Nucleotide sequences of chimpanzee MHC class I alleles: evidence for *trans*-species mode of evolution

W.E.Mayer¹, M.Jonker², D.Klein¹, P.Ivanyi³,
G.van Seventer³ and J.Klein¹

¹Max-Planck-Institut für Biologie, Abteilung Immunogenetik, Corrensstrasse 42, D-7400 Tübingen, FRG, ²TNO-REP-Institutes, Primate Center, Lange Kleiweg 151, 2288GJ Rijswijk and ³Central Laboratory of the Netherlands, Red Cross Blood Transfusion Service and Department of Experimental and Clinical Immunology, University of Amsterdam, Plesmanlaan 125, 1006 AD Amsterdam, The Netherlands

Communicated by J.Klein

To obtain an insight into the evolutionary origin of the major histocompatibility complex (MHC) class I polymorphism, a cDNA library was prepared from a heterozygous chimpanzee cell line expressing MHC class I molecules crossreacting with allele-specific HLA-A11 antibodies. The library was screened with human class I locus-specific DNA probes, and clones encoding both alleles at the A and B loci have been identified and sequenced. In addition, the sequences of two HLA-A11 subtypes differing by a single nucleotide substitution have been obtained. The comparison of chimpanzee and human sequences revealed a close similarity (up to 98.5%). The chimpanzee A locus alleles showed greatest similarity to the human *HLA-A11/A3* family of alleles, one of them being very close to *HLA-A11*. Similarly, segments of the *ChLA-B* alleles displayed greatest similarity to certain *HLA-B* alleles. The calculated evolutionary branch point for the *A11*-like alleles is 7×10^6 to 9×10^6 years, whereas the other A locus alleles diverged between 12×10^6 and 17×10^6 years ago. Since the human and chimpanzee lineages separated 5×10^6 to 7×10^6 years ago, our data support the notion that during evolution, MHC alleles are transmitted from one species to the next.

Key words: MHC/evolution/HLA-A11/ChLA/hominid phylogeny

Introduction

Major histocompatibility complex (MHC) encoded class I molecules are ubiquitous cell-surface glycoproteins that serve as recognition elements for cytotoxic T lymphocytes in allograft rejection or in the destruction of virus-infected cells (Klein, 1986). They are expressed as heterodimers of heavy (α) and light (β) chains. The heavy chain consists of two polymorphic extracellular domains ($\alpha 1$ and $\alpha 2$), a conserved extracellular domain ($\alpha 3$), a transmembrane region, and a short cytoplasmic tail (Steinmetz and Hood, 1983). The light chain, β_2 -microglobulin (β_2m), is non-covalently bound to the $\alpha 3$ domain. Both β_2m and the $\alpha 3$ domains have an immunoglobulin-like fold, whereas the $\alpha 1$ and $\alpha 2$ domains form a platform composed of β -sheets, with two ridges formed by exposed α -helices (Bjorkman *et al.*, 1987a,b).

The latter two domains are believed to interact with the T-cell receptor.

The functional MHC loci display high polymorphism. In humans, for example, 23, 47 and 8 alleles have been defined at the *HLA-A*, *HLA-B* and *HLA-C* loci respectively. Alleles at one locus can differ almost as much as class I loci differ from each other. Most of the observed variability is confined to the $\alpha 1$ and $\alpha 2$ domains.

The origin of the polymorphism is still obscure. Since the evolutionary rate of the MHC loci is not higher than that of other loci (Hayashida and Miyata, 1983; Klein and Figueroa, 1986), the presumed rapid diversification within species has been explained by mechanisms such as gene conversion, that lead to exchanges of whole DNA stretches within a gene (Lalanne *et al.*, 1982; Weiss *et al.*, 1983; Jaulin *et al.*, 1985; Geliebter and Nathanson, 1987; Szöts *et al.*, 1986). However, an alternative explanation has been proposed by one of us (Klein, 1980, 1987). According to this explanation, the evolution of MHC polymorphism is a *trans*-species process, starting from a group of ancestral major alleles that are passed on in phylogeny from one species to another, while accumulating further mutations. The observation that the chimpanzee MHC is organized into two or three functional class I loci like the human MHC (*ChLA-A*, *-B*, *-C*; Balner *et al.*, 1974, 1978), and that the products of these loci show crossreactivity with HLA antibodies, enabled us to test the latter hypothesis by comparing ChLA sequences, obtained from cDNA clones, with their human counterparts.

Results

Characterization of HLA-A11 subtype cDNA clones

Antibodies reacting specifically with HLA-A11 proteins also recognize class I molecules of some chimpanzees. To be able to compare the HLA-A11 and ChLA-A108 sequences we constructed a cDNA library from the human EBV-transformed cell line K.LIE, carrying two subtypes of HLA-A11 (Neefjes *et al.*, 1986). The library was screened with the *HLA-A* locus-specific probe pHLA-2a.1. Full-length inserts of positive clones were subcloned in pUC18. The insert of one subclone (pIII1-1A) was then sequenced completely. Of a second subclone (pV5-1B), only the sequence encoding the extracellular domains was determined. The nucleotide sequences are shown in Figure 1; the deduced amino acid sequences are shown in Figure 2 in comparison with other *HLA-A*-locus alleles. The two sequences differ from each other only by a single nucleotide substitution leading to a non-conservative single amino acid substitution at position 19. In pIII1-1A GAG at this position codes for glutamic acid, whereas in pV5-1B AAG codes for lysine. The two clones thus represent two subtypes of HLA-A11, which we shall provisionally refer to as HLA-A11E and HLA-A11K. The sequence of pIII1-1A is identical to that recently published by Cowan *et al.* (1987).

Leader Peptide

ChLA-A108 .GGCCGTCATGCCGCCCCGAACCTCTCCTGCTACTCTCGGGGGCCCTGGCCCTGACCCAGACCTGGGCG
ChLA-A126 AT-----G-----C-----G-----
HLA-A11 -----G-----C-----

α1 Domain

ChLA-A108 GGCTCCCCTCCATGAGGTATTCTACACCTCCGTGTCGCCGCCCCGGCGCGGGAGCCCCGCTTCATCGCCGTGGGTACGTGGACGAC
ChLA-A126 -----C-----A-----G-----
HLA-A11K -----A-----

ChLA-A108 ACGCAGTTCGTGCGGTTTCGACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGGGTCCGGAGTATTGG
ChLA-A126 -----
HLA-A11 -----G-----

ChLA-A108 GACCAGGAGACCGGAATATGAAGGCCTCCGCACAGACTGACCGAGTGGACCTGGGGACCCTGCGCGGCTACTACAACAGAGCGAGGAC
ChLA-A126 ---G-----G-----G-----
HLA-A11 -----G-----CAGT-----

α2 Domain

ChLA-A108 GGTTCTCACCATCCAGATAATGTATGGCTGCGACGTGGGGTCGGACGGGCGCTTCCTCCGCGGGTACCGGCAGGACGCCTACGACGGCAA
ChLA-A126 -----T-----T-----
HLA-A11 -----C-----

