Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 958–962, February 1989
Genetics

Nucleotide substitution at major histocompatibility complex class II loci: Evidence for overdominant selection

(DNA evolution/polymorphism/nonsynonymous substitution/gene conversion)

AUSTIN L. HUGHES AND MASATOSHI NEI

Center for Demographic and Population Genetics, The University of Texas Health Science Center at Houston, Houston, TX 77225

Communicated by Robert K. Selander, October 14, 1988 (received for review July 15, 1988)

ABSTRACT **To study the mechanism of maintenance of polymorphism at major histocompatibility complex (MHC) loci, synonymous and nonsynonymous (amino acid-altering) nucleotide substitutions in the putative antigen-recognition site (included in the first domain of the MHC molecule) and other regions of human and mouse class II genes were examined. In the putative antigen-recognition site, the rate of nonsynonymous substitution was found to exceed that of synonymous substitution, whereas in the second domain the former was significantly lower than the latter. In light of a previous theoretical study and parallel findings in class I MHC loci, we conclude that the unusually high degree of polymorphism at class II MHC loci is caused mainly by overdominant selection (heterozygote advantage) operating in the antigen-recognition site.**

The major histocompatibility complex (MHC) includes several genetic loci that are highly polymorphic in humans, mice, and other vertebrates (1). Four hypotheses have been proposed to account for this polymorphism: (*i*) an unusually high mutation rate (2), (*ii*) gene conversion or interlocus genetic exchange (3, 4), (*iii*) overdominant selection (5, 6), and (*iv*) frequency-dependent selection (7, 8). To decide which of these hypotheses is correct, we previously studied rates of synonymous and nonsynonymous nucleotide substitution between alleles from human and mouse class I MHC loci and found that the nonsynonymous rate is higher than the synonymous rate in the 57 codons encoding the antigen-recognition site but is lower than the latter in other regions of the gene (9). Since overdominant selection is known to increase the rate of codon substitution as well as the extent of polymorphism (10), we concluded that class I polymorphism is caused mainly by overdominant selection that operates in the antigen-recognition site. Here we extend our analysis to class II MHC genes.

Both class I and class II genes encode glycoproteins that are expressed on cell surfaces and that function to provide a context for recognition of intracellularly processed foreign peptides (antigens) by specific immune-system cells (1). Class I antigens are expressed on all nucleated cells, whereas class II antigens are expressed on antigen-presenting cells of the immune system, which present foreign antigens to helper T cells (1, 11). In the case of class I molecules, the amino acid residues involved in binding the foreign peptide and in T-cell recognition have been identified (12), making it possible to study the pattern of nucleotide substitution in the DNA region encoding the antigen-recognition site (ARS). In the case of class II molecules, the ARS has yet to be identified, but antigen recognition by helper T cells is known to be localized in the first domains of both α and β chains of the molecule, as opposed to the second domains (13, 14). Furthermore, by analogy with the class I molecule, a putative ARS in domain 1 has been reported (15). In this paper we compare the rates of synonymous and nonsynonymous nucleotide substitution in these regions to examine the type of selection operating at class II MHC loci.

MATERIALS AND METHODS

The class II MHC molecule is composed of an α chain and a β chain, which are noncovalently associated. Both α and β chains include domain 1 (D1), domain 2 (D2), the transmembrane portion, and the cytoplasmic tail (1). D1 of the α chain is encoded by the last two codons of exon 1 and by exon 2 of the gene, and D2 by exon 3. D1 of the β chain is encoded by the last four codons of exon 1 and by exon 2, and D2 by exon 3. The transmembrane portion and the cytoplasmic tail of the α chain are encoded by exon 4, whereas those of the β chain are encoded by exons 4 and 5. There are 19 or 20 amino acid residues in D1 of the α chain and 15 or 16 residues in D1 of the β chain in the putative ARS (15).

