Evolution of HLA Class II Molecules: Allelic and Amino Acid Site Variability Across Populations

Hugh Salamon,*,[1] William Klitz,*,[2] Simon Easteal,[†] Xiaojiang Gao,[†,3] Henry A. Erlich,[‡] Marcello Fernandez-Viña,[§] Elizabeth A. Trachtenberg,[‡,] Shannon K. McWeeney,* Mark P. Nelson* and Glenys Thomson***

**Department of Integrative Biology, University of California, Berkeley, California 94720-3140, [†]Human Genetics Group, John Curtin School of Medical Research, The Australian National University, Canberra, ACT 0200, Australia, [‡]Department of Human Genetics, Roche Molecular Systems, Alameda, California 94501, [§]American Red Cross, National Histocompatibility Laboratory, Baltimore, Maryland 21201 and **Children's Hospital Oakland Research Institute, Oakland, California 94609*

Manuscript received July 6, 1998
Accepted for publiction February 18, 1999

ABSTRACT

Analysis of the highly polymorphic β1 domains of the HLA class II molecules encoded by the *DRB1*, *DQB1*, and *DPB1* loci reveals contrasting levels of diversity at the allele and amino acid site levels. Statistics of allele frequency distributions, based on Watterson's homozygosity statistic *F*, reveal distinct evolutionary patterns for these loci in ethnically diverse samples (26 populations for *DQB1* and *DRB1* and 14 for *DPB1*). When examined over all populations, the *DQB1* locus allelic variation exhibits striking balanced polymorphism ($P < 10^{-4}$), *DRB1* shows some evidence of balancing selection ($P < 0.06$), and while there is overall very little evidence for selection of *DPB1* allele frequencies, there is a trend in the direction of balancing selection ($P < 0.08$). In contrast, at the amino acid level all three loci show strong evidence of balancing selection at some sites. Averaged over polymorphic amino acid sites, *DQB1* and *DPB1* show similar deviation from neutrality expectations, and both exhibit more balanced polymorphic amino acid sites than *DRB1*. Across ethnic groups, polymorphisms at many codons show evidence for balancing selection, yet data consistent with directional selection were observed at other codons. Both antigen-binding pocket- and non-pocket-forming amino acid sites show overall deviation from neutrality for all three loci. Only in the case of *DRB1* was there a significant difference between pocket- and non-pocket-forming amino acid sites. Our findings indicate that balancing selection at the MHC occurs at the level of polymorphic amino acid residues, and that in many cases this selection is consistent across populations.

THE function of human leukocyte antigen (HLA) class I and class II molecules is to collect peptide fragments inside the cell and transport them to the cell surface, where the peptide-HLA complex is surveyed by immune system T cells (see, *e.g.*, Monaco 1993). Three regions of HLA class II genes produce functional antigen-presenting heterodimers; these are labeled *DR*, *DQ*, and *DP*. Each class II heterodimer is made up of the noncovalent association of two glycopeptide chains: the α chain and the β chain, encoded by, *e.g.*, for *DQ*, the *DQA1* and *DQB1* loci, respectively. T cells recognize both MHC and foreign antigen presented together as a complex.

Specificity and affinity for peptides by the antigen recognition sites (ARS) are determined by contacts with the amino acid side chains of the bound peptide and thus are due to the amino acid sequences of both the bound peptide and the HLA allelic variant. Definable pockets in the ARS of the HLA molecules bind to the antigenic peptide side chains to create binding specificity.

The genetic variation of HLA, and the major histocompatibility complex (MHC) regions in other species' genomes, is especially striking in the amino acid residues of the ARS (see, *e.g.*, Parham *et al.* 1989; Lawlor *et al.* 1990; Hedrick *et al.* 1991). High heterozygosity at positions critical to antigen recognition suggests positive selection in favor of diversity. This is also indicated by analyses of Hughes and Nei (1988, 1989) and Takahata *et al.* (1992), which show that in the ARS, nonsynonymous differences between alleles are proportionally more common than synonymous ones. In the rest of the molecule, there is an excess of synonymous changes, as is normally observed in protein coding sequences.

The neutral theory of evolution at a locus assumes that all alleles are selectively equivalent, that mutation is the source of new genetic variation, and that genetic drift results in loss of variation. Ewens (1972) developed

Corresponding author: Glenys Thomson, Department of Integrative Biology, 3060 Valley Life Sciences Bldg., MC#3140, University of California, Berkeley, CA 94720-3140.
E-mail: glenys@allele5.biol.berkeley.edu

[1] *Present address:* Division of Infectious Diseases and Geographic Medicine, Stanford University, Stanford, CA 94305.

[2] *Present address:* Department of Public Health, University of California, Berkeley, CA 94720.

[3] *Present address:* Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702.

a sampling theory that predicts the distribution of alleles at a locus drawn from a population at equilibrium under the infinite alleles model of neutrality. Watterson (1978) developed a method, using a homozygosity measure, denoted by *F*, and the observed number of alleles and sample size, to compare a sample of allele frequencies to neutrality expectations. For a given number of observed alleles, a low value of the homozygosity measure *F*, compared to the neutrality expectation, corresponds to all alleles having similar frequencies. Some form of balancing selection is implied when the observed allele frequencies are more even than the neutrality expectation (Watterson 1978). A high value of *F* compared to the neutrality expectation corresponds to one allele predominating with the others at low frequencies. Directional selection is implicated in this case. Note that the fit of a sample of allele frequencies to neutrality expectations does not mean that selection is not operating, but simply that one cannot reject the neutral model.

