Immunogenetics 23: 172–180, 1986

Immunogenetics

An Isolated β_1 Exon Next to the DR_α Gene in the *HLA-D* Region

Helene F. Meunier*, Susan Carson†, Walter F. Bodmer, and John Trowsdale

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom

Abstract. A cosmid clone containing the DR_α gene and a β_1 exon of a DR_β-related gene was isolated from a human cosmid clone bank made from the consanguineous DR7 cell line MANN. No other DR_β-related exons were found on this clone. The β_1 exon was located about 15 kb away from the DR_α gene in a tail-to-tail (3′ to 3′) orientation. The exon contained several deleterious mutations: a defective splice site at the 5′ end, two translational frame shifts (a 1 bp deletion and a 1 bp insertion), and three extra cysteine residues. Nucleotide and amino acid sequence comparisons of the β_1 exon indicated that although it is substantially different from other class II β-chain genes, it is slightly more related to DR than to any other class II gene. The DR_β-related sequence was on a DNA fragment which showed no polymorphism on a panel of cell lines with Eco RI or Pst I. These Southern blots, however, revealed a related, polymorphic sequence in the human genome. Nucleotide sequences in the intron flanking the β_1 exon shared greater sequence homology than the β_1 exon itself when compared with the DR_β genomic sequence. The exon may play a role in the generation of variation in expressed class II β-chain genes and it may be a relic of a different subset of class II products.

Introduction

The class II *HLA-D*-region genes encode a series of cell-surface determinants involved in the immune response. These polymorphic surface glycoproteins are composed of α and β chains with molecular masses of approximately 34 000 and 28 000, respectively (Snary et al. 1976, Springer et al. 1977), which include two extracellular domains (α1, α2 and β1, β2), a transmembrane region, and a cytoplasmic tail (Shackelford et al. 1982, Larhammar et al. 1983).

The three most clearly established regions of class II antigens are called DP, DQ, and DR (Bodmer and Bodmer 1984). So far, there is evidence of at least six α-chain genes and more than seven β-chain genes (Trowsdale et al. 1985, Korman et al. 1985), most of which have already been attributed to the *HLA-D* subregions. The characteristic polymorphism observed in the *HLA-D* region is carried mainly by the β chains, although the DQ_α chain also shows considerable allelic variation (Auffray et al. 1983, Chang et al. 1983, Spielman et al. 1984, Trowsdale et al. 1983a). Lately, much effort has been dedicated to a better understanding of the molecular organization of the *HLA-DR* subregion (Spies et al. 1985, Larhammar et al. 1985). From these studies, as well as from restriction polymorphism analysis of serologically defined DR cell lines (Spielman et al. 1984, Trowsdale et al. 1983a, 1984, So et al. 1984), it is clear that the *HLA-DR* subregion includes more than one functional β-chain gene, and that their number may vary in different haplotypes.

In this paper, we describe a β_1 exon physically linked to DR_α which was isolated on a cosmid clone from a genomic library prepared from a homozygous DR7 cell line and we discuss the relationship of this pseudogene sequence to other *HLA-D*-region β-chain genes.

Materials and Methods

Isolation and characterization of a genomic clone. A genomic library was constructed from DNA prepared from the homozygous, consanguineous B-lymphoblastoid cell line MANN (A29, B12, Bw4, Cw4, DR7, DPw2) which was partially digested with Mbo I and then ligated to the cosmid vector pTCF (Grosveld et al. 1982). The library was screened with a DR_α cDNA probe labeled with [α^{32}P] deoxycytidine triphosphate by nick translation (Rigby et al. 1977). Hybridization was carried out in 6 × NaCl/cit (1 × NaCl/cit is 0.15 *M*

* *Present address:* The Salk Institute, San Diego, California 92138
† *Present address:* Basel Institute for Immunology, Basel, Switzerland

NaCl/0.015 *M* sodium citrate)/10 × Denhardt's solution (1 × Denhardt's solution is 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/10% dextran sulfate/0.1% sodium dodecyl sulfate (SDS)/denaturated salmon sperm DNA (100 μg/ml) for 18 h at 65 °C. Filters were washed twice in 6 × NaCl/cit/0.1% SDS for 1 h and exposed for 2 days to Kodak XAR-5 film. One DR_α-containing clone, M4.2, was identified and characterized by restriction site mapping and Southern blot hybridization.

cDNA and genomic probes. The DR_α cDNA probe used for screening and restriction mapping consisted of two Pst I inserts (A and B) from clone pDRH2 (Lee et al. 1982). The 900 bp Pst I fragment (11–13) derived from cosmid LC11 (Trowsdale et al. 1983b) comprises part of the β_1 exon and the 5′ end intron sequence of the $DP_{\beta 1}$ gene. Pst I/380 and Pst I/530 are two DR_β-specific fragments generated by Pst I digestion of the cDNA clone DRβ-II (Long et al. 1983). The first fragment contains the signal sequence and β_1 exon, while the second includes the β_2 exon, transmembrane and cytoplasmic regions, and part of the 3′ untranslated region. A 0.95 kb Hind III/Pst I subfragment was generated by Hind III digestion of a 1.6 kb Pst I fragment isolated from cosmid M4.2. This subfragment includes most of the M4.2 β_1 exon and approximately 750 bp of the 5′ end intron sequence (see Fig. 1B).

DNA sequence determination. A 1.6 kb Pst I fragment from cosmid M4.2 which cross-hybridizes with the $DP_{\beta 1}$ probe 11–13 was subcloned in plasmid pAT153 and the recombinant plasmid grown in the DH1 strain of *Escherichia coli.* After self ligation, 10 μg of DNA prepared from this recombinant plasmid was sonicated, end-repaired and ligated into Sma I-cut vector M13mp8 (Vieira and Messing 1982). Random clones of approximately 300 bp were then sequenced using the dideoxy chain termination method (Sanger et al. 1980, Bankier and Barrell 1983) and sequences were aligned by computer as described previously (Kelly and Trowsdale 1985). To complete the sequence of the β_1 exon, a 2.3 kb Hind III fragment was subcloned in plasmid pAT153 to isolate the 1.3 kb Hind III/Pst I subfragment contiguous to the 3′ end of the 1.6 kb Pst I fragment (see Fig. 1B). This subfragment was ligated into Hind III/Pst I-cut M13mp9, and the first 240 nucleotide sequence was obtained by the chain termination method.

Northern analysis. Northern blot analysis of poly(A)$^+$-selected mRNA was performed as described by Thomas (1983).