In the mouse, there are eight class II loci in the $H\text{-}2^b$ haplotype (perhaps seven in some other haplotypes) (6). At least four of them (A_β, A_α, E_β, and E_α) are known to be expressed (Table 1). It remains uncertain whether the $A_{\beta 2}$ gene is expressed (16). The human class II MHC can be divided into the following five regions from the centromeric side of the chromosome: *DP*, *DN* (formerly *DZ*), *DO*, *DQ*, and *DR* (44, 45). In the latest nomenclature, α-chain and β-chain genes are designated as *A* and *B* respectively (Table 1). The *DP* region includes two expressed genes, *DPB1* and *DPA1*. The *DN* region contains a single poorly expressed α-chain locus, *DNA* (30), whereas the *DO* region contains a single poorly expressed β-chain gene, *DOB* (31). The *DQ* region includes the two expressed genes *DQB1* and *DQA1* and the probably unexpressed genes *DQB2* and *DQA2*. In the *DR* region, there are a single α-chain locus, *DRA*, and one or two expressed β-chain loci, depending on the haplotype (38); the latter loci are designated as *DRB1* and *DRB3*. In the case of the *DQ*- and *DR*-region genes, it is not always clear from which locus a particular published sequence comes. We therefore treated these regions as though they were single loci.

The α- and β-chain genes are remotely related, having separated probably more than 370 million years ago (6). Certain α- and β-chain genes in mice are more closely related to certain human genes than to other mouse genes (1, 6), indicating orthologous (homologous) relationships (Table 1). In this study we examined the pattern of nucleotide substitution between alleles from polymorphic α- and β-chain loci within the mouse and the human and between alleles from orthologous loci of the two species. We used published DNA sequences excluding those with incomplete D1 and D2 and

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: ARS, antigen-recognition site; D1 and D2, domains 1 and 2; MHC, major histocompatibility complex.

Table 1. Class II MHC genes used in this study

Locus or region	Ortho-logue	Alleles	Ref(s).
Mouse			
$A_{\beta 2}$	*DOB*	Monomorphic	16
A_β ($A_{\beta 1}$)	*DQB*	*b, d, f, k, q, s, u*	17–19
A_α	*DQA*	*b, d, f, k, s, r*	20, 21
E_β	*DRB*	*b, d, k, u*	3, 22–24
E_α	*DRA*	Low polymorphism	25
Human			
DPB1		*w2, w3, w4a*	26
DPA1		Monomorphic	27
DNA		Monomorphic	28
DOB	$A_{\beta 2}$	Monomorphic	29
DQA2 (*DXA*)		Monomorphic?	27
DQB	A_β	*1w12, 3, w1* (Daudi), *w1.2, w1w2, w3* (MCF), *w3* (Swei), *w3* ($Ta10^-$)	30–36
DQA1	A_α	*w1w2, w2, w3, w6w6*	27, 31, 32, 37
DRB	E_β	*1, 2a* (AZH), *2a* (PGF), *2b* (AZH), *2b* (PGF), *3a, 3b, 3w6, 4a, 5a, 6a, 7a, 7b, w53*	38–42
DRA	E_α	Monomorphic	43

The loci are listed in the order in which they appear on the chromosome from the centromeric side. For polymorphic loci, the alleles analyzed are listed. The allele designations in parentheses refer to cell lines.

minor variants of other sequences (defined as having <1% nucleotide difference in D1).

As mentioned earlier, the purpose of this paper is to study the mechanism of maintenance of MHC polymorphism by examining the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N). If there is no selection, d_S and d_N are expected to be more or less the same. If certain amino acids are conserved by purifying selection as in the case of most other proteins, d_N will be lower than d_S. If positive Darwinian selection is responsible for the polymorphism, d_N will be higher than d_S because such selection will accelerate amino acid changes. The distinction between overdominant selection and frequency-dependent selection can be made by other considerations. With these predictions in mind, we estimated d_N and d_S separately for the putative ARS, for the remainder of D1, and for D2, using Nei and Gojobori's method I (46, 47). Since the number of sequence comparisons is large, we present only the means of d_S and d_N for each group of allelic comparisons. For simplicity we denote these means by d_S and d_N rather than by $\bar{d}_S$ and $\bar{d}_N$ in the following.

