*Proc. Natl. Acad. Sci. USA*
Vol. 88, pp. 8480–8484, October 1991
Immunology

# HLA-DR and -DQ gene polymorphism in West Africans is twice as extensive as in North European Caucasians: Evolutionary implications

**(selection/major histocompatibility complex class II/population bottleneck)**

OLLE OLERUP*†, MARITA TROYE-BLOMBERG‡, G. M. T. SCHREUDER§, AND ELEANOR M. RILEY¶‖

*Center for BioTechnology, Karolinska Institute, NOVUM and Department of Clinical Immunology, Karolinska Institute at Huddinge Hospital, Huddinge, Sweden; ‡Department of Immunology, University of Stockholm, Stockholm, Sweden; §Department of Immunohematology and Blood Bank, University Hospital, Leiden, The Netherlands; and ¶Medical Research Council Laboratories, P.O. Box 278, Fajara, The Gambia

*Communicated by Philippa Marrack, June 19, 1991*

**ABSTRACT** **The HLA genes are the most polymorphic coding loci known in humans. *DRB–DQA–DQB* gene polymorphism was investigated by *Taq* I restriction fragment length polymorphism analysis in more than 700 West Africans and found to be almost twice as extensive in West Africans as in North European Caucasians. This finding indicates that Africans comprise the oldest and genetically most diverse human population and supports the hypothesis of the occurrence of a population bottleneck in the emergence of the White race. As in Caucasians, less than one-third of possible cis-encoded *DQA–DQB* combinations were encountered, indicating constraints on the pairing of DQ $\alpha$ and $\beta$ polypeptides. Heterozygote advantage (i.e., positive selection) was found for *DRB*, *DQA*, and *DQB* alleles as well as for *DQA–DQB* combinations. However, in West Africans as well as in North Europeans the observed frequencies of *DRB–DQA–DQB* homozygotes were close to neutrality expectations. Although the hypothesis that HLA polymorphism is maintained by parasite-driven overdominant selection is attractive, there is little evidence to support that view. We propose instead that one of the forces maintaining a low frequency of HLA homozygotes might be a decreased likelihood of potentially autoreactive T-cell clones escaping thymic selection in HLA heterozygotes. This would be consistent with the central role of HLA molecules as self/nonself discriminators.**

Most expressed HLA genes exhibit a remarkable degree of allelic polymorphism, which derives from sequence differences predominantly localized to discrete hypervariable regions in the amino-terminal domains of the molecule. HLA variability influences both self-adjustment of the T-cell repertoire during thymic maturation and immune responsiveness of an individual to antigenic peptides.

Major histocompatibility complex (MHC) polymorphism has several unique features: most loci have many alleles, no allele dominates in frequency, and alleles differ by many amino acid substitutions. Two hypotheses, retention of ancestral polymorphisms and hypermutational diversification, have dominated speculations on the evolutionary origin of MHC diversity. The trans-species evolution theory (1)—i.e., that most major MHC allelic types diverged prior to the origin of the species in which they are found—has been supported by sequence analysis of rodent and primate MHC genes (2–5). The rate of amino acid-altering substitutions exceeds that of silent substitutions in codons of contact amino acids in the antigen-binding site of MHC class I and class II molecules, indicating that selection operates directly on the antigen-binding site (6–8). The high degree of polymorphism, long persistence of alleles, low frequency of homozygotes, and high rate of replacement substitutions can best be explained by overdominant selection. However, the evolutionary force(s) exerting such selection pressure remain unclear.

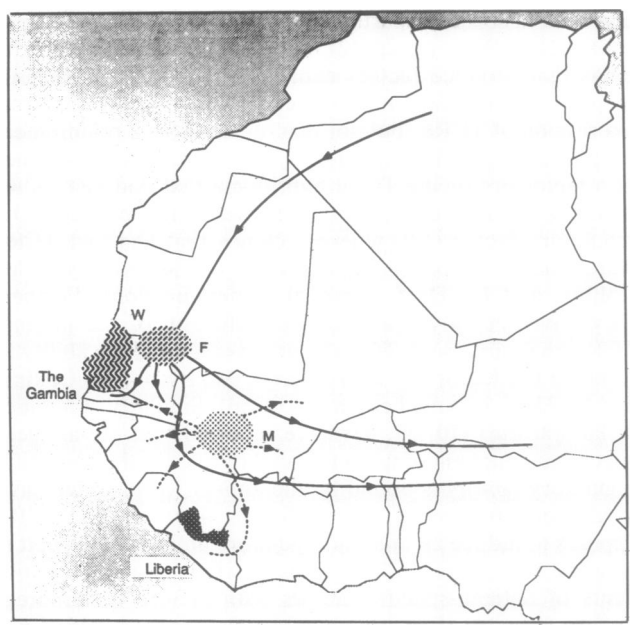


FIG. 1. Map of West Africa showing origins and migrations of the three Gambian ethnic groups: Fulas (F), nomadic descendents of Berber and Arab herdsmen, who settled in north Senegal in the 11th and 12th centuries, migrated widely during the 12th–19th centuries; Mandinkas (M), originated from the Mandings of the Mali empire and migrated during the 13th–15th centuries; Wollofs (W), indigenous population of northwest Senegal, center of the Jolof empire (14th–16th centuries), gradually spreading [Archives of the National Museum of the Gambia, Banjul, The Gambia (1966)]. Present habitat (the Yekepa area) of Liberians included in the study is indicated on the map. Most Liberian donors belonged to the Mano tribe.

