University of Alberta

Allelic diversity of antigen processing genes in wild mallards

by

Kristina Petkau

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Physiology, Cell and Developmental Biology

Biological Sciences

Fall 2012

Edmonton, Alberta

Abstract

The recognition of virus infected cells by cytotoxic T lymphocytes is dependent on antigen presentation by MHC class I on the cell surface. Antigen presentation is critically dependent on the transporter associated with antigen processing (TAP), which facilitates the transport of peptides across the endoplasmic reticulum (ER) membrane. TAP specificity could affect the repertoire of peptides loaded on MHC class I. Here we analyzed the diversity of antigen presentation genes of 12 mallards, and the bidirectional promoter between TAP1 and TAP2. Alleles are from 94 - 100% similar with variation at 32 positions in TAP1 and 81 in TAP2. Non-conservative amino acid substitutions were observed in both TAPs, which align with peptide binding regions of the mammalian TAP. Polymorphism in functional regions of TAP could potentially restrict antigen presentation in ducks. This could contribute to the poor immune response to vaccines and the perpetuation of influenza virus in this host.

Acknowledgements

I would like to thank my supervisor Dr. Kathy Magor for giving me the opportunity to work in her lab and for all her support and guidance throughout my degree. Thank you to all the members from the K.Magor lab, past and present, for all the help. Thank you to Dr. Domingo Navarro for sharing his knowledge of Southern blot analysis and for all his assistance during experiments. I would like to thank Hillary Vanderven for her endless help with various techniques and for being a great friend. I would like to acknowledge Alysson Blaine and Ximena Fleming for their assistance in the lab, especially for including some of my many sequences on their plates. I would like to acknowledge Chelsea Davidson, who started the work on TAP2, and Sarah Zhao, who worked with me for a summer on this project.

I would also like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this work.

I would like to thank my parents, Boris and Tatjana, my sister, Darja, with her family, my brother, Markus, and my partner, Jon. Thank you for always supporting and encouraging me over the last couple of years. I could not have done this without you. Finally I would like to say thank you to all my friends I have made over the last years. You have made grad school a lot more fun and a great experience.

Table of Contents

List of Tables

List of Figures

Abbreviations

ABC	ATP binding cassette
AI	avian influenza
ATP	adenosine-5′-triphosphate
bp	basepairs
β_2m	beta-2 microglobulin
cDNA	complementary DNA
CDS	coding sequence
CMV	cytomegalovirus
CTL	cytotoxic T lymphocyte
dpi	days post infection
ER	endoplasmic reticulum
ERAP	endoplasmic reticulum aminopeptidase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HA	hemagglutinin
HLA	human leukocyte antigen
Ig	immunoglobulin
IFN-γ	interferon gamma
kb	kilobase
LMP	low molecular weight protein
MHC	major histocompatibility complex
NA	neuraminidase
NBD	nucleotide binding domain

nt	nucleotide
PBR	peptide binding region
PBS	phosphate buffered saline
PCR	polymerase chain reaction
qPCR	quantitative PCR
RT-PCR	reverse transcription PCR
TAP	transporter associated with antigen processing
Tapasin	TAP associated glycoprotein
TESS	transcription element search system
TM	transmembrane
TMD	transmembrane domain
UTR	untranslated region

1 INTRODUCTION

1.1 A brief introduction to TAP

In vertebrates, adaptive immunity has evolved as an important defense mechanism against invading pathogens. Antigen presentation plays a prominent role in this defense mechanism. Major histocompatibility (MHC) class I molecules present endogenous antigens to cytotoxic T lymphocytes (CTLs) on the cell surface. CTLs screen the cell surface for foreign peptides and induce the killing of infected cells (Townsend and Bodmer, 1989; Bjorkman and Parham, 1990). The MHC class I and the peptide complex is assembled in the endoplasmic reticulum (ER) and is stable once a peptide binds firmly to the class I molecule. Antigens presented to class I molecules are derived from cytosolic degradation by the proteasome and are then transported across the ER membrane into the lumen (Rock et al., 1994; Grant et al., 1995). Even though the MHC was one of the most intensely studied regions of the genome the transport process was poorly understood. Studies on mutant cell lines that showed decreased MHC class I expression suggested that the antigen complex is dependent on a specific transporter, which translocates peptides from the cytoplasm into the ER. These cell lines had mutations in two genes and showed a deficiency in available peptides in the ER (Anderson et al., 1991; Bahram et al., 1991). Now known as the transporter associated with antigen processing (TAP), the peptide pump is encoded by two genes *TAP1* and *TAP2*, which are located in the class II region of the human MHC. TAP is a member of the superfamily of adenosine triphosphate

(ATP)-binding cassette (ABC) transporters, which are widespread among all living organisms. These transporters are membrane proteins that couple ATP-hydrolysis and the translocation of substrates across biological membranes (Higgins, 1992).

Due to its pivotal role in antigen presentation, studies on TAP have become of great interest. Defects in this transporter are associated with the cause of human disease such as the bare lymphocyte syndrome (BLS). A non-functional TAP is the result of a premature stop codon in TAP2, which leads to decreased surface expression of MHC class I (Gadola et al., 2000). The MHC is the most gene dense and polymorphic region in the genome. The most common form of variation is single nucleotide polymorphisms (SNPs), some of which are associated with immune deficiencies. To name one example, mutations within the *TAP1* gene in humans was associated with impaired MHC class I expression and therefore to recurrent infections and even high mortality (de la Salle et al., 1999).

1.2 The antigen presentation complex

1.2.1 A brief introduction to the ABC transporter

ABC transporters represent one of the largest and most diverse families of membrane proteins. They have been identified in organisms from all three domains of life. Since all cells or intracellular compartments are surrounded by lipid membranes, the transport of substrates across the membranes is essential for many cellular processes, such as nutrient uptake, protein secretion, signal transduction, drug and antibiotic resistance, bacterial pathogenesis and sporulation

and antigen presentation. ABC transporters function either as unidirectional importers or exporters that couple ATP-hydrolysis and the translocation of substrates across biological membranes. Importers pump substrates like sugars or ions against a concentration gradient into the cell, whereas exporters release substances like toxins from the cell. Conformational changes in the subunits of the transporter are induced by ATP hydrolysis and power the translocation of the substrate (as reviewed by Higgins, 1992).

An ABC transporter has four core domains: two transmembrane domains (TMDs) and two nucleotide binding domains (NBDs) (ATP-binding domains). In some transporters these four domains are expressed as individual polypeptides, but often they are fused together to larger multifunctional polypeptides. In a typical ABC transporter each membrane spanning domain spans the membrane six times. Both N- and C-termini are located in the cytoplasm and the ATP hydrolyzing domain stretches into the cytoplasm. One transporter has a total of six loops extending into the extracellular region and four cytoplasmic loops (Figure 1-1a). This was reviewed in detail by Higgins (1992).

The nucleotide binding domains are highly conserved among all ABC transporters and include characteristic features which allow the classification of transporters as members of the ABC superfamily (Higgins et al., 1986). The characteristic features are sequence motifs shared between all ABC transporters. These motifs represent the nucleotide binding sites and have been described as Walker A/P-loop, Walker B (Walker et al., 1982), ABC transporter signature motif (Hyde et al., 1990), Q-loop, D-loop and H-loop/switch region (Figure 1-1b).

The two transmembrane domains of an ABC transporter are highly hydrophobic and together they form the pore for passage of the substrates across the membrane. A standard transport complex consists of 12 transmembrane segments. Therefore some transporters function as homodimers whereas others form heterodimers to establish a functional transport system. The transmembrane domain is also involved in substrate binding (Kuchler et al., 1989; Greenberger et al., 1991). The substrate recognition site can span more than one transmembrane segment and extramembranous loop of each subunit. The transmembrane domains are very diverse among different ABC transporters and therefore define substrate specificity of each transporter.

There is evidence that mutations in the four core domains of ABC transporters can impair the transport of peptides. Mutations within the ATP binding sites, such as the invariant residues K and D in the Walker A and B motifs respectively, can inhibit the attachment of the nucleotides or ATP hydrolysis and therefore impair the activity of the transporter altogether (Shyamala et al., 1991). Variations in the transmembrane domains can alter substrate specificity (Gros et al., 1991). Over the last few years one particular member of the ABC family has drawn attention due to its critical role in antigen presentation. The transporter associated with antigen processing provides the peptides for binding to MHC class I, and is therefore an essential part of the immune response to intracellular pathogens.

1.2.2 The structure of TAP

The TAP transporter is encoded by two genes, *TAP1* and *TAP2*, located within the MHC (Spies et al., 1992). Each gene encodes for a protein of 81 kDa (TAP1) or 76 kDa (TAP2) (Abele and Tampe, 2004), which together form a heterodimer (Spies et al., 1992). The expression of both TAP1 and TAP2 is essential for the formation of a functional peptide pump and the entire loading complex (Kelly et al., 1992; Meyer et al., 1994). Each TAP protein contains of one transmembrane domain at the N-terminus and one nucleotide binding domain at the C-terminus. The nucleotide binding domains of TAP1 and TAP2 are highly homologous and contain the conserved sequence motifs of an ABC transporter, whereas the transmembrane domains share less sequence homology.

Members of TAP transporters differ from other ABC transporters in that they have additional domains. They posses the typical core translocation domain and accessory domains at the N-termini. The core domain is composed of a transmembrane domain at the N-terminus, which spans the membrane six times, and the cytoplasmic nucleotide binding domain at the C-terminus for each of the two subunits. The accessory domains are additional transmembrane domains at the N-termini of both TAP1 and TAP2 (Figure 1-2). The core domain is sufficient for targeting TAP to the ER, assembling the heterodimer, binding and translocating peptides. Therefore a fully functional TAP complex can exist without the extra N-terminal domains of TAP1 and TAP2 (Koch et al., 2004).

The accessory domains (TMD0) are typically four additional transmembrane segments at the N-termini of each TAP (Figure 1-2). The length

of the TMD0s varies between TAPs in different species and determines the location of the N-terminus to be either cytosolic or in the ER lumen. These accessory domains are involved in the association of tapasin with the TAP complex and the assembly of the peptide loading complex (Koch et al., 2004).

The association of a peptide with TAP is determined by the specificity of the transporter and is ATP-independent. In the human TAP, multiple regions have been identified as attachment sites for peptides located within the translocation core. TAP1 has two peptide binding sites. The first one is located in the cytosolic loop between transmembrane segments (TMs) four and five of the core TMD, and the second site overlaps TM six and stretches into the C-terminal tail. TAP2 has also two regions which crosslink to bound peptide, which are localized between TM four to five and TM six and part of the C-terminal tail (Nijenhuis et al., 1996). Due to difficulties in crystallizing membrane proteins, it is a challenge to fully understand and visualize the process of peptide binding and transport via TAP.

1.2.3 Functions of TAP

The transporter associated with antigen processing has been a major focus of research since its role in adaptive immunity, specifically in the translocation of the peptide and the presentation to class I molecules, became clear (Kelly et al., 1992). To date no other transporter is known to perform this function. TAP is essential for the formation of the peptide loading complex and defects in TAP can be deleterious. This can result in deficiencies in antigen presentation and is often the cause of disease.

TAP represents the translocation mechanisms of peptides across the ER membrane and therefore pre-selects the peptide repertoire. TAP preferentially and most effectively transports peptides with a length of 8-13 amino acids (aa), but can translocate peptides that are longer at a slower rate (Momburg et al., 1994a). Peptides are selected by TAP based on their three N-terminal and three C-terminal residues, and are limited by the binding specificity of the transporter (Heemels and Ploegh, 1994). Other results suggest that TAP also interacts with the MHC class I heterodimer, and is involved in the assembly of class I and peptide complexes (Lewis et al., 1996; Peace-Brewer et al., 1996).

Viruses have developed strategies to block antigen presentation at the level of TAP (Hewitt, 2003). For example, the early herpes simplex virus protein ICP47 binds to TAP in the cytoplasm and inhibits peptide association with the transporter (Fruh et al., 1995; Ahn et al., 1996). The US6 protein of human cytomegalovirus (CMV) associates with TAP resulting in deficient translocation of peptides into the ER (Lehner et al., 1997).

1.2.4 ATP dependence in antigen translocation

Substrates bind to the transporter without the consumption of energy, but their translocation is completely dependent on ATP hydrolysis (Neefjes et al., 1993). It has been suggested that binding and hydrolysis of ATP induces conformational changes in TAP, powering peptide transport. Intact nucleotide binding domains in both subunits are crucial for peptide translocation, and they are functionally distinct. It has been suggested that interaction with TAP2

controls peptide binding and release, while interaction with TAP1 leads to structural rearrangements of the TMs (van Endert et al., 2002)

ATP binding to TAP is facilitated by the Q-loop and induces dimerization of the two nucleotide binding domains, with the two ATPs sandwiched at the interface and bridging contact between the two domains (Figure 1-3) (Smith et al., 2002; Procko et al., 2006). Upon peptide binding, structural changes in the transmembrane domains are induced, which lead to the opening of the transporter towards the ER lumen and release of the peptide (Neumann and Tampe, 1999). ATPs bound to the nucleotide binding domains are hydrolyzed and phosphates are released, which weakens the interaction between the two nucleotide binding domains. Once the bond between the nucleotide binding domains weakens, the transporter closes towards the ER and the NBDs dissociate releasing ADPs. This step opens up the transporter for a new peptide to bind in order to repeat the process of translocation (Janas et al., 2003).

1.2.5 The structure of tapasin

The effective loading of peptides to MHC class I molecules and the release of this complex from the ER to the cell surface is dependent on the interaction with many chaperones. One such mediator is a member of the immunoglobulin (Ig) superfamily, the TAP associated glycoprotein (Tapasin). It is a 48 kDa protein found in the ER lumen, which is encoded by the tapasin gene, that maps to the MHC class II region of the genome (Ortmann et al., 1994). The precise location of tapasin is in the centromeric end of the human chromosome

six, and in the MHC for mice, rats and chickens (Grandea et al., 1998; Herberg et al., 1998; Frangoulis et al., 1999; Jacob et al., 2000).

Both human and chicken tapasin proteins show a general structure of 4 domains, one transmembrane at the N-terminus followed by three luminal domains (Herberg et al., 1998; Frangoulis et al., 1999). Based on the presence of cysteine residues and sequence similarities to other Ig family members, two of the three luminal domains were identified as Ig-like domains. These Ig-like domains are followed by a double-lysine motif at the C-terminus that is responsible for ER retention (Jackson et al., 1990; Frangoulis et al., 1999).

1.2.6 The functions of tapasin

Tapasin plays an important role in the assembly of the peptide loading complex in the ER. The formation of a functional peptide loading complex is essential for effective binding of peptides to MHC class I molecules and the expression of these complexes on the cell surface. Tapasin bridges the gap between the transporter and empty MHC class I molecules, bringing them into close proximity to the translocation core where peptides are presented (Sadasivan et al., 1996). In the absence of tapasin, empty MHC class I molecules fail to associate with TAP leading to decreased peptide binding and cell surface expression (Grandea et al., 1995).

Furthermore the functions of tapasin can be divided according to its domains. The TMD and cytoplasmic domains of tapasin interact with TAP for stabilization and transporter function (Garbi et al., 2003), the ER luminal domains

seem to be crucial for MHC class I and ERp57 recruitment and for stable complex assembly (Rizvi and Raghavan, 2006). Tapasin also facilitates the selection of optimal peptides to ensure firm binding. Failure to form stable class I-peptide complexes results in very unstable surface expression of antigens, which are subject to degradation (Chen and Bouvier, 2007). Tapasin appears to be a key player in recruiting ER resident chaperones to the translocation core, where the interaction of all mediators promote the formation of efficient and firm class I-peptide associations (Rizvi and Raghavan, 2010).

Although tapasin seems to play an important role in antigen presentation, certain MHC class I alleles can function in a tapasin-independent manner. A single amino acid substitution in HLA-B from Asp116 to Tyr116 allows the latter to function in a completely tapasin independent manner (Sieker et al., 2007). It also appears that tapasin influences the peptide repertoire presented to MHC class I, favoring certain peptides over others. In the absence of tapasin antigen presentation is altered rather than deficient and is still sufficient to induce in an immune response (Boulanger et al., 2010).

1.2.7 Assembly of the MHC class I and peptide complex

Endogenous antigens are constitutively expressed on the plasma membrane of all nucleated cells. Antigen presentation is important for surveillance by CTLs, which patrol the cell surface to eliminate potentially infected cells. The antigens are derived from cytosolic degradation and presented on MHC class I molecules. $CD8^+$ cells recognize this complex via their T-cell

receptor. Typically cells express self-antigen, which signals a healthy state of the cell, in which case $CD8^+$ detach from cells without any further action. Infected or transformed cells present viral or malignant peptides. Upon detection CTLs are activated, which leads to the release of cytotoxins and the killing of infected cells. MHC class I molecules are ER–resident glycoproteins, which require the help of many other proteins to establish an effective peptide loading complex. This process relies on high quality control to ensure only properly assembled molecules leave the ER and get transported to the cell surface.

Most cytosolic and nuclear proteins are processed by the proteasome. The core is the 20S proteasome, that is composed of four rings each of which is made up of seven subunits. Each ring is formed from α or β subunits. The β subunits form the two inner rings and the α subunits form the outer parts on either end to make a cylindrical shape (Groll et al., 1997). Three of the β subunits are catalytically active and responsible for the degradation of proteins. These are located centrally within the proteasome and therefore proteins must enter the proteasome to reach the catalytic site. This process is controlled by the 19S regulator, which is responsible for the unfolding of proteins, so that they can pass through the proteasome. Together the 20S and 19S make the 26S proteasome complex that efficiently processes proteins to deliver peptides for antigen presentation (Braun et al., 1999).

In an infected cell, the active β subunits are replaced by low molecular mass polypeptides (LMP) 2 and 7 to form the immunoproteasome. These proteins are encoded within the MHC class II region of the genome and are induced by

interferon- γ (IFN-γ) production. The immunoproteasome is more efficient in producing peptides that will strongly interact with class I molecules compared to the regular proteasome (Fehling et al., 1994; Van Kaer et al., 1994).

Peptides that can bind firmly to MHC class I molecules are determined by their anchor residues and have to match the binding motif of class I. MHC class I molecules are polymorphic and therefore have different binding specificities. TAP transports peptides of 8-13 amino acids or longer. Peptides that are too long for binding to class molecules are trimmed in the ER by aminopeptidases like ERAP1 or ERAP2 (Saveanu et al., 2005).

MHC class I molecules are heterodimers composed of a heavy chain and a beta-2 microglobulin (β_2m) chain and are assembled in the ER lumen. MHC class I heavy chains are co-translationally fed into the ER, where they become associated with the transmembrane chaperone calnexin (Degen et al., 1992). Upon association of the β_2m chain with the heavy chain, calnexin is released and a different soluble chaperone, calreticulin, binds to MHC class I (Sadasivan et al., 1996). These proteins are suggested to facilitate the proper assembly and folding of the class I heterodimer. Tapasin associates with the N-terminal domains of the TAP heterodimer (Ortmann et al., 1997) and binds the thiol oxireductase ERp57 (Peaper et al., 2005). The tapasin and ERp57 dimer then recruits calreticulin with the empty MHC class I molecule bringing it into close proximity to TAP to initiate peptide loading (Wearsch and Cresswell, 2007).

Once a stable peptide-MHC class I complex is formed the chaperones dissociate and the complex is released from TAP. It travels via the Golgi apparatus to the cell surface, where it signals to immune cells (Figure 1-4).

1.3 TAP in different species

TAP transporter genes have been identified in the MHC of many non-mammalian as well as mammalian species. These include all vertebrates with an adaptive immune system, such as cartilaginous and bony fish, amphibians, birds, humans and rodents. The structural and phylogenetic study of these genes has provided evidence for linkage and co-evolution of antigen presentation genes in some species.

1.3.1 TAP in lower vertebrates

Many lower vertebrates show a similar gene organization in the MHC, having one MHC class I gene adjacent to two TAP transporter genes. Such an organization indicates the existence of a primordial MHC at the beginning of the evolution of an adaptive immune system. Cartilaginous fish, like sharks, represent the most ancestral vertebrates to have an adaptive immune system. In the nurse shark antigen presentation genes are linked together in one functional haplotype (Ohta et al., 2002). This also appears to be true for bony fish for which antigen presentation genes have been identified. In Atlantic salmon, two *TAP2* and one *TAP1* loci were identified. Both of the *TAP2* loci are expressed and are polymorphic (Grimholt, 1997). Two *TAP2* loci were also identified in the MHC

of rainbow trout. One locus shows polymorphic alleles of the *TAP2* genes (Hansen et al., 1999).

This was extensively studied in the amphibians *Xenopus laevis* and *tropicalis,* and two distinct alleles for both TAPs were identified (Ohta et al., 1999; Ohta et al., 2003). Polymorphisms appear to be in functional regions but have not been tested. Interestingly, in *X. laevis LMP7*, *TAP1* and *TAP2* genes are mapped in the class I region with only one MHC class I gene present. This again indicates the presence of a primordial MHC with tight linkage of antigen presentation genes within a functional haplotype (Ohta et al., 2003).

1.3.2 Mammalian TAP

In mammals, the antigen presentation genes *TAP1* and *TAP2* are distantly removed from the polymorphic MHC class I genes and are located in the class II regions (Beck et al., 1992). Humans express three MHC class genes co-dominantly allowing them to express a variety of class I molecules, which allows binding of many peptides.

Human TAP1 and TAP2 span the membrane ten and nine times, respectively (Figure 1-5a). This places both N- and C-termini of TAP1 in the cytoplasm. The N-terminus of TAP2 on the other hand is located in the ER lumen and the C-terminus stretches into the cytoplasm (Schrodt et al., 2006). The six C-terminal segments and the ABC-cassette of both subunits represent the core of the transporter. The human TAP1 has four accessory TMs and TAP2 has three accessory TMs, which are important for tapasin binding and the assembly of the

peptide loading complex. Two regions within each subunit were identified to be involved in peptide binding (Figure 1-5a). These are localized in the cytosolic loop between TMs 4 and 5 and TM6 spanning into the C-terminal tail of each subunit (Androlewicz and Cresswell, 1994; Nijenhuis et al., 1996).

1.3.3 Chicken TAP

The chicken MHC (B-locus) is simple and small. It is 92 kilobases (kb) in size and contains only 19 genes (Kaufman et al., 1999). This is much smaller than the human MHC with roughly 140 genes. The chicken MHC contains the class I, II and III regions (Guillemot et al., 1988), however their order is different from that in humans. It appears that the chicken MHC still resembles a primordial organization from which the mammalian MHC evolved by rearrangement (Nonaka et al., 1997).

Orthologues of all 19 chicken MHC genes can be found in the human MHC but many of the genes identified in the human MHC are absent in the chicken, such as the proteasomal genes *LMP2* and *LMP7*. In humans *LMP2* and *LMP7* are in very close proximity to the TAP genes, but these have been lost in the chicken MHC and are absent in the genome altogether (Kaufman et al., 1999; International Chicken Genome sequencing Consortium, 2004). The genomic organization of the chicken MHC is strikingly different from the human MHC. Chicken *TAP1* and *TAP2* are located within the class I region, whereas the mammalian TAP is encoded within the class II region. Chickens posses only two classical MHC class I genes, which are polymorphic and flank the two TAP

genes, and two classical class II genes are on either side of the *tapasin* gene (Kaufman et al., 1995).

The chicken TAP genes are in opposite transcriptional direction and share a bidirectional promoter (Kaufman et al., 1999). The gene structure is well characterized for both *TAP1* and *TAP2* in chickens and differs from that of human TAP. The chicken *TAP1* and *TAP2* have 11 and 9 exons, respectively (Walker et al., 2005). In *TAP1* a truncation in exon 1 has been identified, which results in the loss of 150 amino acids (Walker et al., 2005) and causes the loss of the entire TMD0 (Figure 1-5b). Chicken TAP genes are highly polymorphic and polymorphisms are in functional domains.

In chickens the outcome of infection is determined by whether they can or cannot express antigens for that particular pathogen on their cell surface, and defense against pathogens can be life or death (Kaufman 2000). The proximity of TAP to the dominantly expressed MHC class allows for co-evolution of the proteins, where matching alleles of TAP and MHC class I are tightly linked to form functional haplotypes.

1.3.4 Duck TAP

Unlike the chicken, the mallard duck (*Anas platyrhynchos*) does not have a minimal MHC as it carries five MHC class I genes in its genome (Moon et al., 2005). Similar to the chicken, only one of the five genes is dominantly expressed, the gene at the *UAA* locus. Low levels of expression were identified for a second locus (*UDA*). No transcript could be identified for the remaining three loci and

this can be explained by a deletion within the promoter (*UBA* locus), a premature stop-codon in frame for the *UCA* locus and the *UEA* lacks the adenylation signal (Moon et al., 2005). The genomic organization in the duck is similar to that in chickens. In both chickens and ducks, the major MHC class I gene is adjacent to *TAP2* (Mesa et al., 2004). In the duck the two transporter genes, *TAP1* and *TAP2*, are in opposite transcriptional orientation sharing a bi-directional promoter. In ducks all five MHC class I genes are downstream of *TAP2* (Kaufman et al., 1999; Moon et al., 2005); whereas in chickens the two MHC class I genes flank *TAP1* and *TAP2* (Kaufman et al., 1999; Moon et al., 2005).

The predominant expression of one MHC class I gene may limit the duck's ability to present antigens to CTLs, similar to the situation in chickens. A recent study showed that the MHC class I genes are highly polymorphic in wild mallards (Jensen, 2009). A total of 38 animals were analyzed for their expression of MHC class I and 118 MHC variants were detected. Between the alleles, a bias towards one dominantly expressed allele was observed in individual ducks. Many polymorphic residues showed high variation in the wild population and were localized within the binding groove of class I. Genomic analysis suggested that the dominantly expressed class I allele is adjacent to *TAP2* (Jensen, 2009).

Previously two distinct TAP2 alleles were identified in the White Pekin duck (Mesa et al., 2004). To date no sequence information from *TAP1* and *TAP2* has been provided, nor has their diversity been studied in a wild population.

1.3.5 Diversity in TAP

The TAP transporter plays a crucial role in determining the peptide repertoire that is presented to MHC class I. The specificity of TAP pre-selects peptides transported into the ER lumen. Polymorphic residues within the peptide binding region of TAP can influence the specificity and further restrict antigen presentation. This is not only important for the outcome of acute viral infection, but has also been associated with other human diseases. For example, increasingly higher numbers of the TAP2 allele with V577 were observed in patients with Sjoegren' syndrome (Kageyama et al., 2004).

Polymorphisms within the catalytic domains of the transporter can also impact its ability to translocate peptides. The invariant glycines and lysine have been shown to be involved in binding of β- and γ-phosphates of nucleotides and changes to aspartic acid, serine and asparagine results in the loss of ATP binding (Shyamala et al., 1991). Changes of the aspartic acid in the Walker B motif, which is predicted to be important for Mg^{2+} binding, to lysine or arginine equally prevent ATP binding (Shyamala et al., 1991). Mutations of the invariant lysine in the Walker A motif to methionine allow ATP binding but ATP hydrolysis is impaired (Lapinski et al., 2001).

Although polymorphisms were identified within the human transporter, most of them appear not to influence substrate specificity suggesting that the human TAP is functionally monomorphic. This indicates that the substitutions do not alter the specificity allowing TAP to translocate nearly any peptide (Neefjes et al., 1995; Armandola et al., 1996). C213 of the human TAP2 has been shown to

play a role in peptide binding. Although Cys-less TAP complexes appeared to be functional, the substrate specificity was affected (Baldauf et al., 2010).

Interestingly polymorphisms in TAP2 were identified in other species, and were localized to functional regions. In the rat two lineages of TAP2 that differ by 25 amino acids were identified. These differences have been shown to result in a change in the peptide repertoire transported across the membrane. Instead of transporting a broad range of amino acids at the C-terminus only peptides with an aromatic or hydrophobic amino acid at the C-terminus are translocated into the ER. It was observed that each of the TAP alleles confers specificity that matches that of its adjacent MHC class I allele. If the TAP alleles are exchanged, they fail to provide sufficient peptides that can be bound by MHC class I molecules. This phenomenon has been characterized as the *cim*-effect (Momburg et al., 1996; Rudolph et al., 2002).

In the chicken, because of their close proximity to the polymorphic MHC class I genes, *TAP1* and *TAP2* were proposed to also be polymorphic. A model was proposed explaining the consequence of a single dominantly expressed class I gene adjacent to TAP. This model suggested that polymorphic TAPs and MHC class I co-evolve for best fit for one another (Kaufman, 1999). Furthermore it was speculated that the peptide binding specificity of the transporter matches that of the major MHC class I to specifically deliver peptides that can be bound by and expressed on the major MHC class I. A peptide that matches the binding pocket of the minor MHC class I but not that of the transporter, would not be transported

across the ER membrane and would fail to bind to the minor class I (Kaufman, 1999).

By single-strand conformation polymorphism, the chicken *TAP1* and *TAP2* genes were determined to be polymorphic at the nucleotide level (Walker et al., 2005). *TAP1* and *TAP2* were amplified from 16 chicken lines and sequenced to identify SNPs. Thirteen to 17 nucleotide (nt) differences were identified in *TAP1*, of which 14 led to amino acid substitutions. In *TAP2*, 24 nucleotide differences were observed, which led to replacement changes in the amino acid sequence (Walker et al., 2011). An important finding was, that those TAPs obtained from chicken lines with the same MHC class I, were identical. Whereas TAPs from lines with different MHC class I, varied significantly in sequence (Walker et al., 2011).

Analysis of two TAP2 alleles in the White Pekin duck, led to the identification of 26 amino acid differences between the two alleles, of which 18 were non-conservative substitutions (Mesa et al., 2004). These polymorphisms were mapped to the cytosolic loop between TM four and five, and to end of TM six stretching into the C-terminal tail within the core TMD. These regions align to those identified to play a role in substrate binding in the mammalian TAP2 (Figure 1-4a), indicating that the duck TAP2 polymorphisms could influence TAP specificity and result in altered or even limited peptide transport. The tight linkage of TAP to the dominantly expressed class I gene might, similarly to the chicken, suggest co-evolution of the antigen presentation genes to functional pairs in the

duck. The diversity of TAP in wild mallard is not known and has yet to be elucidated.

1.4 Adaptive immunity in the duck

Ducks are the natural reservoir of the influenza A virus, in which all 16 hemagglutinin (HA) and 9 neuraminidase (NA) subtypes can propagate. Ducks often carry the virus asymptomatically and are mostly unharmed by highly pathogenic avian influenza infection. Despite this they show a poor adaptive immune response. Ducks make a truncated IgY antibody as their primary serum Ig indicating an impaired humoral immunity. The truncated antibody lacks the Fc region and can therefore only neutralize the virus. Since IgY lacks the Fc region it cannot bind to Fc receptors on macrophages and ultimately cannot initiate the killing of infected cells resulting in poor humoral immunity to the virus (Magor et al., 1994). Consequently ducks have to rely on cell-mediated immunity to create an effective adaptive immune response. Cell-mediated immunity is induced by antigen presentation. If antigen presentation is limited in the duck due to pre-selection of peptides pumped into the ER by TAP and additionally due to the reduced usage of MHC class I, the cell-mediated response could also be reduced.

1.5 Co-evolution of antigen presentation genes

In mammals, the TAP genes are located in MHC class II region of the genome distantly removed from the MHC class I genes. This wide separation allows for recombination (as reviewed by Carrington, 1999) and likely led to the

emergence of a functionally monomorphic human TAP transporter, which can provide peptides for any gene in the MHC class I multigene family (Momburg et al., 1994b; Obst et al., 1995).

In contrast to this, rats have two distinct TAP2 lineages. Here, the TAP genes are adjacent to the MHC class I genes. The two TAP alleles are functionally distinct with different binding specificities and usually associated with a certain MHC class I haplotype. The TAP2 alleles are distinct in that allele A is permissive and can transport almost any peptide, like the human TAP. Allele B showed specificity for peptides with aliphatic amino acids at the C-terminus (Powis et al., 1992). This is significant when considering the associated MHC class I allele. In fact it was demonstrated in MHC-congenic rats, that TAP from *cim*^{a} could efficiently provide peptides for RT1A^{a} expression, where as *cim*^{b} failed to do so (Powis et al., 1992), suggesting highly specific MHC class I and TAP haplotypes. This provides significant evidence for co-evolution of these genes (Joly et al., 1998).

In lower vertebrates, a simple MHC was found with one MHC class I gene tightly linked to TAP (Nonaka et al., 1997; Flajnik et al., 1999). In the frog *Xenopus leavis* two distinct lineages of TAP were identified, that diverged from a common ancestor 60-100 million years ago. The persistence of two TAP lineages suggests different specificities of the two TAP alleles (Ohta et al., 2003).

In the chicken, only one of the two MHC class I genes is expressed at high levels. Both TAP and MHC class I are highly polymorphic in chickens and it is suggested that the proteins co-evolve due to the close proximity of their genes

(Kaufman, 1999; Walker et al., 2005; Walker et al., 2011). Furthermore the proximity suggests the lack of recombination in the small MHC of the chicken. TAP1 and TAP2 of 16 chicken lines were analyzed for their diversity. Of the 16 lines, seven had identical MHC. Interestingly, it was observed that TAP1 alleles with the same MHC class I are identical. TAP1 alleles with different MHC haplotypes differ in their sequence. This is also true for TAP2. Peptides presented on MHC class I on erythrocytes were isolated from the B4 haplotype. A distinct motif of these peptides was observed, which interestingly was identical to the binding motif of the major MHC class I. Iodinated indicator peptides were used to analyze peptide transport. Peptides with a different charge or other substitutions were transported and clearly showed, that TAP can translocate peptides that will match the binding motif of the major MHC class I. This suggests that TAP and MHC class I form a functional haplotype in chickens, which can be explained by co-evolution of polymorphic genes that are in close proximity (Walker et al., 2011).

In the duck, a similar trend was observed, that indicates the expression of a single MHC class I (Mesa et al., 2004; Jensen, 2009). As in chickens and rats, the major MHC class I gene is adjacent to TAP, which raises the question if the rule of co-evolving MHC class I and TAP haplotypes is true for mallard ducks.

1.6 Research Goals

Ducks, like the chicken, express only one MHC class gene predominantly. This limits the possibilities of binding viral peptides effectively and presenting

them to CTLs. Previous investigations have shown that the class I genes are highly polymorphic in wild mallards. Since TAP is important for antigen presentation the specificity of TAP could additionally impact the repertoire of peptides expressed on the cell surface.

