Experientia 42 (1986), Birkhäuser Verlag, CH-4010 Basel/Switzerland

Immunogenetics

by K. Bender

Institut für Humangenetik and Anthropologie der Universität Freiburg i. Br., Albertstrasse 11, D-7800 Freiburg (Federal Republic of Germany)

Summary. The 1985 Catalog of Mapped Genes (Human Gene Mapping 8; 33) has been used to pick out the known, immunologically important genes; these are then discussed in the following order: 1. genes controlling organs, tissues and cells of the immune apparatus, 2. genes determining 'self' structures, 3. genes determining the structures of immunological specificity, 4. genes determining substances with immunoregulatory and effector properties. The symbols for the genes and the biological functions of their products are explained. The genetics of the ABO blood groups, of the HLA-system and of antibody formation are given in rather more detail.

Key words. Genes with immunological function; human gene map; ABO; HLA; immunoglobulins; immunoglobulin supergene family; interleukin; interferon; complement.

Introduction

Immunogenetics may be regarded as the science that studies the genetic structures of the self-non-self discrimination mechanism, called the immune system. The immune system consists particularly of leukocytes (i.e. lymphocytes, macrophages, and granulocytes) and a series of macrophage-related cells (i.e. dendritic cells, epidermal Langerhans cells, and specialized epithelial cells). These occur in organized tissues and organs (bone marrow, lymph nodes, spleen, Peyer's patches of the intestine, tonsils, and thymus); a very substantial fraction of the leukocytes, however, comprises, a recirculating pool of cells found in the blood and lymph vessels.

The ability of an organism to respond to virtually any antigen is achieved primarily by the lymphocytes. An adult human possesses more than 10^6 different lymphocyte specificities (clones), each bearing receptors for distinct antigens.

In fact, two types of immunologic responses have been found, mediated by two classes of lymphocytes: *humoral immunity* occurs when B lymphocytes (bursa Fabricii-derived or bone-marrow derived lymphocytes) produce immunoglobulins called antibodies, which either inactivate antigens (foreign structures) or mark them for destruction by certain effector cells or effector molecules; *cell-mediated immunity* depends on T lymphocytes (thymus-derived lymphocytes) of which three functionally different subpopulations are known, T_C (cytotoxic), T_H (helper), and T_S (suppressor) cells. T_C cells act as killer cells that attack virus-infected and cancerous cells directly, but also cause the rejection of allotransplants; T_H and T_S cells act as regulator cells in amplifying or suppressing antibody production.

Whereas B cells can recognize foreign (mostly soluble) antigens directly via their specific receptors (i.e. membrane-bound immunoglobulins), T cells can see foreign antigens via their specific receptors (T cell receptors) only in context with certain 'self' molecules, namely those encoded in the major histocompatibility complex (MHC). The T cell response is thus *MHC-restricted.* There exists still another class of lymphocytes, natural killer (NK) cells, which recognize and kill tumor cells which (as is frequently observed) have lost their 'self' (MHC) antigens[26].

A large number of genes (or even complex gene regions) for the determination of various immunogenetic characters have been identified so far, referring to 1) the organs and the histological and cellular architecture of the immunological apparatus, 2) the 'self' markers (erythrocyte and histocompatibility antigens), 3) the clonotypic B cell and T cell receptors, and 4) the molecules with immunoregulatory and effector properties. The table lists those of the genes which, besides being well defined, have also been mapped to certain chromosomes. The list can be considered to be quite representative of our current immunogenetic knowledge, so in order to achieve something like completeness, I thought it convenient to follow the presentation in that catalog.

1. Genes controlling organs, tissues and cells of the immune apparatus

For its adequate functioning, the immune system requires an intact anatomical and histological architecture. A number of gene defects have been identified which cause histopathological disturbances resulting in more or less severe immunodeficiency diseases (table):

DGS (DiGeorge syndrome). The deletion of the DGS gene results in a congenital anomaly in the development of derivatives of the 3rd and 4th pharyngeal pouches. Besides ear, nose, mouth and aortic deformations the major immunogenetic consequences of the gene defect lie in the absence of the thymus, the central organ of T lymphocyte differentiation.

IMD1 (immunodeficiency 1; Bruton type of agammaglobulinemia). This X-chromosome linked disease is characterized by a nearly total loss of immunoglobulins due to the absence of B lymphocytes. The T cell function is, however, normal.

IMD2 (immunodeficiency 2; Wiskott-Aldrich syndrome). This is another fatal X-linked disease, perhaps due to a defect in certain immune cell interactions.

IMD3 (immunodeficiency 3). The male patients with this X-linked disease are deficient in IgA and IgG antibodies but have highly elevated levels of IgM antibodies. The lymphoid tissues show disorganization of the follicular architecture.

Catalog of immunologically important genes (adapted from Human Gene Mapping 8[33])

Gene Symbol	Assigned to chromosome	Marker name
1. *Genes controlling organs, tissues and cells of the immune apparatus (see section 1)*		
DGS	22q11	DiGeorge syndrome
IMD1	X	Bruton agammaglubolinemia
IMD2	X	Wiskott-Aldrich syndrome
IMD3	X	Immunodeficiency 3
IMD4	X	Immunodeficiency 4
ADA	20q13.2- qter	Adenosine deaminase
NP	14q13.1	Nucleoside phosphorylase
CGD	Xpter – p21	Chronic granulomatous disease
CD2	4	Sheep red blood cell receptor (SRBCR)
CD3	11	T3 antigen (on mature T cells); OKT3
CD4	12	T4 antigen (on T helper cells); OKT4
CD5	11	Leukocyte antigen 1; OKT1
CD7	17	Leukocyte antigen 7
CD8	2pter – p12	T8 antigen (on T_C and T_S cells); OKT8
CD9	12p	Monocyte antigen
CD15	11	Granulocyte antigen
TFRC	3q26.2 – qter	Transferrin receptor; OKT9
C3BR	6p21.3	Complement 3b receptor
C3DR	6p21.3	Complement 3d receptor
2. *Genes determining 'self' structures, (see section 2)*		
ABO	9q34	ABO blood groups
CO	2	Colton blood group
FY	1p21 – q23	Duffy blood group
H	19	H blood group antigen
JK	2	Kidd blood group
KEL		Kell blood group
LE	19	Lewis blood group
LU	19	Lutheran blood group
MIC2	Xpter – p22.32 and Yp	Antigen identified by monoclonal antibody 12 E7
MNS	4q28 – q31	MNSs blood groups
P1	22	P blood group
RD	1pter – p22.1	Radin blood group
RH	1p36.2 – p34	Rhesus blood group
SC	1p36.2 – p22.1	Scianna blood group
SE	19	ABH secretion
SF	4	Stoltzfus blood group
XG	Xpter – p22.3	Xg blood group
B2M	15q22	beta-2-Microglobulin
HLA-A	6p21.3	HLA-A heavy chain
HLA-B	6p21.3	HLA-B heavy chain
HLA-C	6p21.3	HLA-C heavy chain
HLA-DP1A	6p21.3	HLA-DP alpha chain
HLA-DP1B	6p21.3	HLA-DP beta chain
HLA-DQ1A	6p21.3	HLA-DQ alpha-1 chain
HLA-DQ2A	6p21.3	HLA-DQ alpha-2 chain
HLA-DQ1B	6p21.3	HLA-DQ beta-1 chain
HLA-DQ2B	6p21.3	HLA-DQ beta-2 chain
HLA-DR1A	6p21.3	HLA-DR alpha chain
HLA-DR1B	6p21.3	HLA-DR beta-1 chain (HLA-DR1 – HLA-DRw 14 antigens)
HLA-DR2B	6p21.3	HLA-DR beta-2 chain (HLA-DRw52 and -DRw53 antigens)
3. *Genes determining the structures of immunological specificity (see section 3)*		
IGKV	2p12	Immunoglobulin (Ig) light chain variable region, kappa
IGLV	22q11	Ig light chain variable region, lambda
IGHV	14q32.3	Ig heavy chain variable region
IGKJ	2p12	Ig kappa light chain, joining region
IGLJ	22q11	Ig lambda light chain, joining region
IGHJ	14q32.3	Ig heavy chain, joining region
IGD	14q32.3	Ig heavy chain, diversity region
IGKC	2p12	IG kappa light chain, constant part
IGLC	22q11	Ig lambda light chain, constant part
IGHD	14q32.3	Ig heavy chain, delta constant part
IGHG1	14q32.3	Ig heavy chain, gamma 1 constant part
IGHG2	14q32.3	Ig heavy chain, gamma 2 constant part
IGHG3	14q32.3	Ig heavy chain, gamma 3 constant part
IGHG4	14q32.3	Ig heavy chain, gamma 4 constant part
IGHA1	14q32.3	Ig heavy chain, alpha 1 constant part
IGHA2	14q32.3	Ig heavy chain, alpha 2 constant part
IGHE	14q32.3	Ig heavy chain, epsilon constant part
TCRA	14pter – q21	T cell receptor, alpha polypeptides
TCRB	7q32 or q35	T cell receptor, beta polypetides
TCRG	7p15	T cell receptor, gamma polypeptides
4. *Genes determining substances with immunoregulatory and effector properties (see section 4)*		
IL1	2q13 – q21	Interleukin 1
IL2	4q26 – q28	Interleukin 2
IL2R	10p15 – p14	Interleukin 2 receptor
IFN2	5p	Interferon 2 (fibroblast)
IFNA	9pter – p13	Interferon, alpha (leukocyte)
IFNB	9p24 – p13	Interferon, beta (fibroblast)
IFNB3	2p23 – qter	Interferon, beta 3 (fibroblast)
IFNG	12q24.1	Interferon, gamma
IFNAR	21q21 – qter	Receptor for interferon alpha
IFNBR	21	Receptor for interferon beta
IFNGR	18	Receptor for interferon gamma
BF	6p23.1	Properdin factor B
C1QB	1p	Complement component 1q
C2	6p23.1	Complement component 2
C3	19p13.3 – p13.2	Complement component 3
C4A	6p23.1	Complement component 4A
C4B	6p23.1	Complement component 4B
C5		Complement component 5
C6	linked to C7	Complement component 6
C7	linked to C6	Complement component 7
C8A	1p36.2 – p22.1	Complement component 8A
C8B	1	Complement component 8B
C9		Complement component 9

