Role of Human Leukocyte Antigens in Studying Population Diversity

Suraksha Agrawal, Sanjeev Srivastava and Piyush Tripathi

INTRODUCTION

The major histocompatibility complex (MHC) is a dense complex of genes with immunological and non-immunological functions and is present in all vertebrates. In humans it is known as human leukocyte antigens (HLA) (Marsh et al., 2000; Gruen and Weissman, 1997). Peter Gorer discovered it during transplantation studies in mice (the H-2 complex) in 1937. Jean Dausset described the first human MHC antigen *Mac* (HLA-A2) followed by the discovery of 4a and 4b. MHC is best known with its role in histocompatibility (Snell, 1981) and in immune regulation (Jorde, 1999). In humans, the genes for the HLA antigens are located on the short arm of chromosome 6 in the band 6p 21.3 (Apanius et al., 1997). It contains approx 4 million nucleotide pairs and contains approx 200 genes The MHC complex is divided into three subgroups called MHC class I, MHC class II, and MHC class III. The MHC class I genes are further divided into classical and non classical. The classical class I genes are located telomeric in the complex and encode the α-chain of the antigen-presenting HLA-A, -B and -C molecules. HLA-E, F and G are termed nonclassical genes and are juxtaposing the classical class I genes. The classical and non classical genes are evolutionarily related but appear to have distinct functions related to the immune response and NK cell recognition. The centromeric class II genes, previously known as immune-response genes include HLA-DR, -DP and -DQ genes, and proteasome (LMP) and peptide-transporter (TAP) genes. The MHC class III region contains many genes with different functions. Among the genes within the MHC are more than 20 loci encoding proteins involved in binding and presentation of the peptide degradation products of proteins to the T cell antigen receptor (Fig. 1).

Regarding the evolution of these genes there is not a definite candidate for the primordial MHC gene. According to one hypothesis the class II MHC evolved first (Hughes and Nei, 1992) whereas another hypothesis holds that the class I MHC originated first as a result of a recombination between an immunoglobulin-like C-domain and the peptide-binding domain of an HSP70 heat-shock protein (Flajnik et al., 1991). A phylogenetic analysis supports a relationship between the class II MHC alpha chain and beta 2-microglobulin and between the class II MHC beta-chain and the class I alpha chain (Hughes and Nei, 1992). Most evidence supports the hypothesis that the ancestral MHC molecule had a class II-like structure and it gave rise to the class I molecule (Hughes and Nei, 1992).

The best-known genes in the MHC region are the subset that encodes cell-surface antigen-presenting proteins. In humans, these genes are referred to as human leukocyte antigen (HLA) genes, although people often use the term MHC to refer to HLA gene products) (Gruen and Weissman, 1997). In mouse it is H-2 (Histocompatibility System – 2), in rabbits it is RLA (Rabbit Leukocyte antigens), in Guinea Pig it is GLA (Guinea Pig leukocyte antigens), in Chimpanzee it is ChLA (Chimpanzee Leukocyte Antigen) and in Cattle it is BoLA (Bovine Leukocyte Antigens).

In humans the most intensely studied HLA genes are the nine so-called classical MHC genes: HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA, and HLA-DRB1. The A, B, and C genes belong to MHC class I, whereas the six D genes belong to class II.

The MHC molecules have a vital role in the complex immunological dialog that must occur between T cells and other cells of the body (Benacerraf and McDevitt, 1972). At maturity, MHC molecules are anchored in the cell membrane, where they display short polypeptides to T cells, via the T cell receptors (TCRs) (Janeway and Travers, 1997). The polypeptides may be self, that is, originating from a protein created by the organism itself, or they may be foreign, originating from bacteria, viruses, pollen, etc.

The immune system has another and equally important method to identify antigen: B cells with their membrane-bound antibodies, also known as B cell receptors (BCRs). However, whereas the

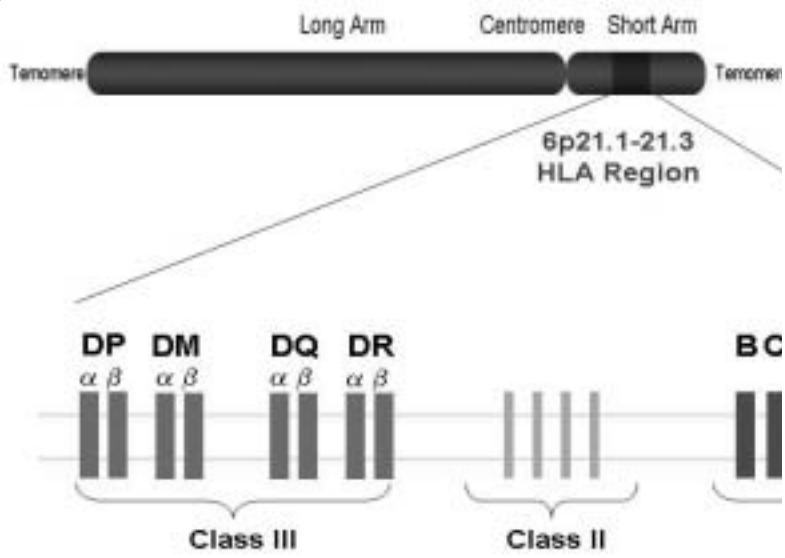

Fig. 1. Gene map of the human leukocyte antigen (HLA) region

BCRs of B cells can bind to antigens without much outside help, the TCRs of T cells require "presentation" of the antigen. It is important to realize that, during the vast majority of the time, MHC is kept busy presenting self-peptides, which the T cells should appropriately ignore. A full-force immune response usually requires the activation of B cells via BCRs and T cells via the MHC-TCR interaction. All MHC molecules receive polypeptides from inside the cells they are part of and display them on the cell's exterior surface for recognition by T cells. However, there are major differences between MHC class I and II in the method and outcome of peptide presentation. Besides being scrutinized by immunologists for its pivotal role in the immune system, the MHC has also attracted the attention of many evolutionary biologists, due to the high levels of allelic diversity found within many of its genes