Because of the proximity of the polymorphic MHC class I to TAP and tapasin we predicted these proteins are polymorphic in wild mallards. We hypothesized that polymorphisms within TAP1 and TAP2 might lead to changes in specificity of the transporter, and therefore restrict the repertoire of peptides that can be transported in ducks. This could contribute to the poor adaptive immune response mounted by ducks following influenza vaccination and the perpetuation of the virus through re-infections. We believe that this weak adaptive immune response results in the duck being the perfect host for the virus.

To elucidate this, our main objective was to analyze the allelic diversity of the TAP transporter in wild mallards. In chickens, TAP is in close proximity to the polymorphic MHC class I genes. Both TAP1 and TAP2 are highly polymorphic as well, which suggests co-evolution of these genes. The genomic organization is strikingly similar between the chicken and the domestic duck. In the duck TAP is adjacent to the predominantly expressed MHC class I gene, which is polymorphic in wild mallards. To determine the diversity of TAP sequences in wild mallards, *TAP1* and *TAP2* were amplified from 12 ducks and alleles were identified. Comparing the amino acid sequences we were able to identify polymorphisms and map them to functional regions. This allowed us to make predictions about potentially functional polymorphisms, which can limit the

peptide repertoire presented to MHC class I molecules. By comparison of our MHC class I and TAP data we will be able to determine if these polymorphic genes co-evolve in ducks. We were also interested in the regulation of both TAP subunits and whether their promoter is also polymorphic. To investigate this we amplified a fragment across the promoter with gene specific primers in *TAP1* and *TAP2*. This also allowed us to identify TAP alleles that are linked on one haplotype. Additionally we tested if *TAP1* and *TAP2* are upregulated during avian influenza (AI) infection. Lastly we wanted to identify whether the antigen presentation complex in the duck includes the association with the chaperone tapasin.

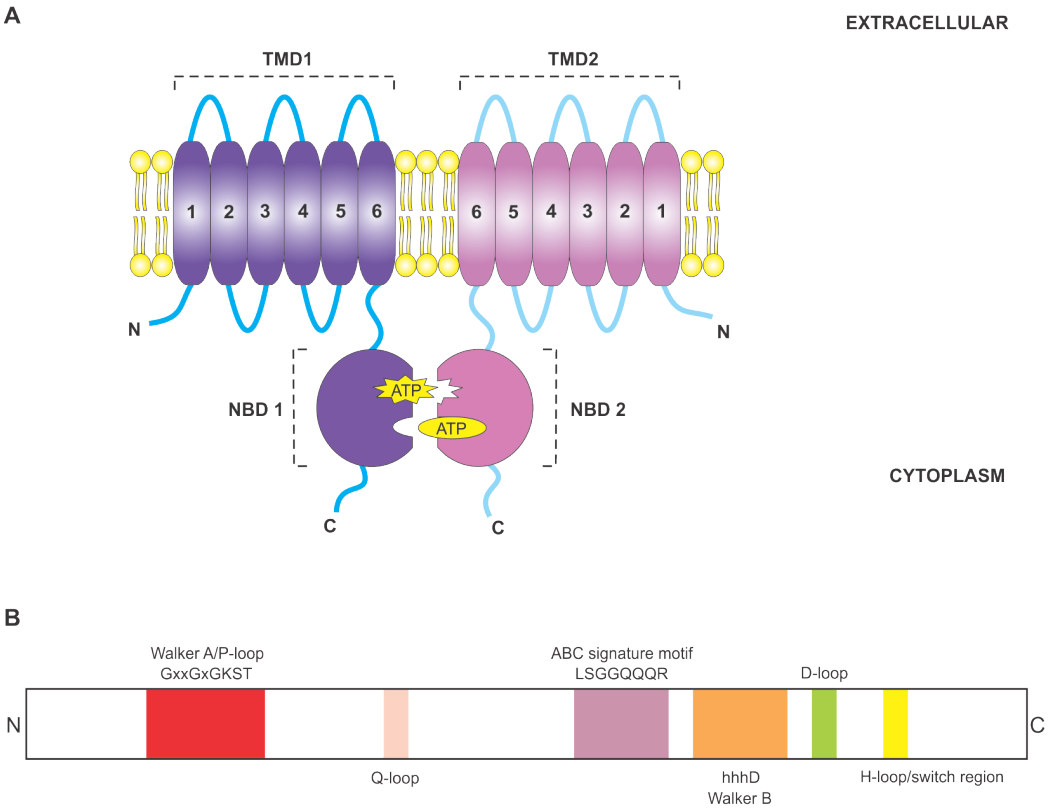

Figure 1-1. Structure of a typical ABC transporter. a) An ABC transporter consists of four functional domains, two transmembrane domains (TMD) and two nucleotide binding domains (NBD), which are the catalytic domains of substrate translocation. A transporter has 6 transmembrane segments (TMs) and functions as a homo or hetero dimer. Illustrated here is a heterodimer that usually consists of two homologous subunits (pink and purple). b) The NBD of a transporter subunit has characteristic motifs, which are conserved between all ABC transporters.

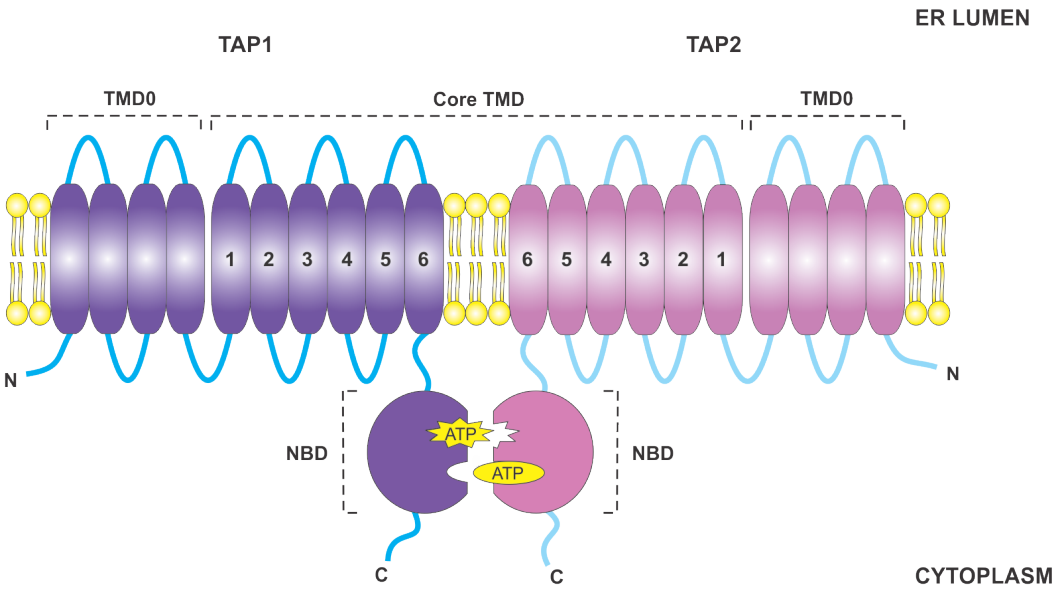

Figure 1-2. Structure of TAP. A typical TAP consists of two subunits, TAP1 and TAP2 shown here in purple and pink respectively. Each subunit is composed of a transmembrane domain and a nucleotide binding domain. The TMDs contain additional transmembrane segments, which form the accessory domains. NBD, nucleotide binding domain; TMD, transmembrane domain; TMD0, accessory TMD.

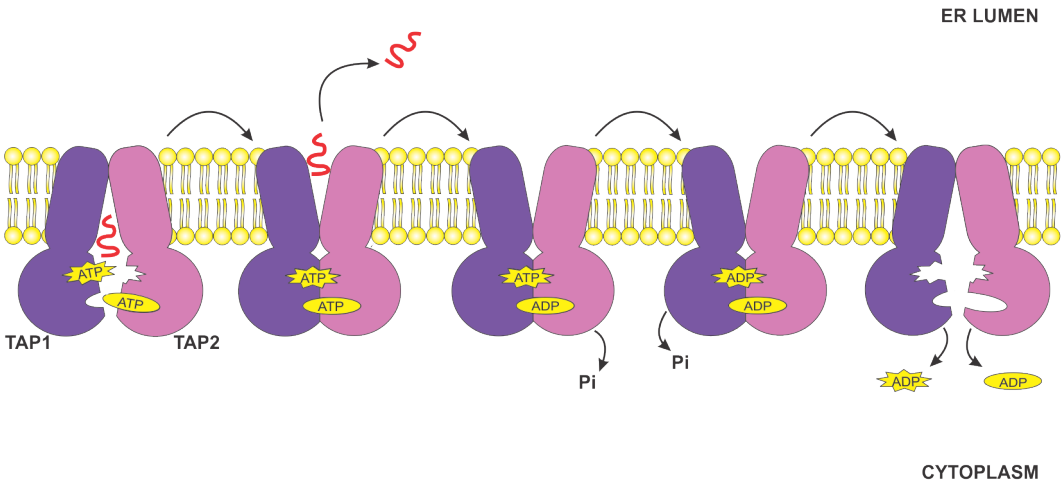

Figure 1-3. Peptide transport is powered by ATP-hydrolysis. Peptides associate with TAP in the cytoplasm and induce ATP binding. The NBDs of the two subunits dimerize trapping the ATP in between them. The TMDs undergo conformational changes leading to the opening of the translocation pore towards the ER lumen and releases the peptide. ATP is hydrolyzed and phosphates are released from first TAP2 and then TAP1. This leads to dissociation of the two NBDs inducing again a conformational change, which allows the transporter to close the pore towards the ER and open the NBDs for new ATP binding.

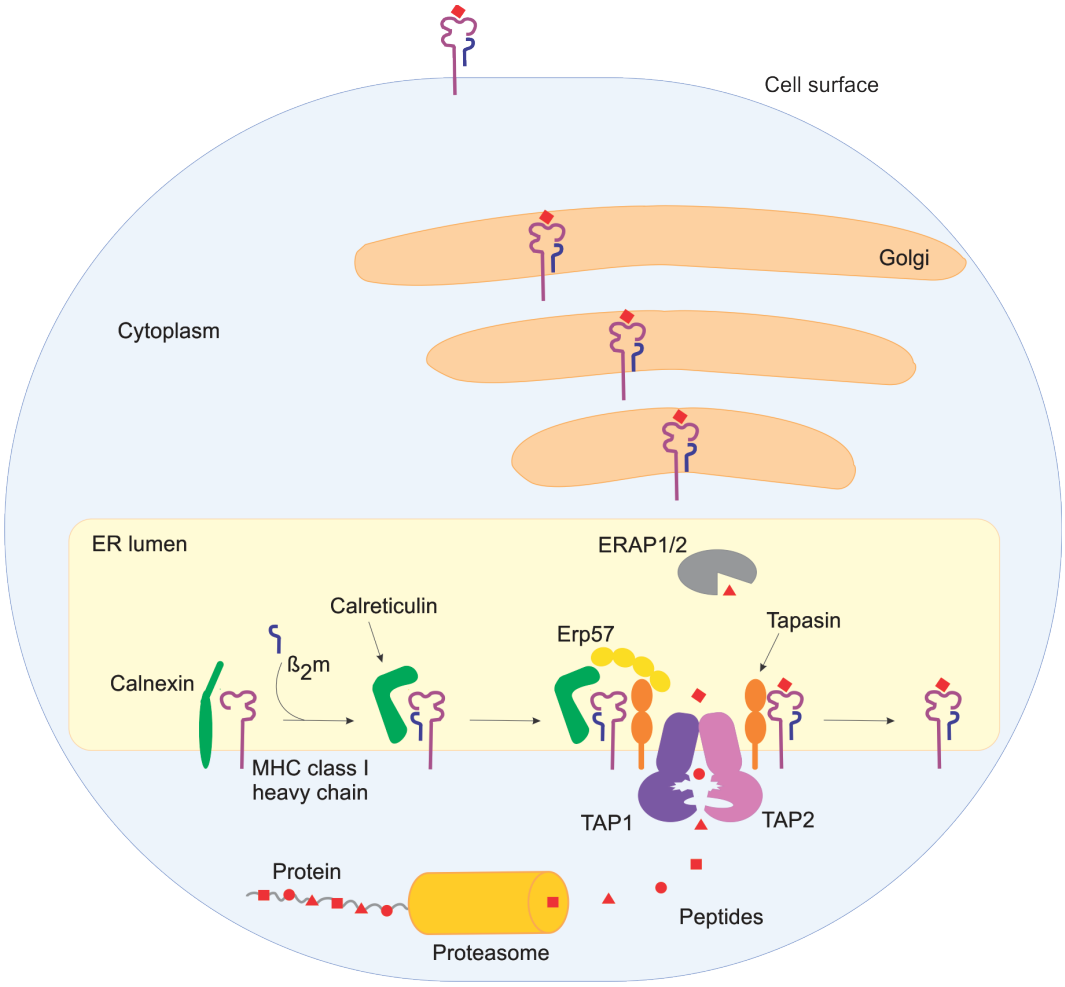

Figure 1-4. The peptide loading complex. Peptides derived from cytosolic degradation are transported into the ER by TAP, where the peptide loading complex forms. This process involves the interaction of many ER resident chaperones and facilitates peptide binding onto MHC class I. A stable MHC class I-peptide complex traffics via the Golgi to the cell surface, where it can be detected by CTLs.

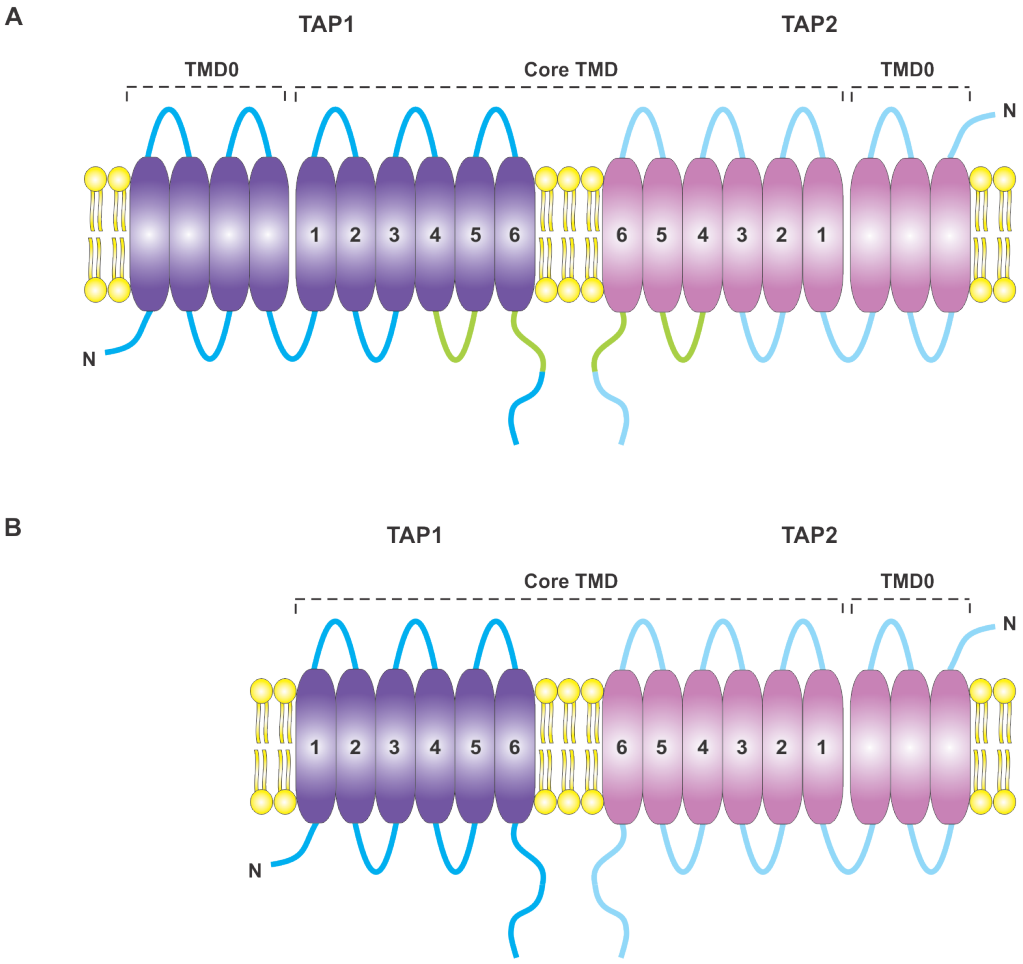

Figure 1-5. The structure of the mammalian and chicken TAP. a) The human TAP is composed of ten and nine TMs for the TAP1 and TAP2 subunits, respectively. The peptide binding region is indicated in green, and is located in the cytosolic loops and C-terminal tails. b) The chicken TAP is composed of six and nine TMs for TAP1 and TAP2, respectively.

2 MATERIALS AND METHODS

2.1 Collection of blood samples from wild mallards

Fifty wild mallard ducks were captured at four sites in Alberta, Canada: Bittern lake (15 mallards), Buffalo Lake (16 mallards), Hay lake and nearby lakes (17 mallards) and Sisib lake (2 mallards). Blood was collected from the leg of the mallard and used for DNA and RNA extraction (Jensen, 2009).

2.2 Genomic DNA and RNA extraction from duck blood and cDNA synthesis

Total DNA and RNA were extracted from the mallard blood using TRIzol® (Invitrogen). Genomic DNA and RNA were stored at 4°C and -80°C, respectively. RNA was quantified using a NanoDrop ND-1000 UV-vis Spectrophotometer (NanoDrop Technologies) and one microgram (μg) of total RNA was used for cDNA synthesis. The samples were DNaseI (Invitrogen) treated to remove any DNA contamination and then used for cDNA synthesis with SuperScript™ III and an Oligo (dT) primer according to the manufacturer's instructions (Invitrogen). To confirm cDNA quality GAPDH was amplified using the gene specific primers GAPDHF1 (5′-CCGTGTGCCAACCCCCAATGTCT-3′) and GAPDHR1 (5′-GCCCATCAGCAGCCTTCACTAC-3′).

2.3 Primer design and amplification of the TAP genes

The first primer set for *TAP2* (DT2E1F3/DT2E93′UR) was designed by Chelsea Davidson and the gene was amplified from genomic DNA from five

mallards (18, 19, 24, 32, 42). Subsequent gene specific primers were designed based on the sequences available in GenBank (accession #AY885227.1) and Ensemble (the duck genome) and *TAP1* and *TAP2* were amplified from cDNA. Several primer sets had to be used yielding different fragment sizes. For a detailed list of primer pairs and amplicon size see Table 2-1. Primers that were designed by Chelsea Davidson are indicated by an asterix. *TAP1* and *TAP2* sequences were obtained from a total of 12 mallards, for some of which multiple primer sets were used to amplify fragments of the gene or cDNA. Table 2-2 shows the primer pairs used for each sample. The primer binding sites are illustrated in Figure 2-1. The amplification was carried out in a total volume of 25 microliters (µl) using the AmpliTaq Gold® 360 master mix (Applied Biosystems) or Kapa HiFi™ PCR kit with fidelity buffer (Kapa Biosystems). Two µl of the GC enhancer was added to all reactions with AmpliTaq®. The cycling conditions for AmpliTaq® Gold PCRs were as follows: 1 min initial denaturation at 95°C, 35 cycles with 30 s at 95°C, 30 s at 50-64°C, 1-2:45 min at 72°C, followed by a final extension of 7 min at 72°C. The cycling conditions used for PCRs with the Kapa HiFi™ polymerase were the following: 1 min initial denaturation at 95°C, 35 cycles with 20 s at 98°C, 15 s at 57-69°C, 0:45-3 min at 72°C. All PCRs were carried out in a Veriti™ thermal cycler or GeneAmp® PCR System 9700 (Applied Biosystems).

2.4 Gel Extraction and adenylation

For amplifications which produced more than a single band, the product of expected size was cut out and purified from the gel using the Qiaquick Kit

(Qiagen) according to the manufacturer's instructions with one exception, the samples were eluted in 30µl of sterile water. PCR amplifications that yielded a single band were used directly for cloning without any purification steps. Reactions performed with the Kapa HiFi™ polymerase produced blunt ends and had to be adenylated post PCR amplification. A separate reaction was set-up with a Taq polymerase and 5µl 1mM dATPs and incubated for 20 min at 72°C.

2.5 Cloning and sequencing of the TAP genes

PCR products or purified fragments were ligated into the TOPO pCR2.1® vector (Invitrogen) for 30 min at room temperature (RT) according to the instruction manual. Chemically competent DH5α cells were transformed by incubating for 30 min on ice, and grown on plates containing kanamycin. Clones with inserts were selected using LacZ blue-white screening. For each mallard a minimum of six LB/kanamycin overnight cultures were grown from white colonies and plasmid DNA was isolated using the GeneJET™ Plasmid Miniprep Kit (Fermentas) following the manufacturer's instructions, except that instead of using elution buffer the plasmid DNA was eluted in 50µl of sterile water at the final step. The plasmid DNA sequence was determined using BigDye Terminators V3.1 (Applied Biosystems) with vector specific primers, M13F (5′-GTAAAACGACGGCCAG-3′) and M13R (5′- CAGCAAACAGCTATGAC-3′) and a 3730 DNA Analyzer (Applied Biosystems). All reactions were carried out in a total volume of 10µl. The addition of 1M [final concentration] betaine to some sequencing PCRs helped to improve sequence quality and length of reads.

Two different alleles were identified for heterozygous mallards and three independent clones for every allele in every animal were sequenced with insert specific primers, until 100% clear coverage of both strands was achieved. The sequences of all the primers used for sequencing of *TAP1* and *TAP2* can be found in Table 2-3 and Table 2-4.

The predicted *TAP1* mRNA sequence from the duck genome spans only exons 2-8 due to gaps in the sequence. These exons were identified in the unannotated genome of the duck and the contig containing TAP1 was searched for a poly-A signal downstream of exon 8. By identifying splice sites and by comparison to the chicken and human TAP1 amino acid (AA) sequences exons 9 to 11 were predicted. Forward primers in exons 6 (DT1E6F3) and 9 (DT1E9F1) and reverse primers in exon 11 (DT1E11-3′UR) and the 3′ untranslated region (UTR) (DT13′UR) were designed. Two primer combinations were used to amplify the fragment from 6 mallards (18, 19, 20, 24, 32, 42) to confirm manually predicted exons 9-11.

2.6 *In silico* sequence analysis

All sequences were edited in ContigExpress (Invitrogen) and a consensus of three clones for each allele in each individual animal was created using CLUSTALW2 (EMBL-EBI). Full length sequences were assembled by alignment of the obtained fragments. The coding sequences (CDS) were translated into AA sequences with the ExPASy translate tool (Swiss Institute of Bioinformatics). Alignments and phylogenetic trees of TAP1 and TAP2 were made using

SeqBuilder and MegAlign by Lasergene 9 (DNASTAR). Hydrophobicity analysis was done to predict transmembrane segments and the topology of TAP by using the TMpred and TopPred2 servers. Functional domains were identified using the Simple Modular Architecture Research Tool (SMART).

2.7 Amplification and analysis of the TAP promoter

A fragment across the TAP promoter was amplified using the Kapa HiFi™ PCR kit with the fidelity buffer (Kapa Biosystems) and gene specific primers for *TAP1* and *TAP2*, DT1E4R1 (5′-CTGCTGGAGCCTGCCGATGGC-3′) and DT2E1R1 (5′-GATGGCAGTGGCAAAGGCAGTG-3′) yielding an amplicon of approximately 2.9 kb in size (Figure 2-2). Approximately 20-30 nanograms (ng) of genomic DNA was used as template in all reactions. The product was amplified using the following cycling conditions: 5 min at 95°C and 30 cycles with 20 s at 98°C, 15 s at 72°C, 2:30 min at 72°C. The TAP promoter was obtained from six mallards (18, 19, 20, 24, 32, 42). All amplifications producing a single band were poly-A tailed as described before for TOPO TA Cloning® and subsequent transformation of chemically competent DH5α cells. Positive clones were screened for different alleles and three clones per allele were sequenced with insert specific primers (Table 2-5). The mallard duck promoters were analyzed with the transcription element search system (TESS) to identify transcription factor binding sites. Additionally, TAP1 and TAP2 alleles that are linked on the same haplotype were identified.

2.8 Infection with avian influenza and RNA extraction

White Pekin ducks were either mock infected or with an H5N2 or an H5N1 virus. A/Viet Nam/1203/04 (H5N1) was a highly pathogenic isolate from a fatal human infection and A/mallard/British Columbia/500/05 (H5N2) was a low pathogenic isolate from birds. Mock infections were performed with phosphate buffered saline (PBS) only. Six ducks were inoculated via the natural route of nose, eyes and trachea and sacrificed either on day one or day three post-infection. For each experiment, lung tissues from three ducks were collected and used for total RNA extraction.

2.9 Expression analysis of the TAP genes by RT-PCR

One μg of total RNA from mock, low and high path infected lung samples was used for cDNA synthesis as described before. Using gene specific primers for *TAP1* and *TAP2* fragments of approximately 150bp were amplified and their expression was analyzed by reverse transcription PCR (RT-PCR) using Platinum® Taq (Invitrogen). The primers used for *TAP1* are DT1E6F1 (5′-CTCGTCACCTTCCTCCTCTAC-3′) and DT1E7R1 (5′-GTGCTAGTGTCCCCGAGGGTG-3′) and the *TAP2* primers used are DT2E6F1(5′-GCGCCCCGGTGAGGTGACAGC-3′) DT2E7R1 (5′-CCCGGATGGAGCCAGAGAAG -3′). RT-PCRs were performed using lung samples from one day and three days post infection. All PCRs included a GAPDH control and were carried out under the same cycling conditions, which are as follows: 1 min at 95°C, 30 cycles (*TAP1*) or 25 cycles (*TAP2*) with 1 min at 96°C,

30 s annealing temperature and 30 s at 72°C, where the annealing temperature and was 52°C and 59°C for *TAP1* and *TAP2*, respectively.

2.10 Expression analysis of the TAP genes by quantitative PCR

To measure the fold upregulation of the TAP genes during avian influenza infection, quantitative PCR (qPCR) was performed using *GAPDH* as an endogenous control. Five hundred ng of RNA was used for cDNA synthesis as explained before. The qPCRs were performed using the FastStart TaqMan® Probe Master Mix (Roche) and gene specific probes and primers. The probes and primers were designed using Primer Express version 3.0 (Applied Biosystems) and sequences are as follows: TAP1 probe: /56-FAM/AGGTCCTGC/ZEN/TCCGCTACTACCCCACTC/3IABkFQ/, primer 1: GGAGCCCACAGCCTTCGT, primer 2: CCTCCTCTACCAGTTGCAGTTCA, TAP2 probe: /56-FAM/AGCGCTTCT/ZEN/ATGAGCCCACGGCTG/3IABkFQ/, primer 1: CGGTGTAGGTACCGGTGCTT, primer 2: TGAACGGCAGTGGGAAGAG, GAPDH probe: /56-FAM/ACCACCAAC/ZEN/TGCCTGGCGCC/3IABkFQ/, primer 1: AAATTGTCAGCAATGCCTCTTG, primer 2: TGGCATGGACAGTGGTCATAA. All qPCRs were carried out in a 7500 Fast Real-Time PCR System (Applied Biosystems) with standard cycling conditions (10 min at 95°C, 40x 15 s at 95°C and 1 min at 60°C).

2.11 PCR amplification of *tapasin*

Previously, the diversity of the avian t*apasin* gene in chickens, turkeys and pheasants was studied by amplifying a fragment of the genes spanning from exon five to six (Sironi et al., 2006). The same primers were ordered, and more primers were designed. Degenerate primers were made based on *tapasin* sequence alignments from human, mouse and chicken. Primers were designed to amplify fragments of conserved regions. Primers were designed to the chicken sequence within the Ig domain encoded by exons 4 and 5. Primers were used in different combinations on mallard cDNA and genomic DNA to attempt to amplify tapasin. The sequences of all primers are listed in Table 2-6.

2.12 Generation of a tapasin probe for southern blot hybridization

One μg of RNA, extracted from chicken lung tissue, was used for cDNA synthesis and 1μl was used to amplify the tapasin probe with TAPBPE5F1 and TAPBPE6R1 with the Platinum® Taq (Invitrogen). The cycling conditions were 1 min at 95°C, 30 cycles with 1 min at 96°C, 30 s at 53°C and 30 s at 72°C. A band 269 bp in size was gel extracted and sequenced to confirm that tapasin was amplified. Probes were labelled using the AlkPhos Direct Labelling and Detection System with CDP-Star (Amersham Biosciences) following the instructions in the manual.

2.13 Southern Blot analysis

Two separate blots were prepared. For the first blot genomic DNA from a chicken and a White Pekin duck was digested with BglII, HaeIII and PstI. Another Southern blot was set-up with PstI digested DNA from duck, chicken, turkey, pheasant and pigeon.

Ten µg of genomic DNA from several avian species was digested with 25 U of restriction enzymes in a total reaction volume of 100µl at 37°C overnight. Another 25 U of enzyme was added the next day and incubated for an additional 1-2 hrs for complete digestion. Three restriction enzymes were used: BglII, HaeIII and PstI (New England BioLabs). DNA was precipitated using 1/10 volume of 3M sodium acetate pH5.2 and 1 volume of isopropanol for 1 hrs at -20°C. After recovering the DNA by centrifugation the pellet was washed in 70% ethanol and resuspended in 10µl ddH_2O. The DNA was separated on a 0.7% agarose gel in 0.5X TBE buffer overnight at 30V. The gel was incubated in 0.2M HCl for 10 min. To denature the gel, it was immersed in 1.5M NaCl, 0.5M NaOH for 15 min and repeated for another 15 min with fresh solution. The reaction was neutralized by incubation the gel 3M NaCl, 0.5M Tris HCl (pH7) for 15 min (twice). The DNA was transferred onto Nytran® SPC 0.45µm Nylon Transfer Membrane (Whatman®) overnight with 2X SSC (20X SSC: 3 M NaCl, 0.3 M sodium acetate) and immobilized by cross-linking using a UV Stratalinker© 2400 (Stratagene). The hybridization buffer was prepared according to the instructions in the AlkPhos Direct Labelling and Detection System with CDP-Star (Amersham Biosciences) manual. The membranes were pre-incubated for 1 hr in hybridization

buffer before labelled probes were added. The blots were hybridized overnight at 55°C and 60°C. The blots were washed in primary wash buffer (2 M Urea, 0.1% (w/v) SDS, 50mM Na phosphate pH7, 150mM NaCl, 1mM $MgCl_2$, 0.2% (w/v) blocking reagent) and secondary wash buffer pH10 (1M Tris base, 2M NaCl, 2mM $MgCl_2$) at 55°C or 60°C. Signal generation and detection was performed by incubation of the membrane with detection reagent for 5 min and exposure to Kodak BioMax XAR Film scientific imaging film.

Table 2-1. Sequences of primer pairs used for *TAP1* and *TAP2* amplification and expected fragment sizes. Primers designed by Chelsea Davidson are indicated by an asterix.

Primer name	Sequence 5′→3′	Fragment size
TAP1F2 DT1E8R4	TGATGGGGGCCTCGTCGCTCG GTCCAAGAGCAGGCGCCCGGC	1.103 kb
DT1E2-3F1 DT1E8R4	AGCGCCGTCACCGAGCTGGCC GTCCAAGAGCAGGCGCCCGGC	981 bp
DT1E3F5 DT1E8R6	GAGCTGGCCTGCGACATCGCC TGCAGGCGCAGCAGCAGGGAC	938 bp
DT1E4F2 DT1E8R5	AGCCTGCTGCTGTGGTACCTGG CGTGGGCTGGTGCAGGCGCAG	762 bp
DT1E6F3 DT1E11-3′UR	TTCTCAGCGCTGGCCCTGAAG TGGTGTCCCCATCCTAGCTCC	1.249 kb
DT1E9F1 DT13'UR	TCATCACCCGCCTGCCCCGAG ATGGTCCCTTGTCACTGCTCC	675 bp
DT2UTF1 DT2E93′UR*	GGCCTTGAACGCTCCTCAGAGC TCAGCTGCTCCATGGGTTGTTAG	2.153 kb
DT2UTF2 DT2E93′UR*	CCAAGCCATGGAGTTGCTGCC TCAGCTGCTCCATGGGTTGTTAG	2.129 kb
DT2UTF2 DT2E9R1	CCAAGCCATGGAGTTGCTGCC GCAGGCGGCTGTAGGGCCCAC	2.099 kb
DT2E1F3* DT2E93′UR*	GAGACCTCCGTTCCCTACTG TCAGCTGCTCCATGGGTTGTTAG	1.617 kb
DT2E1F4 DT2E9R1	ATGGAGTTGCTGCCCACCTTG GCAGGCGGCTGTAGGGCCCAC	2.092 kb

Table 2-2. The primer sets used for amplification of *TAP1* and *TAP2* from various ducks. Primers designed by Chelsea Davidson are indicated by an asterix.

	Mallard number											
Primer set	9	10	12	16	17	18	19	20	24	25	32	42
TAP1F2/ DT1E8R4		+	+	+	+			+		+		
DT1E2-3F1/ DT1E8R4	+	+	+	+	+	+	+	+	+	+		
DT1E3F5/ DT1E8R6												+
DT1E4F2/ DT1E8R5											+	
DT1E6F3 DT1E11-3′UR						+	+	+	+		+	+
DT1E9F1 DT13′UR						+		+				
DT2UTF1 DT2E93′UR*		+								+		
DT2UTF2 DT2E93′UR*	+		+	+	+			+			+	+
DT2UTF2 DT2E9R1		+		+								
DT2E1F3* DT2E93′UR*						+	+		+		+	+
DT2E1F4 DT2E9R1								+				

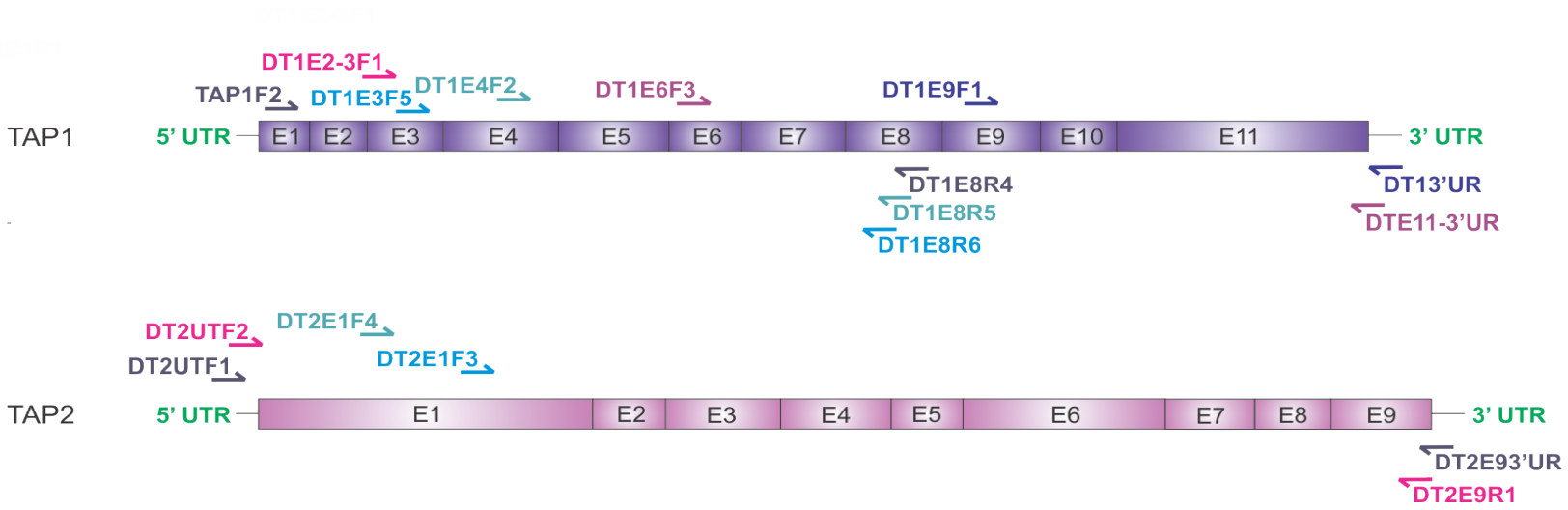

Figure 2-1. TAP primer binding sites for *TAP1* and *TAP2* amplifications from cDNA. The arrows indicate the approximate binding sites of the primers used for amplification of the TAP genes.

