

## Allele and haplotype frequencies of human leukocyte antigen–A, –B, –C, –DRB1, and –DQB1 in Chinese patients with hematological diseases

YanJun Jia, Wei Li, Lijun Wang, Na Liu, Dongmei Wang, Dongmei Li, Yuanyuan Jing,  
Jie Wang, Zhongmei Wang, Xiaoyan Shan\*

Department of HLA, Beijing Red Cross Blood Center, Beijing 100088, China

### ABSTRACT

The human leukocyte antigen (HLA) system plays a central role in the immune response to pathogens, as well as in organ and allogeneic hematopoietic stem cell transplantation (HSCT). Finding a five–locus (*i.e.*, HLA–A, –B, –C, –DRB1, and –DQB1) matched unrelated donor for a patient awaiting HSCT is a major clinical challenge, due to the lack of HLA–identical sibling donors and the high polymorphism of HLA. To date, most studies providing HLA allele frequencies (AF) and haplotype frequencies (HF) in Chinese populations have focused on donors instead of the recipients and have provided data for three loci (HLA–A, –B, and –DR); however, data from five–locus HLA typing in a large sample of patients, especially those with hematological diseases, remains unavailable. Therefore, this study was designed to determine HLA AF and two–, three–, four– and five–locus HF in a large cohort of Chinese Han patients with hematological diseases. The AF and the HF were determined using high–resolution HLA typing data from 2,878 patients. The total number of HLA–A, –B, –C, –DRB1, and –DQB1 alleles was determined to be 48, 92, 49, 52, and 24, respectively. Hardy–Weinberg equilibrium (HWE) analyses indicated significant deviations from HWE for HLA–A, –C, –DRB1, and –DQB1 AF, but not for HLA–B locus. The three most common alleles at each locus were A\*11:01, A\*24:02, A\*02:01; B\*46:01, B\*40:01, B\*13:02; C\*01:02, C\*07:02, C\*06:02; DRB1\*09:01, DRB1\*15:01, DRB1\*07:01; DQB1\*03:01, DQB1\*03:03, and DQB1\*06:01. Our data may help to determine whether the current bone marrow registry contains sufficient diversity to meet the demand.

**Keywords:** human leukocyte antigen (HLA), allele frequencies, haplotype frequencies, hematopoietic stem cell transplantation

### INTRODUCTION

The human leukocyte antigen (HLA) system, located at the short arm of chromosome 6, is the most complex and highly polymorphic genetic system in humans. It plays a central role in the immune response to pathogens, as well as in organ and hematopoietic stem cell transplantation (HSCT)<sup>[1]</sup>. Previous studies have demonstrated the importance of HLA matching between patients and donors for the outcome of allogeneic HSCTs<sup>[2–5]</sup>. Finding a five–locus (*i.e.*, HLA–A,

–B, –C, –DRB1, and –DQB1) matched unrelated donor for a patient is difficult, due to the lack of HLA–identical sibling donors and the high polymorphism of HLA. Therefore, many countries have established large stem cell donor registries such as the National Marrow Donor Program (NMDP) in the USA, with over seven million potential HSC donors, and the Chinese Marrow Donor Program (CMDP), with over two million registered volunteers by 2014. Information about HLA allele frequencies (AF) and haplotype frequencies (HF) in defined ethnic groups may help physicians predict the probability of finding HLA–identical HSC donors for patients<sup>[6–8]</sup>. This information is very useful for the estimation of the number of donors required for a donor registry, in order to meet

\*Correspondence to: Xiaoyan Shan, Department of HLA, Beijing Red Cross Blood Center, No. 37 Beisanhuan Zhong Road, Beijing 100088, China. TEL: +86–10–82807279; E–mail: jiajy321@hotmail.com.

the criterion of providing a fully matched cell donor to a certain percentage of patients<sup>[9]</sup>.

To date, several reports have provided HLA AF and HF data for Chinese populations. For example, Hei *et al.*<sup>[10]</sup> analyzed a set of high-resolution HLA typing data obtained from 718 donors found in CMDP for donor-confirmatory typing, which provided the AF and HF of HLA-A, -B, -C, -DRB1, and -DQB1 in this donor population. In 2014, Liaoning Blood Center, China, reported HLA-A, -B, and -DRB1 AF and HF in the group of 21,918 Han and other ethnic residents living in Liaoning, China<sup>[11]</sup>. Based on the family segregation method, Pan *et al.*<sup>[12]</sup> studied the distribution of HLA-A, -B, and -DRB1 alleles and haplotypes obtained by high-resolution genotyping in 167 families in Jiangsu, China. However, most of these studies focused on the population of donors instead of the recipients, and provided the AF and HF data for three loci, HLA-A, -B, and -DR, except in the study performed by Hei *et al.*<sup>[10]</sup>. Data from five-locus HLA typing in a large sample of patients, especially those with hematological diseases, remains unavailable.

Our laboratory is one of the five CMDP-contracted laboratories that has performed HLA typing since its establishment in 2001, and has carried out high-resolution HLA five-locus typing of over 12,000 patients and donors by a sequence-based typing method. In this study, 2,878 patients with hematological diseases were selected and AF and HF of HLA-A, -B, -C, -DRB1, and -DQB1 were analyzed with the aim of predicting the most common alleles, and two-, three-, four- and five-locus haplotypes in this population of Chinese patients.

## MATERIALS AND METHODS

### Sample population

There were 2,878 Chinese Han patients with different hematological diseases, including acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic myeloid leukemia (CML), severe aplastic anemia (SAA), and myelodysplastic syndrome (MDS), enrolled in this study. The population in the study was composed of patients from the whole country and was mixed in terms of geographic distribution. The study population consisted of 1,100 females and 1,778 males, with a median age of 24 (1-65) years. High resolution HLA typing was performed for all patients, at both HLA class I (HLA-A, -B and -C) and HLA class II (HLA-DRB1 and -DQB1) loci between July 2006 and July 2015, in order to choose suitable related or unrelated HSCT donors. The study protocol and all experimental procedures were ap-

proved by the Ethics Committee of Beijing Red Cross Blood Center, written informed consent was provided by all patients for HLA typing experiments, and all patients agreed with the use of HLA typing data in further research.

### DNA extraction and HLA typing

Genomic DNA was extracted from peripheral whole blood samples using QIAGEN QIAamp DNA Blood Mini Kit (Qiagen Inc., Valencia, CA, USA). For samples obtained before 2012, the Hamilton Microlab Starlet Automatic DNA Extracting System (Hamilton Robotics, Switzerland) was used, while for samples post 2012, the Chemagic DNA Blood Kit (PerkinElmer Genetics, USA) was used. For all samples, HLA typing at HLA-A, -B, -C, -DRB1, and -DQB1 loci was conducted using sequence-specific oligonucleotide probe (SSO) and sequencing-based typing (SBT) methods, in order to reduce the probability of method-related or human errors. Genomic DNA amplification was performed using polymerase chain reaction (PCR) on a GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA). A Life Match Fluoroanalyzer (LuminexMAP Technology, TX, USA) was used for SSO typing with LAB-Type kits (One Lambda, Canoga Park, CA, USA), while ABI 3730 or 3730XL DNA Analyzers (Applied Biosystems, Foster City, CA, USA) were used for SBT typing, together with AlleleSEQR HLA typing kits (Celera Corporation, Alameda, CA, USA) following the manufacturer's instructions.

