# DIRECT OBSERVATION OF THE GENE ORGANIZATION OF THE COMPLEMENT C4 AND 21-HYDROXYLASE LOCI BY PULSED FIELD GEL ELECTROPHORESIS

By IAN DUNHAM, CAROLE A. SARGENT, ROGER L. DAWKINS,* AND R. DUNCAN CAMPBELL

*From the Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom; and the *Department of Clinical Immunology, Royal Perth Hospital, GPO Box X2213, Perth, Western Australia 6001*

The class III region of the human MHC between HLA-B and HLA-DR includes genes encoding the complement components C2, Factor B, C4A and C4B, and the cytochrome P450 steroid 21-hydroxylase (21-OHase) (1–3). The *C2* and *factor B* genes are separated by 421 bp (4) and lie 30 kb from the *C4A* locus, which in turn is ~10 kb from the *C4B* locus (5). One of the two copies of the *21-OHase* gene, *21-OHase A* and *21-OHase B*, lies 3 kb downstream of each *C4* gene (6). Recently, a gene that has been called *RD*, due to the presence of an unusual periodic structure, has been mapped between the *factor B* and *C4A* genes (7).

Although the two C4 isotypes, C4A and C4B, are highly homologous and differ by <1% in their derived amino acid sequence, C4 is exceptionally polymorphic (8). The sequences of four C4A and five C4B cDNA and genomic clones have established the pattern of polymorphism in the C4d fragment of the $\alpha$ chain of C4 (9) and have provided a structural basis for the observed functional (10, 11) and serological differences (12, 13) between the isotypes. There is also heterogeneity in *C4* gene size (9, 14, 15). All *C4A* genes studied at locus I are 22 kb in size (long *C4* gene). *C4B* genes, however, may be either 22 or 16 kb (short *C4* gene), due to the presence or absence of a 6–7-kb intron ~2.5 kb from the 5′ end of the gene. In addition, variation in the number of copies of *C4* genes present on individual chromosomes has been observed. Duplication of the C4B locus resulting in the presence of three *C4* genes has been observed on the extended haplotype B14 C4A2 C4B1,2 C2C BfS DR1, and the gene frequency has been estimated at 1–2% (16–18). Null alleles at either locus are much more common and gene frequencies of 5–15% for *C4AQ0* alleles and 10–20% for *C4BQ0* alleles have been estimated (19–21). Investigation of the molecular basis of null alleles is particularly important as they occur with increased frequency in some HLA-associated diseases, such as SLE (22–24). It has been estimated that about half the null alleles are due to deletion of the *C4* gene together with the flanking *21-OHase* gene (14).

C. A. Sargent and I. Dunham held Medical Research Council and Science and Engineering Research Council studentships, respectively. Address correspondence to R. D. Campbell, MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U. K.

A number of RFLPs are available to follow the differences at the human *C4* loci in the population (see reference 25 for summary, and 26, 27). In particular, the *Taq* I polymorphism at the 5′ end of the genes and the *Nla* IV and *Eco*0 109 polymorphic patterns can give information about the nature and characteristics of the two *C4* loci. These include whether the 6–7-kb intron is present, whether they express the C4A or C4B isotypes, and whether they would express the Rg1 or Ch1 antigenic determinants. However, it is not easy to deduce the gene copy number as it is necessary to compare the intensities of the bands revealed by autoradiography and comparison between different samples can be difficult. Second, family studies are necessary to imply which genes lie on the same chromosome.

We have recently constructed a physical linkage map of the human MHC (2) using the powerful technique of pulsed field gel electrophoresis (PFGE)[1], which allows the separation of large DNA fragments (28, 29). These studies have been extended to compare the DNA organization of different HLA haplotypes using PFGE and homozygous typing lymphoblastoid cell lines. By utilization of large DNA fragment RFLPs, we have been able to observe directly the variations in size of the *C4* loci in different haplotypes. Specifically, the size of the diagnostic *Bss*H II or *Sac* II restriction fragments observed with a C4- or 21-OHase-specific DNA probe indicates both the number of *C4* genes present on a chromosome and their size (C4 long or short). This technique can be applied to PBMC DNA, and together with the previously established RFLPs (14, 25), allows a complete definition of the *C4* gene organization of an individual.

## Materials and Methods

*Preparation of DNA.* The HLA and complement types of the cell lines studied are listed in Table I. The method for preparing high molecular weight DNA in agarose blocks from tissue culture cells has been described previously (2), except that cells were finally resuspended in 1% agarose at $2 \times 10^7$ cells/ml.

Whole blood was isolated from healthy individuals and PBMC were obtained by separation on Histopaque-1077 (Sigma Chemical Co., St. Louis, MO) according to the protocol supplied by the manufacturer. The cells were washed three times in PBS containing 1 mM EDTA before counting and were finally resuspended at $4 \times 10^7$ cells/ml in PBS. DNA in agarose blocks was then prepared as previously described (2).

*Restriction Enzyme Digestion and PFGE Analysis.* Agarose blocks containing ~5 μg DNA were treated before digestion as previously described (2). Restriction enzyme digestions were carried out as specified in reference 2, except that for *Bss*H II (New England Biolabs, Beverley, MA), 10–15 U were used in a 3-h incubation, for *Sac* II (Northumbria Biologicals Ltd., Cramlington, Northumberland, UK), 40 U of enzyme were used for a 3-h incubation, and for *Taq* I (Bethesda Research Laboratories, Gaithersburg, MD), 50 U were used in an overnight incubation at 55°C. After digestion, DNA in agarose blocks was loaded directly into PFGE gels without proteinase K treatment.

The PFGE gel box was constructed in our departmental workshop according to the design of Southern et al. (30). Digested samples were subjected to electrophoresis on 1.5% agarose (type I; Sigma Chemical Co.) gels in 20 mM Tris-acetate, 1 mM EDTA, pH 8.5, at 150 V for 30 h at a 30-s switching interval or for 36 h at a 7.5-s switching interval. Markers were intact yeast chromosomes in agarose blocks (31), concatemers of λ cI857S7 DNA prepared in agarose blocks (32), or λ DNA digested with *Hin*d III (Amersham International, Amersham, UK). Gels were stained, transferred onto Genescreen plus (DuPont Co., Stevenage, UK) membranes, and hybridized with radiolabeled DNA probes as previously described (2).

[1] *Abbreviation used in this paper:* PFGE, pulsed field gel electrophoresis.

*Taq* I-restricted DNA in agarose blocks was subjected to electrophoresis on conventional 0.7% agarose submarine gels, transferred to nylon-backed nitrocellulose membranes (Hybond C extra; Amersham International), and hybridized using standard procedures (33). Southern blots were stripped of probe and rehybridized as described by Dunham et al. (2).