Table 2-3. The primers used for sequencing of *TAP1*.

Primer Name	Primer Sequence 5′→3′
M13F(-20)	GTAAAACGACGGCCAG
M13R	CAGGAAACAGCTATGAC
DT1E3F1	GGATGTGGCAGCACGGGTGACA
DT1E3F2	AGCCTGCTGCTGTGGTACCTGG
DT1E4R1	CTGCTGGAGCCTGCCGATGGC
DT1E6F1	CTCGTCACCTTCCTCCTCTAC
DT1E7R1	GTGCTAGTGTCCCCGAGGGTG
DT1E8F2	TTGGACGGCCATCCCCTCAGC
DT1E8R4	GTCCAAGAGCAGGCGCCCGGC
DT1E10F1	AGTGCCAGCAGCAGGTGGAGC
DT1E10R1	TGGCCACCCCTTGCCGCTGTC

Table 2-4. The primers used for sequencing of *TAP2*.

Primer Name	Primer Sequence 5′→3′
M13F(-20)	GTAAAACGACGGCCAG
M13R	CAGGAAACAGCTATGAC
DT2E1F1	TGAGACCTCCGTGCCCTACTG
DT2E1F2	CACATGGCTGGAGGCTGGG
DT2E1F3	GAGACCTCCGTTCCCTACTG
DT2E1R1	GATGGCAGTGGCAAAGGCAGTG
DT2E1R	GCCCAGCCTCCAGCCATGTG
DT2E2F	CTTCCAGCAGACCACAGCAGG
DT2E2F2	TTCATCATGGCTCGCCTCACC
DT2E3F1	TGCCCCTTGCCGTTGTTGCGC
DT2E3F2	TGATCTGGCCTCCCGGCTGG
DT2E3R	CAAGCAGCGCCAGCAGGGTCAG
DT2I3F	TGGTGGACATGGGACAGGAGAA
DT2E4F	GGTGCAGGAGGCCATCT
DT2E4R2	GGCCCTCTGCACCACCA
DT2E4R	CCCGCCGGATGAGGAGGA
DT2E5F	ACCCTCACTGCTGGCGGCCTTG
DT2E5F1	GTCGCCTTCATCCTCTACCAG
DT2E5R2	GGCCGCAGTACAACACCAGTG
DT2I5RM	AGGTACTGGCACCAAGGAWGG
DT2E6F1	GCGCCCCGGTGAGGTGACAGC
DT2E6R	TCGAAGATTTTGTGGGCAGCTG
DT2E6R2M	CAGGTCRCCGTAGGCGTAYGC
DT2E7F	CCATCCGGGACAACATTGCCTAT
DT2E7F2	CTCTGCACTGGACCAAGGCTT
DT2E7F3	GGAGGAGGAGATCATAGCAG
DT2E7R1	CCCGGATGGAGCCAGAGAAG
DT2E8I8F1	GGGATGTGGCGGTGAGTGGTG
DT2E8R1	CAACACGCTGCTTCTGCCCA
DT2E9R1	GCAGGCGGCTGTAGGGCCCAC

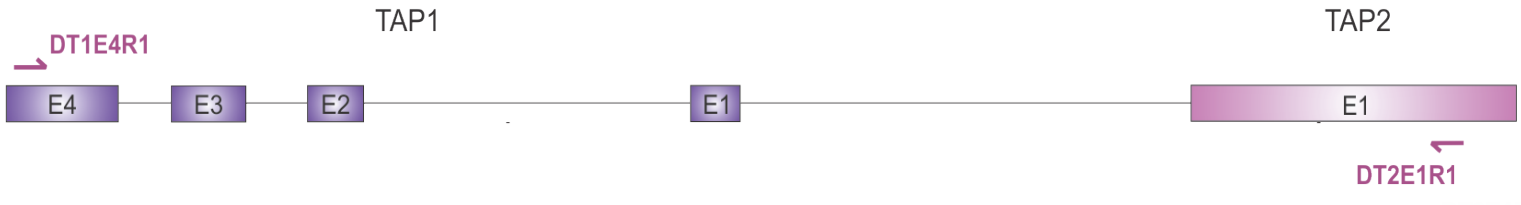

Figure 2-2. TAP primer binding sites for promoter amplification. The arrows indicate the approximate binding sites of the primers used for amplification across the TAP promoter.

Table 2-5. The primers used for sequencing of the TAP promoter.

Primer Name	Primer Sequence 5′→3′
M13F(-20)	GTAAAACGACGGCCAG
M13R	CAGGAAACAGCTATGAC
DT1E1F5	GAGCTGGCCTGCGACATCGCC
DT1E1R1	GGCGCAGCGACTCCAGGAGGC
DT1E1R2	AGCGACGAGGCCCCCATCAGC
DT1E1R4	AGCGACTCCAGGAGGCGGCGG
DT1E1R5	CGCAGCGACTCCAGGAGGCGG
DT1I1F1	CAGAGCAGCATTGATATTTGG
DT1I1R1	GCAACCCTCTACATTTCATGC
DT1I1R3	GATTTATGGCTGGCCCCCCAC
DT1I1R4	ATGGCTGGCCCCCCACAT
DT1I1R5	AGCTCAACGCTGCTCCCCATT
DT1I1R8	TCCCCATTGGGTTGCTCTCTC
DT1I1R9	CTCAATGCAACTCCCCATTGG
DT1I1R10	CTGCTCTGATAGGTGTGCTGC
DT1I1R11	CTCCCCAGATTTTGTGCAGCC
DT1E2R1	ATACCTGCTCAGCCCCAGCAG
DT1E2R2	AGTAGGGGATGGCCATCTCC
DT1E3R1	CCACGAAGGCGATGTCGCAGG
DT1E4R1	CTGCTGGAGCCTGCCGATGGC
DTAP1E1R3	CGAGCGACGAGGCCCCCATCA
DTAP1I1R2	AAGTGGGTGTGCTGCCCCCCT
DT1I1E1R1	CTCCCCTACCGAGCGACGAGG
DTAPpromR1	TGAGCTCCCTGGGACACCAGT
DTAPpromR2	ATCCCCAAACCTCTCAGGAAC
DTAPpromR3	CCATCGTCTCCGGTCAGTGC
DTAP2promR5	AGTTGAGCTGAGCAGGTCTGG
DTAP2promR6	CTCTATGCTCTGGCCCCGCCA
DTAP2promR7	GTTTCATTTTCCGAGAAACGG
DTAP2promR8	TGACGCGCCACCGGCCTCAGC
DTAP2promR9	AAGCTGCTCCGGCCCGATCTG
DTAP2promR10	CGCCCCACTCCCGCCTCACCA
DTAP2promR11	TCCCCCCAGCCTGCTCTGCTT
DT2E1R1	GATGGCAGTGGCAAAGGCAGTG
DT2E1R	GCCCAGCCTCCAGCCATGTG
DT2UTF2	CCAAGCCATGGAGTTGCTGCC
TAP1F1	TGGAGTCGCTGCGCCCCGAGC
TAP1F2	TGATGGGGGCCTCGTCGCTCG

Table 2-6. Sequences of the *tapasin* primers.

Primer name	Sequence 5′→3′	Description
TAPBPE5F1	GGGACACAGTGATGGACAGC	Exon 5 forward
TAPBPE6R1	GTAGAGCCAACGGATGAGGC	Exon 6 reverse
UTAPBPF1	TGGYVCTGRCBGTSCTCACCCA	Universal FW
UTAPBPR1	GYGWCGCCACTCBAGSMCAAAG	Universal REV
UTAPBPF2	TGAGGGCACCTACATCTGCTC	Exon 4 FW
UTAPBPR2	ACGTCACCGTCACATCCAAGG	Exon 5 REV
UTAPBPR3	ACCGTCACATCCAAGGGGTAG	Exon 5 REV
DTAPBPE4F1	ATCCCACTGCACTGCGCCTTC	Exon 4 FW
DTAPBPE4F2	ATCCTCCTTTGTCCTCGAGTG	Exon 4 FW
DTAPBPE4R1	GCTGCAGCACTGTCTGTGTGT	Exon 4 REV
DTAPBPE4R2	AAAGACGTGGAGCTGCAGCAC	Exon 4 REV

3 RESULTS

3.1 Antigen presentation genes in ducks

To investigate the diversity of TAP in a population of wild mallard ducks blood samples were collected and RNA was extracted. The ducks were typed for their MHC, and for the analysis of TAP, ducks with similar MHC alleles and some with very different MHC alleles were chosen. *TAP1* and *TAP2* were amplified from 12 mallards and cloned into the pCR2.1 TOPO vector (Invitrogen). Positive clones containing the gene of interest were sub-cultured and inserts were sequenced to identify single nucleotide polymorphisms and determine different alleles. Amino acid sequences were analyzed for conservative and non-conservative amino acid substitutions. Additionally a fragment across the promoter was amplified for some mallards to analyze the diversity of the bi-directional promoter and to identify TAP alleles that are linked on the same haplotype. The main goal was to elucidate notable differences between TAP alleles that could relate to differences in the specificity of the transporter and may result in constraints on antigen presentation to CD8+ cells in the natural host of influenza A virus. Sequence analysis shows that *TAP1, TAP2* and their shared promoter are diverse in wild mallard ducks. Polymorphisms were identified in functional regions that could potentially restrict the repertoire of peptides transported by TAP. Detailed sequence alignments of all obtained alleles for *TAP1*, *TAP2* and the TAP promoter are found in the Appendix A (Figure A-1 to A-5). It was unclear if the peptide loading complex in the duck functions in a

tapasin dependent or independent manner. To investigate the presence of the gene, low stringency Southern blot analysis was performed and the results suggest that *tapasin* is absent the duck genome.

3.1.1 Amplification of TAP2 from wild mallards

Previously, to investigate the diversity of *TAP2* in wild mallard ducks, the gene was amplified from mallards 18, 19, 24, 32, 42 using genomic DNA. The primer set used in this amplification was DT2E1F3/DT2E93′UR, which yielded a product from mid exon 1 to the 3′UTR of *TAP2*. Approximately 1.6kb of the coding sequence (CDS) was obtained and screened for different alleles. Mallards 18, 24, 32 and 42 were heterozygous and two distinct alleles were identified for each mallard. The primer set DT2E1F3/DT2E93′UR successfully amplified only one allele from mallard 19. Likely the primer set failed to bind to the second allele. Three clones per allele were sequenced with internal primers until completion.

New primers were designed to amplify *TAP2* by RT-PCR from cDNA to obtain the full length coding sequence of *TAP2*. Two primers were designed in the 5′UTR (DT2UTF1 and DT2UTF2) and one at the start of exon 1 (DT2E1F4). A new reverse primer in exon 9 (DT1E9R1) was also designed to attempt the amplification of the second allele of mallard 19.

Several primer sets were used to amplify *TAP2* from 7 additional mallards by RT-PCR. Most amplifications were performed with the primer set DT2UTF2/DT2E93′UR and a product of 2.129kb was successfully amplified

from mallards 9, 12, 16, 17 and 20 (Figure 3-1). The amplification with this primer set was repeated for mallards 19, 32 and 43. To obtain the second allele from mallard 16 a second primer set (DT2UTF2/DT2E9R1) was used for amplification of the gene product. This primer set yielded a product of 2.099kb (Figure 3-1). To obtain both alleles, *TAP2* was amplified from mallard 10 cDNA using two primer sets: DT2UTF1/DT2E93′UR and DT2UTF2/DT2E9R1 yielding PCR products 2.153kb and 2.099kb in size (Figure 3-1). Primer set DT2E1F4/DT2E9R1 was used to obtain the second allele from mallard 20. *TAP2* was amplified from mallard 25 using the primer set DT2UTF1/DT2E93′UR. Only one clone could be obtained for mallard 25 allele B. Maybe the first primer set used failed to amplify the second allele in high copies. Therefore several other primer sets were tested and over 50 clones were screened. All attempts to obtain more clones for allele B failed. This was also the case for mallard 19 and it was speculated that these mallards are homozygous. But TAP promoter analysis clearly shows the presence of two distinct alleles and our primers simply failed to amplify allele B or it is not expressed.

A minimum of six clones was analyzed from every mallard to obtain three clones per allele. The clones were fully sequenced until full coverage on both strands was achieved. Three clones were analyzed to avoid PCR generated artifacts. In the screening process of clones for different alleles a bias towards one dominant allele was observed for most mallards. Often the number of clones of allele A were higher compared to the less abundant second allele B. For a detailed nucleotide sequence alignment of TAP2 see Figure A-1.

The *TAP2* sequences of five mallards (18, 19, 24, 32, 42) were submitted to NCBI with accession numbers of GQ487649-GQ487655, but will be withheld from the public until submission of the remaining sequences and publication of this work.

3.1.2 Amplification of *TAP1* from wild mallards

Initially, when this work was started, only partial sequences of two *TAP1* alleles were available from one White Pekin duck (accession #AY885227.1). This *TAP1* fragment spanned from exons 1-8. Forward and reverse primers were designed at the both ends of the sequence. Later the duck genome was sequenced and a fragment from exons 2-8 of the mRNA sequence of *TAP1* was obtained. An alignment of these sequences was used to design more primers in exons 2-4 and 8 for amplifications of *TAP1*.

Several primer sets were used to amplify a fragment of *TAP1* from mallard cDNA. The first primer set DT1E2-3F1/ DT1E8R4 was used for most RT-PCRs yielding a product of 981bp (Figure 3-2). Using this primer combination *TAP1* was obtained from mallards 9, 10, 12, 16, 17, 18, 19, 20, 24 and 25. The PCR amplification was repeated with a second primer set, TAP1F2/DT1E8R4, for mallards 10, 12, 16, 17, 20 and 25. A longer *TAP1* fragment of 1.103kb starting in exon 1 was amplified. DT1E3F5/DT1E8R6 produced a 938bp amplicon from mallard 42. For amplification of *TAP1* from mallard 32 the primer set DT1E4F2/DT1E8R5 was used and yielded a product 762bp in size (Figure 3-2). The nucleotide sequences were analyzed for SNPs. Most mallards appeared to be

heterozygous. Only one clone with exons 5-8 could be obtained for the second allele with the amplifications from mallard 32. Another primer set was used but failed to produce a product altogether. Exons 1-4 from mallard 32 were obtained from the amplifications across the TAP promoter. Exons 6-11 were amplified in separate step for 32 allele B. All the fragments were assembled to one contig.

Again a minimum of six clones were analyzed for each mallard to obtain three clones per allele. All clones were sequenced until completion. For most mallards more than six clones were screened. As seen for *TAP2*, a bias towards one dominant *TAP1* allele was observed for most mallards.

After translating the first *TAP1* alleles into amino acid sequences and analysis for functional domains, it became clear that all the sequences end at the start of the Walker A motif. This and the fact that there was no stop codon in frame for the fragment obtained indicated the presence of more coding sequence. The chicken and human *TAP1* have 11 exons. The question arose whether ducks also have more than only 8 exons that were predicted by computational analysis. To investigate for the presence of further exons the unannotated duck genome was analyzed.

The unannotated genome was searched for the partial *TAP1* sequence exons 2-8 using a standalone version of the NCBI blast tool (designed by Kieran Ryan-Jean, BSc). The search yielded high similarity to scaffold 3191 and after analysis exons 2-8 were identified within the genomic sequence. Due to a region of low sequence quality (containing many Ns, where N denotes the position of the base that cannot be called) upstream of exon 2 the computational mRNA

prediction failed to identify exon 1. The sequence after the predicted exon 8 also contained Ns and is the reason why the predicted mRNA sequence ends in mid exon 8.

To identify the 3′ end of *TAP1* the scaffold was searched for an adenylation signal sequence downstream of exon 8. An adenylation signal was found located 2178 bp downstream of exon 8. In the next step, this fragment was analyzed for splice sites to identify intron exon junctions. This was done manually and by comparison to GENSCAN (Chris Burge, Stanford University). By comparison to the chicken and human TAP1 amino acid sequences exons 8 -11 and the 3′UTR to the adenylation signal were predicted in scalffold3191 (Figure 3-3). A translation of the predicted full length *TAP1* into an amino acid sequence included a stop codon in frame.

To verify the predicted exons 8-11 of duck *TAP1* primers were designed in exon 6, exon 11 or 3′ UTR. RT-PCRs were performed on cDNA from mallards 18, 19, 20, 24, 32 and 42. The expected size of the PCR product was 1249bp (Figure 3-4).

The amplifications from mallards 18 and 19 yielded a fragment of 1.114kb and the product from mallard 32 was only 1.201kb in size. All other mallards yielded the expected product size. Positive clones were analyzed for SNPs and two alleles for each mallard were identified. Three clones per allele were fully sequenced. The results confirmed the presence of exons 8-11 in the duck *TAP1* for mallards 20, 24 and 42, and exons 8-11 are as predicted from the genomic sequence indicated in grey in (Figure 3-4).

To explain the shorter fragments obtained from mallard 18, 19 and 32 the primer binding sites were analyzed. The reverse primer used in these amplifications was designed to bind at the end of exon 11 and span into the 3′ UTR, overlapping the stop codon. Surprisingly the reverse primer could bind upstream of the actual stop with a mismatch of one nucleotide changing the codon from TGG to TAG at positions 1828-1830bp creating an early stop. A second reverse primer in the 3′ UTR in combination with a forward primer in exon 9 was used to amplify a 675bp fragment from mallards 18 and 20 (Figure 3-4). The products obtained were the same size for both mallards. Sequencing of the fragments provided an exact match to the previous sequences of mallard 20. The sequences obtained from mallard 18 confirmed that the first primer caused a premature stop. The new sequences obtained provided full length sequence information of *TAP1* exons 8-11. A translation of the mallard 18B nucleotide sequence showed a premature stop codon at position 596. This was the case with fragments obtained from both amplifications. Mallard 18B has an SNP at the position 1786bp (A to T) creating a premature stop.

An alignment of the nucleotide sequences showed an unusual insertion for mallards 20A and 24B in exon 11. These alleles have an insertion of 18 nucleotides at position 1819 of the CDS. The insertion does not cause a frame shift or premature stop in the amino acid sequence. The insertions are in frame but accounts for 6 additional amino acids that have been observed for allele 20A and 24B only. All other alleles sequenced did not show any additional nucleotides in exon 11. For a detailed nucleotide sequence alignment of TAP1 see Figure A-2.

3.1.3 The structure and topology of the duck TAP

Full length nucleotide sequences of both TAP genes were assembled by aligning all fragments obtained. Contigs were translated into the amino acid sequence and further analyzed. The duck TAP1 and TAP2 proteins are 654 and 700 aa, respectively, in size. The amino acid sequences were used to predict the structure and topology of the duck peptide transporter.

Previously, 7 transmembrane segments and a cytosolic nucleotide binding domain containing Walker A and B motifs were predicted for TAP2 obtained from a White Pekin duck (Mesa et al., 2004). Hydrophobicity analysis of the wild mallard sequences confirmed seven TMs for TAP2. Furthermore, the TMs of TAP2 can be divided up into a core TMD of six TMs and an accessory TMD with one TM. The N-terminus of TAP2 is localized in the ER whereas the3-5 C-terminus stretches into the cytoplasm (Figure 3-5). TAP1 of wild mallards has only six transmembrane segments. It is composed of the core TMD but lacks the additional N-terminal TMD. Both N- and C-termini are located in the cytoplasm for TAP1. SMART analysis identified C-terminal nucleotide binding domains in both TAP1 and TAP2, which contain the ABC-transporter motifs Walker A and B, ABC signature motif, Q-loop, D-loop and H-loop/switch region.

3.2 Diversity in the duck TAP transporter

3.2.1 Amino acid differences of TAP2

An alignment of all TAP2 sequences obtained was created in MegAlign (DNAStar) using CLUSTALW. For a detailed aa sequence alignment of TAP2

see Figure A-3.The alignment shows that the protein is polymorphic in wild mallard ducks. The alleles share between 95.7-100% amino acid identities. A total number of 81 polymorphic positions were identified. Nine of the polymorphisms fall within the tapasin binding region. All of these polymorphisms are non-conservative substitutions, with the exception of P89S. R26 is substituted with Q in 8 alleles and W in 1 allele. At position 55, 13 of the TAP2 alleles have an M and 11 alleles have an R instead. In 17 alleles an A66 was observed compared to 7 alleles with P66 (Table 3-1).

In the transmembrane domain of TAP2 43 polymorphic regions were identified. Mallard 12 allele B appears to be very unusual with several unique amino acid changes compared to all other alleles. It has two unique cysteines, C229 and C446, which are conserved G229 or R446 in all other alleles. More single amino acid substitutions for this allele are: M230, R263, I264, M297 and E422 (Table 3-2). Interestingly, several substitutions to cysteines were observed in the transmembrane domain, which appeared in more than one allele. At position 155 50% of the alleles have a hydrophobic phenyalanine and the other 50% show cysteines, including the unusual M12B allele. C212 appears in mallards 16A and 42B, where all the other alleles posses a threonine or serine. Several non-conservative substitutions appear in at least two alleles at position H232Q, R259D and G263F (Table 3-2).

Twenty-nine polymorphic positions were observed within the nucleotide binding domain (Table 3-3). Four positions showed an unusual substitution to cysteines: R477C, R/H669C, H/R675C and R700C. The C669 switch was

observed in 3 cases, in which all alleles were from different mallards. C700 appears in alleles 16B and 42B compared to R in all other alleles.

The consensus of the TAP2 mallard sequences was aligned with chicken and human TAP2 (Figure 3-6). Interestingly, many polymorphisms are localized in two regions, which have been shown to be involved in peptide binding in mammals: L299P, A300T, V302A, R304H, I306V, H311Y, M327V, A341P, A343T, G353S, G361R, K370T, I376V, R377Q, D415N, A422E, G428S (Table 3-3). These residues are located in the ER luminal loop between TM5 and 6 of the TAP2 core TMD and TM6 spanning into to the C-terminal tail (Figure 3-7).

Four polymorphic residues of the duck TAP2 align to the positions, that were identified to alter TAP specificity in rats: R259D, P262S, G263F/R and R377Q. In all four cases the amino acids were substituted with amino acid from a different biochemical group. The switch to D259 was observed for mallard 20B and 25B. G263 appeared in 15 alleles, 7 alleles have a phenyalanine instead and mallard 12B has an arginine. P262S and R377Q substitutions appeared only in mallard 12B and 16B, respectively. The locations of these polymorphic residues are in the cytosolic loop between TM4 and 5 of and ER luminal loop between TM5 and 6 of the TAP2 core transmembrane domain (Figure 3-7).

3.2.2 Amino Acid difference of TAP1

The sequences obtained for all TAP1 alleles were aligned in MegAlign (DNAStar) using CLUSTALW. For a detailed aa sequence alignment of TAP1 see Figure A-4.The TAP1 amino acid sequences share 94.1%-100% sequence

similarities. Thirty-two polymorphic positions were identified within TAP1, of which 17 fall within the TMD (Table 3-4) and 15 are found in the NBD (Table 3-5). Twenty-four of the polymorphisms are non-conservative substitutions. In 23 alleles a threonine was observed at position 79 within the TMD. A unique substitution occurred in allele 19B substituting threonine with a methionine resulting in a change of the biochemical group from polar but uncharged to hydrophobic. A unique allele was recorded for mallard 16A with several unique amino acid substitutions: T116M, M119V, A125D, E128K and R153H. These residues are conserved between all other alleles analyzed (Table 3-4). An unusual cysteine was recorded for mallard 16B at the position 355. R355 is conserved between the remaining alleles. C442 appears in the nucleotide binding regions of three alleles, compared to histidine in two alleles and arginine in seven alleles (Table 3-5). R445 is replaced with Q in mallards 19B and 20B. Three other positions were observed with non-conservative amino acid substitution. G593R and E600K appear in more than one allele. In 20 alleles an alanine was recorded at position 368, whereas four alleles posses a valine. This polymorphic position is localized in the Walker A motif of ABC transporter domain (Table 3-5).

The consensus sequence of all mallard TAP1 sequences was aligned to the human and chicken TAP1 sequences (Figure 3-8). In chickens, TAP1 exon 1 is truncated, which led to the loss of the tapasin binding site. Similarly, TAP1 in the duck lacks the tapasin binding region due to a truncation. Two polymorphic residues of TAP1, Q223K and D231N, were identified in a region that aligned with the peptide binding regions of the mammalian TAP1. They are localized in

the cytosolic loop between TM4 and 5 (Figure 3-7). Interestingly, the residue Q223, which is a polar but uncharged aa, appears in 14 of the alleles, whereas the remaining 10 alleles have a positively charged K at the same position. Residue D231, which 19 of the alleles posses, belongs to the group of negatively charged amino acids and the other 5 alleles posses a polar but uncharged N at the same position. Two polymorphic positions in the chicken TAP1, which match the binding specificity of the major MHC class I, do not align with polymorphic positions in duck TAP1 (Figure 3-7).

3.2.3 Phylogenetic analysis of TAP alleles

To investigate the distance of the TAP1 alleles a phylogenetic tree was constructed in MegAlign (DNAStar) for both TAP1 and TAP2. Only partial TAP1 amino acid sequences were used for the analysis. The longest fragments of TAP1 and TAP2 obtained for all alleles were aligned to create the phylogenetic tree. A 697 aa fragment was used for TAP2. Mallard 19B was excluded from the analysis, because only partial exon 1 sequence was obtained. The alleles show 95.7-100% sequence identities (Figure 3-9). Two pairs, mallards 10B-19A and 16A-42B, were identified to share 100% amino acid similarities between the two alleles. To construct a phylogenetic tree for TAP1 a 327 aa long fragment was used, which spans from TM2 to the Walker A motif. For mallard 32 allele B only one allele was obtained for the fragment covering exons 5-8. The sequence quality of this fragment was not 100% and therefore it was excluded from the analysis. The analysis of the other 23 alleles revealed 100% identical pairs(identical for this

fragment of the amino acid sequence): mallards 17B-25A, 17A-20A and 25B-42A (Figure 3-10). By amplifying a fragment across the promoter with gene specific primers in *TAP1* and *TAP2,* linked TAP alleles could be identified. This was performed for mallards 18, 19, 20, 24, 32 and 42. TAP1 and TAP2 alleles A are in the same haplotype and alleles B represent the second haplotype. The phylogenetic trees suggest that the same TAP1 and TAP2 are not linked in a haplotype. Furthermore different TAP2 and TAP1 alleles pair together. Mallards 17A and 20A have identical TAP1 alleles, whereas the TAP2 alleles share 98.4% amino acid identities. TAP2 20A shares greater amino acid identities with 32B. The for TAP2 identical alleles from mallards 16A and 42B fall into different branches of the TAP1 phylogenetic tree suggesting the alleles are significantly different.

3.3 Amplification and sequencing of the TAP promoter of six ducks

To investigate the diversity of the TAP promoter in mallards, gene specific primers in *TAP1* and *TAP2* were used to amplify a fragment of approximately 2.9kb across the promoter (Figure 3-11). The promoter sequences from mallards show between 95.8 to 99.7% nucleotide identities. A detailed alignment of the consensus sequences obtained for each allele can be found in Figure A-5. The consensus of the mallard promoter sequences share 97.1% and 56.8% identities with duck D26 and chicken, respectively. Using TESS, control regions were identified within the duck promoter. The binding sites for interferon-stimulated response element (ISRE) and IFN γ-activated site (GAS) were found at positions

306 and 67, respectively. Binding sites for Y box were also identified in the duck TAP promoter (Figure 3-12). Overall the mallard duck promoter appears to be intact, despite many polymorphisms, and seems to be induced by interferons. The presence of only one ISRE indicates that the promoter is bidirectional regulating both, *TAP1* and *TAP2*.

3.4 Expression analysis of TAP

To elucidate the expression profile of TAP during influenza infection, first RT-PCR was performed on cDNA samples from mock, low and highly pathogenic avian influenza infected lung tissues from one and three dpi. Gene specific primers were used to amplify a fragment of 152bp and 206bp from *TAP1* and *TAP2*, respectively. The amplification of *TAP1* and *TAP2* was carried out for 30 cycles and 25 cycles, respectively. At 1 dpi both *TAP1* and *TAP2* appear to be upregulated in VN1203 infected lung samples. In lung tissues infected with BC500 TAP appears to be less expressed and low expression was detected in mock (PBS only) infected samples (Figure 3-13). Three days post infection *TAP1* and *TAP2* expression appears to be equally low in all samples, indicating TAP is not upregulated (Figure 3-14).

To validate these results the exact fold-increase in expression of TAP transcripts was measured by qPCR with gene specific probe and primer sets. Results show a 36.5 fold upregulation of *TAP1* in VN1203 samples at 1dpi compared to 6.0 fold upregulation in BC500 infected sample and only 2.3 fold increase in mock infected lung tissue. A 22.0 fold increase in expression was

recorded for *TAP2* in VN1203 infected lung tissue, whereas BC500 and mock infection caused only an increase of 8.1 and 4.0 in fold expression, respectively.

At 3dpi the expression levels are equally low for both TAPs in all samples. *TAP1* expression ranges from 1.2, 1.1 and 1.4 fold for mock, BC500 and VN1203 and *TAP2* ranges from 0.9, 0.8 and 1.5 fold (Figure 3-15).

3.5 PCR amplification of tapasin

In an attempt to obtain the fragment of *tapasin* from ducks, several primers were designed in conserved regions of the chicken gene. Several primer combinations were tested on cDNA or genomic DNA. While *tapasin* could be amplified from several galliformes, all attempts failed to obtain a fragment of the gene from a White Pekin duck or wild mallards. The primer set TAPBPE5F1/TAPBPE6R1 yielded clear bands for amplifications from chicken, and turkey, pheasant genomic DNA and a faint band from pigeon, but no product for PCRs on mallard DNA (Figure 3-16). The obtained products were extracted and sequenced. Bands from all galliformes verified the amplification of *tapasin*. Attempts to sequence the product obtained from pigeons failed. Several other primer sets were used and products obtained for ducks were sequenced. None of these products had sequence homology to *tapasin*. Therefore all attempts to PCR amplify *tapasin* from domestic or wild duck DNA failed.

3.6 Southern blot hybridization

To investigate the presence of *tapasin* in the duck genome two separate low stringency Southern blots were performed. The probe used in all hybridizations was amplified with the primer set TAPBPE5F1/TAPBPE6R1 from chicken cDNA yielding a product 269bp in size. For the first blot, duck and chicken DNA was digested with BglII, HaeIII and PstI. The blot was hybridized at low stringency temperatures of 55°C or 60°C overnight and the blot was exposed for four hours. The probe hybridized to distinct bands in the chicken PstI digest indicating the presence of *tapasin*. No hybridizing bands were detected in any of the duck digests suggesting *tapasin* is absent from the duck genome (Figure 3-17). A second blot was prepared to compare these results with other avian species. Duck, chicken, turkey, pheasant and pigeon genomic DNA was PstI digested and blotted onto a nylon membrane. The blot was hybridized with the same tapasin probe obtained from chicken cDNA at 55°C or 60°C overnight followed by a 4 hour exposure. The blot hybridized at 55°C shows many bands, which disappear in the hybridization at 60°C suggesting they are background. Distinct bands, which were previously seen in the chicken can be identified in the chicken, turkey and pheasant samples but are missing in the duck and pigeon indicating again the absence of *tapasin* in the duck genome. The probe hybridized to a band of 20kb in the duck genome and the identity has yet to be determined (Figure 3-18). Most likely this is non-specific hybridization to undigested DNA at the top of the gel.

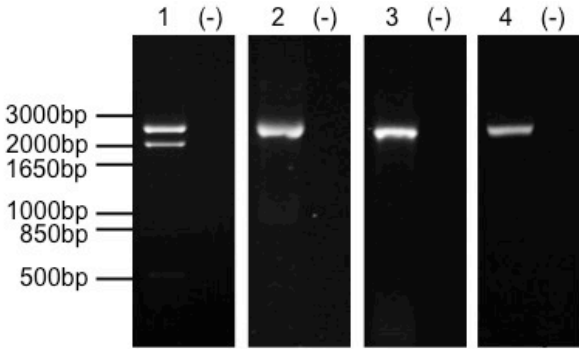

Figure 3-1. PCR products for *TAP2* primer sets. Several primer sets were used for TAP2 amplification. The products obtained and the negative control (-) for each primer set is shown. The bands of expected size are the brightest bands. The numbered lanes represent the following primer sets: (1) DT2UTF1/DT2E93′UR, (2) DT2UTF2/DT2E93′UR, (3) DT2UTF2/DT2E9R1, (4) DT2E1F4/DT2E9

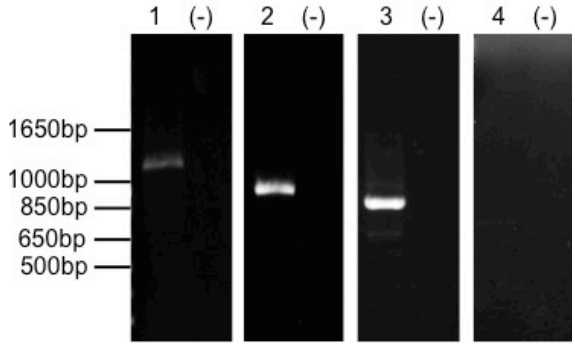

Figure 3-2. PCR products for *TAP1* primer sets. Shown are the products obtained for TAP1 amplifications and the negative control (-) for each set. The bands of expected size are the brightest bands. The numbered lanes represent the following primer sets: (1) TAP1F2/DT1E8R4, (2) DT1E2-3F1/DT1E8R4, (3) DT1E3F5/ DT1E8R6, (4) DT1E4F2/DT1E8R5.

