EXTENSIVE DELETIONS AND INSERTIONS IN DIFFERENT MHC SUPRATYPES DETECTED BY PULSED FIELD GEL ELECTROPHORESIS

By KATSUSHI TOKUNAGA, G. SAUERACKER, P. H. KAY, F. T. CHRISTIANSEN, R. ANAND,* AND R. L. DAWKINS

*From the Departments of Clinical Immunology, Royal Perth Hospital, Queen Elizabeth II Medical Centre and University of Western Australia, Perth, Western Australia; and the *Department of Biochemistry, Oxford University, Oxford, United Kingdom*

The human MHC is located within ~2 cM of the 6p21.3 band (1) and contains at least 20 genes coding "*inter alia*" for three major classes of proteins: HLA A, B, and C (class I), HLA DP, DQ, and DR (class II), and serum complement components, C2, C4, and Bf. In addition, the region contains genes for 21-hydroxylase (21OH) and TNF (2). Thus, the MHC is a chromosomal segment containing several gene clusters of uncertain biological significance, but undoubtedly plays a role in disease susceptibility.

The most striking feature of the MHC is the remarkable degree of polymorphism; there are numerous alleles at multiple loci, including HLA A, B, DR, and C4. Surprisingly, however, some 70% of Caucasoids possess 1 or 2 of only 20 particular sets of MHC alleles (3) variously referred to as supratypes (4), extended haplotypes (5), or preferential allelic associations (6). Each of these supratypes has its own specific disease associations (4). Since we have shown that supratypes define and mark entire chromosomal segments of some 2 megabases (Mb),[1] we postulated that each supratype would contain its own unique set of new genes, deletions, duplications, and insertions (7). Furthermore, at least some of these supratype-specific arrangements (7) may be relevant to function and disease.

The technique of pulsed field gel electrophoresis (PFGE) allows separation of DNA molecules of up to 2 Mb in size (8). Using restriction endonucleases that cut infrequently, long range maps have been produced (7, 9–12), but there are discrepancies. Most of these maps (9–12) can only be approximations, since the importance of supratype-specific patterns was not appreciated. Accordingly, we present here the PFGE profiles of six different MHC supratypes and provide evidence for extensive deletions and insertions. These changes are present in unrelated subjects with the same supratype and are therefore predictable from comprehensive allotyping of the products of the known genes within the HLA B to DQ region.

This is publication 8754 of the Department of Clinical Immunology. This work was funded by grants from the National Health and Medical Research Council of Australia, by the Western Australian Arthritis and Rheumatism Foundation, and by the Immunogenetics Research Foundation. Address correspondence to R. L. Dawkins, Dept. of Clinical Immunology, Royal Perth Hospital, GPO Box X2213, Perth, Western Australia 6001.

[1] *Abbreviations used in this paper:* Mb, megabases; PFGE, pulsed field gel electrophoresis.

Materials and Methods

Cell Lines. High molecular weight DNA was isolated from EBV-transformed cell lines from unrelated subjects who were, in most cases, homozygous for the following supratypes: HLA A3,Cw7,B7,C2C,BfS,C4A3,C4B1,DR2 (3,7,2); HLA A1,Cw7,B8,C2C,BfS,C4AQO, C4B1,DR3 (1,8,3); HLA B18,C2C,BfF1,C4A3,C4BQO,DR3 (18,F1,3); HLA Bw65,C2C,BfS, C4A2,C4B1+2,DR1 (65,1+2, 1); HLA Bw57,C2C,BfS,C4A6,C4B1,DR7 (57,6,7); HLA A2, Cw1+3,Bw46,C2C,BfS,C4A4,C4B2,DRw9 (2,46,9). Diseases associated with these supratypes have been reported previously (4, 13–15).

DNA Preparation and Restriction Enzyme Digestion. Preparation of DNA, restriction enzyme digestion, PFGE, and Southern hybridizations were performed essentially as described (8, 16). Briefly, high molecular weight DNA was prepared from EBV-transformed cells, plugged in agarose of low gelling temperature, and treated with Proteinase K (1 mg/ml) in NDS (0.5 M EDTA, 10 mM Tris-HCl, 1% lauroyl sarcosine, pH 9.5) at 50°C 2× for 18 h. Genomic DNA plugs containing 6×10^5 cells were washed in TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5), equilibrated with the appropriate restriction buffer, and digested as described (8).

