3(t)$, and $f_4(t)$, respectively. The current frequencies of the corresponding haplotypes in our study population were denoted by f_1, f_2, f_3, and f_4. A dominant selection was assumed for *DPB1*04:01* (i.e., relative finesses of *DPB1*04:01/DPB1*04:01*, *DPB1*04:01/DPB1*X*, and *DPB1*X/DPB1*X* are 1, 1, and 1 - s, respectively). The initial haplotype frequencies were set as $f_1(t) = z$, $f_2(t) = 0$, $f_3(t) = (1-z)f_3/(f_3+f_4)$, and $f_4(t) = (1-z)f_4/(f_3+f_4)$. The recombination between *HLA-DPB1* and *HLA-DQB1* loci was assumed to occur at a rate of c. Since the recombination rate between *HLA-DQB1* and *HLA-DPB1* has been estimated to be between 0.004 and 0.012 [41,42], a uniform recombination rate (c) within this range was used as a prior distribution. To estimate suitable parameter sets of z, s, and c, each value was drawn by a random number generator in every simulation run. The random numbers were between 0.0001 (i.e., $2/2N$) and 0.005 (i.e., $100/2N$) for z, between 0 and 0.1 for s, and between 0.004 and 0.012 for c.

Next, to evaluate the similarity between simulated and observed frequencies,

$$e = \sum_{i=1}^{4} \frac{(f_i(t) - f_i)^2}{f_i(t) + f_i}$$

was calculated. As the simulated haplotype frequencies, $f_1(t)$, $f_2(t)$, $f_3(t)$, and $f_4(t)$, approaches values close to the observed frequencies, f_1, f_2, f_3, and f_4, the value of e approaches 0. The rejection method [18,32,33] was used to accept only simulation runs that resulted in (i) e of less than 0.01, (ii) $f_1(t)$ of not less than $f_1 - 0.01$ nor more than $f_1 + 0.01$, and (iii) t of not less than 92 nor more than 115 generations. A total of 2,500 runs were accepted. The mean and 95% credible interval of s were obtained from the 2,500 accepted runs.

Supporting Information

Data S1 Pairwise LD measures for individual HLA allele pairs.
(XLSX)

Table S1 Linkage Disequilibrium between pairs of HLA loci.
(XLSX)

Table S2 Estimated frequencies of 2-locus HLA haplotypes.
(XLSX)

Table S3 Estimated frequencies of 3-locus HLA haplotypes.
(XLSX)

Table S4 Estimated frequencies of 4-locus HLA haplotypes.
(XLSX)

Table S5 Estimated frequencies of 5-locus HLA haplotypes.
(XLSX)

Acknowledgments

We deeply thank all the subjects for their participation in the study. We also thank Ms Yoriko Mawatari, Ms Megumi Sageshima, Ms Yuko Ogasawara, Ms Natsumi Baba, and Ms Rieko Hayashi (University of Tokyo) for technical assistance.

Author Contributions

Conceived and designed the experiments: MK JO. Performed the experiments: MK NN. Analyzed the data: MK JO. Contributed reagents/materials/analysis tools: JO NN KT. Wrote the paper: MK JO. Assembled the data: MK NN. Performed the computer simulation: JO.

References

1. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, et al. (1987) The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature 329: 512–518.
2. Campbell RD, Trowsdale J (1993) Map of the human MHC. Immunology today 14: 349–352.
3. Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124: 967–978.
4. Takahata N (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc Natl Acad Sci U S A 87: 2419–2423.
5. Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335: 167–170.
6. Hughes AL, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci U S A 86: 958–962.
7. Klein J (1987) Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. Hum Immunol 19: 155–162.
8. Hedrick PW, Thomson G (1983) Evidence for balancing selection at HLA. Genetics 104: 449–456.
9. Solberg OD, Mack SJ, Lancaster AK, Single RM, Tsai Y, et al. (2008) Balancing selection and heterogeneity across the classical human leukocyte antigen loci: a meta-analytic review of 497 population studies. Hum Immunol 69: 443–464.
10. Aly TA, Eller E, Ide A, Gowan K, Babu SR, et al. (2006) Multi-SNP analysis of MHC region: remarkable conservation of HLA-A1-B8-DR3 haplotype. Diabetes 55: 1265–1269.
11. Horton R, Gibson R, Coggill P, Miretti M, Allcock RJ, et al. (2008) Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. Immunogenetics 60: 1–18.
12. Alper CA, Larsen CE, Dubey DP, Awdeh ZL, Fici DA, et al. (2006) The haplotype structure of the human major histocompatibility complex. Hum Immunol 67: 73–84.
13. Yunis EJ, Larsen CE, Fernandez-Vina M, Awdeh ZL, Romero T, et al. (2003) Inheritable variable sizes of DNA stretches in the human MHC: conserved extended haplotypes and their fragments or blocks. Tissue Antigens 62: 1–20.
14. Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, et al. (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. PLoS Genet 2: e9.
15. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, et al. (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nature genetics 38: 1166–1172.
16. Awdeh ZL, Raum D, Yunis EJ, Alper CA (1983) Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. Proc Natl Acad Sci U S A 80: 259–263.
17. Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, et al. (2002) Detecting recent positive selection in the human genome from haplotype structure. Nature 419: 832–837.
18. Ohashi J, Naka I, Patarapotikul J, Hananantachai H, Brittenham G, et al. (2004) Extended linkage disequilibrium surrounding the hemoglobin E variant due to malarial selection. Am J Hum Genet 74: 1198–1208.
19. Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Popul Biol 3: 87–112.
20. Excoffier L, Slatkin M (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 12: 921–927.
21. Lewontin RC (1964) The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics 49: 49–67.
22. Mitsunaga S, Kuwata S, Tokunaga K, Uchikawa C, Takahashi K, et al. (1992) Family study on HLA-DPB1 polymorphism: linkage analysis with HLA-DR/DQ and two "new" alleles. Human immunology 34: 203–211.