RESULTS

Fig. 1 shows the nucleotide sequences of D1 for seven alleles from the mouse A_β locus. The pattern of nucleotide substitution between these alleles is quite different from that of most other eukaryotic and prokaryotic genes, where a majority of substitutions are synonymous, often occurring at third nucleotide positions of codons (47). Here we see many substitutions in the first and second positions that are obviously nonsynonymous. Particularly at the codons in the ARS (marked with + in Fig. 1) the rate of nonsynonymous substitution is very high. Note that this unusual pattern is observed only in D1 and that the pattern of nucleotide substitution in D2 is similar to that of most other genes (data not shown).

Table 2 shows d_S and d_N values for various groups of allelic comparisons. In the case of the mouse A_β locus, d_N is significantly higher than d_S in the ARS, as expected from the sequence data in Fig. 1, whereas d_N is much lower than d_S in D2. In D1 outside the ARS, d_N is somewhat higher than d_S, suggesting that some residues outside the ARS might also be involved in T-cell recognition. Nevertheless, d_N in the ARS is significantly higher than d_N in the remainder of D1 or in D2. Similar results are obtained for the four alleles derived from the E_β locus, though d_S in the ARS is greater in this locus than in the A_β locus and the d_S value in D2 is lower. These results are consistent with those obtained for class I MHC genes (9) and strongly suggest that the rate of amino acid substitution in the ARS is enhanced by positive Darwinian selection.

Comparison of the A_β and E_β genes, which apparently separated before the human and mouse divergence (about 75 million years ago), shows that d_N in the ARS is only about 2 times greater than d_S but is still significantly higher than d_N in D2. The d_N/d_S ratio in the ARS is lower in the interlocus comparison than in the intralocus comparison, which probably reflects the fact that some amino acids (e.g., aspartic acid at position 57 and asparagine at position 82) tend to be conserved even in the ARS and the frequency of parallel and backward mutations at the amino acid level increases as d_N

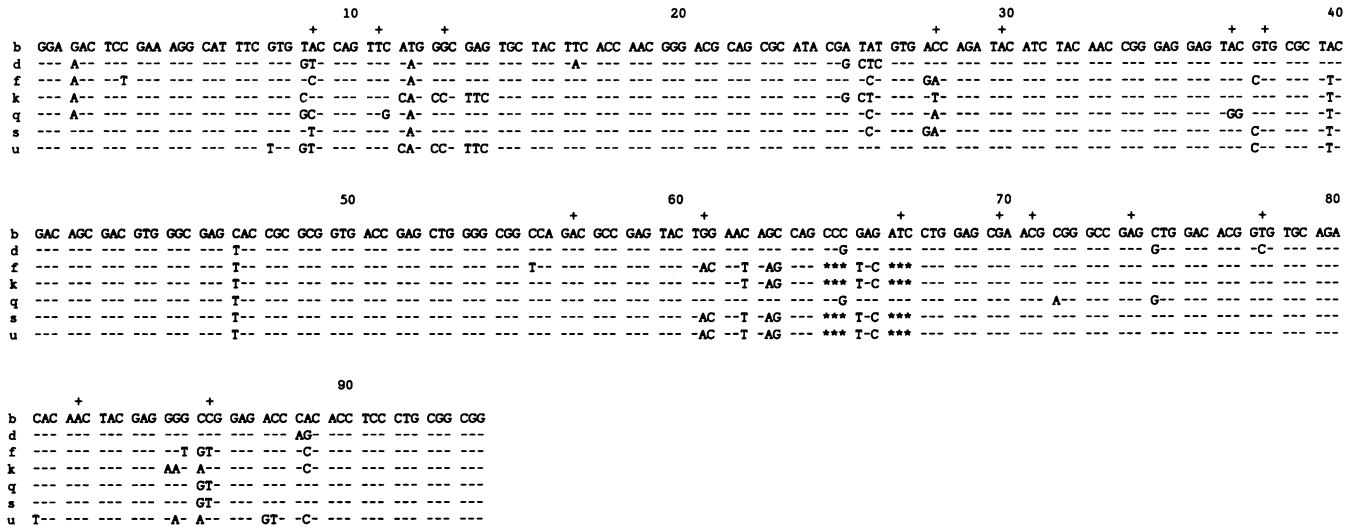

FIG. 1. DNA sequences of D1 of seven alleles at the mouse MHC A_β locus. +, Codons in the putative ARS; −, nucleotide identical with that of the *b* allele; *, gap in sequence.