Previous analyses of serologically defined class I (*A*, *B*, and *C*) and class II (*DR*) alleles and of polymerase chain reaction/sequence-specific oligonucleotide probe (PCR/SSOP) typed class II (*DRB1* and *DQB1*) alleles show that the alleles at these loci are often more even in frequency than expected under neutrality, which implies some form of balancing selection (Hedrick and Thomson 1983; Klitz *et al.* 1986, 1992; Begovich *et al.* 1992). However, in many, but not all, populations studied to date, the distribution of *DPB1* alleles does not differ significantly from neutrality expectations, despite the molecular similarity of the encoded proteins and proximity of the locus to *DR* and *DQ* genes (Begovich *et al.* 1992).

The availability of population data from a number of ethnic groups with PCR/SSOP molecular-defined variation at the HLA class II *DRB1*, *DQB1*, and *DPB1* loci has allowed us to compare and contrast genetic variation among populations at both the allelic and amino acid levels. Using Watterson's homozygosity statistic, the allele frequency distributions were tested for departure from neutral expectations. We also investigated amino acid frequency distributions with respect to neutrality expectations and determined if the results corresponded with the findings for the allele level investigation. In addition, genetic variation at the amino acid level was examined for the different functional categories of polymorphic amino acid sites.

MATERIALS AND METHODS

The Watterson homozygosity test of neutrality: The homozygosity measure for the allele frequencies at a locus is denoted by F_A:

$$F_A = \sum_{i=1}^{k} p_i^2, \tag{1}$$

where p_i is the observed frequency of allelic type i, $i = 1, 2, \ldots, k$, in a sample of size n individuals, *i.e.*, a total of $2n$ alleles (Watterson 1978). Note that the Watterson homozygosity test of neutrality does not examine deviations from Hardy-Weinberg proportions, but uses the expected Hardy-Weinberg homozygosity as a measure of the distribution of allelic frequencies for a given sample size ($2n$) and observed number of alleles (k). Values of F_A range from 0 to 1. Several confidence levels for the distribution of F_A under neutrality are available for combinations of $2n$ and k up to 500 and 40, respectively (Anderson 1979; also see Ewens 1979). Individual values of F_A are tested for significant deviation from neutrality by comparison to quantiles of the neutral distribution (matching for k and $2n$) as calculated in Anderson (1979). If the exact combinations of $2n$ and k in the HLA data are not observed in the table, values are linearly interpolated. In cases in which $2n > 500$, values for 500 are used. This gives a conservative value when the alternative hypothesis is balancing selection, where the homozygosity observed is expected to be lower than the neutrality expectation for the given number of alleles.

To allow comparison of homozygosity *F* statistics across populations with different numbers of alleles and sample sizes, tabled values of the mean and variance of F_A (Anderson 1979) were used to calculate the normalized deviate, Fnd_A;

$$\mathrm{Fnd}_A = \frac{F_{A\ \mathrm{observed}} - F_{A\ \mathrm{expected}}}{[\mathrm{Var}(F_{A\ \mathrm{expected}})]^{1/2}} \tag{2}$$

for each sampled population. A negative value of Fnd_A is in the direction of balancing selection, a value of zero is the neutrality expectation, and a positive value is in the direction of directional selection. The mean of Fnd_A across m populations, $\overline{\mathrm{Fnd}}_A$, was calculated

$$\overline{\mathrm{Fnd}}_A = \frac{\sum \mathrm{Fnd}_{Aj}}{m}, \tag{3}$$

where Fnd_{Aj} is the normalized deviate for the jth population, $j = 1, 2, \ldots, m$. By the central limit theorem,

$$\overline{\mathrm{Fnd}}_A \sim N(0, 1/m). \tag{4}$$

The estimated standard error of Fnd_A is given by s_A, where

$$S_A^2 = \frac{\sum (\mathrm{Fnd}_{Aj} - \overline{\mathrm{Fnd}}_A)^2}{m - 1}. \tag{5}$$

The overall test of $\overline{\mathrm{Fnd}}_A$ is a two-tailed test using $z = \overline{\mathrm{Fnd}}_A \sqrt{m}/S_A$.

Neutrality testing at the amino acid level is analogous to that for allelic variation. The homozygosity statistic for amino acid polymorphism is calculated as

$$F_{aa} = \sum_{i=1}^{t} q_i^2, \tag{6}$$

where q_i is the observed frequency of amino acid residue i, with a total of t amino acids, $i = 1, 2, \ldots, t$, at the site under consideration, in a sample of size n individuals, *i.e.*, a total of $2n$ amino acid "alleles." The normal deviate for f_{aa} denoted Fnd_{aa} is calculated as for allelic variation (see Equation 2). The Fnd_{aa} can be averaged across m populations, or across r polymorphic sites, and the statistical properties and tests of the averages are as in Equations 3–5.

Population samples: Allele and amino acid variation are examined for *DRB1* and *DQB1* β1 domains in 23 human populations, including 5 Melanesian: Papua New Guinea Highlanders (from Goroka) and populations from Madang, Rabaul, New Caledonia, and Fiji (Gao *et al.* 1992a); 3 Polynesian populations, Rarotonga, Niue, and West Samoa (Gao *et al.* 1992b);

2 Micronesian populations, Nauru and Kiribati (Gao *et al.* 1992b); two different Indonesian populations, here designated Indonesia and Java (Gao *et al.* 1992b); a "Southern" Filipino population (Bugawan *et al.* 1994); Caucasian populations, the Centre d'Etude du Polymorphisme Humain (CEPH; Begovich *et al.* 1992), Czech (Cerna *et al.* 1992), and Norwegian (Rønningen *et al.* 1990); 2 Chinese populations, southern Chinese from Gaunzou and northern Chinese from Beijing (both supplied by M. Fernandez-Viña); a Senegalese population (10th HLA workshop data); North Americans of African descent (Fernandez-Viña *et al.* 1991); a population sample from the United States of individuals of Mexican origin (Erlich *et al.* 1993); and 2 different native South American populations, the Cayapa in Ecuador (Trachtenberg *et al.* 1995) and the Toba of the Gran Chaco in Argentina (Cerna *et al.* 1993).

