In: Advances in Medicine and Biology. Volume 53
Editor: Leon V. Berhardt

ISBN: 978-1-62081-565-6

Chapter IV

Major Histocompatibility Complex: Biology, Functions and Roles in Disease

Kwame A. Darfour-Oduro and Lawrence B. Schook[1]
Department of Animal Sciences, University of Illinois, US

Abstract

The Major Histocompatibility Complex (MHC) found only in the gnathostome (jawed vertebrate) encode cell-surface glycoproteins that bind and present short peptides from degraded self and non-self (bacteria, fungal and viral-derived) proteins onto the surface of immune system cells to T cells which in turn initiate immune responses on recognition of the non-self proteins. The organization of the MHC within the jawed vertebrates (with the shark and human occupying the opposite ends of the evolutionary spectrum) appears to be well conserved. MHC occupies a segment of approximately 4 centimorgans on the short arm of chromosome 6 in humans and is divided into three major subfamilies: class I, class II and class III. X-ray crystallographic analysis revealed the structural organization of molecules encoded by these gene subfamilies and showed that class I and class II molecules have peptide binding grooves for presentation of self and foreign antigens to T lymphocytes. Other functions of MHC class I molecules include but not limited to interactions with Natural Killer (NK) cell receptors in protection against cell-mediated toxicity. Another function of MHC class II molecules includes interactions with T cell receptors resulting in cell death. A recently discovered role of intracellular MHC class II molecules is their involvement in signaling of Toll-like receptors (TLRs) by interacting with tyrosine kinase Btk, a protein essential for normal B-cell receptor signaling. The characteristics of the MHC region including high polymorphism, linkage disequilibrium (LD) between haplotype blocks and the presence of copy number variants (CNV) gives the region a complex genetic architecture and implicate alleles and haplotypes in autoimmune, inflammatory and infectious diseases. As the antigen presentation theory has proven to be inconsistent with some MHC haplotype-disease associations, a novel theory known as the cusp theory have been

[1] Department of Animal Sciences, University of Illinois, Urbana, IL 61801, US. Email: schook@illinois.edu.

proposed; A cusp-like feature in one of the two α helices that form the perimeter of the groove of all MHC molecules ligands interact with other cell surface receptors apart from those of the MHC and thus may result in disease in the event of abnormalities in activated pathways. Recent technological advances in genome sequencing and genotyping have resulted in extensive sequence characterization of MHC genes and have provided the tools for Genome wide association studies (GWAS) to investigate genetic variants and their associations with diseases such that novel haplotypes –disease associations have been identified.

Introduction

The MHC is a family of gene clusters that play a pivotal role in how vertebrate distinguish between self and non-self-antigens and elicit immune responses. The MHC and other components (TCR, Ig genes and RAG) of the adaptive immune system are thought to have co-evolved in the gastrointestinal region of primitive jawed vertebrates in order to protect vertebrates against pathogens introduced by the predatory lifestyle (Matsunaga and Rahman, 1998).

MHC genes encode a family of molecules and occupy a segment of approximately 4 centimorgans on the short arm of chromosome 6 in humans and the equivalent region in mice is located on chromosome 17. These genes constitute about 0.1% of the human genome (Yu et al. 2000). MHC genes are divided into three major subfamilies: class I, class II and class III with differences observed in organization and structure of these classes across vertebrates. For example, the genomic organization of the swine MHC is such that the centromere separates the class II region from the class III and class I regions and therefore the class II region is located on the q-arm of chromosome 7 whereas class I and class II are located on the p-arm of chromosome 7 (Smith, 1995). This genomic organization is inconsistent with that of other mammals whose class I, II and III loci exhibit a tight linkage (Smith, 1995). Non mammalian vertebrates such as the chicken, shark and quail do not have class I duplication blocks and conserved flanking framework genes seen in mammals (Kulski et al. 2002). Several large scale studies have revealed an extended human MHC region (Horton et al. 2004) spanning about 7.6 Mb of the genome. The MHC region is characterized by extensive polymorphism (more alleles per locus) and alternative splicing of MHC transcript is more frequent than genome-wide splicing (72.5% vs. 62.1% of genes, $P \leq 1 \times 10^{-4}$) (Vandiedonck et al. 2011). The region is further characterized by strong linkage disequilibrium across loci.

The MHC was first observed as a genetic locus involved in immunological functions following the rapid rejection of grafts of tissues when donor and recipients were mismatched at the MHC in both humans and mice (Simpson, 1988). Later (1975), it was established that T cells required the recognition of both viral antigens and MHC together on the cell surface in order to recognize and lyse an infected cell (Doherty and Zinkernagel, 1975). This was followed by a description of how T cell immune responses co- evolved with MHC genes (Benacerraf, 1981). The mid 1980s and 1990s witnessed the elucidation of the structure of the MHC by X-ray crystallographic analysis. Such studies have shown that significant similarities and differences exist between the various classes of the MHC.

Structural analysis of MHC molecules has given insight into their functions such that alternative theories regarding the classical MHC functions of the presentation of antigenic

peptides to T lymphocytes by MHC molecules, a process that's success depends upon highly coordinated intracellular events has been proposed. The unique MHC genetic architecture implicates the region as a target for association studies with numerous forms of human disease. Indeed, the MHC region have been implicated in numerous autoimmune diseases including insulin dependent diabetes mellitus, multiple sclerosis, systemic lupus erythematosus, myasthenia gravis and rheumatoid arthritis (reviewed by Thomson, 1995), infectious diseases such as ankylosis spondylitis, Reiter's disease, acute anterior uveitis of viral origin, Human immunodeficiency virus 1 (HIV-1), hepatitis C and parasitic infections as malaria, scabies and diffuse cutaneous leishmaniasis (Singh et al. 1997). Ulcerative colitis and Crohn's disease are some inflammatory disease known to be influenced by alleles at the MHC loci (reviewed by Fernando et al. 2008).

Recent advances in molecular genetic technology have unraveled novel single nucleotide polymorphism (SNP) markers within the MHC to more explicitly determine their association with diseases. In particular next generation resequencing projects have expanded the knowledge on the MHC (Traherne, 2008). Genome wide association studies, case-control analysis in which SNP allele and genotype frequencies are contrasted for different phenotypes of a trait with the objective of investigating genetic association between markers and traits have been employed in identifying novel MHC alleles and refining known alleles associated with various diseases.

Below, MHC biology, functions and its roles in disease in the context of both past, current findings and new technologies of MHC research are reviewed.

Evolution of MHC Molecules

Comparisons of MHC genomic sequence across species and mammals give an insight into the molecular evolutionary history of the MHC gene family in which loci are duplicated, lost or made nonfunctional (Kelley et al. 2005). The MHC is thought to have emerged in the shark some 528 million years ago. Table 1 shows the divergence time of vertebrate species and some parameters concerning their MHC region.

Comparative Genomic Analysis of MHC Region across Primates

Across primates, there is similarity in how the MHC region is organized and most MHC loci are common for humans, chimpanzees and rhesus macaques (Table 2). The human MHC class I genes (*HLA-A, B, C, E, F, G*) have functional orthologs in the MHC of the common chimpanzee (Pan troglodytes) (Adams et al. 2001). The *Patr-AL* gene, a non-classical MHC class I gene discovered in chimpanzees is not found in humans and or any other African ape species but is similar to the expressed A locus in the orangutan (Adams et al. 2001). Nucleotide sequence similarity between human and chimpanzees approximates 99% across the genome but is 86% in the MHC Class I region (Anzai et al. 2003).

Table 1. MHC characteristics across mammalian and non-mammalian vertebrates

Species	Estimated divergence time (Mya)*	Chromosomal location of MHC	Length (bp)	Number of genes	Reference
Human	0.2	6	3.6M	224	MHC Sequencing Consortium, 1999
Chicken	310	B locus	92K	19	Kaufmann et al. 1999
Mouse	100	17			Walter et al. 2002
Rat	100	20	3.8M	-	Kelley et al. 2005
Swine	60	7	2.0M	152	Morris, 2009
Cat	55	B2	3.3M	-	Yuhki et al. 2007

*Apart from the estimated divergence time of the cat and human, all other estimated divergence time were obtained from the review by Kulski et al., 2002. Mya is million years ago.

Table 2. Orthologous MHC loci in humans (*HLA*), chimpanzees (*Patr*) and rhesus macaques (*Mamu*)

	Species		
	HLA	*Patr*	*Mamu*
Loci			
Class I	A	A	A[a]
	B	B	B[a]
	C	C	absent
Non-classical Class I	E	E	E
	F	F	F
	G	G	G[b]
absent	absent	AG[a]	
	absent	AL	absent
	absent	absent	I
Class II	DR	DR	DR
	DQ	DQ	DQ
	DP	DP	DP

Source: Bontrop and Watkins, 2005.

[a] Mamu-A and B genes are polygenic and the number of genes might vary according to haplotype or region configuration. Mamu-AG is in a polygenic family.

[b] T he Mamu-G gene is a pseudogene and its function is performed by Mamu-AG.

Rhesus macaques diverged from the common ancestor 30 Mya while the divergence time of chimpanzee is 5 Mya.