Restriction fragment length polymorphism (RFLP) analysis of HLA class II genes, especially *DR* and *DQ*, has been developed into a powerful tool for identifying allelic variability in Caucasians, with a close correlation between specific RFLPs and cellularly and serologically defined specificities (9–12). We have employed this technique in a study of HLA class II gene polymorphism in more than 700 West Africans.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MHC, major histocompatibility complex; RFLP, restriction fragment length polymorphism.

†To whom reprint requests should be addressed.
‖Present address: Institute of Cell, Animal and Population Biology, Department of Genetics, University of Edinburgh, Edinburgh, Scotland.

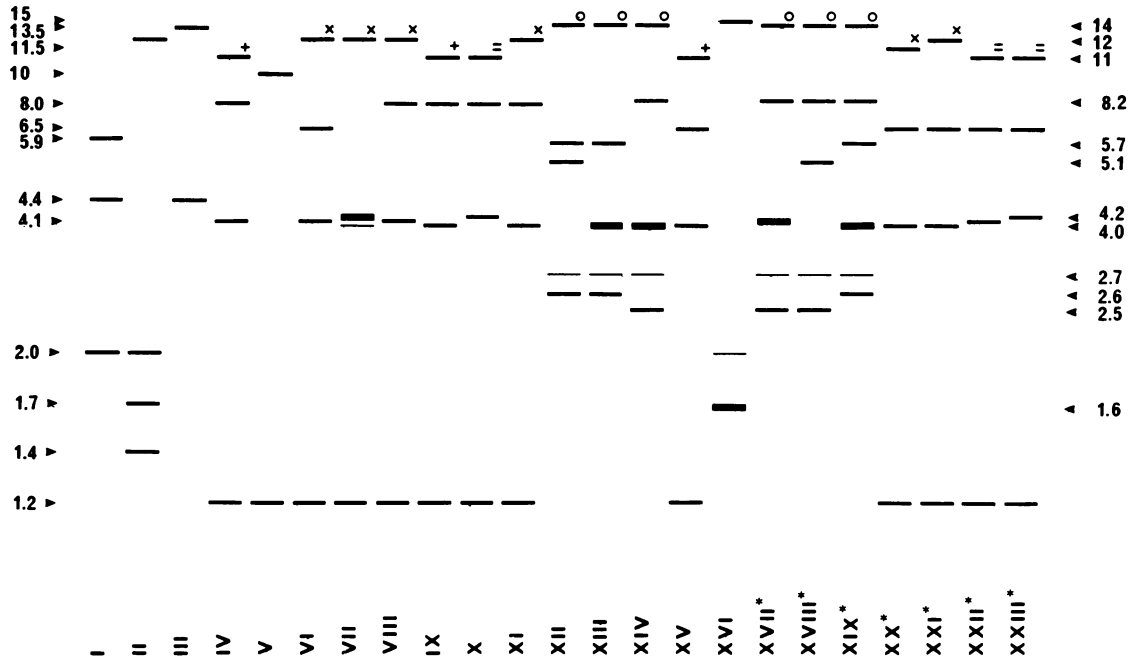


FIG. 2. Schematic illustration of allelic *Taq* I *DRB* restriction fragment patterns found in West Africans. Fragment sizes are given in kilobases. o, DRw53-associated 14-kb band; +, DRw52a; ×, DRw52b; =, DRw52c. Local RFLP nomenclature (Roman numerals) is indicated below each lane. RFLP patterns not observed in Caucasian populations are denoted by asterisks. Individuals homozygous for the non-Caucasian *DRB* patterns XVII, XX, XXI, and XXIII have been found in the present study. *DRB* patterns XVII and XXI have also been confirmed by segregation analysis. *DRB* patterns XVIII, XIX, and XXII were inferred in heterozygous individuals where the other *Taq* I *DRB* bands constituted a well-known allelic pattern.

## MATERIALS AND METHODS

**Study Subjects.** Venous blood samples were obtained from 731 West Africans: 638 Gambians belonging to three different ethnic groups (295 Fulas, 191 Mandinkas, and 152 Wollofs) and 93 Liberians (Fig. 1). The participants of the study came from rural areas, lived in separate villages, and tended to marry within their own ethnic group. Known first-degree relatives were excluded.

***Taq* I RFLP Analysis of *HLA-DRB*, *-DQA*, and *-DQB*.** Southern blot analysis of *Taq* I-cleaved DNA was done by standard techniques. Minor technical modifications and probes used are detailed in ref. 13.