Southern blot analysis of genomic DNA. The high molecular mass DNA derived from eight homozygous and three heterozygous EBV-transformed cell lines was prepared as described previously (So et al. 1984). For details of the cell lines and their HLA types, see Trowsdale

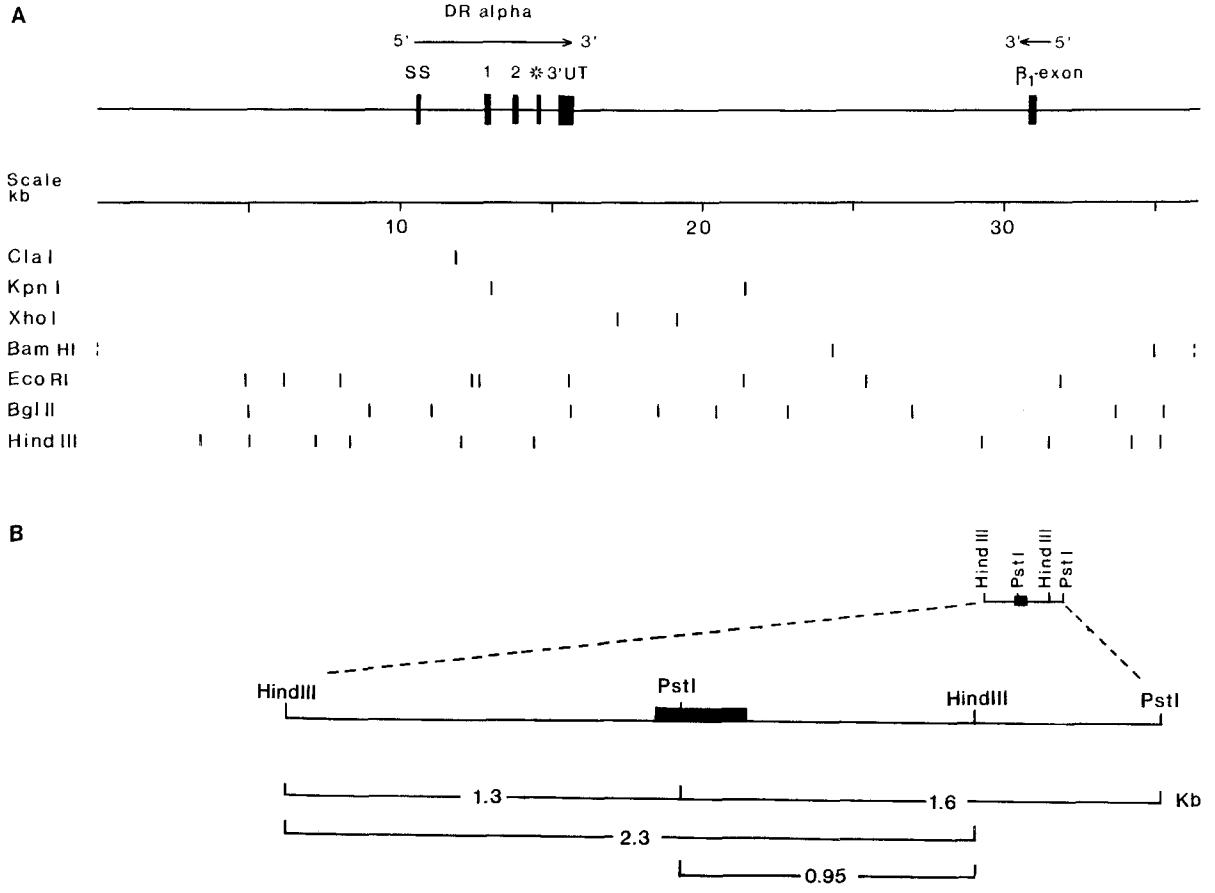

Fig. 1 A and B. A Restriction map of cosmid M4.2 containing the DR_α gene and a β_1 exon. Exons are shown as *filled boxes. Arrows* represent the direction of transcription (5′–3′) for the DR_α gene and the orientation of the β_1 exon obtained by random sequencing of a Pst I fragment (see *Materials and Methods* and the legend to Fig. 2 for further details). Exons are denoted as follows: SS, signal sequence; 1, first domain; 2, second domain; *, transmembrane and cytoplasmic domains; and 3′ UT, 3′ untranslated region. The two Bam HI sites (re)generated at each end of the insert are depicted by a *broken line.* **B** Illustration of the restriction sites which were used in the sequencing strategy of the β_1 exon. The β_1 exon sequence was derived from overlapping M13 clones generated by sonication of a 1.6 kb Pst I fragment and from the Pst I site of a 1.3 kb Hind III/Pst I fragment subcloned in M13 (see *Materials and Methods*). This map agrees with that derived by Spies and co-workers (1985)

and co-workers (1985). DNA was digested with Pst I or Eco RI and transferred to nitrocellulose filters. Filters were hybridized with a 0.95 kb Hind III/Pst I fragment from cosmid M4.2, as described above, for 18 h at 68 °C and washed in 0.1 × NaCl/cit/0.1% SDS at 65 °C.

Results

Isolation of a cosmid containing a DR_α gene and a β_1 exon. The cosmid library, prepared from the homozygous DR7 cell line MANN, was screened with a DR_α cDNA probe under low stringency conditions as already described in *Materials and Methods*. One DR_α clone, M4.2, was isolated after resubmitting the colonies to stringent hybridization conditions. The restriction map of cosmid M4.2 was determined using a combination of restriction enzymes and blot hybridization (Fig. 1A). We had previously shown that a probe containing the β_1 exon of the DP_β gene hybridized to multiple *HLA* class II β-chain genes under nonstringent conditions (Trowsdale et al. 1983b). Under these low stringency conditions, the $DP_{\beta 1}$ probe 11–13 hybridized to the following restriction fragments on cosmid M4.2: Eco RI, 6.3 kb; Pst I, 1.6 kb; and Hind III, 2.3 kb. However, when the hybridization was carried out under stringent conditions, no cross-hybridization could be detected with the DP_β probe.

In order to obtain more information about the nature of this β-like sequence, two DR_β-specific fragments derived from a cDNA clone (Long et al. 1983), Pst I/380 and Pst I/530 (see *Materials and Methods*), were used to probe restriction digests from cosmid M4.2. After repeated low salt washes (0.3 × NaCl/cit), one strongly hybridizing band was observed for each restriction enzyme digest hybridized with Pst I/380. Pst I/530 did not detect any restriction fragments, even when the stringency of hybridization conditions was decreased (data not shown). Under relaxed hybridization conditions, Pst I/380 did not detect any additional band which might have suggested the presence of a DR_β-related signal sequence. These results showed that the β_2 and signal sequence exons may either be missing or may be a large distance away from the β_1 exon.