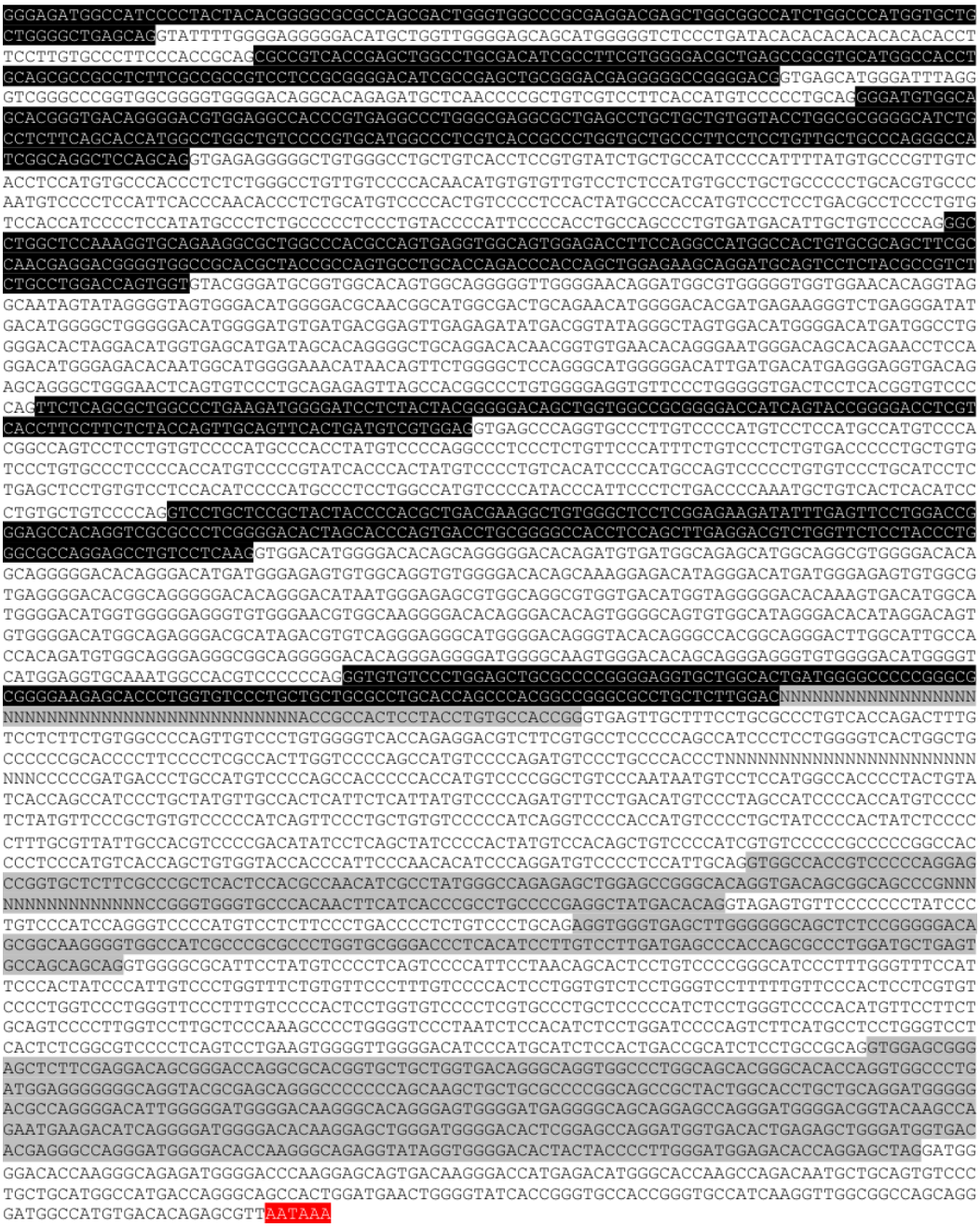

Figure 3-3. Genomic sequence of *TAP1*. Indicated in black are the exons 2 to mid exon 8 that were previously identified. Shown in red is the adenylation signal. In grey are exons 8-11 that were predicted by identification of splice sites and verified by sequencing the fragment from 6 mallards.

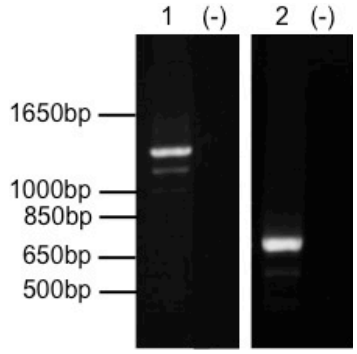

Figure 3-4. PCR products for *TAP1* exon 6-11 amplification. The products obtained for the TAP1 fragment are shown and the negative control (-) for each primer set. The bands of expected size are the brightest bands. The numbered lanes represent the following primer sets: (1) DT1E6F3/ DT1E11-3′UR, (2) DT1E9F1/DT13′UR.

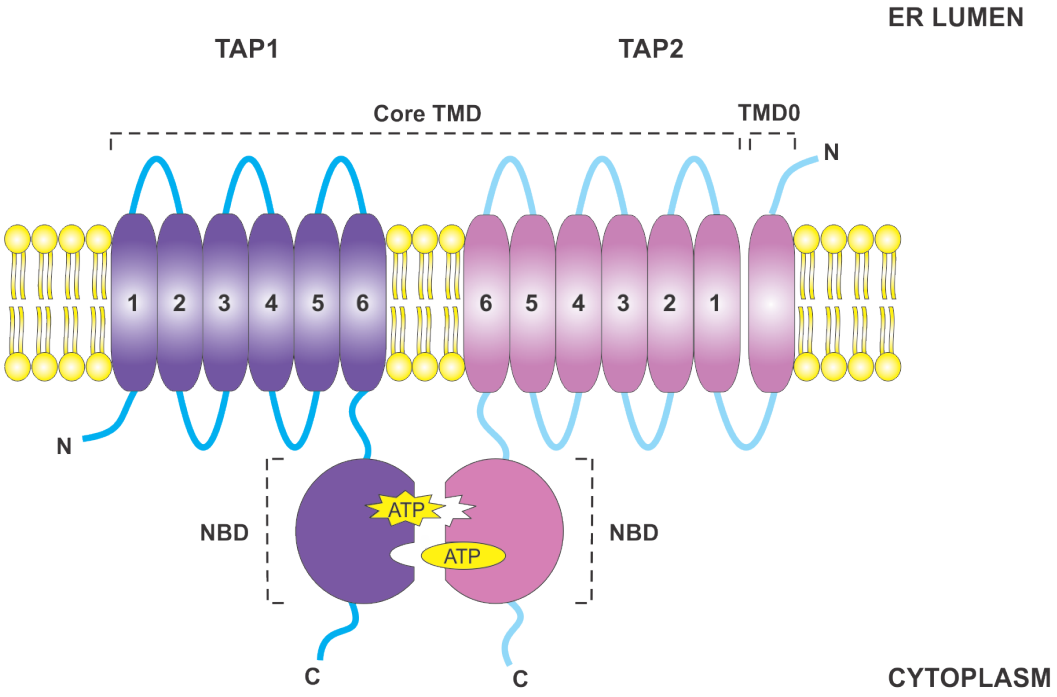

Figure 3-5. Structure of the duck TAP. Shown in purple and pink is the duck *TAP1* and *TAP2* heterodimer. TAP consists of the core transmembrane domain with twelve transmembrane segments and an accessory transmembrane domain in *TAP2* with one transmembrane segment. The location of the N- and C-termini of both TAP subunits are indicated. TMD, transmembrane domain; NBD, nucleotide binding domain; TMD0, accessory transmembrane domain.

Table 3-1. Polymorphism in the tapasin binding domain of TAP2. The amino acids are colored according to their biochemical group: red-hydrophobic, magenta-positively charged, blue-negatively charged, green-polar but uncharged, orange-special case.

	26	55	58	60	63	66	86	89	98
9A	R	R	A	G	G	A	G	P	L
9B	Q	R	A	G	G	A	G	P	P
10A	R	R	D	G	G	A	G	P	P
10B	R	R	A	G	G	P	R	P	P
12A	R	M	A	G	Q	A	G	P	P
12B	R	M	A	G	G	A	G	P	P
16A	R	R	A	G	G	P	G	P	P
16B	Q	M	A	G	G	A	G	S	P
17A	Q	M	A	G	G	A	G	P	P
17B	R	R	A	G	G	P	G	P	P
18A	R	M	A	G	G	A	G	P	P
18B	Q	M	A	G	G	A	G	P	P
19A	R	R	A	G	G	P	R	P	P
19B	R	R	A	G	G	P	G	P	P
20A	Q	R	A	G	G	A	G	P	P
20B	R	M	A	G	G	A	G	P	P
24A	R	R	A	G	G	P	G	P	P
24B	Q	M	A	G	G	A	G	P	P
25A	R	M	A	G	G	A	G	P	L
25B	W	M	A	R	G	A	G	P	P
32A	R	M	A	G	G	A	G	P	P
32B	Q	M	A	G	G	A	G	P	P
42A	Q	R	A	G	G	A	G	P	P
42B	R	M	A	G	G	P	G	P	P

Table 3-2. Polymorphism in the transmembrane domain of TAP2 (continued).

	133	148	155	176	182	188	193	204	209	212	215	229	230	232	243	259	262	263	264	265	269	297
9A	A	F	C	I	G	T	V	A	G	T	Q	G	L	H	A	R	P	G	S	A	L	V
9B	A	F	F	V	G	T	V	A	G	S	R	G	L	H	A	R	P	G	S	A	L	V
10A	A	L	F	V	G	T	V	A	S	T	R	G	L	H	A	R	P	F	S	A	L	V
10B	A	F	F	V	G	T	V	S	G	T	R	G	L	Q	A	R	P	F	S	A	L	V
12A	A	F	F	V	G	T	V	A	G	S	R	G	L	H	A	R	P	G	S	A	L	V
12B	A	F	C	V	G	T	V	A	G	T	R	C	M	H	A	R	S	R	I	A	P	M
16A	A	F	C	V	G	T	V	A	G	C	R	G	L	H	A	R	P	G	S	A	L	V
16B	A	F	C	V	G	T	V	A	G	T	R	G	L	H	A	R	P	G	S	A	L	V
17A	A	F	C	V	G	T	V	A	S	T	R	G	L	H	A	R	P	F	S	V	L	V
17B	A	F	F	V	G	T	M	A	G	T	Q	G	L	H	A	R	P	G	S	A	L	V
18A	A	F	C	V	G	T	V	A	S	T	R	G	L	H	A	R	P	F	S	A	L	V
18B	E	F	C	V	S	T	V	A	S	T	R	G	L	H	A	R	P	F	S	A	L	V
19A	A	F	F	V	G	T	V	S	G	T	R	G	L	Q	A	R	P	F	S	A	L	V
19B	A	F	F	I	G	T																
20A	A	F	F	V	G	T	V	A	G	S	Q	G	L	H	A	R	P	G	S	A	L	V
20B	A	F	C	V	S	T	V	A	G	T	Q	G	L	H	A	D	P	G	S	A	L	V
24A	A	F	F	V	G	T	V	A	G	T	R	G	L	H	A	R	P	F	S	V	L	V
24B	A	F	F	V	G	I	V	A	G	S	R	G	L	H	A	R	P	G	S	A	L	V
25A	A	F	C	I	G	T	V	A	G	T	Q	G	L	H	A	R	P	G	S	A	L	V
25B	A	F	C	V	G	T	V	A	G	T	Q	G	L	H	A	D	P	G	S	A	L	V
32A	A	F	F	V	G	T	M	A	G	T	R	G	L	H	T	R	P	G	S	A	L	V
32B	A	F	F	V	G	T	V	A	G	S	Q	G	L	H	A	R	P	G	S	A	L	V
42A	E	F	C	V	S	T	V	A	S	T	R	G	L	H	A	R	P	G	S	A	L	V
42B	A	F	C	V	G	T	V	A	G	C	R	G	L	H	A	R	P	G	S	A	L	V

Table 3-2. Polymorphism in the transmembrane domain of TAP2. The amino acids are colored according to their biochemical group: red-hydrophobic, magenta-positively charged, blue-negatively charged, green-polar but uncharged, orange-special case. The sequence information past the indicated position was not obtained for 19B.

	299	300	302	304	306	311	327	341	343	353	361	370	376	377	397	407	415	422	428	446	449
9A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
9B	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	V
10A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	H	I	D	A	G	R	A
10B	L	A	A	H	V	H	M	P	A	G	G	T	V	R	H	I	D	A	G	R	A
12A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
12B	L	A	A	H	V	Y	V	A	A	G	R	K	I	R	R	I	N	E	G	C	A
16A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
16B	L	A	A	H	V	H	M	A	A	G	G	K	I	Q	H	I	D	A	G	R	A
17A	L	A	V	R	I	Y	M	A	T	G	G	K	I	R	H	I	D	A	G	R	A
17B	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	S	R	V
18A	L	T	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
18B	L	A	A	H	V	H	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
19A	L	A	A	H	V	H	M	P	A	G	G	T	V	R	H	I	D	A	G	R	A
19B																					
20A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
20B	L	A	V	H	I	Y	V	A	A	S	G	K	I	R	H	I	D	A	G	R	A
24A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	I	N	A	G	R	A
24B	P	A	V	R	I	Y	M	A	A	G	G	K	I	R	H	I	D	A	G	R	A
25A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
25B	L	A	V	R	I	Y	V	A	A	G	G	K	I	R	R	V	N	A	G	R	A
32A	L	A	A	H	V	H	M	A	A	G	G	T	V	R	H	I	D	A	G	R	A
32B	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	H	V
42A	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A
42B	L	A	V	R	I	Y	M	A	A	G	G	K	I	R	R	V	N	A	G	R	A

Table 3-3. Polymorphism in the nucleotide binding domain of TAP2. The amino acids are colored according to their biochemical group: red-hydrophobic, magenta-positively charged, blue-negatively charged, green-polar but uncharged, orange-special case.

	456	460	464	477	478	479	481	535	537	541	558	569	574	579	594
9A	S	T	H	R	P	K	L	K	R	R	R	R	I	A	D
9B	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
10A	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
10B	S	T	H	R	P	E	L	K	R	R	W	R	I	A	D
12A	S	T	R	R	P	E	L	K	R	R	R	R	I	A	N
12B	L	T	H	R	L	E	L	K	L	R	R	R	I	A	D
16A	S	T	H	R	P	E	L	K	R	R	R	R	V	A	D
16B	L	T	H	R	P	E	L	K	R	R	R	R	I	A	D
17A	S	T	H	R	P	E	L	K	R	R	R	R	I	S	D
17B	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
18A	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
18B	L	T	H	R	P	E	L	K	R	R	R	R	I	A	D
19A	S	T	H	R	P	E	L	K	R	R	W	R	I	A	D
19B															
20A	S	T	H	C	P	K	V	K	R	R	R	R	I	A	D
20B	L	I	H	R	P	E	L	K	R	R	R	R	I	A	D
24A	S	T	H	R	P	E	L	K	R	H	R	R	I	A	D
24B	L	T	H	R	P	E	L	K	R	R	R	R	I	A	D
25A	S	T	H	R	P	E	L	R	R	R	R	R	I	A	D
25B	S	T	H	R	P	E	L	R	R	R	R	R	I	A	D
32A	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
32B	S	T	H	R	P	E	L	K	R	R	R	R	I	A	D
42A	S	T	H	R	P	E	L	K	R	C	R	T	I	A	D
42B	S	T	H	R	P	E	L	K	R	R	R	R	V	A	D

	596	623	631	641	648	649	655	669	670	671	675	681	688	700
9A	D	T	T	P	S	G	L	R	V	L	H	M	A	R
9B	D	A	T	A	S	G	L	R	V	L	H	M	A	R
10A	D	T	T	P	S	G	L	C	V	L	H	M	A	R
10B	D	T	N	P	S	A	L	R	V	L	H	M	A	
12A	D	T	T	A	S	G	M	R	V	M	H	M	A	R
12B	D	T	T	P	S	G	L	R	V	L	R	M	A	R
16A	D	T	T	P	T	G	L	R	V	L	H	M	A	
16B	D	T	T	P	S	G	L	R	V	L	R	M	A	C
17A	D	T	T	P	S	G	L	R	V	L	H	M	A	R
17B	D	T	T	P	S	G	L	C	V	L	H	M	S	R
18A	D	T	T	A	S	G	L	R	I	L	H	M	A	R
18B	D	T	T	P	S	G	L	C	V	L	H	M	A	R
19A	D	T	N	P	S	A	L	R	V	L	H	M	A	R
19B														
20A	D	T	T	P	S	G	L	R	V	L	H	M	A	R
20B	D	T	T	A	S	G	M	R	V	M	C	I	A	
24A	D	T	T	P	S	G	L	R	V	L	H	M	A	R
24B	D	T	T	P	S	G	L	H	V	L	R	M	A	R
25A	Y	A	T	A	S	G	L	R	V	L	H	M	A	R
25B	Y	A	T	A	S	G	L	R	V	L	H	M	A	R
32A	D	T	T	P	S	G	L	R	V	L	H	M	A	R
32B	D	T	T	P	S	G	L	R	V	L	H	M	A	R
42A	D	T	T	A	S	G	L	R	V	L	H	M	A	R
42B	D	T	T	P	T	G	L	R	V	L	H	M	A	C

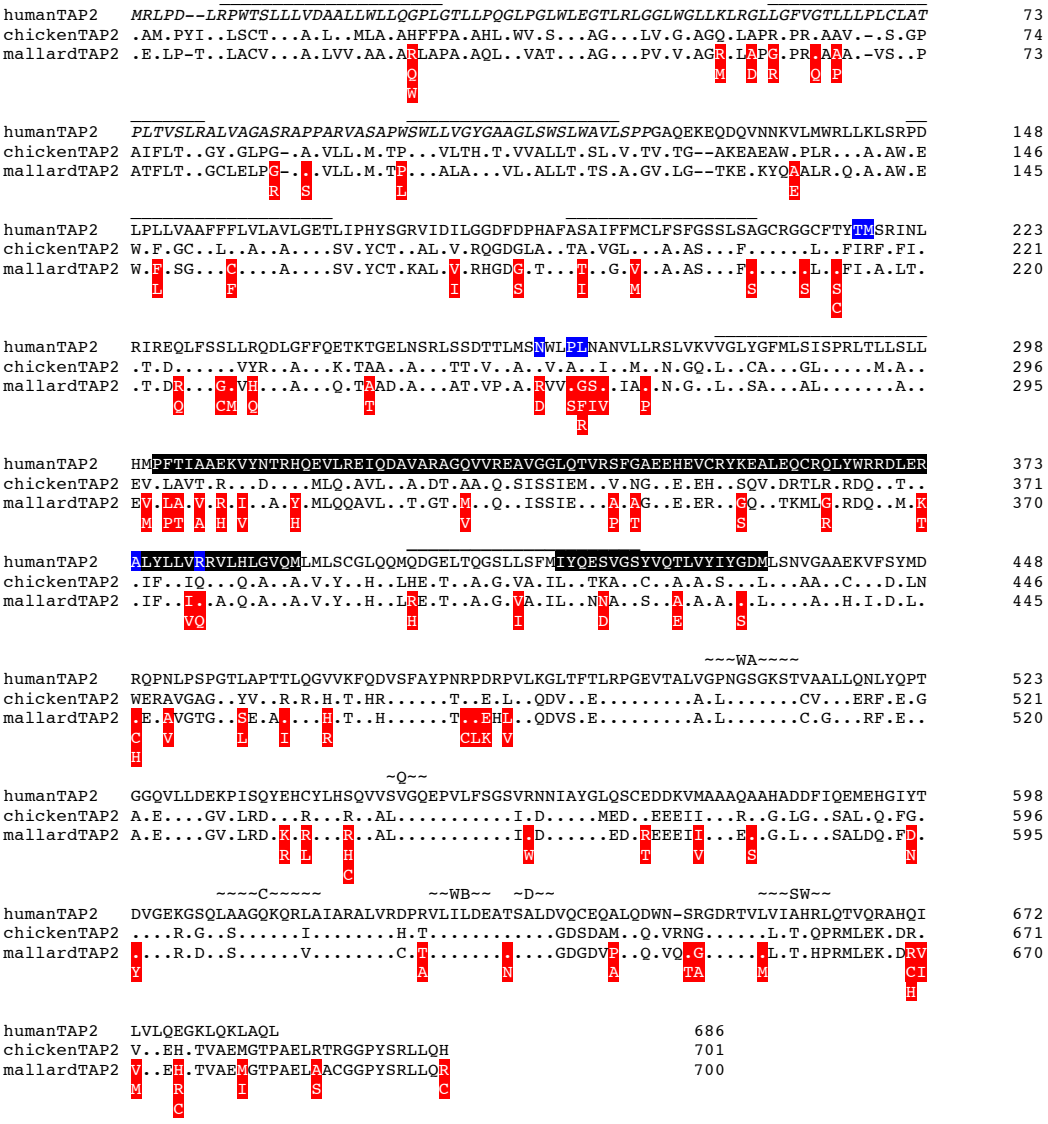

Figure 3-6. Alignment of TAP2 from different species. The TAP2 consensus of the mallard sequences was aligned with the chicken and human TAP2. Dots represent matches to the human sequences, dashes are gaps to optimize the alignments. Italicized letters of the human TAP2 indicate the tapasin binding region. Overlined are the predicted transmembrane segments of the duck TAP2. The ABC motifs are indicated with tildes: WA, Walker A; Q, Q-loop; C, signature motif; WB, Walker B; D, D-loop; SW, switch region/ H-loop. TAP2 polymorphisms are highlighted in red. Sequence in black indicates the peptide binding region in mammals and in blue are the rat polymorphisms, that change TAP specificity.

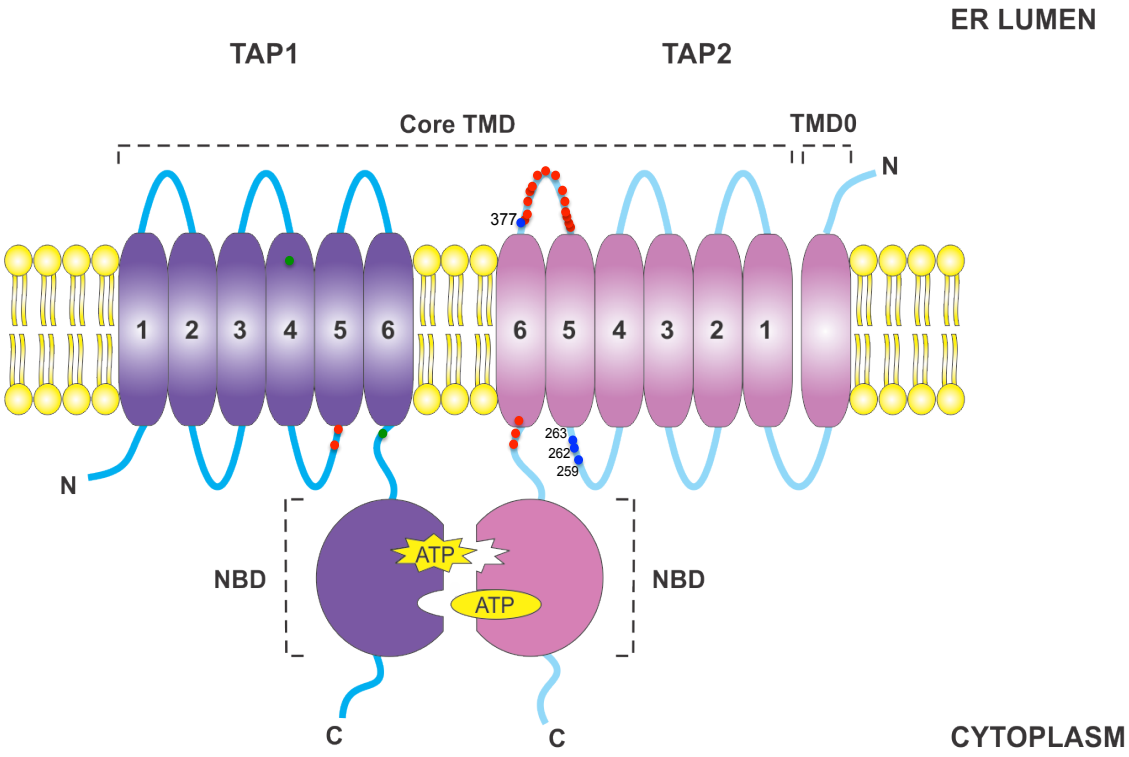

Figure 3-7. Location of polymorphisms within the PBD of duck TAP. Red dots indicate the approximate positions of polymorphisms observed in the duck TAP, which align to the peptide binding region of the mammalian TAP. In blue are the positions within the duck TAP2 that align to polymorphisms shown to alter TAP specificity in rats. Two positions of the chicken TAP1 are indicated in green, which match the binding specificity of the major MHC class I of the same haplotype.

Table 3-4. Polymorphism in the transmembrane domain of TAP1. The amino acids are colored according to their biochemical group: red-hydrophobic, magenta-positively charged, blue-negatively charged, green-polar but uncharged, orange-special case. The first parts of the sequences could not be obtained for some alleles and are represented as gaps.

	19	48	67	79	85	93	95	101	116	119	125	128	153	223	231	324	355
9A			V	T	G	A	V	A	T	V	A	E	R	Q	D	S	R
9B			V	T	S	A	V	A	T	V	A	E	R	Q	D	S	R
10A		H	I	T	G	A	V	A	T	V	A	E	R	Q	D	S	R
10B		R	V	T	G	A	V	A	T	V	A	E	R	Q	D	S	R
12A		R	V	T	G	A	I	A	T	V	A	E	R	K	D	S	R
12B		R	V	T	G	A	V	A	T	V	A	E	R	Q	N	S	R
16A		R	V	T	G	A	V	A	M	M	D	K	H	Q	D	S	R
16B		R	V	T	G	A	V	A	T	V	A	E	R	Q	N	S	C
17A		R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
17B		R	I	T	G	A	V	A	T	V	A	E	R	K	D	S	R
18A	A	R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
18B	V	R	V	T	G	A	V	A	T	V	A	E	R	Q	N	S	R
19A	A	R	V	T	G	A	V	A	T	V	A	E	R	Q	N	S	R
19B	V	R	V	M	G	A	V	A	T	V	A	E	R	Q	N	S	R
20A		R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
20B	V	R	I	T	G	A	V	A	T	V	A	E	R	Q	D	R	R
24A	A	R	I	T	G	A	V	A	T	V	A	E	R	K	D	S	R
24B	A	R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
25A		H	I	T	G	A	V	A	T	V	A	E	R	K	D	S	R
25B		R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
32A	A	R	V	T	G	A	V	T	T	V	A	E	R	Q	D	S	R
32B	V	R	V	T	G	A	V	A	T	V	A	E	R	Q	D	S	R
42A	A	R	V	T	G	A	V	A	T	V	A	E	R	K	D	S	R
42B	V	R	V	T	G	G	V	A	T	V	A	E	R	Q	D	S	R

Table 3-5. Polymorphism in the nucleotide binding domain of TAP1. The amino acids are colored according to their biochemical group: red-hydrophobic, magenta-positively charged, blue-negatively charged, green-polar but uncharged, orange-special case. Most of the sequences obtained from mallards are only partial TAP1.

	368	392	406	442	445	593	594	597	600	606	611	635	641	646	652
9A	A														
9B	A														
10A	A														
10B	A														
12A	V														
12B	A														
16A	A														
16B	A														
17A	A														
17B	A														
18A	V	S	R	R	R	G	T	A	K	D	D	T	G	T	T
18B	A	G	R	C	R	G	T	A							
19A	A	G	R	R	R	G	T	A	K	D					
19B	A	G	Q	H	Q	G	T	A	E	D					
20A	A	G	R	C	R	R	T	A	E	E	D	T	G	I	T
20B	A	G	R	H	Q	G	T	A	K	D	D	T	D	T	T
24A	V	G	R	C	R	G	T	A	E	D	D	T	G	T	T
24B	V	G	R	R	R	R	M	A	E	E	G	T	G	I	S
25A	A														
25B	A														
32A	A	G	R	R	R	G	T	A	E	D	-	A	G	T	T
32B	A	G	R	R	R	G	T	A	E	D					
42A	A	G	R	R	R	G	T	G	E	D	D	T	G	T	T
42B	A	G	R	R	R	G	T	A	K	D	D	T	G	T	T

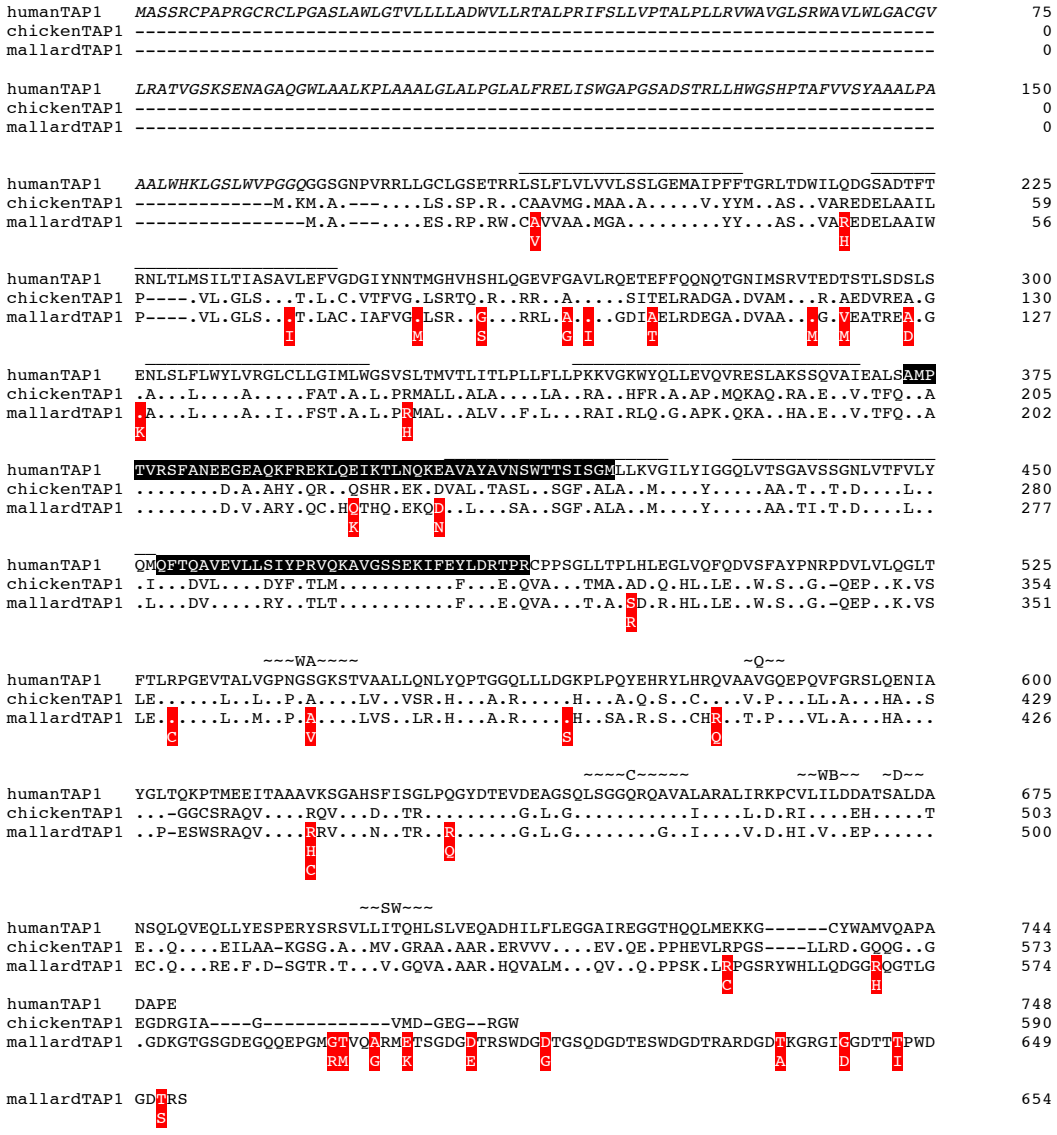

Figure 3-8. TAP1 sequence alignment. The consensus of the TAP1 mallard sequences was aligned with the chicken and human TAP1. Dots represent matches to the human sequences, dashes are gaps to optimize the alignments. Letters in italics in the human TAP1 represent the tapasin binding site. Overlined are the predicted duck TAP1 transmembrane segments. The ABC motifs are indicated with tildes: WA, Walker A; Q, Q-loop; C, signature motif; WB, Walker B; D, D-loop; SW, switch region/ H-loop. The position of polymorphisms are highlighted in red. Sequence in black indicates the peptide binding region in mammals.

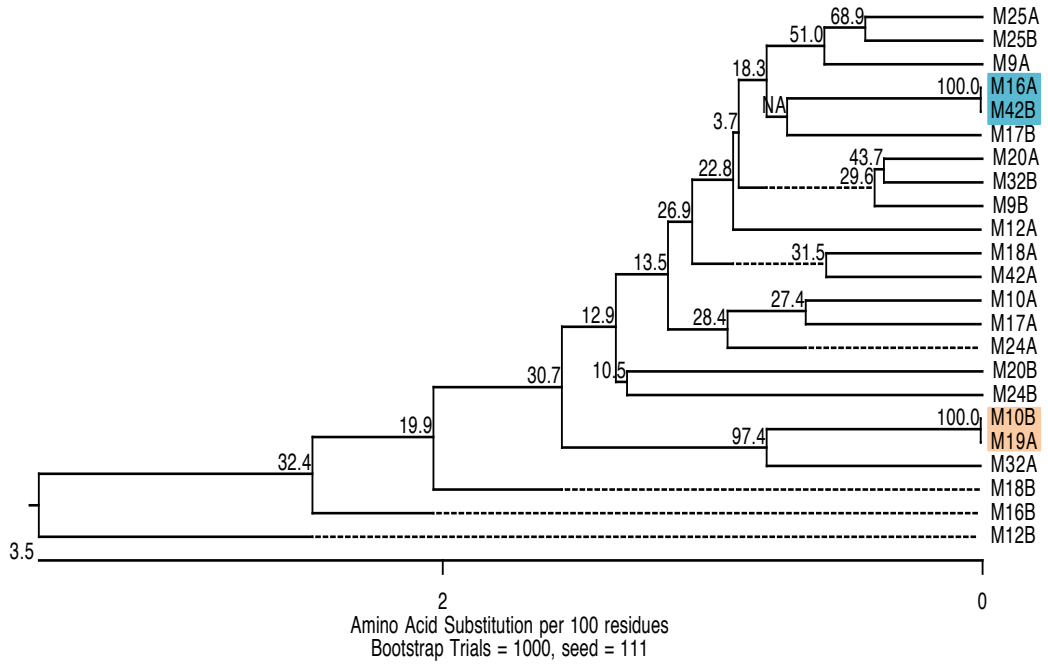

Figure 3-9. Phylogenetic tree of the TAP2 sequences. The amino acid sequences of all mallard TAP2 fragments were aligned and a phylogenetic tree was created using MegAlign by DNAStar. The pairs of alleles, which are 100% similar, are indicated by colored boxes.

