Journal of Medical Genetics 1987, **24**, 649–658

The molecular genetics of human chromosome 6

VINCENT CUNLIFFE AND JOHN TROWSDALE

From the Human Immunogenetics Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX.

SUMMARY Chromosome 6 contains several clinically important markers as well as classical enzyme loci, proto-oncogenes, and a growing number of anonymous DNA restriction fragment length polymorphisms (RFLPs). It is also of unique interest because of the location of the major histocompatibility complex (MHC) on the short arm, at 6p21·3. The MHC is one of the most detailed areas of the human genetic map to date and many important diseases, some of a suspected autoimmune aetiology, are associated with it.

About 70 well characterised genetic loci have been mapped to human chromosome 6 (fig 1). Considerable interest has been focused on those of the HLA system, located on the short arm at 6p21·3. This multigene complex, otherwise termed the major histocompatibility complex (MHC), encodes class I and class II HLA molecules as well as some complement components and is involved in defence against a variety of diseases. Particular alleles of some HLA loci are associated with hereditary predispositions to diseases, for example, insulin dependent diabetes mellitus (IDDM), narcolepsy, ankylosing spondylitis, rheumatoid arthritis, and Hodgkin's lymphoma. The HLA complex spans over 3 × 10^6 bp and to date contains more than 20 expressed genes. The considerable knowledge that has accumulated concerning the organisation, structure, regulation, and functions of the genes in this linkage group forms a paradigm for studies of the molecular biology of a complex region of the human genome and it will be covered in detail in this article.

Chromosome 6 is already known to encode a large number of enzyme and surface antigen loci, several proto-oncogenes, and at least six pseudogenes, four of which have been assigned close to, or within, the HLA region. In addition, there are several fragile sites involved in translocations, inversions, and deletions which are aetiologically linked to some lymphoid and other malignancies. The genetic markers of chromosome 6 are summarised in fig 1 and are referenced in the table.

Received for publication 18 May 1987.
Accepted for publication 21 May 1987.

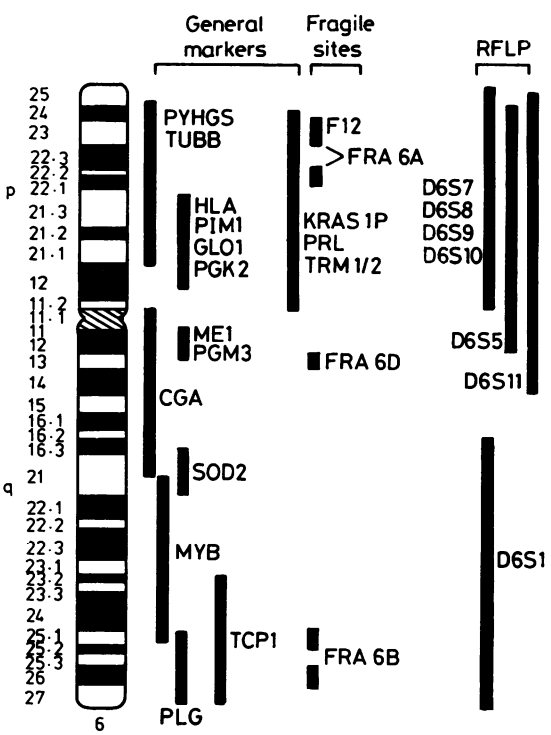

FIG 1 *Regionally localised markers on human chromosome 6. This figure is mostly based on data collated in the 8th Human Gene Mapping Workshop. The HLA complex is shown in detail in fig 2 and the markers are explained in the table. Vertical bars indicate the (approximate) extents of the boundaries within which a particular marker is located.*

TABLE *Markers on chromosome 6.*

Gene symbol	*Location*	*Description*	*Reference and remarks*
ADCP	6	Adenosine deaminase complexing protein 1	1, 17
ARG1	6	Arginase, liver	15, 62
ASSP2	6	Argininosuccinate synthetase pseudogene 2	4, 5, 31, 32
BEV1	6	Baboon M7 virus integration site	6, 84
BKMA1	6q21	Banded krait minor satellite DNA-1	85, 101
CA21HA	6p21·3	Congenital adrenal hyperplasia (21-hydroxylase deficiency) A	HLA region, see fig 2
CA21HB	6p21·3	Congenital adrenal hyperplasia (21-hydroxylase deficiency) B	HLA region, see fig 2
CGA	6q12→q21	Chorionic gonadotrophin, α polypeptide	45, 85
CSC1	6	Corticosterone sidechain isomerase	42–44, 82–84
D6S7	6p21→pter	Anonymous DNA segment from p7H4	19, 83
D6S8	6p21	Anonymous DNA segment from p2C5	18, 83
DHFRP2	6	Dihydrofolate reductase pseudogene 2	34, 77
ER	6	Oestrogen receptor	17, 59, 60, 98
F12	6p23	Clotting factor XII (Hageman factor)	14, 15, 47, 48
F13A	6p23→qter	Clotting factor XIII, A subunit	16
FEA	6	Fg embryonic antigen	14, 92
FRA6A	6p23	Fragile site, folic acid type, rare, fra(6)(p23)	12, 72
FRA6B	6p25·1 or 6q26	Fragile site, aphidicolin type, common, fra(6)(q25·1 or q26)	12, 72
FRA6C	6p22·2	Fragile site, aphidicolin type, common, fra(6)(p22·2)	12, 72
FRA6D	6q13	Fragile site, aphidicolin type, common, fra(6)(q13)	12, 72
FTHL	6p21·3→p12	Ferritin-like sequence, heavy polypeptide	12, 87, may be in HLA region, see note, fig 2
FUCA2	6	α-1-fucosidase 2	13, 49
GLO1	6p21·21→p21·1	Glyoxalase I	6–8, 17–19
HFE	6p21·3	Haemochromatosis	20, 88, may be in HLA region, strongly associated with HLA-A, see fig 2
HLA and complement components			
C2, C4, BF	6p21·3	See detailed map, fig 2	1, 44, 91 (this contains a discussion of HLA class I, class II, and class III genes and diseases associated with the region)
INFR	6q	α interferon receptor	57, 61
INSL	6p23→q12	Insulin-like sequence, AGB6	38, 78
KRAS1	6p23→q12	Proto-oncogenes, c-Ki-*ras*-1 processed pseudogene	21–23, 35–37
LQT	6	Long QT (Romano-Ward syndrome)	24, 93
ME1	6q12	Malic enzyme 1, soluble	11, 25
ME2	6	Malic enzyme 2, mitochondrial	10, 26
MCF3	6q16→q22	Proto-oncogene MCF-3	71, 102
MRBC	6	Monkey red blood cell receptor	27, 57
MSK28	6	Antigen identified by monoclonal antibody MG2	28, 53
MSK29	6	Antigen identified by monoclonal antibody A42	28, 53
MYB	6q15→q24	Avian myeloblastosis viral (v-*myb*) oncogene homologue	13, 29–31, 67–70
NDF	6	Neutrophil differentiation factor	15, 94
NEU	6p21·3	Neuraminidase 1, sialitosis	15, 90
P	6	P blood group (globoside)	32, 33, 55, 56
PGK1P2	6p21→p12	Phosphoglycerate kinase 1, pseudogene 2	20, 28–30, 34, 35, 102
PGM3	6q12	Phosphoglucomutase 3	12, 13, 37, 38
PIM		Proto-oncogene *pim*	54, 56, 64, 65
PLG	6p25→qter	Plasminogen	39, 40, 50, 51
PRL	6p23→q12	Prolactin	39, 79
PUT1	6	Polypeptide m44p in lymphocyte membrane	42, 52
PYHG5	6pter→p21	Protein spot in 2D gels (MW 90K)	43, 95
ROS	6q16→q22	Proto-oncogene *ros*	59, 63
RPL32P1	6p21·3	Ribosomal protein L32 pseudogene	33, 76
SCA1	6	Spinocerebellar ataxia	45, 46, 80, 81
SOD2	6q21	Superoxide dismutase 2, mitochondrial	14, 61
SRBC	6	Sheep red blood cell receptor	48, 58, M Brown, personal communication, 1987
S5	6	Surface antigen 5 (MW 45KD)	49, 54
TCP1	6q23→qter	T complex polypeptide 1	5, 24, 103
TNFA	6p23→q12	Tumour necrosis factor α	18, 78, 79, 97
TNFB	6p23→q12	Tumour necrosis factor β, lymphotoxin	18, 78, 79, 97
TRM1	6p23→q12	Methionine tRNA initiator 1	5
TRM2	6p23→q12	Methionine tRNA initiator 2	5
TS546	6	Temperature sensitivity complementation, ts546	19, 97
TUBB	6pter→p21	β tubulin	4, 52
YES2	6	Yamaguchi sarcoma viral (v-*yes*-2) oncogene homologue 2	53, 66