### Statistical analysis

AF of HLA-A, -B, -C, -DRB1, and -DQB1 were directly counted using Excel software (Microsoft Inc., Redmond, WA, USA). The exact test developed by Guo and Thompson's (1992) was used to estimate the deviations from Hardy-Weinberg equilibrium (HWE) proportions for each locus. HF, including two-, three-, four-, and five-locus frequencies, were evaluated by the expectation maximization algorithm using Arlequin version 3.11 software, allowing the maximum-likelihood estimates of gene frequencies<sup>[13]</sup>.

## RESULTS

### Frequency of five HLA alleles

The number of different HLA alleles detected is presented in **Table 1**. Because of the influence of rare alleles on the total number of alleles observed, we also separately calculated the number of common and well-documented (CWD) alleles and rare alleles, according to the CWD list of CMDP (2014).

The common alleles in this list are defined as alleles with a frequency >0.001, and well-documented alleles are those that appeared at least five times in the CMDP registry. Rare alleles are those not included on the CWD list. In the study population, the total number of HLA-A, -B, -C, -DRB1, and

-DQB1 alleles detected was 48, 92, 49, 52, and 24, respectively. Forty-eight HLA-A alleles consisted of 43 CWD (comprising 91.67% of the total number of alleles, and 99.93% of samples) and five rare alleles (accounting for 8.23% of the total number of alleles, and 0.07% of samples).

**Table 1** The frequency of CWD and rare HLA alleles

Locus	Total number of alleles	CWD alleles			Rare alleles		
		Allele count	% alleles	% samples	Allele count	% alleles	% samples
HLA-A	48	43	89.58	99.8263	5	8.23	0.1737
HLA-B	92	89	96.74	99.8958	3	4.30	0.1042
HLA-C	42	42	100	100	0	0	0
HLA-DRB1	59	55	93.22	99.8610	4	6.78	0.1390
HLA-DQB1	24	23	95.83	99.9653	1	4.17	0.0347

### HWE analysis of HLA-A, -B, -C, -DRB1, and -DQB1 loci

The frequencies of the observed and expected homozygous genotypes, as analyzed by HWE tests, are presented in **Table 2**. A *P*-value was used to determine the magnitude of deviation from HWE in the population sample. As shown in **Table 2**, *P*-values obtained in HWE analysis of HLA-A, -B, -C, -DRB1, and -DQB1 loci were 0, 0.09133, 0.03894, 0.04593, and 0.02633, respectively, which suggests significant deviations from HWE at the HLA-A, -C, -DRB1, and -DQB1 loci, but not at HLA-B locus.

**Table 2** Hardy -Weinberg equilibrium analysis for HLA-A, -C, -B, -DRB1, and -DQB1 loci

Locus	Num. of Genotypes	Observed heterozygosity	Expected heterozygosity	<i>P</i> -value
HLA-A	2,878	0.86449	0.89092	0
HLA-B	2,878	0.92842	0.95553	0.03894
HLA-C	2,878	0.88499	0.91076	0.09133
HLA-DRB1	2,878	0.90584	0.92404	0.04593
HLA-DQB1	2,878	0.86623	0.88625	0.02633

### Allele frequency distributions at HLA-A, -B, -C, -DRB1, and -DQB1 loci

The AF of 48 HLA-A, 92 HLA-B, 49 HLA-C, 52 HLA-DRB1, and 24 HLA-DQB1 in our study population are listed in **Table 3**. Among the 48 HLA-A alleles, the 5 most common ones were A\*11:01 (AF=19.15%), A\*24:02 (AF=16.12%), A\*02:01 (AF=14.85%), A\*02:07 (AF=8.83%), and A\*30:01 (AF=6.34%). Seven alleles (A\*11:01, A\*24:02, A\*02:01, A\*02:07, A\*30:01, A\*33:03, and A\*02:06) had frequencies above 5%, representing 77.59% of the HLA-A diversity in the study population. In addition to 44 CWD alleles, 5 rare alleles were found at HLA-

A locus, including A\*02:51, A\*02:85, A\*30:11, A\*31:02, and A\*31:03. The single occurrences of these alleles were found and listed in the IMGT/HLA database as well. In this database, all of these alleles, except A\*02:51, were defined as CWD. The ethnicity of the individual with the A\*30:11 allele was unknown, while the remaining alleles were identified in individuals with Hispanic or Caucasoid backgrounds. The A2 serological group contained 10 alleles and represented the most diverse allele family. The second most diverse allele family was the A24 serological group, which included eight alleles. Groups consisting of two to four alleles included A03, A11, A26, A29, A30, A31, A33, A68, and A74. Other groups, such as A01, A23, A32, A34, A66, and A69 only comprised of one allele each.

Ninety-two HLA-B alleles were identified at the HLA-B locus, demonstrating its extreme diversity. The most common HLA-B alleles were B\*46:01 (AF=9.85%), B\*40:01 (AF=9.21%), and B\*13:02 (AF=6.83%). Other common alleles occurring at frequencies of 4%~6% included B\*51:01 (AF=5.72%), B\*15:01 (AF=4.99%), B\*13:01 (AF=4.95%), and B\*58:01 (AF=4.81%). Alleles with a frequency between two and three percent included B\*40:06 (AF=3.93%), B\*52:01 (AF=3.25%), B\*35:01 (AF=3.21%), B\*54:01 (AF=3.15%), B\*15:02 (AF=2.95%), B\*48:01 (AF=2.81%), B\*07:02 (AF=2.80%), B\*38:02 (AF=2.45%), B\*44:03 (AF=2.43%), B\*15:11 (AF=2.28%), and B\*55:02 (AF=2.03%). These alleles accounted for 80.90% of the HLA-B alleles identified in this population. Three rare alleles (B\*07:68, B\*35:31, and B\*40:15) were observed at this locus, and were identified only once for each allele. For B\*07:68 and B\*58:01, the ethnic origin of patients was unknown, and CWD status was not defined in the IMGT/HLA database. However, B\*35:31 was defined as a CWD allele,