To confirm the authenticity of our genomic clone, restriction enzyme digests of MANN genomic DNA were probed with a Hind III/Pst I 0.95 kb fragment which contains most of the β_1 exon and some of the 5′ end intron sequence (see Fig. 1B). The following bands were detected on Southern blots: Bam HI 10.5 kb, Eco RI 6.3 kb, and Pst I 1.6 kb, all of which agree with the restriction map of cosmid M4.2 described in Figure 1A. This map agrees with the data from Spies and coworkers (1985) concerning the *DR* region from a *DR4*-containing haplotype.

Nucleotide and predicted amino acid sequences of the β_1 exon. The nucleotide sequence and translated amino acid sequence of the β_1 exon found in cosmid M4.2 is shown in Figure 2. Two translational frame shifts were introduced in order to obtain alignment with some already published DR_β sequences. These were apparently due to a 1 bp deletion and insertion, 86 nucleotides apart at positions 1155 and 1243, respectively. Although no significant match with other known protein sequences could be observed in the absence of such a modification, the succession of nucleic acids remains in an uninterrupted reading frame for 639 bp, which is in frame with the DR_β-related sequence for only part of its length. An apparently functional splice site is present at the 3′ end of the β_1 exon, whereas the expected 'AG' dinucleotide (Breathnach and Chambon 1981) at the intron-exon junction at the 5′ end of the β_1 exon has been changed to 'GG', making it unlikely that it is used as a functional splice site.

One interesting feature of this β_1 exon is the presence of three extra cysteine residues at positions 13, 56, and 65. Free cysteine residues have been described in the first domain of mouse E^b, E^k β-chain genes (Mengle-Gaw and McDevitt 1983, Widera and Flavell 1984) and in the first domain of the DR_β gene conferring DRw53 specificity (Spies et al. 1985). These extra cysteine residues were found at positions 12 and 13, respectively. On the other hand, free cysteine residues are not generally present in the first domain of the DR_β gene conferring the main DR serological specificities.

Northern blots. In an attempt to determine whether the β_1 exon is transcribed in this lymphoblastoid cell line despite the above-mentioned abnormalities, 1 µg of MANN poly(A)$^+$-selected mRNA was transferred to nitrocellulose filters and hybridized with the 0.95 kb Hind III/Pst I β_1-exon probe derived from cosmid M4.2. No signal was detected on a northern blot with this β_1-exon probe. However, the same filter gave a clear 1.3 kb band when hybridized with a $DP_{\beta 1}$-exon probe (11–13) (data not shown). Thus, it appears that this β_1 exon is either not transcribed, or is transcribed at very low levels. The same would be true for any β-chain gene highly related to the exon.

Comparison of nucleotide and amino acid sequences with other HLA-D β-chain genes and pseudogenes. We have compared the nucleotide and amino acid sequences of the β_1 exon from cosmid M4.2 with some already known HLA-D β-chain sequences. The nucleic acid sequence of the β_1 exon from cosmid M4.2 appears to be most similar to expressed DR_β-chain genes with which 71–76% of nucleotides are shared (Table 1). It also shares 70% homology with the β_1 exon of the DR_β pseudogene (Larhammar et al. 1985). Similarly, when

