

HLA class I diversity among rural rainforest inhabitants in Cameroon: identification of *A*2612-B*4407* haplotype

J. N. Torimiro¹, J. K. Carr², N. D. Wolfe³, P. Karacki⁴, M. P. Martin⁴, X. Gao⁴, U. Tamoufe³, A. Thomas⁵, E. M. Ngole^{1,6}, D. L. Birx⁷, F. E. McCutchan², D. S. Burke³ & M. Carrington⁴

1 Walter Reed – Johns Hopkins Cameroon Program, Yaounde, Cameroon

2 Henry M. Jackson Foundation, Rockville, MD, USA

3 Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

4 Basic Research Program, SAIC Frederick Inc., Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD, USA

5 London School of Hygiene and Tropical Medicine, London, UK

6 Military Health Research Center, Yaounde, Cameroon

7 Walter Reed Army Institute of Research, Rockville, MD, USA

Key words

allele frequency; Cameroon; haplotype frequency; heterozygosity; locus; polymorphic; population

Correspondence

Judith N. Torimiro, PhD

Walter Reed – Johns Hopkins Cameroon Program

B.P. 7039 Yaounde
Cameroon

Tel: 237 222 9161/237 772 9270

Fax: 237 221 3382

e-mail: jtorimro@hopkinscameroon.org

Received 25 July 2005; revised 07 September 2005; accepted 30 September 2005

doi: 10.1111/j.1399-0039.2005.00527.x

Abstract

The population distribution of alleles of the classical HLA class I loci in Cameroon has not been well studied but is of particular interest given the AIDS and malarial epidemics afflicting this population. We investigated the genetic diversity of HLA-A, HLA-B and HLA-C alleles in remote populations of Cameroon. Subjects from seven small, isolated, indigenous populations ($N = 274$) in the rainforest of southern Cameroon were typed for HLA-A, HLA-B and HLA-C alleles using a polymerase chain reaction/sequence-specific oligonucleotide probe assay and sequence analysis. Multiple alleles of the HLA-A ($N = 28$), HLA-B ($N = 41$) and HLA-C ($N = 21$) loci were identified, of which *A*2301* [allele frequency (AF) = 12.8%], *B*5802* (AF = 10.9%) and *Cw*0401* (AF = 16.6%) were the most frequent individual alleles and *A*02* (AF = 19.0%), *B*58* (AF = 15.9%) and *Cw*07* (AF = 22.4%) the most common serologically defined groups of alleles. Twenty-six (28.9%) alleles with a frequency of less than 1% (AF < 1%), 39 (43%) with a frequency of 2.0–15.0% (AF = 2.0–15.0%), three globally uncommon alleles [*A*2612* (AF = 2.0%), *B*4016* (AF = 0.7%) and *B*4407* (AF = 1.4%)], and the *A*2612-Cw*0701/06/18-B*4407* haplotype (haplotype frequency = 1.3%) were also identified. Heterozygosity values of 0.89, 0.92 and 0.89 were determined for *HLA-A*, *HLA-B* and *HLA-C*, respectively. The extensive allelic and haplotypic diversity observed in this population may have resulted from varied natural selective pressures on the population, as well as intermingling of peoples from multiple origins. Thus, from an anthropologic perspective, these data highlight the challenges in T-cell-based vaccine development, the identification of allogeneic transplant donors and the understanding of infectious disease patterns in different populations.

Introduction

The *HLA-A*, *HLA-B* and *HLA-C* genes within the major histocompatibility complex (MHC) are highly polymorphic, particularly in the peptide-binding region encoded by exons 2 and 3 (1). Sequences for 396 *HLA-A*, 699 *HLA-B* and 198 *HLA-C* alleles identified worldwide

are available through the European Bioinformatic Institute web site (<http://www.ebi.ac.uk/imgt/hla/docs/release.html>). *HLA* heterogeneity, which reflects a natural selective process for protective immunity, may not only modulate antigen presentation and T-cell recognition but also modify the T-cell repertoire (2). Within human

populations, several polymorphic *HLA* amino acid motifs have persisted to different degrees and at different times (3), resulting in the generation and fixation of new alleles. Consequently, disease susceptibility and progression may vary among populations due to differences in selective pressures and functional adaptations.

Africans display greater *HLA* diversity than other racial groups (4–7). Over 200 tribes with distinct language, art and crafts, tradition, religion and history contribute to the rich culture of the Cameroon population. In addition, migration of Europeans, Arabs and Americans into Cameroon has influenced the population structure of the native Cameroon Bantu-speaking people. However, among certain ethnic groups, such as the pygmies living in Southern and Eastern Cameroon, the cultural heritage has remained unchanged for centuries. The population of Cameroon is estimated to be over 15 million (2001 estimate) and is categorized by geographic origins into five main ethnic groups dispersed into proto-Bantu groups in the south and sudanic settlements in the north (<http://inet01.cm.undp.org/cameroun>).