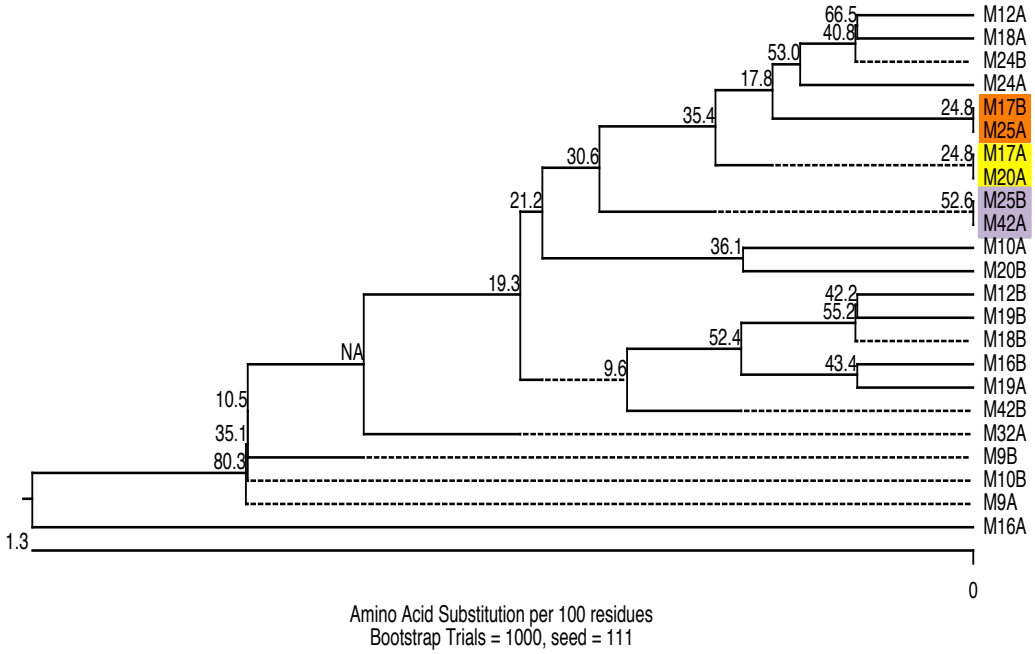

Figure 3-10. Phylogenetic tree of the TAP1 sequences. The amino acid sequences of all mallard TAP1 fragments were aligned and a phylogenetic tree was created using MegAlign by DNAStar. The pairs of alleles, which are 100% similar, are indicated by colored boxes.

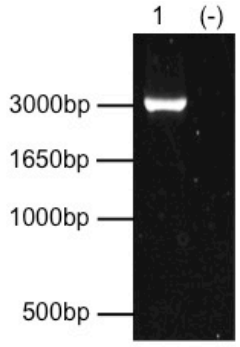

Figure 3-11. PCR product of the TAP promoter fragment. *TAP1* and *TAP2* specific primers were used to amplify across the promoter. The lanes represent the TAP promoter fragment (1) and the negative control.

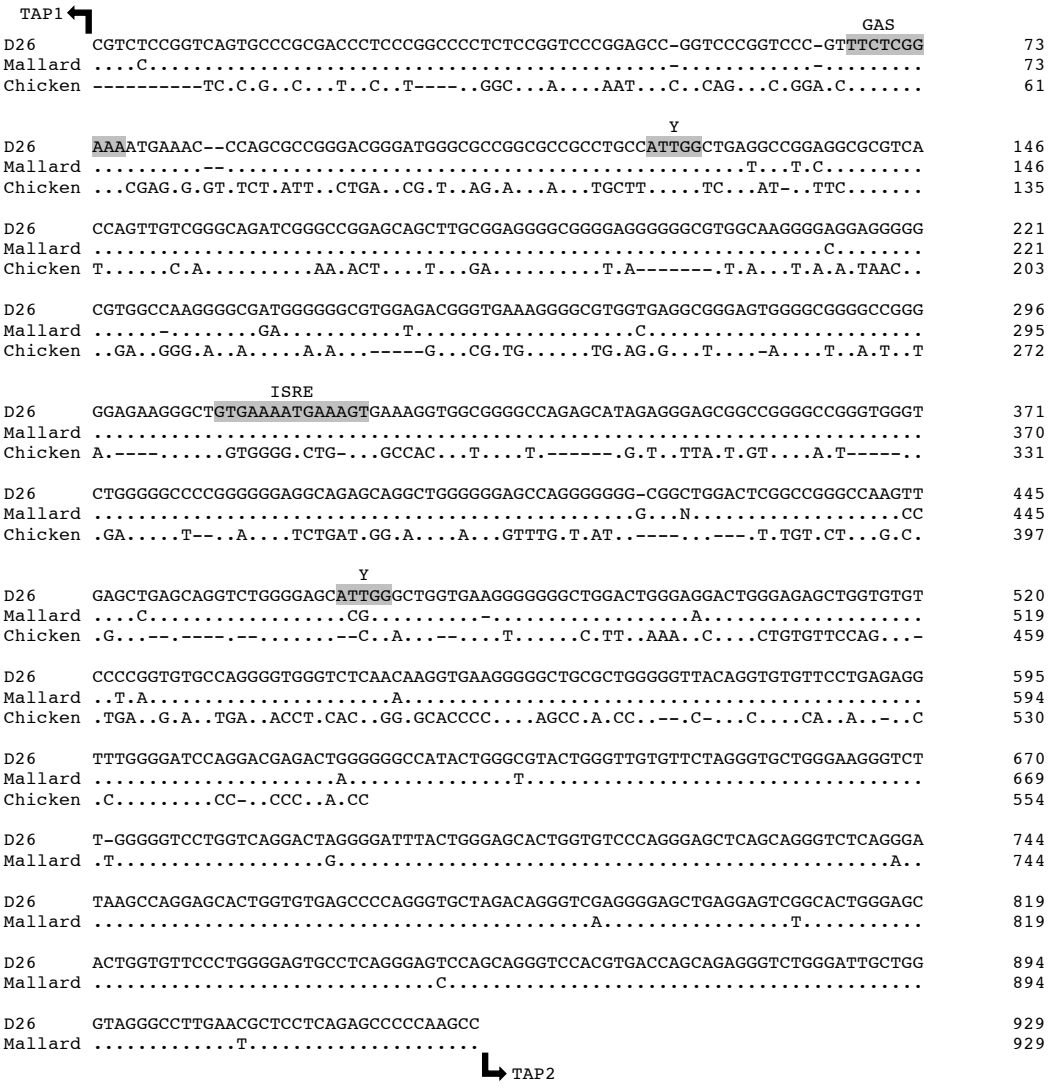

Figure 3-12. The polymorphic TAP promoter. The consensus of the mallard TAP promoter sequences was aligned with the promoter of a White Pekin duck (D26) and the chicken promoter. *TAP1* and *TAP2* start site are indicated by arrows. Transcription factor binding sites for GAS, ISRE and Y are shaded in grey and labelled. GAS, IFN γ-activated site; ISRE, interferon-stimulated response element; Y, Y box.

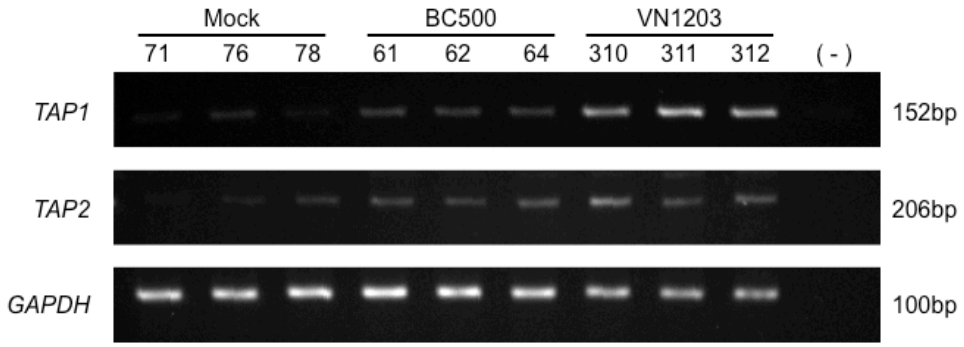

Figure 3-13. RT-PCR of TAP in lung tissues at 1 dpi. Gene expression of *TAP1* and *TAP2* was measured by RT-PCR in lung tissue from mock, BC500 and VN1203 infected ducks. *TAP1* was amplified for 30 cycles and *TAP2* for 25 cycles. Amplification of *GAPDH* was used as a positive control.

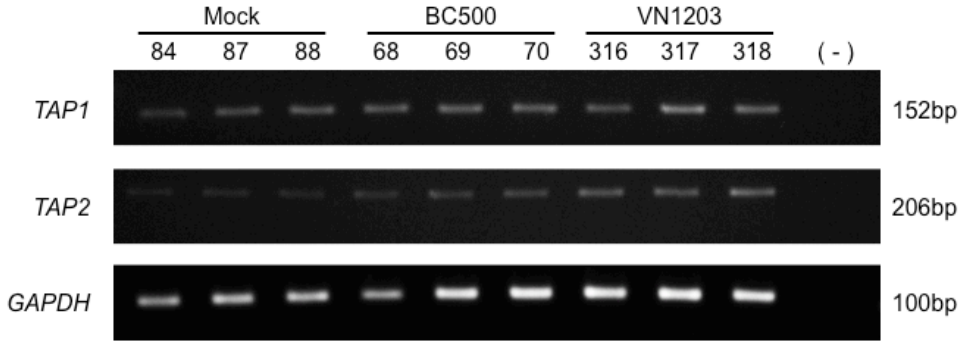

Figure 3-14. RT-PCR of TAP in lung tissues at 3dpi. The gene expression of *TAP1* and *TAP2* was measured by RT-PCR in lung tissues from mock, BC500 and VN1203 infected ducks. *TAP1* was amplified for 30 cycles and *TAP2* for 25 cycles. Amplification of *GAPDH* was used as a positive control.

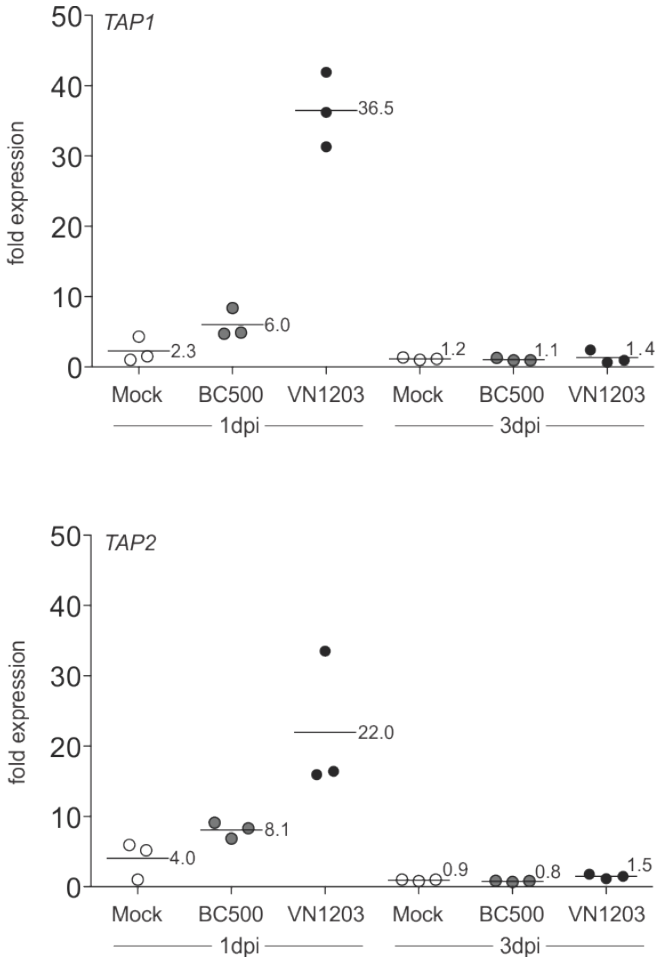

Figure 3-15. Expression of TAP during avian influenza infection. Three ducks were infected with PBS only, BC500 or VN1203. RNA was extracted from lung tissues at 1 and 3 dpi. *TAP1* and *TAP2* levels of expression were measured by qPCR.

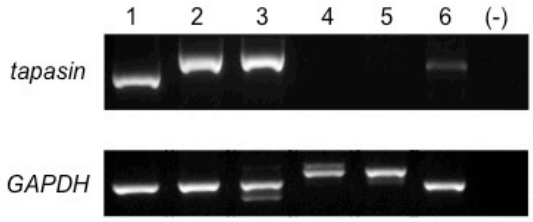

Figure 3-16. PCR amplification of *tapasin* from avian species. The primer set TAPBPE5F1/TAPBPE6R1 was used to amplify *tapasin* from genomic DNA. The numbered lanes represent the following: (1) chicken, (2) turkey, (3) pheasant, (4) mallard 19, (5) mallard 42, (6) pigeon and (-) as the negative control.

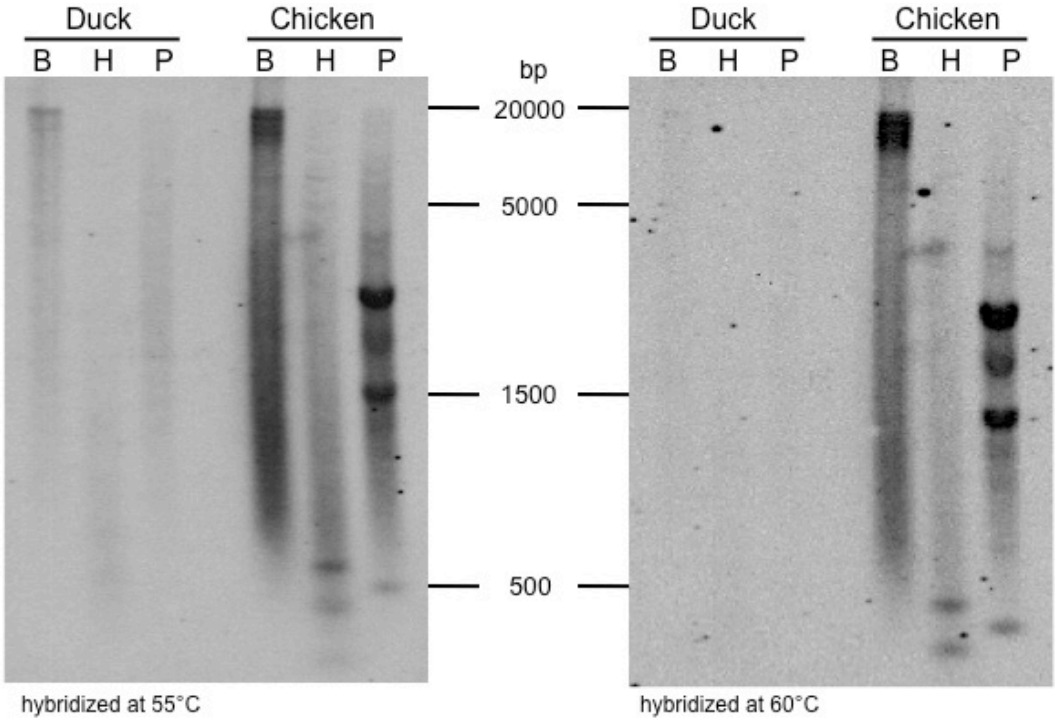

Figure 3-17. Low stringency Southern blot analysis of *tapasin* in the duck and chicken. Genomic DNA was digested with three restriction enzymes and separated by gel electrophoresis. The DNA was transferred to a nylon membrane and hybridized with a chicken tapasin probe overnight at 55°C or 60°C. The sizes are indicated by bands of the GeneRuler 1kb Plus (Fermentas). B, BglII; H, HaeIII; P, PstI.

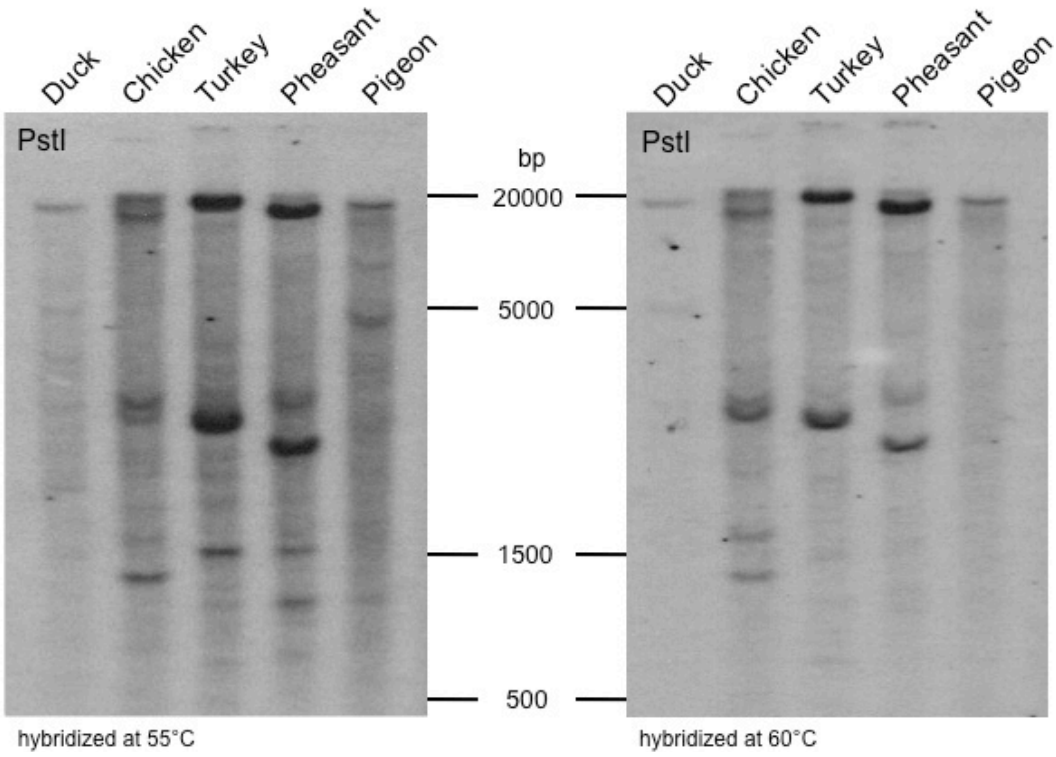

Figure 3-18. Comparison of hybridization patterns of *tapasin* in avian species. Duck, chicken, turkey, pheasant and pigeon genomic DNA was PstI digested and hybridized with a chicken tapasin probe overnight at 55°C or 60°C. Shown here are the x-ray films after 4 hours of exposure. The sizes are indicated by bands of the GeneRuler 1kb Plus (Fermentas).

4 DICUSSION

In wild mallard ducks the transporter associated with antigen processing genes are found next to the polymorphic MHC class I genes. Here we wanted to investigate if the TAP genes are also polymorphic in wild mallards. Additionally we wanted to identify the location of the polymorphisms within TAP. Polymorphisms found in functional regions of the TAP protein may influence antigen processing and ultimately antigen presentation. Furthermore polymorphisms within TAP could potentially restrict the peptide repertoire presented to MHC class I. This could lead to limitations in the duck's adaptive immunity and may allow for the propagation of influenza A virus in the natural host.

To analyze the differences of expressed TAP alleles, TAP1 and TAP2 were amplified by RT-PCR from 12 mallards. In total 32 and 81 polymorphic residues were identified in TAP1 and TAP2, respectively. Some of these polymorphisms are amino acid substitutions from biochemically different groups. They have been mapped to functional regions within the protein and suggest having functional implications on the TAP transporter. Furthermore we speculate that the observed polymorphisms may restrict the repertoire of transported peptides and limit the ability to present viral antigens to $CD8^+$ cells in the duck. Southern blot analysis revealed the absence of tapasin in the duck genome. This discovery could have further implications on the peptide loading complex in the natural host. It could mean that TAP heterodimer stability and the peptide loading can function efficiently in a tapasin independent manner, or the assembly of a

stable TAP heterodimer and peptide binding to MHC class I molecules is additionally impaired.

4.1 Polymorphisms in TAP do not interfere with ATP binding or hydrolysis

Polymorphisms in the nucleotide binding domain of an ABC transporter can affect ATP binding or ATP hydrolysis. Amino acid substitution within the ABC motifs can lead to loss in ATP binding or hydrolyzing activity. The sequences of these motifs are conserved in all mallard sequences except A368, which is located within the Walker motif of TAP1. Twenty alleles contain an alanine, which is substituted by a valine in four alleles. Both amino acids belong to the same biochemical group of hydrophobic amino acids. Therefore such a substitution most likely does not affect the function the peptide. Mutations of invariant residues within Walker A have been shown to affect ATP- binding activity of other ABC transporters. Structure and function analysis of the histidine permease from *Salmonella typhimurium* demonstrated that the invariant glycine and lysine residues in the Walker A motif are involved in binding of β- and γ-phosphates and changes to aspartic acid, serine and asparagine results in the loss ATP binding (Shyamala et al., 1991). Switches of the invariant aspartic acid in the Walker B motif, which is predicted to be important for Mg^{2+} binding, to lysine or arginine equally prevent ATP binding (Shyamala et al., 1991). Mutations of the invariant lysine in the Walker A motif to methionine allow ATP binding but ATP hydrolysis is impaired (Lapinski et al., 2001). These residues are conserved in all

duck TAP1 and TAP2 alleles and indicate that the ability to bind ATP is not affected in any peptides.

4.2 Duck TAP polymorphisms appear in functional regions

Polymorphic residues that differ between the two rat TAP2 alleles play an essential role in the peptide selection and certain substitutions have been shown affect TAP specificity. Two rat TAP2 alleles were described to have 25 allelic differences resulting in altered peptide specificity (Momburg et al., 1996). While most peptides transported by allele B contain hydrophobic C-terminal residues, allele A additionally translocates peptides with polar or uncharged residues at the C-terminus (Heemels et al., 1993; Momburg et al., 1994b). Polymorphic residues were observed mainly in the TMD of the peptide and correlate to our findings. More than 50% of the polymorphic residues found in the duck TAP are located within the TMD.

In TAP1 17 of the 32 polymorphic residues are located in the TMD and 43 out of 81 for TAP2. Four polymorphic residues of the duck TAP2 align with residues that are known to alter peptide specificity in the rat (Momburg et al., 1996). Three of the residues were localized in the cytosolic loop between transmembrane segment five and six of the duck TAP2, and the other residue was found in the ER luminal loop between TM six and seven. Residue G263 is of interest. It is substituted by arginine in mallard 12B, changing the amino acid group from a special case to an amino acid with positive charge. In seven of the duck alleles a switch to phenylalanine was observed, changing the biochemical

group from special to hydrophobic. It is tempting to speculate that these polymorphic residues might restrict peptide transport in the duck similar to rat alleles.

4.3 Evidence against co-evolving duck TAP alleles

Here we show that the TAP proteins are polymorphic in the wild mallards. Despite the presence of five MHC class I genes, only one appears to be extensively expressed in wild mallards (Jensen, 2009). The TAP genes are located adjacent to the dominantly expressed MHC class I gene, which is also polymorphic (Mesa et al., 2004). In lower vertebrates, the gene locations within in the MHC indicate an ancestral organization of this region. Additionally, the expression of a single MHC class I gene points to the original MHC, which expanded through gene duplication events and reorganization of genes in higher vertebrates (Nonaka et al., 1997; Ohta et al., 2002; Ohta et al., 2003). In chickens the expression of MHC class I gene and the fine specificity of TAP2 to the major MHC class I is explained by co-evolution (Walker et al., 2011). In chickens, the close proximity of TAP and MHC class I hinders recombination (Kaufman, 1999). This is in contrast to the big separation that is seen in the human MHC. The close proximity of antigen presentation genes in ducks, suggests co-evolution of functionally effective haplotypes.

In chickens, TAPs with identical MHCs appear to be similar. In fact polymorphic residues within the peptide binding region match the binding motifs of the associated major MHC class I. The chicken major MHC class I allele from

the B4 line shows a binding motif of positive residues in the PBR. The TAP1 allele linked to this particular MHC class I shows two polymorphic residues in the peptide binding regions, substituting for positively charged amino acids compared to other alleles (Walker et al., 2011). These positions were found to be monomorphic in mallard ducks.

If TAP alleles co-evolve to form a functional haplotype with MHC class I and matches its specificity, sequence similarities should be observed. Phylogenetic analysis revealed that similar TAP1 alleles are not linked to similar TAP2 alleles. In fact TAP1 and TAP2 alleles, which are linked in the same haplotype show higher sequence diversity than other alleles. This was unexpected, because if they match in peptide binding specificity, alleles should show higher amino acid identities. Phylogenetic analysis indicates no association of linked TAP alleles. In fact the completely opposite was observed. TAP1 alleles 25B and 42A show 100% sequence identity. This is not true for TAP2 25B and 42A. In fact TAP2 25B is very similar in sequence to 25A and 9A. TAP2 42A share the most sequence identity with 18A. The results indicate the association of different TAP1 and TAP2 alleles. This contradicts the presence of co-evolving haplotypes in ducks. The results suggest that TAP1 may be co-evolving with the major MHC class I, but not TAP2. If this is the case, then the polymorphisms observed in ducks are not functional.

Interestingly we observed a similar grouping pattern when we compared the TAP trees to phylogenetic trees constructed with the alpha-1 and alpha-2 domains of the MHC class I from wild mallard (Jensen, 2009). MHC class I

alleles 17.1 and 20.1 fall within the same branch of the phylogenetic tree indicating high sequence similarity. The same pattern was observed for TAP1 mallard 17A and 20A, which are 100% identical. These two alleles are also located in the same branch of the phylogenetic tree. Another example is 17.2 and 25.1, which show close proximity in the phylogenetic tree of the α1 and α2 domains (Jensen, 2009). TAP1 alleles 17B and 25A are again localized in the same branch of the tree, whereas TAP2 25A and 25B are paired on the same branch, and closer to 9A than to TAP2 17B. Interpreting these results allows us to speculate that a correlation between the peptide binding specificity of TAP1 and MHC class I molecules might exit, whereas polymorphisms in TAP2 most likely "hitch-hike" along within functionally linked and polymorphic MHC-TAP haplotypes.

4.4 Interferon induced TAP promoter

Our TAP promoter analysis identified polymorphic residues in the wild mallard duck promoter that are different from the domestic duck. Identification of transcription factor binding sites reveals the presence of ISRE and GAS elements in the duck promoter and is in accordance with studies performed on the chicken TAP promoter, which also identified the presence of binding site for ISRE and GAS in chicken TAP promoter (Walker et al., 2005). This shows the possibility of induction by interferons (IFNs). We demonstrated here that TAP transcript levels are elevated in lung samples infected with highly pathogenic avian influenza at 1 dpi compared to mock or low path infected samples. This can be explained by the

presence of ISRE and GAS and the robust IFN response mounted by the natural host. More studies on the activity of promoter have to be done to study gene expression and potential impacts of polymorphism.

4.5 Tapasin independence of the peptide loading complex

To investigate the presence of *tapasin* in the duck, several different approaches were conducted to obtain a sequence and provide evidence for the glycoprotein. Most species studied to date show the presence of a *tapasin* gene or a tapasin-like gene within the genome. In the mammalian genome, *tapasin* is localized within the extended MHC class II region (The MHC sequencing consortium, 1999). Other studies showed the presence of tapasin in avian species such as chicken (Frangoulis et al., 1999), turkey and pheasant (Sironi et al., 2006). A fragment within a BAC clone of zebra finch showed high homology to the turkey *tapasin* gene exons 1-4 (Balakrishnan et al., 2010). Primers were designed on the chicken sequence or alignments with other species. All attempts to amplify *tapasin* from mallards or a domestic duck by RT-PCR failed to yield a *tapasin* product. Further trials to obtain *tapasin* from genomic DNA did not produce the desired product. It is possible that the primers used failed to amplify *tapasin* from ducks because of high diversity between species. The human and chicken tapasin gene for example show only 36% homology (Frangoulis et al., 1999). Other studies have also shown that the *tapasin* gene is polymorphic in avian species (Sironi et al., 2006). In these experiments primers were designed within a conserved region and a fragment was successfully amplified from chicken, turkey

and pheasant genomic DNA, but failed to amplify *tapasin* from ducks in our trials. SNPs in the primer binding sites can prevent effective binding of the primer and fail to produce the desired product. Therefore, PCR amplifications may be too sensitive to detect duck *tapasin* if it is present but shows low similarity to the chicken.

A less sensitive approach was performed by low stringency southern blot analysis. Distinctive bands could be detected in several avian species including chicken, turkey and pheasant. However the probe does not hybridize to the duck and pigeon genome. For southern blot hybridizations, the probe has to match 70% of the sequence. The probe was designed to the conserved region of the chicken tapasin gene within the Ig-domain. Sequencing verified the identity of tapasin and top hits of a blastx search against the NCBI database were avian tapasin genes with approximately 90% similarity to turkey and pheasant. Supporting evidence for the absence of *tapasin* from the duck genome is provided. Recently, the duck genome was sequenced and the entire MHC was annotated, but a gene with homology to chicken or human *tapasin* could not be identified (Duck sequencing consortium Beijing, unpublished data). The duck MHC class II region consists of five MHC class II genes and we speculated we would find *tapasin* located upstream from these five genes. Analysis of the unannotated duck genome identified the location of the class II genes but searches for *tapasin* within proximity were unsuccessful. The TAP1 sequence alignment with human and chicken TAP1 reveals a trunctation at the start of the amino acid sequence compared to the human TAP1. The first 165 amino acids found in the human

TAP1 are missing from the duck TAP1. A similar truncation was described in the chicken TAP1 (Walker et al., 2005). This region was described to be involved in tapasin binding. In particular through binding to TAP tapasin stabilizes the heterodimer (Garbi et al., 2003). A loss of this region indicates inability to bind tapasin. In fact, some TAP alleles can function independently from tapasin. The role of tapasin in the peptide loading complex is controversial. Although some sources indicate the crucial role of tapasin in the peptide loading complex (Garbi et al., 2003; Chen and Bouvier, 2007; Rizvi and Raghavan, 2010), others show that the absence of tapasin does not impair antigen transport and binding, but rather alters the peptide repertoire (Boulanger et al., 2010).

The tapasin binding domains of TAP1 and TAP2 function separately from one another. Recruitment of tapasin involves only one of the two tapasin domains of the TAP subunits at a time in the peptide loading complex (Leonhardt et al., 2005). The dependence on the interaction between TAP2 and tapasin, but not TAP1, to form a full peptide loading complex was observed in human T2 cell lines expressing N-terminally truncated TAP (Leonhardt et al., 2005).This suggests that the lack of the tapasin docking domain in TAP1 might not be significant.

4.6 Significance of polymorphisms during avian influenza A virus infection

The expression of a single MHC class I gene is significant for genetic association to diseases. It confers either high susceptibility or resistance to a pathogenic infection. Analysis of antigen presentation genes showed great diversity in *TAP1* and *TAP2* genes in a population of wild mallard ducks.

Polymorphisms were identified in functional regions, which can affect peptide binding. Several non-conservative amino acids substitutions were observed in peptide binding regions indicating that TAP alleles may have different specificities for peptide binding. If the polymorphisms present result in functionally different transporters, this could result in restricted peptide transport and limited presentation on MHC class I. A restricting TAP transporter could limit the number of epitopes a particular duck can present from influenza A to CTLs, and ultimately cause constraint in the cell-mediated immune response to the virus in the natural host. Absence of tapasin could also implement constraint on antigen presentation and affect the formation of stable MHC class I and peptide complexes. Limited peptide diversity or stability could explain the poor immune response to vaccines and the perpetuation of the virus through re-infections.

5 REFERENCES

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

Ahn, K., T. H. Meyer, S. Uebel, P. Sempe, H. Djaballah, Y. Yang, P. A. Peterson, K. Fruh and R. Tampe (1996). "Molecular mechanism and species specificity of TAP inhibition by herpes simplex virus ICP47." The EMBO journal **15**(13): 3247-3255.

Anderson, K., P. Cresswell, M. Gammon, J. Hermes, A. Williamson and H. Zweerink (1991). "Endogenously synthesized peptide with an endoplasmic reticulum signal sequence sensitizes antigen processing mutant cells to class I-restricted cell-mediated lysis." The Journal of experimental medicine **174**(2): 489-492.

Androlewicz, M. J. and P. Cresswell (1994). "Human transporters associated with antigen processing possess a promiscuous peptide-binding site." Immunity **1**(1): 7-14.

Armandola, E. A., F. Momburg, M. Nijenhuis, N. Bulbuc, K. Fruh and G. J. Hammerling (1996). "A point mutation in the human transporter associated with antigen processing (TAP2) alters the peptide transport specificity." European journal of immunology **26**(8): 1748-1755.

Bahram, S., D. Arnold, M. Bresnahan, J. L. Strominger and T. Spies (1991). "Two putative subunits of a peptide pump encoded in the human major histocompatibility complex class II region." Proceedings of the National Academy of Sciences of the United States of America **88**(22): 10094-10098.

Balakrishnan, C. N., R. Ekblom, M. Volker, H. Westerdahl, R. Godinez, H. Kotkiewicz, D. W. Burt, T. Graves, D. K. Griffin, W. C. Warren and S. V. Edwards (2010). "Gene duplication and fragmentation in the zebra finch major histocompatibility complex." BMC biology **8**: 29.