Table 2. Mean numbers of nucleotide substitutions per synonymous site (d_S) and per nonsynonymous site (d_N) expressed as percentages, with their standard errors, between alleles at mouse and human class II MHC loci

Comparisons (no.)	Putative ARS† (n = 15–20)		Rest of D1 (n = 64–78)		D2 (n = 94)	
	d_S	d_N	d_S	d_N	d_S	d_N
			β-Chain loci			
Mouse						
A_β vs. A_β (21)	0.0 ± 0.0	30.0 ± 6.7***	4.0 ± 1.6	6.7 ± 1.2	7.3 ± 2.3	1.3 ± 0.5*
A_β vs. E_β (28)	39.5 ± 20.9	70.7 ± 14.7	29.7 ± 7.9	26.2 ± 3.9	89.5 ± 19.0	25.9 ± 3.9**
E_β vs. E_β (6)	11.6 ± 9.1	41.5 ± 9.5*	1.8 ± 1.3	5.2 ± 1.3	0.9 ± 1.1	0.6 ± 0.4
Human						
DPB vs. *DPB* (3)	3.9 ± 5.5	19.0 ± 6.4	2.4 ± 1.7	2.8 ± 1.1	5.3 ± 2.4	0.6 ± 0.4
DPB vs. *DQB* (24)	54.7 ± 27.2	73.1 ± 17.0	37.3 ± 9.1	22.9 ± 3.6	62.5 ± 13.1	15.8 ± 2.8***
DPB vs. *DRB* (42)	43.4 ± 22.8	74.0 ± 14.7	43.9 ± 10.4	22.0 ± 3.6	66.3 ± 13.4	17.9 ± 3.0***
DQB vs. *DQB* (28)	13.7 ± 7.7	26.5 ± 5.4	8.5 ± 2.3	6.7 ± 1.3	5.6 ± 2.0	1.6 ± 0.5
DQB vs. *DRB* (112)	47.5 ± 21.5	83.9 ± 16.1	37.1 ± 8.5	24.7 ± 3.7	91.3 ± 18.8	19.9 ± 3.2***
DRB vs. *DRB* (91)	15.0 ± 8.5	45.7 ± 6.2**	8.0 ± 1.9	4.5 ± 0.9	8.3 ± 1.8	3.3 ± 0.6**
Mouse vs. human						
$A_{\beta 2}$ vs. *DOB* (1)	33.1 ± 2.0	15.1 ± 6.9*	60.9 ± 15.4	13.5 ± 2.9***	53.3 ± 11.9	11.5 ± 2.5***
A_β vs. *DQB* (49)	51.0 ± 26.1	37.6 ± 5.9	33.6 ± 8.2	22.8 ± 3.5	55.2 ± 11.4	11.1 ± 2.2***
E_β vs. *DRB* (56)	41.8 ± 19.8	56.4 ± 7.7	43.1 ± 10.1	15.8 ± 2.7**	55.4 ± 11.5	9.5 ± 2.0***
			α-Chain loci			
Mouse						
A_α vs. A_α (15)	3.2 ± 3.0	23.7 ± 4.9***	2.8 ± 1.7	2.6 ± 0.8	7.2 ± 2.2	0.7 ± 0.4**
Human						
DQA vs. *DQA* (6)	21.7 ± 11.8	27.0 ± 6.7	8.0 ± 3.2	4.3 ± 1.3	4.0 ± 1.9	2.4 ± 0.8
DPA vs. *DNA* (1)	120.7 ± 67.1	76.9 ± 21.1	130.9 ± 42.9	31.9 ± 5.5*	65.1 ± 14.4	22.6 ± 3.6***
DPA vs. *DRA* (1)	116.7 ± 66.6	43.4 ± 12.5	84.0 ± 24.4	28.4 ± 5.0*	94.7 ± 21.6	19.1 ± 3.3***
DNA vs. *DRA* (1)	98.6 ± 48.5	71.9 ± 19.3	126.8 ± 41.1	40.0 ± 6.4*	77.5 ± 16.9	22.1 ± 3.6***
Mouse vs. human						
A_α vs. *DQA* (24)	68.6 ± 30.3	63.0 ± 12.6	78.1 ± 20.0	18.9 ± 3.5**	33.8 ± 7.9	17.1 ± 2.9*
E_α vs. *DRA (1)*	49.5 ± 26.4	10.1 ± 4.9	118.6 ± 38.3	13.5 ± 3.2**	60.8 ± 13.4	10.8 ± 2.4***

