HLA DRB1, DQA1, DQB1 and DPB1 polymorphisms in Namibian Khoi and San and in the Xhosa and South African mixed-ancestry populations

Jeanne Rousseau[a*], Robert C. Williams[b] and Pauline Creemers[a]

We have analysed the HLA allele distributions in unrelated Namibian Khoi and San and South African Xhosa and mixed-ancestry (so-called Cape coloured) populations. The allelic specificities of the DRB1, DQA1, DQB1 and DPB1 loci were determined. We found loss of allelic diversity and predominance of certain alleles to be more pronounced in the San than in the Khoi. By contrast, the Xhosa indicated wide allelic diversity especially at the DRB1 and DPB1 loci. We found evidence that the Caucasoids may have been derived from an early migration wave, whereas African blacks arose from a later migration wave of an ancestral population pre-dating ethnic diversity. Frequencies in the Xhosa for the DRB1*0302, *1101, *1302 and *1304 alleles revealed clinal variation in a north–south direction across the African continent. For the DRB1, DQA1, DQB1 haplotypes there was greater variation in the Khoi than in the San, in whom certain haplotypes predominated. We found 13 previously unreported haplotypes. In Xhosa we found 30 different DRB1, DQA1, DQB1 haplotypes. Allele and haplotype characteristics and frequencies in the South African mixed-ancestry population were mostly intermediate between those found in Xhosa and reported for Caucasoids, and three possible ancestral southern African DRB1, DQB1, DPB1 haplotypes were found. We discuss the implications of our findings for organ donor transplantation in Xhosa and in the South African mixed-ancestry population.

The human leukocyte antigen (HLA) system, located within the major histocompatibility complex (MHC) on chromosome 6, is highly polymorphic and represents one of the most intensely studied regions of the human genome.[1] HLA molecules bind peptide antigens, which they present to T cells and thus shape the T-cell repertoire. HLA antigens play a key role in allograft rejection and disease susceptibility. HLA-DRB1, -DQA1 and -DQB1 alleles are inherited together in closely associated haplotypes, which are found repeatedly and in characteristic patterns in different populations. This linkage disequilibrium between the HLA genes has been used to characterize population differentiation.[2]

The present-day Namibian San and Khoi represent original southern African populations and have currently been reduced to very small numbers consisting of communities of approximately 50–100 people. On the basis of linguistic, cultural and genetic studies, the Khoi and San are thought to share a common ancestry and to be direct descendants of an early humanoid African population from which all other ethnic populations are derived.[3] This founding population is believed to have lived around 100 000 years ago, at which time they diverged into different ethnic groups. The Khoi and San are thought to have diverged at least 10 000 years ago. African blacks, originating from West Africa, are postulated to represent an early migratory wave of the founding population.[3–5] Xhosa are derived from this population; they migrated to southeast Africa during the past 1000 years. Some later admixture with the San is also assumed to have taken place.[3–5] The C4, BF (properdin factor B) and DPB (vitamin D binding protein) polymorphisms and HLA associations[6,7] have confirmed that: 1) the Khoi and San are closely related; 2) the Khoi share some characteristics with Caucasoids that are not shared with Xhosa; and 3) the combined Khoisan group shares more similarities with Xhosa than with Caucasoids. The founding populations of the South African mixed-ancestry (so-called Cape coloured) population are the Khoi and western Europeans, with an admixture of Xhosa and southeast Asian genes.[7]

Using DNA typing techniques, HLA alleles have been best characterized in Caucasoids,[8,9] Gambians and Malawians,[10] Zimbabweans[11] and African-Americans.[12] Gambians and African-Americans showed a far greater allelic diversity than Caucasoids. Studies in South American Indians showed a similar trend to allelic diversification.[13,14] Many new alleles may have been generated recently[15] in these populations, which arose while the people lived in tropical rain-forests for extended periods. These tropical conditions, with abundant vegetation and animal life, which promote the presence of a large diversity of viral, bacterial and parasitological pathogens, could favour the retention of allelic diversity within the HLA system. By contrast, the habitat of the Khoi has been limited to southwestern Africa, encompassing a semi-desert and Mediterranean climate, while the habitat of the San was limited to the harsh Kalahari semi-desert and desert.[1–3,5–7]

Recently, studies have been conducted on isolated populations living in dry, harsh climates. A study in five different Siberian populations showed limited variability at the allelic level, but selected HLA alleles were present in very high frequencies. However, there was high variability at the DRB1, DQA1, DQB1 haplotype level.[16] In the Chinese Kazak and Han populations, who live in northwestern China in the vicinity of the Gobi desert, the same phenomenon occurred. The DRB1, DQA1, DQB1 haplotype repertoire was limited while other haplotypes tended to predominate.[17]

Information on the molecular biology of the highly polymorphic HLA system in the Khoi, San and Xhosa may reveal clues about the divergence of modern *Homo sapiens* into the various ethnic groups. We report on the HLA allelic specificities and frequencies, as well as the DRB1, DQA1, DQB1 and DRB1, DQB1, DPB1 haplotype frequencies in Namibian Khoi and San, and in the South African Xhosa and the mixed-ancestry populations.

Materials and methods

Populations. The groups of Namibian San (Bushman: Bushman-speaking) and Khoi (Khoi-khoi: Hottentot-speaking) have been described earlier.[5–7,18] The San and Khoi communities

[a]Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town, Observatory 7925, South Africa.
[b]Department of Anthropology, Arizona State University, Tempe, Arizona, U.S.A.
*Author for correspondence. E-mail: jeanne@chempath.uct.ac.za

from which the samples were collected were very small (fewer than 100 individuals) and inbreeding could not be excluded. Individuals were judged to be unrelated on the basis of interviews. The Xhosa [Nguni (Bantu)-speaking African Negroes][18] were mainly recent immigrants from the Eastern Cape region of South Africa. The South African mixed-ancestry population group, which arose from an admixture of western European, Khoisan, southeast Asian and African black genes, was from the greater Cape Town area.

DNA typing. HLA-DNA typing was performed using the PCR-SSOP method. Briefly, DNA, extracted from blood specimens, was amplified by the polymerase chain reaction (PCR) with DRB1, DQA1, DQB1 and DPB1 group-specific primer pairs. Amplified samples were dot-blotted onto nylon membranes and hybridized with sequence-specific oligonucleotide probes (SSOPs) obtained from the XIth International HLA workshop. Alleles were assigned according to the reaction patterns and interpreted by two independent observers.[19,20]

The DRB1, DQA1, DQB1 and the DRB1, DQB1, DPB1 haplotypes were studied independently: the loci of the DRB1, DQA1, DQB1 haplotype are in close proximity, resulting in stable haplotypes which may be used for population comparisons. For successful bone marrow transplantation, the specificities of these loci must be identical. The relatively large distance between the DQB1 and DPB1 loci results in a greater recombination frequency between them and less stable DRB1, DQB1, DPB1 haplotypes across populations. Successful bone marrow transplantation is possible in the presence of a specificity difference in the DPB1 locus. We have therefore focused on the common haplotypes.

Statistical analysis. Allele frequencies were calculated by gene counting, while the variance and 95% confidence interval were calculated by standard methods.[21] Haplotype frequencies and their genetic linkage disequilibria were computed by a modified EM algorithm after the method of Long *et al*.[21] Tests for the significance of the disequilibria between loci were performed by the simulation of 1000 samples using the allele frequency vectors at each locus as the probability for drawing a random allele. Two random alleles were selected for each of the four loci for n phenotypes, where n represents the sample size of the observed population. The EM algorithm was applied to each of the 1000 samples and likelihoods calculated for all sets of disequilibria. The likelihood was calculated as

$$\ln L_k = \sum_{i=1}^{n} \ln \Pr(P_i),$$

where 1n Pr(P_i) is the natural logarithm of the probability of the ith person's phenotype in the kth random sample of size n (ref. 21). Means and 95% confidence intervals for the likelihoods were calculated. The confidence intervals represent the 25th and 975th observations of the ordered set of 1000 likelihoods for each combination of disequilibrium. D is a measure of strength of association of alleles of different loci.