Variation in the *DPB1*-encoded β1 domain is considered in the following 14 populations: Cayapa, CEPH, Czech, Indonesia, Japan (Mitsunaga *et al.* 1992), Mexican (U.S.), Nauru (supplied by S. Easteal), Northern China (Beijing), Norway, Papua New Guinea Highlanders from Goroka (supplied by S. Easteal), Southern China (Gaunzou), Senegal, Filipino, and Toba. *DPB1* populations unreferenced in this list are from the same sources as in the *DRB1* and *DQB1* list. All of the populations were typed using PCR/SSOP molecular typing methods.

Amino acid sequences were obtained from the EMBL ftp server at ftp://FTP.EMBL-Heidelberg.DE/ (sequences as in Marsh and Bodmer 1993, 1995; Bodmer *et al.* 1997). Amino acid sites of the β1 domains were considered: sites 6–90, inclusive for *DQB1* and *DRB1*, and 8–90 for *DPB1*. All alleles studied were defined at the amino acid level, by molecular, not serological typing.

Pocket *vs.* nonpocket categories of amino acid sites: Sites were classified on the basis of a *DRB1* structure (Brown *et al.* 1993) and on the peptide contacts observed in *DR1*-influenza peptide complex (Stern *et al.* 1994). All polymorphic sites that contact antigenic peptide side-chains were assigned to the pocket category. All other sites are considered to be in the nonpocket category.

The *DR1* structure was used as a model for class II peptide binding. Sixteen amino acid sites in *DRB1* and *DQB1* were classified as participating in pockets: 9, 11, 13, 28, 47, 57, 61, 67, 70, 71, 74, 78, 85, 86, 89, and 90. The sites that appear homologous (in the sense of serial homology) to these in *DPB1* are 9, 11, 13, 26, 45, 55, 59, 65, 68, 69, 72, 76, 83, 84, 87, and 88. The differences in site designations are due to a relative deletion in *DPB1* as compared to *DQB1* and *DRB1*. All other amino acid sites were classified as nonpocket.

RESULTS

Allelic variation in *DRB1*, *DQB1*, and *DPB1*: The homozygosity statistic F_A and the normalized deviate Fnd_A for the *DRB1* locus show evidence of balancing selection (Table 1). Two populations (African-American and Norway) show individually significant negative (balancing selection) values for Fnd_A. Only 5 of the 23 populations show positive Fnd_A values, which is significant at $P < 0.005$ by a sign test from the neutrality 50% expectation of an equal number of positive and negative values. Thus more populations show deviation in the direction of balancing selection than expected by chance. For *DRB1* alleles the mean of the normal deviate of the homozygosity statistic $\overline{Fnd}_A$ is marginally significantly different from neutrality expectations ($P < 0.06$) in the direction of balancing selection. However, note that 2 independently sampled Indonesian populations, Indonesia and Java, respectively, show large positive values with Java significantly different from neutrality, which is indicative of directional selection.

DQB1 shows the strongest evidence for balancing selection of the three loci investigated (Table 2). Six of the 23 populations exhibit individually significant homozygosity values below neutral expectations. Only 3 out of the 23 populations exhibit positive Fnd_A values, and the test of signs under neutrality gives $P < 0.0002$. The mean across populations of the normal deviates, $\overline{Fnd}_A$, is well below neutral expectations and highly statistically significant ($P \ll 10^{-4}$). Despite the strong linkage disequilibrium between *DRB1* and *DQB1*, the sample designated Indonesia exhibits a negative value for Fnd_A at *DQB1* (Table 2) and a positive one at *DRB1* (Table 1). On the other hand, the Java sample does in fact exhibit one of the few positive Fnd_A values observed at *DQB1*, as well as at *DRB1*.

For *DPB1*, the normal deviates of homozygosity, Fnd_A, show a scatter of values spanning zero (the neutrality expectation) with five positive and nine negative (Table 3). No single population exhibits a value of homozygosity significantly different from neutrality. However, the mean value (across all populations) of the normal deviates of homozygosity, $\overline{Fnd}_A$, is negative and in the direction of balancing selection ($P \leq 0.08$).

To summarize, at the level of allelic variation, *DQB1* shows strong evidence of balancing selection. Balancing selection is implicated at *DRB1*, although the evidence is much less striking than with *DQB1*. The overall picture for *DPB1* is in agreement with neutral theory predictions, although the mean across populations of the normal deviates shows a trend in the direction of balancing selection.

Variation at amino acid sites: For each population, the frequency distributions of variants at every polymorphic amino acid in the β1 domain encoded by *DRB1*, *DQB1*, and *DPB1* were analyzed and the homozygosity statistics were calculated. The mean homozygosity F_{aa} and its normal deviate $\overline{Fnd}_{aa}$ were calculated including every polymorphic site across all populations sampled, and are presented in Figures 1, 2, and 3 for *DRB1*, *DQB1*, and *DPB1*, respectively. For all three loci many sites show strongly balanced polymorphism in the amino acids across human populations, which is indicated by negative values for $\overline{Fnd}_{aa}$ in Figures 1–3; 17/30 amino acid sites in *DRB1* (57%), 24/30 in *DQB1* (80%), and 13/17 in *DPB1* (76%) exhibit strongly balanced polymorphism. Significantly balanced polymorphism is observed both for amino acid sites classified as ARS and for sites that are not part of the ARS.

