IMMUNOLOGY

Elementary Knowledge of Major Histocompatibility Complex and HLA Typing

Tapasya Srivastava and Subrata Sinha
Department of Biochemistry
All India Institute of Medical Sciences
New Delhi - 110029

CONTENTS

Keywords

Major Histocompatibility Complex, T-cell receptor, MHC-TCR interaction, HLA, HLA typing, allelic diversity

What is MHC?

The major histocompatibility complex (MHC) is a large genomic region or gene family found in most vertebrates. It is the most gene-dense region of the mammalian genome and plays an important role in the immune system, autoimmunity, and reproductive success. The proteins encoded by the MHC are expressed on the surface of cells in all jawed vertebrates, and display both self antigens (peptide fragments from the cell itself) and nonself antigens (e.g. fragments of invading microorganisms) to a type of white blood cell called a T cell that has the capacity to kill or co-ordinate the killing of pathogens, infected or malfunctioning cells.

Broad Functions of MHC Molecules

The MHC proteins act as "signposts" that display fragmented pieces of an antigen on the host cell's surface. These antigens may be self or nonself. If they are nonself, there are two ways by which the foreign protein can be processed and recognized as being "nonself".

If the host is a leukocyte, such as a monocyte or neutrophil, it may have engulfed the particle (be it bacterial, viral, or particulate matter), broken it apart using lysozymes, and displayed the fragments on Class II MHC molecules.

On the other hand, if a host cell was infected by a bacterium or virus, or was cancerous, it may have displayed the antigens on its surface with a Class I MHC molecule. In particular, cancerous cells and cells infected by a virus have a tendency to display unusual, nonself antigens on their surface. These nonself antigens, regardless of which type of MHC molecule they are displayed on, will initiate the specific immunity of the host's body.

It is important to note that cells constantly process endogenous proteins and present them within the context of MHC I. Immune effector cells are trained not to react to self peptides within MHC, and as such are able to recognize when foreign peptides are being presented during an infection/cancer.

Classification of MHC Molecules

In humans, the 3.6 Mb (3 600 000 base pairs) MHC region on chromosome 6 contains 140 genes between flanking genetic markers MOG and COL11A2 Subgroups.

The MHC region is divided into three subgroups called MHC class I, MHC class II, and MHC class III (Table 1).

Class III has a very different function than class I and II, but it has a locus between the other two (on chromosome 6 in humans), so they are frequently discussed together.

Table 1: MHC classification and function

Name	Function	Expression
MHC class I	Encodes heterodimeric peptide binding proteins, as well as antigen processing molecules such as TAP and Tapasin.	All nucleated cells. MHC class I proteins contain an a chain & b2-micro-globulin b chain. They present antigen fragments to cytotoxic T-cells cells and will bind to CD8 on cytotoxic T-cells.
MHC class II	Encodes heterodimeric peptide binding proteins and proteins that modulate peptide loading onto MHC class II proteins in the lysosomal compartment such as MHC II DM, MHC II DQ, MHC II DR and MHC II DP.	On antigen-presenting cells MHC class II proteins contain a & b chains and they present antigen fragments to T-helper cells by binding to the CD4 receptor on the T-helper cells.
MHC class III region	Encodes for other immune components, such as complement components (e.g., C2, C4, factor B) and some that encode cytokines (e.g., TNF-α) and also hsp.	Variable

MHC Presentation and Interaction with TCR

The classical MHC molecules (also referred to as HLA molecules in humans) have a vital role in the complex immunological dialogue that must occur between T cells and other cells of the body. At maturity, MHC molecules are anchored in the cell membrane, where they display short polypeptides to T cells, via the T cell receptors (TCRs). The polypeptides may be "self," that is, originating from a protein created by the organism itself, or they may be foreign ("nonself"), originating from bacteria, viruses, pollen, etc. The overarching design of the MHC-TCR interaction is that T cells should ignore self peptides while reacting appropriately to the foreign peptides.

The immune system has another and equally important method for identifying an antigen: B cells with their membrane-bound antibodies, also known as B cell receptors (BCRs). However, whereas the BCRs of B cells can bind to antigens without much outside help, the TCRs of T cells require "presentation" of the antigen: this is the job of MHC. It is important to realize that, during the vast majority of the time, MHC are 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 (Fig. 1). This duplicity creates a system of "checks and balances" and underscores the immune system's potential for running amok and causing harm to the body (see autoimmune disorders).