*Probe.* The human C4 5′ cDNA probe, $P_A$, is a 476-bp *Bam*H I/*Kpn* I restriction fragment from the full-length C4 cDNA clone, pAT-A (34). The human 21-OHase probe is a 1.2-kb *Cla* I/*Pvu* II genomic DNA restriction fragment isolated from a cloned 21-OHase B gene (35). The other probes were a 1.6-kb *Hin*d III/*Bam*H I genomic DNA fragment located ~6 kb 5′ to the C2 gene (probe K) and a 1.7-kb *Bam*H I genomic DNA fragment lying 10 kb 5′ of the C4 gene, which were both derived from cloned cosmid DNA inserts as described by Dunham et al. (2). The positions of probes relative to the complement loci are shown in Fig. 4.

All probes were labeled with α-[$^{32}$P]dCTP (Amersham International) by random hexanucleotide priming (36).

## Results

*Detection of Deletions at the C4 Loci.* To observe differences between the DNA organization in several common HLA haplotypes, high molecular weight DNA in agarose blocks was prepared from the HLA-homozygous lymphoblastoid cell lines whose HLA types are detailed in Table I. The DNA was digested with restriction enzymes that cut infrequently in the mammalian genome (see reference 31 for review) and were known to give DNA fragments of an informative size around the complement gene cluster in the HLA class III region (2). DNA from each of the cell lines digested with a single restriction enzyme was then separated by PFGE at a switching interval so as to optimally resolve the DNA fragment size of interest. The DNA was depurinated, transferred to nylon membranes, and then hybridized with genomic and cDNA probes specific for the complement gene cluster.

It is important to note that the amount of DNA loaded on a PFGE gel affects the migration of the DNA relative both to other genomic DNA samples and to the DNA size markers (37, I. Dunham, unpublished observations) and that DNA fragment sizing can be inconsistent between different gel runs and switching intervals (38). These problems can be overcome by loading equivalent amounts of DNA for each sample and by comparing the sizes of restriction fragments for different cell lines on the same gel. We also possess a number of internal size markers for our prototype cell line 1, since the exact size of certain genomic DNA fragments produced with specific enzyme and probe combinations in the class III region has been determined by comparison with the cloned cosmid DNA (Sargent, C. A., I. Dunham, and R. D. Campbell, manuscript submitted for publication).

It is also interesting to note that the restriction fragment sizes observed with the PFGE apparatus used here were generally ~10% smaller than our previously published data using the orthogonal field alternation gel electrophoresis apparatus and were found to match more accurately with the sizes predicted from the cloned DNA.

The results of hybridization of *Mlu* I-digested DNA from each of the cell lines with the 21-OHase probe are shown in Fig. 1 and Table I. In cell lines 1–4, a 200-kb *Mlu* I fragment was seen, in cell lines 5 and 8, a 225-kb *Mlu* I fragment was seen, while cell lines 6 and 7 gave a 230-kb *Mlu* I fragment. Thus, there appears to be a 25–30-kb difference in the size of the *Mlu* I fragment that encompasses the complement genes between cell lines 1–4 and 5–8. Analysis of the cell line DNAs for the *Taq* I polymorphism (14, 25) on conventional agarose gels using the probe at the 5′ end of the C4 cDNA, $P_A$, and the 21-OHase probe suggests the basis for the

TABLE I

*Restriction Fragment Sizes (in Kilobases) Observed in Cell Line DNA*

| Cell line | HLA type | *Mlu* I | *Bss*H II | | | *Sac* II* | | | *Taq* I | | C4 Genotype |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 210H/ $P_A$ | 210H/ $P_A$ | 1.7-kb *Bam*H I | Probe K | 210H/ $P_A$ | 1.7-kb *Bam*H I | Probe K | $P_A$ | 210H | |
| 1 | B7 BfS C4A3 C4BQ0 DR2 | 200 | 80 | 12 | 55 | 40 | 12 | 55<br>(25) | 7.0 | 3.7 | Long locus I |
| 2 | B8 BfS C4AQ0 C4B1 DR3 | 200 | 70 | 12 | 55 | 33 | 12 | 55<br>(25) | 6.4 | 3.7 | Short locus I/II recombinant |
| 3 | B8 BfS C4AQ0 C4B1 DR3 | 200 | 70 | 12 | 55 | 33 | 12 | 55<br>(25) | 6.4 | 3.7 | Short locus I/II recombinant |
| 4 | B18 BfF1 C4A3 C4BQ0 DR3 | 200 | 80 | 12 | 55 | 40 | 12 | 55<br>(25) | 7.0 | 3.7 | Long locus I |
| 5 | B18 BfS C4A4 C4B2 DR2 | 225 | 105 | 12 | 55 | 65 | 12 | 55<br>(25) | 7.0<br>5.4 | 3.7<br>3.2 | Long locus I + short locus II |
| 6 | B44 BfS C4A3 C4BQ0 DR4 | 230 | 115 | 12 | 55 | 70 | 12 | 55<br>(25) | 7.0<br>6.0 | 3.7<br>3.2 | Long locus I + long locus II |
| 7 | B35 BfS C4A3 C4B1 DR5 | 230 | 115 | 12 | 55 | 70 | 12 | 55<br>(25) | 7.0<br>6.0 | 3.7<br>3.2 | Long locus I + long locus II |
| 8 | B55 BfS C4A4 C4B5 DR6 | 225 | 105 | 12 | 55 | 65 | 12 | 55<br>(25) | 7.0<br>5.4 | 3.7<br>3.2 | Long locus I + short locus II |

* The *Sac* II fragment of 25 kb (in parentheses) detected by probe K is due to partial digestion.