## RESULTS AND DISCUSSION

***DRB*, *DQA*, and *DQB* RFLPs.** The elucidation of HLA class I and class II polymorphism has been the main objective of the International Histocompatibility Workshops since these were initiated in the 1960s. In recent years, serology, proliferative assays, and biochemical methods have been at least partially replaced by indirect (RFLP) and direct [hybridizing DNA amplified by the polymerase chain reaction with sequence-specific oligonucleotide probes (PCR–SSO) and sequencing] techniques to detect genetic variability. In looking for previously unidentified HLA class II polymorphism in non-Caucasian populations, RFLP analysis is superior to PCR–SSO because conserved primer sequences are not postulated, many SSOs are not allele-specific but rather complementary to conserved sequence motifs, which in non-Caucasian populations may be present in other alleles than in Caucasians, and homozygosity is readily recognized.

Most, if not all, of the functionally important allelic variability of the *DR* and *DQ* subregions in European and North American Caucasians have been characterized. However, in African and American Blacks only part of the polymorphism is known; e.g., only 56% of American Blacks can be HLA-D typed with homozygous typing cells for Dw1–Dw18, compared with 98% of Caucasians (14). There has been a marked increase in studies of HLA polymorphism in populations of African origin during the last few years (15–19). Non-Caucasian alleles as well as new DR–DQ associations have been observed.

Seven *DRB* and one *DQB* allelic *Taq* I RFLP pattern not observed in Caucasians were found in West Africans (Fig. 2; Table 1). This is consistent with the fact that among class II genes *DRB1* has the highest polymorphism. All but two of the new patterns were present in all four ethnic groups. Some of the non-Caucasian RFLP patterns were the most frequent ones in West Africans, with phenotype frequencies as high as 40–57%. All Caucasian and non-Caucasian allelic *DRB*, *DQA*, and *DQB* patterns were observed in >1% of the investigated individuals in one or usually several of the ethnic groups (Table 1), 1% being the conventional arbitrary line between true alleles and rare genetic variants.

**DQ $\alpha$–$\beta$ Pairing.** In West Africans only 15 of the 48 possible cis-encoded combinations of the six *DQA* and eight *DQB* RFLP patterns were observed. The situation was the same for Caucasians: only 29% of the potential *DQA–DQB* combinations were found. These data imply structural or functional constraints on the assembly of DQ $\alpha$ and $\beta$ chains. Intraisotypic (pairing of $\alpha$ and $\beta$ chains encoded in the same subregion but on different chromosomes) and interisotypic (pairing of $\alpha$ and $\beta$ chains encoded by genes in different subregions) transcomplementation has been demonstrated in murine (21, 22) and human (23) *in vitro* systems. Our population data, the evolutionary maintenance of specific $\alpha$–$\beta$ combinations (5), and gene transfer experiments in humans (24) and transgenic mouse experiments (25) showing restricted assembly of HLA-DQ and I-A molecules, respectively, suggest that intraisotypic transcomplementation is likely only for naturally occurring $\alpha$–$\beta$ heterodimers.

***DRB–DQA–DQB* Haplotypes.** The *Taq* I *DRB–DQA–DQB* haplotypes observed in 731 West Africans and 250 Swedes are detailed in Table 1. Also given are the serological DR and DQ specificities associated with the various RFLP patterns. Forty-five *DRB–DQA–DQB* haplotypes were found in West

Table 1. Distribution of *Taq* I *DRB–DQA–DQB* haplotypes in four West African ethnic groups and in North European Caucasians