ChLA-A108 GGATTACATCGCCCTGAACGAGGACCTGCGCTCTTGGACCGCGCGGCCATGGCAGCTCAGATCACCAGCGCAAGTGGGAGGCGGCCATG
ChLA-A126 -----A-----G-----C-----
HLA-A11 -----A-----

ChLA-A108 CGGCGGAGCAGCTGAGAGCTACCTGGAGGGCCGGTGCCTGGAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAGACGCTGCAGCGCACG
ChLA-A126 -----AC-----
HLA-A11 -----A-----T-----

α3 Domain

ChLA-A108 GACCCCCCAAGACACATATGACCCACCACCCCATCTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGCTTCTACCCCTGCGGTGAT
ChLA-A126 -----C-----G-----
HLA-A11 -----A-----

ChLA-A108 CACACTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACGGAGCTCGTGGAGACCAGGCGCCGAGGGGATGGAACCTTCCAGAAGT
ChLA-A126 -----T-----
HLA-A11 -----T-----

ChLA-A108 GGGCGGCTGTGGTGGTACCTTCTGGAGAGGAGCAGAGATACACCTGCCATGTGCAGCATGAGGGTCTGCCCAAGCCCCCTACCCTGAGATGG
ChLA-A126 ---A-----
HLA-A11 -----G-----

Transmembrane Domain

ChLA-A108 GAGCCGCTCTCCAGCCCACCATCCCATTTGTGGGCATCATTGCTGGCCTGGTTCTCCTTGGAGCTGTGATCACTGGAGCTGTGGTTCGCT
ChLA-A126 -----C-----
HLA-A11 -----T-----C-----

ChLA-A108 GCCGTGATGTGGAGGAGGAAGAGCTCA
ChLA-A126 -----
HLA-A11 -----

Cytoplasmic Domains

ChLA-A108 GATAGAAAAGAGGGAGCTACACTCAGGCTGCAAGCAGTGACAGTGCCAGGGCTCTGATGTGTCTCTCACAGCTTGTAAAGTGTGA
ChLA-A126 -----T-----
HLA-A11 -----T-----

3' Untranslated Region

ChLA-A108 GACAGCTGCCTTGTGTGGGACTGAGAGGCAAGAGTTGTTCTGCCCTTCCCTTTGTGACTTGAAGAACCCTGACTTTGTTTCTGCAAAGG
ChLA-A126 -----
HLA-A11 -----

ChLA-A108 CACCTGCATGTGTCTGTGTTTCATGTAGGCATAATGTGTGGAGGTGGGAGACCACCCACCCCATGTCCACCATGACCCCTCTTCCACG
ChLA-A126 -----T-----A-----
HLA-A11 -----G-----A-----G-----

ChLA-A108 CTGACCTGTGCTCCCTCCCAATCATCTTTCTGTTCCAGAGAGGTGGGGCTGAGGTGTCTCCATCTCTGTCTCAACTTCATGGTCACT
ChLA-A126 -----
HLA-A11 -----T-----

ChLA-A108 GAGCTGTAACCTTCTTCCTTCCCTATTAAATAGAACCTGAGTATpoly (A)
ChLA-A126 -----AAATTACTTTCTCAAATCTTGCCATGAGAGGTTGATGAGTTAA
HLA-A11 -----poly (A)

ChLA-A126 TTAAAGGAGAAGATTCTAAAATTTGAGAGACAAATAAATGGAACCACATGAGpoly (A)

Fig.1. Nucleotide sequences of *HLA-A11K* and *ChLA-A* alleles. '-' indicates identity with ChLA-A108 and '.' unidentified nucleotides. The polyadenylation signals are underlined.

Identification and characterization of cDNA clones encoding ChLA sequences

To test whether the available probes could detect class I genes of the chimpanzee, Northern and genomic Southern blot analyses were performed. At stringent hybridization conditions the probes can be expected to give a similar pattern of bands with human and chimpanzee RNA or DNA, but no hybridization with mouse nucleic acids. As demonstrated in Figure 3B, both locus-specific probes hybridized with 17S transcripts in human and chimpanzee poly(A)⁺ RNA (lanes 2, 5, 8 and lanes 1, 4, 7) but not in mouse RNA (lanes 3, 6, 9). In the genomic Southern blot (Figure 3A), at least 16 bands, corresponding to all class I genes present in the genome (including non-expressed genes), were detected when probed with the HLA-A11 cDNA clone pIII1-1A. About half of the bands appeared to be shared between the human and the chimpanzee DNA (lanes 1 and 2), whereas the mouse DNA (from strains DBA/2 and C3H/HeJ; lanes 3 and 4) gave a different pattern and weaker bands. The overall number of bands was about the same in all three species. The *HLA-B* locus-specific probe pHLA-1.1 recognized a single *Pvu*II fragment of ~400 bp in both human and chimpanzee DNA (lanes 9 and 10). There was no hybridization with mouse DNA. Exclusive recognition of primate DNA was also found in the case of the *HLA-A* locus-specific probe pHLA-2a.1. This probe hybridized strongly with a class I gene carried on a 2.3 kb fragment (the *HLA-A* and *ChLA-A* genes) and weakly with two to three other class I genes (lanes 5 and 6).

A cDNA library was then prepared from the chimpanzee cell line Hugo-EBV. Screening of 160 000 clones prior to amplification yielded 24 clones positive with the probe pHLA-2a.1 and 27 clones positive with pHLA-1.1. No clone was recognized simultaneously by both probes. This high number was not unexpected, since it is known that Epstein-Barr virus (EBV) can enhance the expression of class I proteins (McCune *et al.*, 1975). Clones containing full-length inserts were subcloned and tested by restriction enzymes and Southern blot analysis for orientation and polymorphism (data not shown). All tested *A*-locus-like and *B*-locus-like cDNA inserts fell into two groups, corresponding to the two alleles at each locus (Figure 4). The complete nucleotide sequence of one representative of each group was obtained (Figures 1 and 5). Figures 2 and 6 show the alignment of the derived amino acid sequences with several HLA-A or -B sequences obtained by other investigators. Since the product of one of the *ChLA-A*-alleles is recognized by HLA-A11-specific antibodies, we denote here the sequence most similar to HLA-A11 as ChLA-A108 and refer to the allelic sequence as ChLA-A126. In the absence of any serological information about the *ChLA-B* locus products, we refer to the alleles coding for these products as *ChLA-B1* and *ChLA-B2*.

It is evident that the HLA and the ChLA class I proteins are closely related. There exists only a small number of substitutions which are confined mainly to the polymorphic first two extracellular domains. Table I summarizes the amino acid as well as the nucleotide differences. Table II then gives the weighted averages of nucleotide substitutions expressed as percentages of potential silent and replacement substitution sites. As was suggested by Miyata *et al.* (1980, 1985), the degree of divergence in silent (or synonymous) positions of a gene in different species is a measure for the

time that has passed since the species had their last common ancestor. In our case it gives an estimate of the evolutionary distance between related *HLA-A* and *ChLA-A* alleles.