Correlating the formal genetic map with the molecular genetic map of chromosome 6

Morton *et al*[2] mapped the distribution of chiasmata along human chromosomes undergoing meiosis at the pachytene stage. Using data taken from a large number of chromosomes they constructed a map which charted the frequency of crossovers per unit

length of chromosome. On chromosome 6 this frequency was quite variable, particularly near the centromere and, therefore, inconsistencies between the molecular map and the formal genetic map are to be expected.

There are estimated to be about 52 chiasmata per human meiosis. The genetic length of the human autosomes is therefore 52/2 × 100 = 2600 cM (division by two is necessary because only two of the four strands in a bivalent are involved in the chiasma; multiplication by 100 gives the length in centimorgans). Chromosome 6 was estimated to constitute 2·5% of the total autosome length of 2600 cM, that is, 65 cM.[2] However, Francke and Pellegrino[3] estimated that chromosome 6 made up more like 6% of the haploid complement, which gives a much greater value of 150 cM. In the next few years, as more RFLP data accumulate, it will be possible to generate detailed molecular maps of sections of chromosome which are saturated with markers. The HLA region is ideal for this kind of analysis.

Enzymes, surface antigens, and pseudogenes

The combination of enzymatic activity assays, DNA hybridisation on Southern blots, and segregation analysis in somatic cell hybrids have localised genes for β tubulin *(TUBB)*,[4] two different tRNA initiator methionine genes *(TRM1, TRM2)*,[5] glyoxalase *(GLO1)*, [6-8] soluble and mitochondrial malic enzyme *(ME1, ME2)*,[9-11] phosphoglucomutase 3 *(PGM3)*, [12 13] mitochondrial superoxide dismutase (SOD–2),[14] liver arginase (one of two expressed genes, *ARG1)*,[15] fucosidase (FUCA2),[16] adenosine deaminase complexing protein *(ADCP1)*,[17] and numerous pseudogenes.

The glycolytic enzyme phosphoglycerate kinase 1 *(PGK1)* is coded for by a gene located on the X chromosome at q13, close to one of its processed pseudogenes *(PGK1P1*, Xq11→q13). Deficiencies at the functional locus can result in haemolytic anaemia.[18] The murine homologue also lies on the X chromosome.[19] There is, however, a PGK isoenzyme expressed exclusively in the testis that has an autosomal location.[20] In the mouse, this locus is on chromosome 17, close to both the major histocompatibility (H–2) complex and the T/t complex.[21] There is homology of synteny between human chromosome 6 and mouse chromosome 17, largely within and adjacent to the MHC.[22] There are also some reports of developmental defects in the axial skeleton[23] and male transmission ratio distortion effects, cosegregating with HLA, as well as recombination suppression between some extended HLA haplotypes.[25] Such phenomena are reminiscent of the behaviour of the mouse T/t complex[26] and lend some substance to the notion of a human equivalent.[24]

In view of this, and the restricted expression of the human and mouse isoenzymes to the testis,[27] it seemed plausible that this human PGK locus might be linked to HLA. A human genomic clone was isolated containing PGK sequences and mapped to 6p23→p12.[28] However, sequence analysis of this DNA identified a PGK processed pseudogene *(PGK1P2)* lacking all introns, flanked by 11 nucleotide direct repeats, punctuated by a premature in frame termination codon.[29] Recent evidence from a different source has indicated that the intron-less *PGK2* gene has a complete open reading frame and a testis specific pattern of expression.[102] As such it would be a rare example of a functional processed cellular (pseudo)gene. It is likely that both groups have studied the same genomic locus and, if this is so, discrepancies between their sequences have yet to be resolved. A third study showed the cosegregation of a PGK sequence with chromosome 19 in a panel of human–rodent hybrid cell lines. Moreover, sequencing of the DNA suggested that it encoded the autosomal testis specific PGK isoenzyme *(PGK2)*.[30]

Recent isolation of a genomic clone containing the *HSP70* gene facilitated its locus assignment to chromosome 6.[73] This gene encodes a 70K protein whose synthesis is induced in response to a variety of physiological stresses, most notably heat shock. It is extremely well conserved between phyla; this probably reflects its essential role in minimising the damaging effects of such stresses.

A cDNA probe for the functional, chromosome 9 linked argininosuccinate synthetase gene was used to identify 10 or more homologous sequences that were dispersed throughout the genome, with a panel of somatic cell hybrids.[31 32] Since a deficiency at the chromosome 9 locus results in citrullinaemia, the other sequences, including that on chromosome 6 *(ASSP2)*, are likely to be non-functional pseudogenes. Generation of processed pseudogenes is believed to depend on the activity of endogenous reverse transcriptase in germline cells acting on cellular RNAs and converting them into cDNAs capable of inserting into accessible regions of the genome. Providing that insertion does not abrogate the expression of essential functions, these insertion events are likely to persist. A processed pseudogene of the ribosomal protein L32 *(RPL32P1)* has been found in an intron of the DPβ1 gene.[33] By using a panel of human–rodent somatic cell hybrids, a processed pseudogene of dihydrofolate reductase *(DHFRP2)* has been mapped to chromosome 6.[34] An assignment to 6p12→p11 has also been made for a sequence with homology to the functional Kirsten-

ras proto-oncogene *(KRAS1)*.[35–37] Between 6p23 and 6q12 lies a 1·3 kb insulin-like sequence AGB6 *(INSL)* expressed in fetal brain.[38] This interval also contains the prolactin gene *(PRL)*.[39] An association between HLA–B8 and the development of prolactin secreting adenomata has been described,[40] so it would be of interest to determine the distance between the HLA region and the prolactin gene.