IMD4 (immunodeficiency 4; thymic epithelial hyperplasia). This is also an X-linked disease, which is characterized by a malignant proliferation of lymphocytes with concomitant depletion of the thymic-dependent areas of lymph nodes and spleen.

ADA (adenosinedeaminase) deficiency. ADA is an enzyme which shows both polymorphism and deficiency. ADA deficiency is the cause of one of several forms of SCID (severe combined immune deficiency disease), in which there is dysfunction of both B and T lymphocytes

with impaired cellular immunity and decreased antibody production.

NP (nucleoside phosphorylase) deficiency. In a way similar to ADA deficiency, NP deficiency results in SCID, with, however, mostly defective T cell immunity. In that respect the clinical picture resembles that of AIDS (acquired immune deficiency syndrome), but the AIDS virus (LAV/HTLV-III) selectively attacks the T_H subset of T lymphocytes[22].

CGD (chronic granulomatous disease). This severe X-linked disease is characterized by inadequate granulocyte function on the basis of either an enzyme defect or a cell membrane defect. Interestingly, the disease is also associated with no-expression of the erythrocyte antigen Kell due to a defect in the Kell precursor substance XK[31].

Besides these pathological immunogenetic characters, I add here (perhaps somewhat arbitrarily) some histological markers which have only recently been defined by means of monoclonal antibodies and can now be used to differentiate the T lymphocyte subsets. These markers have been designated CD (cluster of differentation antigens) with a current number or, in some cases, also MIC (monoclonal antibody identified cell structures) with a current number (table). They play a crucial role in lymphocyte interaction and triggering[1].

CD2 (synonymous T11, OKT11, Leu5) is the so-called sheep red blood cell receptor (SRBCR) and occurs on all T cells. (It is of practical importance for the separation of T cells from B cells[34]); CD3 (syn. T3, OKT3, Leu4) is present on all mature T cells and is intimately associated with the T cell receptor[1]; CD4 (syn. T4, OKT4, Leu 3a) is a marker of $T_H$cells; CD5 (syn. T1, OKT1, Leu1) as well as CD7 (syn. Leu 9) are on all T cells; CD8 (syn. T8, OKT8, Leu 2a) is a marker of T_C and T_S cells; CD9 (syn. MIC3) is a monocyte antigen; CD15 is a granulocyte antigen; TFRC (syn. OKT9, 5E9) is the transferrin receptor, expressed on only highly proliferating cells; the complement 3 receptors (C3BR and C3DR) are markers of B lymphocytes.

Figure 1. a) The human erythrocyte and its most polymorphic antigens and a schematic demonstration of the genetics of the ABO pathway. b) Allele frequencies in the European population.

Gene	Alleles	Gene	Alleles
ABO	$A_1 = 0.23$	FY	a = 0.41
	$A_2 = 0.06$		b = 0.59
	B = 0.08	JK	a = 0.51
	O = 0.63		b = 0.49
MNS	Ms = 0.30	P1	p = 0.55
	MS = 0.25		p = 0.45
	Ns = 0.37	XG	a+ = 0.64
	NS = 0.08		a− = 0.36
RH	CDe = 0.41	KEL	K = 0.04
	$C^wDe = 0.01$		k = 0.96
	cDE = 0.14	CO	a = 0.96
	cde = 0.41		b = 0.04
	rare combinations = 0.03	LU	a = 0.04
			b = 0.96

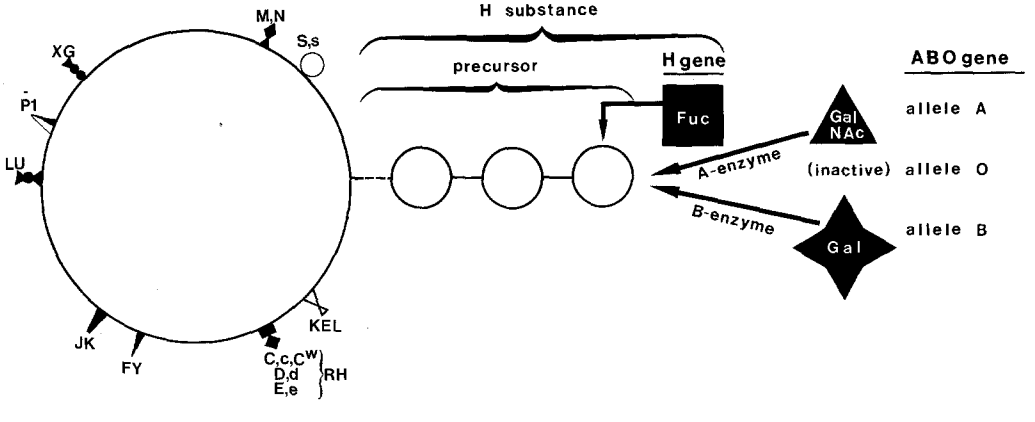
2. Genes determining 'self' markers

It is a trivial but fundamental fact that the recognition of 'foreign' is only possible after the definition of 'self'. The human genome carries many genes for the characterization of self (table), and they all have in common that they are rather polymorphic with allele frequencies resulting in a high degree of heterozygosity. One can divide the genes into two groups. Those in one group are expressed primarily or exclusively on erythrocytes; the others are localized in a complex chromosome region, called the major histocompatibility complex.

