Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

HLA polymorphism of the Zhuang population reflects the common HLA characteristics among Zhuang-Dong language-speaking populations*

Li SHI[§1], Xiao-qin HUANG[§1], Lei SHI[1], Yu-fen TAO[1], Yu-feng YAO[1], Liang YU[1], Ke-qin LIN[1], Wen YI[1], Hao SUN[1], Katsushi TOKUNAGA[†‡2], Jia-you CHU[†‡1]

([1]*Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming 650118, China*)

([2]*Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan*)

[†]E-mail: tokunaga@m.u-tokyo.ac.jp; chujy08@gmail.com

Received July 26, 2010; Revision accepted Dec. 27, 2010; Crosschecked Apr. 27, 2011

Abstract: A study of the human leukocyte antigen (HLA) genetic characteristics in the Zhuang, the largest ethnic population in China, would provide insight into Zhuang history and give a useful tool for disease associations, transplantation, and anthropology. In the present study, we report the comprehensive HLA-A, HLA-B, HLA-C, and HLA-DRB1 alleles and haplotypes in the Zhuang population of southern China for the first time. A total of 13 HLA-A, 24 HLA-B, 22 HLA-C, and 18 HLA-DRB1 were identified in 104 Zhuang individuals. The frequencies of HLA-A*11:01, A*02:07, A*24:02, A*02:03, and A*33:03 on A loci, B*15:02, B*58:01, B*46:01, and B*13:01 on B loci, C*03:04, C*08:01, C*01:02, C*03:02, and C*07:02 on C loci, and DRB1*15:01, DRB1*16:02, DRB1*14:01, DRB1*15:02, and DRB1*03:01 on the DRB1 loci were >10%. The A*33:03-C*03:02-B*58:01-DRB1*03:01 and A*02:07-C*01:02-B*46:01-DRB1*14:01 haplotypes were predominant in the Zhuang. The phylogenetic tree, as well as the analysis of haplotypes, suggested that the Zhuang are genetically similar to southern Chinese populations, especially the Zhuang-Dong language-speaking populations, such as the Bouyei, Dai, and Maonan. Even though the Zhuang and southern Chinese populations shared common alleles and haplotypes, the Zhuang has maintained its unique genetic characteristics.

Key words: Human leukocyte antigen (HLA) alleles, Haplotypes, Zhuang population
doi:10.1631/jzus.B1000285 **Document code:** A **CLC number:** Q987

1 Introduction

Human leukocyte antigens (HLAs) play a central role in immune response and exhibit extensively high levels of polymorphism. As of December 2009, 965 HLA-A, 1543 HLA-B, 626 HLA-C, and 762 HLA-DRB1 alleles had been identified worldwide (Marsh *et al.*, 2010). So, HLA has been used as a valuable tool for tracing the genetic background of human populations (Tokunaga *et al.*, 1996; 1997). Moreover, the high-resolution molecular typing method provides the precise examination of HLA allele distribution in different populations and allows a detailed analysis of the haplotypic relationship between alleles of different loci (Bugawan *et al.*, 1999; Cao *et al.*, 2004).

In addition to the Han population, the Zhuang ethnic group, with a population size of approximately 16.2 million, is the largest ethnic population in China. Most of the Zhuang live in the Guangxi Zhuang

‡ Corresponding authors
§ The two authors contributed equally to this work
* Project supported by the Yunnan Provincial Science and Technology Department (No. 2008CC021), the National Natural Science Foundation of China (No. 30900798), and the CAMS-JSPS China-Japan Medical Exchange Program

Autonomous Region in southern China; some Zhuang also live in the Yunnan, Guangdong, Guizhou, and Hunan provinces. The Zhuang people speak a language that belongs to the Zhuang-Dai branch of Zhuang-Dong (or Kam-Tai) subfamily of the Sino-Tibetan language family, which has northern and southern dialects. The Zhuang's language is related to Thai, Laos, and Dai languages. The Zhuang has recorded their history from the eastern Zhou dynasty (475 to 221 BC) of China. The Zhuang originated from the Xi'ou and Luoyue of the Baiyue ancient tribe, who had inhabited along the southeast coast of China up to the Yunnan province and the northern part of Southeast Asia from the fifth century BC (Guo and Dong, 2000).

Our previous studies based on HLA polymorphisms in several ethnic groups of Southwest China have suggested that the same linguistic populations tend to cluster together (Shi *et al*., 2006; 2008; Ogata *et al*., 2007). As a typical Zhuang-Dong language-speaking population, the analyses of HLA genes and haplotypes in the Zhuang population would be helpful for indentifying the common HLA characteristics of the same linguistic family. Furthermore, it would also provide an insight into Zhuang history and give a useful tool for HLA matching in transplantation and disease associations.

2 Materials and methods

A total of 104 health and unrelated individuals (unrelated at least three generations) were randomly selected in the Tiandeng county of Nanning city of Guangxi Zhuang Autonomous Region in southern China. Blood samples were collected after obtaining informed consent.