Other sites show neutral or even significantly positive values, the latter indicating directional selection (*e.g.*, the *DQB1* amino acid 40 has a positive $\overline{Fnd}_{aa}$ value of 0.081 from the 11 populations polymorphic at this resi-

TABLE 1

***DRB1* allele homozygosity statistics in 23 human populations**

Population[a]	k_A	$2n$	F_A	Fnd_A
Cayapa	13	200	0.175	−0.795
CEPH	26	248	0.091	−0.877
Czech	25	194	0.096	−0.676
Fiji	16	114	0.162	−0.019
Indonesia	18	226	0.278	1.524
Java	17	152	0.323	2.533 ($P < 0.05$)
Kiribati	10	124	0.233	−0.614
Madang	11	130	0.275	0.115
Mexican (U.S.)	32	232	0.057	−1.335
African American	23	121	0.069	−1.330 ($P < 0.01$)
Nauru	14	134	0.169	−0.520
New Caledonia	18	130	0.163	0.156
Northern China	26	182	0.082	−1.004
Niue	13	140	0.173	−0.670
Norway	17	362	0.103	−1.391 ($P < 0.01$)
Papua Highlands	11	114	0.162	−1.043
Rabaul	9	120	0.212	−0.921
Rarotonga	23	156	0.118	−0.134
Southern China	22	203	0.102	−0.831
Senegal	24	347	0.122	−0.501
Southern Filipino	22	210	0.188	0.986
Toba	22	274	0.135	−0.330
Western Samoa	16	102	0.108	−1.068
	$\overline{k_A} = 18.6$		$\overline{\mathrm{Fnd}}_A$	−0.381 ($P = 0.06$)
			$s_A/\sqrt{m}$ =	0.199

[a] For each population, the values of the following parameters are given: the number of alleles k_A in the sample of size $2n$, the homozygosity statistic F_A (Equation 1), and the normalized deviate Fnd (Equation 2). The mean values across populations, $\overline{k}_A$ and $\overline{\mathrm{Fnd}}_A$ are given, and the standard error of the mean estimate $s_A/\sqrt{m}$. Very rare alleles for which complete sequence information was not available have been removed from the analysis so that the $2n$ value may differ slightly from the total sample size.

due). Positive $\overline{\mathrm{Fnd}}_{aa}$ values for individual sites are often found in cases where many populations do not exhibit polymorphism at those sites at all. This is not surprising as the sites that show $\overline{\mathrm{Fnd}}_{aa}$ values consistent with directional selection are the same ones at which low frequency variants are easily missed in sampling.

We also observe that many amino acid sites maintain levels of balanced polymorphism across populations. Two sites within a locus may both present to us homozygosity values significantly and consistently lower than neutrality expectations, with the homozygosity values consistently higher at one site than the other across human populations. This phenomenon of distinct levels of homozygosity being maintained at certain amino acid sites across human populations is observed in all three loci.

DRB1 exhibits the most polymorphism as measured by numbers of variants segregating at a single site. For example, *DRB1* amino acid sites 11, 13, and 30 are observed with up to six amino acid variants in a single population sample. Note, however, that site 30 does not exhibit evidence for balanced polymorphism. Although *DRB1* sites exhibit many variants (mean = 3.26), the variants at *DPB1* (mean = 2.42) and *DQB1* (mean = 2.51) sites tend to have more even frequencies. However, in contrast to this general trend, note that the most balanced polymorphisms, *i.e.*, the lowest $\overline{\mathrm{Fnd}}_{aa}$ values at a single site, are observed in *DRB1* (*e.g.*, sites *DRB1* 12 and 67).

The mean value over amino acid sites of all mean normal deviates of homozygosity across populations in *DQB1* is not statistically different from that observed for *DPB1* (see Figures 2 and 3). *DPB1* and *DQB1* exhibit significantly more balanced amino acid polymorphisms than does *DRB1*. The mean normal deviates plus or minus one standard error of the mean value for each locus are $\overline{\mathrm{Fnd}}_{aa}$ (*DPB1*) = −0.959 ± 0.059 (205 observed polymorphisms), (*DQB1*) = −0.924 ± 0.034 (641 polymorphisms), and (*DRB1*) = −0.642 ± 0.039 (635 polymorphisms). These values are all significantly below neutrality expectations.

Polymorphic sites were found in both categories of sites (pockets and nonpockets), which allowed statistics for each category to be examined. Overall, the contrast

TABLE 2

***DQB1* allele homozygosity statistics in 23 human populations**

Population[a]	k_A	2n	F_A	Fnd_A
Cayapa	4	200	0.301	−1.664 ($P < 0.01$)
CEPH	12	248	0.135	−1.369 ($P < 0.01$)
Czech	12	194	0.163	−1.038
Fiji	11	114	0.242	−0.152
Indonesia	11	226	0.189	−0.970
Java	12	154	0.295	0.493
Kiribati	6	124	0.269	−1.156
Madang	10	130	0.356	0.671
Mexican (U.S.)	14	232	0.179	−0.652
African American	12	121	0.146	−1.097
Nauru	10	134	0.246	−0.425
New Caledonia	10	130	0.201	−0.847
Northern China	13	177	0.118	−1.401 ($P < 0.01$)
Niue	9	140	0.235	−0.768
Norway	10	362	0.139	−1.586 ($P < 0.01$)
Papua Highlands	7	114	0.208	−1.344 ($P < 0.05$)
Rabaul	6	120	0.227	−1.420 ($P < 0.025$)
Rarotonga	13	156	0.218	−0.175
Southern China	13	199	0.152	−1.042
Senegal	12	379	0.344	0.400
Southern Filipino	13	210	0.152	−1.057
Toba	9	268	0.294	−0.507
Western Samoa	11	102	0.167	−0.954
	$\overline{k_A} = 10.4$		$\overline{Fnd}_A$ =	−0.786 ($P \ll 10^{-4}$)
			$s_A/\sqrt{m}$	0.137

[a] As for Table 1.