Discussion

Comparison of chimpanzee and human class I alleles

The main conclusion that can be drawn from the comparison of ChLA and HLA class I sequences is that some of the human alleles are much more similar to certain chimpanzee alleles than they are to certain human alleles. This conclusion is best borne out by the comparison of ChLA-A108 and HLA-A11 sequences. The two DNA sequences differ in 17 nucleotides (overall sequence similarity of 97.5%) and the two corresponding proteins differ in nine amino acid residues, while some of the *HLA-A* alleles differ from *HLA-A11* in 45–54 nucleotides and the corresponding proteins differ in 25–32 amino acid residues. Only related alleles, such as *HLA-A11* and *HLA-A3* or *HLA-A2* and *HLA-A28*, differ from each other in such a small number of nucleotides as the genes in the two primate species. The similarity between *HLA-A11* and *ChLA-A108* genes extends over the entire sequence from the first exon to the 3' untranslated region. It is particularly striking at certain positions. For example, at position 90 of the mature protein, ChLA-A108 (as well as ChLA-A126 and HLA-A11) have Asp, while all other sequenced HLA-A allomorphs have Ala. Similarly, at position 163, ChLA-A108 and HLA-A11 have Arg, whereas all other HLA-A allomorphs have Thr. More significantly, in the 3' untranslated region, *ChLA-A108* and *HLA-A11* use a different polyadenylation site than all other *HLA-A* genes do (ATTAAA instead of AATAAA; see also Birnstiel *et al.*, 1985). The aberrant site is located 101 nucleotides upstream from the site used by other genes so that the ChLA-A108 and HLA-A11 are shorter at the 3' end than other sequenced class I cDNAs.

The second chimpanzee allele, *ChLA-A126*, shows a less striking similarity to *HLA-A11*, although it is clearly related to it. It belongs to the group of *HLA-A11/HLA-A3* alleles, within which the alleles are more related to each other than they are to other *HLA-A* alleles. The *ChLA-A108* allele uses the more common polyadenylation site just as the *HLA-A3* alleles do.

Assuming that all changes in the sequence were caused by point mutations and that transitions and transversions occurred with the same frequency, and using the rate of substitutions at silent positions as a molecular clock (Kimura and Ohta, 1972), we can calculate the time (*T*) of divergence of the individual alleles from the formula $T = K_s^c/2V_s$. In this formula, K_s^c is the average number of silent substitutions per potential site corrected for multiple events and V_s is the mean evolutionary rate. For chimpanzee-human comparisons, Li *et al.* (1987) estimated V_s to be equal to 1.2×10^{-9} per site per year; for Old World monkey-human comparisons Sakoyama *et al.* (1987) estimated V_s to be equal to 1.56×10^{-9} per site per year. Using the two values, we calculate the times of divergence for the individual allelic pairs as follows: ChLA-A108-HLA-A11E, 7.1–9.3 million years (myr) ago; ChLA-A108-HLA-A3.2, 12–15 myr ago; HLA-A11E-HLA-A3.2, 7.3–9.7 myr ago [if we accept the notion that the rate of molecular evolution is 1.2 times (Li *et al.*, 1987) to 1.6 times (Li and Tanimura, 1987) slower

Leader Peptide

```

ChLA-A126      -----V-----
ChLA-A108      .--P-----
                -20      -10
HLA-A11        MAVMAPRTL L L L L LSGALALTQTWA
HLA-A3         MARGDQ-----
HLA-A2.1       -----V-----
HLA-A2.2Y      -----V-----
HLA-A2.3       -----V-----
HLA-Aw24       ---G---V-----
HLA-Aw68.1     -----V-----
HLA-Aw68.2     -----V-----
HLA-Aw69       -----V-----

```

α1 Domain

```

ChLA-A126      -----S-----E--S--SA-----
ChLA-A108      -----M--SA-----
                10      20      30      40      50      60      70      80      90
HLA-A11E      GSHSMRYFYTSVSRPGRGEPRFIAVG YVDDTQFVRFDSDAASQRM EPRAPWIEQEGPEYWDQETRN VKAQSQTDRVDLGT LRGYYNQSED
HLA-A11K      -----K-----
HLA-A3.1      -----F-----A
HLA-A3.2      -----F-----A
HLA-A2.1      -----F-----G--K--H--H-----A
HLA-A2.2F     -----F-----R-----G--K--H--H-----A
HLA-A2.2Y     -----R-----G--K--H--H-----A
HLA-A2.3      -----F-----G--K--H--H-----A
HLA-A2.4.1    -----G--K--H--H-----A
HLA-A2.4.2    -----G--K--H--H-----A
HLA-A28       -----RN-----A
HLA-Aw24      -----S-----E--GK--H----EN-RIALR-----A
HLA-A32       -----F-----H----ES-RIALR-----A
HLA-Aw68.1    -----RN-----A
HLA-Aw68.2    -----M-----RN-----A
HLA-Aw69      -----RN-----A

```

α2 Domain

```

ChLA-A126      -----L-F---S-----Q-----L---T-----
ChLA-A108      -----S-----A-----L-----
                100      110      120      130      140      150      160      170      180
HLA-A11E      GSHTIQIMYGCDVGP DGRFLRGYRQDAYDGKDYIALNEDLRSWTAADMAAQITKRKWEAAHAAEQQRAYLEGR CVEWLRRYLENGKETLQRT
HLA-A11K      -----
HLA-A3.1      -----S-----E--L---D-T-----
HLA-A3.2      -----S-----V-----D-T-----
HLA-A2.1      ---V-R---S-W---H-Y---K---T-H---V--L---T-----
HLA-A2.2F     ---L-R---S-W---H-Y---K---T-H---V--W---T-----
HLA-A2.2Y     ---L-R---S-W---H-Y---K---T-H---V--W---T-----
HLA-A2.3      ---V-R---S-W---H-Y---K---T-H---T-E--W---T-----
HLA-A2.4.1    ---V-R---S-W---H-Y---K---T-H---V--L---T-----
HLA-A2.4.2    ---V-R-.-S-W---H-Y---K---T-H---V--L---T-----
HLA-A28       .-.-S-----K-.-.-.-T-H---V--.-.-T-.-.-.-
HLA-Aw24      ---L-M-F---S-----H-Y---K---V-----T-DG-----A
HLA-A32       ---M-----L---Q-----Q---RV--L---T-----
HLA-Aw68.1    ---M---S-----K---T-H---V--W---T-----
HLA-Aw68.2    ---R-----H-Y---K---T-H---V--W---T-----
HLA-Aw69      ---V-R---S-W---H-Y---K---V--T-H---V--L---T-----