23. Cullen M, Erlich H, Klitz W, Carrington M (1995) Molecular mapping of a recombination hotspot located in the second intron of the human TAP2 locus. American journal of human genetics 56: 1350–1358.
24. Djilali-Saiah I, Benini V, Daniel S, Assan R, Bach JF, et al. (1996) Linkage disequilibrium between HLA class II (DR, DQ, DP) and antigen processing (LMP, TAP, DM) genes of the major histocompatibility complex. Tissue antigens 48: 87–92.
25. Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76: 449–462.
26. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68: 978–989.
27. Miretti MM, Walsh EC, Ke X, Delgado M, Griffiths M, et al. (2005) 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.
28. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR (2011) Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. Nucleic Acids Res 39: D913–919.
29. Abdulla MA, Ahmed I, Assawamakin A, Bhak J, Brahmachari SK, et al. (2009) Mapping human genetic diversity in Asia. Science 326: 1541–1545.
30. Gjesing AP, Andersen G, Burgdorf KS, Borch-Johnsen K, Jorgensen T, et al. (2007) Studies of the associations between functional beta2-adrenergic receptor variants and obesity, hypertension and type 2 diabetes in 7,808 white subjects. Diabetologia 50: 563–568.
31. Song EY, Park MH, Kang SJ, Park HJ, Kim BC, et al. (2002) HLA class II allele and haplotype frequencies in Koreans based on 107 families. Tissue Antigens 59: 475–486.
32. Ohashi J, Naka I, Tsuchiya N (2011) The impact of natural selection on an ABCC11 SNP determining earwax type. Mol Biol Evol 28: 849–857.
33. Tishkoff SA, Varkonyi R, Cahinhinan N, Abbes S, Argyropoulos G, et al. (2001) Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. Science 293: 455–462.
34. Blackwell JM, Jamieson SE, Burgner D (2009) HLA and infectious diseases. Clin Microbiol Rev 22: 370–385, Table of Contents.
35. 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.
36. Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biol 4: e72.
37. Dunbar SA (2006) Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. Clin Chim Acta 363: 71–82.
38. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8: 103–106.
39. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1: 47–50.
40. Watterson GA (1978) The homozygosity test of neutrality. Genetics 88: 405–417.
41. Martin M, Mann D, Carrington M (1995) Recombination rates across the HLA complex: use of microsatellites as a rapid screen for recombinant chromosomes. Hum Mol Genet 4: 423–428.
42. Begovich AB, McClure GR, Suraj VC, Helmuth RC, Fildes N, et al. (1992) Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. J Immunol 148: 249–258.

Genome-wide Association Study Identifies *TNFSF15* and *POU2AF1* as Susceptibility Loci for Primary Biliary Cirrhosis in the Japanese Population

Minoru Nakamura,[1,2,3,30,*] Nao Nishida,[4,5,30] Minae Kawashima,[4] Yoshihiro Aiba,[1] Atsushi Tanaka,[6] Michio Yasunami,[7] Hitomi Nakamura,[1] Atsumasai Komori,[1] Makoto Nakamuta,[2] Mikio Zeniya,[8] Etsuko Hashimoto,[9] Hiromasa Ohira,[10] Kazuhide Yamamoto,[11] Morikazu Onji,[12] Shuichi Kaneko,[13] Masao Honda,[13] Satoshi Yamagiwa,[14] Kazuhiko Nakao,[15] Takafumi Ichida,[16] Hajime Takikawa,[6] Masataka Seike,[17] Takeji Umemura,[18] Yoshiyuki Ueno,[19] Shotaro Sakisaka,[20] Kentaro Kikuchi,[21] Hirotoshi Ebinuma,[22] Noriyo Yamashiki,[23] Sumito Tamura,[24] Yasuhiko Sugawara,[24] Akira Mori,[25] Shintaro Yagi,[25] Ken Shirabe,[26] Akinobu Taketomi,[26] Kuniaki Arai,[13] Kyoko Monoe,[10] Tatsuki Ichikawa,[15] Makiko Taniai,[9] Yasuhiro Miyake,[11] Teru Kumagi,[12] Masanori Abe,[12] Kaname Yoshizawa,[2,18] Satoru Joshita,[18] Shinji Shimoda,[27] Koichi Honda,[17] Hiroki Takahashi,[8] Katsuji Hirano,[16] Yasuaki Takeyama,[20] Kenichi Harada,[28] Kiyoshi Migita,[1] Masahiro Ito,[1] Hiroshi Yatsuhashi,[1] Nobuyoshi Fukushima,[2] Hajime Ota,[2] Tatsuji Komatsu,[2] Takeo Saoshiro,[2] Jinya Ishida,[2] Hirotsugu Kouno,[2] Hirotaka Kouno,[2] Michiyasu Yagura,[2] Masakazu Kobayashi,[2] Toyokichi Muro,[2] Naohiko Masaki,[2] Keiichi Hirata,[2] Yukio Watanabe,[2] Yoko Nakamura,[2] Masaaki Shimada,[2] Noboru Hirashima,[2] Toshiki Komeda,[2] Kazuhiro Sugi,[2] Michiaki Koga,[2] Keisuke Ario,[2] Eiichi Takesaki,[2] Yoshihiko Maehara,[26] Shinji Uemoto,[25] Norihiro Kokudo,[24] Hirohito Tsubouchi,[29] Masashi Mizokami,[5] Yasuni Nakanuma,[28] Katsushi Tokunaga,[4] and Hiromi Ishibashi[1]

For the identification of susceptibility loci for primary biliary cirrhosis (PBC), a genome-wide association study (GWAS) was performed in 963 Japanese individuals (487 PBC cases and 476 healthy controls) and in a subsequent replication study that included 1,402 other Japanese individuals (787 cases and 615 controls). In addition to the most significant susceptibility region, human leukocyte antigen (HLA), we identified two significant susceptibility loci, *TNFSF15* (rs4979462) and *POU2AF1* (rs4938534) (combined odds ratio [OR] = 1.56, $p = 2.84 \times 10^{-14}$ for rs4979462, and combined OR = 1.39, $p = 2.38 \times 10^{-8}$ for rs4938534). Among 21 non-HLA susceptibility loci for PBC identified in GWASs of individuals of European descent, three loci (*IL7R*, *IKZF3*, and *CD80*) showed significant associations (combined $p = 3.66 \times 10^{-8}$, 3.66×10^{-9}, and 3.04×10^{-9}, respectively) and *STAT4* and *NFKB1* loci showed suggestive association with PBC

[1]Clinical Research Center, National Hospital Organization (NHO) Nagasaki Medical Center, and Department of Hepatology, Nagasaki University Graduate School of Biomedical Sciences, Omura, Nagasaki 856-8562, Japan; [2]Headquaters of Primary Biliary Cirrhosis Research in the NHO Study Group for Liver Disease in Japan, Clinical Research Center, NHO Nagasaki Medical Center, Omura, Nagasaki 856-8562, Japan; [3]Headquaters of the gp210 Working Group in Intractable Liver Disease Research Project Team of the Ministry of Health and Welfare in Japan, Department of Hepatology, Nagasaki University Graduate School of Biomedical Sciences, Omura, Nagasaki 856-8562, Japan; [4]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan; [5]The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa 272-8516, Japan; [6]Department of Medicine, Teikyo University School of Medicine, Tokyo 173-8605, Japan; [7]Department of Clinical Medicine, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan; [8]Department of Gastroenterology and Hepatology, Tokyo Jikei University School of Medicine, Tokyo 105-8461, Japan; [9]Department of Medicine and Gastroenterology, Tokyo Women's Medical University, Tokyo 162-8666, Japan; [10]Department of Gastroenterology and Rheumatic Diseases, Fukushima Medical University, Fukushima 960-1295, Japan; [11]Department of Gastroenterology and Hepatology, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences, Okayama 700-8558, Japan; [12]Department of Gastroenterology and Metabology, Ehime University Graduate School of Medicine, Matsuyama 791-0295, Japan; [13]Department of Gastroenterology, Kanazawa University Graduate School of Medicine, Kanazawa 920-0942, Japan; [14]Division of Gastroenterology and Hepatology, Niigata University Graduate School of Medical and Dental Sciences, Niigata 951-8510, Japan; [15]Department of Gastroenterology and Hepatology, Nagasaki University Graduate School of Biomedical Sciences, Nagaski 852-8501, Japan; [16]Department of Gastroenterology and Hepatology, Juntendo University Shizuoka Hospital, Shizuoka 410-2295, Japan; [17]First Department of Internal Medicine, Faculty of Medicine, Oita University, Oita 879-5593, Japan; [18]Division of Gastroenterology and Hepatology, Department of Medicine, Shinshu University School of Medicine, Matsumoto 390-8621, Japan; [19]Division of Gastroenterology, Tohoku University Graduate School of Medicine, Sendai 980-8575, Japan; [20]Department of Gastroenterology and Medicine, Fukuoka University School of Medicine, Fukuoka 814-0180, Japan; [21]Deprtment of Internal Medicine, Teikyo University Mizonokuchi Hospital, Kawasaki 213-8507, Japan; [22]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio Graduate School of Medicine, Tokyo 160-8582, Japan; [23]Organ Transplantation Service, The University of Tokyo, Tokyo 113-8655, Japan; [24]Hepatobiliary and Pancreatic Surgery Division and Artificial Organ and Transplantation Division, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan; [25]Division of Hepato-pancreato-biliary Surgery and Transplantation, Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan; [26]Department of Surgery and Science, Kyushu University Graduate School of Medical Sciences, Fukuoka 812-8582, Japan; [27]Department of Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, Fukuoka 812-8582, Japan; [28]Department of Human Pathology, Kanazawa University Graduate School of Medicine, Kanazawa 920-0942, Japan; [29]Department of Digestive and Lifestyle-Related Disease, Kagoshima University Graduate School of Medical and Dental Science, Kagoshima 890-8520, Japan
[30]These authors contributed equally to this work
*Correspondence: nakamuram@nmc.hosp.go.jp
http://dx.doi.org/10.1016/j.ajhg.2012.08.010.