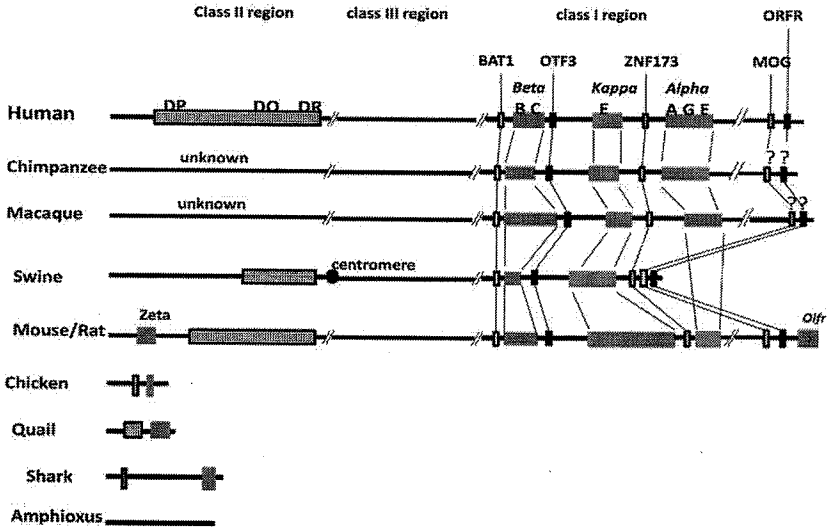
Comparative Genomic Analysis of Mammalian and Non-Mammalian Vertebrate MHC Region

Placental mammals have at least three distinct class I duplication blocks known as α, β and κ blocks (Kulski et al. 2000). The swine has the β and κ blocks but lacks the α block while the mouse and the rat have additional class I duplication blocks referred to as olfr and ζ blocks (Figure 1) (Kulski et al. 2002). Non mammalian vertebrates including chicken, quail and shark lack the mammalian duplication blocks and conserved flanking framework genes (evolutionary conserved non-class I or non-class II genes within the MHC framework) (Figure 1). A striking similarity in gene contents and order of framework genes was observed by comparing MHC sequences of human, mouse, opossum, dog and cat spanning from *KIFC1* gene through *UBD* gene (Yuhki et al. 2007); opossum, human and mouse have one contiguous gene content implying an ancestral form of MHC while the MHC of cat and dog have the same split form at *TRIM39* and *TRIM26* in the class I region as compared with human HLA. Across animal species comparison of the MHC reveal that the class I region is more divergent than the class II region and the class III region is conserved in terms of gene diversity and organization (Kulski et al. 2002). In humans, mice and other mammals, the MHC class I and II regions are characterized by higher incidences of pseudogenes that are intermingled with functional genes than is typical for the human genome (Beck et al. 1999).

In a study comparing the genomic structure of human, dog and cat MHC region, (Yuhki et al. 2007) observed that the human HLA class I region which spans 550 kbp from *HTEX4* to *MOG* has been reduced to 55 kbp in the cat FLA and the *HLA-E* gene in human is also missing in the cat FLA. The MHC class II region of mammals is characterized by duplicated loci and deletions and this has resulted in a different class II region configuration across mammals. For example, the human class II region has duplicated DP, DQ and DR genes while there is a deletion of the entire DQ region in cats (Figure 2). In the cat, the DR region is duplicated apparently to compensate for the loss of the DQ region. Across mammals, the DP region appears to become non-functional and this is observed for dog DLA, cat FLA and equine MHC (Yuhki et al. 2007).

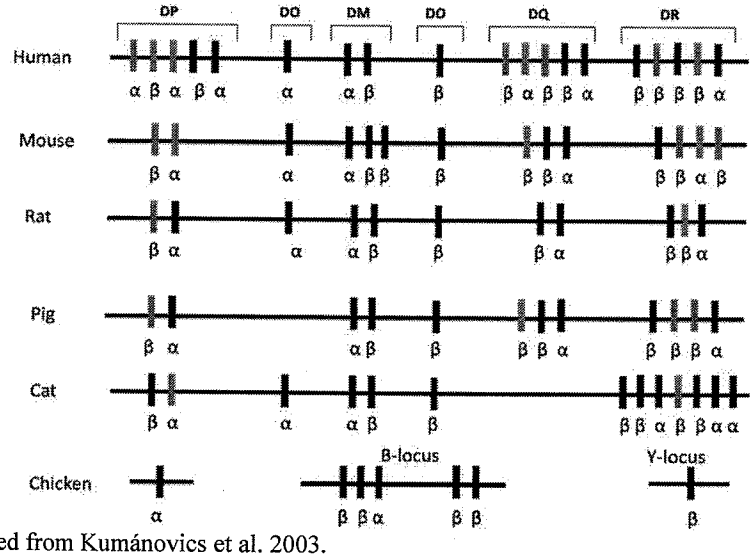
In comparison with mammals, the chicken MHC is smaller in size (9.2 kb DNA sequence containing 19 genes) (Kaufman et al. 1995) and is approximately 20-fold smaller than the human MHC (MHC sequencing consortium, 1999). However, almost all the genes within the chicken MHC have counterparts in the human MHC indicating that these genes have been conserved over 200 million years of divergence between birds and mammals (Kaufman et al. 1995).

For the MHC class III region, an amphibian (Xenopus tropicalis) and a marsupial (Monodelphis domestica) share most of the genes in the human class III region and non-mammalian vertebrates such as the chicken and quail have a class III region represented by only a single gene that codes for complement component gene *C4* (reviewed by Deakin et al. 2006). The organization of the Teleost MHC class III is also different and it is characterized by MHC split over several different chromosomes (Deakin et al. 2006). It is therefore obvious that evolution of the MHC region across species has been characterized by gene duplication, losses and in the case of the swine, centromere invasion.



Source: Adapted from Kulski et al. 2002.

Figure 1. Duplication blocks within the MHC class I region: The deep grey rectangles represent the different class I duplication blocks. The swine is the only mammal that lacks the α block and its' centromere separates the class II region from the class III and class I regions. The olfr and ζ blocks are only found in the MHC region of the mouse and the rat. The duplication blocks are flanked by framework genes (*BAT1, OTF3, ZNF173, MOG, ORFR*). The region of the class II genes are shown as rectangles with diamonds. The non-mammalian vertebrates chicken, quail and shark have different class I and class II genomic structure as compared with mammals while Amphioxus appears to have no class I and class II genes (Kulski et al. 2002).



Source: Adapted from Kumánovics et al. 2003.

Figure 2. Orthologous class II genes in mammals: Black boxes represent genes and grey boxes represent pseudogenes. The class II MHC region is arranged in separate regions of the chromosome and contain α and β chain genes. The cat lacks the entire DQ region and has a duplicated DR region that appears to compensate for the lack of a DQ region. In chicken, the genes are on the same chromosome but not linked genetically (Kumánovics et al. 2003).

How the MHC Genes Emerged in Vertebrates

MHC genes cannot be traced through a gradual evolutionary process but are found only in jawed vertebrates (Bartl et al. 2003) and absent in lower chordates. The complement component system within the class III region however contains molecules such as alpha-2-macroglobulins that have evolved from other functions in invertebrates to become part of both the innate and the adaptive immune systems (Bartl et al. 2003). Two whole-genome wide duplications and the emergence of recombination-activation gene (*RAG*) transposons are two evolutionary events believed to have contributed to the emergence of the adaptive immune system (Flanjnik and Kasahara, 2009). Support for the whole genome duplication event contribution to the emergence of MHC stems from the fact that paralogs thought to have emerged close to the origin of vertebrates normally occur in clusters on multiple chromosomes (Flanjnik and Kasahara, 2009). RAG1 and RAG2 are lymphocyte-specific proteins that create nicks between germline chromosomal variable, diversity and joining components of immunoglobulins and T cell receptors (Laird et al. 2000). RAG1 and RAG2 are believed to have been contained in a retrotransposon which integrated into a large cis promoter element of a vertebrate. Consistent with this belief, MHC receptors have not being isolated in species that lack RAG (Laird et al. 2000) thus implicating RAG in the emergence of MHC. MHC class I and class II include genes that have evolved through the birth and death process (Nei and Hughes, 1992), a process where new genes are created by repeated gene duplication with some genes later becoming pseudogenes or being deleted (Piontkivska and Nei, 2003). Class I molecules appear to have a relatively faster rate of birth and death evolution (Nei and Hughes, 1992). Thus, over evolutionary time, the turnover rate of class I molecules has been rapid (Hughes, 2008). As a consequence of this rapid turnover rate, no orthologous relationships exist between class I genes among orders of different mammals (Hughes and Nei, 1989). In contrast, for class II genes, many orthologous loci are shared between different mammalian orders (Hughes and Nei, 1990).

Evolutionary Forces Maintaining Polymorphism at the MHC Loci

Selective forces partly drive the high polymorphism that characterizes the MHC loci. A long held view suggests that two types of balancing selection are responsible for maintaining this polymorphism. According to the overdominance hypothesis, MHC polymorphisms are maintained by heterozygous individuals recognizing a broad spectrum of pathogens (Doherty and Zinkernagel, 1975). The negative frequency dependent selection hypothesis is based on the assumption that rare MHC alleles offer better parasite resistance because selection favors parasites that are capable of avoiding recognition by the most common MHC variants (Clarke and Kirby, 1966).

A new model that accounts for the effects of peculiar characteristics of the MHC region such as linkage disequilibrium (LD) and the high gene diversity have also been proposed (van Oosterhout, 2009). Termed as 'associative balancing complex evolution', this model suggests that SNPs linked to the MHC region can accumulate as recessive deleterious mutations as a result of high MHC heterozygosity and strong LD that renders purifying selection ineffective.

These mutations could contribute to balancing selection in the MHC when they become fixed (van Oosterhout, 2009) through hitchhiking.