**Table 3 AF of HLA-A, -B, -C, -DRB1, and -DQB1 in 2,878 patients with hematological diseases**

Allele	AF (%)	Allele	AF (%)	Allele	AF (%)	Allele	AF (%)	Allele	AF (%)
A*01:01	3.23	B*13:01	4.95	B*41:01	0.12	C*07:63	0.02	DRB1*12:10	0.05
A*02:01	14.85	B*13:02	6.83	B*41:02	0.04	C*07:66	0.02	DRB1*13:01	1.41
A*02:03	2.69	B*14:01	0.02	B*44:02	1.30	C*08:01	9.43	DRB1*13:02	2.75
A*02:05	0.49	B*14:02	0.31	B*44:03	2.43	C*08:02	0.33	DRB1*13:03	0.04
A*02:06	6.03	B*15:01	4.99	B*45:01	0.17	C*08:03	0.71	DRB1*13:05	0.02
A*02:07	8.83	B*15:02	2.95	B*46:01	9.85	C*08:06	0.04	DRB1*13:07	0.05
A*02:10	0.56	B*15:03	0.02	B*46:06	0.02	C*08:22	0.73	DRB1*13:12	0.68
A*02:17	0.02	B*15:05	0.23	B*47:01	0.02	C*12:02	3.23	DRB1*13:56	0.02
A*02:51	0.02	B*15:07	0.38	B*48:01	2.81	C*12:03	1.86	DRB1*14:01	0.52
A*02:53	0.10	B*15:10	0.02	B*48:03	0.07	C*14:02	4.24	DRB1*14:02	0.02
A*02:85	0.02	B*15:11	2.28	B*49:01	0.24	C*14:03	0.87	DRB1*14:03	0.56
A*03:01	3.46	B*15:12	0.12	B*50:01	0.73	C*15:02	3.72	DRB1*14:04	0.66
A*03:02	0.26	B*15:13	0.12	B*51:01	5.72	C*15:04	0.05	DRB1*14:05	2.36
A*03:77	0.02	B*15:17	0.21	B*51:02	1.16	C*15:05	0.63	DRB1*14:06	0.02
A*11:01	19.15	B*15:18	1.49	B*51:08	0.02	C*15:11	0.02	DRB1*14:07	0.17
A*11:02	1.44	B*15:21	0.07	B*52:01	3.25	C*15:12	0.04	DRB1*14:10	0.04
A*23:01	0.26	B*15:25	0.30	B*53:01	0.04	C*15:21	0.02	DRB1*14:11	0.04
A*24:02	16.12	B*15:27	0.78	B*54:01	3.15	C*16:01	0.02	DRB1*14:12	0.02
A*24:03	0.10	B*15:29	0.05	B*55:01	0.14	C*16:02	0.12	DRB1*14:14	0.02
A*24:04	0.02	B*15:32	0.19	B*55:02	2.03	C*16:04	0.09	DRB1*14:18	0.04
A*24:07	0.10	B*15:35	0.05	B*55:04	0.05	C*17:01	0.12	DRB1*14:25	0.02
A*24:08	0.05	B*15:46	0.14	B*55:07	0.02	C*17:03	0.02	DRB1*14:49	0.02
A*24:10	0.05	B*15:58	0.10	B*55:12	0.05			DRB1*14:54	2.68
A*24:13	0.02	B*18:01	0.23	B*56:01	0.37	DRB1*01:01	1.93	DRB1*15:01	13.13
A*24:20	0.19	B*18:02	0.07	B*56:03	0.04	DRB1*01:02	0.35	DRB1*15:02	3.18
A*26:01	2.76	B*27:02	0.05	B*56:04	0.07	DRB1*03:01	3.84	DRB1*15:04	0.19
A*26:20	0.02	B*27:03	0.02	B*56:10	0.02	DRB1*03:04	0.02	DRB1*16:01	0.09
A*29:01	0.68	B*27:04	0.61	B*57:01	1.13	DRB1*03:27	0.02	DRB1*16:02	2.17
A*29:02	0.07	B*27:05	0.52	B*58:01	4.81	DRB1*04:01	0.97		
A*30:01	6.34	B*27:06	0.04	B*59:01	0.12	DRB1*04:02	0.10	DQB1*02:01	3.84
A*30:02	0.02	B*27:07	0.10	B*67:01	0.75	DRB1*04:03	1.30	DQB1*02:02	8.36
A*30:04	0.05	B*27:24	0.04	B*73:01	0.02	DRB1*04:04	0.57	DQB1*03:01	20.78
A*30:11	0.02	B*27:25	0.02	B*81:02	0.14	DRB1*04:05	4.76	DQB1*03:02	5.23
A*31:01	3.02	B*27:36	0.04			DRB1*04:06	2.50	DQB1*03:03	17.06
A*31:02	0.02	B*35:01	3.21	C*01:02	15.12	DRB1*04:07	0.19	DQB1*03:04	0.02
A*31:03	0.02	B*35:02	0.28	C*01:03	0.76	DRB1*04:08	0.12	DQB1*03:05	0.04
A*32:01	1.37	B*35:03	1.04	C*01:06	0.10	DRB1*04:10	0.37	DQB1*03:13	0.05
A*33:01	0.23	B*35:05	0.16	C*02:02	0.56	DRB1*07:01	10.01	DQB1*03:17	0.02
A*33:03	6.27	B*35:08	0.10	C*03:02	4.86	DRB1*08:01	0.04	DQB1*03:38	0.02
A*34:01	0.07	B*35:31	0.02	C*03:03	7.82	DRB1*08:02	0.63	DQB1*04:01	4.64
A*66:01	0.07	B*37:01	1.11	C*03:04	9.42	DRB1*08:03	6.10	DQB1*04:02	1.18
A*68:01	0.70	B*38:01	0.43	C*03:56	0.02	DRB1*08:04	0.02	DQB1*04:05	0.02
A*68:02	0.02	B*38:02	2.45	C*04:01	5.68	DRB1*08:09	0.14	DQB1*05:01	4.08
A*68:24	0.02	B*39:01	1.88	C*04:03	0.75	DRB1*09:01	15.86	DQB1*05:02	5.75
A*69:01	0.05	B*39:05	0.12	C*04:69	0.02	DRB1*09:04	0.04	DQB1*05:03	4.33
A*74:01	0.02	B*39:09	0.02	C*04:81	0.02	DRB1*09:10	0.02	DQB1*05:04	0.04
A*74:02	0.05	B*39:24	0.04	C*05:01	1.10	DRB1*10:01	1.18	DQB1*06:01	10.32
A*74:13	0.04	B*40:01	9.21	C*06:02	9.83	DRB1*11:01	5.92	DQB1*06:02	10.04
		B*40:02	1.93	C*07:01	0.71	DRB1*11:02	0.05	DQB1*06:03	1.41
B*07:02	2.80	B*40:03	0.26	C*07:02	14.99	DRB1*11:04	0.61	DQB1*06:04	1.23
B*07:05	0.61	B*40:06	3.93	C*07:04	1.16	DRB1*11:06	0.12	DQB1*06:07	0.02
B*07:06	0.09	B*40:15	0.02	C*07:06	0.70	DRB1*12:01	3.35	DQB1*06:09	1.44
B*07:68	0.02	B*40:40	0.16	C*07:18	0.02	DRB1*12:02	7.87	DQB1*06:10	0.04
B*08:01	0.82	B*40:78	0.04	C*07:51	0.02	DRB1*12:06	0.02		

and its ethnic origins were shown to be Caucasoid or Oriental. The B\*15 group was found to be the most diverse allele family. The B27, B35, B40, and B55 serological groups also exhibited high diversity and included nine, six, seven, and five different alleles, respectively.

