Understanding antigen processing in chickens using genome editing technology

Daniel Wise

Darwin College

Department of Pathology

This dissertation is submitted for the degree of Doctor of Philosophy

Preface:

I hereby declare that my dissertation entitled:

Understanding antigen processing in chickens using genome editing technology

- Is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the acknowledgements and specified in the text.

- Is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Preface and specified in the text.

- Does not exceed the prescribed word limit

.....

.....

.....

Summary

The major histocompatibility complex (MHC) plays crucial roles in many biological processes, most especially disease resistance. Classical MHC molecules are extremely polymorphic and encode cell surface glycoproteins that present antigenic peptides to T lymphocytes of the immune system. Several other molecules are involved in loading and optimising the peptides for the two major groups of classical MHC molecules, including the transporter associated with antigen processing (TAP) for MHCI molecules, and the dedicated chaperone DM for the MHCII molecules. In mammals, these critical antigen loading molecules have limited polymorphism, providing peptides for a multigene family of well-expressed classical MHC molecules.

Outside of mammals, the chicken is the only system in which peptide loading and presentation has been studied in molecular detail. Compared to mammals, the chicken MHC is small, simple and arranged differently. Chickens have two classical MHCI genes but strongly express only one at the RNA and protein levels which presents the majority of peptides. This phenomenon is apparently due to co-evolution with the polymorphic TAP genes, among other peptide loading associated genes. The exact polymorphic TAP residues involved in peptide loading of chicken MHCI remain unclear. For the class II system, there are two MHCII B genes and two DMB genes, but only BLB2 and DMB2 are expressed at high levels in haemopoietic cells. One hypothesis is that each MHCII B gene co-evolves with one of the two DMB genes, resulting in differential expression, but the interactions and importance of each DM gene are unknown.

In this thesis, the genome editing approach of clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 was used to begin to understand the peptide loading systems in the chicken. For the first time, the CRISPR-Cas9 system was optimised and implemented in avian cell lines, beginning with the creation of MHCI and TAP gene knock-out (KO) chicken cell lines. TAP-dependent peptide transport was assessed in each line allowing for the first steps to determine the critical peptide binding residues in chicken TAP transport.

To understand the role of DM in chickens, KO cell lines for BLB and DM genes were created and analysed at the RNA and protein levels, additionally the peptide repertoire was determined by immunopeptidomics. Striking and unexpected changes in MHCII expression were seen by deleting different components of MHCII presentation.

Initial findings suggest the chicken MHCII system utilises two DM chaperones for effective MHCII expression, stability and peptide presentation. The work in this thesis provides some of the first detailed insights into the peptide loading systems for MHC molecules in an organism outside of mammals and provides the basis for genome editing in primordial germ cells (PGCs), eventually allowing analysis of MHC presentation in vivo for improved understanding of pathogen resistance and vaccine responses.

Table of contents

List of figures

List of tables

List of appendices

List of abbreviations

AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
ABC	ATP-binding cassette
APS	ammonium persulphate
β2m	β2-microglobulin
BSA	bovine serum albumin
bp	base pairs
CDS	coding sequence
CLIP	class II-associated invariant chain peptide
CIITA	class II transactivator
cM	centimorgans
CREB	cAMP response element-binding protein
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
Ct	threshold cycle
Da	Dalton
EBI	European Bioinformatics Institute
ER	endoplasmic reticulum
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
HDR	homology-directed repair
HIV	human immunodeficiency virus

HLA	human leukocyte antigen
Ig	immunoglobulin
Ii	invariant chain
IP	immunoprecipitation
KD	knock down
KI	knock in
KO	knock out
LSB	Laemmli sample buffer
mAb	monoclonal antibody
MDV	Marek's disease virus
MFI	mean fluorescent intensity
MIIC	MHC class II-containing compartment
MHC	major histocompatibility complex
miRNA	microRNA
NCBI	National Centre for Biotechnology Information
NFY	nuclear factor Y
NF-κB	nuclear factor kappa B
NHEJ	non-homologous end joining
NP40	Nonidet-P 40
NBD	nucleotide binding domain
PAGE	polyacrylamide gel electrophoresis
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PBS	Phosphate-buffered saline
REV	reticuloendotheliosis virus

RGENs	RNA-guided engineered nucleases
RPMI	Roswell Park Memorial Institute
RT-PCR	reverse-transcription polymerase chain reaction
RT-qPCR	reverse-transcription quantitative polymerase chain reaction
SDS	sodium dodecyl sulphate
sgRNA	single guide RNA
siRNA	short interfering RNA
TALEN	transcription activator-like effector nucleases
TAP	transporter associated with antigen processing
TEMED	tetramethylethylenediamine
TMD	transmembrane domains
tracrRNA	trans-activating crRNA
UTR	untranslated region
ZFN	zinc-finger nucleases

1. General Introduction

1. General introduction

1.1 Major histocompatibility complex molecules

In mammals the major histocompatibility complex (MHC) molecules have been shown to play a crucial role in histocompatibility (graft rejection), immune responses and disease resistance and susceptibility. Most of our knowledge of the MHC comes from studies in humans and mice (Ploegh 1981, Kaufman et al 1984, Germain 1994, Jones 1997, Vyas et al 2008, Neefjes 2011). The MHC region is known to have the most associations to disease resistance and susceptibility. In humans, the strong associations are generally with autoimmune diseases and are weak or non-existent with infectious pathogens, with a few exceptions such as HIV progression (Carrington et al 1999, 2003).

The MHC molecules are comprised of two classes, class I (MHCI) and class II (MHCII); both classes are further broken down into classical and non-classical MHC molecules. Classical MHCI and II are incredibly polymorphic and encode cell surface glycoproteins that enable the immune system to recognize foreign antigens. Through the antigen processing pathway, MHCI molecules present mainly cytosolic antigens to CD8 positive T-cells, whereas MHCII molecules present mainly endosomal peptides to CD4 positive T-cells (Neefjes et al. 2011). The small protein fragments (peptides) are presented by MHC molecules and the presentation of these peptides to T lymphocytes of the immune system initiates effector and regulatory cell activation. MHCI and MHCII molecules have differences in their structure and expression that enables them to achieve different functions of antigen presentation. The overall structure of MHCI and MHCII molecules is similar, with both classes having two membrane-distal domains that form a peptide binding groove and two immunoglobulin-like membrane-proximal domains (Kaufman et al 1984, Bjorkman et al 1987, Brown et al 1993, Jardetzky et al 1994, Stern et al 1994). MHCI molecules are present on all nucleated cells and

consist of a large α chain (heavy chain), which contains the peptide binding groove, and a small non-covalently-associated chain named β_2-microglobulin (β2m); MHCI molecules present peptides to cytotoxic T cells (CTLs) expressing the co-receptor CD8. MHCII molecules are comprised of two non-covalently associated α and β chains, with each chain contributing a membrane-proximal immunoglobulin-like domain and a membrane-distal domain. The membrane distal-domain of the α and β chains comprise the peptide groove, and both chains have a transmembrane domain anchoring the MHC to the cell surface. MHC II expression is generally restricted to antigen presenting cells (APCs) such as macrophages, and present peptides to T cells expressing the CD4 co-receptor.

1.1.1 MHC class I structure

The MHCI heavy chain is separated into three domains, with the immunoglobulin-like α3 domain forming the only anchorage point through the transmembrane domain. The α1 and α2 domains together form the platform for peptide binding, which is comprised of an eight stranded β-sheet with two long α-helices (one from α1 and the other from α2) overlaying the β-sheet and forming the edges of the peptide binding groove (Bjorkman et al 1987) (fig 1.1.1). The key binding pockets in MHCI are defined as interacting with the side chains from peptide position P2 (or sometimes P5/6) and P8/9, with strong binding generally occurring at either end of the binding groove causing the bound peptide to bulge in the middle, resulting in peptide side chains pointing out of the groove (Falk et al 1991, Hunt et al 1992). Unlike MHCII, MHCI has conserved residues at either end of the binding groove, leading to a closed binding groove which restricts the length of peptides bound; these peptides are generally between 8-10 amino acids in length (Matsumura et al 1992, Bouvier and Wiley 1994). Polymorphisms of MHCI are mostly localised to that of the peptide binding groove and it is

generally accepted that the residues associated with peptide binding are under selection primarily for pathogen recognition, in what has been termed a 'molecular arms race' (Hughes and Nei 1988).

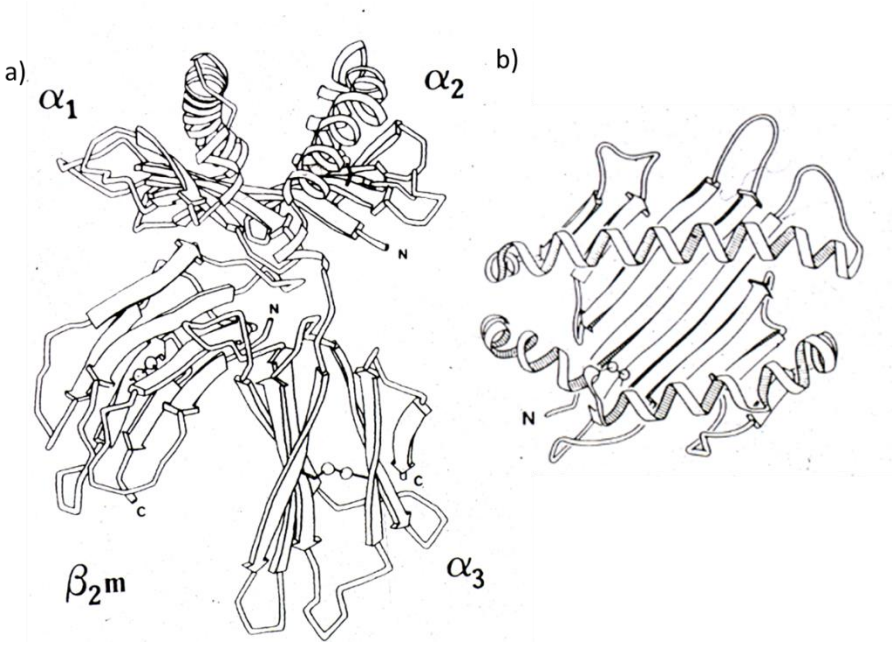


Figure1.1.1 Cartoon representation of human HLA-A crystal structure.

a) MHCI molecules consist of a large α chain (heavy chain), which contains the peptide binding groove, and a small non-covalently-associated chain named β_2-microglobulin (β2m). b) The α1 and α2 domains together form the platform for peptide binding, which is comprised of an eight stranded β-sheet with two long α-helices (one from α1 and the other from α2) overlaying the β-sheet and forming the edges of the peptide binding groove (Bjorkman et al 1987).

1.1.2 MHC class II structure

MHCII molecule expression is generally restricted to APCs including B cells, monocytes/macrophages, dendritic cells and thymic epithelial cells (TECs). MHCII expression can be induced in other cell types by stimulation, including by interferon gamma (IFN-y), in endothelial cells and fibroblasts (Geppert and Lipsky 1985, Umetsu et al 1986), keratinocytes (Czernielewski and Bagot 1986, Baker et al 1988), mesenchymal stromal cells (Chan et al 2006, Romieu-Mourez et al 2007) and intestinal epithelial cells (IECs) (Hershberg et al 1997, Thelemann et al 2014). As previously mentioned, MHCII consists of two non-covalently associated chains (α and β), both of which have a cytoplasmic, transmembrane and immunoglobulin-like domain, and together comprise the peptide binding groove (Springer et al 1977, Kaufman et al 1984) (figure 1.1.2). The α chain has one disulphide bond in the α2 domain whilst the β chain has two disulphide bonds, one in the β1 domain and one in the β2 domain (Kaufman et al 1984). The peptide binding groove consists eight anti-parallel β strands covered by two α helices (Brown et al 1988, 1993). Peptides are bound in the peptide binding groove by a hydrogen bonds between residues in the peptide backbone and the binding groove in a sequence independent manner. Conversely, sequence specific interactions are formed between the side chains of peptide residues and residues in the binding groove that form 'pockets' (Stern et al 1994). Unlike MHCI, MHCII has an open-ended binding groove and can bind peptides of varying lengths, with optimal peptides being of 12-20 amino acids in length (Jardetsky et al 1996, Godkin et al 2001). The most significant pockets for binding are P1, P4, P6 and P9 and bind 'anchor residues' from the peptide; P3, P7 potentially have smaller binding pockets (Suri et al 2006). Stability of MHCII is dependent on the peptide that is bound, with different peptide sequences and lengths effecting stability and flexibility of MHCII (Nelson et al 1996, O'Brien et al 2008, Ferrante 2013). As with MHCI,

polymorphisms of MHCII are mostly localised to the peptide binding groove and are under strong selection for diversity, presumably for response to pathogens (Hughes and Nei 1989).

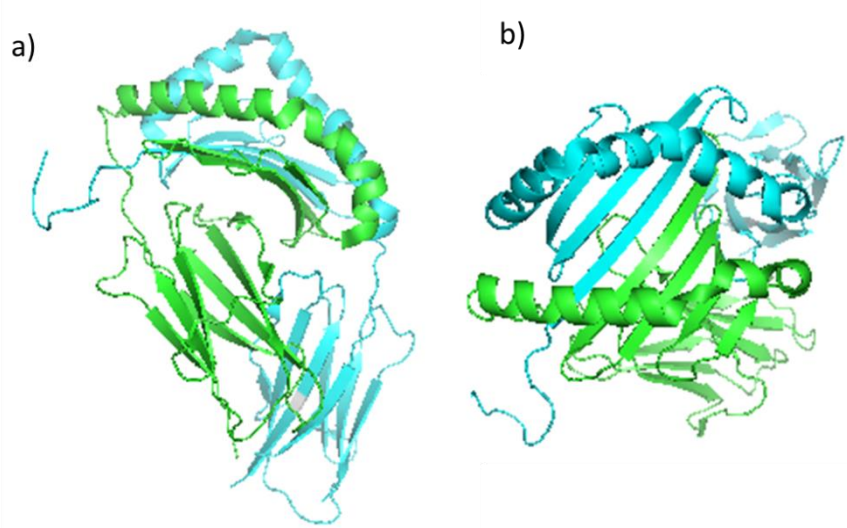


Figure1.1.2 Cartoon representation of DR1 crystal structure.

a) MHCII molecules are comprised of two non-covalently associated α and β chains. The α chain consists of two domains, α1 and α2 domain whilst the β chain consists of β1 and β2 domains. b) the peptide binding groove consists eight anti-parallel β strands covered by two α helices formed by the associated α1 and β1 domains (adapted from Pos 2012)

1.2 MHC class I antigen processing and presentation

Antigen processing and presentation have been primarily and extensively studied in human and murine systems, giving a detailed understanding of MHC antigen processing and presentation. MHCI molecules classically present peptides that are cytosolic in origin that have been generated by the proteasome and are pumped into the lumen of the endoplasmic reticulum (ER) by the transporter-associated with antigen processing (TAP) heterodimer (fig 1.2.1). The TAP1 and TAP2 genes encode the TAP transporter which translocate peptides from the cytosol into the ER, ready for binding to class I molecules (Deverson et al 1990, Trowsdale et al 1990, Spies et al 1990). The TAP heterodimer is a four-domain structure with two hydrophobic transmembrane domains (TMDs) and two hydrophilic nucleotide-binding domains (NBDs). The TMDs provide a passage of peptides to be transported and the

substrate-binding site. The two NBDs are located in the cytoplasm and are involved in ATP binding and hydrolysis. In humans, the TAP genes are located in the class II region, separated from the highly polymorphic class I genes, and are monomorphic, having evolved to transport peptides which will bind to all class I alleles. In the ER, peptides are trimmed further by endoplasmic reticulum aminopeptidase 1 (ERAP1) 1 and ERAP2 after being bound by MHCI.

The MHCI heavy chain folding is assisted by the ER resident chaperones calnexin, calreticulin and the protein disulphide isomerase (PDI), ERp57, which assists disulphide bond formation in the MHCI heavy chain and then associates with the β2m subunit before binding peptide (Sadasivan et al 1996, Leach et al 2002, Zhang et al 2006). MHCI forms a large complex with the TAP transporter bridged by the chaperone tapasin, calreticulin and ERp57 in what is known as the peptide loading complex (PLC) (Blees et al 2017). The order in which the PLC is formed is still not fully understood, though it is thought it is thought to start with ERp57 then forming a disulphide bond with tapasin.

Tapasin interacts with a loop below the α2 helix and a β-strand in the α3 domain of the MHCI; then tapasin binds the TAP transporter through positive residues located in the transmembrane domain, confining the MHCI to the source of peptides as well as acting to stabilise the empty MHCI (Sadasivan et al 1996, Ortmann et al 1997, Petersen et al 2005). ERp57 associates with calreticulin and has been shown to be essential for effective PLC activity (Ortmann et al 1997 and Petersen et al 2005). Stable bound peptide-MHCI complexes then dissociate from the PLC and are subject to a second round of peptide editing by a second peptide editor, TAPBPR. TAPBPR binds the MHCI at similar contact sites to tapasin and catalyses the exchange of low affinity peptide for those with a higher affinity

(Boyle et al 2013). Stable peptide-MHCI complexes loaded with peptides of high affinity are then transported to the cell surface for presentation to $CD8^+$ T cells.

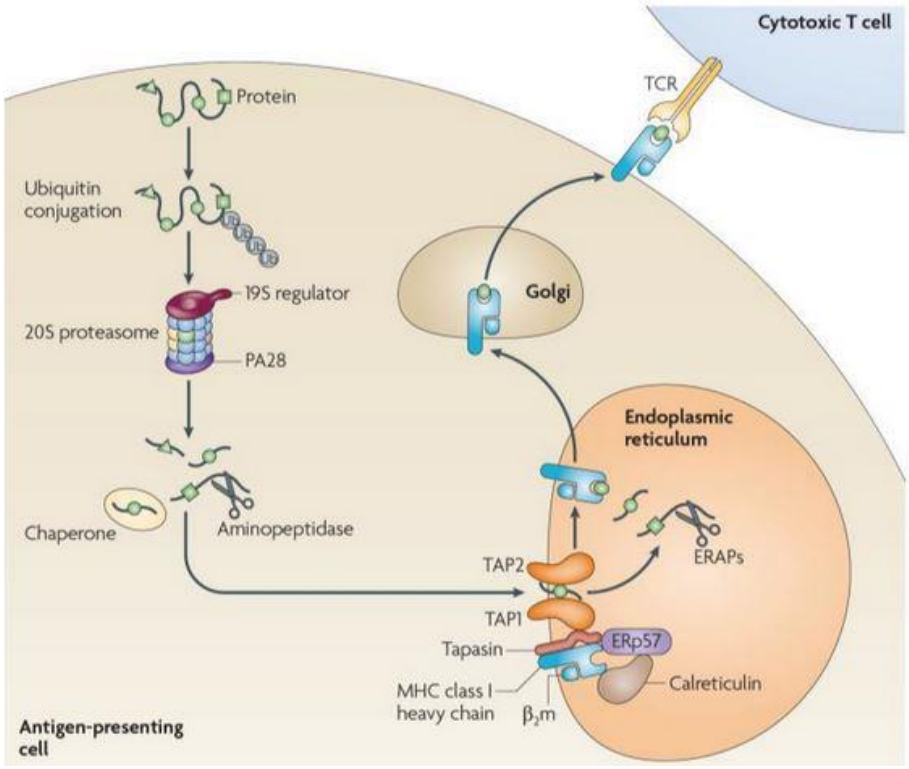


Figure1.2.1 Schematic of MHCI antigen processing and presentation

Intracellular antigens are processed into peptides by the immunoproteasome, which is comprised of multiple subunits, such as LMP2. Peptides are then transported into the lumen of endoplasmic reticulum (ER) via transporter associated with antigen presentation (TAP) proteins, where they are loaded into the binding groove of the MHC class I complex. MHC class I-peptide complexes migrate to the cell surface for presentation to $CD8^+$ T cells (Groettrup et al 2010).

1.3 MHC class II antigen processing and presentation

Initially MHCII α and β chains fold in the ER in association with a third chain, known as the invariant chain (Ii) or CD74 (Cresswell 1996), which is a non-polymorphic transmembrane domain protein that is expressed in conjunction with MHCII (Long 1985). Ii is well conserved from sharks to humans (Cristiello et al 2012), has a critical role in MHCII trafficking and is

required for MHCII transport through the endocytic pathway, with Ii absence resulting in MHCII localisation in the ER (Lamb et al 1991, Viville et al 1993). In addition Ii is vital for immune development, as mice lacking Ii have impaired thymic CD4+ T cell maturation and MHCII presentation (Viville et al 1993, Elliot et al 1994). Biochemical and structural studies have suggested that the Ii forms a trimer and associates with three MHCII heterodimers to form a nonameric complex or that Ii trimers associate with MHCII heterodimer to form a pentameric structure (Koch et al 2011, Marks et al 1990, Roche et al 1991, Park et al 1995). MHCII and Ii associate through the transmembrane and ectodomains (King and Dixon 2010); residues 90-104 of Ii bind to MHCII occupying the peptide binding groove. This binding stops MHCII binding peptides in the ER and early endosomes, as well as stabilising the class II molecule (Roche and Cresswell 1991, Sette et al 1992).

MHCII molecules are then transported through the endocytic pathway to the late endosomal compartment, MHC class II compartment (MIIC) (Peters et al 1991), which is governed by signalling from the Ii cytoplasmic domain (Lamb et al 1991, Benaroch et al 1995). In the MIIC, Ii is cleaved by cathepsin and other proteases (Costantino et al 2008, Hűttl et al 2016) leaving a short peptide (Ii 91-104) named the class II-associated invariant chain peptide (CLIP) occupying the MHCII binding groove (Riberdy et al 1992) (fig 1.3.1).

CLIP is then exchanged for higher affinity peptides, that are derived from proteolytically degraded proteins present in late endocytic compartments. This peptide exchange is catalysed by the non-classical class II molecule HLA-DM in humans (H-2Mb in mice) and is termed a 'peptide editor', with different MHCII alleles being more or less dependent on DM activity for peptide loading (Pious et al 1994, Stebbins et al 1995, Patil et al 2001). Once a peptide is stably bound, the MHCII-peptide complex is transported to the cell surface for presentation to $CD4^+$ T cells.

In mammals, there is another non-classical class II molecule (HLA-DO in humans and H2-O in mice), which binds to DM inhibiting its activity in a competitive manner providing further control of the peptide repertoire (Glazier et al 2002, Mellins and Stern 2014). DO acts in direct competition with MHCII by binding DM at the same contact sites as MHCII binds (Guce et al 2013) and is found in the late endosome, with its transport from the ER being dependent on DM association (Mellins and Stern 2014). DO expression is more cell type restricted than that of DM, being expressed in B cells, thymic epithelial cells and some subsets of dendritic cells (Hornell 2006, Xiu 2011). One role proposed for DO is to reduce the presentation of self antigens, striking a balance between response to pathogens and preventing autoimmunity (Mellins and Stern 2014).

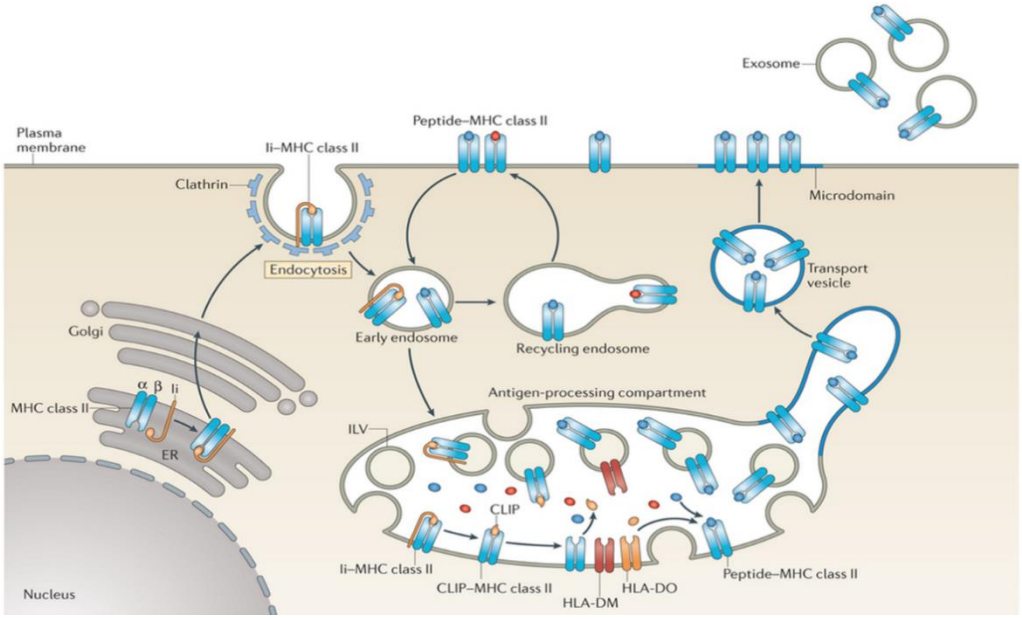


Figure1.3.1 schematic of MHCII antigen processing and presentation

Antigens from endosomal compartments are processed by endolysosomal enzymes into peptides. These peptides are then bound to the MHC class II complex after the class II-associated invariant chain peptide (CLIP) is dissociated. HLA-DM and HLA-DO regulate the antigen-loading process. The MHC class II presents antigens to $CD4^+$ T cells (Roche and Furuata 2015)

1.4 Structure and function of DM

HLA-DM is comprised of an α/β heterodimer and has a similar overall structural fold to MHCII but with noticeable differences (Kelly 1991, Moysak et al 1998, Pos et al 2012). Firstly, unlike MHCII, DM has limited polymorphism. Secondly, structures show that segments of the α-helices in DM are packed more closely together and result in a closed binding groove which prevents peptide binding, with the residues of the binding groove in MHCII not conserved in DM. Additionally, residues that keep the α helices of the α1 and β1 domains of DM close together are conserved in many species (Fremont et al 1998, Mosyak et al 1998, Pos et al 2012), which suggests that the region is not under selection for diversity. Structurally, HLA-DM has three disulphide bonds at conserved positions and two additional disulphides compared to HLA-DR. It has been suggested that these additional disulphides may increase the stability of DM when compared to MHCII (Fremont et al 1998). A crystal structure of HLA-DM in complex with the HLA-DR1 show that HLA-DM mainly contacts the α1 domain of MHCII close to P1 and the membrane-proximal β2 domain, supporting previous mutational analysis (Doebele et al 2000, Painter et al 2011, Anders et al 2011). The mechanism of DM action is now known in molecular detail, starting with the dissociation of the N-terminus of the bound peptide from the P1 pocket, which is either caused by or causes the DRA W43 to flip away from the P1 pocket and become available for interaction with DM. The empty DR is then stabilised by DM whilst the DRα residues F51 and βF89 guard the P1 pocket; peptides then attempt to bind the partially available groove (Pos et al 2012). Peptides that fail to outcompete the DR residues in the P1 and P2 pockets are not bound, but when the N-terminus of a peptide is bound the conformational changes in DR are reversed and DM is dissociated (Pos et al 2012) (fig 1.4.1). The key residues for DR-DM interaction are heavily conserved across species, including the chicken.

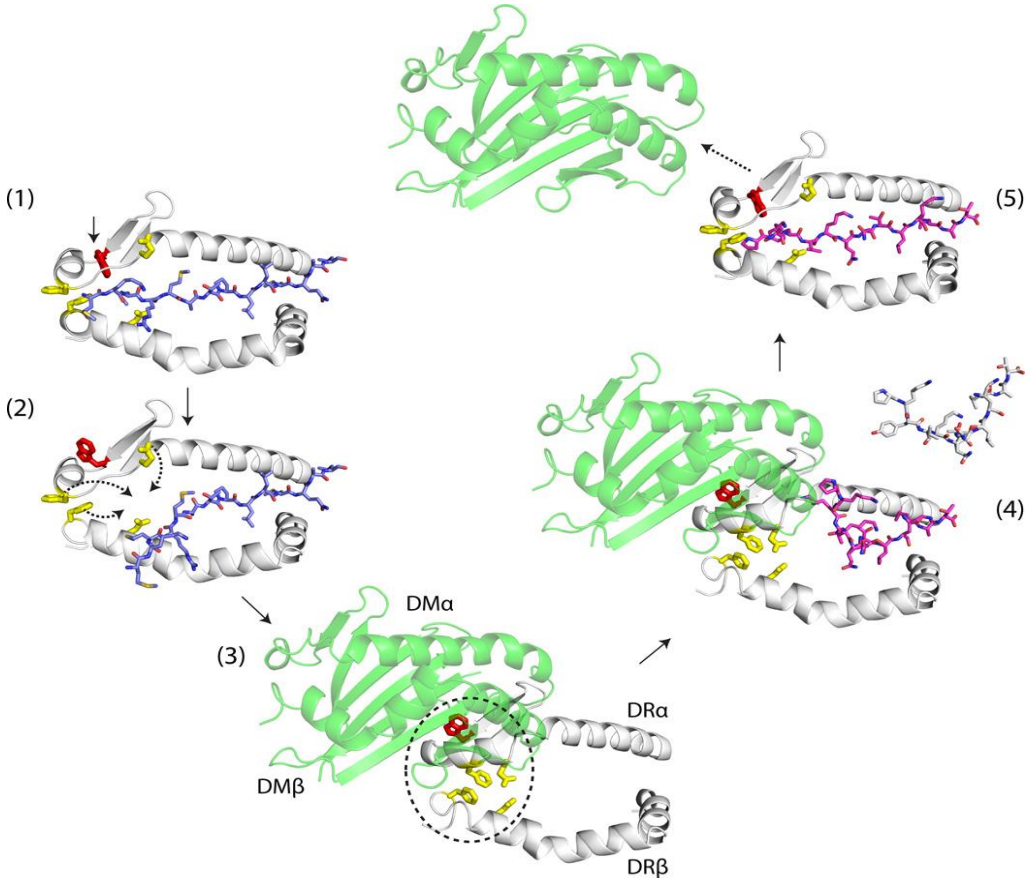


Figure1.4.1 Model proposing the key steps in HLA-DM – HLA-DR action

1) CLIP is bound in DR peptide binding groove whilst DR αW43 (red) stabilises the P1 pocket. 2) The N-terminus of the peptide dissociates from DR binding groove and the DR αW43 rotates away from the P1 pocket and becomes available for interaction with DM. DR residues (arrows) move into the peptide groove during the transition to a DM-bound state. 3) DM stabilizes the empty DR, whilst the DR αF51 and βF89 protect the hydrophobic P1 pocket. 4) Binding of peptides to partially accessible peptide groove of DR result in peptides that do not successfully outcompete DR residues (yellow) for P2 site and P1 pocket are not stably bound or outcompete the DR residues. 5) Peptides that outcompete the DR residues (yellow) result in binding of the peptide N-terminus and reverse the conformational changes of DR and the dissociation of DM (Pos et al 2012).

1.5 Organisation of the mammalian MHC

Classical and non-classical class I and II genes are encoded in the multi-locus region of the genome known as the MHC; this includes the TAPs and tapasin for the class I system and DM and DO for the class II system (The MHC sequencing consortium 1999, Kelley et al 2005). The MHC was originally discovered as the locus responsible for tissue allograft rejection in the mouse and was later described in humans (Gorer 1936, Dausset 1958).

The MHC region of most mammals is large, covering hundreds of kilobases, and contains the multigene families of class I and class II. The MHC also includes many other genes with a variety of functions which are mainly immune associated, as well as numerous pseudogenes and repetitive regions (Kelley et al 2005). The human MHC is 3.8 megabases with over 220 genes. The locus is divided into three regions, such as the class I region, that contains the classical class I genes HLA-A, B and C and some but not all non-classical class I genes (Kelley 2005, Trowsdale et al 2011).

The class II region in humans contains the classical class II genes that encode HLA-DP, DQ and DR, as well as the non-classical class II genes that encode HLA-DM and DO; in each case (except DO), the α and β chains are encoded adjacently (Ting and Trowsdale 2002). For DP and DQ one copy of each gene is well-expressed and functional; DRB is an exception where more than one copy per haplotype can be functional. In mice, one classical class II molecule is encoded by H-2A and another by H-2E, but in many mouse strains the H-2E molecule is lacking due to disruptions in the H-2Eα or H-2Eβ genes. Interestingly, the mouse contains two H2-Mb genes, Mb1 and Mb2, both of which are expressed. Sequence analysis reveals H-2E is most closely related to HLA-DR, whilst H-2A is most similar to HLA-DQ and there is no functional murine homologue of HLA-DP. The organisation of the class II region is similar throughout mammals, although there are many species-specific differences to (van der Poel et al 1990, Yuhki et al 2003). The class II region also contains genes involved in class I antigen

processing such as TAP1 and TAP2, whilst tapasin is located in the extended class II region (Beck and Trowsdale 1999, Kelley et al 2005). The class II region has strong associations with autoimmunity (Beck and Trowsdale 1999, Marsh et al 2000, Kelley et al 2005) (fig 1.5.1). The class III region includes complement component genes, cytokines and many other immune and non-immune associated genes (Kelley et al 2005); interestingly the class III region separates the class I and class II regions.

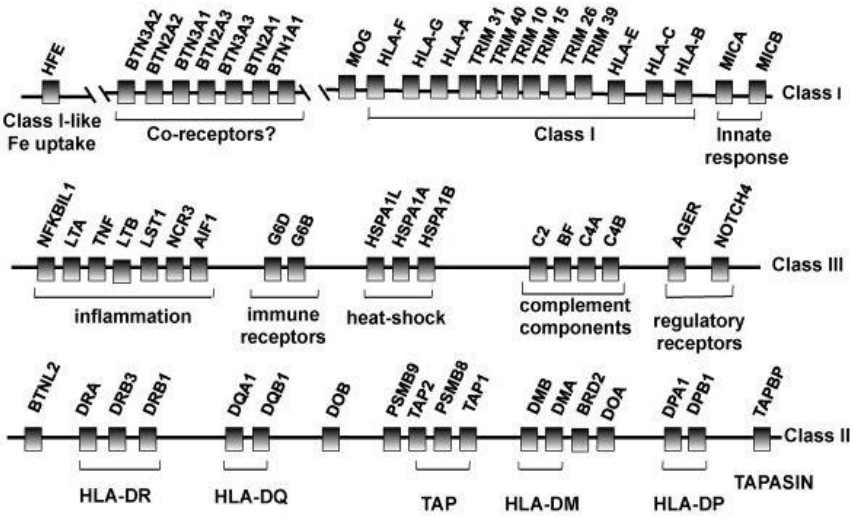


Figure1.5.1 Schematic of the MHC organisation in humans

representation of the class I, class III and class II regions of the human MHC. Classical MHCI genes are shown to be in the class I region and classical and non-classical MHCII genes are seen in the class II region. Genes associated with MHCI peptide loading are found in the class II region (TAP and tapasin). The class I and class II regions are separated by the class III region. Genes are represented as grey boxes (Trowsdale 2011).

1.6 Transcriptional regulation of MHCII expression

MHCII molecules are constitutively expressed on APCs with all three MHCII molecules being expressed at the same time, although some subsets of B cells only express DQ. The promoter elements for MHCII are very conserved, with all classical and non-classical promoters (including Ii) containing the S, X1/X2 and Y boxes, some of which are bound by proteins factors such as cAMP-responsive-element binding protein (CREB), regulatory factor X (RFX)

and nuclear factor Y (NFY), whilst the binding protein of the S box (sometimes called W box) has yet to be defined (Choi et al 2011). Collectively this protein complex recruits the class II transactivator (CIITA) protein, which binds to conserved elements of the proteins that bind the proximal-promoter regions of the class II genes (Choi et al 2011) (fig 1.6.1). Activation of the CIITA is driven by phosphorylation and ubiquitylation, and multiple transcriptional start sites (Sisk et al 2003 Greer et al 2004, Bhat et al 2010). Expression of CIITA in humans is regulated by four distinct promoters (each having a different product), with different isoforms being expressed in different cell types (Smith et al 2011, Choi et al 2011). Studies have also shown that changes in chromatin organisation can affect transcriptional activation or repression in specific cell types, providing a mechanism for differential expression of class II molecules, despite sharing the same transcriptional elements (Choi et al 2011). This additional level of regulation could explain previous findings that show that HLA-DR is generally more highly expressed than HLA-DP and HLA-DQ molecules. The relative levels of different MHCII molecules expressed will change the contribution that each isotype makes to the total peptide repertoire presented by the cell, and is of interest as previous findings have shown that particular HLA-DP, DQ and DR alleles are strongly associated with autoimmune diseases.

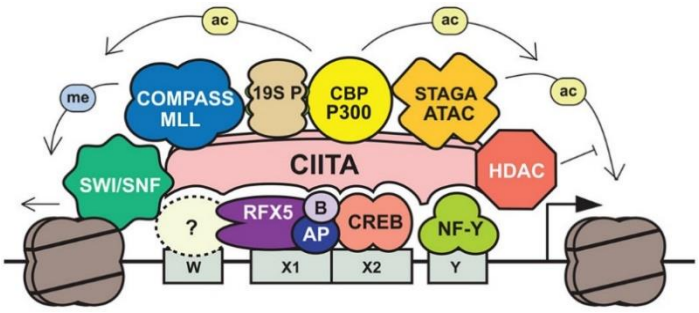


Figure1.6.1 Regulation of MHCII by proximal promoter

The highly conserved W, X and Y box is bound by RFX, CREB, and NFY. This unique structure is recognized by CIITA. The CIITA recruits many additional transcriptional co-activators and their associated complexes that modulate CIITA activation (Choi et al 2011).

1.7 The chicken MHC

The chicken MHC is contained in the B locus, which was first identified as a serological blood group that encodes the highly polymorphic and immunogenic BG genes (Briles et al 1952). The B locus has strong associations with graft rejection, graft-vs-host reactions and mixed lymphocyte reactions (Pazderka et al 1975, Vainio et al 1988), leading to reasoning that the B locus must contain the chicken MHC, which was conclusively shown by full sequencing (Kaufman et al 1999).

The chicken MHC is found on chromosome 16 and separated into two regions by recombination. The BF/BL region contains class I and class II genes as well as the BG1 gene; importantly no recombination has been seen within this region by experimental mating, and only one natural recombinant has been seen (Scandinavian B19 haplotype). The TRIM/BG region is filled with tripartite motif (TRIM) genes and then immediately followed by the BG region, which is shown to undergo huge expansion and contraction between chicken haplotypes. A further locus, the Rfp-Y region was also identified on chromosome 16 and contains a non-classical class I gene (YF), non-classical class IIB (YLB) genes and lectin like genes, though it is genetically unlinked to the B locus, being separated by a repetitive GC region (Briles et al 1993, Miller et al 1994, 1996, Rogers et al 2003, Salomonsen 2014). Sequencing of the BF/BL region was initially shown in the B12 haplotype and revealed striking differences to that of the typical mammalian MHC. The chicken BF/BL region is small (92 kb) and simple, encoding just 19 genes (Kaufman et al 1999), organised as predicted previously (Pink et al 1977, Gulliemot et al 1988) and was shown to be consistent in 15 different haplotypes (Jacob et al 2000, Shaw et al 2007, Hosomichi 2008).

The chicken MHC region is even more compact (44 kb) containing only 11 genes, with many of the mammalian MHC genes being present including MHCI, MHCII B, TAP, tapasin and DM genes. It is important to note that there are only two classical class I genes (BF1 and BF2) and

two classical class II B genes (BLB1 and BLB2). Interestingly, not all typical mammalian MHC genes are found in or near to the chicken MHC, including the classical class II α, inducible proteasome (low molecular mass polypeptide, LMP) and DO genes, and has been described as a minimal essential MHC (Kaufman et al 1999).

The organisational structure of the chicken MHC is vastly different from that of typical mammals. As mentioned previously, for the typical mammalian MHC, the class I and II regions are separated by the class III region. In the chicken, the class II and I regions are located next to each other with the class III region on the outside (fig 1.7.1). Importantly, it is noted that the TAP genes are flanked by the classical class I genes and that tapasin is flanked by the class II B genes (fig 1.7.1). It is also observed that the genes involved in peptide loading for class I and II (TAP, tapasin and DM) are all polymorphic whereas in typical mammalian systems they are functionally monomorphic or have limited polymorphism (Jacob et al 2000, van Hateren et al 2013, Walker et al 2011).

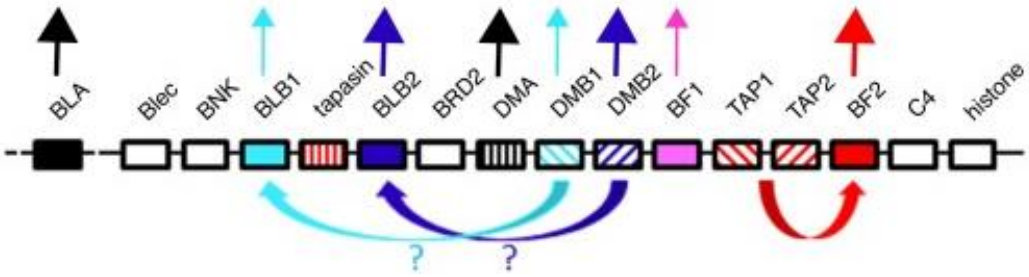


Figure1.7.1 Organisation and co-evolution in the chicken MHC

The chicken MHC is small and simple containing most genes necessary for antigen presentation. This includes the classical class I and class II B genes, the non-classical class I and class II genes TAP1, TAP2, tapasin, DMA, DMB1 and DMB2. The TAP genes are seen to be flanked by the classical class I genes, with co-evolution being indicated (curved red arrow). The possible co-evolution of class II genes is shown by the presence of curved blue arrows. The differences in expression of all genes (in haematopoietic cells) are shown by varying thickness of straight arrows.

1.7.1 The chicken MHC encodes a dominantly expressed classical class I molecule

Cloning and sequencing of cDNA showed the presence of two classical class I genes expressed in most common MHC haplotypes (Kaufman et al 1999). Surprisingly, the number of cDNA clones showed that one MHCI gene (BF2) was found in much greater abundance than the other (BF1), with up to 10-fold more expression being found in many haplotypes (Kaufman et al 1999, Wallny et al 2006), leading BF2 to be referred to as the major and BF1 the minor class I gene. To understand this difference of expression the whole of both BF genes was sequenced, which showed all BF2 genes and promoter elements were intact. BF1 showed a variety of differences between haplotypes, including deletions and divergence in the promoter elements as well as disruptions to the BF1 gene itself (Kaufman et al 1999, Livant et al 2004, Shaw et al 2007). This difference in BF abundance was also seen at the protein level by two-dimensional electrophoresis in peripheral blood lymphocytes and at the peptide level, where isolation of bound peptides showed all peptide motifs were consistent with the binding motif of the BF2 molecule (Wallny et al 2006).

Dominant class I expression is due to co-evolution with class I antigen processing genes. The polymorphic class I genes are in strong linkage disequilibrium with the polymorphic peptide loading genes (tapasin and TAP), resulting in co-evolution. This co-evolution of the class I system results in specific alleles of the TAP genes that specify peptide translocation of peptides with a motif that correlates with that of one particular MHCI, specifically the dominantly expressed BF2 but not BF1 (Walker et al 2011) (fig 1.7.1). This results in the BF2 molecule receiving many suitable peptides and in contrast, BF1 receives much less and is expected to have become less important in antigen presentation (Walker et al 2011).

1.7.2 MHC-association with pathogen response

As previously mentioned, in humans the MHC region is known to have the most associations to disease resistance and susceptibility, with strong associations being generally with autoimmune diseases. In the chicken, there are strong associations with infectious pathogens that were found by poultry researchers over half a century ago. These associations were initially located to the B locus, which was later shown to contain the MHC. Unlike in humans, there is a wealth of data detailing MHC disease associations to a variety of infectious pathogens, including viruses, bacteria and parasites (Briles et al 1977, Bacon 1981, Plachy et al 1994, Lamont 1998, Hudson et al 2002, Lui et al 2002, Schou et al 2007, Miller and Taylor 2016). The hypothesis of a minimal essential MHC provides a potential explanation for these incredibly strong pathogen associations seen in the chicken. In humans, the TAP genes are monomorphic and able to pump a very broad repertoire of peptides to the multigene family of MHCI molecules, which allows the presentation of a wide variety of peptides by HLA-A, B and C molecules. This allows humans to confer more or less protection to a diverse variety of pathogens, which reads out as weak genetic association to most infectious disease. In the chicken, peptide presentation is restricted by the polymorphic TAPs to one particular class I (BF2), which reads out as a strong genetic association, giving rise to the concept that 'chickens live or die by their MHC' (Walker et al 2011). Many studies have shown that the presence of a particular MHC haplotype confers life or death resistance or susceptibility to a particular pathogen, including commercially important pathogens such as Marek's disease virus (MDV) and Rous sarcoma virus (RSV) as well as pathogens central to zoonosis, for example avian influenza virus (AIV) (Briles et al 1977, Plachy et al 1994, Boonyanuwat et al 2006). Moreover, it could be reasoned that one of the six expressed human class I molecules (in a heterozygote) is more likely to bind a self-

peptide and could explain the strong associations with autoimmunity found with the human MHC.

1.7.3 Generalists and specialists

One major question arising from the discovery of one dominantly expressed MHCI molecule is how the chicken survives with only one MHCI molecule in a homozygote, or two in a heterozygote? Unlike in humans, it was demonstrated that the MHCI molecules of some chicken MHC haplotypes are able to bind a very broad spectrum of peptides (termed promiscuous) compared to what is seen for typical mammalian MHC molecules (Chappell et al 2015, Kaufman 2018). Interestingly, the TAP complex found in these haplotypes were also found to be very promiscuous, pumping a huge variety of peptides (Tregaskes et al 2016). It was proposed that these promiscuous MHC haplotypes present such a breadth of peptides that they can functionally act similarly to that of a multigene family found in typical mammals (Kaufman et al 2018). Between the haplotypes analysed, the breadth of peptide presentation by the chicken MHCI was found to be a spectrum, with some MHCI molecules being promiscuous and other presenting a narrow repertoire that is limited by more restrictive TAPs, termed fastidious MHCI molecules (Walker et al 2011, Chappell et al 2015, Tregaskes et al 2016, Kaufman 2018). The more fastidious MHC molecules are thought to have arisen during particularly virulent pathogen outbreaks, resulting in a purifying selection for a particular peptide presented in high abundance by a fastidious MHCI. This notion has been supported by findings that show fastidious MHCI molecules are much more abundant on the cell surface than that of promiscuous MHCI molecules (Chappell et al 2015). The spectrum of MHCI peptide presentation breadth and cell surface expression has led to the hypothesis that chicken MHCI molecules can be split into generalist and specialist MHCI

molecules, with generalists providing a broad range of protection to many pathogens, whilst specialists provide protection to particularly virulent pathogens (Kaufman 2018).

1.7.4 The chicken MHC class II processing and presentation system

The first MHC gene isolated outside of mammals was a chicken MHCII gene, using a human HLA-DQB cDNA probe which isolated chicken gDNA clones and were then mapped to the chicken BF/BL region (Bourlet et al 1988). Initially, five class II β genes were identified by cosmid isolation for chickens of the B12 haplotype and numbered I-V. Further studies revealed that one cosmid cluster mapped BLB-I and II to the BF/BL region, whilst the remaining three BLBs on other cosmids were located to the unlinked Rfp-Y region. Genes derived from the BF/BL region appeared to be classical MHCII B genes, based on their high levels of polymorphism and sequence diversity and were eventually renamed BLB1 and BLB2, whilst the other three BLB genes had little polymorphism and sequence diversity and therefore appeared to be non-classical MHCII genes and were renamed Y-LB genes (Zoorob et al 1990, Miller et al 1994, Kaufman et al 1999). It was demonstrated in the B12 haplotype that the two BLB genes in the BF/BL region flanked the tapasin gene and are in opposite transcriptional orientation (Kaufman et al 1999).

Similarly to what is seen for MHCI, the chicken has two classical MHC II B chains (BLB1 and BLB2) with only one BLB (BLB2) being strongly expressed at the RNA level in haematopoietic cells (Jacob et al 2000). Analysis of the class II α chain (BLA) cDNA revealed it to be non-polymorphic and much like DRA of humans (Salomonsen et al 2003). In contrast to humans, the class II α chain (BLA) is located outside of the MHC, being mapped 5 cM away from the BF/BL region (Jacob et al 2000). Similar to humans, the chicken has genes encoding the peptide editor and chaperone DM. Humans encode one α chain and one β chain for DM (HLA-DMA and DMB), but interestingly the chicken has two DMB chains (DMB1 and DMB2)

giving the possibility of two DM peptide editors to be expressed and notably, only one of the DMB genes is strongly expressed in haematopoietic cells (DMB2) (Kaufman et al 1999, Parker and Kaufman 2017). Moreover, the non-classical MHCII genes are highly polymorphic unlike typical mammals, reflecting what is seen in the MHCI system. The combination of dominantly expressed BLB2 and DMB2 genes as well as polymorphic DM genes suggested that there may be co-evolution between MHCII genes and its chaperones, as is seen for the MHCI system (Parker and Kaufman 2017) (fig 1.7.1).