Wright's[22] fixation index, or inbreeding coefficient, F, is calculated as

$$F = 1.0 - \frac{y}{y_0},$$

where y is the observed proportion of heterozygotes at a locus and y_0 is the proportion expected under Hardy-Weinberg equilibrium when there are k alleles at a locus and p_i and p_j are the frequencies of the alleles in the heterozygote,

$$y_0 = \sum_{i<j}^{k} 2\ p_i\ p_j.$$

Tests for Hardy-Weinberg equilibrium at a locus were performed by the method of Nam and Gart.[21]

Results

Namibian Khoi and San

HLA allele specificities

Statistical analysis. For the Khoi, the observed heterozygosities ranged from 0.756 for the DQA1 and DPB1 loci to 0.837 for the DQB1 locus. In each case, the expected heterozygosity at Hardy-Weinberg equilibrium was not significantly different from the observed values. At the DQB1 locus the Nam-Gart T statistic was significantly different from 1.0:1.832; 95% confidence interval: 1.163–2.501 (results not shown). The San population had observed heterozygosity ranging from 0.493 for the DQA1 locus to 0.783 for DPB1 locus. At the DRB1 locus the expected heterozygosity, H = 0.772, while not significantly different from the observed value, was accompanied by a large, positive value for Wright's F, 0.118, and an estimate for the Nam-Gart T statistic that is significantly different from 1.0:2.501; 95% confidence interval: 1.651–3.352.

DRB1 locus. The allele frequencies for the DRB1, DQA1, DQB1 and DPB1 alleles are listed in Table 1. For comparison of the allele specificities and frequencies for the DRB1 locus, we used European Caucasoids,[8,9] Gambians and Malawians,[10] and Zimbabweans.[11] In the Khoi, 21 different DRB1 alleles were found and 14 in the San. There were some characteristic differences between the Khoi and San. The DRB1*0101 allele, which is found in populations across the world, was present in the Khoi only. The DRB1*0102 allele, which has been reported in Caucasoids, was present at a frequency of 5.2% in the Khoi, and was not found in the San. The DRB1*0301 allele, which is present in Khoi, Caucasoids and Gambians in a rather high frequency, occurred at a significantly lower frequency in San. The DRB1*0302 allele was only sporadically found in both the Khoi and San. This allele has not been reported in Caucasoids, but has been found in African blacks.[10,11] An interesting finding was that the DRB1*0401 allele occurred at a frequency of 40.9% in San, while in Khoi it was only 13.5%, which is closer to the frequencies found in Caucasoids and Gambians. In the San we also found a DRB1*04 allele, which could not be sub-typed. The DRB1*0701 allele, found in Gambians[10] but not in Caucasoids,[8,9] was found in both the Khoi and the San at relatively high frequencies of 4.1 and 6.3%, respectively. The DRB1*0804 allele occurred in both Khoi and San but the DRB1*0901 and DRB1*1101 alleles, which are frequent in other populations, were found only in the San. The DRB1*1001 specificity was found only in Khoi. The rare DRB1*1202 allele, which was found only in Khoi, has been reported in Gypsies[8] but not in Caucasoids [9] or African blacks.[10,11] The DRB1*1301 allele was present at high frequency in both Khoi (7.0%) and San (12.3%). The DRB1*1302 allele frequency was high in both Khoi and San, with frequencies of 11.5% and 11.6%, respectively. This is a characteristic shared with the Japanese.[23] The DRB1*1501 allele had a high frequency of 30.0% in Khoi, but was present in only 1.5% of San. This allele is common in other ethnic groups. The DRB1*1601 allele, which is common in Caucasoids [8,9] but absent in blacks,[10,11] was found only in Khoi.

DQA1 locus. The frequencies for the alleles of the DQA1 locus in the Khoi and San are listed in Table 1. In the Khoi the overall distribution of alleles was similar to that found in African Americans.[12] However, the frequencies of the DQA1*0102 and*0301 alleles were relatively high (42.6% and 21.4%, respectively), whereas that of the DQA1*0401 allele was low. The

Table 1. DRB1, DQA1, DQB1 and DPB1 allele frequencies in Namibian Khoi and San.

	Frequency			Frequency	
Allele	Khoi	San	Allele	Khoi	San
DRB1*			**DQB1***		
0101	0.0174	–	0201	0.1105	0.0841
0102	0.0523[S]	–	0302	**0.1444**	**0.4970**[S]
0103	0.0581	0.1159	0402	0.0553	0.0580
0302	0.0116	0.0072	0501	0.1047[S]	–
04	–	0.0625[S]	0502	0.0174	–
0401	**0.1345**	**0.4088**[S]	0601	0.0058	–
0404	0.0336	0.0441	0602	**0.2725**[S]	0.1304
0405	0.0058	–	0603	0.0407	0.0290
0701	0.0407	0.0625	0604	0.0174	**0.1079**[S]
0804	0.0291	0.0362	0605	**0.1156**[S]	0.0290
0901	–	0.0145	Blank	0.0477	0.0211
1001	0.0174	–			
11	0.0058	–	**DPB1***		
1101	–	0.0145	0101	**0.1752**	**0.2349**
1102	0.0174	0.0217	0201	**0.1636**[S]	0.0290
1104	0.0058	–	0301	0.0568	0.0362
1202	0.0116	–	0401	**0.1142**	**0.2923**[S]
1301	0.0698	**0.1232**	0402	**0.2600**[S]	**0.1159**
1302	**0.1154**	**0.1159**	0501	0.0058	–
1401	0.0116	0.0072	0801	0.0058	0.0217
1501	**0.2997**[S]	0.0145	1101	0.0581[S]	0.0072
1502	0.0116	–	1301	0.0058	–
1601	0.0058	–	1401	0.0233	0.0072
Blank	0.0213	0.0472	1501	0.0174	0.0145
			1801	0.0058	–
DQA1*			1901	0.0058	–
0101	0.1221[s]	0.0145	2201	0.0058	–
0102	**0.4261**[s]	0.1658	2301	0.0058	0.0145
0201	0.0814[S]	0.0199	2701	–	0.0072
03	0.0174	–	3001	0.0116	**0.0797**
0301	**0.2135**	**0.6045**[s]	3401	0.0174	0.0217
0401	0.0233	0.0290	5501	0.0349	**0.1082**
0501	0.1264[S]	0.0362	Blank	0.0267	0.0094
Blank	0.0130	0.0340			

[S]Significantly ($P > 0.0001$) higher in Khoi than in San, or *vice versa*. Figures in bold: high frequencies compared to other populations. Number of individuals tested: DRB1, DQA1, and DQB1: Khoi, 119; San, 109; DPB1: Khoi, 86; San, 69.

Table 2. Three-locus DRB1, DQA1, DQB1 haplotype frequencies in Namibian Khoi and San.

Haplotype			Frequency		[a]Reported
DRB1*	DQA1*	DQB1*	Khoi	San	
0101	0101	0501	0.0174	–	*
0102	0101	0501	**0.0523**	–	*
0301	0501	0201	**0.0523**	–	*
0301	0501	0301	0.0231	0.0072	[b]Xhosa
0301	0501	0605	0.0058	–	No
0302	0401	0402	0.0116	0.0072	AA
04	0301	0302	–	**0.0602**	*
0401	0301	0302	**0.1333**	**0.3906**	*
0404	0301	0402	0.0336	0.0280	No
0405	0301	0302	0.0058		*
0701	0301	0201	0.0407	**0.0530**	Xhosa,[c]CC
0804	0102	0604	–	0.0217	No
0804	0401	0301	–	0.0145	AA
0804	0401	0402	0.0058	–	AA
0804	0501	0301	0.0174	–	AA
0901	0301	0201	–	0.0115	*
1001	0101	0501	0.0174	–	AA
11	0501	0301	0.0058	–	No
1101	0101	0602	–	0.0072	No
1101	0102	0602	–	0.0072	*
1102	0501	0301	0.0174	0.0217	*
1202	0102	0602	0.0058	–	AA
1301	0102	0603	0.0058	–	Japanese
1301	0102	0605	0.0058	–	No
1301	0103	0602	0.0346	**0.0870**	No
1301	0103	0603	0.0178	0.0290	*
1302	0102	0602	–	0.0072	AA
1302	0102	0604	0.0169	**0.0797**	*
1302	0102	0605	**0.0963**	0.0290	*
1401	0101	0602	–	0.0072	AA
1501	0101	0602	0.0101	–	*
1501	0102	0602	**0.2091**	0.0144	*
1501	0102	0603	0.0122	–	CC
1501	03	0201	0.0127	–	No
1502	0101	0501	0.0116	–	No
1601	0102	0502	0.0058	–	*

Number of individuals tested: Khoi, 119; San, 109. [a]Symbols used: *reported in more than one ethnic group; AA, reported only in African-Americans;[12] reported in Japanese only;[23] [b]Xhosa: previously unreported, but described in this investigation in Xhosa (Table 4). [c]CC, previously unreported, but found during this study in the Cape coloured population (Table 6). Bold script indicates unusually high haplotype frequency.