The slow southward movement of the Bantu-speaking group from West-Central Africa throughout sub-Saharan Africa and the migration of people of different origins and languages into Cameroon in the 19th century created the basis for gene flow, which has led to the description of Cameroon as a crossroads in Africa (8). One of the largest migrations in human history of the Bantu linguistic group began from Cameroon and split into the Eastern and Western Bantu groups. Understanding the *HLA* diversity of a predominantly Bantu-speaking population, such as the closed geo-cultural settlements in the southern Cameroon rainforest, coincidentally a region where a broad diversity of HIV-1 has been reported (9), may provide clues to the design of T-cell vaccines (10).

Materials and methods

Study sites and study population

Seven rural locations in five Provinces in the Cameroon rainforest region with minimal access to other human settlements were selected (Figure 1). The study population is predominantly of the equatorial Bantu, which is further subdivided into Beti ($n = 174$), Bamileke ($n = 77$) and Sawa ($n = 13$) non-Pygmy ethnic groups and the Pygmy Bakas ($n = 10$) practicing peasant and subsistence agriculture in the lowland or savannah forest. From the Center, South and East Provinces are found the Betis, from the West plateau and grassfields the Bamilekes who constitute the majority of the Cameroon population, along the Littoral Provinces the Sawas, and in the East and South Provinces the pygmy populations (<http://inet01.cm.undp.org/cameroun>).



Figure 1 Map of Cameroon showing study sites. Study sites (red dots), Bangourein and Massangan, are located in the grassfields in the West plateau and are occupied predominantly by the Bamilekes, while the Betis form a distinct cultural group in Ndikinimeki, Nyabessang, Lomie and Moloundou, and the Sawas inhabit the coastal region in the rainforest in Manyemen.

Methods

Blood specimen collection and processing

Blood was collected after informed consent was given by each of the 274 adult volunteers. DNA was isolated from peripheral blood mononuclear cells using the automated MagNA Pure LC DNA Isolation equipment (Roche Molecular Biochemicals, Mannheim, Germany).

HLA class I genotyping

Polymerase chain reaction (PCR)/sequence-specific oligonucleotide probe (SSOP) assay was used to genotype *HLA-A*, *HLA-B* and *HLA-C* alleles (<http://www.ihwg.org>). Locus-specific primers were used to amplify exons 2 and 3 in a two-round PCR protocol. The amplified DNA was immobilized on nylon membranes and hybridized with SSOP. *HLA* alleles were assigned by the reaction patterns of the SSOP based on known *HLA* sequences. Samples with ambiguous SSOP results were resolved by sequencing using the ABI Big Dye Terminator Cycle Sequence Ready Reaction kit (Applied Biosystems, Foster City, CA, USA) (11). Exons 1, 4 and 5 were not examined, and therefore, it was not possible to resolve

