HLA AND DISEASE ASSOCIATIONS—THE INDIAN SCENE

N. K. MEHRA
Department of Anatomy, All India Institute of Medical Sciences,
New Delhi 110029, India.

ABSTRACT
Besides its prime role in transplantation, the HLA system in man has been implicated in a number of other biological phenomenon which includes the HLA-linked control of immune response to a variety of pathogens, recognition of self antigens, regulation of various cell types involved in the immune response etc. In particular, the extreme degree of polymorphism that exists in the HLA complex has provided the researchers with an efficient genetic marker system. A large number of diseases have been found to be associated and/or linked with a particular HLA allele/haplotype. Such studies have proved useful in the diagnosis and prognosis of many diseases, differentiating them in several cases and in understanding their immunopathogenesis etc. Our recent data on the HLA and disease associations in the native Indian population have been reviewed and discussed in relation to that reported in other major ethnic groups in the world.

INTRODUCTION
The human leucocyte antigen (HLA) system has arisen out of a search for blood groups on the surface of the leucocytes. It is a part of the major histocompatibility complex (MHC) which is known to exist in a number of vertebrate species including man. The system comprises a series of codominant genes lying at several closely linked loci which exert their prime influence in the allograft reaction. The most thoroughly investigated model of a MHC is the H-2 system in mice, the homologue of which in man is the HLA. With the demonstration of immune response genes in the MHC, the subject of histocompatibility has assumed greater biological significance since the MHC products have been directly implicated in the immunoregulatory mechanisms of the body. This knowledge has helped in explaining how various HLA factors cause susceptibility and/or resistance to a variety of diseases.

Most of the data concerning HLA and disease associations are based on studies by European and American workers on patients of Caucasian and Negroid origin. Very little information is available on patients of the Asiatic region and diseases prevalent in the tropical countries. This report presents associations that have been established in North Indians, the various hypotheses about their mechanisms and delineates areas of future research.

Components of the HLA system
General surveys and historical background of the HLA system have been given elsewhere. Only a brief summary will be given here with emphasis on the recent discoveries. The MHC gene products are located on a small region on chromosome 17 in mouse and on short arm of chromosome 6 in man (figure 1). The HLA in man is approximately 2 M (centimorgan) in length lying close to loci coding for red cell enzymes, phosphoglucomutase-3 (PGM3) and glyoxylase. The MHC cell membrane molecules are glycoproteins, clustered in three classes or factors as suggested by Klein. In man, the HLA class I molecules are present on cell membrane of almost all nucleated cells and are coded for genes within the HLA-A, -B and -C loci. The alleles in these loci are numbered numerically i.e. A1, B5 etc and can be determined serologically by complement dependent serological assay. The comparable class I molecules in the mouse, H-2 are coded by K, D and L loci. The class II gene products are structurally, biochemically and functionally different from the class I molecules. These antigens have a rather restricted tissue
Figure 1. A schematic representation of the classes and loci of MHC genes in mouse and man. The HLA in man is represented on a small region on the short arm of chromosome 6 and its homologue in the mouse, H-2, is located on chromosome 17. At least four classes of molecules have been described in both species.

Distribution being expressed preferentially on B lymphocytes (but also on activated T lymphocytes) and macrophages. The class II antigens include the I-A and I-E regions in the mouse, and HLA-D/DR and SB in man. The D-antigens are detectable by cell culture techniques using homozygous typing cells (HTCs) in mixed lymphocyte cultures (MLC) and the DR (D-related) antigens by serologic techniques on B lymphocytes. Recently, the primed lymphocyte typing has proven useful in studying the D/DR region. Using this method, new series of lymphocyte stimulating antigens, SB (now called DR locus) has been described. Also, the previously designated antigens in the MT, MB and DC loci have been grouped into the DO locus.