Pulsed Field Gel Electrophoresis. After digestion, the plugs were equilibrated with STOP (10 mM EDTA, 3% Ficoll, 5 mM Tris-HCl, 0.01% bromophenol blue, pH 8.0) and loaded into the gel slots. The digested DNA was run on 1.5% (wt/vol) agarose gels in 0.5× TAE (1 × TAE: 40 mM Tris-acetate, 2 mM EDTA, pH 8.5) at 20°C using 120-V, 35-h, 90-s pulses for large fragments, and 90-V, 65-h, 60-s pulses for small fragments (<500 kb) (8, 16). λ DNA oligomers and intact yeast chromosomes were used for size markers.

Southern Hybridization. After ethidium bromide staining, the gels were depurinated (0.25 M HCl, 2 × 15 m) and denatured (0.5 M NaOH + 1.5 M NaCl, 2 × 15 m). Southern transfer was performed onto a nylon membrane (Biotrace, Michigan) using 0.4M NaOH. Membranes were prehybridized, hybridized with the probes labeled by the random priming method (17), washed, and exposed according to standard methods for conventional gels. After stripping of the probe, the membranes were checked to ensure complete stripping by overnight exposure, then they were used for reprobing. The probes used are specific for DP α (18), DQ α (19), DR α (20), DR β (21), 21OH (22), Bf (23), TNF (24), and class I (25) genes.

Results

Genomic DNA was digested with the restriction endonucleases Mlu I, Nae I, Not I, Pvu I, and Sfi I, followed by PFGE/Southern blotting. Not I and Pvu I digests proved the most useful for establishing long-range restriction maps of the various supratypes.

Fig. 1 shows typical examples of the experiments. DNA was digested with Not I, Pvu I, or Not I + Pvu I. After separation in a PFGE apparatus and transfer onto a nylon membrane, restriction fragments were sequentially hybridized with a series of ^{32}P-labeled probes specific for MHC genes. Extensive variation between supratypes was observed not only in fragment length but also in band intensity. In Fig. 1 *a*, for example, the intensity of Bf and 21OH bands reflects the gene copy number per genome as in the case of conventional agarose gel electrophoresis. Thus, all three supratypes have two Bf genes per genome. By contrast, 1,8,3 and 18,F1,3 have two 21OH genes per genome, whereas 65,1+2,1 has six genes (26–28).

First, the map of one of the most common Caucasoid supratypes, 3,7,2, was constructed (Fig. 2). A large Not I fragment (1,030 kb) included the DQ and DR regions in addition to the class III genes (10, 12, 29) (Fig. 1 *c*), but DP and TNF genes were detected on different Not I fragments of 370 and 290 kb, respectively. The class I probe detected one strong band (960 kb) and a few weaker bands ranging from 400 to 700 kb. After Pvu I digestion, a large fragment (730 kb) hybridized to both TNF and class I probes (Fig. 1 *b*). These data indicate that the TNF genes are between