DQA1*0201 and *0601 alleles were not found in either the Khoi or San. The frequency distribution in the San differs considerably from that in the Khoi. In San, the frequency of DQA1*0101 was lower than found in most other populations. The frequency of the DQA1*0102 allele was similar to that found in Caucasoids.[8,9] It appears that the DQA1*0301 allele has selectively assumed the extremely high frequency of 60.5%.

DQB1 locus. The frequency values for the DQB1 alleles are also listed in Table 1. The DQB1*0303 allele was absent in both the San and the Khoi. The DQB1*0502 and 0601 alleles were absent in the San and the Xhosa. The DQB1*0502, but not the DQB1*0601, allele has also been found in African-Americans; it may have been derived from a Caucasoid admixture. In the Khoi, the allele distribution largely resembled that of African-Americans. In the San, there was selective expansion of the DQB1*0302 allele to almost 50%. The frequency of the DQB1*0604 allele was also high (10.8%) compared with other populations.

DPB1 locus. The frequencies for the DPB1 alleles are listed in Table 1. In the Khoi 18 different alleles were found, 14 in the San. The alleles not found in the San were DPB1*0501, *1301, *1801. The last two alleles have commonly been found in other ethnic groups. The DPB1*1901 allele, which has been reported in Caucasoids[9] and gypsies,[8] was not found and neither was the DPB1*2201 allele, which has been found in various other ethnic groups, albeit at very low frequencies. The DPB1*2701 allele was found in the San but not in the Khoi.

There were considerable differences in the allele frequencies of this locus between the Khoi and the San. The DPB1*0201, *0402, and *1101 alleles were found at much higher frequencies in the Khoi than in the San. In the San, the DPB1*0401, *3001 and *5501 alleles seemed to have expanded selectively. The frequency of the DPB1*0101 allele was very high in both the Khoi and the San. The frequencies of the DPB1*0401 and *0402 in the Khoi and San were very high and similar to those found in Caucasoids.[8,9]

HLA haplotype characteristics and frequency

Statistical analysis. Tests for genetic disequilibria demonstrated significant likelihood values for all three locus combinations shown in this article (results not shown).

DRB1, DQA1, DQB1 haplotypes. The DRB1, DQA1, DQB1 haplotypes of the Khoi and San are listed in Table 2. In the Khoi, 29 different haplotypes were found, of which 11 were shared with the San. In the San, 19 different haplotypes were found. Fifteen of the Khoisan haplotypes have been described in more than one ethnic group; eight have been described only in African-Americans[12] and one has been reported only in the Japanese.[23] Thirteen haplotypes were found which had not previously been described. Of these, one was found in Xhosa, one in the South African mixed-ancestry population and one in

both Xhosa and the mixed-ancestry population (see below).

If one defines a public or conserved haplotype as one that has been found in more than one ethnic group, and one considers the Khoisan (Khoi and San) a distinct ethnic group, then the Khoisan haplotypes consist mainly of public haplotypes. Our data show that 73.0% (27/36) of all Khoisan haplotypes occurred in other ethnic groups, and are therefore conserved, or public haplotypes.

Twelve haplotypes were found which, to our knowledge, have not been described previously. Of these, only four had a *D*-value of higher than 0.0100 (the *D*-value is a measure of the strength of association of the alleles of the three different loci). They were DRB1*0404, DQA1*0301, DQB1*0402, $D = 0.0242$, with a frequency in Khoi of 3.4%, and in San of 2.8%; DRB1*0701, DQA1*0301, DQB1*0201, $D = 0.0238$, with a frequency of 3.5% in Khoi and of 5.3% in San. This haplotype was also present in Xhosa and in the South African mixed-ancestry population. The haplotype DRB1*0804, DQA1*0102, DQB1*0604 ($D = 0.0132$) was found only in San (frequency = 2.2%) and the haplotype DRB1*1301, DQA1*0103, DQB1*0602 ($D = 0.0165$) occurred with a frequency of 3.5% in Khoi and 8.7% in San.

Of the newly discovered haplotypes with a *D*-value below 0.0100, three were also found in Xhosa and in the South African mixed-ancestry population in this study. They were the DRB1*0301, DQA1*0501, DQB1*0301 haplotype, which was present in Xhosa, and the haplotype DRB1*1501, DQA1*0102, DQB1*0603, which was present in the mixed-ancestry population. Their presence in more than one population may indicate also that these are stable haplotypes

Some haplotypes were present with a high frequency. The highest frequency was shown by the DRB1*0401, DQA1*0301, DQB1*0302 haplotype (39.1% in San and 13.3% in Khoi). Haplotype DRB1*1501, DQA1*0102, DQB1*0602 was the next most common (20.9% in Khoi but only 1.4% in San). Three other common haplotypes were the DRB1*1302, DQA1*0102, DQB1*0605 (frequency of 9.6% in Khoi and 2.9% in San); DRB1*1301, DQA1*0103, DQB1*0602 (frequency of 8.7% in San and 3.5% in Khoi); and DRB1*1302, DQA1*0102, DQB1*0604 (frequency of 1.7% in Khoi and 8.0% in San).

South African Xhosa population

HLA-class II allele specificities

Statistical analysis. The observed heterozygosity ranged from 0.872 for the DQA1 and DQB1 loci to 0.962 for the DRB1 locus. In each case the expected heterozygosity at Hardy-Weinberg equilibrium was less than the observed frequency, which yields a negative value for Wright's *F* for each locus. At DPB1 locus the expected *H*, 0.939, is not significantly different from the observed value but does serve as the lower boundary for its confidence interval. At none of the loci was the Nam-Gart *T* statistic significantly different from 1.0 (results not shown).

DRB1 locus. The allele frequencies for the DRB1, DQA1, DQB1 and DPB1 alleles for the Xhosa are listed in Table 3. A relatively large number of different DRB1 alleles (25) was present in Xhosa. Some DRB1 alleles have not previously been reported in Caucasoids.[8,9] Most striking were the DRB1*0302, DRB1*1503 and DRB1*0701 alleles. Unusual alleles which were found only once were DRB1*0405, *0802, *1305 and *1404. The DRB1*0302, *0405, *0802 and *1305 alleles have been described in other African black populations.[10–12] They may represent specificities which are unique to African blacks. The frequencies of certain DRB1 alleles showed a tendency either to increase or decrease from south to north Africa (clinal variation). The DRB1*1304 specificity, which has an extremely high (27.3%) frequency in Gambian blacks,[10] was not found in Xhosa. It was reported at a frequency of 3% in Malawians.[10] In addition, the frequency of DRB1*1302 was lower in Xhosa than that reported for Gambians[10] (7.1 *vs* 16.4%). The frequency for the DRB1*0301 and *0302 alleles, which is low (6.1%) in Gambians and somewhat higher in Malawians (13.7%),[10] was relatively high in Xhosa (25.6%). This effect was due mainly to an increase in the DRB1*0302 allele (15.4% in Xhosa). In Caucasoids the frequency of the DRB1*0301 allele was 11.5%.[8,9] The DRB1*11 alleles were also present at a high frequency in Xhosa (18.0%); the frequency in Malawians and in Gambians was 15.5%.[10] The frequency of the DRB1*1101 allele was particularly high in Xhosa (10.3%). The frequency in Caucasoids is 11.5%.[8] Furthermore, the frequencies in Xhosa for the DRB1*0101 and *0701 alleles were far lower than in Caucasoids.[8] Overall, the frequencies for the DRB1 alleles in Xhosa reflected those reported for other black Africans.[10–12] The DRB1*0402, *0403, *0408, *0801 and *1502 alleles, which are relatively common in Caucasoids,[8] were not found in Xhosa.