```

α3 Domain

```

ChLA-A126      -----R-----
ChLA-A108      -----V-----

```

	190	200	210	220	230	240	250	260	270
HLA-A11E	DPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELVETRPAGDGTQKWAAVVVP	SGEEQRYTCHVQHEGLPKPLTLRW							
HLA-A11K	-----	-----	-----	-----	-----	-----	-----	-----	-----
HLA-A3.1	-----	-----	-----	-----	-----	-----	-----	-----	-----
HLA-A3.2	-----	-----	-----	-----	-----	-----	-----	-----	-----
HLA-A2.1	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A2.2F	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A2.2Y	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A2.3	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A2.4.1	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A2.4.2	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-A28	-A-...-AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-Aw24	-----	-----	-----	-----	-----	-----	-----	-----	-----
HLA-A32	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
HLA-Aw68.1	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----V-----Q-----	-----	-----
HLA-A68.2	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----V-----Q-----	-----	-----
HLA-A69	-A-----AV-----S-----	-----	-----	-----	-----	-----	-----Q-----	-----	-----
<u>Transmembrane Domain</u>					<u>Cytoplasmic Domains</u>				
ChLA-A126	-P-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
ChLA-A108	-P-----	-----	-----	-----	-----	-----	-----	-----	-----
	280	290	300	310		320	330	340	
HLA-A11E	ELSSQPTIPIVGIIAGLVLLGAVITGAVVAAMWRRKSS					DRKGGSYTQAASSDSAQGSVDVSLTACKV			
HLA-A3.1	-----	-----	-----	-----	-----	-----	-----	-----	-----
HLA-A2.1	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-A2.2Y	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-A2.3	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-Aw24	-P-----V-----N-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-Aw68.1	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-Aw68.2	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----
HLA-Aw69	-P-----F-----	-----	-----	-----	-----	-----S-----	-----	-----	-----

Fig. 2. Amino acid sequences, in the single-letter code, predicted from the nucleotide sequences of *ChLA*- and *HLA-A* alleles. '-' indicates identity with HLA-A11E and '.' unidentified residues. The sequences are from the following sources: ChLA-A108, ChLA-A126, HLA-A11E, HLA-A11K, this report; HLA-A3.1, Strachan *et al.* (1984); HLA-A3.2, Cowan *et al.* (1985); HLA-A2.1, Koller and Orr (1985), Mattson *et al.* (1987); HLA-A2.2F, HLA-A2.3, Mattson *et al.* (1987), Krangel *et al.* (1982, 1983); HLA-A2.2Y, Holmes *et al.* (1987); HLA-A2.4.1, Ezquerria *et al.* (1986a); HLA-A2.4.2, Ezquerria *et al.* (1986b); HLA-Aw24, N'Guyen *et al.* (1985); HLA-A28, Lopez de Castro *et al.* (1982); HLA-A32, Wan *et al.* (1986); HLA-Aw68.1, HLA-Aw69, Holmes and Parham (1985); HLA-Aw68.2, Holmes *et al.* (1987).

in the human than in the chimpanzee lineage]; ChLA-A126—HLA-A3.2, 12–17 myr ago. Since the time of separation of the human and chimpanzee lineages is estimated to be 3.1–8.4 myr ago (Hasegawa *et al.*, 1987; Li and Tanimura, 1987; Sakoyama *et al.*, 1987; Sibley and Ahlquist, 1987), the divergence of even the closely related *HLA-A11* and *HLA-A3.2* must have pre-dated human speciation. The less related *HLA-A* alleles, such as *HLA-A2*, must have diverged from *HLA-A11* much earlier than that. These data thus clearly support the *trans*-species origin of *MHC* alleles proposed by one of us several years ago (Klein, 1980).

Further support for the *trans*-species hypothesis of *MHC* polymorphism is provided by the analysis of the ChLA-B and HLA-B sequences, although here the situation is more complicated than in the case of the ChLA-A and HLA-A sequences. There is no single *HLA-B* allele that would be more similar to *ChLA-B1* or *ChLA-B2* over the entire length of the gene than other alleles. There are, however, sequence segments within the *ChLA-B* genes that show a striking similarity to segments within certain *HLA-B* genes. Thus the first exon of both *ChLA-B1* and *ChLA-B2* is strikingly similar to the first exon of *HLA-B7*; the exons coding for the $\alpha 2$,

$\alpha 3$, transmembrane and cytoplasmic domains of *ChLA-B2* resemble *HLA-Bw58*, and so do the exons coding for the transmembrane and cytoplasmic regions of *ChLA-B1*. The similarity among second exons of the different alleles is more difficult to ascertain. Part of the $\alpha 1$ domain of ChLA-B1 resembles HLA-Bw58 and this resemblance is strengthened by the observation that Hugo cells were typed as cross-reacting with HLA-B17-specific antibodies (Balner *et al.*, 1978) and that the HLA-B17 antigen was later split into HLA-Bw57 and HLA-Bw58 (Albert *et al.*, 1984). The domain, however, also shows some resemblance to HLA-B27.2.

Segments of the second exon of *ChLA-B2* resemble HLA-B40, -Bw60, -B7, -B13 and -B44. These data can best be explained by postulating that some of the exons have been reshuffled by intragenic crossing-over in the ancestral alleles from which the human and chimpanzee *HLA-B* alleles originated. It is also possible that among the *HLA-B* alleles that have not yet been sequenced, there will be some that resemble *ChLA-B1* and *ChLA-B2* along their entire sequence. Since the transfer of genetic information between humans and chimpanzees is not possible, the similarities between

Fig. 3. (A) Southern blot analysis of genomic DNA isolated from human and chimpanzee cell lines and from mouse kidneys. Ten micrograms of genomic DNA from K.LIE-EBV (lanes 1, 5, 9), Hugo-EBV (lanes 2, 6, 10), DBA/2 (lanes 3, 7, 11) and C3H/HeJ (lanes 4, 8, 12) were digested with *PvuII*, electrophoresed in a 0.8% agarose gel and transferred to nitrocellulose filters. The filters were then hybridized with pIII1-1A (lanes 1–4), pHLA-2a.1 (lanes 5–8) and pHLA-1.1 (lanes 9–12) probes. (B) Northern blot analysis of RNA isolated from human and chimpanzee cell lines and from mouse spleen. One microgram of poly(A)⁺ RNA was electrophoresed in a formaldehyde-agarose gel and transferred to nitrocellulose filters. The filters were hybridized with pIII1-1A (lanes 1–3), pHLA-1.1 (lanes 4–6) and pHLA-2a.1 (lanes 7–9) probes. Lanes 1, 4, 7: Hugo-EBV RNA; lanes 2, 5, 8: K.LIE-EBV RNA; lanes 3, 6, 9: mouse spleen RNA.

portions of *ChLA-B* and *HLA-B* alleles cannot be explained by gene conversion. The fact that these segments of *ChLA-B* genes are more similar to the segments of certain *HLA-B* genes than the *HLA-B* alleles are to each other, again suggests that the *HLA-B* alleles diverged before the separation of human and chimpanzee lineages.