In the mouse, the level of corticosteroid side chain isomerase is under the control of a single locus linked to the H–2 complex. Furthermore, both H–2 and HLA are associated with differential sensitivity to glucocorticoids.[41–44]

Human chorionic gonadotrophin has two subunits, α and β. The α subunit is also a component of luteinising hormone, thyroid stimulating hormone, and follicle stimulating hormone, and is encoded by a single gene *(CGA)*. In situ and filter hybridisation have localised *CGA* to 6q12→q21.[45] Other studies, however, have concluded that the gene is on chromosome 18 at p11.[46]

Recent cloning of the gene coding for the oestrogen receptor has enabled its cytogenetic assignment to this autosome,[59 60] and an analysis of somatic cell hybrids showed that the α/β interferon receptor gene was also located on chromosome 6.[61] Interestingly, the synteny of the α/β interferon receptor and mitochondrial superoxide dismutase gene on 6q is recapitulated by synteny of the γ interferon receptor and cytoplasmic superoxide dismutase gene on chromosome 21. This may well reflect a conservation of linkage between sequences now on chromosome 6 during evolution.

Defects in coagulation can be caused by lesions in three loci that map to chromosome 6. Clotting factors XII (Hageman factor)[47 48] and XIII[16] map to 6p23→qter by deletion mapping and the plasminogen gene resides in the 6q25→qter region.[49–51]

Segregation analysis has been particularly useful in mapping the loci of cell surface antigens, several of which lie on chromosome 6[52–56] (see also the table). The gene encoding CDII, a T cell differentiation antigen, was thought to reside here.[57 58] However, recent molecular cloning of its gene facilitated in situ and filter hybridisation that placed the locus at 1p13 (M Brown, 1987, personal communication). Some cystic fibrosis (CF) patients have recently been noted to have deficiencies of vasoactive intestinal peptide. Chromosomal assignment of the locus of the gene encoding this hormone was essential to establish if it coincided with the CF locus defined by RFLP analysis on chromosome 7. Somatic cell hybrid analysis regionally localised this gene to 6p21→qter and refuted the hypothesis that defects at this locus were the primary cause of CF.[62]

Chromosome 6 and neoplasia

The link between genetic aberration and cancer was first established by the oberservation of a peculiarly small chromosome in the leukaemic cells of patients with chronic granulocytic leukaemia. Abnormalities as conspicuous as gross chromosomal rearrangements and as subtle as single point mutations can be equally potent causes of neoplasia. Such lesions can activate proto-oncogenes and thereby change the growth characteristics of the cell. Deletions, translocations, and inversions of chromosome 6 material have been observed in many types of leukaemia. Other rearrangements of this linkage group have been shown in retinoblastoma, malignant melanoma, and colon carcinoma and adenoma. The *ROS*,[63] *PIM*,[64 65] *YES*,[66] *MYB*,[67–70] and *MCF3*[71] proto-oncogenes all reside on chromosome 6. *ROS* and *MCF3* are, in fact, the same locus which has been transduced twice by retroviruses to generate the v-*ros* and *mcf*-3 viral oncogenes. Aberrant expression of any of these genes, as a consequence of any chromosomal rearrangement (including amplifications)[70] in which they are involved, could explain the aetiology of tumours possessing such lesions.

Fragile sites are regions of chromatin particularly prone to deletion or translocation and four such sites lie on chromosome 6: *FRA6A*, *FRA6B*, *FRA6C*, and *FRA6D* (table). Deletions covering some of them have been observed in malignant lymphoma, acute myeloid leukaemia, acute lymphoblastic leukaemia, and adult T cell leukaemia.[72]

The major histocompatibility complex (MHC)

The MHC consists of a cluster of at least 30 genes covering about 3 million base pairs (3 mega bp or 3 mbp) of DNA. It is divided into three regions, containing class I, class II, and class III genes, as shown in fig 2. There is also a fourth category containing all those genes which are not complement or HLA related and therefore include the 21-hydroxylase gene and the loci for TNF α and β.

The MHC has been mapped by conventional linkage analysis and, more recently, by pulsed field gradient gel electrophoresis.[98–100] These two techniques are in broad agreement on the positions of the subregions, assuming that 1 cM is roughly equivalent to 1 mbp. The maps in fig 2a and b are based on these data as well as on information both from overlapping cosmid clones and from B cell deletion mutant lines. There are several recent reviews of the HLA complex that should be consulted for detailed information (cited in reference 91).