Table 3. DRB1, DQA1, DQB1 and DPB1 allele frequencies in the South African Xhosa population.

Allele	Frequency	Allele	Frequency	Allele	Frequency
DRB1*		**DQA1***		**DPB1***	
0101	0.0192	01	0.0640	0102	0.1026
0102	0.0064	0101	0.0705	0201	0.0897
0301	**0.1026**	0102	0.2507	0202	0.0128
0302	**0.1538**	0103	0.1000	0401	0.0577
0401	0.0449	0201	0.0320	0402	0.0385
0404	0.0064	0301	0.0897	0601	0.0064
0405	0.0064	0401	0.1538	1101	0.0192
0701	0.0577	0501	0.1529	1301	0.0449
0802	0.0064	blank	0.0862	1501	0.0128
0803	0.0064			1701	0.0064
0804	0.0064	**DQB1***		1801	0.0641
0901	0.0128	02	0.0064	2001	0.0128
1001	0.0321	0201	0.1665	2101	0.0064
1101	**0.1026**	0301	0.1026	2301	0.0064
1102	0.0321	0302	0.0385	2501	0.0064
1104	0.0449	0303	0.0128	2601	**0.0737**
1201	0.0256	0402	**0.1603**	2701	0.0192
1301	0.0626	0501	**0.1205**	2801	0.0128
1302	**0.0705**	0602	0.2221	3101	0.0128
1303	0.0064	0603	0.1026	3201	**0.0385**
1305	0.0064	0604	0.0321	3301	**0.0321**
1401	0.0320	0605	0.0256	3401	**0.0128**
1404	0.0064	blank	0.0100	3901	0.0064
1501	0.0769			4001	0.0064
1503	0.0576			4101	**0.0264**
Blank	0.0016			4401	0.0064
				4601	0.0064
				4701	**0.0192**
				4901	**0.0192**
				5001	**0.0449**
				5101	0.0128
				Blank	0.1627

Number of individuals tested: DRB1 locus, 83; DQA1 locus, 72; DQB1 locus, 90; DPB1 locus, 78. Bold script indicates frequencies which were substantially different from those reported for other African blacks.

DQA1 locus (Table 3). The frequencies of the DQA1*0102 and *0401 alleles were relatively high compared to Caucasoids. Similar results were found for Zimbabwean and American blacks.[11,12] In agreement with the literature,[9,12] the DQA1*0601 allele was not found in Xhosa.

DQB1 locus (Table 3). The frequency of the DQB1*0402 allele was considerably higher (16.0%) in Xhosa than reported for Zimbabweans[11] and African-Americans[12] (6.0% and 6.6%, respectively). By contrast, the frequency of the DQB1*0501 allele was considerably lower in Xhosa (12.1%) than reported for Zimbabweans (22.0%[11]). The frequencies of the other alleles of

Table 4. Three-locus DRB1, DQA1, DQB1 haplotype frequencies in South African Xhosa.

Haplotype			Frequency	[a]Reported
DRB1*	DQA1*	DQB1*		
0101	0101	0501	0.0192	*
0102	0101	0501	0.0049	*
0301	0501	0201	**0.0759**	*
0301	0501	0301	0.0128	[b]Khoisan
0302	0401	0402	**0.0802**	**
0401	0301	0301	0.0064	*
0401	0301	0302	0.0321	*
0401	0301	0303	0.0064	No
0404	0301	0301	0.0064	*
0405	0301	0302	0.0064	*
0701	0201	0201	0.0256	*
0701	0301	0201	**0.0192**	Khoisan, [c]CC
0804	0401	0301	0.0064	AA
0901	0301	0201	0.0128	*
1101	0102	0602	**0.0780**	*
1101	0501	0301	0.0106	AA
1102	0501	0301	0.0321	*
1104	0103	0603	0.0449	*
1201	0101	0501	0.0256	*
1301	0103	0603	0.0349	*
1302	0102	0501	0.0119	*
1302	0102	0604	0.0321	*
1302	0102	0605	0.0251	*
1303	0201	0201	0.0064	**
1305	0501	0301	0.0064	*
1501	0102	0602	**0.0547**	**
1503	0101	0501	0.0121	**
1503	0101	0602	0.0067	CC
1503	0102	0501	0.0065	**
1503	0102	0602	0.0293	**

Number of individuals tested: 78. [a]Haplotypes marked * have been described in more than one ethnic group; haplotypes marked ** have been described only in African blacks. Haplotypes marked AA have only been reported in African-Americans[12]. [b]Khoisan: previously unreported, but found during this study in the Khoisan population; [c]CC, previously unreported, but found during this study in the Cape coloured population. Bold script indicates frequencies of haplotypes which were found at an unusually high frequency.

Table 5. DRB1, DQA1, DQB1 and DPB1 allele frequencies in the South African mixed-ancestry (Cape coloured) population.

Allele	Frequency	Allele	Frequency	Allele	Frequency
DRB1*		**DQA1***		**DPB1***	
0101	0.0333	01	0.0476	0101	0.1279
0102	0.0167	0101	0.0833	0201	0.1002
0103	0.0056	0102	0.1333	0210	0.0056
0301	0.0833	0103	0.0500	0301	0.0333
0302	0.0278	0201	0.1158	0401	0.1724
0401	0.0719	0301	0.1205	0402	0.1153
0403	0.0111	0401	0.0333	0403	0.0056
0404	0.0056	0501	0.2294	0501	0.0389
0405	0.0167	0601	0.0333	0901	0.0222
0408	0.0056	blank	0.1533	1101	0.0056
0701	0.1387			1201	0.0056
0901	0.0167	**DQB1***		1301	0.0722
1001	0.0328	0201	0.1929	1401	0.0278
1101	**0.1301**	0301	0.2167	1501	0.0056
1102	0.0167	0302	0.0886	1601	0.0111
1104	0.0278	0303	0.0278	1701	0.0056
1201	0.0222	0402	0.0500	1801	0.0167
1202	**0.0500**	0501	0.1356	1901	0.0111
1301	0.0222	0502	0.0333	2001	0.0056
1302	**0.0167**	06	0.0056	2101	0.0056
1303	0.0056	0601	0.0333	2301	0.0056
1308	**0.0056**	0602	0.1317	2601	0.0056
1401	0.0222	0603	0.0500	2701	0.0170
1501	0.0833	0605	0.0056	2801	0.0056
1502	0.0282	0609	0.0056	3101	0.0056
1503	0.0244	blank	0.0234	3901	0.0056
1602	**0.0056**			5001	0.0056
Blank	0.0145			5101	0.0056
				5501	0.0056
				Blank	0.1450

Number of individuals tested: DRB1 locus, 88; DQA1 locus, 73; DQB1 locus, 103; DPB1 locus, 78. Bold script indicates frequencies which were substantially different from those found in the founding Khoi, San, Xhosa and Caucasian populations.

the DQB1 locus in Xhosa are roughly in agreement with those reported for Zimbabweans.[11]. One DQB1*02 allele could not be sub-typed.

