Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 6545–6549, July 1992
Immunology

Evolution of major histocompatibility complex class II allelic diversity: Direct descent in mice and humans

ANTE S. LUNDBERG AND HUGH O. MCDEVITT

Departments of Microbiology and Immunology, and Medicine, Stanford University Medical School, Stanford, CA 94305

Contributed by Hugh O. McDevitt, February 14, 1992

ABSTRACT **The high degree of polymorphism seen at major histocompatibility complex (MHC) class II loci is a feature unique to the MHC. Most of the β-chain polymorphism is localized in "hypervariable" regions (HVRs). HVR amino acid sequence similarity between distantly related species has recently been found. We have employed a Monte-Carlo statistic to show that shared HVR polymorphism between β-chain genes of humans and mice represents direct descent of ancestral sequences rather than convergent evolution. Furthermore, half the sequence polymorphism seen in class II β-chain genes of mice persists in evolution and is encoded by the same DNA sequence in humans. No evidence for increased mutation rate within the HVR was found. We postulate that the HVR can be considered the genetic unit of recombination, with selection for HVR sequences and combinations of HVRs constrained by functional considerations.**

Class II molecules of the major histocompatibility complex (MHC) are cell surface α/β heterodimeric proteins that present processed foreign antigen to T cells. Several class II β-chain loci are extremely polymorphic, a feature unique to the MHC class I and class II genes. The ratio of replacement to silent nucleotide substitutions among these sequences is relatively high, suggesting that these loci are under strong positive selection. At the whole gene level, allelic polymorphisms in different species do not appear related. Polymorphism may have arisen *de novo* in each species, perhaps with evolutionary pressure from environmental factors. However, polymorphic regions from class II alleles of different species of mice (1), and from chimpanzees and humans (2, 3), appear to be related, as do alleles at the MHC class I locus between chimpanzees and humans (4). The observations that allelic polymorphism predates divergence of these species at least 3–10 million years ago, as well as theoretical considerations of the nature of polymorphism in the MHC (5), have given rise to the "trans-species" hypothesis of MHC evolution.

Polymorphism within class II β-chain alleles occurs in limited hypervariable regions (HVRs) of the β1 domain (6). While polymorphism has evolved in part by point mutation and selection (7), multiple examples of genetic exchange between class II β-chain alleles in both rodents (8–10) and primates (6, 11–14) have been found. New class II alleles generated by recombination are often the product of shuffled HVR sequences from other alleles (6, 9, 15).

In class II β-chain genes, we have previously (16) noted similar amino acid and nucleotide sequence polymorphisms between humans and mice. More recent work has demonstrated that shuffling of HVR segments has increased allelic diversity among humans and nonhuman primates (14) and among rodent sequences (9). Recombinational events tend to obscure genealogies (9) and can make phylogenetic comparison of entire alleles misleading. In this paper, we have employed a Monte–Carlo statistic, independent of phylogenetic analysis, to study class II β-chain evolution by analyzing polymorphisms shared between humans and mice.

MATERIALS AND METHODS

Sequence Analysis. Twenty-six human DRβ1,7 DRβ3 (17) and 6 orthologous mouse I-Eβ (18–20) β1 domain amino acid sequences, and 13 DQβ (17) (11 full-length sequences) and 8 orthologous mouse I-Aβ (21–24) β1 domain sequences, were examined. Protein sequence numbering is according to Gregersen *et al.* (17). Amino acid position 2 of I-Eβ (18) was aligned with position 1 of DRβ, and all DRβ/I-Eβ sequence positions are given relative to DRβ (17). Allelic variability at each amino acid position of orthologous mouse and human class II β chain loci was calculated according to Wu and Kabat (25).

Statistical Analysis. The most similar interspecies pairs of orthologous HVR sequences were identified. The nonexpressed alleles at the I-Eβ locus (s, b, w17) are no longer under selection pressure, and we excluded them from the analysis. When more than one pair of sequences were equally similar (counting only identical amino acids), as in DQβ/I-Aβ HVR_{II}, each pair was used in turn to calculate the number of identical amino acids, observed codon match, and mean expected (Monte-Carlo) codon match, and an arithmetical average of these values was obtained. The number of identical codons between pairs of sequences was counted, as a fraction of the number of identical amino acids. Expected codon identity was calculated according to McCaldon and Argos (26). Species-specific codon frequencies (27) for each amino acid were normalized to total unity for each amino acid. The frequency of each codon in species 1 was multiplied by the frequency of the identical codon in species 2, and the products were summed. The distribution of expected codon identity between two sequences under the null hypothesis of convergent evolution is unknown, so a Monte-Carlo computer simulation was run to estimate this distribution.