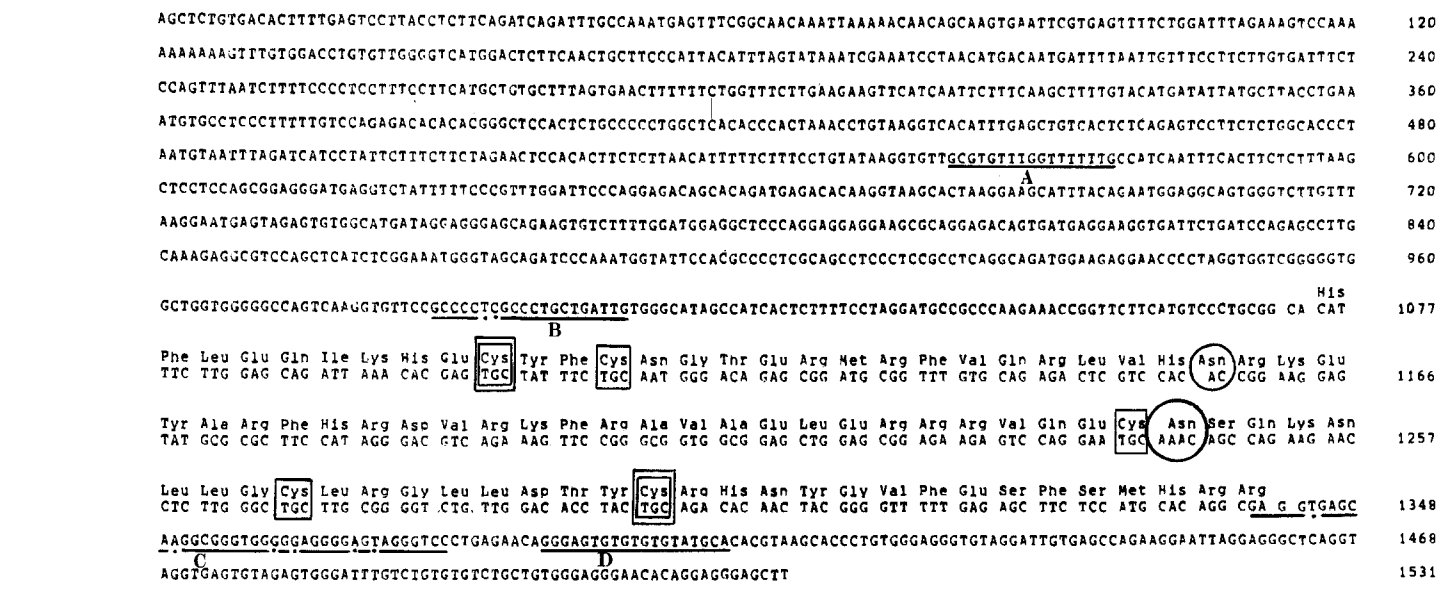

Fig. 2. The complete nucleotide sequence of the β_1 exon isolated from cosmid M4.2. The nucleotide sequence was determined as explained in the legend to Figure 1. The derived amino acid sequence shown above the nucleotide sequence was obtained after the introduction of a 1 bp insertion and a 1 bp deletion, as shown by the two *circled codons*. The extra cysteine residues are indicated by a *single box*, and conserved cysteine residues by a *double box*. The positions of sequences matching other known nucleotide sequences are indicated by *solid lines*. *Dots* indicate unmatched nucleotides. Two 18 bp sequences, A and D, were respectively identical to a short sequence found upstream of the *V. cholera* toxin operon gene (Mekalanos et al. 1983) and to a sequence found in the first intron of the human haptoglobin gene (Maeda et al. 1984). Closer to the β_1 exon, sequence B shares 18 bp with a nucleotide sequence derived from the 5′ end of the kappa immunoglobulin gene (Klobeck et al. 1984). This nucleotide sequence from the kappa immunoglobulin gene contains a putative regulatory element which is found in the promoter region of most class II genes (Das et al. 1983, Hyldig-Nielsen et al. 1983, Malissen et al. 1983, Falkner and Zachau 1984, Gillies et al. 1984, Parslow et al. 1984, Klobeck et al. 1984, Kelly and Trowsdale 1985, Larhammar et al. 1985). At the end of the β_1 exon, sequence C shares 32 bp with a complementary sequence of the mouse H-2K^b first intron (Weiss et al. 1983). Finally, sequence D shares 27 bp with the poliovirus type 3 p3–2b coding sequence (Stanway et al. 1983). The significance of these conserved sequences is not known

the amino acid sequence of the β_1 exon is compared with that of expressed DR_β genes, the homology varies from 52% to 58% (DR$_\beta$ pseudogene; 52%). These values are much lower than those observed when DR_β alleles are compared to each other. In this case, the homology is about 95% and 82–90% for nucleotide and amino acid sequences, respectively. The mouse equivalents of DQ and DR, A and E, share 68–69% sequence homology with the β_1 exon at their nucleotide level and only 46–48% at their amino acid level. Other class II β-chain genes, DP, DX, and DQ, are approximately 62–65% related to the β_1-exon nucleic acid and 45–48% to the amino acid sequences. These results show that although the β_1 exon from cosmid M4.2 is substantially different from published DR$_\beta$ sequences, it is slightly more closely related to DR_β-chain genes than it is to any of the known *HLA-D* β-chain genes. However, the similarity to DR$_\beta$ may not be particularly significant, since comparisons of β_1 domains of DR and DQ or DP and DQ indicate that, on a percentage basis, they are as related to each other as the M4.2 β_1 domain is to DR$_{\beta 1}$ (data not shown). Table 2 gives a breakdown of the relationship between the β_1 exon and other DR$_\beta$ sequences in terms of replacements and silent substitutions. There is a marked increase in both of these parameters when compared to the difference between two DR_β alleles, or between DR$_\beta$ and its pseudogene (Larhammar et al. 1985). This indicates considerable evolutionary divergence and thus a relatively long time of separation, on a par with that for DP, DQ, or DR with each other. However, there are more potential silent substitutions, which argues for the fact that much of the divergence was, at one time at least, of functional relevance.

The complete nucleotide sequence shown in Figure 2 was also compared with the genomic sequence of the DR$_\beta$ pseudogene (Larhammar et al. 1985), since this is the only DR$_\beta$ genomic sequence published so far (Fig. 3). Interestingly, the nucleotide homology extended to the flanking intron sequences on both sides of the β_1 exon. With the exception of one 40 bp insertion and two smaller deletions, significant matches were observed for most of the intron sequences. In a more detailed amino acid analysis, we looked at the distribution pattern of the predicted sequences for various *HLA-D* and mouse *A/E* β-chain genes. It is clear from Figure 4 that the β_1 exon isolated from the DR7 cell line has retained some of the primary structural elements which are character-

Table 1. Percent homology observed between the nucleotide and amino acid sequences of β_1 exons from different class II β-chain genes

M4.2 β_1 exon homology with	Nucleotides*	Amino acids*
DR4	76	57
DR3, 6a	76	57
DR4, 6	75	58
DR3, 6b	75	55
DR1	73	58
DR2, 2	–	55
DRw53	71	52
ψDR4	70	52
DQ3	65	48
DQ4	63	45
DX	65	47
$DPw2^a$	62	45
$DPw2^b$	62	45
E^d	69	46
A^b	68	48

Sequences were obtained from the following sources: DR4 and DRw53 (Spies et al. 1985), DR4, 6 (Long et al. 1983), DR3, 6a and b (Gustafsson et al. 1984), DR1 (Bell et al. 1985), DR2, 2 (Kratzin et al. 1981), ψDR4 (Larhammar et al. 1985), DQ3 (Boss and Strominger 1984), DQ4 (Larhammar et al. 1983a), DX (Okada et al. 1985), $DPw2^a$ (Kelly and Trowsdale 1985), $DPw2^b$ (Kappes et al. 1984), E^d (Saito et al. 1983), and A^b (Larhammar et al. 1983b)

* In these calculations, extra nucleotides or amino acids were excluded at certain locations to permit correct alignment of sequences (see Fig. 4)

istic of the *HLA* class II β-chain genes. The two largest conserved amino acid regions were located near the cysteine residues which are highly characteristic of expressed class II β-chain genes (Kaufman and Strominger 1982). Among other conserved stretches of amino acids in the β_1 exon isolated from cosmid M4.2 were two short regions, mainly specific for DR_β genes located at each extremity of the β_1 exon. It becomes more apparent from this comparison that the β_1 exon described above has diverged considerably from other class II β-chain genes, notably in areas different from those generally described as hypervariable regions. This is suggested mainly by the replacement of certain amino acid sequences which are conserved in HLA-DR, DQ, DP and, to a lesser extent, mouse A/E, as observed at positions 41, 42, 45, 46, 54, 56–61, and 69. Interestingly, the amino acids at positions 16 and 35 correspond to those found in DP and DQ rather than DR, while differences in positions 90–92 are in a region that appears to be subregion-specific.

Southern blot analysis of genomic DNAs. In an attempt to look for related sequences, and to determine whether the isolated β_1 exon was present only in the DR7 cell line MANN or was found in cell lines of other DR types, we submitted genomic DNA isolated from a panel of cell lines of different specificities to Southern blot analysis using a probe specific for the β_1 exon from cosmid 4.2 (Fig. 5A). After digestion of genomic DNAs with Pst I, a 1.6 kb band hybridizing strongly with the Hind III/Pst I probe was observed in each DR-typed cell line tested.

In addition to the 1.6 kb Pst I fragment, other bands ranging in size from 2.3–5 kb also hybridized to the Hind III/Pst I probe. We have independently identified which restriction fragments on these blots correspond to DP_β and DQ_β genes using appropriate probes (data not shown). Under the high stringency conditions in which these hybridizations were carried out, we saw no cross-reaction with these *HLA-DP* or *DQ* genes. Using the same filters and hybridization conditions, the DR_β-specific cDNA probe Pst I/380 detected restriction fragments which did not correspond to the genomic restriction fragments described in Figure 5. We suggest that the additional bands seen in Figure 5A correspond to copies of sequences highly related to the sequence included in the 0.95 kb Hind III/Pst I subfragment. They are in fact, more related than DR_β. This interpretation is supported by the presence of 1–3 extra bands on a Southern blot of genomic DNA digested with the restriction enzyme Eco RI (Fig. 5B) in addition to the constant 6.3 kb band. These related sequences could represent a previously unidentified polymorphic class II β-chain gene.