(combined $p = 1.11 \times 10^{-6}$ and 1.42×10^{-7}, respectively) in the Japanese population. These observations indicated the existence of ethnic differences in genetic susceptibility loci to PBC and the importance of TNF signaling and B cell differentiation for the development of PBC in individuals of European descent and Japanese individuals.

Primary biliary cirrhosis (PBC, MIM 109720) is a chronic and progressive cholestatic liver disease, presumably caused by autoimmune reactions against biliary epithelial cells, leading to liver cirrhosis and hepatic failure.[1] The incidence and prevalence of PBC range from 0.33 to 5.8 and from 2 to 40 per 100,000 inhabitants, respectively, in different geographical areas.[2] This may indicate the contribution of environmental or genetic factors in the development of PBC, whereas the clinical profiles of PBC are thought to be similar between different ethnicities and/or different geographical areas, including European-descent and eastern Asian populations. The high concordance rate in monozygotic twins compared to dizygotic twins[3] and familial clustering of individuals with PBC indicate the involvement of strong genetic factors in the development of PBC; however, the pathogenesis of PBC is still poorly understood. Previous genome-wide association studies (GWASs) and subsequent meta-analyses have identified *HLA* and 21 non-HLA susceptibility loci (*IL12A* [MIM 161560], *IL12RB2* [MIM 601642], *STAT4* [MIM 600558], *IRF5* [MIM 607218], *IKZF3* [MIM 606221], *MMEL1* [MIM 120520], *SPIB* [MIM 606802], *DENND1B* [MIM 613292], *CD80* [MIM 112203], *IL7R* [MIM 146661], *CXCR5* [MIM 601613], *TNFRSF1A* [MIM 191190], *CLEC16A* [MIM 611303], *NFKB* [MIM 164012], *RAD51L1* [MIM 602948], *MAP3K7IP1* [MIM 602615], *PLCL2* [MIM 614276], *RPS6KA4* [MIM 603606], *TNFAIP2* [MIM 603300], 7p14, and 16q24) to PBC in individuals of European descent,[4–7] indicating the important role of several autoimmune pathways (i.e., IL12A signaling, TNF/TLR-NF-κB signaling, and B cell differentiation) in the development of PBC. However, GWASs for PBC have never been reported for ethnicities other than European descent, limiting our knowledge of the genetic architecture of PBC. Here, we conducted a GWAS for PBC in the Japanese population to identify host genetic factors related to PBC, which would not only expand our knowledge of pathogenic pathways in PBC but also lead to the development of rationale for therapies in the future.

Samples from 2,395 individuals (1,295 cases with PBC and 1,100 healthy volunteers working at the National Hospital Organization (NHO) in Japan as a medical staff who declared having no apparent diseases, including chronic liver diseases and autoimmune diseases [healthy controls]) were collected by members of the Japan PBC-GWAS Consortium, which consists of 31 hospitals participating in the NHO Study Group for Liver Disease in Japan (NHOSLJ) and 24 university hospitals participating in the gp210 Working Group in Intractable Liver Disease Research Project Team of the Ministry of Health and Welfare in Japan. Most of the case and control samples were collected from the mainland and the neighboring islands of Japan (Honshu, Kyushu, and Shikoku). Previous studies have shown that there is little genetic heterogeneity in resident populations in these areas.[8] In fact, the genetic inflation factor was close to 1.00, and only a small portion of the samples were identified as outliers in the principal component analysis. The cases were diagnosed with PBC if they met at least two of the following internationally accepted criteria:[9] biochemical evidence of cholestasis based mainly on alkaline phosphatase elevation, presence of serum anti-mitochondrial antibodies, histological evidence of non-suppurative destructive cholangitis, and destruction of interlobular bile ducts. The demographic details of PBC cases are summarized in Table S1, available online. Of the 487 PBC cases in the GWAS, 57 were male and 430 were female, ages ranged from 33 to 90 years, the median age was 66 years, 320 cases had early-stage PBC (a stage without any signs indicating portal hypertension or liver cirrhosis), 110 had late-stage PBC without jaundice (a stage with signs of portal hypertension or liver cirrhosis but without persistent jaundice), and 57 were at the late stage with jaundice (persistent presence of jaundice [total bilirubin >2 mg/dl]). Of the 476 healthy controls in the GWAS, 170 were male and 306 were female, ages ranged from 25 to 87 years, and the median age was 40. Of the 808 PBC cases in the replication set, 120 were male and 688 were female, ages ranged from 24 to 85 years, the median age was 61 years, 646 had early-stage PBC, 121 had late-stage PBC without jaundice, and 39 were at the late stage with jaundice. Of the 624 healthy controls in the replication set, 271 were male and 353 were female, ages ranged from 24 to 74 years, and the median age was 33 years. Concomitant autoimmune diseases are also shown in Table S1. As for inflammatory bowel diseases such as Crohn disease (CD, MIM 266600) and ulcerative colitis (UC, MIM 266600), only one out of 1,274 PBC cases had UC, but none had CD. DNA was extracted from whole peripheral blood with the QIAamp DNA Blood Midi Kit (QIAGEN, Tokyo).