2.1 Polymorphic genes encoding erythrocyte antigens

Erythrocytes have a rather short life span of about 100 days. Their most obvious function is in oxygen transport. They are, however, also equipped with a large number of different antigenic structures. The table gives only those characters usually determined in most laboratories for paternity testing or anthropological studies. One cell carries up to 10^6 ABO antigenic sites[21], and the antigenic sites of the other characters are hardly fewer. But it is not only the multiplicity of different antigens and their density, it is also the polymorphism of the genes that code for them, and the allele frequencies which give rise to a high degree of heterozygosity (fig. 1).

Sheep red cells possess an antigenic system which influences the potassium concentration of the erythrocytes[10]. This suggests that the primary role of red cell antigens might be connected with the membrane structure, and only secondarily be antigenic. However, most researchers are inclined to see in the multiplicity of the gene and their polymorphism an important basic mechanism of 'self'-definition; the more heterozygous an individual is with respect to his blood-group genes, the more unique his



erythrocyte surface will become. An invading pathogen will hardly be able to achieve molecular mimicry[11].

A most interesting hypothesis in favor of the immunological importance of (at least) the ABO blood-group system was advanced by Vogel et al.[48]. Comparing the different frequencies of the alleles A, B, and O between human races and populations, they recognized in most European countries a rather uniform O allele distibution (0.60), but also a remarkable surplus of the O allele (0.75–0.90) in certain islands and isolates (Iceland, Ireland, Corsica, Sardinia, the Basque Provinces, some valleys of Switzerland). Vogel et al.[48] asked the question whether these isolates still represent the primordial European O frequency, which was reduced in most other countries by the plague epidemics of the late Middle Ages. The isolated populations were saved from the pestilence. In fact, it has been shown that *Pasteurella pestis* carries a membrane antigen which is serologically related to the so-called H substance (i.e. the main antigenic structure of blood group O cells, see below). Thus, O type individuals might have been less able than people with A, B, or AB to recognize the pathogen as foreign, and it was perhaps predominantly these individuals who died.

Much could be told about the different blood groups, but I shall confine myself to the ABO blood group and for the other red cell antigens I refer the reader to the excellent text books of Race and Sanger[38] and Spielmann and Kühnl[42].

The ABO blood-group system was not only the first to be discovered but is also the most important one in medicine (transfusion, transplantation). Biochemically, ABO structures are carbohydrate chains mostly attached to membrane lipids. Two transferases in turn attach a certain sugar residue to a carbohydrate moiety, designated precursor substance. The one transferase is coded by the H gene (table, fig. 1) and converts the precursor into the so-called H substance (H refers to human). The H substance is the substrate of the ABO gene coded transferases: in the case of allele A a specific N-acetylgalactosaminyl-transferase (briefly, A-enzyme) will add 'GalNAc'; in the case of allele B a specific galactosyl-transferase (briefly, B-enzyme) will attach 'Gal'; in the case of allele O the H substance will stay unchanged.

Yoshida[53] was able to demonstrate that allele O is not silent but codes for an enzymatically inactive protein with properties enabling it to cross-react with rabbit antibody raised against the purified A-enzyme. The presence of this cross-reacting material (CRM) in the plasma of A or B group individuals identifies them with certainty as being heterozygous AO or BO[53]. (Serologically, homozygous AA and BB persons cannot be differentiated from heterzygous AO and BO).

Interestingly, the precursor is also the substrate of another transferase coded by the Lewis (LE) gene (table). The complicated and not yet completely understood interactions between the LE and another related gene, that of the so-called secretor (SE; table) phenotype, will not be discussed here (for details see Watkins[49] and Race and Sanger[38]). But it should be mentioned that all the genes controlling the conversion of the precursor substance (H, LE, SE) are located on chromosome 19 (table).

A more recent research subject is the biochemistry of the A subgroups A_1 and A_2 as analyzed mostly by Yoshida[21, 54, 55]. The A_1-enzyme; converts all the H substance to A, the A_2-enzyme, however, only one third, keeping most of the H substances unchanged. The A_2-enzyme is thus a subactive N-acetylgalactosaminyl-transferase. In fact, the A_2 blood-group is recognized by just two subsequent positive reactions, with an Anti-A reagent and with an Anti-H reagent. (The typing reagents are mostly lectins from seeds of the plant family *Papilionaceae;* it is suggested that the biological role of these lectins is the retention of nitrogen fixing bacteria).

An interesting fact is the presence of the so-called 'natural' antibodies, i.e. the Anti-A in B individuals, the Anti-B in A persons and the Anti-A+B in O people. Newborns do not possess isoagglutinins. But during the first year of life immunization will occur due to invading bacteria (mostly of the digestive tract) which carry antigens on their membranes which are serologically related to the ABH substances. Interestingly, this type of latent immunization results in antibodies of the macromolecular IgM type. Only an incompatible blood transfusion will induce immune antibodies of the IgG class.

2.2 Genetics of the major histocompatibility complex (MHC)

Many genes are involved in making the erythrocyte surface unique, whatever the biological significance of this individuality may be. The genes have 2–4 alleles and are mostly located on different chromosomes (table).

The genomes of all higher vertebrates comprise, however, still another system, composed of a number of different but closely linked, and multi-allelic genes producing a similar grade of inter-individual variability in all nucleated somatic cells. The gene compound has been given the name 'major histocompatibility complex' (MHC), because it strongly influences graft survival or rejection; its biological role, however, is the surveillance of the body's integrity by governing basic mechanisms of immune responses. The MHC of man is called the HLA system (Human Leucocyte A system). Its cluster of genes is located on the short arm of chromosome 6 (table). The story of the detection and unravelling of the organization, structure and function of this 'supergene family' belongs to the highlights of science[6, 8, 14, 18]. The following consideration will concentrate on only a few outstanding facts.

Figure 2 presents the current genetic concept of the HLA system. (Detailed considerations of the MHCs of other vertebrate species are given by Götze[23].) The chromosome region comprises about 2–3 centimorgans (0.08% of the genome) and includes three classes of genes: class I (HLA-A, HLA-B, HLA-C), class II (HLA-DR, HLA-DQ, HLA-DP), class III (C2, C4A, C4B, BF). The class I antigens are integrated membrane proteins composed of two noncovalently associated molecules: a highly polymorphic glycolysated heavy chain (with 341, 338 and 342 amino acids for, respectively, the HLA-A, HLA-B and HLA-C polypeptides) encoded in the major histocompatibility complex on chromosome 6 and an invariant β_2 microglobulin (B2M) polypeptide (99 amino acids) encoded on chromosome 15.