Genomic DNA was extracted from peripheral lymphocytes using the QIAamp blood kit (Qiagen, Hilton, Germany) according to the manufacturer's protocol. HLA-A, HLA-B, HLA-C, and HLA-DRB1 polymorphisms were detected using the Luminex multi-analyte profiling system (xMAP) with a WAKFlow HLA typing kit (Wakunaga, Hiroshima, Japan) as described before (Yao *et al*., 2009; Shi *et al*., 2010b). Genotype determination and data analyses were performed automatically using the WAKFlow typing software according to the manufacturer's instructions. The ambiguous alleles could not be identified using this method listed in Table 1.

Allele frequencies were calculated based on the HLA genotyping results and after linkage disequilibrium (LD) was analyzed, haplotype frequencies were calculated by the Expectation-Maximization algorithm using PyPop software (Lancaster *et al*., 2003; 2007). For each locus, the Hardy-Weinberg equilibrium test was performed as the method by Guo and Thompson (1992) and the Ewens-Watterson homozygosity test of neutrality was performed using PyPop software too (Ewens, 1972; Watterson, 1978). Nei's standard genetic distances among 18 populations were calculated based on HLA-A, HLA-B, and HLA-DRB1 allele frequencies and a neighbor-joining tree was constructed using Mega4.0 software (Tamura *et al*., 2007). Frequency data for other Asian populations were obtained from previous studies, as follows: Wa and Jinuo (Shi *et al*., 2008), Bouyei (Imanishi *et al*., 1992; Chen *et al*., 2007), Maonan (Ogata *et al*., 2007), Han in Taiwan (Wen *et al*., 2008), Han in Hongkong (Hardy *et al*., 1997), Han in Guangdong (Trachtenberg *et al*., 2007), Menba (Zhang *et al*., 2005), Tibetan (Lai *et al*., 1999; Chen *et al*., 2006), Hui (Yu *et al*., 2002; Hong *et al*., 2007), Han in Beijing (Sun *et al*., 1997; Deng *et al*., 2005), Han in Northern (Yang *et al*., 2006), Dai (Shi *et al*., 2010a), Han in Yunnan (Yao *et al*., 2009), and Bulang and Hani (Shi *et al*., 2010b). Fig. 1 shows the geographic distribution of these 17 populations (including the Zhuang) in China.

3 Results

3.1 Allele frequencies

The frequencies of HLA-A, HLA-B, HLA-C, and HLA-DRB1 alleles in the Zhuang population are shown in Table 1. A total of 13 HLA-A, 24 HLA-B, 22 HLA-C, and 18 HLA-DRB1 in the Zhuang were identified. The HLA-A locus was dominated by A*11:01, which had a frequency of 0.288. The next four most common alleles, A*02:07, A*24:02, A*02:03, and A*33:03, were each observed at frequencies of >10%, and these four alleles accounted for 85.9% of the total for HLA-A in this population. For the HLA-B locus, B*15:02, B*58:01, B*46:01, and B*13:01 were common in the Zhuang at

Table 1 HLA-A, HLA-B, HLA-C, and HLA-DRB1 allele frequencies of the Zhuang population

HLA-A	No.	AF	HLA-B	No.	AF	HLA-C	No.	AF	HLA-DRB1	No.	AF
A*11:01	61	0.288	B*15:02	34	0.160	C*03:04	34	0.160	DRB1*15:01	39	0.184
A*02:07	35	0.165	B*58:01	32	0.151	C*08:01	32	0.151	DRB1*16:02	32	0.151
A*24:02	30	0.142	B*46:01	31	0.146	C*01:02	30	0.142	DRB1*14:01	30	0.142
A*02:03	28	0.132	B*13:01	29	0.137	C*03:02	29	0.137	DRB1*15:02	29	0.137
A*33:03	28	0.132	B*38:02	12	0.057	C*07:02	27	0.127	DRB1*03:01	26	0.123
A*11:02	8	0.038	B*40:01	12	0.057	C*14:02	11	0.052	DRB1*09:01	10	0.047
A*29:01	7	0.033	B*51:01	10	0.047	C*04:03	8	0.038	DRB1*11:01	10	0.047
A*02:06	6	0.028	B*55:02	7	0.033	C*04:01	6	0.028	DRB1*12:02	9	0.042
A*02:01	3	0.014	B*15:25	6	0.028	C*07:14	5	0.024	DRB1*13:03	5	0.024
A*03:01	2	0.009	B*07:05	4	0.019	C*12:03	5	0.024	DRB1*10:01	4	0.019
A*26:01	2	0.009	B*39:01	4	0.019	C*15:05	4	0.019	DRB1*04:05	3	0.014
A*24:07	1	0.005	B*40:02	4	0.019	C*03:03	3	0.014	DRB1*04:06	3	0.014
A*31:01	1	0.005	B*52:01	4	0.019	C*08:04	3	0.014	DRB1*08:03	3	0.014
			B*15:01	3	0.014	C*12:02	3	0.014	DRB1*13:02	3	0.014
			B*27:04	3	0.014	C*01:14	2	0.009	DRB1*14:18	3	0.014
			B*48:01	3	0.014	C*03:05	2	0.009	DRB1*13:01	1	0.005
			B*56:01	3	0.014	C*03:37	2	0.009	DRB1*14:04	1	0.005
			B*56:02	3	0.014	C*08:03	2	0.009	DRB1*14:05	1	0.005
			B*35:01	2	0.009	C*03:07	1	0.005			
			B*35:05	2	0.009	C*04:15	1	0.005			
			B*35:03	1	0.005	C*14:09	1	0.005			
			B*38:01	1	0.005	C*15:02	1	0.005			
			B*56:04	1	0.005						
			B*67:01	1	0.005						