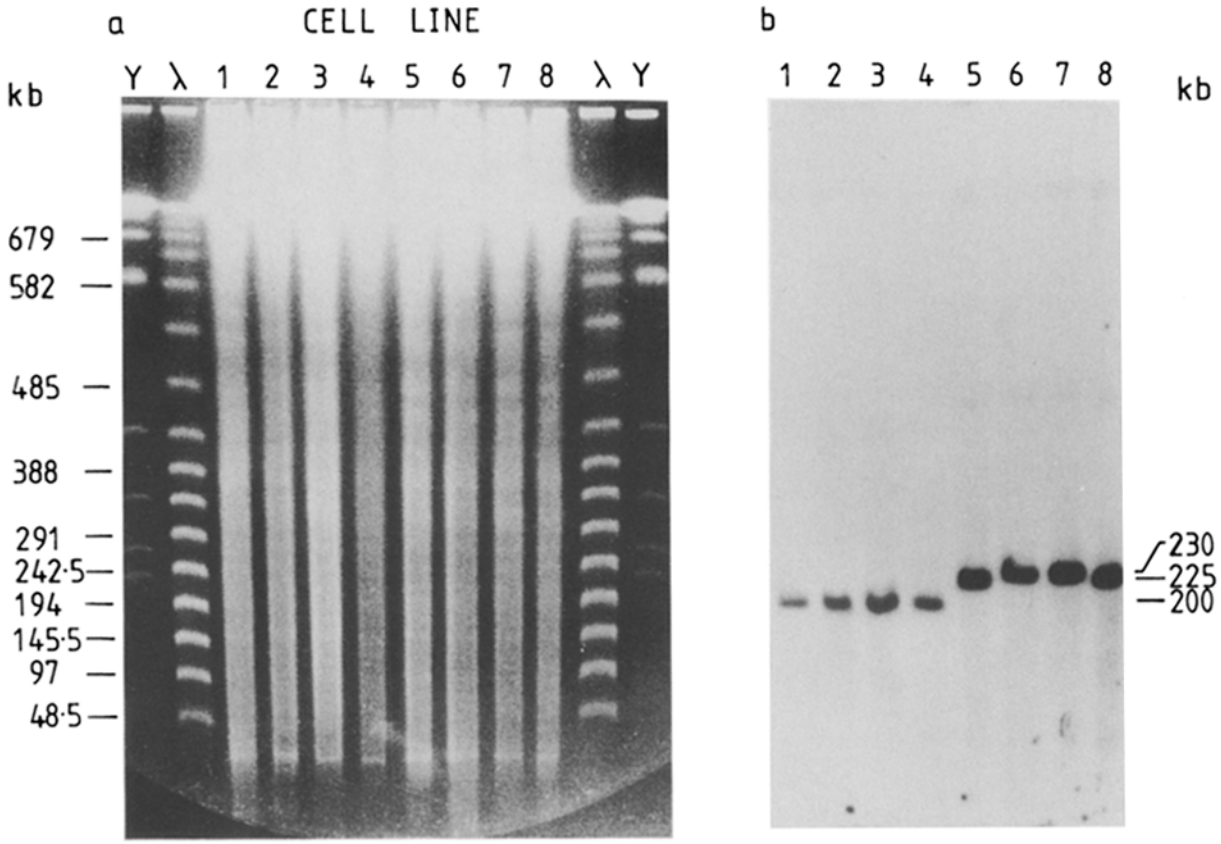


FIGURE 1. *Mlu* I restriction fragment patterns of haplotypes with one or two *C4* loci. (*a*) The EtBr-stained PFGE gel showing separation of *Mlu* I-digested cell line DNA at a 30-s switching interval. The markers are yeast chromosomes (*Y*) and λ concatemers (λ). (*b*) The autoradiograph of the Southern blot after hybridization with the 21-OHase probe showing the presence of a single Mlu I fragment of either 200, 225, or 230 kb, dependent upon the number of *C4* genes present and their size (i.e., long or short). Numbers above tracks refer to the cell lines described in Table I.

differences in size observed using *Mlu* I (Fig. 2 and Table I). Cell lines 1 and 4 had a single 7.0-kb *Taq* I fragment with the $P_A$ probe and a 3.7-kb fragment with the 21-OHase probe. Thus, these cell lines have a single long *C4* gene (which is *C4A* from the complotyping) and a single *21-OHase B* gene on each haplotype. The *Taq* I results for cell lines 2 and 3 indicate that they possess a single short *C4* gene and a *21-OHase B* gene. However, cell lines 5 and 8 gave two *Taq* I fragments of 7 and 5.4 kb with $P_A$, and 3.2- and 3.7-kb fragments with the 21-OHase probe. This suggests that cell lines 5 and 8 have one long *C4* gene, one short *C4* gene, and a copy of both *21-OHase A* and *21-OHase B* genes. Cell lines 6 and 7 had *Taq* I fragments characteristic of two long *C4* genes, and a copy of each *21-OHase* gene. Since the size of the DNA segment containing single *C4* and *21-OHase* genes is ~25 kb when the *C4* gene is short and ~30 kb when there is a long *C4* gene, the observed difference in size of the *Mlu* I restriction fragment appears to be due to the possession of two *C4* genes in cell lines 5–8, one of which is deleted in cell lines 1–4. Indeed, the *Mlu* I fragment in cell lines 6 and 7 was ~5 kb larger than for cell lines 5 and 8. This corresponds to the presence of the 6–7-kb intron at the second *C4* locus in cell lines 6 and 7, but not in cell lines 5 and 8, allowing for the limits of the size resolution of this PFGE gel. However, the equivalent difference due to the presence or absence of this intron could not be detected between cell lines 1 and 4 and 2 and 3. Hybridization of other probes to the same Southern blot of *Mlu* I-digested genomic DNA (results not shown) suggested that the mobility of the DNA fragments for cell lines 2 and 3 was slightly retarded in the PFGE gel, and hence, the size difference due to the 6–7-kb intron has been obscured.

*Pattern of the BssH II and Sac II Polymorphisms.* To confirm these observations, high molecular weight DNA from each of the cell lines was digested with *Bss*H II (Fig. 3 and Table I) and *Sac* II (Table I). The digested DNA was separated by PFGE at

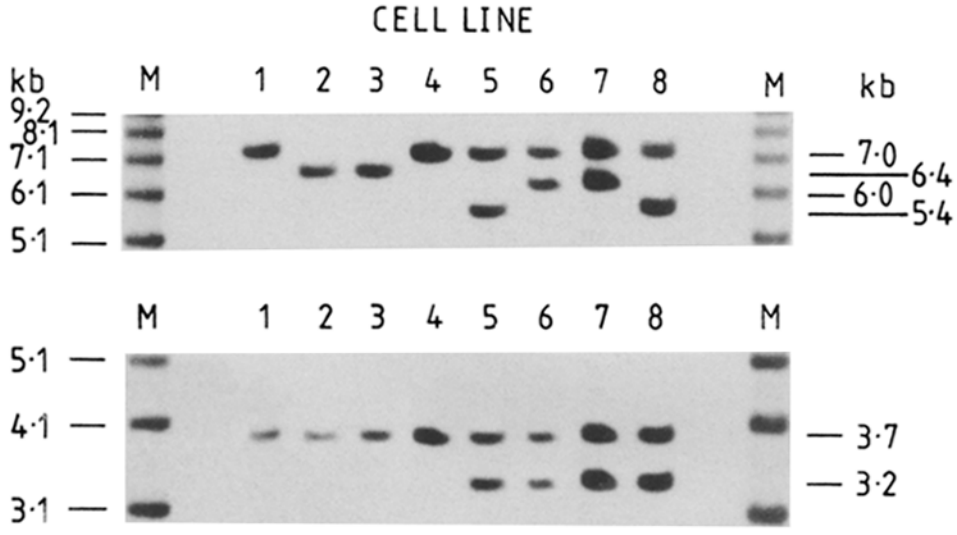


FIGURE 2. *Taq* I restriction patterns of genomic DNA showing the nature of the *C4* and *21-OHase* loci in the different cell lines. Markers are the BRL 1-kb ladder (*M*).

a switching interval of 7.5 s to obtain optimal resolution in the 10–200-kb size range. Fig. 3 shows the results of hybridization of probes from the complement gene cluster of a Southern blot of *Bss*H II-digested DNA. The positions of the probes are indicated in Fig. 4. Hybridization of probe K (1.6-kb *Hin*d III/*Bam*H I fragment from a region telomeric of the *C2* gene) to the *Bss*H II-digested cell line DNA revealed a common 55-kb fragment for all the cell lines. The size of this fragment is different from what had been observed previously (2), where only a larger 150-kb fragment was observed due to partial digestion.