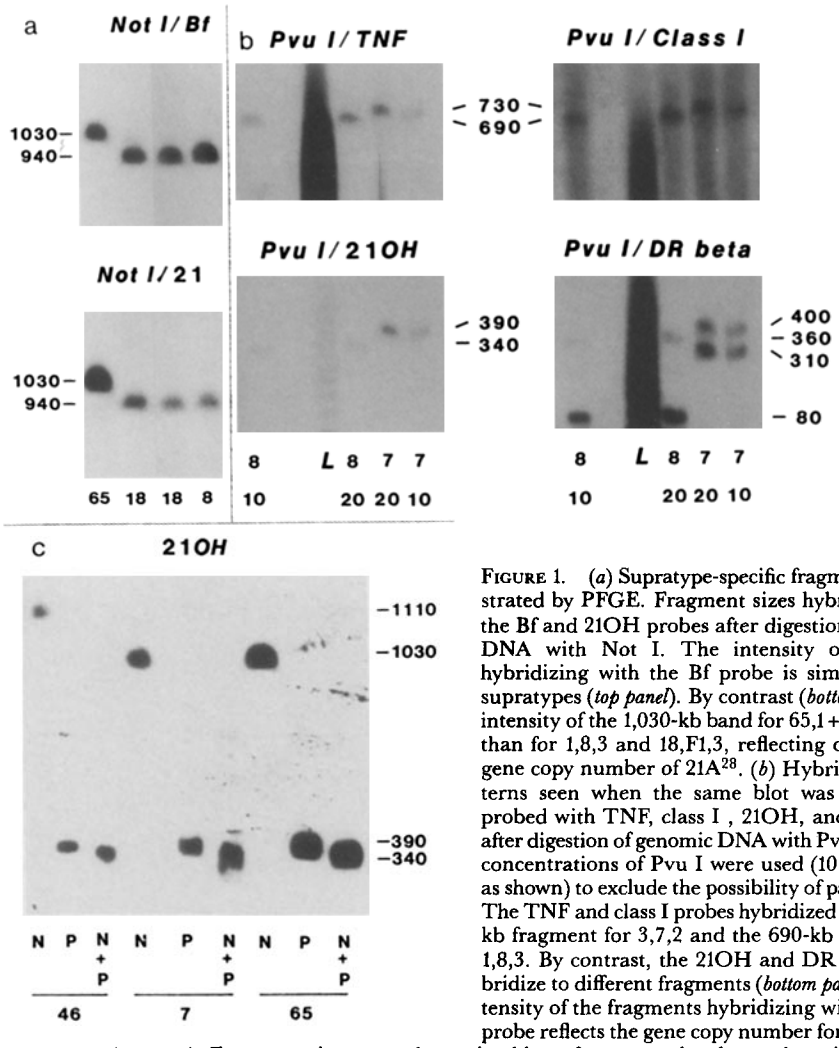

FIGURE 1. (*a*) Supratype-specific fragments demonstrated by PFGE. Fragment sizes hybridizing with the Bf and 21OH probes after digestion of genomic DNA with Not I. The intensity of the bands hybridizing with the Bf probe is similar between supratypes (*top panel*). By contrast (*bottom panel*), the intensity of the 1,030-kb band for 65,1+2,1 is greater than for 1,8,3 and 18,F1,3, reflecting differences in gene copy number of 21A[28]. (*b*) Hybridization patterns seen when the same blot was sequentially probed with TNF, class I , 21OH, and DR probes after digestion of genomic DNA with Pvu I. Different concentrations of Pvu I were used (10 U and 20 U as shown) to exclude the possibility of partial digests. The TNF and class I probes hybridized with the 730-kb fragment for 3,7,2 and the 690-kb fragment for 1,8,3. By contrast, the 21OH and DR β probes hybridize to different fragments (*bottom panels*). The intensity of the fragments hybridizing with the 21OH probe reflects the gene copy number for the different supratypes (see text). Fragment sizes were determined by reference to the electrophoretic mobility of yeast chromosome markers or λ DNA oligomers (L). (*c*) Sizes of fragments hybridizing with 21OH probe after digestion of genomic DNA with Not I (N), Pvu I (P), and Not I + Pvu I (N+P). The double digest allowed precise mapping of the Not I and Pvu I sites between TNF and the class III region. It can be seen that the 3,7,2 and 65,1+2,1 supratypes are similar in length but with the latter the fragments are denser in keeping with the increased gene copy number. For both supratypes N+P yield a fragment of ~340 kb, indicating that there are adjacent P and N sites (see Fig. 3). With the 2,46,9 the N fragment is ~1,110 kb. The reduced intensity is apparently due to a lower DNA concentration rather than a reduction in the number of 21OH genes.

the class III and class I genes. Two Pvu I fragments (310 and 400 kb) hybridized to the DR β probe. The shorter 310-kb band was the more intense, indicating the presence of two DR β genes on this fragment (Fig. 1 *b*), which also included the DQ/DX region. Two Mlu I fragments (460 and 590 kb) hybridized to the DQ α probe. The shorter extends from the middle of the DQ/DX region to DP, whereas the longer extends through the DR region. Not I + Pvu I double digests confirmed