Biology of MHC Class I Molecules

MHC class I molecules are expressed by nearly all nucleated cells. MHC class I molecules differ from MHC class II and III in level of expression, structure of the encoded protein and function (Hughes, 2008). The class I MHC molecules consist of two separate polypeptide chains, a heavier α chain of molecular weight of 44-47, 000 Da made up of three extracellular domains α1, α2 and α3, and the lighter β chain of molecular weight of 12,000 Da. The extracellular domains are noncovalently associated with β2–macroglobulin which is encoded on chromosome 15 in humans. β2–microglobulin is involved in the proper folding and cell surface display of MHC class I molecule. Individuals with a nonfunctional β2–microglobulin gene fail to express any class I antigen and thus have a deficiency of cytotoxic T cells (Sridhar, 2011). The α chain has a cytoplasmic region which contain sites for phosphorylation and binding to skeletal elements, and a transmembrane region containing hydrophobic amino acids by which the molecule is anchored in the cell membrane. The first two α domains create a platform composed of a single β-pleated sheet topped by α-helices (Bjorkman et al. 1987). A long groove between the helices forms the peptide binding site (Bjorkman et al. 1987). The α3 domain associates with T cell co-receptor CD8 during T cell recognition (Collins et al. 1995). Each of the extracellular domains (α1, α2 and α3) is approximately 90 amino acids long and is encoded by a separate exon. X-ray crystallographic studies have shown that the α3 domain is not required for the structural integrity of MHC class I or for the binding of peptides (Collins et al. 1995)., MHC class I molecules are divided into two groups, classical and non-classical groups. In the case of humans, classical MHC class I molecules include HLA-A, HLA-B and HLA-C. The non-classical class I molecules of all mammals studied lack the universal tissue expression and high level of polymorphism associated with the classical molecules and in humans consist of HLA-E, HLA-F, HLA-G, MICA, MICB, HFE, CD1 (located on chromosome 1), fcrN (located on chromosome 19), MRI and zinc α2-glycoprotein (Hughes, 2008). A schematic map depicting the basic organization of the human MHC loci is presented as Figure 3.

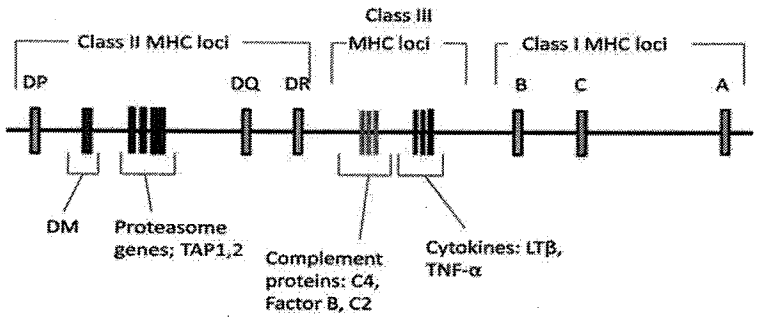


Figure 3. Schematic map of the human MHC loci: The sizes of genes are not drawn to scale. The class III region is between the class II and class I regions. Class II loci consists of several genes.

Biology of MHC Class II Molecules

MHC class II molecules are expressed by antigen-presenting cells such as macrophages, dendritic cells and B lymphocytes. Class II MHC molecules consist of a heavy chain of two extracellular domains α1 and α2 and a light chain of extracellular domains β1 and β2. The α and β chains are of approximately equal length and have molecular weights of 34,000 Da and 28,000 Da respectively (Sridhar, 2011). All domains of MHC class II molecules are stabilized by disulphide bridges except the α1 domain (Sridhar, 2011). The α1 domain is covalently associated with the β1 while α2 domain is covalently associated with β2 domain. The cytoplasmic regions of the α and β chains contain sites for phosphorylation and binding to cytoskeletal elements and a transmembrane region containing hydrophobic amino acids through which the molecule is anchored in the cell membrane (Mayer and Nyland, 2011). CD4 molecule of Helper T lymphocytes binds to β2 domain during antigen presentation (Sridhar, 2011). In humans, the MHC class II region is divided into three regions each harboring distinct molecules. The DR region has HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DRA molecules; the DP region has HLA-DPB1 and HLA-DPA1 while the DQ region has HLA-DQB1 and HLA-DQA1. Two other regions are DM and DO and are less polymorphic as compared with the other regions (Table 3). Comparison of number of alleles at MHC loci across humans, chimpanzees, swine and dogs reveal that more alleles have been reported for humans than the other species (Table 3). Other MHC class II related genes form heterodimers (DMA/DMB and DOA/DOB) whose products regulate the loading of peptides by class II molecules (Hughes, 2008). Studies on the three dimensional structure of HLA-DR1 revealed a peptide binding groove similar to that of MHC class I molecules. However, the peptide binding site of class II MHC are opened at both ends enabling the accommodation of longer peptides (15 vs. 9 residues) as compared with MHC class I binding site (Brown et al. 1993).

Biology of MHC Class III Molecules

The human MHC class III region, encoded between the MHC class I and II regions (Figure 3), contains 57-60 structural genes containing more than 500 exons spanning 654-759 kb of genomic DNA (Yu et al. 2000) with average gene size of approximately 8.5 kb and an average intergenic distance of just 3 kb (Xie et al. 2003). Structural genes present in the human MHC class III region have orthologs in the mouse (Yu et al. 2000). The region contains among other important genes, the complement component genes *C4*, *C2,* and *factor B*, the major heat shock protein HSP70 and the tumor necrosis factors TNFα and TNFβ. *C4 gene* codes a serum glycoprotein composed of three disulphide linked polypeptides (Schreiber and Muller-Eberhard, 1974). Native C4 is a 200,000 Da protein consisting of three polypeptides α, β and γ of approximate weights of Mr 95,000, 78,000 and 31,000 respectively (Schreiber and Muller-Eberhard, 1974). The glycoprotein has an 185,000 Da single chain precursor, pro – C4 whose synthesis is directed by a polyadenylated mRNA of approximately 5 kb (Whitehead et al. 1983). C2 and factor B are single chain glycoproteins of molecular weights of Mr 102,000 and 95,000 respectively (Kerr and Porter, 1978).

Table 3. Number of alleles at MHC class I and II loci of humans (*HLA*), chimpanzees (*Patr*), swine (*SLA*) and dogs (*DLA*)

	MHC Class I alleles									
	Classical genes			Non-classical genes						
Gene	A		B	C		E		F	G	H
Number of alleles (HLA)a	1,729		2,329	1291		10		22	47	-
Number of alleles (Patr)b	38		45	31		1		1	1	1
Gene	SLA-1		SLA-2	SLA-3		SLA-6		SLA-7	SLA-8	
Number of alleles (SLA)b	44		46	26		9		2	2	
Gene	DLA-88					DLA-12		DLA-64	DLA-79	
Number of alleles (DLA)c	51					-		-	-	
	MHC Class II alleles									
Gene	DRA	DRB	DQA1	DQB1	DPA1	DPB1	DMA	DMB	DOA	DOB
Number of alleles (HLA)a	7	1,150	46	160	33	150	7	13	12	13
Number of alleles (Patr)b	2	72	7	10	3	29	-	2	-	1
Number of alleles (SLA)b	13	82	20	44	-	-	5	1	1	1
Number of alleles (DLA)b	-	39	14	27	-	-	-	-	-	-

References for figures in Table 3: [a] Robinson et al., 2011.
[b] IPD-MHC website (http://www.ebi.ac.uk/ipd/mhc).
[c] http://sharedresources.fhcrc.org/content/dla-typing-protocols-and-primers.

Following the activation by a single cleavage event, the catalytic subunit of C2 of Mr 60,000 in association with activated C4 forms the C3 convertase of the classical pathway of the complement system (Reid and Porter, 1981). The *HSP70* genes encode cytosolic molecular chaperons which are involved in the biogenesis of protein including synthesis, folding and translocation in organelles (Hendrick and Hartl, 1993). *TNFα* and *TNFβ* genes encode cachectin and lymphotoxin-molecules, respectively. TNF is secreted as an unglycosylated 157-amino acid polypeptide of relative molecular mass Mr of 17,350 (Jones et al. 1989). The primary structure of TNF is evolutionary conserved and the protein is thought to be trimeric in its active form (Smith and Baglioni, 1987). Genes in the MHC class III region are clustered but represent groups that are very divergent in sequence and structure (Gruen and Weissman, 2001).

The MHC class III region is the most gene dense region in the human genome. A few of the genes overlap at their 3' ends (Yu et al. 2000).

Immunological Functions of MHC Class I Molecules

Classical MHC class I molecules present endogenous peptide antigens to CD8+ T cells. The complex formed by the MHC class I molecules and the peptide interacts with CD8 and the T cell receptor, signaling the T cell to kill a MHC class I-bearing cell (Abele and Tampé, 2004). This classical antigen presenting function of the MHC class I molecules is a highly regulated process involving a plethora of proteins.

Antigenic peptides are generated through the degradation of endogenous proteins in the cytoplasm by multisubunit proteases known as proteosomes. Antigenic peptides are bond by the MHC encoded transporter associated with antigen processing (TAP) and transferred to the lumen of the endoplasmic reticulum (ER). Many of the peptides transported by TAP are longer than the 8-10 residues suitable for binding to MHC class I molecules and therefore ER aminopeptidases trims the peptides to achieve optimal binding to MHC class I molecules (Wearsch and Cresswell, 2008). Newly synthesized heavy chain class I/β2m heterodimers within the ER binds to a peptide loading complex made up of tapasin, ERp57, TAP and calreticulin. Following peptide binding, the MHC class I molecule is released from the complex through the golgi apparatus to the cell surface where cells with exposed peptides derived from non self proteins or mutated self proteins are eliminated by cytotoxic T cells.

The Role of the MHC in Self and Non-Self-Discrimination

The main task of the immune system is distinguishing self-antigens from non-self (foreign) antigens. In order to prevent autoreactivity, a situation where T cells will attack self-peptide-MHC complex, immature thymic T cells expressing T cell receptors during differentiation in the thymus are deleted by negative selection (Lanier and Philips, 1996). Maturation of useful T cells with receptors is a requirement for binding foreign antigens plus self MHC molecules to elicit immune response. Some potentially autoreactive T cells do escape negative selection in the thymus and are controlled by peripheral mechanisms

including active control by a subset of CD4+ T cells, also known as regulatory T (Treg) cells, that dampen immune and inflammatory responses (Sakaguchi et al. 2006).

Other Functions of MHC Class I Molecules

Aside the immunological functions of antigen presentation by MHC class I molecules, additional roles performed by these molecules have been investigated. One such role is discussed below.