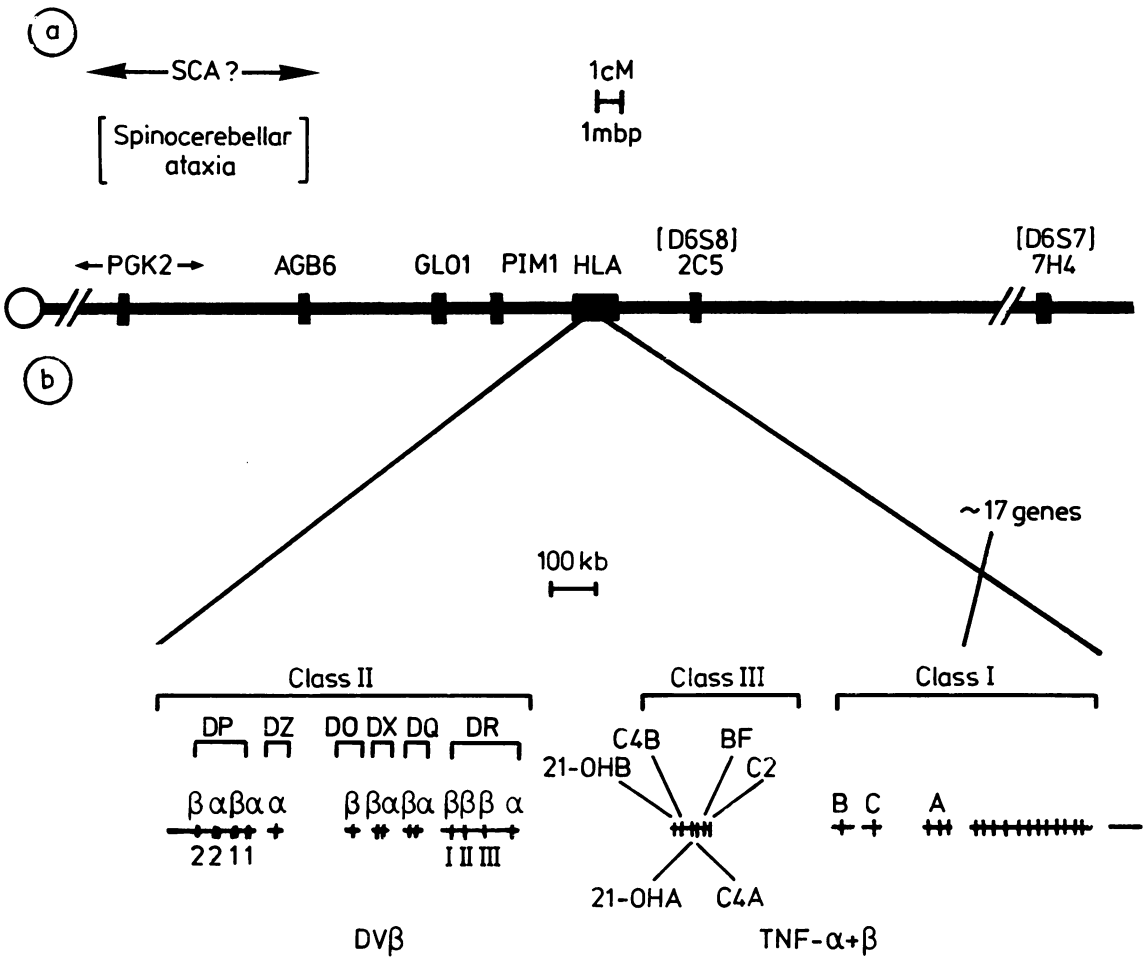

FIG 2 *Map of the region of human chromosome 6 around the HLA complex. The HLA region has been compiled using data from a number of sources as described in the text. The complement (class III) is shown orientated as in mouse although this has not yet been established in humans.*[101] *The tumour necrosis factor (TNF) α and β genes are thought to map within this complex from studies of deletion mutants. In the mouse these genes are proximal to H-2D and so may also be located between the complement region and HLA-B in man.*[101] *The PIM1 oncogene is also positioned centromeric to the HLA complex in view of its position in the mouse (A Berns, 1987, personal communication).*

THE HLA CLASS I REGION

Serological analysis established that this region contains multiple related genes. The best characterised expressed loci are HLA-A, HLA-B, and HLA-C which express glycoprotein chains of approximately 43 000 Daltons, associated with β2-microglobulin at the cell surface. There are other class I loci which may be equivalent to the mouse Qa and TL loci. However, such human homologues may be unlinked to HLA, given the recent assignment to chromosome 1 of a set of HLA class I related sequences, the CD1 family.[74] One of the best estimates of the number of class I genes comes from gene cloning from a haplotype loss mutant LCL 721 (H T Orr, 5th International H-2 and HLA Cloning Workshop) and suggests a minimum of 17.

At least six of these genes, by sequence analysis, are pseudogenes. All of the functionally defined

class I sequences isolated so far map to the same HLA region, largely at the HLA-A end of the MHC. Other genes may reside between the HLA-A and HLA-C loci.

HLA CLASS II GENES

The HLA class II products are heterodimeric glycoproteins present on the surface of antigen presenting cells consisting of an α chain of about 33 kd and a β chain of about 28 kd. Typically the genes for both partners in the heterodimers are close to each other, for example, DQα and β are adjacent as are DPα and β. Evidence from transfection experiments show that heterodimers of mixed isotype may form, for example, DPα coupling with DQβ, and similar events have been documented for the equivalent mouse Ia genes (R Sekaly, 1986, personal communication). At the centromeric end of the class II region is a stretch of 100 kb containing four DP genes, two α and two β, arranged in the order: centromere–DPβ2(pseudogene)–DPα2(pseudogene)–DPβ1–DPα1. The arrangement of these genes was determined by mapping of overlapping cosmid clones. The first intron of the DPβ1 gene contains a processed pseudogene of ribosomal protein L32 *(RPL32P1)*.

A gap of at least 125 kb separates two genes of unknown function, DZα and DOβ, and this could contain other genes. Both DZα and DOβ are transcribed, but so far no transcripts have been found for the DX genes. DXα and β are closely related by sequence to the next two genes, DQα and β, which encode the serologically defined HLA-DQ product.

The DR region contains one α chain gene, DRα, one or two expressed β chain genes, and a variable number of pseudogenes. Notably DR1 persons contain only one expressed gene whereas most of the other DR types contain two. There is a considerable gap, probably over 300 kb (D Campbell, 1987, personal communication), before the class III region. The recently discovered HLA class II gene DVβ remains to be placed on the map but is probably near DX (H Inoko, 1987, personal communication).

HLA CLASS III GENES

Conventional linkage analysis established that this region lay between the HLA-D and HLA-A,B,C regions. Although the order of the genes in the class III region has been established, the orientation of the complex vis à vis the neighbouring class II and class I regions has not been determined. It might be predicted that the order is the same as in mice,[101] given the general structural similarities of the two homologous chromosomes.

The class III region has been mapped in detail using sets of overlapping cosmid clones. It contains genes for complement components C2, factor B, and C4.[75] 21-hydroxylase *(CA21H)* is located very close to C4. Some persons contain a tandem duplication of the C4 and 21-OH genes. Hereditary defects in the 21-OH gene result in the disease congenital adrenal hyperplasia.[76] The class III region is separated by several hundred kb from both the class I and class II region (D Campbell, 1987, personal communication).

OTHER GENES LINKED TO THE MHC

Fig 2a indicates the positions of some markers that flank the HLA region. *GLO1*, *PIM1*, and CA21H have homologous locations on mouse chromosome 17. The genes encoding tumour necrosis factor α *(TNFA)* and lymphotoxin *(TNFB)* both map within the murine and human MHC.[77] The two murine genes are directly arranged in tandem, separated by only 1 kb, and lie 70 kb centromeric to the H-2D locus.[78] Both human genes are contained on a single 42 kb cosmid.[79] Filter hybridisation to a series of haplotype deletion mutants derived from a B cell lymphoblastoid line, showed that they reside on 6p either centromeric to HLA-B or centromeric to HLA-DP.[77 79] They may be in an analogous position to the mouse, namely centromeric to HLA-B (H-2D in mouse).[101]

The murine T/t complex is characterised by two large inversions.[26] Chromosomal abnormalities of this type could explain many of the associated phenotypes and although this structural rearrangement may be peculiar to the mouse, it is still plausible that syntenic loci therein have a homologous location on human chromosome 6. The murine *TCP1* gene is contained within the T/t complex, but its human homologue was recently mapped to 6q23→qter in humans.[103]

Spinocerebellar ataxia *(SCA1)* maps within 12 cM of the MHC by conventional linkage analysis.[80 81] Localising unique DNA sequences to specific positions within the genome is invaluable in mapping genetic disorders, especially if RFLP data exist, and assists the detailed analysis of chromosome structure. Levine *et al*[82] recently showed that *AGB6* mapped between *PGK1P2* and HLA, using mutant cell lines that were deleted for *GLO1* but positive for *PIM1*. *AGB6* is 7 cM away from GLO1, which in turn is 6 cM away from HLA.