Table 1a Frequency of HLA-A alleles in the different ethnic groups

HLA-A allele	Major ethnic group					Frequency (%) (Ellis <i>et al.</i> 2003) (N = 184)
	Bamileke [n (%)] (N = 154)	Beti [n (%)] (N = 348)	Sawa [n (%)] (N = 26)	Baka [n(%)] (N = 20)	Combined [n (%)] (N = 548)	
A*0101	4 (2.6)	4 (1.1)	0	0	8 (1.4)	1.1
A*0102	0	1 (0.3)	0	0	1 (0.2)	–
A*0201	17 (1.1)	39 (11.2)	1 (3.8)	0	57 (10.4)	–
A*0202	16 (10.4)	21 (6.0)	0	0	37 (6.7)	8.2
A*0205	4 (2.6)	4 (1.1)	1 (3.8)	1 (5.0)	10 (1.8)	2.2
A*0301	12 (7.8)	30 (8.6)	1 (3.8)	1 (5.0)	44 (8.0)	–
A*2301/07N	23 (14.9)	47 (13.5)	0	0	70 (12.8)	18.7
A*2402/09N	0	1 (0.3)	2 (7.7)	2 (10.0)	5 (0.9)	–
A*2501	0	1 (0.3)	0	0	1 (0.2)	–
A*2601	1 (0.6)	7 (2.0)	0	0	8 (1.4)	1.7
A*2612	4 (2.6)	7 (2.0)	0	0	11 (2.0)	0.6
A*2902	13 (8.4)	22 (6.3)	2 (7.7)	2 (10.0)	39 (7.1)	10.4
A*3001	5 (3.2)	14 (4.0)	1 (3.8)	0	20 (3.6)	5.5
A*3002	11 (7.1)	33 (9.5)	2 (7.7)	1 (5.0)	47 (8.6)	6.0
A*3004	1 (0.6)	3 (0.9)	3 (11.5)	3 (15.0)	10 (1.8)	1.1
A*3101	0	4 (1.1)	0	0	4 (0.7)	1.1
A*3201	0	7 (2.0)	0	0	7 (1.3)	1.1
A*3301	1 (0.6)	1 (0.3)	0	0	2 (0.4)	–
A*3303	7 (4.5)	8 (2.3)	6 (23.1)	5 (25.0)	26 (4.7)	2.8
A*3402	4 (2.6)	11 (3.2)	0	0	15 (2.7)	2.2
A*3601	4 (2.6)	13 (3.7)	0	0	17 (3.1)	2.2
A*6601	7 (4.5)	16 (4.6)	2 (7.7)	1 (5.0)	26 (4.7)	6.0
A*6602	1 (0.6)	3 (0.9)	0	0	4 (0.7)	1.1
A*6603	0	0	2 (7.7)	2 (10.0)	2 (0.4)	–
A*6801	2 (1.3)	11 (3.2)	1 (3.8)	1 (5.0)	15 (2.7)	0.6
A*6802	7 (4.5)	11 (3.2)	1 (3.8)	1 (5.0)	20 (3.6)	2.8
A*7401/02	7 (4.5)	29 (8.3)	1 (3.8)	0	37 (6.7)	–
A*8001	3 (1.9)	2 (0.6)	0	0	5 (0.9)	–

ambiguous results that were due to differences in these exons. Allele frequencies were calculated using direct gene counting. Two- and three-locus haplotype frequencies and linkage disequilibrium were calculated on the basis of most likelihood estimation using the HAPL_E software package developed in the Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD, USA.

Results

The goal of this study was to determine the level of diversity at the classical *HLA* class I loci in distinct populations of Cameroon. These indigenous populations represent three major non-Pygmy ethnic groups of southern Cameroon and one Pygmy population (Baka). The study population is made up of 63% Betis, 28% Bamilekes, 4.7% Sawas and 4.3% Bakas.

HLA alleles

Of the 274 samples typed, 28 *HLA-A* alleles were identified (Table 1a), of which A*2301 (AF = 12.8%) and A*02

group of alleles (combined allele frequency, AF = 18.9%) were the most common. Eight rare *HLA-A* alleles were also observed (rare alleles refer to those with AF less than 1% in the study population), and overall, the observed heterozygosity at the A locus was 0.89. The *HLA-B* locus showed the greatest allelic diversity with 41 alleles (Table 1b), of which B*5802 (AF = 10.9%) was the most frequent single allele and B*58 (AF = 15.8%) the most common serologically defined *HLA-B* group of alleles. Sixteen rare *HLA-B* alleles were identified, and the observed heterozygosity at the B locus in this population was 0.92. As in other populations, *HLA-C* showed the least amount of diversity with 21 alleles observed (Table 1c), of which Cw*0401 (AF = 16.6%) was most frequent and Cw*07 the most frequent allele family (AF = 22.6%). Only two rare alleles were found at the C locus, and there was an observed heterozygosity of 0.89. Four individuals (1.5%) were homozygous at all three loci. Overall, 40 intermediate frequency alleles (defined as AF = 2–15% in the study population) were observed, of which 15, 14 and 11 were *HLA-A*, *HLA-B* and *HLA-C*