Monte-Carlo analysis was performed by generating sequences of binary ("codon match"/"no codon match") events equal to the number of identical amino acid pairs in the allelic HVRs tested. Each event corresponds to an amino acid pair and has a probability of success equal to the expected codon identity for that given amino acid pair. The number of successes (codon matches) among the binary events was then counted. For each group of allelic HVR pairs of sequences, 10,000 trials were produced to generate a distribution of expected number of identical codons. From this distribution a mean value and one-tailed *P* value were obtained.

RESULTS

The most variable amino acid positions in the class II β-chain locus tend to cluster in limited regions (Table 1). Three HVRs

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; HVR, hypervariable region.

Table 1. Intralocus amino acid sequence variability (25) of class II β1 domain sequences

Position*	Variability: Among DRβ1, DRβ3, and I-Eβ alleles	Variability: Among DQβ and I-Aβ alleles	Putative function†
9	6.1	15.4	Ag binding
10	8.8		
11	24.7		Ag binding
13	14.4		Ag binding
14		8.0	
26	4.7	8.8	Helical region
28	9.9	6.6	Ag binding
30	9.4	4.3	Ag binding
37	16.5		Ag binding
38		7.5	Ag binding
52		5.0	
56		5.0	Helical region
57		5.7	Ag binding
66		5.5	T-cell recognition
67	5.8		Ag binding
70	4.5	6.6	Ag binding
71	7.9	5.7	Ag binding
74	9		Ag binding
86	9.5		Strand region

*Amino acid numbering according to Gregersen *et al.* (17). Amino acid position 2 of I-Eβ (18) was aligned with position 1 of DRβ.
†Positions predicted to be involved in processed antigen (Ag) binding or T-cell recognition or in the helical or strand region of the molecule (28) are as indicated.

are identified in DRβ and IEβ alleles and five in DQβ and IAβ alleles. These HVRs encompass the most variable positions, except for positions 36 and 86 in DRβ/I-Eβ and 38 in DQβ/I-Aβ. Although allelic differences occur outside these HVRs, the HVR sequences are sufficient to characterize almost all alleles. Of a possible 325 pairwise comparisons of DRβ1 alleles, only 5 pairs share identical sequences at all three HVRs and differ elsewhere in the molecule (data not shown).

Limited numbers of allelic HVR sequences are used at the β-chain loci to generate a larger number of alleles (9, 14, 16). Among the 26 DRβ1 β1 domain protein sequences, there are only 8 unique sequences at HVR_I, 9 at HVR_{II}, and 13 at HVR_{III} (Table 2). DQβ allelic HVR sequences are also fewer in number (4–7 for HVR_{I-V}) than the total number of alleles (13) at the DQβ locus. These variable residues in β-chain genes have been predicted to be involved in processed antigen binding or T-cell recognition in a class II structural model (28).

In DRβ and I-Eβ, the three HVRs contain 48% of the β1 domain variability. Yet several amino acid sequences at each HVR show remarkable interspecies similarity (Table 3). This could be due to convergent evolution driven by selective pressures from very similar pathogens (29, 30). These sequences would not then be encoded by the same codons. Alternatively, if sequence similarity reflected a common DNA origin, the HVRs shared between humans and mice would be descendant from a common ancestor. In this case, similarity of nucleotide codon sequence would be greater than expected by chance.