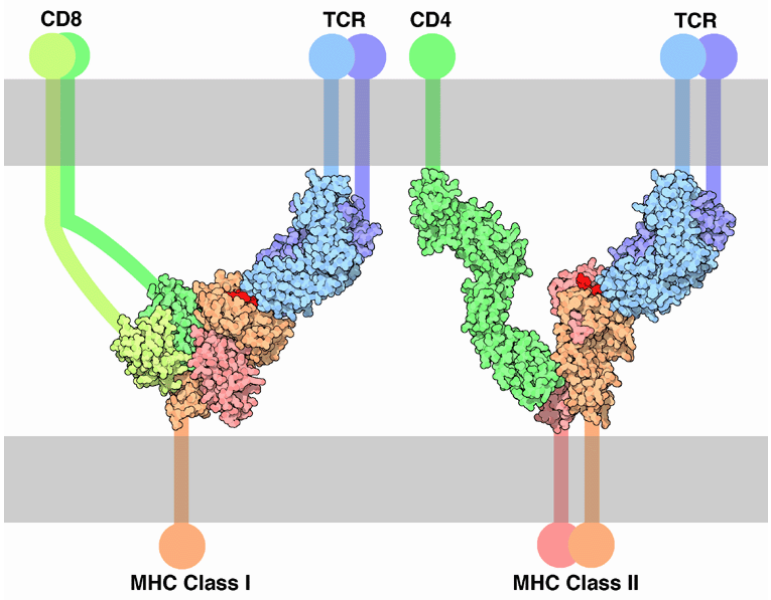

Fig 1: MHC-TCR interaction

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 MHC class II in the method and outcome of peptide presentation.

MHC in Humans: The Human Leukocyte Antigen (HLA)

About HLA

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 abbreviation MHC to refer to HLA gene products. To clarify the usage, some of the biomedical literature uses HLA to refer specifically to the HLA protein molecules and reserves MHC for the region of the genome that encodes for this molecule; however this convention is not consistently adhered to. These genes are present on chromosome 6 (Fig. 2).

They also have a role in:

- Disease defense
- Reproduction - may be involved in mate selection.
- Cancer - May be protective or fail to protect.
- Human disease:
  - In autoimmunity - known to mediate many autoimmune diseases.
  - As antigens - responsible for organ transplant rejection.

Classification of HLA

Aside from the genes encoding the 6 major antigens, there are a large number of other genes, many involved in immune function located on the HLA complex. Diversity of HLA in human population is one aspect of disease defense, and, as a result, the chance of two unrelated individuals having identical HLA molecules on all loci is very low.

The major HLA antigens are essential elements in immune function and different classes have different functions

Class I antigens (A, B & C) - Present peptides from inside the cell (including viral peptides if present)

Class II antigens (DR, DP, & DQ) - Present phagocytosed antigens from outside of the cell to T-lymphocytes

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.

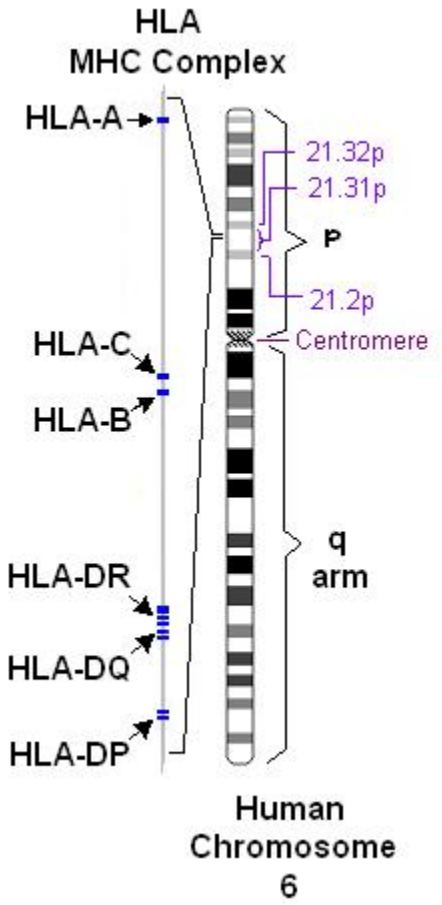

Fig 2: Localization of MHC gene

Functions of HLA

The proteins encoded by HLAs are the proteins on the outer part of body cells that are (effectively) unique to that person. The immune system uses the HLAs to differentiate self cells and non-self cells. Any cell displaying that person's HLA type belongs to that person (and therefore is not an invader).