TABLE 3

***DPB1* allele homozygosity statistics in 14 human populations**

Population[a]	k_A	$2n$	F_A	Fnd_A
Cayapa	6	200	0.406	−0.402
CEPH	17	248	0.212	0.219
Czech	14	193	0.232	0.074
Indonesia	12	272	0.199	−0.780
Japan	12	418	0.25	−0.454
Mexican (U.S.)	17	164	0.204	0.433
Nauru	6	100	0.349	−0.541
Northern China	13	189	0.188	−0.622
Norway	16	362	0.265	0.451
Papua Highlands	6	172	0.286	−1.101
Southern China	12	200	0.212	−0.539
Senegal	10	106	0.172	−1.071
Southern Filipino	12	180	0.236	−0.247
Toba	12	272	0.361	0.754
	$\overline{k_A} = 11.8$		$\overline{Fnd}_A$ =	−0.274 ($P = 0.08$)
			$s_A/\sqrt{m}$	0.154

[a] As for Table 1.

between pockets and nonpockets, with respect to the evenness of amino acid variant distributions, is significant only at *DRB1* (Figure 4). At *DPB1* and *DQB1* the mean normal deviates of homozygosity over all sampled amino acid sites in the nonpockets category are indistinguishable from the levels observed for the sites in the pockets.

DISCUSSION

Deviation from neutral expectations of the homozygosity *F* statistic is evident at the level of amino acid sites in class II HLA β1 domains encoded by *DRB1*, *DQB1*, and *DPB1*; many polymorphic amino acid sites in class II HLA β1 domains are more evenly distributed than expected under neutrality. This is in contrast to allele level variation in which *DQB1* shows very strong evidence for balancing selection, *DRB1* some evidence, and *DPB1* very little evidence. In one population, the Java, there was significant evidence of directional selection at the allele level. This may indicate that selection is most intense at the amino acid level.

It appears that different evolutionary histories have shaped the three β1 domains. Two possible explanations for the contrast between *DRB1* nonpocket poly-

Site	N_p	$\bar{k}$	$\bar{F}_{aa}$	$\overline{Fnd}_{aa}$
9	23	2.78	0.58	-0.89
10	23	2.48	0.53	-1.44
11	23	5.09	0.39	-0.87
12	23	2.00	0.53	-1.74
13	23	5.52	0.30	-1.21
14	16	2.00	0.88	0.28
16	23	2.17	0.72	-0.50
25	16	2.00	0.88	0.28
26	22	2.82	0.65	-0.49
28	22	2.82	0.61	-0.70
30	22	4.55	0.62	0.22
31	19	2.58	0.83	0.39
32	23	2.00	0.62	-1.20
33	23	2.00	0.75	-0.46
37	23	4.61	0.34	-1.26
38	20	2.45	0.79	0.07
40	11	2.00	0.97	0.81
47	23	2.00	0.58	-1.47
57	23	3.91	0.50	-0.60
58	23	2.00	0.80	-0.13
60	23	2.91	0.63	-0.53
67	23	3.00	0.40	-1.77
70	23	3.00	0.45	-1.52
71	23	3.43	0.61	-0.30
73	18	2.00	0.82	-0.03
74	23	4.39	0.49	-0.51
77	16	2.00	0.90	0.44
78	19	2.00	0.81	-0.12
85	20	2.00	0.80	-0.16
86	23	2.00	0.56	-1.60

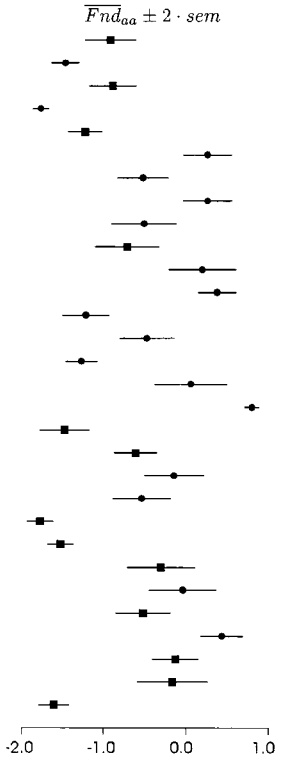

Figure 1.—*DRB1* amino acid site statistics calculated across 23 human populations. Each site that is observed to be polymorphic in at least one of the 23 populations is shown. The mean number of variants observed across N_p populations polymorphic for each site, $\bar{k}_{aa}$, is given as well as the mean homozygosity, $\bar{F}_{aa}$, and the mean normal deviate of homozygosity $\overline{Fnd}_{aa}$. (■) Pocket and (●) nonpocket sites are distinguished.