The standard errors of mean d_S and d_N were calculated by an extension of the method of Nei *et al.* (48). (A computer program is available upon request.) The difference between mean d_S and mean d_N is significant at 5% level (*), 1% level (**), or 0.1% level (***). In the ARS, the number of codons compared (n) is 15 in A_β and $A_{\beta 2}$, 16 in E_β and in all human β-chain genes, 19 in *DPA*, and 20 in all other α-chain genes. In the remainder of D1, n is 64 in E_α and *DRA*; 65 in *DPA* and *DNA*; 67 in *DQA*; 68 in A_α; 76 in *DPB*; 77 in A_β; and 78 in $A_{\beta 2}$, E_β, *DOB*, *DQB*, and *DRB*.

†The amino acid positions involved in the ARS are 11, 13, 14, 26, 28, 35, 41, 51, 53, 54, 56, 57, 62, 66, 69, 70, 73, 76, 77, and 80 in D1 of the α chain and 9, 11, 13, 28, 30, 37, 38, 57, 61, 67, 70, 71, 74, 78, 82, and 86 in D1 of the β chain (from figure 3 of ref. 15). Amino acid positions are numbered as in refs. 18, 20, 31, and 37 and in Fig. 1.

increases. In the comparison of A_β and E_β, d_S in D2 is much higher than that in D1. This seems to be partly due to the difference in G+C content at third nucleotide positions of codons between D1 and D2 (77–88% vs. 70–77%). Since the observed rate of nucleotide substitution tends to be lower as G+C content increases (49), d_S is expected to be lower in D1 than in D2. Another possible explanation for the high value of d_S in D2 is exon shuffling, which would have exchanged exon 3 of either the A_β or the E_β gene with that of another class II MHC locus.

The results from intralocus comparisons of human β gene sequences are similar to those from mouse β gene comparisons. In the ARS, d_N is 2–4 times greater than d_S for all three loci examined, whereas it is again smaller than d_S in D2. The results from interlocus comparisons of human sequences are also similar to those from mouse sequence comparisons. In the ARS, d_N is again greater than d_S, whereas in D2 it is significantly lower than d_S. The d_S value in D2 is again higher than that in D1; part of the reason for this seems to be a lower G+C content in the third nucleotide positions in D2 (66–78%) than in D1 (81–84%).

In the case of α-chain genes there are only a few sequences that can be used for our purpose, and some of the sequences come from monomorphic loci. The pattern of nucleotide substitution for intralocus comparisons of polymorphic A_α sequences in mice is nearly the same as that for β-chain genes. In the ARS, d_N is much higher than d_S, whereas in D1 the reverse is true. A similar pattern is observed at the *DQA* locus in humans, though d_S is much higher in the ARS than in other regions at this locus.

Table 2 includes the interlocus comparisons of monomorphic loci (*DPA*, *DNA*, and *DRA*) from humans. The pattern of nucleotide substitution in these comparisons is quite different from that for polymorphic loci. The d_N value is still greater in the ARS than in D2 but is lower than d_S. This indicates that amino acids are conserved in the entire gene region considered. The function of these monomorphic loci is yet to be discovered, but this finding suggests that the function might be different from that of polymorphic loci. The fact that d_N is higher in the ARS than in D2 might indicate that positive selection once existed before the function changed.