CLASSIFICATION OF HLA

MHC Class I

MHC class I is further categorised in to classical and non classical antigens. MHC class I classical antigens i.e. A, B and C are found on almost every nucleated cell of the body except central nervous system, skeletal and smooth muscle cells, parathyroid cells, pancreatic cells and corneal epithelium. Both male and female germinal cells are also devoid of classical class I antigens. The non classical MHC class I antigens i.e. G, E and F are present on the placenta and extra villous membranes (Parham and Ohta, 1996). MHC class I molecules are heterodimers, consisting of a single transmembrane polypeptide chain (the α-chain) which is about 44 kD and a β_2 microglobulin is 12 KD protein. The alpha chain is non-covalently linked with Beta chain (Fig. 2A). The heavy α chain comprises of domains, α1, α2, α3 domains, a transmembrane region and a cytoplasm domain.

In all the three domains 90 amino acids are present seperately. α2 and α3 domains are linked with inter chain disulphide bonds. In α3 domains at position 86, a glycosylated asparagine residue is situated. In transmembrane region 23 hydrophobic amino acid residues are present due to which α-helical conformation is achieved. To the C terminal of the membrane arginine and lysine forms a cluster. These then get linked with the polypeptide chain in the membrane by interacting with negatively charged phospholipid groups of the inner membrane. The hydrophobic cytoplasmic domain consists of 30 amino acids; of which 50 are polar amino acids particularly serine and some are phosphorylated by a cAMP dependent protein kinase.

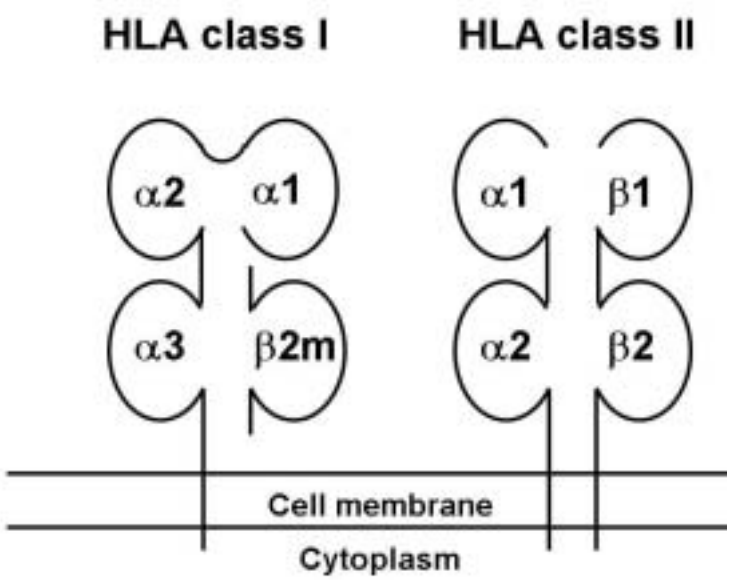

Fig. 2. Structure of MHC class I and class II antigens

Nomenclature of HLA

HLA molecules play central role in the immunesurveillance hence these are highly polymorphic in the nature. Conventionally the HLA typing was dependent on serology and limited polymorphism but with the advent of molecular techniques more and more alleles have been discovered. The total number of known alleles is shown in the Table1.

The new nomenclature as been established according to which the first two digits of an allele name refer to the underlying serological specificity and the third and fourth digits indicate a specific allele sequence. For example, HLA-A*0205 and A*0210 are alleles encoding distinct polypeptides within the A2 serotype. These two alleles both encode the epitope recognized by the anti-A2 antisera but have 5 nucleotide differences elsewhere in exons 2–3 resulting in amino acid variations. Alleles within a serological group may vary from each other by a single or by several nucleotides. For Class II molecules, both the A and the B genes may contribute to antigen variability.

Thus a DR15 serotype may be found in an individual with one of the DRB1*15 alleles such as DRB1*1501 and DRA*0101. When necessary, a fifth digit is used to identify silent polymorphisms and the sixth and seventh digits are used to denote variation occurring outside of coding regions, such as the promoter and introns. Null allele sequences which result in either no or reduced levels of functional HLA molecules because of transcription changes, aberrant RNA splicing, and frame shift and nonsense mutations, or in frame termination codons are designated by an allele number appropriate to the group and the letter N. Individuals with null alleles may have discrepancies between serological and DNA-based typing.