1.8 Genome editing

Programmable nucleases that make permanent effects to the genome are a desirable technique and have been developed in multiple forms. Programmable nucleases are designed to make site-specific DNA double-stranded breaks (DSBs); these DSBs are then repaired in one of two ways (fig 1.8.1). The first is the error-prone DNA repair pathway non-homologous end joining (NHEJ) which often results in an insertion or deletion (indel) event that potentially disrupts the target genes function, making a knockout (KO). The second, homology-directed repair (HDR) relies on supplying a 'DNA repair template' that has homology to the surrounding genomic region of the DSB break allowing a specific sequence to be 'knocked in' (KI) to that region (Kim et al 2014).

Zinc-finger nucleases (ZFN) were the first technology for successful precise genome editing. ZFNs have been successfully used to modify genes of various organisms, ranging from viruses and bacteria to mammals such as mice and pigs as well as cultured avian cell lines (Segal and Meckler 2013, Perez-Pinera et al 2013, Urnov et al 2010). ZFNs are comprised of two domains: a DNA-binding zinc-finger protein (ZFP) domain and a DNA nuclease domain that is derived from the FokI restriction enzyme (Kim et al 1996). The DNA-binding domain of FokI is replaced with ZFPs to create a ZFN, two ZFN monomers are needed to create an active

nuclease as the dimerization of the FokI nuclease domains allows DNA cleavage. Each ZFN monomer must bind to adjacent half recognition sites that are separated by the 5-7 bp linker from the FokI domain. The FokI interacting interfaces were modified to require heterodimers to form, as wild-type FokI domains can only form homodimers allowing monomeric binding of DNA to result in cleavage and off-target effects.

ZFN sequence specificity is determined by ZFPs, which consist of tandem arrays of C2H2 zinc-fingers; the most common DNA-binding motif in higher eukaryotes. Each zinc-finger recognizes 3-bp of DNA sequence, and 3–6 zinc-fingers are used to generate a single ZFN subunit that binds to DNA sequences of 9–18 bp. Importantly, the DNA-binding specificities of zinc-fingers can be altered by mutagenesis, which is a key feature of constructing a programmable nuclease.

The next major breakthrough in genome editing came through the development of transcription activator-like effector nucleases (TALENs). Like ZFNs TALENs have been used to edit a whole multitude of organisms including viruses, plants, insects, frogs, fish and mammals, as well as in cultured mammalian cells (Nemudryi et al 2014). The structure of TALENs is similar to that of ZFNs, containing a FokI nuclease domain at the C-terminus.

The differences are found at the DNA binding domains, where they use transcription activator-like effectors (TALEs), which are derived from the plant pathogen *Xanthomonas spp*. TALEs are comprised of tandem arrays of 33–35 amino acid repeats, each of which recognizes a single base-pair. The nucleotide specificity of each repeat domain is determined by two amino acids located at positions 12 and 13, these two amino acids are termed repeat variable diresidues (RVDs). Four different RVD combinations are used to recognise guanine, adenine, cytosine and thymine, the most commonly used are; Asn-Asn, Asn-Ile, His-Asp and Asn-Gly, respectively. As

with ZFN two monomers are needed for dimerization of the FokI nuclease domains to create a functional DNA nuclease and allow DNA cleavage.

TALENs were followed by RNA-guided engineered nuclease (RGEN) technology, which was derived from the adaptive immune system found in some prokaryotic organisms. The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) locus was discovered in 1987 as an unusual repeating cluster that, at the time, had little homology to known sequences. The CRISPR region was found to create small RNAs (Tang et al 2002). From here it was discovered that in bacteria and archaea, an RNA-guided DNA cleavage system provides adaptive immunity against invading phages or plasmids (Makarova et al. 2006). These organisms regularly capture small DNA fragments from the foreign DNA of an invading phage or plasmid and then insert a sequence termed as a protospacer into their own genome to form a CRISPR.

In type II CRISPR systems, these CRISPR regions are transcribed as a pre-CRISPR RNA (pre-crRNA) and are processed to give rise to target-specific crRNA fragments. In addition, an invariant target-independent *trans*-activating crRNA (tracrRNA) is also transcribed from the same locus, which adds to the processing of pre-crRNA (Deltcheva *et al*. 2011). Both the crRNA and tracrRNA are then complexed with CRISPR-associated protein 9 (Cas9), forming an active DNA endonuclease. The resulting targeted DNA nuclease cleaves a 23-bp target DNA sequence, which is comprised of the 20-bp guide sequence in the crRNA (the protospacer) and the NGG sequence known as protospacer adjacent motif (PAM), which is recognized by the Cas9 itself facilitating DNA cleavage (Hsu et al. 2013, Mali et al. 2013, Mojica et al. 2009). Cas9 proteins derived from species other than *Streptococcus pyogenes* recognise different PAM sequences (Cong et al. 2013, Shah et al 2013, Hou et al 2013).

The two RNA components (crRNA and tracrRNA) can be fused to form a chimeric single-chain guide RNA (sgRNA), giving a simplified component for RGEN mediated genome editing (Jinek et al 2012). The most successful CRISPR-Cas9 guided genome editing method uses a plasmid containing the sgRNA with cloning sites to simply insert a 20-bp guide site of choice to specify the editing location (Ran et al 2013) (fig 1.8.1).

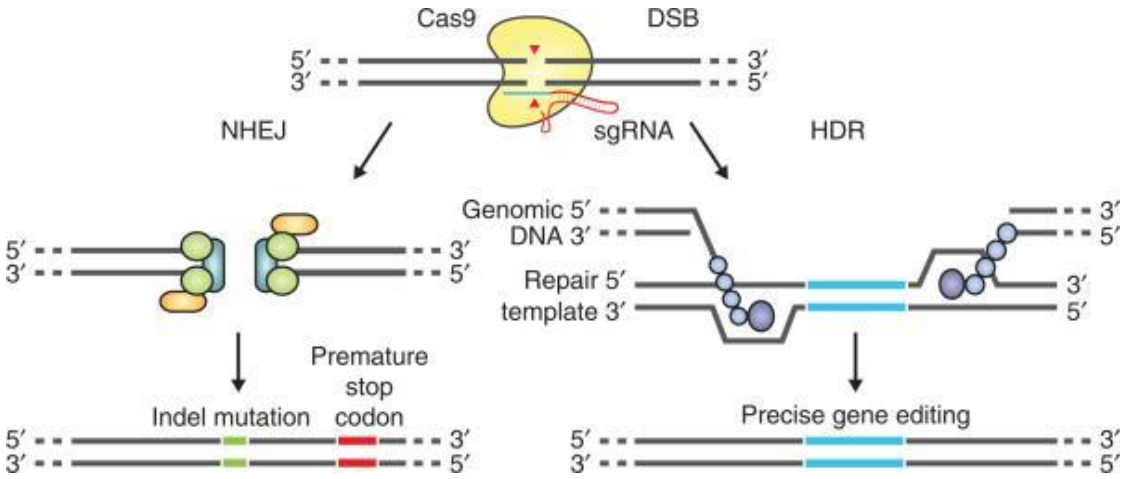


Figure1.8.1 Schematic of targeted genome cleavage and editing by CRISPR-Cas9

The Cas9 nuclease from *S. pyogenes* (in yellow) is targeted to desired genomic DNA by a chimeric sgRNA consisting of a 20-nt guide sequence (blue) and an RNA scaffold (red). The guide sequence pairs with the DNA target directly upstream of a required 5′-NGG adjacent motif. Cas9 mediates a DSB ~3 bps upstream of the PAM site (red triangle). This cleavage is then repaired in either two ways; the error prone NHEJ pathway, often resulting in indel mutations (green bar) or by HDR that results in a precise gene editing event (Ran et al 2011).

1.9 Project aims

The main aim of the project is to gain a deeper understanding of the antigen processing systems in the chicken. This includes antigen processing and selection from both the class I and class II antigen presentation pathway.

1. Assess the importance of particular chicken TAP1 and TAP2 polymorphisms in the activity and specificity of peptide transport and class I loading, with the goal of understanding which residues are important for function.

2. Assess the importance of DMA, DMB1 and DMB2 for expression and peptide loading of the classical class II molecules including BLB1 and BLB2, with the goal of understanding whether there is co-evolution between DM and class II isotypes.

2. Materials and Methods

2. Material and Methods

2.1 Chickens

White leghorn chicken derived lines 6_1 MHC haplotype (B2), C-B4 (B4), C-B12 (B12), 15I (B15), P2a (B19) and 0 (B21) (Shaw *et al.* 2007) were bred and maintained under defined pathogen-free environment at the Institute for Animal Health (Compton, UK). The inbred MHC-homozygous lines 6_1, 15I, and 0 were developed by the Regional Poultry Research Laboratory (East Lansing, MI). Lines C-B4 and C-B12 are sub-lines derived from the Reaseheath line C developed at the Northern Poultry Breeding Station (Reaseheath, UK). The P2a line was acquired from the Institute for Animal Science and Health (Lelystad, The Netherlands) and originated from the line P2a from Cornell University (Ithaca, NY). The histories of some of these lines are reviewed in Schat *et al.* 1980, Simonsen *et al.* 1982, Hála *et al*. 1987, and Briles *et al.* 1982.

2.2 Development of REV cell lines and culture conditions

Chicken cell lines were transformed with replication-defective avian reticuloendotheliosis virus strain T (REV-T). The cell line TG12A5 was generated by *in vitro* REV-T transformation of concanavalin A-activated splenocytes from B12 birds and the TG15 cell line was made from REV-T transformed bursa (E. K. Meziane, personal communication. 2014). These cell lines were generated at the Basel Institute for Immunology by Thomas Göbel (Marmor *et al.* 1993 and Walker *et al*. 2011). REV cell lines were cultured in suspension in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 0.3 g/l L-glutamine and 10 % fetal bovine serum (FBS) (Gibco) at 37 °C, 5 % CO_2.

2.3 Agarose gel electrophoresis

Agarose gels were prepared as 1 % UltraPure™ Agarose (Invitrogen) in 1x Tris acetate EDTA (TAE) buffer (40 mM Tris-acetate pH 8.0, 1 mM EDTA) containing GelRed dye (Biotium) at 1/40,000 and visualised using a FluorChem fluorescent imager (AlphaInnotech).

2.4 Genomic DNA extraction and molecular cloning

Genomic DNA from REV-T cell lines was extracted using GenElute™ Genomic DNA Miniprep Kit (Sigma-Aldrich). Genomic regions of interest were amplified by PCR using the "high fidelity" buffer in the VELOCITY™ DNA Polymerase kit (Bioline) or “high fidelity” in the Phusion® DNA polymerase kit (New England Biolabs). The following PCR mixture was prepared

Component	Amount (μl)
gDNA	1
100mM dNTP's	0.5
5x high fidelity buffer	5
Forward primer (10 μM)	1
Reverse primer (10 μM)	1
DNA polymerase	0.5
Nuclease free H_2O	15.5

The samples were placed in a thermocycler under the following conditions: 5 min at 98 °C, followed by 30 cycles of 60 s at 98 °C, 60 s at 60-70 °C and 240 s at 72 °C followed by 10 min at 72 °C.

The PCR product was run on a 1 % agarose gel, the correct sized band was excised from the gel and purified using a MinElute Gel Extraction Kit (Qiagen). Purified amplicons were ligated into the pJET1.2/blunt vector (Fermentas) which contains a lethal restriction enzyme gene that is disrupted by ligation of a DNA insert into the cloning site. Ligated plasmids were transformed into DH5α chemically competent *E.coli*, prepared by Alicia Martín-López (protocol was adapted from Molecular Cloning, 3rd edition, Sambrook J and Russell D, 2001). Colonies were picked in 2 ml cultures of LB containing 100 μg/ml ampicillin and grown overnight at 37°C and shaking at 250 rpm. Plasmid DNA was purified from bacterial cultures using the QIAprep Spin Miniprep Kit (Qiagen) (based on the method of Vogelstein and Gillespie. 1979). Purified plasmid DNA was screened by PCR and electrophoresis on a 1 % agarose gel to identify plasmids containing an insert of the expected size. Plasmids with an insert of the expected size were confirmed by dideoxy/chain termination sequencing (Sanger *et al.* 1977). This was performed on an Applied Biosystems 3730xl DNA Analyser by the University of Cambridge Biochemistry Department Sequencing Facility, using a T7 sequencing primer (5' - TAA TAC GAC TCA CTA TAG GG - 3').

Primer	Sequence	Gene
UC54	GTGCCCGCAGCGTTCTTC	BLB
UC55	TCCTCTGCACCGTGAAGG	
DMA indel FP	TCTGTGCCTGCAGAGCCAT	DMA
DMA indel RP	TGGCGATTCTGGTGAGCAA	
C321	TCCCCATGCCCCCAACGCAG	DMB1
C322	CCACTTGCATGCAGCGGTGC	
DMB2 indel FP	TGCCTTCATGGTGCATGTG	DMB2
DMB2 indel RP	TCTGTGCCCAGAACTGTGC	
b2m indel FP	CAAGGTGCAGGTGTACTCC	B2m
b2m indel RP	ACTTGTAGACCTGCGGCTC	
T1_intFor	CACCACAGATGCAGAAGGC	TAP1
T1_intRev	ATAGGAGATGTTGGCGTGGAG	
T2_intFor	GCTGCTGTCGGCCTCATG	TAP2
T2_intRev	GGTAGCGGTGCTCGTAGTCC	

Table 2.1 Primers used to assess genome editing.

This table details the primers that were used to amplify the regions surrounding the guide sites for CRISPR-Cas9 editing.

2.5 RNA extraction

Cell lines were harvested and $3x10^6$ cells were used for RNA extraction. RNA extraction was performed using the NucleoSpin® RNA II RNA extraction kit (Machery-Nagel), following the manufacturer's protocol. The area where the RNA extraction was performed, and the pipettes used were first treated with the RNase removing agent RNAseZAP® (Sigma).

2.6 Reverse transcription of RNA to cDNA

RNA was reverse-transcribed to cDNA using an oligo-dT primer and the Maxima H Minus First Strand cDNA Synthesis Kit (ThermoFisher) as per manufacturer's instructions. One microgram of RNA was incubated with dNTPs, oligo-dT, RNase inhibitors and the reverse transcriptase enzyme mix for 45 min at 55°C. Reactions were inactivated at 85°C for 45 min and stored at -20°C.

2.7 Reverse transcription quantative polymerase chain reaction (RT-qPCR)

RT-qPCR primers were designed to amplify products between 100–150 base pairs (bp) with an optimal annealing temperature of ~60°C. Primers were checked for specificity against the NCBI database using the BLAST algorithm (Altschul *et al.* 1990). Targets were amplified from 2, 10 and 100 ng of cell line derived cDNA in 10 µl reactions of *Power Up™ SYBR® Green* Master Mix (Thermo Fisher) as per the manufacturer's instructions. Primer pairs were optimised with primers having an amplification efficiency of >80% were considered for further use. All amplifications were performed on a *7500 Real-Time PCR* System (applied Biosystems). Primer pairs were subject to melt curve analysis using the 7500 software (applied biosystems). At 95°C the melt curve was checked for the presence of multiple products. Primer pairs producing one single peak were used for further experimentation. Target genes were amplified under the following conditions: 15 min at 95 °C, followed by 40 cycles of 60 s at 95 °C, 60 s at 60°C and 30 s at 72 °C. Fluorescence data was collected after each cycle and analysed using the 7500 software (applied biosystems), with subsequent analysis preformed in Microsoft Excel.

The amplification baseline threshold was automatically determined by the *7500 Real-Time PCR* System as the point where amplification curves are at the start of the log phase. The cycle at which the amplification curves crossed the baseline threshold were recorded, generating the Ct value. Each sample was normalised to the slope and intercept of the standard curve generated by the 3 x dilution series for every experimental plate. Samples were run in triplicate before results were analysed by the delta delta ct method. The averages of the triplicates for each sample were then represented as fold-increases and fold-decreases in expression between each sample. Error bars represent the standard error of the mean.

Primer	Gene	Sequence
UC111 FP	DMA	ATGCCTGAAGCCAAGGGC
UC112 RP	DMA	TAGGAGAAGCGCATGAAGCC
UC116 FP	DMB2	ATGTTGGTGCTATTGGGGCTG
UC118 RP	DMB2	AGTCGCAGGCGTTGAAGAG
UC230 FP	BLB2	TCTTCGTGTTCCTGCGCGGT
UC232 RP	BLB2	ATCGTGGCGACATCTCCAACG
UC29 FP	BLA	CTCAGACCGTATGGAGGCTG
UC30 RP	BLA	CAGATGAGGACGTTGGGTTC
qPCR DMB1 FP1.2	DMB1	ATGGCCTTCAACAAGAA
qPCR DMB1 RP1.2	DMB1	AAACATGGCAGATGAGGT
qPCR BLB1 FP1	BLB1	ATGCTGAATTAGCTGCTG
qPCR BLB1 RP1	BLB1	GTCTGAGCGACTTCTTGG
HPRT FP	HPRT	GATGAACAAGGTTACGACCTGGA
HPRT RP	HPRT	TATAGCCACCCTTGAGTACACAGAG
GAPDH FP	GAPDH	GTGGTGCTAAGCGTGTTATCATC
GAPDH RP	GAPDH	GGCAGCACCTCTGCCATC
18S FP	18S	CGAAAGCATTTGCCAAGAAT
18S RP	18S	GGCATCGTTTATGGTCGG

Table 2.2 RT-qPCR primers.

This table details the primers that were used for all RT-qPCR experimentation.

2.8 Clustered random interspaced palindromic repeats (CRISPR)-cas9 mediated genome editing

Selected REV-T cell lines were edited using the CRISPR-cas9 genome editing technique (Rang *et al*. 2013). CRISPR-cas9 sgRNA plasmids (table 2.1) were assembled to include targeted guide sites (Rang *et al*. 2013).

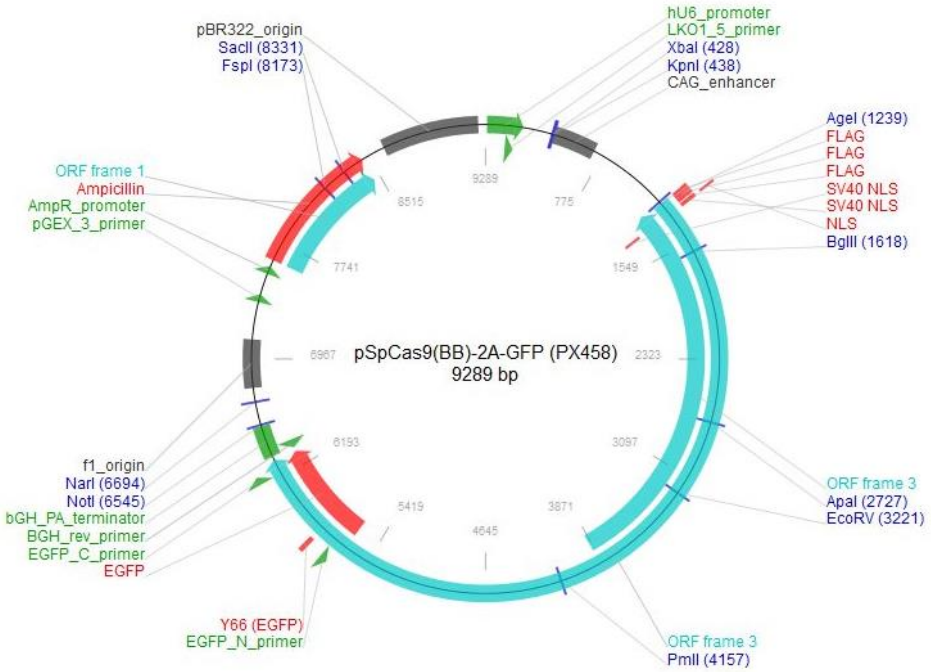


Figure 2.1. Sequence map of the CRISPR-Cas9 sgRNA plasmid (https://www.addgene.org/48138/).

This figure shows the structural organisation of a sgRNA plasmid. Including features such as the human U6 promoter, eGFP and triple flagged regions.

Plasmid name	Selection marker	Addgene reference
PX330	None	42230
PX458	GFP	48138
PX459	Puromycin	48139

Table 2.3 CRISPR-Cas9 sgRNA plasmids.

This table details the sgRNA plasmids used currently.

Initially guide sites were designed to target *DMA*, *DMB1*, *DMB2* and *BLB2* genes using the optimised CRISPR online design tool (http://crispr.mit.edu/). For each gene, early exons were selected to increase the chance of gene KO by indel generation. The genomic DNA sequence for each chosen exon was analysed for viable guide sites using the "optimised CRISPR design" tool. Potential guides sites are assigned a quality score. The quality score

ranges from 0 (many off target effects) to 100 (no off-target effects). Guide sites with the highest quality score were selected for cloning into chosen sgRNA plasmids.

Plasmid name	Sequence 5' to 3'	Location
DMA sgRNA1	gtttggcccggtggcgattc	Exon 2
DMA sgRNA2	catcagtgaggacacatcgt	Exon 2
DMB1 sgRNA1	ctgtaggtccagtccccgtt	Exon 3
DMB1 sgRNA2	ccaacggggactggacctac	Exon 3
DMB2 sgRNA1	gtgcggggacgatgcggacc	Exon 2
DMB2 sgRNA2	ccagatgatggtcacctcgg	Exon 2
BLB2 sgRNA1	ctggtggcactgctggcgct	Exon 1
BLB2 sgRNA2	cccggcggcgggggccgtgc	Exon 1

Table 2.4 Guide sites selected for DMA, DMB1, DMB2 and BLB2. This figure details the guide sites selected for genome editing. The guides were selected from the optimised CRISPR design tool analysis.

2.9 Guide site design using the “sgRNA designer” online tool

Subsequently, guide sites were designed using the sgRNA designer online design tool. Guides were made to target *DMA*, *DMB1*, *DMB2*, *BLB1*, *BLB2* and β_2m genes. Again, early exons were selected to increase the chance of successful gene KO. The genomic DNA sequence for each chosen exon was analysed for potential guide sites by the “sgRNA designer online tool" (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design).

Available guide sites were given an 'on-target efficacy' score. The 'on-target efficacy' score ranges from 0 (low efficacy) to 1 (high efficacy). Guide sites with the highest 'on-target efficacy' score and fewest off target effects were selected for cloning into sgRNA plasmids.

Oligo Name	Sequence 5' to 3'	Location	Broad institute score (0-1)
BI B2m sgRNA1	tcttggtgcccgcagaggcg	Exon 2	0.757
BI B2m sgRNA1	tcttggtgcccgcagaggcg	Exon 2	0.757
BI B2m sgRNA2	gccgacttcacgcccagcag	Exon 2	0.638
B2m sgRNA2	gccgacttcacgcccagcag	Exon 2	0.638
BI B2m sgRNA3	tccatcacgctgatgaagga	Exon 2	0.629
BI B2m sgRNA3	tccatcacgctgatgaagga	Exon 2	0.629
BI B2m sgRNA4	gaacgtcctcaactgcttcg	Exon 2	0.549
BI B2m sgRNA4	gaacgtcctcaactgcttcg	Exon 2	0.549
BI DMA sgRNA1	gagcagctcacggcacagcg	Exon 2	0.834
BI DMA sgRNA2	tcgatgtccccaactcgcag	Exon 2	0.661
BI DMB1 sgRNA1	gcagcagcccccaatcgcaa	Exon 2	0.940
BI DMB1 sgRNA2	gcttctacccttgcgattgg	Exon 2	0.594
BI DMB2 sgRNA1	gcagccaatggctctctgcg	Exon 1	0.770
BI DMB2 sgRNA2	gcggggcttcgacctcaccg	Exon 1	0.633
BI BLB1 sgRNA1	cagcgacgtggggaaatacg	Exon 2	0.677
BI BLB1 sgRNA2	tacgtggctgatacaccgct	Exon 2	0.596
BI BLB2 sgRNA1	ttgcagatacctcacccgct	Exon 2	0.756
BI BLB2 sgRNA2	tttgtggccgattcaccgct	Exon 2	0.730

Table 2.5 Guide sites selected using the sgRNA designer tool. This figure details the 'new' guide sites selected for genome editing for the following genes; β_2m, *DMA*, *DMB1*, *DMB2*, *BLB1* and *BLB2*.

2.10 Assembly of CRISPR -cas9 sgRNA expression construct

The selected sgRNA plasmid was digested with the restriction enzyme Bbs I at 37 ^{o}C for 120 min. The linearised plasmid DNA was run on a 1 % agarose gel; the correct sized band was excised from the gel and purified using MinElute Gel Extraction Kit (Qiagen). Selected guide site oligonucleotide sequences were synthesised (Sigma-Aldrich) for assembly into the sgRNA plasmid. The following mixture was prepared to phosphorylate the guide site insert and to anneal the guide site insert together ready for ligation into the sgRNA plasmid. The prepared mixture was placed into a thermocycler under the following conditions: 37 °C for 30 min; 95 °C for 5 min; ramp down to 25 °C at 5 °C per min.

Component	Amount (µl)
sgRNA top (100 µM)	1
sgRNA bottom (100 µM)	1
T4 Polynucleotide Kinase	1
Milli-Q H_2O	7
Total	10

The guide sites were phosphorylated and annealed together during the treatment, ready for ligation. The guide site inserts were diluted in milli-Q H_2O at a ratio of 1:200. The annealed guide sites were ligated into the linearised sgRNA plasmid by preparation of the following mixture.

Component	Amount (μl)
sgRNA plasmid (5 fM)	1
Guide site insert (50 μM)	1
10x T4 DNA ligase buffer	1
T4 DNA ligase	0.5
Milli-Q H_2O	6.5
Total	10

The ligation mixture was left at room temperature for 60 min. Ligated plasmids were transformed into DH5α chemically competent *E.coli*. Plasmids were incubated with aliquots of bacterial cells on ice for 30 min, incubated 45 s at 42°C, chilled on ice then incubated at 37 °C, shaking at 250 rpm in 500 μl Luria-Bertani (LB) medium (1% w/v Bacto tryptone, 0.5% w/v Bacto yeast extract, 0.5% w/v NaCl). Aliquots (100 μl) of cultures were spread onto LB-agar plates containing 100 μg/ml ampicillin and colonies were allowed to grow overnight at 37°C. Colonies were picked in 2 ml cultures of LB containing 100 μg/ml ampicillin and grown overnight at 37°C and shaking at 250 rpm. Plasmid DNA was purified from bacterial cultures using the QIAprep Spin Miniprep Kit (Qiagen) (based on the method of Vogelstein and Gillespie. 1979). Plasmid DNA was screened for insertion of the guide site by sequencing. The primers designed to sequence the guide site of the sgRNA plasmids used are found in table 2.3. Dideoxy/chain termination sequencing (Sanger *et al.* 1977) was performed on an Applied Biosystems 3730xl DNA Analyser by the University of Cambridge Biochemistry Department Sequencing Facility.

Primer name	Sequence
PX459 FP	GACTATCATATGCTTACCGT
PX458 FP	GGCTGTTAGAGAGATAATTGG
PX330 FP	CACATGTGAGGGCCTATTTC

Table 2.6 Primers designed for Dideoxy/chain termination sequencing of sgRNA plasmid guide sites.

This table shows the primers used to verify the ligation of the chosen guide site sequence into the sgRNA plasmid.

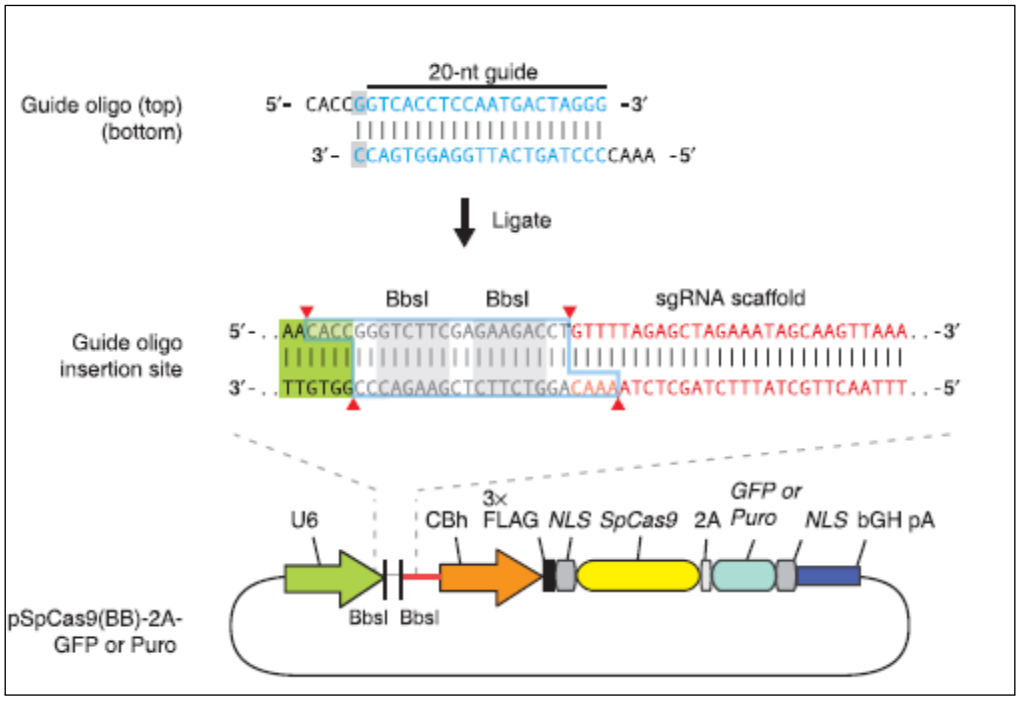


Figure 2.2 (Ran et al. 2013) Schematic for scarless cloning of the guide sequence into a sgRNA expression construct.

The guide oligos for the top strand (blue) have overhangs for ligation into the pair of BbsI sites in pSpCas9(BB) plasmid. The top and bottom strand orientations match those of the genomic target with the top oligo being the 20-bp sequence preceding 5'-NGG in genomic DNA. Digestion of pSpCas9(BB) with BbsI allows the replacement of the Type II restriction sites (blue outline) with direct insertion of annealed oligos. A G-C base pair (grey rectangle) is added at the 5' end of the guide sequence for U6 transcription, this does not adversely affect targeting efficiency.

2.11 Transfection of sgRNA expression constructs

Purified assembled sgRNA construct was used at 5 μg for 3.5 x 10^6 REV-cells diluted in 100 μl Amaxa™ Solution T (Lonza Bio). Cells were transfected by electroporation using the Nucleofector™ II device (Lonza Bio) set to program X-001. Transfected cells were transferred to 6-well plates and maintained in 5 ml RPMI 1640 supplemented with 0.3 g/l L-glutamine and 10 % FBS (Gibco) at 37 °C, 5 % CO_2. Thirty hours post-transfection, depending on the sgRNA plasmid transfected, cells were either sorted for positive GFP expression using an S3 cell sorter (Bio-Rad) or the selection marker puromycin (Sigma-Aldrich) was added at 1 μg/ml. Cells were left to grow to confluence at 37 °C, 5 % CO_2.

2.12 Flow cytometry

Approximately 3 x 10^6 REV-cells were washed twice in PBS (spins at 400 *g* for 3 min at 4°C) and re-suspended in pre-chilled FACS buffer [PBS plus 0.5% bovine serum albumin (BSA) and 0.05% sodium azide (Sigma-Aldrich)]. For intracellular staining, cells were first fixed in 200 μl of 2 % paraformaldehyde (PFA) for 10 min at room temperature. Then permeabilised in 200 μl of permeabilization buffer [PBS plus 0.5% bovine serum albumin (BSA), 0.5 % Polysorbate-20, NP-40 (Sigma-Aldrich) and 0.05% sodium azide (Sigma-Aldrich)] for 15 min at room temperature and were washed twice.

Aliquots (200 μl) of cells were added to a round-bottomed 96-well plate (Greiner) and incubated at room temperature for 1 hour with 200 μl primary antibody dilution. After, cells were spun down at 400 *g* for 3 min at 4°C, the supernatant was removed and cells were washed by re-suspension in 200 μl FACS buffer, spinning and washing was repeated three times. Cells were then incubated for 45 min with the secondary antibody goat anti-mouse IgG-PE (Dako) at 1/200 in FACS buffer, and then washed a further three times. Fluorescence

was then detected using a FACScan II Analyser (Becton-Dickinson) and CellQuest software (BD). Data was post analysed by FloJo software (www.flowjo.com).

2.13 Membrane enriched lysates

Cell lysates were prepared according to a method described in Walker *et al.* 2011. Cultured REV-cells were spun down (at 400 *g* for 3 min at 4 °C), washed twice in PBS and re-suspended in 1 ml PBS. Twenty microlitres of cells were mixed with 20 µl Trypan Blue solution (Sigma-Aldrich) for live/dead discrimination, the live cell count was estimated using an improved Neubauer haemocytometer (Thermo Scientific). Cells were re-suspended in 1 ml cold 'freeze/thaw buffer' [1 mM $MgCl_2$ in PBS solution with 0.1 mM 4-(2-amino ethyl) benzenesulfonyl fluoride (AEBSF; Pefabloc; Roche)], frozen on dry ice and thawed at room temperature twice and centrifuged at 17,900 *g*, 4 °C for 1 hour. Pellets were solubilised on ice for 30 min in digitonin lysis buffer [150 mM NaCl, 1 mM $MgCl_2$, 10 mM TrisCl, pH 8, with 1% digitonin (Calbiochem) and 0.1 mM AEBSF] to give 2 x 10^8 cells equivalent per ml and centrifuged as described above for 10 min. Supernatants were used immediately or stored at -80 °C.

2.14 SDS-PAGE and Western blotting

Western blotting was performed according to an optimised method detailed in Walker *et al.* 2011. Ten microlitre aliquots of membrane-enriched REV-cell lysates, prepared as above, were heated at 95 °C for 5 min in 15 µl 1 x Laemmli sample buffer (2.5 % SDS, 50 mM TrisCl, pH 8, 20 % glycerol, 0.1 % bromophenol blue, with or without 5 % β-mercaptoethanol as a reducing agent). Proteins were electrophoresed on 4 % stacking/12 % separating

polyacrylamide/bis-acrylamide (37.5:1) gels. Standards were XP Western Blot Standards (Invitrogen) and BlueEye Prestained Protein Standards (Jena Bioscience). Gels were soaked for 2 min with rotation in 'transfer buffer' (25 mM Tris, 192 mM glycine, 20 % methanol, 0.0375 % SDS). Gels were transferred in the same buffer to Hybond-C Extra supported nitrocellulose membranes (GE Healthcare) for 40 min using the Trans-Blot SD semidry transfer apparatus (BioRad). Membranes were blocked overnight with 'block buffer' [TBS-T (150 mM NaCl, 50 mM TrisCl pH 8, 0.05 % Tween-20) plus 3% milk powder]. Membranes were then incubated with primary mAb diluted in block buffer for 1–2 hours at room temperature or overnight at 4 °C. Membranes were washed three times for 1 min and one time for 15 min with TBS-T and incubated with secondary antibody (goat anti-mouse IgG-HRP, Sigma-Aldrich) at 1/2500 in block buffer for 30 min at room temperature before washing as previously described. Membranes were incubated with chemiluminescent detection reagent (ECL reagents, GE Healthcare) and exposed to XAR-5 film (Konica Minolta) or the membrane was directly measure using the G:BOX Chemi XX6 (syngene).

2.15 Antibodies

Monoclonal antibodies (mAbs) to chicken class I heavy chain (F21-2), β2m (F21-21) and MHC class II β chain (2G11) were raised against chicken erythrocytes and lymphocytes (Crone *et al*. 1985). mAbs were raised to the cytoplasmic tail of DMA (DMA-8 and DMA-31), DMB1 (DMB1$_2$-3) and DMB2 (DMB2-19). Synthetic peptides were generated by Lawrence Hunt, and hybridoma supernatants were generated and provided by Karsten Skjødt and colleagues.

2.16 Radiolabelled Transport Assay

For the transport assays the cells were harvested, washed twice in PBS, resuspended in ICT buffer (78 mM KCl, 4 mM MgCl2, 8.37 mM CaCl2, 10 mM EGTA, 1 mM DTT, 1 mg/mL BSA, 50 mM Hepes, pH 7) and then were divided into aliquots of 5×10^6 cells. Each sample was treated with 2 µg (1,500 hemolytic units) streptolysin O (Sigma) at 37 °C for 45 min to permeabilize the cells. Cells were washed twice in ice-cold ICT buffer, before being resuspended in 750 µL ICT buffer. 1 µL of radiolabelled transport peptide (with or without competitor peptide) and ATP to 2 mM final concentration was added, before samples were incubated at 41 °C for 5 min. The cells were then lysed by the addition of 400 µL of lysis buffer (1% Triton X-100, 150 mM NaCl, 50 mM TrisCl, pH 8) with 1 mM AEBSF and 1 mM iodoacetamide and incubated on ice for 10 min. The subcellular debris was removed by centrifugation at 17000g for 10 min at 4 °C. The supernatant was added to silica spin columns containing 50 µL of 50% Con A–Sepharose (Sigma) in lysis buffer. The Con A–Sepharose was washed three times with 1 mL wash buffer (TBS, 0.1% Triton X-100) by centrifugation for 10 min at 17000g. Columns were then place in γ-counter tubes (Perkin-Elmer 1270-401) and counted in a γ-counter. Each Sample was performed in in triplicate and results were analysed in Microsoft Excel.

3. Establishing CRISPR Cas-9 mediated genome editing in avian cells

3. Establishing CRISPR Cas-9 mediated genome editing in avian cells

3.1 Introduction

The modification of the genome at a precise locus, referred to as targeted genome editing, is a broadly applicable technology capable of furthering research in many fields. The ability to specifically alter aspects of a genomic region can be used to answer many biological questions, for example the disruption of targeted genes in a view to determining function. Techniques such as small interfering RNAs (siRNAs) and microRNAs (miRNAs) have been employed to silence protein coding of specific genes, although gene 'knockdown' (KD) by these means are often incomplete and nonspecific, as well as being transient in their effect (Krueger *et al.* 2007, Jackson *et al.* 2003). The field was drastically changed with the development of the CRISPR-Cas9 system for precision gene editing, using a method derived from the adaptive immunity system found in some prokaryotic organisms (Ishino et al 1987).

The CRISPR systems fall into two classes, class I and class II, but it is the class II system that has primarily been ulitlised as a genome editing tool. The composition of the class II system sees the CRISPR regions transcribed as a pre-CRISPR RNA (pre-crRNA) and are processed to give rise to target-specific crRNA fragments. In addition, an invariable target-independent *trans*-activating crRNA (tracrRNA) is also transcribed from the same locus, which adds to the processing of pre-crRNA (Deltcheva *et al*. 2011). Both the crRNA and tracrRNA are then complexed with CRISPR-associated protein 9 (Cas9), forming an active DNA endonuclease. The resulting targeted DNA nuclease cleaves at a 23-bp target DNA sequence, which is comprised of the 20-bp guide sequence in the crRNA (the protospacer) and the NGG sequence known as protospacer adjacent motif (PAM) The PAM is recognized by the Cas9 itself facilitating DNA cleavage (Hsu *et al*. 2013, Mali *et al.* 2013, Mojica *et al.* 2009). This system was quickly adapted into a versatile tool following the publication of a method illustrating simple and efficient

assemble of specific single guide RNAs (sgRNAs) which are comprised of the crRNA and tracrRNA fused together. A simple cloning system was introduced enabling the 20 nucleotide protospacer or guide site to be changed to a target of choice, allowing genome editing to become easily accessible (Zhang *et al* 2015).

The original aim was to understand the immunological role of the two MHC classical class II molecules, BLB1 and BLB2, as well as the three non-classical class II molecules; DMA, DMB1 and DMB2. In mammals, the class II peptide editor DM is comprised of one α and one β chain pairing. In the chicken it was predicted that DMα is encoded for by DMA and DMβ by DMB2 but there is the potential of a second DM editor comprised of DMα and DMβ1 from DMB1. We began editing these genes as a starting point for disentangling the mechanism of class II presentation. This chapter details the initial teething problems when establishing the CRISPR-Cas9 procedure in chicken cell lines and ultimately the development of a robust protocol for genome editing in chicken lymphocytes.

3.2 Attempts to edit DMA and DMB1 in the TG12 cell line

TG12 REV-cells are fast growing cells that were derived from CB B12 chickens (that have been fully sequenced for the BF/BL region) and were initially selected for targeted genome editing using the CRISPR-Cas9 system. In view to understanding the role of DM in peptide loading and presentation, the genes *DMA* and *DMB1* were targeted for a DSB in exon 2 to disrupt the gene by indel generation from NHEJ. Guide sites complementary to viable cut sites were identified from the genomic DNA sequence using the online design tool "optimised CRISPR design" (http://crispr.mit.edu/); guides with the fewest off-target effects were chosen and assembled in the PX459 plasmid.

Two sgRNAs were created to target DMA, DMA sgRNA1 and DMA sgRNA2 and two sgRNAs were created to target DMB1, DMB1 sgRNA1 and DMB1 sgRNA2. Four TG12 cell aliquots were separately transfected with the four guides, two to DMA (DMA sgRNA1 and DMA sgRNA2) and two guides to DMB1 (DMB1 sgRNA1 and DMB1 sgRNA2). The cells transfected with guides to DMA were then analysed by flow cytometry of permeabilised cells. Both cell pools showed no reduction in DMA expression with the DMA expression of transfected pools matching mock transfected samples (fig 3.2.1).

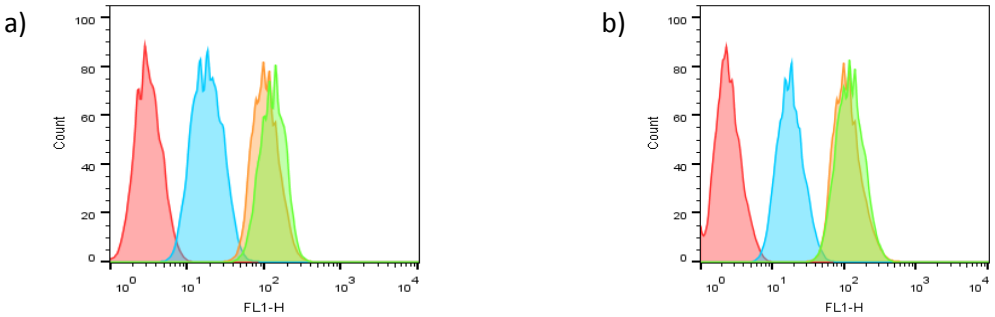


Figure 3.2.1 TG12 REV cell pool transfected with a DMA targeting sgRNA show no reduction is DMA expression.

TG12 cells were separately transfected with a sgRNA:cas9 plasmid targeted to a genomic DNA sequence of DMA. Cell pools were analysed via flow cytometry using the antibody DMA-31, which was raised to the membrane-proximal peptide of the cytoplasmic tail. TG12 cells transfected with PX459 DMA sgRNA1 (a) and PX459 DMA sgRNA2 (b) were stained after permeabilisation for DMA (green) and compared to untransfected stained cells (orange). Untransfected and unstained cells (red) and untransfected cells stained with only the secondary antibody (blue) were included.

To confirm this result, detergent lysates were made from membranes prepared from the pools of cells transfected with TG12 DMA sgRNA 1 and 2, and untransfected TG12 cells. The lysates were assessed for amount of protein by western blotting. The TG12 DMA sgRNA 1 and 2 showed no reduction in protein expression compared to the negative control (fig 3.2.2). These results indicate that the sgRNAs had not disrupted the *DMA* gene to a discernible level in the transfected pools. Transfected cell pools were cloned, and the single cells were expanded, enabling the assessment of DMA expression in individual clones.

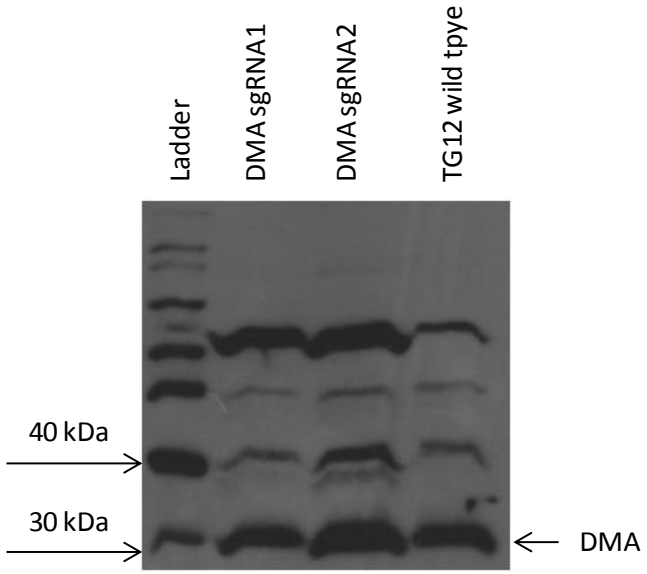


Figure 3.2.2 TG12 REV cell pools transfected with a DMA targeting sgRNA show no reduction is DMA protein expression.

TG12 cells were transfected with two sgRNA:cas9 plasmids (DMA sgRNA1 and sgRNA2) targeted to a genomic DNA sequence of DMA. Cell pools were lysed using digitonin after a membrane enriching protocol. The lysates were analysed via western blotting using the antibody DMA-31.

In addition to analysing the DMA transfected pools, the transfected pools targeting DMB1 (TG12 DMB1 sgRNA1 and 2) were cloned as single cells by limiting dilution. These clones were screened by western blotting for DMB1 protein expression. No screened clones displayed significant reduction in DMB1 protein levels (fig3.2.3).

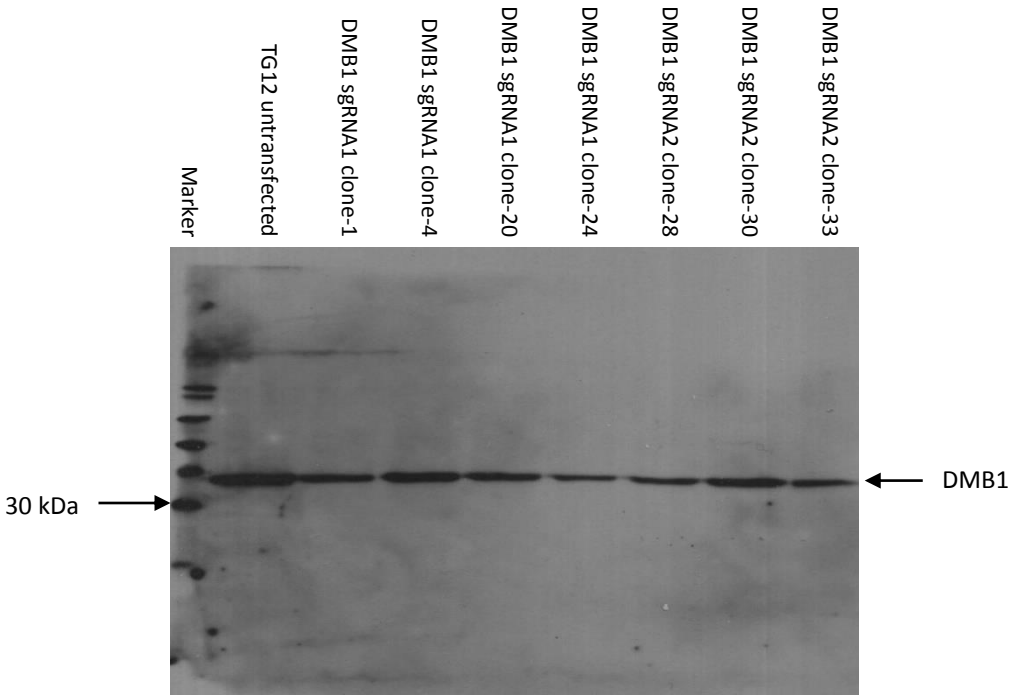


Figure 3.2.3 Single clones of TG12 cells transfected with DMB1 targeting sgRNA1 and sgRNA2 show no strong reduction in DMB1 protein expression.

TG12 cell pools separately transfected with DMB1 sgRNA1 and sgRNA2 were cloned via limiting dilution and clones were expanded. Clones were lysed using digitonin after a membrane enriching protocol. The clone lysates were probed for DMB1 protein expression via western blotting using the antibody $DMB1_2$-3.

3.2.1 Screening of DMA clonal lines

The clones from TG12 cells that had been transfected with DMA sgRNA1 or sgRNA2 were screened by western blot. One of the screened clones showed almost a complete reduction in DMA expression, DMA sgRNA1 clone 13 (DMA.1 cl-13) (fig3.2.4).