The cluster of substitutions in the transmembrane domain of ChLA-B1 can best be explained by assuming a duplication of codons 302–304, possibly by a slippage event during replication (Streisinger *et al.*, 1966), followed by a deletion at codons 294–296, and mutation of Ile to Thr and of Gly to Val (Figure 7). Therefore, the nucleotide sequence of codons 295–301 of ChLA-B1 can be aligned to the allelic sequence of codons 298–304 with just two nucleotide mismatches in the codons for Thr-297 and Val-298.

The pre-speciation divergence of *HLA* alleles has important implications for the origin of man (J.Klein *et al.*, submitted). If all major alleles pre-date the rise of modern man, it means that at any time during the evolution of the species, the population could not have been smaller than the size required to maintain these alleles at appreciable frequencies. Once the number of major *HLA* alleles is determined, it should be possible to calculate the minimal size of the population from which human evolution began.

Materials and methods

Cell lines

K.LIE-EBV is an EBV-transformed human B-cell line derived from peripheral blood lymphocytes of an Indonesian individual (*HLA* haplotypes *A11.1*, *B27D*, *Cw3*, *DRw12/A11.2*, *B27C*, *Cw-*, *DR4*). Hugo-EBV is an EBV-transformed B-cell line from a chimpanzee in the Dutch Primate Center at Rijswijk, The Netherlands (*ChLA* haplotypes *A108*, *B118/A126*, *B125*). Serological crossreactivities: ChLA-A108–HLA-A11; ChLA-A126–HLA-A31; ChLA-B125–HLA-B17 (Bw57 and Bw58).

Preparation of cDNA libraries

Total RNA was prepared from the cell lines by homogenization in guanidine

Fig. 4. Restriction maps of the HLA-A11, ChLA-A108, ChLA-A126, ChLA-B1 and ChLA-B2 cDNA clones. The coding region (solid bar) and the 3' untranslated region (thin line) are indicated. Restriction sites are K: *KpnI*, P: *PstI*, Pv: *PvuII*, R: *RsaI*.

isothiocyanate (Chirgwin *et al.*, 1979) and sedimentation in 5.7 M CsCl (Glisin *et al.*, 1974). The cDNA was synthesized from the poly(A)⁺ fraction according to the procedure of Gubler and Hoffman (1983) using Moloney murine leukemia virus reverse transcriptase (BRL) for first-strand synthesis and RNase H (Pharmacia) and *Escherichia coli* polymerase I (Boehringer Mannheim) as well as *E. coli* DNA ligase (Pharmacia) for second-strand synthesis. After the methylation of internal *EcoRI* sites with *EcoRI* methylase (New England Biolabs), linker addition and digestion with *EcoRI*, the cDNA was inserted into λ gt10 (K.LIE-EBV cDNA) or NM 1149 (Hugo-EBV cDNA). Following *in vitro* packaging with Gigapack Plus (Stratagene) recombinant phages were plated directly (using *E. coli* C600hfl or POP 13b) and screened by hybridization with the *HLA-A* locus-specific probe pHLA-2a.1 and the *HLA-B* locus-specific probe pHLA-1.1 (Koller *et al.*, 1984). Both probes are derived from 3' untranslated regions. The screening procedures were as described in Maniatis *et al.* (1982). The hybridization was carried out in 5 $\times$ SSPE (1 $\times$ SSPE is

Leader Peptide

ChLA-B1 ATGCAGGTACGGCGCCCCGAACGTCCTCTGCTGCTCTCGGCGGCCCTGGCCCTGACCGAGACCTGGGCC
 ChLA-B2-----

$\alpha 1$ Domain

ChLA-B1 GGCTCCCACTCCATGAAGTATTTCTACACCTCCGTGTCCCGCCCCGGCCGCGGGAGCCCCGCTTCATCACCGTGGGTACGTGGACGAC
 ChLA-B2 -----G-----G-----T-A-----

ChLA-B1 ACCCAGTTCGTGAGGTTTCGACAGCGACGCCGCGAGTCCGAGCATGGAGCCCCGGCGCCATGGATAGAGCAGGAGGGCCGGAGTATTGG
 ChLA-B2 --G-----T-----AGA-----G-----

ChLA-B1 GACCGGGAGACACGGAATATGAAGGCCTCCGCACAGACTGACCGAGAGAACCTGCGGATCGCGCTCCGCTACTACAACCAGAGCGAGGCC
 ChLA-B2 -----A-----A--TCTCC--A--AA-----T-----G-----A-CT--G-G-----

$\alpha 2$ Domain

ChLA-B1 GGGTCTCACACCTGGCAGACGATGTATGGCTGCGACATGGGGCCGGACGGGCGCCTCCTCCGCGGTATGGCCAGTACGCCTACGACGGCAA
 ChLA-B2 --T-----T-ATC---G-----C-----AG-----

ChLA-B1 GGATTACATCGCCCTGAACGAGGACCTGAGCTCCTGGACCGCGCGGACACCGCGGCTCAGATACCCAGCGCAAGTGGGAGGCGGCCCGTG
 ChLA-B2 -----T-----

ChLA-B1 AGGCGGAGCAGCGGAGAGCCTACCTGGAGGGCAGTGCCTGGAGTGGCTCCGCAGATACCTGGAGAAGCGGAAGGAGACGCTGCAGCGCGTG
 ChLA-B2 G-----T-----

$\alpha 3$ Domain

ChLA-B1 GACCCCCCAAAGACACACGTGACCCACCACCCCATCTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGCTTCTACCTGCGGAGAT
 ChLA-B2 -----T-----T-----

ChLA-B1 CACACTGACCTGGCAGCGGGATGGCGAGGACCAACTCAGGACACCGAGCTTGTGGAGACCAGGCCAGAAGGAGATAGAACCTTCCAGAAGT
 ChLA-B2 -----A-----C-----

ChLA-B1 GGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGATACACATGCCATGTGCAGCATGAGGGGTGCCGAAGCCCCCTACCCCTGAGATGG
 ChLA-B2 -----A-----

Transmembrane Domain

ChLA-B1 GAGCCATCTTCCAGTCCACCATCCCATTTGTGGGCATTGTTGTGCTGGCTGGCTGCTCTTGTGGTACCGTAGCTGTGGTTCGCTGTGGTG
 ChLA-B2 -----C-----A-CA--T --G-TCA-C-GA-----

ChLA-B1 GCTGCTGTGATGTGTAGGAGGAAGAGCTCA
 ChLA-B2 -----

Cytoplasmic Domains

ChLA-B1 GGTGGAAAAGGAGGGAGCTACTCTCAGGCTGCATCCAGCGACAGTGGCCAGGGCTCTGATGTGTCTCTCACAGCTTGA
 ChLA-B2 -----G-----

3' Untranslated Region

ChLA-B1 AAAGCCTGAGACAGCTGTCTTGTGAGGGACTGAGATGCAGGATTCTTCTACGCCTCCCC TTTGTGACTTCAAGAGCCTCTGGCATCTC
 ChLA-B2 -----A-----CC-----