Table 1b Frequency of HLA-B alleles in the different ethnic groups

HLA-B allele	Major ethnic group				Combined [<i>n</i> (%)] (<i>N</i> = 548)	Frequency (%) (Ellis <i>et al.</i> 2003)(<i>N</i> = 184)
	Bamileke [<i>n</i> (%)] (<i>N</i> = 154)	Beti [<i>n</i> (%)] (<i>N</i> = 348)	Sawa [<i>n</i> (%)] (<i>N</i> = 26)	Baka [<i>n</i> (%)] (<i>N</i> = 20)		
<i>B*0702</i>	11 (7.1)	30 (8.6)	0	0	41 (7.5)	–
<i>B*0705/06</i>	2 (1.3)	1 (0.3)	0	0	3 (0.5)	–
<i>B*0801/19N</i>	1 (0.6)	4 (1.1)	1 (3.8)	1 (5.0)	7 (1.3)	5.4
<i>B*1302</i>	3 (1.9)	8 (2.3)	0	0	11 (2.0)	2.7
<i>B*1401</i>	2 (1.3)	5 (1.4)	0	0	7 (1.3)	1.1
<i>B*1402</i>	1 (0.6)	8 (2.3)	0	0	9 (1.6)	0.5
<i>B*1403</i>	6 (3.9)	6 (1.7)	0	0	12 (2.2)	1.6
<i>B*1503</i>	7 (4.5)	13 (3.7)	2 (7.7)	2 (10.0)	24 (4.4)	4.9
<i>B*1510</i>	2 (1.3)	10 (2.8)	1 (3.8)	0	13 (2.4)	1.1
<i>B*1516</i>	5 (3.2)	4 (1.1)	0	0	9 (1.6)	1.1
<i>B*1517</i>	0	2 (0.6)	0	0	2 (0.4)	1.1
<i>B*1531</i>	0	2 (0.6)	0	0	2 (0.4)	–
<i>B*1801</i>	2 (1.3)	8 (2.3)	0	0	10 (1.8)	2.7
<i>B*2705</i>	1 (0.6)	1 (0.3)	0	0	2 (0.4)	–
<i>B*3501</i>	6 (3.9)	38 (10.9)	3 (11.5)	0	47 (8.6)	7.1
<i>B*3701</i>	0	0	2 (7.7)	3 (15.0)	5 (0.9)	0.5
<i>B*3902</i>	0	1 (0.3)	0	0	1 (0.2)	–
<i>B*3910</i>	1 (0.6)	5 (1.4)	0	0	6 (1.1)	0.5
<i>B*4001</i>	0	1 (0.3)	0	0	1 (0.2)	–
<i>B*4016</i>	1 (0.6)	3 (0.9)	0	0	4 (0.7)	1.1
<i>B*4101</i>	0	0	4 (15.4)	4 (20.0)	8 (1.4)	0.5
<i>B*4102</i>	0	0	1 (3.8)	1 (5.0)	2 (0.4)	–
<i>B*4201</i>	6 (3.9)	14 (4.0)	1 (3.8)	0	21 (3.8)	4.9
<i>B*4202</i>	2 (1.3)	1 (0.3)	0	0	3 (0.5)	–
<i>B*4402</i>	0	4 (1.1)	0	0	4 (0.7)	–
<i>B*4403</i>	13 (8.4)	38 (10.9)	1 (3.8)	1 (5.0)	53 (9.7)	–
<i>B*4407</i>	2 (1.3)	6 (1.7)	0	0	8 (1.4)	2.2
<i>B*4501/07</i>	10 (6.5)	16 (6.4)	1 (3.8)	1 (5.0)	28 (5.1)	3.3
<i>B*4701</i>	4 (2.6)	0	0	0	4 (0.7)	1.1
<i>B*4703</i>	3 (1.9)	4 (1.1)	0	0	7 (1.3)	0.5
<i>B*4901</i>	3 (1.9)	11 (3.2)	0	0	14 (2.5)	5.4
<i>B*5001</i>	1 (0.6)	1 (0.3)	0	0	2 (0.4)	0.5
<i>B*5101</i>	5 (3.2)	5 (1.4)	0	0	10 (1.8)	–
<i>B*5301</i>	13 (8.4)	34 (9.8)	3 (11.5)	2 (10.0)	52 (9.5)	10.9
<i>B*5703</i>	5 (3.2)	8 (3.2)	0	0	13 (2.4)	2.7
<i>B*5801/10N</i>	8 (5.2)	13 (3.7)	3 (11.5)	3 (15.0)	27 (4.9)	5.4
<i>B*5802</i>	22 (14.3)	34 (9.8)	2 (7.7)	2 (10.0)	60 (10.9)	10.9
<i>B*7301</i>	1 (0.6)	4 (1.1)	0	0	5 (0.9)	–
<i>B*7801</i>	0	2 (0.6)	0	0	2 (0.4)	–
<i>B*8101/02</i>	2 (1.3)	6 (1.7)	0	0	8 (1.4)	4.4
<i>B*8201</i>	0	1 (0.3)	0	0	1 (0.2)	–

alleles, respectively. Among the Bakas, 11, 10 and 7 alleles of *HLA-A*, *HLA-B* and *HLA-C*, respectively, were identified, of which *A*3303* (AF = 25%), *B*4101* (AF = 20%), *Cw*0202* (AF = 25%) and *Cw*0701/06/18* (AF = 25%) were most frequent.