DPB1 locus (Table 3). The DPB1 alleles yielded the most interesting results, in that a large number of relatively rare alleles were found. Out of 56 alleles tested, 31 different alleles were found. The most frequently occurring was DPB1*0102 (10.3%). The next most frequent was DPB1*0201 (9.0%). The DPB1*0101 allele, which is commonly found in Caucasoids,[8] was absent in Xhosa. The frequencies of the DPB1*0401 and *0402 alleles, which have a very high incidence in most other populations, were only 5.8% and 3.9% in Xhosa, respectively. This is similar to findings in African-Americans.[12] The relatively rare alleles DPB1*2601 and *1801 had frequencies of 7.4% and 6.4%, respectively. Other rare alleles that were repeatedly found in Xhosa were the DPB1*2001, *2701, *3201, *3301, *3401, *4101, *4701, *4901, *5001 and *5101 alleles. Of these only the DPB1*2001, *2701 and *5101 were found in African-Americans.[12]

HLA DRB1, DQA1, DQB1 haplotype characteristics and frequencies

The DRB1, DQA1, DQB1 haplotypes are listed in Table 4. Most of these have been described in family studies.[9] However, we found some which have, to our knowledge, not been described. The DRB1* 0301, DQA1*501, DQB1*301 haplotype was found repeatedly. It was also present in the Khoisan population. The DRB1*0701, DQA1*0301, DQB1*0201 haplotype was also found repeatedly. This haplotype was present in the Khoisan as well as in the South African mixed-ancestry population. Two other new haplotypes which occurred at a relatively low frequency were the DRB1*0401, DQA1*0301, DQB1*0303 and the DRB1*1503, DQA1*0101, DQB1*0602. The last was also found in the South African mixed-ancestry population. Two haplotypes were found in Xhosa that have been reported only in other (non-black) ethnic groups: the DRB1*0404, DQA1*0301,QB1*0301 and the DRB1*1104, DQA1*0103, DQB1*0603 haplotypes.[8,9] In total, 30 different haplotypes were found, of which 12 (40.0%) were specific for African blacks. In African-Americans, 73 different haplotypes were found.[12] This may be due to the larger number (241) of individuals in their study. In our study 7/30 (23.3%) Xhosa haplotypes were found only once.

South African mixed-ancestry population

HLA allele specificities and frequencies

Statistical analysis. For the South African mixed-ancestry population the observed heterozygosity ranged from 0.822 for the DQB1 locus to 0.922 for DQA1. Expected heterozygosity for the DQA1 locus is 0.823, which is significantly smaller than the observed value and yields a Wright's F of –0.120. The Nam-Gart T statistic was not significantly different from 1.0, except for the DRB1 locus, for which the T statistic was 2.246 (95% confidence interval: 1.152–3.339) (results not shown).

DRB1, DQA1, DQB1 and DPB1 loci. The results are depicted in Table 5. Allele frequencies for all four loci were mostly intermediate between those found in Xhosa and Caucasoids.[8,9] There were 27 DRB1 alleles. The DRB1*1308 and *1602 alleles were not present in Khoisan, Xhosa or Caucasoids.[8,9] The frequency for

Table 6. Three-locus DRB1-DQA1-DQB1 haplotype frequencies in the Cape coloured population.

Haplotype			Frequency	[a]Reported
DRB1*	DQA1*	DQB1*		
0101	0101	0501	0.0333	*
0102	0101	0501	0.0165	*
0301	0501	0201	**0.0706**	*
0302	0401	0402	0.0222	**
0401	0301	0302	**0.0599**	*
0403	0301	0302	0.0056	*
0405	0102	0502	0.0056	No
0405	0301	0401	0.0111	Chinese
0408	0301	0301	0.0056	Jews
0701	0201	0201	**0.0756**	*
0701	0201	0303	0.0222	*
0702	0301	0201	0.0167	[b]Khoisan, [c]Xhosa
0901	0301	0201	0.0111	*
1101	0102	0602	0.0190	*
1101	0301	0302	0.0056	No
1101	0501	0301	**0.1055**	AA
1102	0501	0301	0.0167	*
1104	0103	0603	0.0167	*
1104	0501	0301	0.0111	*
1201	0101	0501	0.0222	*
1202	0501	0301	0.0166	No
1202	0601	0301	0.0333	Chinese
1301	0103	0603	0.0056	*
1302	0102	0605	0.0056	*
1501	0102	0502	0.0167	No
1501	0102	0602	0.0556	**
1501	0102	0603	0.0056	Khoisan
1502	0103	0601	0.0222	*
1503	0101	0602	0.0056	Xhosa
1602	0102	0502	0.0056	*

Number of individuals tested: 88. [a]Haplotypes marked * have been described in more than one ethnic group; haplotypes marked ** have been described only in African blacks. Chinese, Jews: Only described in these populations. Haplotypes marked AA have been reported only in African-Americans[12]. [b]Khoisan: previously unreported, but found during this study in the Khoisan population; [c]Xhosa, previously unreported, but found during this study in the Xhosa population. Bold script indicates frequencies of haplotypes which were found in an unusually high frequency.

Table 7. Possible ancestral Southern African DRB1, DQB1, DPB1 haplotypes: Frequencies in Namibian Khoi and San and in South African Xhosa with a frequency higher than 1% and a [a]*D* value higher than 0.0100. Also included were haplotypes which were found in both Khoi and San, or shared with the Xhosa or the Cape coloured population.

Haplotype			Frequency			
DRB1*,	DQB1*,	DPB1*	Khoi	San	Xhosa	Cape coloured
0101	0501	0401	–	0.0058	0.0064	–
0102	0501	0401	–	0.0229	–	–
0301	0201	0101	–	0.0117	–	0.0385
0301	0201	0201	–	–	0.0315	–
0301	0201	0402	–	0.0291	–	–
0301	0301	0101	–	0.0173	0.0064	–
0302	0402	0101	–	0.0115	0.0064	0.0142
0401	0302	0101	**0.1194**	0.0287	–	–
0401	0102	0301	0.0176	0.0102	–	0.0062
0401	0302	0401	**0.1530**	–	–	0.0127
0401	0302	0402	0.0253	**0.0861**	0.0064	0.0155
0404	0401	0402	0.0144	0.0266	–	–
0404	0402	0801	0.0217	–	–	–
0701	0201	0101	–	0.0175	–	0.0056
0701	0201	0401	0.0426	0.0116	0.0531	–
0701	0303	0401	–	–	0.0167	–
0804	0301	0201	–	0.0163	–	–
0804	0604	3001	0.0145	–	–	–
1001	0501	0201	–	0.0174	–	–
1101	0301	0101	–	–	0.0269	–
1101	0301	0402	–	–	0.0309	–
1102	0301	0402	0.0217	–	–	–
1301	0602	0101	0.0078	0.0232	–	–
1301	0602	5501	0.0207	0.0058	–	–
1301	0603	0101	0.0280	–	0.0113	–
1302	0604	0401	0.0228	0.0116	–	–
1302	0604	5501	**0.0563**	–	–	–
1302	0605	0201	0.0145	**0.0561**	–	0.0056
1302	0605	0301	–	0.0292	–	–
1501	0602	0101	–	0.0258	–	0.0140
1501	0602	0201	0.0072	0.0197	0.0069	0.0089
1501	0602	0401	–	0.0381	–	0.0091
1501	0602	0402	–	**0.0906**	–	–

Number of individuals tested: Khoi, 86; San, 69; Xhosa, 78; Cape coloureds, 78. [a]*D*, measure of the strength of association of alleles. Bold script indicates an unusually high haplotype frequency.

the DRB1*1101 allele was higher than in Khoisan, Xhosa or Caucasoids;[8,9] the frequency for the DRB1*1302 allele was considerably lower than in these populations. These differences may reflect the southeast Asian ancestry of the Cape coloured population.[29]

For the DQA1 locus, the allele frequencies in the Cape coloured population were closer to those of Caucasoids[8,9] than to those of the Xhosa. However, the frequency for the DQA1*0601 allele was higher than in Caucasians or Khoisan. This may also be due to admixture of populations from Southeast Asia. For the DQB1 locus, an interesting finding was the presence of the rare DQB1*0609 allele, which was not present in Khoisan, Xhosa or Caucasoids. Twenty-nine different DPB1 alleles were found, of which several were not present in Khoisan, Xhosa or Caucasoids [8,9]: DPB1*1401 and *1601 alleles were repeatedly found, whereas DPB1*2601 and *5501 alleles were found only once.