These antigens are expressed on all nucleated cells and have the role of self-definition ('identity card of the entire organism')[18] or, in immunological terms, the role of re-

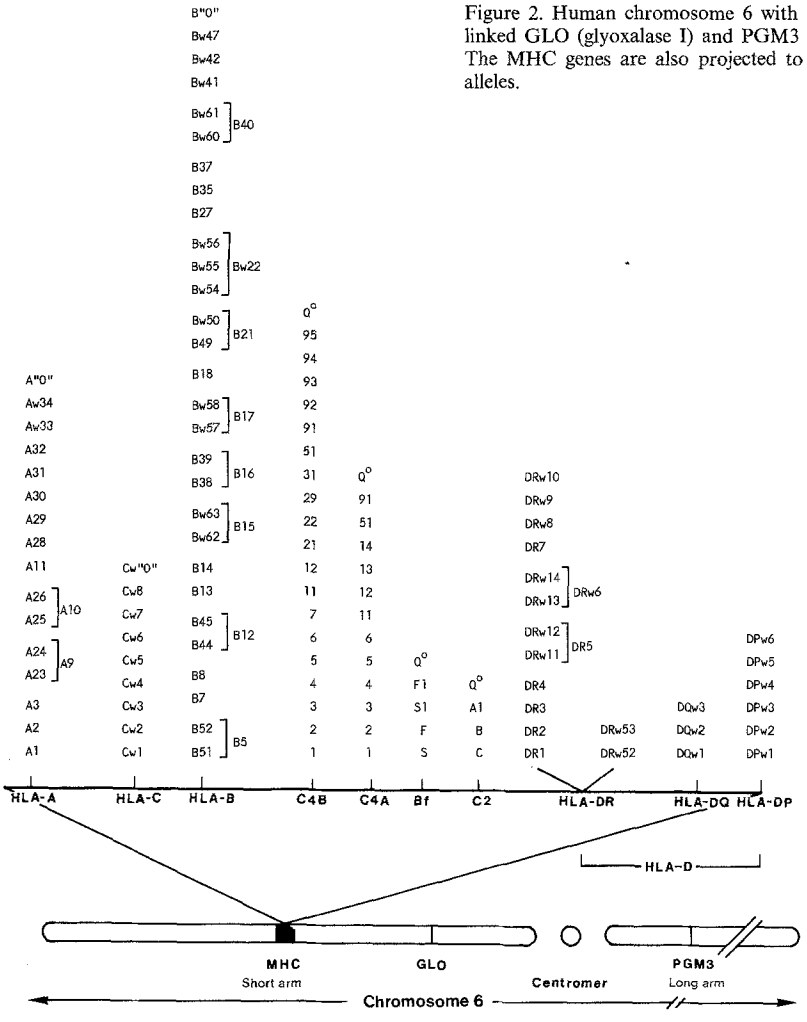


Figure 2. Human chromosome 6 with the MHC gene cluster and the linked GLO (glyoxalase I) and PGM3 (phosphoglucomutase 3) genes. The MHC genes are also projected to show details on their multiple alleles.

stricting the T_C cell activity to only altered self cells: Thus T cells will only kill those altered (virus-infected) cells which share the same class I antigens. (The fact that T_C cells can also kill allotransplantated cells, i.e. cells with non-identical class I antigens, is considered not to be against the rule but to be due to cross-reaction[20]).

The class II antigens (also transmembrane proteins) are dimers of one α and one β chain, each 229 amino acids in length[28, 52], but they are confined to macrophages, B lymphocytes and the series of macrophage-related cells (see Introduction), i.e. cells coming first into contact with foreign structures.

The human genome comprises at least four different α genes (HLA-DR_α, HLA-$DQ_{\alpha 1}$, HLA-$DQ_{\alpha 2}$ and HLA-DP_α) and at least five different β genes (HLA-$DR_{\beta 1}$, HLA-$DR_{\beta 2}$, HLA-$DRQ_{\beta 1}$, HLA-$DQ_{\beta 2}$, HLA-DP_β), all located within the HLA region on chromosome 6 (MHC). The respective polypeptides combine to form several HLA-DR, HLA-DQ and HLA-DP dimers (see below fig. 6). The function of the class II antigens lies in a (loose) binding of foreign structures, followed by the presentation of the complex of self-molecule plus foreign antigen to the T cells with the appropriate T cell receptor. The specific T cell will thus be stimulated to perform its immune responses.

The class III genes synthesize complement proteins, i.e. effector molecules participating in the destruction of foreign particles marked by antibody binding. Most remarkable is the polymorphism of the C4 genes[32].

The first surprise was to see so many genes, all involved in the immune response, concentrated in one short chromosome segment. The second surprise was the extreme polymorphism of most of the HLA-genes, surpassing that of all other known loci; moreover each allele has a low frequency in the population (0.01–0.20) and differs from its counterparts in at least 10–40 sites[5]. Still more surprising, protein sequence analyses[17, 30, 37, 44] have shown that the amino acid substitutions are not scattered all over the molecule but are confined to certain segments (fig. 3). That raises the question of how such mutational events can happen. It is now generally assumed that rather than point mutations (as occur in most other genes) a special kind of event must have occurred, namely interallelic or even intergenic exchanges of genetic material (term: gene conversion[13, 39]).

The close linkage of the MHC genes implies that a given arrangement is generally inherited 'en bloc' (fig. 4).

Such a complex unit of inheritance is called a haplotype. Because of the many alleles at each single locus a very large number of different haplotypes exists. Surprisingly,

a few haplotypes are rather frequent. In the European population for instance, the arrangement A1, Cw7, B8, DR3, C2[1], C4A[0], C4B[1], Bf[S] has a frequency of about 7%. (It is also present in the pedigree of fig. 4). The reason for the over-representation of certain haplotypes (term: linkage disequilibrium) has been widely discussed[7, 12, 14]. They can hardly be relics of the founder population, since the observed *crossing over* rate is sufficient to disrupt all founder haplotypes. It is more likely that these 'optimally composed' arrangements carried some selective advantage, at least in former times. Interestingly enough, carriers of the 'A1, Cw7, B8, ...' haplotype under present day conditions have a higher risk of acquiring a number of late-onset immunopathological diseases, apparently due to an excessive immune responsiveness. ('Either one is not able to respond to pathogens and dies of infection, or one has too much surveillance and develops autoimmune diseases')[16]. The association of certain HLA alleles

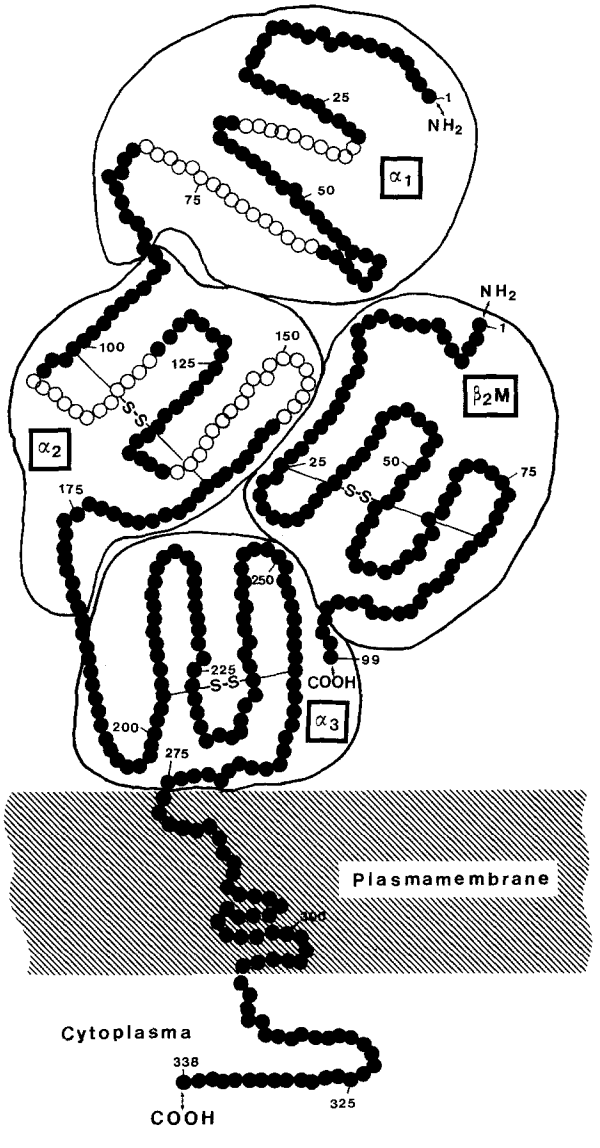


Figure 3. Schematic presentation of an HLA class I antigen molecule. Each dot represents an amino acid residue. The open circles mark the regions with preferential allelic differences[17].

