This is to certify that the

thesis entitled

SEQUENCING OF CANINE LOW MOLECULAR MASS PROTEASOMES (LMPs) AND TRANSPORTERS ASSOCIATED WITH ANTIGEN PROCESSING (TAPs)

presented by

Sharon Keely Palmer

has been accepted towards fulfillment
of the requirements for

Master of Science degree in Clinical Laboratory Science

John A. Gerlach

Major professor
JOHN A. GERLACH

Date 4/14/99

SEQUENCING OF CANINE LOW MOLECULAR MASS PROTEASOMES (LMPs) AND TRANSPORTERS ASSOCIATED WITH ANTIGEN PROCESSING (TAPs)

By

Sharon Keely Palmer

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degreee of

MASTER OF SCIENCE

Department of Medical Technology

1998

Abstract

SEQUENCING OF CANINE LOW MOLECULAR MASS PROTEASOMES (LMPs) AND TRANSPORTERS ASSOCIATED WITH ANTIGEN PROCESSING (TAPs)

By

Sharon Keely Palmer

The immune system is responsible for recognition of self vs. non-self and the elimination of detected foreign matter (Germain, Janeway). In humans, it is the Major Histocompatibility Complex (MHC) genetic products that regulates the immune responses. The MHC consists of three separate genetic regions, MHC-I, -II and -III (Trowsdale). Canines possess the Dog Leukocyte Antigens (DLA). The DLA is the canine equivalent to the human MHC and can be divided into the DLA-A, -B, -C and -D (Bull, Deeg).

Previous studies have shown genetic conservation can be noted between the DLA and MHC. In humans the endogenous immune response involves the genetic II region of the MHC, more specifically the low molecular mass proteasomes (LMPs) and transporter associated with antigen processing (TAPs) genes. Using human primers to the two genes, we were able to amplify canine genomic DNA via PCR. Purified amplicons were sequenced using LPCR (linear polymerase chain reaction). It was found that the canine LMP-2 has a 74% homology to that of the human LMP-2.

Dedication

In loving memory of Minnie O. Keely, Isabel Dement and Rose Starwasz.

May I always make you proud

Acknowledgments

I would like to express my gratitude to Dr. John Gerlach. His constant support, encouragement and valued suggestions meant more than he could possibly realize.

I would like to say a special thanks to Dr. Robert Bull and Doug Estry for their time and support. Their being a part of my research was greatly appreciated.

Lastly, I would like to thank my husband for his unconditional love and support. In some ways, this degree is just as much his as it is mine.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The immune system is responsible for the recognition of self, non-self and the elimination of foreign material (Germain, Janeway). Detection of foreign material within a host stimulates a sequence of signaling events which assists/aids in the processing of antigens and invokes the immune response. Transplantation of organs or invasion of the host by pathogens can set into motion the immune response.

The Major Histocompatibility Complex (MHC) plays a central role in the immunological response to foreign proteins (Trowsdale) as well as autoimmune disorders. The MHC is the genetic region where the class I, II and III genes are located. Genes in the MHC encode proteins that process and present immunogenic peptides (Trowsdale). It is believed that all vertebrates possess MHC genes (Trowsdale, Powis).

Several comparative studies have shown that the MHC is well conserved within vertebrates (Trowsdale, Polvi). The species comparison includes cattle (Anderson), sheep (Anderson), rat (Howard), goat (Anderson), rabbit (Chabbane) chicken (Kroemer), red deer (Swarbrick), and canine (Sarmiento). The MHC, however, is best known in humans and mice (Trowsdale).

In humans, the MHC is also referred to as the human leukocyte antigen complex or HLA (Klein). The human MHC is located on the short arm of chromosome 6. It is estimated the entire MHC spans over four million base pairs (Klein). There are over one hundred known genes within the MHC. On average, a gene can be located every 40 kilobase (kb), though the kilobase may vary depending on the region of the MHC (Marshall). As advances are made in detecting coding sequences (Marshall), it is likely more genes will be located.

The human MHC can be divided into three separate class regions: MHC-I, MHC-II and MHC-III. The HLA genes are important in tissue and organ transplantations as well as both autoimmune and non-autoimmune disease associations (Marshall). The highly polymorphic HLA genes A, B and C are located within the MHC I region. The HLA DP, DQ and DR genes, which are also highly polymorphic, are within the MHC II region (Trowsdale). Other genes within region II include TAP, transporter associated with antigen processing, and LMP, low molecular mass proteasome (Monaco). Class I and class II gene families are separated by approximately 1100 kb encoding gene of the class III region (Cross). Class III region encodes proteins also involved in immune responses (Trowsdale, Cross). Genes such as complement proteins C2,

Factor B and C4 and the tumor necrosis factors (TNF) A and B (Cross).

The H2 complex, the MHC in mice, is organized similar to that of the human MHC (Trowdales). Other than lacking the DP loci and having two Mb genes, the class II region resembles the human class II region.

Cattle MHC is termed BoLA or bovine lymphocyte antigen and is divided into four regions: class I, IIa, IIb and III (Anderson). In class IIa region, there are the DRA and DRB genes as well and DQ genes. The class IIb region is where the LMP-7 gene can be located (Trowsdale).

The RLA or rabbit leukocyte antigen maintains a class II region. Again like the species mentioned earlier the DP and DR genes are represented (Trowsdale).

The MHC in the canine is termed DLA or Dog Leukocyte antigens. The DLA is characterized by three serologically defined loci, DLA-A, -B, & -C (Bull) and a fourth loci, DLA-D (Deeg). The class I and class II DLA studies have discovered genes which are homologous to the human MHC as well as other species (Polvi, Trowsdale, Burnett).

The human class II region consist of the DP, DR and DQ genes with the TAPs and LMPs genes being located between DP and DQ as shown in Figure 1. Previous studies have shown canines possess DR, DP and DQ genes similar to that of humans (Sarmiento, Wagner). Seeing that the MHC is believed

Figure 1

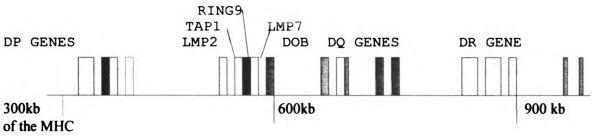


Figure 1. Human MHC Class II. Above is an illustration of the MHC class II region. One can note the close proximity of the genes. Comparison studies have shown that the canine possess homologous the DP, DP and DQ genes in humans. Our study was to detect and sequence the TAP and LMP genes of the canine.

to be a conserved region processed by all vertebrates, it is our hypothesis that canines also do maintain the genetic backbone for the class II encoded region for TAPs and LMPs.

Processing and Presentation of Antigens to T Cells

T cells require antigens to be recognized in the context of self-encoded cell surface molecules produced from genes in the MHC (Monaco). Two immunological pathways, endogenous and exogenous, exist within vertebrate cells to generate antigenic peptides for recognition by T cells (Monaco, Rudensky). The generation of antigenic peptides, presenting molecules and lymphocyte recognition vary between the two pathways (Neefjes).

Class II-Exogenous Pathway

Exogenous pathogens, such as *Mycobacteria*, enter the antigen presenting cells (APC) via extracellular domains. This process begins with the exogenous antigen being engulfed by endocytosis (Janeway). The exogenous antigen are degraded into antigenic peptides, usually 13-18 amino acids in length, by the endosomal/lysosomal degradative process (Monaco). The acidic pH in the endosome/lysosomes contributes not only to the antigen degradation but also the efficiency of the peptide binding to the MHC class II molecule groove (Neefjes).

The MHC class II molecules are constructed in the endoplasmic reticulum (ER). The Class II Molecules are assembled from α, β and Ii (invariant) chains. The Ii chain is associated transiently to the heterodimer and prohibits the binding of peptides in the ER to the Class II molecule. In the ER, the Ii chains associate as a homotrimer. The homotrimer forms a complex with three $\alpha\beta$ heterodimers. (Neefjes). The nine subunit complex is transported from the ER through the golgi apparatus to the trans-golgi reticulum (Neefjes). As the unit is traveling from the ER to the trans-golgi reticulum, Ii chain is degraded by endosomal proteases. The MHC class II molecule remains associated with a fragment of the Ii chain, referred to as CLIP (Fremont). From the trans-golgi reticulum, the MHC II molecule-CLIP travels via vesicular transport to the lysosome where it can bind to the antigenic peptide (class II-peptide). Foreign proteins enter this vesicular system either from the golgi or via extracellular endocytosis (Fremont) and are degraded into peptides. The peptide must have the correct sequence motif in order to compete with the CLIP for the binding groove. Once the peptide-class II complex is formed, it then travels to the APC surface and is presented to the CD4+ T lymphocytes. The CD4+ T helper cells do not kill the infected cells directly (Janeway). T helper cells, either Th1 or Th2, recognize the class II-

bound peptide and activate the immune response. Th1 cells will activate the APC and the APC in turn will destroy the invader. Th2 cells may activate B cells in order to have antibodies produced.

Class I-Endogenous Pathway

The class I immune response involves the processing of cytoplasmic pathogenic proteins by proteasomes, a multi-subunit protease located in the cytoplasm of most cells. The 26s protease complex is made up in part by the 20s proteasome. This 20s proteasome has a cylindrical structure with four layers of ring, each which is composed of seven subunits (Kandil). The outer rings are the α subunits and the inner two rings are composed of the β subunits. This is where the proteolytic core/catalytic sites reside. The β subunits are also known as the low molecular mass polypeptide or LMP-2 and LMP-7.

Expression of LMPs is increased by γ-interferon or IFN-γ (Kandil, Robertson). Cells, which have been stimulated by IFN-γ, will incorporate the two MHC encoded LMP-2 and LMP-7 subunits (Kandil). This is thought to occur by the displacement of the housekeeping subunits δ (subunit 2) and MB1 (subunit 10) within the proteasome (Kandil).

Proteasomes containing LMPs are able to cleave antigenic peptides (Monaco). Antigenic peptides are usually cleaved to be 8-15 amino acids in length with a hydrophobic carboxy-terminal residue (Monaco, Kandil, Carreno). This cleavage permits the binding in the F pocket of the MHC class I molecule. The F pocket is the groove, which peptide fragments of antigen interact with MHC molecule. More specifically, the groove consists of the $\alpha 1$ and $\alpha 2$ heavy chains. It is here that the antigen peptide will be positioned so it will be available for recognition by the CD8+ cells.