2*n*=212. No.: allele numbers; AF: allele frequencies. A*11:01=A*11:01/11:03/11:05; C*03:04=C*03:04/03:06/03:09/03:10; C*07:02=C*07:02/07:05; DRB1*14:01=DRB1*14:01/14:54

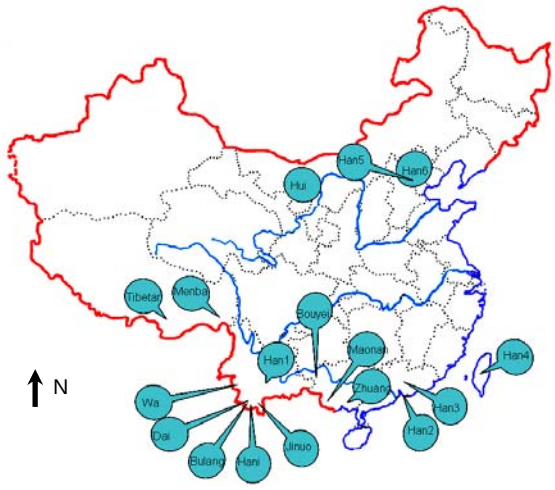

Fig. 1 Geographic distribution of the areas inhabited by 17 Chinese populations

Han1: Han in Yunnan; Han2: Han in Hongkong; Han3: Han in Guangdong; Han4: Han in Taiwan; Han5: Han in Beijing; Han6: Han in Northern

frequencies of >10% and accounted for 59.4% of the total. Similarly, for the HLA-C locus, C*03:04, C*08:01, C*01:02, C*03:02, and C*07:02 occurred at frequencies >10% and accounted for 71.7% of the total. For the HLA-DRB1 locus, DRB1*15:01, DRB1*16:02, DRB1*14:01, DRB1*15:02, and DRB1*03:01 were at frequencies >10% and accounted for 73.7% of the total. All of the alleles in this population were common and existed in other East Asia populations, except a rare allele, DRB1*14:18 (0.014), which was only identified in the Han in Yunnan (0.005), Lahu (0.018) and Yao (0.024) in Yunnan province (Jia *et al*., 2002; Liu *et al*., 2006; Yao *et al*., 2009).

3.2 Hardy-Weinberg equilibrium test and Ewens-Watterson homozygosity test of neutrality

The distribution of genotypes at each locus did not show significant deviations from those expected

under the Hardy-Weinberg equilibrium (HLA-A, P=0.518; HLA-B, P=0.657; HLA-C, P=0.330; HLA-DRB1, P=0.165). Ewens-Watterson homozygosity test of neutrality selection for all the loci could not reject the neutral model (Table 2).

3.3 Haplotype frequencies

The likelihood ratio test of linkage disequilibrium demonstrated that all pair-wise associations were statistically significant in the Zhuang (P<0.0001).

Table 2 Ewens-Watterson homozygosity test for neutrality in the Zhuang population

Locus	F_O	F_E	Fnd	Watterson F^*
A	0.169	0.249	−0.884	0.156
B	0.102	0.129	−0.644	0.266
C	0.111	0.142	−0.658	0.265
DRB1	0.119	0.178	−0.950	0.109

F_O: observed F; F_E: expected F; Fnd: normalized deviate of the homozygosity. $^*P$ value

Therefore, we estimated HLA haplotypes bearing 2–4 loci. Tables 3 and 4 show 2–4 loci haplotypes in which the frequencies were >2%.

Some predominant haplotypes even occurred at frequencies >10%, and included A*33:03-B*58:01 (12.3%), A*11:01-B*15:02 (10.7%), C*08:01-B*15:02 (15.1%), C*03:02-B*58:01 (13.7%), C*01:02-B*46:01 (13.2%), C*03:04-B*13:01 (12.2%), B*58:01-DRB1*03:01 (11.3%), A*33:03-C*03:02-B*58:01 (11.3%), and A*11:01-C*08:01-B*15:02 (10.0%). Other predominant haplotypes occurred at frequencies >5%, and included A*33:03-C*03:02-B*58:01-DRB1*03:01 (7.5%), A*02:07-C*01:02-B*46:01-DRB1*14:01 (6.6%), A*33:03-B*58:01-DRB1*03:01 (8.5%), A*02:07-B*46:01-DRB1*14:01 (7.0%), A*11:01-B*15:02-DRB1*15:01 (5.7%), C*03:02-B*58:01-DRB1*03:01 (9.9%), C*01:02-B*46:01-DRB1*14:01 (8.9%), C*08:01-B*15:02-DRB1*15:01 (5.9%), and C*03:04-B*13:01-DRB1*15:01 (5.4%).