Baldauf, C., S. Schrodt, M. Herget, J. Koch and R. Tampe (2010). "Single residue within the antigen translocation complex TAP controls the epitope repertoire by stabilizing a receptive conformation." Proceedings of the National Academy of Sciences of the United States of America **107**(20): 9135-9140.

Beck, S., A. Kelly, E. Radley, F. Khurshid, R. P. Alderton and J. Trowsdale (1992). "DNA sequence analysis of 66 kb of the human MHC class II

region encoding a cluster of genes for antigen processing." Journal of molecular biology **228**(2): 433-441.

Bjorkman, P. J. and P. Parham (1990). "Structure, function, and diversity of class I major histocompatibility complex molecules." Annual review of biochemistry **59**: 253-288.

Boulanger, D. S., R. Oliveira, L. Ayers, S. H. Prior, E. James, A. P. Williams and T. Elliott (2010). "Absence of tapasin alters immunodominance against a lymphocytic choriomeningitis virus polytope." Journal of immunology **184**(1): 73-83.

Braun, B. C., M. Glickman, R. Kraft, B. Dahlmann, P. M. Kloetzel, D. Finley and M. Schmidt (1999). "The base of the proteasome regulatory particle exhibits chaperone-like activity." Nature cell biology **1**(4): 221-226.

Carrington, M. (1999). "Recombination within the human MHC." Immunological reviews **167**: 245-256.

Chen, M. and M. Bouvier (2007). "Analysis of interactions in a tapasin/class I complex provides a mechanism for peptide selection." The EMBO journal **26**(6): 1681-1690.

de la Salle, H., J. Zimmer, D. Fricker, C. Angenieux, J. P. Cazenave, M. Okubo, H. Maeda, A. Plebani, M. M. Tongio, A. Dormoy and D. Hanau (1999). "HLA class I deficiencies due to mutations in subunit 1 of the peptide transporter TAP1." The Journal of clinical investigation **103**(5): R9-R13.

Degen, E., M. F. Cohen-Doyle and D. B. Williams (1992). "Efficient dissociation of the p88 chaperone from major histocompatibility complex class I molecules requires both beta 2-microglobulin and peptide." The Journal of experimental medicine **175**(6): 1653-1661.

Fehling, H. J., W. Swat, C. Laplace, R. Kuhn, K. Rajewsky, U. Muller and H. von Boehmer (1994). "MHC class I expression in mice lacking the proteasome subunit LMP-7." Science **265**(5176): 1234-1237.

Flajnik, M. F., Y. Ohta, A. S. Greenberg, L. Salter-Cid, A. Carrizosa, L. Du Pasquier and M. Kasahara (1999). "Two ancient allelic lineages at the single classical class I locus in the Xenopus MHC." Journal of immunology **163**(7): 3826-3833.

Frangoulis, B., I. Park, F. Guillemot, V. Severac, C. Auffray and R. Zoorob (1999). "Identification of the Tapasin gene in the chicken major histocompatibility complex." Immunogenetics **49**(4): 328-337.

Fruh, K., K. Ahn, H. Djaballah, P. Sempe, P. M. van Endert, R. Tampe, P. A. Peterson and Y. Yang (1995). "A viral inhibitor of peptide transporters for antigen presentation." Nature **375**(6530): 415-418.

Gadola, S. D., H. T. Moins-Teisserenc, J. Trowsdale, W. L. Gross and V. Cerundolo (2000). "TAP deficiency syndrome." Clinical and experimental immunology **121**(2): 173-178.

Garbi, N., N. Tiwari, F. Momburg and G. J. Hammerling (2003). "A major role for tapasin as a stabilizer of the TAP peptide transporter and consequences for MHC class I expression." European journal of immunology **33**(1): 264-273.

Grandea, A. G., 3rd, M. J. Androlewicz, R. S. Athwal, D. E. Geraghty and T. Spies (1995). "Dependence of peptide binding by MHC class I molecules on their interaction with TAP." Science **270**(5233): 105-108.

Grandea, A. G., 3rd, P. G. Comber, S. E. Wenderfer, G. Schoenhals, K. Fruh, J. J. Monaco and T. Spies (1998). "Sequence, linkage to H2-K, and function of mouse tapasin in MHC class I assembly." Immunogenetics **48**(4): 260-265.

Grant, E. P., M. T. Michalek, A. L. Goldberg and K. L. Rock (1995). "Rate of antigen degradation by the ubiquitin-proteasome pathway influences MHC class I presentation." Journal of immunology **155**(8): 3750-3758.

Greenberger, L. M., C. J. Lisanti, J. T. Silva and S. B. Horwitz (1991). "Domain mapping of the photoaffinity drug-binding sites in P-glycoprotein encoded by mouse mdr1b." The Journal of biological chemistry **266**(31): 20744-20751.

Grimholt, U. (1997). "Transport-associated proteins in Atlantic salmon (Salmo salar)." Immunogenetics **46**(3): 213-221.

Groll, M., L. Ditzel, J. Lowe, D. Stock, M. Bochtler, H. D. Bartunik and R. Huber (1997). "Structure of 20S proteasome from yeast at 2.4 A resolution." Nature **386**(6624): 463-471.

Gros, P., R. Dhir, J. Croop and F. Talbot (1991). "A single amino acid substitution strongly modulates the activity and substrate specificity of the mouse mdr1 and mdr3 drug efflux pumps." Proceedings of the National Academy of Sciences of the United States of America **88**(16): 7289-7293.

Guillemot, F., A. Billault, O. Pourquie, G. Behar, A. M. Chausse, R. Zoorob, G. Kreibich and C. Auffray (1988). "A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer." The EMBO journal **7**(9): 2775-2785.

Hansen, J. D., P. Strassburger, G. H. Thorgaard, W. P. Young and L. Du Pasquier (1999). "Expression, linkage, and polymorphism of MHC-related genes in rainbow trout, Oncorhynchus mykiss." Journal of immunology **163**(2): 774-786.

Heemels, M. T. and H. L. Ploegh (1994). "Substrate specificity of allelic variants of the TAP peptide transporter." Immunity **1**(9): 775-784.

Heemels, M. T., T. N. Schumacher, K. Wonigeit and H. L. Ploegh (1993). "Peptide translocation by variants of the transporter associated with antigen processing." Science **262**(5142): 2059-2063.

Herberg, J. A., J. Sgouros, T. Jones, J. Copeman, S. J. Humphray, D. Sheer, P. Cresswell, S. Beck and J. Trowsdale (1998). "Genomic analysis of the Tapasin gene, located close to the TAP loci in the MHC." European journal of immunology **28**(2): 459-467.

Hewitt, E. W. (2003). "The MHC class I antigen presentation pathway: strategies for viral immune evasion." Immunology **110**(2): 163-169.

Higgins, C. F. (1992). "ABC transporters: from microorganisms to man." Annual review of cell biology **8**: 67-113.

Higgins, C. F., I. D. Hiles, G. P. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell and et al. (1986). "A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria." Nature **323**(6087): 448-450.

Hyde, S. C., P. Emsley, M. J. Hartshorn, M. M. Mimmack, U. Gileadi, S. R. Pearce, M. P. Gallagher, D. R. Gill, R. E. Hubbard and C. F. Higgins (1990). "Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport." Nature **346**(6282): 362-365.

International Chicken Genome sequencing Consortium (2004)."Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution." Nature **432**(7018): 695-716.

Jackson, M. R., T. Nilsson and P. A. Peterson (1990). "Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum." The EMBO journal **9**(10): 3153-3162.

Jacob, J. P., S. Milne, S. Beck and J. Kaufman (2000). "The major and a minor class II beta-chain (B-LB) gene flank the Tapasin gene in the B-F /B-L region of the chicken major histocompatibility complex." Immunogenetics **51**(2): 138-147.

Janas, E., M. Hofacker, M. Chen, S. Gompf, C. van der Does and R. Tampe (2003). "The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p." The Journal of biological chemistry **278**(29): 26862-26869.

Jensen, S. (2009). Allelic diversity of MHC I in wild mallards. M.Sc, University of Alberta.

Joly, E., A. F. Le Rolle, A. L. Gonzalez, B. Mehling, J. Stevens, W. J. Coadwell, T. Hunig, J. C. Howard and G. W. Butcher (1998). "Co-evolution of rat TAP transporters and MHC class I RT1-A molecules." Current biology : CB **8**(3): 169-172.

Kageyama, G., S. Kawano, S. Kanagawa, S. Kondo, M. Sugita, T. Nakanishi, A. Shimizu and S. Kumagai (2004). "Effect of mutated transporters associated with antigen-processing 2 on characteristic major histocompatibility complex binding peptides: analysis using electrospray ionization tandem mass spectrometry." Rapid communications in mass spectrometry : RCM **18**(9): 995-1000.

Kaufman, J. (1999). "Co-evolving genes in MHC haplotypes: the "rule" for nonmammalian vertebrates?" Immunogenetics **50**(3-4): 228-236.

Kaufman, J., S. Milne, T. W. Gobel, B. A. Walker, J. P. Jacob, C. Auffray, R. Zoorob and S. Beck (1999). "The chicken B locus is a minimal essential major histocompatibility complex." Nature **401**(6756): 923-925.

Kaufman, J., H. Volk and H. J. Wallny (1995). "A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism." Immunological reviews **143**: 63-88.

Kelly, A., S. H. Powis, L. A. Kerr, I. Mockridge, T. Elliott, J. Bastin, B. Uchanska-Ziegler, A. Ziegler, J. Trowsdale and A. Townsend (1992). "Assembly and function of the two ABC transporter proteins encoded in the human major histocompatibility complex." Nature **355**(6361): 641-644.

Koch, J., R. Guntrum, S. Heintke, C. Kyritsis and R. Tampe (2004). "Functional dissection of the transmembrane domains of the transporter associated with antigen processing (TAP)." The Journal of biological chemistry **279**(11): 10142-10147.

Kuchler, K., R. E. Sterne and J. Thorner (1989). "Saccharomyces cerevisiae STE6 gene product: a novel pathway for protein export in eukaryotic cells." The EMBO journal **8**(13): 3973-3984.

Lapinski, P. E., R. R. Neubig and M. Raghavan (2001). "Walker A lysine mutations of TAP1 and TAP2 interfere with peptide translocation but not peptide binding." The Journal of biological chemistry **276**(10): 7526-7533.

Lehner, P. J., J. T. Karttunen, G. W. Wilkinson and P. Cresswell (1997). "The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation." Proceedings of the National Academy of Sciences of the United States of America **94**(13): 6904-6909.

Leonhardt, R. M., K. Keusekotten, C. Bekpen and M. R. Knittler (2005). "Critical role for the tapasin-docking site of TAP2 in the functional integrity of the MHC class I-peptide-loading complex." Journal of immunology **175**(8): 5104-5114.

Lewis, J. W., A. Neisig, J. Neefjes and T. Elliott (1996). "Point mutations in the alpha 2 domain of HLA-A2.1 define a functionally relevant interaction with TAP." Current biology : CB **6**(7): 873-883.

Magor, K. E., D. A. Higgins, D. L. Middleton and G. W. Warr (1994). "One gene encodes the heavy chains for three different forms of IgY in the duck." Journal of immunology **153**(12): 5549-5555.

Mesa, C. M., K. J. Thulien, D. A. Moon, S. M. Veniamin and K. E. Magor (2004). "The dominant MHC class I gene is adjacent to the polymorphic TAP2 gene in the duck, Anas platyrhynchos." Immunogenetics **56**(3): 192-203.

Meyer, T. H., P. M. van Endert, S. Uebel, B. Ehring and R. Tampe (1994). "Functional expression and purification of the ABC transporter complex associated with antigen processing (TAP) in insect cells." FEBS letters **351**(3): 443-447.

Momburg, F., E. A. Armandola, M. Post and G. J. Hammerling (1996). "Residues in TAP2 peptide transporters controlling substrate specificity." Journal of immunology **156**(5): 1756-1763.

Momburg, F., J. Roelse, G. J. Hammerling and J. J. Neefjes (1994a). "Peptide size selection by the major histocompatibility complex-encoded peptide transporter." The Journal of experimental medicine **179**(5): 1613-1623.

Momburg, F., J. Roelse, J. C. Howard, G. W. Butcher, G. J. Hammerling and J. J. Neefjes (1994b). "Selectivity of MHC-encoded peptide transporters from human, mouse and rat." Nature **367**(6464): 648-651.

Moon, D. A., S. M. Veniamin, J. A. Parks-Dely and K. E. Magor (2005). "The MHC of the duck (Anas platyrhynchos) contains five differentially expressed class I genes." Journal of immunology **175**(10): 6702-6712.

Neefjes, J., E. Gottfried, J. Roelse, M. Gromme, R. Obst, G. J. Hammerling and F. Momburg (1995). "Analysis of the fine specificity of rat, mouse and human TAP peptide transporters." European journal of immunology **25**(4): 1133-1136.

Neefjes, J. J., F. Momburg and G. J. Hammerling (1993). "Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter." Science **261**(5122): 769-771.

Neumann, L. and R. Tampe (1999). "Kinetic analysis of peptide binding to the TAP transport complex: evidence for structural rearrangements induced by substrate binding." Journal of molecular biology **294**(5): 1203-1213.

Nijenhuis, M., S. Schmitt, E. A. Armandola, R. Obst, J. Brunner and G. J. Hammerling (1996). "Identification of a contact region for peptide on the TAP1 chain of the transporter associated with antigen processing." Journal of immunology **156**(6): 2186-2195.

Nonaka, M., C. Namikawa, Y. Kato, M. Sasaki, L. Salter-Cid and M. F. Flajnik (1997). "Major histocompatibility complex gene mapping in the amphibian Xenopus implies a primordial organization." Proceedings of the National Academy of Sciences of the United States of America **94**(11): 5789-5791.

Obst, R., E. A. Armandola, M. Nijenhuis, F. Momburg and G. J. Hammerling (1995). "TAP polymorphism does not influence transport of peptide variants in mice and humans." European journal of immunology **25**(8): 2170-2176.

Ohta, Y., E. C. McKinney, M. F. Criscitiello and M. F. Flajnik (2002). "Proteasome, transporter associated with antigen processing, and class I genes in the nurse shark Ginglymostoma cirratum: evidence for a stable class I region and MHC haplotype lineages." Journal of immunology **168**(2): 771-781.

Ohta, Y., S. J. Powis, W. J. Coadwell, D. E. Haliniewski, Y. Liu, H. Li and M. F. Flajnik (1999). "Identification and genetic mapping of Xenopus TAP2 genes." Immunogenetics **49**(3): 171-182.

Ohta, Y., S. J. Powis, R. L. Lohr, M. Nonaka, L. D. Pasquier and M. F. Flajnik (2003). "Two highly divergent ancient allelic lineages of the transporter associated with antigen processing (TAP) gene in Xenopus: further evidence for co-evolution among MHC class I region genes." European journal of immunology **33**(11): 3017-3027.

Ortmann, B., M. J. Androlewicz and P. Cresswell (1994). "MHC class I/beta 2-microglobulin complexes associate with TAP transporters before peptide binding." Nature **368**(6474): 864-867.

Ortmann, B., J. Copeman, P. J. Lehner, B. Sadasivan, J. A. Herberg, A. G. Grandea, S. R. Riddell, R. Tampe, T. Spies, J. Trowsdale and P. Cresswell (1997). "A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes." Science **277**(5330): 1306-1309.

Peace-Brewer, A. L., L. G. Tussey, M. Matsui, G. Li, D. G. Quinn and J. A. Frelinger (1996). "A point mutation in HLA-A*0201 results in failure to bind the TAP complex and to present virus-derived peptides to CTL." Immunity **4**(5): 505-514.

Peaper, D. R., P. A. Wearsch and P. Cresswell (2005). "Tapasin and ERp57 form a stable disulfide-linked dimer within the MHC class I peptide-loading complex." The EMBO journal **24**(20): 3613-3623.

Powis, S. J., E. V. Deverson, W. J. Coadwell, A. Ciruela, N. S. Huskisson, H. Smith, G. W. Butcher and J. C. Howard (1992). "Effect of polymorphism of an MHC-linked transporter on the peptides assembled in a class I molecule." Nature **357**(6375): 211-215.

Procko, E., I. Ferrin-O'Connell, S. L. Ng and R. Gaudet (2006). "Distinct structural and functional properties of the ATPase sites in an asymmetric ABC transporter." Molecular cell **24**(1): 51-62.

Rizvi, S. M. and M. Raghavan (2006). "Direct peptide-regulatable interactions between MHC class I molecules and tapasin." Proceedings of the National Academy of Sciences of the United States of America **103**(48): 18220-18225.

Rizvi, S. M. and M. Raghavan (2010). "Mechanisms of function of tapasin, a critical major histocompatibility complex class I assembly factor." Traffic **11**(3): 332-347.

Rock, K. L., C. Gramm, L. Rothstein, K. Clark, R. Stein, L. Dick, D. Hwang and A. L. Goldberg (1994). "Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules." Cell **78**(5): 761-771.

Rudolph, M. G., J. Stevens, J. A. Speir, J. Trowsdale, G. W. Butcher, E. Joly and I. A. Wilson (2002). "Crystal structures of two rat MHC class Ia (RT1-A) molecules that are associated differentially with peptide transporter alleles TAP-A and TAP-B." Journal of molecular biology **324**(5): 975-990.

Sadasivan, B., P. J. Lehner, B. Ortmann, T. Spies and P. Cresswell (1996). "Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP." Immunity **5**(2): 103-114.

Saveanu, L., O. Carroll, V. Lindo, M. Del Val, D. Lopez, Y. Lepelletier, F. Greer, L. Schomburg, D. Fruci, G. Niedermann and P. M. van Endert (2005).

"Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum." Nature immunology **6**(7): 689-697.

Schrodt, S., J. Koch and R. Tampe (2006). "Membrane topology of the transporter associated with antigen processing (TAP1) within an assembled functional peptide-loading complex." The Journal of biological chemistry **281**(10): 6455-6462.

Shyamala, V., V. Baichwal, E. Beall and G. F. Ames (1991). "Structure-function analysis of the histidine permease and comparison with cystic fibrosis mutations." The Journal of biological chemistry **266**(28): 18714-18719.

Sieker, F., S. Springer and M. Zacharias (2007). "Comparative molecular dynamics analysis of tapasin-dependent and -independent MHC class I alleles." Protein science : a publication of the Protein Society **16**(2): 299-308.

Sironi, L., B. Lazzari, P. Ramelli, C. Gorni and P. Mariani (2006). "Single nucleotide polymorphism discovery in the avian Tapasin gene." Poultry science **85**(4): 606-612.

Smith, P. C., N. Karpowich, L. Millen, J. E. Moody, J. Rosen, P. J. Thomas and J. F. Hunt (2002). "ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer." Molecular cell **10**(1): 139-149.

Spies, T., V. Cerundolo, M. Colonna, P. Cresswell, A. Townsend and R. DeMars (1992). "Presentation of viral antigen by MHC class I molecules is dependent on a putative peptide transporter heterodimer." Nature **355**(6361): 644-646.

The MHC sequencing consortium. (1999). "Complete sequence and gene map of a human major histocompatibility complex." Nature **401**: 921-923.

Townsend, A. and H. Bodmer (1989). "Antigen recognition by class I-restricted T lymphocytes." Annual review of immunology **7**: 601-624.

van Endert, P. M., L. Saveanu, E. W. Hewitt and P. Lehner (2002). "Powering the peptide pump: TAP crosstalk with energetic nucleotides." Trends in biochemical sciences **27**(9): 454-461.

Van Kaer, L., P. G. Ashton-Rickardt, M. Eichelberger, M. Gaczynska, K. Nagashima, K. L. Rock, A. L. Goldberg, P. C. Doherty and S. Tonegawa (1994). "Altered peptidase and viral-specific T cell response in LMP2 mutant mice." Immunity **1**(7): 533-541.

Walker, B. A., L. G. Hunt, A. K. Sowa, K. Skjodt, T. W. Gobel, P. J. Lehner and J. Kaufman (2011). "The dominantly expressed class I molecule of the chicken MHC is explained by coevolution with the polymorphic peptide transporter (TAP) genes." Proceedings of the National Academy of Sciences of the United States of America **108**(20): 8396-8401.

Walker, B. A., A. van Hateren, S. Milne, S. Beck and J. Kaufman (2005). "Chicken TAP genes differ from their human orthologues in locus organisation, size, sequence features and polymorphism." Immunogenetics **57**(3-4): 232-247.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." The EMBO journal **1**(8): 945-951.

Wearsch, P. A. and P. Cresswell (2007). "Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer." Nature immunology **8**(8): 873-881.