Discussion

The most striking feature of this isolated β_1 exon is its extent of divergence, not only from *DP* and DQ_β genes, but also from the known DR_β sequences. Although it resembles the latter slightly more than the former, the difference between the isolated β_1-exon sequence and available $DR_{\beta 1}$-exon sequences is almost as great as that between DR and DP, DR and DQ, or DP and DQ.

Table 2. Sequence divergence in the β_1 exon of DR_β genes and pseudogene(s)

	Percent potential silent changes	Percent potential replacement changes
M4.2 β_1 exon/ψDR4*	53.8	33.7
M4.2 β_1 exon/DR4	42.5	25.7
DR4/DR1	23.1	7.6
DR4/ψDR4*	22.8	11.4

Percents of silent changes and replacement changes were calculated according to Perler and co-workers (1980). Nucleotide sequences were obtained as follows: M4.2 β_1 exon (this paper, Fig. 2); ψDR4 (Larhammar et al. 1985); DR4 (Spies et al. 1985); DR1 (Bell et al. 1985)

* To obtain proper alignment of the sequences, the two missing codons in the ψDR4 β_1 exon were replaced by the nucleotide sequence CTG GAA, which is most commonly found in $DR_{\beta 1}$ genes

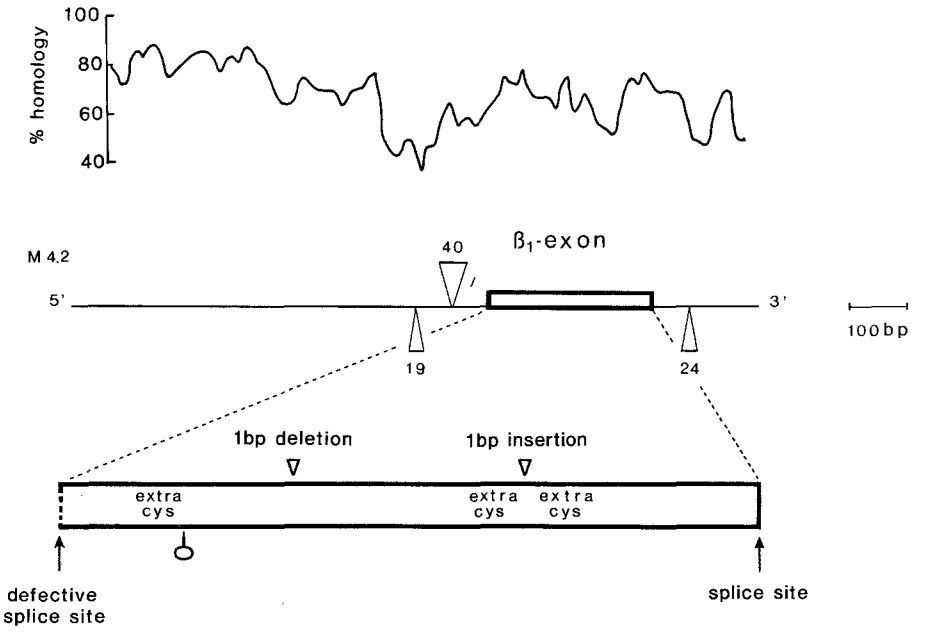

Fig. 3. Schematic representation of the β_1 exon isolated from cosmid M4.2. Above it is a comparison with the intron sequence from the DR_β pseudogene (Larhammar et al. 1985). The 5′ end of the DR_β pseudogene genomic sequence corresponds to the Hind III site described in the legend to Figure 2. One 40 bp insertion, as well as two deletions, 19 and 24 bp, are indicated on cosmid M4.2. The curved line at the top of the figure shows the distribution of homologous regions, as calculated every 10 nucleotides for 50 consecutive nucleotides. ♁ represents an intact putative glycosylation site

This suggests that the M4.2 exon sequence is very old, dating back to the time of separation of the major subregions of the *HLA-D* region. A second puzzle is the surprisingly high sequence homology of the flanking regions of this exon with DR_β intron sequences. Flanking and intron sequences between subregions generally differ much more than the corresponding exon sequences of the various subregions. The overall data, including the pattern of amino acid sequence differences and similarities of the β_1 exon with other β-exon amino acid sequences, suggest that it is derived from a different, as yet unknown, subregion, that may have a slightly more recent common origin with the DR_β than with the other subregions. The similarity of the flanking regions to DR_β intron sequences could have arisen if the exon as a whole was transposed to the *DR* subregion at some point in evolution after its divergence from DR. The surrounding sequences would then be related directly to DR_β. It is even possible that once a sequence has been inserted in the *DR* subregion, subsequent gene conversion-like events lead to a tendency to increased similarity with DR_β. In this case, the close relationship with DR_β may postdate rather than predate the transposition of the sequence to its present location. The troughs in the homology profile immediately flanking the β_1 exon may have resulted from genetic events involved in shifting the exon to its present location (Fig. 3). Clearly, it will be of interest to try and identify sequences corresponding to this new subregion elsewhere within HLA-D. Functional versions may correspond to the extra bands detected on Southern blots by the M4.2 sequence. Alternatively, it may be that the subregion now only exists in a relic form in the human genome, and in that case it will be interesting to try and identify its presence in other species and to trace its evolutionary history.