Table 3 Two- and four-locus haplotype frequencies of HLA in the Zhuang population

A-B	HF	C-B	HF	B-DRB1	HF	A-C-B-DRB1	HF
33:03-58:01	0.123	08:01-15:02	0.151	58:01-03:01	0.113	33:03-03:02-58:01-03:01	0.075
11:01-15:02	0.107	03:02-58:01	0.137	46:01-14:01	0.099	02:07-01:02-46:01-14:01	0.066
02:07-46:01	0.099	01:02-46:01	0.132	15:02-15:01	0.070	11:01-08:01-15:02-15:01	0.047
11:01-13:01	0.070	03:04-13:01	0.122	13:01-15:01	0.054	11:01-03:04-13:01-16:02	0.033
02:03-38:02	0.038	07:02-38:02	0.052	13:01-16:02	0.051	24:02-08:01-15:02-15:02	0.028
11:01-51:01	0.035	14:02-51:01	0.042	15:02-15:02	0.045	02:03-07:02-38:02-16:02	0.025
24:02-13:01	0.033	04:03-15:25	0.024	40:01-16:02	0.032	02:06-01:02-46:01-14:01	0.024
02:03-15:02	0.025	07:02-40:01	0.024	51:01-11:01	0.028	11:01-14:02-51:01-11:01	0.024
02:06-46:01	0.024			38:02-16:02	0.025	24:02-03:04-13:01-15:01	0.023
11:01-15:25	0.024			15:02-16:02	0.022	11:01-03:04-13:01-15:01	0.021
24:02-15:02	0.023						
02:03-55:02	0.021						
02:07-13:01	0.021						

n=106. HF: haplotype frequency

Table 4 Three-locus haplotype frequencies of HLA in the Zhuang population

A-B-DRB1	HF	A-C-B	HF	C-B-DRB1	HF
33:03-58:01-03:01	0.085	33:03-03:02-58:01	0.113	03:02-58:01-03:01	0.099
02:07-46:01-14:01	0.070	11:01-08:01-15:02	0.100	01:02-46:01-14:01	0.089
11:01-15:02-15:01	0.057	02:07-01:02-46:01	0.094	08:01-15:02-15:01	0.059
11:01-13:01-16:02	0.043	11:01-03:04-13:01	0.055	03:04-13:01-15:01	0.054
11:01-51:01-11:01	0.028	02:03-07:02-38:02	0.038	08:01-15:02-15:02	0.046
24:02-15:02-15:02	0.028	24:02-03:04-13:01	0.035	03:04-13:01-16:02	0.040
11:01-13:01-15:01	0.025	11:01-14:02-51:01	0.033	07:02-38:02-16:02	0.026
02:03-38:02-16:02	0.025	24:02-08:01-15:02	0.027	14:02-51:01-11:01	0.024
02:06-46:01-14:01	0.024	02:06-01:02-46:01	0.024	08:01-15:02-16:02	0.023
24:02-13:01-15:01	0.023	11:01-04:03-15:25	0.024		

n=106. HF: haplotype frequency

3.4 Phylogenetic tree analysis

We constructed a neighbor-joining tree using HLA-A, HLA-B, and HLA-DRB1 allele frequencies in 18 populations because HLA-C genotyping results were not found in some populations of interest (Fig. 2). On this tree, the Han in Northern and Han in Beijing were clustered together with the Hui, Menba, and Tibetan as the northern branch. The southern Han, which included the Han in Yunnan, Han in Hongkong, Han in Singapore, Han in Guangdong, and Han in Taiwan, was clustered together and close to the Zhuang, Bouyei, Dai, and Maonan, while the Wa, Bulang, Hani, and Jinuo were clustered together and showed some genetic distance to the other southern Chinese populations.

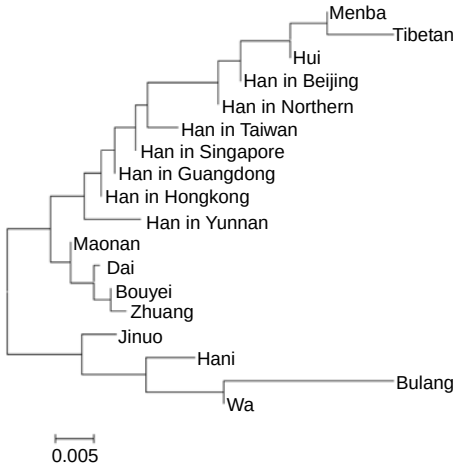

Fig. 2 Neighbor-joining tree of 18 populations in China constructed by Nei's genetic distance based on HLA-A, HLA-B, and HLA-DRB1 frequencies

4 Discussion

The study of the HLA characteristics of the Zhuang is useful for medical and forensic applications, as well as in anthropology. This is the first report on the distributions of HLA-A, HLA-B, HLA-C, and HLA-DRB1 alleles and haplotypes at high resolution in the Zhuang.