ChLA-B1 TTTCTGCAAAGGCACCTGAATGTGTCTGCGTCCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCACCCTTGTGTCCACTGTGACC
 ChLA-B2 -----T-----

ChLA-B1 CCTGTTCCCATGCTGACCTGTGTTTCTCTCCCGAGTCATCTTCTTGTTCAGAGAGGTGGGGCTGGATATCTCCATCTCTGTCTCAACTT
 ChLA-B2 -----G-----

ChLA-B1 TATGTGCACTGAACTGCAACTTCTTACTTCCCTACTGAAAAATAAGAATCTGAATATAAATTTGTTTTCTCAAATATTGTGATGAGAGGT
 ChLA-B2 -----G-----

ChLA-B1 TGATGGATTAATTAATAAGTCAATTCTGGAATTTGAGAGAGCAATAAAGACCTGAGAACCTTCCpoly (A)
 ChLA-B2 -----poly (A)

Fig. 5. Nucleotide sequences of *ChLA-B* alleles. '-' indicates identity with *ChLA-B1* and '.' unidentified nucleotides. The polyadenylation signal is underlined.

150 mM NaCl, 10 mM NaH_2PO_4 , 1 mM EDTA, pH 7.4), 50% formamide, $5 \times$ Denhardt's solution, 250 $\mu\text{g}/\text{ml}$ denatured salmon sperm DNA, 1 $\mu\text{g}/\text{ml}$ poly(C) and 0.1% SDS at 42°C for 16 h. The concentration of the probes, labelled to $1-2 \times 10^9$ c.p.m./ μg by the hexamer priming method (Feinberg and Vogelstein, 1984), was 1×10^6 c.p.m./ml. The final wash was in $0.2 \times$ SSPE, 0.1% SDS at 65°C. Positive plaques were rescreened and then subcloned in the plasmid vector pUC18 (Yanisch-Perron *et al.*, 1985).

DNA and RNA blot analysis

Digested DNA was electrophoresed in 1% agarose gels in Tris-acetate buffer (Maniatis *et al.*, 1982), except for genomic DNA for which 0.8% agarose gels were used. The DNA was transferred to nitrocellulose membranes (Millipore) by capillary blotting with $20 \times$ SSPE (Southern, 1975) and hybridized as above.

Electrophoresis of RNA was carried out in 2.2 M formaldehyde, 1% agarose gels. Prior to the transfer to nitrocellulose membranes the gels were

Leader Peptide

ChLA-B1 .----V----SA-L-----
 ChLA-B2 -Q-----V----SA-L-----
 -20 -10
 HLA-B27.1 MRVTAPRTL L L L L L L L L WGAVALTETWA
 HLA-B7 -L-M----V----SA-L-----
 HLA-B13 -----
 HLA-B44 .-----
 HLA-Bw47 -----
 HLA-Bw58 -----V-----

 α 1 Domain

ChLA-B1 -----Y-----Q-----M-----RNM-S-----N--IA-----
 ChLA-B2 ----K--Y-A-----S-----Q-W-----S-TN--Y--S--N-RG-----
 10 20 30 40 50 60 70 80 90
 HLA-B27.1 GSHSMRYFHTSVSRPGRGEPFITVGYVDDTLFVRFDSDAASPREEPRAPWIEQEGPEYWDRETQICKAKAQTDREDLRTLRLRYNQSEA
 HLA-B27.2 -----N--IA-----
 HLA-B27.3 -----S-----
 HLA-B27.4 -----S-----
 HLA-B27f -----Y--N--A-----
 HLA-B7 -----Y-----S-----Q-----N--Y--Q-----S--N-RG-----
 HLA-B13 -----Y-AM-----Q-----T--MA-----S-TNT--Y--N--A-----
 HLA-B40 -----AM-----T--K-----S-TNT--Y--S--N-RG-----
 HLA-B40* -----T--K-----S-TNT--Y--S--N-RG-----
 HLA-B44 -----Y-AM-----T--K-----S-TNT--Y--N--AA-----
 HLA-Bw47 -----Y-AM-----T--K-----S-TNT--Y-----
 HLA-Bw58 -----Y-AM-----A-----Q-----T-----G--RNM--S--Y--N--IA-----
 HLA-Bw60 -----AM-----T--K-----S-TNT--Y--S--N-RG-----

 α 2 Domain

ChLA-B1 ----W-T----M-----G-Y-----E--R-----T-----
 ChLA-B2 ---II-R---M-S-----E-Y-----W-----T-----
 100 110 120 130 140 150 160 170 180
 HLA-B27.1 GSHTLQNMYGCDVGPDPGRLLRGYHQDAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAARVAEQLRAYLEGECEVWLRRLRYLENGKETLQRA
 HLA-B27.2 -----
 HLA-B27.3 -----E-----
 HLA-B27.4 -----D-Y-----E-----
 HLA-B27f -----
 HLA-B7 -----S-----HD-Y-----R-----E--R-----DK-E--
 HLA-B13 ---W-T---L-----HN-L-----L-----
 HLA-B40 -----R-----HN-Y-----R-----L-----DK-E--
 HLA-B40* -----S-----HN-Y-----R-----
 HLA-B44 ---II-R-----D-----D-----L--S-IA-----
 HLA-Bw47 -----R-F-----
 HLA-Bw58 ---II-R---L-----HD-S-----L-----
 HLA-Bw60 -----R-----HN-Y-----R-----S--L-----DK-E--

 α 3 Domain

ChLA-B1 -----E-----
 ChLA-B2 -----
 190 200 210 220 230 240 250 260 270
 HLA-B27 DPPKTHVTHHPISDHEATLRCAWALGFYPAEITLTWQRDGEDQTQDTELVEVTRPAGDRTFQKWAAVVVP SGEEQRYTCHVQHEGLPKPLTLRW
 HLA-B7 -----
 HLA-B13 -----
 HLA-B40 -----
 HLA-B40* -----
 HLA-B44 -----S-----
 HLA-Bw47 -----
 HLA-Bw58 -----V-----
 HLA-Bw60 -----

Transmembrane Domain					Cytoplasmic Domains	
ChLA-B1	-----I-----V-T-AVV-----				-----S-----	
ChLA-B2	-----I-----				-----S-----	
	280	290	300	310	320	330
HLA-B27	EPSSQSTVPVIGIVAGLAVLAV VVIGAVVAVMCRKSS				GGKGGSYSQAACSDSAQGSVDVSLTA	
HLA-B7	-----PS-----P-----				-----	
HLA-B13	-----				-----	
HLA-B44	-----				-----	
HLA-Bw47	-----				-----V-----	
HLA-Bw58	-----I-----T-----				-----S-----	

Fig. 6. Amino acid sequences, in the single-letter code, predicted from the nucleotide sequences of *ChLA*- and *HLA-B* alleles. '-' indicates identity with HLA-B27.1 and '.' unidentified residues. The sequences are from the following sources: ChLA-B1, ChLA-B2, this report; HLA-B7, Biro *et al.* (1982); HLA-B13, Zemmour *et al.* (1988); HLA-B27.1, Ezquerro *et al.* (1985), Weiss *et al.* (1985); HLA-B27.2, Vega *et al.* (1985a); HLA-B27.3, Vega *et al.* (1985b); HLA-B27.4, Vega *et al.* (1986); HLA-B27f, Aparicio *et al.* (1987), Rojo *et al.* (1987); HLA-B40, Lopez de Castro *et al.* (1983); HLA-B40*, Ways *et al.* (1987); HLA-B44, Kottmann *et al.* (1986); HLA-Bw47, Zemmour *et al.* (1988); HLA-Bw58, Ways *et al.* (1985); HLA-Bw60, Ways *et al.* (1987).