The class III gene products include complement components, C2, C4 and Bf of the properdin factor system. Some of these, particularly C4 are genetically polymorphic and mapping of their structural loci within the HLA complex makes them useful additional genetic markers of this region. Finally, although biochemically similar to class I products, a number of so called lymphocyte differentiation antigens are suggested to form a distinct class of antigens, class IV products.

Molecular structure of HLA Antigens

Both HLA class I as well as class II antigens consist of one α and one β polypeptide chain (figure 2). The class I α chain (molecular weight = 44,000 daltons) consists of three extracellular domains (α1, α2 and α3), each of approximately 90 amino acid residues, one transmembrane region of about 30 residues and a cytoplasmic region of approximately 30 residues. The corresponding genes in these encompass 4–6 kilobases (Kb) and consist of 7–8 exons. The first exon encodes a leader sequence, the 2nd, 3rd and 4th each encode one of the extracellular domains, the 5th encodes the transmembrane region, while the last 2–3 exons encode the cytoplasmic and 3' untranslated regions. The β chain of class-I (Mw = 12,000) is the β2-microglobulin consisting of only one extracellular domain.

In the class II molecules, each of the α (Mw = 34,000) and β (Mw = 20,000) chains consist of two extracellular domains; α1 and α2 or β1 and β2 respectively, one transmembrane portion and a short cytoplasmic region. The α2 and β2 domains are very similar in structure to the class-I α3 domains, β2-microglobulin, Thy-I T-cell differentiation antigens and constant domains of Immunoglobulin. Structural similarities are also observed between class-I α1 and class II β domains as well as between class-II α1 and the constant domain of 1g. There is thus reason to believe that the MHC and immunoglobulin molecules may have an evolutionary common ancestor gene.

Recombinant DNA technology and HLA Allelic Polymorphism

One of the most characteristic features of the MHC in man is its extraordinary polymorphism. Using serological procedures, more than 100
different alleles have been identified in the HLA class-I and class-II molecules alone. Accordingly, several million genotypes are possible within a given population and the chance for two unrelated individuals to be HLA compatible is indeed very remote. Technically, one often encounters difficulties in detecting 'workshop' specificities (i.e. those having a 'w' prefix), the recently identified splits in most antigens, the serologically cross reactive antigens etc. The identification of class-II HLA antigens is even more difficult primarily because of the non-availability of potentially 'good' antisera as well as other technical reasons.

The HLA system can now be precisely characterized at the DNA level, as has already been done for the genes of the immunoglobulins, and perhaps others. A very important breakthrough for this was brought about by the discovery of a new class of enzymes called 'restriction endonucleases' which possess the ability to cut the DNA double helix at specific sites. For example, the restriction endonuclease isolated from E. coli and named EcoRI splits the DNA whenever the sequence GAATTG/CCTAGG is present. Further, due to the availability of additional DNA modifying enzymes (viz exonuclease, DNA-ligase, DNA polymerase, terminal deoxynucleotidyl transferase, reverse transcriptase etc), a new methodology called recombinant DNA technology has been created. It enables one to isolate almost any gene from prokaryotic or eukaryotic cells in order to gain a considerable amount of functional information. The subject of DNA recombinant technology has been extensively reviewed by many investigators. The first cloning of a HLA gene was reported by Ploegh and coworkers who isolated a DNA segment consisting of 525 nucleotide base pairs (bp) with the reading frame from the terminal part of an HLA antigen obtained from the LK cells line with phenotype HLA-A1, B8. By applying similar techniques, a number of other HLA genes have been isolated and characterized. These include class-I α genes, class-II α genes,
and class-II β genes. The MHC polymorphism revealed by the molecular genetics procedures appears to correlate well with that seen by serological reagents.

**HLA AND DISEASE ASSOCIATIONS**

The subject of HLA and disease associations assumes importance chiefly because of two reasons: (a) The extreme degree of polymorphism that exists in this system makes it an extremely valuable tool in immunogenetics. Compared to the ABO blood groups in which there is only one locus, three alleles and a maximum of six possible phenotypes, the HLA is a multi-locus, multi-allelic system which envisages to have several million possible phenotypes in a given population. (b) The existence of the polymorphic Ir genes which contain information for the Ir gene products provide important knowledge concerning the immunopathogenesis and susceptibility to certain diseases.

Ever since the demonstration by Amiel (17) of an association of antigen '4c' with Hodgkin's disease in humans, the usefulness of HLA studies in clarifying the genetics of various diseases has been substantiated. Also, our knowledge of the biologic function of the HLA gene products and of the homologous MHC's in other species has increased considerably although there are still a lots of unanswered questions. In particular, consensus regarding the exact mechanism underlying HLA-controlled susceptibility to various diseases is yet to be reached. The newest data and trends have been published from time to time (18-19).

Whereas the association of some diseases with a particular HLA allele has been confirmed in other major population groups (B27 with ankylosing Spondylitis), a few diseases showing doubtful or no associations previously have now been found to be definitely HLA associated e.g. rheumatoid arthritis and systemic lupus erythematosus. In addition, the list has grown immensely with the inclusion of some more conditions. Besides, many of the diseases previously known to be HLA-B associated have now been shown to have stronger association with HLA-D/DR antigens e.g. DR3 with insulin dependent diabetes mellitus etc.

In the native Indian population, most of the data have been compiled by our group in New Delhi. The patients studied belonged to the North Indian states of Punjab, Haryana, Uttar Pradesh and Delhi. Broadly; they represent the Khatris, Guptas, Brahmins and Jats. A summary of the data on some of the diseases studied by us is given in table I.

**Spondyloarthropathies**

The discovery in 1973 of a remarkable association between ankylosing spondylitis (AS) and HLA-B27 has provided the researchers with an excellent model for studying the relationship between a disease process and an inherited antigen (20). Except for Australian Aborigins and Blacks in which there is a complete absence of both B27 as well as AS, the association is indeed strong in most other populations. Our data in the native North Indian population indicates B27 occurring in 92–95% of patients compared to a frequency of 5.8% in the normal population giving a very high relative risk of 1941-23. A similar increase in the frequency of B27 is seen in patients suffering from the related spondyloarthropathies viz. acute anterior uveitis (AAU) and Reiter's syndrome (RS). Incidentally, Reiter's syndrome in India is mainly of the post-dysenteric variety in contrast to the post-veneral RS observed in the west (24). Also, whereas the presence of A9 (A24) in combination with B27 confers greater susceptibility to AAU in North Indian patients (25), antigen B35 affords protection in RS (26). Thus apart from B27, other MHC linked factors have also been found to increase susceptibility to AS and related spondyloarthropathies. Based on the available data in various racial groups, a working hypothesis for susceptibility to spondylosis is proposed. According to this, HLA-B27 or linked gene acts as a primary susceptibility factor which in itself may be insufficient to cause the disease directly. Instead, it requires interaction of a second and third gene influencing infection and sex related factors respectively for the development of
Table 1 Summary of the data on HLA and diseases in the North Indian population.

<table>
<thead>
<tr>
<th>Disease</th>
<th>HLA antigen</th>
<th>Percent Frequency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankylosing spondylitis</td>
<td>B27</td>
<td>92.2</td>
<td>21, 22, 26</td>
</tr>
<tr>
<td>Reiter’s disease</td>
<td>B27</td>
<td>80.2</td>
<td>24, 26</td>
</tr>
<tr>
<td>Juvenile chronic polyarthritis</td>
<td>B27</td>
<td>83.3</td>
<td>21, 22, 26</td>
</tr>
<tr>
<td>‘Unclassifiable’ arthritis</td>
<td>B27</td>
<td>82.5</td>
<td>21, 22, 26</td>
</tr>
<tr>
<td>Colitic arthritis</td>
<td>B27</td>
<td>87.5</td>
<td>21, 22, 26</td>
</tr>
<tr>
<td>Acute anterior uveitis</td>
<td>B27</td>
<td>55.0</td>
<td>23, 25</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>DR4</td>
<td>67.7</td>
<td>27</td>
</tr>
<tr>
<td>Insulin dependent</td>
<td>BW21</td>
<td>35.0</td>
<td>30, 32</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>DR3</td>
<td>80.4</td>
<td>31, 32</td>
</tr>
<tr>
<td>Myasthenia gravis</td>
<td>BW21</td>
<td>19.0</td>
<td>37</td>
</tr>
<tr>
<td>Rheumatic heart disease</td>
<td>BW35</td>
<td>36.5</td>
<td>47</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>DR3</td>
<td>62.5</td>
<td>unpublished</td>
</tr>
</tbody>
</table>

Leprosy
- population data
- Family data
  | HLA antigen | Percent | Reference |
  | A9          |         | 38, 39, 44 |
  | DR2         |         | 41, 42    |

Pulmonary tuberculosis
- population data
  | HLA antigen | Percent | Reference |
  | DR2         | 50.8%   | 45        |

Classical AS in males. Besides, other HLA linked factors at the same or distant locus might influence the course and severity of the future spondylytic disease. Thus whereas A2, B27 supertype confers greater susceptibility to AS and “unclassifiable arthritis”, the existence of A9 in B27 positive AS patients signifies the presence of A9. Also, the presence of B35 along with B27 could confer protection from acquiring severe variety of Reiter’s syndrome. Ultimately, a clinician may be able to design appropriate treatment in a patient based on his initial HLA phenotypic expression and early manifestation of symptoms. However, in order to answer definitely the question of B27 involvement in the disease pathogenesis, it will be necessary to study the B27 molecule at the DNA level.

Rheumatoid Arthritis

Alleles in the HLA class I molecules do not reveal any significant deviation in patients of classical rheumatoid arthritis (RA). However, a significant excess of HLA-DR4 has been reported in almost all racial groups including the Asian Indians. That the disease in fact may be caused by a defective immune response gene in linkage disequilibrium with DR4 has been amply demonstrated by studies involving multiple case families and HLA haplotype segregation analysis.

Insulin-dependent diabetes mellitus

Studies in different ethnic groups have conclusively shown that genetic susceptibility to type I diabetes Mellitus (IDDM) is HLA linked. The most
extensively studies are the European and the North American Caucasoid populations in whom DR3 (and B8) or DR4 (and B15) increase the risk of developing IDDM while DR2 (and B7) exerts a protective influence\textsuperscript{28}. Interestingly, whereas the association of DR3 and/or DR4 is common to most populations including the Japanese, the B allele has been reported to be variable being Bw54 in the Japanese\textsuperscript{29} and B21 (B49) in the North Indians\textsuperscript{30}. This indicates that the putative diabetogenic gene conferring susceptibility to IDDM is located nearer to the D-locus (in strong linkage disequilibrium with DR3 or DR4); however, the axis of susceptibility varies in different population groups. Amongst the North Indians, we have recently demonstrated that IDDM is associated with DR3 and not DR4\textsuperscript{31}. Data from the multiple case families of IDDM indicates that the affected sibs share most commonly the HLA haplotypes among themselves as compared to the unaffected, healthy sibs. Most of this data favours a recessive mode of inheritance as opposed to the dominant mode\textsuperscript{32,33}. In some of the autoimmune diseases, typing of the class III complement components has been particularly useful in identifying the whole haplotype of linked genes conferring susceptibility and/or resistance. The best examples are IDDM and RA. Linkage studies in multiple case IDDM families living in the Paris area have helped in identifying three high risk haplotypes or segments of them\textsuperscript{33} viz, A2, Cw3, B15, BFS, DR4; Aw30, Cw5, B18, BfF1, DR3 and A1, B8, BFS, DR3. Our recent data in the IDDM patients of North Indian origin revealed the 'high risk' haplotype to be A28, B21, BFS1, DR3 (unpublished observation). Curiously, the 'low risk' haplotype conferring resistance or protection in this disease is found to be common in all populations viz A3, B7, DR2. Such data could be useful in understanding the model of inheritance of IDDM and other diseases in various populations. Recently, with the demonstration of a rare C4B allele (C4B2.9), Dawkins and co-workers\textsuperscript{34} have successfully identified a "supratype" with alleles at multiple loci to be associated with RA. Interestingly, the same supratype is also associated with IDDM suggesting some relationship between this disease and RA. Indeed, evidence has accumulated suggesting that these two diseases do cluster within families.

**NEUROLOGIC DISEASES**

HLA association studies have been extremely helpful in understanding the immunopathogenesis of neurologic disorders particularly those with an unknown etiology. Multiple sclerosis (MS) in which both population as well as family studies have been carried out in most caucasian populations reveals an association with A3 and B7. Typing of HLA-D determinants showed an even stronger association with DR2\textsuperscript{35}. In myasthenia gravis (MG), the association of B8 in caucasian patients is well established\textsuperscript{36}. In patients of Asian Indian origin, our studies have suggested that this allele is associated more strongly with thymic hyperplasia suggesting a possible heterogeneity in the disease\textsuperscript{37}. This therefore is a good example to suggest that different pathological processes can result in the same syndrome. Typing of HLA-DR determinants in 25 North Indian patients with MG indicates a strong association with HLA-DR3 which is compatible with the data in European Caucasoids.

**Leprosy**

One of the most fascinating recent developments in HLA and disease associations has been that concerning the infectious diseases. In leprosy, as many as 15 publications including two from our group\textsuperscript{38,39} have sought an association of HLA class-I antigens at the population level. However, no single antigen was reported to be associated amongst studies carried out in different population groups. This led workers to think in terms of conducting multiple case family studies of which presently at least seven reports are available in four different population groups viz Surinam\textsuperscript{40}, India\textsuperscript{41,42} Venezuela\textsuperscript{43} and China (personal communication). These investigations suggested an appreciably increased sharing of HLA haplotypes by the affected sibs as compared
to the unaffected healthy sibs. Also, a preferential segregation of HLA-DR2 from healthy parents to children affected with TT leprosy was observed. On the contrary, HLA-DRW6 showed decreased frequency in the TT affected sibs. It was therefore, concluded that susceptibility to atleast polar tuberculoid leprosy (TT) was controlled by HLA encoded genes. Assuming DR2 to be the marker for TT leprosy, we performed another population study in Wardha to test whether, this association could also be shown in nonfamilial sporadic cases. However, no such association could be recorded indicating that perhaps there could be different immunopathogenetic mechanisms operative for predisposition to familial versus sporadic leprosy. Another interesting facet has emerged after evaluating in detail, the total family data from various populations. Accordingly, whereas the segregation of HLA haplotypes amongst the affected sibs was non-random, that amongst the healthy sibs in these families was random indicating that susceptibility to leprosy per se may not be controlled by HLA linked genes; however these genes might act to 'modulate' the type of immune response to *M. leprae*. Also, whereas the Surinam as well as the Indian data pointed towards the recessive mode of inheritance, the Venezuela data indicates the dominant mode thus suggesting the race specificity of the disease and HLA associated factors. Recently, the data compiled at the Second Asian and Oceanian Histocompatibility Workshop (2AOH) revealed that polar lepromatous leprosy in most populations may be associated with a supertypic DR specificity, MTI.

**Pulmonary Tuberculosis**

The situation as regards pulmonary tubercu-
losis (PTB) is similar to leprosy. We have for the first time reported that while there is no association with any antigen of HLA-A, -B and -C loci, the frequency in the DR locus revealed significant deviations in the patients as compared to controls. At least two genes in the DR locus seem to be involved in governing susceptibility to PTB viz DR2 conferring susceptibility and DRw6 resistance. Encouraged by these findings, we undertook multiple case family studies involving 25 families obtained from the Delhi area. The affected sibs in these families shared haplotypes most frequently amongst each other as compared to the healthy sibs suggesting that PTB is controlled by HLA linked genes. Once again, HLA-DR2 was implicated.

**CARDIOVASCULAR DISEASES**

In this category, only rheumatic fever and rheumatic heart disease (RHD) seem to have been studied in some detail chiefly by our group in New Delhi and that of Zabriskie in Columbia. Before this, most of the studies were conducted only on alleles at the class I HLA loci and these did not take into consideration the D/DR locus. No single HLA-A or -B antigen was implicated in most of these studies. A preliminary study conducted by us revealed an appreciably increased frequency of A28 and B35, the relative risk for the latter being 3.13 in RHD. Recently, the group in Columbia demonstrated that 75% of all rheumatic fever (RF) patients expressed an antigen on their B cells called 883+ as identified by the use of a serum obtained from a single multiparous women. Using hybridoma technology, the same investigators have now been able to create two monoclonal antibodies labelled respectively as 83S19.23 and 25S10, the former being identical to 883+ while the latter identifying those patients who appeared negative with the 883+ antiserum. Since these markers are present in only 15% of the normal population, the relative risk of contracting RF in those individuals bearing these markers is calculated to be approximately 15 times. Recently, we studied RF and RHD patients of Indian origin and found HLA-DR3 to be appreciably increased in the patient group as compared to controls; 62.5% versus 20%, *p* < 0.001, relative risk 6.6 (unpublished observation). These data suggest that susceptibility to RF and RHD is HLA class-II antigen mediated and is a result of an abnormal Ia gene to streptococcal antigen. In coronary artery disease (CAD) also, our data indicate a strong association with DR3 which was particularly pronounced in those group of patients.
having a positive history. Interestingly DR3 did not reveal any association with the known risk factors in CAD viz smoking, diabetes mellitus, atherosclerosis etc.

CONCLUSIONS

HLA studies and new genetic methodology have significantly advanced our knowledge about the inheritance of some diseases. The discovery of the Ir genes controlling immune responsiveness to a variety of pathogens is perhaps the best thing that has happened to the field of immunogenetics and with an almost 'disease-a-month phenomenon' continuing, the validity of such reports to clinicians is beginning to be realized. Firstly, the associations thus far reported have contributed to the proper understanding of the aetiopathogenesis and prognosis of several diverse and hitherto complex disease entities. Secondly, they have helped in de-differentiating a number of disorders and suggested a common pathogenetic link between them. The knowledge thus provided the physician with a system of rationale classification of many complex disorders and helped in understanding the heterogeneity in most of them. Nevertheless, we still do not know whether the associations reported thus far are due to the direct effect of the HLA antigens themselves or whether these are due to other as yet unknown HLA factors present on the surface and controlled by genes at a distant but linked locus. Another point where controversy still prevails is the exact mechanism by which susceptibility and/or resistance to a disease occurs. In this regard, recent discovery of the supertypic DR loci viz. MT, MB and SB and the alleles encoding them is worth exploring. Finally, with the development of DNA probes for various parts of the HLA system, a new era has begun in the HLA and disease studies. Henceforth, it may be possible to know the exact DNA sequence in the vicinity of the HLA genes which will serve as markers for discerning anomalies of susceptibility genes. Efforts in this direction have already borne fruits. Using CDNA probes, distinct differences in hybridization patterns between DNA from healthy individuals and diabetic patients was observed after digestion with restriction endonucleases. This has opened up a new chapter in the subject of HLA and disease associations.

28 January 1985; Revised 20 April 1985

21. Malaviya, A. N., Mehra, N. K., Adhar, G. C.,