Does the isolated β_1 exon that we have identified have any functional relevance? It has been argued that

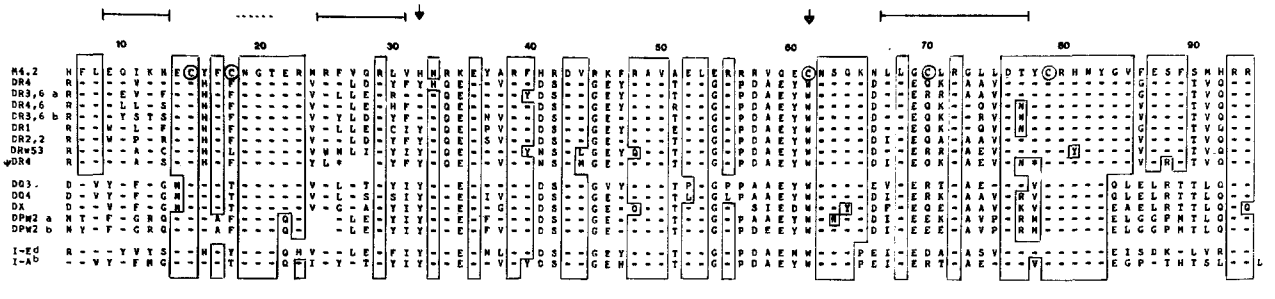

Fig. 4. Comparison of the β_1-exon amino acid sequences predicted for a series of *HLA-D* β-chain genes and mouse *A/E* β-chain genes. DR2, 2 is derived from a protein sequence. Sequences are from the same sources as in the footnote to Table 1. The *boxed areas* represent sequences of amino acids which are conserved in other class II β-chain genes. *Arrows* indicate the position of a 1 bp deletion and a 1 bp insertion. The *solid lines* represent the hypervariable regions in the β_1 exon of the *HLA-DR*$_\beta$-chain genes. The *dotted line* indicates the glycosylation site in the β_1 domains

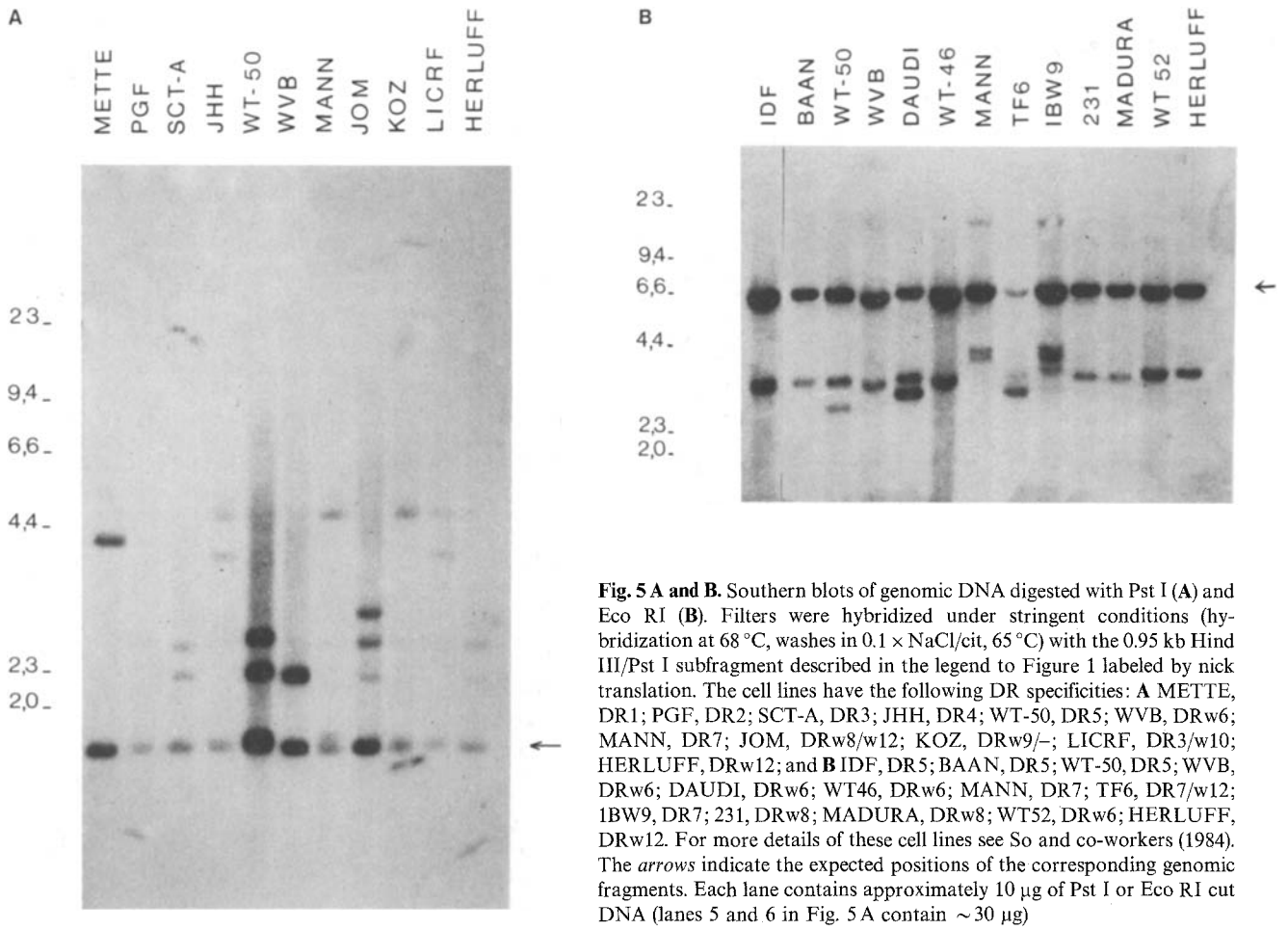

Fig. 5 A and B. Southern blots of genomic DNA digested with Pst I (**A**) and Eco RI (**B**). Filters were hybridized under stringent conditions (hybridization at 68 °C, washes in 0.1 × NaCl/cit, 65 °C) with the 0.95 kb Hind III/Pst I subfragment described in the legend to Figure 1 labeled by nick translation. The cell lines have the following DR specificities: **A** METTE, DR1; PGF, DR2; SCT-A, DR3; JHH, DR4; WT-50, DR5; WVB, DRw6; MANN, DR7; JOM, DRw8/w12; KOZ, DRw9/–; LICRF, DR3/w10; HERLUFF, DRw12; and **B** IDF, DR5; BAAN, DR5; WT-50, DR5; WVB, DRw6; DAUDI, DRw6; WT46, DRw6; MANN, DR7; TF6, DR7/w12; 1BW9, DR7; 231, DRw8; MADURA, DRw8; WT52, DRw6; HERLUFF, DRw12. For more details of these cell lines see So and co-workers (1984). The *arrows* indicate the expected positions of the corresponding genomic fragments. Each lane contains approximately 10 μg of Pst I or Eco RI cut DNA (lanes 5 and 6 in Fig. 5 A contain ~30 μg)

pseudogenes may be maintained as reservoirs of donor sequences for gene conversion or other related genetic mechanisms (Bodmer et al. 1984, Larhammar et al. 1985). Since the β_1 exon is by far the most polymorphic domain of the class II β-chain genes, and the rest of the sequence remains relatively constant, it is not necessary to maintain the whole pseudogene in order for it to act as an effective sequence donor source.