For the GWAS, we genotyped 1,015 samples (515 Japanese PBC cases and 500 Japanese healthy controls) using the Affymetrix Axiom Genome-Wide ASI 1 Array, according to the manufacturer's instructions. After excluding three PBC samples with a Dish QC of less than 0.82, we recalled the remaining 1,012 samples (512 cases and 500 controls) using the Genotyping Console v4.1 software. Here, Dish QC represents the recommended sample quality control (QC) metric for the Axiom arrays.[10] Of the 600,000 SNPs embedded in the array, samples with an overall call rate of less than 97% were also excluded. As a result, 508 cases and 484 controls were subjected to further analysis. All samples used for GWAS passed a heterozygosity check, and no duplicated and related samples were identified in identity by descent testing. Moreover, principal component analysis found 29 outliers to be excluded via the Smirnov-Grubbs test

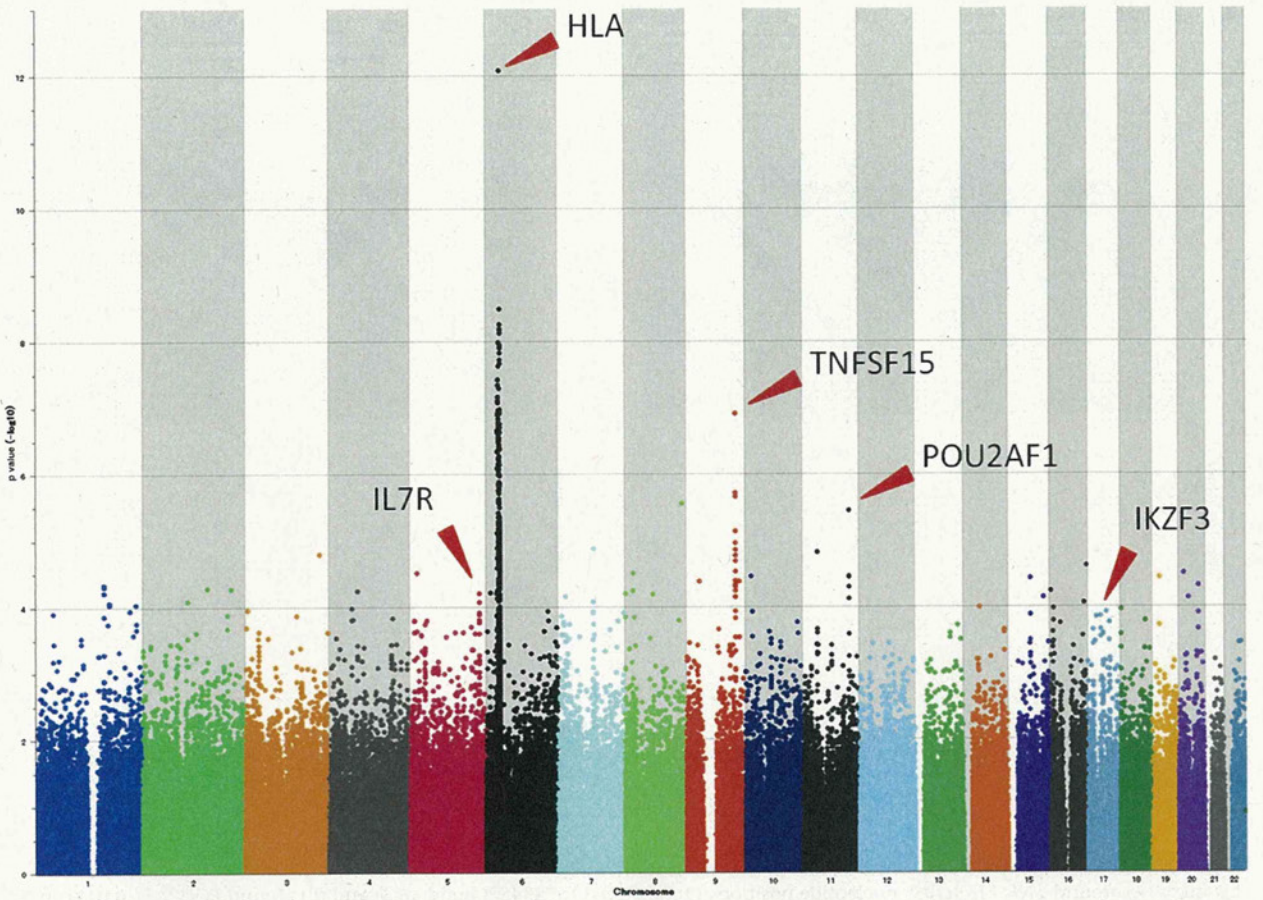


Figure 1. GWAS Results
From 963 samples (487 Japanese PBC cases and 476 Japanese healthy controls), p values were calculated with a chi-square test for allele frequencies among 420,928 SNPs.

and finally showed that all PBC cases (n = 487) and healthy controls (n = 476) formed a single cluster together with the HapMap JPT (Japanese in Tokyo from the CEPH collection), but not with CHB (Han Chinese in Beijing) samples (Figure S1, Table S2). These results indicate that the effect of population stratification was negligible. The average overall call rates of the remaining 487 PBC cases and 476 healthy controls were 99.38% (97.15–99.80) and 99.27% (97.01–99.81), respectively.[11] We then applied the following thresholds for SNP quality control during the data cleaning: SNP call rate ≥95%, minor allele frequency ≥5% in both PBC cases and healthy controls, and Hardy-Weinberg Equilibrium (HWE) p value ≥0.001 in healthy controls.[12] Of the SNPs on autosomal chromosomes and in the pseudoautosomal regions on the X chromosome, 420,928 and 317 passed the quality control filters and were used for the association analysis, respectively (Table S3). A quantile-quantile plot of the distribution of test statistics for the comparison of genotype frequencies in PBC cases and healthy controls showed that the inflation factor lambda was 1.039 for all the tested SNPs, including those in the HLA region, and was 1.026 when SNPs in the HLA region were excluded (Figures S2A and S2B). Table S4 shows the 298 SNPs with $p < 0.0001$ in the GWAS. All cluster plots for the SNPs with a $p < 0.0001$ from a chi-square test of the allele frequency model were checked by visual inspection, and SNPs with ambiguous genotype calls were excluded. For the GWAS and replication study, a chi-square test was applied to a two-by-two contingency table in an allele frequency model.

Figure 1 shows a genome-wide view of the single-point association data, which are based on allele frequencies. We found that the *HLA-DQB1* locus (MIM 604305) had the strongest association with susceptibility to PBC (rs9275175, odds ratio [OR] = 1.94; 95% confidence interval [CI] = 1.62–2.33, $p = 8.30 \times 10^{-13}$) (Figure 1 and Table S4); this finding was consistent with findings from previous studies.[4–7] In addition to the HLA class II region, loci *TNFSF15* and *POU2AF1* showed evidence indicative of association with PBC (rs4979462, OR = 1.63; 95% CI = 1.36–1.95, $p = 1.21 \times 10^{-7}$ for *TNFSF15*; rs4938534, OR = 1.53; 95% CI = 1.28–1.83, $p = 3.51 \times 10^{-6}$ for *POU2AF1*).