The next step for antigen processing/presentation involves the class II genes TAP-1 and TAP-2 (transporter associated with antigenic processing). The antigenic peptides are translocated from the cytosol of the cell to the lumen of the endoplasmic reticulum and possibly the cis-Golgi. This action is completed by a TAP heterodimer. The heterodimer contains two subunits, TAP-1 and TAP-2. Each subunit consist of a transmembrane segment, which spans the membrane 6-8 times, and a cytosolic domain containing ATP-binding cassette (Androlewicz, Carreno, Suh).

The TAP transport of the ligand is ATP dependent. Studies have shown that the peptides ranging from 8-15 residues are transported efficiently with some sequence preferences (Suh, Carreno). Peptides with H (histidine) or

E (Glutamic acid) as the COOH-terminus residue are translocated more efficiently than peptides with the residue of T or Threonine (Momburg, Neefjes). Chemical characteristics of the ligand, such as distance between N and C terminus of the peptide, also influence TAP transport (Germain). Each of these specific findings suggest that TAP has some polymorphic features which influence the binding of the ligands to the class I molecules.

Class I molecules are assembled in the ER. One important characteristics of the class I molecule is that the cell surface expression is distinguished by extraordinary protein polymorphisms (Burnett). The heterodimer consists of a polymorphic transmembrane heavy chain (45kD) and a noncovalently associated $\beta_2$microglobulin (Germain). The $\beta_2$microglobulin is a 12kd non-MHC encoded, soluble protein. The heavy chain consists of three α domains, each of which has interactions with the $\beta_2$microglobulin. The α_1 and α_2 domains together form the antigenic peptide groove. Class I molecules will undergo a complex series of interactions prior to presenting the antigenic peptides to the CD8+ cells. These interactions include chaperones molecules calnexin, tapasin and calreticulin (Sadasivan).

The free class I molecule first assoiciates with calnexin. Calnexin, also known as p88 molecule, is a

resident protein of the ER (Solheim, Song). Its role it to assist with the assembly and/or folding of the heavy chain (Solheim). It is also thought that calnexin plays a protective role in that it stabilizes the heavy chain and prevents intracellular degradation. Studies, however, have shown that calnexin negative cell lines are able to process and present antigens, proving calnexin may not be essential (Sadasivan). The current opinion is that another chaperone molecule similar to that of calnexin, such as BiP (immunoglobulin binding protein) may functionally replace calnexin when calnexin is not present.

Tapasin is a chaperone protein that is required for class I assembly. Studies have shown that tapasin bridges the heavy chain/$\beta_2$microglobulin dimers and TAP (Sadasivan). It has also been suggested that tapasin assists with peptide loading and/or monitoring of the folding of the class I molecule (Solheim).

Another chaperone molecule is calreticulin. It is structurally related to calnexin. Calreticulin associates with the heavy chain/$\beta_2$microglobulin and TAP (Solheim). An association of the class I heavy chain/$\beta_2$microglobulin dimers with calreticulin is a prerequisite for the association with TAP via tapasin (Sadasivan).

Once the class I heavy chain/$\beta_2$microglobulin-antigenic peptide complex is associated within the ER, the complex is

transported from the ER to the surface of the cell via a constitutive secretory vesicle. The role of the class I heavy chain/$\beta_2$microglobulin is to present the antigenic peptide on the surface of the APC to cytotoxic lymphocyte CD8+ cells. The α3 domain of the heavy chain associates with the APC for specificity recognition. For further illustration, see Figure 2. Polymorphic residues are positioned so that when presented by the MHC class I molecule, the antigenic peptide will line the floor and sides of the peptide binding groove (Newton-Nash). The groove itself consists of the α1 and α2 domains. Polymorphisms of peptide sequences concerning the anchor amino acids as well as side chains and the stoichiometry all play a significant role in the antigen presentation and recognition by APC.

It is possible for the endogenous antigenic peptides to be presented to CD8+ without certain aspects of the pathway, though the efficiency and effectiveness of the immune response may be hindered. For instance, cells with impaired TAP function express low levels for surface MHC class I complexes (Wolpert). Most of the MHC class I molecules are retained in the ER, few are transported to the surface. These MHC class I molecules are often referred to as empty or peptide receptive (Wolpert). Case studies have shown

Figure 2

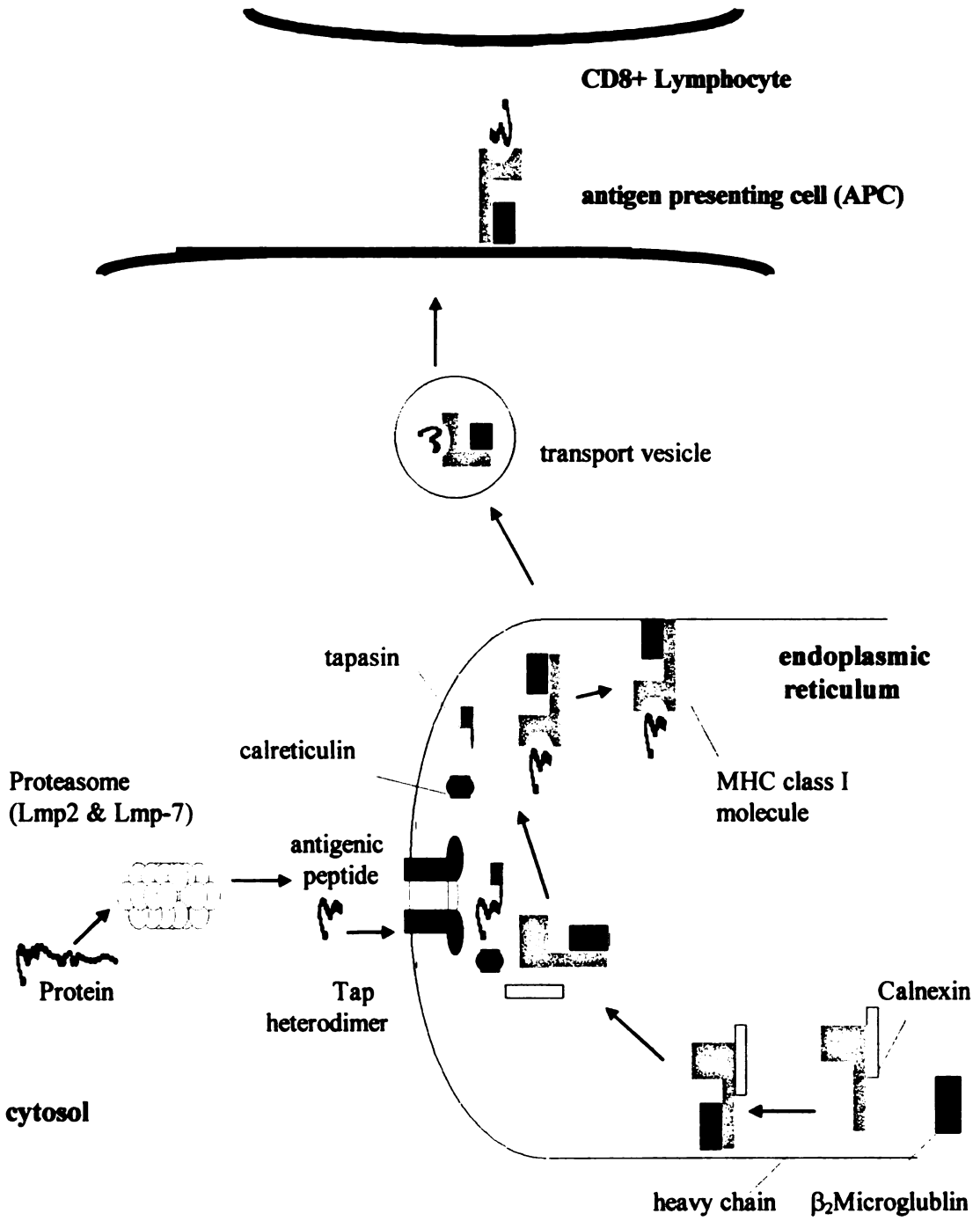


Figure 2. Endogenous Pathway. Above is an illustration of the endogenous pathway, representing the processing of antigenic protein and antigen presentation. The protein is degraded within cytosolic proteasome, transported into the ER via TAP heterodimer, binds to MHC class I molecule, which travels to the cells surface to present the antigenic peptide to the CD8+ cells.

that persons with defected or absent TAP heterodimer are more susceptible to certain pathogens (Schmitt).

The 20S proteasome consist of 14 subunits, one of which is LMP-7, another of which is LMP-2 (Groettrup). It is possible for the 20s proteasome to cleave antigenic peptides without LMP genes or with defective LMP genes. The degradation and cleavage priority, however, may be altered (Groettrup).

These types of defects in the MHC genes can lead to severe immunodeficiency syndromes or predispose the person to autoimmune afflictions. It is also possible however, that the defects in the genes are so minor, that no phenotypic changes can be noted in the patient.

Comparison Studies

In 1967 the first report of HLA associated disease was announced by Amiel (Thorsby). Today, the references include more than 500 disease states. For many, the disease associations are weak. Several, however, are strong. The pathogenesis for the substantial HLA association may involve one or more of the MHC genes (Thorsby). Strong HLA associated disease states include: ankylosing spondylitis (Ploski), Reiter's disease (Thorsby), acute anterior uveitits (Thorsby), subacute thyroiditis (Thorsby),

psoriasis vulgaris (Thorsby), narcolepsy(Thorsby) , Graves' disease (Thorsby), myasthenia gravis (Thorsby), Addison's Disease (Thorsby), rheumatoid arthritis (Thorsby), juvenile rheumatoid arthritis (Ploski, Thorsby), celiac disease (Tighe, Thorsby), multiple sclerosis (Moins-Teisserenc, Middleton, Thorsby) and diabetes, type I (Thorsby).

Though canines are not subject to all of the afflictions given above, advances in human genetics can affect the progress in other species. Concerning the disease states that do cause pathological conditions in canines, such as diabetes, knowing the disease HLA association may lead to advances in veterinary medical diagnosis and treatments.

Each animal species has a specific number of chromosomes, the arrangement of genes are similar for all species. This is known as genetic conservation. Because of genetic conservation, researchers are able to use animals for human medical models and humans for animal medical models.