(or haplotypes) with certain diseases[43, 46] has in the meantime achieved a much higher scientific impact than the original topic of tissue transplantation.

Most of the illnesses correlate more strongly with HLA class II antigens. But there is one exception, the class I antigen HLA-B27. Carriers of this antigen have an extremely high risk of contracting various forms of rheumatoid disorders, among which ankylosing spondylitis predominates. Szöts et al.[44] have probably localized the two amino acids of HLA-B27 that cause the disfunction.

3. The structures of immunological specificity

In 1959 Burnet[15] proposed his *clonal selection hypothesis* which stated that an unimmunized individual is a mosaic of lymphocyte clones, each clone being derived from a single progenitor cell. Lymphocytes of the same clone can produce only one kind of antibody and display this type also on their surfaces as an antigen receptor. When an antigen enters the body, it is bound by cells of that clone which carries the receptor specific for this antigen. The binding will stimulate the clone, which then expands and begins to secrete the antibody type. An antigen thus selects, from a vast variety of lymphocytes, that clone which is capable of producing antibodies to this antigen[27]. The finding of two types of immunity (humoral and cell-mediated) meant that this hypothesis had to be extended to the two classes of lymphocytes, B *and* T cells. In fact, recent research has resulted in the detection of both the B cell receptor (including immunoglobulins) and the T cell receptor.

3.1 Genetics of the B cell receptor and the immunoglobulins

This topic belongs to the most fascinating of immunogenetics and has given rise to much discussion and controversy during this century. It has been reviewed recently by Tonegawa[47] and Milstein[35].

Antibodies are tetrameric Y shaped molecules, composed of two identical *light chains* (L chains) of about 220 amino acids in length (consisting of two homologous domains, each with 110 amino acids) and two identical heavy chains (H chains) of about 440 amino acids (consisting of four homologous domains, each with about 110 amino acids, see fig. 6). The L chains can be of two types, kappa (κ) and lambda (λ); there are even 9 different H chains (μ, δ, $\gamma 1$, $\gamma 2$, $\gamma 3$, $\gamma 4$, $\alpha 1$, $\alpha 2$ and ε) for the 9 different antibody classes (IgM, IgD, IgG1, IgG2, IgG3, IgG4, IgA1, IgA2, and IgE), all possessing different effector properties.

The antibody specificity is fixed by the amino acid sequence. But interestingly, the *NH_2-terminal domains* of the L and of the H chains alone were found to carry all the variability which gives rise to the 10^6 (or even 10^8) different antibody specificities. This domain was therefore called the *variable* part (V_L and V_H). The other domains are (with respect to the different L chain and H chain types) *constant* (term: constant part, C_L, C_H).

Already in 1965 Hilschmann and Craig[24] and Dreyer and Bennett[19] postulated that each L and each H chain must be the product of two gene elements, a 'V' and a 'C' one. The problem, apart from the question of the 'V-C' joining mechanism, was whether the genome carries thou-

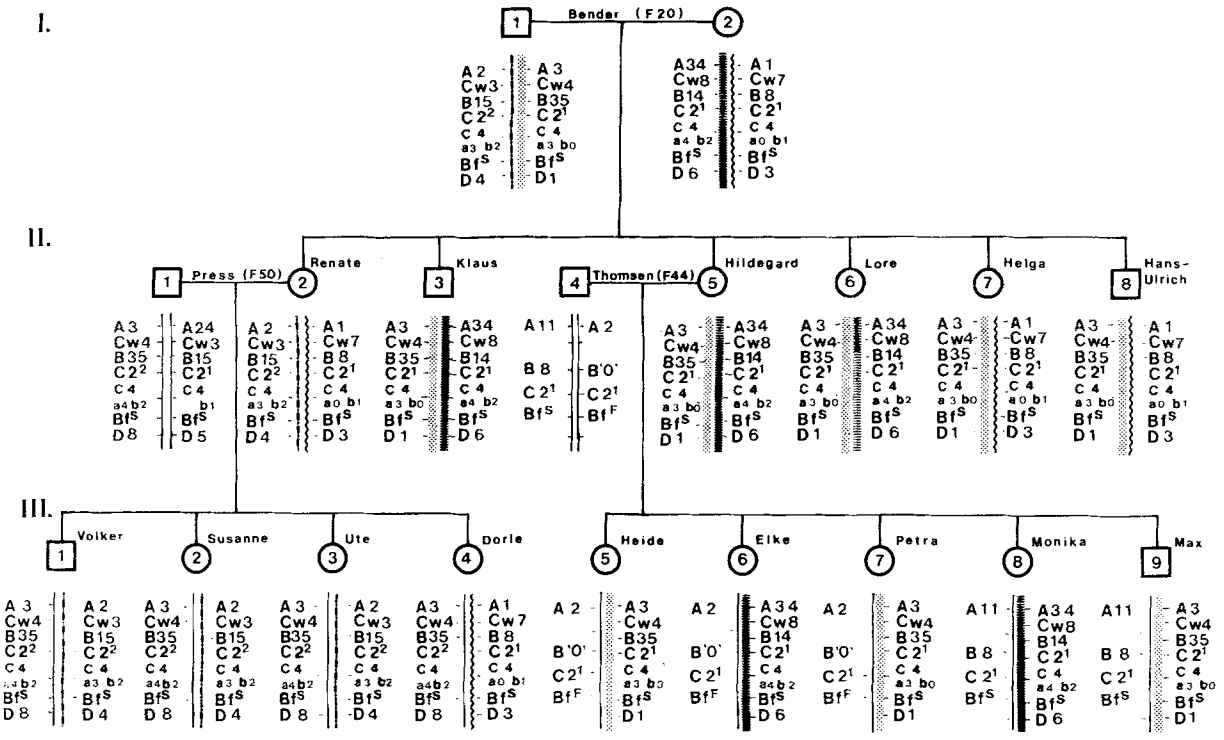


Figure 4. The 'en bloc' inheritance of HLA haplotypes over three generations, demonstrated within the author's own kinship[6].

sands of different V gene elements (of which one is fused with the C gene segment) or whether there is only one V gene with exceptional (i.e. hypermutable) properties. The contrary hypotheses were called the *germline theory* and the *somatic mutation theory*.

By using newly worked-out methods of DNA technology, Tonegawa found out that nature solves the problem of making so many B cell clones (and thus antibody specificities) in a rather economical way. The human genome carries (table) *three* clusters of antibody genes, one for the L chains of the κ-type (located on chromosome 2), one for the L chains of the λ-type (located on chromosome 22), and one for the H chains (located on chromosome 14). Each cluster of genes comprises (table) firstly a gene region with multiple V gene elements (IGKV, IGLV, IGHV), each for encoding the first 97