Table I. Substitutions in *A*-locus sequences

	ChLA-A108	ChLA-A126	HLA-A11E	HLA-A3.1	HLA-A3.2	HLA-A2.1	HLA-Aw24
							<i>Amino acids</i>
ChLA-A108		14	9	12	13	28	32
ChLA-A126	24		15	16	17	25	27
HLA-A11E	17	27		8	7	27	29
HLA-A3.1	25	27	14		2	24	29
HLA-A3.2	25	27	12	3		24	27
HLA-A2.1	49	47	45	40	41		30
HLA-Aw24	54	43	47	44	43	52	
	<i>Nucleotides</i>						

Number of substitutions between *HLA-A* and *ChLA-A* locus alleles of allomorphs. The lower left-hand part shows nucleotides, the upper right-hand part amino acid substitutions.

Table II. Corrected percentages of replacement and silent substitutions in the A3/A11 group

	ChLA-A108	ChLA-A126	HLA-A11E	HLA-A3.1	HLA-A3.2
					<i>Replacement substitutions</i>
ChLA-A108		1.94	1.38	2.00	2.14
ChLA-A126	3.97		2.21	2.36	2.50
HLA-A11E	2.23	4.19		1.17	1.04
HLA-A3.1	4.17	3.83	1.89		0.42
HLA-A3.2	3.71	3.67	1.46		
	<i>Silent substitutions</i>				

Corrected percentages of replacement (upper right-hand corner) and silent (lower left-hand corner) substitutions between pairs of *HLA-A* and *ChLA-A* alleles belonging to the A11/A3 group.

A)

	293			296			298							306
	Val	Leu	Ala	Val		Val	Val	Ile	Gly	Ala	Val	Val	Ala	Ala
ChLA-B2	GTC	CTA	GCA	GTT	...	GTG	GTC	ATC	GGA	GCT	GTG	GTC	GCT	GCT
		*	*	*	*	*	*	*	*	*	*	*	*	*
ChLA-B1	GTC	CTT	GTG	GTC	ACC	GTA	GCT	GTG	GTC	GCT	GTG	GTC	GCT	GCT
	Val	Leu	Val	Val	Thr	Val	Ala	Val	Val	Ala	Val	Val	Ala	Ala
	293													306

B)

	293				298					304				306		
ChLA-B2	GTC	CTA	GCA	GTT	GTG	GTC	ATC	GGA	GCT	GTG	GTC	...	...	...	GCT	GCT
							*	*								
ChLA-B1	GTC	CT.	...	..T	GTG	GTC	ACC	GTA	GCT	GTG	GTC	GCT	GTG	GTC	GCT	GCT
	293				295					301						306

Fig. 7. Two alignments of codons 293–306 of ChLA-B1 and ChLA-B2. The direct alignment of the two sequences is shown in (A). In (B) a gap of six nucleotides in codons 294–296 of ChLA-B1 and a gap of nine nucleotides between codons 304 and 305 of ChLA-B2 have been introduced. Asterisks indicate mismatches.

soaked twice in distilled water and once in $20 \times$ SSPE. Transfer and hybridization of RNA were as for DNA.

DNA sequencing

The deletion subcloning procedure of Frischauf *et al.* (1980) was applied using *Sa*I linkers to obtain a set of plasmids with ordered deletions in the insert. The chemical sequencing method of Maxam and Gilbert (1980) was used for both strands after endlabelling of the DNA at the unique *Sa*I site either at the 3' end with the Klenow fragment or at the 5' end with the T4 polynucleotide kinase. Ambiguities were ruled out by the chain termination method (Sanger *et al.*, 1977) after subcloning in M13mp18, using the Sequenase enzyme (UBS) in the presence of dITP.

Calculations of silent and replacement substitutions

The rate for silent and replacement substitutions was calculated as described by Perler *et al.* (1980) and Salser (1977). The correction for multiple hits was made assuming equal rates for transitions and transversions (Kimura and Ohta, 1972).

Acknowledgements

We thank Dr H.Jäckle for his gift of the vector NM 1149, Drs H.Orr and E.Weiss for kindly providing us with the locus-specific probes, Mr K.Lamberty for preparing the line drawings, Mr R.Lorenz for preparing the photographs, and Ms Lynne Yakes for editorial assistance. This work was supported in part by grant AI23667 from the National Institutes of Health, Bethesda, MD, USA.