Canines are an excellent model for medical comparative studies. This is true for several reasons. 1. Canines can possess phenotypic traits that might not be seen in smaller animals (Das). 2. There are known genetic similarities between humans and canines (Wagner) as well as other species. 3. Certain pathological states, which are thought

to be genetically related to MHC mutations, can be seen in human and as well as canines (Dean).

Canines maintain class I molecules which have the arrangement of exons and introns, as well as expected protein sequences (Burnett). Since the class I molecules sustains an important role with regards to antigenic presentation and the furthering of the immune response, structural integrity is essential. Structural analysis found 3 complete MHC class I genes. Each of the 3 genes was shown to be transcribed in canine peripheral blood leukocytes.

Other studies used HLA-monoclonal antibodies to canine lymphocytes (Chouchane). With the use of flow cell cytometry, they noted that a cross reactivity was detectable to canine lymphocytes using anti-human CD8+, thus suggesting that canines have CD8+ cells.

Other MHC related canine studies have been concentrated in the regions of the DR, DP and DQ genes. These studies have confirmed that canines posses the DR, DP and DQ genes (Williamson).

Since canines have both CD8+ cells and functional MHC class I molecules, we know that the end results of the endogenous pathway is comparable to that in humans. There is also further evidence of genetic conservation within the DR, DP and DQ regions. With the DR, DP and DQ genes being

located within close proximity to the TAPs and LMPs in human, the theory of genetic conservation would lead us to believe that canines should indeed have both the genes for TAPs and LMPs.

Given our background knowledge, the objective of this study was to detect and sequence canine TAPs and LMPs. This was to be completed by the use of PCR. The PCR utilized human primers specific for the TAP and LMP genes. These primers were to amplify homologues in the canine species. Amplicons generated were sequenced and compared to DNA in Genbank to verify fidelity to that of the known sequences. The percent homology was then calculated.

MATERIAL AND METHODS

Obtaining Blood Samples

Blood samples from both canine and human subjects were obtained via venipuncture. Samples were collected in ACD tubes. Blood samples were centrifuged in a swinging bucket rotor (IEC,USA) for 20 minutes at 750x gravity (g). Buffy coats were transferred into 15 ml (milliliter) conical polypropylene screw cap tubes (Corning, Inc.; Corning, NY). Red cell lysis was accomplished by adding 100 millimolar (mM) Tris ammonium chloride to the tubes (total volume of 15 ml) and incubated for 5 minutes at 37 degrees Celsius (°C). After incubation, the samples were centrifuged at 200x g for 20 minutes. The supernatant was discarded and the white cell pellet was retained for genomic deoxyribonucleic acid (DNA) isolation.

DNA Isolation

A DNA salt extraction method (Trucco) was used to obtain genomic DNA from purified peripheral white blood cells. WBCs were washed once with DPBS (DPBS without calcium or magnesium) and centrifuged at 225x g for 20 minutes. The supernatant was removed without disturbing the pellet and the cells were resuspended in the residual liquid. Lysis solution (0.375ml of per ml of whole blood

collected), 20% SDS (18.75ml per ml of lysis solution) and 1.5ul of Proteinase K were added to the cells. The tubes were inverted and placed on a rotator for 30 minutes at 37 °C. The tubes were removed from the rotator and were incubated 2-8 hours at 37 °C.

After the long incubation, 100 micrograms/milliliter (ug/ml) of DNase free RNase was added and the samples incubated for 1 hour at 37 °C. Next 0.125ml of saturated 6 Molar (M) NaCL solution (per ml of whole blood collected) was added and the tubes were shaken for 10 second. The samples were centrifuged at 850x g for 10 minutes. The supernatant (DNA) was poured into sterile tube and the remaining pellet (protein) was discarded. The supernatant was centrifuged again at 850x g for 10 minutes.

The supernatant was transferred into a glass borosilicate tube. A 2x volume of cold, absolute ETOH was added to each sample to precipitate out the DNA. The DNA samples were transferred to microtube with plastic transfer pipette and washed twice with 90% ETOH. The pellet was dried by dry vacuum for 5-17 minutes without heat. The DNA was resuspended in 100 millimolar (mM) Tris HCL and incubated overnight at room temperature.

The quantity and quality of DNA was determined by spectrophotometric analysis. One microliter (ul) DNA was diluted 1:200 with distilled deionized water (ddH_2O).

Absorbance (A) readings of the DNA solution were taken at wavelengths of 280 nanometers (nm) and 260 nm. Quality of DNA was determined by the absorbance ratio at 280 to 260 nm. DNA was quantified at 260 nm. DNA samples were diluted with ddH_2O for a final concentration of 1.0 microgram per microliter (ug/ul) as calculated by the formula:

[(A260)/(200/1)(50 ug)]/1000 ul= Xug/ul.

Genomic DNA Amplification by Polymerase Chain Reaction

Both human and canine genomic DNA were amplified using the same procedures and standard amounts of reagents. Human DNA was the positive control. The negative control used ddH_2O in replacement of the genomic DNA. 1.0 ug of DNA was utilized for each of the polymerase chain reactions (PCR). A combination of Hot Starts (Perkin Elmer) and Touchdown methods (Hecker, 1996) were used to produce higher yields of specific PCR amplified products. PCR amplification was performed with the Gene-Amp PCR system Perkin Elmer 9600 (Perkin Elmer Cetus; Norwalk, VA). The primers used are illustrated in Table 1.

PCR amplification was completed in 0.2 ml thin walled reaction tubes with a total volume of 100 ul. Reaction mixture was set-up as follows: Part I 1). 10 ul of genomic DNA (1 ug) samples or 10 ul ddH_2O for the negative control

Table 1

Human Primers	Sense	Antisense	Reference
LMP 2	tctcctttgcatatcccaat	cacagctctagggaaactc	VanEndert
LMP 7	cagaaagggcacgctcttgt	agagaacacgcagaacat	Fruh
TAP 1			
Set A	ggcgtccgagtgccaatg	gttctgttggaaaaactccg tc	Colonna
Set B	gggtgacgggatctataaca	cagcccctgtagcactaag	
Set C	agtggtctgttgactccctt	gcgcaggtctgacaaggc	
TAP 2			
Set A	agagacggagcggacctc	atccgcaagttgattcgaga	Colonna
Set B	ccgaggaggctgcttcac	accaggacgtagggtaaac	
Set C	tctcctttgcatatcccaat	cacagctctagggaaactc	

Table 1. PCR Primers. The above primers were used in the PCR process. TAP-1 and TAP-2 primers expand the entire gene, whereas the LMP-2 and LMP-7 do not. The LMP primers concentrate on a certain region of each gene. Both sense and antisense are in 5′ to 3′ form. Protocol for cycling for each primer listed above can be found in Table 2.

2). 65 ul of reaction mix consisting of 10 ul of 10x reaction buffer (500 mM KCl; 100 mM Tris-HCl, pH 8.3; 15 mM MgCl2: 0.1% gelatin), 8 ul of deoxynucleotide 5' triphosphate (dNTP mix), consisting of 100 mM deoxyadenosine 5'-triphosphate (dATP), 100 mM deoxycytosine 5' triphosphate (dCTP), 100 mM deoxyguanosine 5' triphosphate (dGTP), 100 mM deoxythymidine 5' triphosphate (dTTP), 27 ul of ddH_2O and 5 ul of each of the 20 uM paired sense and anti sense primers. Part II Gem 50 wax beads (Perkin Elmer, Branchburg, New Jersey), were then added to the each 0.2 ml reaction tubes. Tubes were capped and placed in a heating block at 80 °C until the wax melted (approximately 5 minutes). Tubes were removed from the heating unit. Wax was allowed to cool at room temperature. Part III Using the hot start methods, an additional mixture of 34 ul ddH_2O and 1.0 ul (5 units) *Thermus aquaticus* (Taq) DNA polymerase (Perkin Elmer, Branchburg, New Jersey) were placed on top of the harden wax in each reaction tube.

Prior to cycling, tubes were incubated at 95 °C for 5 minutes to allow melting of the wax barrier and the mixing of reagents. The Denaturing temperature was programmed for 94 °C for 30 seconds. Annealing temperatures and time varied depending on the primer pairs (see table 2) in order to obtain specific and high yielded products. Each primer

Table 2

Initial Incubation	Denaturation	Annealing	Extension	Cycles
95 C hold - 5 min	94 C - 11 sec	59 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	58 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec *	12
95 C hold - 5 min	94 C - 11 sec	59 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	58 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec *	12
94 C hold - 5 min	94 C - 30 sec	53 C - 60 sec	72 C - 60 sec	12
	94 C - 30 sec	52 C - 60 sec	72 C - 60 sec	12
	94 C - 30 sec	51 C - 60 sec	72 C - 30 sec *	12
95 C hold - 5 min	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	56 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	55 C - 30 sec	72 C - 30 sec *	12
95 C hold - 5 min	94 C - 11 sec	59 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	58 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec *	12
95 C hold - 5 min	94 C - 11 sec	59 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	58 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec *	12
94 C hold - 5 min	94 C - 11 sec	59 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	58 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec *	12
95 C hold - 5 min	94 C - 11 sec	57 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	56 C - 30 sec	72 C - 30 sec	12
	94 C - 11 sec	55 C - 30 sec	72 C - 30 sec *	12

Table 2. PCR Cycling Protocols. This table represents the PCR cycling procedures that were used to obtain amplified products for both the human and canine genomic DNA. * A final extension, 72°C for 10 minutes, was done to insure completion of LPCR products.

amplification profile used the touchdown method (Hecker) with 3 separate temperatures for annealing. An example of the annealing temperatures is: 57 °C for 10 cycles, 56 °C for 10 cycles and 55 °C for 10 cycles. Extension for each of the touchdown cycles was completed at 72 °C. Following the 30th cycle of PCR, an extension of 72 °C for 10 minutes occurred.