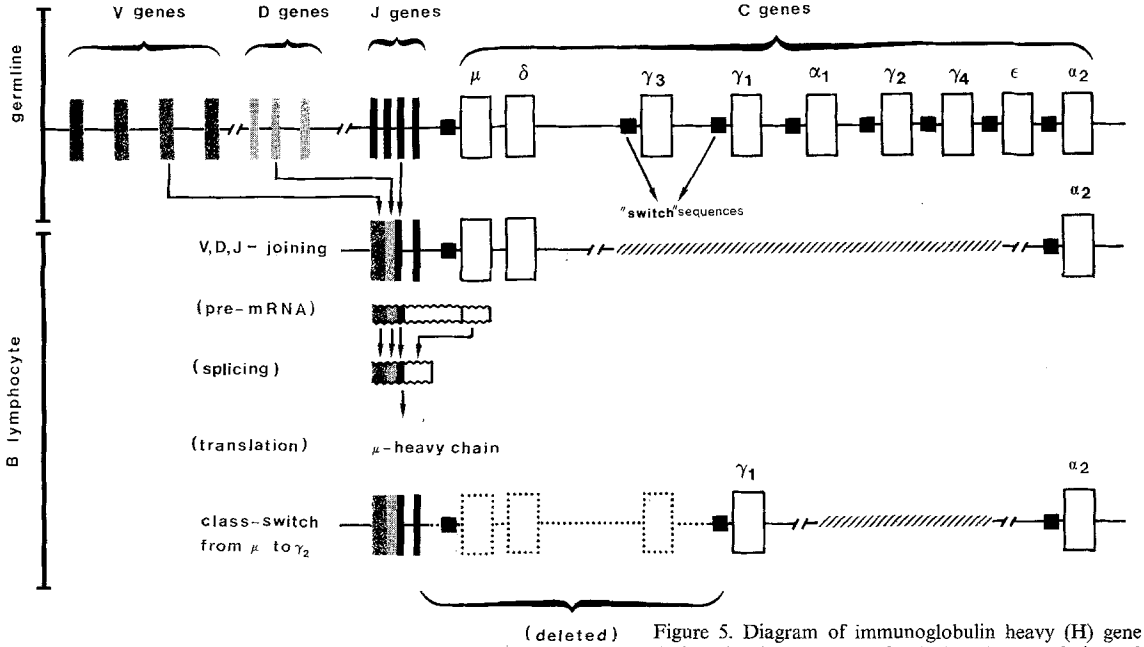


Figure 5. Diagram of immunoglobulin heavy (H) gene arrangements during the development of a B lymphocyte: fusion of V, D, J gene elements and class switch.

amino acids of V_K, V_λ and V_H polypeptides; secondly, a gene region with several J (joining) gene elements (IGKJ, IGLJ, IGHJ) encoding about 13 further amino acids of the variable domain; only the H gene region carries in addition several D (diversity) gene segments for encoding about 5-7 amino acids intermediate to the V_H and J_H polypeptide; thirdly, a few constant gene elements (see table: IGKC; IGLC; IGHM, IGHD, IGHG1, IGHG2, IGHG3, IGHG4, IGHA1, IGHA2, and IGHE).

The variable part of an L gene is created by bringing together a V_L and a J_L segment. One element of each fragment is combined by the deletion of intervening DNA regions to create a collinear V domain gene. A continuous messenger RNA (called pre-mRNA) spanning the joint gene up to the distantly located C_L gene is then transcribed. The pre-mRNA will then be spliced to the collinear $V_LJ_LC_L$ transcript which will be translated into a complete L chain polypeptide. The variable domain of an H chain is created by the fusion of a V_H, D_H and J_H fragment (fig. 5).

An intriguing molecular puzzle is presented by the long-standing observation that both the light and the heavy chains expressed by a given lymphocyte are the product of only one of the two alleles (term: allelic exclusion). Preferentially, two models can be envisioned. The first (restricted rearrangement) invokes a special mechanism to limit recombination to a single chromosome in a given cell, whereas stochastic models include *inefficient rearrangements,* in which each allele rearranges independently but often aberrantly, thereby inactivating an allele. Surprisingly, each model is favored by some results[2]. The juxtaposing of *one* out of 30–300 different variable (V), of *one* out of some twenty diversity (D), and *one* out of about four joining (J) segments represents, however, only a first source of variation, the co-called *combinatorial* choice. A second source of variation is the junctional diversity, created by the formation of new codons due to imprecise recombination as these elements are fused together[40]. A third diversification can occur around the D segment by insertion of new sequences, conducted by the enzyme terminal transferase[3].

The mature B cell (with membrane bound immunoglobulin = B cell receptor of the same specificity as the antibodies to be secreted by it) has distinctive activation requirements. This activation generally requires the interaction of the proper antigen in conjunction with an HLA class II antigen and a specific helper T cell and has two distinct phases: proliferation and differentation. The role of proliferation is to expand the number of cells capable of reacting against antigenic substances that have been introduced into the individual, since the frequency of B cells specific for any individual antigen in an unimmunized animal is very low. A fraction of the cells that proliferate on stimulation with antigen will differentiate into antibody-secreting cells (plasma cells), another fraction of the cells will lead to a state of immunological memory. A second immunization with the same antigen will thus lead to a prompter response of greater magnitude than the primary response.

Interestingly, the primary response results largely in the secretion of IgM antibodies. But after the second challenge with antigen, further stages of differentiation can be observed. One is the so-called class-switch (fig. 5), i.e. the new arrangement of the heavy chain variable gene (V_HDJ_H complex) to another C_H gene (the C_μ and C_δ genes being deleted); the B cell will now produce IgG_1 (or IgG_2, ..., IgE) antibodies with, however, the same light chain, so that the antibody specificity will not be changed (A recent paper of Blattner and Tucker[9], gives some insight into the switch mechanism). The other differentiation step results in the accumulation of somatic mutations, preferentially within the variable gene segment, giving

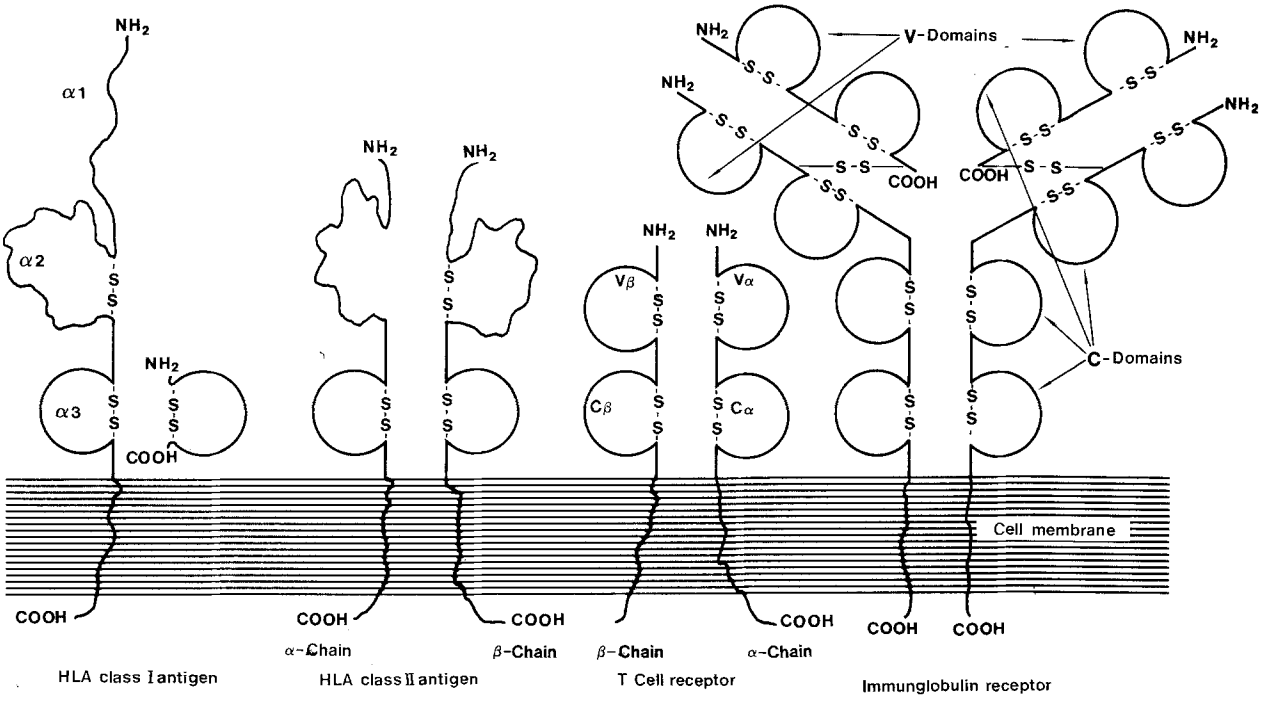


Figure 6. Models of membrane proteins with homology to immunoglobulin.

the chance of the creation of improved affinity of the antibody for the antigen[35].