Anonymous DNA markers on the telomeric side of the HLA complex include *D6S8* and *D6S7*, which are 4 cM and 45 cM away from HLA, respectively.[83] The *BEVI* locus maps to chromosome 6,[84] and it is of interest that the *BKMA* satellite sequence, which is a characteristic of the sex determining chromosome in the heterogametic sex of

other vertebrates, shows an 80% concordance with chromosome 6.[85]

Southern blot analyses of human–rodent cell hybrids identified sequences homologous to a ferritin heavy subunit cDNA probe on at least eight chromosomal locations, one of which mapped to 6p21→cen *(FTHL)*.[86 87] However, only a sequence that mapped to chromosome 11 appeared to be expressed in the hybrids. A family study of sibs affected with hereditary haemochromatosis analysed recombination data which placed the haemochromatosis locus *(HFE)* between HLA–A and HLA–B.[88]

DISEASES ASSOCIATED WITH MHC

As well as haemochromatosis and congenital adrenal hyperplasia there are some 200 diseases that show a statistically significant association, though often weak, with the MHC.[96] The more detailed studies considered insulin dependent diabetes mellitus (IDDM), rheumatoid arthritis (RA), myasthenia gravis, Graves disease, coeliac disease, and narcolepsy. The autoimmune aetiology suspected for many of these diseases implicates the HLA class I or class II genes directly in the susceptibility, although there is no direct evidence for this. A disease such as narcolepsy is not obviously of the autoimmune type, yet it is almost exclusively associated with a particular HLA–DR allele, DR2. A new gene closely associated with the DRβ chain gene may be invoked, although DR2 itself may be one of a number of factors that together cooperate to produce the pathology. Bare lymphocyte syndrome and some types of severe combined immunodeficiency (SCID) are characterised by the absence of low levels of HLA class II and sometimes class I products on lymphocytes. They are caused by lesions in unlinked genes coding for *trans* acting factors which normally promote expression of the HLA–D genes.[89]

The recent report of a patient with combined deficiencies of neuraminidase and 21-hydroxylase lends credibility to the argument that the *NEU* gene lies at 6p21·3 between *HLA–A* and *GLO*.[90]

RECOMBINATION IN THE MHC

Having a closely linked set of markers permits detailed analysis of recombination in the HLA region. There is evidence for recombination hotspots that separate the DQ–DRβ region from the rest of class II, since DX and class III markers do not exhibit a strong association with DQ–DR. A recombination hotspot is defined as a region at which recombination occurs at a frequency higher than the average (for a discussion see reference 1). Different haplotypes* may differ in the location of these hotspots.

The new HLA region map generated by physical techniques agrees more or less with the recombination data (assuming 1 cM is roughly equivalent to 1 mbp). However, now that restriction fragment length polymorphisms (RFLPs) of unique sequences have been described in most regions, it would be valuable to revise the serologically determined recombination frequencies, particularly for DP, where typing has previously been both arduous and inaccurate.

EVOLUTION OF CHROMOSOME 6

As more genetic markers are mapped in human, rodent, and other species, and chromosome banding techniques become more sophisticated, it appears that relatively large linkage groups may be conserved between species.[22] In the primates, chromosomal evolution has been relatively conservative and homologies between the different species are easy to discern. Almost all of the chromosomes of the primates were conserved intact throughout 12 million years of evolution.

Comparison of human chromosome 6 with the homologous chromosome from chimpanzee, gorilla, or orang-utan indicates almost identical banding patterns.[91]

In rodents, on the other hand, there have been obvious chromosomal rearrangements, such that complete linkage groups are not conserved. The HLA linkage group, extending as far as the *GLO* marker, is conserved in man and mouse, as well as other species including rabbits, hares, dogs, and rats. However, *HLA–GLO* linkage is not observed in chickens.[91]

Conclusions

Human chromosome 6 contains one of the most detailed areas of the human gene map to date, the MHC, and its structural analysis will be invaluable in the effort to understand the behaviour of the human genome, particularly recombination, in the context of a detailed molecular map. Moreover, it will help to establish the genetic basis of many HLA associated diseases and provide material for analysing the regulation and function of the MHC in the immune response. The rest of this autosome contains relatively few markers, a situation that will change radically over the next few years.

We wish to thank Dr Peter N Goodfellow for useful discussions and critical reading of this paper, and we are grateful to Kim Richardson for typing the manuscript.

*Haplotype is a term used to refer to the collection of the HLA markers on one chromosome, for example, HLA–A1, B8, DR3 is a common haplotype.

References

[1] Trowsdale J, Young JAT, Kelly AP, *et al.* Structure, sequence and polymorphism in the HLA-D region. *Immunol Rev* 1985; **85**:8–43.

[2] Morton NE, Rao DC, Lindsten J, Hultén M, Yee S. A chiasma map of man. *Hum Hered* 1977;**27**:38–51.

[3] Francke U, Pellegrino M. Assignment of the major histocompatibility complex to a region of the short arm of human chromosome 6. *Proc Natl Acad Sci USA* 1977;**74**:1147–51.

[4] Floyd-Smith GA, de Martinville B, Francke U. A beta-tubulin expressed gene M40 *(TUBB)* is located on human chromosome 6 and two related pseudogenes are located on chromosomes 8 *(TUBBP1)* and 13 *(TUBBP2)*. *Cytogenet Cell Genet* 1985;**40**:629.

[5] Naylor SL, Sakaguchi AY, Shows TB, Grzeschik KM, Holmes M, Zasloff M. Two nonallelic $tRNA_i^{Met}$ genes are located on the q23–q12 region of human chromosome 6. *Proc Natl Acad Sci USA* 1983;**80**:5027–31.

[6] Bender K, Grzeschik KH. Possible assignment of the glyoxalase I (GLO) to chromosome 6 using man-mouse somatic cell hybrids. *Hum Genet* 1976;**31**:341–5.

[7] Weitkamp LR. Linkage of GLO with HLA and Bf. Effect of population and sex on recombination frequency. *Tissue Antigens* 1976;**7**:273–9.

[8] Ziegler A, Fonatsch C, Kompf J. Mapping of the locus for glyoxalase-1 *(GLO1)* on human chromosome 6 using mutant cell lines. *Cytogenet Cell Genet* 1985;**40**:787.