References

- Albert, E.D., Baur, M.P. and Mayr, W.R. (eds) (1984) *Histocompatibility Testing*. Springer Verlag, Berlin, pp. 163–165.
- Aparicio, P., Rojo, S., Jaraquemada, D. and Lopez de Castro, J.A. (1987) *J. Immunol.*, **139**, 837–841.
- Balner, H., Gabb, B.W., D'Amaro, J., van Vreeswijk, W. and Visser, T.P. (1974) *Tissue Antigens*, **4**, 313–328.
- Balner, H., van Vreeswijk, W., Roger, J.H. and D'Amaro, J. (1978) *Tissue Antigens*, **12**, 1–18.
- Birnsteil, M.L., Busslinger, M. and Strub, K. (1985) *Cell*, **41**, 349–359.
- Biro, P.A., Pan, J., Sood, A.K., Kole, R., Reddy, V.B. and Weissman, S.M. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 1079–1086.
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley, D.C. (1987a) *Nature*, **329**, 506–512.
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley, D.C. (1987b) *Nature*, **329**, 512–518.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294–5299.
- Cowan, E.P., Jordan, B.R. and Coligan, J.E. (1985) *J. Immunol.*, **135**, 2835–2841.
- Cowan, E.P., Jelachich, M.L., Biddison, W.E. and Coligan, J.E. (1987) *Immunogenetics*, **25**, 241–250.
- Ezquerria, A., Bragado, R., Vega, M.A., Strominger, J.L., Woody, J.A. and Lopez de Castro, J.A. (1985) *Biochemistry*, **24**, 1733–1741.
- Ezquerria, A., Domenech, N., van der Poel, J.J. and Lopez de Castro, J.A. (1986a) *Sixth International Congress of Immunology*, Abstract 2.31.4
- Ezquerria, A., Domenech, N., van der Poel, J.J., Strominger, J.L., Vega, M.A. and Lopez de Castro, J.A. (1986b) *J. Immunol.*, **137**, 1642–1649.
- Feinberg, A.P. and Vogelstein, B. (1984) *Analyt. Biochem.*, **137**, 266–267.
- Frischauf, A.M., Garoff, H., and Lehrach, H. (1980) *Nucleic Acids Res.*, **8**, 5541–5549.
- Geliebter, J. and Nathanson, S.G. (1987) *Trends Genet.*, **3**, 107–112.
- Glisin, V., Crkvenjakov, R. and Byus, C. (1974) *Biochemistry*, **13**, 2633–2637.
- Gubler, U. and Hoffman, B.J. (1983) *Gene*, **25**, 263–269.
- Hasegawa, M., Kishino, H. and Yano, T. (1987) *J. Mol. Evol.*, **26**, 132–147.
- Hayashida, H. and Miyata, T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2671–2675.
- Holmes, N. and Parham, P. (1985) *EMBO J.*, **4**, 2849–2854.
- Holmes, N., Ennis, P., Wan, A.M., Denney, D.W. and Parham, P. (1987) *J. Immunol.*, **139**, 936–941.
- Jaulin, C., Perrin, A., Abastado, J.-P., Dumas, B., Papamatheakis, J. and Kourilsky, P. (1985) *Immunogenetics*, **22**, 453–470.
- Kimura, M. and Ohta, T. (1972) *J. Mol. Evol.*, **2**, 87–90.
- Klein, J. (1980) In Fougereau, M. and Dausset, J. (eds), *Immunology 80*. Academic Press, London, pp. 239–253.
- Klein, J. (1986) *Natural History of the Major Histocompatibility Complex*. John Wiley and Sons, New York.
- Klein, J. (1987) *Human Immunol.*, **19**, 155–162.
- Klein, J. and Figueroa, F. (1986) *CRC Crit. Rev. Immunol.*, **6**, 295–386.
- Klein, J., Kasahara, M., Mayer, W.E. and Figueroa, F. (1988) *Science*, in press.
- Koller, B.J., Sidwell, B., DeMars, R. and Orr, H.T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5175–5178.
- Koller, B.H. and Orr, H.T. (1985) *J. Immunol.*, **134**, 2727–2733.
- Kottmann, A.H., Seemann, G.H.A., Guessow, H.D. and Roos, M.H. (1986) *Immunogenetics*, **23**, 396–400.
- Krangel, M.S., Taketani, S., Biddison, W.E., Strong, D.M. and Strominger, J.L. (1982) *Biochemistry*, **21**, 6313–6321.
- Krangel, M.S., Biddison, W.E. and Strominger, J.L. (1983) *J. Immunol.*, **130**, 1856–1862.
- Lalanne, J.L., Bregegere, F., Delarbre, C., Abastado, J.P., Gachelin, G. and Kourilsky, P. (1982) *Nucleic Acids Res.*, **10**, 1039–1049.
- Li, W.-H. and Tanimura, M. (1987) *Nature*, **326**, 93–96.
- Li, W.-H., Tanimura, M. and Sharp, P.M. (1987) *J. Mol. Evol.*, **25**, 330–342.
- Lopez de Castro, J.A., Strominger, J.L., Strong, D.M. and Orr, H.T. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3813–3817.
- Lopez de Castro, J.A., Bragado, R., Strong, D.M. and Strominger, J.L. (1983) *Biochemistry*, **22**, 3961–3969.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Mattson, D.H., Handy, D.E., Bradley, D.A., Coligan, J.E., Cowan, E.P. and Biddison, W.E. (1987) *Immunogenetics*, **26**, 190–192.
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- McCune, J.M., Humphreys, R.E., Yocum, R.R. and Strominger, J.L. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3206–3209.
- Miyada, C.G., Klofeld, C., Reyes, A.A., McLaughlin-Taylor, E. and Wallace, R.B. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2890–2894.
- Miyata, T., Yasunaga, T. and Nishida, T. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7328–7332.
- Neeffes, J.J., Briur-Vriefendorp, B.F., van Seventer, G.A., Ivanyi, P. and Ploegh, H.L. (1986) *Human Immunol.*, **16**, 169–181.
- N'Guyen, C., Sodoyer, R., Trucy, J., Strachan, T. and Jordan, B.R. (1985) *Immunogenetics*, **21**, 479–489.
- Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980) *Cell*, **20**, 555–566.
- Rojo, S., Aparicio, P., Choo, S.Y., Hansen, J.A. and Lopez de Castro, J.A. (1987) *J. Immunol.*, **139**, 831–836.
- Sakoyama, Y., Hong, K.J., Byun, S.M., Hisajima, H., Ueda, S., Yaoita, Y., Hayashida, H., Miyata, T. and Honjo, J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1080–1084.
- Salser, W. (1977) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 985–1002.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Sibley, C.G. and Ahlquist, J.E. (1987) *J. Mol. Evol.*, **26**, 99–121.
- Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
- Steinmetz, M. and Hood, L. (1983) *Science*, **222**, 727–733.
- Strachan, T., Sodoyer, R., Damotte, M. and Jordan, B.R. (1984) *EMBO J.*, **3**, 887–894.
- Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol.*, **31**, 77–84.
- Szöts, H., Riethmüller, G., Weiss, E. and Meo, T. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1428–1432.
- Vega, M.A., Ezquerria, A., Rojo, S., Aparicio, P., Bragado, R. and Lopez de Castro, J.A. (1985a) *Proc. Natl. Acad. Sci. USA*, **82**, 7394–7398.
- Vega, M.A., Wallace, L., Rojo, S., Bragado, R., Aparicio, P. and Lopez de Castro, J.A. (1985b) *J. Immunol.*, **135**, 3323–3332.
- Vega, M.A., Bragado, R., Ivanyi, P., Pelaez, J.L. and Lopez de Castro, J.A. (1986) *J. Immunol.*, **137**, 3557–3565.
- Wan, A.M., Ennis, P., Parham, P. and Holmes, N. (1986) *J. Immunol.*, **137**, 3671–3674.
- Ways, J.P., Coppin, H.L. and Parham, P. (1985) *J. Biol. Chem.*, **260**, 11924–11933.
- Ways, J.P., Lawlor, D.A., Wan, A.M. and Parham, P. (1987) *Immunogenetics*, **25**, 323–328.
- Weiss, E.H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. and Flavell, R.A. (1983) *Nature*, **301**, 671–674.
- Weiss, E.H., Kuon, W., Dörner, C., Lang, M. and Riethmüller, G. (1985) *Immunobiology*, **170**, 367–380.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, **33**, 103–119.
- Zemmour, J., Ennis, P.D., Parham, P. and Dupont, B. (1988) *Immunogenetics*, **27**, 281–287.

Received on April 28, 1988; revised on June 10, 1988