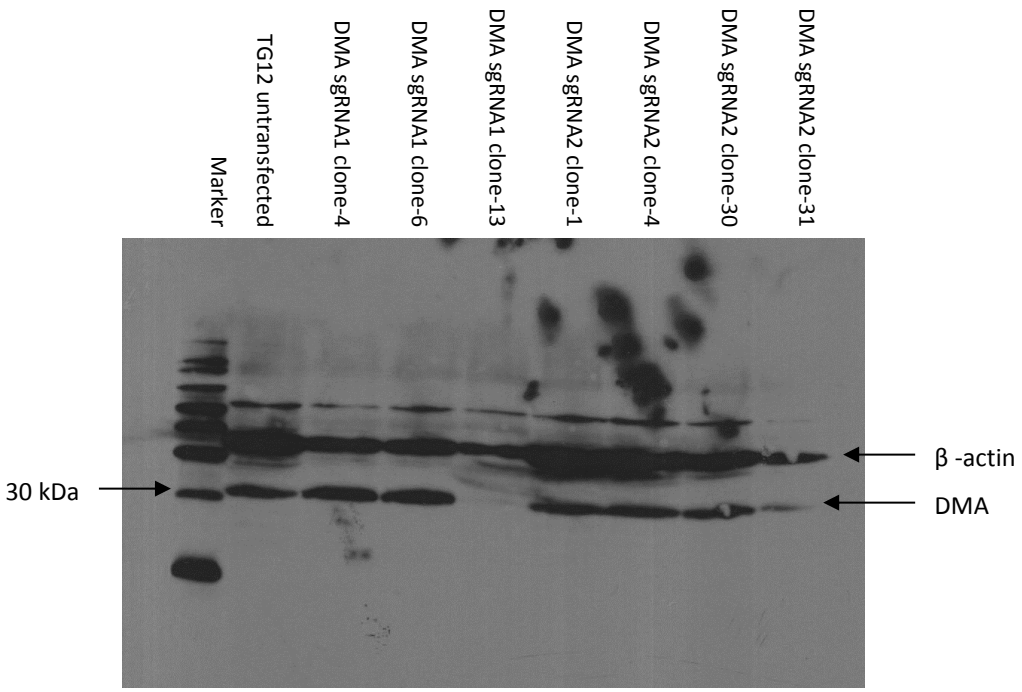


Figure 3.2.4 TG12 DMA sgRNA1 clone-13 shows a dramatic reduction in DMA protein expression.

TG12 cell pools separately transfected with DMA sgRNA1 and sgRNA2 were single cell cloned via limiting dilution and clones were expanded. Clones were lysed using digitonin after a membrane enriching protocol. The clone lysates were probed for DMA protein expression via western blotting using the antibody DMA-31. Clones were subsequently stained for β -actin levels as a loading control.

In mammalian systems, *DMA* comprises the α chain of the peptide editor DM, which mediates the dissociation of CLIP; a loss of DMA would disrupt DM function and the exchange of CLIP in class II molecules. Class II molecules bound with only CLIP would be expected to have a less stable confirmation, resulting in sensitivity to SDS.

Non-boiled class II from normal cells run at a mobility of 50 kDa by SDS gel electrophoresis, whereas non-boiled class II from DMA KOs do not as the SDS in Laemmli sample buffer can destabilise class II molecules that are stabilised by CLIP (Pious *et al*. 1994). Therefore, clone-13 class II protein levels were compared to TG12 untransfected cells and two negative clones (clones with an undiminished level of DMA protein) with and without boiling in SDS.

Non-boiled untransfected TG12 and negative clone lysates showed a similar level of MHCII protein at a mobility of ~50 kDa while clone-13 showed a significant reduction (fig3.2.5). For most samples, the boiled and reduced negative clone samples showed a ~28 kDa band indicating the presence of free class II β chain. However, clone-13 shows significantly reduced free class II β chain in boiled samples, so it is not stability but the amount of MHCII β chain that is reduced in clone 13. One possibility to explain this finding is that this particular MHCII allele is able to dissociate CLIP without DM but DM is still integral in stabilising the MHCII. In this case with DMA absent the MHCII could degrade more quickly resulting in less MHCII overall. Another possibility was that CRISPR caused a large deletion in clone 13, removing not only DMA but other nearby genes including class II β chain genes.

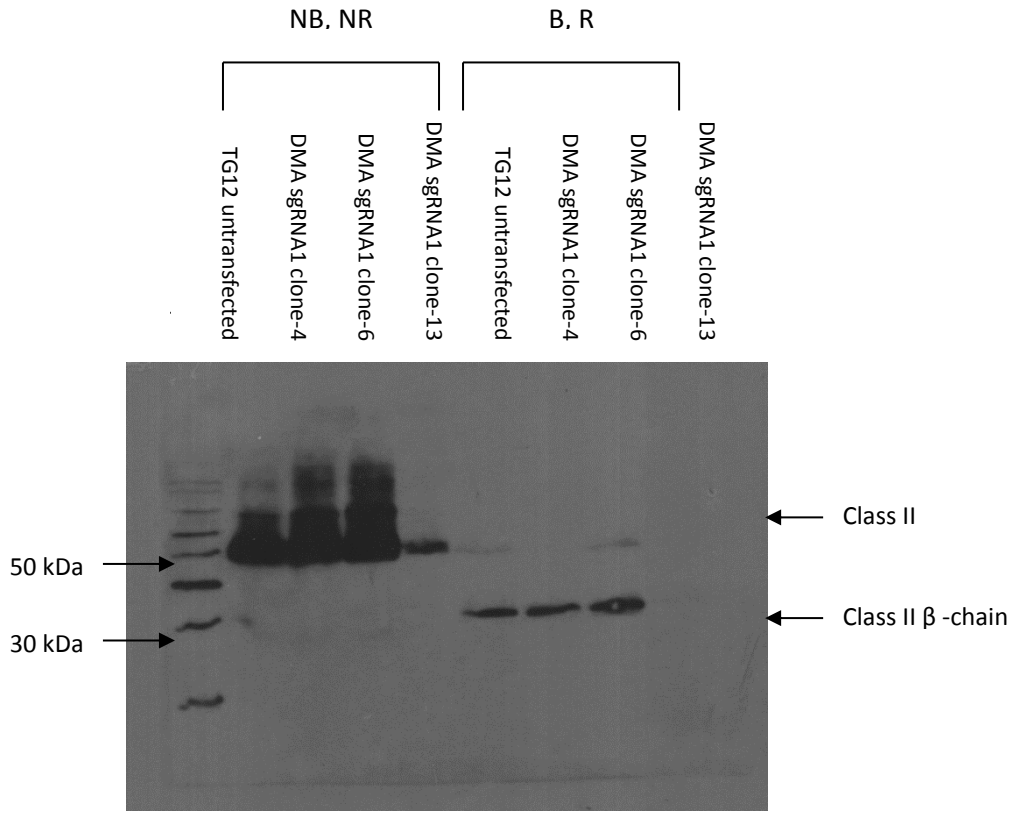


Figure3.2.5 TG12 DMA sgRNA1 clone-13 shows a significant reduction in BLB2 protein expression.

TG12 DMA sgRNA1 clone-13 lysate was probed for BLB2 protein expression using the antibody 2G11. The protein expression of clone-13 was compared to two negative clones, 4 and 6 as well as untransfected TG12 cell lysate. Samples were either 'non-boiled ' and ' non-reduced' (NB, NR) or 'boiled' and 'reduced' (B, R).

When a CRISPR-Cas9 KO is generated by NHEJ, the size of the insertion or deletion can vary dramatically. Current literature suggests that indels are usually small, between 1-10bp, although indels over 1 kb have been reported. To assess whether a large deletion had taken place during CRISPR-Cas9 editing, untransfected TG12, two negative clones and clone-13 lysates were probed for tapasin expression by western blotting. All samples showed the same level of tapasin expression (fig3.2.6), demonstrating that any successful indel generation had not disrupted the *tapasin* gene which is located next to *BLB2*, suggesting that a large scale deletion has not occurred.

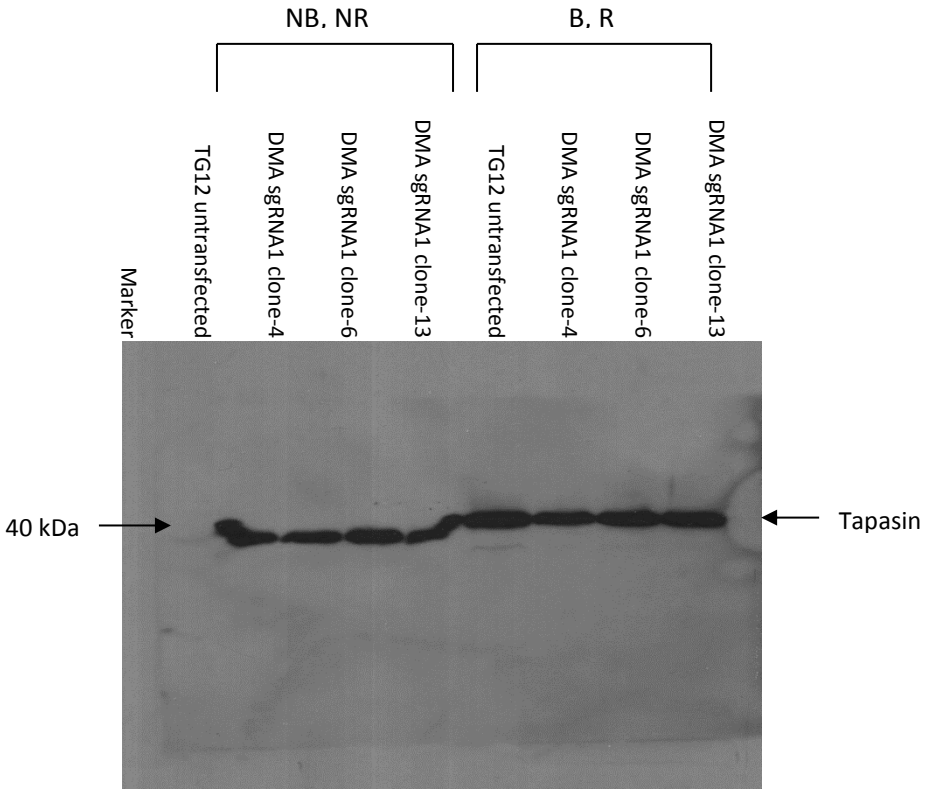


Figure 3.2.6 TG12 DMA sgRNA1 clone-13 shows no significant reduction in tapasin protein expression.

TG12 DMA sgRNA1 clone-13 lysate was probed for tapasin protein expression using the antibody TPN. The protein expression of clone-13 was compared to two negative clones, 4 and 6 as well as untransfected TG12 cell lysate. Samples were either 'non-boiled ' and ' non-reduced' (NB, NR) or 'boiled' and 'reduced' (B, R).

To further examine the reduced protein expression of class II in TG12 DMA.1 cl-13 found by western blotting, the levels of class II on the cell surface were assessed by flow cytometry.

Clone-13 showed a clear reduction of class II cell surface expression compared to untransfected TG12 cells (fig 3.2.7). In contrast, no change can be observed in class I expression, with levels of β_2m being the same in clone-13 and untransfected TG12 cells (fig 3.2.7).

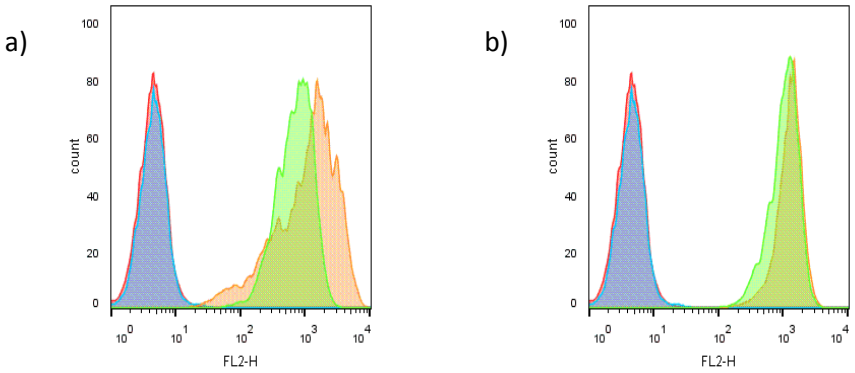


Figure 3.2.7 TG12 DMA sgRNA1 clone-13 shows a reduction in cell surface class II but not class I expression.

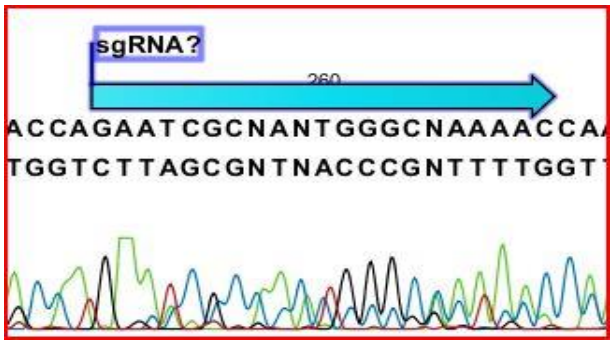
TG12 DMA clone-13 cells were analysed for extracellular B_2m and MHCII protein expression using the antibodies F2-21 and 2G11 respectively. a) DMA clone-13 (green) class II protein expression were compared to untransfected TG12 cells (orange). b) DMA clone-13 (green) β_2m cell surface levels were compared to untransfected TG12 cells (orange). In both histograms untransfected and unstained cells (red) and untransfected cells stained with only the secondary antibody (blue) were included.

In order to confirm that clone-13 was a CRISPR-Cas9 derived KO, genomic DNA was extracted from clone-13 and untransfected TG12 cells, the *DMA* region was amplified by PCR and the amplicons were sequenced. In the first instance the *DMA* PCR product was sequenced directly. If an indel event had occurred, the indel should be seen in the chosen 20 bp targeted guide site as the cas9 enzyme cuts at the 18th bp of the guide site.

Direct sequencing of the amplicon from DMA.1 cl-13 cells shows a messy trace sequence possibly with a modified guide site (fig 3.2.8). The sequence shows the addition of an extra thymine in the guide site at the 3' end (position 2-5) of the antisense guide site, making the chosen guide site 21 bp long on one chromosome; the chromatograph shows conflicting

sequence, which would be expected if editing had occurred (fig 3.2.8). This insertion is not present in untransfected TG12 cells, which has a much clearer trace sequence (fig 3.2.8).

a)



b)

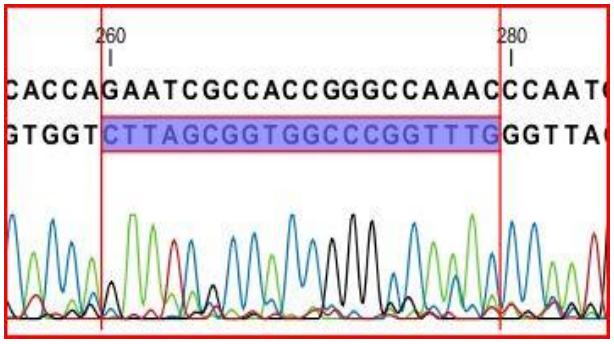


Figure 3.2.8 TG12 DMA sgRNA1 clone-13 shows a potential indel in the sgRNA target site in the DMA gDNA sequence.

Genomic DNA was extracted from TG12 DMA clone-13 and untransfected TG12 cells, DMA was amplified and directly sequenced via Sanger sequencing using the primers C532 and C534. a) gDNA sequence of DMA clone-13 showed the addition an of a thymine in position 2-5 in the targeted guide site. b) gDNA sequence of untransfected TG12 cells showed no indel generation in the targeted guide site.

To confirm that an indel had been generated in DMA.1 cl-13, the DMA amplicon that was sequenced previously was cloned into the pJET cloning vector. A total of 20 colonies where picked, cultured and the DNA extracted, the size of the insert was then checked by PCR before sending 10 clones for sequencing. The guide site remained intact for all 10 clones selected for sequencing, a representative selection of these clones are shown below (fig 3.2.9). These results shows that an indel had not been generated in DMA.1 cl-13 and that CRISPR mediated genome editing had not occurred. This result does not agree with figure 3.2.7, the previous result could have stemmed from a mixed sequence being present in DMA.1 cl-13.

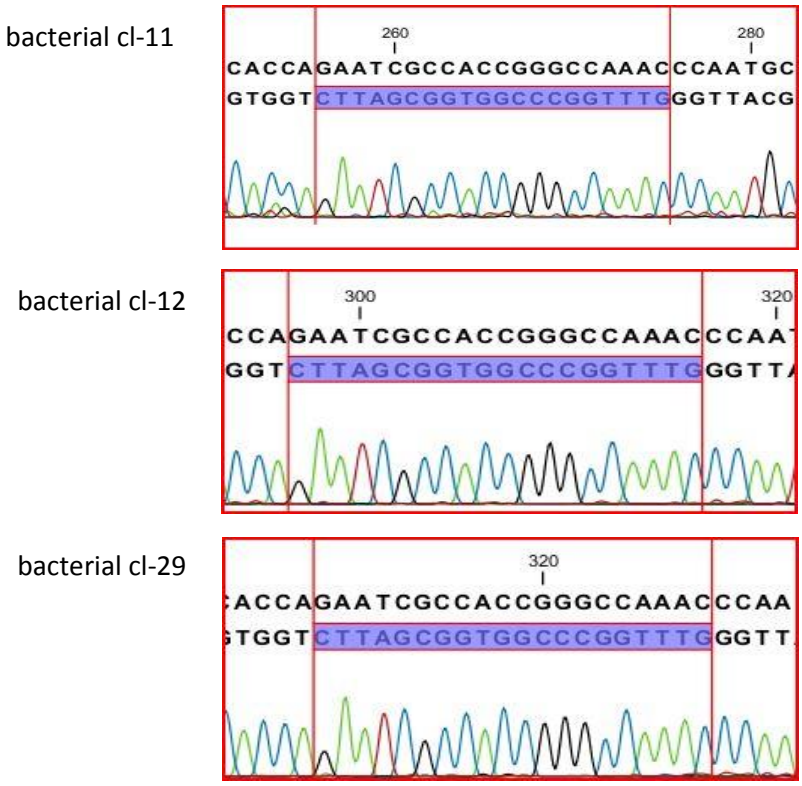


Figure 3.2.9 TG12 DMA sgRNA1 clone-13 shows no indel generation in DMA gene.

DMA amplicons from TG12 DMA.1 clone-13 was cloned in to the pJET 1.2 vector. Bacterial clones were screened via PCR and then sequenced via Sanger sequencing. No indels were found in the targeted guide site for DMA clone-13 pJET clone-11 , DMA clone-13 pJET clone-12 and DMA clone-13 pJET clone-29. The guide site is highlighted in purple.

As no indel event in the cloned amplicons from clone-13 cells could be detected, DMB1 and DMB2 protein expression was assessed, to check for down regulation of other class II associated proteins. A clear reduction of DMB1 and DMB2 protein expression in clone-13 cells was seen by flow cytometry (fig 3.2.10). No reduction could be seen in the class I related protein β_2m, suggesting a defect in the class II pathway of clone-13 rather than successful genome editing. A possible explanation for this global down regulation of class II proteins could be changes in the class II, major histocompatibility complex, transactivator (*CIITA*) gene. The *CIITA* gene encodes a protein that regulates class II gene expression; therefore a random mutation in this region would result in reduced expression of class II related genes in DMA.1 cl-13. The possibility of a naturally occurring mutation in clones derived from single cells

illustrates the importance of having a homogeneous cell population when beginning genome editing. This would reduce the risk of identifying cells with varying protein expression that could be mistaken for genome edited cells.

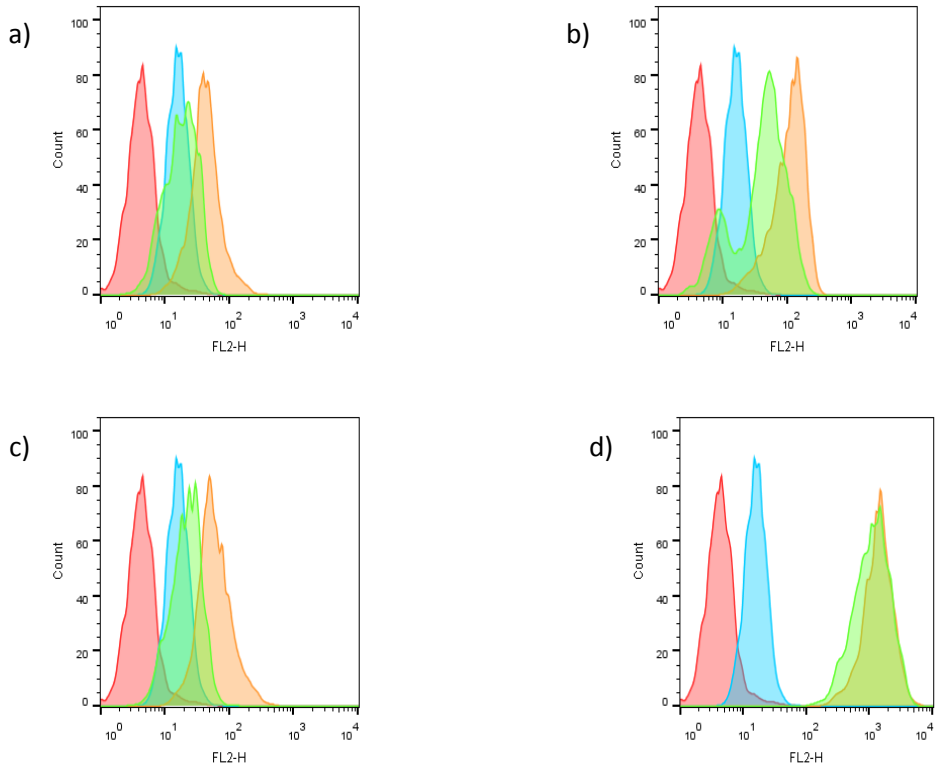


Figure 3.1.10 TG12 DMA clone-13 cells were stained intracellularly for DMA, DMB1 and DMB2 protein expression, using the antibodies DMA-8, DMB1$_1$-19 and DMB2-19.

DMA clone-13 DMA (a), DMB1 (b), DMB2 (c) protein expression (green) was compared to untransfected TG12 cells (orange). d) β_2m protein levels in DMA clone-13 (green) were compared to untransfected TG12 cells (orange). In all histograms untransfected and unstained cells (red) and untransfected cells stained with only the secondary antibody (blue) were included.

3.2.3 Co-transfection of DM targeting sgRNAs with a GFP reporter plasmid

In the previous experiments, transfected cell pools were under no selection meaning that the proportion of transfected cells in the pool was undetermined. To ensure that cell pools only contained transfected cells, two previously designed sgRNAs each targeted to the *DMA, DMB1, DMB2* and *BLB2* genes were co-transfected with a green fluorescent protein (GFP) reporter plasmid. The GFP reporter plasmid allowed GFP positive cells to be isolated by cell sorting. The transfected pools contained ~30 % GFP positive transfected cells with sgRNAs

before sorting. These GFP positive pools were assessed for expression of the respective protein. No reduction in the desired proteins was discernible in any of the transfected cell pools (fig3.2.11). In fact, the targeted protein showed a distinct increase in expression, illustrating that a mock transfected control would be essential for all subsequent experiments.

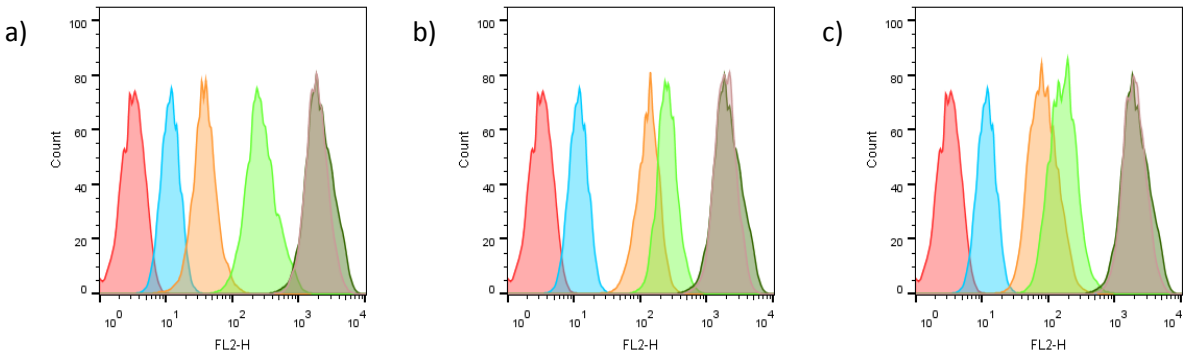


Figure 3.2.11 TG12 REV cells separately transfected with a DMA, DMB1 and DMB2 targeting sgRNA show no reduction is DM expression.

TG12 cells were co-transfected with a GFP expression plasmid and GFP positive cell pool were collected via cell sorting. TG12 cells transfected with either a) DMA.1, b) DMB1.1 and c) DMB2.1 sgRNA shown in green were compared to untransfected TG12 cells (orange) in all histograms. The levels of the targeted protein, DMA (a), DMB1 (b) and DMB2 (c) were stained with the respective antibodies DMA-31, DMB1$_1$-19 and DMB2-19. Untransfected unstained cells (red), untransfected cells stained with only the secondary antibody (blue) and untransfected cells stained for β_2m (brown) were included in all histograms.

3.3 Creation of CRISPR-Cas9 mediated β_2m KO in the IS19 cell line

The previously described attempts to create a CRISPR-Cas9 mediated KO line had failed. The literature regarding CRISPR-Cas9 genome editing details many considerations that may affect the efficiency of genome editing and indel generation. These considerations include the stage of cell cycle, the origin of the cell line, the cell type, the genomic region targeted, the guide site chosen, the nucleotides immediately flanking the guide and PAM site as well the GC content of the chosen region. The previous experiments targeted multiple different genes making it seem unlikely that it was the genomic region selected that was responsible for the previous failed attempts.

Three changes were made in the revised strategy to optimise the CRISPR-Cas9 mediated editing based on the following reflections. One consideration was the transfection efficiency of the sgRNA, as previous experiments had either no selection for transfected cells or relied on the co-transfection of a GFP reporter, which is an indirect method for selecting transfected cell populations. To ensure that cells transfected with the sgRNA were isolated, a new sgRNA plasmid was used that contained a GFP reporter marker (PX458). Secondly, it is possible that all the guide sites selected happened to be low efficiency in their ability to mediate cutting, although many different guide sites had been tested. A third revision was to design the targeted guide site using a different protocol, the "sgRNA designer" online tool (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design) that calculates the on-target efficacy of potential guide sites.

Thirdly, the chosen target for KO was changed to β_2m, as β_2m cell surface levels can be conveniently assessed by flow cytometry as opposed to the DM proteins, which are intracellular. Using the new plasmid and designs detailed above, four β_2m targeting sgRNAs were assembled and IS19 cells were transfected with two of these sgRNAs (one and four). A clear difference was the low transfection efficiency of the GFP containing sgRNAs, with only ~10 % of the cell population being GFP positive (data not shown). Previously ~30 % of the cells were GFP positive when transfected with a GFP reporter plasmid. This discrepancy in transfection efficiency between the GFP tagged sgRNA and co-transfection with a GFP reporter shows that the sorted GFP positive population of cells co-transfected with the GFP expressing plasmid were not highly enriched with the sgRNA. This could explain why CRISPR-cas9 mediated KOs could not be seen in figure 3.2.10, as the sorted population may not have contained many cells transfected with the targeting sgRNA.

The IS19 cells transfected with the sgRNAs, β_2m sgRNA1 and sgRNA4, were sorted for positive GFP expression. The IS19 β_2m sgRNA1 (fig 3.3.1) and β_2m sgRNA4 (fig 3.3.2) cell pools showed a subset of cells that had a reduction in β_2m cell surface expression compared to mock transfected cells. In addition, there was a similar reduction in class I cell surface expression, as would be expected in the absence of β_2m (fig 3.3.1 and 3.3.2). The cell lines IS19 β_2m sgRNA1 and sgRNA4 were both stained for surface expression of class II as a control; unexpectedly a small reduction in class II levels can be seen in both cell lines.

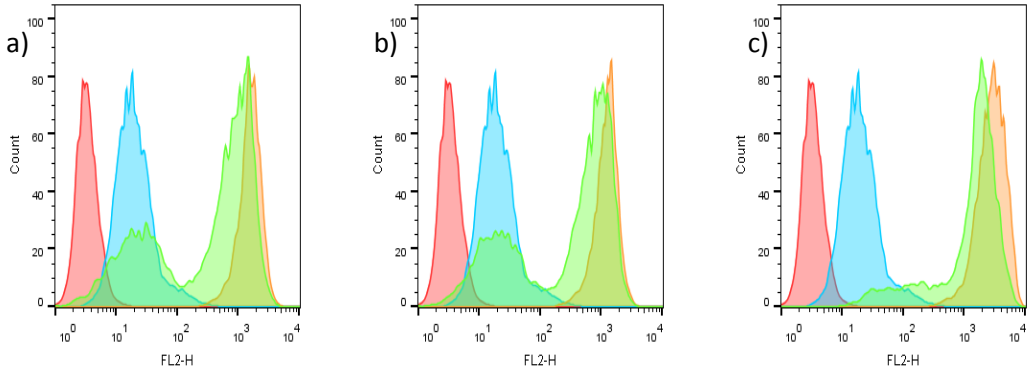


Figure 3.3.1 IS19 REV-cell pool transfected with the β_2m targeting sgRNA PX458 sgRNA1 show a clear reduction in β_2m expression.

IS19 PX458 β_2m sgRNA1 cells (green) β_2m (a), MHC class I (b) and MHC class II (x) protein levels were compared to mock transfected stained cells (orange). The antibodies F21-21, F21-2 and 2G11 were used respectively. Mock transfected and unstained cells (red) and mock transfected cells stained with only the secondary antibody (blue) were included in all histograms.

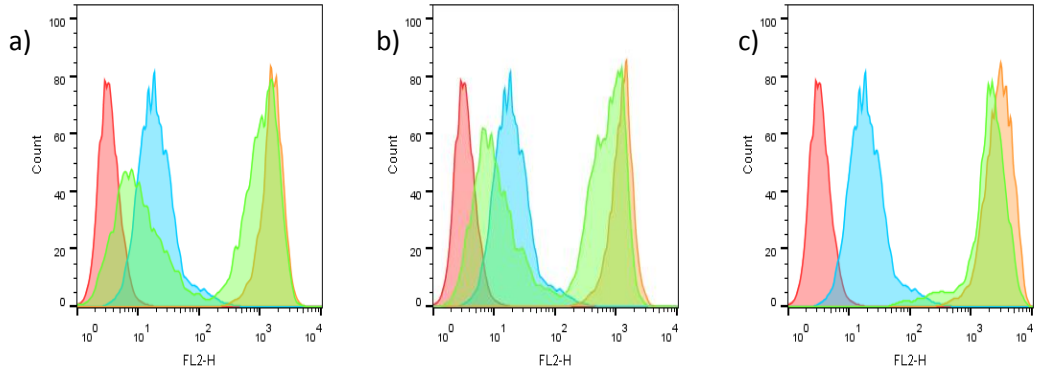


Figure 3.2.2 IS19 REV-cell pool transfected with the β_2m targeting sgRNA PX458 sgRNA4 show a clear reduction in β_2m expression.

IS19 PX458 β_2m sgRNA4 cells (green) β_2m (a), MHC class I (b) and MHC class II (c) protein levels were compared to mock transfected stained cells (orange). The antibodies F21-21, F21-2 and 2G11 were used respectively. Mock transfected and unstained cells (red) and mock transfected cells stained with only the secondary antibody (blue) were included in all histograms.

Next clonal lines of the B2m KOs were made by limiting dilution and assessed for β2m levels (fig 3.3.3). All clonal lines (B2m.1 cl-4, B2m.1 cl-7, B2m.4 cl-2 and B2m.4 cl-4) were shown to have no β2m expression when compared to wild type staining and β2m staining in all KO lines overlaid with the secondary antibody control.

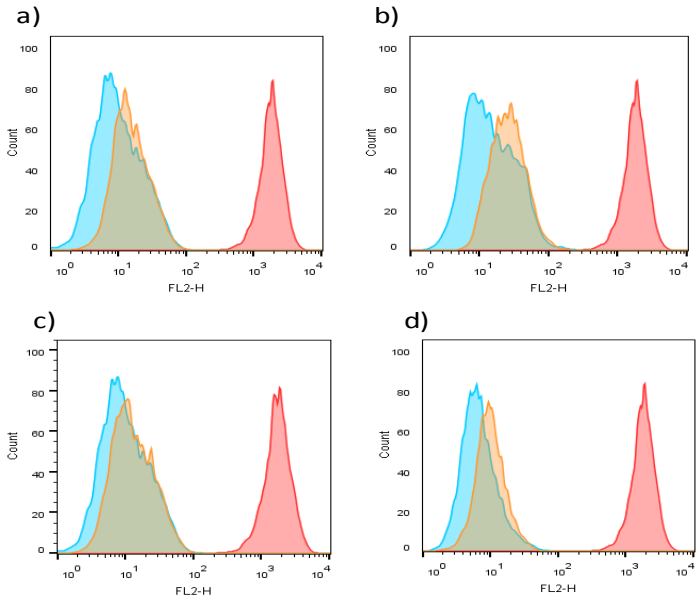


Figure 3.3.3. Flow cytometry cell surface staining of β2m in IS19 β2m KO clonal lines.

IS19 wild type and B2m KO clonal lines B2m,1 cl-4 (a), B2m.1 cl-7 (b), B2m.4 cl-2 (c) and B2m.4 cl-4 (d) were analysed via flow cytometry using the antibody F21-21 to B2m. B2m KO clonal lines were either derived from IS19 cells transfected with PX458 B2m sgRNA1 (a & b) or PX458 b2m sgRNA4 (c & d). B2m KO clonal lines were stained for b2m (orange) and secondary antibody only (blue) and compared to wild type cell expression (red). IS19 PX458 β_2m sgRNA4 cells (green) β_2m (A), MHC class I (B) and MHC class II (C) protein levels were compared to mock transfected stained cells (orange).

β2m KO clonal lines were then screened by PCR, cloning and sequencing of the β2m locus revealed that successful editing had taken place. Figure 3.3.4 shows the B2m KO clonal line, with the PAM site at position 100-102, the edits consist of a two base pair deletion on one allele and a one base pair insertion on the second allele (fig 3.3.4a).

The effect of the editing on the coding sequence when compared to the wild type sequence show disruptions to the gene begin at amino acid position 36 for the two different edits. Although the indels generated a frameshift the amino acid coding sequence, no premature stop codon was generated (fig 3.3.4b).

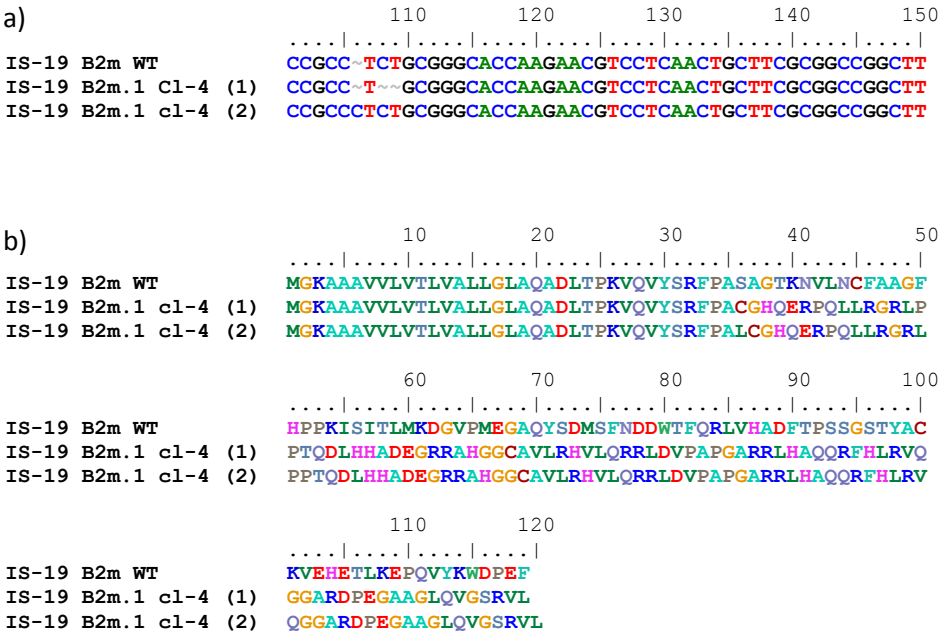


Figure 3.3.4 DNA and amino acid sequence alignment of clonal line B2m.1 cl-4 shows editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of B2m were mapped onto the whole of the B2m DNA sequence and compared to the wild type, editing events can be seen in both alleles immediately before the PAM site at position 100-102. b) The effect of these editing events on the whole amino acid sequence were then shown for each of the two edited alleles. Changes in the coding sequence begin at position 36. The three sequences are labelled on the left-hand margin; wild type sequence (IS19 b2m WT), the first edited allele of clonal line b2m.1 cl-4 (IS19 b2m.1 cl-4 (1)) and the second allele (IS19 b2m.1 cl-4 (2)).

Figure 3.3.5 shows the B2m KO clonal line b2m.1 cl-7 has one successful editing events which can be seen immediately downstream of the PAM site at position 100-102. The edit consists of a two base pair deletion (fig 3.3.5a).

The effect of the editing on the coding sequence when compared to the wild type sequence show disruptions to the gene begin at amino acid position 36. Although the indels generated a frameshift, the amino acid coding sequence has no premature stop codon (fig 3.3.5b).

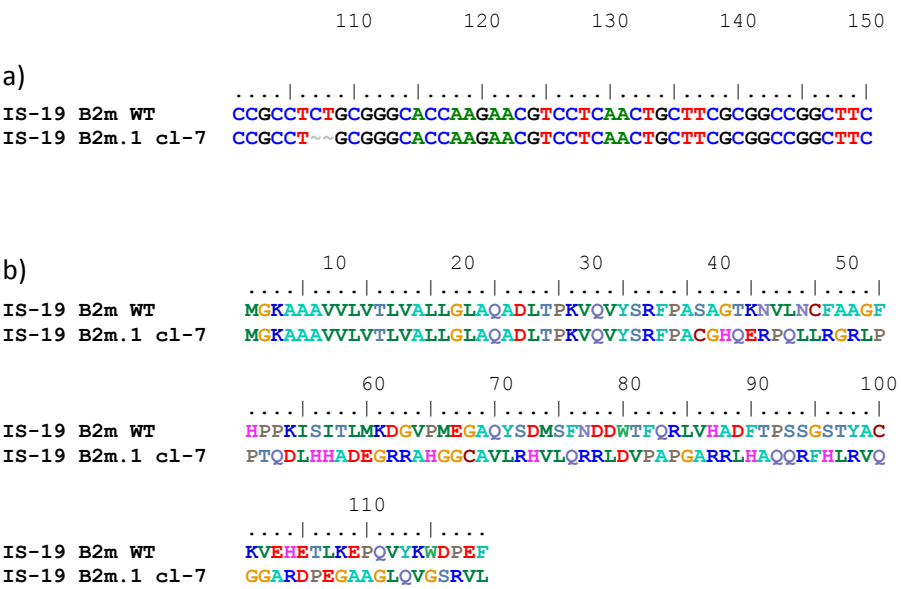


Figure 3.3.5 DNA and amino acid Sequence alignment of clonal line B2m.1 cl-7 shows editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of B2m were mapped onto the whole of the B2m DNA sequence and compared to the wild type, the editing events can be seen in both alleles immediately before the PAM site at position 100-102. b) The effect of the editing events on the whole amino acid sequence were suggested to be the same for both alleles. Changes in the coding sequence begin at position 36. The two sequences are labelled on the left-hand margin; wild type sequence (IS19 b2m WT) and edited alleles of clonal line b2m.1 cl-7 (IS19 b2m.1 cl-7).

3.4 Discussion

The starting point of the project was to establish CRISPR-Cas9 KO lines in chicken cell lines, which had yet to be demonstrated in the literature. The cell line selected was the TG12A5-REV cell line and the target genes were *DMA, DMB1* and *DMB2*. The attempts detailed in section 3 failed to create a KO cell pool for the three genes. Consulting the literature and reviewing the previous failed attempts, I identified what seemed to be the most likely causes.

The first cause was the variability in on-target efficacy of prospective guide sites. Current reports have stated that some guide sites work and that some do not; the mechanism

behind target guide site efficacy is still unclear (Zhang *et al*. 2013). The guide sites that do work vary in their efficiency to generate an indel derived KO. A recent report assessed the variation in sequence including 5' of the guide site, 3' of the PAM site and the guide site itself to derive an on-target efficacy score for potential guide sites. The report details how the on-target efficacy score is experimentally defined based on a logistic regression classifier trained on 1841 sgRNAs targeting multiple genes (Doench *et al*. 2014). An online guide site design tool ("sgRNA designer") was created from these findings, which allows the analysis of potential guide sites for on-target efficacy.

I then analysed my guide sites, previously designed in the "Optimised CRISPR design tool", to *DMA, DMB1*, *DMB2* and *BLB2* using the "sgRNA designer". Surprisingly, I found that all of the previously designed guides received a low on-target efficacy score. Therefore, I decided to design new guide sites to genes of interest using the "sgRNA designer" and to select guides with the highest on-target efficacy scores (detailed in materials and methods).

The second cause of my initial failures was the delivery and isolation of sgRNAs in target cells. Previously, cells were transfected without selection or were co-transfected with a GFP expressing plasmid to allow GFP positive cells to be sorted. It was clear that it would be advantageous if not essential to isolate the sgRNA transfected population and I was concerned that the transfection of the GFP reporter plasmid did not represent the transfection of the sgRNA. To address this issue a different sgRNA plasmid (PX458) was sourced, that contained a GFP selection marker. Using this new plasmid I could be sure that cell sorting GFP positive cells reflected cells transfected with the sgRNA.

I decided to use a different target gene (β_2m), one for which I could easily assess the cell surface levels, to facilitate the optimisation of the CRISPR-Cas9 system. Taking into consideration both points detailed above, I decided to design guide sites using the sgRNA

designer and selected guide sites with a high on-target efficacy score. These sites were cloned into the PX458 sgRNA plasmid. Of the four sgRNAs designed to target β_2m, two were transfected into IS19 REV-cells and the cell pools were sorted for positive GFP expression. The two sorted cell lines, named IS19 b2m sgRNA1 and 2, were found to have reductions in β_2m cell surface expression. This result strongly suggests that the guide site on-target efficacy and/or the isolation of the transfected population was responsible for the previous failed attempts to generate a CRISPR-Cas9 mediated KO, although further work would be required to confirm this.

B2m KO clonal lines were then derived from the polyclonal edited populations IS19 B2m.1 and IS9 b2m.4 These lines were initially screened by flow cytometry. Figure 3.2.3 shows four clones B2m.1 cl-4, B2m.1 cl-7, B2m.4 cl-2 and B2m.4 cl-4, all of which have B2m levels overlaying with the secondary antibody control suggesting that no B2m is present. To confirm this result B2m KO clonal lines were further screened by PCR of the B2m locus, cloning and sequencing. Figure 3.2.4 shows sequencing results of the clonal lines B2m.1 cl-4 and B2m.1 cl-7, which showed that indels had been generated in the expected region (between position 16 and 18 of the PAM site). B2m.1 cl-4 was shown to have two different indels on the two different alleles, a two base pair deletion and a one base pair insertion. B2m.1 cl-7 was shown to only have one edit of a two base pair deletion. To check that both alleles of B2m were not edited 10 bacterial clones were screened from two independent PCRs, as well as direct sequencing of the PCR projects. All of which showed one clear trace sequence with a two base pair deletion which suggests that editing has occurred on both chromosomes. With the CRISPR-Cas9 system optimised in chicken lymphocytes we were then ready to begin making class II KOs to begin our study of class II antigen presentation in the chicken.

4. IS19 MHCII single KO generation and screening

4. IS19 Class II single KO generation and screening

4.1 Introduction

The main goal of this project was to begin to understand how the chicken MHC class II antigen presentation functions, and to see how similar or different this system is when compared to mouse and human class II presentation for which there is an abundance of data. The chicken MHC loci has a number of differences to the human MHC when we look at MHC class II classical and non-classical genes. Notably: the chicken only has two classical class II B genes (BLB1 and BLB2), whereas humans have three (DP, DQ and DR), the chicken has three non-classical DM class II genes DMA, DMB1 and DMB2, whereas humans have DMA and DMB, and lastly no chicken non-classical class II genes for DO have been found, whereas DO is present in humans (comprised of two genes DOA and DOB). The differences in genes found in the MHC (not to mention the differences in organisation) suggest that the way class II presentation may be achieved is different between chickens and classic model organisms. For example, the fact that two DM beta chains are found in the chicken suggests that chickens have two different DM chaperones, and as we know that DM is responsible for the ensuring the class II molecule is loaded with high affinity peptide, this suggest that chicken class II molecules could be subject to two different types of peptide editing and therefore have more possibilities for the peptide repertoire presented.

To even begin to understand how these two different DMs effect the peptide repertoire displayed on chicken MHC class II molecules and more broadly to understand the mechanisms involved in antigen presentation, we need a strategy of how to look at the effect of each component of the chicken class II presentation system. To start to answer this question, we took the approach of using CRISPR-Cas9 to KO each of the key components in the class II presentation. For this study we initially started by making KOs of the following

genes: BLB1, BLB2, DMA, DMB1 and DMB2. Detailed in the first data chapter were the initial teething problems we had in establishing CRISPR-Cas9 mediated editing and how we optimised this technique for reliable and accurate editing; with the technique working consistently in our hands we were able to start the class II KOs in earnest. We decided to start our study in the REV transformed cell line IS19, which is homozygous for the B19 haplotype.

This chapter will give full details of how the class II KO clonal lines were generated, how the cell lines were characterised; the identification of the indel at the DNA level, how this effects the translation and ultimately the effect it has on protein expression for the target gene. In addition, we touch upon initial results at the protein level that warrant more thorough investigation and describe optimisation and preparation for analysis at the RNA level by RT-qPCR, both of which are detailed in chapter 5.

4.2 Generation of MHCII specific sgRNAs

Initially two guide sites each were designed (detailed in material and methods) to the following genes: BLB1, BLB2, DMA, DMB1 and DMB2. The nomenclature for naming each guide was as follows BLB1 sgRNA1 = BLB1.1, BLB1 sgRNA2 = BLB1.2 and so on. This naming convention is used for all sgRNA guides designed. All guides designed were tested for their ability to make successful edits *in vitro* and the guides which yielded successful editing results were chosen for future work. Unfortunately, the two guides designed to target DMB1 (guides DMB1.1 and DMB1.2) showed no editing events. A further two guides were designed for DMB1 (DMB1.3 and DMB1.4); initial screening showed successful editing with DMB1.4.

The guide sgRNAs which showed successful editing events for each class II gene are shown in figure 4.2.1, these guides were selected for future experiments.

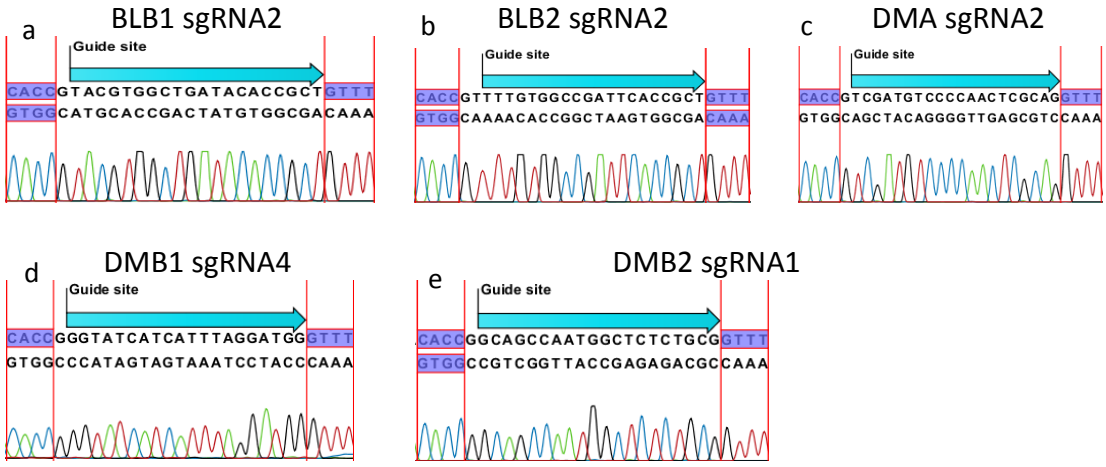


Figure 4.2.1 cloning of MHC class II specific guide sites into PX458 and PX459V2.

Guide sites cloned into PX458 were sequenced and aligned against PX458 to show that the sequence is that of the of the vector. Sequences were then checked for the presence of the expected guide site. Only guides that produced successful edits after screening are shown above. a) BLB1 sgRNA2 b) BLB2 sgRNA2 c) DMA sgRNA2 d) DMB1 sgRNA4 and e) DMB2 sgRNA1. The blue arrow shows the sequence of the cloned guide site.

4.3 Identification of indels and frameshifts in class II single KO clonal lines

IS19 cells that were transfected with guides that showed some successful editing events in the bulk population were then sorted for a single cell per well of a 96 well plate; two plates were sorted per bulk population of each gene KO (BLB1, BLB2, DMA, DMB1 and DMB2). Initially, 10 clonal lines for each gene KO line were screened by PCR, molecular cloning and sequencing. Two to five edited clones were found for each cell line. The sequencing results for each clonal line were aligned to the appropriate gene, the successful editing events are shown below.

Starting with the selected BLB1 KO clonal line BLB1.2 cl-1, two successful editing events can be seen adjacent to the PAM site at position 238-240, the edits consist of a one base pair insertion on one allele and a two base pair insertion on the second allele (figure 4.3.1a). It is important that editing events are not a multiple of three bases pairs deleted or insert, this is to ensure that the target gene is out of frame and that the edit disrupts the gene.