In a subsequent replication analysis, 27 SNPs with $p < 0.0001$ in the initial GWAS were also studied, in addition to SNPs at the *TNFSF15* and *POU2AF1* loci. Tagging SNPs were selected from the regions surrounding *TNFSF15* and

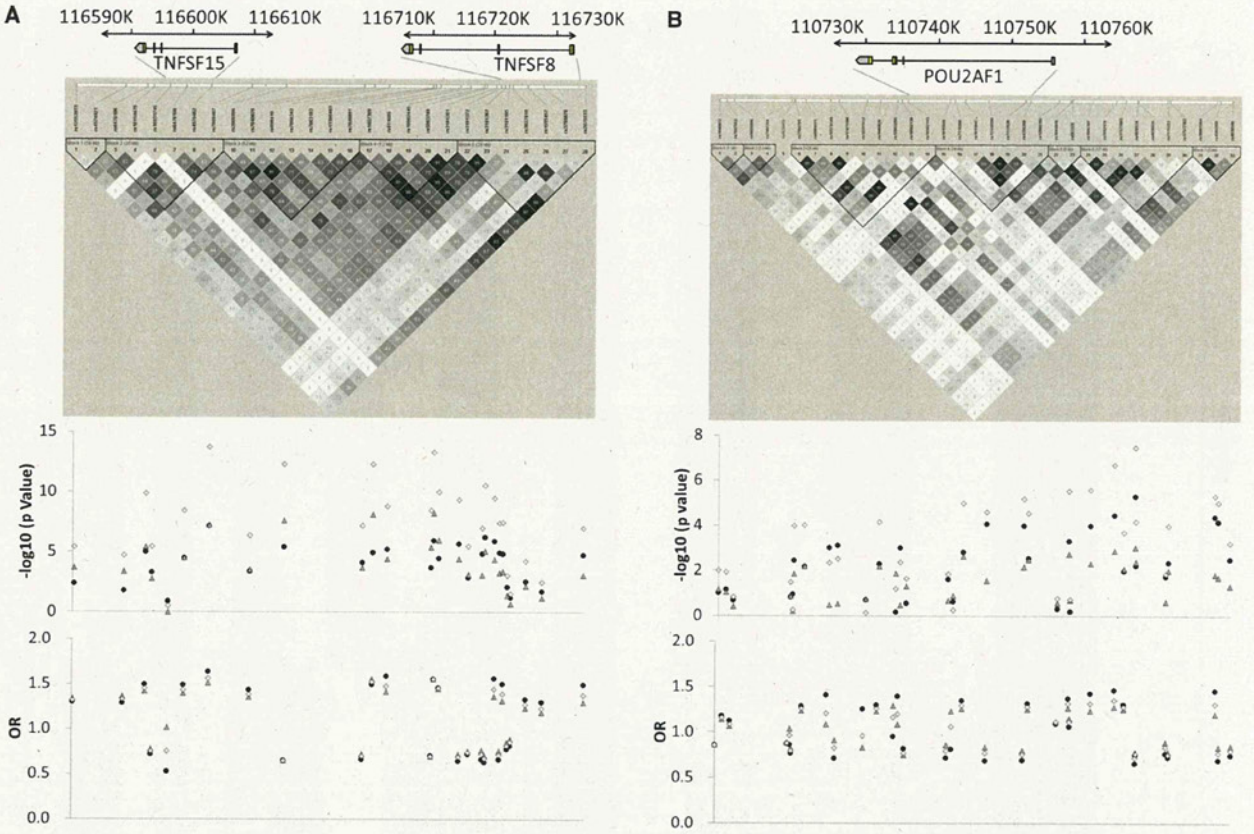


Figure 2. LD Structure, p Values, and OR Plots in the Association Analysis
LD maps (A) around *TNFSF15* (chr9: nucleotide position: 116561403–116733452; build 36.3) and (B) around *POU2AF1* (chr11: nucleotide position: 110684600–110802128; build 36.3). The middle panels show estimates of pairwise r^2 for (A) 28 SNPs and (B) 33 SNPs in the high-density mapping with a total of 2,365 samples used. The bottom panels show p values and OR-based chi-square tests for the allelic model for the left panels of 963 samples in the GWAS (●), the right panels of 1,402 samples in the replication study (▲), and the combined analysis (◇).

POU2AF1 (28 and 33, respectively) for high-density association mapping (Table S5, Figures 2A and 2B). For this follow-up replication analysis, an independent set of 1,402 samples (787 Japanese PBC cases and 615 Japanese healthy controls) and the original set of 963 samples (487 PBC cases and 476 healthy controls) were genotyped with the DigiTag2[13] and custom TaqMan SNP genotyping assays (Applied Biosystems, Foster City, CA, USA) on the LightCycler 480 Real-Time PCR System (Roche, Mannheim, Germany). The strongest associations identified in the initial GWAS were replicated in the independent set of 1,402 samples (OR = 1.52, $p = 5.79 \times 10^{-8}$ for rs4979462; OR = 1.29, $p = 9.32 \times 10^{-4}$ for rs4938534, Table 1). The combined p values were 2.84×10^{-14} (OR = 1.56; 95% CI = 1.39–1.76) for rs4979462 and 2.38×10^{-8} (OR = 1.39; 95% CI = 1.24–1.56) for rs4938534 (Table 1), both of which reached the genome-wide significance level of $p < 5 \times 10^{-8}$. In contrast, the other 27 weakly associated SNPs identified in the initial GWAS (p values <0.0001) were not found to have significant associations with PBC (Table S5). Moreover, no strongly associated SNPs were observed when comparing PBC cases between the early and late stages (Table S5).

A haplotype analysis of the *TNFSF15* and *POU2AF1* regions was conducted with the use of the genotype data from all 2,365 samples (1,274 PBC cases and 1,091 healthy controls). Linkage disequilibrium (LD) blocks were analyzed with Gabriel's algorithm,[14] and five blocks were observed in the *TNFSF15* region and seven blocks in the *POU2AF1* region (Figures 2A and 2B). There were no differences in the LD blocks between PBC cases and healthy controls. The risk haplotypes in each region showed a lower level of association than did the individual SNPs ($p = 8.26 \times 10^{-14}$ for *TNFSF15* and $p = 1.00 \times 10^{-4}$ for *POU2AF1*) (Tables S6 and S7).