HLA haplotypes

Forty-nine *HLA-A-B* (Table 2a), 28 *HLA-B-C* (Table 2b) and 20 *HLA-A-C-B* haplotypes (Table 2c), with

frequencies greater than 1.0%, were identified in the study population. At the four-digit specificity level, *A*3002-B*0702* [haplotype frequency (HF) = 5.1%], *A*6801-B*3501* (HF = 5.1%) and *A*0201-B*3501* (HF = 5.1%) were the most frequently identified *HLA-A-B* haplotypes (Table 2a), while *B*5802-Cw*0602* (HF = 10.8%), *B*5301-Cw*0401* (HF = 8.1%) and *B*0702-Cw*0702* (HF = 6.4%) were the three most common *HLA-B-C* haplotypes observed (Table 2b). A three-locus haplotype analysis of *HLA-A-C-B* showed *A*6601-*

Table 1c Frequency of HLA-C alleles in the different ethnic groups

HLA-C allele	Major ethnic group				Combined [<i>n</i> (%)] (<i>N</i> = 548)
	Bamileke [<i>n</i> (%)] (<i>N</i> = 154)	Beti [<i>n</i> (%)] (<i>N</i> = 348)	Sawa [<i>n</i> (%)] (<i>N</i> = 26)	Baka [<i>n</i> (%)] (<i>N</i> = 20)	
<i>Cw</i> *0202	17 (11.0)	16 (4.6)	5 (19.2)	5 (25.0)	43 (7.8)
<i>Cw</i> *0302	2 (1.3)	8 (2.3)	0	0	10 (1.8)
<i>Cw</i> *0304	2 (1.3)	11 (3.2)	1 (3.8)	0	14 (2.5)
<i>Cw</i> *0401	20 (13.0)	65 (18.7)	4 (15.4)	2 (10.0)	91 (16.6)
<i>Cw</i> *0407	4 (2.6)	5 (1.4)	0	0	9 (1.6)
<i>Cw</i> *0501	0	8 (2.3)	0	0	8 (1.4)
<i>Cw</i> *0602	27 (17.5)	44 (12.6)	4 (15.4)	3 (15.0)	78 (14.2)
<i>Cw</i> *0701/06/18	20 (13.0)	45 (12.9)	5 (19.2)	5 (25.0)	75 (13.7)
<i>Cw</i> *0702	11 (7.1)	23 (6.6)	3 (11.5)	3 (15.0)	40 (7.3)
<i>Cw</i> *0704/11	0	1 (0.3)	0	0	1 (0.2)
<i>Cw</i> *0705	0	8 (2.3)	0	0	8 (1.4)
<i>Cw</i> *0802	11 (7.1)	28 (8.0)	0	0	39 (7.1)
<i>Cw</i> *0804	0	6 (1.7)	0	0	6 (1.1)
<i>Cw</i> *1203	2 (1.3)	6 (1.7)	0	0	8 (1.4)
<i>Cw</i> *1402	6 (3.9)	4 (1.1)	0	0	10 (1.8)
<i>Cw</i> *1403	9 (5.8)	17 (4.9)	1 (3.8)	1 (5.0)	28 (5.1)
<i>Cw</i> *1505	1 (0.6)	8 (2.3)	0	0	9 (1.6)
<i>Cw</i> *1601	7 (4.5)	21 (6.0)	1 (3.8)	0	29 (5.3)
<i>Cw</i> *1604	0	1 (0.3)	0	0	1 (0.2)
<i>Cw</i> *1701/02/03	8 (5.2)	14 (4.0)	2 (7.7)	1 (5.0)	25 (4.6)
<i>Cw</i> *1801/02	7 (4.5)	9 (2.6)	0	0	16 (2.9)

*Cw**0602-*B**5802 to be the most frequent haplotype (HF = 3.3%) (Table 2c).

Rare HLA alleles and haplotypes

Globally uncommon HLA class I alleles (alleles that are uncommon across all populations) identified in the study population include *HLA-A**2612 (AF = 2.0%), *HLA-B**4016 (AF = 0.7%) and *HLA-B**4407 (AF = 1.4%), none of which were identified in the Baka pygmy population. Overall, 26 rare alleles were identified in this study population: *HLA-A* = 8, *HLA-B* = 16 and *HLA-C* = 2. *A**2612-*Cw**0701/06/18-*B**4407 (HF = 1.3%) and *A**0202-*Cw**1402-*B**1516 (HF = 1.3%) (Table 2c) were observed in this population, haplotypes that are uncommon worldwide and have hitherto not been reported in Cameroon. These results also indicate that Cameroon alleles *A**2612 and *B**4407 are found on the same haplotype.