To distinguish convergent evolution from descent, we examined the DNA of the most-similar interspecies pairs of amino acid sequences at each allelic HVR (Table 3). Sixteen of the 18 amino acids in the most-similar interspecies pairs of sequences at the three HVRs in DRβ1, DRβ3, and I-Eβ contained the same amino acid (Table 4). Of the 16 identical amino acids, 12 were coded for by the same nucleotide sequence, where theoretically 6.07 would be expected if these variable region sequences had evolved independently ($P = 0.005$). Among the most similar interspecies pairs of allelic HVR sequences from DQβ and I-Aβ, 19 of 23 positions contained identical amino acids, and 13.5 amino acids were coded for by the same codon, whereas 7.19 would be expected ($P = 0.007$). (Sequences at HVR_V in DQβ and I-Aβ showed no interspecies similarity, and they were excluded.) However, HVR regions contain several invariant positions that may bias statistical analysis. Excluding these positions (DRβ positions 27, 29, 68, 69, and 72 from DRβ and I-Eβ and positions 10, 27, 29, 54, 68, and 69 from DQβ and I-Aβ), nucleotide identity is still somewhat greater than expected by chance ($P = 0.115$, $P = 0.064$, respectively).

It is possible that only these pairs of HVR sequences represent direct descent from a common ancestor. To examine whether other polymorphisms reflect common ancestry, we identified the amino acids at the 10 most-variable positions in the β1 domain that are identical in DRβ and expressed I-Eβ sequences (DRβ positions 9, 10, 11, 13, 28, 30, 37, 71, 74, and 86; of these, only positions 37 and 86 were not included in the three HVRs). There are 22 different amino acids among the expressed I-Eβ sequences at these 10 positions, 9 of which do not occur at the analogous positions in DRβ1 and DRβ3 alleles (Table 5). All three amino acids at position 87 in the mouse I-Eβ sequences do not occur in the analogous position 86 in the human DRβ1 and DRβ3 alleles. Of the remaining 13 amino acids, 2 are coded for by all possible codons in either humans or mice and were excluded. The other 11 are each encoded by a single codon, and 10 of these 11 amino acids are encoded by the same codon in both humans and mice, where 4.5 would be expected by calculation and by Monte-Carlo analysis ($P < 0.0001$). Approximately half (11 of 22) of the amino acids that occur in the variable positions of expressed I-Eβ alleles occur in, and are encoded by the same codon in, DRβ1 and DRβ3 alleles. Similar comparisons in miniature swine (30) led to a similar result ($P = 0.026$; data not shown).

At the 9 most variable positions in the β1 domain of DQβ and I-Aβ (9, 14, 26, 28, 38, 57, 66, 70, and 71, excluding

Table 2. Usage of allelic HVR sequences in β-chain alleles

Locus	Total no. alleles	Number of unique amino acid sequences: HVR_I 9–13	HVR_I 9–14	HVR_{II} 26–30	HVR_{III} 52–57	HVR_{III} 67–74	HVR_{IV} 66–71	HVR_V 85–89
$DR\beta_{I+III}$	~33	12		11		14		
$DR\beta_I$	26	8		9		13		
I-Eβ	6	4		5		5		
I-Eβ expressed	3	3		3		3		
DQβ	11*		5	6	7		6	4
I-Aβ	8		8	5	3		3	6

Amino acid positions are given for each HVR.
*Eleven full-length sequences, 13 sequences containing HVR_{II-V}.