Interaction between MHC Class I Molecules and Natural Killer (NK) Cells

NK cells are innate immune lymphocytes with potent effector functions against infected and tumor cells (Elliot et al. 2010). They represent 5-10% of peripheral blood lymphocytes and the proportion varies according to the individual's age (Shearer et al. 2003). NK cells responses are mediated through inhibitory and activating cell surface receptors, some of which monitor MHC class I molecules expression on surrounding cells (Yokohama and Plougastel, 2003). Negative signals from inhibitory receptors usually dominate over positive signals from activating receptors (Ewald and Livant, 2004). NK cell receptors that recognize MHC class I molecules include the LIR, Ly49 (in mice) and the KIRs in humans (Deng and Maruzza, 2006). Structural homologs of the MHC class I molecules, such as MICA and MICB that are up-regulated in stressed cells (Groh et al. 1998) are recognized by NK cell NKG2D receptors (Vivier et al. 2002). NK cells attack target cells that express low or no MHC class I proteins on the cell surface (Kärre et al. 1986). MHC class I molecules are down-regulated in the event of viral infection or cellular transformation (Medzhitov and Janeway, 2002). Thus, the recognition of a missing self-ligand by NK cells allows NK cells to kill infected and transformed cells and to spare normal, healthy cells (Medzhitov and Janeway, 2002).

Immunological Functions of MHC Class II Molecules

MHC classic class II molecules (DR, DQ and DP), heterodimers of α and β chains are expressed on antigen presenting cells of the immune system and function to present peptides to helper T cells. In mice, classic class II molecules, H2-A and H2-E are constitutively expressed on the surface of antigen presenting cells. They present extracellular pathogens to helper CD4+ T cells.

The antigen processing step involves endocytosis of exogenous proteins by acidic proteases such as Legumain, Cathepsin L and Cathepsin S that are found in endosomes and lysosomes (Hsing and Rudensky, 2005). Unfolding of endocytosed antigens in MHC class II molecules containing compartments, a step necessary for subsequent proteolysis is achieved

by Gamma-Interferon-inducible Lysosomal Thiol Reductase (GILT) (Phan et al. 2000). GILT achieves folding by catalyzing disulfide bond reduction. CD74, an integral membrane protein (Starlets et al. 2006) assembles MHC class II in the endoplasmic reticulum and escorts it into the endosomal system. Following assembly of MHC class II, CD74 bound to MHC class II molecule is cleaved by proteases and digested into CLIP (Class II-associated invariant chain peptide) which becomes unbound with the aid of another MHC encoded molecule, DM (Denzin and Cresswell, 1995). DM also aids in the in the rapid loading of high affinity antigen-derived peptides into the MHC class II binding groove. The MHC class II molecule and the peptide complex leave the endocytic compartments to present peptide antigen to T helper cells at the cell surface (Weenik and Gautam, 1997). MHC class II molecules are expressed on thymic stromal cells where they direct the process of positive and negative T cell selection, shaping the repertoire during T cell maturation (McDevitt, 2000).

Other Functions of MHC Class II Molecules

Aside MHC classic class II molecules role in antigen presentation, cross-linking of MHC class II cell surface molecules by T cell receptors lead to cell death or apoptosis (Truman et al. 1996). MHC class II mediated death is induced via Fas/Fas ligand interaction in human splenic B lymphocytes (Truman et al. 1997). A study investigating the binding site and mechanism that mediates cell death showed that following the binding of HLA-DQA1 to MHC class II molecules outside the peptide binding groove, there is MHC class II signaling resulting in the initiation of protein kinase signaling which ultimately leads to cell death (Zang et al. 2006).

Another function of MHC class II molecules in TLR signaling has been proposed following observations by Piani et al. (2000) that macrophages lacking MHC class II molecules or low expression of it, on stimulation by lipopolysaccharide (LPS) elicit a defective inflammatory response. Indeed, recent studies by Lui et al. (2011) have demonstrated that intracellular MHC class II promotes TLR signaling in mice. The authors observed that this activated signaling was not as a result of a direct interaction between TLR and MHC class II, but rather MHC class II molecules, owing to their short cytoplasmic domains associates indirectly with Btk, a tyrosine kinase, through direct association with CD40 which mediates Btk activation. After stimulation with TLR ligands such as LPS, MHC class II molecules interact with Btk through CD40 and maintain Btk activation. Activated Btk induces the production of proinflammatory cytokines and type I interferons by interacting with MYD88 and TRIF which in turn promote the activation of MyD88-dependent and TRIF-dependent pathway.

The Roles of the MHC in Disease

Genetic Architecture of the MHC Region and Its Role in Disease

Given the myriad of pathogens that confront hosts, over evolutionary time, polymorphism has been generated to ensure enough MHC diversity to deal with diseases presented by these pathogens. Indeed, it has been suggested that species which show limited or no MHC diversity e.g. cheetahs (Aconyx jubatus) and Asiatic lions (Panthera leo persica) may be susceptible to epidemic diseases in the future (O'Brien and Evermann, 1988). Interaction between MHC and non MHC gene families has been observed to influence diseases. One such interaction is that between MHC and the KIR gene family which have been implicated in cancer, reproductive failure, autoimmune diseases and viral infection (Rajagopalan and Long, 2005). MHC class II region bounded by HLA-DRB1 and HLA-DRB9 and the RCCX region of class III region have confirmed copy number variants (CNVs) (Traherne, 2008). CNVs are structural variants that represents >1 kb of duplication or deletion of the genome and may influence gene expression, affect several metabolic traits and have been associated with mendelian and complex genetic disorders (Oroczo et al. 2009). CNVs lead to disease conditions through dosage sensitive genes and disruption of functional genes (Zhang et al. 2009). Another characteristic of the MHC region mentioned earlier is strong linkage disequilibrium (LD) that exists between alleles at loci close to each other. Though such strong LDs have aided in the identification of genetic variants associated with diseases, often it becomes difficult to determine whether these variants-disease associations are really due to the particular variant or the association is due to another variant that happens to be in LD with the variant in question.

The Cusp Theory of the MHC

MHC molecules are implicated in disease susceptibility and resistance based on the hypothesis that MHC molecules present antigens to T cells to elicit immune response. Alternative hypotheses have however emerged in view of the fact that the antigen presentation theory fails to explain certain allele-disease association, for example copy number variation and disease susceptibility and the fact that some MHC disease association as seen in narcolepsy do not involve antigen presentation (de Almeida and Holoshitz, 2011).

An alternative theory is the MHC Cusp theory (de Almeida and Holoshitz, 2011) which proposes a role in disease outcome for a tri-dimensional cusp-like structural motif common to all products of the MHC gene family located in the $\alpha 2$ domain of MHC class I molecules and a similar structure located in the $\beta 1$ domain class II molecules (de Almeida and Holoshitz, 2011). It is thought that ligands produced by these cusp-like structures encoded by their HLA alleles interact with non-adaptive immune system receptors and activate important pathways, aberrations or amplification of which could lead to MHC related diseases. Empirical support for this theory comes from studies on rheumatoid arthritis. It was observed that the cusp shared epitope ligand encoded by HLA-DRB1 alleles interacts at a low affinity with cell surface receptors, allowing the normal activation of desirable biological function (de Almeida

and Holoshitz, 2011). However, the authors propose that abnormal activation of the pathway could occur in the event of an environmental trigger and this could lead to disease.

Recent Technological Advances and Their Roles in MHC-Disease Association

Next generation sequencing (NGS) have led to extensive sequence characterization at the MHC. In NGS, whole genome or specific regions within the genome are randomly digested with restriction enzymes into small fragments (short reads) that are sequenced and are aligned to a reference genome in search of genetic variants. NGS approaches generate a number of such short reads of contiguous sequences in a short time and it takes few days to sequence hundreds of millions of bases from amplified single DNA molecules (Bentley et al, 2009). Recently, Bentley et al. (2009) have employed NGS in a very economic way in the complete typing of 7-locus HLA Class I and II in a single GS FLX sequencing run. Their method uses pooled DNA samples of individuals yet sequences of individuals are identifiable because the process makes use of multiplex identification tags. Among the achievements of the method is the detection of rare variants even in chimeric mixtures. New genotyping technologies such as dense genotyping chips that have hundreds of thousands of SNPs and large clinical samples now in place (The Welcome Trust Consortium, 2007) combined with NGS have provided the ingredients for Genome-wide association studies (GWAS) to test HLA variants for disease associations. GWAS tests for association between the frequency of thousands of common SNPs and a given disease. SNPs that exceed a conservative genome-wide threshold for association ($P < 5 \times 10^{-8}$) are tested for evidence of replication in independent cohorts (Gibson, 2010). In order to increase the power to identify genetic variants, GWAS data from single studies are combined (meta-analysis).

As particular alleles associated with diseases can be located in more than one haplotype (Vandiedonck et al. 2011), which may be in LD with another, identification of the specific haplotype on which a risk allele is located may be problematic. To deal with the issue, an approach which links sequence variation (genotypes) within the MHC to different heritable gene expression levels and accounts for differences in gene expression among different haplotypes have been proposed (Vandiedonck et al. 2011); The authors used MHC microarray with alternate allele probes capable of detecting different MHC haplotypes and observed that 96 genes were differentially expressed in the 3 haplotypes in their study. Thus, with this approach, candidate genes influencing a disease phenotype can be easily traced to the particular haplotype.

Roles of Specific MHC Genes in Disease

The Human Leukocyte Antigens (HLA) genes, part of the hundreds of genes at the MHC region, constitute the gene family within the MHC that has been extensively studied in terms of gene-disease associations and they are passed on to the next generation as haplotypes. In the following section, we discuss the role of HLA and other vertebrates MHC alleles, haplotypes and other genes within the MHC in specific diseases.

Autoimmune Diseases

Autoimmune diseases occur as results of elevated immune response against host self tissues and cells. The resulting inflammation and tissue damage (Lettre and Rioux, 2008) lead to such autoimmune diseases as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and type 1 diabetes (T1D).