The d_N and d_S values for comparisons of alleles from the mouse and human orthologous loci are also presented in Table 2. Comparisons A_β vs. *DQB*, E_β vs. *DRB*, and A_α vs. *DQA* are for polymorphic loci, whereas the other comparisons are for monomorphic loci. In all comparisons, d_N is much smaller than d_S in D2, again suggesting rather strong purifying selection in this region. In the ARS, d_N is of the same order of magnitude as d_S for polymorphic loci. This seems to be due to a saturation effect in amino acid substitution mentioned earlier. In comparisons of monomorphic loci, however, d_N is much lower than d_S for both the ARS and D2.

DISCUSSION

We have presented evidence that the rate of amino acid substitution in the ARS is enhanced by positive Darwinian

selection compared with the neutral rate. There are several forms of selection that may enhance the rate of amino acid substitution. The most important one is selection for advantageous mutations. In the present case, however, this form of selection can be ruled out because it cannot explain the unusually high degree of polymorphism (70–90% heterozygosity) and the unusually long persistence of polymorphic alleles in the population (at least 3–10 million years) at MHC loci (50–52). Selection for advantageous mutations is known to reduce the level of polymorphism and shorten the persistence time of polymorphic alleles well below the level of neutral alleles (47).

A popular explanation for MHC polymorphism is frequency-dependent selection. Several authors (e.g., refs. 7 and 8) proposed the hypothesis that an individual carrying a new mutant allele has a selective advantage because pathogens will not have had the time to evolve the ability to infect host cells carrying a new mutant antigen. This model generates a higher rate of nonsynonymous substitution than of synonymous substitution, but it can explain neither the high degree of polymorphism nor the unusually long persistence of polymorphic alleles (ref. 9; N. Takahata and M.N., unpublished observation). There are several other models of frequency-dependent selection, but most of them are too simplistic, being based on the assumption of two alleles per locus without mutation (53). When these models are modified by considering more realistic situations, they are not necessarily powerful in maintaining polymorphism. For example, in the case of molecular mimicry (54), which is supposed to generate frequency-dependent selection, the selection advantage of a rare MHC allele will depend not only on the frequency of the allele but also on the frequency of the pathogen peptide (allele) that mimics the MHC allele and the pattern of mutation to which the MHC and pathogen genes are subject. A preliminary study on this subject suggests that molecular mimicry is less efficient than overdominant selection in maintaining polymorphism (M.N. and N. Takahata, unpublished work).

Our observations concerning MHC polymorphism can be explained most easily by the overdominance hypothesis. Overdominant selection is known to enhance the rate of amino acid substitution and increase the heterozygosity and persistence time of polymorphic alleles tremendously compared with those of neutral alleles (10, 47). Therefore, all three observations concerning MHC polymorphism (high rate of nonsynonymous nucleotide substitution, high polymorphism, and long persistence of polymorphic alleles) can be explained by this hypothesis as long as mutation occurs at a normal rate. Recent observations about the trans-specific mode of MHC polymorphism (50–52) can also be explained by this hypothesis.

The hypothesis of an unusually high mutation rate in the MHC can be rejected, since the mutational differences between human and mouse MHC genes are not particularly high compared with those of other genes (9). The hypothesis of interlocus genetic exchange also seems to have some problems (6). The most serious is that it cannot explain the differences between d_S and d_N in different regions of MHC genes. In the case of class II genes, there are additional observations that contradict this hypothesis. First, both α and β gene complexes include loci that are more similar between humans and mice than they are to other loci from the same species. Second, the two gene complexes include both polymorphic and monomorphic loci, and it is hard to imagine how gene conversion operates only on polymorphic loci. It is also difficult to understand why polymorphism is concentrated in the ARS, if gene conversion is the major cause of MHC polymorphism. Nevertheless, one cannot rule out rare events of gene conversion or exon shuffling, as mentioned earlier.

