Liver INTERNATIONAL
OFFICIAL JOURNAL OF THE INTERNATIONAL ASSOCIATION FOR THE STUDY OF THE LIVER

Liver International ISSN 1478-3223

CLINICAL STUDIES

Percutaneous ethanol injection for hepatocellular carcinoma: 20-year outcome and prognostic factors

Shuichiro Shiina, Ryosuke Tateishi, Masatoshi Imamura, Takuma Teratani, Yukihiro Koike, Shinpei Sato, Shuntaro Obi, Fumihiko Kanai, Naoya Kato, Haruhiko Yoshida, Masao Omata and Kazuhiko Koike

Department of Gastroenterology, University of Tokyo, Tokyo, Japan

Keywords
ablation – hepatocellular carcinoma – percutaneous ethanol injection – prognostic factor – recurrence – survival – treatment outcome

Abbreviations
AFP-L3, lectin-reactive AFP; AFP, α-fetoprotein; BCLC, Barcelona Clinic Liver Cancer; CI, confidence intervals; CT, computed tomography; DCP, des-γ-carboxy-prothrombin; HBs-Ag, hepatitis B surface antigen; HCC, hepatocellular carcinoma; HCV, hepatitis C virus.

Correspondence
Shuichiro Shiina, MD, PhD
Department of Gastroenterology, University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
Tel: +81 3 3815 5411
Fax: +81 3 3814 0021
e-mail: sshiina-tky@umin.ac.jp

Received 4 April 2012
Accepted 22 May 2012

DOI:10.1111/j.1478-3231.2012.02838.x

Abstract

Background: Ethanol injection is the best-known image-guided percutaneous ablation for hepatocellular carcinoma (HCC) and a well-tolerated, inexpensive procedure with few adverse effects. However, there have been few reports on its long-term results. *Aims:* We report a 20-year consecutive case series at a tertiary referral centre. *Methods:* We performed 2147 ethanol injection treatments on 685 primary HCC patients and analysed a collected database. *Results:* Final computed tomography demonstrated complete ablation of treated tumours in 2108 (98.2%) of the 2147 treatments. With a median follow-up of 51.6 months, 5-, 10- and 20-year survival rates were 49.0% [95% confidence interval (CI) = 45.3–53.0%], 17.9% (95% CI = 15.0–21.2%) and 7.2% (95% CI = 4..5–11.5%) respectively. Multivariate analysis demonstrated that age, Child–Pugh class, tumour size, tumour number and serum alpha-fetoprotein level were significant prognostic factors for survival. Five-, 10- and 20-year local tumour progression rates were 18.2% (95% CI = 15.0–21.4%), 18.4% (95% CI = 15.2–21.6%) and 18.4% (95% CI = 15.2–21.6%) respectively. Five-, 10- and 20-year distant recurrence rates were 53.5% (95% CI = 49.4–57.7%), 60.4 (95% CI = 56.3–64.5%) and 60.8% (95% CI = 56.7–64.9%) respectively. There were 45 complications (2.1%) and two deaths (0.09%). *Conclusions:* Ethanol injection was potentially curative for HCC, resulting in survival for more than 20 years. This study suggests that new ablation therapies will achieve similar or even better long-term results in HCC.

Hepatocellular carcinoma (HCC) is the fifth most common malignant neoplasm in the world. Only 20% of HCC patients are candidates for resection (1). Furthermore, recurrence is frequent even after curative resection. Liver transplantation is restricted by donor shortage. Thus, various non-surgical therapies have been introduced (2). Among these, image-guided percutaneous ablation is considered best for early-stage HCC.

The most studied percutaneous ablation is ethanol injection. Ethanol injection is a well-tolerated, inexpensive procedure with few adverse effects and has been considered the standard against which any new ablation therapy should be compared (2). Although ethanol injection was introduced into clinical practice in the 1980s (3, 4), few reports of its long-term results have been published (5–8). We report here a 20-year consecutive case series at a tertiary referral centre. This study documents the largest number of ethanol injection treatments at a single institution. Findings in this 20-year experience may be extrapolated to other ablation therapies, such as radiofrequency ablation, in which such long-term outcomes are not yet available (9).

Patients and methods

Indications for ethanol injection

Ethanol injection was performed in patients satisfying the following criteria: (i) ineligible for resection or transplantation, or had refused surgery; (ii) no extrahepatic metastasis or vascular invasion. Exclusion criteria were as follows: (i) tumour was not visualized

by ultrasonography or not accessible percutaneously; (ii) total bilirubin level ≥ 3.0 mg/dl; (iii) platelet count $<40 \times 10^9$/L; (iv) prothrombin activity <35%; (v) refractory ascites. In general, we performed ethanol injection on patients with Child–Pugh class A or B, with 3 or fewer tumours ≤ 3 cm in diameter. We performed ethanol injection on patients beyond these conditions, however, who were likely to benefit from the procedure for possible cure or prolongation of life. No patients were excluded solely because of tumour location (10). Informed consent was obtained from each patient. This study was conducted according with the Helsinki Declaration of 1975 and approved by the Institutional Review Board.

Patients

In this cohort study, we analysed a prospectively collected computerized database. Between 1985 and 2005, 2735 HCC patients were admitted to the Department of Gastroenterology, University of Tokyo (Fig. 1). At initial hospitalization, 1615 had primary HCC and the remaining 1120 had recurrent HCC. The recurrent HCC patients had undergone therapies other than ethanol injection for primary HCC.

Of the 1615 patients with primary HCC, 1459 (90.3%) underwent percutaneous ablation as the initial treatment, including ethanol injection. The remaining 156 patients received other therapies: transarterial chemoembolization for 123 patients with multinodular or large tumours that could not be treated by ablation therapies; hepatic resection for 18 with good liver function who consented to an operation; chemotherapy for four with vascular invasion or extrahepatic metastasis; and best supportive care for 11 with decompensated cirrhosis or poor general condition.