In the case of MHC loci, the well-known theory for their role in autoimmune diseases is the escape of autoreactive T lymphocytes from tolerance due to aberrant class II presentation (MHC forming low-affinity or weak interactions with the peptides they present) of self or foreign antigens. This theory implicates only the MHC molecule as a cause of abnormal peptide presentation. Recent findings by way of the manner in which chromagranin A, an autoantigenic peptide is presented and recognized by autoreactive CD4+ T cell receptors thus leading to T1D (Stadinski et al. 2010) suggests an alternative hypothesis for MHC molecules involvement in T1D and other autoimmune disease. It has been observed that the chromogranin A peptide fills only part of the MHC binding groove and this MHC-peptide interaction is weak. This decreases the presentation of MHC-chromogranin A peptide in the thymus increasing T cell ability to evade central tolerance. Elsewhere in the body, high concentrations of the source protein produce more of these aberrant MHC-peptide complexes. The encounter between T cells and these complexes generate an autoimmune response.

In the following section, we present examples of known MHC alleles or haplotypes that have been associated with some autoimmune diseases.

Rheumatoid arthritis (RA) is a chronic inflammatory systemic disorder that affects joints, tissues and organs within the human body. The *HLA-DRB1* locus been shown to be associated with the disease and accounts for less than half of the genetic susceptibility (Newton et al. 2004). *HLA-DRB1 alleles* that are associated with RA susceptibility include *DRB1*0401*, **0404*, **0408*, **0101*, **0102*, **1402*, **09* and **1001* (Milicic et al. 2002). The *DRB1*0401* allele shows a high risk level with a relative risk (RR) of approximately 3 while the *DRB1*0101*, **0404*, **1001* and **0901* alleles exhibit a moderate RR around 1.5 (Fernando et al. 2008). Other alleles (*DRB1*0103*, *DRB1*0402*, *DRB1*0802* and *DRB1*1302*) are known to provide protection against RA (Newton et al. 2004).

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease affecting young women more than men. MHC class II alleles *HLA-DRB1*0301* and *HLA-DRB1*1501* and their haplotypes have been identified as risk factors for SLE in Caucasian populations (Fernando et al. 2007). Copy number variation within the *C4A* locus of MHC class III has been implicated as a risk factor of SLE. Specifically, studies in European American cohorts have shown that low C4A copy number is a risk factor while a high C4A copy number gives protection against SLE (Yang et al. 2007).

Type 1 diabetes (T1D) is a chronic autoimmune disease with a phenotypic manifestation of insulin deficiency and hyperglycemia resulting from immune-mediated selective destruction of pancreatic islet cells (Bluestone et al. 2010). *HLA-B*39* allele has been implicated in T1D susceptibility with a relative risk of 3.55 (Nejentsev et al. 2007). Some MHC loci harbor alleles that have opposing effects on T1D and other diseases. For example, two strong risk alleles for T1D, *rs9275383* and *rs477515* are strong protective alleles for ulcerative colitis (Wang et al. 2010) and this partly explains why risk alleles continue to be present in populations and are not selected against. The role of copy number in T1D has been

confirmed by a deletion on chromosome 6p21, close to an *HLA-DQ* allele in patients that had T1D or were at a high risk of acquiring the disease (Grayson et al. 2010).

Infectious Diseases

The polymorphic nature of the MHC loci ensures differences in susceptibility to infectious disease thus preventing the occurrence of the type of epidemics that have the potential to wipe out large proportions of a population. Though allelic diversity at the MHC loci may partly be influenced by pathogens, associations between infectious diseases and specific MHC alleles or haplotypes are seldom found (Goto et al. 2009). Difficulties in observing such associations arise partly from the fact that many infectious diseases are under the control of polygenes, genotype-environment interactions also influence such diseases and experimental designs required to pinpoint mechanisms of allele or haplotype disease associations in humans raises ethical issues (Nikolich-Zugich et al. 2004). Two infectious diseases and how HLA haplotypes predispose or protect an individual against these diseases are discussed below.

Malaria

Malaria infection caused by Plasmodium falciparum causes more than 2 million deaths annually (Hananantachai et al. 2005) and mainly affects children in sub-Saharan Africa (Diakite et al. 2009). The *HLA-B53* and *DRB1*04* have been associated with protection from P.falciparum malaria (Cserti-Gazdewich et al. 2011). An effective presentation of a particular T lymphocyte epitope on P. falciparum antigen 1 is believed to give *HLA-B53* allele this protective effect (Hill et al. 1992). A study involving children (6-60 months old) with severe malaria in multiple populations in Ghana confirmed *HLA-DRB1*04* allele as a risk factor in severe malaria (Osafo-Addo et al. 2008). The low frequency of *HLA-DRB1 *04* allele in sub-Saharan African populations as compared with other populations elsewhere in the world (Hill et al. 1992) may indicate that selection is acting against this allele in the sub region.

HIV-1 Related Disease

HIV positive patients show acute viremia followed by a decline to a relatively stable virus load set point (McMichael et al. 2010). GWAS have identified a polymorphism (rs2395029) within the HLA complex P5 (*HCP5*) gene which has an association with low viral load by virtue of being in LD with *HLA-B*5701* (Fellay et al. 2007), an allele known to be associated with low veremia. It has thus being suggested that the *HCP5* gene variant may be a marker for the *HLA-B*5701* protective effect (Fellay et al. 2008). Another polymorphism (rs9264942) located 35kb from the transcription start site in the *HLA-C* gene has being shown to be associated with low viremia and higher *HLA-C* gene expression (Fellay et al. 2007).

A very comprehensive GWAS involving HIV controllers (subjects able to control viral load in the absence of any medical intervention) and subjects with advanced form of AIDS from Europe, African-American and Hispanic groups identified a SNP (rs2523590) in a

noncoding site 2kb upstream of *HLA-B* which was associated with suppression of viral load in African-Americans and Hispanics (The International HIV Controllers Study, 2010). In the same study (The International HIV Controllers Study, 2010) SNP rs9264942 at 35kb upstream of *HLA-C*, SNP rs4418214 near *MICA* and rs3131018 in *PSORSIC3* all had protective effects on viral load in Europeans. Furthermore, in the European subjects, amino acid positions 67, 70 and 97 in the peptide binding groove of HLA-B were strongly associated with viremia with protective and risk disposition dependent on the type of amino acid present in these positions. Epistatic interaction between *KIR3DS1*, an activating *KIR* allele and a subset of the HLA-Bw4 cluster have been implicated in delayed progression to AIDS (Blackwell et al. 2009).

Two things are clear from these observations. First, GWAS has confirmed the role of the MHC region in HIV-1 by identifying novel variants. Second, lessons learnt from GWAS applications on HIV could be applied with modifications to investigate variants at MHC loci for association with other infectious diseases and therefore bridge the gap between autoimmune diseases and infectious diseases in terms of ease of GWAS applicability.

Inflammatory Diseases

Two chronic diseases that cause intestinal inflammation, Crohn's disease (CD) and ulcerative colitis (UC) have been extensively investigated in terms of their association with the MHC region. MHC region contribute about up to a third of the total genetic risk of CD and over half of that of UC (Muro et al. 2009). Studies by Ahmad et al. (2006) have confirmed the *HLA-DRB1*15* as a risk factor in UC patients of Europe, North Americans and Koreans. *HLA-B27* is an allele that leads to susceptibility to many infectious diseases including reactive arthritis, uveitis, ankylosing spondylitis, psoriasis and inflammatory bowel diseases (UC and CD), collectively called spondyloarthropathies (SPA). The allele can be found in 6-8% of Caucasian populations in the US (Khan, 1995). In the UK, 95% of patients with ankylosing spondylitis are *HLA-B27* positive (Antoniou et al. 2011). In the case of reactive arthritis, the prevalence of the disease is five times higher in *HLA-B27* positive people than the general populace (Colmegna et al. 2004). Several mechanisms have been proposed for the involvement of *HLA-B27* subtypes (*HLA-B*2701 to HLA-B*2725*) in diseases susceptibility.

One such mechanism is molecular mimicry where foreign pathogens stimulate self-reactive T cells that can respond to *HLA-B27* derived peptide. HLA-B27 misfolding hypothesis also suggests that misfolding of HLA-B27 in the ER induces cellular stress responses that interfere with normal cellular function and results in the production of proinflammatory cytokines which lead to SPA (Antoniou et al. 2011). A positive role for *HLA-B27* includes eliciting strong CTL responses against the influenza, EBV or HIV viruses (de Castro, 2005).

Table 4. Diseases of vertebrates associated with alleles at the MHC loci

Species	Disease	Associated MHC allele	Effect	Reference
Rhesus macaques	SIV disease[a]	*DQB1*0601,*	Susceptibility	Sauermann et al. 1997
		*DQB1*0601, 1806**		
Chicken	Marek's Disease	*B21*	Protective	Briles et al. 1977
		B19	Susceptibility	
Dog	SLO[b]	*DRB1*01801/DQA1*00101/ DQB1*00802*	Susceptibility	Wilbe et al. 2010
		*DRB1*02001/DQA1*00401/ DQB1*01303*	Protective	
Cattle	Dermatophilosis	*DQB*1804*	Susceptibility	Maillard et al. 1999
		*DRB3.2*09/*45*		

[a] simian immunodeficiency virus.
[b] Symmetrical lupoid onychodystrophy.

Association of the MHC Region with Diseases in Other Vertebrates

MHC alleles associations with diseases have been observed in non-human primates and mammals (Table 4). Briles et al. (1977) in a study on lines of chickens selected for either resistance or susceptibility to Marek's disease observed that chicken with the *B21* haplotype were resistant to Marek's disease while the counterparts with the *B19* haplotype were susceptible to the disease. In simian immunodeficiency virus (SIV) infected rhesus macaques, Mamu-*DQB1*0601* and a combination of *DQB1*0601*, *1806** alleles were significantly associated with rapid disease progression (Sauerman et al., 1997). Symmetrical lupoid onychodystrophy (SLO) is an immune mediated disease in dogs that affect their claws. The haplotype *DRB1*01801/DQA1*00101/DQB1*00802* is a risk haplotype that increased the risk of dogs developing SLO while the haplotype *DRB1*02001/DQA1*00401/DQB1*01303* conferred strong protection in dogs (Wilbe et al., 2010). The BoLA - *DQB1*01303* and *DRB3.2*09/*45* are strongly associated with susceptibility to dermatophilosis, a skin disease in livestock caused by the bacterium dermatophilosis congolensis.