Table 3. Most-similar interspecies pairs of sequences at each HVR

Loci	HVR	Position	Alleles	Sequences
DRβ and I-Eβ	HVR_I	10–14	I-Eβ^d	**E Y V T S** **GAA TAC GTT ACA TCT** GAG TAC TCT ACG TCT
		9–13	DR3β1; DR3DQwa; DR5Dw5β1; DR5JVM; DRw6aβ1; DRw6bβ1; DRw6amala	E Y S T S
		10–14	I-Eβ^u	**G Y S T S** **GGA TAT TCT ACA TCT** GAG TAC TCT ACG TCT
		9–13	DR3β1; DR3DQwa; DR5Dw5β1; DR5JVM; DRw6aβ1; DRw6bβ1; DRw6amala	E Y S T S
	HVR_{II}	27–31	I-Eβ^u	**F L D R Y** **TTT CTG GAC AGA TAC** TTC CTG GAC AGA TAC
		26–30	DR4Dw4,10,13,14,15; DR4KT2, Cetus; DR5Dw5, JVM; DRw6aβ1,6bβ1; DR2Dw2β3,Dw12β3, AZHβ3	F L D R Y
	HVR_{III}	68–75	I-Eβ^u	**I L E Q T R A A** **ATC CTG GAG CAA ACG CGG GCC GCG** ATC CTG GAG CAG GCG CGG GCC GCG
		67–74	DR2Dw2β3; DR2Dw12β3; DR2AZHβ1	I L E Q A R A A
DQβ and I-Aβ	HVR_I	9–14	I-Aβ^s	**F Q F K G E** **TTC CAG TTC AAG GGC GAG** TTC CAG TTT AAG GGC ATG
			DR2DQw1.2; DR4DQwa; DR8DQwa	F Q F K G M
	HVR_{II}	26–30	I-A$\beta^{d,nod}$	**L V T R Y** **CTC GTG ACC AGA TAC** CTT GTG ACC AGA TAC
			DR2DQw1.2; DR4DQw3.2; DR9DQw3.3	L V T R Y
			I-A$\beta^{u,b}$	**Y V T R Y** **TAT GTG ACC AGA TAC** TAT GTG ACC AGA TAC
			DR2DQw1.12; DR4DQw3.1	Y V T R Y
	HVR_{III}	52–57	I-A$\beta^{d,q,k,u,s,b}$	**E L G R P D** **GAG CTG GGG CGG CCA GAC** CCG CAG GGG CGG CCT GAT
			DR2DQw1.2; DR6DQw1.18	P Q G R P D
			I-A$\beta^{d,q,k,u,s,b}$	**E L G R P D** **GAG CTG GGG CGG CCA GAC** CCG CAG GGG CGG CCT GAC
			DR2DQw1.12; DR6DQw1.9	P Q G R P D
	HVR_{IV}	66–71	I-A$\beta^{d,q,b}$	**E I L E R T** **GAG ATC CTG GAG CGA ACG** GAC ATC CTG GAG AGG ACC
			DR2DQw1.12	D I L E R T
			I-A$\beta^{d,q,b}$	**E I L E R T** **GAG ATC CTG GAG CGA ACG** GAA GTC CTG GAG AGG ACC
			DR4DQw3.1; DR4DQw3.2; DR9DQw3.3; DR6DQw1.19	E V L E R T

Mouse sequences, boldface type; human sequences, lightface type.

HVR_V), I-Aβ sequences contain 22 amino acids (Table 5). Only one position (14) shares no amino acids with known DQβ sequences. Of the 22 amino acids, 11 occur in the analogous position in both DQβ and I-Aβ. One amino acid (Asp-57) is coded for by both codons in humans and was excluded. One amino acid (Leu-26) is encoded by two (of six possible) codons in mice, and all human sequences with the same amino acid at the analogous position are encoded by one of those two codons. Of the 10 remaining amino acids identical in the analogous variable positions in both DQβ and I-Aβ, 5.5 are coded for by the same codon in both species, whereas 3.74 would be expected if these variable position amino acids had arisen independently ($P = 0.12$). Surprisingly, the variable positions within the β1 domain of DQβ that show interspecies codon identity with I-Aβ all reside in the NH_2-terminal half of the β1 domain, at positions 9, 26, 28, and 38, while positions 57, 66, 70, and 71 (and HVR_V) show no interspecies codon identity at all. All 6 amino acids shared between humans and mice at these 5 positions in the NH_2-terminal half of the β1 domain are encoded by the same DNA

Table 4. Interspecies codon identity at allelic HVR regions

Comparison		No. of amino acids		Codon identity		Monte Carlo	
Loci	Most-similar pairs	Total	Identical	Obs.	Exp.	Mean	Significance
DRβ vs. I-Eβ	HVR_I: avg. of I-Eβ^u–DR3β1 I-Eβ^d–DR3β1	18	16	12	6.07	6.05	$P = 0.005$
	HVR_{II}: I-Eβ^u–DR4Dw4						
	HVR_{III}: I-Eβ^u–DR2Dw2β3						
	Variable positions only	13	11	7	4.66	4.62	$P = 0.115$
DQβ vs. I-Aβ	HVR_I: I-Aβ^s–DR2DQw1.2	23	19	13.5	7.19	7.18	$P = 0.007$
	HVR_{II}: avg. of I-Aβ^d–DR2DQw1.2 I-Aβ^u–DR2DQw1.12						
	HVR_{III}: avg. of I-Aβ^d–DR2DQw1.2 I-Aβ^d–DR2DQw1.12						
	HVR_{IV}: avg. of I-Aβ^d–DR2DQw1.12 I-Aβ^d–DR4DQw3.1						
	Variable positions only	13	9	7.5	4.99	4.95	$P = 0.064$