Of the 1459 patients treated by percutaneous ablation, 685 underwent ethanol injection, 122 underwent microwave ablation, and the remaining 652 radiofrequency ablation. The type of percutaneous ablation performed varied with the date of treatment. We started ethanol injection in December 1985, microwave ablation in October 1995 and radiofrequency ablation in February 1999 (11). Between October 1995 and February 1999, both ethanol injection and microwave ablation were performed. Microwave ablation was chosen for patients who had better liver function and whose tumour was located in a position where the electrode could be inserted and held safely. Since February 1999, both ethanol injection and radiofrequency ablation have been performed. Between April 1999 and January 2001, 232 patients with three or fewer tumours, each ≤ 3 cm in diameter, and Child–Pugh class A or B were entered into a randomized controlled trial (12). Patients outside these inclusion criteria were mostly treated by radiofrequency ablation. After this trial, radiofrequency

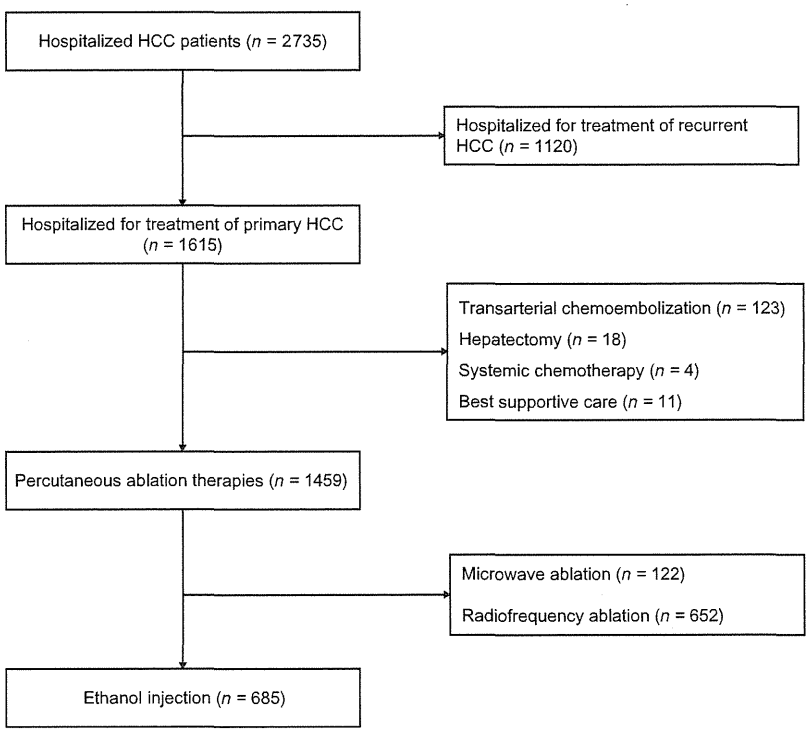

Fig. 1. Flow of patients in this study. HCC, hepatocellular carcinoma

ablation was generally the treatment of choice, and ethanol injection was used only in those unsuitable for radiofrequency ablation: those with either enterobiliary reflux or tumour adhesion to the gastrointestinal tract.

Hepatocellular carcinoma was diagnosed based on typical imaging findings of early phase enhancement and late phase contrast washout on computed tomography (CT) (13). HCC diagnosis was also confirmed by biopsy in 630 (92.0%) of the 685 patients with primary HCC treated by ethanol injection. A total of 587 (85.7%) were diagnosed as having cirrhosis.

In general, chemoembolization was combined with ethanol injection in patients with either ≥ 4 tumours or those with two or three tumours at least one of which is >3.0 cm in diameter. The combination of chemoembolization with ethanol injection was performed in 186 patients.

Treatment methods

Preoperative planning including ultrasound examination and evaluation of all imaging findings was performed to identify the tumours and to determine the access route. The procedure was performed according to an institutional protocol and under the supervision of experienced physicians who had performed this treatment more than 200 times. The precise techniques of ethanol injection are described elsewhere (12). Briefly, all procedures were performed percutaneously under ultrasound guidance. Artificial pleural effusion or artificial ascites method is much less frequently used in ethanol injection compared with radiofrequency ablation, because the procedure is necessary to be repeated several times. Since 1990, we have used two or three needles to inject ethanol into several sites in one procedure (12). Ethanol injection was performed twice per week. The procedure was repeated until ethanol appeared to have been injected throughout the tumour. To judge a timing to stop repetition of injecting ethanol and to order a CT scan, we considered total volume of injected ethanol and change of echogenicity. The general guideline for the necessary volume of injected ethanol was calculated according to the following numerical expression, $V = (4/3)\ \pi\ (r + 0.5)^3$, where V (in millilitres) is the volume of ethanol and r (in centimetres) is the radius of the tumour; 0.5 is added to provide a safety margin, which is based on the concept that some surrounding liver parenchyma all around the tumour as well as the tumour itself must be ablated (5).

A CT scan was then performed 1–3 days after the procedure to evaluate technique effectiveness (14). Complete ablation was defined as hypoattenuation of the entire tumour. When the presence of unablated tumour portions was suspected, a few more procedures were performed. We did not predefine the number of procedures in a treatment. The ethanol injection treatment was generally continued until CT demonstrated the entire tumour necrosis.

Follow-up

Follow-up investigations consisted of CT, ultrasonography and measurement of serum α-fetoprotein (AFP), des-γ-carboxy-prothrombin (DCP) (since April 1993) levels and lectin-reactive AFP (AFP-L3) (since July 1997) every 4 months. Local tumour progression was defined as appearance of viable tumour touching the original tumour (14) and distant recurrence as emergence of tumour(s) separate from the primary site. Ethanol injection was used for recurrence if the patient still met the indication criteria. If multiple recurrences were not treatable with ethanol injection, chemoembolization was generally performed.

Statistical analyses

This study is a report of a consecutive case series. All ethanol injection treatments performed on primary HCC patients at the Department of Gastroenterology, University of Tokyo between 1985 and 2005 were included. Data are presented as mean ± SD for quantitative variables, and as absolute frequencies for qualitative variables.

A 'procedure' was defined as a single intervention episode that consisted of one or more ablations performed on tumours, and a 'treatment' as the completed effort to ablate tumours. A treatment consisted of several procedures (14). 'Technique effectiveness' rate was defined as the percentage of successfully eradicated macroscopic tumours as evidenced at CT scan after the last procedure (14). In cases in which there was Lipiodol deposit inside the tumour because of the combination of chemoembolization with ethanol injection, we judged that the tumour had been successfully eradicated if it was surrounded with completely non-enhanced tissue in final CT.

Overall survival was calculated in the 685 primary HCC patients. Survival curves were generated using the Kaplan–Meier method. In addition to overall survival, subgroup analyses were performed with clinical characteristics including tumour size, tumour number and Child–Pugh class. Recurrence was evaluated in 591 patients in whom ethanol injection was performed with curative intent. All tumours were treated by ethanol injection in those patients. The remaining 94 patients were excluded from the recurrence analysis because some small tumours had been left untreated by ethanol injection on account of detection failure by ultrasonography. Recurrence rates were calculated using the Gaynor method (15). All time estimates were made from the date of the first ethanol injection. The follow-up was finalized at either death or the last visit to the outpatient clinic before December 31 2010. Transplanted patients were censored from this study at the date of transplantation.