| Haplotype* | | | Confirmation of haplotype† | Associated serological specificity‡ | | Frequency of haplotypes§ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Gambians | | | | | |
| *DRB* | *DQA* | *DQB* | | DR | DQ | Fulas ($n$ = 590) | Mandinkas ($n$ = 382) | Wollofs ($n$ = 304) | Liberians ($n$ = 186) | West Africans¶ ($n$ = 1462) | Swedes ($n$ = 500) |
| I | I | I | $C^{s,hz}$, $A^{hz}$ | 1, "Br" | w5 | 0.0186 | 0.0236 | 0.0164 | 0.1075 | 0.0308 | 0.0960 |
| I | IV | V | $C^{hz}$ | "Br" | w7 | – | – | – | – | – | 0 |
| II | II | II | $C^{s,hz}$, $A^{hz}$ | w15 | w6 | 0.0119 | 0.0131 | 0.0097 | 0.0645 | 0.0185 | 0.1720 |
| II | II | I | $C^{hz}$ | w15 | w1 | 0.0026 | 0 | 0 | 0.0007 | 0.0020 | 0.0020 |
| II | III | VIII | $C^{hz}$ | w15 | w6 | 0 | 0 | 0 | 0.0054 | 0.0007 | 0 |
| XVI | II | I | $C^{s,hz}$, $A^{hz}$ | w16 | w5 | 0.0186 | 0 | 0.0033 | 0.0430 | 0.0137 | 0.0060 |
| IV | IV | III | $C^{s,hz}$ | w17 | w2 | 0.0017 | 0.0026 | 0 | 0.0161 | 0.0034 | 0.0860 |
| VIII | IV | III | $C^{hz}$, $A^{hz}$ | w17 | w2 | 0.0475 | 0.0340 | 0.0297 | 0.0484 | 0.0404 | 0.0040 |
| XV | VI | IV | $A^{s,hz}$ | w18 | w4 | 0.0136 | 0.0157 | 0.0197 | 0.2043 | 0.0397 | – |
| XXIII | VI | IV | $A^{hz}$ | 3 | w4 | 0 | 0.0340 | 0.0164 | 0 | 0.0123 | – |
| XII | V | V | $C^{s,hz}$ | 4 | w7 | 0 | 0 | 0.0066 | 0 | 0.0014 | 0.0640 |
| XII | V | IV | $C^{s,hz}$, $A^{hz}$ | 4 | w8 | 0.0356 | 0.0183 | 0.0230 | 0.0323 | 0.0280 | 0.1420 |
| XII | V | VI | $A^{hz}$ | 4 | w2 | 0.0712 | 0.0026 | 0.0033 | 0 | 0.0301 | – |
| XVIII | V | IV | $A^{hz}$ | 4 | w8 | 0.0051 | 0.0079 | 0.0033 | 0.0054 | 0.0055 | – |
| XIII | V | IV | $C^{s,hz}$ | 7 | w9 | – | – | – | – | – | 0.0240 |
| XIII | V | VI | $C^{s,hz}$, $A^{hz}$ | 7 | w2 | 0.0237 | 0.0445 | 0.0164 | 0.0269 | 0.0280 | 0.0120 |
| XIV | V | VI | $C^{s,hz}$, $A^{hz}$ | 7 | w2 | 0.1153 | 0.0236 | 0.0099 | 0.0054 | 0.0554 | 0.0220 |
| XIV | V | IX | $A^{hz}$ | 7 | w2 | 0.0051 | 0.0026 | 0.0033 | 0 | 0.0034 | – |
| XIX | V | VI | —‖ | 7 | w2 | 0 | 0 | 0 | 0.0054 | 0.0007 | – |
| V | VI | IV | $C^{s,hz}$ | w8 | w4 | 0 | 0 | 0 | 0.0054 | 0.0007 | 0.0480 |
| V | VI | V | $A^{hz}$ | w8 | w7 | 0.0678 | 0.0236 | 0.0164 | 0.0215 | 0.0397 | 0.0020 |
| V | IV | V | $A^{hz}$ | w8 | w7 | 0.0186 | 0.0131 | 0.0395 | 0.0215 | 0.0219 | – |
| V | II | II | $A^{hz}$ | w8 | w6 | 0.0034 | 0.0157 | 0.0164 | 0.0054 | 0.0096 | – |
| V | II | I | $A^{hz}$ | w8 | w1 | 0.0017 | 0 | 0.0033 | 0 | 0.0014 | – |
| XIII | V | IV | $C^{s,hz}$ | 9 | w9 | – | – | – | – | – | 0.0220 |
| XVII | V | IX | $A^{s,hz}$ | 9 | w2 | 0.0475 | 0.0445 | 0.1020 | 0.0430 | 0.0575 | – |
| III | I | I | $C^{s,hz}$, $A^{hz}$ | w10 | w5 | 0.0966 | 0.0471 | 0.1151 | 0.0161 | 0.0773 | 0.0100 |
| VI | IV | V | $C^{s,hz}$, $A^{hz}$ | w11 | w7 | 0.0169 | 0.0314 | 0.0329 | 0.0108 | 0.0233 | 0.0700 |
| VI | III | II | — | w11 | w6 | – | – | – | – | – | 0 |
| VI | II | II | $A^{hz}$ | w11 | w6 | 0 | 0 | 0.0066 | 0.0215 | 0.0041 | – |
| VI | II | I | —‖ | w11 | w1 | 0.0017 | 0.0026 | 0 | 0 | 0.0014 | – |
| XX | IV | V | $A^{hz}$ | w11 | w7 | 0.0254 | 0.0052 | 0.0197 | 0.0054 | 0.0164 | – |
| XX | II | II | $A^{hz}$ | w11 | w6 | 0.0051 | 0.0026 | 0 | 0.0108 | 0.0041 | – |
| XX | II | I | $A^{hz}$ | w11 | w6 | 0.0085 | 0.0105 | 0.0099 | 0 | 0.0082 | – |
| XXI | IV | V | $A^{s,hz}$ | w11 | w7 | 0.1305 | 0.3586 | 0.2105 | 0.0484 | 0.1963 | – |
| XXI | II | II | $A^{hz}$ | w11 | w6 | 0.0017 | 0 | 0.0066 | 0.0108 | 0.0034 | – |
| XXII | II | I | $A^{hz}$ | w11 | w6 | 0.0017 | 0.0079 | 0.0132 | 0.0108 | 0.0068 | – |
| VII | IV | V | $C^{s,hz}$, $A^{hz}$ | w12 | w7 | 0.0085 | 0.0079 | 0.0066 | 0.0269 | 0.0103 | 0.0300 |
| VIII | III | II | $C^{s,hz}$ | w13 | w6 | 0.0017 | 0 | 0 | 0 | 0.0007 | 0.0420 |
| VIII | I | II | $A^{s}$ | w13 | w6 | 0 | 0.0026 | 0 | 0.0108 | 0.0021 | – |
| IX | III | II | $C^{s,hz}$, $A^{hz}$ | w13 | w6 | 0.0373 | 0.0079 | 0.0197 | 0.0161 | 0.0233 | 0.0500 |
| IX | III | I | $A^{hz}$ | w13 | w1 | 0.0288 | 0 | 0 | 0 | 0.0116 | – |
| IX | V | IV | $C^{s}$ | w13 | w3 | 0.0017 | 0 | 0 | 0 | 0.0007 | 0 |
| X | II | I | $C^{s,hz}$, $A^{hz}$ | w13 | w6 | 0.0542 | 0.1152 | 0.1645 | 0.1237 | 0.1019 | 0.0560 |
| XI | III | II | $A^{hz}$ | w13 | w6 | 0.0288 | 0.0052 | 0.0033 | 0 | 0.0137 | – |
| XV | IV | V | $C^{hz}$, $A^{hz}$ | w13 | w7 | 0.0153 | 0.0314 | 0.0329 | 0.0161 | 0.0233 | 0.0080 |
| XV | V | VI | —‖ | w13 | w2 | 0 | 0.0052 | 0 | 0 | 0.0014 | – |
| XI | I | I | $C^{s,hz}$, $A^{hz}$ | w14 | w5 | 0.0017 | 0.0157 | 0 | 0 | 0.0048 | 0.0320 |
| IX | IV | V | $C^{s,hz}$ | w14 | w7 | 0 | 0 | 0.0066 | 0 | 0.0014 | 0 |
| Impossible to type by *Taq* I RFLP analysis** | | | | | | 0.0271 | 0.0209 | 0.0132 | 0.0108 | 0.0205 | 0 |