Conclusion

The MHC molecules provide the surface for non-self antigen recognition by T cells. Negative and positive selection of T cells in the thymus achieves this discrimination. Across species, the MHC has evolved through loci loss and duplications. Investigations on the biology of the MHC region have given insight into the functions of MHC molecules such that apart from their immunological function of antigen presentation, they have been implicated in numerous other functions including a new role for MHC class II in TLR signaling. The numerous genes found in the MHC and the complexity of the region have made it a target for extensive research and alternative theories have been propounded to explain some hallmarks of the region including the high polymorphism and the association of genes with diseases. Two such new theories mentioned in this review are the 'associative balancing complex evolution' theory that suggests that accumulated recessive mutations near the MHC region contributes to the observed polymorphism and the cusp theory which proposes that ligands of specific HLAs in the cusp region of the MHC interact with non MHC receptors leading to the activation of pathways involved in immune response. MHC haplotypes may lead to disease susceptibility or may offer protection in numerous autoimmune, inflammatory and infectious diseases as well as other forms of diseases in humans, other placental mammals and non-mammalian vertebrates. Currently, new technologies in sequencing and genotyping together with meta-analysis that combine data form different GWAS studies are leading efforts in identifying new MHC alleles and haplotypes associated with complex diseases.

References

Abele, R. and Tampé, R. (2004). The ABCs of immunology: Structure and function of TAP, the transporter associated with antigen processing. *Physiology*, 19(4), 216-224.

Adams, E. J. Cooper, S. and Parham, P. (2001). A novel, nonclassical MHC class I molecule specific to the common chimpanzee. *Journal of Immunology*, 167, 3858-3869.

Ahmad, T. Marshall, S. E. and Jewell, D. (2006). Genetics of inflammatory bowel disease: The role of the HLA complex. *World Journal of Gastroenterology*, 12, 3628-3635.

Antoniou, A.N. Lenart, I. and Guiliano, D. B. (2011). Pathogenicity of misfolded and dimeric HLA-B27 molecules. *International Journal of Rheumatology*. doi. 10.1155/486856.

Anzai, T. Shiina, T. Kimura, N. Yanagiya, K. Kohara, S. Shigenari, A. Yamagata, T. Kulski, J. K. Naruse, T. K. Fujimori, Y. Fukuzumi, Y. Yamazaki, M. Tahiro, H. Iwamoto, C. Umehara, Y. Imanishi, T. Meyer, A. Ikeo, K. Gojobori, T. Bahram, S. and Inoko, H. (2003). Comparative sequencing of human and chimpanzee MHC class I regions unveils insertions/deletions as the major path to genomic divergence. *Proceedings of the National Academy of Science*, 100 (13), 7708-7713.

Bartl, S. Baish, M. Weissman, I. L. and Diaz, M. (2003). Did the molecules of adaptive immunity evolve from the innate immune system? *Integrative and Comparative Biology*, 43, 338-346.

Beck, S. Geraghty, D. Inoko, H. Rowen, L. Aguado, B. Bahram, S. Campbell, R. D. Forbes, S. A. Guillaudeux, T. Hood, L. Horton, R. Janer, M. Jasoni, C. Madan, A. Milne, S. Neville, M. Oka, A. Qin, S. Ribas-Despuig, G. Rogers, J. Shiina, T. Spies, T. Tamiya, G. Tahiro, H. Trowsdale, J. Vu, Q. Williams, L. and Yamazaki, M. (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature*, 401, 921-923.

Benacerraf, B. (1981). Role of MHC gene products in immune regulation. *Science*, 212, 1229-1238.

Bentley, G. Higuchi, R. Hoglund, B. Goodridge, D. Sayer, D. Trachtenberg, E. A. and Erlich, H. A. (2009) High-resolution, high-throughput HLA genotyping by next generation sequencing. *Tissue antigens*, 7(5), 393-403.

Bjorkman, P.J. Saper, M. A. Samraoui, B. Bennett, W. S. Strominger, J. L. and Wiley, D. C. (1987). Structure of the human class I histocompatibility antigen. HLA-A2. *Nature*, 329, 506–511.

Blackwell, J. M. Jamieson, S. E. and Burgner, D. (2009). HLA and infectious diseases. *Clinical Microbiology Reviews*, 22(2), 370–385.

Bluestone, J. A. Herold, K. and Eisenbarth, G. (2010). Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature*, 464, 1293-1300.

Bontrop, R. E. and Watkins, D. I. (2005). MHC polymorphism. AIDS susceptibility in non-human primates. *Trends in Immunology*, 26(4), 227-233.

Briles, W. E. Stone, H. A. and Cole, R. K. (1997). Marek's disease: effects of B-histocompatibility alloalleles in resistant and susceptible chicken lines. *Science*, 195, 193-195.

Brown, J. H. Jardetzky, T. S. Gorga, J. C. Stern, L. J. Urban, R. G. Strominger, L. C. and Wiley, D. C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, 364, 33-39.

Clarke, B. and Kirby, D.R.S. (1966). Maintenance of histocompatibility polymorphisms. *Nature*, 211, 999-1000.

Collins, E. J. Garboczi, D. N. Karpuss, M. N. and Wiley, D.C. (1995). The three-dimensional structure of a class I major histocompatibility complex molecule missing the α3 domain of the heavy chain. *Proceedings of the National Academy of Science*, 92, 1218-1221.

Colmegna, I. Cuchacovich, R. and Espinoza, L. E. (2004). HLA-B27-associated reactive arthritis: pathogenetic and clinical considerations. *Clinical Microbiology Reviews*, 17(2), 348-369.

Cserti-Gazdewich, C. M. Mayr, W. R. and Dzik, W. H. (2011). Plasmodium falciparum malaria and immunogenetics of ABO, HLA and CD36 (platelet glycoprotein IV). *Vox Sanguinis*, 100(1), 99-111.

de Almeida, D. E. and Holoshitz, J. (2011). MHC molecules in health and disease. At the cusp of a paradigm shift. *Self Nonself*, 2(1), 43–48.

de Castro, J. A. L. (2005). HLA-B27: portraying immuno dominant viral epitopes. *European Journal of Immunology*, 35, 336-340.

Deakin, J. E. Papenfuss, A. T. Belov, K. Cross, J. G. R. Cogill, P. Palmer, S. Sims, S. Speed, T. P. Beck, S. and Graves, J. A. M. (2006). Evolutionary and comparative analysis of the MHC class III inflammatory region. *BMC Genomics*, 7, 281.

Deng, L. and Maruzza, R. A. (2006). Structural basis for recognition of MHC and MHC-like ligands by natural killer cell receptors. *Seminars in Immunology*, 18, 159-166.

Denzin L. K. and Cresswell, P. (1995). HLA-DM induces CLIP dissociation from MHC class II alpha beta dimers and facilitates peptide loading. *Cell*, 82(1), 155-165.

Diakite, M. Clark T. G. Auburn, S. Campino, S. Fry, A. E. Green, A. Morris A. P. Richardson, A. Jallow, M. Sisay-Joof, F. Pinder, M. Kwiatkowski, D. P. and Rockett, K. A. (2009). A genetic association study in the Gambia using tagging polymorphisms in the major histocompatibility complex class III region implicates a HLA-B associated transcript 2 polymorphism in severe malaria susceptibility. *Human Genetics*, 125(1), 105-109.

Doherty, P. C. and Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, 256, 50-52.

Elliott, J. M. Wahle, J. A. and Yokohama, W. M. (2010). MHC class I – deficient natural killer cells acquire a licensed phenotype after transfer into an MHC class I-sufficient-environment. *JEM* (Brief Definitive Report), 207(10), 2073-2079.

Ewald, S. J. and Livant, E. J. (2004). Distinctive polymorphism of chicken B-FI (major histocompatibility complex I) molecules. *Poultry Science*, 83, 600-605.

Fellay, J. Shianna, K. V. Ge, D. Colombo, S. Ledergerber, B. Weale, M. Zhang, K. Gumbs, C. Castagna, A. Cossarizza, A. Cozzi-Lepri, A. De Luca, A. Easterbrook, P. Francioli, P. Mallal, S. Martinez-Picado, J. Miro, J. M. Obel, N. Smith, J. P. Wyniger, J. Descombes, P. Antonarakis, S. E. Letvin, N. L. McMichael, A. J. Haynes, B. F. Telenti, A. and Goldstein, D. B. (2007). A whole-genome association study of major determinants for host control of HIV-1. *Science*, 317, 944-947.

Fellay, J. Ge, D. Shianna, K. V. Colombo, S. Ledergerber, B. Cirulli, E. T. Urban, T. J. Zhang, K. Gumbs, C. E. Smith, J. P. Antonella Castagna, A. Cozzi-Lepri, A. De Luca, A. Easterbrook, P. Gu, H. F. Mallal, S. Mussini, C. Dalmau, J. Martinez-Picado, J. Miro, J. M. Obel, N. Wolinsky, S. M.. Martinson, J. J. Detels, R. Margolick, J. B. Jacobson, L. P. Descombes, P. Antonarakis, S. E. Beckmann, J. S. O'Brien, S. J. Letvin, N. L. McMichael, A. J. Haynes, B. F. Carrington, M. Feng, S. Telenti, A. Goldstein, D. B. NIAID Center for HIV/AIDS Vaccine Immunology (CHAVI) (2009). Common genetic variation and the control of HIV-1 in humans. *PLoS Genetics*, 5(12), e1000791.

Fernando, M. M. A. Stevens, C. R. Walsh, E. C. De Jager, P. L. Goyette, P. Plenge, R. M. Vyse, T. J. and Rioux, J. D. (2008). Defining the role of the MHC in autoimmunity: A review and pooled analysis. *PLoS Genetics*, 4 (4), e1000024.