HLA DRB1, DQA1, DQB1 haplotype characteristics and frequencies

We found 30 different haplotypes (Table 6), seven of which have not been reported previously: The DRB1*0701, DQA1*0301, DQB1*0201, and the DRB1*1503, DQA1*0101, DQB1*0602 haplotypes were found repeatedly; these haplotypes were also found in Xhosa; the former was also found in Khoisan. The DRB1*1501, DQA1*0102, DQB1*0603 was found only once but was also observed in Khoisan. Thus, four of these haplotypes are well documented. Three other haplotypes which had not been described previously, and were present at a low frequency were the DRB1*0405, DQA1*0102, DQB1*0502, the DRB1*1101, DQA1*0301, DQB1*0302, and DRB1*1202, DQA1*0501, DQB1*0301. Of the total of 30 haplotypes, 17 were shared with Xhosa, of which 12 (40.0%) may be specific for African blacks.

The most frequently found haplotypes were the DRB1*1101, DQA1*0501, DQB1*0301 (10.6%); the DRB1*0701, DQA1*0201, DQB1*0201 (7.6%) and the DRB1*0301, DQA1*0501, DQB1*0201 (7.1%).

HLA DRB1, DQB1, DPB1 haplotypes in southern African populations

For the DRB1, DQB1, DPB1 haplotype, the results for the Namibian Khoi and San and the South African Xhosa and Cape coloured population are presented together in Table 7. Owing to a cross-over frequency between the DQB1 and DPB1 loci of approximately 1%, the DRB1, DQB1, DPB1 haplotypes are usually extremely varied and of limited value for anthropological purposes. We found 30 different DRB1, DQB1, DPB1 haplotypes in the Khoi and 47 in the San; of these, 10 were shared. In both the Xhosa and the South African mixed-ancestry populations we found 62 different DRB1, DQB1, DPB1 haplotypes, of which only five were shared between the two groups.

The factors which determine the stability of a haplotype in different populations are poorly understood. Although increased distance between loci is generally accepted to be an important factor, other, less well-defined characteristics of the haplotype are believed to play a role.[24] Some haplotypes may be relatively more stable than others, and therefore of more interest for anthropological purposes.[25,26]

To determine which haplotypes could possibly be derived from a hypothetical ancestral founding population, we selected DRB1, DQB1, DPB1 haplotypes on the basis of the following two arbitrary criteria:

1) Khoi, San and Xhosa haplotypes which were found in a frequency of higher than 1% and had a *D*-value greater than 0.0100. The *D*-value is a measure of the strength of association of the DRB1, DQB1 and DPB1 alleles. Since the Cape coloured population is derived from recent and varied ethnic admixture, this source was excluded.

2) Haplotypes which were found in more than one southern African population. Since the Khoisan is an important founding population of the Cape coloureds,[1–3] the latter was included for this criterion.

The results are depicted in Table 7. In the Khoi and San many of the DRB1, DQB1, DPB1 haplotypes were found at a high frequency (>2.0%). This phenomenon was more pronounced in the Khoi; haplotypes DRB1*0401, DQB1*0302, DPB1*0101 and DRB1*0401, DQB1*0302, DPB1*0401 were found at frequencies of 11.9% and 15.3%, respectively. Haplotype DRB1*1302, DQB1*0604, DPB1*5501 was found at a frequency of 5.6%. In San, haplotypes DRB1*0401, DQB1*0302, DPB1*0402 (8.6%) and DRB1*1302, DQB1*0605, DPB1*0201 (5.6%) were relatively common.

In both Xhosa and Cape coloureds, the highest frequency for any one haplotype was 4.0%. In the Xhosa the DRB1, DQB1, DPB1 haplotype repertoire included seven haplotypes which were also found in the Khoisan; three of these were found only in San, three were found in both Khoi and San, and one only in Khoi. The South African mixed-ancestry population DRB1, DQB1, DPB1 haplotypes included 10 haplotypes which occurred in the Khoisan; of these, five were found only in San, four in both Khoi and San, and one only in Khoi.

Discussion

Anthropological considerations

The primary aim of this investigation was to elucidate the anthropological relationship between the Khoi and San, and present-day populations with regard to the HLA system. Because of its central role in immune reactions, the HLA system is an important target of therapeutic drugs for autoimmune diseases. Its diversity is also expected to affect the efficacy of vaccines and of immuno-therapeutic intervention. The Khoisan are genetically the closest representatives still in existence of a founding population pre-dating ethnic diversity. Their HLA allele and haplotype specificities may encompass a broad spectrum of diverse populations. Understanding the anthropological relationships of the HLA system in different ethnic groups is therefore of prime importance. It should be noted that small numbers of individuals were investigated and in the Khoi and San inbreeding could not be excluded. This makes it difficult to draw specific conclusions regarding human evolution. Owing to their recent origin and ethnic admixture, the results of the South African mixed-ancestry population were not included in this discussion.

In the Khoi and San, a selective expansion of certain alleles appears to have taken place, which is more evident in the San than in the Khoi. This is similar to findings in the Chinese Kazak and Han populations.[17] Taking into account the overwhelming anatomical, linguistic and cultural evidence for a common ancestry of the Khoi and the San,[3] it is plausible that the loss of allelic repertoire and selective expansion of alleles is a relatively recent development, which has taken place since the divergence of the Khoi and the San. This divergence coincided with the gradual desiccation of southern Africa. Climatological change may have been the driving selection force for the patterns of the HLA class II allelic repertoire observed in the present-day Khoi and San. The more profound loss of allelic diversity and the predominance of certain alleles in the San reflect the more extreme desert and semi-desert habitat of this population. The ancestral founding population of *Homo sapiens* may have lived under tropical conditions and may have displayed a far greater allelic diversity in the HLA loci than the present-day Khoi and San. The Khoi, with a greater allelic diversity, are more likely to reflect the genetic repertoire of this ancestral population than the San.

There are several indications that Caucasoids may be derived from an early, and African blacks from a later, migration wave of an ancestral founding population: The DQA1*0601 allele was not found in Xhosa. This is in agreement with most reports,[10,12] which indicate that this allele has not been found in African blacks. Only a Zimbabwean study reports an incidence of 1% for this allele[11] and this may have been incidental. This allele was also not found in Namibian San and Khoi but is present in Caucasoids. Alternatively, this allele may have evolved in Caucasoids after migration. Furthermore, there are several alleles which were present in both the Khoi and Caucasoids but were not found in Xhosa and San. They are the DRB1*1601, DQB1*0502 and *0601 alleles, and the DPB1*0501, *1901 and *2201 alleles. Thus, the African blacks may have been derived from the ancestral population after they had already lost these alleles, or from the San. The allele frequency distribution for the HLA loci in the Khoi is more similar to that found in African blacks than in Caucasoids,[8,9] suggesting a longer separate development of the latter.

The biological advantage of allelic predominance is not known. An interesting finding was the relatively high frequency (40.9%) of the DRB1*0401 allele in the San. This allele, which has been implicated in the aetiology of rheumatoid arthritis (RA) in the Cape coloured population and in Xhosa,[27] may have had a survival advantage in the extremely hot and dry conditions in which the San live. The DRB1*1001 specificity, which has been implicated in the aetiology of juvenile chronic arthritis in the mixed-ancestry population,[28] as well as in the development of RA of other ethnic groups,[29] was found only in Khoi. It should be noted that the specificities of alleles which had selectively expanded in the Khoi and San were, in some cases, the same as those which had expanded in other ethnic groups. The DRB1*0401 was also found at a high frequency in some Siberian Eskimos.[16] The DRB1*1301 and *1302 alleles, which were very prevalent in both Khoi and San, are also prevalent in the Japanese.[23] The DRB1*1501 allele, which had a frequency of only 1.5% in the San but of 30.0% in Khoi, was also found with a high frequency in Chinese Han.[17] The DQA1*0102 and *0301, which occurred frequently in Khoi and San, did so also in Chinese Kazak[17]. The DQB1*0302 and *0602 alleles occurred frequently in the Khoi and San and in some Siberian populations.[16] Predominance of the DPB1*0401 and *0402 alleles, as in the Khoi and San, also occurred in the Chinese Kazak[17] population. By contrast, many characteristics of the Khoi and San haplotypes were not shared with the Han, Kazak and Siberians.