Gel Electrophoresis for Fidelity of Amplification

Fidelity of the PCR procedure was determined by polyacrylamide gel electrophoresis (PAGE). 5 ul of the PCR amplified product was added to 1 ul of 6x loading buffer (0.25% (percent) bromophenol blue, 0.25% xylene cyanol, 15% ficoll in ddH_20 (Sigma Chemicals; St. Louis, MO). The product-loading buffer mixtures, along with molecular weight marker VIII were loaded into separate lanes of a 6% polyacrylamide gel (29 part acrylamide to 1 part n,n'methylene bisacrylamide, Bio Rad; Hercules, CA). 1X Tris acetate EDTA (TAE) buffer (40 mM Tris-acetate, 1 mM EDTA; pH 8.0) was used in the mini-protean II gel system (Bio Rad; Hercules, CA) for PAGE. Samples were electrophesed by applying 200 V to the system for approximately 15 minutes. Gels were stained with ethidium bromide (0.5 ug/ml) for 15 minutes and destained with ddH20 for 5 minutes. DNA bands were visualized at 254 nm using

ultraviolet trans-illumination light source (UVP, Inc. ts36: San Gabriel, CA). A photograph was taken of the gel using FCR-10 camera (Fotodyne, Inc.: New Berlin, WI) with 667 Polaroid film (Polaroid Corporation: Cambridge, MA).

DNA Elution of Amplicon

Two different procedures, DNA electroelution and mini-column elutions, were used to purify the PCR amplified products. DNA electroelution was as follows: 6% PAGE, yhe product bands of interest were excised from the gel. The gel fragments were electroeluted using deairated 1x TAE buffer in the Micro Centrilutor System (Amicon; Beverly, MA) with Centricon-100 concentrators (Amicon; Beverly, MA) at 275 V for 2-8 hours. Purified DNA was rinsed 3 times in a Centricon-100 with ddH_2O by centrifugation using a fixed angle rotor for 20-30 minutes at 1000x g. In order to collect the DNA the Centricon-100 assembly was inverted and centrifuged for 5 minutes at 1000x g. Eluted samples were then transferred to a microcontainer and stored at -20 °C.

The procedure using mini-column elutions was as follows: PCR products were gel purified by using 1% low-melting temperature agarose gel with 1x TAE buffer in an agarose gel electrophoresis system (Bio-Rad, Hercules, CA). The amplified bands of interest were excised from the gel and purified using Promega Wizard PCR Preps Purification

Resin Kit (Madison, WI). Eluted samples were stored at -20 °C.

Linear Cycle Sequencing

The elution purified dsDNA PCR products were used with the appropriate primer(s) with the AmpliCycle™ sequencing Kit (Perkin Elmer, Foster City, CA) to perform Linear Cycle sequencing. Components of the sequencing kits were: 10x buffer (4 units AmpliTaq® DNA Polymerase, CS, 500 mM (millimolar) Tris-HCl; pH 8.9, 100 mM KCl, 25 mM MgCl2 solution, 0.25% (v/v) Tween® 20), G termination mix (22.5 uM (micromolar) C^7 dGTP, 10 uM each dATP. dCTP, dTTP and 80 uM ddGTP), A termination mix (22.5 uM C^7 dGTP, 10 uM each dATP. dCTP, dTTP and 600 uM ddATP), C termination mix (22.5 uM C^7 dGTP, 10 uM each dATP, dCTP, dTTP and 300 uM ddCTP), and T termination mix (22.5 uM C^7 dGTP, 10 uM each dATP, dCTP, dTTP and 900 uM ddTTP). The sequencing set up was as follows: 3-7 ul of dsDNA eluted product, 4 ul of sequencing 10x buffer, anti-sense or sense primer (20 uM) calculated at 0.0588 ul/base volume added, and 1.0 ul (10 mCi/ml or millicurie/milliliter) of alpha-Phosphorus-33 (α-33P),(NEN Life Science Products) and ddH_2O to achieve total volume of 30 ul. The sequencing master mixture was mixed well. 6 ul of the master mixture was added to 2 ul of each of the ddNTPs termination mixtures. The sequence cycling was completed with the Gene-AMP PCR system Perkin Elmer 9600

with 0.2 ml thin walled reaction tubes. The Denaturing step was 96 °C for 30 seconds. The extension step was 72 °C for 1 minute 30 seconds. The annealing temperature was set at the lowest of touchdown profile annealing temperature used during PCR amplification. The LPCR was completed in a total of 25 cycles. 4ul of stop solution (95% formamide, 20 mM EDTA, 0.05% brompheol Blue and 0.02% Xylene cyanal FF) was added to the tubes upon completion of the cycles (volume totaled 12 ul). Reactions were stored at -20 °C (Celsius).

Sequencing Gel Electrophoresis

The 40 cm (centimeters) long by 20 cm wide by 0.4 mm (millimeters) thick sequencing gels consisted of 7.2% polyacrylamide (35% acrylamide: 1 n,n'methylene bisacrylamide), 6M urea and 1x Tris borate EDTA (TBE). The gels were run on the Sequi-gen sequencing cell (Bio Rad, Hercules, CA) with 1x TBE buffer. The gels were preheated to a temperature of 55 °C by applying a voltage of 1700-1800 V (volts). Prerun time was approximately 45 minutes. Tubes containing the sequencing reactions were heated at 100 °C for 1 minute to denature dsDNA. 2 ul of each of the four reactions were added to separate lanes on the gel. The voltage was adjusted to maintain a constant temperature of 55 °C. After loading the gel, the voltage was applied for approximately 90 minutes. This amount of time would allow

the loading dye front to completely run off the bottom of the gel. At that point another 2 ul of each of the reactions was loaded onto the gel into separate, unused lanes. This process would be repeated until a total of 3 to 4 loadings had been completed. The gels were removed from the glass plates and transferred to blotting paper and dried using a heat vacuum dryer (Fisher-Biotech FB-GDS-4050; Springfield, NJ) at 80 °C for 1 hour. The heat unit of the vacuum dryer was turned off and the gel was allowed to cool under vacuum for 30-45 minutes.

Autoradiographs

The α-33P radiolabelled gels were exposed to x-ray films (Kodak; Rochester, New York) for several hours to several days, in order to obtain optimal signals. The films were developed using Kodak developer and Kodak fixer and replinisher (Kodak; Rochester, New York), then rinsed well with distill H_2O. The wet radiographs were allowed to air dry for at least one hour before being analyzed. The sequences were obtained by reading the gels manually by use of a light box (Science Accessories Corporation; Statford, CT).

Secondary Primers and Sequence Walking

Individual sequences were compiled to build a large contiguous sequence. Secondary primers were needed to

compare the sense-antisense overlap. The DNA code for the secondary primers was obtained from the autoradiograph sequences. These secondary primers were used in the linear sequencing reactions in place of the original PCR primers. The same protocol was used as previously stated.

Sequence Analysis and Comparison

Individual sequences were compiled to build a large contiguous sequence. The sequencing data obtained was analyzed using the Genetic Computer Group software (GCG; Madison, WI). Programs used within GCG included: Seq Ed (sequence data, reverse and complement), FastA (fidelity of the sequences), Best Fit (comparison of known to unknown sequence) and Pile Up (multiple sequence comparisons). GCG incorporates known sequence data from Genbank. The data base allows for the comparison of unknown sequence to that of known sequence of several Prokaryotic and Eukaryotic species.

RESULTS

Venipunture procedures for peripheral blood collection were completed to obtain white blood cells (WBCs). WBCs were used in the DNA salt extraction method for DNA isolation (see material and methods). Purity and quantity of each DNA sample was analyzed using a spectrophotometer. DNA samples that had a purity ratio of 1.8 or higher were used for the PCR processes. DNA purity average was 1.84.

PCR amplified products were analyzed utilizing 6% PAGE. Typical PCR amplification results can be found in Figure 3. Any contamination of the fragment of interest, whether it be because of excess primer, primer dimer, nonspecific amplification of secondary PCR products, or residual genomic DNA could corrupt the LPCR process. For these given reasons, amplified products were purified by one of the elution processes as described in material and methods.

LPCR's were completed using 3-7ul of purified PCR product. Each sequencing run, 3 to 4 loadings of the sequencing reaction, gave approximately 275-350 base pairs of readable sequence. Figure 4 is representative of a typical autoradiograph.

PCR Results

Each set of the Homo sapien primers amplified canine genomic DNA. Though partial sequencing data was obtained

Figure 3

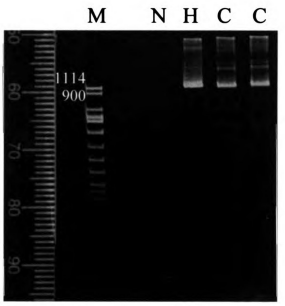
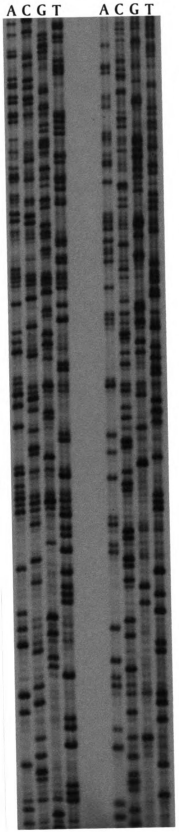


Figure 3. 6% PAGE. Above is a represenetation of a typical 6% PAGE of LMP-2 PCR products. The symbols are as follows: M - Marker VIII. N - negative control (blank). H - human genomic DNA (positive control), amplicon 1163 bp. C - canine genomic DNA, amplicon approximately 1000 bp.

Figure 4. Sequencing Autoradiograph. This is a represenation of a typical sequencing autoradiograph. Sequencing gels were completed with LPCR and radiolabelling. The radiolabelling was achieved via use of α-P^{33}. Sequencing gels were loaded left to right - A, C, G and T

Figure 4



for what is believed to be the both the canine TAPs and LMPs genes, LMP-2 was the only sequences in which a sense/anti-sense overlap was established.

PCR products the human and canine DNA were obtained for all of the given primer sets (see material and methods). Each canine PCR product detected was equivalent in length to that of the corresponding human product, except for primer sets of TAP1-3 and TAP2-3, note Table 3.

An interesting observation was noted with both the TAP1-3 and the TAP2-3 primer sets. The canine amplified product was significantly shorter than that of the comparative human products. DNA length of the genes will vary from each species. A good example of this is a comparison of the Homo sapien to that of the murine seen with the MHC gene LMP2. The length of the genomic DNA of murine is considerably longer than the human. Given the same logical thought, it is possible that the length of the canine TAP-1 and TAP-2 genes are significantly shorter in length than that of the human.