[9] Kömpf J, Schunter F, Wernet P, Ritter H. Linkage between loci for mitochondrial malic enzyme *(ME2)* and coagulation factor XIIIA subunit *(F13A)*. *Hum Genet* 1985;**70**:43–4.

[10] Kömpf J, Ritter H. Linkage between the loci for mitochondrial malic enzyme *(ME2)* and coagulation factor XIIIA subunit *(F13A)*. HGM8. *Cytogenet Cell Genet* 1985;**40**:672.

[11] Chen TR, McMorris FA, Creagan R, Ricciuti F, Tischfield J, Ruddle FH. Assignment of the genes for malate oxidoreductase decarboxylating to chromosome 6 and peptidase B and lactate dehydrogenase B to chromosome 12 in man. *Am J Hum Genet* 1973;**25**:200–7.

[12] Lamm LU, Friedrich U, Bruun Petersen G, *et al.* Assignment of the major histocompatibility complex to chromosome no 6 in a family with pericentric inversion. *Hum Hered* 1974;**24**: 273–84.

[13] Jongsma A, Someren H, van Westerveld A, Hagemeijer A, Pearson P. Localization of genes on human chromosomes by studies of human-Chinese hamster somatic cell hybrids. *Humangenetik* 1973;**20**:195–202.

[14] Creagan R, Tischfield J, Ricciuti F, Ruddle FH. Chromosome assignments of genes in man using mouse-human somatic cell hybrids: superoxide dismutase (Indophenoloxidase-B tetrameric) to chromosome 6. *Humangenetik* 1973;**20**:203–9.

[15] Sparkes RS, Dizikes CG, Grodey WW, *et al.* The human gene for liver arginase *(ARG1)* is assigned to chromosome 6. *Cytogenet Cell Genet* 1985;**40**:750.

[16] Olaisen B, Gedde-Dahl T, Teisberg P, *et al.* A structural locus for coagulation factor XIIIA (F13A) is located distal to the HLA region on chromosome 6p in man. *Am J Hum Genet* 1985;**37**:215–20.

[17] Koch G, Shows TB. A gene on human chromosome 6 functions in assembly of tissue-specific adenosine deaminase isozymes. *Proc Natl Acad Sci USA* 1978;**75**:3876–80.

[18] Valentine WN, Hseih HS, Paglia DE, *et al.* Hereditary hemolytic anaemia associated with phosphoglycerate kinase deficiency in erythrocytes and leukocytes. A probable X-chromosome linked syndrome. *N Engl J Med* 1968;**280**:528–34.

[19] Kozak LP, MacLean GK, Eicher EM. X-linkage of phosphoglycerate kinase in the mouse. *Biochem Genet* 1974;**11**:41–7.

[20] Michelson AM, Bruns GAP, Morton CC, Orkin SH. Characterization of the human phosphoglycerate kinase multigene family. *Am J Hum Genet* 1983;**35**:180A.

[21] Eicher EM, Cherry M, Flaherty L. Autosomal phosphoglycerate kinase linked to mouse major histocompatibility complex. *Mol Gen Genet* 1978;**158**:225–8.

[22] Lalley PA, McKusick VA. Report of the committee on comparative mapping. *Cytogenet Cell Genet* 1985;**40**:536–66.

[23] Fellous M, Boue J, Malbrunot C, *et al.* A five-generation family with sacral agenesis and spina bifida: possible similarities with the mouse T-locus. *Am J Hum Genet* 1982;**12**:465–87.

[24] Willison K, Dudley K, Spurr N, Goodfellow P. Chromosomal assignment of *TCP1* the human homologue of a mouse-t-complex locus. *Cytogenet Cell Genet* 1985;**40**:779–80.

[25] Alper CA, Awdeh ZL, Yunis EJ. Complotypes, extended haplotypes, segregation distortion and disease markers. *Hum Immunol* 1986;**15**:366–73.

[26] Frischauf AM. The T/t complex of the mouse. *Trends Genet* 1985;**2**:100–3.

[27] Chen SH, Donahue RP, Scott CR. Characterisation of phosphoglycerate kinase from human spermatozoa. *Fertil Steril* 1976;**27**:699–701.

[28] Szabo P, Grzeschik KH, Siniscalco M. A human autosomal phosphoglycerate kinase locus maps near the HLA cluster. *Proc Natl Acad Sci USA* 1984;**81**:3167–9.

[29] Tani K, Singer-Sam J, Munns M, Yoshida A. Molecular cloning and structure of an autosomal processed gene for human phosphoglycerate kinase. *Gene* 1985;**35**:11–8.

[30] Gartler S, Riley D, Eddy R, Shows T. A human autosomal phosphoglycerate kinase gene has been assigned to chromosome 19. *Cytogenet Cell Genet* 1985;**40**:635–6.

[31] Beaudet AL, Su TS, O'Brien WB, D'Eustachio P, Barker PE, Ruddle FH. Dispersion of argininosuccinate-synthetase like human genes to multiple autosomes and the X-chromosome. *Cell* 1982;**30**:287–93.

[32] Su TS, Nussbaum RL, Airhart S, *et al.* Human chromosomal assignments for 14 argininosuccinate synthetase pseudogenes: cloned DNAs as reagents for cytogenetic analysis. *Am J Hum Genet* 1984;**36**:954–64.

[33] Young JAT, Trowsdale J. A processed pseudogene in an intron of the HLA-DPβ1 chain gene is a member of the ribosomal protein L32 gene family. *Nucleic Acids Res* 1985;**13**:8883–91.

[34] Anagnou NP, Antonarakis SE, O'Brien SJ, Nienhuis AW. Chromosomal localisation and racial distribution of the polymorphic hDHFR-psi-1 pseudogene. *Clin Res* 1985;**33**:328.

[35] O'Brien SJ, Nash WG, Goodwin JL, Lowry DR, Chang EH. Dispersion of the *ras* family of transforming genes to four different chromosomes in man. *Nature* 1983;**302**:839–42.

[36] McBride OW, Swand DC, Tronick SR, Regional chromosomal localisation of N-*ras*, K-*ras*-1, K-*ras*-2 and *myb* oncogenes in human cells. *Nucleic Acids Res* 1983;**11**:8221–36.

[37] Sakaguchi AY, Zabel BU, Grzeschik KH, *et al.* Regional localization of two human cellular Kirsten ras genes on chromosomes 6 and 12. *Mol Cell Biol* 1984;**4**:989–93.

[38] Shows TB, Eddy KL, Grzeschik KH, Gonzalez A, Villa-Komaroff L. A DNA sequence bearing homology to the human insulin gene is located in the p23→q12 region of chromosome 6. *Cytogenet Cell Genet* 1984;**37**:582.

[39] Owerbach D, Rutter WJ, Cooke NE, Martial JA, Shows TB. The prolactin gene is located on chromosome 6 in humans. *Science* 1981;**212**:815–6.

[40] Farid NR, Noel EP, Sampson L, Russell NA. Prolactin-secreting adenomata are possibly associated with HLA-B8. *Tissue Antigens* 1980;**15**:333–5.