Next, we focused on data from our initial GWAS in 21 loci that are reportedly associated with susceptibility to PBC in populations of European descent.[4–7] We found that three such loci (*IL7R*, *IKZF3*, and *STAT4*) had p values of less than 0.001 and eight other such loci (*RAD51L1*, *CXCR5*, *PLCL2*, *IL12RB2*, *NFKB1*, *CD80*, *DENND1B*, and 7p14) showed evidence of marginal associations ($p < 0.05$) in the initial GWAS in 487 Japanese PBC cases and 476 Japanese healthy controls (data not shown). We genotyped three SNPs (rs6890503 for *IL7R*, rs9303277 for *IKZF3*, and rs7574865 for *STAT4*) in an independent set

Table 1. *TNFSF15* SNP rs4979462 and *POU2AF1* SNP rs4938534 Associated with Susceptibility to PBC

dbSNP rsID	Nearest Gene	Risk Allele	Allele (1/2)	Stage	PBC Cases				Healthy Controls				OR[a]	
					11	12	22	RAF	11	12	22	RAF	95% CI	p Value[b]
rs4979462	*TNFSF15*	T	T/C	GWAS	154 (31.8)	244 (50.4)	86 (17.8)	0.57	98 (20.7)	230 (48.5)	146 (30.8)	0.45	1.63 (1.36–1.95)	1.21×10^{-7}
				Replication	253 (32.3)	390 (49.7)	141 (18.0)	0.57	131 (21.6)	305 (50.3)	170 (28.1)	0.47	1.52 (1.30–1.76)	5.79×10^{-8}
				Combined	407 (32.1)	634 (50.0)	227 (17.9)	0.57	229 (21.2)	535 (49.5)	316 (29.3)	0.46	1.56 (1.39–1.76)	2.84×10^{-14}
rs4938534	*POU2AF1*	A	G/A	GWAS	114 (23.6)	229 (47.3)	141 (29.1)	0.53	151 (31.8)	247 (52.0)	77 (16.2)	0.42	1.53 (1.28–1.83)	3.51×10^{-6}
				Replication	179 (22.8)	391 (49.8)	215 (27.4)	0.52	179 (29.4)	299 (49.2)	130 (21.4)	0.46	1.29 (1.11–1.50)	9.32×10^{-4}
				Combined	293 (23.1)	620 (48.9)	356 (28.1)	0.52	330 (30.5)	546 (50.4)	207 (19.1)	0.44	1.39 (1.24–1.56)	2.38×10^{-8}

Parenthetical numbers indicate the percentage of allele 11, 12, or 22 among total alleles in PBC cases or healthy controls. The following abbreviations are used: PBC, primary biliary cirrhosis; RAF, risk allele frequency; and GWAS, genome-wide association study.
[a]Odds ratio (OR) of minor allele from the two-by-two allele frequency table.
[b]p value of Pearson's chi-square test for the allelic model.

of 1,402 samples (787 Japanese PBC cases and 615 Japanese healthy controls) and the original set of 963 samples (487 PBC cases and 476 healthy controls) using the DigiTag2[13] and custom TaqMan SNP genotyping assays. Two SNPs, rs6890853 and rs9303277 located in loci *IL7R* and *IKZF3*, respectively, showed significant associations and the *STAT4* locus (rs7574865) showed suggestive association with PBC in 2,365 Japanese samples (1,274 PBC cases and 1,091 healthy controls) (rs6890853, combined p value = 3.66×10^{-8}, OR = 1.47 for *IL7R*; rs9303277, combined p value = 3.66×10^{-9}, OR = 1.44 for *IKZF3*; rs7574865, combined p value = 1.11×10^{-6}, OR = 1.35 for *STAT4*) (Tables S5 and S8).

Moreover, we genotyped 16 additional associated SNPs, all of which were the same SNPs as identified in previous studies,[4–7] and revealed that six out of 16 SNPs (located on *CXCR5*, *NFKB1*, *CD80*, *DENND1B*, *MAP3K7IP1*, and *TNFAIP2*) were replicated ($p < 0.05$) in 2,365 Japanese samples (Table S8). The SNP rs2293370, located in the *CD80* locus, showed a significant association and the *NFKB1* locus (rs7665090) showed a suggestive association with PBC in the Japanese population (rs2293370, combined p value = 3.04×10^{-9}, OR = 1.48 for *CD80*; rs7665090, combined p value = 1.42×10^{-7}, OR = 1.35 for *NFKB1*). Although further study for determining the primary SNP at each locus is necessary, the remaining ten loci (*RAD51L1*, *PLCL2*, *IL12RB2*, *IRF5*, *SPIB*, *RPS6KA4*, *CLEC16A*, *TNFRSF1A*, *IL12A*, and *MMEL1*) did not show significant association ($p < 0.05$) with PBC in the Japanese population (Table S8).

In the current GWAS in the Japanese population, we identified two significant susceptibility loci for PBC, *TNFSF15* (rs4979462) and *POU2AF1* (rs4938534), which had not been identified in the previous GWAS in populations of European descent. In addition, of the 21 PBC susceptibility loci that have been identified in populations of European descent, three loci (*IL7R*, *IKZF3*, and *CD80*) showed significant associations and two loci (*STAT4* and *NFKB1*) showed suggestive associations with PBC in the Japanese population. Eight other loci (*RAD51L1*, *CXCR5*, *PLCL2*, *IL12RB2*, *DENND1B*, *MAP3K7IP1*, *TNFAIP2*, and 7p14) also showed marginal associations with PBC in the Japanese population. These results indicate the presence of additional important disease pathways (via TNFSF15 and POU2AF1)—differentiation to T helper 1 (Th1) cells (via IL7R and STAT4), B cell differentiation (via IL7R and IKZF3), T cell activation (via CD80), and NF-κB signaling—in addition to the previously reported disease pathways in the development of PBC in Japanese populations.

TNFSF15 is a newly described member of the TNF superfamily that interacts with death receptor 3 (*DR3* [MIM 603366], also known as *TNFRSF25*) not only to promote effector T cell expansion (i.e., Th1 and Th17 cells) and cytokine production (i.e., interferon-γ [IFN-γ, MIM 147570]) at the site of inflammation, but also to induce apoptosis in cells that overexpress DR3.[15] Interestingly, genetic polymorphisms in *TNFSF15* are associated with susceptibility to CD, UC, ankylosing spondylitis (AS, MIM 106300), and leprosy (MIM 609888)[16–20] (Table S8). Strong association of five SNPs (rs3810936, rs6478108, rs6478109, rs7848647, and rs7869487) in the *TNFSF15* region with CD was first reported for a Japanese population,[16] and the finding was replicated in an independent Japanese population and in European-descent and Korean populations.[21–25] Another SNP within *TNFSF15* (rs4263839) is also associated with susceptibility to CD in populations of European descent.[17,20,26] In addition, the risk alleles of the SNPs were significantly associated with *TNFSF15* mRNA expression in peripheral blood.[27,28] Given that there exists strong LD among SNPs in *TNFSF15*, including those in the promoter region (rs6478109 and

rs7848647) and introns (rs4263839 and rs4979462), it is very probable that the PBC susceptibility haplotype containing rs4979462 also influences *TNFSF15* mRNA expression. Additionally, TNFSF15 signaling via DR3 synergizes with interleukin-12 (IL-12) and IL-18 to promote IFN-γ production.[15] The IL-12 signaling pathway includes *IL12A* and *IL12RB* (MIM 601604), variants of which have been identified as PBC susceptibility loci in previous GWASs of peoples of European ancestry, and has been implicated as a key player in the pathogenesis of PBC.[4–7] STAT4 is essential for IL-12 signal transduction via the IL-12 receptor (IL12R) for IFN-γ production and Th1 polarization.[29] Thus, the evidence that *TNFSF15* and *STAT4* were identified and confirmed as PBC susceptibility loci in the present study might indicate that the IL-12 signaling pathway via IL12R is also operative in PBC pathogenesis in Japanese populations, as it is in populations of European descent.