Sequencing Results

LMP-2

The PCR product amplified by the selected Human LMP-2 primers selected was known to be 1163 base pairs in length. Canine LMP-2 was calculated to be between 1000-1032 base

Table 3

Primer Set	Human Amplicon (bp)	Canine Amplicon (bp)
TAP 1-1	1494	1400
TAP 1-2	3647	1150
TAP 1-3	3196	350
TAP 2-1	3022	1800
TAP 2-2	5323	5000
TAP 2-3	1952	350
LMP-2	1163	1000
LMP-7	3454	3300 & 3600

Table 3. PCR Fragmentation Size, Human and Canine. Each of the primer sets listed above has the given amplicon for both human and canine genomic DNA. The human amplicons are known size fragments, confirmed by gel determinations. The canine amplicon bp sizes were determined by use of PAGE fragmentation via use of the marker and the corresponding human amplicon (Maniatis).

pairs in length via use of the gel calculations, actual obtained sequence was 1008 base pairs in length. In order to have sense and anti-sense strand overlap, secondary primers were designed for sequence walking.

Primer 1 AAA TGA GTA TAT GGG
Primer 2 AGT CTA ACT TGG ATC ATT AC

The GCG program, SeqEd, was use to compile the 1000 plus base pairs of canine LMP-2 sequence. The SeqEd program permits a given sequence to be edited, reversed and complemented. This provided the computer software that was needed to determine the sense/anti-sense overlap. The compiled Canine LMP-2 sequence can be seen in Figure 5.

Once the sequence was arranged, the GCG FastA program was utilized for the comparison of the canine LMP-2 sequence to that of known sequences. GCG FastA program identifies the designated sequence to that of the known/published sequence in Genbank data bases. The comparison(s) is/are based on the similarity between sequences. The results of the FastA LMP-2 Canine can be seen in Figure 6.

As expected, Homo sapiens LMP-2 gene corresponded to the Canis familiaris were alignments of repetitive units found in the canine genome. This led to the conclusion that a portion of the canine LMP-2 gene was sequenced.

Figure 5

```
   1  AATGTACCAG TCAAGGTGAA TGCTTGTXXA AGAAAGCCAA TGGCCCCAAG
  51  TACATGATTT TTTCTTACCC ATTCATGAAC TTATTAGTAA AGTGGAGCTT
 101  CGGCAGAAAC TGAGTTAACC TTTTAGAAAG GCCCCTATAA TGAGATACTG
 151  GGCAAGTGCT TGGGTCCCTA AATTCATGGA GAATCTGGGG CCTAAAAGCT
 201  TATGTGGCCT GTCCCTGCAG GGATGGGTAT AGAACGATAG GGTATGAGAA
 251  GAAGAAGTAG GATGCAAAGA TTTCAGGGGC CTGCCTTCCT TTTATGTTTA
 301  GAGATAGAGA AGACGGCTCT GCCAGGAAGA TATTCTGGTT TTTCATCCAG
 351  CACATGTTGA TCTGGTTCTG TCTCCTCCAT TCTGGCCAGT TTCCCCCTAT
 401  GCAGAATGAG TATATGGGAT TAGGCAATAT CACTTACTTT ACATCCCATA
 451  GAACACGAGA CTTTCTCARG ATGGATAGAG CGAGGAATTT CTGAGGAATA
 501  TTATTCTGAA TTTACTATTT TTTTGTTTTA TTATGGAATC GAAGCCGXXX
 551  AAGAAGTCTA ACTTGGATCA TTACTTTCTG TTTCATTTAT TGGAGATTAT
 601  TTTTTTCTTC CTGTTTGAAT CCAAGTATCC AATATCATGG ACTCAGGGAT
 651  CCCTGGGTGG CTCGGCGGTT TGGCACCTGC CTTTGGCCCA GGGCGCGATC
 701  CTGGAGTCCT GGGATCGAGT CCCGCCTTGG GCTCCCGGCA TGGAGCCTGC
 751  TTCTCCCTCC TCCTGTGTCT CTGCCTCTCT CTCTCTCTGT CTATGTCTAT
 801  CGTAAATAAA TGAGTAAATT TTAAAAAAAA AAAAAATCAT GGACTCCAGT
 851  GCTCTTCTCT GTCAAATAAA TTTGGGAAAC ACTGATTAAC CTAAAAAATT
 901  CTGGCGTTGC TAATGTTGCT GCCTGCTGAC TCCTTCCCTG CAGAGAAGCA
 951  GTGGGTAAAC CGAGCATTTX ACAAGCTGTC CCCACTACAC CAACACATGT
1001  ACTGTGCT
```

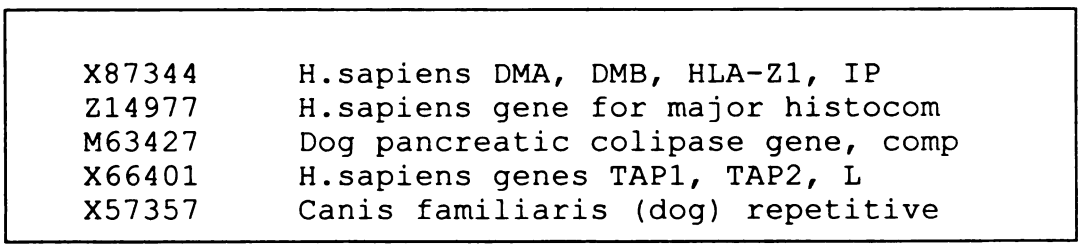
Figure 5. Canine LMP2 Sequence. The canine LMP-2 sequence obtained by LPCR with α-P^{33} radiolabeling.

To verify the findings, the NIH web site was used. A Blast search was completed using the same Canine LMP-2 sequence. Blast search is not as restrictive as the GCG FastA program. This would allow for detection of any other less limited results. The results obtained from the Blast search, however, were comparable to that of the GCG FastA findings. The finding led us to the same conclusions, that a portion of the canine LMP-2 had been sequenced. These results can be noted in Figure 7.

Another comparison study was completed with the assistance of GCG using the Bestfit program. The Bestfit program of GCG aligns and compares sequences based on the similarity between the specified segments. Once the comparisons are completed, the alignments are able to be visualized. The percent similarity, sequence gaps and sequence contrasts with the two given sequences can be noted. The LMP-2 Canine to LMP-2 Human sequence comparison resulted in a percent similarity of 73.997. The results can be noted in Figure 8.

It was observed that the Human and Canine sequence had dissimilarity at the canine sequence, bases 639 to 893. This promoted an inquiry into the known human LMP-2 sequence. It was found that the human sequence contained an ALU repetitive unit within the LMP-2 region of PCR amplified product. Because of the discoveries of the original Bestfit

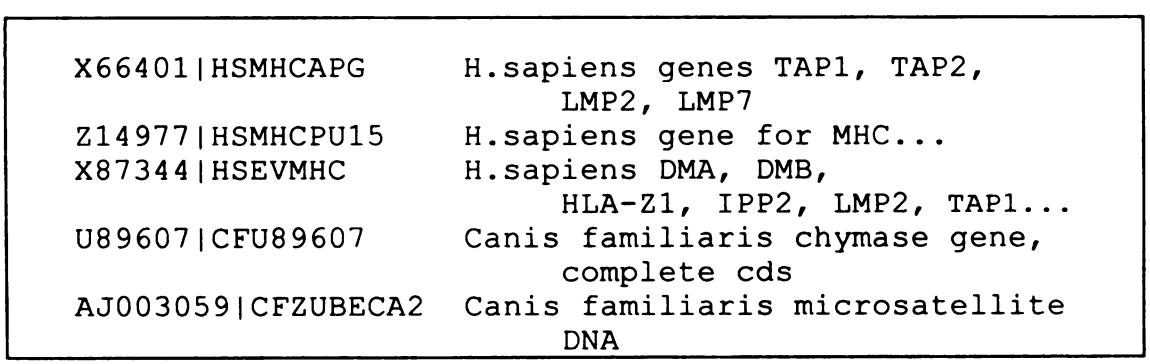
Figure 6



X87344	H.sapiens DMA, DMB, HLA-Z1, IP
Z14977	H.sapiens gene for major histocom
M63427	Dog pancreatic colipase gene, comp
X66401	H.sapiens genes TAP1, TAP2, L
X57357	Canis familiaris (dog) repetitive

Figure 6. Canine LMP-2 FastA Results. Above are the GCG FastA results for canine LMP-2. It should be noted, the highest results of homology were detected to the human MHC class II region, more specifically the human LMP-2 and to the canine SINE (repetitive unit).