In Caucasians, *DR–DQ* haplotypes can be identified without family data due to the tight linkage disequilibrium among *DRB*, *DQA*, and *DQB* loci (10–12). In West Africans, two haplotypes were assigned if the *DRB*, *DQA*, and *DQB* patterns constituted two haplotypic combinations that were well-known in Caucasians or (in the case of new, non-Caucasian *DRB–DQA–DQB* haplotypes) if these could be confirmed in the African population by segregation analysis or homozygosity (with a few exceptions, see footnote ‖).

*Allelic *Taq* I *DRB*, *DQA*, and *DQB* restriction fragment patterns, with local RFLP nomenclature in Roman numerals, are illustrated in Fig. 2 (*DRB*) and in ref. 10 (*DQA* and *DQB*), except *DQA* pattern VI (6.4-kb fragment) and *DQB* patterns VIII and IX (4.7-kb fragment and 3.9-, 2.4-, and 1.5-kb fragments, respectively).

†*DRB–DQA–DQB* haplotypes were confirmed by segregation analysis ($^{s}$) or homozygosity ($^{hz}$), in Caucasians (C) and Africans (A).

‡The serologically defined DR and DQ specificities associated with specific *Taq* I *DRB–DQA–DQB* haplotypes in Caucasians are given in refs. 10–12. In Africans, the association between allelic RFLP patterns and serological specificities was determined by the microlymphocytotoxicity typing technique in individuals representing all the new *Taq* I RFLP patterns and most of the new haplotypic combinations. Several of the non-Caucasian DRw11-associated *DRB–DQA–DQB* haplotypes also showed some reactivity with DRw13 sera. Our data confirmed the serological associations found by other investigators (17–20).

§The distribution of *Taq* I *DRB–DQA–DQB* haplotypes in 250 randomly selected healthy Swedes is given for comparison; $n$ denotes the number

Africans but only 26 in North Europeans. All but 3 of the Caucasian haplotypes were observed in West Africans. The 22 *DRB–DQA–DQB* haplotypes seen only in West Africans have not been observed in North Europeans, South Europeans, North American Caucasians, or Asians (of whom 3000, 300, 100, and 200 have been investigated, respectively). The great difference in the degree and distribution of DR–DQ polymorphism in West Africans and Caucasians can be exemplified by the observation that of the 10 most frequent West African *DRB–DQA–DQB* haplotypes, only 6 have been encountered in North European Caucasians, and only 2 of these 6 shared *DRB–DQA–DQB* haplotypes belong to the 10 most frequently occurring Swedish haplotypes.