It is plausible that certain allele specificities have a tendency to predominate under certain climatological conditions. Thus, this similarity in allele frequencies would be the result of parallel development, rather than a measure of kinship between these ethnic groups.

In Xhosa, this investigation revealed a high allelic diversity for the DRB1 and DPB1 loci. This is similar to findings in other black African populations[10,12,25,26] and compatible with findings for Brazilian Indians,[14,15] who also have a tropical rainforest as habitat. Compared to reports on Gambians, Malawians and Zimbabweans,[10,11] our results for the Xhosa reveal clinal variation for certain DRB1 alleles across the African continent: the frequencies of the DRB1*0302 and *1101 alleles show a gradual increase from north to south, whereas the frequencies for the DRB1*1302 and *1304 alleles decrease gradually from north to south. Frequencies for these alleles in Caucasians did not fit into this pattern. Thus, these tendencies appear to be restricted to the African continent. Similar clinal variation has also been described for the vitamin D binding protein and Properdin Factor B.[7] This variation may have arisen during the most recent migration of west African Bantu-speaking tribes to central and southern Africa.[3]

Our results show that the repertoire of the DRB1, DQA1, DQB1 haplotypes of the San was more limited than that of the Khoi, which is to be expected as a result of the larger allelic diversity of the latter. A high percentage (73.0%) of the Khoisan haplotypes were public haplotypes, since they were shared with other ethnic groups. By contrast, only 24% of public haplotypes in the combined Han and Kazak population[17] and only 17% of public haplotypes in the Siberian populations[16] were found. This supports the hypothesis that the Khoisan are closely related to an ancestral founding population.

Since the Khoisan did not move to different climatic zones, no bottleneck or founder effects could have arisen. It is therefore likely that the DRB1, DQA1, DQB1 haplotypes found in the Khoisan group were selected for their stability. In the Khoi we found 28, and in the San only 19 different DRB1, DQA1, DQB1 haplotypes. The Siberian populations,[16] as well as the Han and Kazak[17] people, possessed a surprisingly high number of different DRB1, DQA1, DQB1 haplotypes. A frequency of 39.1% for the DRB1*0401, DQA1*0301, DQB1*0302 haplotype in the San is, to our knowledge, the highest ever described for a DRB1, DQA1, DQB1 haplotype.

Haplotype diversity is achieved by intergenic recombination. Once a founder effect has taken place, this diversity may represent a faster mechanism to achieve diversification of the HLA system than the generation of new alleles if this is required for the adaptation to a change in habitat.

In Xhosa, we found 30 different DRB1, DQA1, DQB1 haplotypes. The African-American[12] population with 73 different haplotypes is thus far more diverse. This diversity is likely to be due to the heterogeneity of their African ancestral populations and seems to have arisen from recombination between DRB1, DQA1, DQB1 haplotypes with different linkage disequilibrium in various black African populations.

We also investigated the DRB1, DQB1, DPB1 haplotypes in the Khoi, San and Xhosa populations, as well as in the Cape coloured population. Although the Khoi expressed more allelic diversity than the San, the number of DRB1, DQB1, DPB1 haplotypes in the San (47) was far greater than in the Khoi (30). The recombination frequency of approximately 1% between the DQB1 and the DPB1 loci may generate a larger variation of these haplotypes in older populations. It is interesting to note that despite a high recombination frequency, 14 of the Khoisan haplotypes were also present in the Xhosa and/or South African mixed-ancestry population. Furthermore, a number of DRB1, DQB1, DPB1 Khoisan haplotypes were present at high frequencies (15–16%). In both the Xhosa and Cape coloured populations a much higher number (62) of DRB1, DQB1, DPB1 haplotypes was found. Comparing the 34 relatively well-documented DRB1, DQB1, DPB1 haplotypes reported in this study with other ethnic populations may provide additional insight into the ethnic anthropology of the HLA system.

Implications for organ donor transplantation

Documentation of the HLA system of populations is of prime importance for the optimization of donor organ transplantation. The HLA allelic specificities and haplotypes of the Xhosa and the South African mixed-ancestry population as determined by DNA typing have not previously been described. There is extreme allelic variation of the DRB1 and DPB1 loci in the South African Xhosa and the Cape coloured populations. The relatively high number of alleles which could not be typed stresses the need for further DNA analysis in these populations. The high degree of allelic diversity has a direct bearing on the feasibility of establishing a programme of matching unrelated donors to tissue transplant recipients[31] in South Africa, as is presently performed for unrelated bone marrow transplantation.[32]

Our findings indicate that there are considerable differences between Xhosa and African-Americans, especially at the HLA haplotype level. As for finding African matched donors, it should be kept in mind that Xhosa are only one of the many different African black tribes, even within South Africa. Since many of these tribes have been geographically isolated until recently, there may be considerable differences in linkage disequilibrium between the HLA loci among tribes. Along with modernization, these tribes can be expected to admix, which would result in added variation at the haplotype level.

We gratefully acknowledge Berenice Arendse, Jonathan Glaser, Genee Harms, Muazzam Jacobs and Kathy Meehan for skilled technical and computational assistance. The Health Branch of the Provincial Administration of the Western Cape provided financial support for this investigation.

Received 20 February. Accepted 10 Novenber 2002.

1. Begovich A., McClure G.R., Suraj V.C. *et al.* (1992). Polymorphism, recombination and linkage disequilibrium within the HLA region. *J. Immunol.* **148**, 249–258.
2. Andersson L., Gustafsson K., Jonsson A.K. and Rask L. (1991) Concerted evolution in a segment of the first domain exon of polymorphic MHC B loci. *Immunogenetics* **33**, 235–242.
3. Cavalli-Sforza L.L. and Bodmer W.F. (1971). *The Genetics of Human Populations.* W.H. Freeman, San Francisco.
4. Du Toit E.D., Taljaard D., Marshall I., Ritchie C. and Oudshoorn M. (1990). The HLA genetic constitution of the Bushmen (San*). Hum. Immunol.* **28**, 406–415.
5. Arendse B., Martell R.W., Jacobs M., Oosthuizen O. and du Toit E.D. (1992). HLA-DP polymorphism in the present-day San (Bushmen) and Khoi (Hottentots) using polymerase chain reaction and sequence-specific oligonucleotide typing. *Tissue Antigens* **40**, 38–40.
6. Creemers P.C. and du Toit E.D. (1996). C4 polymorphism and extended HLA haplotypes in Namibian San and Khoi and in South African Xhosa. *Tissue Antigens* **47**, 111–116.
7. Creemers P.C., du Toit E.D. and Kriel J. (1995). DBP (vitamin D binding protein and BF properdin factor B) allele distribution in Namibian San and Khoi and in other South African populations. *Gene Geog.* **9**, 185–189.
8. Cerna M., Fernandez-Vina M., Ivaskova E. and Stastny P. (1992). Comparison of HLA alleles in gypsy and Czech populations by DNA typing with oligonucleotide probes. *Tissue Antigens* **39**, 111–116.
9. Fernandez-Vina M.A., Gao X., Moraes M.E. *et al.* (1991). Alleles at four HLA loci determined by oligonucleotide hybridization and their associations in five ethnic groups. *Immunogenetics* **34**, 299–312.
10. Hill A., Allsopp C., Kwiatkowski D. *et al.* (1992). Extensive genetic diversity in the HLA region of Africans, with a focally predominant allele, DRB1*1304. *Proc. Natl. Acad. Sci. USA* **89**, 2277–2281.
11. Garcia-Pacheco J.M., Herbut B., Cutbush S. *et al.* (1992). Distribution of