HLA-C locus comprised a total of 42 alleles, all of them were CWD alleles. Seven HLA-C alleles had a frequency higher than 5%, accounting for 72.29% of the total CWD alleles. The C07 serological group showed the most complex diversity, and was shown to be comprised of eight alleles. The C15 group was found also to be highly diverse and included six alleles.

At the HLA-DRB1 locus, 59 alleles were identified, including 55 CWD and 4 rare alleles (DRB1\*03:04, DRB1\*12:06, DRB1\*13:56, and DRB1\*14:14). These rare alleles were found one time each in the studied population. In the IMGT/HLA database, DRB1\*03:04 and DRB1\*14:14 were defined as well-documented alleles, but the CWD v2.0 status of the DRB1\*12:06 and DRB1\*13:56 alleles was unknown. DRB1\*03:04 was found in an individual with Caucasoid background, DRB1\*12:06 in an Oriental individual, DRB1\*13:56 in an individual with a Caucasoid or Hispanic background, and DRB1\*14:14 in an individual with an Australian aboriginal background. The most common DRB1 alleles were DRB1\*09:01 (AF=15.86%), DRB1\*15:01 (AF=13.13%), and DRB1\*07:01 (AF=10.01%). The six most common alleles (DRB1\*09:01, DRB1\*15:01, DRB1\*07:01, DRB1\*12:02, DRB1\*08:03, and DRB1\*11:01), were all found with a frequency above 5%, and accounted for 58.89% of all DR alleles. DR14 was demonstrated

to be the most diverse serological group, comprising of 15 alleles. DR04, DR13, and DR08 groups included 9, 7, and 5 alleles, respectively.

The HLA-DQ locus showed the lowest diversity, and 24 HLA-DQB1 alleles were detected, of which 4 alleles had a frequency higher than 10%. DQB1\*03:01, DQB1\*03:03, DQB1\*06:01, DQB1\*06:02, DQB1\*02:02, DQB1\*03:02, and DQB1\*05:02, with frequencies higher than 5%, which comprised of 77.54% of the total DQB1 alleles. The most diverse serological group was DQB1\*03, which included eight alleles. Seven alleles were detected in the DQ06 group and four alleles were detected in the DQ05 group. These twenty-four HLA-DQB1 alleles included twenty-three CWD and only one rare allele, DQB1\*04:05, which was observed in a single individual in this sample population. We were unable to find any additional information about this allele in the IMGT/HLA database.

### Two-, three-, four- and five-locus HF of HLA Two-locus HLA haplotypes

HLA HF were estimated by the expectation-maximization (EM) method. Common two-locus HLA haplotypes among the 2,878 patients with hematological diseases are listed in **Table 4**. A total of 12 A-C, 7 A-B, 15 C-B, 4 B-DRB1, and 14 DRB1-DQB1 haplotypes with frequencies higher than 2% were observed in the population. The common A-C haplotypes (with frequencies of >2%), accounted for 40.36% of all A-C haplotypes, common A-B haplotypes (HF>2%) comprised 25.67% of all A-B haplotypes, common C-B haplotypes (HF>2%) contributed to 57.00% of all C-B haplotypes, common B-DRB1

**Table 4 Common (HF>2%) HLA two-locus haplotypes in 2,878 patients with hematological diseases**

Haplotypes	HF (%)	Haplotypes	HF (%)	Haplotypes	HF (%)
A*02:07-C*01:02	6.13	C*01:02-B*46:01	8.78	B*13:01-DRB1*12:02	2.30
A*30:01-C*06:02	5.60	C*06:02-B*13:02	6.64		
A*11:01-C*07:02	4.32	C*03:02-B*58:01	4.73	DRB1*09:01-DQB1*03:03	15.26
A*33:03-C*03:02	3.75	C*03:04-B*13:01	4.40	DRB1*15:01-DQB1*06:02	9.75
A*11:01-C*08:01	3.02	C*07:02-B*40:01	4.35	DRB1*07:01-DQB1*02:02	8.30
A*11:01-C*03:04	2.95	C*14:02-B*51:01	4.05	DRB1*12:02-DQB1*03:01	7.51
A*02:01-C*03:03	2.72	C*08:01-B*40:06	3.13	DRB1*08:03-DQB1*06:01	5.85
A*24:02-C*07:02	2.65	C*01:02-B*54:01	2.99	DRB1*11:01-DQB1*03:01	5.80
A*11:01-C*01:02	2.51	C*08:01-B*15:02	2.90	DRB1*04:05-DQB1*04:01	4.48
A*24:02-C*03:04	2.48	C*03:04-B*40:01	2.86	DRB1*03:01-DQB1*02:01	3.79
A*24:02-C*01:02	2.22	C*07:02-B*07:02	2.76	DRB1*12:01-DQB1*03:01	3.18
A*02:01-C*03:04	2.01	C*12:02-B*52:01	2.59	DRB1*04:06-DQB1*03:02	2.47
A*02:07-B*46:01	6.43	C*07:02-B*38:02	2.38	DRB1*15:02-DQB1*06:01	2.21
A*30:01-B*13:02	5.47	C*03:03-B*15:11	2.24	DRB1*16:02-DQB1*05:02	2.17
A*33:03-B*58:01	3.74	C*03:03-B*35:01	2.20	DRB1*15:01-DQB1*06:01	2.15
A*11:01-B*40:01	3.12			DRB1*14:05-DQB1*05:03	2.12
A*24:02-B*40:01	2.63	B*13:02-DRB1*07:01	5.65		
A*11:01-B*15:02	2.17	B*46:01-DRB1*09:01	5.07		
A*11:01-B*13:01	2.11	B*58:01-DRB1*03:01	2.36		

haplotypes (HF>2%) comprised 15.36% of the total B-DRB1 haplotypes, and DRB1-DQB1 haplotypes (HF>2%) accounted for 75.04% of all DRB1-DQB1 haplotypes. Fourteen predominant haplotypes, with frequencies higher than 5% at each locus were found in the study population.

### Three-locus HLA haplotypes

A total of 956 different HLA-A-C-B haplotypes and 1,777 different HLA-A-B-DRB1 haplotypes with frequencies higher than 0.01% were identified in the 2,878 samples estimated by the EM method. **Table 5** lists all three-locus haplotypes with frequencies higher than 1%, which comprised 27.45% of all A-C-B haplotypes, and 13.43% of all A-B-DRB1 haplotypes. The most frequent haplotypes in each set were A\*02:07-C\*01:02-B\*46:01 (HF=5.92%), A\*30:01-C\*06:02-B\*13:02 (HF=5.43%), A\*30:01-B\*13:02-DRB1\*07:01 (HF=4.62%), and A\*02:07-B\*46:01-DRB1\*09:01 (HF=3.52%).