The genetic distances between Swedes and the four African ethnic groups were, as expected, larger than the genetic differences between the latter. Each of the West African groups exhibit more DR–DQ polymorphism than North European Caucasians (Table 1). Thus, the observed pronounced DR–DQ gene polymorphism in West Africans is not an effect of adding limited, but different, variability in several ethnic isolates.

The linkage disequilibrium between the *DR* and *DQ* subregions is extremely tight and recombinations between them or between the *DQA1* and *DQB1* loci have not been documented. It has been suggested that the *DR–DQ* region might have evolved as a "frozen block" since the separation of the human races (26). Our RFLP data indicate that recombinations indeed have occurred between *DR* and *DQ* after the emergence of class II alleles. Recombinations may well have occurred after speciation and the African/non-African split and might also have taken place within the *DR* and *DQ* subregions. In the latter case, recombination is probably restricted by structural and/or functional constraints on DQ $\alpha$–$\beta$ association.

Most RFLPs are located in noncoding regions. Consequently, new RFLP patterns and new RFLP-defined haplotypes might not represent expressed allelic differences. However, in Caucasians all *Taq* I *DRB–DQA–DQB* haplotypes (10–12) are associated with coding differences in at least one locus. Further, concordance between the extent of RFLP-defined variability and functional polymorphism has been described for MHC class I loci (27). These observations, the findings of the present study, and serological (16) and cellular (14, 15) typing results indicate that many new HLA class II sequences or combinations of allelic sequences within a class II haplotype will be found in populations of African origin.

**A Population Bottleneck in the Emergence of the White Race?** Current archaeological data place the origin of modern humans in southern and eastern Africa >100,000 years ago (28). Subsequent migration out of Africa resulted in the separation of the human races. The reconstruction of human phylogeny from contemporary genetic information as well as linguistic data confirms the separation of Africans from non-Africans (29). Africans exhibit the greatest diversity in mitochondrial DNA (30) and, as shown in the present study, in HLA class II RFLPs, which further supports the concept that Africans are the oldest and genetically most diverse human population.

In West Africans we have identified almost 50% more RFLP-defined *DRB* alleles and close to twice as many *DRB–DQA–DQB* haplotypes as in North European Caucasians. Considering the old age of the major MHC allelic types (2–5), it seems unlikely that Africans alone would have accumulated a considerable amount of HLA class II polymorphism during the last 50,000–100,000 years. As there has been no documented decrease in the population size of Caucasians drastic enough to cause loss of a large number of alleles/haplotypes, our findings can best be explained by a bottleneck in the emergence of the White race—i.e., the founding population of Caucasians was not large enough to maintain the total pool of HLA class II alleles.

**Positive Selection for HLA Class II Polymorphism.** Two main lines of evidence support the idea that HLA polymorphism is maintained by balancing selection: (*i*) the allelic frequency distributions are not neutral since observed homozygosity values for HLA class I and II loci are below neutrality expectations (31), and (*ii*) codons of contact amino acids of the antigen-binding site of MHC class I and II molecules accumulate more replacement substitutions than silent substitutions, whereas the opposite is true for other codons of MHC genes (6–8). Consequently, speculations on the evolutionary forces driving MHC polymorphism must take into account that selection acts directly on the antigen-binding site (i.e., probably operates via the main immunological functions of MHC molecules).

Our large groups of genotyped West Africans and Swedes, in which homozygosity can be directly observed, allow a more firm rejection of the neutrality hypothesis than previous studies, in which the Hardy–Weinberg estimate of homozygosity in phenotyped populations has been used. *DRB*, *DQA*, and *DQB* alleles as well as *DQA–DQB* combinations were found to experience positive selection ($P < 10^{-3}$–$10^{-5}$), whereas *DRB–DQA–DQB* haplotypes were selectively neutral (Fig. 3). The absence of positive selection for *DRB–DQA–DQB* homozygotes suggests genetic and/or functional constraints on *DR–DQ* recombinations on the evolutionary conserved *DRB–DQA–DQB* haplotypes.

**What Drives the Selection of HLA Class II Polymorphism?** Four mechanisms, not mutually exclusive, have been suggested to maintain HLA polymorphism: (*i*) overdominant selection (33, 34), the selective advantage of a heterozygote over either homozygote; (*ii*) frequency-dependent selection (35), the decreased fitness of an allele with increased frequency; (*iii*) mating preference (36); and (*iv*) selective abortion of histocompatible fetuses (37). Empirical findings (6–8, 31, 33, 34) as well as computer simulations (38) favor the overdominant model. However, frequency-dependent selection by minority advantage—i.e., the selective advantage of rare alleles—cannot be ruled out (38).