Next, we looked at the effect of the edit on the coding sequence when compared to the original sequence. Disruptions to the gene begin at amino acid positions 78 and 79 for the two different edits, these changes in codon frame result in an early stop codon at positions 81 and 98 terminating the protein (figure 4.3.1b).

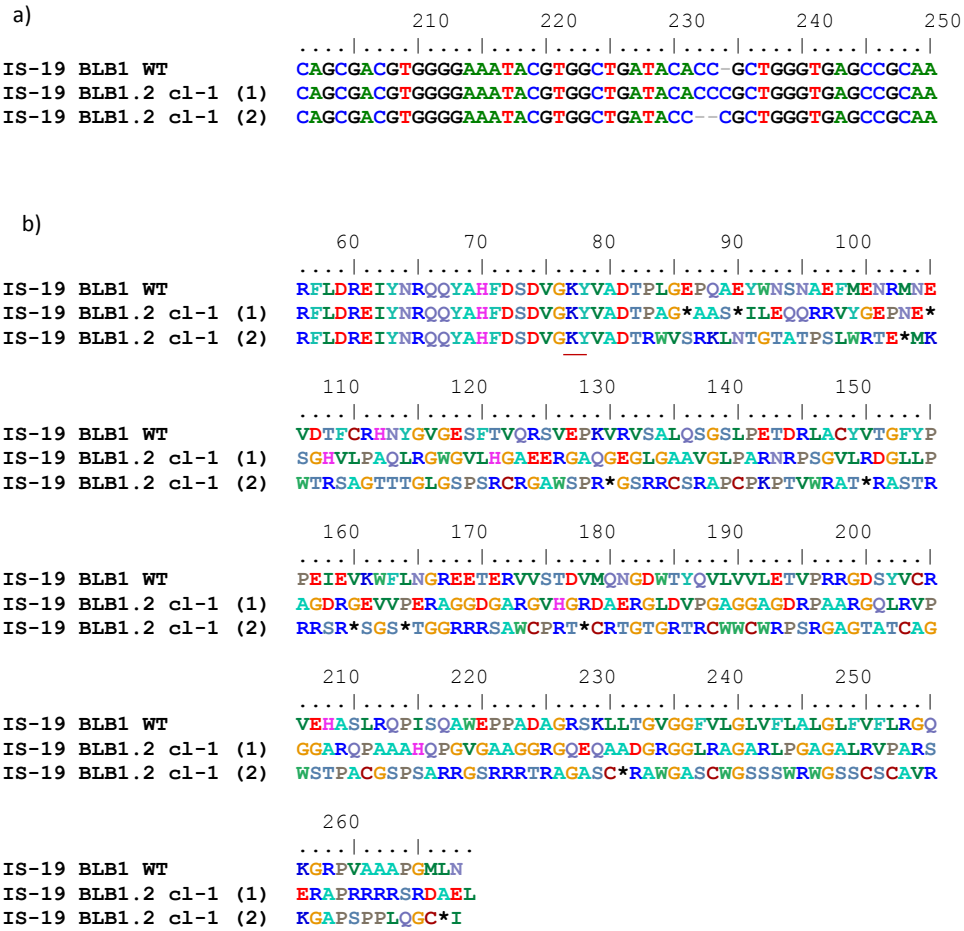


Figure 4.3.1 DNA and amino acid Sequence alignment of clonal line BLB1.2 cl-1 show editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of BLB1 were mapped onto the whole of BLB1 cDNA sequence and compared to the wild type, editing events can be seen in both alleles immediately before the PAM site (red line) at position 238-240. b) The effect of these editing events on the whole amino acid sequence were then shown for each of the two edited alleles. Changes in the coding sequence begin at position 88 and 89, premature stop codons can be seen down stream of these positions.

The three sequences are labelled on the left-hand margin; wild type sequence (IS19 BLB1 WT), the first edited allele of clonal line BLB1.2 cl-1 (IS19 BLB1.2 cl-1 (1)) and the second allele (IS19 BLB1.2 cl-1 (2)).

Due to the high DNA sequence similarities between BLB1 and BLB2 the primers to amplify BLB1 amplify both BLB1 and BLB2. This enabled us to check that we did not have off-target editing events in BLB2. From cloned PCR product a total of 20 clones were screened, two different editing events were found in BLB1 as shown above and eight clones of BLB2 showed no editing event had occurred (not shown) suggesting that our targeting was specific.

The selected clonal line for BLB2 KO, BLB2.1 cl-1 also showed two distinct successful editing events again in the expected position, slightly downstream of the PAM site at position 238-240. The edits consist of a one base pair insertion on one allele and a two base pair deletion on the second allele (figure 4.3.2a), both of which result in a frameshift. The effect of the mutation on the coding sequence showed disruptions beginning at amino acid positions 78 and 79 for the two different edits, this change in translation frame results in early stop codons at positions 80 and 81 terminating the protein (figure 4.3.2b).

A total of 20 bacterial clones were screened 11 of which showed wild type BLB1 sequence (not shown) giving confidence that off-target editing had not taken place in BLB1.

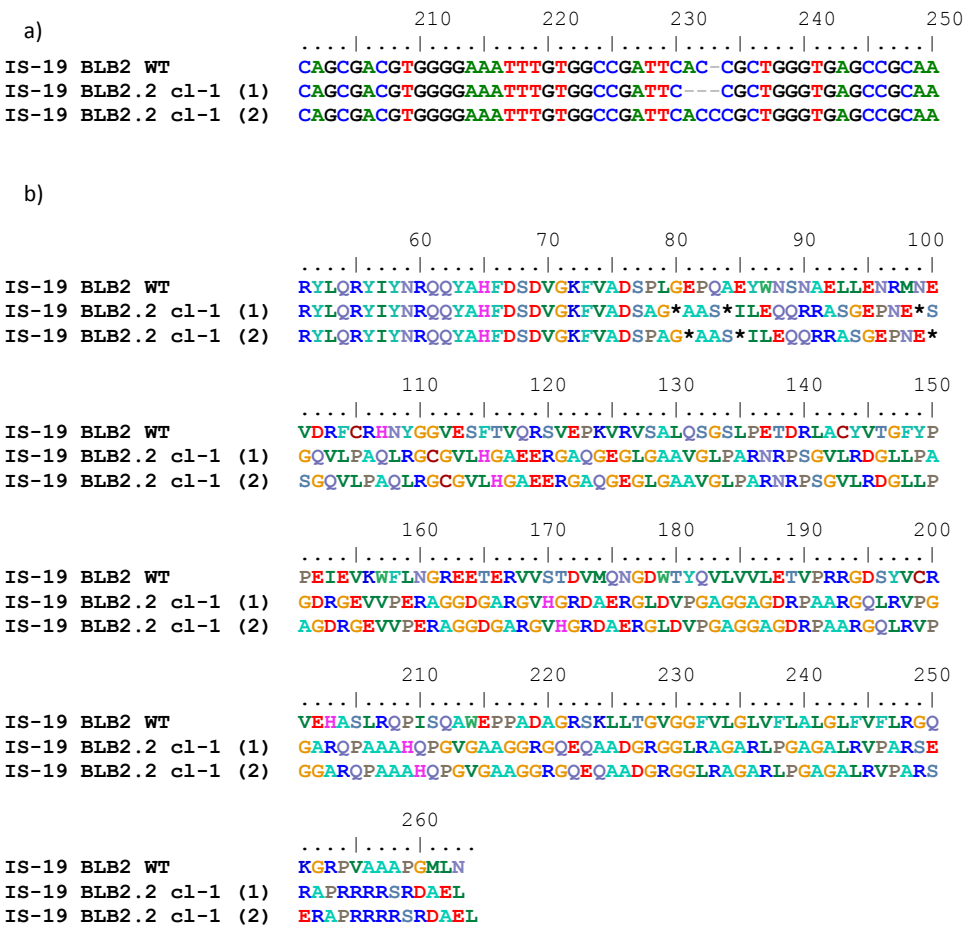


Figure 4.3.2 DNA and amino acid sequence alignment of clonal line BLB2.2 cl-1 show editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of BLB2 were mapped onto the whole of BLB1 DNA sequence and compared to the wild type, editing events can be seen in both alleles immediately before the PAM site (red line) at position 238-240. b) The effect of these editing events on the whole amino acid sequence were then shown for each of the two edited alleles. Changes in the coding sequence begin at position 78 and 79, premature stop codons can be seen down stream of these positions.

The three sequences are labelled on the left-hand margin; wild type sequence (IS19 BLB2 WT), the first edited allele of clonal line BLB2.2 cl-1 (IS19 BLB2.2 cl-1 (1)) and the second allele (IS19 BLB2.2 cl-1 (2)).

The clonal line selected for DMA KO, DMA.2 cl-1, showed only one editing event in all bacterial clones screened. The edit is a one base pair deletion immediately downstream of the PAM site located at position 193-195 (figure 4.3.3a). The effect of the mutation on the coding sequence shows disruption beginning at amino acid position 63, this change in

translation frame results in an early stop codon at position 134 terminating the protein (figure 4.3.3b).

A total of 14 bacterial clones were screened from two separate PCRs. The uncloned PCR product was also sequenced and showed one clear trace with the one base pair deletion at position 190.

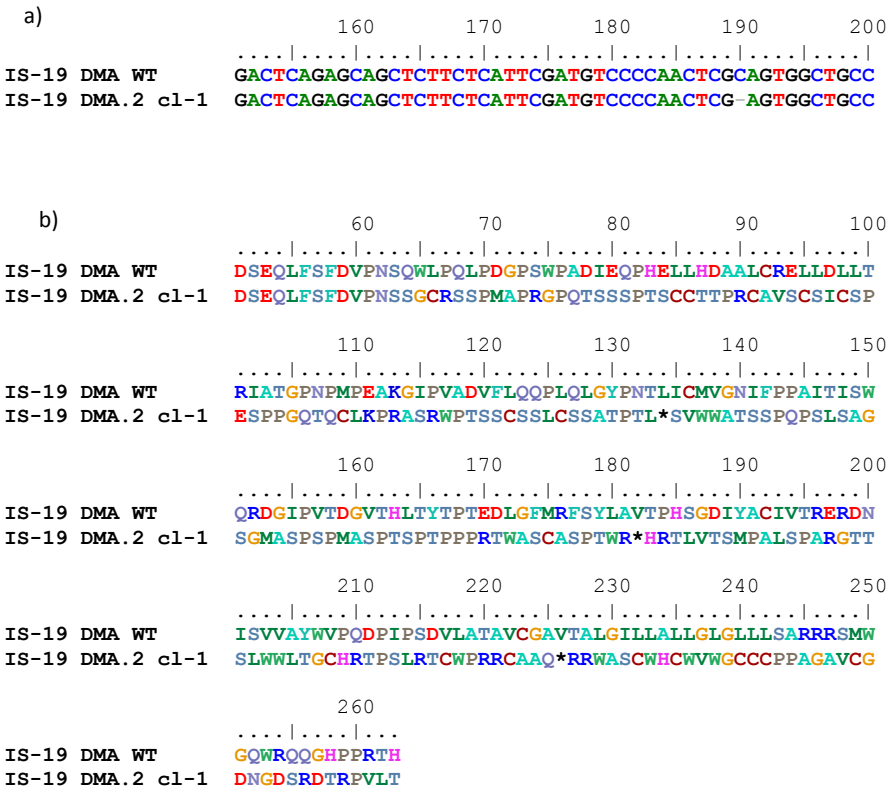


Figure 4.3.3 DNA and amino acid sequence alignment of clonal line DMA.2 cl-1 show editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of DMA were mapped onto the whole of DMA cDNA sequence and compared to the wild type, the editing event appears to be the same on both alleles immediately before the PAM site (red line) at position 193-195. B) The effect of this editing event on the whole amino acid sequence is shown. The change in the coding sequence begins at position 63, a premature stop codon can be seen down stream of this position.

The two sequences are labelled on the left-hand margin; wild type sequence (IS19 DMA WT), the edited alleles are labelled (IS19 DMA.2 cl-1).

The clonal line selected for DMB1 KO, DMB1.4 cl-10, showed only one editing event in all bacterial clones screened. The edit is a two base pair deletion immediately downstream of

the PAM site located at position 239-241 (figure 4.3.4a). The effect of the mutation on the coding sequence shows disruption beginning at amino acid position 82, this change in translation frame results in an early stop codon at position 84 terminating the protein (figure 4.3.4b).

A total of 16 bacterial clones were screened from two separate PCRs. The uncloned PCR product was also sequenced and showed one clear trace with a two base pair deletion at position 243-244.

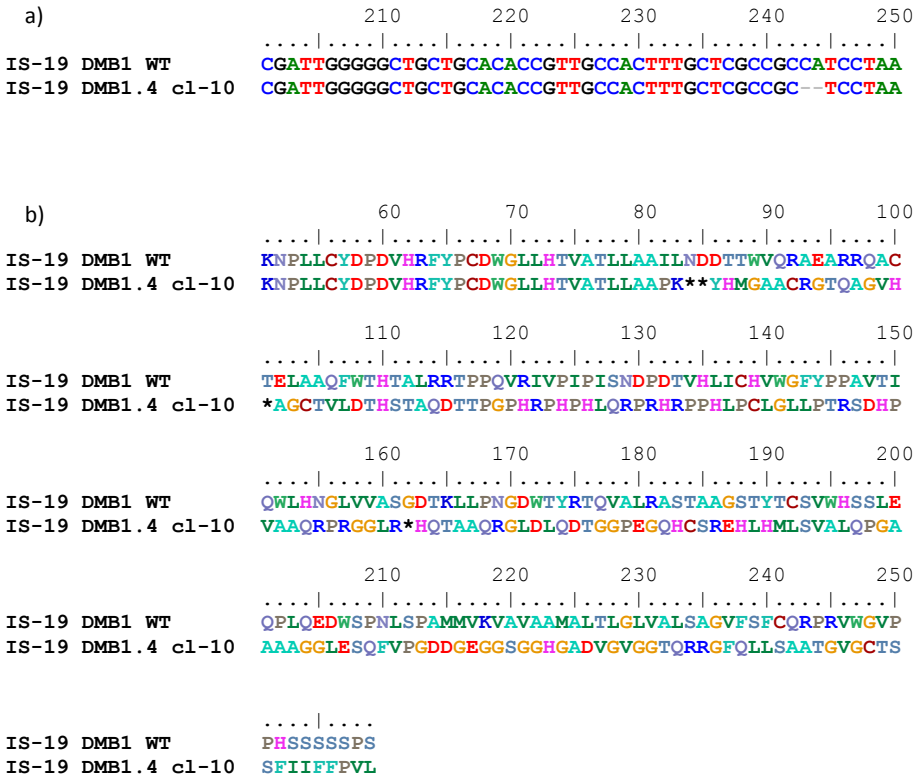


Figure 4.3.4 DNA and amino acid sequence alignment of clonal line DMB1.4 cl-10 show editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 2 of DMB1 were mapped onto the whole of DMA cDNA sequence and compared to the wild type, the editing event appears to be the same on both alleles immediately before the PAM site (red line) at position 239-241. b) The effect of this editing event on the whole amino acid sequence is shown. The change in the coding sequence begins at position 82, a premature stop codon can be seen down stream of this position.

The two sequences are labelled on the left-hand margin; wild type sequence (IS19 DMB1 WT), the edited alleles are labelled (IS19 DMB1.4 cl-10).

The final class II clonal line, selected for DMB2 KO (DMB2.1 cl-11), showed only one editing event in all bacterial clones screened. The edit is a one base pair insertion immediately downstream of the PAM site located at position 102-104 (figure 4.3.5a). The effect of the mutation on the coding sequence shows disruption beginning at amino acid position 34, this change in translation frame results in an early stop codon at position 79 terminating the protein (figure 4.3.5b).

A total of 15 bacterial clones were screened from two separate PCRs. The uncloned PCR product was also sequenced and showed one clear trace with a one base pair insertion at position 98.

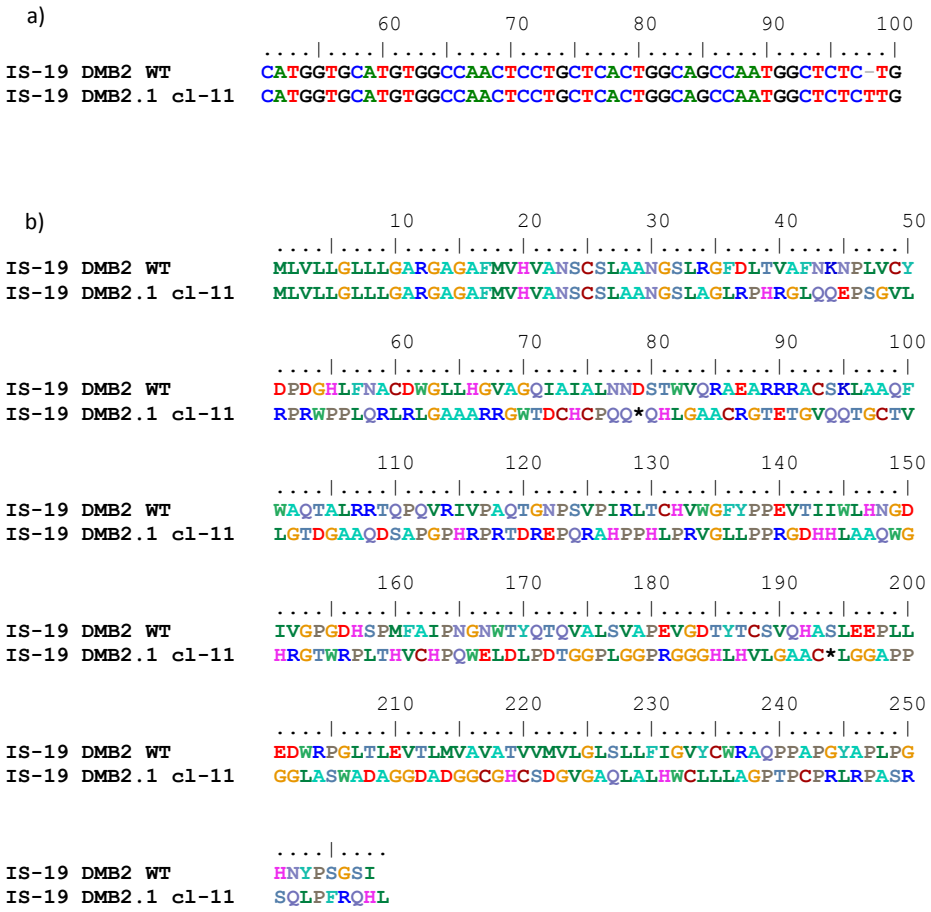


Figure 4.3.5 DNA and amino acid sequence alignment of clonal line DMB2.1 cl-11 show editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 1 of DMB2 were mapped onto the whole of DMA cDNA sequence and compared to the wild type, the editing event appears to be the same on both alleles immediately before the PAM site (red line) at position 102-104. b) The effect of this editing event on the whole amino acid sequence is shown. The change in the coding sequence begins at position 34, a premature stop codon can be seen down stream of this position.

The two sequences are labelled on the left-hand margin; wild type sequence (IS19 DMB2 WT), the edited alleles are labelled (IS19 DMB2.1 cl-11).

From the five selected clones of MHC class II gene KO lines it is noted that three of the five clonal lines appear to have the same indel event on both alleles, which initially seemed like a strange occurrence if the indel generated is thought to be largely random. Upon deeper reading of the exploding literature of CRISPR-Cas9 it seems that this may not be an uncommon phenomenon. When an allele is cut and edited and then an indel is generated, a subsequent cut to the second allele may be repaired by homology directed repair using the

first edited allele as the 'repair template'. This would result in the same indel being present on both alleles. The likelihood of this occurring depends on the cellular preference for either NHEJ or HDR as the repair method; this preference can be completely dependent on the cell type used and the stage of cell cycle when the edit occurs. The literature has many detailed examples of genes similar to the target gene being unexpectedly used as a repair template for the targeted gene, resulting in the presence of chimeric proteins, an issue that was encountered when making double KO mutations of the class II genes.

The next step was to show that our edited cell lines had no protein expression of our target genes, proving that our frameshift mutation had resulted in disruption of the gene and that our cell lines were viable for further characterisation and study.

4.4 Confirmation that indel generation in class II single KO lines results in no protein expression.

Membrane enriched lysates were prepared at a concentration of $2x10^8$ cells per mL for wild type IS19 cells and all five class II KO clonal lines to study protein expression by Western blotting. Initially, the cell lines BLB1.2 cl-1 and BLB2.2 cl-1 were examined by denaturing SDS-PAGE with samples being boiled before loading. Samples were stained with the primary antibody BLBcyt10, targeted to the conserved cytoplasmic tail of both BLB proteins enabling us to detect both BLB chains.

Figure 4.4.1 shows the wild type IS19 has two bands present, with the top band being BLB1 at 35 kDa and the bottom BLB2 at 31 kDa. BLB1.2 cl-1 shows a complete loss of the upper band (BLB1) and BLB2.2 cl-1 shows a complete loss of the lower band (BLB2), as is expected from the change in translation detailed in the sequencing results (figures 4.4.1 and 4.4.2). Unexpectedly, BLB2.2 cl-1 appeared to show reduced protein expression of BLB1 in addition to the complete loss of BLB2 despite our sequencing results showing that the BLB1 locus remained unedited; this result will be addressed further later in this chapter.

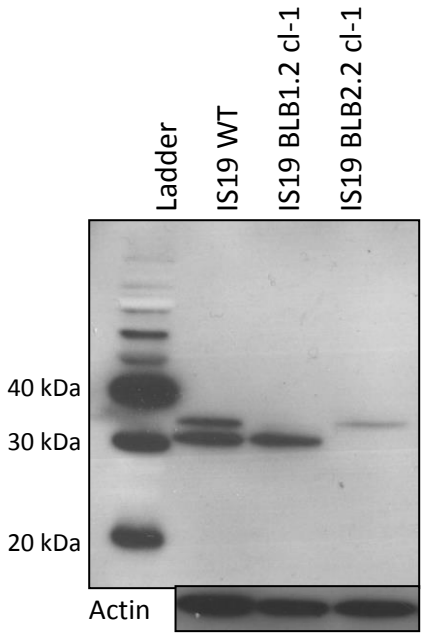


Figure 4.4.1 BLB1 and BLB2 KO clonal lines lack protein expression of the targeted BLB.

IS19 WT cells and clonal derived cell lines BLB1.2 cl-1 and BLB2.2 cl-1 were lysed using digitonin after a membrane enriching protocol. The clone lysates were probed for BLB protein expression by western blotting using the antibody BLBcyt10. Membranes were subsequently stained for β -actin levels as a loading control.

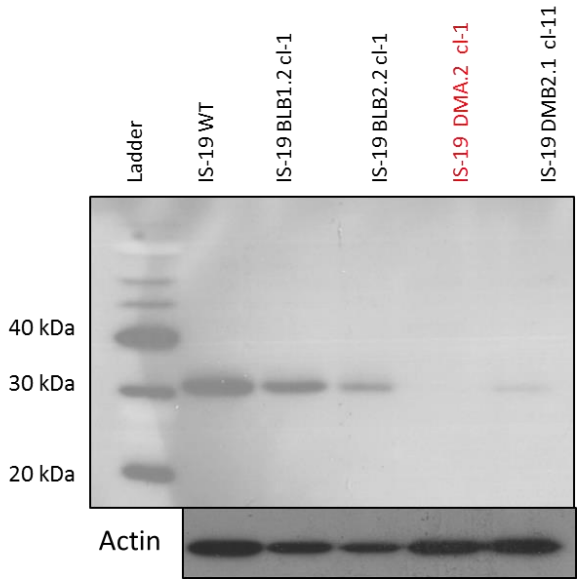


Figure 4.4.2 DMA KO clonal line has no protein expression of the targeted DMA.

IS19 WT cells and the clonal derived cell line DMA.2 cl-1 were lysed using digitonin after a membrane enriching protocol. The lysates were probed for DMA protein expression by western blotting using the antibody DMA-8. Clones were subsequently stained for β -actin levels as a loading control.

The cell line DMA.2 cl-1 was examined by denaturing SDS-PAGE with samples being boiled before loading and samples were stained with the primary antibody DMA-8; targeted to the cytoplasmic tail. Figure 4.4.2 shows wild type IS19 DMA present at a mobility of 31 kDa. IS19 DMA.2 cl-1 shows a complete loss of DMA, as expected from the change in translation detailed in the sequencing results (figure 4.3.3). In addition, BLB1, BLB2 and DMB2 KO lines were also stained for the presence of DMA. It was seen that BLB1 and BLB2 seemed to show comparable levels of DMA when compared with the actin staining, but it was noted that the DMB2 KO line had a considerable reduction in DMA expression. This was not unexpected and has been seen in human and mouse class II literature but warranted closer investigation (see chapter 5).

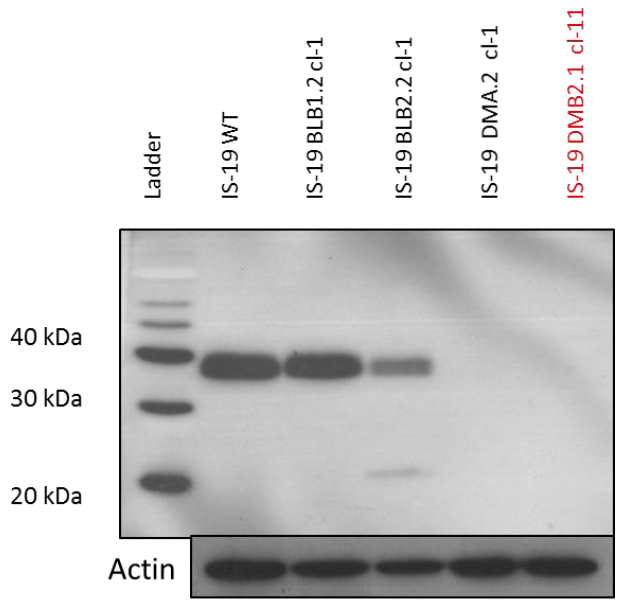


Figure 4.4.3 DMB2 KO clonal line has no protein expression of the targeted DMB2.

IS19 WT cells and the clonal derived cell line DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol. The lysates were probed for DMB2 protein expression by western blotting using the antibody DMB2-11. Clones were subsequently stained for β -actin levels as a loading control.

The cell line DMB2.1 cl-11 was examined by denaturing SDS-PAGE with samples being boiled before loading and samples were stained with the primary antibody DMB2-11, targeted to the cytoplasmic tail. Figure 4.4.3 shows the wild type IS19 has DMB2 present at an approximate mobility of 38 kDa. IS19 DMB2.1 cl-11 shows a complete loss of DMB2, as expected from the change in translation detailed in the sequencing results (figure 4.3.5).

As before, BLB1, BLB2 and DMB2 KO lines were stained for the presence of DMB2. It was seen that BLB1 KO showed comparable levels of DMB2 when compared with the actin staining. The BLB2 KO line appeared to show a reduction in DMB2 expression and the presence of a smaller band at ~20 kDa, which was unexpected and is investigated further in chapter 5. It was also seen that the DMA KO line showed the complete loss of DMB2

expression. Human and mouse class II literature has shown that the loss of DMA can result in the reduction of DMB, but it was unexpected to see a complete loss of DMB2 protein.

Next, we determined if the chicken class II protein in IS19 is resistant to SDS and can remain as a heterodimeric complex when examined by reducing PAGE and the samples are not boiled before loading onto the gel.

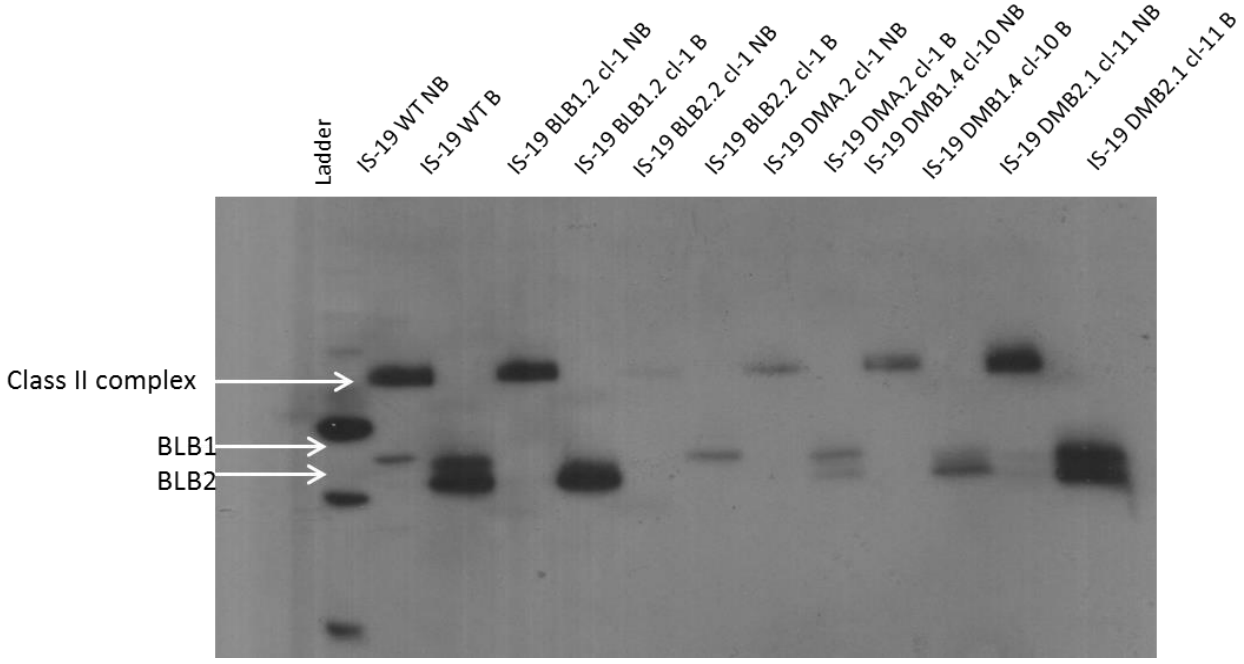


Figure 4.4.4 Assessment of SDS-sensitivity in all class II KO clonal lines.

IS19 WT cells and the clonal derived cell lines BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol. The lysates were probed for BLB protein expression by western blotting using the antibody BLBcyt10. Before loading lysates were either not boiled (NB) or boiled (B).

IS19 WT was seen to have the phenomena of SDS-resistance as described in other species, illustrated by the class II protein remaining in complex and running at ~50 kDa. Once the WT sample is boiled the complex disappears and instead is seen as the two single class II β-chains (fig 4.4.4). For the BLB1 and BLB2 KO lines we observe the same pattern as the wild type. The DMA KO line shows no sensitivity to SDS but appears to show reduced protein expression of BLB1 and BLB2. The DMB1 KO line shows heterodimeric MHCII complex and is concluded to be SDS-insensitive, but again shows reduced BLB1 and BLB2 protein

expression, in particular BLB1. The DMB2 KO line also shows heterodimeric MHCII complex and is concluded to be SDS-insensitive.

Interestingly, in these preliminary results we saw not only the loss of the gene targeted for KO but what appeared to be reductions in protein expression of other class II associated proteins. These results will be explored in greater detail in chapter 5.

4.5 Identification and optimisation of housekeeping gene standards for RT-qPCR

In the class II single KO lines generated, we began to look at the effect of protein expression of the components of class II antigen presentation, but also to characterise these lines at the peptide (chapter 6) and RNA level (chapter 5). Before we could begin to look at effects on the RNA level, suitable housekeeping genes had to be tested and optimised as well as primers for each of the class II genes of interest.

Three housekeeping genes were tested as standards for RT-qPCR. Hypoxanthine phosphoribosyltransferase (HPRT), a transferase which mainly functions to recover purines from DNA that has been degraded to introduce them back into the purine synthetic pathways. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a dehydrogenase enzyme from glycolysis. 18S, a component of the small eukaryotic ribosomal subunit (40S). The primers tested were from a publication optimising qRT-primers in chicken tissue (Olias et al 2014).

Primers were tested for HPRT with successful results. Average Cts ranged from ~31 to ~27 (figure 4.5.1a) and producing a standard curve that has an R^2 value of 0.993 indicating a high efficiency of amplification by the primers (figure 4.5.1b). Importantly, the primers also showed a 'clean' melt curve with only one clear peak for all replicates of all cDNA

concentrations (figure 4.5.1c).

R^2= 0.993

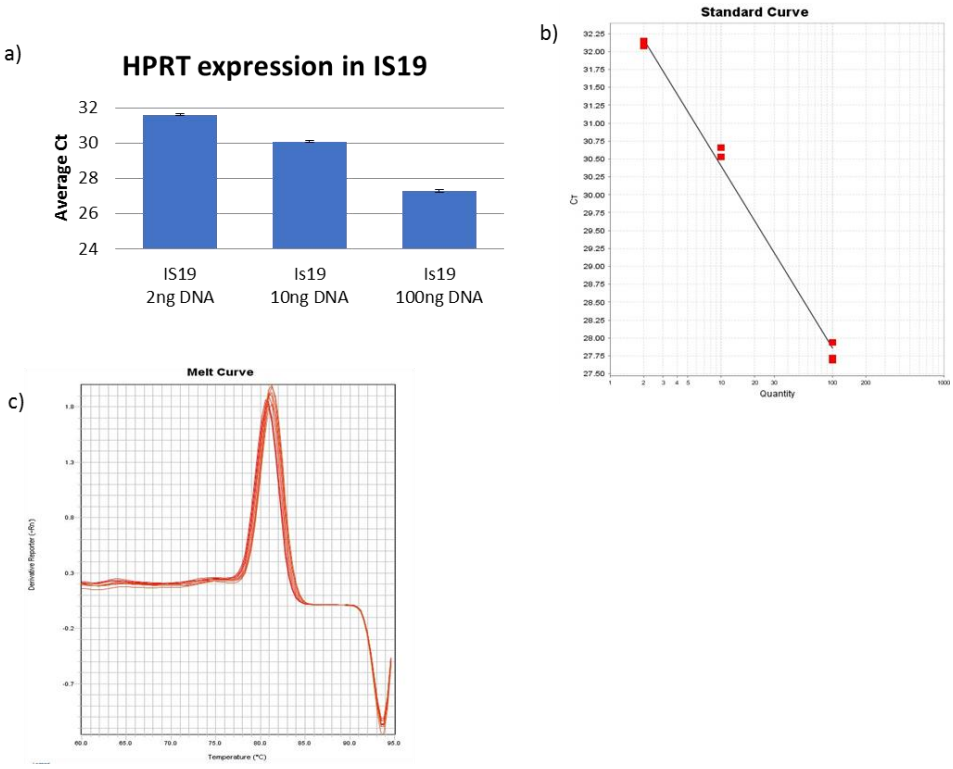


Figure 4.5.1 optimisation of housekeeping HPRT showing the average Ct, standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the housekeeping gene HPRT, 3 technical replicates were used for each concentration. a) the average Ct was plotted as a bar graph for each concentration. b) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.993 (replicates are shown as red boxes). c) the melt curve of each replicate at each cDNA concentration.

Next GAPDH was tested, giving average Ct values from ~24 to ~17 (figure 4.5.2a) and producing a standard curve that has an R^2 value of 0.998 indicating a high efficiency of amplification by the primers (figure 4.5.2b). The melt curve showed one clear peak for all replicates of all cDNA concentrations (figure 4.5.2c), showing specific primers that amplified one product.

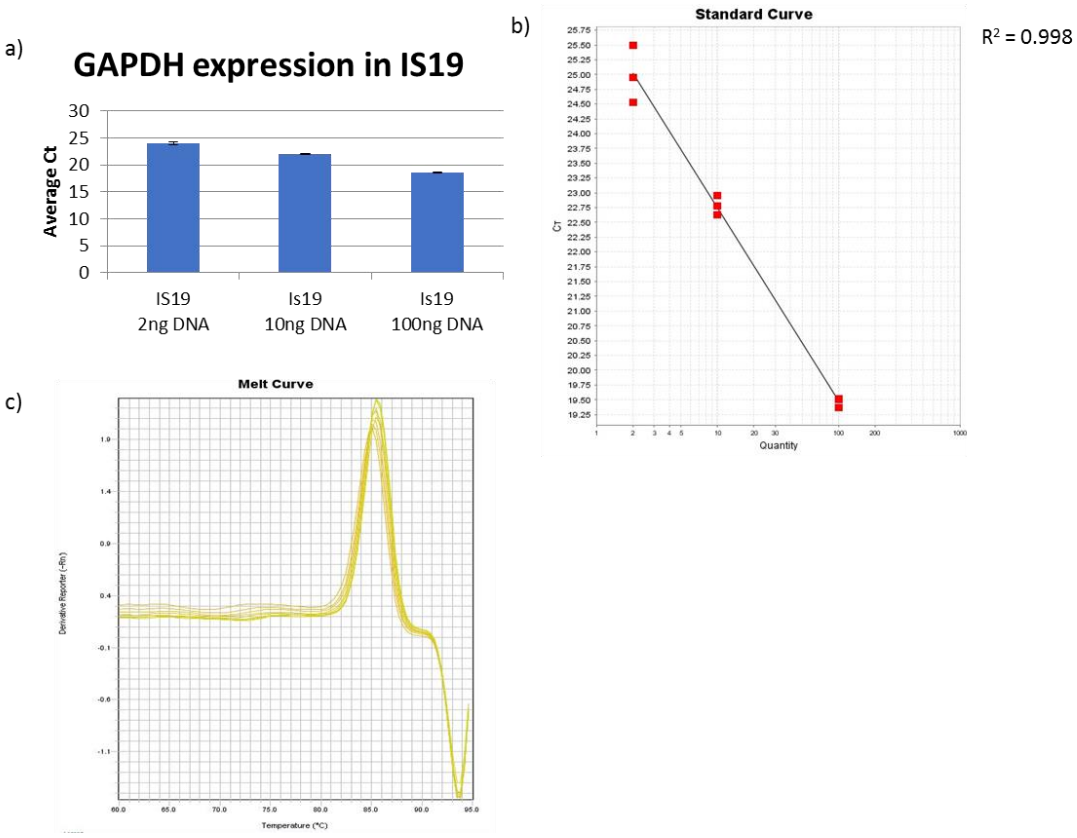


Figure 4.5.2 optimisation of housekeeping GAPDH showing the average Ct, standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the housekeeping gene GAPDH, 3 technical replicates were used for each concentration. a) the average Ct was plotted as a bar graph for each concentration. b) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.998 (replicates are shown as red boxes). c) the melt curve of each replicate at each cDNA concentration.

The last set of primers tested were for 18S. Average Ct values were seen from ~21 to ~17 (figure 4.5.3a) and produced a standard curve with an R^2 value of 0.919 indicating a high efficiency of amplification by the primers (figure 4.5.3b). The melt curve showed one clear peak for all replicates of all cDNA concentrations (figure 4.5.2c), showing specific primers that amplified one product.

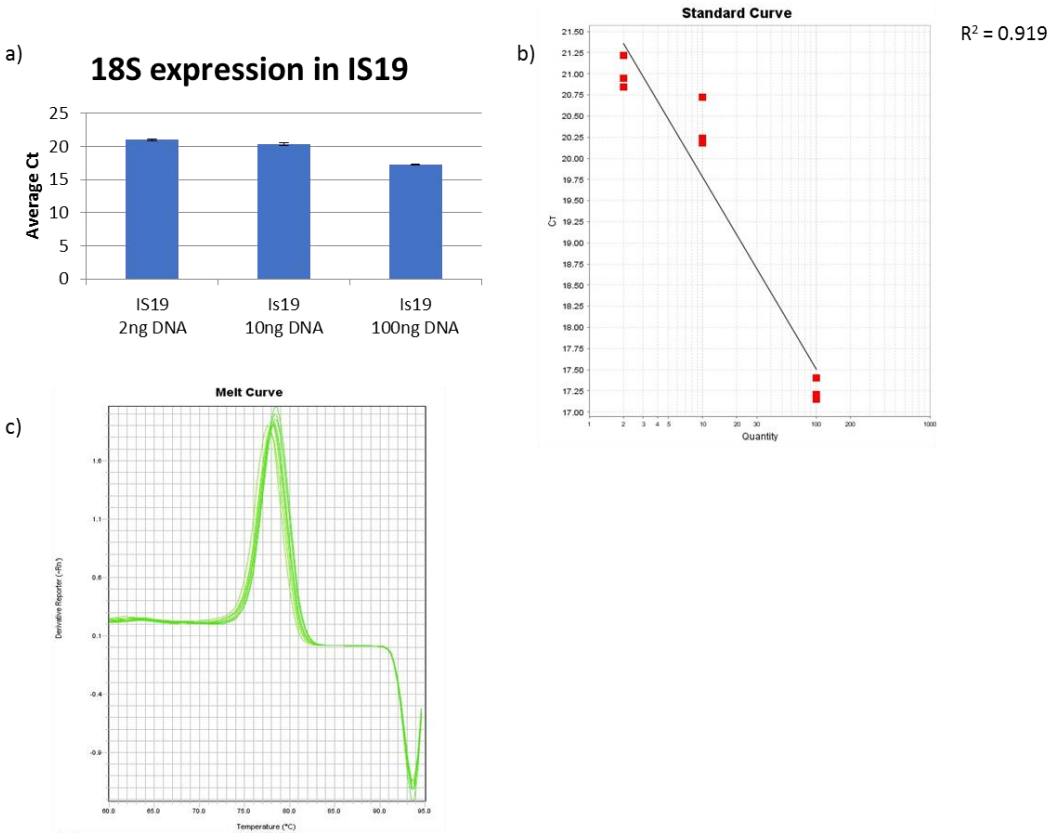


Figure 4.5.3 optimisation of housekeeping 18S showing the average Ct, standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the housekeeping gene 18S, 3 technical replicates were used for each concentration. a) the average Ct was plotted as a bar graph for each concentration. b) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.919 (replicates are shown as red boxes). c) the melt curve of each replicate at each cDNA concentration.

4.6 Identification and optimisation class II primers for RT-qPCR

With appropriate housekeeping genes identified and optimised next appropriate primers were tested for each of the class II genes of interest; BLB1, BLB2, DMA, DMB1 and DMB2. BLB1 proved to be difficult to design primers that would amplify with a high efficiency (R^2 above 0.8). The best primer combination had an R^2 value of 0.883 and a melt curve that showed the presence of one major peak indicating largely specific amplification (figure 4.6.1).

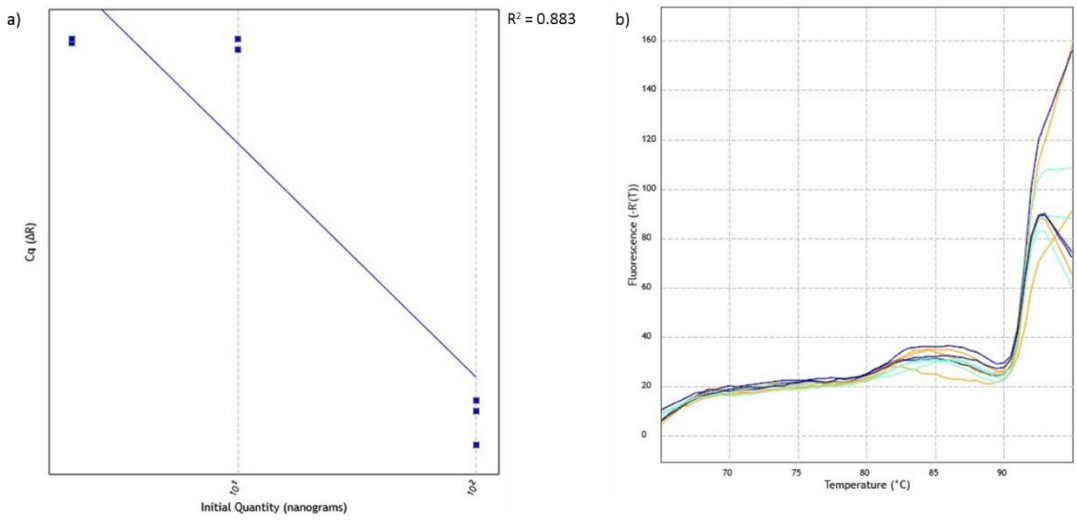


Figure 4.6.1 optimisation of qRT-PCR primers for BLB1 showing the standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the BLB1; 3 technical replicates were used for each concentration. a) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.883 (replicates are shown as blue boxes). b) the melt curve of each replicate at each cDNA concentration.

The best set of primers for BLB2 gave an R^2 value of 0.995 and a melt curve with one peak shown in figure 4.6.2.

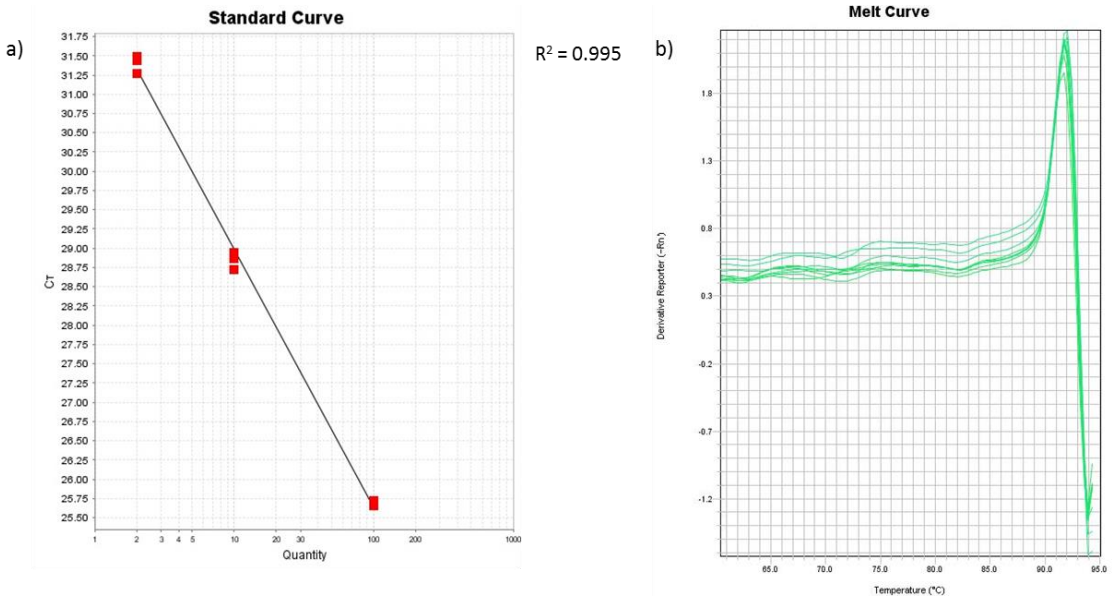


Figure 4.6.2 optimisation of qRT-PCR primers for BLB2 showing the standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the BLB2, 3 technical replicates were used for each concentration. a) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.995 (replicates are shown as red boxes). b) the melt curve of each replicate at each cDNA concentration.

The best set of primers for DMA gave an R^2 value of 0.986 and a melt curve with one peak shown in figure 4.6.3.

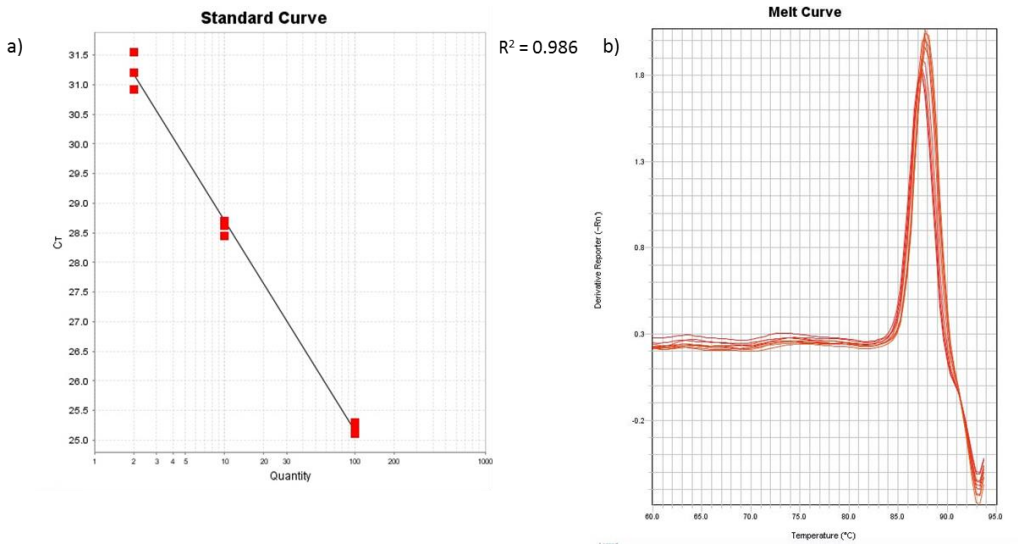


Figure 4.6.3 optimisation of qRT-PCR primers for DMA showing the standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the DMA; 3 technical replicates were used for each concentration. a) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.986 (replicates are shown as red boxes). b) the melt curve of each replicate at each cDNA concentration.

Primers designed to DMB1 gave an R^2 value of 0.837 and a melt curve with one peak, shown in figure 4.6.4.