HLA-DQA1, -DQB1 and DRB1 alleles in black IBBM patients and controls from Zimbabwe. *Tissue Antigens* **40**, 145–149.
12. Just J.J., King M.C., Thomson G. and Klitz W. (1997). African-American HLA allele and haplotype diversity. *Tissue Antigens* **49**, 547–555.
13. Belich M.P., Madrigal J.A., Hildebrand W.H. *et al.*(1992). Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature* **357**, 326–329.
14. Watkins D.I., McAdam S.N., Liu X., Strang C.R. *et al.* (1992). New recombinant HLA-B alleles in a tribe of South Amerindians indicate rapid evolution of MHC class I loci. *Nature* **357**, 329–333.
15. Mack S.J. and Erlich H.A. (1998). HLA polymorphism in the Ticuna of Brazil: evolutionary implications of the DRB1*0807 allele. *Tissue Antigens* **51**, 41–50.
16. Grahovac B., Sukernik R.I., O'hUigin C. *et al.* (1998). Polymorphism of the HLA loci in Siberian populations. *Hum. Genet.* **102**, 27–43.
17. Mizuki M., Ohno S., Ando H., Sato T. *et al.* (1997). Major histocompatibility complex alleles in Kazak and Han populations in the Silk Route of north-western China. *Tissue Antigens* **50**, 527–534.
18. Jenkins T. and Tobias P.V. (1977). Nomenclature of population groups in Southern Africa. *Afr. Stud.*, **36**, 50–55.
19. Bidwell J.L. (1992). Applications of the polymerase chain reaction to HLA typing. *Vox Sang.* **63**, 83–89.
20. Andersson G., Andersson L., Larhammer D., Rask L. and Sigurdardottir S. (1994). Simplifying genetic locus assignment of HLA-DRB genes. *Immunol. Today* **15**, 58–61.
21. Long J.C., Williams R.C. and Urbanek M.(1995). An E-M algorithm and testing strategy for multiple-locus haplotypes. *Am. J. Hum. Genet.* **56**, 799–10.
22. Nam J. and Gart J.J. (1987). On two tests of fit for HLA data with no double blanks. *Am. J. Hum. Genet.* **41**, 70–76.
23. Kinoshita T., Hashimoto M., Yamasaki M., Ihara H., Ichikawa Y. and Fukunishi T. (1994). Striking conservation of three extended HLA-DR13 haplotypes in the Japanese population. *Tissue Antigens* **44**, 294–299
24. Salamon H., Klitz W., Easteal S. *et. al.* (1999). Evolution of HLA molecules. Allelic and amino acid site variability across populations. Genetics **152**, 393–400.
25. Begovich A.B., Moonsamy, P.V., Mack, S.J. *et al.* (2001). Genetic variability and linkage disequilibrium within the HLA-DP region: analysis of 15 different populations. *Tissue Antigens* **57**, 429–439.
26. Sanchez-Mazas A. (2001) African diversity from the HLA point of view: influence of genetic drift, geography, linguistics and natural selection. *Human Immunology* **62**, 937–948.
27. Rousseau J., Pokorny L., Glaser J., Creemers P.C. (2003). Predisposing and protective HLA-DR and -DQ alleles for rheumatoid arthritis in South African mixed-ancestry and Xhosa populations. *S. Afr. J. Sci.* **99**, 89–91.
28 Bhettay E., Martell R., Creemers P.C.(1994). Association of HLA-DR10 with juvenile chronic arthritis in South Africans of South African mixed ancestry population. *Tissue Antigens* **43**, 47–49.
29. Wordsworth B.P., Stedeford J., Rosenberg W.M.C., Bell J.I. (1991). Limited heterogeneity of the HLA contribution to susceptibility to rheumatoid arthritis is suggested by positive association with HLA-DR4, -DR1 and Drw10. *Br. J. Rheumatol.* **30**, 178–180.
30. Rousseau J., Mathew, C.G.P., Rees J.S., Du Toit E., Botha, M.C & Harley E.H. (1983). Incidence of Hb Barts and alpha thalassaemia genotypes in a South African population. *Acta Haematologica* **73**, 159–169**.**
31. Lonjou C., Clayton J., Cambon-Thomsen A., Raffous C. (1995). HLA-A, -B, -DR haplotype frequencies in France – implications for recruitment of potential bone marrow donors. *Transplantation* **60**, 375–378.
32. Creemers P., Jacobs P., Bird, A., du Toit E.D. (1993). The South African bone marrow donor registry. *S. Afr. Med. J.* **83**, 819–821.

Predisposing and protective HLA-DR and -DQ alleles for rheumatoid arthritis in South African mixed-ancestry and Xhosa populations

J. Rousseau*†, L. Pokorny*, J. Glaser* and P.C. Creemers*

We have investigated the distribution of the HLA-DRB1, -DQA1 and -DQB1 alleles in rheumatoid arthritis (RA) by comparing the allele frequencies in blood from 65 Cape coloured (mixed-ancestry) RA patients and 114 controls, and from 25 Xhosa RA patients and 94 controls. The strongest positive association with RA was found for the DRB1*0401 allele, followed by the DQA1*0301 and DQB1*0302 alleles, which are strongly linked with DRB1*0401. Data for both populations were statistically significant. In addition, DQB1*0501, which is in linkage disequilibrium with DR1 and DR10, showed a positive association with RA. These findings are in agreement with those for Caucasoids; they indicate that haplotypes that predispose for RA are highly conserved during evolution. Negative associations, that is, a protective effect for RA, were also found, but only for broad specificities; the associations were generally weaker. New findings were negative associations for DRB1*03, DRB1*0701, DQA1*0501 and DQB1*06. The DRB1*0301 and DQA1*0501 alleles are in linkage disequilibrium; a negative association was found in both populations. The negative association of DRB1*0701 was found only in the mixed-ancestry population and was absent in Xhosa. The effect of DQA1*06 was significant in both populations. Thus, the protective HLA-DR and DQ alleles show a greater ethnic diversity.

*Department of Clinical Laboratory Sciences, Faculty of Health Sciences Core Facility, Falmouth Building, University of Cape Town Medical School, Observatory 7925, South Africa.

†Author for correspondence. E-mail: jeanne@chempath.uct.ac.za

Introduction

The human leukocyte antigen (HLA) class II genes play an important role in the genetic predisposition to many autoimmune diseases. In the case of rheumatoid arthritis (RA), predisposing or susceptibility as well as protective genes[1] have been identified. Studies of HLA disease associations are complicated by the great extent of ethnic differences in HLA alleles and linkage disequilibrium between the DR and DQ loci.[2–4] Since HLA alleles may be targets for future therapeutic intervention, defining the HLA disease associations in every population is important. The Cape coloured population is unique in that it is of mixed origin, that is, it has Khoisan, South African Negroid, western European and southeast Asian (Malaysian, Indonesian and Indian) ancestry. South African Xhosa are Bantu-speaking African blacks in whom miscegenation with Khoisan has frequently occurred. Serologically determined frequencies of HLA-A, -B, -DR and -DQ in these populations differ considerably from those described for western Europeans.[5] The majority of previous disease association studies in the Cape coloured and southern African black populations have been performed using serological HLA typing for the HLA-DR and DQ loci; in the present investigation, high-resolution DNA class II typing was performed. DRB1*04 has been implicated as a predisposing factor for RA in many populations.[6–8] Similar results were found for a population of Zimbabwean blacks.[9,10] In Mediterranean populations, DRB*01 and DRB1*10 were predominantly found to be predisposing factor;[11] we found that DRB1*10 was a predisposing factor for juvenile RA in the mixed-ancestry population.[12]

Various theories have been proposed to explain the mode of action of HLA alleles. The 'shared epitope' hypothesis was developed to explain that different HLA-DRB1 alleles have been identified as susceptibility genes for RA: similar motifs within the third hyper variable (HV3) region of some HLA-DRB1 alleles have been found, which may affect the peptide binding characteristics of the HLA molecules. The HV3 region proteins induce stronger proliferative T-cell responses when derived from non-predisposing DRB1 chains,[13] suggesting an abnormal T-cell response. Another study found that the HV3 region of predisposing DRB1 chains constitutes a binding motif for the highly