What evolutionary forces exert the selection pressure? Doherty and Zinkernagel (33) suggested that the overdominant selection of MHC polymorphism was parasite-driven. To us this seems unlikely to be the main explanation for the following reasons. (*i*) Many microorganisms are able to kill infected individuals before an MHC-restricted specific immune response can be mounted. (*ii*) Despite many attempts, associations between HLA alleles and susceptibility to infectious diseases have not been demonstrated. (*iii*) Although malaria is a major cause of child death in Africa and exerts a strong selection pressure (e.g., in the maintenance of the sickle cell

of chromosomes investigated. 0, haplotype has been observed once or a few times in the more than 3000 Caucasians of North European extraction that have been typed in our laboratory (O.O.); −, haplotype has never been observed in these individuals.

¶Gambians plus Liberians.

‖In five individuals, representing three new rare haplotypes, the *Taq* I RFLP patterns made up one established, frequently occurring Caucasian or African haplotype, as well as a new haplotypic combination. However, these three new haplotypes were not confirmed by segregation or homozygosity.

**The heterozygous combination of *DRB* patterns IV/XI, IV/XXI, and IX/XXI could not always be distinguished from VIII/IX, VIII/XV and XI/XV, respectively, even when *DQA* and *DQB* linkage disequilibria were considered.

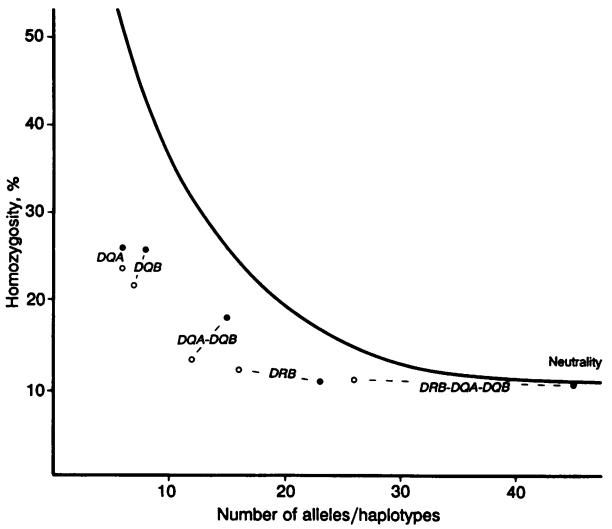


FIG. 3. Frequencies (%) of homozygosity for *Taq* I *DRB*, *DQA*, and *DQB* allelic patterns, *DQA–DQB* combinations, and *DRB–DQA–DQB* haplotypes in West Africans (●) and North European Caucasians (○). Values below neutrality expectations indicate heterozygote advantage: $P < 10^{-3}$ for *DRB* alleles; $P < 10^{-5}$ for *DQA* and *DQB* alleles; $P < 10^{-3}$ and $P < 10^{-5}$ for *DQA–DQB* combinations in West Africans and Swedes, respectively; and *P* not significant for *DRB–DQA–DQB* haplotypes. Tables of the expected values of homozygosity under neutrality (32) were compared with the observed frequencies of homozygosity. The homozygosity is conditional on the number of alleles ($k$) and the sample size ($2n$). Values were linearly extrapolated for $k > 30$. The neutrality curve is for $2n = 500$. With increasing sample size the expected homozygosity increases; i.e., for West Africans the deviations from neutrality were underestimated.

trait), no close association between HLA alleles and responsiveness to immunodominant B- and T-cell epitopes of major malaria antigens has been found (reviewed in ref. 39). Furthermore, data from monozygotic twins, HLA-identical siblings, and HLA-compatible individuals show that the B- and T-cell responses to immunodominant malaria antigens are genetically influenced, but by non-HLA-linked genes (41).

We suggest instead that one of the evolutionary forces causing overdominant selection of HLA polymorphism might be a decreased probability of potentially autoreactive T-cell clones escaping negative thymic selection during T-cell maturation in HLA heterozygotes. This hypothesis is supported by the observation that negative selection is a codominant genetic trait (40). A theory of overdominance based on thymic selection is compatible with the notion that evolutionary forces maintaining HLA polymorphism act directly on the antigen-binding site.

We thank L. Ohlson for expert technical assistance. This study was supported by grants from the United Nations Development Program/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases, the Swedish Agency of Research Cooperation with Developing Countries, and the Swedish Medical Research Council.