Hybridization of the same blot with a 1.7-kb *Bam*H I fragment lying ~10 kb upstream of the *C4A* gene identified a 12-kb *Bss*H II fragment also in all the cell lines. Hence, there appears to be no difference between these cell lines in ~67 kb of DNA containing the *C2* and *factor B* genes. Hybridization of these probes to *Sac* II-digested DNA confirmed this result (Table I). However, when either the 21-OHase or $P_A$ probes were hybridized to the Southern blot of *Bss*H II-digested DNA, the sizes of the fragments observed were different between the cell lines depending on the organization of the *C4* loci. Cell lines 2 and 3, which possess a single short *C4B* gene and a *21-OHase B* gene, gave a 70-kb *Bss*H II fragment with the $P_A$ probe. For cell lines 1 and 4, which instead have a long *C4A* gene, the $P_A$ probe hybridized to an 80-kb *Bss*H II fragment. Thus, by using PFGE at this resolution, the difference in size of the C4 genes due to the 6–7-kb intron can be readily detected. Similarly, in the cell lines with one long and one short *C4* gene (cell lines 5 and 8), the *Bss*H II fragment was ~10 kb shorter than in those with two long *C4* genes (cell lines 6 and 7). In addition, the difference between the *Bss*H II fragment sizes for haplotypes with one *C4* gene or two *C4* genes corresponded to the size of a C4-21-OHase unit that is deleted or present. Therefore, the size of the *Bss*H II fragment that hybridizes to the $P_A$ or 21-OHase probes is directly related to the amount of DNA that is present at the *C4* and *21-OHase* loci.

It is also apparent that the size differences between these haplotypes that were observed in the *Mlu* I fragment can be completely accounted for by the difference in the size of the *Bss*H II fragment that contains the *C4* genes. In these haplotypes, at least, it seems that there are no other changes in the amount of DNA present in the region surrounding the *C4* genes as defined by the *Mlu* I fragment. The results using *Bss*H II were completely reflected using *Sac* II (Table I). With $P_A$ or the 21-OHase probe, a smaller fragment with *Sac* II was observed (70 kb for two long *C4* genes) than with *Bss*H II, but the differences in size of the *Sac* II fragments between haplotypes were again fully accounted for by the number and type of *C4* genes present. Therefore, the extent of the differences between these cell lines due to deletion of *C4* and *21-OHase* genes is completely defined to be within the limit of the *Sac* II fragment that hybridizes to the $P_A$ or 21-OHase probes.

The region of genomic DNA contained in the 200-kb *Mlu* I fragment that hybridizes with the 21-OHase probe has been cloned in a series of overlapping cosmid clones from a cosmid library prepared using DNA from cell line 1 (2; Sargent, C. A., I. Dunham, and R. D. Campbell, manuscript submitted for publication). These cosmid clones were mapped for rare cutting restriction enzymes and the data are presented in Fig. 4. The restriction enzyme sites that are cleaved in the genomic DNA of cell line 1 were identified using a series of probes derived from the cloned cosmid inserts (Sargent, C. A., I. Dunham, and R. D. Campbell, manuscript submitted for publi-

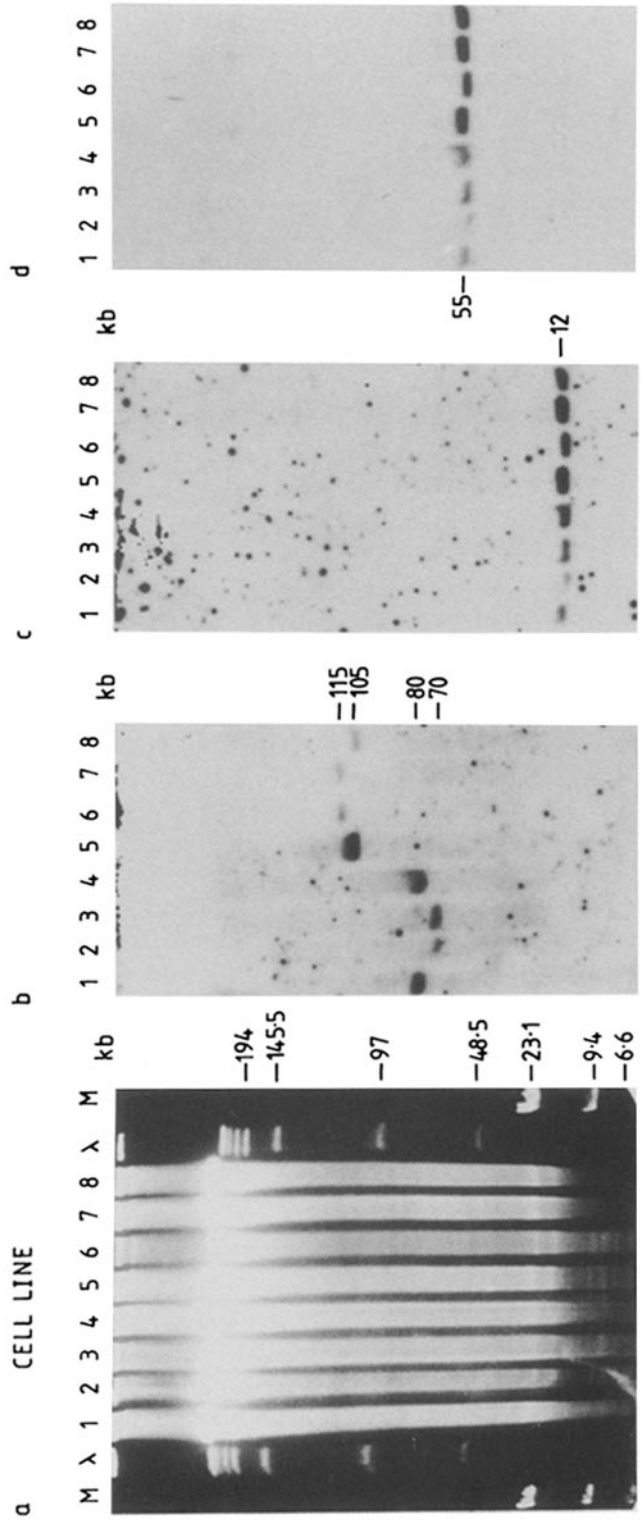


FIGURE 3. *Bss*H II restriction fragment patterns of haplotypes with differences in the number and size of the *C4* loci present. (*a*) The EtBr-stained PFGE gel showing separation of *Bss*H II–digested cell line DNA at a 7.5-s switching interval. (*b–d*) The same filter hybridized with the $P_A$ probe, the 1.7-kb *Bam*H I fragment, and probe K (see Fig. 4 for position of probes), respectively. Numbers above each track refer to the cell lines described in Table I. Markers are concatemers of λ DNA (λ) and a *Hind* III digest of λ DNA (*M*).