Discussion

We examined the allele frequency of the HLA class I classical loci in rural Cameroon populations using high-resolution genotyping methods. HLA class I molecules are constitutively expressed by all cell types, HLA-C being expressed at lower levels than HLA-A and HLA-B (12). Twenty-eight *HLA-A*, 41 *HLA-B* and 21 *HLA-C* alleles

were identified with frequencies ranging from <1 to 16.4%. The higher number of *HLA-B* alleles is in accordance with the more rapid evolution of the *HLA-B* locus relative to *HLA-A* alleles (13–15), the significance of which remains unknown (10). Twenty-nine percent of the alleles identified in this study were rare alleles, suggesting dynamic changes in the HLA repertoire in this population perhaps due to pathogen-mediated selection and/or the introduction of new alleles by admixture. Of the 26 rare alleles identified, 16 were of *HLA-B*. Although the HLA class I alleles show similar roles in pathogen control, it is possible that each class I locus contributes differently to disease outcome. For instance, HIV-specific CD8⁺ T-cell activity has been shown among southern Africans to be strongly influenced by *HLA-B* genes (10). In this study, *HLA-B* showed the highest heterozygosity (0.92). Among African-Americans, one of the five major ethnic groups in the United States, Cao and colleagues reported levels of heterozygosity of *HLA-A*, *HLA-B* and *HLA-C* loci greater than 0.90 (16). A comparison of *HLA-A*, *HLA-B* and *HLA-C* allelic distribution among sub-Saharan African populations reveals a remarkable level of diversity (17). For example, from Zambia in East Africa, *A**3002 (AF = 23.3%) was the most frequently identified allele (17) compared with a frequency of 8.6% observed in this study. *A**2301 was most frequently observed in this study (AF = 12.8%) as well as that of Ellis *et al.* (AF = 18.7%) (18).

Table 2a Frequency of predicted *HLA-A-B* haplotypes

A*-B*	Frequency (%)
3002-0702	5.1382
6801-3501	5.1282
0201-3501	5.1282
6601-5802	3.8462
3001-4201	3.8462
0301-4403	3.8462
0205-5801/10N	2.5641
2612-4407	2.5641
3002-4403	2.5641
7401/02-4501/07	2.5641
6802-1302	2.5641
3002-3501	2.5641
2301/07N-5802	2.5641
2301/07N-4901	2.5641
2301/07N-1503	2.5641
0201-1503	2.5641
3601-5301	2.5641
0201-5301	2.5641
0202-4403	2.5641
2301/07N-5801/10N	1.2821
3601-1302	1.2821
2301/07N-0702	1.2821
0202-1510	1.2821
3002-3910	1.2821
2501-5802	1.2821
7401/02-3501	1.2821
2902-7301	1.2821
2902-4201	1.2821
3001-4101	1.2821
3201-0702	1.2821
2301/07N-3501	1.2821
2902-3501	1.2821
2902-4403	1.2821
2301/07N-4403	1.2821
2612-4703	1.2821
7401/02-5301	1.2821
8001-1801	1.2821
7401/02-1510	1.2821
3303-4201	1.2821
2902-5301	1.2821
3201-4901	1.2821
6802-1503	1.2821
0202-5301	1.2821
56601-4101	1.2821
8001-5801/10N	1.2821
2902-8101/02	1.2821
3004-3501	1.2821
3004-5101	1.2821
0101-4403	1.2821
6802-3501	1.2821

The only other study reporting *HLA* class I diversity in Cameroon among unrelated Bantu-speaking individuals was carried out by Ellis and colleagues. The authors identified *A*03012*, *A*2612*, *A*3006*, *B*1403*, *B*4016* and *B*4703* as 'novel' alleles (18). We also identified *A*2612*

Table 2b Frequency of predicted *HLA-B-Cw* haplotypes

B*-Cw*	Frequency
5802-0602	0.1077
5301-0401	0.0807
0702-0702	0.0638
4403-1403	0.0474
3501-0401	0.0470
4201-1701/02/03	0.0383
1503-0202	0.0383
4901-0701/06/18	0.0255
4403-0401	0.0229
4403-0701/06/18	0.0227
1510-0304	0.0219
1403-0802	0.0219
5801-0701/06/18	0.0210
4501/07-1601	0.0182
1516-1402	0.0164
1402-0802	0.0164
5101-1601	0.0164
4501/07-0407	0.0146
4407-0701/06/18	0.0146
3501-0705	0.0146
4501/07-0602	0.0146
4101-0701/06/18	0.0146
5801-0302	0.0128
5703-1801/02	0.0128
4703-0701/06/18	0.0128
3501-0802	0.0124
1801-0501	0.0109
3910-1203	0.0109