Figure 7

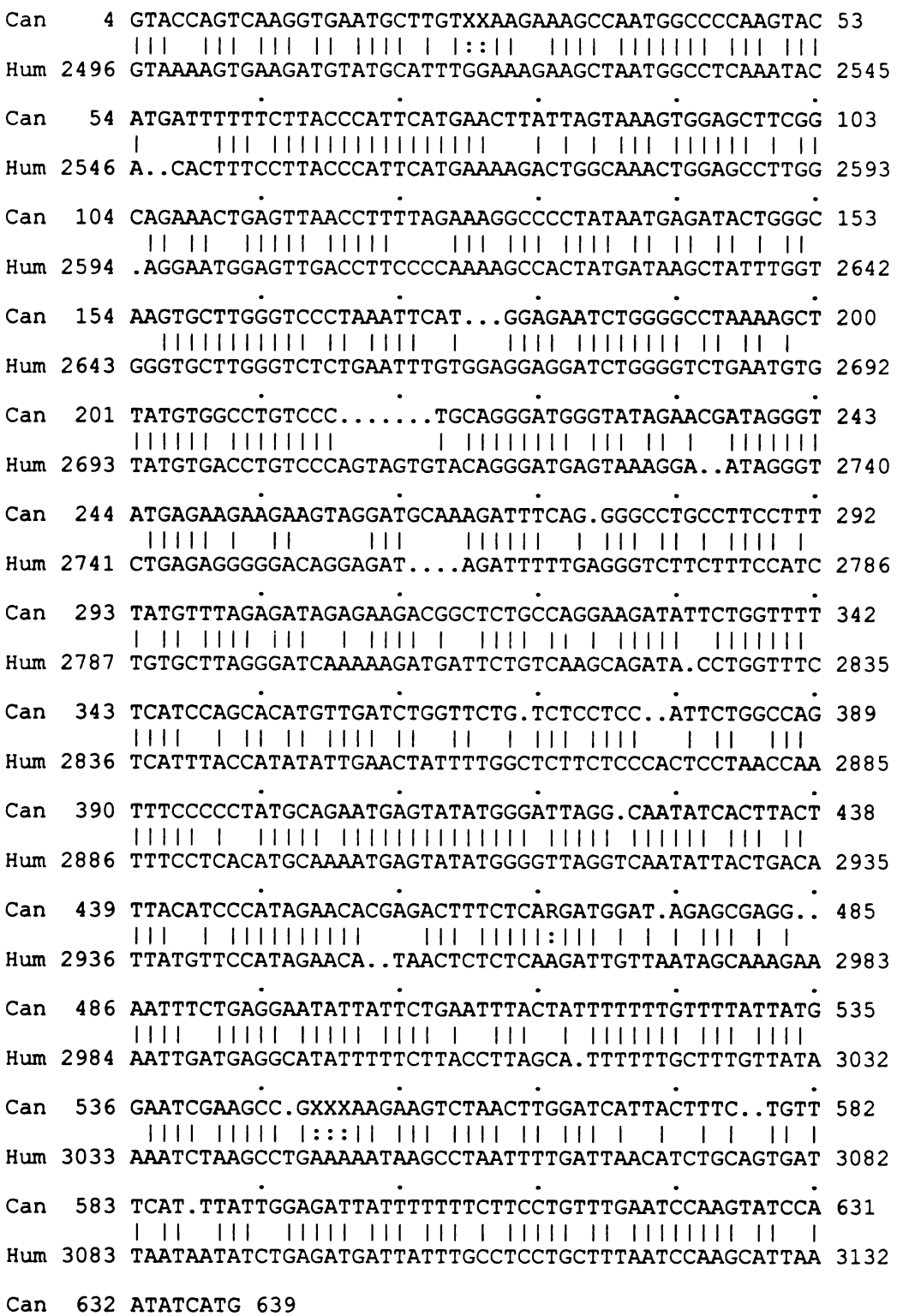


X66401|HSMHCAPG	H.sapiens genes TAP1, TAP2, LMP2, LMP7
Z14977|HSMHCPU15	H.sapiens gene for MHC...
X87344|HSEVMHC	H.sapiens DMA, DMB, HLA-Z1, IPP2, LMP2, TAP1...
U89607|CFU89607	Canis familiaris chymase gene, complete cds
AJ003059|CFZUBECA2	Canis familiaris microsatellite DNA

Figure 7. Canine LMP-2 NIH Blast Results. Above are the NIH Blast result of the Canine LMP-2. AS with the FastA results (Figure 6), the highest results of homology were found to be to the human MHC class II region, more specifically the human LMP-2, and to the canine SINE (repetitive unit).

Figure 8. Bestfit Canine LMP-2 to Human LMP-2 Part I. Bestfit GCG program was utilized in the alignment of the complied canine LMP-2 to the human LMP-2 sequence (Z14977). The percent homology (from base pair 1 to 639 of the canine LMP-2) was 73.997. Primers used were from exon 2 to exon 3 of the human LMP-2 gene.

Figure 8


Can 4 GTACCAGTCAAGGTGAATGCTTGTXXAAGAAAGCCAATGGCCCCAAGTAC 53
Hum 2496 GTAAAAGTGAAGATGTATGCATTTGGAAAGAAGCTAATGGCCTCAAATAC 2545
Can 54 ATGATTTTTTCTTACCCATTCATGAACTTATTAGTAAAGTGGAGCTTCGG 103
Hum 2546 A..CACTTTCCTTACCCATTCATGAAAAGACTGGCAAACTGGAGCCTTGG 2593
Can 104 CAGAAACTGAGTTAACCTTTTAGAAAGGCCCCTATAATGAGATACTGGGC 153
Hum 2594 .AGGAATGGAGTTGACCTTCCCCAAAAGCCACTATGATAAGCTATTTGGT 2642
Can 154 AAGTGCTTGGGTCCCTAAATTCAT...GGAGAATCTGGGGCCTAAAAGCT 200
Hum 2643 GGGTGCTTGGGTCTCTGAATTTGTGGAGGAGGATCTGGGGTCTGAATGTG 2692
Can 201 TATGTGGCCTGTCCC.....TGCAGGGATGGGTATAGAACGATAGGGT 243
Hum 2693 TATGTGACCTGTCCCAGTAGTGTACAGGGATGAGTAAAGGA..ATAGGGT 2740
Can 244 ATGAGAAGAAGAAGTAGGATGCAAAGATTTCAG.GGGCCTGCCTTCCTTT 292
Hum 2741 CTGAGAGGGGGACAGGAGAT....AGATTTTTGAGGGTCTTCTTTCCATC 2786
Can 293 TATGTTTAGAGATAGAGAAGACGGCTCTGCCAGGAAGATATTCTGGTTTT 342
Hum 2787 TGTGCTTAGGGATCAAAAAGATGATTCTGTCAAGCAGATA.CCTGGTTTC 2835
Can 343 TCATCCAGCACATGTTGATCTGGTTCTG.TCTCCTCC..ATTCTGGCCAG 389
Hum 2836 TCATTTACCATATATTGAACTATTTTGGCTCTTCTCCCACTCCTAACCAA 2885
Can 390 TTTCCCCCTATGCAGAATGAGTATATGGGATTAGG.CAATATCACTTACT 438
Hum 2886 TTTCCTCACATGCAAAATGAGTATATGGGGTTAGGTCAATATTACTGACA 2935
Can 439 TTACATCCCATAGAACACGAGACTTTCTCARGATGGAT.AGAGCGAGG.. 485
Hum 2936 TTATGTTCCATAGAACA..TAACTCTCTCAAGATTGTTAATAGCAAAGAA 2983
Can 486 AATTTCTGAGGAATATTATTCTGAATTTACTATTTTTTTGTTTTATTATG 535
Hum 2984 AATTGATGAGGCATATTTTTCTTACCTTAGCA.TTTTTTGCTTTGTTATA 3032
Can 536 GAATCGAAGCC.GXXXAAGAAGTCTAACTTGGATCATTACTTTC..TGTT 582
Hum 3033 AAATCTAAGCCTGAAAAATAAGCCTAATTTTGATTAACATCTGCAGTGAT 3082
Can 583 TCAT.TTATTGGAGATTATTTTTTTCTTCCTGTTTGAATCCAAGTATCCA 631
Hum 3083 TAATAATATCTGAGATGATTATTTGCCTCCTGCTTTAATCCAAGCATTAA 3132
Can 632 ATATCATG 639
Hum 3133 ACTTCATG 3140

and the ALU repeat, a second Bestfit was completed using the canine LMP-2 sequence from the three prime end. The second Bestfit found a percent similarity of 75.893 of canine to human comparison. The results can be seen in Figure 9.

The region of nonidentity/nonsimilarity for the canine LMP-2 sequence was found to the match the intronic canine repetitive unit. Another Bestfit program was completed using the canine LMP-2 to that of the canine intron repetitive unit. The results of the Bestfit of the canine LMP-2 to the Canine intron repeat (SINE) can be seen in Figure 10.

Figure 9

```
                 .         .         .         .         .
Can  893 AAAAAATTCTGGCGTTGCTAATGTTGCTGCCTGCTGACTCCTTCCCTGCA 942
         ||||||  |  | |||  | |  ||| || ||| ||||||| |||   ||
Hum 3498 AAAAAAAACCTGAGTT..TTAACTTGGTGACTGTTGACTCCCTCCTGACA 3545
                 .         .         .         .         .
Can  943 GAGAAGCAGTGGGTAAACCGAGCATTTXACAAGCTGTCCCCACTACACCA 992
         | || || || ||| ||||||| |||:|||||||||||||| || ||| |
Hum 3546 GCGAGGCGGT.GGTGAACCGAGTGTTTGACAAGCTGTCCCCGCTGCACGA 3594
                 .
Can  993 ACACATGTACTGTGC 1007
          | ||| ||||||||
Hum 3595 GCGCATCTACTGTGC 3609
```

Figure 9. Bestfit Canine LMP-2 to Human LMP-2 Part II. Bestfit GCG program was utilized in the alignment of the complied canine LMP-2 to the human LMP-2 sequence (Z14977). The percent homology (from base pair 893 to 1007 of the canine LMP-2) was 75.893. Primers used were from exon 2 to exon 3 of the human LMP-2 gene.

Figure 10

```
                .         .         .         .         .
LMP   656 GGTGGCTCGGCGGTTTGGCACCTGCCTTTGGCCCAGGGCGCGATCCTGGA 705
          |||||| | | ||  |||| | |||  |  | |||| ||||||||||||
SINE    6 GGTGGC.CAGTGGAATGGCGCTTGC..TCCGGCCAGAGCGCGATCCTGGA 52
                .         .         .         .         .
LMP   706 GTCCTGGGATCGAGTCCCGCCTTGGGCTCCCG..GCATGGAGCCTGCTTC 753
            ||  ||||||||||||   | |||||||||  ||||||||||||||||
SINE   53 AACCCAGGATCGAGTCCCATGTCGGGCTCCCGGTGCATGGAGCCTGCTTC 102
                .         .         .         .         .
LMP   754 TCCCTCCTCCTGTGTCTCTGCCTCTCTCTCTCTCTGTCTATGT..CTATC 801
          ||||||  |||||||||| |||||||||||||||||| ||| |||  |||||
SINE  103 TCCCTCTGCCTGTGTCTGTGCCTCTCTCTCTCTCTCTCTCTGTGACTATC 152
                .         .
LMP   802 GTAAATAAATGAGTAAATTTT 822
           |||| ||||   || || ||
SINE  153 ATAAAAAAATATATATATATT 173
```

Figure 10. Bestfit of Canine LMP-2 and Canine SINE. Bestfit GCG program was utilized in the alignment of the complied canine LMP-2 to the canine SINE sequence (U17996). The percent homology was 81.707. It should be noted that the location of the canine SINE is comparable to the location of the human ALU repeat seen in the human LMP-2 gene (Z14977).

DISCUSSION

The optimization of all the sets of primer, proved to be troublesome. The reason that only the LMP-2 sequence obtained was due to the fact that the LMP-2 primers seems to be less problematic than the other sets of primers. With the size of the some of the PCR fragments being greater than 1.5kb, time became an important issue. Sequencing of the LMP-2 fragment (1008bp) maintained a time frame of just over 6 months, that does not include the time required for optimization the PCR cycling conditions. For this reason, a decision was made to concentrate on obtaining the LMP-2 sequence exclusively.