Site	N_p	$\bar{k}$	$\bar{F}_{aa}$	$\overline{Fnd}_{aa}$
9	23	2.83	0.64	-0.52
13	23	2.00	0.58	-1.46
14	22	2.00	0.72	-0.63
23	11	2.00	0.93	0.62
26	23	3.00	0.43	-1.64
28	17	2.00	0.83	0.02
30	22	2.77	0.60	-0.79
37	22	2.64	0.70	-0.32
38	22	2.00	0.57	-1.55
45	23	2.00	0.61	-1.27
46	17	2.00	0.83	0.02
47	17	2.00	0.83	0.02
52	17	2.00	0.83	0.02
53	22	2.00	0.57	-1.50
55	23	2.74	0.50	-1.37
56	22	2.00	0.88	0.30
57	23	3.61	0.54	-0.58
66	23	2.00	0.64	-1.13
67	23	2.00	0.64	-1.13
70	23	2.91	0.57	-0.85
71	23	3.65	0.55	-0.51
74	23	2.74	0.57	-1.00
75	23	2.00	0.60	-1.32
77	22	2.00	0.65	-1.05
84	22	2.00	0.57	-1.50
85	22	2.00	0.57	-1.50
86	22	2.59	0.57	-1.13
87	22	3.00	0.50	-1.23
89	22	2.00	0.57	-1.50
90	22	2.00	0.57	-1.50

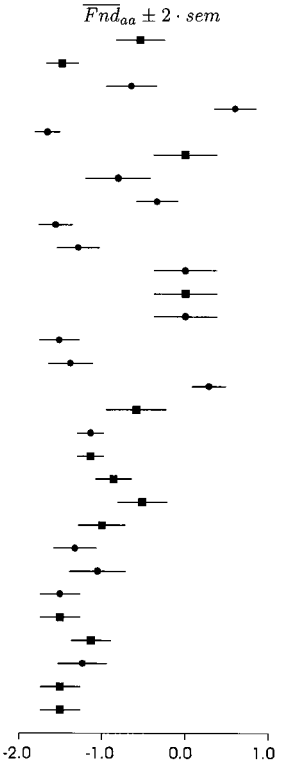

Figure 2.—*DQB1* amino acid site statistics calculated across 23 human populations, as described in Figure 1.

morphic sites and *DPB1* and *DQB1* nonpocket sites are: (1) the designations of site functions are incorrect for *DQB1* and *DPB1* despite the strong similarity in the sequences and molecules, which indicates that the functional use of sites has diverged in the different β1 domains; and (2) the pocket and nonpocket categories of sites in *DRB1* are uncoupled by intraexonic recombination and/or segmental exchange, while *DQB1* and *DPB1* categories of sites do not evolve independently in this fashion.

To test hypothesis 1, we used another test for selection, the ratio of nonsynonymous substitutions per nonsynonymous site (d_n) to synonymous substitutions per synonymous site (d_s) (Hughes and Nei 1988, 1989). Under neutral expectations, this ratio should equal 1. Values >1 are seen as evidence for overdominant selection. The ratio of d_n/d_s was examined for sites in *DQB1* on the basis of both the current pocket/nonpocket designation, as well as sites grouped by their negative $\overline{Fnd}_{aa}$ values. The ratio for *DQB1* pocket sites was 1.5, while the negative $\overline{Fnd}_{aa}$ sites had a ratio of 3.9. This indicates that the current designation for pocket/nonpocket may not be inclusive of all sites.

When we look at only *DQB1* and *DRB1* sequences, hypothesis 2 also appears plausible, as far fewer *DQB1* than *DRB1* distinct sequences have been observed, and *DQB1* shows larger blocks of amino acid variants that segregate together than does *DRB1*. Recent work by Jakobsen *et al.* (1998) lends support for hypothesis 2. She used compatibility matrices to show that, although *DRB1* and *DPB1* both have high levels of recombination relative to *DQB1*, they behave very differently. *DRB1* is essentially scrambled so that the pocket and nonpocket sites may be uncoupled. Although there is more recombination at *DP* than at *DR*, the allelic diversity at *DP* is almost entirely due to shuffling of a small number of motifs, and there are consequently relatively few variable sites despite the large number of alleles. The larger number of monomorphic sites at *DPB1* than at *DQB1* and *DRB1* suggests that the persistence time of *DPB1* alleles is much shorter than that of *DRB1* and *DQB1* alleles, due to weaker selection at *DPB1* or frequent recombination. This may also provide an explanation for the difference between the allele and amino acid analyses of *DP*. If selection is acting at the amino acid or motif level it may have no discernible effect at the allele level because of the extent of shuffling of motifs.

The observations presented here can be interpreted as predicting that either the antigenic peptide-binding pockets in *DPB1* and *DQB1* employ different sets of sites than is the case for *DRB1*, or the evolution of individual amino acid sites in *DRB1* is considerably more independent than it is within *DPB1* and *DQB1* second exons.

Site	N_p	$\bar{k}$	$\bar{F}_{aa}$	$\overline{Fnd}_{aa}$
8	3	2.00	0.66	-0.99
9	14	2.93	0.61	-0.66
11	14	2.00	0.69	-0.86
33	5	2.00	0.95	0.75
35	14	2.86	0.61	-0.72
36	14	2.00	0.62	-1.27
55	14	2.86	0.48	-1.44
56	14	2.00	0.60	-1.36
57	14	2.00	0.77	-0.35
65	14	2.00	0.81	-0.15
69	14	2.36	0.62	-0.99
72	1	2.00	0.92	0.46
76	14	2.93	0.66	-0.41
84	14	2.29	0.57	-1.35
85	14	2.00	0.57	-1.53
86	14	2.00	0.57	-1.53
87	14	2.00	0.57	-1.53

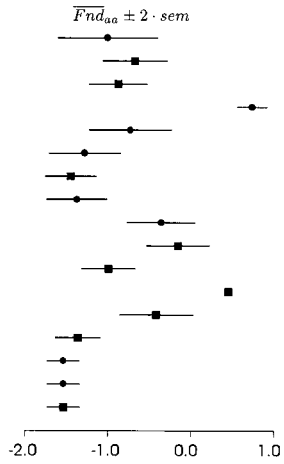

Figure 3.—*DPB1* amino acid site statistics calculated across 14 human populations, as described in Figure 1.