Table 2c Frequency of predicted *HLA-A-C-B* haplotypes

A*-Cw*-B*	Frequency
6601-0602-5802	0.0328
0201-0401-5301	0.0281
3001-1701/02/03-4201	0.0237
0301-1403-4403	0.0235
3601-0401-5301	0.0180
0202-0602-5802	0.0175
2301/07N-0702-0702	0.0172
2902-0701/06/18-4901	0.0162
3402-0701/06/18-4403	0.0146
2612-0701/06/18-4407	0.0128
0202-1402-1516	0.0128
0301-0602-5802	0.0128
7401/02-0401-3501	0.0128
7401/02-0202-1503	0.0127
6802-0702-0702	0.0125
0202-0702-0702	0.0116
2301/07N-0407-4501	0.0109
6801-0802-3501	0.0109
3002-0401-5301	0.0109
0201-1601-5101	0.0103

(AF = 2.0%) and *B*4016* (AF = 0.7%) along with 25 additional rare *HLA* alleles (Table 1a,b). Bruges Armas *et al.* (19) studied *HLA* class I variation in West African pygmies, including the Bakola from Cameroon. They identified two alleles (*B*37* and *B*41*) and three haplotypes (*A*30-B*37*, *A*66-B*41* and *A*68-B*58*) that appeared to be typical of western Pygmies. In this study, *B*37* was identified in the Bakas (AF = 15%), as well as the non-Pygmy Sawa population (AF = 7.7%). Similarly, *B*41* was found in the Bakas (AF = 25%) and in the non-Pygmy Sawas (AF = 19.2%). The overall combined frequency in the study population was 0.9% for *B*37* and 1.8% for *B*41*. Of the three Pygmy haplotypes observed by Bruges Armas *et al.* (19), only *A*66-B*41* was observed at a frequency of 1.3% in the total study population. *HLA-A*8001*, a divergent allele which may be characteristic of populations with African admixture (20) was identified in five individuals (AF = 0.9%). Of the 16 rare *HLA-B* alleles identified in this study, two of these (*B*3701* and *B*4001*) are more commonly found in Caucasians than Blacks (20). *B*3910*, which has been reported only in Black populations (20) (<http://www.allelefrequency.net>), was also identified in this study with an allele frequency of 1.1%. *B*7301*, a rare allele in the world population and also a structurally divergent allele (21), showed a frequency of 0.9% in this study. Among Blacks, Caucasoid and Asian populations, an allele frequency <0.1% for *B*7301* has been reported (21). Fourteen *Cw*12* alleles have been officially assigned (22), of which only *Cw*1203* (AF = 1.4%) was observed in this study, an allele that has also been identified in the Bubi tribe from West Africa (23). Of the three most commonly identified allotypes in this study population, *B*58* (AF = 15.9%) has been reported to have the highest frequency in the Black population (20). *Cw*0706* and *Cw*1802* are unusual alleles which have been reported from the Bubi tribe in West Africa (24). The typing algorithm used in this study did not allow us to discriminate between *Cw*0701/06/18* and *Cw*1801/1802*; therefore, we were unable to determine whether or not *Cw*0706* and *Cw*1802* were present in this study population. Multiple intermediate frequency haplotypes are found in populations of Mali, Uganda, Kenya and Zambia (17) that were not identified in this study group except for *A*3601-Cw*0401-B*5301* (HF = 1.75%). The remarkable *HLA* genetic diversity and distribution observed in sub-Saharan African populations compared with Europeans may provide clues to the understanding of infectious disease patterns reported from different regions of the world due to migration and evolution (25,26).

The PCR/SSOP technique used to identify *HLA* alleles in this study utilizes primers and probes specific for exons 2 and 3. Thus, there could be polymorphisms in other exons that may have been undetected in this study and

that are relevant to the understanding of their peptide-binding and peptide-presenting functions (27). Identification of polymorphisms in other exons might also lead to the identification of novel alleles (27). *HLA* class I molecules can be classified on the basis of their functional binding specificities into nine supertypes designated A1, A2, A3, A24, B7, B27, B44, B58 and B62 (28). The different motifs and ligand-binding specificities of the alleles that constitute the supertypes may modulate the outcome of disease (29) by providing a broad immune response in recognition of a foreign peptide (28,30). *HLA* B7 supertype was predominant in our study population, and a similar finding has been reported among American Blacks (28). Although the overall population structure (i.e., the ethnic make-up) of the communities where the subjects were recruited is not precisely known, these data are undoubtedly relevant in understanding MHC polymorphism in humans and also a useful tool in designing effective T-cell-based vaccines.

Acknowledgments

We thank the Government of Cameroon for permission to undertake this study. This work was supported by an award from the U.S. Military HIV Research Program (to Donald Burke) and in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-12400. We thank the Walter Reed Johns Hopkins Cameroon Program staff and the U.S. Embassy in Yaounde for assistance in carrying out this work.

Disclaimer: The views and opinions expressed here do not necessarily reflect those of the US Army or the Department of Defense.