[41] Waker MC, Marandici A, Martin K, Monder C. Genetic control of corticosteroid side chain isomerase activity in the mouse. *Endocrinology* 1983;**112**:924–30.

[42] Gupta C, Goldman A. H-2 histocompatibility region: influence on the murine glucocorticoid receptor and its response. *Science* 1982;**216**:994–6.

[43] Tyan ML. Genetic control of hydrocortisone-induced thymus atrophy. *Immunogenetics* 1979;**8**:177–81.

[44] Erickson RP, Heidel L, Kapur JJ, *et al.* HLA antigens, phytohemagglutinin stimulation, and corticosteroid response. *Am J Hum Genet* 1985;**37**:761–70.

[45] Naylor SL, Chin WW, Goodman HM, Lalley PA, Grzeschik KH, Sakaguchi AY. Chromosome assignment of genes encoding the alpha and beta subunits of the glycoprotein hormones in man and mouse. *Somat Cell Genet* 1983;**9**:757–70.

[46] Hardin JW, Riser ME, Trent JM, Kohler PO. The gonadotropin alpha subunit gene is on human chromosome 18 in JEG cells. *Proc Natl Acad Sci USA* 1983;**80**:6282–5.

[47] Pearson PL, van der Kamp J, Veldtkamp J. Reduced Hageman factor level in 6p– patient. *Cytogenet Cell Genet* 1982;**32**:309.

[48] Niebuhr E, Eiberg H, Schonsboe I. Localisation of human *F12, F13A* and *ACP1*. *Cytogenet Cell Genet* 1985;**40**:714.

[49] Eiberg H, Mohr J. Synteny of plasma-L-fucosidase (FUCA2) and the plasminogen (PLG) system. *Cytogenet Cell Genet* 1984; **37**:460.

[50] Swisshelm K, Dyer K, Sadler E, Disteche C. Localisation of plasminogen gene *(PLG)* to the distal portion of the long arm of human chromosome 6 by *in situ* hybridisation. *Cytogenet Cell Genet* 1985;**40**:756.

[51] Murray JC, Sadler E, Eddy RL, Shows TB, Buetow KH. Evidence for assignment of plasminogen *(PLG)* to chromosome 6 but not chromosome 4. *Cytogenet Cell Genet* 1985;**40**:709.

[52] Hamaguchi H, Yamada MM, Yamakawa K, *et al.* Studies on the linkage relationship between a new polymorphic membrane polypeptide and HLA loci. *Cytogenet Cell Genet* 1984;**37**:485.

[53] Rettig WJ, Murty VVVS, Dracopoli NC, Mattes MJ, Chaganti RSK, Old LJ. Human chromosome 6 controls A42 (gp65) and MG2 (gp200/100) cell surface expression in inter-species hybrids. Expression of A42 but not MG2, also requires the presence of human chromosome 14. *Cytogenet Cell Genet* 1985;**40**: 732.

[54] Knowles BB, Mausner R, Aden DP. Preliminary characterization of human cell surface molecules controlled by human chromosomes 6 and 7. *Cytogenet Cell Genet* 1978;**22**:250–4.

[55] Fellous M, Billardon C, Dausset J, Frézal J. Genetique - linkage probable entre les locus HLA et P. *CR Acad Sci [D] (Paris)* 1971;**272**:3356–9.

[56] Nielsen LS, Mohr J, Eiberg H. Data concerning the linkage relationship of the HLA and P systems. *Cytogenet Cell Genet* 1984;**37**:555.

[57] Pellegrino MA, Curry RA, Pellegrino AG, Hoch JA. Linkage between the B-cell receptor for monkey red blood cells and HLA. *Immunogenetics* 1975;**2**:543–9.

[58] Schroder J, Suomalainen HA, Nikinmaa B, Enns J, Herzenberg LA, Sussman MH. Chromosomal assignment of genes coding for human cell surface antigens with mouse/human lymphocyte hybrids. *Cytogenet Cell Genet* 1984;**37**:577–8.

[59] Walter P, Green S, Greene G, *et al.* Cloning of the human estrogen receptor cDNA. *Proc Natl Acad Sci USA* 1985;**82**: 7889–93.

[60] Gosden JR, Middleton PG, Rout D. Localisation of the human oestrogen receptor gene to chromosome 6q24–q27 by *in situ* hybridisation. *Cytogenet Cell Genet* 1986;**43**:218–20.

[61] Rashidibaigi A, Langer JA, Jung V, *et al.* The gene for the human immune interferon receptor is located on chromosome 6. *Proc Natl Acad Sci USA* 1986;**86**:384–8.

[62] Gozes I, Avidor R, Yahav Y, Katznelson D, Croce CM, Huebner K. The gene encoding vasoactive intestinal peptide is located on human chromosome 6p21–6qter. *Hum Genet* 1987; **75**:41–4.

[63] Nagarajan L, Louie E, Tsujimoto Y, Balduzzi PC, Heubner K, Croce CM. The human c-*ros* gene (ROS) is located at chromosome region 6q16→6q22. *Proc Natl Acad Sci USA* 1986;**83**: 6568–72.

[64] Nagarajan L, Louie E, Tsujimoto Y, ar-Rushdi A, Heubner K, Croce CM. Localisation of the human *pim* oncogene *(PIM)* to a region of chromosome 6 involved in translocations in acute leukaemias. *Proc Natl Acad Sci USA* 1986;**83**:2556–60.

[65] Selten G, Cuypers TH, Boelens W, *et al.* The primary structure of the putative oncogene *pim*-1 shows extensive homology with protein kinases. *Cell* 1986;**46**:603–11.

[66] Semba K, Yamanishi Y, Nishizawa M, *et al.* Location of the c-yes gene on the human chromosome and its expression in various tissues. *Science* 1985;**227**:1038–40.

[67] Dalla-Favera R, Franchini G, Martinotti S, Wong-Staal F, Gallo RC, Croce CM. Chromosomal assignment of the human homologues of feline sarcoma virus and avian myeloblastosis virus *onc* genes. *Proc Natl Acad Sci USA* 1982;**79**:4714–7.

[68] Harper MF, Franchini C, Love J, Simon MI, Gallo RC, Wong-Staal F. Chromosomal sublocalisation of human c-*myb* and c-*fes* cellular *onc* genes. *Nature* 1983;**304**:169–71.

[69] Zabel BU, Naylor SL, Grzeschik KH, Sakaguchi AY. Regional assignment of human protooncogene c-*myb* to 6q21–qter. *Somatic Cell Mol Genet* 1984;**10**:105–8.

[70] Pelicci PG, Lanfrancone L, Brathwaite MD, Wolman SR, Dalla-Favera R. Amplification of the c-*myb* gene in a case of human acute myelogenous leukaemia. *Science* 1984;**224**: 1117–21.