6 Appendix A

```
Majority ATGGAGTTGCTGCCCACCTTGCGCCTGGCCTGTGTCCTGCTCCTGGCTGACCTGGTCGTG

9A       ............................................................      60
9B       ............................................................      60
10A      ............................................................      60
10B      ............................................................      60
12A      ............................................................      60
12B      ............................................................      60
16A      ............................................................      60
16B      ............................................................      60
17A      ............................................................      60
17B      ............................................................      60
18A      ............................................................      60
18B      ............................................................      60
19A      ......C.....................................................      60
19B      ............................................................      60
20A      ............................................................      60
20B      ............................................................      60
24A      ............................................................      60
24B      ............................................................      60
25A      ............................................................      60
25B      ............................................................      60
32A      ...........................................................T...      60
32B      ............................................................      60
42A      ............................................................      60


Majority CTGGCAGCACTGGCCCGGTTGGCCCCGGCACTGGCCCAGCTGGGTCTAGTGGCCACATGG

9A       ........................A...................................     120
9B       ................A.........A.................................     120
10A      .....G......................................................     120
10B      ............................................................     120
12A      ............................................................     120
12B      ............................................................     120
16A      ............................................................     120
16B      ................A...........................................     120
17A      ................A...........................................     120
17B      ............................................................     120
18A      ............................................................     120
18B      ................A...........................................     120
19A      ............................................................     120
19B      ............................................................     120
20A      ................A.........A.................................     120
20B      ............................................................     120
24A      ............................................................     120
24B      ................A...........................................     120
25A      ........................A...................................     120
25B      ................T...........................................     120
32A      ........................A...................................     120
32B      ................A.........A.................................     120
42A      ................A...........................................     120
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority CTGGAGGCTGGGCTGCGGCTACCAGTGCTGGTGGGAGCTGGGAGGCTGTTGGCCCCCGGA

9A       ..........................................................        180
9B       ..........................................................        180
10A      .....................................................A....        180
10B      ..........................................................        180
12A      ..........................................T...............        180
12B      ..........................................T...............        180
16A      .....................................................T...        180
16B      ..........................................T...............        180
17A      ..........................................T...............        180
17B      ..........................................................        180
18A      ..........................................T...............        180
18B      ..........................................T...............        180
19A      ..........................................................        180
19B      ..........................................................        180
20A      ..........................................................        180
20B      ..........................................T...............        180
24A      ..........................................................        180
24B      ..........................................T...............        180
25A      ..........................................................        180
25B      .....................................................A....        180
32A      ..........................................................        180
32B      ..........................................................        180
42A      ..........................................T...............        180


Majority GGACCCCGGGGAGCCGCGGCCCTGGTGAGCCTGGCCCCTGCCACCTTCCTTACCCTGCGG

9A       ..........................................................        240
9B       ................A.........................................        240
10A      ................A.........................................        240
10B      ..............C...........................................        240
12A      .......A..................................T.............        240
12B      ..........................................................        240
16A      ..............C...........................................        240
16B      ..........................................T.............        240
17A      ..........................................................        240
17B      ..............C...........................................        240
18A      ..........................................T.............        240
18B      ..........................................................        240
19A      ..............C...........................................        240
19B      ..........................................................        240
20A      ................A.........................................        240
20B      ................A.........................T.............        240
24A      ..............C...........................................        240
24B      ..........................................T.............        240
25A      ..........................................................        240
25B      ................A.........................................        240
32A      ..........................................................        240
32B      ................A.........................................        240
42A      ..........................................................        240
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GGCTGCCTGGAGCTGCCTGG-GGCTCCACCAGTGCTGCTGGCCATGGCCACACCGTCCTG

9A       ..................-................................T......        299
9B       ..................-................................A.....        299
10A      ..................-...................................        299
10B      .................A.-...................................        299
12A      ..................-...................................        299
12B      ..................-...................................        299
16A      ..................-...................................        299
16B      ..................-....T.............A...................        299
17A      ..................-...................................        299
17B      ..................-................................A.....        299
18A      ..................-.............A.....................        299
18B      ..................-.............T.....................        299
19A      C-TG..A.T.GC..ATGG..CA.TCTTG.TG.CC.....CA..TG.A..T.C.T.G.AC-        298
19B      ..................-...................................        299
20A      ..................-................................A.....        299
20B      ..................-................................A.....        299
24A      ..................-...................................        299
24B      ..................-................................A.....        299
25A      ..................-................................T......        299
25B      ..................-...................................        299
32A      ..................-...................................        299
32B      ..................-................................A.....        299
42A      ..................-.............T.....................        299


Majority GCTGGCATTGGCCTATGGGGCAGTCTTGCT--GGCCCTGCTCACCTGGACCTCCCTGGCA

9A       ............................--............................        357
9B       ............................--............................        357
10A      ............................--............................        357
10B      ............................--............................        357
12A      ..............C.............--................G.........        357
12B      ..............C.............--............................        357
16A      ............................--............................        357
16B      ..............C.............--............................        357
17A      ..............C.............--............................        357
17B      ............................--............................        357
18A      ..............C.............--................G.........        357
18B      ..............C.............--............................        357
19A      -....GG.G.C...GG..AC..AGGAG.TCAA............................        357
19B      ............................--............................        357
20A      ............................--............................        357
20B      ..............C.............--............................        357
24A      ............................--............................        357
24B      ............................--............................        357
25A      ............................--............................        357
25B      ............................--............................        357
32A      ..............C.............--............................        357
32B      ............................--............................        357
42A      ..............C.............--............................        357
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority CCTGGGGTGGCCCTGGGGACCAAGGAGGTCAAGTACCAGGCGGCCCTGCGCCGGCAGCTG

9A       ............................................................        417
9B       ............................................................        417
10A      ............................................................        417
10B      ............................................................        417
12A      ............................................................        417
12B      ............................................................        417
16A      ............................................................        417
16B      ............................................................        417
17A      ............................................................        417
17B      .....A.....................................A..................        417
18A      ............................................................        417
18B      ........................................A...................        417
19A      ............................................................        417
19B      ............................................................        417
20A      ............................................................        417
20B      ............................................................        417
24A      ............................................................        417
24B      ...................................................A........        417
25A      ............................................................        417
25B      ............................................................        417
32A      ............................................................        417
32B      ............................................................        417
42A      ........................................A...................        417


Majority GCCCTGGCCTGGCCTGAGTGGCCCTTCCTCAGCGGAGCCTTCTTCTXCCTCGTGCTGGCT

9A       ..............................................G.............        477
9B       ..............................................T.............        477
10A      ........................C.......T.............T.............        477
10B      ..............................................T.............        477
12A      ..............................................T.............        477
12B      ..............................................G.............        477
16A      ..............................................G.............        477
16B      ..............................................G.............        477
17A      ..............................................G.............        477
17B      ..............................................T.............        477
18A      ..............................................G.............        477
18B      ..............................................G.............        477
19A      ..............................................T.............        477
19B      ..............................................T.............        477
20A      ......................................T.......T.............        477
20B      ..............................................G.............        477
24A      ..............................................T.............        477
24B      ..............C...............................T.............        477
25A      ..............................................G.............        477
25B      ..............................................G.............        477
32A      ..............................................T.............        477
32B      ..............................................T.............        477
42A      ..............................................G.............        477
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GCATTGGGTGAGACCTCCGTGCCCTACTGCACTGGGAAGGCCTTGGATGTCCTCCGCCAT

9A       ..............................................CA...........        537
9B       ................................................................T...        537
10A      ..............................................C.............        537
10B      ................................................................        537
12A      ........................................T.....C.............        537
12B      ..............................................A...............        537
16A      ................................................................        537
16B      ..............................................A...............        537
17A      ................................................................T...        537
17B      ..............................................C.............        537
18A      ..............................................A...............        537
18B      ..............................................A...............        537
19A      ................................................................        537
19B      ..............................................CA...........        537
20A      ................................................................T...        537
20B      ..............................................A...............        537
24A      ................................................................        537
24B      ................................................................T...        537
25A      ..............................................CA...........        537
25B      ..............................................A...............        537
32A      ..............................................C.............        537
32B      ................................................................T...        537
42A      ..............................................A...............        537


Majority GGGGACGGCCCCACTGCCTTTGCCACTGCCATCGGCTTTGTGTGCCTXGCCTCCGCCAGC

9A       ........................................C........C.............        597
9B       .....T..................................C.....T......        597
10A      .....T..................................T.............        597
10B      ..............................................C.............        597
12A      .....T..................................T.............        597
12B      ..............................................T.............        597
16A      ..............................T.............C.............        597
16B      ..............................................T.............        597
17A      .....T..................................T.............        597
17B      .....T.........................T......A.......C.............        597
18A      ........................................T.....T.............        597
18B      ......A..................................T.............        597
19A      ..............................................C.............        597
19B      ..................................                           570
20A      .....T..................................C.....T......        597
20B      ......A..................................T.............        597
24A      ..............................T.............T.............        597
24B      .....T...................T.....................C.....T......        597
25A      ........................................C........C.............        597
25B      ................................................T.............        597
32A      .....T.........................T......A.......C.............        597
32B      .....T..................................C.....T......        597
42A      ......A..................................T.............        597
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority AGTTCGCTGTTTGCTGGCTGCCGTGGGGGCCTCTTCACCTTCATCATGGCTCGCCTCACC

9A       ..........................................................       657
9B       .....................................G.....................A       657
10A      ........................A.................................       657
10B      .....A......T.............................................G       657
12A      ..................................G.....................A       657
12B      ..........................................................       657
16A      ....................................TG.....................G       657
16B      ..........................................................G       657
17A      ........................A.................................       657
17B      ..........................................................       657
18A      ........................A.................................       657
18B      ........................A.................................       657
19A      .....A......T.............................................G       657
20A      ....................................G.....................       657
20B      ..........................A.............................       657
24A      ........................A.................................       657
24B      ....................................G.....................A       657
25A      ..........................................................       657
25B      ..........................................................       657
32A      ..........................................................T..G       657
32B      ....................................G.....................       657
42A      ........................A.................................       657


Majority CTGCGCACCCGTGACCGGCTCTTCTCTGGCCTGGTACACCAGGATCTGGCCTTCTTCCAG

9A       ..........C....A..........................................       717
9B       ..........C..........................................................       717
10A      ..........................................................       717
10B      .....T.................................G.....................       717
12A      ..........C..........................................................       717
12B      ..........C.............T..A.............................       717
16A      ..............................A...........................       717
16B      ..........C.......................G.......................       717
17A      ..........................................................       717
17B      .............A.............A..G.......................       717
18A      ..........................................................       717
18B      ..........................................................       717
19A      .....T.................................G.....................       717
20A      ..........C....A..........................................       717
20B      .............A.................G.......................       717
24A      ..........................................................       717
24B      ..........C..........................................................       717
25A      ..........C....A..........................................       717
25B      .............A.............A..G.......................       717
32A      ..........C..........................................................       717
32B      ..........C....A..........................................       717
42A      ..........................................................       717
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority CAGACCACAGCAGCTGATCTGGCCTCCCGGCTGGCCACTGACGTGCCACTGGCGAGCAGG

9A       ..........................................................        777
9B       ..........................................................        777
10A      .........................................T..................        777
10B      ..........................................................A.....        777
12A      ..........................................................        777
12B      .........................................T..................        777
16A      ..........................................................        777
16B      ..........................................................        777
17A      .........................................T..................        777
17B      ..........................................................        777
18A      .........................................T..................        777
18B      .........................................T..................        777
19A      ..........................................................A.....        777
20A      ..........................................................        777
20B      ..........................................................GAC        777
24A      .........................................T..................        777
24B      ..........................................................        777
25A      ..........................................................        777
25B      ..............................................G.........GAC        777
32A      .........A................................................        777
32B      ..........................................................        777
42A      ..........................................................A.....        777


Majority GTGGTGCCAGGCAGTGCCAACATCGCCCTGAGGAACCTGGGGAAGGTTCTGGGGCTCAGT

9A       ..........................................................        837
9B       .....T..................T.....................................        837
10A      .........TT.....................................A.....C        837
10B      ..T......TT.................................................C        837
12A      .....T..................T.....................................        837
12B      ......T..A.A.T............T.C..................................        837
16A      ..........................................................        837
16B      ..........................................................        837
17A      .........TT.....T.....................................C        837
17B      .....T..................T.....................................        837
18A      .........TT.................................................C        837
18B      .........TT.................................................        837
19A      ..T......TT.................................................C        837
20A      ..........................................................        837
20B      ..........................................................        837
24A      .........TT.....T.....................................        837
24B      ....................T..T.....................................        837
25A      ..........................................................        837
25B      ..........................................................        837
32A      ......................T.....................................        837
32B      ..........................................................        837
42A      .....T..................T.....................................        837
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GCCTTCATGCTGGCGCTCTCACCACGCCTGACCCTGCTGGCGCTGCTTGAGGTGCCCCTT

9A       ..............A.............................................       897
9B       .....................................................T......       897
10A      ............................................................       897
10B      ..............A.............................................       897
12A      .....................................................T......       897
12B      ...................................................A........       897
16A      ..............A..................T..........................       897
16B      ............................................................       897
17A      ..............A..........................A..................       897
17B      .....................................................T......       897
18A      ..................................................A.........       897
18B      ...........................................................C       897
19A      ..............A.............................................       897
20A      ..............A.............................................       897
20B      ............................................................       897
24A      ............................................................       897
24B      .........................................A................C.       897
25A      ..............A.............................................       897
25B      .................................T..........................       897
32A      .........................................A..................       897
32B      ..............A.............................................       897
42A      .....................................................T......       897


Majority GCCATTGTTGCACGCAAGATCTACAATGCCCGATACCAGATGCTGCAGCAGGCCGTGCTG

9A       ............................................................       957
9B       ...........G................................................       957
10A      ...................................T........................       957
10B      .......C...G.A....G.............GC..........................       957
12A      ...........G................................................       957
12B      .......C...T.A....G.........................................       957
16A      ............................................................       957
16B      .......C...G.A....G.............GC..........................       957
17A      ...................................T........................       957
17B      ...........G................................................       957
18A      A............................T.....T........................       957
18B      .......C...G.A....G.............GC..........................       957
19A      .......C...G.A....G....T........GC..........................       957
20A      ............................................................       957
20B      .............A..............................................       957
24A      ............................................................       957
24B      ...........G................................................       957
25A      ............................................................       957
25B      ............................................................       957
32A      .......C...G.A....G.............GC..........................       957
32B      ............................................................       957
42A      ...........G................................................       957
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GATGCAACAGCTGGCACTGGGATGGTGGTGCAGGAGGCCATCTCCTCCATCGAGACAGTG

9A       ..C.....................................................       1017
9B       ..........................................................T.....       1017
10A      ..........................................................       1017
10B      ..........................................................       1017
12A      ..........................................................T.....       1017
12B      ...........G.........G.....................................       1017
16A      ..........................................................       1017
16B      ..........................................................T.....       1017
17A      ..........................................................T.....       1017
17B      ..........................................................T.....       1017
18A      ..........................................................       1017
18B      ..C.....................................................       1017
19A      ..........................................................       1017
20A      ..........................................................T.....       1017
20B      ....................G.....................................       1017
24A      ..C.....................................................       1017
24B      ..C.....................................................       1017
25A      ...................................T..........T.....       1017
25B      ....................G.....................................       1017
32A      ..........................................................       1017
32B      ..........................................................T.....       1017
42A      ..........................................................       1017


Majority CGGGCCTTCGCTGGGGAGGAGGAGGAAGAGCGCCGGTACGGCCAGGCACTGACCAAGATG

9A       ........T...............................................       1077
9B       ........T...............................................       1077
10A      ........T...............................................       1077
10B      ...C....................................................       1077
12A      ..........................................A.............       1077
12B      ..........................................................       1077
16A      ..........................................................       1077
16B      ..................................A.....................       1077
17A      .........A..........................A.....................       1077
17B      ..........................................................       1077
18A      ..........................................................       1077
18B      ..........................................................       1077
19A      ...C....................................................       1077
20A      ..........................................................       1077
20B      ........................................A.................       1077
24A      ..........................................................       1077
24B      ..........................................................       1077
25A      ..........................................................       1077
25B      ..........................................................       1077
32A      ..........................................................       1077
32B      ..........................................................       1077
42A      ..........................................................       1077
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority CTGGGGCTGCGGGACCAGCGGGACATGGAGAAGGCAATCTTCCTCCTCATCCGGCGGGCA

9A       ..........................................................      1137
9B       ..........................................................      1137
10A      ..T.................A.............................T.........      1137
10B      ...............................CA..............G...........      1137
12A      ..........................................................      1137
12B      ...A......................................................      1137
16A      ..........................................................      1137
16B      ..T.....................................................T.A.....      1137
17A      ..T......A...........................T...........T.........      1137
17B      ..........................................................      1137
18A      ..........................................................      1137
18B      ..........................................................      1137
19A      ...............................CA..............G...........      1137
20A      ..........................................................      1137
20B      ..........................................................      1137
24A      ..........................................................      1137
24B      ..........................................................      1137
25A      ..........................................................      1137
25B      ..........................................................      1137
32A      ...............................CA..............G...........      1137
32B      ..........................................................      1137
42A      ..........................................................      1137


Majority CTGCAGCTGGCCGTGCAGGCACTGGTGTTGTACTGTGGCCACCAGCAGCTCCGTGAGGGG

9A       ..........................................................      1197
9B       ..........................................................      1197
10A      ......................G.............................A....A.......      1197
10B      ....................................C..............T.A.......      1197
12A      ..........................................................      1197
12B      ...........T.......................C..........................      1197
16A      ..........................................................      1197
16B      ..................................C..............T.A.......      1197
17A      ..................................C..............T.A.......      1197
17B      ..........................................................      1197
18A      ..........................................................      1197
18B      ..........................................................      1197
19A      ..................................C..............T.A.......      1197
20A      ..........................................................      1197
20B      ..................................................A...A...      1197
24A      ..........................................................      1197
24B      ..................................................A...A...      1197
25A      ...........T..............................................      1197
25B      ...........T..............................................      1197
32A      ......................G.........................A....A.......      1197
32B      ..........................................................      1197
42A      ..........................................................      1197
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority ACCCTCACTGCTGGXGGCCTCGTCGCCTTCATCCTCTACCAGAATAATGCCGGCAGCAGT

9A       ..............T.....T.......................................       1257
9B       ..............T.............................................       1257
10A      ..............T......A.......................G..............       1257
10B      ..............C......A.......................G..............       1257
12A      ..............C.....T.............................T.........       1257
12B      ..............C......A......................................       1257
16A      ..............C.....T.......................................       1257
16B      ..............C......A.......................G..............       1257
17A      ..............C......A.......................G..............       1257
17B      ..............T.............................................       1257
18A      ..............C.....T.......................................       1257
18B      ..............C.....T.......................................       1257
19A      ..............C......A.......................G..............       1257
20A      ..............T.............................................       1257
20B      ..............T......A.......................G..............       1257
24A      ..............T......A......................................       1257
24B      ..............T......A.......................G..............       1257
25A      ..............C.....T.......................................       1257
25B      ..............C.....T.......................................       1257
32A      ..............T......A.......................G..............       1257
32B      ..............T.............................................       1257
42A      ..............T.............................................       1257


Majority GTGCAGGCGCTGGCGTACGCCTACGGCGACCTGCTGAGCAATGTGGCAGCTGCCCACAAG

9A       .......................T...................................A       1317
9B       ..............A.............................................       1317
10A      ............................................................       1317
10B      ..........................T.................................       1317
12A      ..........................T.................................       1317
12B      .......A....................................................       1317
16A      ..........................T.................................       1317
16B      ............................................................       1317
17A      ..........................T.................................       1317
17B      ..............A.........A...................................       1317
18A      ..............A...........T.................................       1317
18B      .......................T....................................       1317
19A      ..........................T.................................       1317
20A      ........A...................................................       1317
20B      ........A...................................................       1317
24A      ..............A...........T.................................       1317
24B      ............................................................       1317
25A      ..............A...........T.................................       1317
25B      ..............A...........T.................................       1317
32A      ........A.....A.............................................       1317
32B      ..............A.............................................       1317
42A      ..............A.............................................       1317
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority ATCTTCGACTACCTGGACCGTGAGCCXGCTGTGGGCACTGGGGGCACCTCGGAGCCTGCC

9A       ........................G.............................       1377
9B       ........................G.T.....................A.........       1377
10A      ...........T.............A.............................       1377
10B      ........................G.............................       1377
12A      ........................G................T.............       1377
12B      ..................T.......A.......................T.........       1377
16A      ........................G.......................A.........       1377
16B      ........................A.....................T.........       1377
17A      ........................A.................A.............       1377
17B      ........................A.T.............................       1377
18A      ........................A.............................       1377
18B      ........................G.....................T.........       1377
19A      ........................G.............................       1377
20A      ........................G.............................       1377
20B      ........................G.....................T.........       1377
24A      ........................A.............................       1377
24B      ........................G.....................T.........       1377
25A      ........................A.............................       1377
25B      ........................A.............................       1377
32A      ...........T.............A.............................       1377
32B      ..................A......G.T...........................       1377
42A      ........................A.............................       1377


Majority ACACTGCAAGGCCACGTCACCTTCCAGCACGTCTCCTTTGCCTACCCCACGCGCCCCGAG

9A       ...............................................A......A..       1437
9B       ...........................................................       1437
10A      ...........................................................       1437
10B      ...........................................................A      1437
12A      ............G..............................................       1437
12B      ..............T.................T.................T....       1437
16A      ...........................................................       1437
16B      ...........................................................       1437
17A      ...........................................................       1437
17B      ...........................................................       1437
18A      ...........................................................A      1437
18B      ...........................................................       1437
19A      ...........................................................A      1437
20A      ...............................................T.....A..       1437
20B      .T............T.............................................       1437
24A      ...........................................................A      1437
24B      ...............................................A.........       1437
25A      ...........................................................T..A    1437
25B      ...........................................................T..A    1437
32A      ..............T.............................................       1437
32B      ...........................................................       1437
42A      ..............T.............................................A      1437
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority CACCTTGTCCTGCAGGATGTCTCCTTCGAGCTGCGCCCCGGTGAGGTGACAGCACTGGCA

9A       ............................................................       1497
9B       ........................T...................................       1497
10A      ............................................................       1497
10B      ......................T.....................................       1497
12A      ............................................................       1497
12B      ............................................................       1497
16A      ............................................................       1497
16B      ............................................................       1497
17A      ........................T...................................       1497
17B      ............................................................       1497
18A      ............................................................       1497
18B      ............................................................       1497
19A      ......................T.....................................       1497
20A      ...G........................................................       1497
20B      ........T...................................................       1497
24A      ............................................................       1497
24B      ........................T...................................       1497
25A      ............................................................       1497
25B      ............................T...............................       1497
32A      ............................................................       1497
32B      ............................................................       1497
42A      ............................................................       1497


Majority GGGCTGAACGGCAGTGGGAAGAGCACCTGTGCTGGGCTGCTGCAGCGCTTCTATGAGCCC

9A       ............................................................       1557
9B       ........T...................................................       1557
10A      ................................T...........................       1557
10B      ........T...................................................       1557
12A      ............................C...............................       1557
12B      ............................C...............................       1557
16A      ............................................................       1557
16B      ........T...................................................       1557
17A      ..........................................................C.....       1557
17B      ............................C...............................       1557
18A      ............................C...............................       1557
18B      ................................T...........................       1557
19A      ........T...................................................       1557
20A      ........T...................C...............................       1557
20B      ........T...................C...............................       1557
24A      ........T...................C...............................       1557
24B      ............................C.................C.....       1557
25A      ........T...................................................       1557
25B      ........T...................................................       1557
32A      ..........................................................C.....       1557
32B      ............................C...............................       1557
42A      ............................................................       1557
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority ACGGCTGGGGAGGTGCTGCTGGATGGGGTGCCACTGCGTGATTACAAGCACCGGTACCTA

9A       ..........................................................       1617
9B       ..........................................................       1617
10A      .....................................................A....       1617
10B      ..........................................................       1617
12A      ..A.......................................................       1617
12B      .....................................................T....       1617
16A      ..A.......................................................       1617
16B      ..........................................................       1617
17A      ................T.........................................       1617
17B      ..........................................................       1617
18A      ..A.......................................................       1617
18B      .....................................................A....       1617
19A      ..........................................................       1617
20A      ..........................................................       1617
20B      ..........................................................       1617
24A      ..........................................................       1617
24B      ..........................................................       1617
25A      ..........................................G...............       1617
25B      ..........................................G...............       1617
32A      ..........................................................       1617
32B      ..........................................................       1617
42A      ................T.........................................       1617


Majority CACCGCCAGGTGGCACTGGTGGGGCAGGAGCCXGTGCTCTTCTCTGGCTCCATCCGGGAC

9A       ..............................T...........................       1677
9B       ..............................C...........................       1677
10A      ..............................C...........................       1677
10B      ..............................T..................T........       1677
12A      ..............................T...........................       1677
12B      ..............................C...........................       1677
16A      ..............................T...........................       1677
16B      ..............................C...........................       1677
17A      ..............................T...........................       1677
17B      ..............................C...........................       1677
18A      ..............................T...........................       1677
18B      ..............................C...........................       1677
19A      ..............................T..................T........       1677
20A      ..............................C...........................       1677
20B      ..............................T...........................       1677
24A      ....A.........................T...........................       1677
24B      ..............................T...........................       1677
25A      ..................A...........C...........................       1677
25B      ..................A...........C...........................       1677
32A      ..............................C...........................       1677
32B      ..............................C...........................       1677
42A      ...T..........................T...........................       1677
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority AACATTGCCTATGGGCTGGAGGACTGCAGGGAGGAGGAGATCATAGCAGCTGCAGAGGCT

9A       ..........................................................      1737
9B       ..........................................................      1737
10A      ..........................................................      1737
10B      ..........................................................      1737
12A      ..........................................................      1737
12B      ..........................................................      1737
16A      ........................................G....G............      1737
16B      ..........................................................      1737
17A      ..........................................................T..      1737
17B      ..........................................................      1737
18A      ..........................................................      1737
18B      ..........................................................      1737
19A      ..........................................................      1737
20A      ..........................................................      1737
20B      ..............T...........................................      1737
24A      ..........................................................      1737
24B      ..............T...........................................      1737
25A      ..........................................................      1737
25B      ..........................................................      1737
32A      ..........................................................      1737
32B      ..........................................................      1737
42A      ............................C.............................      1737


Majority GCTGGTGCCTTGGATTTTATCTCTGCACTGGACCAAGGCTTTGACACTGATGTAGGGGAG

9A       ..............C.............................................      1797
9B       ..............C.............................................      1797
10A      ..........................................................      1797
10B      ..............C.............................................      1797
12A      ........................................A.................      1797
12B      ..........................................................      1797
16A      ..........................................................      1797
16B      ..........................................................      1797
17A      ..........................................................      1797
17B      ..............C.............................................      1797
18A      ..........................................................      1797
18B      ..........................................................      1797
19A      ..............C.............................................      1797
20A      ..............C.............................................      1797
20B      ..........................................................      1797
24A      ..................................................C.........      1797
24B      ..........................................................      1797
25A      ..............................................AT.C.........      1797
25B      ..............................................AT.C.........      1797
32A      ..............C.............................................      1797
32B      ..........................................................      1797
42A      ..............C.............................................      1797
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority AGAGGAGATCAGCTCTCGGCTGGGCAGAAGCAGCGTGTTGCCATTGCCCGGGCCTTGGTG

9A       ................A.......................................       1857
9B       ................................................................       1857
10A      ................................................................       1857
10B      ....................................A.........................       1857
12A      ................A.......................................       1857
12B      ................................................................       1857
16A      ................................................................       1857
16B      ................................................................       1857
17A      ................................................................       1857
17B      ................................................................       1857
18A      ................................................................       1857
18B      ................................................................       1857
19A      ....................................A.........................       1857
20A      ................................................................       1857
20B      ................................................................       1857
24A      ................A.......................................       1857
24B      ................................................................       1857
25A      ..............T.........................................       1857
25B      ..............T.........................................       1857
32A      ................A.....................A.....................       1857
32B      ................................................................       1857
42A      ....................................A.........................       1857


Majority CGATGCCCCACTGTCCTCATCCTCGACGAGGCCACCAGTGCTCTGGATGGGGATGGGGAT

9A       ......................T.................................       1917
9B       .........G..............................................       1917
10A      ......................T.................................       1917
10B      ..................................A.....................       1917
12A      ................................................................       1917
12B      ................................................................       1917
16A      ..................................T.....................       1917
16B      ................................................................       1917
17A      ................................................................       1917
17B      ................................................................       1917
18A      ................................................................       1917
18B      ......................T.................................       1917
19A      ..................................A.....................       1917
20A      ................................................................       1917
20B      ................................................................       1917
24A      ................................................................       1917
24B      ................................................................       1917
25A      .........G..............................................       1917
25B      .........G..............................................       1917
32A      ................................................................       1917
32B      ................................................................       1917
42A      ................................................................       1917
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GTGCCGCTGCAGCAGTGGGTGCAGAGTGGGGGGACCGGACGGTGCTGCTCATCACCCAC

9A       .....................................................T......       1977
9B       ..AG........................................................       1977
10A      ............................................................       1977
10B      ............................CA..............................       1977
12A      ...G............................A........A...A..............       1977
12B      ............................................................       1977
16A      .........................C..................................       1977
16B      ............................................................       1977
17A      ............................................................       1977
17B      ............................................................       1977
18A      ...G........................................................       1977
18B      ............................................................       1977
19A      ............................CA..............................       1977
20A      ............................................................       1977
20B      ...G............................A........A...A..............       1977
24A      ............................................................       1977
24B      .....................................................T......       1977
25A      ..AG........................................................       1977
25B      ..AG..---------.............................................       1968
32A      .........................................A..................       1977
32B      ............................................................       1977
42A      ...G........................................................       1977


Majority CACCCGCGGATGCTGGAGAAGGCCGACCGCGTCGTGGTGCTGGAGCATGGCACCGTGGCT

9A       ............................................................       2037
9B       .......................T........T...........................       2037
10A      ...........................T.........................T......       2037
10B      ............................................................       2037
12A      ............................................................       2037
12B      ..............................................G.............       2037
16A      ............................................................       2037
16B      ..............................................G.............       2037
17A      .............................T..............................       2037
17B      ...........................T................................       2037
18A      ..............................A.............................       2037
18B      ...........................T.........................T......       2037
19A      ............................................................       2037
20A      .............................T..............................       2037
20B      .................................A...........TG.............       2037
24A      ............T...............................................       2037
24B      ............................A.................G.............       2037
25A      ............T...............................................       2037
25B      ............T...............................................       2028
32A      .....................................................T......       2037
32B      ................................T...........................       2037
42A      ............................................................       2037
```

Figure A-1: TAP2 nucleotide sequence alignment (continued).

```
Majority GAGATGGGGACGCCCGCTGAGCTCGCGGCCTGTGGTGGGCCCTACAGCCGCCTGCTGCAG

9A       ..............T...........A.................................       2097
9B       ...........A.................................................       2097
10A      ..............T..............................................       2097
10B      ..................................................... ......       2092
12A      ...........................................................       2097
12B      ..............T..............................................       2097
16A      ........................A.....................................       2092
16B      ..............T..............................................       2097
17A      ...........................................................       2097
17B      ...........T..T.........T.....................................       2097
18A      ........A..A..............A...................................       2097
18B      ..............T..............................................       2097
19A      ...........................................................       2097
20A      ...........................................................       2097
20B      .....T.................T.............................. ......       2092
24A      ........................A.....................................       2097
24B      ...........................................................       2097
25A      ...........................................................       2097
25B      ...........................................................       2088
32A      ..............T..............................................       2097
32B      ...........................................................       2097
42A      ...........................................................       2097


Majority CGCTAA

9A       ......                                                             2103
9B       ......                                                             2103
10A      ......                                                             2103
10B                                                                         2092
12A      ......                                                             2103
12B      ......                                                             2103
16A                                                                         2092
16B      T.....                                                             2103
17A      ......                                                             2103
17B      ......                                                             2103
18A      ......                                                             2103
18B      ......                                                             2103
19A      ......                                                             2103
20A      ......                                                             2103
20B                                                                         2092
24A      ......                                                             2103
24B      ......                                                             2103
25A      ......                                                             2103
25B      ......                                                             2094
32A      ......                                                             2103
32B      ......                                                             2103
42A      ......                                                             2103
```

Figure A-1: TAP2 nucleotide sequence alignment. The CDS of all TAP2 alleles were aligned with CLUSTALW in MegAlign (DNAStar). Dots represent nucleotides that match the majority. Dashes represent gaps in the sequences.

```
Majority ATGGGCGCCGGCCGCCGCCTCCTGGAGTCGCTGCGCCCCGAGCGCTGGCGCTGCGCGGTG

9A       ------------------------------------------------------------          0
9B       ------------------------------------------------------------          0
10A      ------------------------------------------------------------          0
10B      ------------------------------------------------------------          0
12A      ------------------------------------------------------------          0
12B      ------------------------------------------------------------          0
16A      ------------------------------------------------------------          0
16B      ------------------------------------------------------------          0
17A      ------------------------------------------------------------          0
17B      ------------------------------------------------------------          0
18A      ............................................................         60
18B      ...........................................................T....         60
19A      ............................................................         60
19B      ..........................................................T.T....         60
20A      ------------------------------------------------------------          0
20B      ...........................................................T....         60
24A      ............................................................         60
24B      ............................................................         60
25A      ------------------------------------------------------------          0
25B      ------------------------------------------------------------          0
32A      ............................A...........................         60
32B      ...........................................................T....         60
42A      ............................................................         60
42B      ...........................................................T....         60


Majority GTGGCGGCGCTGATGGGGGCCTCGTCGCTCGGGGAGATGGCCATCCCCTACTACACGGGG

9A       ------------------------------------------------------------          0
9B       ------------------------------------------------------------          0
10A      ----------..................................................         50
10B      ----------..................................................         50
12A      ----------..................................................         50
12B      ----------..................................................         50
16A      ----------..................................................         50
16B      ----------..................................................         50
17A      ----------..................................................         50
17B      ----------..................................................         50
18A      ............................................................        120
18B      .....A..............T.......................................        120
19A      ...................T........................................        120
19B      .....A..............T.......................................        120
20A      ----------..................................................         50
20B      ...................T........................................        120
24A      ...................T........................................        120
24B      ............................................................        120
25A      ----------..................................................         50
25B      ----------..................................................         50
32A      ............................................................        120
32B      ............................................................        120
42A      ............................................................        120
42B      ...................T........................................        120
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority CGCGCCAGCGACTGGGTGGCCCGCGAGGACGAGCTGGCGGCCATCTGGCCCATGGTGCTG

9A       ------------------------------------------------------------          0
9B       ------------------------------------------------------------          0
10A      ......................A......T.............................        110
10B      ..T.....................................................              110
12A      ........................................................              110
12B      ........................................................              110
16A      ........................................................              110
16B      ........................................................              110
17A      ............................T.............................        110
17B      ..T.........................T.............................        110
18A      ..T.....................................................              180
18B      ........................................................              180
19A      ............................T.............................        180
19B      ............................T.............................        180
20A      ........................................................              110
20B      ........T.................................................A...        180
24A      ........................................................              180
24B      ..T.....................................................              180
25A      ......................A......T.............................        110
25B      ........T.....................T.............................        110
32A      ............................T.............................        180
32B      ..T.........................T.............................        180
42A      ........................................................              180
42B      ........................................................              180


Majority CTGGGGCTGAGCAGCGCCGTCACCGAGCTGGCCTGCGACATCGCCTTCGTGGGGACGCTG

9A       ------------..............................................           48
9B       ------------..............................................           48
10A      ..................A.......................................          170
10B      ..................................T........G...........A...        170
12A      ........................................................              170
12B      ........................................................              170
16A      ........................................................              170
16B      ........................................................              170
17A      ........................................................T..          170
17B      ..................A.......................................          170
18A      ....................................T.....................T..        240
18B      ........................................................              240
19A      ........................................................              240
19B      ........................................................T....        240
20A      ........................................................              170
20B      ..................A.......................................          240
24A      ..............T...A.......................................          240
24B      ..................................................C.......          240
25A      ..................A.......................................          170
25B      ........................................................              170
32A      ..............T.........................................              240
32B      ..............................................T.............        240
42A      ........................................................              240
42B      ........................................................              240
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majoity AGCCGCGTGCACGGCCACCTGCAGCGCCGCCTCTTCGCCGCCGTCCTCCGCGGGGACATC

9A        ...........T................................................        108
9B        ...........TA..............................................T        108
10A       .........................................T........T.........        230
10B       ......................................T.....................        230
12A       ...........T..............................A.................        230
12B       .........................................T..................        230
16A       .....T...................................T........T.........        230
16B       .....T.....T................................................        230
17A       ............................................................        230
17B       ............................................................        230
18A       ............................................................        300
18B       .........................................T..................        300
19A       ......................................T.....................        300
19B       ............................................................        300
20A       ...........T..........................T.....................        230
20B       ............................................................        300
24A       ...........T................................................        300
24B       ...........T..........................T.....................        300
25A       ...........T................................................        230
25B       .........................................T..................        230
32A       ............................................................        300
32B       ...........T.......................T........................        300
42A       ...................................T.GG.....................        300
42B       ............................................................        300


Majority GCTGAGCTGCGGGACGAGGGGGCCGGGGATGTGGCAGCACGGGTGACAGGGGACGTGGAG

9A        ............................................................        168
9B        ..C.........................................................        168
10A       ...............................................G............        290
10B       ..C.........................................................        290
12A       ..C.........................................................        290
12B       ..C............................................G............        290
16A       ..............T..........................T..A.TG......A.....        290
16B       .......................T.......................G............        290
17A       ..C......................................A.....G............        290
17B       ..C.........................................................        290
18A       ..C......................................A.....G............        360
18B       ..C............................................G............        360
19A       .......................T....................................        360
19B       ............................................................        360
20A       ............................................................        290
20B       .......................T....................................        360
24A       ..C......................................A.....G............        360
24B       ............................................................        360
25A       ..C.........................................................        290
25B       ...............................................G............        290
32A       A........................................A..................        360
32B       ............................................................        360
42A       ..........................A.................................        360
42B       ......................................G.....................        360
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority GCCACCCGTGAGGCCCTGGGCGAGGCGCTGAGCCTGCTGCTGTGGTACCTGGCGCGGGGC

9A      ..........................................................      228
9B      ..........................................................      228
10A     ..........................................................      350
10B     ..........................................................      350
12A     ..........................................................      350
12B     ..........................................................      350
16A     ............A.......A.....................................      350
16B     ..........................................................      350
17A     ..........................................................      350
17B     ..........................................................      350
18A     ..........................................................      420
18B     ..........................................................      420
19A     ..........................................................      420
19B     ..........................................................      420
20A     ..........................................................      350
20B     ..........................................................      420
24A     ..........................................................      420
24B     ..........................................................      420
25A     ..........................................................      350
25B     ..........................................................      350
32A     ..........................................................      420
32B     ..................................................A.......      420
42A     ..................................................A.......      420
42B     ..........................................................      420


Majority ATCTGCCTCTTCAGCACCATGGCCTGGCTGTCCCCGCGCATGGCCCTCGTCACCGCCCTG

9A      ..........................................................      288
9B      ..........................................T.................      288
10A     ..........................................T.................      410
10B     ..................................A.......................      410
12A     ..........................................................      410
12B     ..........................................................      410
16A     ....................................A.....................      410
16B     ................................G..............T.............      410
17A     ..........................................................      410
17B     ..................................A.......................      410
18A     ..........................................................      480
18B     ..........................................................      480
19A     ..........................................A...............      480
19B     ..........................................T.................      480
20A     ..................................A.......................      410
20B     ..................................................T.......      480
24A     ..................................................T.......      480
24B     ..................................................T.......      480
25A     ..........................................................      410
25B     ..........................................................      410
32A     ..........................................................      480
32B     ..........................................T.................      480
42A     ..........................................................      480
42B     ..................................................T.......      480
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority GTGCTGCCCTTCCTCCTGTTGCTGCCCAGGGCCATCGGCAGGCTCCAGCAGGGCCTGGCT

9A       ............................................................       348
9B       ............................................................       348
10A      ............................................................       470
10B      ............................................................       470
12A      ............................................................       470
12B      ............................................................       470
16A      ............................................................       470
16B      ............................................................       470
17A      ............................................................       470
17B      ............................................................       470
18A      ............................................................       540
18B      ............................................................       540
19A      ............................................................       540
19B      ............................................................       540
20A      ............................................................       470
20B      ............................................................       540
24A      ............................................................       540
24B      ............................................................       540
25A      ............................................................       470
25B      ............................................................       470
32A      ............................................................       540
32B      ............................................................       540
42A      ............................................................       540
42B      ............................................................       540


Majority CCAAAGGTGCAGAAGGCGCTGGCCCACGCCAGCGAGGTGGCAGTGGAGACCTTCCAGGCC

9A       ..............................T.............................       408
9B       ............................................................       408
10A      ............................................................       530
10B      ............................................................       530
12A      ............................................................       530
12B      ............................................................       530
16A      ............................................................       530
16B      ............................................................       530
17A      ............................................................       530
17B      ........................................G...................       530
18A      ............................................................       600
18B      ............................................................       600
19A      ................A...........................................       600
19B      ............................................................       600
20A      ............................................................       530
20B      ..............................T.............................       600
24A      ............................................................       600
24B      ............A.........................G...................       600
25A      ............................................................       530
25B      ..............................T.............................       530
32A      ............................................................       600
32B      ..............................T.............................       600
42A      ........................T...................................       600
42B      ............................................................       600
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority ATGGCCACTGTGCGCAGCTTCGCCAACGAGGACGGGGTGGCCGCACGCTACCGCCAGTGC

9A       ............................................................       468
9B       ............................................................       468
10A      ............................................................       590
10B      ..................T...........T...........G..............       590
12A      ............................................................       590
12B      ............................................................       590
16A      ............................................................       590
16B      ..............................T.............................       590
17A      ............................................................       590
17B      ............................................................       590
18A      ............................................................       660
18B      ..................T...........T...........G..............       660
19A      ............................................................       660
19B      ............................................................       660
20A      ..................T.....................T..................       590
20B      ........C..............................T..G..............       660
24A      ............................................................       660
24B      ............................................................       660
25A      ............................................................       590
25B      ............................................................       590
32A      ........................T...................................       660
32B      ........C...................................................       660
42A      ............................................................       660
42B      ........C...........T.......................G..............       660


Majority CTGCACCAGACCCACCAGCTGGAGAAGCAGGATGCAGTCCTCTACGCCGTCTCTGCCTGG

9A       ............................................................       528
9B       ............................................................       528
10A      ............................................................       650
10B      ..................................G.........................       650
12A      ................................................T...........       650
12B      ............................AA..............................       650
16A      .....T.....................................................       650
16B      ............................A.............T................       650
17A      ................................................T...........       650
17B      ............................................................       650
18A      ............................................................       720
18B      ............................A.............T................       720
19A      ............................A.................T............       720
19B      ............................A...............................       720
20A      ............................................................       650
20B      ..................................G.........................       720
24A      ............................................................       720
24B      ............................................................       720
25A      ............................................................       650
25B      ............................................................       650
32A      ................................................T...........       720
32B      ..................................G.........................       720
42A      ......A.....................................................       720
42B      ............................................................       720
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority ACCAGTGGTTTCTCAGCGCTGGCCCTGAAGATGGGGATCCTCTACTACGGGGGACAGCTG

9A       ..........................................................       588
9B       ..........................................................       588
10A      ................A.........................................       710
10B      ..........................................................       710
12A      ......................T...................................       710
12B      ..........................................................       710
16A      ..........................................................       710
16B      ..........................................................       710
17A      ..........................................................       710
17B      ..........................................................       710
18A      ..........................................................       780
18B      ..........................................................       780
19A      ..........................................................       780
19B      ..........................................................       780
20A      ..........................................................       710
20B      ..........................................................       780
24A      ..........................................................       780
24B      ..........................................................       780
25A      ..........................................................       710
25B      ..........................................................       710
32A      ..........................................................       780
32B      ..........................................................       780
42A      ..........................................................       780
42B      ..........................................................       780


Majority GTGGCCGCGGGGACCATCAGTACCGGGGACCTCGTCACCTTCCTCCTCTACCAGTTGCAG

9A       ..........................................................       648
9B       ..........................................................       648
10A      ........................A.....T.............................       770
10B      ....................C..T.....................................       770
12A      ....................C..T........T...........T...............       770
12B      ..........................................................       770
16A      ..........................................................       770
16B      ..........................................................       770
17A      ..........................................................       770
17B      ..........................................................       770
18A      ....................C...........T...........................       840
18B      ..........................................................       840
19A      ..........................................................       840
19B      ......................T...................................       840
20A      ..........................................................       770
20B      ..........................................................       840
24A      ....................C..T........T...........T...............       840
24B      .....T..............C.....................................       840
25A      .........A..............................................           770
25B      ..........................................................       770
32A      ..........................................................       840
32B      ..........................................................       840
42A      ..........................................................       840
42B      ..........................................................       840
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority TTCACTGATGTCGTGGAGGTCCTGCTCCGCTACTACCCCACGCTGACGAAGGCTGTGGGC

9A       ....................................T.................       708
9B       ....................................T.................       708
10A      ........................................................       830
10B      ........................................................       830
12A      ....................................T.................       830
12B      ........................................................       830
16A      ........................................................       830
16B      ........................................................       830
17A      ........................................................       830
17B      ....................................T.................       830
18A      ....................................T.................       900
18B      ........................................................       900
19A      ........................................................       900
19B      ........................................................       900
20A      ........................................................       830
20B      ........................................................       900
24A      ....................................T.................       900
24B      ........................................................       900
25A      ........................................................       830
25B      ........................................................       830
32A      ........................................................       900
32B      ....................................T.................       900
42A      ..........T.............................................       900
42B      ........................................................       900


Majority TCCTCGGAGAAGATCTTTGAGTTCCTGGACCGGGAGCCACAGGTCGCACCCTCGGGGACA

9A       ......................T.................................       768
9B       ........................................................       768
10A      ........................................................       890
10B      ............A.............................G.....A......       890
12A      ........................................................       890
12B      ........................................................       890
16A      ..............................................G.........       890
16B      ......................T.................................       890
17A      ..................................................A.....       890
17B      ......................T.................................       890
18A      ........................................................       960
18B      ........................................................       960
19A      ........................................................       960
19B      ........................................................       960
20A      ..............................................G.........       890
20B      ..............................................G.........       960
24A      ........................................................       960
24B      ........................................................       960
25A      ........................................................       890
25B      ..............................................G.........       890
32A      ........................................................       960
32B      ........................................................       960
42A      .....A..................................................       960
42B      ..............................................G.........       960
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority CTAGCACCCAGTGACCTGCGGGGCCACCTCCAGCTTGAGGACGTCTGGTTCTCCTACCCT

9A       ........G.....................................................       828
9B       ........G.....................................................       828
10A      ............T.................................................       950
10B      ............................................................       950
12A      ............T.................................................       950
12B      ............................................................       950
16A      ............................................................       950
16B      ............................................................       950
17A      ............T.................................................       950
17B      ........G.....................................................       950
18A      ............T.................................................      1020
18B      ............T.................................................      1020
19A      ............................................................      1020
19B      ............................................................      1020
20A      ......................................T.......................       950
20B      .........C....................................................      1020
24A      ............T.................................................      1020
24B      ............................................................      1020
25A      ............................................................       950
25B      ............................................................       950
32A      ............................................................      1020
32B      ........G.....................................................      1020
42A      ............................................................      1020
42B      ......................................T.......................      1020


Majority GGGCGCCAGGAGCCTGTCCTCAAGGGTGTGTCCCTGGAGCTGCGCCCCGGGGAGGTGCTG

9A       ..............................................T.............       888
9B       ..............................................T.............       888
10A      ............................................................      1010
10B      ............................................................      1010
12A      ............................................................      1010
12B      ............................................................      1010
16A      ..............................................T.............      1010
16B      ......................................T.......................      1010
17A      ............................................................      1010
17B      ............................................................      1010
18A      ............................................................      1080
18B      ............................................................      1080
19A      ............................................................      1080
19B      ............................................................      1080
20A      ............................................................      1010
20B      ............................................................      1080
24A      ............................................................      1080
24B      ............................................................      1080
25A      ............................................................      1010
25B      ..............................................................T..      1010
32A      ..............................................T.........T..      1080
32B      ............................................................      1080
42A      ............................................................      1080
42B      ............................................................      1080
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority GCACTGATGGGGCCCCCGGGCGCGGGGAAGAGCACCCTGGTGTCCCTGCTGCTGCGCCTG

9A       ...T.....................................................        948
9B       ...T.....................................................        948
10A      .........................................................       1070
10B      .........................................................       1070
12A      ......................T..................................       1070
12B      .........................................................       1070
16A      .........................................................       1070
16B      .........................................................       1070
17A      .........................................................       1070
17B      .........................................................       1070
18A      ..................A..T.T.................................       1140
18B      ..................A..T...................................       1140
19A      .........................................................       1140
19B      .........................................................       1140
20A      ....................T....................................       1070
20B      .........................................................       1140
24A      ......................T..................................       1140
24B      ......................T..................................       1140
25A      .........................................................       1070
25B      .........................................................       1070
32A      ........................A................................       1140
32B      .........................................................       1140
42A      .........................................................       1140
42B      .........................................................       1140


Majority CACCAGCCCACGGCCGGGCGCCTGCTCTTGGACGGCCATCCCCTCAGCGCCTACCGCCAC


9A       ..............................                                  981
9B       ..............................                                  981
10A      ..............................                                 1103
10B      ..............................                                 1103
12A      ..............................                                 1103
12B      ..............................                                 1103
16A      ..............................                                 1103
16B      ..............................                                 1103
17A      ..............................                                 1103
17B      ..............................                                 1103
18A      ...................................A.........................       1200
18B      .........................................................       1200
19A      .........................................................       1200
19B      .........................................................       1200
20A      ..............................................T............       1130
20B      ...........................................T................       1200
24A      .........................................................       1200
24B      .........................................................       1200
25A      ..............................                                 1103
25B      ..............................                                 1103
32A      ..............................................T............       1200
32B      .........................................................       1200
42A      .........................................................       1200
42B      .........................................................       1200
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority TCCTACCTGTGCCACCGGGTGGCCACCGTCCCCCAGGAGCCGGTGCTCTTTGCCCGCTCA

18A      ..........................................................      1260
18B      ..................................................C.....      1260
19A      ..........................................................      1260
19B      ................A.........................................C.....      1260
20A      ..........................................................      1190
20B      ..................................................C.....      1260
24A      ..................................................C.....      1260
24B      ..................................................C.....      1260
32A      ..........................................................      1260
32B      ..................................................C.....      1260
42A      ..........................................................      1260
42B      ..........................................................      1260


Majority CTCCACGCCAACATCGCCTATGGGCCAGAGAGCTGGAGCCGGGCACAGGTGACAGCGGCA

18A      ..................C.....................G.............      1320
18B      .....T.....................................................      1320
19A      ..................C.....................G.............      1320
19B      ..........................................................      1320
20A      .....T.....................................................      1250
20B      ..........................................................      1320
24A      ..........................................................      1320
24B      ..........................................................      1320
32A      ..........................................................      1320
32B      ..........................................................      1320
42A      ..........................................................      1320
42B      ..................C.....................................      1320


Majority GCCCGCCGGGTGGGTGCCCACAACTTCATCACCCGCCTGCCCCGAGGCTATGACACAGAG

18A      ............C........T.....................................      1380
18B      ...T.......................................................      1380
19A      ............C........T.....................................      1380
19B      ....A.....................................A.............      1380
20A      ...T.......................................................      1310
20B      ....A.....................................A.............      1380
24A      ...T.................................T.....................      1380
24B      ..........................................................      1380
32A      ......................T...................................      1380
32B      ..........................................................      1380
42A      ......A....................................................      1380
42B      ............C..............................................      1380


Majority GTGGGTGAGCTTGGGGGGCAGCTCTCCGGGGGACAGCGGCAAGGGGTGGCCATCGCCCGT

18A      ..........................................................      1440
18B      ................A.........................................      1440
19A      ..........................................................      1440
19B      ..........................................................      1440
20A      ................A.........................................      1370
20B      ..........................................................      1440
24A      ..........................................................      1440
24B      ..........................................................C     1440
32A      ................A.........................................      1440
32B      ................A.........................................      1440
42A      ..........................................................      1440
42B      ..........................................................      1440
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority GCCCTGGTGCGGGACCCTCACATCCTTGTCCTTGACGAGCCCACCAGCGCCCTGGATGCT

18A      ...............................................T............     1500
18B      ..T................................T........................     1500
19A      ...............................................T............     1500
19B      ...................................T........................     1500
20A      ..T................................T........................     1430
20B      ...................................T........................     1500
24A      ........................................................C...     1500
24B      ............................................................     1500
32A      ........................................................C...     1500
32B      ...............................................T............     1500
42A      ........................................................C...     1500
42B      ..T................................T........................     1500


Majority GAGTGCCAGCAGCAGGTGGAGCGGGAGCTCTTCGAGGACAGCGGGACCAGGCGCACGGTG

18A      ............................................................     1560
18B      ..................................................A.........     1560
19A      ............................................................     1560
19B      ............................................................     1560
20A      ..................................................A.........     1490
20B      ........................................................A...     1560
24A      ..................................................A.........     1560
24B      ............................................................     1560
32A      ............................................................     1560
32B      ............................................................     1560
42A      ..................................................A.........     1560
42B      ........................................................A...     1560


Majority CTGCTGGTGACAGGGCAGGTGGCCCTGGCAGCACGGGCACACCAGGTGGCCCTGATGGAG

18A      ............................................................     1620
18B      ............................................................     1620
19A      ............................................................     1620
19B      ............................................................     1620
20A      ............................................................     1550
20B      ............................................................     1620
24A      ............................................................     1620
24B      ............................................................     1620
32A      .....................................................A......     1620
32B      ............................................................     1620
42A      ............................................................     1620
42B      ............................................................     1620


Majority GGGGGGCAGGTACGCGAGCAGGGCCCCCCCAGCAAGCTGCTGCGCCCCGGCAGCCGCTAC

18A      ........A.............................T.....................     1680
18B      .......................T.......................T............     1680
19A      ........A.............................T.....................     1680
19B      ..........................................T.................     1680
20A      .......................T.......................T............     1610
20B      ........A...................................................     1680
24A      .......................T....................................     1680
24B      ..............T................................T............     1680
32A      ............................................................     1680
32B      ......................................T.....................     1680
42A      .......................T.......................T............     1680
42B      ........A...................................................     1680
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority TGGCACCTGCTGCAGGATGGGGGACGCCAGGGGACATTGGGGGATGGGGACAAGGGCACA

18A      ..........................................................      1740
18B      ......................A...................................      1740
19A      ..........................................................      1740
19B      ..........................................................      1740
20A      ....................G.....................................      1670
20B      ..........................................................      1740
24A      ..........................................................      1740
24B      ..........................................................      1740
32A      ..........A...............................................      1740
32B      ..........................................................      1740
42A      ..........................................................      1740
42B      ..........................................................      1740


Majority GGGAGTGGGGATGAGGGACAGCAGGAGCCAGGGATGGGGACGGTACAAGCCAGAATGGAG

18A      ..........................................................A..      1800
18B      ........................................A...T..............      1800
19A      ..........................................................A..      1800
19B      ................G.....................................G......      1800
20A      ................G.................A.....................      1730
20B      ..........................................................A..      1800
24A      ..........................................................      1800
24B      ................G.................A...T..................      1800
32A      ................G.....................A..................      1800
32B      ................G.............A.........................      1800
42A      ..................................................G.........      1800
42B      ..........................................................A..      1800


Majority ACATCAGGGGATGGGGACXXXXXXXXXXXXXXXXXXACAAGGAGCTGGGATGGGGACACT

18A      ................-----------------.......................      1842
18B      ................-----------------.......................      1842
19A      ................-----------------.........A...........C      1842
19B      ............A..-----------------..G.......A...........C      1842
20A      ................GCTGGGCAGCACTGAGCAG.......................      1790
20B      ................-----------------.......................      1842
24A      ................-----------------.......................      1842
24B      ................GCTGGGCAGCACTGAGCAG.............G.........      1860
32A      ............T..-----------------------------------------      1817
32B      ................-----------------.........A...........A      1842
42A      ................-----------------.......................      1842
42B      ................-----------------.......................      1842


Majority GGGAGCCAGGATGGTGACACTGAGAGCTGGGATGGTGACACAAGGGCCAGGGATGGGGAC   \

18A      ....................................G.....G...........      1902
18B      ................................C.....................      1902
19A                                                                  1843
19B      A                                                           1843
20A      ..........................................................      1850
20B      ....................................G.....G...........      1902
24A      ..........................................................      1902
24B      ..........................................................      1920
32A      ----------------.....................................A....      1860
32B      ---------------------------------------.............A....      1835
42A      ....................................G.....G...........      1902
42B      ....................................G.....G...........      1902
```