The prognostic relevance of baseline variables (Table 1), the combination of chemoembolization,

Table 1. Baseline characteristics of the 685 Patients undergoing percutaneous ethanol injection for primary hepatocellular carcinoma

Variable	
Age (years)	64.0 ± 8.9
Males, *n* (%)	502 (73.3)
Viral infection*	
HBs-Ag positive, *n/N* (%)	64/685 (9.3)
Anti-HCV positive, *n/N* (%)	570/673 (84.7)
Both positive, *n/N* (%)	11/673 (1.6)
Both negative, *n/N* (%)	52/673 (7.7)
Alcohol consumption >80 g/day, *n* (%)	143 (20.9)
Ascites, *n* (%)	122 (17.9)
Encephalopathy, *n* (%)	44 (6.5)
Albumin (g/dl)	3.55 ± 0.50
Total bilirubin (mg/dl)	0.96 ± 0.536
Prothrombin time (%)	71.6 ± 15.9
Platelet count ($\times 10^4$/mm^3)	10.3 ± 4.6
AST (IU/L)	80.6 ± 48.2
ALT (IU/L)	79.2 ± 61.9
Child–Pugh class, *n* (%)	
A	425 (62.1)
B	228 (33.3)
C	32 (4.6)
Tumour size (cm)	2.83 ± 1.47
Tumour number	2.0 ± 1.7
Serum AFP (ng/ml), *n* (%)	
≤100	525 (76.6)
101–400	95 (13.9)
>400	65 (9.5)
Serum DCP (mA U/ml), *n* (%)†	
≤100	428 (82.8)
101–400	49 (9.5)
>400	40 (7.7)
Serum AFP-L3 (%), *n* (%)‡	
≤15	193 (86.2)
15.1–40	16 (7.1)
>40	15 (6.7)

*Anti-HCV was not tested in 12 patients.
†Serum DCP level was not measured in 168 patients.
‡Serum AFP-L3 level was not measured in 461 patients.
HBs-Ag, hepatitis B surface antigen; HCV, hepatitis C virus; AFP, α-fetoprotein; DCP, des-gamma-carboxy-prothrombin; AFP-L3, lectin-reactive α-fetoprotein.
Data are expressed as mean ± standard deviation.

HCC recurrence and the number of ethanol injection sessions to survival was analysed by univariate and multivariate models. The prognostic relevance of baseline variables (Table 1), the combination of chemoembolization and the number of ethanol injection sessions to local tumour progression and distant recurrence was also analysed by univariate and multivariate models. In multivariate analysis, we evaluated models including Child–Pugh class and excluding its components to avoid multicollinearity. Serum DCP and AFP-L3 levels were excluded from the multivariate model because of absence of data from 168 and 461 patients respectively. Some continuous variables in which log-linearity could not be assumed were transformed into categorical variables. Variables with a P value <0.05 determined by univariate comparison were subjected to multivariate analysis. A stepwise variable selection was performed with Akaike Information Criteria in multivariate analysis. Results were expressed as hazard ratios with corresponding 95% confidence intervals (CI), with P values from the Wald test. All significance tests were two-tailed, and differences with a P value <0.05 were considered statistically significant.

Complications were defined according to the guidelines of the Society of Interventional Radiology (16).

Results

Antitumour effect

We performed 2147 ethanol injection treatments, comprising 13 526 procedures. Thus, procedure number per treatment was 6.3 ± 2.6. The total volume of injected ethanol per treatment was 40.9 ± 16.3 ml. Many patients received iterative ethanol injection treatments for recurrence. A total of 108 patients underwent ethanol injection treatment once, 118 patients twice, 196 patients 3 times, 153 patients 4 times, 71 patients 5 times, 28 patients 6 times, 8 patients 7 times and 3 patients 8 times.

Technique effectiveness rate was 98.2% (2108/2147 treatments). It was similar between the initial ethanol injection treatments and the other ethanol injection treatments for recurrence ($P = 0.397$). Complete ablation of the tumour was achieved in 675 (98.5%) of the 685 initial treatments and in 1433 (98.0%) of the 1462 other treatments. However, technique effectiveness rate significantly differed with tumour size ($P = 0.002$). No apparent viable portions remained in 758 (99.0%) of 766 treatments for tumours ≤2.0 cm in diameter, in 704 (98.4%) of 717 treatments for tumours 2.1–3.0 cm, in 570 (97.9%) of 582 treatments for tumours 3.1–5.0 cm and in 76 (92.7%) of 82 treatments for tumours >5.0 cm.

Survival

Table 1 shows clinical characteristics of the 685 patients. A total of 136 patients (19.9%) were older than 75 years. In all, 180 patients had tumours ≤2.0 cm in diameter, 274 had tumours 2.1–3.0 cm, 192 had tumours 3.1–5.0 cm and 39 had tumours >5.0 cm. A total of 367 patients had one tumour, 238 patients had 2 or 3 tumours and 80 had 4 or more tumours.

As of December 2010 (with a median follow-up of 51.6 months), 70 patients (10.2%) remained alive, 52 (7.6%) were lost to follow-up and 563 (82.2%) had died. Of the 685 patients, two were transplanted. The number of patients who survived longer than 5, 10 and 20 years after the first ethanol injection treatment was 305, 97 and 3 respectively. The cause of death was HCC

in 297 patients (52.8%), liver failure in 129 (22.9%), upper gastrointestinal bleeding in 30 (5.3%), complications related to the procedure in 2 (0.4%), liver-unrelated diseases in 84 (14.9%) and undetermined in 21 (3.7%).

The 1-, 3-, 5-, 10-, 15- and 20-year survival rates of all 685 patients were 91.0% (95% CI = 88.9–93.2%), 67.6% (95% CI = 64.1–71.3%), 49.0% (95% CI = 45.3–53.0%), 17.9% (95% CI = 15.0–21.2%), 8.6% (95% CI = 6.4–11.7%) and 7.2% (95% CI = 4.5–11.5%) respectively (Fig. 2; Table 2). Survival rates significantly differed with tumour number ($P = 0.0001$), tumour size ($P = 0.0001$) and Child–Pugh class ($P = 0.0001$). In patients with 1–3 tumours, all ≤ 3 cm, and in Child–Pugh class A or B, the 5-year survival rate was 59.5% (95% CI: 54.7–64.7%).