Fernando. M. M. Stevens. C. R. Sabeti, P. C. Walsh, E. C. McWhinnie, A. J. Shah, A. Green, T. Rioux, J. D. and Vyse, T. J. (2007). Identification of two independent risk factors for Lupus within the MHC in United Kingdom Families. *PLoS Genetics*, 3(11), e192.

Flanjnik, M. F. and Kasahara, M. (2009). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews*, 47-59.

Gibson, G. (2010). Hints of hidden heritability in GWAS. *Nature Genetics*, 42, 558-560.

Goto, R. M. Wang, Y. Taylor Jr, R. L. Wakenell, P. S. Hosomichi, K. Shiina, T. Blackmore, C. S. Briles, W. E. and Miller, M. M (2009). BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proceedings of the National Academy of Science*, 106(39), 16740-16745.

Grayson, B. L. Smith, M. E. Thomas, J. W. Wang, L. Dexheimer, P. Jeffrey, J. Fain, P. R. Nanduri, P. Eisenbarth, G. S. and Aune, T. M. (2010). Genome-wide analysis of copy number variation in type 1 diabetes. *PLoS One*, 5 (11), e15393.

Groh, V. Steinle, A. Bauer, S. Spies, T. (1998). Recognition of stress induced MHC molecules by intestinal epithelial γδ T cells. *Science*, 279, 1737-1740.

Gruen, J. R. and Weissman, S. M. (2001). Human MHC class III and IV genes and disease association. *Frontiers in Bioscience*, 6, d960-972.

Hananantachai, H. Patarapotikul, J. Ohashi, J. Naka, I. Looareesuwan, S. Tokunaga, K. (2005). Polymorphisms of the HLA-B and HLA-DRB1 genes in Thai malaria patients. *Japanese Journal of Infectious Disease*, 58(1), 25-28.

Hendrick, J. P. and Hartl, F-U. (1993). Molecular chaperone functions of heat-shock proteins. *Annual Review of Biochemistry*, 62, 349-384.

Hill, A. V. S. Elvin, J. Willis, A. C. Aidoo, M. Allsopp, C. E. M. Gotch, F. M. Ming, G. X. Takiguchis, M. Greenwood, B. M. Townsend, Al. R. M. McMichael, A. J. and Whittle, H. C. (1992). Molecular analysis of the association of HLA-B53 and resistance to severe malaria. *Nature*, 360, 434-439.

Horton, R. Wilming, L. Rand, V. Lovering, R. C. Bruford, E. Khodivar, V.K. Lush, M. J. Povey, S. Talbot, M.C. Jr. Wright, M. W. Wain, H. M. Trowsdale, J. Ziegler, A. and Beck, C. (2004). Genome map of the extended human MHC. *Nature Review Genetics*, 5(12) 889-899.

Hsing, L. C. and Rudensky, A. Y. (2005). The lysosomal cystein proteases in MHC Class II antigen presentation. *Immunological reviews*, 207, 229-241.

Hughes, Austin L (May 2008) Major Histocompatibility Complex (MHC) Genes: Evolution. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net [doi: 10.1002/ 9780470015902.a0005133.pub2].

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

Hughes, A. L. and Nei, M. (1990). Evolutionary relationships of class II major histocompatibility genes in mammals. *Molecular Biology and Evolution*, 7, 491-514.

IPD-MHC (2011). Available at *http://www.ebi.ac.uk/ipd/mhc*. (accessed 17 December 2011).

Jones, E. Y. Stuart, D. I. and Walker, N. P. C. (1989). Structure of tumor necrosis factor. *Nature*, 338, 225-228.

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

Kaufman, J. Milne, S. Gobel, T. W. F. Walker, B. A. Jacob, J. P. Auffray, C. Zoorob, R. and Beck, S. (1999). The chicken B locus is a minimal essential major histocompatibility complex. *Nature*, 401, 923-925.

Kaufman, J. Volk, H. and Wallny, H. J. (1995). A minimal-essential-MHC and an unrecognized-MHC-2 extremes in selection for polymorphism. *Immunological Reviews*, 143, 63-68.

Kelley, J. Walter, L. and Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics,* 56, 683-695.

Kerr, M. A. and Porter, R. R. (1978). The purification and properties of second components of human complement. *Biochemistry Journal*, 171, 99-107.

Khan, M. A. (1995). HLA-B27 and its subtypes in world populations. *Current Opinion in Rheumatoloy*, 7, 263-269.

Kulski, J. K. Shiina, T. Anzai, T. Kohara, S. and Inoko, H. (2002). Comparative genomic analysis of the MHC: the evolution of Class I duplication blocks, diversity and complexity from shark to man. *Immunological Reviews*, 190, 95-122.

Kulski, J. K. Gaudieri, S. and Dawkins, R. L. (2000). Transposable elements and the metamerismatic evolution of the HLA class I region. In: Major histocompatibility complex. *Evolution, structure, and function.* (ed. M. Kasahara), pp. 158-177. Springer, Tokyo.

Kumánovics, A. Takada, T. and Lindahl, K. F. (2003). Genomic organization of the mammalian MHC. *Annual Review of Immunology*, 21, 629-657.

Laird, D. J. De Tomaso, A. W. Cooper, M. D. and. Weissman, I. L (2000). 50 million years of chordate evolution: Seeking the origins of adaptive immunity. *Proceedings of the National Academy of Science*, 97(13), 6924-6926.

Lanier, L. L. and Philips, J. H. (1996). Inhibitory MHC class I receptors on NK cells and T cells. *Immunology Today*, 17(2), 86-91.

Lettre, G. and Rioux, J. D. (2008). Autoimmune diseases: insights from genome-wide association studies. *Human Molecular Genetics*. 17 Review. Issue 2, R116–R121.

Lui, X. Zhan, Z. Li, D. Xu, L. Ma, F. Zhang, P. Yao, H. and Cao, X. (2011). Intracellular MHC class II molecules promote TLR-triggered innate immune responses by maintaining activation of kinase Btk. *Nature Immunology*, 12(5), 416-424.

Maillard, J.C. Chantal, I. Berthier, D. Stachurski, F. and Elsen, J.M. (1987). Molecular markers of genetic resistance and susceptibility to bovine dermatophilosis, *Archives of Animal Breeding*; Arch. Tierz, Dummerstorf, 42, 93–96.

Matsunaga, T. and Rahman, A. (1998). What brought the adaptive immune system to vertebrates?. The jaw hypothesis and the sea horse. *Immunological Reviews*, 166, 177-186.

Mayer, G. and Nyland, J. Major Histocompatibility Complex (MHC) and T cell receptors-Role in immune responses. *http://www.microbiologybook.org/bowers/mhc.htm.* (Accessed 25 October, 2011).

McDevitt, H. O. (2000). Discovering the role of the major histocompatibility complex in the immune response. *Annual Review of Immunology*, 18, 1-17.

McMichael, A. J. Borrow, P. Tomara, G. D. Goonetilleke, N. and Haynes, B. F. (2010). The immune response during acute HIV-1 infection. Clues for vaccine development. *Nature Reviews. Immunology*, 10(1), 11-123.

Medzhitov, R. and Janeway (Jr), C. A. (2002). Decoding the pattern of self and non-self by the innate immune system. *Science*, 296-298.

MHC Sequencing Consortium (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature*, 401, 921-923.

Milicic, A. Lee, D. Brown, M. A. Darke, C. and Wordsworth, B. P. (2002). HLA-DR/DQ haplotype in rheumatoid arthritis: novel allelic associations in UK Caucasians. *The Journal of Rheumatology*. 29, 1821-1826.

Morris, K. (2009). The feline Major Histocompatibility Complex. Orbit: *The University of Sydney undergraduate research Journal*, 1(1) 76-96.

Muro, M. López, R. Campillo, J. A. Salama, H. Martínez, P. Lucas, D. Valdés, M. Botella, C. Minguela, A. Miras, M. Carballo, F. and Álvarez-López, R. (2009). The role of Human Leukocyte Antigen (HLA) complex in BD: Crohn's Disease and Ulcerative Colitis. *Advanced Study in Biology*. 1(1), 37-41.

Nei, M. and Hughes, L. A. (1992). Balanced polymorphism and evolution by the birth-and-death process in the MHC loci. In: 11th Histocompatibility Workshop and Conference.. pp. 27-38, (eds. K. Tsuji, M. Aizawa, and T. Sasazuki) Oxford University Press, Oxford.

Nejentsev, S. Howson, J. M. M. Walker, N. M. Szeszko, J. Field, S. F. Stevens, H. E. Reynolds, P. Hardy, M. King, E. Masters, J. Hulme, J. Maier, L. M. Smyth, D. Bailey, R. Cooper, J. D. Ribas, G. Campbell, R. D. The Wellcome Trust Case Control Consortium, Clayton, D. G. and Todd, J. A. (2007). Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature*, 450, 887-892.

Newton, J. L. Harney, S. M. J. Wordsworth, B. P. and Brown, M. A. (2004). A review of the MHC genetics of rheumatoid arthritis. *Genes and Immunity*, 5, 151-157.

Nikolich-Zugich, J. Fremont, D. H. Miley, M. J. and Messaoudi, I. (2004) The role of MHC polymorphism in anti-microbial resistance. *Microbes and Infection*. 6(5), 501-512.

O'Brien, S. J. and Evermann, F. F. (1988). Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology and Evolution*, 3, 254-259.

Oroczo, L. D. Cokus, S. J. Ghazalpour, A. Ingram-Drake, L. Wang, S. van Nas, A. Che, N. Araujo, J. A. Pellegrini, M. and Lusis, A. J. (2009). Copy number variation influences gene expression and metabolic traits in mice. *Human Molecular Genetics*, 18(21), 4118–4129.

Osafo-Addo, A. D. Koram, K. A. Oduro A. R. Wilson, M. Hodgson, A. and Rogers. W. O. (2008). HLA-DRB1*04 allele is associated with severe malaria in northern Ghana. *American Journal of tropical medicine and hygiene*, 78(2), 251-255.