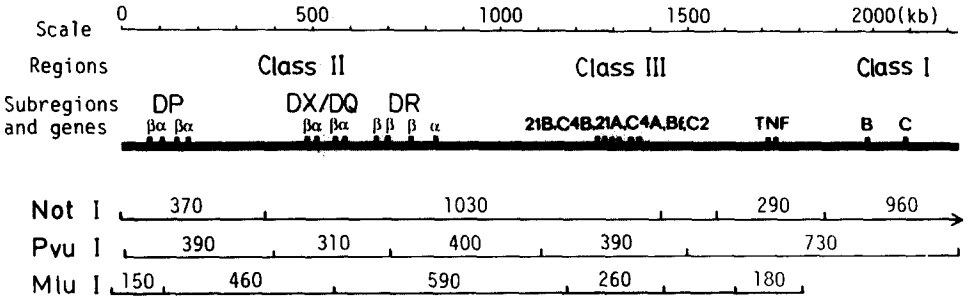

FIGURE 2. Genomic map of the human MHC in 3,7,2 supratype. The published maps of the class II (9) and III (26) regions were combined with our data. The two Not I sites between class III and TNF genes were localized by using Not I + Pvu I double digests. No probe was available to determine sizes of the Not I and Mlu I fragments between class III and TNF. Sizing errors are ~10 kb for the fragments of <500 kb, and 20 kb for >600 kb.

that the 290-kb Not I fragment hybridizing with the TNF probe is included in the 730-kb Pvu I fragment. Mapping of the Not I and Pvu I sites between class III and TNF genes and between DP and DQ subregions was determined by Not I + Pvu I double digestion (Fig. 1 *c*).

Having characterized the 3,7,2 supratype as a putative standard, we extended the analysis to other supratypes. At least three different cell lines from unrelated individuals or six unrelated haplotypes were examined for each supratype. Each supratype showed an identical hybridization pattern, with the exception of some bands hybridizing with the class I probe. This confirms that each supratype reflects a fixed structure throughout the entire HLA B to DQ region of more than 1 Mb (7). Further analysis (data not shown) has indicated that some supratypes "fragment" between DQ and DP.

In the 1,8,3 supratype (see Fig. 3), the large Not I fragment (940 kb) that includes the DQ, DR, and class III genes was smaller by ~90 kb than the corresponding fragment of 3,7,2 (1,030 kb). This is consistent with the finding that 1,8,3 has shorter Pvu I/class III (340 kb) and the Pvu I/DR (360 kb) fragments by 40–50 kb when compared with 3,7,2 (Fig. 1 *b*). The Pvu I fragment hybridizing with both the TNF and class I probes (690 kb) was also shorter by 40 kb when compared with 3,7,2 (Fig. 1 *b*). These findings indicate that the 1,8,3 supratype contains at least three and apparently four deletions. One of these is known to encompass 21A, most of C4A, some of C4B, and to result in the complotype C4AQO,C4B1 (26–28). The approximate locations of the other three deletions are shown in Fig. 3. Moreover, two further variations were identified. The 1,8,3 supratype has one extra Pvu I site between the DQ and DR regions, which results in an 80-kb fragment most likely including two out of three DR β genes (Fig. 1 *b*) and a 230-kb fragment including only DQ/DX genes. On the other hand, this supratype lacks the Not I site between TNF and class I genes, because a large Not I fragment (1,200 kb) hybridizes to both TNF as well as the class I probe. Furthermore, after Not I + Pvu I double digestion the 610-kb fragment also hybridized with the class I probe. It is possible that the Not I site is included in the 40-kb deletion within the class I/TNF region