The biological basis of overdominant selection for class II MHC loci seems to be similar to that for class I MHC loci. It is now well established that a particular class II MHC molecule preferentially binds to a particular foreign peptide, thus providing improved recognition of that peptide by helper T cells and enhanced immune response (55). A heterozygote for two different alleles at a locus will therefore have resistance to two different types of pathogens and consequently have a higher fitness than a homozygote for either allele. If this view is correct, heterozygotes for many loci would have a higher fitness than those for a small number of loci, since there are a variety of pathogens. One might then expect that a genetic system with many polymorphic loci would evolve. In practice, however, the number of polymorphic loci is relatively small. There seem to be two reasons for this. (*i*) An allelic product can recognize many different foreign peptides sharing a certain structural motif, and thus a relatively small number of MHC antigens per individual is sufficient for protection from various pathogens (56). (*ii*) If there are too many polymorphic loci, the proportion of T cells that are eliminated to avoid autoimmunity becomes intolerably high (57).

This study was supported by grants from the National Institutes of Health and the National Science Foundation.

1. Klein, J. (1986) *Natural History of the Major Histocompatibility Complex* (Wiley, New York).
2. Bailey, D. W. & Kohn, H. I. (1965) *Genet. Res. Cambridge* **6**, 330–340.
3. Widera, G. & Flavell, R. A. (1984) *EMBO J.* **3**, 1221–1225.
4. Mengle-Gaw, L. & McDevitt, H. O. (1985) *Annu. Rev. Immunol.* **3**, 367–396.
5. Doherty, P. C. & Zinkernagel, R. M. (1975) *Nature (London)* **256**, 50–52.
6. Klein, J. & Figueroa, F. (1986) *CRC Crit. Rev. Immunol.* **6**, 295–386.
7. Snell, G. D. (1968) *Folia Biol. (Prague)* **14**, 335–358.
8. Bodmer, W. F. (1972) *Nature (London)* **237**, 139–145.
9. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335**, 167–170.
10. Maruyama, T. & Nei, M. (1981) *Genetics* **98**, 441–459.
11. Carson, S. & Trowsdale, J. (1987) *Oxford Surv. Eukaryotic Genet.* **3**, 63–94.
12. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987) *Nature (London)* **329**, 512–518.
13. Folsom, V., Gay, D. & Tonegawa, S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1678–1682.
14. Germain, R. N., Ashwell, J. D., Lechler, R. I., Margulies, P. H., Nickerson, K. M., Suzuki, G. & Ton, J. Y. L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2940–2944.
15. Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J. & Wiley, D. C. (1988) *Nature (London)* **322**, 845–850.
16. Larhammar, D., Hammerling, V., Rask, L. & Petersen, P. A. (1985) *J. Biol. Chem.* **260**, 14111–14119.
17. Estess, P., Begovich, A. B., Koo, M., Jones, P. P. & McDevitt, H. O. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3594–3598.
18. Larhammar, D., Hammerling, V., Denaro, M., Lund, T., Flavell, R. A., Rask, L. & Peterson, P. A. (1983) *Cell* **34**, 179–188.
19. Malissen, M., Hunkapiller, T. & Hood, L. (1983) *Science* **221**, 750–752.
20. Benoist, C. O., Mathis, D. J., Kanter, M. R., Williams, V. E. & McDevitt, H. O. (1983) *Cell* **34**, 169–177.
21. Landais, D., Matthes, H., Benoist, C. & Mathis, C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2930–2934.
22. Saito, H., Maki, P. A., Clayton, L. K. & Tonegawa, S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5520–5524.
23. Mengle-Gaw, L. & McDevitt, H. O. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7621–7625.
24. Ayane, M., Mengle-Gaw, L., McDevitt, H. O., Benoist, C. & Mathis, D. (1986) *J. Immunol.* **137**, 6948–6951.
25. Mathis, D. J., Benoist, C. O., Williams, V. E., Kanter, M. R. & McDevitt, H. O. (1983) *Cell* **32**, 745–754.