Total amino acids = total number of amino acids in the pairs of HVR sequences examined. Identical amino acids = number of identical amino acids in the same pairs of sequences. Codon identity = observed and expected number of identical codons in the same pairs of sequences. Monte-Carlo mean and significance = mean number of identical amino acids generated in 10,000 trials of Monte-Carlo analysis and one-tailed P value obtained from analysis of the distribution. In the comparison of HVR sequences from DRβ1 and DRβ3 with I-Eβ, the invariant positions that are excluded in the second line are DRβ positions 27, 29, 68, 69, and 72. In the comparison of HVR sequences from DQβ and I-Aβ, the invariant positions excluded are 10, 27, 29, 54, 68, and 69.

sequence ($P = 0.006$). At both the β1 domain of DRβ and I-Eβ and the NH_2 terminus of the β1 domain of DQβ and I-Aβ, allelic variability appears highly conserved in evolution, with approximately half the sequence polymorphism occurring prior to divergence of the human and mouse lineages.

The most similar interspecies pairs of HVR sequences may represent orthologous nucleotide sequences in evolution, arising directly from a common ancestral sequence. Differences in nucleotides between these sequences presumably occurred since human–mouse divergence, and synonymous (silent) substitutions between these sequences can be used to estimate the rate of mutation, at these loci, a molecular clock (31). Evolutionary distance, d_s (estimated number of synonymous nucleotide substitutions per synonymous site), is related to substitution rate, λ, and time, t, by the equation (31)

$$d_s = 2\lambda t.$$

Given that humans and mice diverged approximately 75 million years ago, the calculated rate of synonymous nucleotide substitution for the most similar interspecies pairs of HVR sequences between DRβ1, DRβ3, and I-Eβ is 2.89×10^{-9} substitutions per base pair per year (Table 6). Similarly, DQβ and I-Aβ HVR have a calculated base pair substitution rate of 7.33×10^{-9} per year. By the Monte-Carlo analysis above, there is stronger evidence to support direct evolution for the NH_2 terminus of these sequences. Omitting HVR_{IV} and HVR_V from the analysis results in a nucleotide substitution rate of 2.67×10^{-9} per year. These rates approximate published rates of 1.56×10^{-9} and 3.1×10^{-9} per year for primate and nonprimate mammals, respectively (32), and are less than the 5.49×10^{-9} per year observed in silent positions between class I molecules of human and mice (15). The rate of nucleotide substitution in rodents may be twice as high as in humans (33). Evolutionary distances compare favorably with previously published estimates of interspecies d_s from the class II putative antigen recognition site (0.418–0.51 substitution per site) and elsewhere in the β1 domain (0.336–0.431 substitution per site) of some of the same loci analyzed here (34), corresponding to rates of nucleotide substitution of $2.78–3.4 \times 10^{-9}$ and $2.24–2.87 \times 10^{-9}$ per year, respectively. These calculations represent limited sequence data, so little confidence can be placed in the numbers. However, they provide an approximation of evolutionary rates at the HVR in these loci, and suggest, together with the conserved nature of allelic sequences, that the HVRs are not subject to an increased rate of evolution.

DISCUSSION

We have identified similar HVR sequences in human and mice and shown that while allelic polymorphism is high, the repertoire of HVR sequences at these loci is limited. Evidence for segmental interallelic recombination at these loci has previously been noted (9, 14). Since interallelic recombination tends to obscure genealogies, limiting phylogenetic studies, a Monte-Carlo simulation was employed to analyze specifically the HVRs within these alleles. A greater than expected nucleotide sequence similarity was found, excluding the possibility of convergent evolution. Furthermore, no evidence for an increased mutational rate at these loci was found, implying that interallelic assortment of HVR sequences, rather than segmental "hypervariability," is a major mechanism for generating allelic polymorphism at these loci.