### Four-locus HLA haplotypes

A total of 2,169 A-C-B-DRB1 and 1,339 C-B-DRB1-DQB1 haplotypes (HF>0.01%, determined by EM) were observed in the study population. Four-locus haplotypes with frequencies >1% are presented in **Table 6**. Eleven A-C-B-DRB1 haplotypes (HF>1%) comprised 18.7% of all A-C-B-DRB1 haplotypes, and 13 C-B-DRB1-DQB1 haplotypes (HF>1%) made up 26.74% of all C-B-DRB1-DQB1 haplotypes. The most frequent A-C-B-DRB1 haplotype was A\*30:01-C\*06:02-B\*13:02-DRB1\*07:01, with a frequency of 4.58%, which was clearly the predominant haplotype in this hematological disease patient population. The second most frequent haplotype was A\*02:07-C\*01:02-B\*46:01-DRB1\*09:01, with a frequency of 3.26%. The two most common C-B-DRB1-DQB1 haplotypes were C\*06:02-B\*13:02-DRB1\*07:01-DQB1\*02:02 (HF=5.41%) and C\*01:02-B\*46:01-DRB1\*09:01-DQB1\*03:03 (HF=4.42%).

**Table 5 Common (HF>1%) HLA three-locus haplotypes in 2,878 patients with hematological diseases**

Haplotypes	HF (%)	Haplotypes	HF (%)
A*02:07-C*01:02-B*46:01	5.92	A*02:03-C*07:02-B*38:02	1.17
A*30:01-C*06:02-B*13:02	5.43	A*24:02-C*07:02-B*40:01	1.13
A*33:03-C*03:02-B*58:01	3.74		
A*11:01-C*08:01-B*15:02	2.07	A*30:01-B*13:02-DRB1*07:01	4.62
A*11:01-C*07:02-B*40:01	2.06	A*02:07-B*46:01-DRB1*09:01	3.52
A*11:01-C*03:04-B*13:01	1.91	A*33:03-B*58:01-DRB1*03:01	1.89
A*02:01-C*03:03-B*15:11	1.51	A*11:01-B*15:02-DRB1*12:02	1.22
A*02:01-C*03:04-B*13:01	1.27	A*02:07-B*46:01-DRB1*08:03	1.16
A*11:01-C*14:02-B*51:01	1.24	A*02:01-B*13:01-DRB1*12:02	1.02

**Table 6 Common (HF>1%) HLA four-locus haplotypes in 2,878 patients with hematological diseases**

Haplotypes	HF (%)	Haplotypes	HF (%)
A*30:01-C*06:02-B*13:02-DRB1*07:01	4.58	C*03:04-B*13:01-DRB1*12:02-DQB1*03:01	2.09
A*02:07-C*01:02-B*46:01-DRB1*09:01	3.26	C*01:02-B*46:01-DRB1*08:03-DQB1*06:01	1.84
A*33:03-C*03:02-B*58:01-DRB1*03:01	1.89	C*12:02-B*52:01-DRB1*15:02-DQB1*06:01	1.61
A*11:01-C*08:01-B*15:02-DRB1*12:02	1.22	C*08:01-B*15:02-DRB1*12:02-DQB1*03:01	1.52
A*02:07-C*01:02-B*46:01-DRB1*08:03	1.19	C*14:02-B*51:01-DRB1*09:01-DQB1*03:03	1.39
A*02:01-C*03:04-B*13:01-DRB1*12:02	1.04	C*07:02-B*07:02-DRB1*15:01-DQB1*06:02	1.38
		C*01:02-B*54:01-DRB1*04:05-DQB1*04:01	1.26
C*06:02-B*13:02-DRB1*07:01-DQB1*02:02	5.41	C*08:01-B*40:06-DRB1*09:01-DQB1*03:03	1.19
C*01:02-B*46:01-DRB1*09:01-DQB1*03:03	4.42	C*03:02-B*58:01-DRB1*13:02-DQB1*06:09	1.18
C*03:02-B*58:01-DRB1*03:01-DQB1*02:01	2.31	C*07:02-B*40:01-DRB1*09:01-DQB1*03:03	1.14

### Five-locus HLA haplotypes

**Table 7** lists the 20 most common HLA-A-B-C-DRB1-DQB1 haplotypes as determined by EM. These haplotypes accounted for 20.99% of all A-C-B-DRB1-DQB1 haplotypes. The most frequent five-locus haplotype was A\*30:01-C\*06:02-B\*13:02-DRB1\*07:01-DQB1\*02:02, with a frequency of 4.41%.

## DISCUSSION

HSCT was developed as an effective treatment for many malignant and non-malignant hematological disorders. Due to the importance of HLA in the rejection of transplants and graft-versus-host disease (GvHD), finding an HLA-matched HSC donor is crucial for the success of a transplantation. The first

**Table 7 Twenty most common HLA-A-B-C-DRB1-DQB1 five-locus haplotypes in 2,878 patients with hematological diseases**

Five-locus Haplotypes	HF (%)
A*30:01-C*06:02-B*13:02-DRB1*07:01-DQB1*02:02	4.41
A*02:07-C*01:02-B*46:01-DRB1*09:01-DQB1*03:03	3.17
A*33:03-C*03:02-B*58:01-DRB1*03:01-DQB1*02:01	1.85
A*02:07-C*01:02-B*46:01-DRB1*08:03-DQB1*06:01	1.20
A*11:01-C*08:01-B*15:02-DRB1*12:02-DQB1*03:01	1.17
A*02:01-C*03:04-B*13:01-DRB1*12:02-DQB1*03:01	1.04
A*33:03-C*03:02-B*58:01-DRB1*13:02-DQB1*06:09	0.93
A*11:01-C*03:04-B*13:01-DRB1*15:01-DQB1*06:01	0.78
A*02:01-C*03:03-B*15:11-DRB1*09:01-DQB1*03:03	0.68
A*11:01-C*04:01-B*15:01-DRB1*04:06-DQB1*03:02	0.66
A*11:01-C*07:02-B*40:01-DRB1*09:01-DQB1*03:03	0.59
A*11:01-C*01:02-B*46:01-DRB1*09:01-DQB1*03:03	0.58
A*01:01-C*06:02-B*57:01-DRB1*07:01-DQB1*03:03	0.55
A*24:02-C*08:01-B*40:06-DRB1*09:01-DQB1*03:03	0.52
A*11:01-C*08:01-B*15:02-DRB1*15:01-DQB1*06:01	0.51
A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04	0.50
A*33:03-C*07:06-B*44:03-DRB1*07:01-DQB1*02:02	0.48
A*03:01-C*07:02-B*07:02-DRB1*15:01-DQB1*06:02	0.47
A*02:03-C*07:02-B*38:02-DRB1*16:02-DQB1*05:02	0.45
A*24:02-C*14:02-B*51:01-DRB1*09:01-DQB1*03:03	0.45