Fig. 2. Overall survival in 685 primary hepatocellular carcinoma patients who underwent ethanol injection.

Univariate analysis indicated that 13 of the 22 variables were relevant to survival. In multivariate analysis, a model that contained age, antibody to hepatitis C virus (anti-HCV), Child–Pugh class, tumour size, tumour number and serum AFP level was selected (Table 3).

Survival rates significantly differed with the time period in which the first ethanol injection was performed ($P < 0.0001$; Fig. 3). In 109 patients who underwent ethanol injection between 1985 and 1991, the 5-year survival rate was 30.3% (95% CI = 22.7–40.5%), whereas it was 51.2% (95% CI = 46.8–55.9%) in 476 patients between 1992 and 1998, and 61.1% (95% CI = 51.3–72.8%) in 100 patients between 1999 and 2005.

Recurrence

Recurrence developed in 449 patients. Local tumour progression alone was found in 61 patients, local tumour progression with distant recurrence in 44 and distant recurrence alone in 344. Of these 344 patients, eight had recurrence in extrahepatic sites: five had lymph node metastasis, one had lung metastasis, one had bone metastasis and the remainder had both lymph node and lung metastasis. Of the 449 patients, the first recurrence was treated by iterative ethanol injection in

Table 2. Survival of patients undergoing ethanol injection, based on tumour number, tumour size and Child–Pugh class

Grading	*n*	Survival (%) 3-Year	5-Year	10-Year	15-Year	20-Yar	Median (years)	*P* value
Overall survival	685	67.6	49.0	17.9	8.6	7.2	4.9	–
Tumour number								
Solitary	367	72.0	56.5	24.6	12.1	9.7	5.8	0.0001
2–3	232	71.5	46.3	12.9	5.9	–	4.7	
≥4	86	37.6	23.8	2.5	1.3	–	2.6	
Tumour size								
≤2.0 cm	240	83.6	63.8	27.6	12.3	6.1	6.9	0.0001
2.1–3.0 cm	221	68.0	47.9	15.0	10.7	10.7	4.8	
>3.0 cm	224	50.2	34.4	10.1	3.5	3.5	3.1	
Child–Pugh class								
A	425	77.3	58.7	24.4	12.5	10.4	6.2	0.0001
B	228	53.9	35.5	8.1	3.0	–	3.5	
C	32	37.5	18.8	3.1	–	–	1.9	
Combination of tumour number, tumour size, and Child–Pugh class								
Solitary, ≤3 cm	275	77.5	62.2	28.8	14.5	10.8	6.8	–
Solitary, ≤3 cm, Child–Pugh A	185	84.9	69.2	36.7	20.2	15.1	7.6	–
1–3 tumours, ≤3 cm	419	78.6	58.0	23.5	12.2	9.1	6.1	–
1–3 tumours, ≤3 cm, Child–Pugh A/B	402	80.5	59.5	24.3	12.8	9.6	6.2	–
Satisfied the indication criteria of surgical resection proposed in the BCLC protocol*	121	86.3	72.8	31.1	14.8	–	7.2	–

*Child–Pugh class A with a normal level of bilirubin, no significant portal hypertension and a single HCC.
BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma.

Table 3. Multivariate analysis of variables relevant to survival, local tumour progression and distant recurrence

Variable	Multivariate analysis Hazard ratio (95% CI)	*P* value
Survival		
Age (per year)	1.03 (1.02–1.04)	<0.0001
Anti-HCV-positive	0.81 (0.69–0.94)	0.006
Child–Pugh class		
A	1	
B	2.01 (1.66–2.44)	<0.0001
C	3.11 (2.08–4.65)	<0.0001
Tumour size (cm)		
≤2.0	1	
2.1–3.0	1.26 (1.00–1.58)	0.051
3.1–5.0	1.51 (1.18–1.93)	0.001
>5.0	2.31 (1.61–3.31)	<0.0001
Tumour number		
solitary	1	
2–3	1.10 (0.90–1.35)	0.34
≥4	2.11 (1.59–2.78)	<0.0001
Serum AFP (ng/dl)		
≤100	1	
101–400	1.47 (1.14–1.90)	0.003
>400	2.16 (1.57–2.97)	<0.0001
Local tumour progression		
Tumour size (cm)		
≤2.0	1	
2.1–3.0	1.47 (1.15–1.88)	0.002
3.1–5.0 vs. ≤2.0	1.30 (0.97–1.75)	0.08
>5.0 vs. ≤2.0	2.81 (1.64-4.82)	0.0002
Distant recurrence		
Tumour size (cm)		
≤2.0	1	
2.1–3.0	1.42 (1.11–1.82)	0.006
3.1–5.0	1.28 (0.95–1.72)	0.10
>5.0	2.48 (1.43–4.28)	0.001
Tumour number		
solitary	1	
2–3	1.47 (1.16–1.85)	0.001
≥4	2.12 (1.36–3.28)	0.0008

AFP, α-fetoprotein; CI, confidence interval; HCV, hepatitis C virus.

399 (88.8%), chemoembolization in 44 (9.8%), systemic chemotherapy in three (0.7%) and best supportive care in three (0.7%).

The 1-, 3-, 5-, 10-, 15- and 20-year rates of local tumour progression with or without distant recurrence were 7.9% (95% CI = 5.7–10.0%), 15.6% (95% CI = 12.6–18.6%), 18.2% (95% CI = 15.0–21.4%), 18.4% (95% CI = 15.2–21.6%), 18.4% (95% CI = 15.2–21.6%) and 18.4% (95% CI = 15.2–21.6%) respectively (Fig. 4). Univariate analysis demonstrated that three variables were relevant to local tumour progression, whereas multivariate analysis indicated that only tumour size was significantly related to local tumour progression (Table 3).

The 1-, 3-, 5-, 10-, 15- and 20-year rates of distant recurrence without local tumour progression were 17.1% (95% CI = 14.0–20.1%), 42.6% (95% CI = 38.6–46.7%), 53.5% (95% CI = 49.4–57.7%), 60.4% (95% CI = 56.3–64.5%), 60.8% (95% CI = 56.7–64.9%) and 60.8% (95% CI = 56.7–64.9%) respectively. Univariate analysis demonstrated that five variables were relevant to distant recurrence, whereas multivariate analysis indicated that tumour size and tumour number were significantly related to distant recurrence without local recurrence (Table 3).