Table 1: HLA loci and known alleles

Generic locus	*Antigen or associated specificity*	*Number of known alleles*
HLA-A	A1 to A80	489
HLA-B	B7 to B81	830
HLA-C	Cw1 to Cw10	266
HLA-E	-	9
HLA-F	-	20
HLA-G	-	23
DRA	DR1 to DR18	3
DRB1	-	463
DQA1	DQ1 to DQ9	34
DQB1	-	78
DPA1	DPw1 to DPw6	23
DPB1	-	125

Source: http://www.ebi.ac.uk/imgt/hla/

Diversity of MHC Alleles

The major histocompatibility complex (MHC)

loci are known to be highly polymorphic in humans, mice and certain other mammals, with heterozygosity as high as 80-90%. These properties make HLA genes a good candidate for studying population diversity. Four different hypotheses have been considered to explain this high degree of polymorphism:

(1) A high mutation rate
(2) Gene conversion or inter locus genetic exchange
(3) Over dominant (balancing) selection
(4) Frequency-dependent selection.

The distribution of the pattern of sequence polymorphism in human and mouse class I genes provides evidence for four co-ordinate factors that contribute to the origin and sustenance of abundant allele diversity that characterizes the MHC in the species. These include: (a) a gradual accumulation of spontaneous mutational substitution over evolutionary time but not an unusually high mutation rate; (b) selection against mutational divergence in regions of the class I molecule involved in T cell receptor interaction and also in certain regions that interact with common features of antigens; (c) positive selection pressure in favor of persistence of polymorphism and heterozygosity at the antigen recognition site; and (d) periodic intragenic (interallelic) and more rarely, intergenic, recombination within the class I genes.

Evolutionary interplay between mutation and recombination varies with MHC locus, and even for subregions of the same gene (Parham and Ohta, 1996; Hughes and Nei, 1998; Hughes and Nei, 1989). For example, phylogenetic inferences based on the exon 2 region of HLA-DRB loci are complicated by selection and recombination (gene conversion). Noncoding region analysis may help clarify patterns of allele evolution usually with contrasting results to those obtained from coding region analyzes (Hickson and Rebecca, 1997). The main source for the variability in the HLA gene sequences is point mutation but the mutation rate is by no means higher in the MHC than elsewhere in the genome (Lawlor et al., 1988, Parham and Ohta, 1996). Because of trans species polymorphism, accumulation of point mutations over millions of years results in extensive polymorphism. In contrast, gene conversions have produced at least 80 new class I alleles since the separation of the Homo lineage and the rate of conversion is much higher than that of point mutation (Little and Parham 1999; Marsh et al., 2000) **(**Fig. 3). Alleles arise from existing alleles through several postulated**.** DRB1*1120 likely arose via interallelic recombination between DRB1*1302 and a *DRB1**11 allele (Cizman et al., 1996) DRB1*0811 probably is derived from DRB1*0802 by a point mutation (Williams et al., 1994).

Nature of Class I and II Gene Polymorphism

Multiple alleles are found within most of the

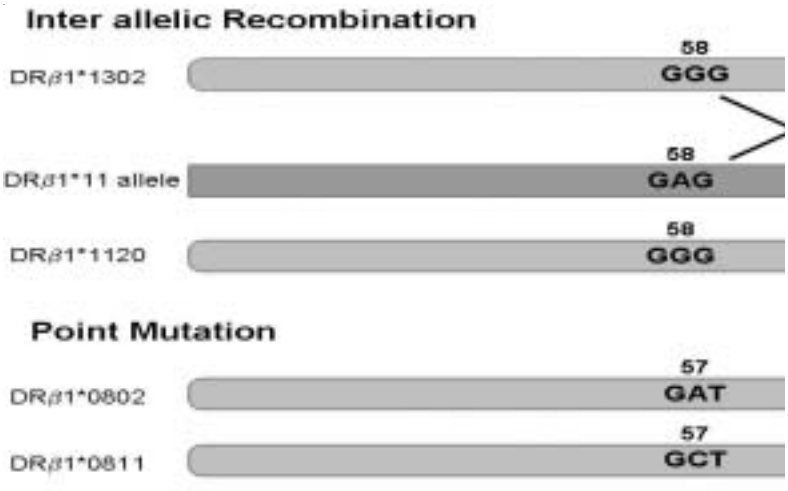

Fig. 3. Mechanism of generation of new alleles

known serotypes, although a few serotypes (an example is DR9, *DRB1**09012) are accounted for solely by a single allele. For example the B35 serotype has more than 39 alleles. HLA allele frequencies exhibit ethnic variation, with some alleles found widely distributed among populations and others almost exclusively within a particular ethnic group. The number of different phenotypes that are possible from all combinations of the known HLA alleles is greater than the earth's population. However, the Class I and II loci reside on a relatively small region of chromosome 6 and specific haplotypes were apparently present at high frequencies in founding populations or were selected for or against by infectious organisms. In this setting, linkage disequilibrium results in a significant over representation of certain haplotypes (Alper et al 1992) (Fig. 4).

Single letter amino acid codes are shown for DRB1 exon 2 codons 6–94. About half of the positions are invariant while the remainder displays polymorphism with a few codons encoding as many as seven different amino acids. For example, all *DRB1* alleles have glycine encoded by position 20 while alleles may encode glycine, valine, or aspartic acid at codon 86. *DRB1* alleles arise through the many possible combinations of these polymorphisms.

Mechanisms Maintaining the Extreme Polymorphism of the MHC

1. Pathogen Driven Mechanism

Pathogen-driven selection favors genetic diversity of the MHC through both heterozygote advantage (over dominance) and frequency-dependent selection (Potts and Wakeland, 1993). Selection is thought to favor rare MHC genotypes, since pathogens are more likely to have developed mechanisms to evade the MHC-dependent immunity encoded by common MHC genotypes. Six molecular models of pathogen-driven selection have been presented.