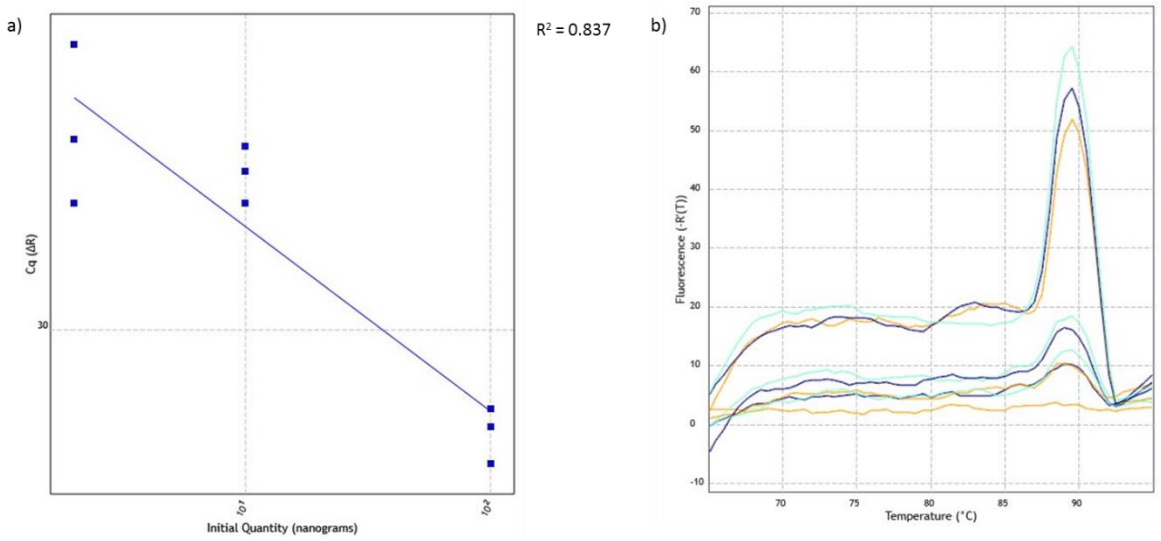


Figure 4.6.4 optimisation of qRT-PCR primers for DMB1 showing the standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the DMB1, 3 technical replicates were used for each concentration. a) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.837 (replicates are shown as blue boxes). b) the melt curve of each replicate at each cDNA concentration.

The best set of primers for DMB2 gave an R^2 value of 0.994 and a melt curve with one peak shown in figure 4.6.5.

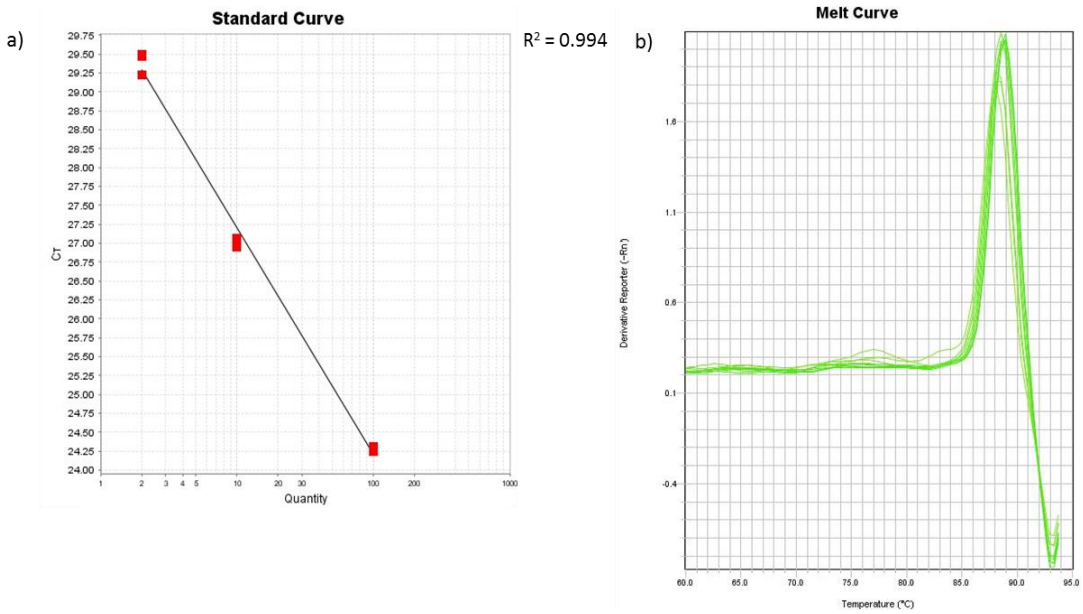


Figure 4.6.5 optimisation of qRT-PCR primers for DMB2 showing the standard curve and melt curve.

RNA was extracted from IS19 WT cells, followed by cDNA synthesis. 2, 10 or 100 ng of cDNA was then used for qRT-PCR analysis to amplify the DMB2; 3 technical replicates were used for each concentration. a) the standard curve was plotted from 3 replicates of each concentration generating an R^2 value of 0.994 (replicates are shown as blue boxes). b) the melt curve of each replicate at each cDNA concentration.

4.7 Discussion

Successful indel events were generated in the IS19 cell type for BLB1, BLB2, DMA, DMB1 and DMB2. The screening of the class II KO cell lines showed a clear predictability for the position of the indel generated in relation to the guide RNA used, with indels occurring four nucleotides downstream of the PAM site. Analysis of these clonally generated KO lines showed that the indel event of each chosen line resulted in a frameshift event that disrupted the coding of the gene and in each case led to the generation of a premature stop codon. Targeting of BLB1 and BLB2 was initially a concern as there is high sequence homology between the two genes. To try and circumvent off-target effects guides were designed with a minimum of 3 bp mismatches between the targeted BLB and its counterpart, based on previous publications on mismatch tolerance in CRISPR-Cas9 mediated gene editing (Hsu et al 2013). This approach appeared to yield no detectable off-target effects when both BLB loci were screened for edits in each BLB KO line.

One aspect of editing which was unexpected was that in multiple cases (DMA fig 4.3.3, DMB1 fig 4.3.4 and DMB2 4.3.5) the same indel was detected on both alleles, which seemed counter intuitive as the indel generation is assumed to be random. Upon further reading and through personal communication I found that this does not seem to be uncommon and a current hypothesis to explain this is that during the cutting and editing process one allele is edited first, subsequently a second allele is cut and instead of the second cut being repaired by NHEJ, it is repaired by HDR using the first allele as a template. After the HDR event has taken place this will then lead to the same indel on both alleles.

To confirm that the indel events have disrupted the protein expression and that in fact both alleles have been edited (in the case where the indels are the same on both alleles) western blotting was performed to check that no protein was present for the targeted gene. This

proved to be successful with the clonal lines having no protein for the targeted gene, (except for DMB1, where protein expression could not be measure) although we discovered some interesting effects on other class II protein expression in our clonal KO lines.

Firstly, we found that in the BLB1 KO line there is no effect on BLB2 expression as expected, but in the BLB2 KO line BLB1 expression is reduced (fig 4.4.1). It was also seen in the DMB2 KO line that DMA expression is reduced (fig 4.4.2). This was not unexpected, as we assume DMA and DMB2 to function as an α-β pair and in the absence of one of the two chains, one could anticipate that the remaining chain is targeted for degradation as it is lacking function in the absence of its cognate chain, similar results have been seen for DM KOs in human and mice cell lines. Interestingly, in the DMA KO line DMB2 expression disappears entirely (fig 4.4.3), which is previously unseen in publications investigating class II antigen presentation in other model species. It could be assumed that because DMA is the common alpha chain for both DMB1 and DMB2, if DMA is removed then DMB2 is completely degraded without its functional partner, where as if DMB2 is removed DMA expression is reduced as there is surplus DMA to requirement, but some DMA remains present to pair with DMB1. It also needs to be considered that the chicken MHC class II differs in genes present and organisation when compared to other model systems. It appears to have two functional DM chaperones, comprised of DMA paired with either beta chain, DMB1 or DMB2 (Parker 2012) suggesting that the system may function differently from other models studied.

A more unexpected result was seen in figure 4.4.3 where the BLB2 KO line appears to show a reduction in DMB2 expression. Current models we have put forward (previously discussed chapter 1, 1.7.5) suggest a handful of ways in which the two DMs could act on the two classical class IIs, one of which proposes that each DM could have a partial or full preference for one of the class II molecules. If this were to be the case and DMB2 functions primarily

with BLB2, then if BLB2 is removed DMB2 expression would reduce. We also saw that when DMA is deleted both BLB1 and BLB2 expression is reduced (fig 4.4.4) and when DMB1 is deleted both BLB1 and BLB2 expression is decreased but with a more pronounced effect on BLB1. Previous studies in humans have shown that deletion of DMA or DMB can result in a reduction on total class II protein expression but the effects are not as pronounced as the initial results seen in the chicken. The findings that DMA deletion shows a large reduction of both BLB1 and BLB2, whereas DMB1 deletion has a more prominent effect on BLB1 expression than BLB2 could also support the idea of preferential editing of a particular BLB for each DMB.

Upon initial investigation of SDS-sensitivity, we found that like previous published data for human and mice class II alleles the B19 class II molecules are able to stay in complex when run by reducing PAGE if the samples are not boiled before loading (fig 4.4.4). B19 class II molecules did not show obvious DM-dependence (fig 4.4.4). When DMA, DMB1 or DMB2 was deleted, the class II molecules were able to stay in complex if the samples were not boiled. Published works for humans and mice have shown that DM-dependence varies between alleles, with some alleles remaining in complex without DM in non-boiled conditions whilst others cannot. It is more than possible that the B19 haplotype is a DM-independent class II allele and we may see DM-dependent alleles if we were to investigate different haplotypes.

It is important to consider that, although all western blots shown in this chapter were completed three or more times with each showing the same result, the comparisons of class II protein expression are on X-ray film which has a very low dynamic range and is inaccurate for quantification and could conceivably be misleading. In chapter 5 the class II KO lines are fully characterised for effects on class II protein expression and to address the issue with X-

ray film, all subsequent experiments for quantitative western blotting were carried out using an infrared antibody system or chemiluminescence directly from the membrane.

Lastly, in preparation for analysis at the RNA level primers were tested for candidate housekeeping genes and for the class II genes of interest. For housekeeping standards HPRT, GAPDH and 18S were screened, all of which gave high efficiency amplification and melt curves with a single peak (figures 4.5.1 - 3). GAPDH was selected for future experiments as it gave the largest range of Ct values from 21 – 17 (fig 4.5.2). Next primers were screened for all class II genes of interest BLB1, BLB2, DMA, DMB1 and DMB2. BLB2, DMA and DMB2 gave very high efficiency amplification and melt curves with a single peak (figures 4.6.2, 4.6.3 and 4.6.5). BLB1 and DMB1 proved to be more difficult and required more primers to be designed and tested (detailed in materials and methods). Primers were found with acceptable amplification efficiencies, with R values of 0.883 and 0.837 for BLB1 and DMB1 respectively and melt curves showing one major peak (figures 4.6.1 and 4.6.4). With a housekeeping gene optimised and selected, and with primers that specifically amplify the targeted class II genes of interest, quantification of gene expression can begin in the class II single KO lines.

5. Characterisation of class II single KOs at the RNA and protein level

5. Characterisation of class II single KOs at the RNA and protein level

5.1 Introduction

Chapter four detailed the creation, screening, and validation of five IS19 class II KO cell lines; BLB1, BLB2, DMA, DMB1 and DMB2. The cell lines chosen for further experimentation were BLB1.1 cl-1, BLB2.1 cl-1, DMA.2. cl-1, DMB1.4 cl-10 and DMB2.1 cl-11; for simplicity from now on each cell line will be referred to as the relevant gene name plus KO e.g. BLB1 KO. Additionally in chapter four, RT-qPCR primers and conditions were optimised to allow characterisation of the KO cell lines at the mRNA level. In this chapter we analyse the expression of all class II genes in the IS19 class II KO cell lines at the RNA level and the protein level and look at the effect of DM KOs on the temperature stability of BLB. Furthermore, initial data on TG21 BLB temperature stability is assessed in DMA and DMB2 KO lines and compared to that of IS19.

5.2 mRNA expression of class II genes, in IS19 single class II KO lines

Initially we looked at RNA expression in all the single class II KOs, with RT-qPCR data being analysed by the delta delta Ct (2–ΔΔ*Ct*) method and plotted as either a fold change compared to wild type expression. Due to sensitivity of the RT-qPCR machine, it is assumed that changes less than one cycle (two-fold change) cannot be detected reliably. Therefore any change two-fold or above is considered as an appreciable change in mRNA expression. Furthermore, the fold changes between class II gene expression and the wildtype expression were analysed by unpaired t-test with significance denoted as follows; * P <0.05, ** P <0.01, *** P <0.001.

Figure 5.2.1 shows *BLA* RNA expression fold changes. No significant upregulation is seen in any of the class II KO lines but it is observed that BLA undergoes a significant downregulation of ~15 fold in the DMB2 KO line.

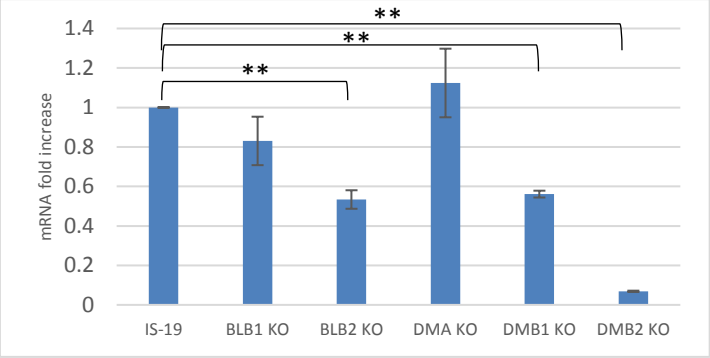


Figure 5.2.1 Fold change of BLA RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for RNA expression; expression was normalised to wild type IS19 GAPDH expression. mRNA was extracted for all cell lines, followed by cDNA synthesis, and BLA expression was measured by RT-qPCR using the UC29 (FP) and UC30 (RP) primers. Relative fold changes were generated using the 2–ΔΔ*Ct* method. a) BLA expression as relative fold increases compared to IS19 wild type cells. b) BLA expression as relative fold decreases compared to IS19 wild type cells.

BLB1 RNA expression is measured in figure 5.2.2 and shows that there is no significant upregulation at the RNA level in any of the class II KOs, though the BLB2 KO line shows a

modest ~ 3-fold decrease in BLB1 expression. It is noted that BLB1 RNA expression is comparable to wild type in the BLB1 KO line, but due to the nature of CRISPR-Cas9 indel loss of function mutants it is common for the edited gene to still be transcribed into mRNA (if the RT-qPCR primer binding site is unaffected) even though the indel results in a loss of functional protein.

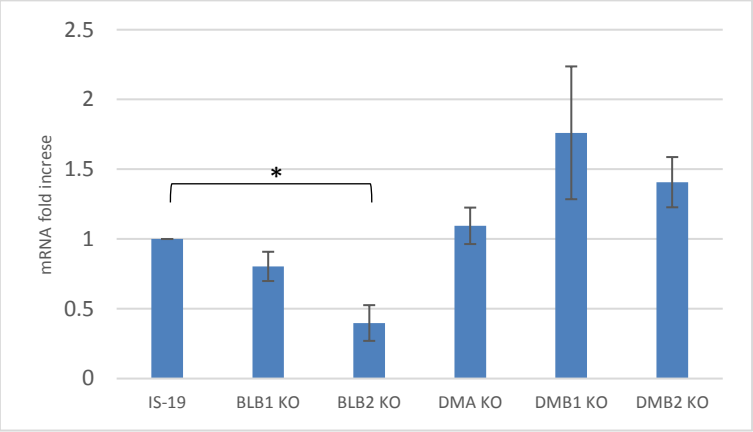


Figure 5.2.2 Fold change of BLB1 RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for RNA expression, expression was normalised to wild type IS19 GAPDH expression. mRNA was extracted for all cell lines, followed by cDNA synthesis, and BLB1 expression was measured by RT-qPCR using the qPCR BLB1 FP1 and qPCR BLB1 RP1 primers. Relative fold changes were generated using the $2^{-\Delta\Delta Ct}$ method. a) BLB1 expression as relative fold increases compared to IS19 wild type cells. b) BLB1 expression as relative fold decreases compared to IS19 wild type cells.

Figure 5.2.3 shows changes in RNA expression of BLB2 in the class II KO lines. A small upregulation of BLB2 can be seen in the BLB1 KO but no significant upregulation can be seen

in any other line. A small but significant downregulation is seen in the DMB2 KO line.

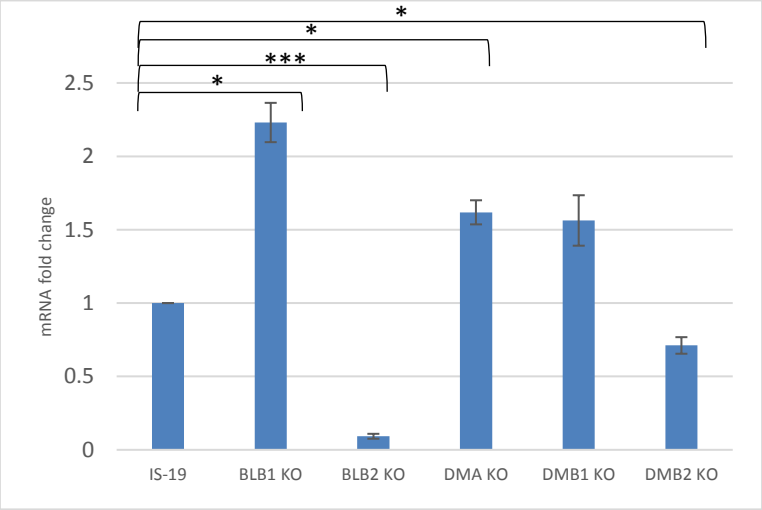


Figure 5.2.3 Fold change of BLB2 RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for mRNA expression, expression was normalised to wild type IS19 GAPDH expression. RNA was extracted for all cell lines, followed by cDNA synthesis, and BLB2 expression was measured by RT-qPCR using the UC116 (FP) and UC118 (RP) primers. Relative fold changes were generated using the $2–\Delta\Delta Ct$ method. a) BLB2 expression as relative fold increases compared to IS19 wild type cells. b) BLB2 expression as relative fold decreases compared to IS19 wild type cells.

Figure 5.2.4 shows the RNA expression of DMA, with no significant upregulation seen in any of the cell lines. Interestingly, DMA RNA is significantly down regulated in the DMB1 KO line, with a ~12 and ~4-fold reduction seen respectively.

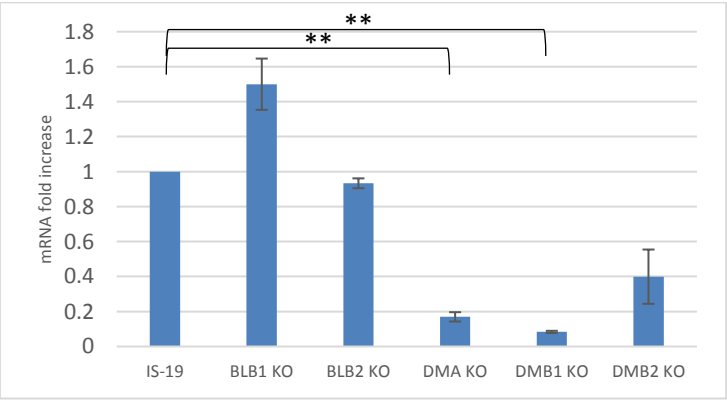


Figure 5.2.4 Fold change of DMA RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for mRNA expression, expression was normalised to wild type IS19 GAPDH expression. RNA was extracted for all cell lines, followed by cDNA synthesis, and DMA expression was measured by RT-qPCR using the UC116 (FP) and UC118 (RP) primers. Relative fold changes were generated using the $2–\Delta\Delta Ct$ method. a) DMA expression as relative fold increases compared to IS19 wild type cells. b) DMA expression as relative fold decreases compared to IS19 wild type cells.

DMB1 RNA expression was analysed in figure 5.2.5 with no significant fold increase being seen in any of the KO lines but decreases in expression were seen in the BLB2, DMA and DMB2 KO lines with reductions of ~40, ~50 and ~25 fold respectively.

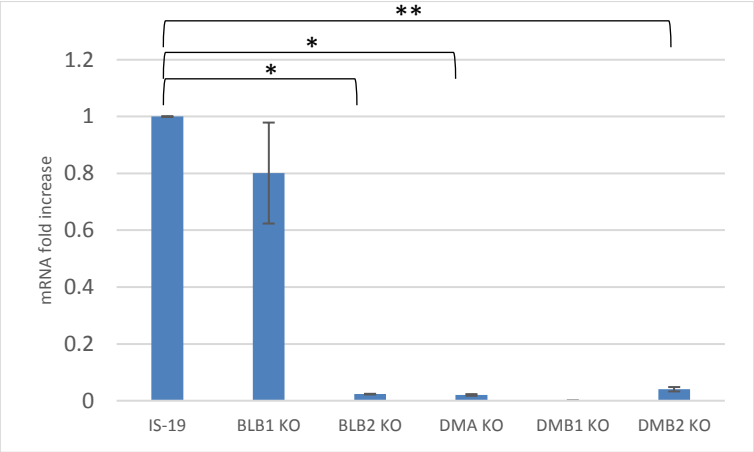


Figure 5.2.5 Fold change of DMB1 RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for mRNA expression, expression was normalised to wild type IS19 GAPDH expression. RNA was extracted for all cell lines, followed by cDNA synthesis, and DMB1 expression was measured by RT-qPCR using the qPCR DMB1 FP1.2 and qPCR DMB1 RP1.2 primers. Relative fold changes were generated using the $2^{-\Delta\Delta Ct}$ method. a) DMB1 expression as relative fold increases compared to IS19 wild type cells. b) DMB1 expression as relative fold decreases compared to IS19 wild type cells.

DMB2 RNA expression remains consistent in all KO lines with no significant increases or decreases seen in any of the KO cell lines (fig 5.2.6).

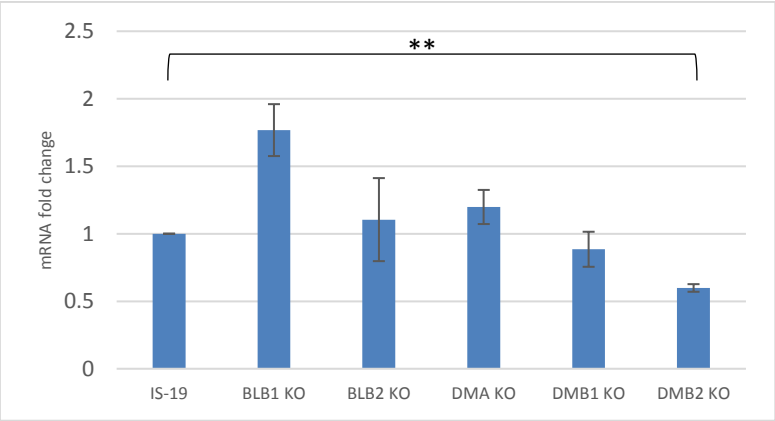


Figure 5.2.6 Fold change of DMB2 RNA of single class II KO cell lines by RT-qPCR

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were measured for mRNA expression, expression was normalised to wild type IS19 GAPDH expression. RNA was extracted for all cell lines, followed by cDNA synthesis, and DMB2 expression was measured by RT-qPCR using the qPCR DMB1 FP1.2 and qPCR DMB1 RP1.2 primers. Relative fold changes were generated using the $2^{-\Delta\Delta Ct}$ method. a) DMB2 expression as relative fold increases compared to IS19 wild type cells. b) DMB2 expression as relative fold decreases compared to IS19 wild type cells.

5.3 Quantification of class II protein expression, in IS19 single class II KO lines

Next the IS19 class II KO lines were characterised for protein expression, starting with flow cytometry analysis of BLB1 and BLB2 surface expression and total expression (intracellular staining of permeabilised cells). Figure 5.3.1a shows the cell surface levels of MHCII, with the BLB1 KO showing a ~55 % reduction in MHCII and the BLB2 KO showing a ~75 % reduction with the remaining expression presumably being BLB2 and BLB1, respectively. The DMA KO line shows a reduction of cell surface MHC II with a ~75 % reduction, whereas the DMB1 KO and the DMB2 KO lines have the same MHCII levels as the wildtype (WT). Analysis of total MHC II levels (permeabilised cells) (fig 5.3.1b) show a different pattern, with a ~25 % reduction of MHC II in the BLB1 KO and a ~90 % reduction in the BLB2 KO, indicating that proportionally more BLB2 compared to BLB1 remains trapped inside the cell than makes it to

the cell surface and vice versa for BLB1. For the DMA KO a ~50 % reduction is seen intracellularly compared to a 75 % reduction on the cell surface. In contrast, both the DMB1 KO and DMB2 KO lines show a ~25 % reduction in total MHCII.

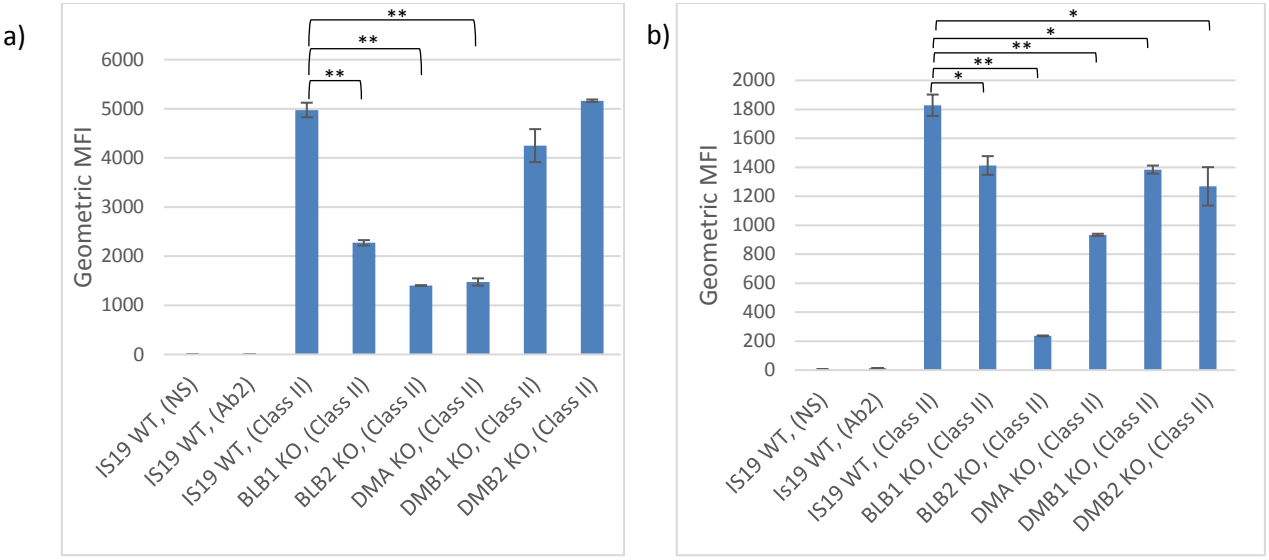


Figure 5.3.1 MHCII staining in IS19 KO cell lines by flow cytometry.

MHCII KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were analysed for MHCII expression by flow cytometry and compared to the wild type. a) Non-permeabilised samples were assessed for cell surface expression of MHCII using the 2G11 antibody. b) Samples were fixed and permeabilised for staining of intracellular MHCII using the BLBcyt10 antibody. Fluorescence is plotted as the geometric mean fluorescence intensity and error bars represent standard error of the mean.

BLB1 and BLB2 protein expression was also measured by quantative-Western blot (q-WB) using densitometry. Figure 5.3.2a shows a western blot that was used for quantification, where all class II KO cell lines were probed for BLB1 and BLB2 expression using the antibody BLBcyt10. The β-actin band can be seen at ~45 kDa, with BLB1 seen at ~35 kDa and BLB2 at 32 kDa when compared to the molecular mass markers. Quantification of BLB1 and BLB2 normalised to wild type level protein expression can be seen in figure 5.3.2.b and c, respectively. For BLB1 expression, BLB1 is completely absent in the BLB1 KO, a ~60 % reduction is seen in the BLB2 KO, a ~70 % reduction in the DMA KO, a ~90 % reduction in the DMB1 KO and a ~60 % reduction in the DMB2 KO. It is expected that DM KOs may result in

less BLB1, if BLB1 is dependent on either of the DM complexes for peptide editing. The absence of DM would be expected to result in a lower stability of BLB1 and therefore degradation. BLB1 protein expression is significantly reduced in the BLB2 KO cell line, suggesting that BLB1 protein expression or stability is linked to the presence of BLB2.

For BLB2 quantification in the class II KO lines, a ~25 % reduction is seen in the BLB1 KO, BLB2 is completely absent in the BLB2 KO, an ~80 % reduction is seen in the DMA KO line, a ~60 % reduction is seen in the DMB1 KO line and a ~75 % reduction in the DMB2 KO line. As mentioned above, it is expected that BLB2 expression is reduced in DM KO lines if BLB2 is dependent on DM editing. BLB2 expression is decreased in the BLB1 KO line, which again suggests that BLB2 expression or stability is linked to BLB1 expression.

It is interesting to note that the DMB1 KO has a slightly more pronounced effect on BLB1 compared to BLB2 and that the DMB2 KO has a slightly more pronounced effect on BLB2 when compared to BLB1.

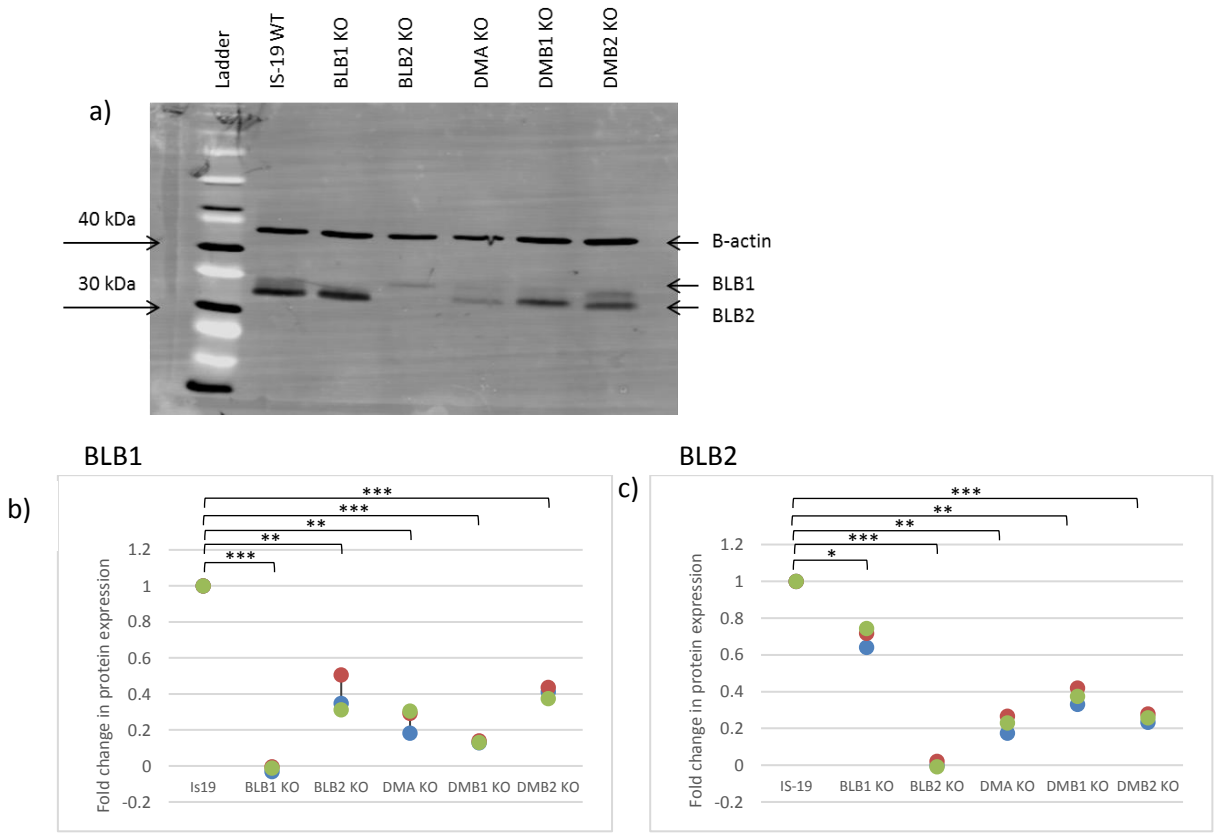


Figure 5.3.2 q-Western blots staining for BLB in IS19 single class II KO clonal lines.

IS19 WT cells and the clonal derived cell lines BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol, lysates were boiled in DTT before loading. a) The lysates were probed for BLB protein expression by western blotting using the antibody BLBcyt10. Relative BLB1 (b) and BLB2 (c) expression was quantified by densitometry, results are plotted as a relative fold change compared to wild type expression, normalised to β-actin staining. Blots were performed in triplicate for quantification.

Next DMA protein expression was assessed; Figure 5.3.3a shows a western blot that was used for quantification. DMA is seen at ~31k Da when compared to the molecular mass marker. Quantification of DMA is normalised to wild type level protein expression in figure 5.3.3.b. DMA expression is reduced by ~70 % in the BLB1 KO, a ~95 % reduction is seen in the BLB2 KO, DMA is completely absent in the DMA KO and a ~95 % reduction is seen in both the DMB1 and DMB2 KO lines. It is proposed that DMB1 and DMB2 are the β chains that associate with DMα to form two DM peptide editors, and so it is expected that a reduction in DMA would be seen in the DMB1 and DMB2 KO lines.

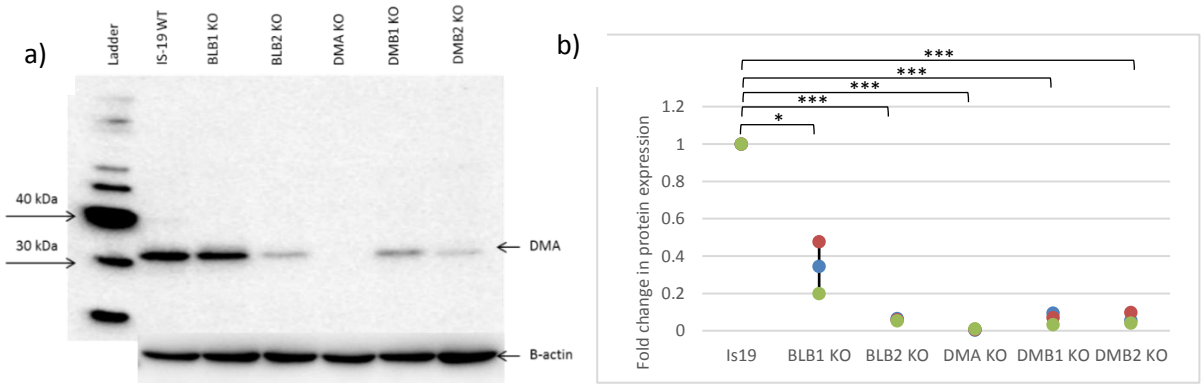


Figure 5.3.3 q-Western blots staining for DMA in IS19 single class II KO clonal lines.

IS19 WT cells and the clonal derived cell lines BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol, lysates were boiled in DTT before loading. a) The lysates were probed for DMA protein expression by western blotting using the antibody DMA-8. b) Relative DMA expression was quantified by densitometry, results are plotted as a relative fold change compared to wild type expression, normalised to β-actin staining. Blots were performed in triplicate for quantification.

Lastly, DMB2 protein expression was analysed. Figure 5.3.4a shows a western blot that was used for quantification. DMB2 is seen at ~38k Da when compared to the molecular mass marker; quantification of DMB2 is normalised to wild type level protein expression in figure 5.3.4.b. DMB2 expression is the same as the wild type in the BLB1 KO, DMB2 is absent in the DMA KO, and DMB2 is absent DMB2 KO line. It is expected that a reduction in DMB2 would be seen in the DMA KO line as DMA is the cognate chain of DMB2.

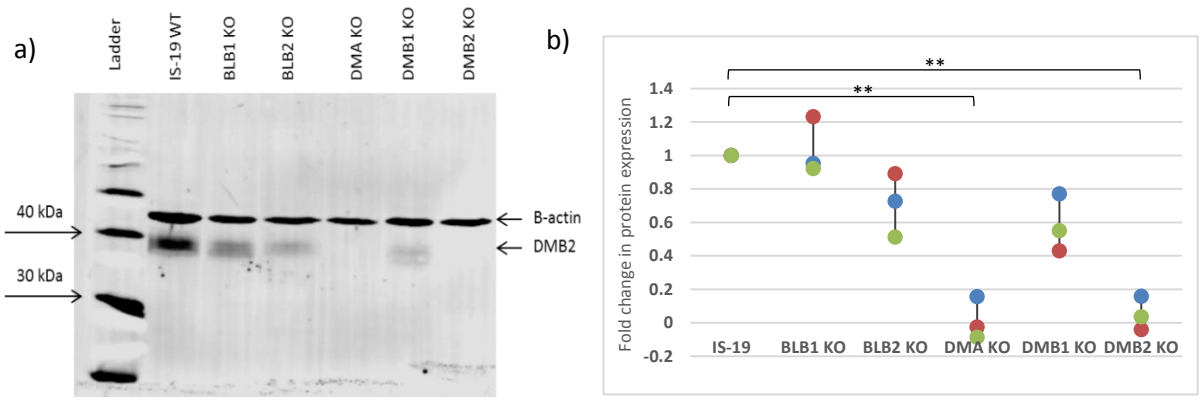


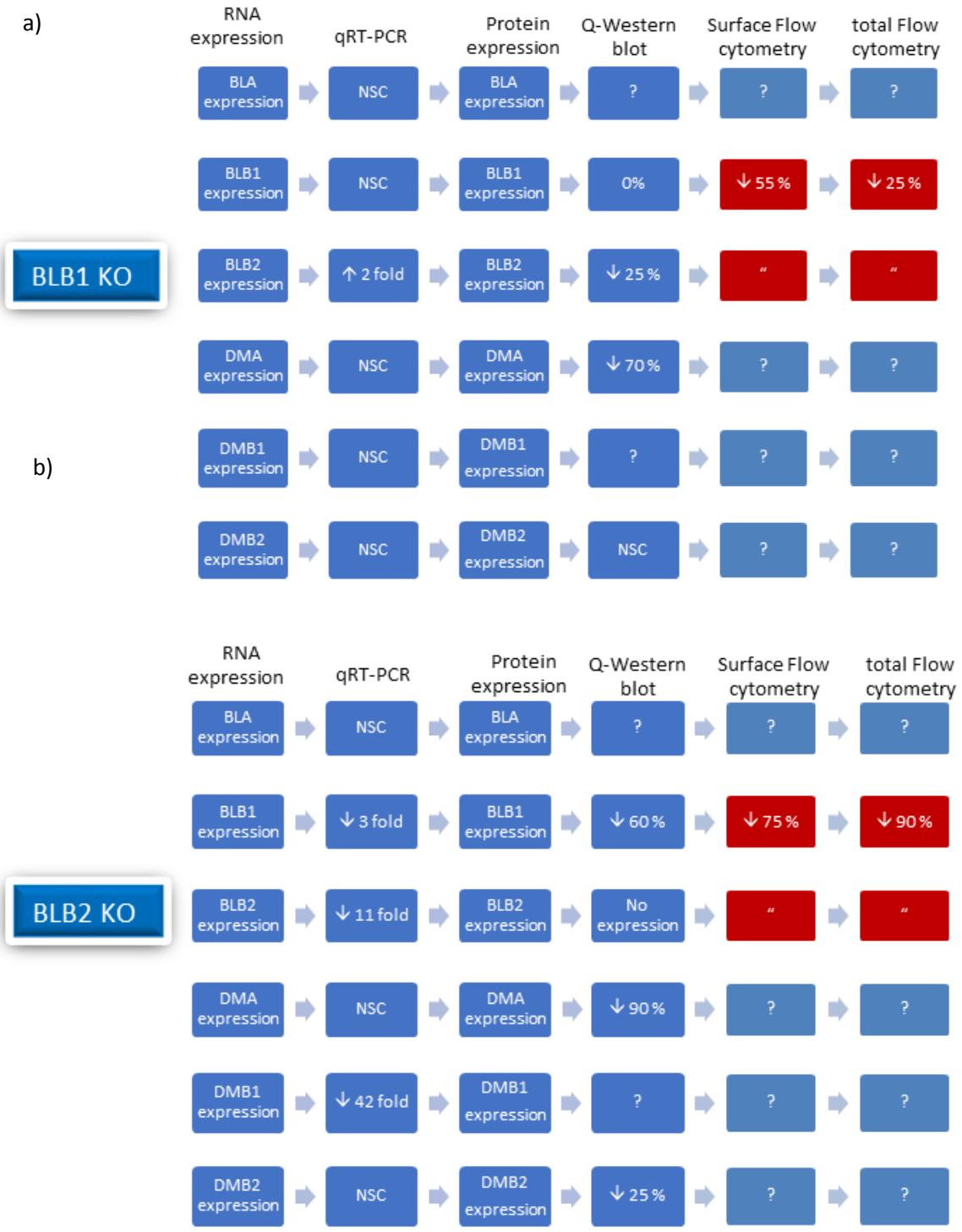
Figure 5.3.4 q-Western blots staining for DMB2 in IS19 single class II KO clonal lines.

IS19 WT cells and the clonal derived cell lines BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol, lysates were boiled in DTT before loading. a) The lysates were probed for DMB2 protein expression by western blotting using the antibody DMB2-11. b) Relative DMA expression was quantified by densitometry, results are plotted as a relative fold change compared to wild type expression, normalised to β-actin staining. Blots were performed in triplicate for quantification.

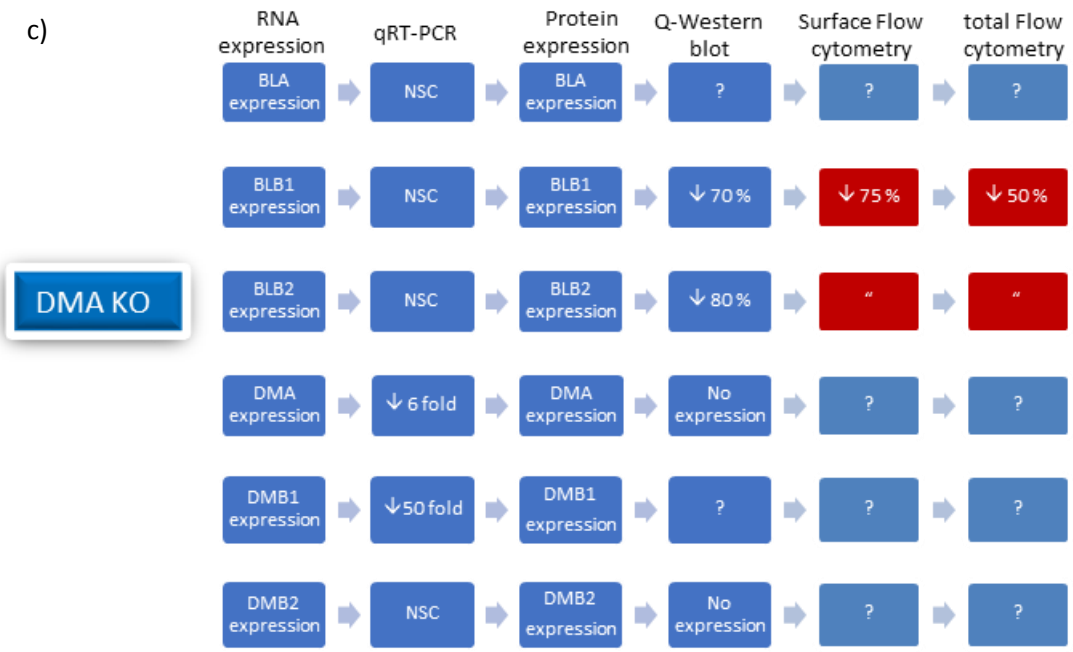
A summary of the analysis at the mRNA and protein level in all class II KO cell lines is described in figure 5.3.5. This summary will be used for discussion in section 5.4.

a)

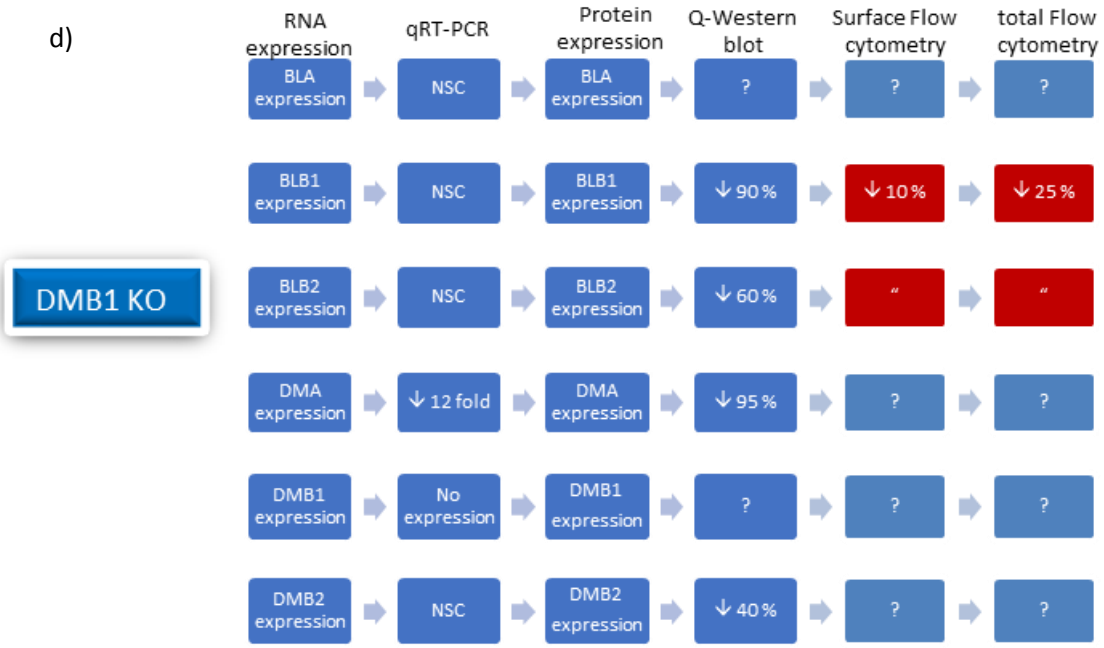
b)



c)


DMA KO
RNA expression
qRT-PCR
Protein expression
Q-Western blot
Surface Flow cytometry
total Flow cytometry
BLA expression
NSC
BLA expression
?
?
?
BLB1 expression
NSC
BLB1 expression
↓70%
↓75%
↓50%
BLB2 expression
NSC
BLB2 expression
↓80%
"
"
DMA expression
↓6 fold
DMA expression
No expression
?
?
DMB1 expression
↓50 fold
DMB1 expression
?
?
?
DMB2 expression
NSC
DMB2 expression
No expression
?
?

d)


DMB1 KO
RNA expression
qRT-PCR
Protein expression
Q-Western blot
Surface Flow cytometry
total Flow cytometry
BLA expression
NSC
BLA expression
?
?
?
BLB1 expression
NSC
BLB1 expression
↓90%
↓10%
↓25%
BLB2 expression
NSC
BLB2 expression
↓60%
"
"
DMA expression
↓12 fold
DMA expression
↓95%
?
?
DMB1 expression
No expression
DMB1 expression
?
?
?
DMB2 expression
NSC
DMB2 expression
↓40%
?
?

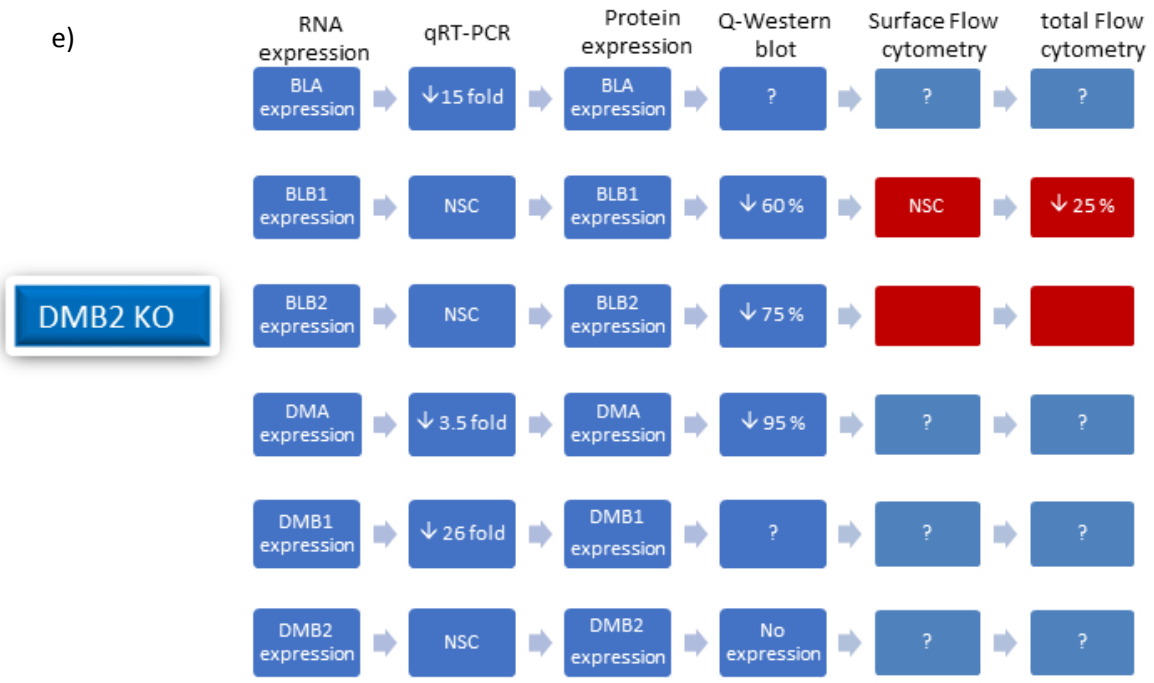


Figure 5.3.5 Summary of mRNA and protein expression in IS19 class II KO lines.