choice is a donor among the patient's siblings, who have a 1/4 probability of sharing the same HLA phenotype. However, it was reported that only 30% of patients in need of HSCT are able to find an HLA-identical related donor<sup>[14]</sup>. Such data are not available for Chinese patients. Nevertheless, using the data presented above, we can infer that the probability of finding an HLA-identical sibling would be even lower in Chinese individuals, since many do not have siblings due to the Chinese one-child policy, which has been in place for over 30 years. Another option is selecting an HLA matched unrelated donor (MUD) from an HSC donor registry, such as NMDP and CMDP. This selection is based on high-resolution typing for HLA-A, -B, -C, -DRB1, and -DQB1 alleles (10/10 matching) in most transplant centers. Previous studies have demonstrated that the degree of matching of these HLA alleles have great impact on the incidence of GvHD and patient mortality<sup>[3, 15-18]</sup>. Disease-free survival for patients who received HSC from fully matched unrelated donors was significantly longer than for patients with HLA allele-mismatched ones<sup>[19]</sup>. In order to improve the success rate of identifying an MUD for patients needing HSCT, knowing the HLA AF and HF in a certain population is very important. Additionally, our data are useful for the evaluation of the optimum size and composition of the donor registry. Statistical data reported earlier on HLA allele and haplotype frequencies in Chinese populations were mostly focused on healthy volunteers, instead of

patients, but there is little data showing information regarding HLA AF and HF in patients. Some research groups have suggested that patients with common HLA alleles and haplotypes have a higher probability of finding an MUD, compared with patients with rare HLA alleles or unusual haplotypes<sup>[20, 21]</sup>.

An additional rationale for conducting this study is the association that exists between HLA alleles, haplotypes and some diseases<sup>[22]</sup>. A relationship between HLA class II genes and multiple sclerosis (MS) has been reported repeatedly in Caucasians and other ethnic groups worldwide<sup>[23]</sup>. For example, HLA-DRB1\*15:01 and DQB1\*06:02 alleles are the main alleles associated with an increased risk of developing MS<sup>[24, 25]</sup>. In the Chinese Han population, the relevance of HLA-DR2/DRB1\*15 for MS was confirmed by a meta-analysis<sup>[26]</sup>. Several studies have also indicated that the HLA gene complex may mediate susceptibility to certain hematological malignancies, such as Hodgkin's disease. In Caucasians, patients with the HLA-DPB1\*03:01 allele are more likely to develop Hodgkin's disease, whereas the HLA-DPB1\*0401 allele exerts protective effects against this disease in Orientals<sup>[27]</sup>. A large study conducted by two centers in United Kingdom suggested that the susceptibility to Hodgkin's disease may correlate with a specific HLA-DPB1 hyper-variable sequence motif, rather than with the DBP1\*03:01 allele perse<sup>[28]</sup>. Furthermore, the association of HLA with other hematological malignancies was demonstrated by several research groups. Also, the frequency of the HLA class II DRB1\*11(DR5) phenotype and HLA-DQB1\*03 allele was shown to be increased in patients with cutaneous T-cell lymphoma<sup>[29]</sup>, and the HLA-DRB4 gene was shown to be associated with a predisposition to childhood acute lymphoblastic leukemia (ALL)<sup>[30]</sup>, together with the HLA-DPB1\*02:01, HLA-DQA1\*01:01/01:04, and HLA-DQB1\*05:01 alleles, which also were shown to represent risk factors for ALL<sup>[31, 32]</sup>. A recent case-control study<sup>[33]</sup>, involving 222 non-Hispanic patients with white follicular lymphoma (FL) and 220 matched controls, indicated that the HLA-DPB1 allele represented a protective factor for resistance to FL, while the haplotypes DRB1\*01:01-DQA1\*01:01-DQB1\*05:01 and DRB1\*15-DQA1\*01-DQB1\*06 were shown to affect the development of FL.

Here, we performed a detailed analysis of five-locus polymorphisms, as well as two-, three-, four-, and five-locus HF in Chinese Han patients with various hematological diseases, by high-resolution DNA typing method, *i.e.*, the SBT method. SBT is widely used for HLA typing, and it represents the current gold-

standard high-resolution method. In order to compare HLA AF and HF in the hematological disease study population, and healthy donor population, we selected data reported by Hei *et al.*<sup>[10]</sup> and the CWD list reported by CMDP in 2014 as our reference sources of information. The former represents the profile of HLA polymorphisms in 718 healthy Chinese marrow donors in the CMDP registry, and the latter represents five-locus HLA genotyping results from over 160,000 CMDP-enrolled HSC donors. **Table 8** lists the frequencies and the rank of the 10 most common alleles at HLA-A, -B, -C, -DRB1, and -DQB1 loci, according to the results of our study and the other two sources. Our results are consistent with these, as demonstrated by the overall similarity in AF between the study populations. However, some differences in the incidence of alleles at each locus remain. For the HLA-A locus, the four most common alleles are the same in all populations. The most pronounced difference is observed for A\*03:01, which was ranked eighth (AF=3.46%) in this study, while it is ranked 11th and 13th in the CWD list and the report by Hei *et al.*<sup>[10]</sup>, respectively. A\*02:03 was ranked eighth in both the CWD list (AF=3.48%) and the study by Hei *et al.* (AF=4.18%), but was not among the 10 most common alleles in this study (AF=2.69%). Discrepancies were observed for the HLA-B locus as well, and, for example, B\*46:01 was shown to be the most frequent allele in this study and the CWD list, while it was the second most common allele found by Hei *et al.*<sup>[10]</sup>. The frequencies of B\*52:01 (rank 9, AF=3.25%) and B\*35:01 (rank 10, AF=3.21%) in our study population were different from those reported by Hei *et al.* (rank 19, with AF=2.02%, and rank 18, with AF=2.16%, respectively). The 10 most common HLA-C alleles were the same in the three studies, but their ranking was different. The biggest difference in allele frequency was found for the C\*03:04 allele. For HLA Class II alleles, the results were more consistent between these three study populations. At the HLA-DRB1 locus, the six most common alleles had the same ranking and similar frequencies. The major difference between the three studies was the frequency of DRB1\*15:02, which was ranked 10th (AF=3.18%) in our study population, but 13th (AF=3.05%) in the CWD list and 14th (AF=2.65%) in the Hei *et al.* study. DQB1\*02:01 was shown to be a common allele, ranked eighth (AF=4.92%) in the CWD list and ninth (AF=3.76%) in the Hei *et al.*<sup>[10]</sup> study. However, in our study this allele had a frequency of 3.16%, and was not among the 10 most common ones.