A. Pathogen Evasion Models Escape of a single T-cell clone recognition
   Escape into holes in the T-cell repertoire produced by T-cells energized by pathogen variants
   Escape into holes in the T-cell repertoire induced by self-tolerance
   Escape of MHC presentation
B. Host-Pathogen Interactions:
   Heterozygote advantage
   Pathogens bearing allo-MHC antigens

MHC associations with specific infectious diseases have been difficult to demonstrate. The best known one is malaria in humans (Hill et al.,

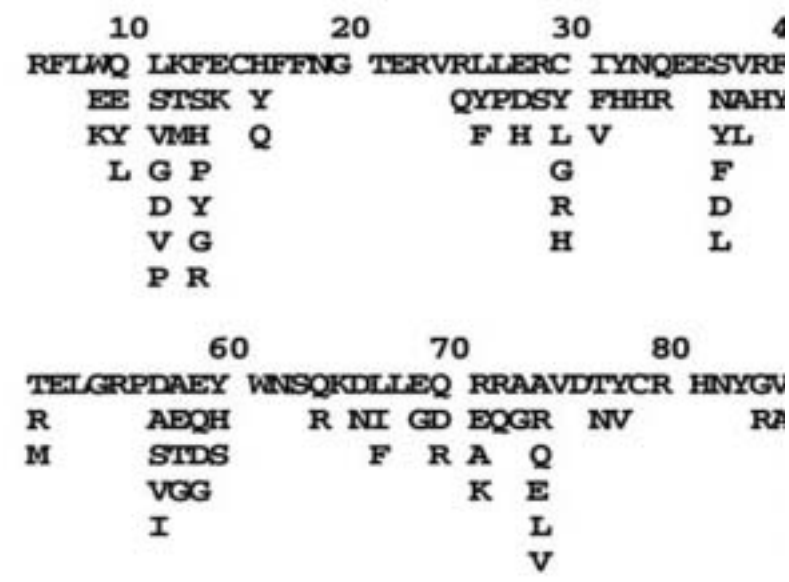

Fig. 4. DRB1 Gene polymorphism (DRB1 locus exon-2 codons)

1991). Since most infectious agents have multiple epitopes which MHC has to deal (Fienberg, 1970), Rather than resistance of specific heterozygous genotypes to specific. In human heterozygote advantage have been reported for a specific genotype in HIV infection (Carrington et al., 1999) and in HBV infection (Thursz et al., 1997)

2. Non-pathogen Driven Mechanisms

MHC is exploited to discriminate against genetic similarity at highly polymorphic loci to avoid inbreeding. MHC-based disassortative mating would produce heterozygous progeny at least at the MHC which would result in increased fitness **(**Ober and van der Ven, 1997**).** Progeny derived from MHC-dissimilar parents would have high fitness because of reduced levels of inbreeding depression and increased resistance to infectious diseases due to high MHC heterozygosity. This selection contributes to the high levels of genetic polymorphism observed at the MHC loci.

Evidence for Selection on MHC Alleles

1. One important feature of the MHC genes is that the ratio of non-synonymous (replacement) to synonymous (silent) substitutions (dn /ds ratio) is very high in the codons encoding the antigen recognition site of polymorphic class II molecules compared to other codons (Hughes et al., 1994). This pattern is evidence that the polymorphism at the antigen recognition sites is maintained by over dominant selection of which the most common form is heterozygote advantage. This kind of selection has been noted for all expressed DRB genes including DRB3 and DRB4 (Klein et al., 1991). This feature and the others such as (1) an extremely large number of alleles; (2) ancient allelic lineages that predate contemporary species (trans-species evolution) and; (3) extremely high sequence divergence of alleles make the MHC a unique system in the whole genome.
2. The expected number of alleles under neutrality is far lower than the number of MHC alleles observed in natural population which indicates that some form of balancing (diversifying) selection is also acting (Hedrick, 1983, 1994). For a subdivided population over a large range of migration rates, it appears that the number of self-incompatibility alleles (or MHC-alleles) observed can provide a rough estimate of the total number of individuals in the population but it underestimates the neutral effective size of the subdivided population (Schierup, 1998; Schierup et al., 2000).
3. The large number of alleles showing a relatively even distribution is against neutrality expectations and indicates that diversifying, and not simply directional, selection operates in contemporary populations.
4. The observed deficiency of homozygotes in humans indicates that selection favors heterozygotes, because of high immune response. When the amino acid heterozygosities per site for HLA-A and -B loci were determined, for the 54 amino acid sites thought to have functional importance, the average heterozygosity per site was 0.301. Sixteen positions have heterozygosities greater than 0.5 at one or both loci and the frequencies of amino acids at a given position are very even, resulting in nearly the maximum heterozygosity possible. High heterozygosity is concentrated in the peptide-binding sites, whereas the sites that interact with the T-cell receptor have lower heterozygosity. Overall, these results indicate the importance of some form of balancing selection operating at HLA loci, maybe even at the individual amino acid level (Hedrick, 1994).
5. The observed linkage disequilibrium among tightly linked MHC genes suggests that the strength of selection is uneven within the MHC (Apanius et al., 1997).
6. Studies in West Africa showed that resistance against malaria is HLA-B53 associated and this is the reason for an increased frequency of B53 in that area. The selection differential for HLA-B*5301 is estimated to be 0.028 (Hill et al., 1991). Apanius et al suggested that MHC haplotypes can confer resistance to infectious diseases that outweighs the deleterious effects from autoimmunity (Apanius et al., 1997). Another hypothesis for the maintenance of autoimmune-predisposing MHC haplotypes is that these alleles protect against initial infection, but the pathogen triggers autoimmunity through molecular mimicry.