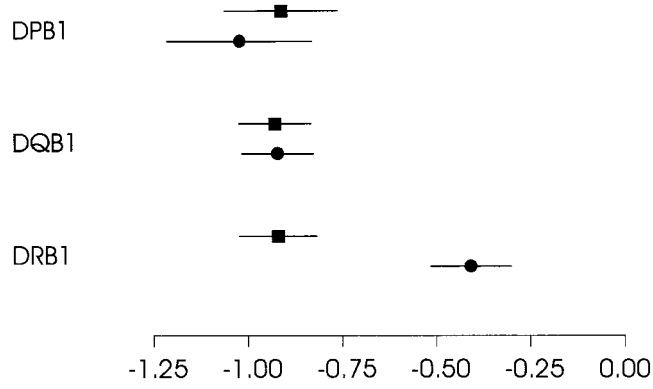

Figure 4.—Normal deviates of homozygosity in pocket *vs.* nonpocket sites. All polymorphic antigenic side-chain-binding amino acid sites (■, pocket category) were used to calculate the mean normal deviates of homozygosity and two standard error of the mean error bars for each locus. Also shown is the same calculation for polymorphic nonside-chain-binding sites (●).

Structural analysis of *DQB1* and *DPB1* molecules will allow these predictions from analysis of population genetic data to be examined.

A technical criticism that initially would suggest that we have not estimated our normal deviates in a conservative fashion is that amino acid sites are limited to 20 "allelic" states, at most. In fact, it is reasonable to assume that fewer amino acid variants could be tolerated even at the most polymorphic HLA amino acid sites. The objection is based on the fact that the neutrality model we used in calculating normal deviates of homozygosity assumed that infinite alleles existed; that is, the models to which we compared data assumed that each mutation produces a state not previously existing in a population. The infinite alleles model is not strictly appropriate for investigating amino acid site polymorphism. A preliminary examination of neutral models with finite allelic states demonstrates that, given reasonable mutation rates and population sizes for human data, the effect of finite alleles on the neutral model (decreasing the homozygosity statistic relative to that expected for the infinite alleles model) is negligible (Salamon 1995). However, estimating the neutrality parameter θ ($4N_e\mu$) from empirically derived mutation rates and population sizes is difficult, given the uncertainty in estimates of both parameters. Therefore, further simulations are required and are in progress to address the issue of the appropriate value of θ and consequently the robustness of the application of infinite alleles tests to finite allele data. At this stage, interpretations are descriptive at best; the Watterson test provides us with a useful reference point to compare allele and amino acid frequency distributions at the various loci and sites, and among populations.

A criticism of the means and variances of the homozygosity statistic is that although the population samples are statistically independent, they are not historically independent. This is most certainly true in a strict sense as all ethnic groups have a common ancestor several thousand generations ago, and in the case of many of those studied here, far fewer generations ago than that. One could sacrifice sample size to obtain a set of populations that are deemed "more independent" than the entire set analyzed here. However, it is not obvious that there exist sets of ethnic groups that are in some respect equivalent, such that one cluster can be compared fairly to another; nor is it obvious that time since divergence is a good measure of evolutionary independence. When we focus on *DRB1* pockets, the normal deviates of homozygosity for Caucasians—Czech, CEPH, and Norwegian samples—are very similar, whereas the values for Melanesians—samples from Fiji, Madang, Papua New Guinea Highlands, and New Caledonia—are widely scattered (data not shown). Clearly Melanesians have diversified into a range of environments and relatively small isolated populations. At the amino acid level, caution must be taken in interpreting *P* values due to nonindependence of amino acid sites. A neutral site linked to a selected site may deviate from neutrality as a result of hitchhiking.

Despite difficulties with independence of genetic information, there exists strong evidence of balanced polymorphisms at particular sites in HLA class II amino acid sequences, a pattern that is found across diverse human populations. The reproducibility of homozygosity statistics lower than neutral expectation across human populations for some polymorphic sites in the β1 domains of *DPB1*, *DQB1*, and *DRB1* is clearly demonstrated here.

This research was supported by National Institutes of Health grant GM35326 (H.S., W.K., S.K.M., M.P.N., G.T.).

LITERATURE CITED

Anderson, R., 1979 Some stochastic models in population genetics. M. Sc. Thesis, Monash University, Clayton, Victoria, Australia.

Begovich, A. B., G. R. McClure, V. C. Suraj, R. C. Helmuth, N. Fildes *et al.* 1992 Polymorphism, recombination and linkage disequilibrium within the HLA class II region. J. Immunol. **148:** 249–258.

Bodmer, J., S. G. E. Marsh, E. D. Albert, W. F. Bodmer, R. E. Bontrop *et al.*, 1997 Nomenclature for factors of the HLA system, 1996. Tissue Antigens **49:** 297–321.

Brown, J. H., T. S. Jardetzky, J. C. Gorga, L. J. Stern, R. G. Urban *et al.*, 1993 Three-dimensional structure of the human class II histocompatibility antigen HLA-*DR*1. Nature **364:** 33–39 [see comments].