In summary, neither of the two contradictory hypotheses (germline and somatic mutation theories) was right alone; both of them contained at least a grain of truth. Antibodies are preferentially expressed in the serum. The C genes of the kappa L chain and of some H chains are polymorphic (Km, Gm polymorphisms).

3.2 Genetics of the T cell receptor

The specificity of the B cells is given by their membrane-bound immunoglobulins, also referred to as B cell receptors. But there was a long search before the problem of the specific structure of the T cells was also unravelled. It was only in 1984 that the unique clonotypic epitopes were found[29,54] which (as the HLA antigens and the immunoglobulins) have immunoglobulin-like structures. Three gene regions (table) have been identified, each with a similar gene-topographic architecture to the immunoglobulin clusters: TCRA (coding for the T cell receptor alpha polypeptide; located on chromosome 14), TCRB (coding for the T cell receptor beta polypeptide; located on chromosome 7) and TCRG (coding for the T cell receptor gamma polypeptide; located on chromosome 7). One α and one β chain associate to form the heterodimeric T cell receptor (fig. 6). The function of the γ chain is unknown.

Each gene region is comprised of V, D, J and C gene elements. The combinatorial choice and rearrangement occurs largely in the same manner as that of the immunoglobulins, including junctional diversity creation and the random addition of nucleotides to either end of the D gene segment in the process of joining to the V and J segment[25].

The fact that class I and class II MHC antigens, β_2 microglobulin, antibodies and the T cell receptor as well as the CD4 and CD8 differentiation antigens are composed of domains and each of the domains is highly homologous stands out clearly (fig. 6). Many of the genes involved in the vertebrate immune response may therefore have a common evolutionary origin[25].

4. Molecules with immunoregulatory and effector properties

Especially from experiments with long-term cultures of immunocompetent cells it was learned that certain growth factors exist which were secreted from activated monocytes *(monokines)* and lymphocytes *(lymphokines)*. The substances keep the cells which secrete them, and also other cells in a state of proliferation. It took a long time before the substances were characterized[50]. They are now designated interleukins, because they act as communication signals between different populations of leukocytes. One of the well-defined interleukins is produced by macrophages and is called interleukin 1 (IL1; table), the other is produced by T lymphocytes and is called interleukin 2 (IL2; table).

According to Acuto and Reinherz[1] a resting T cell possesses many T cell receptors but very few IL-2 receptors (IL2R; table). After activation the T cell receptor is modulated (capping phenomenon[45]) and largely pinocytosed; instead, a high expression of IL-2 receptors is observed followed by IL2 secretion. IL2 can bind to its own cell and to other T cells, and as long as a high density binding is achieved, DNA synthesis and mitosis will occur. In the absence of additional stimulation, the T cell receptor is re-expressed and the T cell has then reached the resting stage again (memory T cell).

Interferons (IFN) are a family of proteins (143–166 amino acids in length) with the biological function of inhibiting both viral replication and cellular proliferation. In addition, it was found that they enhance the expression of HLA antigens. Each of these effects has a potential clinical importance in regulating cell function and response to neoplasia[41]. There are at least 13 genes controlling the production of interferons. The table lists a number of them and also the genes for the interferon receptor.

Another important class of soluble substances is represented by the complement proteins. They comprise a system of mediators of inflammation and can lead to the direct lysis of foreign cells marked by antibody binding (antigen-antibody-complexes). Eighteen plasma proteins have been identified[36] with cascade-like activating, regulatory and lytic properties (The table lists only those for which gene mapping information was available). Three functional pathways have been worked out: two different activation mechanisms and one common effector cascade. The complement proteins Clqrs, C4 and C2 on the one hand and DF, C3 and BF on the other hand are involved in the activation of C3, which is central in the system and is also critical for the opsonization (i.e. preparation for phagocytosis) of bacteria and other particles. C5, C6, C7, C8 and C9 are involved in the effector cascade which results in the lysis of the cell. In addition, fragments of C3 and C5 (C3a and C5a) have powerful inflammatory effects.

1 Acuto, O., and Reinherz, E. L., The human T-cell receptor. Structure and function. New Engl. J. Med. *312* (1985) 1100–1111. (see also: Reinherz, E. L., Acuto, O., Fabby, M., Bensussan, A., Milanese, C., Royer, H. D., Meuer, S. C., and Schlossman, S. F., Clonotypic surface structure on human T lymphocytes: Functional and biochemical analysis of the antigen receptor complex. Immun. Rev. *81* (1984) 95–129).

2 Adams, J. M., The organization and expression of immunoglobulin genes. Immun. Today *1* (1980) 10–17.

3 Alt, F. N., and Baltimore, D., Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-J_H fusions. Proc. natn. Acad. Sci. USA *79* (1982) 4118–4122.

4 Alt, F. W., Yancopoulos, G. D., Blackwell, T. K., Wood, C., Thomas, E., Boss, M., Coffmann, R., Rosenberg, N., Tonegawa, S., and Baltimore, D., Ordered rearrangement of immunoglobulin heavy chain variable region segments. EMBO J. *3* (1984) 1209–1219.

5 Arden, B., and Klein, J., Biochemical comparisons of major histocompatibility complex molecules from different subspecies of *Mus musculus:* Evidence for trans-specific evolution of alleles. Proc. natn. Acad. Sci. USA *79* (1982) 2342–2346.

6 Bender, K., Grundlegende Konzepte der Immunologie. Struktur und Funktion des HLA-Systems. Umschau *81* (1981) 300–303.

7 Bender, K., The HLA-A: HLA-B crossovers and their contribution in analysing possible haplotype-specific recombination rates. Hum. Genet. *61* (1982) 127–134.

8 Bender, K., Das HLA-System, pp. 1–106. Biotest-Serum-Institut GmbH, Frankfurt 1983.

9 Blattner, F. R., and Tucker, P. W., The molecular biology of immunoglobulin D. Nature *307* (1984) 417–422.

10 Blostein, R., Lauf, P.K., and Tosteson, D.C., Characteristics of Na^+-ATPase of low-K^+ sheep red cell membranes stimulated by blood group L antiserum. Biochim. biophys. Acta *249* (1971) 623–627.

11 Bodmer, W.F., Evolutionary significance of the HL-A system. Nature *237* (1982) 139–145.

12 Bodmer, W.F., Population genetics of the HL-A system: Retrospect and prospect. Histocompatilibility Testing 1972, pp. 611–617. Munksgaard, Copenhagen 1973.

13 Bodmer, W.F., The William Allan Memorial Award Address: Gene clusters, genome organization, and complex phenotypes. When the sequence is known, what will it mean? Am. J. hum. Genet. *33* (1981) 664–682.

13 Bodmer, W.F., and Bodmer, J.G., Evolution and function of the HLA system. Br. Med. Bulletin *34* (1978) 309–316.

15 Burnet, F.M., The Clonal Selection Theory of Aquired Immunity. Cambridge University Press, London 1959.

16 Ceppellini, R., Discussion remark, in: Genetic Control of Immune Responsiveness, p. 365. Eds H.O. McDevitt and M. Landy. Academic Press, New York 1972.

17 Coligan, J.E., Kindt, T.J., Uehara, H., Martinko, J., and Nathenson, S.G., Primary structure of a murine transplantation antigen. Nature *291* (1981) 35-39.

18 Dausset, J., The major histocompatibility complex in man. Past, present, and future concepts. Science *213* (1981) 1469–1474.

19 Dreyer, W.J., and Bennett, C.J., The molecular basis of antibody formation: a paradox. Proc. natn. Acad. Sci. USA *54* (1965) 864–869.