Several pieces of evidence indicate that the number of DR_β-related genes differs in different haplotypes (So et al. 1984). However, from blotting studies it appears that all the haplotypes possess the β_1 exon described here. The proximity to the relatively constant DR_α gene may explain the lack of associated restriction polymorphism, since polymorphic regions are associated with the highly polymorphic β_1 and $DQ_{\alpha 1}$ exons (see for example, Trowsdale et al. 1985).

The analysis of 5′ untranslated regions and intron sequences flanking the β_1 exon of expressed DR_β genes, as well as sequences found on the related polymorphic genomic fragments described in this paper, could provide some information with respect to the origin and possible function of this isolated β_1 exon.

Acknowledgments. We wish to thank the other members of the laboratory, in particular Messrs. Adrian Kelly, John Young, and David Wilkinson for helpful discussions. We are grateful to Dr. Alex So for kindly providing genomic DNA from serologically defined DR cell lines and for making the DR_β cDNA probes available to us. We also thank Ms. Susan Simmons for her help in preparing this manuscript. Helene Meunier is a recipient of a fellowship from the Medical Research Council of Canada.

References

Auffray, C., Ben-Nun, A., Roux-Dossetto, M., Germain, R. N., Seidman, J. G., and Strominger, J. L.: Polymorphism and complexity of the human DC and murine I-A α chain genes. *EMBO J. 2:* 121–124, 1983

Bankier, A. T. and Barrell, B. G.: Shotgun DNA sequencing *In* H. L. Kornberg (ed.): *Techniques in Life Sciences, Nucleic Acid Biochemistry*, pp. 1–34, Elsevier Scientific Publishers, County Clare, Ireland, 1983

Bell, J. I., Estess, P., St.-John, T., Saiki, R., Watling, D. L., Erlich, A. A., and McDevitt, H. O.: DNA sequence and characterization of human class II major histocompatibility complex β chains from the DR1 haplotype. *Proc. Natl. Acad. Sci. U.S.A. 82:* 3405–3409, 1985

Bodmer, J. and Bodmer, W. F.: Histocompatibility 1984. *Immunol. Today 5:* 251–254, 1984

Bodmer, W. F., Bodmer, J., Lee, J., Spielman, R., Travers, P., and Trowsdale, J.: Structure, evolution and polymorphism of the HLA-D region. *In* T. Sasazuki and T. Tada (eds.): *Immunogenetics – Its Applications to Clinical Medicine*, pp. 263–281, Academic Press, Tokyo, 1984

Boss, J. M. and Strominger, J. L.: Cloning and sequence analysis of a DC-3β gene. *Proc. Natl. Acad. Sci. U.S.A. 81:* 5199–5203, 1984

Breathnach, R. and Chambon, P.: Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem. 50:* 349–383, 1981

Chang, H. C., Moriuchi, T., and Silver, J.: The heavy chain of human β-cell alloantigen HLA-DS has a variable N-terminal region and a constant immunoglobulin-like region. *Nature 305:* 813–815, 1983

Das, H. K., Lawrance, S. K., and Weissman, S. M.: Structure and nucleotide sequence of the heavy chain gene of HLA-DR. *Proc. Natl. Acad. Sci. U.S.A. 80:* 3543–3547, 1983

Falkner, F. G. and Zachau, H. G.: Correct transcription of an immunoglobulin K gene required an upstream fragment containing conserved sequence elements. *Nature 310:* 71–74, 1984

Gillies, S. D., Folsom, V., and Tonegawa, S.: Cell specific enhancer element associated with a mouse MHC gene, Eβ. *Nature 310:* 594–597, 1984

Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M., and Flavell, R. A.: The construction of cosmid libraries which can be used to transform eukaryotic cells. *Nucleic Acids Res. 10:* 6715–6732, 1982

Gustafsson, K., Wiman, K., Emmoth, E., Larhammar, D., Böhme, J., Hyldig-Nielsen, J. J., Ronne, H., Peterson, P. A., and Rask, L.: Mutations and selection in the generation of class II histocompatibility antigen polymorphism. *EMBO J. 3:* 1655–1661, 1984

Hyldig-Nielsen, J. J., Schenning, L., Hammerling, U., Widmark, E., Heldin, E., Lind, P., Servinius, B., Lund, T., Flavell, R., Lee, J. S., Trowsdale, J., Schreier, P. H., Zablitzky, F., Larhammar, D., Peterson, P. A., and Rask, L.: The complete nucleotide sequence of the I-Eα^d immune response gene. *Nucleic Acids Res. 11:* 5055–5071, 1983

Kappes, D. J., Arnot, D., Okada, K., and Strominger, J. L.: Structure and polymorphism of the HLA class II SB light chain genes. *EMBO J. 3:* 2985–2993, 1984

Kaufman, J. F. and Strominger, J. L.: HLA-DR light chain has a polymorphic N-terminal region and a conserved Ig-like C-terminal region. *Nature 297:* 694–697, 1982

Kelly, A. and Trowsdale, J.: Complete nucleotide sequence of a functional HLA-DP$_\beta$ gene and the region between the DP$_{\beta 1}$ and DP$_{\alpha 1}$ genes: Comparison of the 5′ ends of HLA class II genes. *Nucleic Acids Res. 13:* 1607–1621, 1985

Klobeck, H. G., Combriato, G., and Zachau, H. G.: Immunoglobulin genes of the kappa light chain type from two human lymphoid cell lines are closely related. *Nucleic Acids Res. 12:* 6995–7006, 1984

Korman, A. J., Boss, J. M., Spies, T., Sorrentino, R., Okada, K., and Strominger, J. L.: Genetic complexity and expression of human class II histocompatibility antigens. *Immunol. Rev. 85:* 45–86, 1985

Kratzin, H., Yang, C., Götz, H., Pauly, E., Köbel, S., Egert, G., Thinnes, F., Wernet, P., Altevogt, P., and Hilschmann, N.: Primärstruktur menschlicher Histokompatibilitätsantigen der Klasse II. *Hoppe-Seylers Z. Physiol. Chem. 362:* 1665–1669, 1981

Larhammar, D., Hyldig-Nielsen, J. J., Servenius, B., Andersson, G., Rask, L., and Peterson, P. A.: Exon-intron organization and complete nucleotide sequence of a human major histocompatibility antigen DCβ gene. *Proc. Natl. Acad. Sci. U.S.A. 80:* 7313–7317, 1983a

Larhammar, D., Hammerling, U., Denaro, M., Lund, T., Flavell, R. A., Rask, L., and Peterson, P. A.: Structure of the murine immune response I-A$_\beta$ locus: Sequence of the I-A$_\beta$ gene and an adjacent β-chain second domain exon. *Cell 34:* 179–188, 1983b