26. Lair, B., Alber, C., Yu, W.-Y., Watts, R., Bahl, M. & Karr, R. W. (1988) *J. Immunol.* **141,** 1353–1357.
27. Auffray, C., Lillie, J. W., Arnot, D., Grossberger, D., Kappes, D. & Strominger, J. L. (1984) *Nature (London)* **308,** 327–333.
28. Tonelle, C., DeMars, R. & Long, E. O. (1985) *EMBO J.* **4,** 2839–2847.
29. Trowsdale, J. & Kelley, A. (1985) *EMBO J.* **9,** 2231–2237.
30. Boss, J. M. & Strominger, J. L. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5199–5203.
31. Schenning, L., Larhammar, D., Bill, P., Wiman, K., Jonsson, A.-K., Rask, L. & Peterson, P. A. (1984) *EMBO J.* **3,** 447–452.
32. Schiffenbauer, J., Didier, D. K., Klearman, M., Rice, K., Shuman, S., Tieber, V. L., Kittlesen, D. J. & Schwartz, B. D. (1987) *J. Immunol.* **139,** 228–233.
33. So, A. K. L., Lindsay, J., Bodmer, J. & Trowsdale, J. (1987) *J. Immunol.* **139,** 3506–3511.
34. Tsukamoto, K., Yasunami, M., Kimura, A., Inoko, K., Ando, A., Hirose, T., Inayama, S. & Sasazuki, T. (1987) *Immunogenetics* **25,** 343–346.
35. Turco, E., Care, A., Compagnone-Post, P., Robinson, C., Cascino, I. & Trucco, M. (1988) *Immunogenetics* **26,** 282–290.
36. Lock, C., So, A. K. L., Welsh, K. I., Parkes, J. D. & Trowsdale, J. (1988) *Immunogenetics* **27,** 449–455.
37. Chang, H.-C., Moriuchi, T. & Silver, J. (1983) *Nature (London)* **305,** 813–815.
38. Bell, J. I., Denney, D., Foster, L., Belt, T., Todd, A. & McDevitt, H. O. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6234–6238.
39. Lee, B. S. M., Rust, N. A., McMichael, A. J. & McDevitt, H. O. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4591–4595.
40. Curtsinger, J. M., Hilden, J. M., Cairns, J. S. & Bach, F. H. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 209–213.
41. Young, J. A. T., Wilkinson, D., Bodmer, W. F. & Trowsdale, J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4929–4933.
42. 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.
43. Lee, J. S., Trowsdale, J., Travers, P. J., Carey, J., Grosveld, F., Jenkins, J. & Bodmer, W. F. (1982) *Nature (London)* **299,** 750–752.
44. Bodmer, W. F., Trowsdale, Y., Young, J. & Bodmer, J. (1986) *Philos. Trans. R. Soc. London Ser. B* **312,** 303–315.
45. Trowsdale, J. & Campbell, R. D. (1988) *Immunol. Today* **9,** 34–35.
46. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
47. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York).
48. Nei, M., Stephens, J. C. & Saitou, N. (1985) *Mol. Biol. Evol.* **2,** 66–85.
49. Mouchiroud, D. & Gautier, G. (1988) *Mol. Biol. Evol.* **5,** 192–194.
50. McConnell, T. J., Talbot, W. S., McIndoe, R. A. & Wakeland, E. K. (1988) *Nature (London)* **332,** 651–654.
51. Figueroa, F., Günther, E. & Klein, J. (1988) *Nature (London)* **335,** 265–267.
52. Lawlor, D. A., Ward, F. G., Ennis, P. D., Jackson, A. D. & Parham, P. (1988) *Nature (London)* **335,** 268–271.
53. Clarke, B. (1976) in *Genetic Aspects of Host-Parasite Relationships*, eds. Taylor, A. E. R. & Muller, R. M. (Blackwell, Oxford), pp. 78–103.
54. Damian, R. T. (1964) *Am. Nat.* **98,** 129–149.
55. Buus, S., Sette, A., Colon, S. M., Miles, C. & Grey, H. M. (1987) *Science* **235,** 1353–1358.
56. Sette, A., Buus, S., Colon, S., Miles, C. & Grey, H. M. (1988) *J. Immunol.* **141,** 45–48.
57. Howard, J. C. (1982) in *Evolution and Vertebrate Immunity*, eds. Kelsoe, G. & Schulze, D. H. (Univ. of Texas Press, Austin), pp. 397–411.