POU2AF1 is a B cell-specific transcriptional factor that coactivates octamer-binding transcriptional factors *POU2F1* (MIM 164175) and *POU2F2* (MIM 164176) on B cell-specific promoters; thus, *POU2AF1* is essential for B cell maturation and germinal center formation.[30] The E-twenty six transcription factor *Spi-B* was recently identified as a direct target of the coactivator *POU2AF1*.[31] *Spi-B* is an important mediator of both B cell receptor signaling and early T cell lineage decisions.[32,33] *Spi-B* also induces IL7R-induced CD40 (MIM 109535, MIM 300386) expression.[34] Given that *Spi-B* has been identified as a PBC susceptibility gene in previous GWASs of peoples of European ancestry,[6,7,35] variation of *POU2AF1* might function along with *Spi-B* in this pathway of B cell signaling and differentiation. The lack of *POU2AF1* reportedly prevents the development of autoimmunity in *Aiolos* (also known as *IKZF3*) mutant mice, which have a systemic lupus erythematosus (MIM 152700)-like phenotype, and in MRL-lpr mice.[36,37] *IKZF3* and *IL7R* were both replicated and confirmed as PBC susceptibility loci in this study; *IKZF3* functions as a transcription factor that participates in the generation of high-affinity bone marrow plasma cells responsible for long-term immunity, and *IL7R* participates in pre-B cell expansion.[38,39] Collectively, these results strengthen the notion that the B cell signaling pathway is involved in the development of PBC.

In conclusion, *TNFSF15* and *POU2AF1* were identified as significant susceptibility loci for PBC in a Japanese population. Our results provide further evidence for the presence of (1) ethnic differences in genetic susceptibility loci (i.e., *TNFSF15*, *IL12A*, and *IL12RB2*), (2) a new autoimmune pathway (i.e., TNFSF15 signaling) shared with other autoimmune diseases (CD, UC, and AS), and (3) common pathogenic pathways such as B cell differentiation (i.e., *POU2AF1*, *IKZF3*, and *SPIB*), IL-12 signaling (i.e., *IL12A*, *IL12RB2*, and *STAT4*), and T cell activation (i.e., *CD80*) for the development of PBC in individuals of European descent and Japanese individuals (Table S8). Functional analysis of these genetic loci, as well as the identification of additional susceptibility loci associated with PBC in eastern Asian populations, should facilitate the analysis of the pathogenesis of PBC worldwide and aid the development of rationale for therapies in the future.

Supplemental Data

Supplemental Data include two figures, eight tables, and Supplemental Acknowledgments and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

The study was approved by the ethics committees of Nagasaki Medical Center and all institutes and hospitals throughout Japan that participated in this collaborative study. All participants provided written informed consent for participation in this study. This study was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (#20590800, #23591006) to M.N., by a Grant-in-Aid for Clinical Research from the NHO to M. Nakamura, by a grant from the Research Program of Intractable Disease provided by the Ministry of Health, Labor, and Welfare of Japan to H.I., and by a grant for Scientific Research on Innovative Areas (Genome Science) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT) to K.T. We thank Yoriko Mawatari, Megumi Sageshima, Yuko Ogasawara, Natsumi Baba, and Rieko Hayashi (University of Tokyo) for technical assistance. We also thank Shinya Nagaoka and Seigo Abiru (NHO Nagasaki Medical Center, Omura, Japan) and Shigeki Hayashi, Hiroshi Mano, Yukio Ohara, Haruhiro Yamashita, Kouki Matsushita, Takeaki Sato, Tsutomu Yamashita, Masahiko Takahashi, Tetsuo Yamamoto, Hironori Sakai, Michio Kato, Fujio Makita, Hitoshi Takaki, and Hideo Nishimura (members of NHOSLJ) for collecting clinical data and blood samples and for obtaining informed consent from PBC cases.

Received: April 15, 2012
Revised: June 3, 2012
Accepted: August 8, 2012
Published online: September 20, 2012

Web Resources

The URLs for data presented herein are as follows:

MEXT Integrated GWAS Database, https://gwas.biosciencedbc.jp/cgi-bin/gwasdb/gwas_top.cgi
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org

References

1. Kaplan, M.M., and Gershwin, M.E. (2005). Primary biliary cirrhosis. N. Engl. J. Med. *353*, 1261–1273.
2. Boonstra, K., Beuers, U., and Ponsioen, C.Y. (2012). Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J. Hepatol. *56*, 1181–1188.
3. Selmi, C., Mayo, M.J., Bach, N., Ishibashi, H., Invernizzi, P., Gish, R.G., Gordon, S.C., Wright, H.I., Zweiban, B., Podda, M., and Gershwin, M.E. (2004). Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology *127*, 485–492.