As the largest ethnic group in China, the Zhuang has had a close cultural affiliation with the Han for centuries (Guo and Dong, 2000). On the neighbor-joining tree in the present study, the Zhuang belonged to the southern Chinese population. Most of the alleles identified in this study have been frequently found in other southern Chinese populations. The predominant alleles in the Zhuang, such as A*11:01, A*02:07, A*02:03, B*15:02, B*46:01, B*13:01, DRB1*16:02, and DRB1*14:01, were predominant in the Chinese, especially in southern Chinese populations. In addition to alleles, the predominant haplotypes in the Zhuang were also predominant in the southern Chinese population and southern East Asia populations. For example, the predominant A-C-B haplotype, A*02:07-C*01:02-B*46:01, was predominant in Vietnamese as well as in some southern Chinese populations, such as the Dai, Hani, Miao, Bouyei, Shui, Han in Yunnan, and Han in Guangdong, but existed at low frequencies in Chinese Hui, Japanese, and Koreans. Another predominant haplotype, A*11:01-C*08:01-B*15:02, was also predominant in Vietnamese and some Chinese populations, such as the Bulang, Dai, Hani, Shui, Han in Yunnan, and Han in Guangdong, who were located in southern China (Tokunaga *et al*., 1997; Middleton *et al*., 2004; Liu *et al*., 2006; Chen *et al*., 2007; Hong *et al*., 2007; Ogata *et al*., 2007; Trachtenberg *et al*., 2007; Hoa *et al*., 2008; Shi *et al*., 2008; 2010a; 2010b; Wen *et al*., 2008; Yao *et al*., 2009). Based on the HLA alleles, haplotypes, and neighbor-joining tree analysis, we concluded that the Zhuang is a typical southern Chinese population.

Despite the Zhuang and southern Chinese populations sharing the common alleles and haplotypes, the Zhuang still keeps its unique genetic characteristics. For example, A*11:01-B*15:02-DRB1*12:02 was common in southern populations of East Asia (Ogata *et al*., 2007; Trachtenberg *et al*., 2007; Hoa *et al*., 2008; Shi *et al*., 2008; 2010a; 2010b; Wen *et al*., 2008; Gendzekhadze *et al*., 2009; Yao *et al*., 2009); however, it was not identified in the Zhuang. Instead, A*11:01-B*15:02-DRB1*15:01 was predominant in the Zhuang, but was not identified in other East Asian populations. Except the different frequencies of predominant haplotypes in different populations, some unique alleles were only indentified in the Zhuang, such as A*02:06-B*46:01-DRB1*14:01, A*11:01-B*51:01-DRB1*11:01, A*24:02-B*13:01-DRB1*15:01, A*24:02-B*15:02-DRB1*15:02, and A*02:06-C*01:02-B*46:01. Therefore, the Zhuang is a typical southern Chinese population, but maintains its unique HLA characteristics as well.

From the neighbor-joining tree, we also found

that the Zhuang clustered together with the Bouyei (Buyi or Bouyi), Dai, and Maonan, who all speak the Zhuang-Dong language. Our previous studies based on HLA distributions suggested that the same linguistic-speaking populations were genetically similar to each other (Shi *et al*., 2006; 2008; Ogata *et al*., 2007). The other studies of Y chromosomes, mitochondrial DNA, as well as short tandem repeat, have also demonstrated that the Zhuang-Dong language-speaking population showed close affinity (Yao *et al*., 2002; Shi *et al*., 2005; Tian *et al*., 2008). For example, the analysis of the hypervariable segment I sequences of the mitochondrial DNA control region in different Chinese ethnic populations and Thai has showed that the haplotype diversities of three Zhuang-Dong language-speaking populations including the Zhuang, Dai and Thai were similar, and the genetic distance among them even showed negative estimated values (Yao *et al*., 2002). When we compared the distribution of the HLA alleles in the Zhuang, Bouyei, Dai, Maonan, and Shui, we found that the HLA characteristics were very similar to each other. Not only were the predominant alleles with high frequencies similar, but also the common alleles with moderate frequencies did. For example, on the HLA-A locus, the predominant alleles with frequencies >10% commonly existed in these four populations, including A*02:07, A*02:03, A*11:01, and A*24:02, while the common alleles with frequencies around 1%–5% included A*02:01, A*02:06, and A*11:02, and all of these alleles comprised 80.66%–95.37% of the total HLA-A alleles, respectively. On the HLA-B locus, the commonly predominant alleles included B*13:01, B*15:02, B*38:02, B*40:01, B*46:01, and B*58:01, while the alleles with moderate frequencies included B*15:01, B*15:25, B*27:04, B*39:01, B*40:02, B*51:01, B*55:02, and B*56:01, and all of them comprised 84.27%–94.45% of the total HLA-B alleles. On the HLA-C locus, the commonly predominant alleles included C*01:02, C*03:02, C*03:04, C*07:02, and C*08:01, while the alleles with moderate frequencies included C*03:03, C*04:03, C*12:02, C*12:03, and C*15:02, and all of them comprised 84.91%–97.30% of the total HLA-C alleles (Chen *et al*., 2007; Ogata *et al*., 2007; Shi *et al*., 2010a). When we compared the HLA-DRB1 locus, two Zhuang-Dong language-speaking populations (Tai Dam and Dai Lue in Thailand) were added. Similarly, the commonly predominant alleles existed in these six populations including DRB1*14:01, DRB1*16:02, DRB1*15:01, and DRB1*15:02, while the alleles with moderate frequencies included DRB1*04:05, DRB1*04:06, DRB1*08:03, DRB1*11:01, and DRB1*13:03. However, DRB1*03:01 was predominant in the Zhuang but moderate in other populations, while DRB1*09:01 and DRB1*12:02 were moderate in the Zhuang but predominant in other populations (Imanishi *et al*., 1992; Chandanayingyong, 1997; Stephens *et al*., 2000). The haplotypes were also similar among the Zhuang-Dong language-speaking populations. For example, A*02:07-B*46:01, A*11:01-B*15:02, A*33:03-B*58:01, B*46:01-DRB1*14:01, and B*58:01-DRB1*03:01 were predominant in the Zhuang, Dai, and Maonan; however, A*11:01-B*40:01, the most common haplotype existing in other Chinese populations, was not identified in the Zhuang, Dai, and Maonan. A*02:07-C*01:02-B*46:01 and A*11:01-C*08:01-B*15:02 were predominant in the Zhuang, Dai, Bouyei, and Shui (Chen *et al*., 2007; Ogata *et al*., 2007; Shi *et al*., 2010a). Zhuang-Dong language-speaking populations include the Zhuang, Bouyei, Dai, Dong, Shui, Mulam (Mulao), Maonan, and Li, who live in southern China. According to the historical record, all of them originated from the ancient Baiyue tribe, who had inhabited along the southeast coast of China up to the Yunnan province and the northern part of Southeast Asia from the fifth century BC (Guo and Dong, 2000; You, 1994). Some Pai-Yuei people later migrated to the North Thailand and contributed to the ancient gene pool of the Thai (Chandanayingyong, 1997; Stephens *et al*., 2000). The Zhuang-Dong language-speaking populations used to live in semi-tropic areas and live on planting rice. The similar surrounding, similar diet shift, similar language, together with similar origination make these populations to be genetically close to each other even though they live in different geographic places.