Figure A-2: TAP1 nucleotide sequence alignment (continued).

```
Majority ACCAAGGGCAGAGGTATAGGTGGGGACACTACTACCCCTTGGGATGGAGACACCAGGAGC

18A      ............................................................      1962
18B      ............................................................      1962
20A      ...............................T............................      1910
20B      ...................A........................................      1962
24A      ............................................................      1962
24B      ...............................T..............G.......      1980
32A      G...........................................................      1920
32B      G...........................................................      1895
42A      ............................................................      1962
42B      ............................................................      1962
```

Figure A-2: TAP1 nucleotide sequence alignment. The CDS of all TAP1 alleles were aligned with CLUSTALW in MegAlign (DNAStar). Dots represent nucleotides that match the majority. Dashes represent gaps in the sequences.

```
Majority MELLPTLRLACVLLLADLVVLAALARLAPALAQLGLVATWLEAGLRLPVLVGAGRLLAPGGPRGAAALVSLAPATFLTLRGCLELPGAPPVLLAMATPSWLALAY

9A       ..............................................................................................L.......     105
9B       .......................Q.......................................................................     105
10A      ..................................................D.....................................     105
10B      ..........................................................P.....................R..........     105
12A      ..................................................M.......Q.....................................     105
12B      ..................................................M.....................................     105
16A      ..........................................................P.....................................     105
16B      .......................Q.......................M.................................S..............     105
17A      .......................Q.......................M.....................................     105
17B      ..........................................................P.....................................     105
18A      ..................................................M.....................................     105
18B      .......................Q.......................M.....................................     105
19A      ..........................................................P.....................R..........     105
19B      ..........................................................P.....................................     105
20A      .......................Q.......................................................................     105
20B      ..................................................M.....................................     105
24A      ..........................................................P.....................................     105
24B      .......................Q.......................M.....................................     105
25A      ..............................................................................................L.......     105
25B      .......................W.............................R.....................................     105
32A      .......................................................................................     105
32B      .......................Q.......................................................................     105
42A      .......................Q.......................M.....................................     105
42B      ..........................................................P.....................................     105
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

```
Majority GAVLLALLTWTSLAPGVALGTKEVKYQAALRRQLALAWPEWPFLSGAFFCLVLAALGETSVPYCTGKALDVLRHGDGPTAFATAIGFVCLASASSSLFAGCRGGL

9A       ..............................................................I.................................     210
9B       ..........................................F.....................................................     210
10A      ......................................L.....F.....................................................S.    210
10B      ..........................................F.....................................................S.....     210
12A      ..........................................F.....................................................     210
12B      ..........................................F.....................................................     210
16A      ..........................................F.....................................................     210
16B      ..........................................F.....................................................     210
17A      ..........................................F.....................................................S.    210
17B      ..........................................F.....................................M.................     210
18A      ..........................................F.....................................................S.    210
18B      ........................E.....................................S.....................S.    210
19A      ..........................................F.....................................................S.....     210
19B      ..........................................F.....................I.............                          190
20A      ..........................................F.....................................................     210
20B      ..........................................F.....................S.....................     210
24A      ..........................................F.....................................................S.    210
24B      ..........................................F.....................................I.....................     210
25A      ..........................................F.....................I.....................     210
25B      ..........................................F.....................................................     210
32A      ..........................................F.....................................M.................     210
32B      ..........................................F.....................................................     210
42A      ........................E.....................................S.....................S.    210
42B      ..........................................F.....................................................     210
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

```
Majority FTFIMARLTLRTRDRLFSGLVHQDLAFFQQTTAADLASRLATDVPLASRVVPGSANIALRNLGKVLGLSAFMLALSPRLTLLALLEVPLAIVARKIYNARYQMLQ

9A       ..............Q..........................................................................................   315
9B       .S.......................................................................................................   315
10A      ....................................................F....................................................   315
10B      .....................Q..............................F......................................A.H.V....H....   315
12A      .S.......................................................................................................   315
12B      ..................CM...............................SRI....P...........................M....A.H.V.........   315
16A      .C.......................................................................................................   315
16B      ...........................................................................................A.H.V....H....   315
17A      ....................................................F.V..................................................   315
17B      ..............Q..........................................................................................   315
18A      ....................................................F....................................T...............   315
18B      ....................................................F......................................A.H.V....H....   315
19A      .....................Q..............................F......................................A.H.V....H....   315
20A      .S............Q..........................................................................................   315
20B      ..............Q.................................D............................................H...........   315
24A      ....................................................F.V..................................................   315
24B      .S......................................................................................P................   315
25A      ..............Q..........................................................................................   315
25B      ..............Q.................................D........................................................   315
32A      ................................T..........................................................A.H.V....H....   315
32B      .S............Q..........................................................................................   315
42A      .........................................................................................................   315
42A      .........................................................................................................   315
42B      .C.......................................................................................................   315
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

```
Majority QAVLDATAGTGMVVQEAISSIETVRAFAGEEEEERRYGQALTKMLGLRDQRDMEKAIFLLIRRALQLAVQALVLYCGHQQLREGTLTAGGLVAFILYQNNAGSSV

9A       ....................................................................................................  420
9B       ....................................................................................................  420
10A      ..........................................................................H.........I.......D.....  420
10B      ........................P.............................T.....V.....................H.........I.......D.....  420
12A      ....................................................................................................  420
12B      ...........V.................................R.....................................I.............  420
16A      ....................................................................................................  420
16B      ..................................................................Q...................H.........I.......D.....  420
17A      ............................T.......................................................H.........I.......D.....  420
17B      ....................................................................................................  420
18A      ....................................................................................................  420
18B      ....................................................................................................  420
19A      ........................P.............................T.....V.....................H.........I.......D.....  420
20A      ....................................................................................................  420
20B      ...........V.........................S.......................................H.........I.......D.....  420
24A      ..................................................................................I.............  420
24B      ..........................................................................H.........I.......D.....  420
25A      ....................................................................................................  420
25B      ...........V........................................................................................  420
32A      ..................................................T.....V.....................H.........I.......D.....  420
32B      ....................................................................................................  420
42A      ....................................................................................................  420
42B      ....................................................................................................  420
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

```
Majority QALAYAYGDLLSNVAAAHKIFDYLDREPAVGTGGTSEPATLQGHVTFQHVSFAYPTRPEHLVLQDVSFELRPGEVTALAGLNGSGKSTCAGLLQRFYEPTAGEVL

9A       ..........................................................K.....................................................     525
9B       ............................V.....................................................................................     525
10A      ..........................................................................................................     525
10B      ..........................................................................................................     525
12A      ........................................R.................................................................     525
12B      .E.......................C.........L.....................L.....................................................     525
16A      ..........................................................................................................     525
16B      ..................................L.......................................................................     525
17A      ..........................................................................................................     525
17B      .......S...................V.....................................................................................     525
18A      ..........................................................................................................     525
18B      ..................................L.......................................................................     525
19A      ..........................................................................................................     525
20A      ..........................................................C.K.V...........................................     525
20B      ..................................L...I...................................................................     525
24A      ..........................................................................................................     525
24B      ..................................L.......................................................................     525
25A      ..........................................................................................................     525
25B      ..........................................................................................................     525
32A      ..........................................................................................................     525
32B      ........................H..V.....................................................................................     525
42A      ..........................................................................................................     525
42B      ..........................................................................................................     525
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

```
Majority LDGVPLRDYKHRYLHRQVALVGQEPVLFSGSIRDNIAYGLEDCREEEIIAAAEAAGALDFISALDQGFDTDVGERGDQLSAGQKQRVAIARALVRCPTVLILDEA

9A       ..........................................................................................................  630
9B       ..........................................................................................A.......  630
10A      ..........................................................................................................  630
10B      ...............................W..........................................................................  630
12A      ..............................................................N...........................................  630
12B      ...........L..........................................................................................  630
16A      ...........................................V..............................................................  630
16B      ..........................................................................................................  630
17A      ..................................................S.......................................................  630
17B      ..........................................................................................................  630
18A      ..........................................................................................................  630
18B      ..........................................................................................................  630
19A      ...............................W..........................................................................  630
20A      ..........................................................................................................  630
20B      ..........................................................................................................  630
24A      ..............H...........................................................................................  630
24B      ..........................................................................................................  630
25A      .........R.....................................................Y.....................A.......  630
25B      .........R.....................................................Y.....................A.......  630
32A      ..........................................................................................................  630
32B      ..........................................................................................................  630
42A      ..............C.........................T.....................................................................  630
42B      ...........................................V..............................................................  630
```

Figure A-3: TAP2 amino acid sequence alignment (continued).

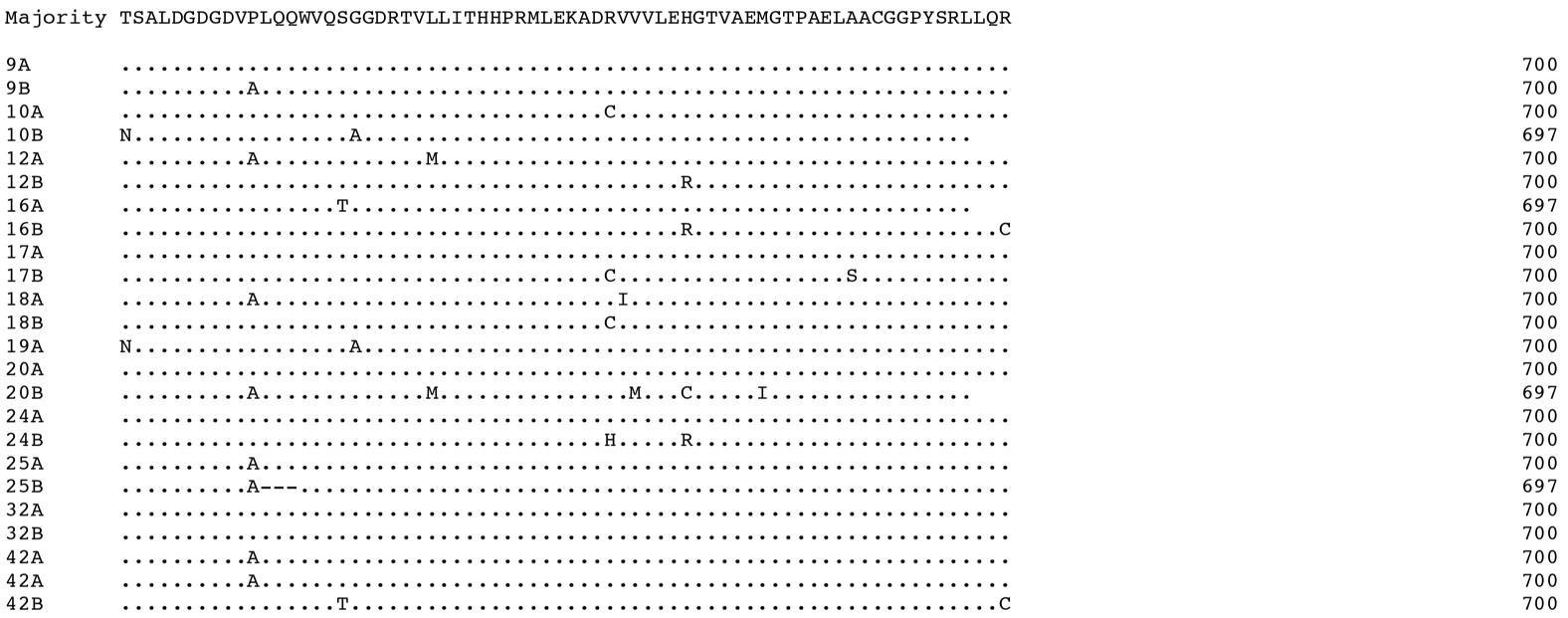

Figure A-3: TAP2 amino acid sequence alignment. The amino acid sequences of TAP2 alleles were aligned with CLUSTALW in MegAlign (DNAStar). Dots represent nucleotides that match the majority. Dashes represent gaps in the sequences.

```
Majority MGAGRRLLESLRPERWRCVVVAALMGASSLGEMAIPYYTGRASDWVAREDELAAIWPMVLLGLSSAVTELACDIAFVGTLSRVHGHLQRRLFAAVLRGDIAELRDEGAGDV

9A       ------------------------------------------------------------.....................................................      47
9B       ------------------------------------------------------------.....................S...............................      47
10A      ------------------------.....................H.................I.....................................................      87
10B      ------------------------.....................................................................................      87
12A      ------------------------.....................................................................I.................      87
12B      ------------------------.....................................................................................      87
16A      ------------------------.....................................................................................      87
16B      ------------------------.....................................................................................      87
17A      ------------------------.....................................................................................      87
17B      ------------------------.....................................I.....................................................      87
18A      ..................A.....................................................................................      111
18B      .....................................................................................................      111
19A      ..................A.....................................................................................      111
19B      ...........................................................................M.........................      111
20A      ------------------------.....................................................................................      87
20B      .....................................................I.....................................................      111
24A      ..................A.....................................................I.....................................      111
24B      ..................A.....................................................................................      111
25A      ------------------------.....................H.................I.....................................................      87
25B      ------------------------.....................................................................................      87
32A      ..................A.....................................................................................T.....      111
32B      .....................................................................................................      111
42A      ..................A.....................................................................................      111
42B      ..................V.....................................................................G...............      111
```

Figure A-4: TAP1 amino acid sequence alignment (continued).

```
Majority AARVTGDVEATREALGEALSLLLWYLARGICLFSTMAWLSPRMALVTALVLPFLLLLPRAIGRLQQGLAPKVQKALAHASEVAVETFQAMATVRSFANEDGVAARYRQCLH

9A       ..........................................................................................................   158
9B       ..........................................................................................................   158
10A      ..........................................................................................................   198
10B      ..........................................................................................................   198
12A      ..........................................................................................................   198
12B      ..........................................................................................................   198
16A      ....M..M.....D..K.....................H.....................................................................   198
16B      ..........................................................................................................   198
17A      ..........................................................................................................   198
17B      ..........................................................................................................   198
18A      ..........................................................................................................   222
18B      ..........................................................................................................   222
19A      ..........................................................................................................   222
19B      ..........................................................................................................   222
20A      ..........................................................................................................   198
20B      ..........................................................................................................   222
24A      ..........................................................................................................   222
24B      ..........................................................................................................   222
25A      ..........................................................................................................   198
25B      ..........................................................................................................   198
32A      ..........................................................................................................   222
32B      ..........................................................................................................   222
42A      ..........................................................................................................   222
42B      ..........................................................................................................   222
```

Figure A-4: TAP1 amino acid sequence alignment (continued).

```
Majority QTHQLEKQDAVLYAVSAWTSGFSALALKMGILYYGGQLVAAGTISTGDLVTFLLYQLQFTDVVEVLLRYYPTLTKAVGSSEKIFEFLDREPQVAPSGTLAPSDLRGHLQLE

9A       ..........................................................................................................     269
9B       ..........................................................................................................     269
10A      ..........................................................................................................     309
10B      ..........................................................................................................     309
12A      K.........................................................................................................     309
12B      ........N.................................................................................................     309
16A      ..........................................................................................................     309
16B      ........N.................................................................................................     309
17A      K.........................................................................................................     309
17B      K.........................................................................................................     309
18A      K.........................................................................................................     333
18B      ........N.................................................................................................     333
19A      ........N.................................................................................................     333
19B      ........N.................................................................................................     333
20A      K.........................................................................................................     309
20B      ..........................................................................................................     333
24A      K.........................................................................................................     333
24B      K.........................................................................................................     333
25A      K.........................................................................................................     309
25B      K.........................................................................................................     309
32A      ..........................................................................................................     333
32B      ..........................................................................................................     333
42A      K.........................................................................................................     333
42B      ..........................................................................................................     333
```

Figure A-4: TAP1 amino acid sequence alignment (continued).

```
Majority DVWFSYPGRQEPVLKGVSLELRPGEVLALMGPPGAGKSTLVSLLLRLHQPTAGRLLLDGHPLSAYRHSYLCHRVATVPQEPVLFARSLHANIAYGPESWSRAQVTAAARRV

9A       ..................................................                                                  327
9B       ..................................................                                                  327
10A      ..................................................                                                  367
10B      ..................................................                                                  367
12A      ..............................V...................                                                  367
12B      ..................................................                                                  367
16A      ..................................................                                                  367
16B      ..................C...............................                                                  367
17A      ..................................................                                                  367
17B      ..................................................                                                  367
18A      ..............................V.....................S.............................................    444
18B      ..................................................................................................C..    444
19A      ..................................................................................................    444
19B      ..............................................................Q...................................H..    444
20A      ..................................................................................................C..    420
20B      ..................................................................................................H..    444
24A      ..............................V...................................................................C..    444
24B      ..............................V...................................................................    444
25A      ..................................................                                                  367
25B      ..................................................                                                  367
32A      ..................................................................................................    444
32B      ..................................................................................................    444
42A      ..................................................................................................    444
42B      ..................................................................................................    444
```

Figure A-4: TAP1 amino acid sequence alignment (continued).

```
Majority GAHNFITRLPRGYDTEVGELGGQLSGGQRQGVAIARALVRDPHILVLDEPTSALDAECQQQVERELFEDSGTRRTVLLVTGQVALAARAHQVALMEGGQVREQGPPSKLLR

18A      ...............................................................................................................      555
18B      ...............................................................................................................      555
19A      ...............................................................................................................      555
19B      ..........Q...................................................................................................C      555
20A      ...............................................................................................................      531
20B      ..........Q....................................................................................................      555
24A      ...............................................................................................................      555
24B      ...............................................................................................................      555
32A      ...............................................................................................................      555
32B      ...............................................................................................................      555
42A      ...............................................................................................................      555
42B      ...............................................................................................................      555


Majority PGSRYWHLLQDGGRQGTLGDGDKGTGSGDEGQQEPGMGTVQARMKTSGDGD-------TRSWDGDTGSQDGDTESWDGDTRARDGDTKGRGIGGDTTTPWDGDTRS

18A      ...................................................------.................................................      654
18B      .............H..........................                                                                      595
19A      ...................................................------...                                                   609
19B      ............................................E......------...                                                   609
20A      .....................................R......E.....ELGSTEQ........................................I........      636
20B      ...................................................------...................................D.............      654
24A      ............................................E......------.................................................      654
24B      .....................................RM.....E.....ELGSTEQ....G...................................I.....S..      660
32A      ............................................E......------.--------------..............A...................      640
32B      ............................................E......------...                                                   609
42A      .........................................G..E......------.................................................      654
42B      ...................................................------.................................................      654
```

Figure A-4: TAP1 amino acid sequence alignment. The amino acid sequences of TAP1 alleles were aligned with CLUSTALW in MegAlign (DNAStar). Dots represent nucleotides that match the majority. Dashes represent gaps in the sequences.

```
Majority CGTCCCCGGTCAGTGCCCGCGACCCTCCCGGCCCCTCTCCGGTCCCGGAGCCGGTCCCGGTCCCGTTTCTCGGAAAATGAAACCCAGCGCCGGGACGGGATGGGCGCCGG

18A      ..................................................................................................     110
18B      ................................G.................................................................     110
19A      ....T.............................................................................................     110
19B      ................................G.................................................................     110
20B      ..............................................................A...................................     110
24A      ............................G...G.................................................................     110
24B      ................................G.................................................................     110
32A      ....T.....................................................................................A.....     110
32B      ........................................................................C.........................     110
42A      ..................................................................................................     110
42B      ....T.............................................................................................     110


Majority CGCCGCCTGCCATTGGCTGATGCCTGCGGCGCGTCACCAGTTGTCGGGCAGATCGGGCCGGAGCAGCTTGCGGAGGGGCGGGGAGGGGGGC-GTGGCAAGGGGCGGAGGG

18A      ..................................................................................-.....G.....     219
18B      ..................G...G.T.....................................................-.............T....     219
19A      ..................G...G.T.............................A.....................C...........A......     220
19B      ..................G...G.T.....................................................-.............T....     219
20B      ................................T.....................................................-.....G.....     219
24A      ..................G.T.G.....................................T.............C.......-.G.............     219
24B      ..................G...G.T.....................................................-.............T....     219
32A      ..........................................................A.A.....................-...........A......     219
32B      ..................................................................................-.....G.....     219
42A      ......................A.....................................................-...........A......     219
42B      ..........................................................A.............A.....................-.....     219
```

Figure A-5: TAP promoter sequence alignment (continued)

```
Majority GGCGTGGCAAGGGGCGGAGGGGGGCGTGGTGACGGG----------TGAAAGGGGCGTGGCGAGGCGGGAGTGGGGCGGGGCCGGGGGAGAAGGGCTG

18A      ..........................C......----------......A.......T.....................................     307
18B      ...................T............AT..----------.....................................................     307
19A      ..............-.G..............----------.....................................................     307
19B      ...................T............AT..----------.....................................................     307
20B      ..----------.....C.....----------.....................................................     286
24A      ...................T............A...----------.............T.....................................     307
24B      ...................T............AT..----------.....................................................     307
32A      .....T.....................C......----------.....................................A     307
32B      ..............T.............C......----------......A.......T.....................................     307
42A      ..........................CA.G...CGGAGGGGGGCGTGGAGACGGG......A.....................................     329
42B      ..................................----------.....................................................     307


Majority TGAAAATGAAAGTGAA------AGGTGGCGGGGCCAGAGCATAGAGGGAGCGGCCGGGGCCGGGTGGGTCTGGGGGCCCCGGGGGGAGGCAGAGCAGGCTGGGGGGAGCC

18A      ...............AGTGAA.....................................A.................T.....................     417
18B      ......C.........------.....................................T.....................A.....................     411
19A      ...............AGTGAA.....................................................................................     417
19B      ......C.........------.....................................T.....................A.....................     411
20B      ...............------.....................................................T.....................     390
24A      ...............------.....................................................................................     411
24B      ......C.........------.....................................T.....................A.....................     411
32A      ...............------.....................................................................................     411
32B      ...............AGTGAA.....................................A.................T.....................     417
42A      ...............------.............................................T.....................T.............     433
42B      ...............------.....................................................................................A.....     411
```

Figure A-5: TAP promoter sequence alignment (continued).

```
Majority AGGGGGGGGCGGXTGGACTCGGCCGGGCCAAGC-----CGAGCCGAGCAGGTCTGGGGAGCACGGGGCTGGTGAGGGGGGG-CTGGACTGGGAAGACTGGGAGAGCTGGT

18A      -...........C..T........A........CGAGC...........................................-............................    525
18B      ............T....................TGAGC.....................A.....................-............................    520
19A      -T..........T....................CGAGCT.-----....................................-............A.........T.....    520
19B      ........-...T....................-----...........................................-............................    514
20B      ............C..T.................CAAGC..-----....................................-............................    494
24A      -...........C......T.....A.......-----...........................................-...........G................    514
24B      ........-...T....................-----...........................................-............................    514
32A      -.........A.C............A.......-----...........................................G...........GT...............    515
32B      -...........C..T.................CGAGC...........................................-............................    525
42A      ........-...T.............A......-----T..........................................-............................    536
42B      ...A....-G..G....................-----T....T..............................C......-............................    514


Majority GTGTCCTCAGTGTGCCAGGGGTGGGTCTCAAAAAGGTGAAGGGGGCTGCGCTGGGGGTTACAGGTGTGTTCCTGAGAGGTTTGGGGATCCAGGACGAGACTAGGGGGCCA

18A      ........................................................T.....................................................    635
18B      ..............................................................................................G...............    630
19A      ........................................................T.....................................................    630
19B      ...............................................................................C..............................    624
20B      ..............................................................................................................    604
24A      ................G..............G.................A...........................A................................    624
24B      ..............................................................................................................    624
32A      ..........C.....G..............G.............................................A................................    625
32B      ..............................................................................................................    635
42A      ..............................................................................................G...............    646
42B      .........T....................................................................................................    624
```

Figure A-5: TAP promoter sequence alignment (continued).

```
Majority TACTGGGTGTACTGGGTTGTGTTCTAGGGTGCTGGGAAGGGTCTTTGGGGGTCCTGGTCAGGACTGGGGGATTTACTGGGAGCACTGGTGTCCCAGGGAGCTCAGCAGGG

18A      .................A.............................................................T.................................     745
18B      .......CA.....................................................T.....................................................     740
19A      .......CA.....................................................................................................A.     740
19B      ..............................................................................................................     734
20B      ..............................................................................................................     714
24A      .......---------.........A.....................................................................................     725
24B      .......CA.....................................................................................................     734
32A      ..............................................................................C...............................     735
32B      .................A.............................................................................................     745
42A      .......CA.....................................................T.....................................................     756
42B      ..............................................................................................................     734


Majority TCTCAGAGA--TAAGCCAGGAGCACTGGTGTGAGCCCCAGGGTGCTAGACAGGGTCAAGGGGAGCTGAGGAGTCTGCACTGGGAGCACTGGTGTTCCCTGGGGAGTGCCT

18A      ......G..--.............................................G.............................................................     853
18B      .........--..............................................................................................     848
19A      ......G..--.............................................G.............................................................     848
19B      .........--..............................................................................................     842
20B      .........--.............................................G.....................................A.....................     822
24A      .........--..............................................................................................A.....     833
24B      .........--..............................................................................................     842
32A      ......G..GC.............................................G.............................................................     845
32B      .........--............T.................................G.............................................................     853
42A      .........--..............................................................................................     864
42B      .........--..............................................................................................     842
```

Figure A-5: TAP promoter sequence alignment (continued).

```
Majority CAGGGAGCCCAGCAGGGTCCACGTGACCAGCAGAGGGTCTGGGATTGCTGGGTAGGGCCTTGAATGCTCCTCAGAGCCCCCAAGCC

18A      ..................................................................................  939
18B      ..........................................G.......................................  934
19A      ......................A...................A.G...................C.....................  934
19B      ..................................................................................  928
20B      ......................A...................A.G...................C.....................  908
24A      ..................................................................................  919
24B      ..........................................G.......................................  928
32A      ................................A.............................C.....................  931
32B      ..A.....................................................CT.....................  939
42A      ..........................................G.......................................  950
42B      ..................................................................................  928
```

Figure A-5: TAP promoter sequence alignment. The nucleotide sequences of TAP promoter alleles were aligned with CLUSTALW in MegAlign (DNAStar). Dots represent nucleotides that match the majority. Dashes represent gaps in the