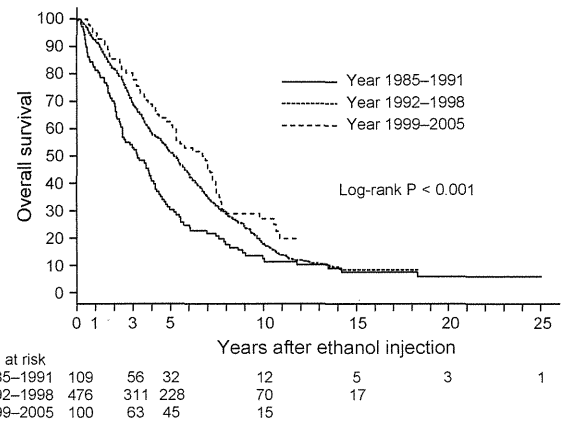

Fig. 3. Survival according to the time period in which the first ethanol injection was performed (1985–1991 vs. 1992–1998 vs. 1999–2005)

Complications

Table 4 shows complications encountered. The incidence rates per treatment and per procedure were 2.1% (45 of 2147) and 0.33% (45 of 13 526) respectively. A patient died of multiple organ dysfunction syndrome caused by procedure-related hemoperitoneum. The tumour was not on the surface but inside the liver. The patient did not have marked bleeding tendency. The other developed myocardial infarction, resulting in death during the procedure. The treatment mortality rate was 0.06%.

Discussion

This study describes a 20-year experience with ethanol injection at a high-volume centre. We performed 2147 ethanol injection treatments on the 685 primary HCC patients, showing that ethanol injection has a high antitumour effect. Tumours were judged to have been completely ablated by final CT imaging in 98.2% of the treatments. The complete response rate may be higher in this study than others (17, 18), probably because we did not predefine the number of procedures in a treatment. We generally repeated the procedure until CT demonstrated complete tumour necrosis. Many other studies limited the procedure number of ethanol injection. Complete tumour ablation has been reported to relate to improved survival (19).

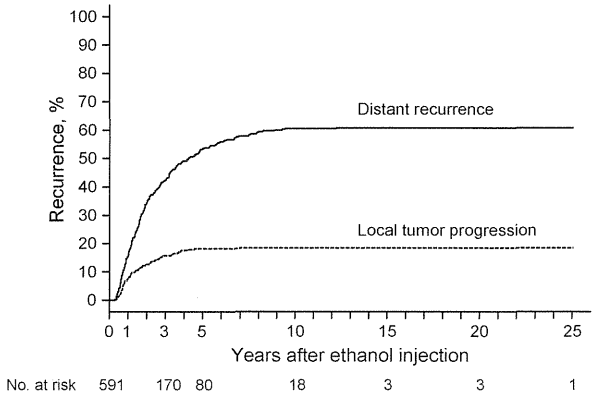

Fig. 4. Local tumour progression or distant recurrence in patients who underwent ethanol injection.

This study showed that ethanol injection could achieve long-term survival over 20 years. Ninety-seven patients survived for more than 10 years and three for more than 20 years. Both tumour factors and liver function were relevant to survival. In addition, age was among the prognostic factors. In this study, 19.9% were older than75 years, which may have resulted in the higher percentage (14.9%) of liver-unrelated deaths compared with other studies. Ant-HCV positivity was a good prognostic factor in this study.

Survival in ethanol injection appears to have improved with times. This is probably because of advances in imaging techniques, such as ultrasound and CT, more refined skills and greater experience in ablation and innovations in the treatment of underlying liver diseases.

Hepatocellular carcinoma frequently recurred after ethanol injection. Most recurrences were, however, not local tumour progression but distant recurrence. Frequent recurrence is not specific to ethanol injection. After hepatic resection, the tumour recurrence rate exceeds 70% at 5 years (20, 21). In this study, periodic follow-up detected most recurrence at limited stage. Ethanol injection was performed again for first recurrence in 88.8% of the cases. In hepatic resection, the rate of repeat resection for first recurrence has been reported to range from 10.4 to 30.6% (21, 22). As ethanol injection is less invasive than hepatic resection, iterative ethanol injection can be performed for recurrence more easily.

Ethanol injection was a safe procedure, although many patients in this study were at risk for surgical treatment because of advanced cirrhosis or other comorbidities. Only 121 (17.7%) of the 685 patients satisfied the indication criteria of surgical resection proposed in the BCLC (Barcelona Clinic Liver Cancer) protocol (23) and were, thus, considered good candidates for surgical resection. Other investigators also reported low complication rates of 0–3.2% (6–8, 24).

Table 4. Complications in 2147 treatments of ethanol injection for hepatocellular carcinoma

Complication	Number
Neoplastic seeding	9
Hemoperitoneum	9
Hemobilia	6
Liver abscess	6
Symptomatic pleural effusion	3
Massive hepatic infarction	3
Biliary cast	2
Hemothorax	2
Abnormal decrease in blood coagulation factor VIII	2
Biloma	1
Biliary bronchial fistula	1
Myocardial infarction	1

For hepatic resection, morbidity rates have been reported to be 38–47% even in recent studies (25–27).

Radiofrequency ablation has steadily replaced ethanol injection (11). At our institution, radiofrequency ablation is currently the first option for percutaneous ablation (28). Several randomized controlled trials including ours (12, 18, 29, 30) demonstrated more reliable local antitumour effect and higher survival. Our 10-year outcome of radiofrequency ablation (28) appears superior to this 20-year outcome of ethanol injection. In addition, radiofrequency ablation requires fewer treatment sessions and shorter hospitalization.

A meta-analysis showed, however, that ethanol injection did not differ from radiofrequency ablation for tumours ≤ 2 cm in diameter (31). A recent randomized controlled trial also demonstrated similar 5-year survival between the two ablations (32). Ethanol injection is at least more feasible and cheaper than radiofrequency ablation.

Surgical resection has been considered the treatment of first choice for HCC. Our first option for resectable tumours was also surgery. However, most patients who came to our department declined surgical resection. Thus, some patients in this study underwent ethanol injection not because of unresectable tumour but because of refusal of surgery. Those who preferred surgery would have gone directly to the surgical department, which has extensive experience in hepatic resection (27).