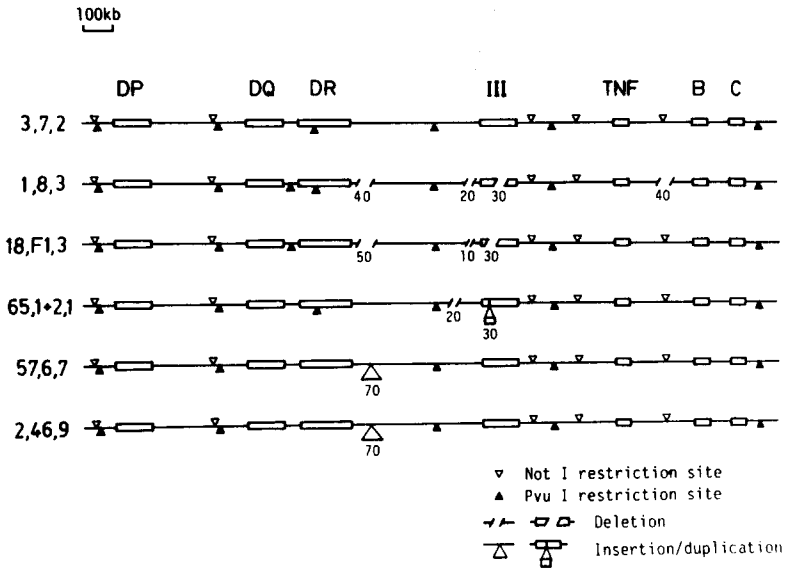

FIGURE 3. Genomic maps of the human MHC in different supratypes. This figure shows the restriction maps, possible deletions, and insertions/duplications specific for six different supratypes. There are other possibilities, including several smaller deletions and duplications. The extent of the deletions and duplications is estimated to ~10 kb for the 1,8,3, 18,F1,3, and 65,1+2,1 and 20 kb for the 57,6,7 and 2,46,9 supratypes.

of 1,8,3. This deletion and possibly others could result in the absence of unidentified genes.

Similar analyses were also performed on the other supratypes as shown in Fig. 3. In 18,F1,3 the large Not I fragment with DQ, DR, and class III genes was also smaller than 3,7,2 by ~90 kb. This agrees with the mapping of the Pvu I/DR (430 kb) and Pvu I/class III (350 kb) fragments. Although lacking the Pvu I site within the DR region, 18,F1,3 shares with 1,8,3 the extra Pvu I site between the DR and DQ genes. Thus, in addition to the known 30-kb deletion within the class III region (28), we now demonstrate that 18,F1,3 has further 50-kb and 10-kb deletions close to the DR and class III regions respectively.

Previously (28) we have reported that 65,1+2,1 has a duplication containing 21A and C4B and this has been confirmed by the demonstration of extra 21OH genes (see above). However, the Not I (1,030 kb, Fig. 1 *a*) and Pvu I (400 kb) fragments, including class III genes suggest that there must also be a deletion of 20–30 kb to explain the similarity in total length to 3,7,2. Thus the hybridization patterns in 65,1+2,1 were similar to those for 3,7,2 (Fig. 1 *c*).

Although different from other supratypes, the maps of 57,6,7 and 2,46,9 cannot be distinguished as yet. As with 18,F1,3, both supratypes lacked a Pvu I site within the DR region. Moreover, both the Not I/DQ + DR + class III fragment (1,110 kb, Fig. 3) and the Pvu I/DQ + DR fragment (780 kb) suggest that the 57,6,7 and 2,46,9 supratypes have extensive duplication(s) or insertion(s) of ~70–80 kb near the DQ or DR region.

Discussion

For the first time it has been possible to compare and contrast multiple haplotypes bearing particular MHC supratypes. Using PFGE we have shown that different supratypes have their own characteristics in terms of deletions, insertions, and duplications. It should be emphasized that our conclusions are based on differences in length and gene copy number. Restriction site polymorphism as might be produced by a single base mutation cannot explain our major findings, but could explain some differences in the class II region. Accordingly, these findings emphazise the importance of regarding the MHC in terms of chromosomal segments rather than as a complex of separated genes.

These newly recognized chromosomal arrangements provide new approaches to explaining the genetic susceptibility of diseases associated with MHC supratypes (4). For some years the standard approach to explaining such associations has involved the detailed study of the class II genes and their products. It has been assumed that specific immune response genes within the class II region are of primary importance in determining disease susceptibility and that other MHC associations are secondary to linkage disequilibrium. The present data suggest that this approach may be too restricted.

A particular supratype is a marker for a chromosomal segment that contains specific alleles at multiple loci over at least 1 Mb. Accordingly it will be important to consider the possible role of all polymorphic genes within the segment. In addition to class I and class II genes controlling antigen-specific responses, there are other (class III) genes with effects on immunoregulation, amplification, inflammation, and even metabolism. Further as yet unidentified genes may be involved. The possibility of *cis*-interaction between two or more genes cannot be discounted.

A potentially important implication of the present findings is that disease-associated supratypes mark the presence of deletions/insertions and differences in gene copy number. It follows that supratypes will reflect quantitative as well as qualitative differences in gene products.