1. Klein, J. (1980) in *Immunology 80*, eds. Fongereau, K. & Dausset, J. (Academic, London), pp. 239–253.
2. Figueroa, F., Günther, E. & Klein, J. (1988) *Nature (London)* **335**, 265–267.
3. McConnell, T. J., Talbot, W. S., McIndoe, R. A. & Wakeland, E. K. (1988) *Nature (London)* **332**, 651–653.
4. Lawlor, D. A., Ward, F. E., Ennis, P. D., Jacksom, A. P. & Parham, P. (1988) *Nature (London)* **335**, 268–271.
5. Gyllensten, U. B., Lashkari, D. & Erlich, H. A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1835–1839.
6. Gustafsson, K., Wiman, K., Emmoth, E., Larhammar, D., Böhme, J., Hyldig-Nielsen, J.-J., Ronne, H., Peterson, P. A. & Rask, L. (1984) *EMBO J.* **3**, 1655–1661.
7. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335**, 167–170.
8. Hughes, A. L. & Nei, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 958–962.
9. Wake, C. T., Long, E. O. & Mach, B. (1982) *Nature (London)* **300**, 372–374.
10. Carlsson, B., Wallin, J., Böhme, J. & Möller, E. (1987) *Hum. Immunol.* **20**, 95–113.
11. Bidwell, J. L., Bidwell, E. A., Savage, D. A., Middleton, D., Klouda, P. T. & Bradley, B. A. (1988) *Transplantation* **45**, 640–646.
12. Bidwell, J. L., Bidwell, E. A. & Bradley, B. A. (1990) in *Blood Transfusion: The Impact of New Technologies*, ed. Contreras, M. (Baillière, London), pp. 355–384.
13. Olerup, O., Hillert, J., Fredrikson, S., Olsson, T., Kam-Hansen, S., Möller, E., Carlsson, B. & Wallin, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7113–7117.
14. Johnson, A. H., Rosen-Bronson, S. & Hurley, C. K. (1989) *Transplant. Proc.* **21**, 3872–3873.
15. Oudshoorn, M., du Toit, E. D., Taljaard, D. G. & MacGregor, K. J. (1986) *Hum. Immunol.* **17**, 273–287.
16. du Toit, E. D., MacGregor, K. J., Taljaard, D. G. & Oudshoorn, M. (1987) *Tissue Antigens* **31**, 109–125.
17. Lee, K. W., Hurley, C. K., Hartzman, R. & Johnson, A. H. (1990) *Hum. Immunol.* **29**, 202–219.
18. Martell, R. W., Oudshoorn, M., May, R. M. & du Toit, E. D. (1989) *Hum. Immunol.* **26**, 237–244.
19. Martell, R. W., Oudshoorn, M., Arendse, B. & du Toit, E. D. (1990) *Hum. Immunol.* **28**, 32–38.
20. Jenkins, D., Fletcher, J., Mijovic, C., Bradwell, A. R. & Barnett, A. H. (1990) *Mol. Immunol.* **3**, 297–302.
21. Germain, R. N., Bentley, D. M. & Quill, H. (1985) *Cell* **43**, 233–242.
22. Sant, A. J., Braunstein, N. S. & Germain, R. N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8065–8069.
23. Giles, R., Demars, R., Chang, C. C. & Capra, J. D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1776–1780.
24. Kwok, W. W., Thurtle, P. & Nepom, G. T. (1989) *J. Immunol.* **143**, 3598–3601.
25. Le Meur, M., Waltzinger, C., Gerlinger, P., Benoist, C. & Mathis, D. (1989) *J. Immunol.* **142**, 323–327.
26. Baur, M. P., Neugebauer, M., Deppe, H., Sigmund, M., Luton, T., Mayr, W. R. & Albert, E. D. (1984) in *Histocompatibility Testing 1984*, eds. Albert, E. D., Baur, M. P. & Mayr, W. R. (Springer, Berlin), pp. 333–341.
27. Yuhki, N. & O'Brien, S. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 836–840.
28. Stringer, C. B. & Andrews, P. (1988) *Science* **239**, 1263–1268.
29. Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. & Mountain, J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6002–6006.
30. Cann, R. L., Stoneking, M. & Wilson, A. C. (1987) *Nature (London)* **325**, 31–36.
31. Klitz, W., Thomson, G. & Baur, M. P. (1984) in *Histocompatibility Testing 1984*, eds. Albert, E. D., Baur, M. P. & Mayr, W. R. (Springer, Berlin), pp. 330–332.
32. Anderson, R. (1979) M.Sc. thesis (Monash University, Clayton, Victoria, Australia).
33. Doherty, P. C. & Zinkernagel, R. M. (1975) *Lancet* **i**, 1406–1409.
34. Doherty, P. C. & Zinkernagel, R. M. (1975) *Nature (London)* **256**, 50–52.
35. Bodmer, W. F. (1972) *Nature (London)* **237**, 139–145.
36. Yamazaki, K., Boyse, E. A., Mike, V., Thlaer, H. T., Mathiesen, B. J., Abbott, J., Boyse, J., Zayas, Z. A. & Thomas, L. (1975) *J. Exp. Med.* **144**, 1324–1335.
37. Clarke, B. & Kirby, D. R. S. (1966) *Nature (London)* **211**, 999–1000.
38. Takahata, N. & Nei, M. (1990) *Genetics* **124**, 967–978.
39. Riley, E. M., Olerup, O. & Troye-Blomberg, M. (1991) *Parasitol. Today* **7**, 5–11.
40. Blackman, M., Kappler, J. & Marrack, P. (1990) *Science* **248**, 1335–1341.
41. Troye-Blomberg, M., Olerup, O., Larsson, Å., Sjöberg, K., Perlmann, H., Lepers, J.-P. & Perlmann, P. (1991) *Int. Immunol.* **3**, in press.