HLA haplotype analysis is of practical importance for human population genetics, anthropological stud-

ies, and HLA-matching for HSCT. The data on common haplotypes can be used to estimate the probability of finding a fully matched donor or closely matched donors in the HSCT donor registry and to delineate donor search strategies<sup>[34]</sup>. A number of studies have reported HLA allele and haplotype distribution in specific areas of China<sup>[35-40]</sup>. Unfortunately, most of them cannot be used as a reference, due to the low-resolution or small sample sizes, resulting in low levels of AF and HF accuracy. Even though the study conducted by Hei *et al.*<sup>[10]</sup> enrolled only 718 patients, it provided high-resolution AF and HF data at five HLA loci, and represented the HLA polymorphism profile at a national level, instead of merely within a certain geographical area or population, which led us to select this study as a comparison for our data. Our results regarding common two-, three-, four-, and five-locus haplotype frequencies are in agreement with this report. For HLA two-locus haplotypes, the five most common A-B haplotypes had similar frequencies and they were ranked the same in the two studies. The five most frequent C-B haplotypes were the same, but their ranking was different. Hei *et al.* provided the frequencies of only 2 B-DRB1 haplotypes, which are consistent with the results obtained in our study in terms of HF and ranking. Top 5 of DRB1-DQB1 haplotypes in these two populations were also similar except for DRB1\*15:01-DQB1\*06:02 which was ranked 2nd in our study but ranked 4th in Hei *et al.* The three-locus A-C-B haplotypes with a frequency higher than 1% were almost the same in the two studies, and the three most common haplotypes were ranked in the same order. Furthermore, the composition and ranking of A-B-DR haplotypes and of four-locus haplotypes A-C-B-DRB1 or C-B-DRB1-DQB1 (with a frequency over 1% in these two studies) were the same.

To the best of our knowledge, here, for the first time we have presented frequency data obtained from a relatively large sample of Chinese Han patients with hematological diseases, and our results may help physicians find a fully matched HSC donor more easily and quickly. Additionally, our research demonstrated that no significant differences exist in AF and haplotype diversity between healthy donors and patients with hematological diseases in China, although several previous studies have suggested relationships between HLA alleles and such diseases. We think that this may be because our study population consisted of patients with different hematological diseases, such as ALL, AML, CML, MDS, and SAA. Subsequently we will focus further research on the determina-

**Table 8** Comparison of frequency and ranking of the most common HLA alleles in this study, WD list, and the study conducted by Hei *et al.*

Allele	This study		CWD		Hei <i>et al.</i>	
	Rank	Freq	Rank	Freq	Rank	Freq
A*11:01	1	19.15	1	21.08	1	21.66
A*24:02	2	16.12	2	15.59	2	16.43
A*02:01	3	14.85	3	12.26	3	13.79
A*02:07	4	8.83	4	8.55	4	9.75
A*30:01	5	6.34	6	5.60	7	6.20
A*33:03	6	6.27	5	7.94	5	7.24
A*02:06	7	6.03	7	5.17	6	7.03
A*03:01	8	3.46	11	2.97	13	1.74
A*01:01	9	3.23	9	3.40	10	2.16
A*31:01	10	3.02	10	3.27	9	2.58
B*46:01	1	9.85	1	10.27	2	9.54
B*40:01	2	9.21	2	9.88	1	11.49
B*13:02	3	6.83	3	5.94	3	7.10
B*51:01	4	5.72	5	5.66	7	5.15
B*15:01	5	4.99	7	4.87	4	5.85
B*13:01	6	4.95	6	4.98	5	5.57
B*58:01	7	4.81	4	5.88	6	5.36
B*40:06	8	3.93	9	3.21	8	3.76
B*52:01	9	3.25	12	2.85	19	2.02
B*35:01	10	3.21	11	2.87	18	2.16
C*01:02	1	15.03	1	15.63	2	15.32
C*07:02	2	14.99	2	15.08	1	15.74
C*06:02	3	9.83	4	9.11	5	8.98
C*08:01	4	9.43	5	8.49	4	9.61
C*03:04	5	9.42	3	9.87	3	11.56
C*03:03	6	7.82	6	7.07	6	7.66
C*04:01	7	5.66	8	5.83	7	6.48
C*03:02	8	4.80	7	5.86	8	5.50
C*14:02	9	4.24	9	4.25	9	4.25
C*15:02	10	3.68	10	3.37	10	3.55
DRB1*09:01	1	15.86	1	14.37	1	14.69
DRB1*15:01	2	13.13	2	11.61	2	12.40
DRB1*07:01	3	10.10	3	9.11	3	9.89
DRB1*12:02	4	7.87	4	8.67	4	8.98
DRB1*08:03	5	6.10	5	6.23	5	6.76
DRB1*11:01	6	5.92	6	5.69	6	5.64
DRB1*04:05	7	4.76	8	4.72	7	4.74
DRB1*03:01	8	3.84	7	4.93	8	3.69
DRB1*12:01	9	3.35	9	3.26	10	3.69
DRB1*15:02	10	3.18	13	3.05	14	2.65
DQB1*03:01	1	20.78	1	21.03	1	20.54
DQB1*03:03	2	17.06	2	15.78	2	15.81
DQB1*06:01	3	10.32	3	10.20	3	10.79
DQB1*06:02	4	10.04	4	7.82	5	8.08
DQB1*02:02	5	8.36	5	7.75	4	8.57
DQB1*05:02	6	5.75	6	7.18	7	6.20
DQB1*03:02	7	5.23	7	5.75	6	7.66
DQB1*04:01	8	4.64	10	4.49	9	4.39
DQB1*05:03	9	4.33	11	4.20	8	5.64
DQB1*05:01	10	4.08	9	4.62	10	2.58

tion of HLA AF and HF of larger and more varied population samples to address this issue.