A summary of the findings of mRNA and protein expression is depicted for all IS19 class II KO lines. The data for each cell line is condensed into a flow diagram for BLB1 KO (a), BLB2 KO (b), DMA KO (c), DMB1 KO (d) and DMB2 KO (e). The KO cell line noted to the right of each diagram, with rows indicating the gene measured and columns describing how the data was measured. Effects on the gene are described as fold change or percentage. Red boxes indicate where class II could only be measured as total class II and BLB1 or BLB2 expression could not be individually discerned.

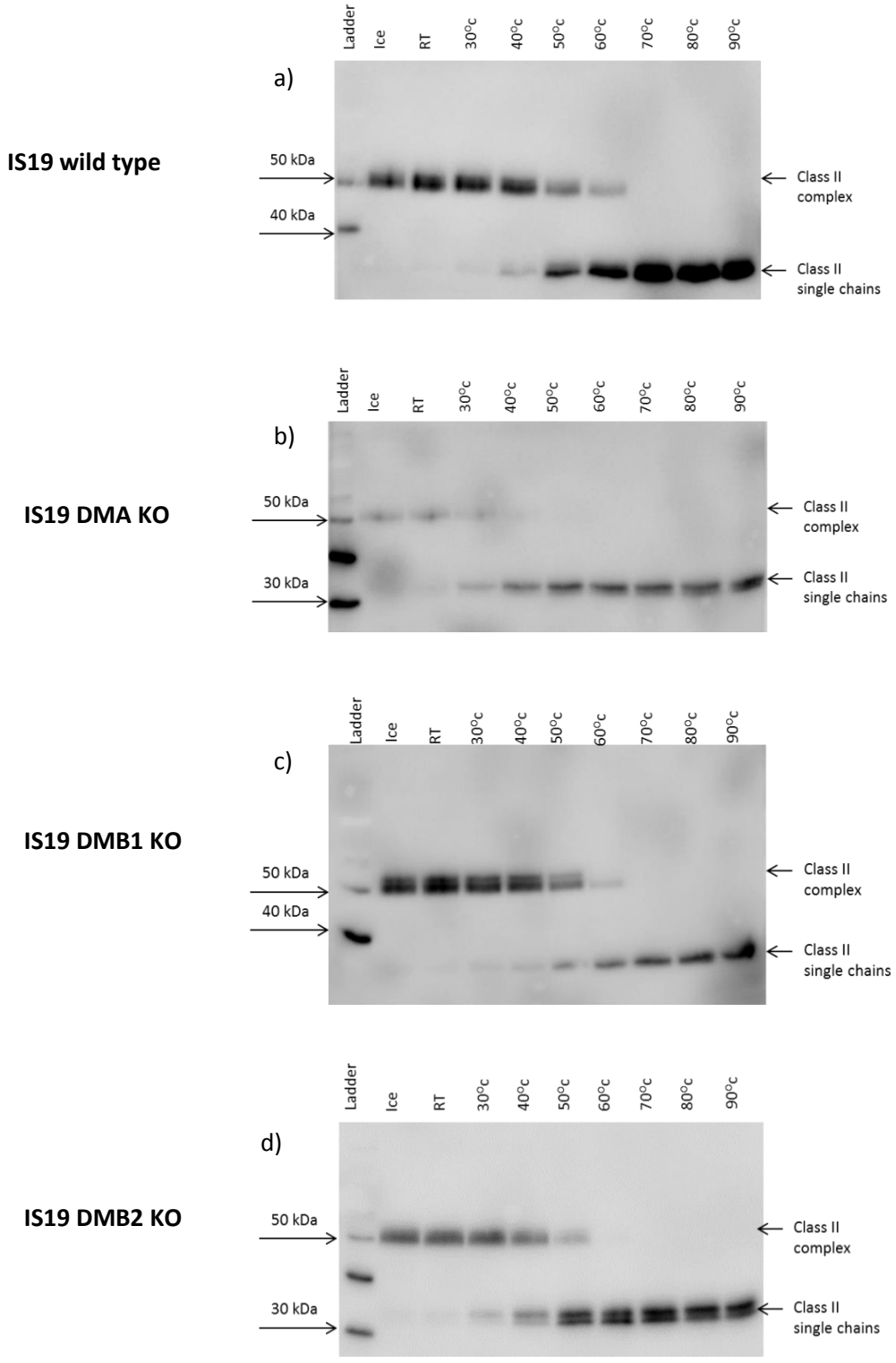


Figure 5.3.6 q-Western blots assessing temperature stability of BLB in IS19 WT and DM KO clonal lines.

IS19 WT cells and the clonal derived cell lines DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol. Lysates were incubated on ice, room temperature (RT) 30°C, 40°C, 50°C, 60°C, 70°C, 80°C or 90°C before loading. All blots were probed for BLB protein expression by western blotting using the antibody BLBcyt10. The temperature stability was assessed in the wild type (A), DMA KO (B), DMB1 KO (C) and DMB2 KO (D). Blots were performed in triplicate for quantification.

Next temperature stability of BLB was assessed in the three DM KO lines. Figure 5.3.6 shows western blots of IS19 WT (a), DMA KO (b), DMB1 KO (c) and DMB2 KO (d), to assess the ability of the MHCII to stay in heterodimeric complex over a temperature gradient between the three different KO lines. The wildtype line shows that the complexed class II begins to predominately fall apart at ~ 50 ^{o}C. Comparing this observation to the DMA KO line where we see that the class II complex begins to fall apart at a much lower temperature ~30 ^{o}C, the DMB1 KO shows dissociation at ~50 ^{o}C (comparable to the WT) and the DMB2 KO shows major dissociation starting at ~40 ^{o}c.

Western blots assessing temperature stability of the DM KOs were carried out in triplicate to allow for quantification, so the effect that the DM KOs have on stability could be quantified.

Figure 5.3.7 shows the quantification of MHCII complex and MHCII single β chains plotted over the temperature gradient, for IS19 (a), DMA KO (b), DMB1 KO (c) and DMB2 KO (d). To determine the point at which the class II complex predominately falls apart we take the value at which the MHCII complex and the MHCII single chains intersect. For IS19 WT stability is ~55 ^{o}C, the DMA KO stability is ~30 ^{o}C, ~50 ^{o}C for the DMB1 KO and ~45 ^{o}C for the DMB2 KO. The DMA KO has the strongest effect on stability, presumably as DMA is required for the formation of both DM peptide editors it has the strongest effect on stability of BLB. In contrast, the DMB1 KO has the subtlest of effects dropping from ~55 ^{o}C in the wild type to ~50 ^{o}C; again, this fits with the hypothesis that DMB1 is the poorly expressed DMB and comprises the minor DM editor. For the DMB2 KO we see a phenotype in between the DMA and DMB1 KO with temperature stability falling to 45 ^{o}c. This was expected as DMB2 is the β chain for the dominantly expressed DM peptide editor and is expected to be the prominent peptide editor, at least in haematopoietic cells.

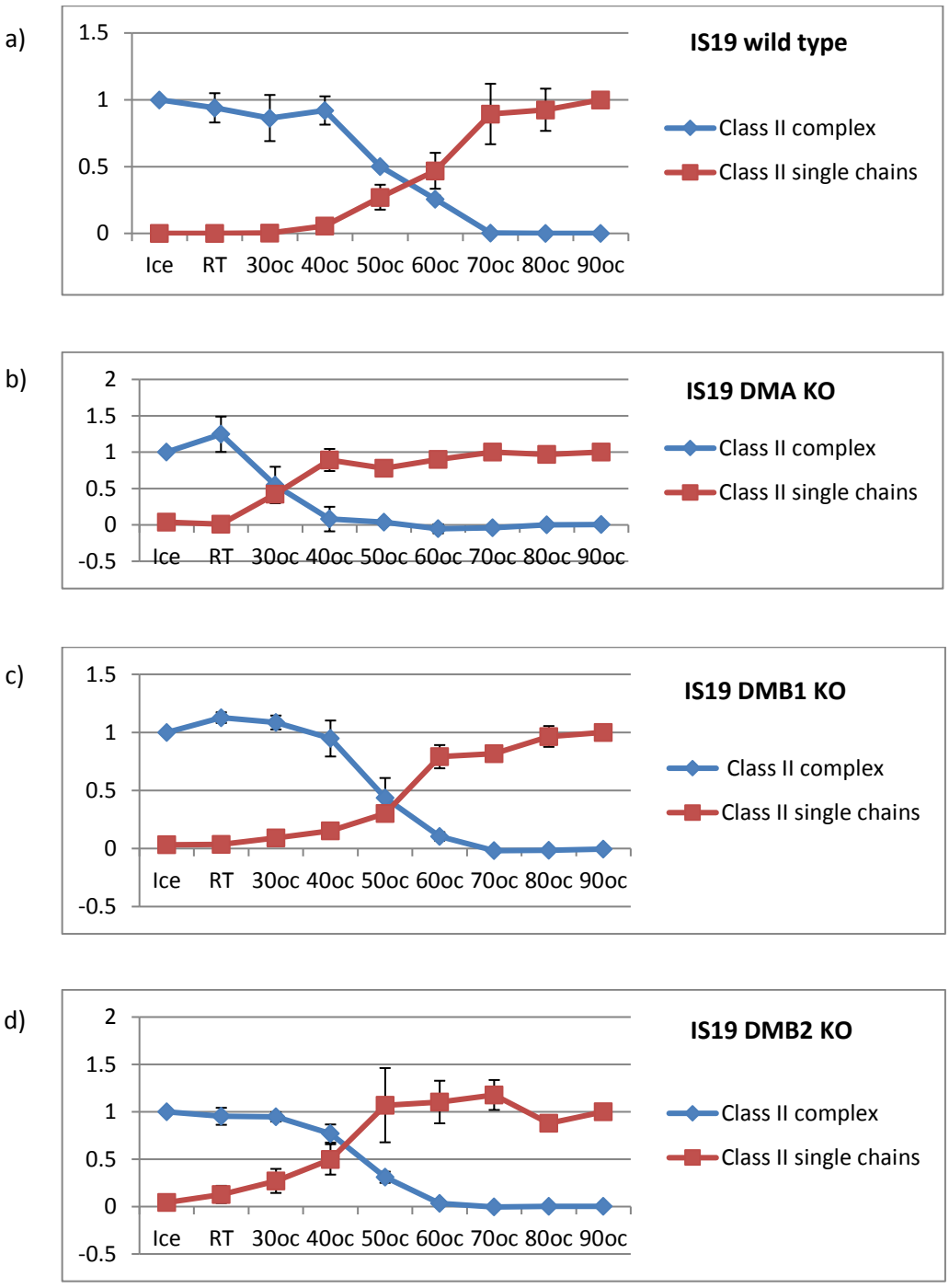


Figure 5.3.7 q-Western blot quantification for temperature stability of BLB class II complex and single beat chains in IS19 WT and DM KO clonal lines.

IS19 WT cells and the clonal derived cell lines DMA.2 cl-1, DMB1.4 cl-10 and DMB2.1 cl-11 were lysed using digitonin after a membrane enriching protocol. Lysates were incubated on ice, room temperature (RT) 30°C, 40°C, 50°C, 60°C, 70°C, 80°C or 90°C before loading. Al blots were probed for BLB protein expression by western blotting using the antibody BLBcyt10. The temperature stability was assessed in the wild type (A), DMA KO (B), DMB1 KO (C) and DMB2 KO (D). Blots were performed in triplicate for quantification from one lysate and reveal similar results of blots from an independent lysate. BLB complex and single chains were quantified by densitometry, results are plotted as a relative fold change normalised to either ice for class II complex or 90 °C for single chains.

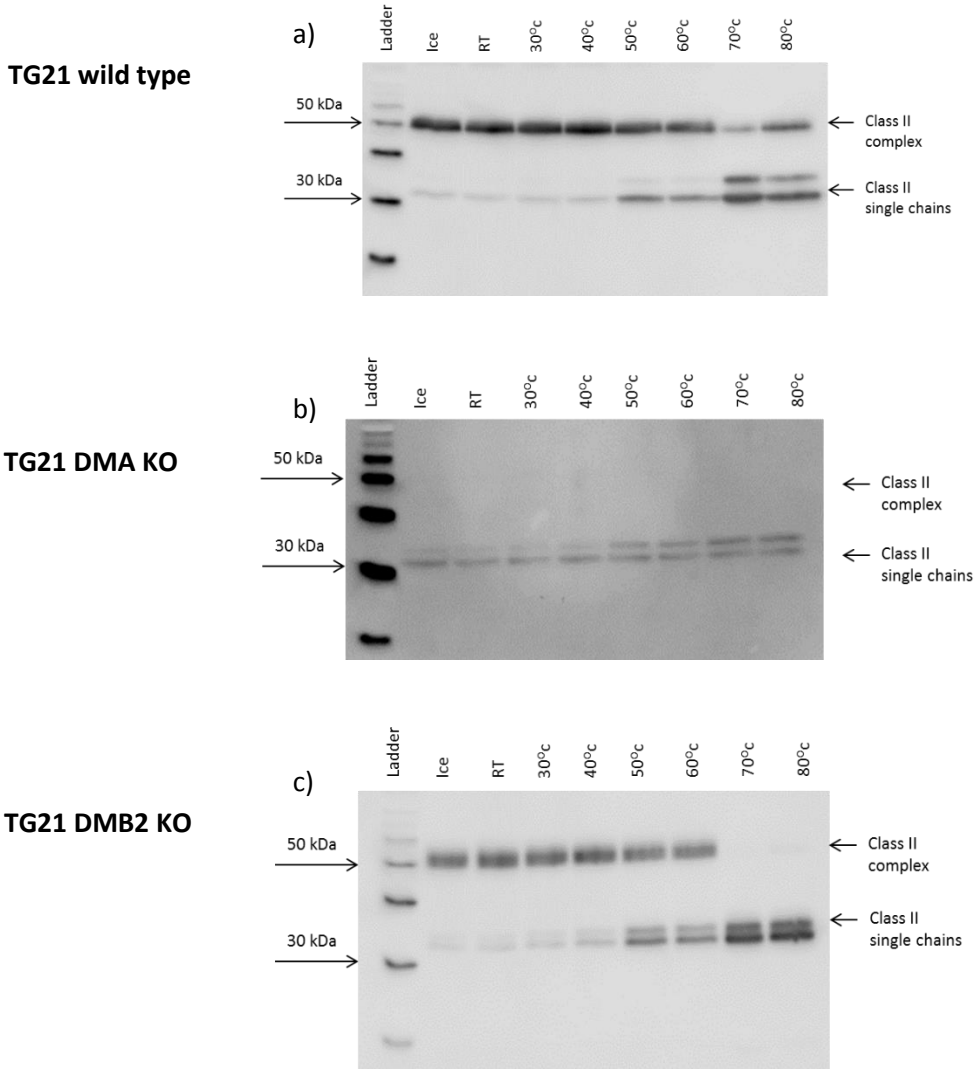


Figure 5.3.8 q-Western blots assessing temperature stability of BLB in TG21 WT, DMA KO and DMB2 KO clonal lines.

TG21 WT cells and the clonal derived cell lines DMA.2 cl-2 and DMB2.1/2 cl-14 were lysed using digitonin after a membrane enriching protocol. Lysates were incubated on ice, room temperature (RT) 30°c, 40°c, 50°c, 60°c, 70°c or 80°c before loading. Al blots were probed for BLB protein expression by western blotting using the antibody BLBcyt10. The temperature stability was assessed in the wild type (A), DMA KO (B) and DMB2 KO (C).

The effect of DM KOs on IS19 were compared to another haplotype, starting with the B21 haplotype using the B21 cell line, TG21. To compare the findings in TG21 with IS19 BLB1, BLB2, DMA, DMB1 and DMB2 KOs, replicate KO cell lines are under construction in TG21. All genes of interest have been knocked out at a polyclonal level with clonal lines derived for DMA and DMB2 KOs. Figure 5.3.8 shows preliminary data for the effect of DMA and DMB2 KOs on temperature stability (as done for IS19) compared to the wild type.

Figure 5.3.8a shows the temperature stability of wild type TG21, which begins to fall apart between 70-80 ^{o}C but is still not fully dissociated by 80 ^{o}C. This is in contrast to IS19 which begins to fall apart around 50 ^{o}C, showing that TG21 BLB has intrinsically higher thermostability. The TG21 DMA KO (fig 5.3.8b) shows full dependence on DM expression for a stable BLB complex to be expressed, with lysates incubated on ice not forming a stable heterodimeric complex. When this is compared to the DMB2 KO (fig 5.3.8c) we see an intermediate phenotype with the class II complex falling apart around 50-60 ^{o}C, suggesting that the absence of DMB2 results in less stable class II complex, presumably as the class II is unable to be edited by the DM2 peptide editor. The stable class II complex present compared to the DMA KO suggests that DM comprised of DMB1 can compensate to some degree. It is important to note that this preliminary experiment has only been done once and will need to be repeated with at least two different lysate preparations and in triplicate for quantification.

5.4 Discussion

In this chapter we undertook a thorough analysis of the IS19 class II KO cell lines at the RNA and protein level, although we were unable to analyse BLA or DMB1 protein levels due to reagent limitations. Moreover, we looked at the effect that DM KOs have on the ability of IS19 to form thermostable class II complexes. Lastly, we show preliminary thermostability data in TG21 WT, DMA and DMB2 KO cell lines.

5.4.1 Characterisation of IS19 BLB1 KO at the RNA and protein level

The RNA expression in the single class II KOs was compared to the protein expression to see how well the two results correlate (fig 5.3.5). Starting with the BLB1 KO, for BLA RNA expression we see no significant change (fig 5.2.1) and unfortunately due reagent limitations we are unable to assess what happens to BLA at the protein level. For BLB1 RNA expression no significant change was seen (fig 5.2.2) which as explained previously is not unexpected even though it is a BLB1 KO cell line, the indels generated in BLB1 do not necessarily disrupt the stability of the mRNA transcript and it is still able to be transcribed and remain stable. Though this is not always the case, and the mRNA can become unstable and be degraded to varying degrees. When this is compared to protein expression from q-Western blotting we see that no BLB1 protein is expressed (fig 5.3.2b). For BLB2 RNA expression we see a modest two-fold increase compared to the WT (fig 5.2.3), interestingly when this is compared to BLB2 protein expression the two results are in contention as q-western blotting reveals a ~25 % reduction in BLB2 (fig 5.3.2c). There are multiple potential explanations for such a result. Firstly, reduction of BLB2 mRNA is at the edge of detection and it is possible that the mRNA expression could be a non-significant change. Secondly, the mRNA upregulation could be accurate but BLB2 is less stable when BLB1 is not present and is turned over at a level that still results in a net 25 % loss of BLB2. Although it is unclear why a lack of BLB1 expression would result in unstable BLB2 expression, as we do not expect the two class II molecules to interact or to be dependent on each other for expression. It is noted that BLB1 and BLB2 are in opposite transcriptional orientations and separated by the tapasin gene (Jacob et al 2000), so it is unexpected that the mRNA expression of one BLB would affect the other.

Flow cytometry was also used to look at cell surface and total protein expression of BLB, but one limitation of this experiment is that the class II antibodies (2G11 and BLBcyt10) detect

both BLB1 and BLB2. In these experiments BLB1 and BLB2 expression can only be assumed by subtraction of class II expression between the BLB1 and BLB2 KO, though this method is flawed as the data suggests that BLB1 KO effects the regulation of BLB2 and *vice versa*.

What is seen from the cytometry data is that the BLB1 KO loses ~55 % surface expression and ~25 % total expression (fig 5.3.1) suggesting that a large amount of BLB1 is present on the surface compared to the total amount of BLB1 found in the cell. This is interesting as it suggests that proportionally more BLB1 is trafficked to the cell surface than BLB2.

For the DM proteins, starting with DMA, we see no significant change in DMA RNA (fig 5.2.4) but at the protein level we see a reduction of 70 % (fig 5.3.3b). This was unexpected, and it was expected that the expression of DMA would be independent of BLB1 expression. One explanation is that DM is expected to interact with BLB1 and in the absence of BLB1, less DM is required and therefore DMA is degraded to some extent. DMB1 shows no significant change at the RNA level (fig 5.2.5) and unfortunately due to limitations with DMB1 antibodies we were unable to look at the effect of BLB1 deficiency on DMB1 protein.

Lastly, we looked at DMB2, with no significant change being seen at the RNA level (fig 5.2.6) or at the protein level (fig 5.3.4b). This seemed inconsistent as DMA is reduced at the protein level, so it was expected that a reduction would be seen in DMB2 as well. Though it is noted that ~30 % of DMA is still expressed in the BLB1 KO and this could be enough substrate to interact with DMB2, and hence maintain DMB2 levels as wild type and not result in DMB2 being degraded.

5.4.2 Characterisation of IS19 BLB2 KO at the RNA and protein level

BLA mRNA expression shows no significant change (fig 5.2.1). BLB1 mRNA expression was down regulated three-fold (fig 5.2.2) and the protein expression is reduced ~60 % (fig 5.3.2b), suggesting that the decrease in BLB1 expression comes from downregulation, rather than degradation. BLB2 RNA expression has an 11-fold decrease (fig 5.2.3), suggesting that the indels generated in BLB2 effect the overall stability of BLB2 mRNA, at the protein level BLB2 is completely absent by q-western blot (fig 5.3.2c).

Flow cytometry analysis of cell surface and total protein expression of BLB shows that the BLB2 KO loses ~75 % surface expression and ~90 % total expression (fig 5.3.1), suggesting that more BLB2 is present inside the cell compared to the cell, which is the opposite of what is seen with BLB1. These two results could indicate that the poorly expressed class II, BLB1, is preferentially transported to the cell surface when compared to BLB2 and may compensate for the lower expression of BLB1.

DMA shows no significant change in mRNA expression (fig 5.2.4), but a 90 % reduction is seen at the protein level (5.3.3). As with the BLB1 KO, this reduction of DMA expression was unexpected, as it was thought that DMA expression would be independent of BLB expression. Again, this could be explained by the fact that DM is expected to interact with BLB and in the absence of BLB2 less DM is required, and excess DMA is degraded. It is interesting that deletion of BLB1 or BLB2 results in decreased DMA at the protein level and is consistent with the notion that DMA expression is dependent on BLB protein expression. It is also noted that knocking out BLB1 (the more poorly expressed BLB) has a reduction of 70 % DMA protein and knocking out the dominantly expressed BLB, BLB2, results in a greater reduction of DMA (90 %). This suggests that the proportion of DMA protein that is reduced depends on the amount of BLB (1 or 2) present in the cell.

DMB1 shows a ~42-fold decrease at the RNA level in the BLB2 KO (fig 5.2.5) but cannot be compared to the protein level. Lastly, we looked at DMB2, with no significant change being seen at the mRNA level (fig 5.2.6) but a 25 % reduction is seen at the protein level (fig 5.3.4). It is curious that a reduction in DMB2 protein is only seen in the BLB2 KO and not the BLB1 KO, it has been hypothesised that the two different DM peptide editors (comprised of either DMB1 and DMB2) might have a preference to the BLB they target. It was suggested that DMB1 may mainly work with BLB1 and DMB2 with BLB2. One interpretation of this result is that DMB2 protein levels are affected by the absence of BLB2 as this is the major substrate of DMB2. Another possibility is that the 90 % reduction of DMA results in degradation of DMB2 as there is not enough DMA for DMB2 to associate with.

5.4.3 Characterisation of IS19 DMA KO at the RNA and protein level

BLA RNA expression shows no significant change in the DMA KO (fig 5.2.1). BLB1 RNA expression shows no significant change (fig 5.2.2) and the protein expression is reduced ~70 % (fig 5.3.2b). BLB2 mRNA expression shows no significant change (fig 5.2.3) and the protein expression is reduced ~80 % (fig 5.3.2c). It is thought that BLB reduction is the result of lower stability leading to degradation, as BLB1 and BLB2 cannot be loaded with high affinity peptide by the DM complexes.

Flow cytometry analysis of cell surface and total protein expression of BLB shows that the DMA KO loses ~75 % surface expression and ~50 % total expression (fig 5.3.1). More class II is absent from the cell surface than in total. In the absence of DMA, BLB is unable to be edited by DM and is expected to result in class II being more frequently loaded with sub-optimal peptides, which is would lower BLB stability. This could explain why less class II is

seen at the surface either because the class II is trafficked to the cell surface less efficiently or because unstable class II molecules on the surface are internalised more rapidly.

DMA RNA expression is reduced by six-fold (fig 5.2.4), with no protein being expressed (fig 5.3.3). DMB1 shows a ~50-fold decrease at the mRNA level (fig 5.2.5) and in future work will be validated in independent DMA KO clones as well as determining DMB1 protein expression when reagents become available.

DMB2 shows no significant change at the mRNA level in the DMA KO (fig 5.2.6) but is completely absent at the protein level (fig 5.3.4). It was expected that DMB2 protein would be decreased without DMA to associate with, but it was unexpected that DMB2 would be undetectable. This is different from what is seen in mammals, where DMA deletion results in a small reduction in DMB protein or no change at all.

5.4.4 Characterisation of IS19 DMB1 KO at the RNA and protein level

BLA RNA expression shows no significant change in the DMB1 KO (fig 5.2.1). BLB1 RNA expression shows no significant change (fig 5.2.2) but the protein expression is reduced ~90 % (fig 5.3.2b). BLB2 mRNA expression shows no significant change (fig 5.2.3) and the protein expression is reduced ~60 % (fig 5.3.2c). In concordance with the DMA KO It is thought that BLB reduction is the result of lower stability leading to degradation, as BLB1 and BLB2 cannot be loaded with high affinity peptide in the absence of DM complexes. In this case only DMB1 is absent and it was unexpected that DMB1 deletion would have a pronounced effect on both BLB1 and BLB2. It is also noted that the DMB1 KO has a larger effect on BLB1 than BLB2 again providing some evidence that DMB1 and DMB2 may have preferentially edit different BLB molecules.

Flow cytometry analysis of BLB shows that the DMB1 KO loses ~10 % surface expression and ~25 % total expression (fig 5.3.1). More class II is absent from the cell surface than in total. This is the opposite pattern of what is seen in the DMA KO, where more total class II is reduced than at the cell surface.

DMA RNA expression is reduced by 12-fold (fig 5.2.4), with a 95 % reduction in protein being expressed (fig 5.3.3). This suggests that DMB1 editing effects the RNA expression of DMA and in turn results in less protein. It is not conclusive that all or some of the DMA reduction is not due to degradation of DMA as it cannot associate with DMB1 and results in turnover of DMA.

DMB1 shows no expression at the RNA level (fig 5.2.5), suggesting that the RNA is completely unstable due to the indels generated. DMB2 shows no significant change at the RNA level in the DMB1 KO (fig 5.2.6) but has a 40 % reduction at the protein level (fig 5.3.4). It was unexpected that a KO of DMB1 would affect the protein levels of DMB2, as DMB1 and DMB2 are not expected to interact. A possible explanation for this is that the DMB1 KO has greatly reduced DMA protein expression, and with such a large reduction in DMA it is possible that DMB2 is in excess compared to its cognate chain (DMA) and is therefore degraded.

5.4.5 Characterisation of IS19 DMB2 KO at the RNA and protein level

BLA RNA expression is reduced 15-fold in the DMB2 KO (fig 5.2.1). This is the only class II KO that effects BLA RNA. The reasons for this remain unclear and will need to be validated in independent DMB2 KO clones as well as assessing the protein level of BLA if and when reagents become available. BLB1 RNA expression shows no significant change (fig 5.2.2) but

the protein expression is reduced ~60 % (fig 5.3.2b). BLB2 mRNA expression shows no significant change (fig 5.2.3) and the protein expression is reduced ~75 % (fig 5.3.2c). As with the DMA KO, it thought that BLB reduction is the result of lower stability leading to degradation, as BLB1 and BLB2 cannot be loaded with high affinity peptide in the absence of DM complexes but in this case only DMB2 is absent. It is also noted that the DMB2 KO has a larger effect on BLB2 than BLB1, again providing some evidence that DMB1 and DMB2 may have preferentially edit different BLB molecules.

Flow cytometry analysis of cell surface and total protein expression of BLB shows that the DMB2 KO has no significant change in surface expression and a ~25 % decrease in total expression (fig 5.3.1). As with the DMB1 KO changes in MHCII expression is seen dominantly for total expression and in this case no change in MHCII protein expression is seen at the cell surface.

DMA mRNA expression is reduced by 3.5-fold (fig 5.2.4), with a 95 % reduction in protein being expressed (fig 5.3.3). This suggests that DMB2 deletion effects the mRNA expression of DMA, though not as much as the DMB1 KO, but the protein is reduced to the same staggering degree of 5 % of the total. It is unlikely that the reduction of DMA is purely down to the down regulation at the mRNA level and that the DMB2 KO is affecting the stability of DMA protein either because DMA does not have DMB2 to associate with or through other means.

DMB1 shows a large down regulation at the RNA level of 26-fold in the DMB2 KO (fig 5.2.5), which was not expected. DMB2 shows no significant change at the mRNA level (fig 5.2.6) but is completely absent at the protein level (fig5.3.4).

5.4.6 Summary of IS19 class II KO characterisation at the mRNA and protein level

Many interesting and currently unexplained results were obtained during this work, the results of which are summarised in figure 5.3.5. The trends that were consistently noticed were that the BLB1 KO reduces BLB2 expression and *vice versa*. It is also seen that BLB KOs effect the levels of DMA protein expressed and this is more pronounced in BLB2 than BLB1 KOs. It is also seen that the BLB2 KO effects DMB2 expression whereas the BLB1 KO does not. Equally, the DMB KOs effect both BLBs but DMB1 has a more pronounced effect on BLB1 and the DMB2 KO on BLB2, which suggests there is a preference between DMBs for which BLB they edit. It is also seen by flow cytometry that BLB1 appears to have a higher percentage of the total go to the cell surface when compared to percentage of total BLB2 that goes to the surface. This is interesting and suggests a possible preferential transport of BLB1 to the cell surface which could be a mechanism to compensate for the low expression of BLB1.

DM expression also appears to be complex, with deletions in any DM gene effecting the expression of the other DMs. It was expected that deletions in DMA would affect DMB1 and DMB2 expression, and deletions in DMB1 and DMB2 not to affect the expression of each other. What we observed was that a deletion in DMA, DMB1 or DMB2 effects (where measurable) the protein expression of the other two DMs and editing of either DMA or DMB2 result in down regulation of DMB1 RNA. There is a strong dependence of expression between the three DMs and it is currently unknown to what level this is down to interactions at the protein or transcriptional level.

These results need to be validated by repeating the experiments in independently made class II KO cell lines, which have already been made and need to be tested. We also intend to

transfect in the knocked-out class II gene into the appropriate KO cell line to see if wild type expression is restored and the effects seen are rescued.

5.4.7 Characterisation of temperature stability of IS19 DM KO cell lines

After seeing such pronounced changes in class II expression in the DM KO cell lines, we looked at the effect each DM KO has on class II stability. As previously mentioned (general introduction) MHCII has been shown to run as a complex by western blot if lysates are run in a reducing gel but are not boiled before loading. To build upon this we treated the lysates at different temperatures ranging from 30 ^{o}C – 90 ^{o}C to assess at what temperature the MHCII complex begins to fall apart. This was completed in WT IS19 cells and the three DM KOs (DMA, DMB1 and DMB2); the results were quantified and compared (fig 5.3.6 and 5.3.7). It was observed that the WT IS19 MHCII complex predominately dissociated at 55 ^{o}C, the DMA KO at 30 ^{o}C , the DMB1 KO at 50 ^{o}C and the DMB2 KO at 40 ^{o}C (fig 5.3.7). This result was as expected, as IS19 appeared to be dependent on DM, the DMA KO would have the biggest effect on temperature stability, the DMB2 KO would have the second largest effect being the predominantly expressed DMB and DMB1 would have the smallest effect as it is poorly expressed. Part of the future work would be to devise an experiment where we can quantify the effect the DM KOs have on the temperature stability of each BLB. As the changes of the class II temperature stability are expected to be a result of the peptide that the class II is loaded with, we expect a change in the peptide repertoire of the three DM KO cell lines. In the future we hope to determine the changes in the DM KO cell line peptide repertoires by immunopeptidomics.

5.3.7 Initial characterisation of temperature stability of TG21 DMA and DMB2 KO cell lines

We set out to make class II KOs in the TG21 cell line to allow comparison with the results shown in IS19. Preliminary results are shown of TG21 WT, DMA KO and DMB2 KOs temperature stability. Initial observations show that BLB in TG21 wildtype cells have a higher native temperature stability when compared to IS19. IS19 majorly dissociates at 55 ^{o}C, whereas BLB can still be seen in complex at 80 ^{o}C in TG21 and starts to dissociate between 70-80 ^{o}c. In the DMA KO BLB is completely dissociated at all temperatures, showing that TG21 is fully dependent on DMA for SDS-stable class II dimers, which is not seen in IS19. The DMB2 KO has reduced temperature stability compared to the wild type with complexed class II falling apart between 50-60 ^{o}C and no complex being seen at 70 ^{o}c. This again indicates an important role for DMB1 in peptide editing of class II, as seen in IS19. Future work will consist of repeating this experiment in triplicate for quantification alongside the DMB1 KO, as well as assessing the peptide repertoire in TG21 DMA, DMB1 and DMB2 KO cell lines.

6. Characterisation of IS19 WT, BLB1, BLB2 and DMA KO cell lines at the peptide level

6. Characterisation of IS19 WT, BLB1, BLB2 and DMA KO cell lines at the peptide level

6.1 Introduction

In chapter 5 the MHCII KO cell lines (BLB1, BLB2, DMA, DMB1 and DMB2) were analysed at the RNA and protein level revealing complex and interesting results, with temperature stability data inferring changes in the peptide repertoire of DM KO lines. Previously, no work has been published detailing the peptide motifs of chicken class II molecules.

In this chapter we use immunopeptidomics to assess the peptide repertoire of IS19 WT, BLB1 KO, BLB2 KO and DMA KO cell lines and use Gibbs clustering to elucidate the peptide motif. In addition, we look at the length distribution of peptides in each line and compare the numbers of shared and unique peptides between these lines.

Firstly, one important consideration when analysing immunopeptidomics data needs to be discussed. The list of de novo sequenced peptides that are collected are not quantative and each unique peptide is only recorded once and does not reflect abundance of a particular peptide that is seen. Depending on the frequency of a peptide in the sample, this peptide may not be seen every time that sample is analysed by immunopeptidomics. For example, if there are dominant peptides in the sample, that peptide will be seen in each run, whereas peptides that are low abundance may only be seen over multiple runs. This sampling issue generally results in groups of peptides that not seen in the same sample over the course of multiple immunopeptidomics runs and needs to be considered when characterising what peptides are unique to different samples.

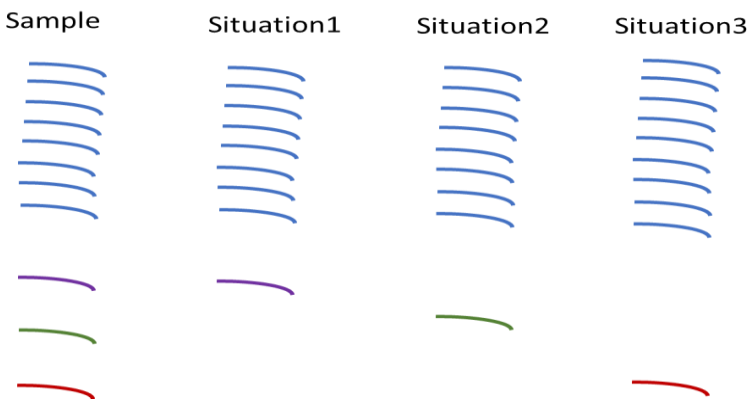


Figure 6.1.1 Illustration of sampling discrepancies found during immunopeptidomics analysis

One sample is shown to contain four different peptides, shown as blue, purple, green and red. Three situations are shown representing example results that can be seen after immunopeptidomics analysis, with the three situations representing three immunopeptidomics runs of the sample. Different peptide subsets are only seen in some runs but not in all.

6.2 Assessing length distribution of peptides eluted from class II KO lines

Firstly, the number of peptides and the number of peptides at a particular length obtained from each cell line were assessed. Figure 6.2.1 shows the length distribution and total peptide numbers for IS19 WT, BLB1 KO, BLB2 KO and DMA KO, for all peptides (fig 6.2.1a), just peptides with post-translational modifications (PTM) (fig 6.2.1b) and peptides without PTM (fig 6.2.1c). For all three plots the wild type has the most peptides, followed by the BLB1 and the BLB2 KOs with roughly equal number of peptides, whilst the DMA KO has the fewest peptides overall (fig 6.2.1). The majority of peptides are 15-19 amino acids in length for all four lines (fig 6.2.1). Interestingly we notice that for total peptides and peptides without PTM an increased number of peptides are seen between 8-9 amino acids in length (fig6.2.1a and c).

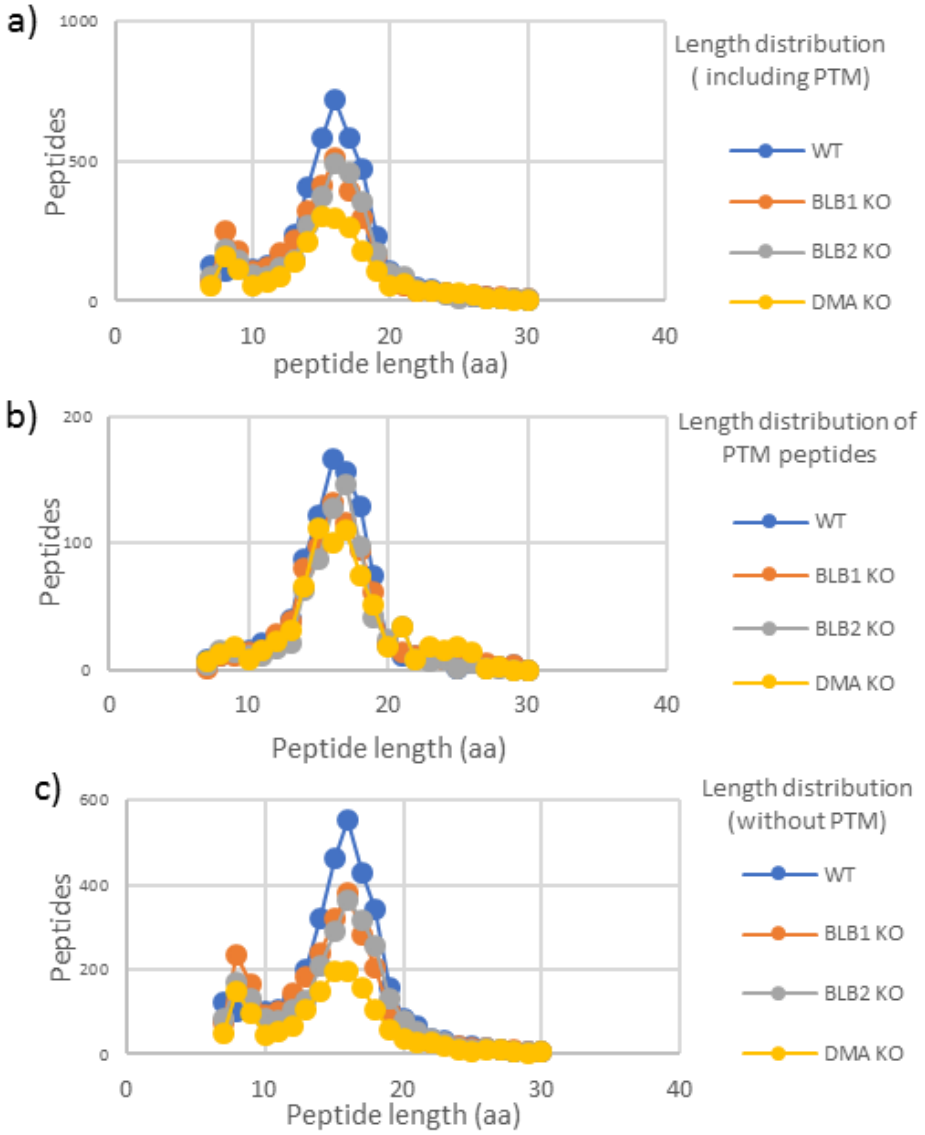


Figure 6.2.1 Length distribution of eluted peptides from WT and MHCII KO lines

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides were plotted by length against peptide number. a) total peptide numbers, b) peptides with post-translational modifications and c) peptides without post-translational modifications. WT is shown in blue, BLB1 KO (orange), BLB2 KO (grey) and DMA KO (yellow).

6.3 Determining peptide motifs of class II KO lines

Peptides between 15-19 amino acids long were analysed for a potential nine amino acid motif for all four cell lines by Gibbs clustering. Figure 6.3.1 shows that the wild type has a motif of mainly large hydrophobic amino acids, leucine, isoleucine phenylalanine and valine at position one, small hydrophobic and polar amino acids at position four, serine, alanine and threonine, with mainly hydrophobic amino acids at positions six, seven and nine (fig 6.3.1a). When comparing the WT motif to the BLB1 KO (peptides from BLB2 alone), we see a similar motif, with the nearly the same amino acids at every position. The main differences are seen at position one where phenylalanine and tyrosine are missing in the BLB1 KO and at position four where valine is missing (fig 6.3.1a and b). Equally, when comparing the WT motif to that of the BLB2 KO (peptides from BLB1 alone), the motif is largely the same at positions six, seven and eight, with most differences being seen at positions one and four. The BLB2 KO has a dominant phenylalanine and tyrosine at position, whereas these residues are much less pronounced in the WT, similarly alanine and valine and isoleucine are dominant at position four for the BLB2 KO and the WT has a larger proportion of serine and threonine (fig 6.3.1 a and c). Interestingly, the DMA KO has a motif that is virtually identical to the BLB1 KO and is missing the phenylalanine and tyrosine seen in the WT and the BLB2 KO at position one. It has the aspartic acid seen at position four in the WT and BLB1 KO but is absent in the BLB2 KO (fig 6.3.1).

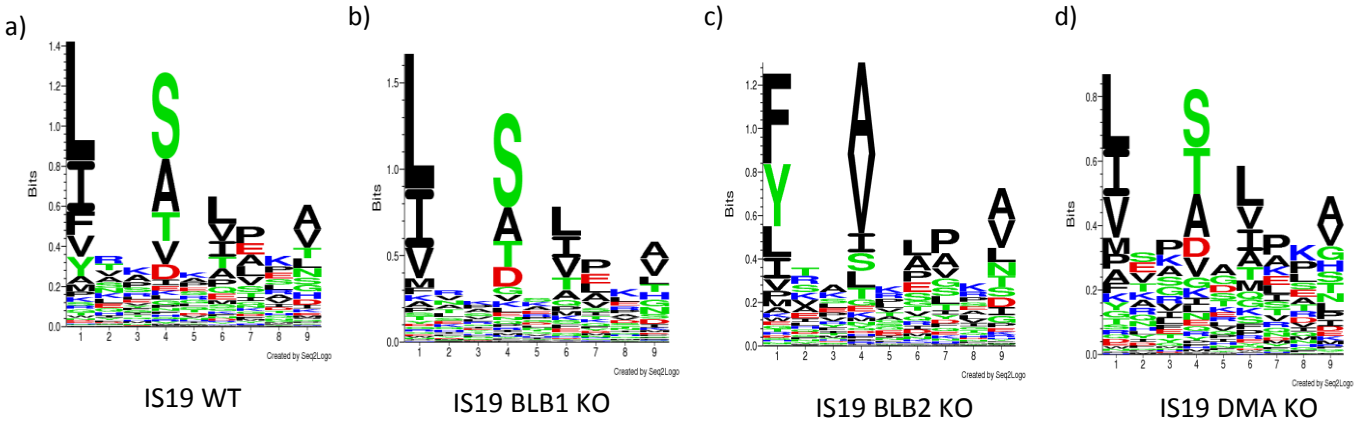


Figure 6.3.1 Peptide motif of eluted peptides from WT and MHCII KO lines

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from a) WT, b) BLB1 KO, c) BLB2 KO and d) DMA KO were analysed by Gibbs clustering to determine a motif.

Figure 6.3.1 shows clearly different motifs between the BLB1 and BLB2 KO. Moreover the motif for both BLBs appears to be present in the WT. To confirm that both BLB motifs are seen in the WT, the WT was subject to further Gibbs clustering to see if the motif of both BLBs is present. Figure 6.3.2 shows additional clustering of the WT which reveals two distinct motifs, IS19 WT cluster 1 shows a motif that is almost identical to the BLB1 KO, with the same residues seen at position one, four, six, seven and nine in the WT and BLB1 KO but with some variations in frequency at position four. Importantly, phenylalanine and tyrosine are missing at position one, showing that the BLB2 KO motif is not present (fig 6.3.2a and b). Conversely, the second WT cluster has a motif that is virtually identical to the BLB2 KO, with the same amino acids shared between the WT and the BLB2 KO at position one, four, six, seven and nine but again with some variations in amino acid frequency (fig 6.3.2c and d). Notably, at position one phenylalanine and tyrosine are seen and at position four aspartic acid is absent and valine is dominant (fig 6.3.2c and d). The two distinct motifs seen in the two WT clusters confirm that the motifs seen for the BLB1 and BLB2 KO lines are representative of motifs seen in the WT.

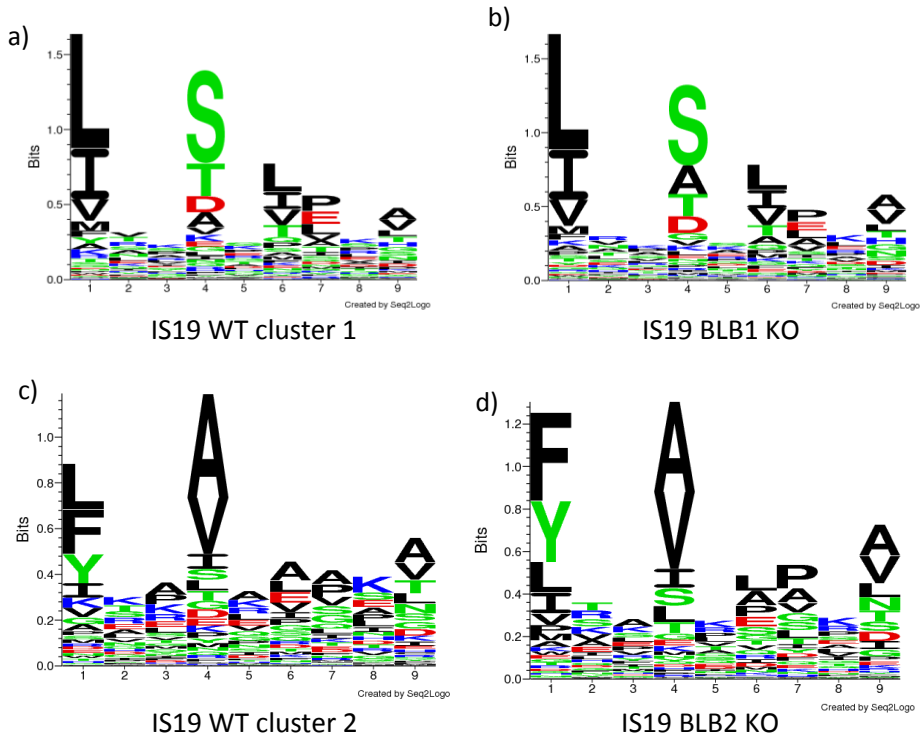


Figure 6.3.2 Two peptide motifs found in WT by further Gibbs clustering

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from WT were analysed by Gibbs clustering to reveal two different motifs a) & c) and compared to the motifs found in b) BLB1 KO and the d) BLB2 KO.

6.4 Analysis of unique peptides shared between class II KO lines

To further confirm that the motif seen in the BLB1 KO is shared with the WT, the eluted peptide repertoire from the WT and the BLB1 KO were compared for shared and uniquely presented peptides by Venn diagram (fig 6.4.1a). This comparison revealed 42.5 % of peptides between the two lines were shared, the motif of these shared peptides was then obtained by Gibbs clustering (fig 6.4.1b), as expected this motif revealed the BLB1 KO motif as seen in figure 6.3.1b, which is the same as the motif seen in WT motif cluster 1 (fig 6.3.2a). A proportion of peptides (43.8 %) are seen to be presented in the WT only (fig 6.4.1.a), which is not unexpected as the WT expresses BLB1 too and therefore will have peptides that are not seen in the BLB1 KO. Interestingly, 13.8 % of peptides from the BLB1

KO are not found in the WT, whereas we would expect that all peptides found in the BLB1 KO to be present in the WT.