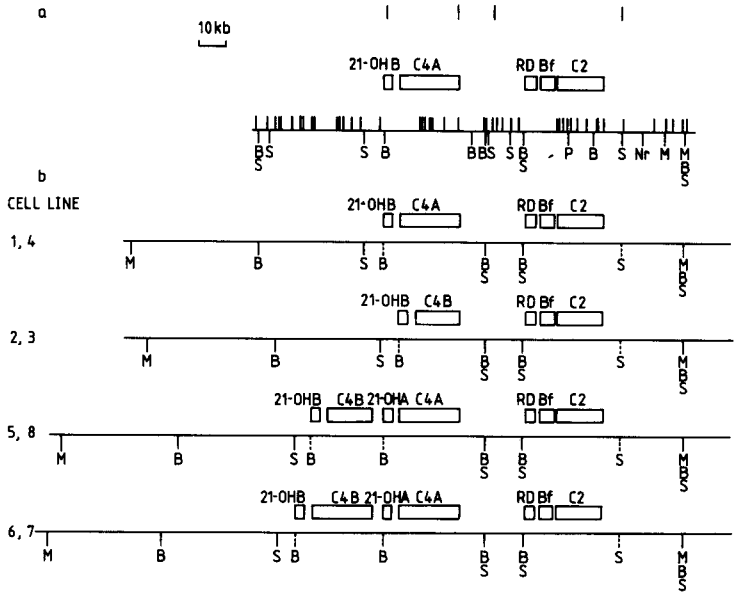


FIGURE 4. Molecular maps of the haplotypes analyzed showing the location of the *Bss*H II, *Sac* II, and *Mlu* I sites that cut in genomic DNA. (*a*) Restriction enzyme map of overlapping cosmid clones (from cell line 1 DNA) encompassing the *C2*, *Factor B*, *RD*, *C4A*, and *21-OHase* B genes (2, 7). The locations of the probes used in this study are illustrated by the four vertical lines at the top of the figure and from left to right are: the 21-OHase probe, probe $P_A$, the 1.7-kb *Bam*H I fragment, and probe K. The upper bars in the restriction map represent the position of the *Bam*H I sites, while the lower bars represent the positions of the sites for the restriction endonucleases: B, *Bss*H I, M, *Mlu* I; Nr, *Nru* I; S, *Sac* II; P, *Pvu* I. (*b*) Positions of the *Mlu* I, *Bss*H II and *Sac* II sites that cut in genomic DNA. A dotted line represents a site that cuts partially. Abbreviations for the restriction endonucleases are as above.

cation) and these data are also shown in Fig. 4. It is obvious that some sites that are present in the cell line DNA are not cleaved by the restriction enzymes in genomic DNA. This is presumably because the sites are methylated in the genomic DNA as both *Bss*H II and *Sac* II are methylation sensitive (39). Taking this information together with the *Mlu* I, *Bss*H II, *Sac* II, and *Taq* I data presented above, it is possible to construct genomic restriction maps for the cell lines analyzed (Fig. 4), depicting the differences in size and number of the *C4* and *21-OHase* genes present. It is not possible to say whether sites that are not restricted in cell line 1 genomic DNA are present in the other cell lines.

*Analysis of the C4/21-OHase Genes in PBMC DNA.* The results presented above suggested that it might be possible to apply the *Bss*H II and *Sac* II RFLPs to the analysis of the *C4* loci of different individuals. To this end, high molecular weight DNA from five healthy C4-typed and, in three cases, HLA-typed individuals was prepared in agarose blocks from PBMC. One possible pitfall of this approach might be that the pattern of methylation of restriction enzyme sites in PBMC is different from that observed for the EBV-transformed lymphoblastoid cell lines. The DNAs were digested with *Taq* I, *Bss*H II, and *Sac* II and separated on conventional 0.7% agarose gels or PFGE gels as appropriate. Southern blots of the separated, digested

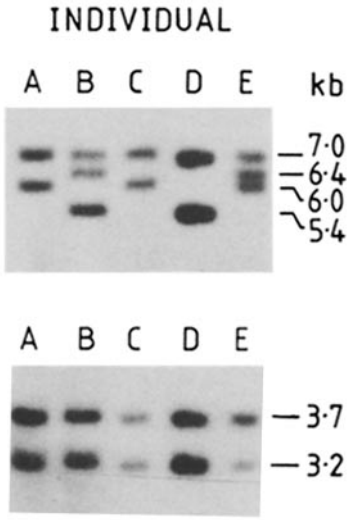


FIGURE 5. *Taq* I restriction patterns showing the nature of the *C4* and *21-OHase* loci in the five unrelated individuals summarized in Table II.

DNAs were hybridized with the $P_A$ and 21-OHase probes. The results are shown in Fig. 5 (*Taq* I) and Fig. 6 (*Bss*H II and *Sac* II) and are summarized in Table II.

Individual A had a single *Bss*H II fragment of 115 kb identified by hybridization with probe $P_A$ migrating in the same position as the *Bss*H II fragment for cell line 7, which was also subjected to electrophoresis on this gel as a standard. This fragment is characteristic of a chromosome possessing two long *C4* genes and the information from the *Taq* I digest confirms the presence of long *C4* genes at locus I and locus II (Fig. 5) and one copy of each of the *21-OHase A* and *B* genes.

Individual B possesses a 70-kb *Bss*H II fragment characteristic of a short *C4* gene. This corresponds in size to the 70-kb fragment of cell line 2, which was also subjected to electrophoresis on this gel as a standard. In addition, this individual has a novel 135-kb *Bss*H II fragment. Examination of the *Taq* I restriction enzyme data for this individual showed the presence of a 6.4-kb *Taq* I fragment, which is due to a short *C4* gene on the SCOI complotype (14, 15, 25). This *Taq* I fragment corresponds to the presence of the 70-kb *Bss*H II fragment, which also goes with this complotype. In addition, comparison of the band intensities in the *Taq* I digest hybridized with $P_A$ for individual B shows the presence of a single long *C4* gene (7.0-kb *Taq* I fragment) and two short *C4* genes (5.4-kb *Taq* I fragment). Since only one *C4* gene lies on the 70-kb *Bss*H II fragment, it can be deduced that the 135-kb *Bss*H II fragment contains one long *C4* gene and two short *C4* genes. The size of the fragment is consistent with this.