References

1. Bodmer JG, Marsh SG, Albert ED *et al.* Nomenclature for factors of the HLA System, 1996. *Hum Immunol* 1997; **53**: 98–128.
2. Macdonald WA, Purcell AW, Mifsud NA *et al.* A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition. *J Exp Med* 2003; **198**: 679–91.
3. Watkins DI, Hodi FS, Letvin NL. A primate species with limited major histocompatibility complex class I polymorphism. *Proc Natl Acad Sci USA* 1988; **85**: 7714–8.
4. Chen YS, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet* 1995; **57**: 133–49.
5. Jorde LB, Bamshad MJ, Watkins WS *et al.* Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. *Am J Hum Genet* 1995; **57**: 523–38.

6. Watson E, Forster P, Richards M, Bandelt HJ. Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet* 1997; **61**: 691–704.
7. Stephens JC, Schneider JA, Tanguay DA et al. Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 2001; **293**: 489–93.
8. Spedini G, Destro-Bisol G, Mondovi S, Kaptue L, Taglioli L, Paoli G. The peopling of sub-Saharan Africa: the case study of Cameroon. *Am J Phys Anthropol* 1999; **110**: 143–62.
9. Carr JK, Torimiro JN, Wolfe ND et al. The AG recombinant IbNG and novel strains of group M HIV-1 are common in Cameroon. *Virology* 2001; **286**: 168–81.
10. Kiepiela P, Leslie AJ, Honeyborne I et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* 2004; **432**: 769–75.
11. Gao X, Nelson GW, Karacki P et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N Engl J Med* 2001; **344**: 1668–75.
12. Snary D, Barnstable CJ, Bodmer WF, Crumpton MJ. Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur J Immunol* 1977; **7**: 580–5.
13. Belich MP, Madrigal JA, Hildebrand WH et al. Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature* 1992; **357**: 326–9.
14. Watkins DI, McAdam SN, Liu X et al. New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature* 1992; **357**: 329–33.
15. McAdam SN, Boyson JE, Liu X et al. A uniquely high level of recombination at the HLA-B locus. *Proc Natl Acad Sci USA* 1994; **91**: 5893–7.
16. Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernandez-Vina MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* 2001; **62**: 1009–30.
17. Cao K, Moormann AM, Lyke KE et al. Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. *Tissue Antigens* 2004; **63**: 293–325.
18. Ellis JM, Mack SJ, Leke RF, Quakyi I, Johnson AH, Hurley CK. Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, *2612, *3006 and HLA-B*1403, *4016, *4703. *Tissue Antigens* 2000; **56**: 291–302.
19. Bruges Armas J, Destro-Bisol G, Lopez-Vazquez A et al. HLA class I variation in the West African Pygmies and their genetic relationship with other African populations. *Tissue Antigens* 2003; **62**: 233–42.
20. Marsh SGE, Barber LD, Parham P. *The HLA Factsbook*. San Diego, California, London: Academic, 2000.
21. Parham P. Functional polymorphism in class I major histocompatibility complex genes. *Soc Gen Physiol Ser* 1994; **49**: 93–103.
22. Marsh SG, Albert ED, Bodmer WF et al. Nomenclature for factors of the HLA system, 2004. *Tissue Antigens* 2005; **65**: 301–69.
23. Vilches C, Bunce M, van Dam M, de Pablo R. A new pair of HLA-C alleles, Cw*12042 and Cw*1203, differing at the KIR-related dimorphism of codons 77–80. *Tissue Antigens* 1998; **51**: 101–5.
24. van der Vlies SA, Voorter CE, de Pablo R, Vilches C, van den Berg-Loonen EM. Strong association between HLA-Cw*0706 and HLA-B*44032 in the Bubi population from Equatorial Guinea. *Tissue Antigens* 2000; **55**: 57–60.
25. Hughes AL, Hughes MK. Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics* 1995; **42**: 233–43.
26. Hughes AL, Yeager M, Carrington M. Peptide binding function and the paradox of HLA disease associations. *Immunol Cell Biol* 1996; **74**: 444–8.
27. van der Vlies SA, Voorter CE, van den Berg-Loonen EM. There is more to HLA-C than exons 2 and 3: sequencing exons 1, 4 and 5. *Tissue Antigens* 1999; **54**: 169–77.
28. Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* 1999; **50**: 201–12.
29. Swelsen WT, Voorter CE, van den Berg-Loonen EM. Polymorphism of intron 4 in HLA-A-B and -C genes. *Tissue Antigens* 2003; **61**: 475–83.
30. Burrows SR, Elkington RA, Miles JJ et al. Promiscuous CTL recognition of viral epitopes on multiple human leukocyte antigens: biological validation of the proposed HLA A24 supertype. *J Immunol* 2003; **171**: 1407–12.