Recently, a putative chi-like recombination sequence was identified in the second exon of DRβ (14). A similar sequence exists in the analogous position of I-Eβ, while the analogous region in DQβ and I-Aβ encompasses HVR_{III} and contains no chi-like sequences. Recent data have suggested that some DQβ sequences have diverged subsequent to human–primate divergence (35). Our data show that it is primarily the COOH

Table 5. Interspecies codon identity at variable positions of the β1 domain

Comparison	Positions	No. of amino acids			Codon identity		Monte Carlo	
		Mouse	Human	Analyzed	Obs.	Exp.	Mean	Significance
DRβ vs. I-Eβ	9, 10, 11, 13, 28, 30, 37, 71, 74, 86	22	13	11	10	4.49	4.47	$P < 0.0001$
DQβ vs. I-Aβ	9, 14, 26, 28, 38, 57, 66, 70, 71	22	11	10	5.5	3.74	3.73	$P = 0.12$
DQβ vs. I-Aβ (NH_2 end)	9, 14, 26, 28, 38	15	6	6	5.5	2.53	2.55	$P = 0.006$

Table 6. Rate of nucleotide substitution between orthologous HVR sequences of mice and humans

Sequences	d_s	λ
DRβ vs. I-Eβ HVR_{I-III}	0.433	2.89×10^{-9}
DQβ vs. I-Aβ HVR_{I-IV}	1.11	7.33×10^{-9}
DQβ vs. I-Aβ HVR_{I-III}	0.401	2.67×10^{-9}

d_s, Number of synonymous substitutions per silent site; λ, nucleotide substitution rate, in substitutions per base pair per year.

terminus of the β1 domain that shares no homology, or ancestral relationship, with mouse I-Aβ sequences, while the variable amino acid sequences in the NH_2 half of the β1 domain are conserved. The divergent sequences encode the α-helical region of the molecule, according to a class II model (28). The lack of evolutionary constraint on this region of the molecule may represent divergent function of DQβ and I-Aβ or less structural constraint on pairing of the α and β chains at these positions.

Generation of new sequences, by either mutation or recombination, and subsequent fixation in a population by selection are distinct processes. Recombination between sequences without selection will tend to homogenize recombinants, decreasing sequence variability, while at the same time increasing differences between recombinants and those sequences not involved in recombination. Overdominant selection (heterozygote advantage), postulated as a mechanism for maintenance of polymorphism at both class I (36) and class II (34) loci, would tend to maintain polymorphism among recombinants. Among class II loci some form of balancing selection (31) at the MHC may be present, with selection of new alleles constrained by structural and functional requirements. Interallelic recombination appears to occur relatively rarely, with selection for recombinants with intact functional HVR sequences or combinations of HVR sequences. Peptide binding requires key residues with charged or hydrophobic functional groups, and it is likely that only certain HVR sequences and combinations will function together in α/β chain pairing, antigen binding, and recognition by T cells. Indeed, an inverse relationship between the variability of the α chain between species and the diversity of the associated β chains has been observed (3).

We have provided evidence that allelic sequences in HVRs are ancient in evolution, highly conserved, and evolve at the same rate as the rest of the genome. It is likely that similar HVR sequences will be found in lower vertebrates. Previously published interspecies similarities of β1 domain class II sequences between humans and mice (16), mice and rats (9), and humans and other primates (2, 14) are all wholly or in part contained within an HVR, rather than across the entire β1 domain. Each HVR can be considered a functional "corner" of the class II molecule, with selection at the class II loci for particular HVR sequences and combinations supported by a highly conservative framework. The HVR represents a unit of recombination between class II alleles, and it may be the genetic unit subject to selection in evolution.

The authors thank Drs. Doug Brutlag, Samuel Karlin, and Peter Parham and the Department of Statistics for much helpful advice.