Individual C was essentially the same as individual A with respect to the polymorphic fragments seen with *Bss*H II, *Sac* II, and *Taq* I and, therefore, has the same *C4* gene organization (Table II).

Individual D has a different combination of *Bss*H II fragments revealed with probe $P_A$. The 105-kb fragment contains one long *C4* gene and one short *C4* gene on one chromosome, while the 135-kb fragment again represents a chromosome with three *C4* genes, one long and two short. Examination of the results from the *Taq* I digest

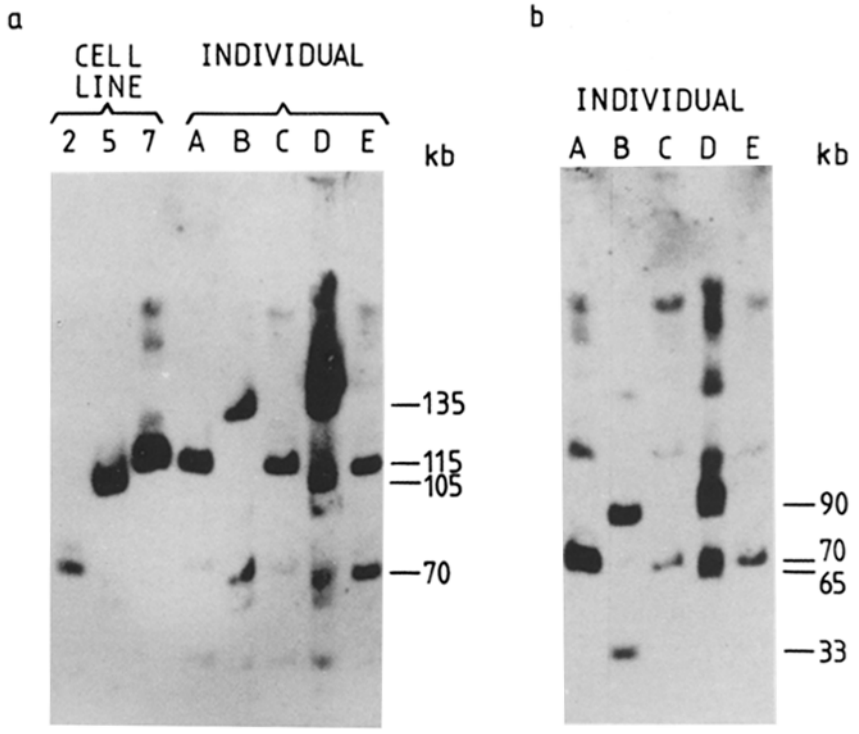


FIGURE 6. (*a*) *Bss*H II and (*b*) *Sac* II patterns showing the number and type of *C4* loci present in five unrelated individuals. Numbers above each track refer to the cell lines described in Table I, while letters refer to the individuals described in Table II. Markers are concatemers of λ DNA.

TABLE II

*Restriction Fragment Sizes (in Kilobases) Observed in PBMC DNA*

| | | *Taq* I | | *Bss*H II | *Sac* II | |
|---|---|---|---|---|---|---|
| Individual | HLA/complement type | $P_A$ | 210H | $P_A$/210H | $P_A$/210H | C4 Genotype |
| A | B7 C4A3 C4B1 DR5 | 7.0 | 3.7 | 115 | 70 | 2 × long locus I + long |
| | B44 C4A3 C4BQ0 DR6 | 6.0 | 3.2 | | | locus II |
| B | B14 C4A2 C4B1,2 DR13 | 7.0 | 3.7 | 135 | 90 | 1 × long locus I + short |
| | | 6.4 | 3.2 | | | locus II + short locus III |
| | B8 C4AQ0 C4B1 DR3 | 2 × 5.4 | | 70 | 33 | 1 × short locus I/II |
| | | | | | | recombinant |
| C | B45 C4A3 C4BQ0 DR11 | 7.0 | 3.7 | 115 | 70 | 2 × long locus I + long |
| | B18 C4A3 C4B1 DR4 | 6.0 | 3.2 | | | locus II |
| D | C4A3, A4; C4B1, B2 | 2 × 7.0 | 2 × 3.7 | 135 | 90 | 1 × long locus I + short |
| | | 3 × 5.4 | 3 × 3.2 | | | locus II + short locus III |
| | | | | 105 | 65 | 1 × long locus I + short |
| | | | | | | locus II |
| E | C4A3, AQ0; C4B1, B1 | 7.0 | 2 × 3.7 | 115 | 70 | 1 × long locus I + long |
| | | 6.4 | 3.2 | | | locus II |
| | | 6.0 | | 70 | 33 | 1 × short locus I/II |
| | | | | | | recombinant |

confirm that this is the case. With $P_A$, the band intensities suggested the presence of two long *C4* genes (7.0-kb *Taq* I fragment) and three copies of the short *C4* genes (5.4-kb *Taq* I fragment). Similarly there are two *21-OHase B* genes and three *21-OHase A* genes. This case clearly demonstrates that since the size of the *Bss*H II fragment is altered with the gene copy number present in an individual rather than the band intensity, it is easy to identify the number of *C4* genes present. In addition, the size of the *Bss*H II fragment is also diagnostic of the *C4* gene organization on an individual chromosome.

Individual E has *Bss*H II fragments consistent with two long *C4* genes on one chromosome and one short *C4* gene on the other SCOI haplotype, and this is confirmed by the *Taq* I digests. In all five cases the *Sac* II fragment sizes were entirely consistent with the organization of the *C4* loci suggested by the *Bss*H II and *Taq* I results (Fig. 6 *b*, Table II).

In Fig. 6 *a* it can also be seen that in addition to the series of strongly hybridizing bands revealed with $P_A$, a number of less intense *Bss*H II fragments were observed. These additional bands are probably the result of partial digestion of *Bss*H II sites, perhaps due to partial methylation at these sites or to differences in the methylation pattern of different cell types in the PBMC population. In particular, it appears that the *Bss*H II sites that lie in the 3′ end of the *21-OHase* genes (Fig. 4 *b*) are being partially digested. This phenomenon does not affect the information obtained from the autoradiographs because the informative fragments are still the major products. Therefore, it appears that the *Bss*H II and *Sac* II polymorphisms can be applied to DNA isolated from PBMC and are diagnostic of the organization of the *C4* loci on each chromosome.

## Discussion

Susceptibility to a number of diseases has been shown to be associated with the possession of certain HLA haplotypes (24, 40). Having established a physical map of the HLA region (2), it was of interest to know whether there are extensive differences in the organization of the DNA in different haplotypes (deletions, insertions, or RFLPs). The existence of deletions at the *C4* and *21-OHase* loci has been documented (14, 41), but because of the size of the *C4* genes, analysis of the extent of the deletions has required cosmid cloning studies. We set out to define whether other changes in the DNA organization exist around the complement genes in the HLA region and whether the documented differences at the *C4* and *21-OHase* loci could be observed in uncloned DNA.