Phan, U. T. Arunachalam, B. and Cresswell, P. (2000). Gamma-Interferon-inducible Lysosomal Thiol Reductase (GILT). Maturation, activity and mechanism of action. *Journal of Biology and Chemistry*, 275, 34, 25907-25914.

Piani, A. Hossle, J. Birchler, T. Siegrist, C. A. Heumann, D. Davies, G. Loeliger, S. Seger, R. and Lauener, R. P. (2000). Expression of MHC class II molecules contributes to lipopolysaccharide responsiveness. *European Journal of Immunology*, 30, 3140-3146.

Piontkivska, H. and Nei, M. (2003). Birth-and-Death evolution in primate MHC class I genes. Divergence time estimates. *Molecular Biology and Evolution*, 20(4), 601-609.

Rajagopalan, S. and Long, E. O. (2005). Understanding how combinations of HLA and KIR genes influence disease. *The Journal of experimental Medicine*, 201 (7), 1025-1029.

Reid, K. B. M. and Porter, R. R. (1981). The proteolytic activation systems of complement. *Annual Review of Biochemistry*, 50, 433-464.

Robinson, J. Mistry, K. McWilliam, H. Lopez, R. Parham, P. and Marsh, S. G. E (2011): The IMGT/HLA database. *Nucleic Acids Research*, 39 Suppl. 1:D1171-1176.

Sakaguchi, S. Ono, M. Setoguchi, R. Yagi. H. Hori. S. Fehervari, Z. Shimizu. J. Takahashi, T. and Nomura, T. (2006). Foxp3+ CD25+ CD4+ natural regulatory T cells in dominant self-tolerance and autoimmune disease. *Immunological Reviews*, 212, 8-27.

Sauermann, U. Krawczak, M. Hunsmann, G. Stahl-Hennig, C.(1997). Identification of the Mhc-Mamu-DQB1 allele combinations associated with rapid disease progression in rhesus macaques infected with simian immunodeficiency virus. *AIDS*, 11, 1196–1198.

Schreiber, R. D. and Muller-Eberhard, H. J. (1974). Fourth component of the human component: Description of a three polypeptide chain structure. *Journal of Experimental Medicine*, 140, 1324-1335.

Shearer, W. T. Rosenblatt, H. M. Gelman, R. S. Oyomopito, R. Plaeger, S. Stiehm, E. R. Wara, D. W. Douglas, S. D. Luzuriaga, K. McFarland, E. J. Yogev, R. Rathore, M. H. Levy, W. Graham, B. L, Spector, S. A; Pediatric AIDS Clinical Trials Group (2003). Lymphocytes subsets in healthy children from birth through 18 years of age. *The Journal of Allergy and Clinical Immunology*, 112(5), 973-980.

Simpson, E. (1988). Function of the MHC. *Immunology*, 1, 27-30.

Singh N. Agrawal, S. and Rastogi, A. R. (1997). Infectious diseases and immunity: Special reference to major histocompatibility complex. *Emerging Infectious Diseases*, 3(1), 41-49.

Smith, T. P. (1995). Directed integration of the physical and genetic linkage maps of swine chromosome 7 reveals that the SLA spans the centromere. *Genome Research*, 5, 259-271.

Smith, R. A. and Baglioni, C. (1987). The active form of tumor necrosis factor is a trimer. *The Journal of Biological Chemistry*, 262, 6951-6954.

Sridhar, R. P. N. Major histocompatibility complex. *www.microrao.com*. (Accessed 25 October, 2011).

Stadinski, B. Kappler, J. and Eisenbarth, G. S. (2010). Molecular targeting of islet autoantigens. *Immunity*, 32, 446-456.

Starlets, D. Gore, Y. Binsky, I. Haran, M. Harpaz, N. Shvidel, L. Becker-Herman, S. Berrebi, A. and Shachar, I. (2006). Cell-surface CD74 initiates a signaling cascade leading to cell proliferation and survival. *Blood*, 107 (12), 4807-4816.

The International HIV Controllers Study (2010). The major genetic determinants of HIV-1 control affect HLA Class I peptide presentation. *Science*, 330, 1551-1557.

The Welcome Trust Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447, 661-678.

Thomson, G. (1995). HLA disease associations: models for the study of complex human genetic disorders. Critical review of Clinical Laboratory. *Science*, 32, 183–219.

Traherne, J. A. (2008). Human MHC architecture and evolution: implications for disease association studies. *International Journal of Immunogenetics*, 35, 179-192.

Truman J. P. Choqueux, C. Charron, D. and Mooney, N. (1996). HLA class II molecule signal transduction leads to either apoptosis or activation via two different pathways. *Cell Immunology*, 172,149-157.

Truman, J. P. Choqueux, C. Tschopp, J. Vedrenne, J. Le Deist, F. Charron, D. and Mooney, N. (1997). HLA class II-mediated death is induced via Fas/Fas ligand interactions in human splenic B lymphocytes. *Blood*, 89 (6), 1996-2007.

van Oosterhout, C. (2009). A new theory of MHC evolution: beyond selection on the immune genes. *Proceedings, Biological Science/The royal Society*, 276 (1657), 657-665.

Vandiedonck, C. Taylor, M. S. Lockstone, H. E. Plant, K. Taylor, J. M. Durrant, C. Broxholme, J. Fairfax, B. P. and Knight, J. C. (2011). Pervasive haplotypic variation in the spliceo-transcriptome of the human major histocompatibility complex. *Genome Research*, 21, 1042-1054.

Vivier, E. Tomasello, E. and Paul, P. (2002). Lymphocyte activation via NKG2D: towards a new paradigm in immune recognition? *Current opinion in Immunology*, 14, 306-311.

Walter, L. Hurt, P. Himmelbauer, H. Sudbrak, R. and Gunther, E. (2002). Physical mapping of the major histocompatibility complex class II and class III regions of the rat. *Immunogenetics*, 54, 268-275.

Wang, K. Baldassano, R. Zhang, H. Qu, H. Q, Imielinski. M, Kugathasan, S. Annese, V. Dubinsky, M. Rotter, J. I. Russell, R. K. Bradfield, J. P. Sleiman. P. M. Glessner, J. T. Walters, T. Hou, C. Kim, C. Frackelton, E. C. Garris, M. Doran, J. Romano, C. Catassi, C. Van Limbergen, J. Guthery, S. L. Denson, L. Piccoli, D. Silverberg, M. S, Stanley, C. A. Monos, D, Wilson, D. C, Griffiths, A. Grant, S. F. Satsangi, J. Polychronakos, C. and Hakonarson, H. (2010). Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes implicates multiple loci with opposite effects. *Human Molecular Genetics*, 19(10), 2059-2067.

Wearsch, P. A. and Cresswell, P. (2008). The quality control of MHC class I peptide binding. *Current Opinion in Cell Biology*, 20 (6), 624–631.

Weenik, S. W. and Gautam, A. (1997). Antigen presentation by MHC class II molecules. *Immunology and Cell Biology*, 75, 69-81.

Whitehead, A. S. Goldberger, G. Woods, D. E. Markham, A. F. and Colten, H. R. (1983). Use of human C4 cDNA clone for analysis of genetic deficiency of the fourth component (C4) in guinea pig. *Proceedings of the National Academy of Science*. 80, 5387-5391.

Wilbe, M. Ziener, M. L. Aronsson, A. Harlos, C. Sundberg, K. Norberg, E. Andersson, L. Lindblad-Toh, K. Hedhammar, A. Andersson, G. and Lingaas, F. (2010). DLA class II alleles are associated with risk for canine symmetrical lupoid onychodystrophy (SLO). *PLoS One,* 5(8), e12332. doi:10.1371/journal.pone.0012332.

Xie, T. Rowen, L. Aguodo, B. Ahearn, M. E. Madan, A. Qin, S. Campbell, R. D. and Hood, L. (2003). Analysis of the gene dense major histocompatibility complex class III region and its comparison to mouse. *Genome Research*, 13, 2621-2636.

Yang, Y. Chung, E. K. Wu, Y. L. Savelli, S. L. Nagaraja, H. N. Zhou, B. Hebert, M. Jones, K. N. Shu, Y. Kitzmiller, K. Blanchong, C. A. McBride, K. M. Higgins, G. C. Rennebohm, R. M. Rice, R. R. Hackshaw, K. V. Roubey, R. A. S. Grossman, J. M. Tsao, B. P. Birmingham, D. J. Rovin, B. H. Hebert, L. A. and Yu, C. Y. (2007). Gene copy number variation and associated polymorphisms of complement C4 in human system lupus erythematosus (SLE). Low copy number is a risk factor for and high copy number is a protective factor against SLE Susceptibility in European Americans. *The American Journal of Human Genetics*, 80, 1037-1054.

Yokoyama, W. M. and Plougastel, B. F. M. (2003). Immune function encoded by the natural killer gene complex. *Nature Reviews, Immunology*, 3, 304-316.

Yu, C. Y. Yang, Z. Blanchong, C. A. and Miller, W. (2000). The human and mouse MHC class III region: a parade of 21 genes at the centromeric segment. *Immunology Today*, 21(7), 320-328.

Yuhki, N. Beck, T. Stephens, R. Neelam, B. and O'Brien, S. J. (2007). Comparative genomic structure of human, dog, and cat MHC: HLA, DLA, and FLA. *Journal of Heredity,* 98(5), 390-399.

Zang, W. Kalache, S. Lin, M. Schroppel, B. and Murphy, B. (2006). MHC Class II–Mediated Apoptosis by a Nonpolymorphic MHC Class II Peptide Proceeds by Activation of Protein Kinase C. *Journal of the American Society of Nephrology*, 16, 3661–3668.

Zhang, F. Gu, W. Hurles, M. E. and Lupski, J. R. (2009). Copy number variation in human health, disease and evolution. *Annual Review of Genomics and Human Genetics*, 10, 451-481.