1. Figueroa, F., Eberhardt, G. & Klein, J. (1988) *Nature (London)* **335**, 265–267.
2. Fan, W., Kasahara, M., Gutknecht, J., Klein, D., Mayer, W. E., Jonker, M. & Klein, J. (1988) *Hum. Immunol.* **26**, 107–122.
3. Gyllensten, U. B. & Erlich, H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9986–9990.
4. Lawlor, D., Ward, F. G., Ennis, P. D., Jackson, A. D. & Parham, P. (1988) *Nature (London)* **335**, 268–271.
5. Klein, J. (1987) *Hum. Immunol.* **19**, 155–162.
6. Bell, J. I., Denny, D., Foster, L., Belt, T., Todd, J. A. & McDevitt, H. O. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6234–6238.
7. Gustafsson, K., Wiman, K., Emmoth, E., Larhammar, D., Bohme, J., Hyldig-Nielsen, J. J., Ronne, H., Peterson, P. A. & Rask, L. (1984) *EMBO J.* **3**, 1655–1661.
8. Mengle-Gaw, L. & McDevitt, H. O. (1985) *Annu. Rev. Immunol.* **3**, 367–396.
9. She, J. X., Boehme, S. A., Wang, T. W., Bonhomme, F. & Wakeland, E. K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 453–457.
10. McConnell, T. J., Talbot, W. S., McIndoe, R. A. & Wakeland, E. K. (1988) *Nature (London)* **332**, 651–654.
11. Gorski, J. & Mach, B. (1986) *Nature (London)* **322**, 67–70.
12. Lee, B. S. M., Bell, J. I., Rust, N. A. & McDevitt, H. O. (1987) *Immunogenetics (New York)* **26**, 85–91.
13. Wu, S., Sanders, T. & Bach, F. (1986) *Nature (London)* **324**, 676–679.
14. Gyllensten, U. B., Sundvall, M. & Erlich, H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3686–3690.
15. Hayashida, H. & Miyata, T. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2671–2673.
16. McDevitt, H. O., Wraith, D. A., Smilek, D. E., Lundberg, A. S. & Steinman, L. (1989) *Cold Spring Harbor Symp. Quant. Biol.* **54**, 853–857.
17. Gregersen, P. K., Todd, J. A., Erlich, H. A., Long, E., Servenius, B., Choi, E., Kao, H. T. & Lee, J. S. (1989) in *Immunobiology of HLA*, ed. Dupont, B. (Springer, New York), pp. 1027–1031.
18. Mengle-Gaw, L. & McDevitt, H. O. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2910–2914.
19. Mengle-Gaw, L. & McDevitt, H. O. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7621–7625.
20. Saito, H., Maki, R. A., Clayton, L. K. & Tonegawa, S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5520–5524.
21. Malissen, M., Hunkapiller, T. & Hood, L. (1983) *Science* **221**, 750–753.
22. Larhammer, D., Hammerling, G., Denaro, M., Lund, T., Flavell, R. A., Rask, L. & Peterson, P. (1983) *Cell* **34**, 179–188.
23. Estess, P., Begovich, A. B., Koo, M., Jones, P. P. & McDevitt, H. O. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3594–3598.
24. Acha-Orbea, H. & McDevitt, H. O. (1988) *Proc. Natl. Acad. Sci. USA* **84**, 2435–2439.
25. Wu, T. T. & Kabat, E. A. (1970) *J. Exp. Med.* **132**, 211–250.
26. McCaldon, P. & Argos, P. (1988) *Proteins* **4**, 99–122.
27. Aota, S., Gojobari, T., Ishibashi, F., Maroyama, T. & Ikemura, T. (1988) *Nucleic Acids Res.* **16** Suppl., 315–402.
28. Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J. & Wiley, D. C. (1988) *Nature (London)* **332**, 845–850.
29. Stuart, C. B., Schilling, J. W. & Wilson, A. C. (1987) *Nature (London)* **330**, 401–404.
30. Gustafsson, K., Germana, S., Hirsch, F., Pratt, K., LeGuern, C. & Sachs, D. H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9798–9802.
31. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York).
32. Sakoyama, Y., Hong, K.-J., Byun, S. M., Hisajima, H., Ueda, S., Yaoita, Y., Hayashida, H., Miyata, T. & Honjo, T. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1080–1084.
33. Wu, C.-I. & Li, W.-H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1741–1745.
34. Hughes, A. L. & Nei, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 958–962.
35. Gyllensten, U. B., Lashkari, D. & Erlich, H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1835–1839.
36. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335**, 167–170.