It is not easy to compare outcomes between ethanol injection and surgical resection. Indications are different between the two treatments. Furthermore, indications for each treatment are different from institution to institution. Thus, a case adjudged to be treatable by ethanol injection or surgical resection at an institution may not be given the same treatment at another. The best-known indication criteria may be those proposed in the BCLC protocol (23), which states that surgical resection should be restricted to patients with performance status 0, Child–Pugh class A, single HCC, normal portal pressure and normal serum bilirubin level. In patients satisfying

those criteria, the 5-year survival rate is expected to be >70% (20). In this study, 5-year survival rate of the patients satisfied the criteria was 72.8%, which appears satisfactory when compared with outcomes following surgical resection. Furthermore, in patients with solitary HCC, ≤ 3 cm in diameter, and Child–Pugh A, 5- and 10-year survival rates were 69.2% and 36.7% respectively. In patients treated by surgical resection, 5- and 10-year survival rates were 34.4–70.0% and 10.5–52.0% respectively (22, 33–39). Although this is an observational study with no control, survivals following ethanol injection appear comparable to those reported following surgical resection.

A randomized controlled trial showed no significant difference in survival between ethanol injection and surgical resection (40). Several non-randomized controlled trials also reported similar overall survival between the two treatments (5–7, 40–43), whereas others reported higher survival with resection (44). Further studies are necessary to resolve this issue of comparing ablation with resection.

We made strenuous efforts to standardize the procedure of ethanol injection because many physicians performed ethanol injection at our institution. In addition to proficient practice of ethanol injection, detailed preoperative planning, cautious postoperative evaluation of therapeutic effect and careful follow-up are vital to achieve satisfactory outcomes.

Source population in this study may represent selection bias, as we performed ethanol injection on most patients who were hospitalized at our department; however, many patients with unfavourable tumour conditions for ethanol injection might not have been referred to us. Therefore, caution is required when extrapolating our findings to the general population of HCC patients.

A second limitation is that study population cannot be clearly defined. This study was based on daily clinical practice over a 20-year period. Indication criteria of ethanol injection changed over time, mainly because of the introduction of the other ablations: microwave ablation and radiofrequency ablation. Furthermore, various treatments besides percutaneous ablations were available for HCC, such as surgical resection and chemoembolization, with frequently overlapping indications.

One further limitation is the fact that this was a single-centre study. To extrapolate the findings in this study to patients at other institutions, consideration should be given to differences in the indications, methods, expertise, performance of available ultrasound and CT equipment and others. Treatment outcome may be influenced by the physicians' expertise and the institution's volume of care. We performed over 2000 ethanol injection treatments, which may represent a much greater number of treatments than those in most other institutions.

In conclusion, our 20-year experience shows that ethanol injection was potentially curative, resulting in long-term survival over 20 years. Findings in this study may suggest that other ablation therapies, such as radiofrequency ablation, will achieve similar or even better long-term results in HCC.

Acknowledgements

The work had no specific funding.

Financial disclosures: There are no financial disclosures from any authors.

References

1. Borie F, Bouvier AM, Herrero A, *et al.* Treatment and prognosis of hepatocellular carcinoma: a population based study in France. *J Surg Oncol.* 2008; **98**: 505–9.
2. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology.* 2005; **42**: 1208–36.
3. Livraghi T, Festi D, Monti F, Salmi A, Vettori C. US-guided percutaneous alcohol injection of small hepatic and abdominal tumors. *Radiology.* 1986; **161**: 309–12.
4. Shiina S, Yasuda H, Muto H, *et al.* Percutaneous ethanol injection in the treatment of liver neoplasms. *AJR Am J Roentgenol.* 1987; **149**: 949–52.
5. Shiina S, Tagawa K, Niwa Y, *et al.* Percutaneous ethanol injection therapy for hepatocellular carcinoma: results in 146 patients. *AJR Am J Roentgenol.* 1993; **160**: 1023–8.
6. Livraghi T, Giorgio A, Marin G, *et al.* Hepatocellular carcinoma and cirrhosis in 746 patients: long-term results of percutaneous ethanol injection. *Radiology.* 1995; **197**: 101–8.
7. Lencioni R, Pinto F, Armillotta N, *et al.* Long-term results of percutaneous ethanol injection therapy for hepatocellular carcinoma in cirrhosis: a European experience. *Eur Radiol.* 1997; **7**: 514–9.
8. Ebara M, Okabe S, Kita K, *et al.* Percutaneous ethanol injection for small hepatocellular carcinoma: therapeutic efficacy based on 20-year observation. *J Hepatol.* 2005; **43**: 458–64.
9. 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.
10. Teratani T, Yoshida H, Shiina S, *et al.* Radiofrequency ablation for hepatocellular carcinoma in so-called high-risk locations. *Hepatology.* 2006; **43**: 1101–8.
11. Shiina S, Teratani T, Obi S, *et al.* Nonsurgical treatment of hepatocellular carcinoma: from percutaneous ethanol injection therapy and percutaneous microwave coagulation therapy to radiofrequency ablation. *Oncology.* 2002; **62**(Suppl. 1): 64–8.
12. Shiina S, Teratani T, Obi S, *et al.* A randomized controlled trial of radiofrequency ablation with ethanol injection for small hepatocellular carcinoma. *Gastroenterology.* 2005; **129**: 122–30.
13. Araki T, Itai Y, Furui S, Tasaka A. Dynamic CT densitometry of hepatic tumors. *AJR Am J Roentgenol.* 1980; **135**: 1037–43.
14. Goldberg SN, Grassi CJ, Cardella JF, *et al.* Image-guided tumor ablation: standardization of terminology and reporting criteria. *Radiology.* 2005; **235**: 728–39.

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.

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]	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
*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

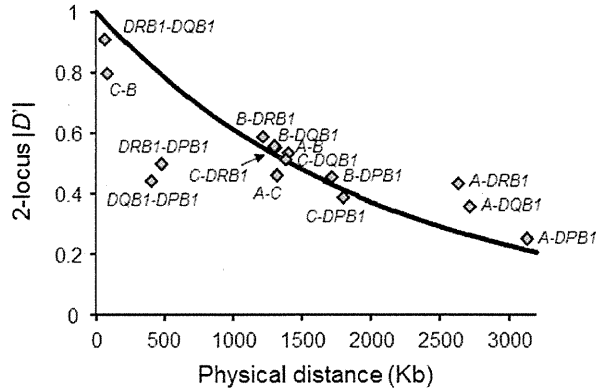

Figure 2. Relationship between two-locus |D'| and physical distance (Kb). A solid-line curve, $2-locus|D'|=(1-0.67\times10^{-5}\times x)^{75,13}$, was obtained using the least-squares method, where x represents the physical distance (Kb). The recombination rate in the *HLA* region was assumed to be 0.67 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