## References

- [1] Sun YC. The HLA system: genetics, immunology, clinical testing and clinical implications. *Yonsei Med J*, 2007, 48: 11–23.
- [2] Lee SJ, Klein J, Haagenson M, et al. High resolution donor–recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*, 2007, 110: 4576–83.
- [3] Morishima Y, Sasazuki T, Inoko H, et al. The clinical significance of human leukocyte antigen (HLA) allele compatibility in patients receiving a marrow transplant from serologically HLA–A, HLA–B, and HLA–DR matched unrelated donors. *Blood*, 2002, 99: 4200–6.
- [4] Petersdorf EW. Optimal HLA matching in hematopoietic cell transplantation. *Curr Opin Immunol*, 2008, 20: 588–93.
- [5] Spellman SR, Eapen M, Logan BR, et al. A perspective on the selection of unrelated donors and cord blood units for transplantation. *Blood*, 2012, 120: 259–65.
- [6] Mori M, Graves M, Milford EL, et al. Computer program to predict likelihood of finding and HLA–matched donor: methodology, validation, and application. *Biol Blood Marrow Transplant*, 1996, 2: 134–44.
- [7] Schipper RF, D’Ama OJ, Oudschoorn M. The probability of finding a suitable related donor for bone marrow transplant in extended families. *Blood*, 1996, 87: 800–4.
- [8] Ghosh K, Shankarkumar U, Mohanty D. Common human leukocyte antigen haplotypes in Indians—its implications in finding unrelated compatible bone marrow donors. *Transf Med*, 2002, 12: 43–8.
- [9] Hurley CK, Fernandez VM, Setterholm M. Maximizing optimal hematopoietic stem cell donor selection from registries of unrelated adult volunteers. *Tissue Antigens*, 2003, 61: 415–24.
- [10] Hei AL, Li W, Deng ZH, et al. Analysis of high–resolution HLA–A, –B, –Cw, –DRB1, and –DQB1 alleles and haplotypes in 718 Chinese marrow donors based on donor–recipient confirmatory typings. *Int J Immunogenet*, 2009, 6: 275–82.
- [11] Li XF, Zhang X, Chen Y, et al. An analysis of HLA–A, –B, and –DRB1 allele and haplotype frequencies of 21,918 residents living in Liaoning, China. *PLoS ONE*, 2014, 9: e93082.
- [12] Pan QQ, Fan S, Wang XY, et al. The distribution of human leukocyte antigen–A, –B, and –DRB1 alleles and haplotypes based on high–resolution genotyping of 167 families from Jiangsu Province, China. *Hum Immunol*, 2011, 72: 872–6.
- [13] Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform*, 2005, 1: 47–50.
- [14] Gourraud PA, Balère ML, Faucher C, et al. HLA phenotypes of candidates for HSCT: comparing transplanted versus non–transplanted candidates, resulting in the predictive estimation of the probability to find a 10/10 HLA matched donor. *Tissue Antigens*, 2014, 83: 17–26.
- [15] Dehn J, Arora M, Spellman S, et al. Unrelated donor hematopoietic cell transplantation: factors associated with a better HLA match. *Biol Blood Marrow Transplant*, 2008, 14: 1334–40.
- [16] Petersdorf EW, Hansen JA, Martin PJ, et al. Major histocompatibility–complex class I alleles and antigens in hematopoietic–cell transplantation. *N Engl J Med*, 2001, 345: 1794–800.
- [17] Furst D, Müller C, Vucinic V, et al. High–resolution HLA matching in hematopoietic stem cell transplantation: a retrospective collaborative analysis. *Blood*, 2013, 122: 3220–9.
- [18] Woolfrey A, Klein JP, Haagenson M, et al. HLA–C antigen mismatch is associated with worse outcome in unrelated donor peripheral blood stem cell transplantation. *Biol Blood Marrow Transplant*, 2011, 17: 885–92.
- [19] Ottinger HD, Ferencik S, Beelen DW, et al. Impact of HLA–A, B, C allele mismatches on outcome after unrelated blood stem cell transplantation in whites. *Transplantation*, 2004, 78: 1077–80.
- [20] Mack SJ, Cano P, Hollenbach JA, et al. Common and well–documented HLA alleles: 2012 update to the CWD catalogue. *Tissue Antigens*, 2013, 81: 194–203.
- [21] Grubic Z, Jankovic KS, Maskalan M, et al. HLA allele and haplotype polymorphisms among Croatian patients in an unrelated hematopoietic stem cell donor search program. *Transpl Immunol*, 2014, 31: 119–24.
- [22] Bateman AC, Howell WM. Human leukocyte antigens and cancer: is it in our genes? *J Pathol*, 1999, 188: 231–6.
- [23] Dymant DA, Ebers GC, Sadovnick AD. Genetics of multiple sclerosis. *Lancet Neurol*, 2004, 3: 104–10.
- [24] Rasmussen HB, Kelly MA, Clausen J. Additive effect of the HLA–DR15 haplotype on susceptibility to multiple sclerosis. *Mult Scler*, 2001, 7: 91–3.
- [25] Wu JS, James I, Qiu W, et al. HLA–DRB1 allele heterogeneity influences multiple sclerosis severity as well as risk in Western Australia. *J Neuroimmunol*, 2010, 219: 109–13.
- [26] Qiu WI, James I, Carroll WM, et al. HLA–DR allele polymorphism and multiple sclerosis in Chinese populations: a meta–analysis. *Mult Scler*, 2011, 17: 382–8.
- [27] Bodmer JG, Tonks S, Oza AM, et al. Hodgkin’s disease study. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991, Vol 1*. Oxford: Oxford University Press, 1992: 701–9.
- [28] Taylor GM, Gokhale DA, Crowther D, et al. Increased frequency of HLA–DPB1\*0301 in Hodgkin’s disease suggests that susceptibility is HVR–sequence and subtype associated. *Leukemia*, 1996, 10: 854–9.
- [29] Jackow CM, McHam JB, Friss A, et al. HLA–DR5 and DQB1\*03 class II alleles are associated with cutaneous T–cell lymphoma. *J Invest Dermatol*, 1996, 107: 373–6.
- [30] Dorak MT, Owen G, Galbraith I, et al. Nature of HLA–associated predisposition to childhood acute lymphoblastic leukemia. *Leukemia*, 1995, 9: 875–8.

- [31] Taylor GM, Robinson MG, Binchy A, et al. Preliminary evidence of an association between HLA-DPB1\*0201 and childhood common acute lymphoblastic leukemia supports an infectious aetiology. *Leukemia*, 1995, 9: 440–3.
- [32] Taylor GM, Dearden S, Payne N, et al. Evidence that an HLA-DQA1-DQB1 haplotype influences susceptibility to childhood common acute lymphoblastic leukemia in boys provides further support for an infection related aetiology. *Br J Cancer*, 1998, 78: 561–5.
- [33] Skibola CF, Akers NK, CondeL, et al. Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. *Tissue Antigens*, 2012, 79: 279–86.
- [34] Pedron B, Yakouben K, Adjaoud D, et al. Listing of common HLA alleles and haplotypes based on the study of 356 families residing in the Paris, France, area: implications for unrelated hematopoietic stem cell donor selection. *Hum Immunol*, 2005, 66: 721–31.
- [35] Zeng J, Wang JX, Sun SX, et al. HLA-A, -B, -DRB1 alleles and haplotypes frequencies in Sichuan, China. *China J Blood Transfusion*, 2005, 18: 384–8.
- [36] Miao RH, Pan QQ, Tang RC, et al. The polymorphism and haplotype analysis of HLA-A, -B and -DRB1 genes of population in Jiangsu province of China. *Int J Immunogenet*, 2007, 34: 419–24.
- [37] Wang CQ, Zhu XF, Zheng Y. Allele frequency of human leukocyte antigen among hematopoietic stem cells donors in Fujian province. *J Mol Diagn Ther*, 2009, 1: 165–7.
- [38] Trachtenberg E, Vinson M, Hayes E, et al. HLA class I (A, B, C) and class II (DRB1, DQA1, DQB1, DPB1) alleles and haplotypes in the Han from southern China. *Tissue Antigens*, 2007, 70: 455–63.
- [39] Shen C, Zhu B, Liu M, et al. Genetic polymorphisms at HLA-A, -B, and-DRB1 loci in Han population of Xi'an city in China. *Croat Med J*, 2008, 49: 476–82.
- [40] Yao Y, Shi L, Shi L, et al. Distribution of HLA-A,-B, -Cw, and -DRB1 alleles and haplotypes in an isolated Han population in Southwest China. *Tissue Antigens*, 2009, 73: 561–8.

**(Received 31 December 2017, Revised 25 January 2018, Accepted 04 February 2018)**