Using PFGE, large DNA fragment RFLPs for the enzymes *Bss*H II, *Mlu* I, and *Sac* II have been observed at the *C4* loci in the HLA class III region. The sizes of the fragments hybridizing with the $P_A$ or 21-OHase probes is directly related to the number and length of the *C4* genes present in the DNA samples analyzed (see Tables I and II). These RFLPs, in particular, those with *Bss*H II and *Sac* II, can be used to directly observe the *C4* gene organization on both chromosomes using high molecular weight DNA isolated from the whole blood of an individual. In combination with the previously described *Taq* I polymorphisms (14, 15, 25), these RFLPs can give a complete picture of the *C4* gene organization, and by implication, the *21-OHase* gene organization for an individual without the need for family studies or DNA cloning.

In addition, it has been demonstrated that in the cell lines studied, representing at least seven different haplotypes (cell lines 2 and 3 carry the same haplotype although they are derived from different sources), the only alterations in the DNA content at the ~5-kb resolution of these PFGE gels are the deletions/duplications associated with the *C4* and *21-OHase* loci. The extent of the differences found in the *Mlu* I fragment between the haplotypes studied are confined to the *Sac* II fragment, which contains the *C4* and *21-OHase* genes. This *Sac* II fragment starts ~10 kb 5′ to the first *C4* gene and ends ~8 kb from the 3′ end of the last *21-OHase* gene (Fig. 4). No other differences in the DNA content of the haplotypes studied over the extent of the *Mlu* I fragment that contains all the HLA class III complement genes were detected using PFGE.

We believe that the RFLPs described here will be invaluable in screening individuals for their *C4* and *21-OHase* gene organization. It is already known that *C4* null alleles are associated with certain HLA-associated diseases (22–24), and this method in combination with the other available RFLPs should give a rapid screening procedure for splitting deleted from nondeleted null alleles. It is also interesting to note that of the five individuals that were analyzed (Table II), there were two chromosomes with deleted *C4AQ0* alleles and two chromosomes with nondeleted *C4BQ0* null alleles. In addition, there were two examples of chromosomes with three *C4* genes present, both having one long *C4* gene and two short *C4* genes. Although there have been previous reports of three *C4* genes on one chromosome (14, 18), there is little data as yet as to how frequently these duplications occur in the population. Using the RFLPs described here, it should be possible to assess the frequency of these deletions and duplications directly.

Both *Bss*H II and *Sac* II give the same information as to the organization of the *C4* loci. The fragments produced with *Sac* II are 40–45 kb smaller than those observed with *Bss*H II using the $P_A$ or 21-OHase probes, and this may make it easier to resolve size differences between different gene organizations. Also, since there are no other *Sac* II sites within the observed fragments, at least from our observations of cell line 1, no smaller partial fragments are produced. However, in our experience, *Bss*H II has given more consistent results and gives complete digestion with less units of enzyme, whereas *Sac* II can give very partial digestion.

Finally, it may be possible to make use of these RFLPs along with isotype and Rg/Ch-specific oligonucleotides to give a complete C4 genotype for an individual with a single PFGE gel.

## Summary

Pulsed field gel electrophoresis and enzymes that cut genomic DNA infrequently have been used to define large RFLPs at the human *C4* loci. With the enzymes *Bss*H II or *Sac* II, and C4 or 21-hydroxylase DNA probes, it has been possible to observe directly the number of *C4* genes present on a haplotype, and also whether the *C4* genes are long (6–7-kb intron present) or short (6–7-kb intron absent). Haplotypes that have either two long *C4* genes or one long and one short *C4* gene generate *Bss*H II fragments of ~115 or ~105 kb, respectively. Haplotypes that have either a single long or a single short *C4* gene generate *Bss*H II fragments of ~80 or ~70 kb, respectively. This technique has been used to analyze the DNA isolated from PBMC and allows

the complete definition of the *C4* gene organization of an individual without the need for family studies.

We thank Dr. Ken Reid for critical review of the manuscript, Ken Johnson and Carolyn Brooks for photographic and secretarial help, and Sandra Smith for excellent technical assistance.

*Received for publication 17 October 1988.*

## References

1. Lamm, L. U., and B. Olaisen. 1985. Report of the committee on the genetic constitution of chromosome 6. *Cytogenet. Cell Genet.* 40:128.
2. Dunham, I., C. A. Sargent, J. Trowsdale, and R. D. Campbell. 1987. Molecular map of the human major histocompatibility complex by pulsed field gel electrophoresis. *Proc. Natl. Acad. Sci. USA.* 84:7237.
3. Carroll, M. C., P. Katzman, E. M. Alicot, B. H. Koller, D. Geraghty, H. T. Orr, J. L. Strominger, and T. Spies. 1987. A linkage map of the human major histocompatibility complex including the tumor necrosis factor genes. *Proc. Natl. Acad. Sci. USA.* 84:8535.
4. Wu, L. C., B. J. Morley, and R. D. Campbell. 1987. Cell specific expression of the human complement protein Factor B gene: evidence for the role of two distinct 5' flanking elements. *Cell.* 48:331.
5. Carroll, M. C., R. D. Campbell, D. R. Bentley, and R. R. Porter. 1984. A molecular map of the human major histocompatibility complex class III region linking complement genes *C4*, *C2* and *Factor B. Nature (Lond.).* 307:237.
6. Carroll, M. C., R. D. Campbell, and R. R. Porter. 1985a. Mapping of steroid *21-hydroxylase* genes adjacent to complement *C4* genes in HLA, the major histocompatibility complex in man. *Proc. Natl. Acad. Sci. USA.* 82:521.
7. Levi-Strauss, M., M. C. Carroll, M. Steinmetz, and T. Meo. 1988. A previously undetected MHC gene with an unusual periodic structure. *Science (Wash. DC).* 240:201.
8. Mauff, G., C. A. Alper, Z. Awdeh, J. R. Batchelor, J. Betrams, G. Bruun-Petersen, R. L. Dawkins, P. Demant, J. Edwards, G. Grosse-Wilde, G. Hauptmann, P. Klouda, L. Lamm, E. Mollenhauer, C. Nerl, B. Olaisen, G. O'Neill, C. Rittner, M. H. Roos, V. Skanes, P. Teisberg, and L. Wells. 1983. Statement on the nomenclature of human C4 allotypes. *Immunobiology.* 164:184.
9. Yu, C. Y., K. T. Belt, C. M. Giles, R. D. Campbell, and R. R. Porter. 1986. Structural basis of the polymorphism of human complement components C4A and C4B: gene size, reactivity and antigenicity. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2873.
10. Law, S. K. A., A. W. Dodds, and R. R. Porter. 1984. A comparison of the properties of two classes, C4A and C4B, of the human complement component C4. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1819.
11. Isenman, D. E., and J. R. Young. 1984. The molecular basis for the difference in immune haemolysis activity of the Chido and Rodgers isotype of human complement component C4. *J. Immunol.* 132:3019.
12. Giles, C. M., B. Uring-Lambert, J. Goetz, G. Hauptmann, A. H. L. Fielder, W. Ollier, C. Rittner, and T. Robson. 1988. Antigenic determinants expressed by human C4 allotypes: a study of 325 families provides evidence for the structural antigenic model. *Immunogenetics.* 27:442.
13. Yu, C. -Y., R. D. Campbell, and R. R. Porter. 1988. A structural model for the location of the Rodgers and Chido antigenic determinants and their correlation with the human complement component C4A/C4B isotypes. *Immunogenetics.* 27:399.
14. Schneider, P. M., M. C. Carroll, C. A. Alper, C. Rittner, A. S. Whitehead, E. J. Yunis,