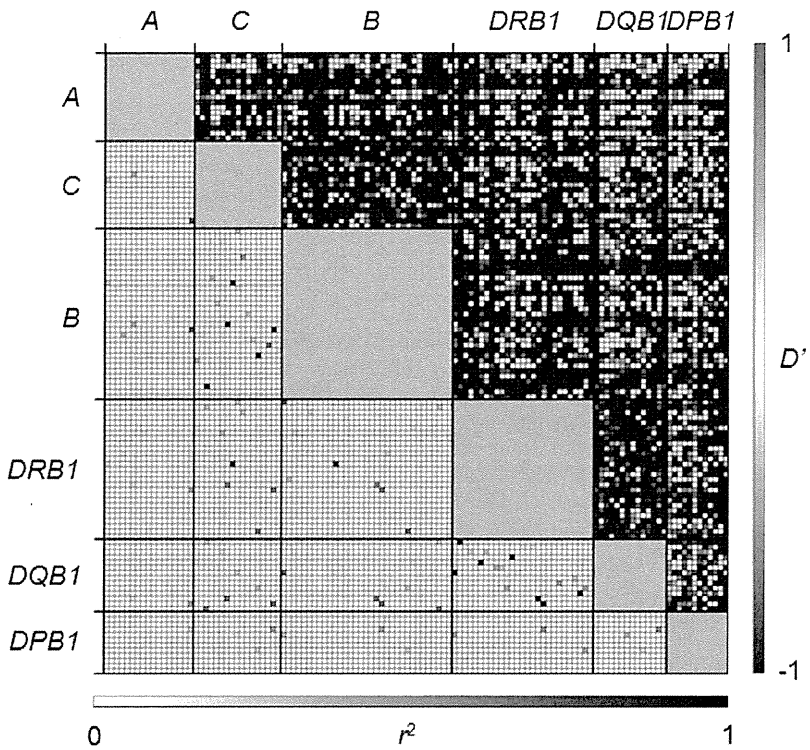

Figure 1. Pairwise estimates of LD parameters, |D'| (upper diagonal) and r^2 (lower diagonal) for every pair of *HLA* alleles. The name of each allele is presented in Data S1.
doi:10.1371/journal.pone.0046806.g001

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*.

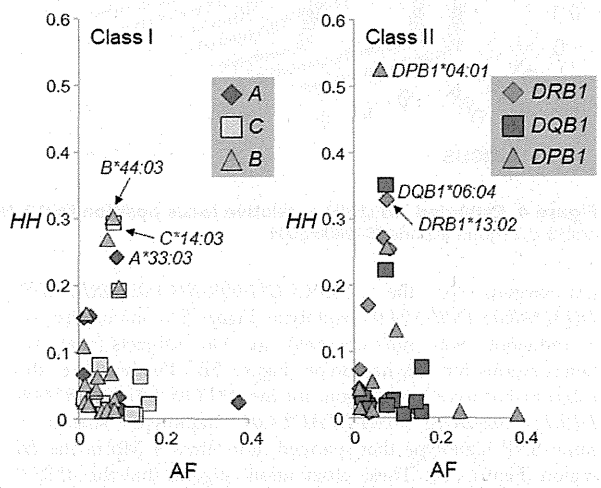

Figure 3. Haplotype homozygosity (*HH*) × allele frequency (AF) of each *HLA* allele. The left and right panels show *HH* values of *HLA* class I alleles and *HLA* class II alleles, respectively. The class I alleles were designated as follows: *HLA-A* (red diamond), *HLA-C* (yellow square), and *HLA-B* (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

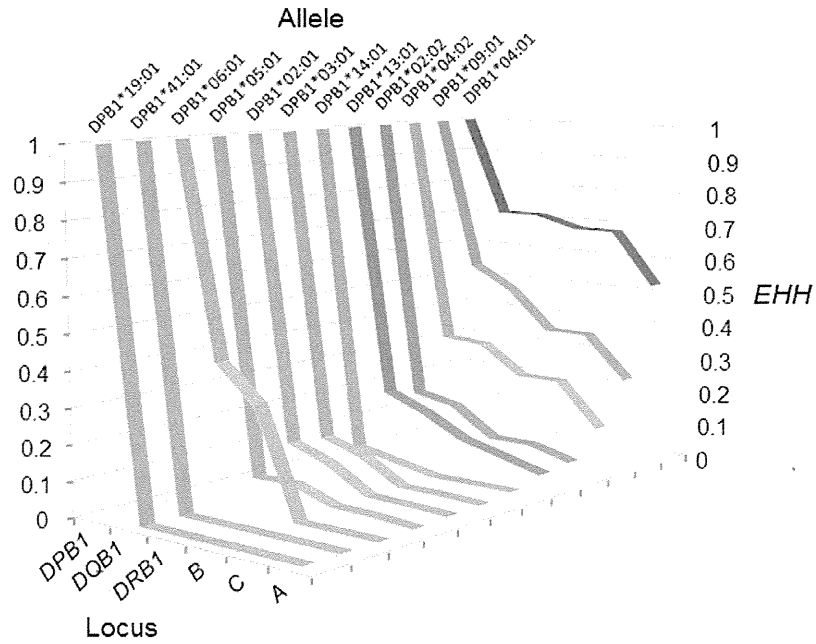

Figure 4. Extended *HH* (*EHH*) × relative locus position for12 *HLA-DPB1* alleles.
doi:10.1371/journal.pone.0046806.g004