4. Hirschfield, G.M., Liu, X., Xu, C., Lu, Y., Xie, G., Lu, Y., Gu, X., Walker, E.J., Jing, K., Juran, B.D., et al. (2009). Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. N. Engl. J. Med. *360*, 2544–2555.
5. Hirschfield, G.M., Liu, X., Han, Y., Gorlov, I.P., Lu, Y., Xu, C., Lu, Y., Chen, W., Juran, B.D., Coltescu, C., et al. (2010). Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. Nat. Genet. *42*, 655–657.
6. Liu, X., Invernizzi, P., Lu, Y., Kosoy, R., Lu, Y., Bianchi, I., Podda, M., Xu, C., Xie, G., Macciardi, F., et al. (2010). Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. Nat. Genet. *42*, 658–660.
7. Mells, G.F., Floyd, J.A., Morley, K.I., Cordell, H.J., Franklin, C.S., Shin, S.Y., Heneghan, M.A., Neuberger, J.M., Donaldson, P.T., Day, D.B., et al.; UK PBC Consortium; Wellcome Trust Case Control Consortium 3. (2011). Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat. Genet. *43*, 329–332.
8. Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M., Nakamura, Y., and Kamatani, N. (2008). Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. Am. J. Hum. Genet. *83*, 445–456.
9. Lindor, K.D., Gershwin, M.E., Poupon, R., Kaplan, M., Bergasa, N.V., and Heathcote, E.J.; American Association for Study of Liver Diseases. (2009). Primary biliary cirrhosis. Hepatology *50*, 291–308.
10. Hoffmann, T.J., Kvale, M.N., Hesselson, S.E., Zhan, Y., Aquino, C., Cao, Y., Cawley, S., Chung, E., Connell, S., Eshragh, J., et al. (2011). Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. Genomics *98*, 79–89.
11. Nishida, N., Koike, A., Tajima, A., Ogasawara, Y., Ishibashi, Y., Uehara, Y., Inoue, I., and Tokunaga, K. (2008). Evaluating the performance of Affymetrix SNP Array 6.0 platform with 400 Japanese individuals. BMC Genomics *9*, 431.
12. Miyagawa, T., Nishida, N., Ohashi, J., Kimura, R., Fujimoto, A., Kawashima, M., Koike, A., Sasaki, T., Tanii, H., Otowa, T., et al. (2008). Appropriate data cleaning methods for genome-wide association study. J. Hum. Genet. *53*, 886–893.
13. Nishida, N., Mawatari, Y., Sageshima, M., and 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.
14. Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., et al. (2002). The structure of haplotype blocks in the human genome. Science *296*, 2225–2229.
15. Meylan, F., Richard, A.C., and Siegel, R.M. (2011). TL1A and DR3, a TNF family ligand-receptor pair that promotes lymphocyte costimulation, mucosal hyperplasia, and autoimmune inflammation. Immunol. Rev. *244*, 188–196.
16. Yamazaki, K., McGovern, D., Ragoussis, J., Paolucci, M., Butler, H., Jewell, D., Cardon, L., Takazoe, M., Tanaka, T., Ichimori, T., et al. (2005). Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum. Mol. Genet. *14*, 3499–3506.
17. Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M., et al.; NIDDK IBD Genetics Consortium; Belgian-French IBD Consortium; Wellcome Trust Case Control Consortium. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat. Genet. *40*, 955–962.
18. Zhang, F.R., Huang, W., Chen, S.M., Sun, L.D., Liu, H., Li, Y., Cui, Y., Yan, X.X., Yang, H.T., Yang, R.D., et al. (2009). Genomewide association study of leprosy. N. Engl. J. Med. *361*, 2609–2618.
19. Zinovieva, E., Bourgain, C., Kadi, A., Letourneur, F., Izac, B., Said-Nahal, R., Lebrun, N., Cagnard, N., Vigier, A., Jacques, S., et al. (2009). Comprehensive linkage and association analyses identify haplotype, near to the TNFSF15 gene, significantly associated with spondyloarthritis. PLoS Genet. *5*, e1000528.
20. Latiano, A., Palmieri, O., Latiano, T., Corritore, G., Bossa, F., Martino, G., Biscaglia, G., Scimeca, D., Valvano, M.R., Pastore, M., et al. (2011). Investigation of multiple susceptibility loci for inflammatory bowel disease in an Italian cohort of patients. PLoS ONE *6*, e22688.
21. Kakuta, Y., Kinouchi, Y., Negoro, K., Takahashi, S., and Shimosegawa, T. (2006). Association study of TNFSF15 polymorphisms in Japanese patients with inflammatory bowel disease. Gut *55*, 1527–1528.
22. Tremelling, M., Berzuini, C., Massey, D., Bredin, F., Price, C., Dawson, C., Bingham, S.A., and Parkes, M. (2008). Contribution of TNFSF15 gene variants to Crohn's disease susceptibility confirmed in UK population. Inflamm. Bowel Dis. *14*, 733–737.
23. Yang, S.K., Lim, J., Chang, H.S., Lee, I., Li, Y., Liu, J., and Song, K. (2008). Association of TNFSF15 with Crohn's disease in Koreans. Am. J. Gastroenterol. *103*, 1437–1442.
24. Michelsen, K.S., Thomas, L.S., Taylor, K.D., Yu, Q.T., Mei, L., Landers, C.J., Derkowski, C., McGovern, D.P., Rotter, J.I., and Targan, S.R. (2009). IBD-associated TL1A gene (TNFSF15) haplotypes determine increased expression of TL1A protein. PLoS ONE *4*, e4719.
25. Thiébaut, R., Kotti, S., Jung, C., Merlin, F., Colombel, J.F., Lemann, M., Almer, S., Tysk, C., O'Morain, M., Gassull, M., et al. (2009). TNFSF15 polymorphisms are associated with susceptibility to inflammatory bowel disease in a new European cohort. Am. J. Gastroenterol. *104*, 384–391.
26. Wang, K., Baldassano, R., Zhang, H., Qu, H.Q., Imielinski, M., Kugathasan, S., Annese, V., Dubinsky, M., Rotter, J.I., Russell, R.K., et al. (2010). Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes implicates multiple loci with opposite effects. Hum. Mol. Genet. *19*, 2059–2067.
27. Kakuta, Y., Ueki, N., Kinouchi, Y., Negoro, K., Endo, K., Nomura, E., Takagi, S., Takahashi, S., and Shimosegawa, T. (2009). TNFSF15 transcripts from risk haplotype for Crohn's disease are overexpressed in stimulated T cells. Hum. Mol. Genet. *18*, 1089–1098.
28. Zucchelli, M., Camilleri, M., Andreasson, A.N., Bresso, F., Dlugosz, A., Halfvarson, J., Törkvist, L., Schmidt, P.T., Karling, P., Ohlsson, B., et al. (2011). Association of TNFSF15 polymorphism with irritable bowel syndrome. Gut *60*, 1671–1677.
29. Lleo, A., Gershwin, M.E., Mantovani, A., and Invernizzi, P. (2012). Towards common denominators in primary biliary cirrhosis: the role of IL-12. J. Hepatol. *56*, 731–733.
30. Strubin, M., Newell, J.W., and Matthias, P. (1995). OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. Cell *80*, 497–506.

31. Bartholdy, B., Du Roure, C., Bordon, A., Emslie, D., Corcoran, L.M., and Matthias, P. (2006). The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. Proc. Natl. Acad. Sci. USA *103*, 11665–11670.
32. Garrett-Sinha, L.A., Su, G.H., Rao, S., Kabak, S., Hao, Z., Clark, M.R., and Simon, M.C. (1999). PU.1 and Spi-B are required for normal B cell receptor-mediated signal transduction. Immunity *10*, 399–408.
33. David-Fung, E.S., Yui, M.A., Morales, M., Wang, H., Taghon, T., Diamond, R.A., and Rothenberg, E.V. (2006). Progression of regulatory gene expression states in fetal and adult pro-T-cell development. Immunol. Rev. *209*, 212–236.
34. Nguyen, V.T., and Benveniste, E.N. (2000). Involvement of STAT-1 and ets family members in interferon-gamma induction of CD40 transcription in microglia/macrophages. J. Biol. Chem. *275*, 23674–23684.
35. Hirschfield, G.M., Xie, G., Lu, E., Sun, Y., Juran, B.D., Chellappa, V., Coltescu, C., Mason, A.L., Milkiewicz, P., Myers, R.P., et al. (2012). Association of primary biliary cirrhosis with variants in the CLEC16A, SOCS1, SPIB and SIAE immunomodulatory genes. Genes Immun. *13*, 328–335.
36. Sun, J., Matthias, G., Mihatsch, M.J., Georgopoulos, K., and Matthias, P. (2003). Lack of the transcriptional coactivator OBF-1 prevents the development of systemic lupus erythematosus-like phenotypes in Aiolos mutant mice. J. Immunol. *170*, 1699–1706.
37. Zuo, J., Ge, H., Zhu, G., Matthias, P., and Sun, J. (2007). OBF-1 is essential for the generation of antibody-secreting cells and the development of autoimmunity in MRL-lpr mice. J. Autoimmun. *29*, 87–96.
38. Cortés, M., and Georgopoulos, K. (2004). Aiolos is required for the generation of high affinity bone marrow plasma cells responsible for long-term immunity. J. Exp. Med. *199*, 209–219.
39. Mackall, C.L., Fry, T.J., and Gress, R.E. (2011). Harnessing the biology of IL-7 for therapeutic application. Nat. Rev. Immunol. *11*, 330–342.