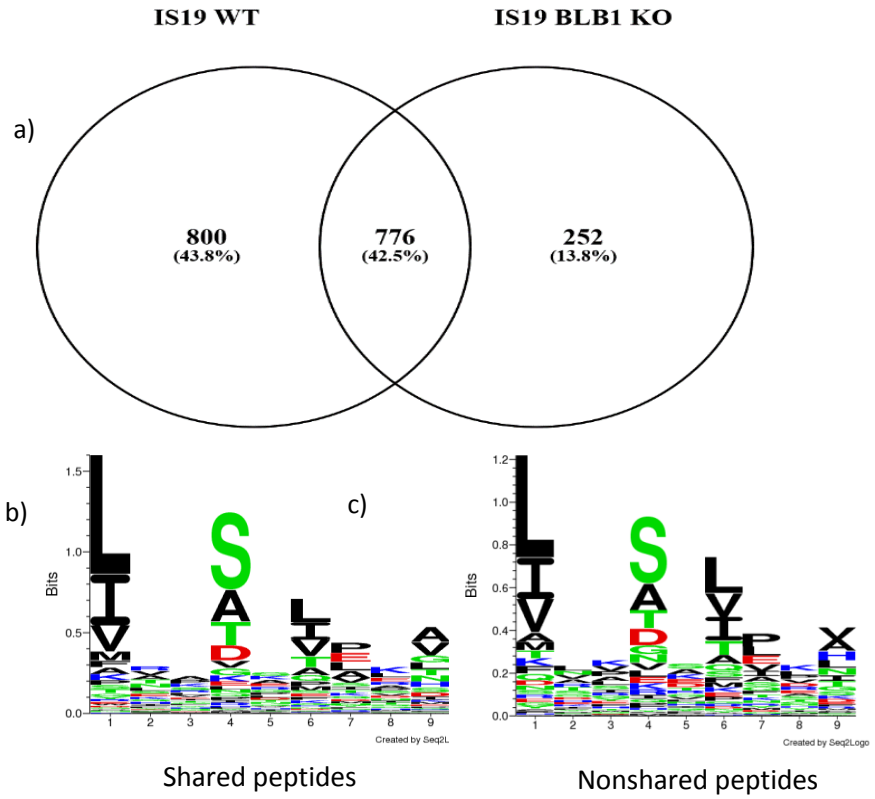


Figure 6.4.1 Comparison of Peptides derived from WT and BLB1 KO

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from WT and BLB1 KO were compared for shared peptide and unique peptide using a Venn diagram a). Peptides that are shared by the WT and the b) BLB1 KO and c) unique to the BLB1 KO were analysed by Gibbs clustering.

To understand if the proportion of peptides that are unique to the BLB1 KO have a unique motif, these peptides were analysed by Gibbs clustering (fig 6.4.1c). The motif of peptides unique to the BLB1 KO revealed a motif that is almost identical to that of peptides shared with the WT. The same dominant residues are seen at position one, four, six, seven and nine; showing that these unique peptides still adhere to the same motif found for the BLB1 KO and is not an exclusive motif seen only in the BLB1 KO. Likely, these unique peptides have

differences in the more variable positions (two, three, five and eight) and are present at low numbers and may not have been identified in the WT sample.

The same analysis was carried out to examine whether the BLB2 KO motif is shared with the WT. This comparison revealed 21.5 % of peptides between the two lines were shared (fig 6.4.2a). The motif of these shared peptides was then obtained by Gibbs clustering (fig 6.4.2b). As expected this motif revealed the BLB1 KO motif as seen in figure 6.3.1c, which is the same as the motif seen in WT motif cluster 2 (fig 6.3.2c). As seen before, a large proportion of peptides (48.2 %) are seen to be presented in the WT only (fig 6.4.2.a), which is not unexpected as the WT expresses BLB2 and therefore will have peptides that are not seen in the BLB2 KO. Again, a subset of peptides (30.3 %) from the BLB2 KO are not found in the WT, against expectations.

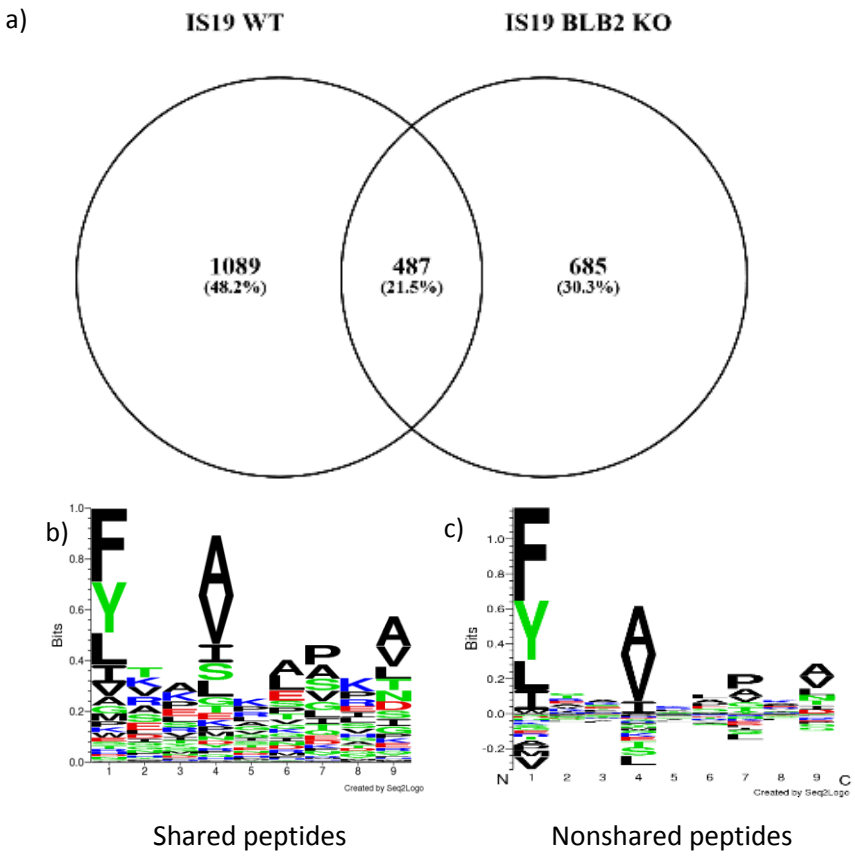


Figure 6.4.2 Comparison of Peptides derived from WT and BLB2 KO

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from WT and BLB2 KO were compared for shared peptide and unique peptide using a Venn diagram a). Peptides that are shared by the WT and the b) BLB2 KO and c) unique to the BLB2 KO were analysed by Gibbs clustering.

The proportion of peptides that are unique to the BLB2 KO were analysed by Gibbs clustering (fig 6.4.2c). The motif of peptides unique to the BLB2 KO revealed a motif that is almost identical to that of peptides shared with the WT, with largely the same dominant residues are seen at position one, four, six, seven and nine. Showing that these unique peptides still adhere to the same motif found for the BLB2 KO and is not an exclusive motif seen only in the BLB2 KO, mimicking the results we saw for the BLB1 KO (fig 6.4.1). It is expected that these unique peptides have differences in the more variable positions (two, three, five and eight) that are present at low numbers and may not have been identified in the WT sample.

The same analysis was carried out to examine the DMA KO motif is shared with the WT. This comparison revealed only 16 % of peptides between the two lines were shared (fig 6.4.3a). The motif of these shared peptides was then obtained by Gibbs clustering (fig 6.4.3b). As expected this motif revealed the DMA KO motif as seen in figure 6.3.1d.

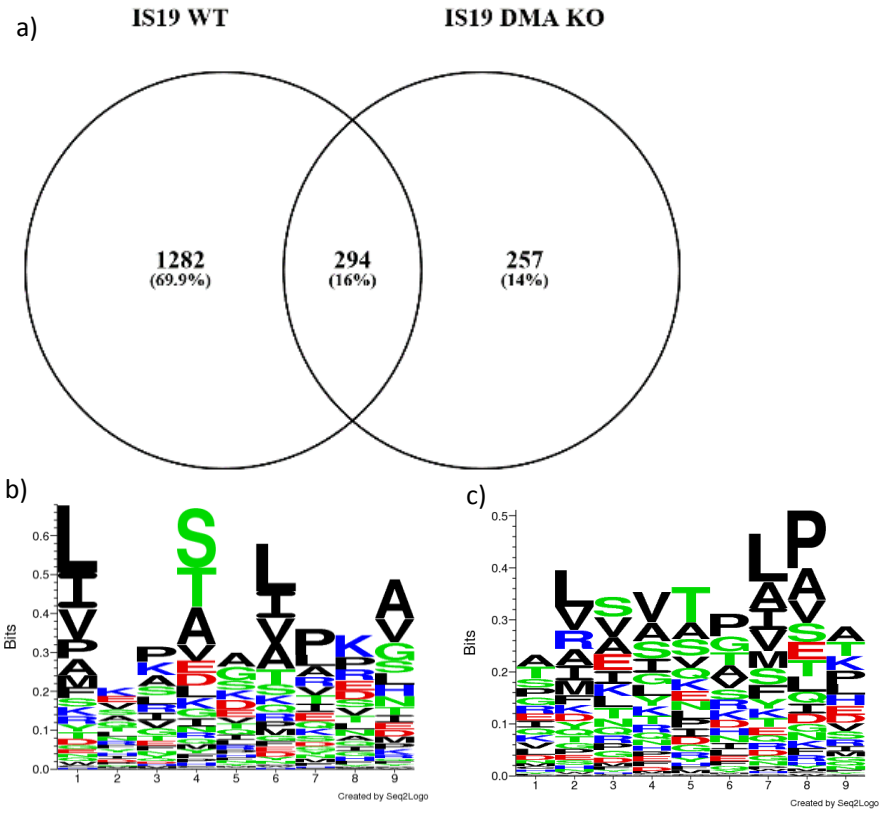


Figure 6.4.3 Comparison of Peptides derived from WT and DMA KO

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from WT and BLB2 KO were compared for shared peptide and unique peptide using a Venn diagram a). Peptides that are shared by the WT and the b) DMA KO and c) unique to the DMA KO were analysed by Gibbs clustering.

Comparable to both BLB KOs, a large proportion of peptides (69.9 %) are seen to be presented in the WT only (fig 6.4.3.a), although the DMA KO expresses both BLB1 and BLB2. This is likely due to reduced levels of BLB protein expression in the DMA KO (as shown in chapter five), which would mean that fewer peptides are being sampled. This notion is confirmed in figure 6.2.1 where it is seen that the DMA KO yielded significantly fewer

peptides than the WT, BLB1 and BLB2 KO lines. Again, unexpectedly a subset of peptides (14 %) from the DMA KO are not found in the WT.

This proportion of peptides that are unique to the DMA KO were analysed by Gibbs clustering (fig 6.4.3c). Unfortunately, these peptides failed to show a clear motif which could suggest that the peptides found to be unique to the DMA KO are diverse in sequence or that the number of unique peptides is too few for a clear motif to be discerned. It is interesting to note that the BLB1 KO has 252 peptides that are exclusive from the WT and this number of peptides is enough to see a clear motif (fig 6.4.1c). Similarly, the DMA KO has 257 peptides that are not shared with the WT but a clear motif cannot be identified (fig 6.4.3c), this suggests that the absence of DMA results in a greater diversity of peptide sequence.

Lastly, the BLB1 and BLB2 KO lines were compared for shared and unique peptides. Figure 6.4.4a shows a small number of peptides in common (7.6 %), with the majority of peptides being unique to the BLB1 KO (42.7 %) and the BLB2 KO (49.7 %), suggesting that a specific peptide repertoire is present by BLB1 and BLB2 as supported in figures 6.3.1 and 6.3.2. Gibbs clustering analysis of the BLB1 and BLB2 KO unique peptides (fig 6.4.4 c and d) shows two distinct motifs, these motifs are identical to what is seen in figures 6.3.1 and 6.3.2. The motif of the shared peptides was also assessed by Gibbs clustering and unfortunately, these peptides failed to show a clear motif (fig 6.4.4b). This suggests that the shared peptides between the BLB1 and BLB2 KO are diverse in their sequence or that the number of unique peptides is too small for a clear motif to be discerned.

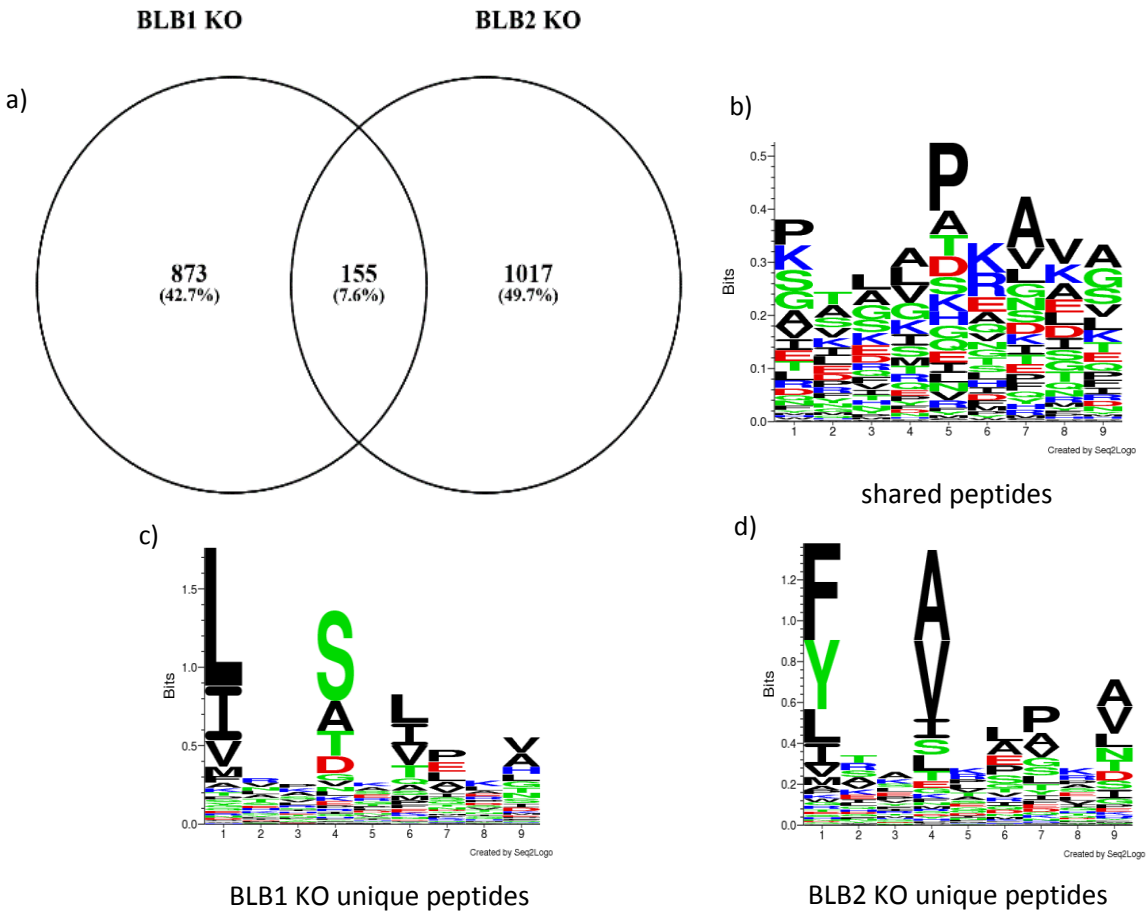


Figure 6.4.4 Comparison of Peptides derived from BLB1 and BLB2 KO

Class II KO cell lines; BLB1.2 cl-1, BLB2.2 cl-1, DMA.2 cl-1 and WT IS19 were lysed and class II molecules were isolated by immunoprecipitation, using the antibody 2G11. Samples were then prepared and analysed by LC-MS/MS, generating a list of de novo sequenced unique peptides. Peptides of 15-19 amino acids in length from BLB1 and BLB2 KO were compared for shared peptide and unique peptide using a Venn diagram a). Peptides that are shared between the b) BLB1 and the BLB2 KO , c) unique to the BLB1 KO and d) unique to the BLB2 KO were analysed by Gibbs clustering.

6.5 Discussion

Initial immunopeptidomic analysis of IS19 WT, BLB1 KO, BLB2 KO and DMA KO revealed many interesting differences between the peptide diversity and total number of peptides found for each line. The first observation was that peptides of 15-19 amino acids in length were the most frequent for all four cell lines (fig 6.2.1). It was also noted that there is an increase in peptides seen in all lines between 8-9 amino acids in length, the motif of these peptides is the same in all lines and has the motif expected for MHCI (data not shown). It is

suggested that during the immunoprecipitation of MHCII during the sample preparation some MHCI molecules are also isolated from the cell membrane.

The BLB1 and BLB2 KO lines have fewer peptide sequences than the WT, which is not unexpected as the WT expresses both BLB1 and BLB2. Interestingly, the DMA KO has significantly fewer peptide sequences than all other lines, although it expresses both BLB1 and BLB2. This is expected to be because the DMA KO has a severe reduction in BLB1 and BLB2 protein expression as discussed in chapter 5. Reduced expression of BLB would drastically change the amount of peptide sequences that are seen by immunopeptidomics. This experiment will be repeated to validate the results shown in this chapter. One adjustment is to increase the cell numbers used for the DMA KO to attempt to normalise the amount of BLB expression in each LC-MS/MS run. This adjustment may result in more peptide sequences for the DMA KO.

The peptide motif was assessed for peptides of 15-19 amino acids in length for each cell line. The clearest observations were that BLB1 and BLB2 have two distinct motifs, with clear differences being seen at positions one, four and six. The BLB1 KO line lacks phenylalanine and tyrosine at position one when compared to the BLB2 KO, at position four isoleucine is not seen in the BLB1 KO whereas aspartic acid is not seen the BLB2 KO, and at position six the BLB1 KO lacks glutamic acid and proline whereas the BLB2 KO lacks isoleucine, valine and tyrosine (fig 6.3.1 b and c). When the BLB1 and BLB2 KO motifs are compared to that of the WT, it is seen that the dominant motif of the WT is largely the same as the BLB1 KO, indicating that the BLB2 motif is the major component of the WT motif (fig 6.3.1a). This was expected as previous data shows that BLB2 is expressed 100X more than BLB1 at the RNA level in spleen and blood (Jacob et al 2000). Chapter 5 details protein expression of ~5X more BLB2 than BLB1. However, the BLB2 KO (BLB1) motif can still be seen in the WT motif, with the presence of phenylalanine and tyrosine at position one and valine at position four

(fig 6.3.1a). The presence of both the BLB1 and BLB2 KO motifs in the WT was supported by further clustering of the WT, revealing two motifs that are very similar to that of the BLB1 and BLB2 KOs (fig 6.3.2).

To further assess the similarities of the BLB1 and BLB2 KO peptides with the WT, Venn diagrams were used to illustrate how many shared and unique peptides were found in the BLB KOs compared to the WT. For both BLB KOs (fig 6.4.1 and 6.4.2); the peptides shared with the WT have the same motif as seen in figure 6.3.1 and 6.3.2. The most interesting result was that both BLB KOs had a proportion of peptides that are not seen in the WT, which could suggest a motif that is unique to the KO lines. Gibbs clustering of the peptides that are unique to the BLB KOs showed both BLB1 and BLB2 KO lines had a motif that is largely the same motif each BLB KO had for the peptides shared with the WT (fig 6.4.1 and 6.4.2). This shows there isn't a second motif that is expressed only in the BLB KO lines and the changes seen are in more variable positions of the peptide sequence and it is likely that these peptides are not seen in the WT because they are at low abundance and were not detected during the experiment. This hypothesis can be tested my repeating the experiment multiple times and looking to see if the unique peptides in the BLB KO lines are found in the WT of replicate experiments.

It was observed that the DMA KO line had a motif that was almost identical to that of the BLB1 KO line (fig 6.3.1b and d) and the amino acids associated with the BLB2 KO (BLB1 motif) were not present. It noted that the DMA KO had much more variation in all positions. This result could suggest that BLB1 has a higher dependence on DMA for peptide loading than that of BLB2, and in the absence of DMA BLB1 fails to bind the peptides that fit the BLB1 motif. This result could also be explained by the fact that the DMA KO has much less BLB1 and BLB2 protein expressed and BLB1 is expressed at relatively low levels when compared to BLB2 generally. This reduction in BLB1 could lead to BLB1 associated peptides not being

seen. Increasing cell numbers of the DMA KO to compensate for this reduced MHCII expression would elucidate which explanation is likely to be true.

The second observation is that the peptides unique to the DMA KO when compared to the WT have no discernible motif with lots of variation at every position (fig 6.4.3c). The first explanation of why no clear motif was found by Gibbs clustering is that the number of peptides unique to the DMA KO is relatively low (257) (fig 6.4.3a) and therefore cannot resolve a motif. Interestingly, the BLB1 KO has only 252 peptides that are not found in the WT but this number was still enough to clearly resolve a sequence motif by Gibbs clustering, suggesting that in the absence of DMA there is a subset of MHCII peptides that have much more sequence variation. This is consistent with the notion in mammals that DM functions to provide high affinity peptide to MHCII and without DM a broader peptide sequence diversity is seen. Further analysis needs to be carried out on our immunopeptidomics data, beginning with a comparison of the number of shared and unique peptides of the BLB1 and BLB2 KOs with the DMA KO and what motifs, if any, are seen. In the future these results need to be confirmed by replicating the experiment. In addition, the DMB1 and DMB2 KO line peptide repertoire will also be assessed to understand how individual DM KOs effect the peptide repertoire of BLB1 and BLB2 and if the two DMs have differing effects on BLB.

7. Understanding TAP transporter-mediated peptide transport in the chicken

7. Understanding TAP transporter-mediated peptide transport in the chicken

7.1 Introduction

The TAP transporter is comprised of a heterodimer of TAP1 and TAP2 and is known to be the key player in the transport of cytosolic peptides into the lumen of the ER. This peptide transport provides most peptides that are subsequently loaded onto the MHC class I for CD-8 t-cell recognition (as discussed in detail in the general introduction). Despite the importance of the TAP complex in immunity and its critical role in class I presentation, much is still unknown about the function and mechanisms of peptide mediated TAP transport. Many studies have proposed mechanisms of how the TAP transporter physically transports peptides from the cytosol to the ER and what changes in confirmation are required, but the question of which residues within the TAP membrane spanning domain (MSD) are require for specific peptide transport is still largely unknown.

TAP1 and TAP2 are functionally monomorphic in humans and mice (Trowsdale et al 2011), which has been a major issue when trying to understand which residues are important for peptide binding to the TAP complex as there aren't obvious residues under selection that could be proposed as candidates. Besides taking the approach of randomly mutating a selection of residues in the MSD or cross-linking peptides, a means of identifying potential residues of interest would need to be devised. The chicken however, has the unique property (compared to mammals) that the TAPs are polymorphic and as previously discussed are thought to have co-evolved with a specific MHCI gene, with the TAPs tailoring the peptides pumped for a specific MHC I allele (Walker et al 2011). This TAP polymorphism is assumed to have occurred by the selection of residues that determine the specificity of peptides pumped by the TAP transporter and potentially highlights candidate residues for peptide transport. To further narrow down candidate residues, the polymorphic amino acids

were overlaid on to a model of human TAPs, based on the structure of a bacterial ABC transporter. Polymorphic residues that fit with peptide translocation specificity were selected for mutagenesis. They can then be tested for their effect on the transport of peptides known to transport that fit the binding motif of chicken class I for a given haplotype (Jim Kaufman, unpublished data). It was further hypothesised that the polymorphic residues found in the chicken when mapped onto the bacterial ABC transporter could also be used to infer which residues may be important for peptide transport in the human and mouse when compared.

Our lab previously attempted to look at the chicken residues of interest by using insect cells to express both chicken TAP1 and TAP2, isolate microsomes and use the microsomal fraction for radiolabelled peptide transport assays. This method yielded mixed results and was deemed to be unsuitable for the study. This chapter details the initial work to develop a system capable of answering the question of which TAP residues determine the specificity of peptide transport. With current advances in genome editing we decided to utilise CRISPR-Cas9 to make TAP1 KO, TAP2 KO and TAP1 and TAP2 double KO (referred to as TAPD KO) lines, in the characterised IS19 cell line and to assess whether transport is abolished in single TAP and double TAP KO lines.

7.2 Generation and screening of TAP1, TAP2 and TAPD KO lines

Two sgRNAs were designed to target largely conserved regions of TAP1 and TAP2, so the guides could be used in other REV cell lines of different MHC haplotypes. Details of the guides can be found in the material and methods section. Two guides targeting TAP1 and two targeting TAP2 were separately transfected into IS19 cells giving four transfected lines, TAP1.1, TAP1.2, TAP2.1 and TAP2.2. The four lines were cell sorted for high GFP expression, GFP positive polyclonal populations were expanded and class I levels of these polyclonal populations were analysed by flow cytometry

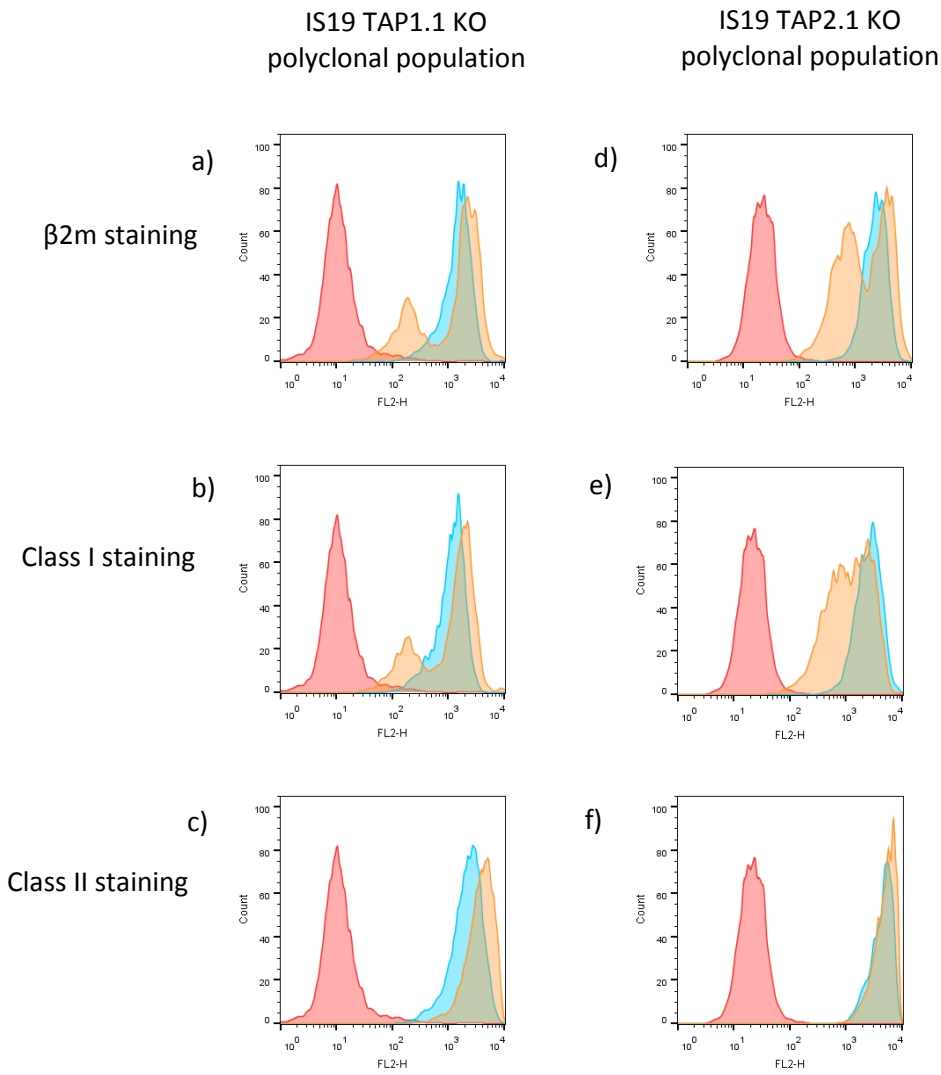


Figure 7.2.1 Flow cytometry cell surface staining of B2m and class I heavy chain in IS19 TAP1 and TAP2 polyclonal populations.

IS19 cells were separately transfected with a sgRNA:cas9 plasmid targeted to a genomic DNA sequence of TAP1 and TAP2. Cell pools were analysed via flow cytometry using the antibody F21-21 and F21-2, to β2m and class I heavy chain respectively. IS19 cells transfected with PX458 TAP1 sgRNA1 (a, b & c) and PX458 TAP2 sgRNA1 (d, e & f). Transfected cells were stained for β2m (a & d), class I heavy chain (b & e) and class II β-chain (c & f) in orange and compared to wild type cell expression (blue). All histograms include transfected cells stained with secondary only (red).

Figure 7.2.1 shows flow cytometry staining of TAP1.1 and TAP2.1 polyclonal KO populations for β2m, class I heavy chain and class II β chains. In the TAP1.1 KO polyclonal population stained for β2m and class I (fig 7.2.1 a & b, orange) we see two distinct populations one that has largely normal β2m and class I expression when compared to wild type expression (blue) and one that has over a log decrease in cell surface expression. A similar result was seen for the TAP2.1 KO polyclonal population. These results were in accordance with previous

findings in human and murine studies that also show that class I expression is decreased when one or both TAP molecules loss function due to mutation (Van Kaer et al 1992, de la Salle et al 1993). Additionally, both TAP KO polyclonal lines were stained for class II expression to check that this reduction was restricted to MHC class I, rather than being a result of greater loss in the MHC by a large deletion. Figure 7.2.1c and f show class II expression in the TAP KOs (orange) remains largely unchanged compared to wild type expression (blue). These initial findings gave confidence that our targeting of TAP was successful and give a phenotype that could be used for cell sorting to hasten the process of making clonal derived TAP KO lines whilst also showing that the lack TAP presentation effects MHC class I cell surface expression in the chicken.

The next task was to isolate TAP1 KO and TAP2 KO clonal lines for further experimentation. It is noted in figure 7.2.1 that the TAP1 KO population is smaller than in TAP2. This difference is likely to be due to differences in transfection efficiency and/or the efficiency of editing after transfection. We decided to use the low class I phenotype in the TAP1 KO population as a target for cell sorting to enrich for class I KO cells.

TAP1 KO cells with low class I were sorted as single cells into a 96 well plate, whilst TAP2 KO cells were cloned into 96 well plates by limiting dilution. Cell lines were then screened by PCR, cloning and sequencing for mutations in the relevant TAP. In addition, wild type IS19 cells were simultaneously transfected with sgRNAs directed to TAP1 (TAP1.1) and TAP2 (TAP2.1) in order to create a double TAP KO line.

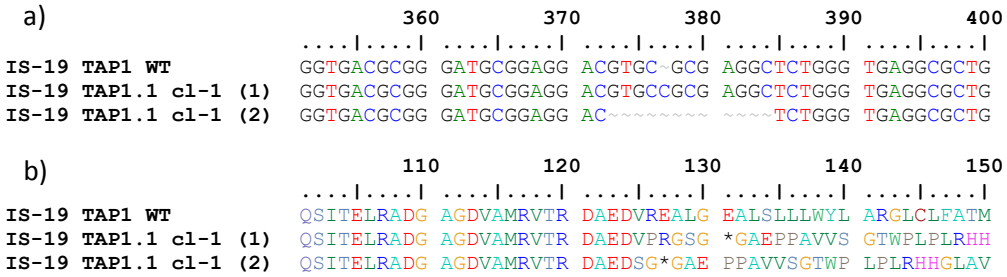


Figure 7.2.2 DNA and amino acid sequence alignment of clonal line TAP1.1 cl-1 shows editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 3 of TAP1 were mapped onto a partial TAP1 cDNA sequence and compared to the wild type. Editing events can be seen in both alleles immediately before the PAM site (red line) at position 381-383. b) The effect of these editing events on the whole amino acid sequence are shown for each of the two edited alleles. Changes in the coding sequence begin at position 125 and 126, premature stop codons can be seen downstream of these positions.

The three sequences are labelled on the left-hand margin; wild type sequence (IS19 TAP1 WT), the first edited allele of clonal line TAP1.1 cl-1 (IS19 TAP1.1 cl-1 [(1)] and the second allele (IS19 TAP1.1 cl-1 [(2)].

The selected TAP1 KO clonal line TAP1.1 cl-1, has two successful editing events that can be seen slightly downstream of the PAM site at position 381-383, the edits consist of a one base pair insertion on one allele and an eleven base pair deletion on the second allele (fig 7.2.2a).

The effect of the editing on the coding sequence when compared to the original sequence showed disruptions to the gene begin at amino acid positions 126 and 125 for the two different edits. These changes in codon frame result in an early stop codon at positions 131 and 127 terminating the protein (fig 7.2.2b).

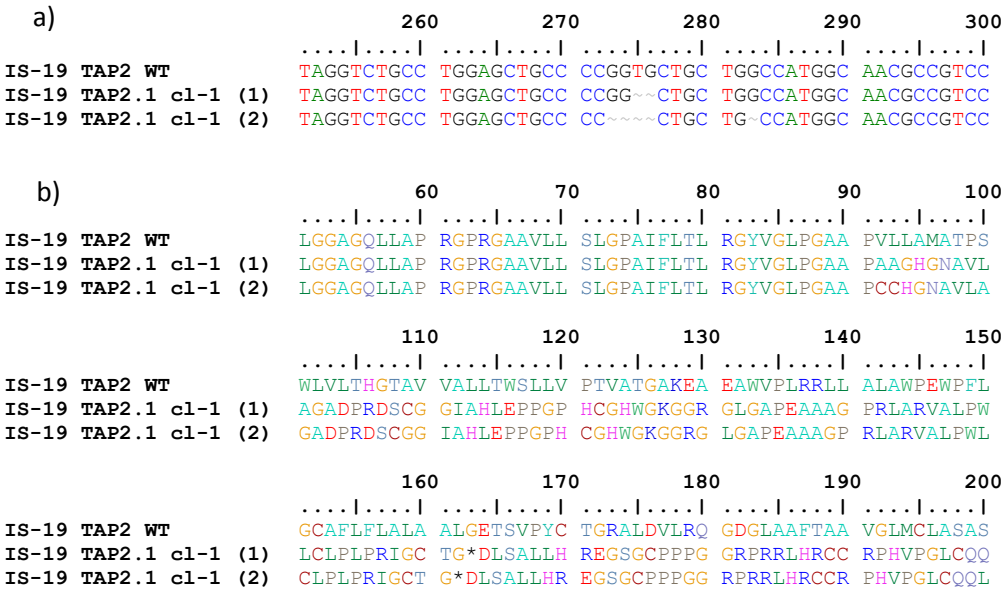


Figure 7.2.3 DNA and amino acid sequence alignment of clonal line TAP2.1 cl-1 shows editing events and the generation of a premature stop codon.

a) The editing events revealed by sequencing exon 1 of TAP2 were mapped onto partial TAP2 cDNA sequence and compared to the wild type. Editing events can be seen in both alleles immediately before the PAM site at position 269-271. b) The effect of these editing events on the whole amino acid sequence were then shown for each of the two edited alleles. Changes in the coding sequence begin at position 92, premature stop codons can be seen downstream of these positions.

The three sequences are labelled on the left-hand margin; wild type sequence (IS19 TAP2 WT), the first edited allele of clonal line TAP2.1 cl-1 (IS19 TAP2.1 cl-1 [(1)]) and the second allele (IS19 TAP2.1 cl-1 [(2)]).

The selected TAP2 KO clonal line TAP2.1 cl-1 has two successful editing events which can be seen immediately downstream of the PAM site at position 269-271, the edits consist of a two base pair deletion on one allele and a five base pair deletion on the second allele (fig 7.2.3a).

The effect of the editing on the coding sequence when compared to the wild type sequence show disruptions to the gene begin at amino acid position 92 for the two different edits. These changes in codon frame result in an early stop codon at positions 163 and 162 terminating the protein (fig 7.2.3b).

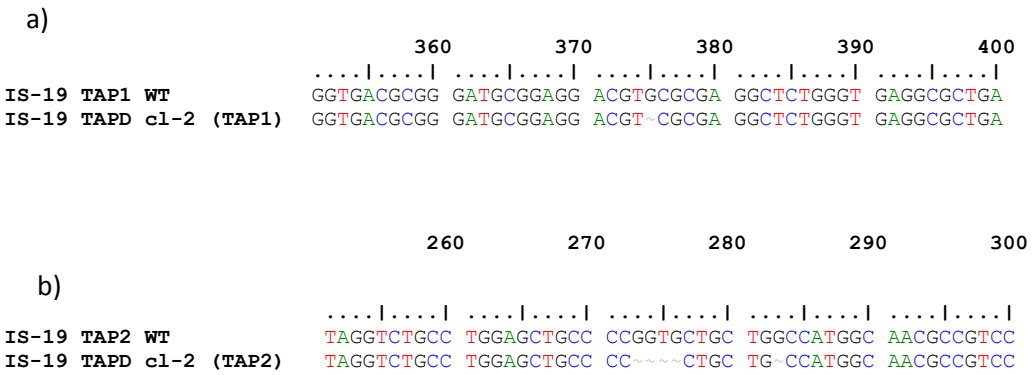


Figure 7.2.4 DNA sequence alignment of clonal line TAPD cl-2 shows editing events and in TAP1 and TAP2.

The editing events revealed by sequencing exon 3 of TAP1 and exon 1 of TAP2 were mapped onto partial TAP1 and TAP2 cDNA sequence respectively and compared to the wildtype. a) editing events can be seen in TAP1 immediately before the PAM site at position 381-383. b) editing events can be seen in TAP1 immediately before the PAM site at position 269-271.

Sequences are labelled on the left-hand margin; wild type sequences (IS19 TAP1 WT or IS19 TAP2 WT), the edited alleles of TAPD cl-2 are labelled IS19 TAPD cl-2 (TAP1) for TAP1 or TAD cl-2 (TAP2) for TAP2.

The TAP double KO clonal line TAPD cl-2 shows successful editing in TAP1 which consists of a one base pair deletion immediately downstream of the PAM site at position 381-383 (fig 7.2.4a).

Editing can also be seen in TAP2 downstream of the PAM site a position 269-271; the edits consist of a five base pair deletion (fig 7.2.4b). Both edits appear to be same on both alleles for TAP1 and TAP2, these edits were confirmed by cloning and sequencing of two independent PCRs and direct sequencing of the PCR product.

The effect of the TAP1 editing on the coding sequence when compared to the original sequence show disruptions to the gene begin at amino acid position 126 and results in a premature stop codon at position 133 (fig 7.2.5). The TAP2 editing shows disruptions to the amino acid sequence at position 92 and a premature stop codon is seen a position 162.

a)

```
                                110        120        130        140        150
                       ....|....| ....|....| ....|....| ....|....| ....|....|
IS-19 TAP1 WT          QSITELRADG AGDVAMRVTR DAEDVREALG EALSLLLWYL ARGLCLFATM
IS-19 TAPD cl-2 (TAP1) QSITELRADG AGDVAMRVTR DAEDVARLWV RR*ASCCGIW HVASASSPPW
```

b)

```
                                 60         70         80         90        100
                       ....|....| ....|....| ....|....| ....|....| ....|....|
IS-19 TAP2 WT          LGGAGQLLAP RGPRGAAVLL SLGPAIFLTL RGYVGLPGAA PVLLAMATPS
IS-19 TAPD cl-2 (TAP2) LGGAGQLLAP RGPRGAAVLL SLGPAIFLTL RGYVGLPGAA PCCHGNAVLA

                                110        120        130        140        150
                       ....|....| ....|....| ....|....| ....|....| ....|....|
IS-19 TAP2 WT          WLVLTHGTAV VALLTWSLLV PTVATGAKEA EAWVPLRRLL ALAWPEWPFL
IS-19 TAPD cl-2 (TAP2) GADPRDSCGG IAHLEPPGPH CGHWGKGGRG LGAPEAAAGP RLARVALPWL

                                160        170        180        190        200
                       ....|....| ....|....| ....|....| ....|....| ....|....|
IS-19 TAP2 WT          GCAFLFLALA ALGETSVPYC TGRALDVLRQ GDGLAAFTAA VGLMCLASAS
IS-19 TAPD cl-2 (TAP2) CLPLPRIGCT G*DLSALLHR EGSGCPPPGG RPRRLHRCCR PHVPGLCQQL
```

Figure 7.2.5 Amino acid sequence alignment of clonal line TAPD cl-2 shows effects of editing on TAP1 and TAP2 coding sequence.

The effect of the DNA editing of TAP1 and TAP2 were shown on the entire amino acid sequence. a) TAP1 edits show changes in the coding sequence begin at position 126, premature stop codons can be seen down stream of these positions. b) TAP2 edits show changes in the coding sequence begin at position 92, premature stop codons can be seen down stream of these positions.

Sequences are labelled on the left-hand margin; wild type sequences (IS19 TAP1 WT or IS19 TAP2 WT), the edited alleles of TAPD cl-2 are labelled IS19 TAPD cl-2 (TAP1) for TAP1 or TAD cl-2 (TAP2) for TAP2.

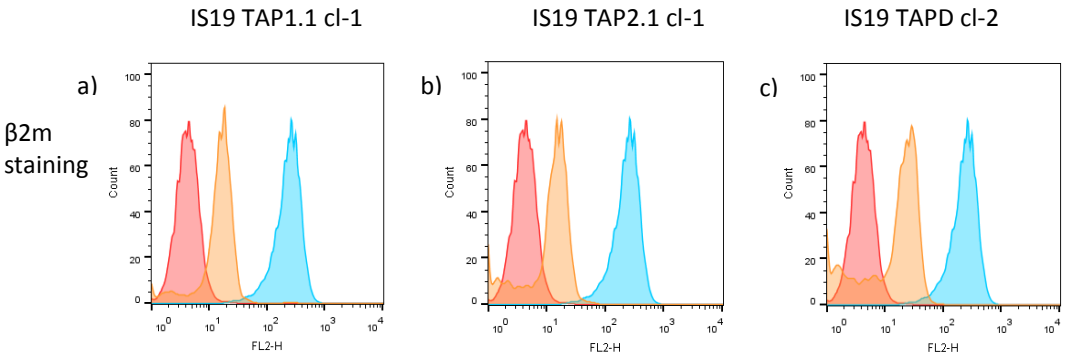


Figure 7.2.6 Flow cytometry cell surface staining of β2m in IS19 TAP1, TAP2 and TAPD KO clonal lines.

Clonal KO lines were derived and screened by sequencing for TAP1, TAP2 and TAPD. These cell lines were analysed via flow cytometry using the antibody F21-21, to β2m. B2m expression is seen in orange for TAP1 KO (a), TAP2 KO (b) and TAPD KO (c), compared to wild type expression (blue). All histograms include transfected cells stained with secondary only (red).

The selected KO clones for TAP1, TAP2 and TAPD were analysed for β2m expression by flow cytometry. Clear reductions of β2m can be seen in all KO lines, with reductions being of the same magnitude, just over a one log reduction. All KO cell lines appear to have one peak for

β2m expression indicating that all cells have impaired class I presentation, something which we would not expect to see if the cell line had unedited alleles or a secondary population of cells that had not been edited. The flow cytometry results are as expected based on the results obtained from the cloning and sequencing of these lines.

Lastly, Western blotting of the TAPs was attempted to confirm that no TAP protein was being produced in the KO lines. This proved difficult, especially when blotting for TAP1 as the western blots were often too messy to decipher. Blots of TAP2 were slightly better but again often too messy for interpretation. Figure 7.2.7 shows screening of TAP2 KO and TAPD KO clone lysates by western blotting for TAP2. TAP2 has a theoretical molecular weight of ~75 kDa and can be seen in the wild type IS19 cell line in both blots (figure 7.2.7). The absence of the TAP2 band is seen in the TAP2 KO (figure 7.2.7a) and the TAPD KO (figure 7.2.7b), as expected from the flow cytometry and sequencing results.

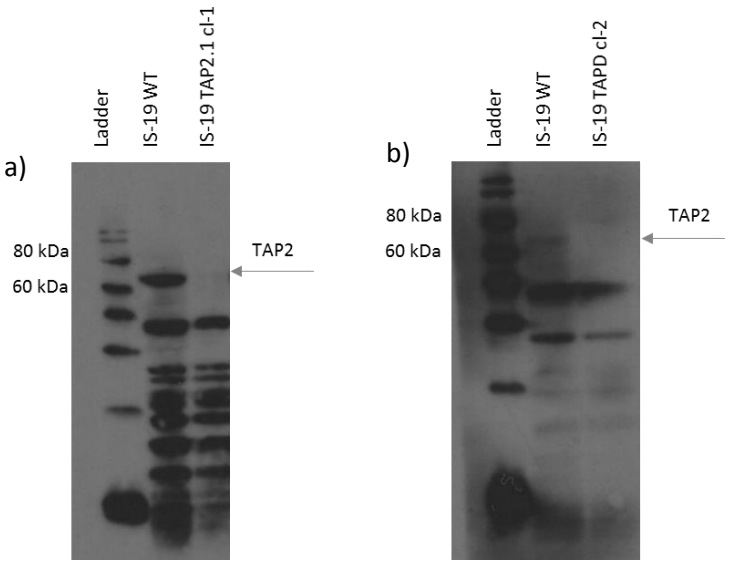


Figure 7.2.7 Western blots staining for TAP2 in IS19 TAP2 and TAPD KO clonal lines.

IS19 WT cells and clonal derived cell lines TAP1.1 cl-1 and TAPD cl-2 were lysed using digitonin after a membrane enriching protocol. The clone lysates were probed for TAP2 protein expression by western blotting using the antibody F1-3.

7.3 Assessing peptide transport in TAP KO lines by radiolabelled peptide transport assays

Before all clonal lines were derived and fully confirmed for all required TAP KO lines, transport assays were undertaken to assess the characterised clonal lines but also as part of the screening process for TAPD KO lines.

The first transport assay (fig 7.3.1) shows transport and inhibition with the peptide KRYNASAY (P198), a peptide that has previously been shown in the lab to transport effectively in IS19 cells (Walker et al 2011). IS19 WT shows a high level of radioactivity as shown by the counts per minute (Y axis). This transportation is greatly inhibited in the addition of excess unlabelled P198 abolishing the transport. The TAP KO clones used at the time of this experiment were TAP1.1 cl-1 (the characterised and selected TAP1 KO line), TAP2.1 cl-4 and TAPD cl-1, neither of which are the selected cell lines for TAP2 KO and TAPD KO detailed and characterised above.

TAP1.1 cl-1 showed a ~75% reduction in peptide transport with a further ~25% that could be inhibited with cold peptide. TAP2.1 cl-4 showed a ~50% reduction in transport, again that could be fully inhibited. TAPD cl-1 showed a ~75% reduction in peptide transport with the remaining transport being fully inhibited. These results had two unexpected outcomes; firstly, that transport is seen in all KO lines between 25 and 50%, as human and mouse data suggests that when one of the two TAP proteins are removed, transport is abolished. Secondly, that the TAP2 KO line shows 50% transport compared to 25% in the TAP1 KO line.

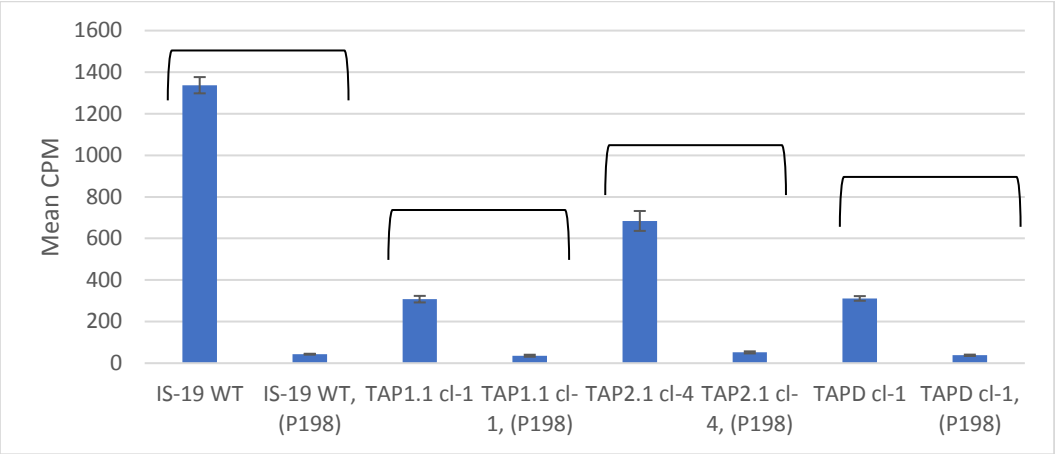


Figure 7.3.1 Radiolabelled transport and inhibition assay of P198 peptide initial IS19 TAP1, TAP2 and TAPD KO clonal lines.

IS19 WT cells and clonal derived cell lines TAP1.1 cl-1, TAP2.1 cl-4 and TAPD cl-2 were assayed for the transport of ^{125}I radiolabelled P198 peptide (KRYNASAY) and inhibited with excess unlabelled P198. Radioactive material present is represented as mean counts per minute (CPM) as assessed by γ-counter. Each condition is performed in triplicate with error bars representing standard error of the mean. X-axis shows labelling of cell line used and the peptide number used for inhibition in brackets.