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

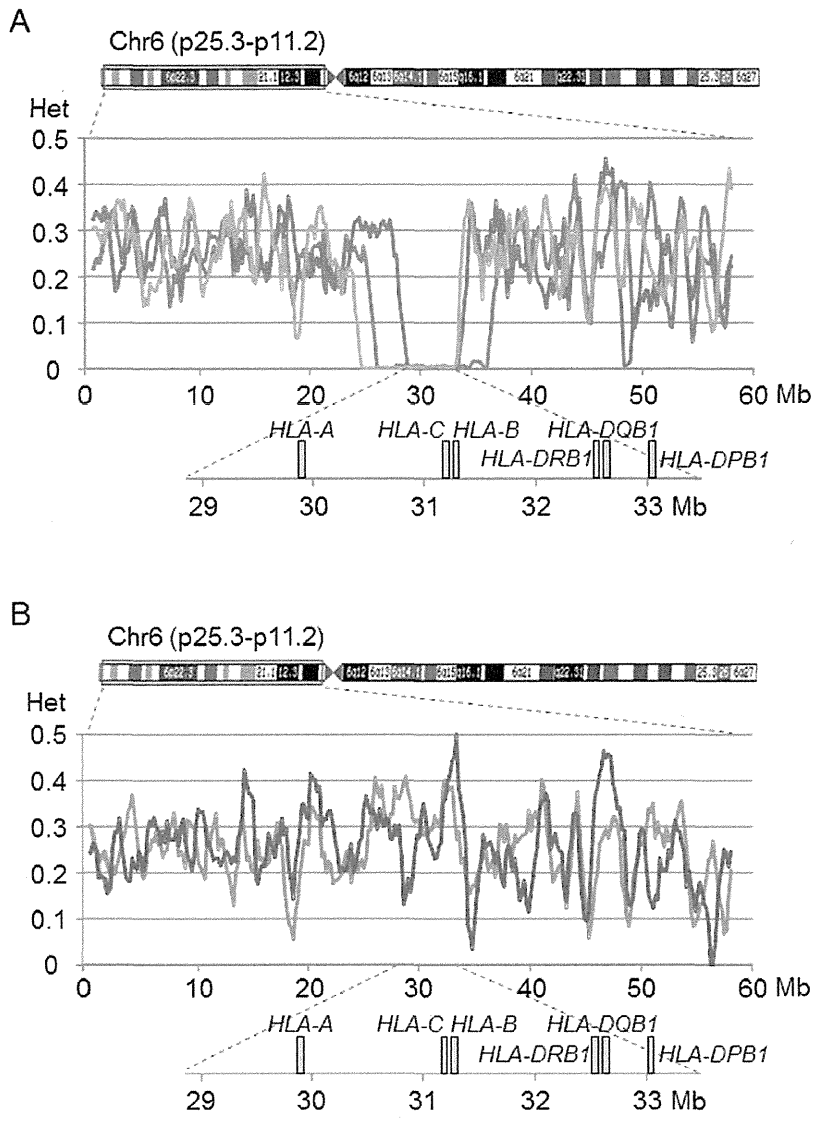

Figure 5. Individual heterozygosity of each subject with the most common 6-locus *HLA* haplotype. The individual heterozygosity in the genomic region on the short arm of chromosome 6 was assessed using the sliding window analysis; in this analysis, the window and step sizes were set to be 1 Mb and 200 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*

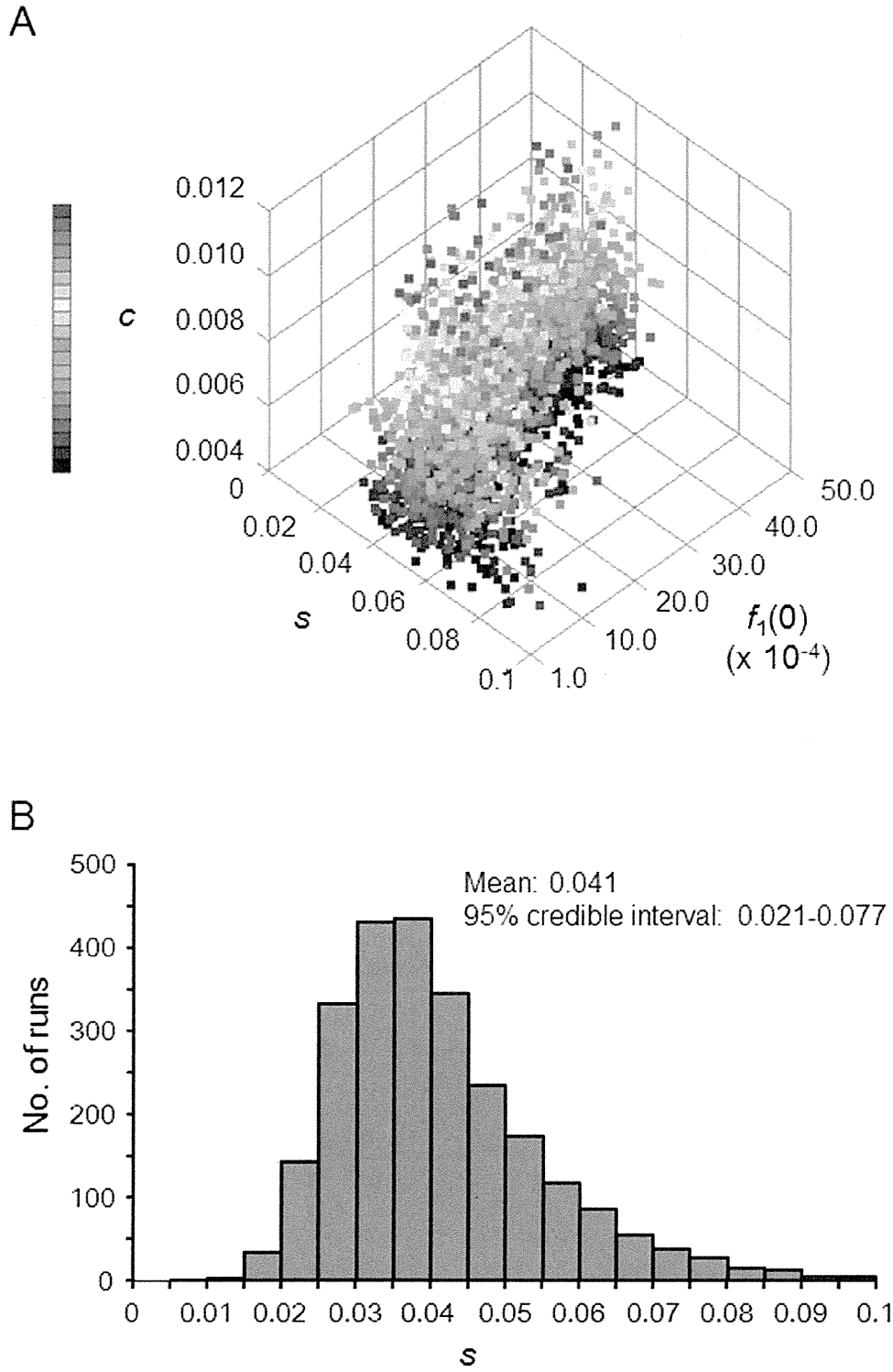

Figure 6. Estimation of model parameters for positive selection acting on *DPB1*04:01*. The recombination rate (*c*), initial haplotype frequency ($f_1(0)$), and selection coefficient (*s*), were estimated by comparing the four haplotype frequencies observed in our study population with the respective 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.