[71] Rabin M, Birnbaum D, Wigler M, Ruddle FH. *MCF3* oncogene mapped to region on chromosome 6 associated with malignant transformation. *Am J Hum Genet* 1985;**37**:36A.

[72] Berger R, Bloomfield CD, Sutherland GR. Report of the committee on chromosome rearrangements in neoplasia and on fragile sites. *Cytogenet Cell Genet* 1985;**40**:490–535.

[73] Goate AM, Cooper DN, Hall C, Leung TKC, Solomon E, Lim L. Localisation of a human heat-shock HSP70 gene sequence to chromosome 6 and detection of two other loci by somatic-cell hybrid and restriction fragment length polymorphism analysis. *Hum Genet* 1987;**75**:123–8.

[74] Calabi F, Milstein C. A novel family of human major histocompatibility complex-related genes not mapping to chromosome 6. *Nature* 1986;**323**:540–3.

[75] Hatch JA, Atkinson JP, Suarez BK, Dykman TR. Evaluation of linkage of the human C3-b/C4b receptor to HLA. *J Immunol* 1984;**132**:2168–9.

[76] Dupont B, Oberfield SE, Smithwick EM, Lee TD, Levine LS. Close genetic linkage between HLA and congenital adrenal hyperplasia (21-hydroxylase deficiency). *Lancet* 1977;**ii**: 1309–11.

[77] Nedwin GE, Naylor SL, Sakaguchi AY, *et al.* Human lymphotoxin and tumour necrosis factor genes: structure, homology and chromosomal localisation. *Nucleic Acids Res* 1985;**13**: 6363–73.

[78] Muller U, Jongeneel CV, Nedospasov SA, Lindahl KF, Steinmetz M. Tumor necrosis factor and lymphotoxin genes map close to H–2D in the mouse major histocompatibility complex. *Nature* 1987;**325**:265–7.

[79] Spies T, Morton CC, Nedospasov SA, Fiers W, Pious D, Strominger JL. Genes for the tumor necrosis factors α and β are linked to the human major histocompatibility complex. *Proc Natl Acad Sci USA* 1986;**83**:8699–702.

[80] Jackson JF, Currier RD, Terasaki PA, Morton NE. Spinocerebellar ataxia and HLA linkage risk prediction by HLA typing. *N Engl J Med* 1977;**296**:1138–41.

[81] Morton NE, Lalouel JM, Jackson JF, Currier RD, Yee S. Genetics of spinocerebellar ataxia (SCA). *Am J Med Genet* 1980;**6**:251–7.

[82] Levine F, Erlich H, Mach B, Leach R, White R, Pious D. Deletion mapping of HLA and chromosome 6p genes. *Proc Natl Acad Sci USA* 1985;**82**:3741–5.

[83] Leach R, DeMars R, Hasstedt S, White R. Construction of a map of the short arm of human chromosome 6. *Proc Natl Acad Sci USA* 1986;**83**:3909–13.

[84] Cohen JC, Murphey-Corb M. Targeted integration of baboon endogenous virus in the *BEVI* locus on human chromosome 6. *Nature* 1983;**301**:129–32.

[85] Kiel-Metzger K, Warren G, Wilson GN, Erickson RP. Evidence that the human Y chromosome does not contain clus-

tered DNA sequences (BKM) associated with heterogametic sex determination in other vertebrates. *N Engl J Med* 1985;**313**: 242–4.

[86] Cragg SJ, Drysdale J, Worwood M. Genes for the 'H'-subunit of human ferritin are present on a number of human chromosomes. *Hum Genet* 1985;**71**:108–12.

[87] McGill JR, Moore CM, Sakaguchi WY, *et al.* Ferritin genes map to multiple sites in the human genome. *Cytogenet Cell Genet* 1985;**40**:696–7.

[88] Edwards CQ, Griffen LM, Dadone MM, Skolnick MH, Kushner JP. The locus for hereditary hemochromatosis maps between *HLA-A* and *HLA-B*. *Cytogenet Cell Genet* 1985;**40**:620.

[89] de Preval C, Lisowska-Grospierre B, Loche M, Griscelli C, Mach B. A *trans*-acting class II regulatory gene unlinked to the MHC controls expression of HLA class II genes. *Nature* 1985; **318**:291–3.

[90] Harada F, Nishimura Y, Suzuki K, *et al.* The patient with combined deficiency of neuraminidase and 21-hydroxylase. *Hum Genet* 1987;**75**:91–2.

[91] Klein J. *Natural history of the major histocompatibility complex*. London: Wiley-Interscience 1986:756–61.

[92] Buc-Caron MH, Gachelin G, Hofnung M, Jacob F. Presence of a mouse embryonic antigen on human sperm. *Proc Natl Acad Sci USA* 1974;**71**:1730–1.

[93] Weitkamp IR, Moss AJ. The long QT (Romano-Ward) syndrome locus, *LQT*, is probably linked to the HLA loci. *Cytogenet Cell Genet* 1985;**40**:775.

[94] Hansen JA, Dupont B, L'Esperance PL, Good RA. Congenital neutropaenia: abnormal neutrophil differentiation associated with HLA. *Immunogenetics* 1977;**4**:327–32.

[95] Shows TB, McAlpine PJ, Miller RL. The 1983 catalogue of mapped human genetic markers and report of the nomenclature committee. *Cytogenet Cell Genet* 1984;**37**:340–93.

[96] Tiwari JL, Terasaki PI, *HLA and disease associations*. New York: Springer-Verlag, 1985.

[97] Baserga R, Potten C, Ming PML. Cell fusion and introduction of new information into temperature sensitive mutants of mammalian cells. In: Nicolini C, ed. *Cell growth*. New York: Plenum Press, 1982:69–81.

[98] Hardym DA, Bell JI, Long EO, Lindstein T, McDevitt HO. Mapping of the class II region of the human major histocompatibility complex by pulsed-field gel electrophoresis. *Nature* 1986;**323**:453–5.

[99] Ragoussis J, von der Blick A, Trowsdale J, Ziegler A. Mapping of HLA genes by pulsed-field gradient electrophoresis. *FEBS Lett* 1986;**204**:1–4.

[100] Lawrance SK, Smith CL, Srivastava R, Cantor CR, Weissman SM. Megabase-scale mapping of the HLA gene complex by pulsed-field gel electrophoresis. *Science* 1987;**235**:1388–90.

[101] Müller U, Stephan D, Philippsen P, Steinmetz M. Orientation and molecular map position of the complement genes in the mouse MHC. *EMBO J* 1987;**6**:369–73.

[102] McCarrey JR, Thomas K. Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. *Nature* 1987;**326**:501–5.

[103] Willison K, Kelly A, Dudley K, *et al.* The human homologue of the mouse t-complex gene *TCP1* is located on chromosome 6 but is not near the HLA region. *EMBO J* 1987;**6**:1967–74.

Correspondence and requests for reprints to V Cunliffe, Human Immunogenetics Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX.