THE HLA GENES AND THEIR DIVERSE POLYMORPHISM

N.K. Mehra

Department of Histocompatibility & Immunogenetics, All India Institute of Medical Sciences, Ansari Nagar, New Delhi-110029

ABSTRACT

Advanced DNA level studies based on HLA class II sequence analysis have revealed considerable diversity in HLA among Asian Indians. High resolution typing of specific alleles such as DR2 and DR4 in the HLA class II region by PCR-SSP or SSOP hybridization and their associated DR-DQ haplotypes have helped to detect unique haplotypes and novel alleles which have subsequently been confirmed by sequencing. Incidentally, remarkable stability has been maintained in several other DRB1 alleles viz. DR1, DR7, DR9 and DR10. The ARMS-PCR technology has been found to be particularly useful for typing HLA-A, HLA-B and HLA-Cw alleles. These technologies are far superior over serological methods. Our studies have shown remarkable heterogeneity of common HLA-A and B alleles in Asian Indians. Molecular subtyping of HLA-A2 revealed that subtype A*0211 is found only in Indian population and may be the result of selection pressure in this population. Investigations into polymorphism in the HLA-B27 gene revealed that subtypes common both to the western caucasians and orientals occur in the Indian population. It is apparent that the population of the Indian subcontinent, placed as it is between the Caucasoids and Negroids on one hand and Australoids and Mongoloids on the other, provides a rich source of many HLA haplotypes. While the most frequent Caucasian haplotypes occur with a reasonable frequency in Asian Indians, those found predominantly in other ethnic groups (e.g., Australian Aborigines and populations of Oceania, China and Japan) are also detected. Knowledge on this is most important for donor selection during organ and bone marrow transplantation and for designing MHC targeted vaccines in specific diseases.

The major histocompatibility complex (MHC) encompasses two major classes of molecules: the MHC class I and class II, both of which appear to have originated from a common ancestor gene. Though only about one thousandth of the total genome, the HLA system bears extreme degree of functional polymorphism with a nucleotide diversity

no less than 0.04 to 0.07 as compared to 0.0002 to 0.011 for most of the other genetic loci. The major biological function of the MHC is to bind peptide fragments derived from protein fragments (viruses, peptides etc.) and display them on the surface of antigen presenting cells (APCs), evoking effector responses upon recognition by the TCRs. This process requires an efficient intracellular machinery to fragment the protein antigens into smaller peptides capable of binding to a host MHC molecule. Since the number of peptides that can theoretically be generated is very large, there is need for an extensive MHC gene pool. Thus the HLA system which is the MHC of man is extremely polymorphic.

Address for correspondence:

Prof N.K. Mehra, Head, at the above address
Tel: 6967588, 6594638
Fax: (91 i1) 6862663
Email: narin98@hotmail.com
nkmehra@medinst.ernet.in

Although the general rules for peptide-MHC interactions for both classes of MHC molecules are essentially similar, there are two fundamental differences: i) MHC class I ligands originate from endogenous sources, mainly from proteins of the cytosol or the nucleus and are delivered by the 'endogenous processing pathway'. In contrast, the MHC class II ligands are generated by the degradation of proteins from the extracellular compartment. These distinctions are also reflected in the responding T cells: CD8 positive T cells being restricted by class I-peptide complexes, and CD4 positive T cells by class II-peptide complexes, ii) in accordance with the distribution of hydrogen bonds in the peptide binding groove of the MHC, the anchor residues are placed at the terminal ends of the class I groove. Contrarily, the binding forces are distributed throughout the class II groove and ensure bonds between peptide's backbone and class II molecule (1-3). The specificity of the interaction is determined by pockets in the MHC groove that have a fixed spacing from each other and that also have a specificity for anchoring particular side chains of the peptide's amino acids.

HUMAN MHC

Although the MHC was first described in the mouse as the H-2 system by Snell (4) in 1948, it was subsequently shown in almost all vertebrate species. Knowledge about the human MHC began in 1954 when Jean Dausset first described the existence of leucoagglutinins in the sera of multi-transfused individuals. The HLA system constitutes the human MHC whose gene cluster spans a region of about 4000 kb (4×10^6 nucleotides) on the short arm of chromosome 6 in the distal portion of the 6p 21.3 band (5). Data on physical mapping, DNA cloning and sequencing of MHC region from a number of laboratories has revealed the presence of a large number of genes with variable expression and functions arranged in the form of three regions: Class I (36 genes), Class II (27 genes) and Class III or central genes (39 genes) (6). These genes are arranged close together and are generally inherited together as a bloc (haplotype). Thus MHC system can be considered as a single genetic unit. The class I and class II regions each encode highly polymorphic

HLA loci

Fig. 1. Number of alleles in various HLA loci determined by molecular methods and their comparison with those defined by Serology.

families of cell-surface glycoproteins involved in immune regulation. Class III genes (central genes) placed between class I and class II genes code for factors involved in the complement system and various non-immune functions.

A striking feature of the products of the MHC is their extreme degree of polymorphism which means that an exceptional inter-individual variability exists as far as HLA antigen profile of a population is considered. With the introduction of DNA based methods of HLA typing, extensive molecular polymorphism has been discovered in each of the relevant HLA class I and class II loci, which is far in excess of the polymorphism obtained by serological methods (Fig.1). For example, currently at least 361 alleles have been described for HLA-B, 181 for HLA-A, 102 for HLA-C, 251 for HLA-DRB1, 22 for DQA1, 45 for HLA-DQB1, 21 for DPA1 and 91 for DPB1 (<http://www.ebi.ac.uk/imgt/hla>). The high degree of polymorphism in HLA appears to have resulted from recombination and exchange of genetic material between alleles of the same locus and also from point mutations and other genetic events. Theoretically, several million genotypic combinations (approximately 150 billion or even more) are possible in the HLA system. According to Klein (7), such a polymorphism is not only advantageous for an individual, but even more for the survival of the species surrounded by many different and often changing pathogens. Thus the polymorphism is probably essential for efficient functioning of the system.

The class I region is the most telomeric part of the MHC complex. Although 30-40 genes have been defined in this region, HLA-A, -B and -C are the only products that are well defined as 'classical transplantation antigens'. Other human class I genes with less defined gene products have recently been identified: HLA-E, -F, -G, -H and -J. HLA class I antigens are expressed on all nucleated cells and platelets. They are heterodimeric glycoproteins consisting of a heavy alpha chain (mol wt. 45 Kda) and a light chain β -2 microglobulin (mol.wt 12 Kda). The β -2 microglobulin (β -2m) is encoded by a gene located on chromosome 15, and associates noncovalently with the large heavy chain. Class I heavy chain is folded into three

globular domains, each of which contains stretches of about 90 aa : α -1, α -2 and α -3. The amino acid sequence differences that account for the rich polymorphism of HLA class I molecules are not randomly distributed throughout the α -1 and α -2 domains, but occur in specific regions. X-ray diffraction study of HLA-A2 molecule has revealed that α 1 and α 2 domains interact to form the 'antigen binding groove' whose floor is constituted of eight antiparallel β -pleated strands while the margins are formed by two α -helixes in the form of helical ribbons (Fig.2). The top α -helix and the four β strands to the left are from the α 1 domain and the lower helix and four β strands to the right are from the α 2 domain (8,9).

Fig. 2. Schematic representation of the antigen binding groove of human MHC class I molecules

Class II region of the human MHC extends over 1000-1200 kb with atleast six subregions termed DR, DQ, DP, DO, DN and DM. These molecules are constitutively expressed as heterodimers primarily on immunocompetent cells such as antigen presenting cells (monocytes, macrophages, dendritic cells, Langerhan's cells), B cells, activated T cells and bone marrow derived

precursor cells. Structurally, class II HLA molecules are similar to class I molecules by having an alpha and a beta chain of integral membrane glycoproteins. Three-dimensional structure of HLA-DR1 molecule indicates specific changes in the peptide binding groove so that its ends are more open (10). The locus specific sequences that distinguish DR, DQ and DP molecules are located in the $\alpha 1$ and $\beta 1$ domains. The sequences that distinguish between allelic variants are localized within 3-4 hypervariable regions within the $\beta 1$ domain of DRB1, DQA1 and DQB1 molecules.

In addition to the classical HLA class II genes, two groups of non-HLA genes have been identified in the MHC class II region. The first group is ABC transporter genes called TAP-1 (Transporter associated with Antigen Processing) and TAP-2 genes whose products are involved in the transport of antigen fragments produced in the cytoplasm into the lumen of the endoplasmic reticulum (11,12). The second group is proteasome-related genes that include LMP2 (Low Molecular mass Polypeptide or Large Multifunctional Protease) and LMP7 genes (13) whose products are large cytoplasmic proteolytic complex molecules that contain multiple catalytic sites. LMP complex is involved in the production of multiple peptides simultaneously from the same substrate (the 'bread slicer' model) to produce peptides better suited for MHC class I binding.

Genetic Diversity of HLA: Population Studies

Population is defined as an entity of individuals, who are very similar in many aspects, but yet differ to a considerable extent from one another. The most important characteristic of a population is the genetic constitution and genetic structure. Genetic constitution is essentially the pool of genes, which make up a population. The gene pool of a population can be specified in terms of gene frequencies. The substructure of a particular population is restricted to the gene flow among the individuals. This could be due to social, geographical, ecological or even biological barriers. The gene frequencies among various populations can vary quite considerably. With passage of time, these can also change within a particular

population. Mutations are responsible for generating new alleles, even whole new genes, when regions are duplicated. The fate of a mutation depends on three evolutionary forces - natural selection, migration and random genetic drift - all of which can alter the allele frequency in a population. Thus genetic diversity and origin of the population can be best studied by using a highly polymorphic genetic system. It is generally believed that MHC polymorphism reflects protection of the species during evolution and may be the result of natural selection and/or genetic drift. Certain class I and class II alleles aid in immune responsiveness to parasitic and bacterial infections. Infectious diseases are the most obvious candidates as the responsible selective force. Although this intensive polymorphism is of great value in the elucidation of genetic aspects of HLA associated diseases, it poses practical difficulties in the selection of a donor for organ and bone marrow transplantation.

India with the second largest population in the world is known for its vast ethnic diversity and cultural traits. Indians can be divided into several groups based on factors such as physiographical region (north, south, east, west), religion (Hindu, Muslim, Christian, Sikh, etc.), caste (>3,700 endogamous groups) and language (Hindi, Urdu, Tamil, Telugu, etc.). Anthropological and historical evidence classifies Indians broadly as Dravidians and Aryans. The former are considered as the earliest settlers who were driven southwards following invasion by Aryans from the Northwest during 2000-3000 BC. Considerable intermingling of Aryans and Dravidians has resulted in a new race, Aryo-or Crypto-Dravidians. Between the 12 and 16 centuries AD, waves of Arabics and Moghuls invaded the Indian subcontinent followed by Mongols and Tibeto-Burmans (14). Several studies using different genetic marker systems including HLA antigens in various communities of India have revealed the existence of a considerable and unique heterogeneity in the allelic and haplotypic frequencies in these populations which might be due to the 'founder effect' but not isolation-dependent divergence (15,16). Using PCR based DNA technologies for identification of both HLA class I and class II alleles, we have determined genetic diversity of HLA in the mixed population of

North India as well as a highly endogamous Brahmin caste group.

Significance of HLA studies in Anthropology

Being the most highly polymorphic genetic system, HLA has been used to great advantage for the definition of various racial groups, their migration pattern, possible admixture etc. The distribution of HLA antigen frequencies among populations shows marked differences. Most of the antigens detected in the Indian population are found in caucasoids suggesting a close kinship of the two population groups. The presence or lack of specific alleles for some ethnic groups is characteristic, e.g., presence of HLA-B42 in Negroids, lack of B8 and DR3 among Orientals, lack of B14 and very low frequency of B16 among Asian Indians (17). HLA and Anthropology studies are advantageous to answer the following questions:

1. How has the ethnic diversity occurred?
2. Why and how have HLA polymorphism and linkage disequilibrium been established and maintained i.e., natural selection or genetic drift?
3. How often has the HLA gene mutation occurred?
4. When and how have the HLA alleles that show cross reaction with one another separated?

The phenotype frequency of most HLA class I alleles in south Indians is similar to that reported in North Indians. However, some differences with regard to the haplotype frequencies occur in the two populations. The North Indian population has A10-B8, A1-B37, A30-B13 as characteristic haplotypes, whereas the South Indian population have A2-B40 and A24-B7 (18). Haplotypes, A1-B17 and A3-B7 have been identified in almost all studies on the Indian populations. Interestingly, the most typical Caucasian haplotype A1-B8-DR3 does not occur prominently in the Indian population. Incidentally, this haplotype is also present in Ashkenazi Jews, Chinese, American Indians, Mexicans and African Blacks and is absent in the Japanese. It is associated strongly with insulin dependent diabetes mellitus (IDDM) in the west. Although it occurs to some extent in Indian patients

with IDDM, the haplotype showing strongest association with this disease in India is A26-B8-DR3 or B21-DR3. It is possible that this haplotype is a recent introduction to European and North American population and its absence in the Japanese and low prevalence in the Middle East and Asian Indian populations represents natural selection based elimination. In India, A26-B8 is favoured possibly due to a 'founder' effect. On the contrary, haplotypes A1-B17 and A3-B7 both occur in the Indian subcontinent with even greater frequency and delta value. It is possible that the climatic and geophysical conditions, the resultant microenvironment and epidemics, and prevailing infections probably played a role in the selection of these haplotypes.

Molecular Studies in Asian Indians

Polymorphism in HLA results not only due to allelic differences in individual loci but also due to haplotypic combinations with neighboring loci. Molecular and genetic analysis of HLA class II polymorphism has been facilitated by the polymerase chain reaction and sequence specific oligonucleotide probe (PCR-SSOP) hybridization technologies. These methods allow identification of alleles differing by even one single amino acid. We have studied the molecular complexity of several class I and class II alleles in Asian Indians of the Delhi area. The data suggests that although the Indian population is essentially caucasoid, in reality it represents a mixture of both Caucasians as well as oriental alleles/haplotypes. As representative example, we present here data on HLA-A2 and B27 and DR2 and DR4. The details have been published separately elsewhere (19-23).

A considerable discrepancy was observed when the molecular data is compared with frequencies estimated by serology during the 11th IHW held in Yokohama (15), particularly with regard to HLA-2, A3, B7, DR3 and DR4. The most common HLA class I alleles in the Indian population were found to be HLA-A*02 (34.9%), A*24 (24.33%), A*11 (34.5%), A*33 (18.1%) and A*3201 (10.2%) in the A locus and B*07 (20.12%), B*35 (23.78%), B*40 (17.68%), B*57 (9.74%) and B*58 (4.26%) in the B locus. There was complete

absence of B*14 in this population. The ARMS-PCR technology used for the determination of HLA class I alleles was not only more sensitive, but it also helped to discriminate several alleles with doubtful serological assignment. The principle of Amplification Refractory Mutation System (ARMS) is that oligonucleotides with 3' mismatched residues will not function as PCR primer under appropriate conditions. DNA amplification will only occur if the ARMS primer target sequence is contained in the template DNA of interest. Using these methods and the reverse blot hybridisations (RBH), a far greater level of diversity could be ascertained in each of the HLA-A, -B, and -C loci than was possible using serological methods. It appears that more automated methods with higher resolution will soon become available to type large number of samples at allele level with minimum effort.

HLA-Cw locus polymorphism

Polymorphism in HLA-Cw locus alleles in the population around Delhi area was studied by DNA technique of ARMS-PCR. A total of 19 Cw locus alleles have been tested, of which Cw*07 was found to be the most predominant (41.9%). The Cw*04 occurred at a frequency of 36.7% showing strong

linkage disequilibrium with HLA-B*35 (Fig 3). Other common alleles include Cw*03 (12.9%). Amongst the five serologically undetectable Cw locus alleles (Cw*12-17), Cw*13 occurred at a frequency of 3.27% while Cw*14 and *15 occurred with a frequency of 4.3% each in this population.

HLA-A*02 subtyping by ARMS PCR and RBH

During the 12th International Histocompatibility Workshop (12 IHW), Paris 1996, an ARMS-PCR A*02 subtyping kit was produced as an extension of the class I kit. This is a nested PCR with 4 first round primers and 19 second round primers allowing the identification of HLA-A*02 subtypes. The more recent reverse blot hybridisation (RBH) technique is the forerunner of the DNA chip technology. The RBH technique involves using nylon membrane strips with an array of allele specific probes immobilised in the form of bands. The DNA to be identified is amplified using biotinylated generic primers for HLA-A or -B loci followed by hybridisation onto the probe immobilized membrane strips. Positive bands are identified by the development of BCIP/NBT based color and scored according to predefined allele specific patterns. HLA-A2 is the most heterogeneous HLA-A allele that occurs with a frequency of nearly

Fig. 3. Frequencies of HLA-Cw locus alleles as defined by the ARMS-PCR.

Table 1. HLA-A*02 allelic frequencies in different populations

HLA-A*02	Caucasian (n=76)	Mongolian (n=56)	N. Indians (n=74)	Han Chinese (n=45)	Japanese (n=53)
A*0201	96.1	55.4		40	47.2
A*0202					
A*0203		3.6	1.4	2.2	
A*0204					
A*0205	1.3	14.3	11.2		
A*0206		23.2	7.04	26.7	35.8
A*0207		3.6	38.02	31.1	17
A*0210					
A*0211	1.3	14	30.98		
A*0213	1.3				
A*0214					
A*0217					

35% in Asian Indians. The high resolution typing of HLA-A*02 on the mixed population of Delhi area revealed that among the 17 subtypes, only six (A*0203, *0205, *0206, *0207, *0209 and *0211) occurred in this population of which *0211 and *0207 are predominant alleles (Table 1). The caucasian classical subtype, A*0201 was found to be absent

in this population. Incidentally, A*0201 is the most ubiquitous allele present in most populations except the Thais. Subtypes, A*0207 and *0206 that occur predominantly among the Chinese occurred at an appreciable frequency among Indians. On the other hand, A*0211 appears to be the unique A*02 subtype in Asian Indians (Fig 4). The allele, A*0202

Fig. 4. A comparison of selected molecular subtypes of HLA-A2 and B27 in Asian Indians, Orientals and Caucasians.

classical subtype, A*0201 was found to be absent in this population. Incidentally, A*0201 is the most ubiquitous allele present in most populations except the Thais. Subtypes, A*0207 and *0206 that occur predominantly among the Chinese occurred at an appreciable frequency among Indians. On the other hand, A*0211 appears to be the unique A*02 subtype in Asian Indians (Fig 4). The allele, A*0202 is found in all African populations with the exception of Sudan but so far not in any of the other populations tested during the 12th IHW. Similarly, A*0203 is the most frequent allele in the Thai population and the Dai Chinese. Perhaps no other population is known to carry this allele.

Molecular subtypes of HLA-B27

HLA-B27 shows remarkable association with ankylosing spondylitis (AS) and related spondyloarthropathies in most populations around the world. Studies carried out by us in the population around Delhi has shown that it occurs with a frequency of 6% in healthy individuals, and about 95% in patients with AS and around 80-90% in other related spondyloarthropathies. HLA-B27 is a serological specificity that represents a family of nine closely related alleles defined by DNA sequences. These alleles differ in a limited number of amino acid substitutions clustered in exon 2 and exon 3 of the class I molecule and have been designated as B*2701-B*2709. These differences occur in the regions 77-83, 113-116 and at residues 59, 74, 131 and 152. Different techniques like serology, ID-isoelectric focusing (ID-IEF), cytotoxic T lymphocyte (CTL) typing and PCR-SSOP have been used to define subtypes in different populations. The functional significance of minor structural variations in the subtypes is evident from the observed differences in the reactivity of CTL against these subtypes.

We have analyzed heterogeneity of the B27 molecule in the North Indian population (22). The PCR-SSOP technology was used to ascertain polymorphism in exon 2 and 3 of the HLA-B27 gene. Four different subtypes were encountered: B*02702, 04, 05, and 07. Other subtypes of B27

viz B*2701, 03, 06 and 08 were absent in this sample of 60 HLA-B27 positive patients and 17 B27 positive controls. B*2704 (common oriental subtype) and B*2705 (common caucasian subtype) were the most common subtypes in the control and patient groups. B*02707 was less frequently encountered in both groups and B*2702 was found in only one individual. Both B*2704 as well as B*2705 occurred predominantly with a frequency of around 45% (Fig 4). Presence of B*2704 and B*2705, the classical subtypes of orientals and caucasians in Asian Indians in almost equal frequencies is a sequel to the racial admixture experienced in this region over centuries.

DNA typing of HLA class II alleles

The most frequent HLA-DR alleles observed in the Indian population are DRB*15/16 and DRB1*07 with DRB1*09 occurring with the least frequency (21). Comparison of DR frequencies between Asian Indians and Caucasians did not reveal significant differences confirming our earlier report that the populations native to the Indian subcontinent are essentially Caucasoid (17). Individually, the distribution of DRB1*01 in Maharashtrians was found to be similar to that of Western caucasoids although this allele is generally low in other Indian groups. Data on the occurrence of DRB1*07 revealed interesting information. This allele occurred most predominantly (45-50%) in the three Brahmin groups of Northern India viz Bhargavas, Chaturvedies and Kashmiri Brahmins (Fig 5).

Results on the distribution of HLA-DR2 and its splits in the Indian population highlights the racial admixture experienced in this region. For example, we observed a complete absence of DRB1*1601 among Asian Indians as compared to the known presence of this allele in Western caucasians (Table 2). Incidentally, this allele is also absent amongst Chinese and populations of Oceania. Similarly DRB1*1502 occurs relatively infrequently in western caucasians, but is a predominant subtype of DR2 both in Asian Indians as well as Orientals. Furthermore, the distribution of DR2 associated DRB1, DQA1 and DQB1 haplotypes revealed that the Indian population is composed of haplotypes that are characteristic both

Fig. 5. A comparison of percent frequencies of HLA-DRB1 alleles in populations of Delhi and Kashmiri Brahmin.

Table 2: Frequency distribution of selected molecular subtypes of HLA-DR2 and DR4 in Asian Indians and other major populations

Alleles/ Haplotypes	Caucasoids	Asian Indians	Orientals
DRB1*1501	25-29	73	40-42
DRB1*1502	3-5	51	60-70
DRB1*1601	5-10	0	0
DRB1*0401	62-70	9	7-9
DRB1*0403	29-35	36.4	10-12
DRB1*0404	3-4	18.1	2-3
DRB1*0405	0	18.1	60-75
DRB1*0406	0	4.5	16-20
DRB1*0407	12-15	0	4-6

polymorphism at the DNA level (Table 2). For example, DRB1*0405, a subtype of DR4 that occurs predominantly among the Japanese and is almost absent in caucasoids and occurs with a frequency of

nearly 18% in Asian Indians and is one of the major subtypes of DR4 encountered in Asian Indian patients with rheumatoid arthritis (23). This population is also characterized by the increased occurrence of DRB1*0403 in comparison to the western caucasian groups.

Non classical HLA class I genes

A group of HLA class I genes, named non-classical alleles or MHC class Ib genes, includes HLA-E, -F, G and -H which code for molecules whose functions have become more prominent in the recent past. These genes have been distinguished from classical MHC class I molecules by their limited polymorphism and lower level of expression on the cell surface. The HLA - E gene is located between HLA-C and HLA-A loci while the HLA-G is located telomeric to the HLA-A locus, showing high degree of sequence similarity to HLA-A2. Molecular studies have shown how HLA-G and HLA-E bind to NK cells, immunoglobulin and lectin-type inhibitory receptors. The HLA-E and HLA-G specialise in presenting peptides to NK cells which normally function to eliminate HLA negative cell types. By interacting with NK receptors, these class I molecules turn off the NK cell signaling and thereby prevent lysis of HLA-intact cells. The absence of classical HLA-A, B and C allotypes on

placental tissue prevents attack by maternal CTLs, however in return, this tissue could be subjected to NK cell mediated cytotoxicity. The HLA-G molecules expressed on the placental tissue thus protect against NK cells through the CD94/NKG2 receptor. Further the HLA-G molecules may present fetal-derived viral peptides to maternal CTLs for elimination of infected tissue. The HLA-H (Hfe) is located telomeric of HLA-A and point mutations in this gene have been clearly defined in familial hemochromatosis.

Among the four HLA-G alleles tested, only 3 occurred with appreciable frequency in Asian Indians: G*01011 (53.3%) G*01012 (93.3%), G*01013 (40%) (Table 3). No example of HLA-G*0103 was found in a sample of 30 individuals tested. Incidentally, G*01012 occurs with high frequency in Negroids (79.2%) and Orientals (75.1%) as compared to the caucasoids (58%). The Caucasian classical allele G*01011 occurred at a comparatively less frequency in Asian Indians (24). Similar observations were made with regard to HLA-E polymorphism, the most frequent alleles being E*0101 (74.2%), E*0103 (45.7%) and E*0104 (34.2%) (Table 4). These frequencies are similar to those observed in other major ethnic groups. No example of HLA-E*0102 was observed among Asian Indians (25).

Table 3: HLA-G allelic frequencies in the major populations of the world

Alleles	% Frequency			
	Caucasoids n=299	Orientals n=161	Asian Ind. n=30	Negroids n=24
G*01011	70.5	54.0	53.3	79.2
G*01012	57.9	75.1	93.3	79.2
G*01013	27.4	32.9	40.0	12.5
G*0103	1.3	0.0	0.0	0.0

Table 4. HLA-E allelic frequencies in the major populations of the world.

Allele	%Frequency			
	Caucasoids n=269	Oriental n=151	Asian Ind. n=35	Negroids n=24
E*0101	78.4	71.5	74.2	75.0
E*0102	9.6	5.3	0.0	0.0
E*0103	54.6	45.0	45.7	37.5
E*0104	14.9	52.3	34.2	33.3
N1	0.4	0.6	0.0	0.0
N2	0.4	0.6	0.0	0.0
N3	0.4	9.2	5.7	12.5
N4	0.4	0.0	0.0	0.0
N5	0.4	0.0	0.0	0.0

Conclusions

HLA region is unique as it possesses extreme degree of nucleotide polymorphism whose functional importance lies in combating a whole range of pathogens by presenting them to the immune system. Information on HLA polymorphism in distinct racial groups is important in understanding the pattern of origin of HLA haplotypes through superimposed effects of racial admixture and environment. An important question that needs to be answered is why and how have such an extensive polymorphism and linkage

disequilibrium been established and maintained i.e. by natural selection or by genetic drift? Also how often have the HLA mutations occurred? With the advent of molecular based techniques, the clinical applications of HLA diversity have been highlighted.

Acknowledgements

Financial assistance for the study was provided by the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India. The author wishes to thank Dr Gurvinder Kaur for helpful suggestions and Mr Sanjay Kumar for typing the manuscript.

REFERENCES

1. Madden, D.R., Garboczi D.N. and Wiley, D.C. (1993) The antigenic identity of peptide-MHC complexes- a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* 75, 693-708.
2. Stern, L.J., Brown, J.H., Jardetzky, T.S. Gorga, J.C., Urban, R.G., Strominger, J.L. and Wiley D.C. (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 368, 215-221.
3. Stern, L.J. and Wiley D.C. (1994) Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure* 2, 245-251.
4. Snell, G. (1948) Methods for the study of the histocompatibility genes. *J. Genet.* 49, 87-108.
5. Lamm, L.U. and Olaisen B. (1985) Report of the committee on the genetic constitution of chromosome 5 and 6. *Cytogenet. Cell Genet.* 40, 128-155.
6. Campbell, R.D. and Trowsdale J. (1993) Map of the human MHC. *Immunology Today*, 14, 349-352.
7. Klein J. (1987) Origin of major histocompatibility complex polymorphism. The transspecies hypothesis. *Human Immunol.* 19, 155-162.

8. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley D.C. (1987a) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329, 506-512.
9. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., Wiley, D.C. (1987b) The foreign antigen binding site and T-cell recognition regions of class I histocompatibility antigens. *Nature* 329, 512-518.
10. Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. and Wiley, D.C. (1993) Three dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364, 33-39.
11. Spies, T., Bresnahan, M., Sands, J., Strominger, J.L. (1989) A new cluster of genes within the human major histocompatibility complex. *Science* 243, 214-217.
12. Trowsdale, J., Hanson, I., Mockridge, I., Beck, S., Townsend, A. and Kelly, A. (1990) Sequences encoded in the class II region of the MHC related to the 'ABC' super-family of transporters *Nature*. 348, 741-743.
13. Monaco, J.J. and McDevitt, H.O. (1986) The LMP antigens: a stable MHC-controlled multisubunit protein complex. *Hum Immunol.* 15, 416-426.
14. Bhasin, M.K., Walter, H., Danker-Hopfe, H. (1994) People of India: an investigation of biological variability in ecological, ethnoeconomic and linguistic groups New Delhi: Kamla Raj Enterprises.
15. Singal, D.P., Pitchappan, R.M., Agarwal, S., Mehra, N.K., Dupont, B., Mackintosh, P. and Suzuki, K. (1992) Distribution of HLA antigens in Asian Indians. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991*, Oxford: Oxford University Press. 658-660.
16. Mehra, N.K., Rajalingam, R., Giphart, M.J. (1996) Generation of DR51 associated DQA1, DQB1, haplotypes in Asian Indians. *Tissue Antigens* 47-85-89.
17. Mehra, N.K., Taneja, V., Kailash, S., Raizada N., Vaidya, M.C. (1986) Distribution of HLA antigens in a sample of the North Indian Hindu population. *Tissue Antigen* 27, 64-67.
18. Pitchappan, R.M., Manickasundari, M., Kakkanaih, V.N., Mahendren, V., Brahamajothi, V., Chandramohan, R., Subash Chander, M. and Rajaseker R. (1986) HLA antigens in Sourastrans of Madurai. In: *HLA in Asia Oceania, proceedings of III Asia Oceania Histocompatibility Workshop and conference*, (Ed) Aizawa M, Natori T, Wakisaka A, Konoeda Y. Sapporo, Japan: Hokkaido University Press. 539-542.
19. Mehra, N.K., Verduijn, W., Taneja, V., Drabbels, J., Singh, S.P.N. and Giphart M. (1991) Analysis of HLA-DR associated polymorphisms by oligonucleotide hybridization in an Asian Indian population. *Human Immunol.* 32, 241-247.
20. Mehra, N.K., Bouwens, A.G.M., Naipal, A., Rajalingam, R., Grubic, Z., Taneja, V., Tilanus, M.G.J. and Giphart M.J. (1994) Asian Indian HLA-Dr2, DR4 and DR52 related DR-DQ genotypes analysed by PCR based non-radioactive oligonucleotide typing: unique haplotypes and a novel DR4 subtypes. *Human Immunol.* 39, 202-210.
21. Mehra, N.K., Rajalingam, R., Kanga U., McEnemy, L., Cullen, C., Agarwal S., Middleton D., Pollack M.S., Amirzargar A. and Singal D.P. (1997) Genetic diversity of HLA in the populations of India, Sri Lanka and Iran. In: *Genetic Diversity of HLA: Functional and Medical Implications*, eds. Dominique Charron, EDK Publishers, Paris, Vol I, 314-320.
22. Kanga, U., Mehra, N.K., Larrea, C.L., Lardy, N.M., Kumar, A. and Feltkamp T.E.W. (1996) Seronegative spondyloarthropathies and HLA-B27 subtypes: A study in Asian Indians. *Clinical Rheumatology* 15, 13-18.

23. Taneja, V., Giphart, M.J., Verduijn, W., Naipal, A., Malaviya, A.N. and Mehra N.K. (1996) Polymorphism of HLA-DRB, -DQA1 and -DQB1 in rheumatoid arthritis in Asian Indians. Association with DRB1*0405 and DRB1*1001. *Human Immunol.*, 46,35-41.
24. Arnaiz-Villena, A., Chandanayingyong, D., Chiewsilp, P., Fan, L., Fauehet, R., Fumiaki, N., Gomez-Casado, E., Guo, S.S., Gyodi, E., Hammond, M., Jersild, C., Juji, T., Lepage, V., Martinetti, M., Martinez-Laso, J., Mehra, N., Penzes, M., Rajalingam, R., Rajczy, K. and Tokunaga. (1997) HLA-G polymorphism: 12th International Histocompatibility workshop study. In: *Genetic Diversity of HLA: Functional and Medical Implications*, eds. Dominique Charron, EDK Publishers, Paris, Vol. 1, 155-159.
25. Arnaiz-Villena, A., Chandanayingyong, D., Chiewsilp, P., Diaz-Campos., Fan, L., Fauehet, R., Fumiaki, N., Guo, S.S., Gyodi, E., Hammond, M., Jersild, C., Juji, T., Lepage, V., Martinetti M., Martinez-Laso, J., Mehra, N., Penzes, M., Rajalingam, R., Rajczy, K. and Tokunaga. (1997) HLA-E polymorphism : 12th International Histocompatibility workshop study. In: *Genetic Diversity of HLA: Functional and Medical implications*, eds. Dominique Charron, EDK Publishers, Paris, Vol. 1, 151-154.