In conclusion, according to our HLA data, we observed that the Zhuang was genetically close to southern Chinese populations. Zhuang-Dong language-speaking populations, like the Bouyei, Dai, and Maonan, have similar HLA characteristics. Even though the Zhuang and southern Chinese populations shared common alleles and haplotypes, the Zhuang still keeps its unique genetic characteristics.

References

Bugawan, T.L., Mack, S.J., Stoneking, M., Saha, M., Beck, H.P., Erlich, H.A., 1999. HLA class I allele distributions in six Pacific/Asian populations: evidence of selection at the HLA-A locus. *Tissue Antigens*, **53**(4):311-319. [doi:10.1034/j.1399-0039.1999.530401.x]

Cao, K., Moormann, A.M., Lyke, K.E., Masaberg, C., Sumba, O.P., Doumbo, O.K., Koech, D., Lancaster, A., Nelson, M., Meyer, D., *et al*., 2004. Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. *Tissue Antigens*, **63**(4):293-325. [doi:10.1111/j.0001-2815.2004.00192.x]

Chandanayingyong, D., 1997. Thai Black, Thai Dai Lue, Thai Normal. *In*: Gjertson, D.W., Terasaki, P.I. (Eds.), HLA 1998. American Society for Histocompatibility and Immunogenetics, Lenexa, p.308-313.

Chen, S., Hong, W., Shao, H., Fu, Y., Liu, X., Chen, D., Xu, A., 2006. Allelic distribution of HLA class I genes in the Tibetan ethnic population of China. *Int. J. Immunogenet.*, **33**(6):439-445. [doi:10.1111/j.1744-313X.2006.00637.x]

Chen, S., Ren, X., Liu, Y., Hu, Q., Hong, W., Xu, A., 2007. Human leukocyte antigen class I polymorphism in Miao, Bouyei, and Shui ethnic minorities of Guizhou, China. *Hum. Immunol.*, **68**(11):928-933. [doi:10.1016/j.humimm.2007.09.006]

Deng, Y.J., Yang, G., Wu, D.Y., Hu, S.N., Li, S.B., Zhu, B.F., Zhu, J., Liu, Y., 2005. Study on the HLA-DRB1 polymorphism determined by high-resolution PCR-SBT in a population of Beijing region. *J. Forensic Med.*, **21**(4):280-285 (in Chinese).

Ewens, W.J., 1972. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, **3**(1):87-112. [doi:10.1016/0040-5809(72)90035-4]

Gendzekhadze, K., Norman, P.J., Abi-Rached, L., Graef, T., Moesta, A.K., Layrisse, Z., Parham, P., 2009. Co-evolution of KIR2DL3 with HLA-C in a human population retaining minimal essential diversity of KIR and HLA class I ligands. *PNAS*, **106**(44):18692-18697. [doi:10.1073/pnas.0906051106]

Guo, D.L., Dong, J.Z., 2000. Summarization of Chinese Nationalities (Zhonghua Mingzu Zhishi Tonglan). Yunnan Education Press, Kunming, p.98-100 (in Chinese).

Guo, S.W., Thompson, E.A., 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**(2):361-372. [doi:10.2307/2532296]

Hardy, S., Siauw, P.A., Scotland, R., Scotland, R., Gu, W.T., Fajatin, M., Luman, S., Shu, M., Song, J., Fueduska, N., Urban, V., 1997. Chinese Hong Kong Normal. *In*: Gjertson, D.W., Terasaki, P.I. (Eds.), HLA 1997. UCLA Tissue Typing Laboratory, Los Angeles, p.261.