Allelic Diversity and MHC Evolution

MHC gene families are found in essentially

all vertebrates, though the gene composition and genomic arrangement vary widely. Chickens, for instance, have one of the smallest known MHC regions (19 genes), though most mammals have an MHC structure and composition fairly similar to that of humans. Gene duplication is almost certainly responsible for much of the genic diversity. In humans, the MHC is littered with many pseudogenes.

One of the most striking features of the MHC, particularly in humans, is the astounding allelic diversity found therein, and especially among the nine classical genes. In humans, the most conspicuously-diverse loci, HLA-A, HLA-B, and HLA-DRB1, have 489, 830 and 463 known alleles respectively this allelic diversity is truly exceptional in the human genome. The MHC gene is the most polymorphic in the genome. One remarkable featureof HLA Loci is that many of these alleles are quite ancient. It is often the case that an allele from a particular HLA gene is more closely related to an allele found in chimpanzees than it is to another human allele from the same gene.

Phylogenetically the marsupial MHC lies between eutherian mammals and the minimal essential MHC of birds, although it is closer in organization to non-mammals. It's Class I genes have amplified within the Class II region, resulting in a unique Class I/II region.

The allelic diversity of MHC genes has created fertile grounds for evolutionary biologists (Arnaiz-Villena et al., 1999). The most important task for theoreticians is to explain the evolutionary forces that have created and maintained such diversity. Most explanations invoke balancing selection, a broad term that identifies any kind of natural selection in which no single allele is absolutely most fit. Frequency-dependent selection and heterozygote advantage are two types of balancing selection that have been suggested to explain MHC allelic diversity.

HLA and Population Diversity

Since long geneticists are involved in the Human genetic variation studies among the individuals forming a species, but the remarkable extent of this variation was not appreciated until about 25 years ago (Cavalli-Sforza and Feldman, 2003). Conspicuous human traits like hair and eye color clearly vary from one individual to the other in many populations; these differences are easily perceived by the layman, as are variation in height, weight, body build, and facial traits, which are also genetically determined to some extent. Their hereditary transmission, however, is complex, and these traits contribute little to understand the extent of variation. The first example of clear-cut genetic variation is that of ABO blood groups which was described at the beginning of the century (Landsteiner and Levine, 1927). Dissimilarities between individuals regarding ABO blood-group variation are due to small chemical differences between molecules found at the surface of red blood cells.

The existence of genetic variation among human population was first demonstrated by Hirszfeld and Herszfeld (1919), in the first human gene to be described-ABO that determines the ABO blood groups (Hirszfeld and Herszfeld, 1919). The subsequent identification of blood group protein markers like MNS and Rh expanded the repertoire of polymorphic markers that were analyzed using immunological techniques. Pauling et al. (1949), has introduced the electrophoresis technique for the analysis of protein variant among human groups (Pauling et al., 1949). Dramatic improvements in genotyping technologies over the past 3 decades, with the emergence of PCR, have facilitated the development of many types of DNA markers. Depending upon the type these could be STRs, SNPs and indels, where as based on their chromosomal origin these could be X, Y, autosomal, and mitochondrial.

Considerable attention has been devoted to both uniparental and autosomal genetic markers. Because of their lack of recombination, uniparental markers mtDNA and the nonrecombining region of the Y chromosome are perhaps easier for tracing maternal and paternal lineages than are recombining markers (Ingman et al., 2000; Underhill et al., 2000). Further, regarding the autosomal markers, recombination may mislead about their history but simultaneous use of an array of autosomal polymorphic loci spread across the genome provides more general inference about demographic history and population relationships compared to gender specific markers of Y chromosome and mtDNA. However, in last 3 decades, with the emergence of PCR, restriction enzyme technology and sequencing methods has revolutionized the studies of genetic variation. Now DNA based markers like RFLPs, Alu insertions, STRs, SNPs and indels are preferred

for their neutral aspect of information. Every molecular marker has certain advantage and disadvantage. Since the discovery of role of HLA polymorphism in population diversity these highly polymorphic and rarely recombining HLA class I and II loci remained most commonly used markers in the population studies.

HLA: A Useful Marker for Population Studies

Molecular HLA-typing has proved to be an invaluable tool in studying the evolutionary origin of human populations (Arguellolo et al., 1998; Luo et al., 1999; Albis-Camps and Blasczyk, 1999). This information, in turn, contributes to the understanding of cultural and linguistic relationships and practices among and within various ethnic groups. This has become possible due to the continual discovery of new HLA alleles using DNA technology which has increased the power of HLA to distinguish individuals. It has been postulated that gene conversion events are the main mechanisms for distributing and reshuffling sequences among alleles. In addition reciprocal recombination and point mutations have been suggested to be responsible for the generation of alleles over evolutionary time (Mason and Parham, 1998).

Human past can be read through a history written in HLA. The HLA polymorphism has proven to be useful for singling out individuals and populations. The discoveries of new loci and presently available DNA typing and sequencing of new alleles have dramatically increased the variety of HLA alleles. Certain alleles are frequent only in specific populations e.g. A36 and A43 in African – Americans. Strong linkage disequilibrium between HLA neighboring loci demonstrates that certain combinations of contiguous alleles (HLA haplotypes) also occur more frequently in a population compared to others. (Gottelli et al., 1994; Paetkau and Strobeck, 1994; Taylor et al., 1994). They can be used to estimate effective population size and to gain insight into the degree of population substructure including both the amount of migration between subpopulations (Gottelli et al., 1994) and genetic relationships among the various subpopulations (Bowcock et al., 1994; Estoup et al., 1996; Bamshad et al., 2001).