The fact that apparently unrelated individuals with the same supratypes have similar if not identical chromosomal segments of >1 cM implies conservation of certain ancestral haplotypes. We have shown that there are substantial differences in the actual length of segments marked by specific supratypes and such differences could affect the frequency of recombination. On the other hand, length cannot be the sole explanation for conservation. As we have shown elsewhere, one of the longest segments (57,6,7) is common to Caucasoids and Chinese and must be highly conserved (30, 31). It seems likely that supratypes mark ancestral haplotypes that have some selective advantage conferred by a particular combination of alleles at multiple loci.

Finally, the present findings demonstrate the value of characterizing specific supratypes by studying subjects who are homozygous as revealed by comprehensive allotyping of the products of multiple genes between HLA B and DR. Maps of heterozygous and recombinant supratypes (9–12) will be different and can be confusing.

Summary

The genomic organization of the human MHC was examined in multiple examples of six different supratypes using pulsed field electrophoresis (PFGE) after digestion of genomic DNA with infrequently cutting restriction endonucleases. Differences

in restriction fragment length and band intensity were shown to be specific for each supratype. Mapping of the MHC revealed that each supratype contains previously undescribed deletions and insertions between HLA B and DQ regions.

We are grateful to Drs. R. D. Campbell and M. C. Carroll for helpful advice.

Received for publication 4 January 1988 and in revised form 23 May 1988.

References

1. Lamm, L. U., and B. Olaisen. 1985. Report of the committee on the genetic constitution of chromosomes 5 and 6. *Cytogenet. Cell Genet.* 40:128.
2. Spies, T., C. C. Morton, S. A. Nedospasov, W. Fiers, D. Pious, and J. L. Strominger. 1986. Genes for the tumour necrosis factors alpha and beta are linked to the human major histocompatibility complex. *Proc. Natl. Acad. Sci. USA.* 83:8699.
3. Wilton, A. N., F. T. Christiansen, and R. L. Dawkins. 1985. Supratype matching improves renal transplant survival. *Transplant. Proc.* 17:2211.
4. Dawkins, R. L., and F. T. Christiansen, P. H. Kay, M. J. Garlepp, J. McCluskey, P. N. Hollingsworth, and P. J. Zilko. 1983. Disease association with complotypes, supratypes and haplotypes. *Immunol. Rev.* 70:1.
5. Awdeh, Z. L., L. D. Raum, E. J. Yunis, and C. A. Alper. 1983. Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc. Natl. Acad. Sci. USA.* 80:259.
6. Festenstein, H., J. Awad, G. A. Hitman, S. Cutbush, A. V. Groves, P. Cassell, W. Ollier, and J. A. Sachs. 1986. New HLA DNA polymorphisms associated with autoimmune diseases. *Nature (Lond.).* 322:64.
7. Martin, E., G. Saueracker, K. Tokunaga, P. Kay, F. T. Christiansen, and R. L. Dawkins. 1987. Complotypes and class III gene rearrangements can be assigned from supratype specific restriction fragment length polymorphism. 10th International Histocompatibility Workshop Newsletter, New York.
8. Anand, R. 1986. Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. *Trends Genet.* 23:278.
9. Hardy, D. A., J. I. Bell, E. O. Long, T. Lindsten, and H. O. McDevitt. 1986. Mapping of the class II region of the human major histocompatibility complex by pulsed-field gel electrophoresis. *Nature (Lond.).* 323:453.
10. Lawrance, S. K., C. L. Smith, R. Srivastava, C. R. Cantor, and S. M. Weissman. 1987. Megabase scale mapping of the HLA gene complex by pulsed field gel electrophoresis. *Science (Wash. DC).* 235:1387.
11. Dunham, I., C. A. Sargent, J. Trowsdale, and R. D. Campbell. 1987. Molecular mapping of the human major histocompatibility complex by pulsed field gel electrophoresis. *Proc. Natl. Acad. Sci. USA.* 84:7237.
12. Carroll, M. C., P. Katzman, E. M. Alicot, B. H. Koller, D. E. Geraghty, D. E. Orr, J. L. Strominger, and T. Spies. 1988. Linkage map of the human major histocompatibility complex including tumour necrosis factor genes. *Proc. Natl. Acad. Sci. USA.* In press.
13. Wilton, A. N., T. J. Cobain, and R. L. Dawkins. 1985. Family studies of IgA deficiency. *Immunogenetics.* 21:333.
14. Kelly, H., V. L. McCann, P. H. Kay, and R. L. Dawkins. 1985. Susceptibility to IDDM is marked by MHC supratypes rather than individual alleles. *Immunogenetics.* 22:643.
15. Dawkins, R. L., B. R. Hawkins, W. Gin, W. J. Zhang, E. Martin, and F. Christiansen. 1988. Association between myasthenia gravis and HLA Bw 46, BfS, C4 A4, C4B2, DRw9 in Chinese. *In* Immunobiology of HLA. Vol. 2. Immunogenetics and Histocompatibility. Springer-Verlag, N.Y. In Press.