Larhammar, D., Servenius, B., Rask, L., and Peterson, P. A.: Characterization of an HLA-DRβ pseudogene. *Proc. Natl. Acad. Sci. U.S.A. 82:* 1475–1479, 1985

Lee, J. S., Trowsdale, J., Travers, P., Carey, J., Grosveld, F., Jenkins, J., and Bodmer, W. F.: Sequence of an HLA-DR α-chain cDNA clone and intron-exon organization of the corresponding gene. *Nature 299:* 750–752, 1982

Long, E., Wake, C. T., Gorski, J., and Mach, B.: Complete sequence of an HLA-DRβ chain deduced from a cDNA clone and identification of multiple non allelic DRβ chain genes. *EMBO J. 2:* 389–394, 1983

Maeda, N., Yang, F., Barnett, D. R., Bowman, B. H., Smithies, O.: Duplication within the haptoglobin Hp^2 gene. *Nature 309:* 131–135, 1984

Malissen, M., Hunkapiller, T., and Hood, L.: Nucleotide sequence of a light chain gene of the mouse I-A subregion Aβ^d. *Science 221:* 750–754, 1983

Mekalanos, J. J., Swartz, D. J., Pearson, G. D. N., Harford, N., Groyne, F., and de Wilde, M.: Cholera toxin genes: Nucleotide sequence deletion analysis and vaccine development. *Nature 306:* 551–557, 1983

Mengle-Gaw, L. and McDevitt, H. O.: Isolation and characterization of a cDNA clone for the murine I-Eβ polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A. 80:* 7621–7625, 1983

Okada, K., Boss, J. M., Prentice, H., Spies, T., Mengler, R., Auffray, C., Lillie, J., Grossberger, D., and Strominger, J.: Gene organization of DC and DX subregions of the human major histocompatibility complex. *Proc. Natl. Acad. Sci. U.S.A. 82:* 3410–3414, 1985

Parslow, T. G., Blair, D. L., Murphy, W. J., and Granner, D. K.: Structure of the 5′ ends of immunoglobulin genes: A novel conserved sequence. *Proc. Natl. Acad. Sci. U.S.A. 81:* 2650–2654, 1984

Perler, F., Efstratiadis, A., Lomedico, P., Giblert, W., Kolodner, S., and Dodgson, J.: The evolution of genes: The chicken preproinsulin gene. *Cell 20:* 555–566, 1980

Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P.: Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymelase I. *J. Mol. Biol. 113:* 237–251, 1977

Saito, H., Maki, R. A., Clayton, L. K., and Tonegawa, S.: Complete primary structure of the Eβ chain and gene of the mouse major histocompatibility complex. *Proc. Natl. Acad. Sci. U.S.A. 80:* 5520–5524, 1983

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A.: Cloning in single stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol. 143:* 161–178, 1980

Shackelford, D. A., Kaufman, J. F., Korman, A. J., and Strominger, J. L.: HLA-DR antigens: Structure, separation of subpopulations, gene cloning and function. *Immunol. Rev. 66:* 133–187, 1982

Snary, D., Barnstable, C., Bodmer, W. F., Goodfellow, P., and Crumpton, M. J.: Human Ia antigen purification and molecular structure. *Cold Spring Harbor Symp. Quant. Biol. 41:* 379–386, 1976

So, A. K. L., Trowsdale, J., Bodmer, J. G., and Bodmer, W. F.: DNA polymorphism of HLA-DRβ and HLA-DPβ genes. *In* E. D. Albert, M. P. Baur, and W. R. Mayr (eds.): *Histocompatibility Testing 1984*, pp. 565–568, Springer-Verlag, Berlin, 1984

Spielman, R. S., Lee, J. S., Bodmer, W. F., Bodmer, J. G., and Trowsdale, J.: Six HLA-Dα chain genes on human chromosome 6: Polymorphism and association of DCα-related sequence with DR types. *Proc. Natl. Acad. Sci. U.S.A. 81:* 3461–3465, 1984

Spies, T., Sorrentino, R., Boss, J. M., Okada, K., and Strominger, J. L.: Structural organization of the DR subregion of the human major histocompatibility complex. *Proc. Natl. Acad. Sci. U.S.A. 82:* 5165–5169, 1985

Springer, T. A., Kaufman, J. F., Terhorst, C., and Strominger, J. L.: Purification and structural characterization of human HLA-linked B-cell antigens. *Nature 268:* 213–218, 1977

Stanway, G., Cann, A. J., Hauptmann, R., Hughes, P., Clarke, L. D., Mountfort, R. C., Minor, P. D., Schild, G. C., and Almond, J. W.: The nucleotide sequence of poliovirus type 3 loc 12 a_1b: Comparison with polivirus type 1. *Nucleic Acids Res. 11:* 5629–5643, 1983

Thomas, P. S.: Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. *Methods Enzymol. 100:* 255–265, 1983

Trowsdale, J., Lee, J., Carey, J., Grosveld, F., Bodmer, J., and Bodmer, W. F.: Sequences related to HLA-DRα chain on human chromosome 6: Restriction enzyme polymorphism detected with DCα chain probes. *Proc. Natl. Acad. Sci. U.S.A. 80:* 1972–1976, 1983a

Trowsdale, J., Lee, J., and Bodmer, W. F.: Genes in the HLA-D region of human chromosome 6. *In* J. W. Parker and R. L. O'Brien (eds.): *Intercellular Communication in Leukocyte Function*, pp. 367–372, John N. Wiley and Sons Ltd., New York, 1983b

Trowsdale, J., Kelly, A., Lee, J., Carson, S., Austin, P., and Travers, P.: Linkage map of two HLA-SBβ and two HLA-SBα-related genes: An intron in one of the SBβ genes contains a processed pseudogene. *Cell 38:* 241–249, 1984

Trowsdale, J., Young, J., Kelly, A., Austin, P., Carson, S., Meunier, H., So, A., Erlich, H., Spielman, R., Bodmer, J., and Bodmer, W. F.: Structure, sequence and polymorphism in the HLA-D region. *Immunol. Rev. 85:* 5–43, 1985

Vieira, J. and Messing, J.: The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene 19:* 259–268, 1982

Weiss, E., Golden, L., Zakut, R., Mellor, A., Fahrner, K., Kvist, S., and Flavell, R. A.: The DNA sequence of the H-2K^b gene: Evidence for gene conversion as a mechanism for the generation of polymorphism in histocompatibility antigens. *EMBO J. 2:* 453–462, 1983

Widera, G. and Flavell, R. A.: The nucleotide sequence of the murine I-Eβ^b immune response gene – evidence for gene conversion events in class II genes of the major histocompatibility complex. *EMBO J. 3:* 1221–1225, 1984

Received November 6, 1985