Hoa, B.K., Hang, N.T., Kashiwase, K., Ohashi, J., Lien, L.T., Horie, T., Shojima, J., Hijikata, M., Sakurada, S., Satake, M., *et al*., 2008. HLA-A, -B, -C, -DRB1 and -DQB1 alleles and haplotypes in the Kinh population in Vietnam. *Tissue Antigens*, **71**(2):127-134. [doi:10.1111/j.1399-0039.2007.00982.x]

Hong, W., Chen, S., Shao, H., Fu, Y., Hu, Z., Xu, A., 2007. HLA class I polymorphism in Mongolian and Hui ethnic groups from Northern China. *Hum. Immunol.*, **68**(5):439-448. [doi:10.1016/j.humimm.2007.01.020]

Imanishi, T., Akaza, T., Kimura, A., Tokunaga, K., Gojobori, T., 1992. Allele and Haplotype Frequencies for HLA and Complement Loci in Various Ethnic Groups. *In*: Tsuji, K., Aizawa, M., Sasazuki, T. (Eds.), HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference. Oxford University Press, Oxford, p.1065-1220.

Jia, Z.J., Fu, Y.G., Pan, D.J., Liu, Z.H., Chen, W.M., Lin, J.H., Chen, R.X., Li, A.S., Zhu, Y.F., Zhou, D.M., *et al*., 2002. Determination of HLA-DRB1 gene polymorphism by PCR-SBT in Lahu ethnic group of Yunnan, China. *Hereditas*, **24**(2):131-136 (in Chinese).

Lai, S.P., Ren, H.M., Hu, H.T., Li, S.B., Lai, J.H., Yan, C.X., Zhang, H.B., Zhao, J.H., 1999. HLA-DRB alleles polymorphism in Han, Hui, Uygur and Tibetan populations in Northwestern China. *Acta Genet. Sin.*, **26**(5):447-457 (in Chinese).

Lancaster, A., Nelson, M.P., Meyer, D., Thomson, G., Single, R.M., 2003. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. *Pac. Symp. Biocomput.*, **2003**:514-525.

Lancaster, A.K., Single, R.M., Solberg, O.D., Nelson, M.P., Thomson, G., 2007. PyPop update—a software pipeline for large-scale multilocus population genomics. *Tissue Antigens*, **69**(Suppl. 1):192-197. [doi:10.1111/j.1399-0039.2006.00769.x]

Liu, Y., Liu, Z., Fu, Y., Jia, Z., Chen, S., Xu, A., 2006. Polymorphism of HLA class II genes in Miao and Yao nationalities of Southwest China. *Tissue Antigens*, **67**(2):157-159. [doi:10.1111/j.1399-0039.2006.00510.x]

Marsh, S.G., Albert, E.D., Bodmer, W.F., Bontrop, R.E., Dupont, B., Erlich, H.A., Fernandez-Vina, M., Geraghty, D.E., Holdsworth, R., Hurley, C.K., *et al*., 2010. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*, **75**(4):291-455. [doi:10.1111/j.1399-0039.2010.01466.x]

Middleton, D., Hawkins, B.R., Williams, F., Meenagh, A., Moscoso, J., Zamora, J., Arnaiz-Villena, A., 2004. HLA class I allele distribution of a Hong Kong Chinese population based on high-resolution PCR-SSOP typing. *Tissue Antigens*, **63**(6):555-561. [doi:10.1111/j.0001-2815.2004.00234.x]

Ogata, S., Shi, L., Matsushita, M., Yu, L., Huang, X.Q., Shi, L., Sun, H., Ohashi, J., Muramatsu, M., Tokunaga, K., *et al*., 2007. Polymorphisms of human leucocyte antigen genes in Maonan people in China. *Tissue Antigens*, **69**(2):154-160. [doi:10.1111/j.1399-0039.2006.00698.x]

Shi, H., Dong, Y.L., Wen, B., Xiao, C.J., Underhill, P.A., Shen, P.D., Chakraborty, R., Jin, L., Su, B., 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *Am. J. Hum. Genet.*, **77**(3):408-419. [doi:10.1086/444436]

Shi, L., Xu, S.B., Ohashi, J., Sun, H., Yu, J.K., Huang, X.Q., Tao, Y.F., Yu, L., Horai, S., Chu, J.Y., *et al*., 2006.

HLA-A, HLA-B, and HLA-DRB1 alleles and haplotypes in Naxi and Han populations in southwestern China (Yunnan province). *Tissue Antigens*, **67**(1):38-44. [doi:10.1111/j.1399-0039.2005.00526.x]

Shi, L., Ogata, S., Yu, J.K., Ohashi, J., Yu, L., Shi, L., Sun, H., Lin, K., Huang, X.Q., Matsushita, M., *et al*., 2008. Distribution of HLA alleles and haplotypes in Jinuo and Wa populations in Southwest China. *Hum. Immunol.*, **69**(1):58-65. [doi:10.1016/j.humimm.2007.11.007]

Shi, L., Yao, Y.F., Shi, L., Matsushita, M., Yu, L., Lin, Q.K., Tao, Y.F., Oka, T., Chu, J.Y., Tokunaga, K., 2010a. HLA alleles and haplotypes distribution in Dai population in Yunnan province, Southwest China. *Tissue Antigens*, **75**(2):159-165. [doi:10.1111/j.1399-0039.2009.01407.x]