Some indigenous populations (e.g. groups of Native Americans or from Papua New Guinea) show a very restricted diversity of alleles at DRB1 as well as other HLA loci (Inman and Rudin 1997). The extensive polymorphism of the major histocompatibility complex (MHC) genes in humans and the differential allelic distribution in ethnic populations of varied origin has been major focus of immuno genetic research. The presence of certain alleles with high frequency only in specific populations (e.g. A36, A43 African Americans) and the strong linkage disequilibrium between HLA neighboring loci, demonstrates that certain combinations of contiguous alleles (HLA haplotypes) show a characteristic frequency or are distinctive in certain living populations (Browning and McMichael, 1996). The wide range of allelic diversity and the conserved combinations of different alleles are used as genetic markers and anthropological data is based on the information supplied by population studies (Saruhan-Direskeneli et al., 2000). Population studies indicate that there are many alleles and DR-DQ haplotypes that appear to be specific for given ethnic group. The existence of ancestral haplotypes implies conservation of large chromosomal segments.

The extra ordinary power of this small segment of the human genome clusters population in a manner expected from linguistic, anthropological and archaeological evidences. Sequences reveal a dramatic level of diversity. Species specific residues i.e. residues that identify a MHC molecule as belonging to a particular species, are extremely rare. In contrast, species unique residues i.e. residues that were not characteristic of a species are unique to individuals of that species appear somewhat more frequently. The allele frequency distributions and patterns of variability within the molecule, suggest strong selective forces acting on class II loci and I. The form of selection is unknown and potential selective mechanisms should be examined in the light of classical population genetic theory, which states that it is very difficult to maintain so many alleles even with strong balancing selection. Thus, HLA variability at the level of DNA is useful in unraveling the evolutionary relationships between populations and in investigating the evolutionary forces which shaped the genetic profiles of contemporary populations (Inman and Rudin, 1997).

The HLA polymorphisms have been created because of balancing selections, which maintain a few allelic lines over very long period (Harpending et al., 1998). The extensive variation in HLA

markers makes the system highly useful for determining genealogical relationships between populations. Monsalve et al. in 1999 have successfully compared the relationship between linguistic and genetic data in Native Americans and Asian populations (Monsalve et al., 1999). They have concluded that gene flow and genetic drift are important factors in shaping the genetic landscape of Native American populations. The results are most congruent with the single migration model. In addition the understanding of the events contributing to MHC class II evolution requires comparison of species that are very closely related. The distribution of alleles in different populations can be used to construct a matrix of genetic distances between populations and a phylogenetic tree or unrooted network in order to examine the historical / evolutionary relationships between these groups.

The high polymorphism, tight linkage among loci and the random association of alleles make the system of particular interest from the perspective of population genetics. Information on the dynamic evolutionary forces that have acted on a locus can be inferred from the number and distribution of alleles that it carries. The major histocompatibility complex (MHC) is unique in the number of highly polymorphic loci spread over such a small chromosomal region. This creates a context for interpreting HLA region variation in both evolutionary and in clinical terms. The extensive allelic variation among the HLA class I and class II genes distinguishes these as the most polymorphic coding sequence.

It is known that HLA Class II is highly polymorphic as at DRB1 locus 463 alleles, at DQA1 locus 34 alleles and at DQB1 locus 78 alleles have been identified worldwide. The world map as per different continents is shown in Figure 5.

Various alleles of class II are widely distributed in different continents of worlds like Asians, Europeans, Africans, Americans and Australians. Our study (unpublished data) has also shown high diversity of HLA class I and II alleles among Indian populations. This is in accordance to other studies on different world populations (Sanchez-Velasco et al., 2003, Zhou et al., 2005, Lee et al., 2005, Farjadian et al., 2006, Abdennaji et al., 2006).

Thus it could be concluded that the amalgamation of the genotypic and haplotypic organization of HLA Class II loci is able to decipher vital information about the amount, pattern and distribution of genetic variation in different populations. The perspective correlation of the genetic profile of studied populations with the past human movements and other historical records facilitated determination of their genetic ancestry/origin.

Fig. 5. World map showing different continents

REFERENCES

Abdennaji, G.B., Yacoub, I., Loueslati, B., Buhler, S., Hmida, S. et al.: Sanchez-Mazas HLA class II genetic diversity in southern Tunisia and the Mediterranean area *International Journal of Immunogenetics*, **33:** 93 (2006).

Albis-Camps, M. and Blasczyk, R.: Fluorotyping of HLA-DRB by sequence specific priming and fluorogenic probing. *Tissue Antigens*, **53:** 301-307 (1999).

Alper, C.A., Awdeh, Z. and Yunis, E.J.: Conserved, extended MHC haplotypes. *Exp. Clin. Immunogenet.*, **9:** 58–72 (1992).

Apanius, V., Penn, D., Slev, P.R. and Ruff, Potts: The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology*, **17:** 179-224 (1997).

Arguellolo, J.R., Little, A.M., Bohan, E., Goldman, J.M., Marsh, S.G.E. and Madrigal, J.A.: High esolution HLA class I typing by Reference strand mediated conformation analysis (RSCA). *Tissue Antigens*, **52:** 57-66 (1998).