In infectious disease

When a foreign pathogen enters the body, specific cells called antigen-presenting cells (APCs) engulf the pathogen through a process called phagocytosis. Proteins from the pathogen are digested into small pieces (peptides) and loaded onto HLA antigens (specifically MHC class II). They are then displayed by the antigen presenting cells for certain cells of the immune system called T cells, which then produce a variety of effects to eliminate the pathogen.

Through a similar process, proteins (both native and foreign, such as the proteins of viruses) produced inside most cells are displayed on HLA antigens (specifically MHC class I) on the cell surface. Infected cells can be recognized and destroyed by components of the immune system (specifically CD8+ T cells).

Once a T-cell recognizes a peptide within a MHC class II molecule it can stimulate B-cells that also recognize the same molecule in their sIgM antibodies. Therefore these T-cells help B-cells make antibodies to proteins they both recognize. There are billions of different T-cells in each person that can be made to recognize antigens, many are removed because they recognize self antigens. Each HLA can bind many peptides, and each person has 3 HLA types and can have 4 isoforms of DP, 4 isoforms of DQ and 4 Isoforms of DR (2 of DRB1, and 2 of DRB3,DRB4, or DRB5) for a total of 12 isoforms. In such heterozygotes it is difficult for disease related proteins to escape detection.

In graft rejection

Any cell displaying some other HLA type is "non-self" and is an invader, resulting in the rejection of the tissue bearing those cells. Because of the importance of HLA in transplantation, the HLA loci are among of the most frequently typed by serology or PCR relative to any other autosomal alleles.

In autoimmunity

HLA types are inherited, and some of them are connected with autoimmune disorders and other diseases. People with certain HLA antigens are more likely to develop certain autoimmune diseases, such as Type I Diabetes, Ankylosing spondylitis, Celiac Disease, SLE (Lupus erythematosus), Myasthenia Gravis, inclusion body myositis and Sjögren's syndrome. HLA typing has led to some improvement and acceleration in the diagnosis of Celiac Disease and Type 1 diabetes; however for DQ2 typing to be useful it requires either high resolution B1*typing (resolving *0201 from *0202), DQA1*typing, or DR serotyping. Current serotyping can resolve, in one step, DQ8. HLA typing in autoimmunity is being increasingly used as a tool in diagnosis. In GSE it is the only effective means of discriminating between 1st degree relatives who are at risk from those who are not at risk, prior to the appearance of sometimes irreversible symptoms such an allergies and secondary autoimmune disease.

In cancer

Some HLA mediated diseases are directly involved in the promotion of cancer. Gluten sensitive enteropathy is associated with increased prevalence of enteritus associated T-cell Lymphoma, and DR3-DQ2 homozygotes are within the highest risk group with close to 80% of gluten sensitive EATL cases. More often; however, HLA molecules play a protective role, recognizing the increase in antigens that were not tolerated because of low levels in the normal state. Abnormal cells may be targeted for aptosis mediating many cancers before clinical diagnosis. Prevention of cancer may be a portion of heterozygous selection acting on HLA.

Genetics of the HLA molecule

MHC class I

MHC class I form a functional receptor on most nucleated cells of the body. There are 3 major and 3 minor MHC class I genes in HLA:

HLA-A, HLA-B, HLA-C are the major genes
HLA-E, HLA-F and HLA-G are the minor genes

β2-microglobulin binds with major and minor gene subunits to produce a heterodimer.

MHC class II

There are 3 major and 2 minor MHC class II proteins encoded by the HLA. The genes of the class II combine to form heterodimeric (αβ) protein receptors that are typically expressed on the surface of antigen presenting cells.

Major MHC class II

HLA-DP
α-chain encoded by HLA-DPA1 locus
β-chain encoded by HLA-DPB1 locus

HLA-DQ
α-chain encoded by HLA-DQA1 locus
β-chain encoded by HLA-DQB1 locus

HLA-DR
α-chain encoded by HLA-DRA locus
4 β-chains (only 3 possible per person), encoded by HLA-DRB1, DRB3, DRB4, DRB5 loci

Other MHC class II

The Other MHC class II proteins, DM and DO are used in the internal processing of antigens, loading the antigenic peptides generated from pathogens onto the HLA molecules of antigen-presenting cell.