and H. R. Colten. 1986. Polymorphism of human human complement *C4* and steroid *21-hydroxylase* genes. Restriction fragment length polymorphisms revealing structural deletions, homo-duplications, and size variants. *J. Clin. Invest.* 78:650.

15. Palsdottir, A., A. Arnason, R. Fossdal, J. H. Edwards, and O. Jensson. 1987. Heterogeneity of human *C4* gene size; a large intron (6.5 kb) is present in all *C4A* and some *C4B* genes. *Immunogenetics.* 25:299.
16. Rittner, C., C. M. Giles, M. L. H. Roos, P. Demant, and E. Molenhauer. 1984. Genetics of human C4 polymorphism: detection and segregation of rare and duplicated haplotypes. *Immunogenetics.* 19:321.
17. Uring-Lambert, B., J. Goetz, M. M. Tongio, S. Mayer, and G. Hauptmann. 1984. C4 haplotypes with duplications at the C4A or C4B loci: frequency and associations with Bf, C2 and HLA-A, B, C, DR alleles. *Tissue Antigens.* 24:70.
18. Carroll, M. C., K. T. Belt, A. Palsdottir, and R. R. Porter. 1984b. Structure and organisation of *C4* genes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 306:379.
19. Schendel, D. J., G. J. O'Neill, and R. Wank. 1984. MHC-linked class III genes. Analyses of *C4* gene frequencies, complotypes and associations with distinct HLA haplotypes in German Caucasians. *Immunogenetics.* 20:23.
20. Tokunaga, K., K. Omoto, T. Akaza, N. Akiyama, H. Amemiya, S. Naito, T. Sasazuki, H. Satoh, and T. Juji. 1985. Haplotype study on C4 polymorphism in Japanese. Associations with MHC alleles, complotypes, and HLA-complement haplotypes. *Immunogenetics.* 22:359.
21. Partanen, J., and S. Koskimies. 1986. Human class III genes, Bf and C4. Polymorphism complotypes and association with MHC class I genes in the Finnish Population. *Hum. Hered.* 36:269.
22. Fielder, A. H. L., M. J. Walport, J. R. Batchelor, R. I. Rynes, C. M. Black, I. A. Dodi, and G. R. U. Hughes. 1983. A family study of the MHC of patients with SLE. Null alleles of C4A and C4B may determine the disease. *Br. Med. J.* 186:425.
23. Hauptmann, G., J. Goetz, B. Uring-Lambert, and E. Grosshans. 1986. C4 deficiency. *Prog. Allergy.* 39:232.
24. Dawkins, R. L., F. T. Christiansen, P. H. Kay, M. Garlepp, J. McCluskey, P. N. Hollingsworth, and P. J. Zilko. 1983. Disease associations with complotypes, supratypes and haplotypes. *Immunol. Rev.* 70:5.
25. Yu, C. Y., and R. D. Campbell. 1987. Definitive RFLPs to distinguish between the human complement C4A/C4B isotypes and the major Rodgers/Chido determinants. Application to the study of C4 null alleles. *Immunogenetics.* 25:383.
26. Whitehead, A. S., D. E. Woods, E. Fleischnick, J. E. Chin, E. J. Yunis, A. J. Katz, P. S. Gerald, C. A. Alper, and H. R. Colten. 1984. DNA polymorphism of the *C4* genes: a new marker for analysis of the major histocompatibility complex. *N. Engl. J. Med.* 310:88.
27. Palsdottir, A., R. Fossdal, O. Jensson, and A. Arnason. 1987. BglII polymorphic site upsteam to the human complement component C4A gene. *Nucleic Acids Res.* 15:2395.
28. Schwartz, D. C., and C. R. Cantor. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell.* 37:67.
29. Carle, G. F., and M. Olson. 1984. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* 12:5647.
30. Southern, E. M., R. Anand, W. R. A. Brown, and D. S. Fletcher. 1987. A model for separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* 15:5925.
31. van Ommen, G. J., and J. M. H. Verkerk. 1986. Restriction analysis of chromosomal DNA in a size range up to two million base pairs by pulse field gradient electrophoresis. *In* Human Genetic Disease, A Practical Approach. K. E. Davies, editor. IRL Press Limited, Oxford. 113–133.

32. Waterbury, P. G., and Lane, M. J. 1987. Generation of lambda phage concatemers for use as pulsed field electrophoresis size markers. *Nucleic Acids Res.* 15:3930.
33. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
34. Belt, K. T., M. C. Carroll, and R. R. Porter. 1984. The structural basis of the multiple forms of human complement component C4. *Cell.* 36:907.
35. Rodrigues, N. R., I. Dunham, C. Y. Yu, M. C. Carroll, R. R. Porter, and R. D. Campbell. 1987. Molecular characterisation of the HLA-linked steroid *21-hydroxylase B* gene from an individual with congenital adrenal hyperplasia. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:1653.
36. Feinberg, A. P., and B. Vogelstein. 1984. Addendum to "A technique for radio labeling DNA restriction endonuclease fragments to high specific activity." *Anal. Biochem.* 137:266.
37. Michiels, F., M. Burmeister, and H. Lehrach. 1987. Derivation of clones close to *met* by preparative field inversion gel electrophoresis. *Science (Wash. DC).* 236:1305.
38. Bernards, A., J. M. Kofter, P. A. M. Michels, R. M. P. Moberts, and P. Borst. 1986. Pulsed field gradient electrophoresis of DNA digested in agarose allows the sizing of the large duplication unit of a surface antigen gene in trypanosones. *Gene (Amst.).* 42:313.
39. Lindsay, S., and A. P. Bird. 1987. Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature (Lond.).* 327:336.
40. Batchelor, J. R., and A. J. McMichael. (1987). Progress in understanding HLA and disease associations. *Br. Med. Bull.* 43:156.
41. Carroll, M. C., A. Palsdottir, K. T. Belt, and R. R. Porter. 1985b. Deletion of complement *C4* and steroid *21-hydroxylase* genes in the HLA class III region. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2547.