Arnaiz-Villena, A., Iliakis, P., Gonzalez-Hevilla, M., Longas, J., Gomez-Casado, E., Sfyridaliv, K., Trapaga, J., Silvera-Redondo, C., Matsouka, C. and Martinez-laso, J.: The origin of Cretan populations as determined by characterisation of HLA alleles. *Tissue Antigens*, **53:** 213-226 (1999).

Bamshad, R.M., Naidu, M., Prasad, B.V.R., Reddy, P.G., Rasanayagam, A. and Papiha, S.S., et al.: Genetic Evidence on the Origins of Indian Caste Populations. *Genome. Res. J.*, **11:** 994-1004 (2001).

Benacerraf, B. and McDevitt, H.O.: Histocompatibility linked immune response genes. *Science*, **175:** 273-279 (1972).

Bowcock, A.M., Riuz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R. and Cavalli-Sfroza, L.L. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368:** 455-457 (1994).

Browning, M. and McMichael, A.: In *HLA and MHC-genes, molecules and function*. BIOS Scientific Publishers (1996).

Carrington, M., Nelson, G.W., Martin, M.P., Kissner, T., Vlahov, D. and Goedert, J.J., et al.: HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage. *Science*, **283:**1748–1752 (1999).

Cavalli-Sforza, L.L. and Feldman, M.W.: The Application of Molecular Genetic Approaches to the Study of Human Evolution. *Nature Genetics*, **33:** 266-275 (2003).

Cizman, B.B., McKean, M.E., Kearns, D.J., Heron, S.D., Wu, J. and Griffith, B.B. et al.: DRB1*1120 allele; another example of the transition between the DR11 and DR13 families of alleles. *Tissue Antigens*, **48:** 52–54 (1996).

Estoup, A., Solignac, M., Cornuet, J.M., Goudet, J. and Scholl, A.: Genetic differentiation of continental and island populations of Bombus terrestris in Europe. *Mol. Ecol.*, **5:** 19-31 (1996).

Farjadian, S., Moqadam, F.A. and Ghaderi, A.: HLA class II gene polymorphism in Parsees and Zoroastrians of Iran. *International Journal of Immunogenetics*, **33:** 185-191 (2006).

Fienberg, S.E.: The analysis of multidimensional contingency tables. *Ecology*, **51:** 419-433 (1970).

Flajnik, M.F., Canel, C., Kramer, J., Kasahara, M.: Evolution of the major histocompatibility complex: molecular cloning of major histocompatibility complex class I from the amphibian Xenopus. *Proc Natl. Acad. Sci. U. S. A.*, **88:** 537–541 (1991).

Gottelli, D., Sillero-Zubiri, C., Applebaum, G.D., Roy, M.S., Girman, D.J. and Garcia-Moreno, J., et al.: Molecular genetics of the most endangered Canid: the Ethiopian wolf Canis simensis. *Mol. Eco.*, **3:** 301-312 (1994).

Gruen, J.R. and Weissman, S.M.: Evolving views of the major histocompatibility complex. *Blood*, **90:** 4252-4265 (1997).

Harpending, H.C., Batzer, M.A., Gurven, M., Jorde, L.B., Rogers, A.R. and Sherry, S.T.: Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. U.S.A.*, **1**7**:** 1961-1967 (1998).

Hedrick, P.W.: Neutrality or selection at HLA? *Am. J. Hum. Genet.*, **35:** 1055–1057 (1983).

Hedrick, P.W.: Evolutionary genetics of the major histocompatibility complex. *Am. Nat.*, **143:** 945-964 (1994).

Hickson, R. and Rebecca, L.C.: Mhc Allelic Diversity and Modern Human Origins *Journal of Molecular Evolution*, **45:** 245-250 (1997).

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A. et al.: Common west African HLA antigens are associated with protection from severe malaria. *Nature*, **352:** 595-600 (1991).

Hirszfeld, L. and Herszfeld, H. Essal d'application des methods au probleme des races. *Anthropologie*, **29:** 505-17 (1919).

Hughes, A.L. and Nei, M.: Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. *Mol. Biol. Evol.*, **6:** 559-579 (1989).

Hughes, A.L., Hughes, M.K., Howell, C.Y. and Nei, M.: Natural selection at the class II major histocompatibility complex loci of mammals. *Philos. Trans. R. Soc. Lond. B.* **346:**359-367 (1994).

Hughes, A.L. and Nei, M.: Maintenance of MHC polymorphism. *Nature*, **355:** 402-403 (1992).

Hughes, A.L. and Nei, M.: Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, **335:** 167-170 (1998).

Ingman, M., Kaessmann, H., Paabo, S. and Gyllensten, U.: Mitochondrial genome variation and the origin of modern humans. *Nature*, **408:**708–713 (2000).

Inman, K. and Rudin, N.: *In An Introduction to Forensic DNA analysis*, CRC Press (1997).

Janeway, C.A. and Travers P.: *Antigen recognition by T lymphocytes. Immunobiology.* Pp. 41–46 3 ed . Garland Publishing,New York and London (1997).

Jorde, L.B.: "*Immunogenetics" in Medical Genetics.* 2nd ed. Moseby, St. Louis (1999).

Klein, J., et al., : Frozen haplotypes in MHC evolution in Molecular Evolution of the Major Histocompatibility Complex J. Klein and D. Klein (Eds.) (1991).