HLA serotype and allele names

There are two parallel systems of nomenclature that are applied to HLA. The, first, and oldest system is based on serological (antibody based) recognition. In this system antigens were eventually assigned letters and numbers (e.g. HLA-B27 or, shortened, B27). A parallel system was developed that allowed more refined definition of alleles, in this system a "HLA" is used in conjunction with a letter * and four or more digit number (e.g. HLA-B*0801, A*68011, A*240201N N=Null) to designate a specific allele at a given HLA locus. HLA loci can be further classified into MHC class I and MHC class II (or rarely, D locus). Every two years a nomenclature is put forth to aid researchers in interpreting serotypes to alleles.

HLA serotyping

In order to create a typing reagent, blood from animals or humans would be taken, the blood cells allowed to separate from the sera, and the sera diluted to its optimal sensitivity and used to type cells from other individuals or animals. Thus serotyping became a way of crudely identifying HLA receptors and receptor isoforms. Over the years serotyping antibodies became more refined as techniques for increasing sensitivity improved and new serotyping antibodies continue to appear. One of the goals of serotype analysis is to fill gaps in the analysis. It is possible to predict based on 'square root','maximum-likelihood' method, or analysis of familial haplotypes to account for adequately typed alleles. These studies using serotyping techniques frequently revealed, particularly for non-European or north East Asian populations a large number of null or blank serotypes. This was particularly problematic for the Cw locus until recently, and almost half of the Cw serotypes went untyped in the 1991 survey of the human population.

There are several types of serotypes. A broad antigen serotype is a crude measure of identity of cells. For example HLA A9 serotype recognizes cells of A23 and A24 bearing individuals, it may also recognize cells that A23 and A24 miss because of small variations. A23 and A24 are split antigens, but antibodies specific to either are typically used more often than antibodies to broad antigens.

HLA Phenotyping

Gene typing is different from gene sequencing and serotyping. With this strategy PCR primers specific to a variant region of DNA are used (called SSP-PCR), if a product of the right size is found, the assumption is that the HLA allele has been identified. New gene sequences often result in an increasing appearance of ambiguity. Because gene typing is based on SSP-PCR it is possible that new variants, particularly in the class I and DRB1 loci may be missed.

For SSP-PCR within the clinical situation is often used for identifying HLA phenotypes. An example of an extended phenotype for a person might be:
A*0101/*0301, Cw*0701/*0702, B*0702/*0801, DRB1*0301/*1501, DQA1*0501/*0102, DQB1*0201/*0602

This is generally identical to the extended serotype: A1,A3,B7,B8,DR3,DR15(2), DQ2,DQ6(1)

For many populations such as the Japanese or European populations so many patients have been typed that new alleles are relatively rare, and thus SSP-PCR is more than adequate for allele resolution. Haplotypes can be obtained by typing family members. In areas of world where SSP-PCR is unable to recognize alleles and typing requires the sequencing of new alleles. Areas of the world were SSP-PCR or serotyping may be inadequate include Central Africa, Eastern Africa, parts of southern Africa, Arabia and S. Iran, Pakistan and India.

Allelic diversity of MHC molecules

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. Indeed, much theory has been devoted to explaining why this particular region of the genome harbors so much diversity, especially in light of its immunological importance.

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 roughly 250, 500, and 300 known alleles respectively -- diversity truly exceptional in the human genome. The MHC gene is the most polymorphic in the genome. Population surveys of the other classical loci routinely find tens to a hundred alleles -- still highly diverse. 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.

Suggested Reading

1. P. Parham and T. Ohta (1996). "Population Biology of Antigen Presentation by MHC class I Molecules.". Science 272. .
2. Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Geraghty DE, Hansen JA, Hurley CK, Mach B, Mayr WR, Parham P, Petersdorf EW, Sasazuki T, Schreuder GM, Strominger JL, Svejgaard A, Terasaki PI, and Trowsdale J. (2005). "Nomenclature for factors of the HLA System, 2004.". Tissue antigens 65: 301-369.
3. MHC Sequencing Consortium (1999). "Complete sequence and gene map of a human major histocompatibility complex". Nature 401: 921–923.