Sequencing Canine LMP-2

The comparison of the canine LMP-2 determined there is 74% identity between the canine (base 4-639 LMP-2) and human LMP-2 (base 2496-3609) homologue. A second region of identity of the canine and human LMP-2 resulted in 76% homology. The primers for the human genomic DNA were from the second exon to that of the third exon for Human LMP-2 (see material methods of primer information).

One of the objectives for this study was to sequence the LMP-2 Homologue in canines. The homology was found to be 75 %. However, during the comparison studies, a region of nonidentity between the human and canine LMP-2 was

encountered. The nonidentity was found to be located at the human ALU repeat. This prompted us investigate further.

It was noted that the 173 bases of the nonidentity canine sequence corresponded to a previously reported canine SINE (U17996). The canine LMP-2 was found to have 81% homology to the previously reported canine SINEs (see Figure 10).

The interspersed repetitive sequences can be divided into two groups, SINEs and LINEs (Nomura). SINEs are short interspersed repetitive elements, LINEs, on the other hand, are long interspersed repetitive elements (Thomsen). The major difference between SINEs and LINEs, besides the length of the repeat, is that LINEs are conserved over a wide range of animal species whereas, SINEs are conserved only in the animals or related species (Nomura). Our discussion will concentrate specifically with SINEs since our research is most clearly related to that repetitive unit

SINEs are found thought the mammalian genome as well as being detected in plants (Okada). Originally SINEs were thought to have evolved through DNA transposition, unequal crossing over or gene conversion (Kass). Current evidence suggests, however, that SINEs are results of retroposons being integrated into the genome (Okada).

SINEs may vary from 70 to 300 base pairs in length (Alexander) and can be present in more than 10^5 copies per

genome (Okada). Individual SINEs are conserved within animals of related species (Nomura) and can be utilized to determine the order of speciation (Murata).

Previous studies completed by Manjula et al determined that a canine SINE can be detected at least once every 5-8.3 kbp, constituting 360,000-600,000 copies in the canine genome. Their studies also located repeats approximately 150 base pair in length, including a 3' $(CT)_n$ tandom and a poly A region. Both the $(CT)_n$ and the poly A tail can be noted in the canine LMP-2 SINE that we have reported.

Canine SINEs have been found to be closely related to the gray wolf-*Canis lupus*, side striped jakal-*Canis adustus*, black jackal-*Canis mesomelas* and distantly related to the arctic fox-*Alopex lagopus* (Das). These finding support the theory of speciation of SINEs to relate animals.

LMP-2 SINEs

It has been shown that both the canine and the human contain the given species SINE within the LMP-2 gene. While it has been demonstrated that other species, such as zebrafish, have the LMP-2 gene, the only other species that has known genomic sequence LMP-2 is the mouse. Using the GCG program of Bestfit, the mouse LMP-2 and mouse repetitive unit were aligned to detect if there was indeed an intron repeat within LMP-2. The mouse specific SINE again can be

detected within the LMP-2 gene. This can be noted in Figure 11.

As more sequencing projects detect the LMP-2 gene, it is possible to complete other species comparison of the LMP-2 SINE positioning. The function of this LMP-2 SINE is unknown at this time. Again, with further research, the function of the SINE may be defined.

Figure 11

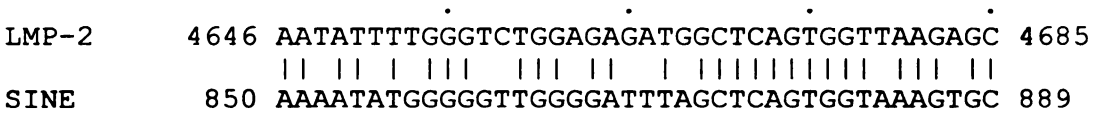


Figure 11. Bestfit of Mouse LMP-2 to Mouse Repetitive Unit. Mouse LMP-2 (D14566) to Mouse repetitive DNA (Z49228. With this alignment the detection of the mouse SINE can be noted in the LMP-2 gene of the mouse.

RECOMMENDATIONS

Future TAP and LMP Studies

I would recommend for future studies sequencing the TAP and LMP genes using cDNA. This would give the opportunity to determine the entire gene sequence as well as the opportunity for the protein sequence to be determined. Once the cDNA sequence is determined, further investigations can be completed concerning the intronic studies, such as repetitive unit locations.

I would also recommend that one sequence the chaperons: calexnin, tapasin and calreticulin. A comparison study would be interesting to determine just how conserved the endogenous pathway across species.

If one would like to undertake the sequencing project(s) as suggested above, I would suggest exploring the use of automated sequencing. It seems to be the most productive route. When comparing the time and cost issue of manual sequencing to that of automated sequencing, the issue seems mute. One can have automated sequencing results as quickly as 24 hours, once the PCR product is purified. Walking 2 kb can be completed as quickly as a week. Another benefit of automated sequencing is that there is less error in determining the actual sequence. Human error is possible when determining sequence from a manual sequencing film.

With automated sequencing. the sequence is determined by the computer program. The program usually will derive a sequence that is usually error free to 500 bases. Granted problems can be encountered with automated sequencing, but sequencing facilities are staff with professionals, which can be of assistance when trouble-shooting is needed.

Future Studies of SINEs

Another project to consider is that of SINEs and MHC linkage. It is known that mouse, human and canine all maintain a species specific repeat within an intron of LMP-2. It is possible that other animals also possess this same feature. It would be interesting fact to determine.

REFERENCES

REFERENCES

Alexander, L.J., Rohrer, R.T. and Beattie, C.W. 1995. Porcine SINE-associated microsatellite markers: evidence for new artiodactyl SINEs. Mammalian Genome. 6,464-468.

Anderson, L. and Davies, C.J. The major histocompatibility complex. In B.M. Goddeeris and W.I. Morrison (eds): Cell Mediated Immunity in Ruminants, CRC press, in press.

Belich, M.P., Glynne, R.J. et al. 1994. Proteasome components with reciprocal expression to that of the MHC-encoded LMP proteins. Current Biology. 4(9), 769-776.

Blanchet, O., Bourge, J.F. et al. 1991. DNA binding of regulatory factors interacting with MHC-class-I gene enhancer correlates with the MHC-class-I transcriptional level in class-I defective cell lines. International Journal of Cancer. 6, 138-145.

Bull, R.W., Vriesendorp, et al. 1987. Joint report of the third international workshop on canine immunogenetics. II. Analysis of the serologicel typing of cells. Transplantation. 43, 154-161.

Burnett, R.C. and Geraghty, D.E. 1995. Structure and expression of a divergent canine class I gene. The Journal of Immunology. 155, 4278-4285.

Burnett. R.C., DeRose, S.A. et al. 1997. Molecular analysis of six dog leukocyte antigen class I sequences including three complete genes, two truncated genes and one full-length processed gene. Tissue Antigens. 49, 484-495.

Carreno, B.M., Solheim, J.C. et al. 1995. TAP associates with a unique class I conformation, whereas calnexin associates with multiple class I forms in mouse and man. Journal of Immunology. 155, 4726-4733.

Chabanne, L., Kaplanski, C. et al. 1995. Screening of 78 monoclonal antibodies directed against human leukocyte antigens for cross-reactivity with surface markers on canine lymphocytes. Tissue Antigens. 43, 202-205.

Chouchane, L. and Kindt, T.J. 1992. Mapping of the rabbit MHC reveals that class I genes are adjacent to the DR subregion and defines an insertion/deletion-related polymorphism in the class II region. Journal of Immunology. 149, 1216-1222.

Cibotti, R., Punt, J.A. et al. 1997. Surface molecules that drive T cell development in vitro in the absence of thymic epithelium and in the absence of lineage-specific signals. Immunity. 6, 245-255.

Colonna, M., Bresnhan, M. et al. 1992. Allelic variants of the human putative peptide transporter involved in antigen processing. Immunology. 89, 3932-3936.

Cook, D.N., Beck, M.A. et al. 1995. Requirement of MIP-1α for an inflammatory response to viral infection. Science. 269, 1583-1587.

Cross, S., Tonks, S. et al. 1991. Novel detection of restriction fragment length polymorphisms in the human major histocompatibility complex. Immunogenetics. 34, 376-384.

Das, M., Lee Chu, L. et al. 1998. Characterization of an abundant short interspersed nuclear element (SINE) present in *Canis familiaris*. Mammalian Genome. 9, 64-69.

Davies, J.L., Kawaguchi, Y. et al. 1994. A genome-wide search for human type I diabetes susceptibility genes. Nature. 371, 130-136.

Dean, R.R., Kildsuff, T.S. et al. 1989. Nacrolepsy without unique MHC class II antigen association: studies in the canine model. Human Immunology. 25, 27-35.

Deeg, H.J., Raff, R.F. et al. 1986. Joint report of the third international workshop on canine immunogenetics. I. Analysis of homozygous typing cells. Transplantation. 41, 111-117.

Fearon, D.T. and Locksley, R.M. 1996. The instructive role of innate immunity the acquired immune response. Science. 272(5258), 50-53.

Fraser, J.D., Straus, D. and Weiss, A. 1992. Signal transduction events leading to T cell lymphokine gene expression. Immunology Today. 14(357), .

Fremont, D.H., Crawford, F. et al. 1998. Crystal Structure of mouse H2-M. Immunity. 9, 385-393.

Fruh, K., Yang, Y. et al. 1992. Alternative exon usage and processing of the major histocompatibility complex-encoded proteasome subunits. The Journal of Biological Chemistry. 267, 22131-22140.

Germain, R.N. 1994. MHC dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell. 76, 287-299.

Gorman, S.D., Frewin, S.P. et al. 1994. Isolation and expression of cDNA encoding the canine CD4 and CD8α antigens. Tissue Antigens. 43, 184-188.

Hecker, K.H. and Roux, K.H. 1996. High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. Biotechniques. 20(3), 478-85.

Hidenori, T. 1996. A population genetic study of the evolution of SINEs II. Sequence evolution under the master copy model. Genetics. 143, 1033-1042.

Hong, S.C., Chelouche, A. et al. 1992. An MHC interaction site maps to the amino-terminal half of the T cell receptor α chain variable domain. Cell. 69, 999-1009.

Howard, J.C. 1993. Restrictions on the use of antigenic peptides by the immune system. Proc Natl Acad Sci. USA. 90, 3777-3779.

Huang, J.H., Getty, R.R. et al. 1994. Protein transfer of preformed MHC-peptide complexes sensitizes target cells to T cells cytolysis. Immunity. 1, 607-613.