Landsteiner, K. and Levine, P.: Further observations on individual differences of human blood. *Proc. Soc. Exp. Biol. Med.*, **77:** 402-405 (1927).

Lawlor, D.A., Ward, F.E., Ennis, P.D., Jackson, A.P. and Parham, P.: HLA-A and B polymorphisms predate the divergence of humans and chimpanzees. *Nature*, **335:** 268-271 (1988).

Lee, K.W., Oh, D.H., Lee, C. and Yang, S.Y.: Allelic and haplotypic diversity of HLA-A, -B, -C, -DRB1, and -DQB1 genes in the Korean population. *Tissue Antigens*, **65:** 437-447 (2005).

Little, A.M. and Parham, P.: Polymorphism and evolution of HLA class I and II genes and molecules. *Reviews in Immunogenetics*, **1:** 105-123 (1999).

Luo, M., Blanchard, J., Pan, Y., Brunham, K., Brunham, R.C.: High resolution sequence typing of HLA-DQA1 and DQB1 Exon2 DNA with taxonomy based sequence analysis (TSBA)allele assignment. *Tissue Antigens*, **54:** 69-82 (1999).

Marsh, S.G.E., Bodmer, J.G., Albert, E.D., Bodmer, W., Bontrop, R.E., Dupont, B., et al.: Nomenclature for factors of the HLA system. *Eur. J. Immun.*, **28:** 377 (2000).

Mason, P. and Parham, P.: HLA class I region sequences. *Tissue Antigens*, **51:** 417-466 (1998).

Monsalve, M.V., Booth, P.B. and Cavalli-Sforza, L.L.: Languages, Geography and HLA haplotypes in native American and Asian Populations. *Proc. R. Soc. Lond. B. Biol. Sci.*, **266:** 2209-2216 (1999).

Ober, C. and van, der, Ven, K.: Immunogenetics of reproduction: an. overview. In: *Current Topics in Microbiology and Immunology.* L. B. Olding (Ed.) (1997).

Paetkau, D. and Strobeck, C.: Microsatellite analysis of genetic variation in Black bear populations. *Mol. Ecol*., **3:** 489-495 (1994).

Parham, P. and Ohta, T.: Population Biology of Antigen Presentation by MHC Class I Molecules. *Science*, Volume No. 272 -273 (1996).

Pauling, L., Itano, H.A., et al.: Sickle cell anemia a molecular disease. *Science.* **110:** 543-8 (1949).

Potts, W.K., Wakeland, E.K.: Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends Genet.* **9(12):** 408-12 (1993).

Sanchez-Velasco, P., Gomez-Casado, E., Martinez-Laso, J., Moscoso, J., Zamora, J., Lowy, E., et al.: HLA alleles in Basques and neighbouring populations. *Tissue Antigens*, **61:** 384-392 (2003).

Saruhan-Direskeneli, G., Uyar, F.A., Bakar, S. and Eraksoy, M.: Molecular analysis of HLA-DRB1, DQA1 and DQB1 polymorphism in Turkey. *Tissue Antigens*. **2:** 171 (2000).

Schierup, M.H., Charlesworth, D. and Vekemans, X.: The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. *Genetical Research*, **76:** 63-73 (2000).

Schierup, M.H.: The number of self-incompatibility alleles in a finite, subdivided population. *Genetics*, **149:** 1153-1162 (1998).

Snell, G.D.: Studies in histocompatibility. *Science*, **10:** 172-8 (1981).

Taylor, A.C., Sherwin, W.B. and Wayne, R.K.: Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat Lasiorhinus krefftii. *Mol. Ecol.* **3:** 277-290 (1994).

Thursz, M.R., Thomas, H.C., Greenwood, B. M. and Hill, A.V.S.: Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nature Genetics*, **17:** 11-12 (1997).

Underhill, P.A., Shen, P., Lin, A.A., Jin, L., Passarino, G., Yang, W.H. et al.: Y chromosome sequence variation and the history of human populations. *Nat Genet.*, **26:** 358-61 (2000).

Williams, T.M., Wu, J., Foutz, T., McAuley, J., Troup, G.M.: A new DRB1 allele (DRB1*0811) identified in Native Americans. *Immunogenetics*, **40:** 314 (1994).

Zhou, L., Lin, B., Xie, Y., Liu, Z., Yan, W. and Xu, A.: Polymorphism of human leukocyte antigen-DRB1, -DQB1, and -DPB1 genes of Shandong Han population in China. *Tissue Antigens*, **66:** 37-43 (2005).

KEYWORDS MHC. HLA. Population Diversity. Polymorphism

ABSTRACT Human Leukocyte Antigen (HLA) loci widely known for their role in generation of immune responses are often considered to be effective in reconstructing human phylogenies due to high degree of polymorphism and rarity of recombination observed at HLA loci. In this review, we have made an attempt to discuss the role of HLA in immune response. Further we have tried to highlight the evolutionary significance of HLA diversity. The reason of high degree of HLA diversity is because of gene conversion and recombination. These evolutionary forces lead the HLA molecule to be most the most suitable marker to study the population diversity.

Authors' Address: **Suraksha Agrawal, Sanjeev Srivastava and Piyush Tripathi,** Department of Medical Genetics, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Raebareli Road, Lucknow (UP) 226 014, Uttar Pradesh, India

Telephone: 091-522 -2668004-8 Ext 2338, 2346, 2347, *Fax:* 091-522 -2668973/2668017

E-mail: suraksha@sgpgi.ac.in