Bugawan, T. L., J. D. Chang, W. Klitz and H. A. Erlich, 1994 PCR/oligonucleotide probe typing of HLA class II alleles in a Filipino population reveals an unusual distribution of HLA haplotypes. Am. J. Hum. Genet. **54:** 331–340.

Cerna, M., M. Fernandez-Viña, E. Ivaskova and P. Stastny, 1992 Comparison of HLA class II alleles in Gypsy and Czech populations by DNA typing with oligonucleotide probes. Tissue Antigens **39:** 111–116.

Cerna, M., M. Falco, H. Friedman, E. Raimondi, A. Maccagno *et al.*, 1993 Differences in HLA class II alleles of isolated South American Indian populations from Brazil and Argentina [see comments]. Hum. Immunol. **37:** 213–220.

Erlich, H. A., A. Zeidler, J. Chang, S. Shaw, L. J. Raffel *et al.*, 1993 HLA class II alleles and susceptibility and resistance to insulin dependent diabetes mellitus in Mexican-American families. Nat. Genet. **3:** 358–364.

Ewens, W. J., 1972 The sampling theory of selectively neutral alleles. Theor. Popul. Biol. **3:** 87–112.

Ewens, W. J., 1979 *Mathematical Population Genetics*. Springer-Verlag, Berlin-New York.

Fernandez-Viña, M. A., X. J. Gao, M. E. Moraes, J. R. Moraes, I. Salatiel *et al.*, 1991 Alleles at four HLA class II loci determined by oligonucleotide hybridization and their associations in five ethnic groups. Immunogenetics **34:** 299–312.

Gao, X., K. Bhatia, R. J. Trent and S. W. Serjeantson, 1992a HLA-*DR,DQ* nucleotide sequence polymorphisms in five Melanesian populations. Tissue Antigens **40:** 31–37.

Gao, X., P. Zimmet and S. W. Serjeantson, 1992b HLA-*DR,DQ* sequence polymorphisms in Polynesians, Micronesians and Javanese. Hum. Immunol. **34:** 153–161.

Hedrick, P. W., and G. Thomson, 1983 Evidence for balancing selection at HLA. Genet. **104:** 449–456.

Hedrick, P. W., W. Klitz, W. P. Robinson, M. K. Kuhner and G. Thomson, 1991 Population genetics of HLA, pp. 248–271 in *Evolution at the Molecular Level*, edited by R. K. Selander, A. G. Clark and T. S. Whittam. Sinauer Associates, Sunderland, MA.

Hughes, A. L., and M. Nei, 1988 Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335:** 167–170.

Hughes, A. L., and M. Nei, 1989 Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci USA **86:** 958–962.

Jakobsen, I. B., S. R. Wilson and S. Easteal 1998 Patterns of reticulate evolution for the classical class I and II HLA loci. Immunogenetics **48:** 312–323.

Klitz, W., G. Thomson and M. P. Baur, 1986 Contrasting evolutionary histories among HLA-region loci. Am. J. Hum. Genet. **39:** 340–349.

Klitz, W., G. Thomson, N. Borot and A. Cambon-Thomsen, 1992 Evolutionary and population perspective of the human HLA complex. Evol. Biol. **26:** 35–72.

Lawlor, D. A., J. Zemmour, P. D. Ennis and P. Parham, 1990 Evolution of class I MHC genes and proteins: from natural selection to thymic selection. Annu. Rev. Immunol. **8:** 23–63.

Marsh, S. G., and J. G. Bodmer, 1993 HLA class II nucleotide sequences, 1992. Eur. J. Immunogenet. **20:** 47–79.

Marsh, S. G., and J. G. Bodmer, 1995 HLA class II nucleotide sequences, 1995. Tissue Antigens. **45:** 258–280.

Mitsunaga, S., S. Kuwata, K. Tokunaga, C. Uchikawa, K. Takahashi *et al.*, 1992 Family study on HLA-*DPB1* polymorphism: linkage analysis with HLA-*DR*/*DQ* and two "new" alleles. Hum. Immunol. **34:** 203–211.

Monaco, J. J., 1993 Structure and function of genes in the MHC class II region. Curr. Opin. Immunol. **5:** 17–20.

Parham, P., D. A. Lawlor, C. E. Lomen and P. D. Ennis, 1989 Diversity and diversification of HLA-A, B, C alleles. J. Immunol. **142:** 3937–3950.

Rønningen, K. S., A. Spurkland, G. Markussen, T. Iwe, F. Vardal *et al.*, 1990 Distribution of HLA class II alleles among Norwegian caucasians. Hum. Immunol. **29:** 275–281.

Salamon, H., 1995 Pattern, process, function and disease in the evolution of protein sequences: a study of antigen presenting molecules in human populations. Ph.D. Thesis, University of California, Berkeley.

Stern, L. J., J. H. Brown, T. S. Jardetzky, J. C. Gorga, R. G. Urban *et al.*, 1994 Crystal structure of the human class II MHC protein HLA-*DR*1 complexed with an influenza virus peptide. Nature **368:** 215–221.

Takahata, N., T. Satta and J. Klein, 1992 Polymorphism and balancing selection at major histocompatibility complex loci. Genetics **130:** 925–938

Trachtenberg, E. A., H. A. Erlich, O. Rickards, G. F. DeStafano and W. Klitz, 1995 HLA class II linkage disequilibrium and haplotype evolution in the Cayapa Indians of Ecuador. Am. J. Hum. Genet. **57:** 415–424.

Watterson, G. A., 1978 The homozygosity test of neutrality. Genetics **88:** 405–417.

Communicating editor: A. G. Clark