Shi, L., Shi, L., Yao, Y.F., Matsushita, M., Yu, L., Huang, X.Q., Yi, W., Oka, T., Tokunaga, K., Chu, J.Y., 2010b. Genetic link among Hani, Bulang and other Southeast Asian populations: evidence from HLA-A, -B, -C, -DRB1 genes and haplotypes distribution. *Int. J. Immunogenet.*, **37**(6):467-475. [doi:10.1111/j.1744-313X.2010.00949.x]

Stephens, H.A., Chandanayingyong, D., Kunachiwa, W., Sirikong, M., Longta, K., Maneemaroj, R., Wongkuttiya, D., Sittisombut, N., Rungruang, E., 2000. A comparison of molecular HLA-DR and DQ allele profiles forming DR51-, DR52-, and DR53-related haplotypes in five ethnic Thai populations from mainland Southeast Asia. *Hum. Immunol.*, **61**(10):1039-1047. [doi:10.1016/S0198-8859(00)00172-5]

Sun, X., Sun, Y., Gao, X., 1997. Diversity of HLA Class I Genes in Northern and Southern Chinese. *In*: Charron, D. (Ed.), HLA: Genetic diversity of HLA. Functional and Medical Implication. EDK, Paris, p.155-157.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, **24**(8):1596-1599. [doi:10.1093/molbev/msm092]

Tian, C., Kosoy, R., Lee, A., Ransom, M., Belmont, J.W., Gregersen, P.K., Seldin, M.F., 2008. Analysis of East Asia genetic substructure using genome-wide SNP arrays. *PLoS ONE*, **3**(12):e3862. [doi:10.1371/journal.pone.0003862]

Tokunaga, K., Imanishi, T., Takahashi, K., Juji, T., 1996. On the Origin and Dispersal of East Asian Populations as Viewed from HLA Haplotypes. *In*: Akazawa, T., Szathmary, E.J. (Eds.), Prehistoric Mongoloid Dispersals. Oxford University, Oxford, p.187-197.

Tokunaga, K., Ishikawa, Y., Ogawa, A., Wang, H., Mitsunaga, S., Moriyama, S., Lin, L., Bannai, M., Watanabe, Y., Kashiwase, K., *et al*., 1997. Sequence-based association analysis of HLA class I and II alleles in Japanese supports conservation of common haplotypes. *Immunogenetics*, **46**(3):199-205. [doi:10.1007/s002510050262]

Trachtenberg, E., Vinson, M., Hayes, E., Hsu, Y.M., Houtchens, K., Erlich, H., Klitz, W., Hsia, Y., Hollenbach, J., 2007. HLA class I (A, B, C) and class II (DRB1, DQA1, DQB1, DPB1) alleles and haplotypes in the Han from southern China. *Tissue Antigens*, **70**(6):455-463. [doi:10.1111/j.1399-0039.2007.00932.x]

Watterson, G.A., 1978. The homozygosity test of neutrality. *Genetics*, **88**(2):405-417.

Wen, S.H., Lai, M.J., Yang, K.L., 2008. Human leukocyte antigen-A, -B, and -DRB1 haplotypes of cord blood units in the Tzu Chi Taiwan Cord Blood Bank. *Hum. Immunol.*, **69**(7):430-436. [doi:10.1016/j.humimm.2008.05.012]

Yang, G., Deng, Y.J., Hu, S.N., Wu, D.Y., Li, S.B., Zhu, J., Zhu, B.F., Liu, Y., 2006. HLA-A, -B, and -DRB1 polymorphism defined by sequence-based typing of the Han population in Northern China. *Tissue Antigens*, **67**(2):146-152. [doi:10.1111/j.1399-0039.2006.00529.x]

Yao, Y., Shi, L., Shi, L., Matsushita, M., Yu, L., Lin, K., Tao, Y., Huang, X., Yi, W., Oka, T., *et al*., 2009. Distribution of HLA-A, -B, -Cw, and -DRB1 alleles and haplotypes in an isolated Han population in Southwest China. *Tissue Antigens*, **73**(6):561-568. [doi:10.1111/j.1399-0039.2009.01237.x]

Yao, Y.G., Nie, L., Harpending, H., Fu, Y.X., Yuan, Z.G., Zhang, Y.P., 2002. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *Am. J. Phys. Anthropol.*, **118**(1):63-76. [doi:10.1002/ajpa.10052]

You, Z., 1994. History of Yunnan Nationalities. Yunnan University Press, Kunming, p.1-11 (in Chinese).

Yu, H.L., Dai, G.Y., Xia, Y.Z., 2002. Investigation of the polymorphism of HLA-DRB1 alleles in minority and Han nationality in Inner Mongolia. *Acta Acad. Med. Nei Mongol.*, **24**:226-228 (in Chinese).

Zhang, H.B., Gao, F., Kang, L.L., Li, S.B., 2005. Determination of HLA-A, B, and DRB1 gene polymorphism in Monba ethnic group of Xizang Autonomous Region, China. *Chin. J. Med. Genet.*, **22**(3):344-346 (in Chinese).