16. Southern, E. M., R. Anand, W. R. A. Brown, and D. S. Hetcher. 1987. A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Res.* 15:5925.
17. Feinberg, A. P. and B. Vogelstein. 1983. A technique for radidabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6.
18. Erlich, H., D. Stetler, R. Sheng-Dong, and R. Saiki. 1984. Analysis by molecular cloning of the human class II genes. *Fed. Proc.* 43:3025.
19. Trowsdale, J., J. Lee, J. Carey, F. Grosveld, J. Bodmer, and W. Bodmer. 1983. Sequences related to HLA DR alpha chain on human chromosome 6: restriction enzyme polymorphism detected with DR alpha chain probes. *Proc. Natl. Acad. Sci. USA.* 80:1972.
20. Gustafsson, K., K. Wiman, D. Larhammar, L. Rask, and P. A. Peterson. 1984. Signal sequences distinguish class II histocompatibility antigen beta chains at different loci. *Scand. J. Immunol.* 19:91.
21. Long, E. O., C. T. Wake, J. Gorski, and B. Mach. 1983. Complete sequence of an HLA-DR beta chain deduced from a cDNA clone and identification of multiple non-allelic DR beta chain genes. *EMBO (Eur. Mol. Biol. Organ.) J.* 2:389.
22. White, P. C., M. I. New, and B. Dupont. 1986. Structure of human steroid 21-hydroxylase genes. *Proc. Natl. Acad. Sci. USA.* 83:5111.
23. Morley, B. J., and R. D. Campbell. 1984. Internal homologies of the Ba fragment from human complement component Factor B, a class III MHC antigen. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:153.
24. Pennica, D., G. E. Nedwin, J. S. Hayflick, P. H. Seeburg, R. Derynck, M. A. Palladino, W. J. Kohr, B. B. Aggarwal, and D. V. Goeddel. 1984. Human tumour necrosis factor: precursor structure, expression and homology to lymphotoxin. *Nature (Lond.).* 312:724.
25. Sood, A. K., D. Pereira, and S. M. Weissman. 1981. Isolation and partial nucleotide sequence of a cDNA clone for human histocompatibility antigen HLA-B by use of an oligodeoxynucleotide primer. *Proc. Natl. Acad. Sci. USA.* 78:616.
26. Carroll, M. C., K. T. Belt, A. Palsdottir, and Y. Yu. 1985. Molecular genetics of the fourth component of human complement and steroid 21-hydroxylase. *Immunol. Rev.* 87:39.
27. White, P. C., M. I. New, and B. Dupont. 1985. Adrenal 21-hydroxylase cytochrome P-450 genes within the MHC class III region. *Immunol. Rev.* 87:123.
28. Garlepp, M. J., A. N. Wilton, R. L. Dawkins, H. Kelly, and P. C. White. 1986. Rearrangement of 21-hydroxylase genes in disease associated MHC supratypes. *Immunogenetics.* 23:100.
29. Ragoussis, J., A. Van der Bliek, J. Trowsdale, and A. Ziegler. 1986. Mapping of HLA genes using pulsed-field gradiant electrophoresis. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 204:1.
30. Kay, P. H., E. Martin, R. L. Dawkins, and P. Charoenwong. 1988. Class III gene rearrangements in Thai/Chinese supratypes containing null or defective C4 alleles. *Immunogenetics.* 27:46.
31. Kay, P. H., R. L. Dawkins, J. Williamson, K. Tokunaga, F. T. Christiansen, and P. Charoenwong. 1988. Co-existence of an MHC chromosomal segment marked by HLA B17,BfS,C4A6,B1,DR7 and DQw3.2 in different ethnic groups. *Hum. Immunol.* In press.