20 Forman, J., T cells, the MHC, and function. Immun. Rev. *81* (1984) 203–219.

21 Fujii, H., and Yoshida, A., Multiple components of blood group A and B antigens in human erythrocyte membranes and their difference between A_1 and A_2 status. Proc. natn. Acad. Sci USA *77* (1980) 2951–2954.

22 Gibblett, E.R., ADA and PNP deficiencies: how it all began. Ann. New York Acad. Sci. *451* (1985) 1–8.

23 Götze, D. (Ed.), The Major Histocompatibility System in Man and Animals. Springer-Verlag Berlin, Heidelberg, New York 1977.

24 Hilschmann, N., and Craig, L.C., Amino acid sequence studies with Bence-Jones proteins. Proc. natn. Acad. Sci. USA *53* (1965) 1403–1409.

25 Hood, L., Kronenberg, M., and Hunkapiller, T., T cell antigen receptors and the immunoglobulin supergene family. Cell *40* (1985) 225–229.

26 Kärre, K., Ljunggren, H.G., Piontek, G., and Kiessling, R., Selective rejection of H-2 deficient lymphoma variants suggests alternative immune defence strategy. Nature *319* (1986) 675–678.

27 Klein, J., Immunology. The Science of Self-Nonself Discrimination. John Wiley & Sons, New York 1982.

28 Kratzin, H., Yang, C., Götz, H., Pauly, E., Kölbel, S., Egert, G., Thinnes, F.P., Wernet, P., Altevogt, P., and Hilschmann, N., Primärstruktur menschlicher Histokompatibilitätsantigene der Klasse II. 1. Mitteilung: Aminosäuresequenz der N-terminalen 198 Reste der β-Kette des HLA-DW2,2; DR2,2 Alloantigens. Hoppe-Seyler's Z. Physiol. Chem. *362* (1981) 1665-1669.

29 Mak, T.W., and Yanagi, Y., Genes encoding the human T cell antigen receptor. Immun. Rev. *81* (1984) 221-233.

30 Malissen, M., Malissen, B., and Jordan, B.R., Exon/intron organization and complete nucleotide sequence of an HLA gene. Proc. natn. Acad. Sci. USA *79* (1982) 893–897.

31 Marsh, W.L., Oyen, R., and Nichols, M.E., Ky antigen, the McLeod phenotype and chronic granulomatous disease: further studies. Vox Sang. 31 (1976) 356–362.

32 Mauff, G., Alper, C.A., Awdeh, Z., Batchelor, J.R., Bertrams, J., Bruun-Petersen, G., Dawskins, R.L., Demant, P., Edwards, J., Grosse-Wilde, H., Hauptmann, G., Klouda, P., Lamm, L., Mollenhauer, E., Nerl, C., Olaisen, B., O'Neill, G.O., Rittner, C., Roos, M.H., Skanes, V., Teisberg, P., and Wells, L., Statement on the nomenclature of human C4 allotypes. Immunobiology *164* (1983) 184–191.

33 McAlpine, P.J., Shows, T.B., Miller, R.L., and Pakstis, A.J., The 1985 catalog of mapped genes and report of the nomenclature committee. Human Gene Mapping 8 (1985). Eighth International Workshop on Human Gene Mapping. Cytogenet. Cell Genet. *40* (1985) 9–66.

34 Mendes, N.F., Tolnai, M.E.A., Silveira, N.P.A., Gilbertsen, R.B., and Metzger, R.S., Technical aspects of the rosette tests used to detect human complement receptor (B) and sheep erythrocyte-binding (T) lymphocytes. J. Immun. *111* (1973) 860–867.

35 Milstein, C., From antibody structure to immunological diversification of immune response. Science *231* (1986) 1261–1268.

36 Müller-Eberhard, H.J., Complement. A. Rev. Biochem. *44* (1975) 697–724.

37 Orr, H.T., Lopez de Castro, J.A., Parham, P., Ploegh, H.L., and Strominger, J.L., Comparison of amino acid sequences of two human histocompatibility antigens, HLA-A2 and HLA-B7. Location of putative alloantigenic sites. Proc. natn. Acad. Sci. USA *76* (1979) 4395–4399.

38 Race, R.R., and Sanger, R., Blood Groups in Man. 6th edn. Blackwell, Oxford 1975.

39 Robertson, M., The evolutionary past of the major histocompatibility complex and the future of cellular immunology. Nature *297* (1982) 629–632.

40 Sakano, H., Huppi, K., Heinrich, G., and Tonegawa, S., Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature *280* (1979) 288–294.

41 Smalley, R.V., and Borden, E.C., Interferons: Current status and future directions of this prototypic biological. Springer Semin. Immunopath. *9* (1986) 73–83.

42 Spielmann, W., and Kühnl, P., Blutgruppenkunde. Thieme Verlag, Stuttgart 1982.

43 Svejgaard, A., Platz, P., and Ryder, L.P., HLA and disease 1982 – a survey. Immun. Rev. *70* (1983) 192–218.

44 Szöts, H., Riethmüller, G., Weiss, E., and Meo, T., Complete sequence of HLA-B27 cDNA identified through the characterization of structural markers unique to the HLA-A, -B, and -C allelic series. Proc. natn. Acad. Sci. USA *83* (1986) 1428–1432.

45 Taylor, R.B., Duffus, W.P.H., Raff, M.C., and de Petris, S., Redistribution and pinocytosis of lymphocyte surface immunoglobulin molecules induced by anti-immunoglobulin antibody. Nature (New Biol.) *223* (1971) 225–229.

46 Tiwar, J.L., and Terasaki, P.I., HLA and Disease Associations. Springer Verlag New York, Berlin, Heidelberg, Tokyo 1985.

47 Tonegawa, S., Somatic generation of antibody diversity. Nature *302* (1983) 575–581.

48 Vogel, F., Pettenkofer, H.J., and Helmbold, W., Über die Populationsgenetik der ABO-Blutgruppen. 2. Mitteilung. Genhäufigkeit and epidemische Erkrankungen. Acta genet. *10* (1960) 267–294.

49 Watkins, W.M., Biochemistry and genetics of the ABO, Lewis, and P blood group systems, in: Advances in Human Genetics 10, pp. 1–116. Eds H. Harris and K. Hirschhorn. Plenum Press, New York 1980.

50 Watson, J., Mochizuki, D., and Gillis, S., T-cell growth factors: interleukin 2. Immun. Today *1* (1980) 113–117. (see also: Dinarello, C.A., and Mier, J.W., Interleukins. A. Rev. Med. *37* (1986) 173–178).

51 Yanagi, Y., Yoshikai, Y., Leggett, K., Aleksander, I., and Mak, T., A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature *308* (1984) 145–149.

52 Yang, C., Kratzin, H., Götz, H., Thinnes, F.P., Kruse, T., Egert, G., Pauly, E., Kölbel, S., Wernet, P., and Hilschmann, N., Primärstruktur menschlicher Histokompatibilitätsantigene der Klasse II. 2. Mitteilung: Aminosäuresequenz der N-terminalen 179 Reste der α-Kette des HLA-Dw2/DR2-Alloantigens. Hoppe-Seyler's Z. Physiol. Chem. *363* (1982) 671–676.

53 Yoshida, A., Identification of genotypes of blood group A and B. Blood *55* (1980) 119–123.

54 Yoshida, A., Biochemical genetics of human blood group ABO system. Am. J. hum. Genet. *34* (1982) 1–14.

55 Yoshida, A., The existence of atypical blood group galactosyltransferase which causes an expression of A_2 character in A^1B red blood cells. Am. J. hum. Genet. *35* (1983) 1117–1125.

0014-4754/86/101138-10$1.50 + 0.20/0