Innis, M.A., Mayambo, K.B. et al. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. Proc. Natl. Acad. Sci. USA. 85, 9436-9440.

Janeway,C.A. 1993. How the Immune system recognizes invaders. Scientific American. Sept, 74-79.

Kandil, E., Namikawa, C. et al. 1996. Isolation of low molecular mass polypeptide complementary DNA clones from primitive vertebrates. Journal of Immunology. 156, 4245-4253.

Kass, D.H., Batzer, M.A. and Deininger, P.L. 1995. Molecular and Cellular Biology. 15 (1), 19-25.

Kaufman, J., Volk, H. and Wallny, H.J. 1993. A "minimal essential Mhc" and an "unrecognized Mhc"; two extremes in selection for polymorphism. Immunological Reviews. 143, 63-88.

Kelly, A., Powis, S.H. et al. 1992. Assembly and function of the two ABC transporter proteins encoded in the human major histocompatibility complex. Nature. 355, 641-644.

Klein, J., Takahata, N. and Ayala, F. 1993. MHC polymorphism and human origins. Scientific American. Dec., 78-85.

Kroemer, G.B., Behar, G.H. et al. 1990. Molecular genetics of the chicken MHC: current status and evolutionary aspects. Immunological Reviews. 113, 119-145.

Laud, P.R., Loflin, P.T. et al. 1996. Transporter associated with antigen-processing-1 (TAP1) alleles in *Gorilla gorilla*: diversification of the locus postspeciation. Human Immunology. 50, 91-102.

Lobigs, M., Rothenfluh, H.S. et al. 1995. Polymorphic peptide transporter in MHC class I monomorphic Syrian hamster. Immunogenetics. 42, 398-407.

Lupas, A., Koster, A. et al. 1993. Structural Features of 26s and 20s proteasomes. Enzyme Protein. 47, 252-273.

Maniatis, T., Fritsch, E.F., and Sambrook, R. 1989. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, New York.

Marshall, B., Leelayuwat, C. et al. 1993. New major histocompatibility complex genes. Human Immunology. 38, 24-29.

Middleton, D., Megaw, G. et al. 1994. TAP1 and TAP2 polymorphism in multiple sclerosis patients. Human Immunology. 40, 131-134.

Minnick, M.F., Stillwell, L.C. et al. 1992. A highly repetitive DNA sequence possibly unique to canids. Gene. 110, 235-238.

Monaco, J.J. 1992. A molecular model of MHC class-I-restricted antigen processing. Immunology Today. 13(5), 174-179.

Monaco, J.J. 1995. Pathways for the processing and presentation of antigens to T cells. Journal of Leukocyte Biology. 57, 543-547.

Moins-Teisserence, H., Semana, G. et al. 1995. TAP 2 gene polymorphism contributes to genetic susceptibility to multiple sclerosis. Human Immunology. 42, 195-202.

Neefjes, J.J., Momburg, F. and Hammerling, G.J. 1993. Selective an ATP-depedent translocation of peptides by the MHC-encoded transporter. Science. 261, 769-771.

Neefjes, J.J. and Ploegh, H.L. 1992. Intracellular transport of MHC class II molecules. Immunology Today. 13(5), 179-183.

Neefjes, J.J. 1996. Antigen presentation by MHC class I and II molecules. Immunobiology. 195, 436-460.

Newell, W.R., Trowsdale, J. and Beck, S. 1996. MHCDB-database of the human MHC (release 2). Immunogenetics. 45(1), 6-8.

Nomura, O., Lin, Z.H. et al. 1998. A SINE species from hippopotamus and its distribution among animal species. Mammalian Genome. 9, 550-555.

Okada, N., Mitsuhiro, H. et al. 1997. SINEs and LINEs share common 3' sequences: a review. Gene. 205, 229-243.

Ortmann, B., Copeman, J. et al. A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. Science. 277, 1305-1308.

Ortmann, B., Androlewicz, M.J. et al. 1994. MHC class 1 β_2-microglobulin complexes associate with TAP transporter before peptide binding. Nature. 368, 864-867.

Ploski, R., Flato, B. et al. 1995. Assoication to HLA-DRB1 *08, HLA-DPB1*0301 and homozygosity for an HLA-linked proteasome gene in Juvenile ankylosing spondylitis. Human Immunology. 44, 88-96.

Ploski, R., Undlien, D. et al. 1994. Polymorphism of human major histocompatibility complex-encoded transporter associated with antigen processing (TAP) genes and susceptibility to juvenile rheumatoid arthritis. Human Immunology. 39, 54-60.

Polvi, A., Garden, O.A. et al. 1997. Canine major histocompatibility complex genes DQA and DBQ in irish setter dogs. Tissue Antigens. 49, 236-43.

Powis. S.H., Mockridge, I. et al. 1992. Polymorphism in a second ABC transporter gene located within the class II region of the human major histocompatibility complex. Proc. Natl. Acad. Sci. USA. 89, 1463-1467.

Powis, S.H., Tonks, S. et al. 1993. Alleles and haplotypes of the MHC encoded ABC transports TAP1 and TAP2. Immunogenetics. 37, 373-380.

Powis, S.H., Cooper, M.D. et al. 1994. Major histocompatibility haplotypes associated with immunoglobulin-A deficiency and common variable immunodeficiency: Analysis of the peptide transported genes TAP1 and TAP2. Tissue Antigens. 43, 261-265.

Powis, S.H. and Geraghty, D. 1995. What is the MHC?. Immunology Today. 16 (10), 466-468.

Rast, J.P., Anderson, M.K. et al. α, β, γ, and δ T cell antigen receptor genes arose early in vertebrate phylogeny. Immunity. 6, 1-11.

Robertson, M. 1991. Proteasomes in the pathway. Nature. 353, 300-301.

Rudensky, A.Y., Preston-Hurlburt, P. et al. 1991. Seuquence analysis of peptides bound to MHC class II molecules. Nature. 353, 622-627.

Rudensky, A.Y., Macic, M. et al. 1994. Intracellur assembly and transport of endogenous peptide-MHC class II complexes. Immunity. 1, 585-594.

Sadasivan, B., Lehner, P.J. et al. 1996. Immunity. Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. Immunity. 5, 103-114.

Sarmiento, U.M. and Rainer, S. 1990. Nucleotide sequence of dog DRB cDNA clone. Immunogenetics. 31, 396-399.

Sarmiento, U.M, Sarmiento, J.I. et al. 1990. Allelic variation in the DR subregion of the canine major histocompatibility complex. Immunogenetics. 32, 13-19.

Schmitt, T.H., Dunbar, B.N. et al. 1997. A case of primary immunodeficiency due to a defect of the major histocompatibility gene complex class I processing and presentation pathway. Immunology. 57(1-3), 183-7.

Suh, W.K., Cohen-Doyle, M.F. et al. 1994. Interaction of MHC class I molecules with the transporter associated with antigen processing. Science. 64, 1322-1326.

Suh, W.K., Mitchell, E.K. et al. 1996. MHC class I molecules form ternary complexes with calnexin and TAP and undergo peptide-regulated interaction with TAP via their extracellular domains. J. Exp. Med. 184, 337-348.

Solheim, J.C., Carreno, B.M. and Hansen, T.H. 1997. Are transporter associated with antigen processing (TAP) and tapasin class I MHC chaperones? Journal of Immunology. 158, 541-543.

Song, R. and Harding, C.V. 1996. Roles of proteasomes, (LMP) transporters for antigen presentation (TAP) and $\beta_2$microglobulin in the processing of bacterial or particulate antigens via an alternate class I MHC processing pathway. Journal of Immunology. 156, 4182-4190.

Spies, T., Bresnahan, M. et al. 1990. A gene in the human major histocompatibility complex class II region controlling the class I antigen presentation. Nature. 348, 744-747.

Swarbrick, P.A., Schwaiger, F.W. et al. 1995. Cloning and sequencing of expressed DRB genes of the red deer (*Cervus elaphus*) MHC. Immunogenetics. 42, 1-9.

Tighe, M.R., Hall, M.A. et al. 1994. Associations between alleles of the major histocompatibility complex-encoded ABC transporter gene TAP2, HLA class II alleles, and celiac disease susceptibility. Human Immunology. 39, 9-16.

Thomsen, P.D. and Miller, J.R. 1996. Pig genome analysis: differential distribution of SINE and LINE sequences is less pronounced than in the human and mouse genomes. Mammalian Genome. 7,42-46.

Thorsby, Erik. 1997. Invited anniversary review: HLA associated disease. Human Immunology. 53, 1-11.

Trowsdale, J. 1995. "Both man and bird and beast": comparative organization of MHC genes. Immunogenetics. 41, 1-17.

Trucco, Massimo. Personal Communication. DNA Salt Extraction. University of Pittsburgh Medical Center. December 1991.

Udaka, K., Tsomides, T.J. et al. 1992. A naturally occurring peptide recognized by alloreactive CD8+ cytotoxic T lymphocytes in association with a class I MHC protein. Cell. 69, 989-998.

VanEndert, P.M., Liblau, R.S. et al. 1994. Major histocompatibility complex-encoded antigen processing gene polymorphism in IDDM. Diabetes. 43, 110-117.

Wange, R.L. and Samelson, L.E. 1996. Complex complexes: signaling at the TCR. Immunity. 5, 197-205.

Wanger, J.L., Burnett, S.A., et al. 1996. Molecular analysis and polymorphism of the DLA-DQA gene. Tissue Antigens. 48, 199-204.

Wanger, J.L., Burnett, S.A. et al. 1996. Molecular analysis and polymorphism of the DLA-DQA gene. Tissue Antigens. 48, 554-561.

Watson, C.J., Cobbold, S.P. et al. 1994. Immunosuppression of canine renal allograft recipients by CD4 and CD8 monoclonal antibodies. Tissue Antigens. 43, 155-162.

White, L.C., Wright, K.L. et al. 1996. Regulation of LMP2 and TAP1 genes by IRF-1 explains the paucity of CD8+ T cells in IRF-1 -/- mice. Immunity. 5, 365-376.

Wolpert, E.Z., Peterson, M. et al. 1997. Generation of CD8+ T cells specific for transporter associated with antigen processing deficient cells. Proc. Natl. Acad. Sci. USA. 94. 11496-11501.

Wright, H., Redmond, J. et al. 1995. The nucleotide sequence of the sheep MHC II DNA gene. Immunogenetics. 41, 131-133.