The differences in transport between TAP1 and TAP2 KO lines was quickly understood, as upon further screening it was seen that one allele of the TAP2.1 cl-4 KO line had an indel that was a multiple of three nucleotides which did not result in a frame shift mutation and presumably did not result in loss of function. Also, the TAPD cl-1 line was found to be a TAP2 only KO and was renamed TAP2.1 cl-1 in section 7.2, but for the remainder of section 7.3 will continue to be called TAPD cl-2. Further screening was required to identify a TAPD KO, as well as continued transport assay analysis of the identified TAP KO lines TAP1.1 cl-1 (TAP1) and TAPD cl-1 (TAP2).

Due to the unexpected result of seeing some transport in TAP KO lines we wanted to determine if this transport could be specifically inhibited. To test this, we used radiolabelled P198, which was either inhibited with unlabelled P198 or the B4 peptide ADYNDSAE (P167). Simultaneously, an aliquot of the TAP cell lines were stained for class I heavy chain expression (fig 7.3.2).

In figure 7.3.2a, we observed that IS19 wild type showed expected transport with P198. This transport was fully inhibited with cold P198 but was barely inhibited by the B4 peptide (P167) suggesting specific transport through the TAPs. The exact same pattern was observed with the TAP1 and TAP2 KO lines (TAP1.1 cl-1 and TAPD cl-1), with transport of radiolabelled P198 being ~25% of the WT that is fully inhibited with unlabelled P198 but not inhibited by P167. Three additional potential TAP KO lines were also tested (TAPD cl-2, cl-4 and cl-5). After screening by sequencing, these new TAP KO clones were seen to be a TAPD KO (TAPD cl-2), a TAP1 KO (TAPD cl-4) and wild type (TAPD cl-5). The newly isolated TAPD KO clone behaved as expected and similarly to the TAP1 and TAP2 single KO lines, with ~25% of WT transport which is inhibited by P198 but not P167. The additional TAP1 KO (TAPD cl-4) behaved as TAP1.1 cl-1, and the clonal line that was found to be WT (TAPD cl-5) showed the same phenotype as WT IS19 cells (fig 7.3.2a).

The transport assay data was then compared with the class I heavy chain staining of the TAP KO clones (fig 7.3.2). IS19 wild type staining was the same when compared with TAPD cl-5 which had no edits in the TAPs. The two TAP1 KOs (TAP1.1 cl-1 and TAPD cl-4) showed the same ~90% reduction of class I. The TAP2 KO showed a ~85% reduction in class I expression and the TAPD a ~95% reduction in class I (fig 7.3.2.b).

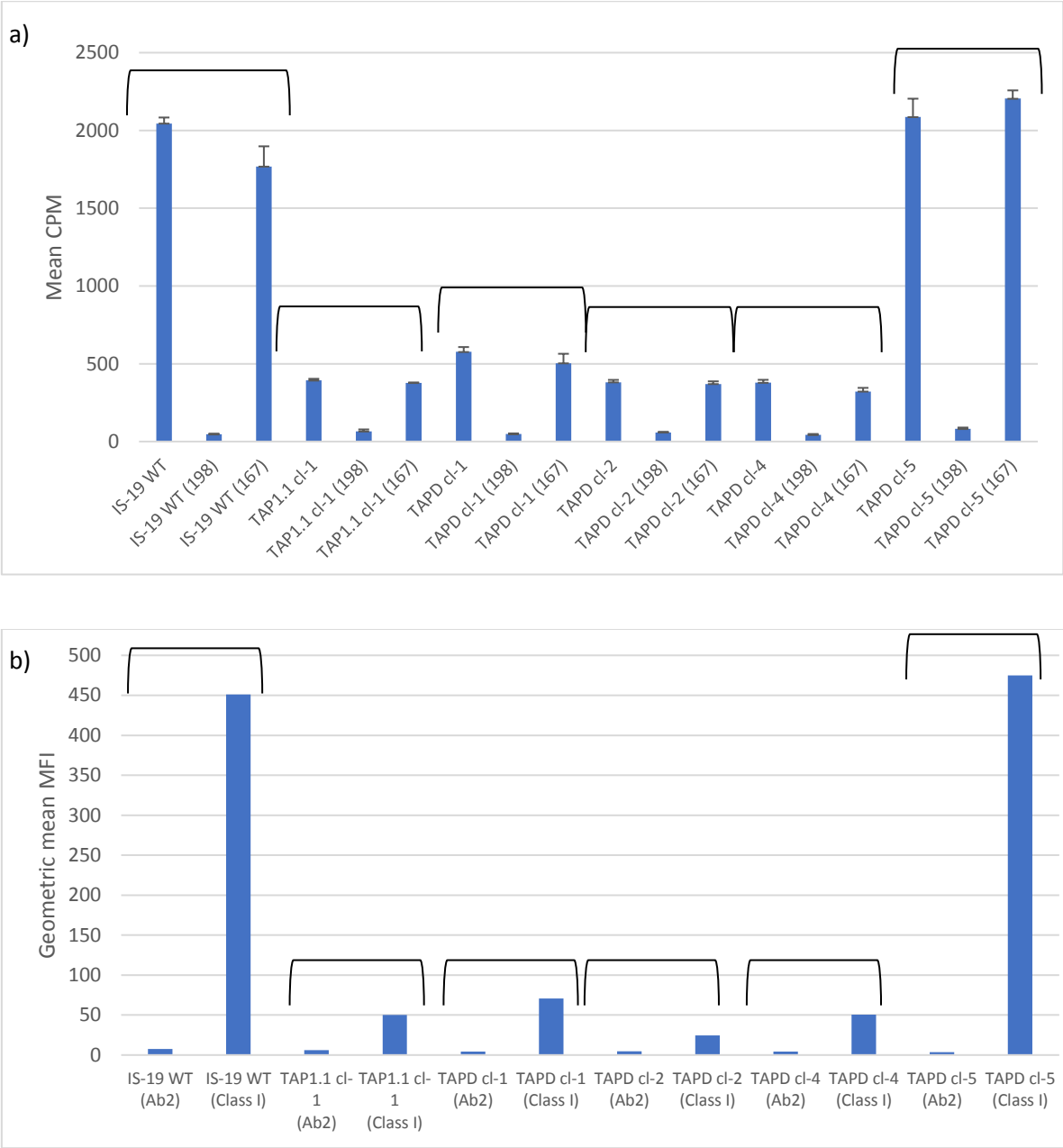


Figure 7.3.2 Transport of radiolabelled P198 peptide in initial IS19 TAP1, TAP2 and TAPD KO clonal lines with P198 and P167 inhibition and flow cytometry staining of class I heavy chain.

a) IS19 WT cells and clonal derived cell lines TAP1.1 cl-1, TAPD cl-1, 2, 4 and 5 were assayed for the transport of ^{125}I radiolabelled P198 peptide (KRYNASAY). Inhibition was performed with excess unlabelled peptide, either P198 or P167 (ADYNDSAE). Radioactive material present is represented as mean counts per minute (CPM) measured by γ-counter. Each condition was performed in triplicate with error bars representing standard error of the mean. b) aliquots of the above cell lines were taken before preparation for transport assays and stained for class I heavy chain expression using the antibody F21-2 for flow cytometry analysis. Fluorescence is plotted as the geometric mean fluorescence intensity (MFI). a) X-axis shows labelling of cell line used and the peptide number used for inhibition in brackets. b) X-axis shows labelling of cell line used and how the sample was stained in brackets, secondary antibody only (Ab2) and F21-2 (class I).

To begin to understand the ~25% of the WT transport that is in the TAP KO lines, the initial tests examined the possibility that peptides are being transported by an additional transporter or that the transport seen is due to peptides binding MHCI molecules. To test if the peptides capacity to bind MHCI effected inhibition of transport the transport assay in figure 7.3.2 was repeated using just the characterised and selected TAP1, TAP2 and TAPD clones (detailed in section 7.2), but additionally inhibited with the peptide RRRREQTVY (P149), that was shown to bind to IS19 MHCI by refolding but is not transported as assessed by transport assay (Tregaskes et al, 2016).

Figure 7.3.3 shows transport of radiolabelled P198 that is fully inhibited by unlabelled P198, a small amount of inhibition is seen with P167 and P149 acts similarly to P167. All TAP KO lines show the same pattern in phenotype, with TAP KO having ~12.5 - 25% reduction in peptide transport of radiolabelled P198 that is fully inhibited by unlabelled P198 but not inhibited at all by P167 or P149.

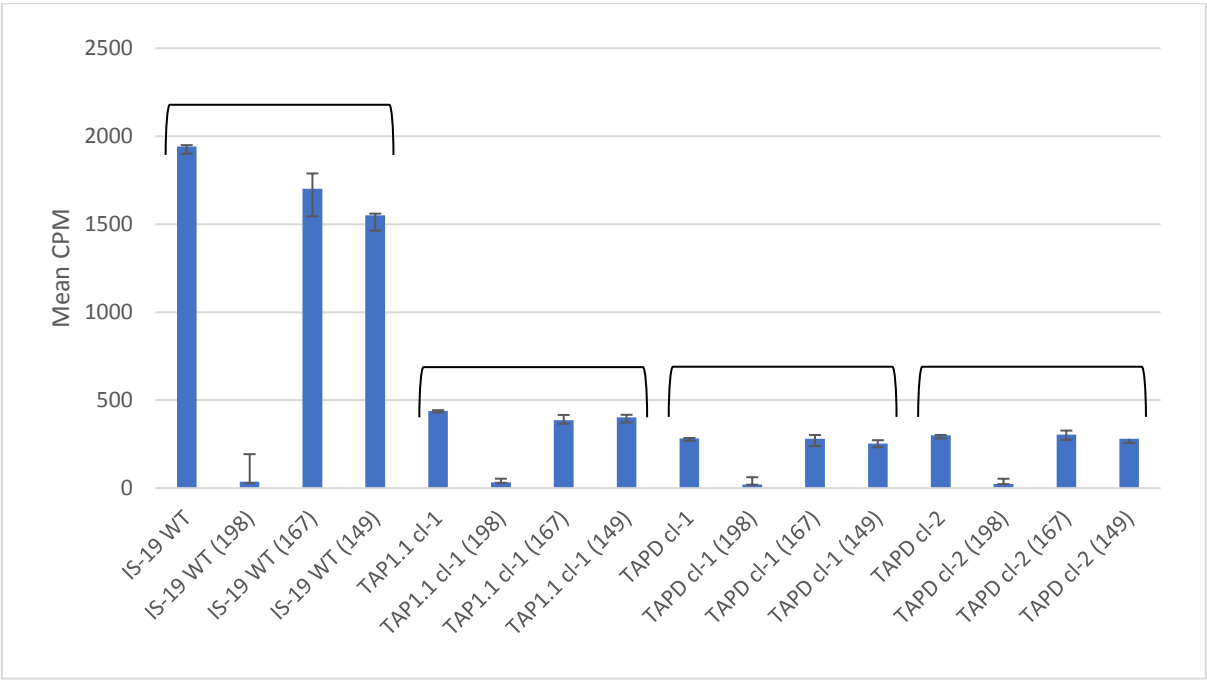


Figure 7.3.3 Transport of the radiolabelled P198 peptide in selected IS19 TAP1 KO, TAP2 KO and TAPD KO clonal lines, with P198, P167 and P149 inhibition.

IS19 WT cells and selected clonal derived cell lines TAP1.1 cl-1, TAPD cl-1 and 2 were assayed for the transport of ^{125}I radiolabelled P198 peptide (KRYNASAY). Inhibition is performed with excess unlabelled peptide, either P198, P167 (ADYNDSAE) or P149 (RRRREQTVY). Radioactive material present is represented as mean counts per minute (CPM) measured by γ-counter. Each condition is performed in triplicate with error bars representing standard error of the mean. X-axis shows labelling of cell line used and the peptide number used for inhibition in brackets.

7.4 Creation of B2m and TAP KO cell lines for transport assay analysis.

To understand the residual ~25% transport of the WT in the TAP KO cell lines, β2m was knocked out in the TAP KO lines to see whether the transport still remained. The selected TAP KO cell lines, TAP1 KO (TAP1.1 cl-1) TAP2 KO (TAPD cl-1, renamed to TAP2.1 cl-1) and TAPD KO (TAPD cl-2), were transfected with the β2m targeting sgRNA B2m.1 (detailed in chapter 3). Clonal lines were derived and initially screened by flow cytometry of β2m and class I heavy chain expression. Figure 7.4.1a shows the MFI of two TAP1+β2m KO clones (cl-1 & 3). IS19 WT expression of β2m and class I heavy chain can be seen compared to the two TAP1+β2m KO clones. Both KO lines show β2m and class I expression at the level of the secondary antibody control, providing strong evidence that these clones are β2m KOs. These clones were confirmed to be β2m KOs by cloning and sequencing (data not shown). Figure 7.4.1b, shows three TAP2+β2m KO clones (cl-1, 2 & 3). Again, all KO clones showed β2m and class I expression at the same level of the secondary antibody control, providing strong evidence that the clones are β2m KOs. These clones were also confirmed to be β2m KOs by cloning and sequencing (data not shown). The same process was used make TAPD+β2m KO clonal lines.

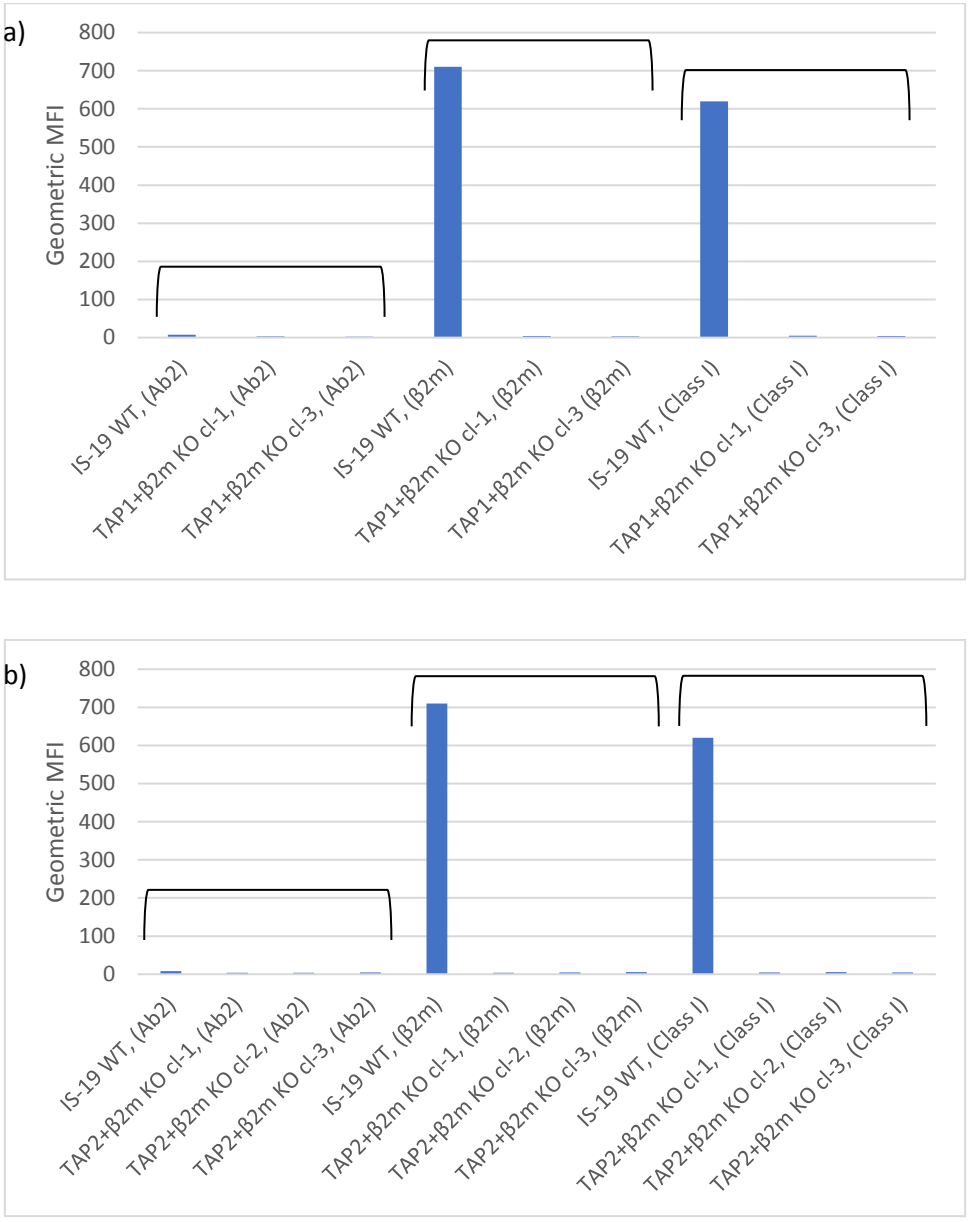


Figure 7.4.1 Screening of TAP1+B2m and TAP2+B2m KO clonal lines by flow cytometry.

a) Clonal derived cell line TAP1.1 cl-1 was transfected with the β2m targeting sgRNA B2m.1 and clonal lines were derived. Clones TAP1+β2m cl-1 and 3 were stained for B2m (F21-21) and class I heavy chain (F21-2) expression.
b) Clonal derived cell line TAP2.1 cl-1 was transfected with the β2m targeting sgRNA B2m.1 and clonal lines were derived. Clones TAP2+B2m cl-1, 2 and 3 were stained for β2m (F21-21) and class I heavy chain (F21-2) expression.
Fluorescence is plotted as the geometric mean fluorescence intensity. X-axis shows labelling of cell line used and how the sample was stained in brackets, secondary antibody only (Ab2), F21-21 (β2m) and F21-2 (class I)

Once the three new clonal lines (TAP1+β2m, TAP2+β2m and TAPD+β2m) were created, all the clonal cell lines were analysed together to check if the transport is dependent on MHCI expression and in conjunction an aliquot of each cell line was taken and stained for β2m expression by flow cytometry.

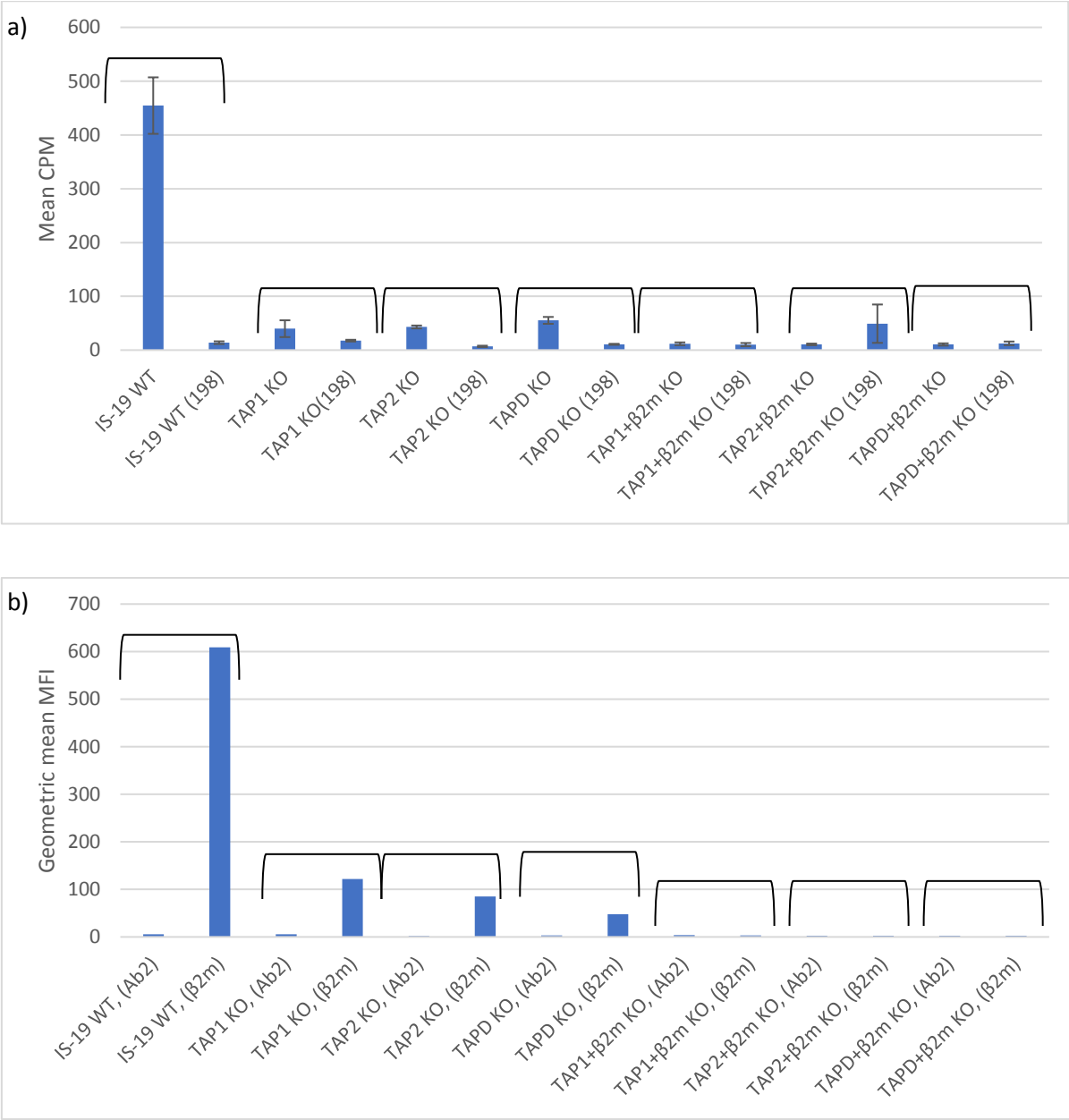


Figure 7.4.2 Transport of radiolabelled P198 peptide in IS19 TAP1 KO, TAP2 KO and TAPD KO with and without β2m KO clonal lines with P198 and P167 inhibition and flow cytometry staining of class I heavy chain.

a) IS19 WT cells and clonal derived cell lines TAP1, TAP2, TAPD KOs and TAP1+β2m, TAP2+β2m and TAPD+β2m KOs were assayed for the transport of ^{125}I radiolabelled P198 peptide (KRYNASAY). Inhibition is performed with excess unlabelled P198 peptide. Radioactive material present is represented as mean counts per minute (CPM) measured by γ-counter. Each condition is performed in triplicate with error bars representing standard error of the mean. X-axis shows labelling of cell line used and the peptide number used for inhibition in brackets. b) Aliquots of the above cell lines were taken before preparation for transport assays and stained for β2m expression using the antibody F21-21 for flow cytometry analysis. Fluorescence is plotted as the geometric mean fluorescence intensity. X-axis shows labelling of cell line used and how the sample was stained in brackets, secondary antibody only (Ab2) and F21-21 (β2m).

Figure 7.4.2a shows the transport assays for the TAP KO lines compared to the TAP+β2m KO lines. All cell lines are able to transport P198 and inhibited with P198. The wild type cell line shows decreased counts compared to the previous transport assays, but this was expected as the radiolabelled iodine had been through multiple half life cycles, but the transport is fully inhibited by unlabelled P198. Results for the TAP1 KO, TAP2 KO and double TAP KO lines show a similar pattern to previous results (fig 7.3.2, 7.3.3) with residual transport being seen in all three lines that can be inhibited by unlabelled peptide. Although, it is noted that the residual transport seen is at ~10% of the wild type compared to ~25% previously. When we compare this result to that of the three TAP+β2m KO lines, we see that the residual transport is completely abolished in the absence of MHCI.

7.5 Discussion

TAP sequences were analysed for guide sites that were largely conserved in TAP sequences for the B2, B4, B12, B15, B19 and B21 haplotypes. Initially, two guides were selected for TAP1 and TAP2. These guides were then cloned in PX458 for testing in the cell line IS19. The scientific literature describing human and mouse TAP KO lines shows that cell surface class I levels are reduced, when either or both TAPs are knocked out (Van Kaer et al 1992, de la Salle et al 1993). We decided to see if this was true for the chicken, so that we could use cell surface MHCI to identify successful TAP KOs.

Figure 7.2.1 shows the identification of two TAP-targeted sgRNAs (TAP1.1 & TAP2.1) that result in populations of reduced class I when transfected into IS19 cells. The IS19 cells were transfected with either TAP1.1 or TAP2.1, then sorted for GFP expression, and the polyclonal GFP positive population was expanded before being stained for MHCI expression. Both samples showed a distinct population of cells that have reduced β2m and class I heavy chain.

To check that this down regulation of MHC is restricted to MHCI molecules; transfected cells were also stained for MHCII, which had one clear population that was comparable to the WT. As the TAP targeting appeared to be a success, clonal cell lines were then derived from the transfected cells and analysed for edits in the appropriate TAP by PCR, cloning and sequencing.

Multiple clonal derived lines for TAP1, TAP2 and TAPD KO were screened, with one successful KO clonal being selected for TAP1, TAP2 and TAPD for further analysis. Figures 7.2.2 -7.2.4 show sequence analysis of the selected TAP KO lines (TAP1, TAP2 and TAPD). To confirm the reduced expression of class I the three KO cell lines were stained for β2m protein expression at the cell surface. Figure 7.2.6 shows a large reduction of β2m as seen in the polyclonal KO populations. To fully validate that these lines were indeed TAP KOs the KO lines were probed for TAP1 or TAP2 protein expression by western blotting. Figure 7.2.7 shows the TAP2 KO line and the TAPD KO line blotted for TAP2, with the TAP2 band being absent in both lines, whilst present in the wild type IS19 line. It can be seen that the TAP2 blots have many bands present and it was challenging to obtain a clear western blot. Unfortunately, TAP1 was more challenging and after multiple attempts a reliable blot for TAP1 could not be produced. We relied on the sequencing of the TAP1 locus that shows the generation of indels and the presence of a frameshift, in conjunction with the reduction in MHC I to confirm that TAP1 had been knocked out. Since undertaking the TAP KO work, the western blotting procedure has been optimised and improved, and in the future the TAP KO lines will be blotted for TAP1 and TAP2 using the optimised western blotting protocol.

With the TAP KO lines made and characterised, we decided to investigate how these KO lines effect transport of designated peptides. We used the radiolabelled P198 peptide (KRYNASAY) for transport and initially tried to inhibit this transport with unlabelled P198,

shown in Figure 7.3.1. It is important to note that this experiment was carried out while the TAP KO lines were still being characterised and validated. The chosen TAP KO lines detailed in section 7.2 were not all used for the experiment in figure 7.3.1 as they had not been characterised yet. The samples used in were TAP1.1 cl-1, which is the chosen TAP1 KO. The TAP2 KO used was TAP2.1 cl-4, which after further analysis was found to have an indel that was in frame for one allele so was not used for further experimentation and a new TAP2 KO was used for final experiments. The creation of the TAPD line was created by transfecting wild type IS19 cells with two sgRNAs at once, one targeted to TAP1 and one to TAP2 and the TAPD line used in figure 7.3.1 was TAPD cl-1 which after complete analysis was shown to be a homozygote TAP2 KO. This clone (TAPD cl-1) was then selected to be used as the TAP2 KO clonal line and was renamed as TAP2.1 cl-1. In summary, the results in figure 7.3.1 show the transport of a TAP1 KO (TAP1.1 cl-1), a TAP2 mutant with a one allele KO and one allele with an in-frame mutation (TAP2.1 cl-4) and a TAP2 KO (TAPD cl-1). Figure 7.3.1 shows that wild type IS19 cells efficiently transport the P198 peptide, giving ~1300 CPM that can be inhibited to ~30 CPM with the addition of cold P198. TAP1.1 cl-1 loses ~75 % of P198 transport compared to wild type and interestingly the remaining ~25 % transport can be inhibited with cold P198. The pseudo TAP2 KO (TAP2.1 cl-4) shows a ~50 % reduction in P198 transport, again this transport can be fully inhibited. The TAP2 KO (TAPD cl-1) shows a similar result to the TAP1 KO, with ~25 % residual transport that can be fully inhibited. At the time of experimentation the TAP2.1 cl-4 result was confusing as to why the TAP2 KO had ~50 % transport remaining when compared to the TAP1 KO, and suggested that perhaps TAP1 could operate as a homodimer complex. This result was particularly interesting as previous data in human models showed that TAP1 could form a homodimer, be it non-functional. This idea is likely to be false as the TAP2.1 cl-4 line was only fully disabled on one allele.

Regardless, the question remained as to why both the TAP1 KO and the TAP2 KO lines both show ~25 % transport and could this be transport from TAP homodimer complexes?

To address this question we took two approaches, firstly to carry on screening for a TAPD KO to test whether transport remains, secondly to test whether the transport that was seen was specifically inhibited by the P198 peptide or could a peptide that has been shown not to transport also inhibit the 25% transport seen.

Figure 7.3.2a shows transport assays of P198 either inhibited with P198 or P167 (ADYNDSAE) a B4 peptide that is not has been shown to not transport in IS19. The lines used were the TAP1 KO, TAP2 KO (TAPD cl-1) and three new potential double KOs, TAPD cl-2, 4 and 5. IS19 showed complete inhibition with P198 and was only marginally affected by inhibition with the B4 peptide, which was as expected and seen previously (Walker et al 2011). The TAP1 KO and the TAP2 KO clones both showed the same result as seen in 7.3.1, with ~ 25% transport of the P198 peptide that was fully inhibited with cold P198. Interestingly, the B4 peptide was unable to inhibit the transport which suggests that the transport seen is specific transport through a transporting system. Two of the new TAPD clones screened (2 & 4) both showed the same results as the TAP1 KO and TAP2 KO, whilst TAPD cl-5 showed results the same as the wild type. When the three new TAPD clones were fully analysed it was seen that TAPD cl-2 was in fact a TAP1 and TAP2 KO (TAPD), TAPD cl-4 was a homozygous TAP1 KO and TAPD cl-5 was unedited for TAP1 and TAP2 (WT). The most surprising aspect of this result is that the TAP double KO, TAPD cl-2 still has residual transport of ~25% that is specifically inhibited by P198, showing that the residual transport is not due to homodimeric TAPs or is in fact due to TAP transport which suggests that the transport could be from separate peptide transporter, such as TAPL.

In conjunction to the transport assay, each cell line was stained for class I heavy chain cell surface expression (fig 7.3.2b). As expected and previously shown, TAP1, TAP2 and TAPD KOs show a ~90 % decrease in class I expression and TAPD cl-5 which was identified as wild type by cloning and sequencing had class I levels comparable with wild type. To try to understand why ~25 % transport is still seen in all TAP KO clones transport assays of P198 were repeated but with an additional inhibition condition, using a peptide that is known to bind class I (Walker et al 2011) but not to transport, RRRREQTVY (P149). Figure 7.3.3 shows that P147 was unable to inhibit transport in the TAP KO lines.

It was unclear whether this residual transport was dependent on the expression of MHCI, so to test whether MHCI was necessary for peptide exchange, B2m was knocked out in the TAP KO lines. Flow cytometry analysis of B2m and class I heavy chain expression is shown in figure 7.4.1, with all identified TAP1+β2m and TAP2+β2m KO clones showing a complete loss of β2m and heavy chain. All lines were used for transport assays of P198 and inhibited with unlabelled P198, shown in figure 7.4.2a. Transport of the wild type and TAP KO lines showed the same results as previously seen and described, but all TAP KO lines that had β2m additionally knocked out lose the residual ~25 % remaining transport. An important control that is missing from this experiment is the β2m KO line which was left out due to technical limitations, though transport of the β2m KO line has been assessed previously and shows no difference to the wild type (Tregaskes, unpublished data). With these considerations in mind, the current data suggests that there is ~25 % transportation of the P198 peptide that is not dependent on the TAPs but is dependent on class I expression. To fully understand this further experimentation is required, such as immunopeptidomic analysis of the TAP KO cell lines.

To summarise, the aim of this project was to create a system in which we could assess which polymorphisms in the TAPs are important for specific peptide transport, with the first stage being to create and characterise TAP KO lines, then assess the effect this has on the ability of the cell line to transport a B19 peptide. We have provided strong evidence that TAPs in chicken do not function as a homodimer in the B19 haplotype and found that even when both TAPs are absent we see ~25% remaining transport that is dependent on MHCI expression. The next stage of the project is to optimise transfection of TAPs (of varying haplotypes) into TAP KO lines (that still contain β2m) to see if MHCI protein levels can be restored on the cell surface, if these cells have restored peptide transport and if so for which peptides.

8. General discussion

8. General discussion

The aim of this project was to gain a deeper understanding of the antigen processing system of the chicken, utilising recent advances in genome editing, in particular the CRISPR-Cas9 system. The aim was broken down into two main sections, peptide loading of MHCI and MHCII.

8.1 Investigating TAP transport specificity in the chicken

The MHCI antigen processing has multiple components such as the TAP transporters, tapasin and the relatively newly identified peptide editor TAPBPR. As mentioned in the general introduction, there are many differences between the chicken and human MHC loci. Key differences include that there is one dominantly expressed MHCI in the chicken and that the MHCI peptide loading genes (TAP1, TAP2 and tapasin) contained in the chicken MHC are polymorphic. The polymorphic MHCI genes are in strong linkage disequilibrium with the polymorphic peptide loading genes (tapasin and TAP), resulting in co-evolution. This co-evolution of the class I system results in specific alleles of the TAP genes that specify peptide translocation of peptides with a motif that correlates with that of one particular MHCI, specifically the dominantly expressed BF2 but not BF1 (Walker et al 2011). As previously discussed it still unknown as to which polymorphic residues of the TAP complex MSD are responsible for the transport of peptides with a specific motif.

The aim of this project was to understand which of these polymorphic residues of the TAP complex MSD are required for specific peptide transport. The initial experimental design was to use CRISPR-Cas9 to make IS19 TAP1 KO, TAP2 KO, TAPD KO lines, assess peptide transport

using a radiolabelled peptide and then proceed to transfect the TAP KO lines with the respective TAP(s) that have mutated polymorphic MSD residues of interest. It was expected that no transport of radiolabelled peptide would be seen in any of the TAP KO lines. Surprisingly, ~25% residual transport of the WT was observed for the P198 peptide, in the TAP1 KO, TAP2 KO and TAPD KO, which was inhibited with unlabelled peptide (fig 7.3.1). Further experimentation was required to understand the remaining transport. It was possible that the transport seen was due to non-specific peptide entry into the ER, for instance if the ER had become permeabilised during sample preparation. If peptide was able to enter the ER in a non-specific manner, then the residual transport could be inhibited by excess unlabelled peptide that is not transported by the TAP transporters. Additional experiments showed that the transport could not be inhibited by peptides known to not be transported by the TAP transporters of IS19, including a peptide that is able to bind to IS19 MHCI by MHC refolding but cannot be transported (fig 7.3.3).

The next step was to examine whether MHCI expression was required for the remaining transport seen. MHCI KOs were made in the TAP KO lines by editing β_2m, producing three more KO lines (TAP1+ β_2m KO, TAP2+ β_2m KO and TAPD+ β_2m KO) and were analysed by radiolabelled transport assay. Surprisingly, the ~25% residual transport of the WT seen in the TAP KOs disappeared when β_2m was additionally edited (fig 7.4.2). One explanation for this result is that the remaining MHCI molecules are escaping the ER and migrating to the cell surface in the TAP KOs (fig 7.4.2), and then bind the radiolabelled P198 peptide. These MHCI molecules binding the radiolabelled peptide are then pulled down with concanavalin A and are therefore seen during analysis by γ-counter. If this is the case, then the radioactivity measured is not from transport and instead by peptide exchange. In the future we plan to

test this by immunoprecipitating MHCI from the samples before concanavalin A treatment, followed with assessment by γ -counter. Alternatively, the 25% residual transport of the WT could be from a second peptide transportation system, such as TAPL.

8.1.2 Future work

The second stage of this project will be to transfect constructs expressing WT TAPs and TAPs with point mutations, and assess MHCI surface expression and peptide transport to see if they are restored. This work has begun in conjunction with a visiting PhD student Magdalena Migalska, with preliminary data suggesting that MHCI protein expression can be restored when WT TAP constructs are transfected into IS19. Interestingly, MHCI protein expression is restored after transfection of WT TAPs from a different haplotype B15 that has similar peptide transport specificity but is not restored when TAPs from a haplotype (B4) with very different peptide transport specificity (9. Appendix).

8.2 Understanding MHCII antigen processing

The second aim focused on MHCII antigen processing and presentation in the chicken, which is much less understood than that of MHCI. What is known is that the chicken has two MHCII B genes (BLB1 and BLB2) and two DMB genes (DMB1 and DMB2), with BLB2 and DMB2 being dominantly expressed in hematopoietic cells, whilst BLB1 and DMB1 are dominantly expressed in intestinal epithelial cells (Jacob et al 2000, Parker and Kaufman 2017). The dominant expression of BLB and DMB pairings suggested that BLB and DMB may have co-evolved for specific antigen processing functions, as seen for the MHCI system.

The main aim was to understand how each DMB interacts and effects each BLB molecule. To do this single KOs were made of most antigen processing and presentation genes in the chicken (BLB1 KO, BLB2 KO, DMA KO, DMB1 KO and DMB2 KO). Double KOs were made for each combination of BLB and DM (BLB1+DMA KO, BLB1+DMB1 KO, BLB1+DMB2 KO, BLB2+DMA KO, BLB2+DMB1 KO and BLB2+DMB2 KO). The experimental design was to use the double KO lines to assess the effect of each DM molecule (or no DM molecules) on each BLB molecule. The first step was to assess the protein expression of BLB1 and BLB2 in the single BLB1 KO, BLB2, DMA KO, DMB1 KO and DMB2 KO lines. It was expected that BLB1 and BLB2 expression would be independent of each other but surprisingly, the first observations showed that the BLB1 KO affected the protein expression of BLB2 and *vice versa* (fig 5.3.2). It was also seen that each DM KO affected protein expression of both BLB molecules, whereas we expected DMB1 to only affect BLB1 and DMB2 to only affect BLB2 (fig 5.3.2). It was also observed that DMA protein expression was reduced in both the DMB1 KO and the DMB2 KO (fig 5.3.3); and that DMB2 protein expression was reduced in the DMA KO and the DMB1 KO (fig 5.3.4). It was unexpected that DM protein expression would be strongly affected by the editing of any DM gene, particularly that DMB2 protein expression could be affected in the DMB1 KO, as the two DMB molecules are not expected to interact in any way. The nature of MHCII protein expression appears to be heavily influenced by all MHCII proteins involved in peptide processing and presentation. Much further work is required to understand the effects of MHCII protein regulation and stability in the MHCII system.

Secondly, temperature stability of MHCII was assessed in the IS19 DMA KO, DMB1 KO and DMB2 KO lines. All DM KO lines were shown to reduce MHCII thermostability (fig 5.3.6, 5.3.7). The DMA KO was found to decrease the thermostability of MHCII, with a lesser effect

found for the DMB1 KO and DMB2 KO lines. Interestingly, the DMB2 KO had a greater effect on thermostability than the DMB1 KO, suggesting that DMB2 plays a more prominent role in editing both BLB molecules (fig 5.3.7). It was noted that the DMB1 KO affects the thermostability of both BLBs. These initial findings suggest that both DMB molecules are important for peptide loading of both BLB molecules, which is in opposition to the hypothesis that DMB editing is restricted to just one BLB molecule. Similarly, in the B21 haplotype the DMB2 KO was able to form thermostable MHCII complexes for BLB1 and BLB2, whereas the DMA KO was unable to form a stable complex at any temperature (fig 5.3.8). This supports the notion that both DMBs have editing functions of both BLB molecules and is likely to be consistent between haplotypes. Interestingly, DM activity varies between haplotypes with B19 able to form stable MHCII complexes in the absence of DM at low temperatures, whereas B21 is not, suggesting a spectrum of DM-dependence.

Lastly, the peptide repertoire of BLB1 and BLB2 was elucidated by immunopeptidomics. BLB1 and BLB2 were shown to have two distinct repertoires expressed, with very peptides shared between the two molecules (fig 6.3.1, 6.4.4). DM was also shown to influence the repertoire of MHCII molecules. DM deficient cells were shown to have a more diverse motif, that overall was very similar to that of BLB2 (fig 6.3.1).

8.2.1 Future work

Current and future work includes reconstitution of the edited MHCII protein to assess whether WT phenotypes are restored. Independent MHCII KO clones have all been created for the IS19 MHCII KO lines, for which protein and RNA levels will be assessed and compared

to the data presented in this thesis. BLB1 KO, BLB2 KO and DMB1 KO clonal lines need to be derived for the B21 haplotype. The analysis of DM-dependent thermostability of BLB1 and BLB2 in B21 needs to be continued, including characterization of the DMB1 KO and replication of all DM KO lines. Ultimately, characterization at the RNA and protein level needs to be carried out for all B21 MHCII KO lines for comparison with the B19 study. Furthermore, MHCII KO lines could be generated in a variety of common MHC haplotype cell lines, such as B2, B4 and B15. Comparisons at the RNA and protein level, and thermostability assays could then be compared between all haplotypes. This would allow unifying themes of the chicken MHCII processing and presentation system to be identified, as well as highlighting the variability of MHCII processing between haplotypes. Immunopeptidomics affords the opportunity to assess the diversity of peptides presented by BLB1 and BLB2 between haplotypes and how these repertoires are affected by the absence of each or both DM molecules, starting with a comparison of B19 and B21.

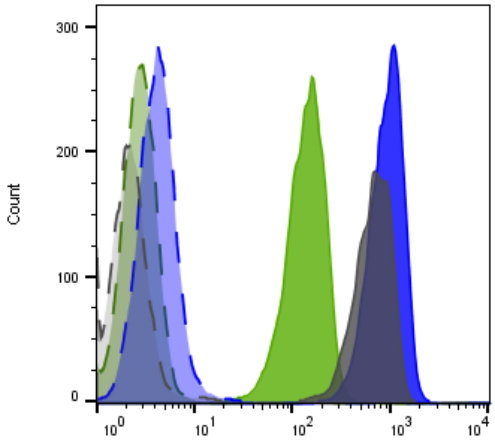
8.3 Future of genome editing in chickens

Recent advances have enabled the culture of chicken primordial germ cells (PGCs). It has been demonstrated that cultured PGCs can be genetically edited and successfully transplanted back into the chicken embryo to generate transgenic chickens. This breakthrough drastically changes the landscape of chicken immunology, allowing the targeted KO of immune genes, and for the effects to be assessed *in vivo*. In collaboration with two groups at the forefront of this technology (Prof Benni Schusser and Dr Mike McGrew), our lab has commenced work to make MHCI and MHCII deficient birds. In the future, such birds could be studied in a variety of ways by the scientific community, including pathogen challenge experiments and vaccine responses. This work has the potential to make

groundbreaking progress, that will not only advance our understanding of immunology but also impact vaccination, breeding strategies and food security.

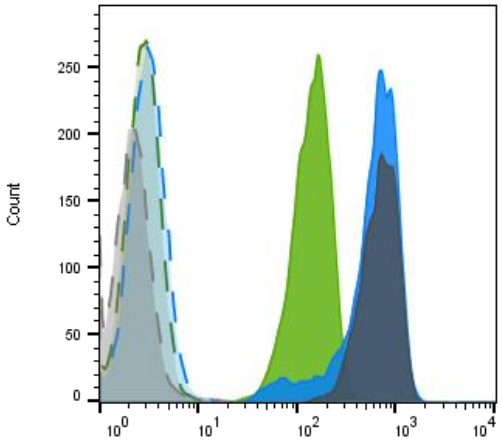
9. Appendices

9.1 Appendix i



Appendix ia. Flow cytometry cell surface staining of B2m n IS19 TAD KO transfected with IS19 TAP1 and TAP2 polyclonal populations.

IS19 TAPD cell line was co-transfected with pcDNA plasmids expressing IS19 TAP1 and TAP2. IS19 WT, TAPD KO and TAPD KO+TAP1 and TAP2 cells were analysed via flow cytometry using the antibody F21-21, to β2m. IS19 cells are shown with secondary antibody only (brown dotted line) and F21-21 (brown). TAPD KO cells are shown with secondary antibody only (green dotted line) and F21-21 (green). TAPD KO+TAP1 and TAP2 cells are shown with secondary antibody only (blue dotted line) and F21-21 (blue). Experiment carried out by Magdalena Migalska.



Appendix ib. Flow cytometry cell surface staining of B2m n IS19 TAD KO transfected with TG15 TAP1 and TAP2 polyclonal populations.

IS19 TAPD cell line was co-transfected with pcDNA plasmids expressing TG15 TAP1 and TAP2. IS19 WT, TAPD KO and TAPD KO+TAP1 and TAP2 cells were analysed via flow cytometry using the antibody F21-21, to β2m. IS19 cells are shown with secondary antibody only (brown dotted line) and F21-21 (brown). TAPD KO cells are shown with secondary antibody only (green dotted line) and F21-21 (green). TAPD KO+TAP1 and TAP2 cells are shown with secondary antibody only (blue dotted line) and F21-21 (blue). Experiment carried out by Magdalena Migalska.

9.2 Appendix ii

```
B2m_RG5             1 CAAGGTGCAGGTGTACTCCCGCTTCCCCGCCTCTGCGGGCACCAAGAACG
50
                      ||||||||||||||||||||||||||||          ||||||||||||
CB_PGC_B2m.1_       1 CAAGGTGCAGGTGTACTCCCGCTTCCCC----------GCACCAAGAACG
40

B2m_RG5            51 TCCTCAACTGCTTCGCGGCCGGCTTCCACCCACCCAAGATCTCCATCACG
100
                      ||||||||||||||||||||||||||||||.||||||||||||||||||
CB_PGC_B2m.1_      41 TCCTCAACTGCTTCGCGGCCGGCTTCCACCCGCCCAAGATCTCCATCACG
90

B2m_RG5           101 CTGATGAAGGACGGCGTGCCCATGGAGGGTGCGCAGTACTCCGACATGTC
150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.1_      91 CTGATGAAGGACGGCGTGCCCATGGAGGGTGCGCAGTACTCCGACATGTC
140

B2m_RG5           151 CTTCAACGACGACTGGACGTTCCAGCGCCTGGTGCACGCCGACTTCACGC
200
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.1_     141 CTTCAACGACGACTGGACGTTCCAGCGCCTGGTGCACGCCGACTTCACGC
190

B2m_RG5           201 CCAGCAGCGGTTCCACCTACGCGTGCAAGGTGGAGCACGAGACCCTGAAG
250
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.1_     191 CCAGCAGCGGTTCCACCTACGCGTGCAAGGTGGAGCACGAGACCCTGAAG
240

B2m_RG5           251 GAGCCGCAGGTCTACAAGT    269
                      ||||||||||||||||||
```

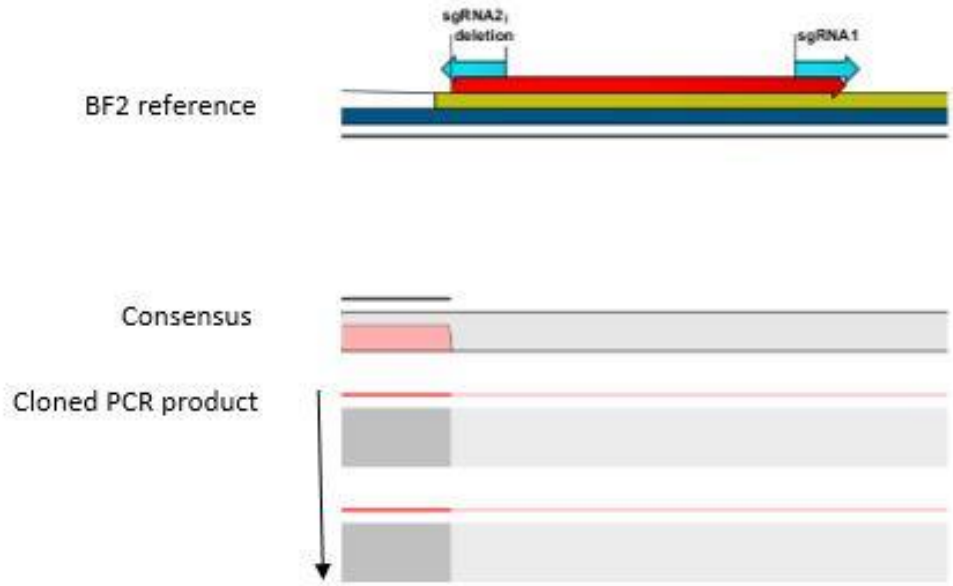
Appendix iia DNA sequence alignment of B12 PGC clonal line B2m.1 cl-1 shows editing.

The editing events revealed by sequencing exon 2 of B2m were mapped onto the exon 2 sequence of B2m and compared to the WT, the editing events can be seen in at position 29. The two sequences are labelled on the left-hand margin; WT sequence (B2m_RG5) and edited alleles of clonal line b2m.1 cl-1 (CB_PGC_B2m.1).

```
B2m_RG5             1 CAAGGTGCAGGTGTACTCCCGCTTCCCCGCCTCTGCGGGCACCAAGAACG
50
                      ||||||||||||||||||           ||||||||||||||||||||
CB_PGC_B2m.7_       1 CAAGGTGCAGGTGTACTC-----------CCTCTGCGGGCACCAAGAACG
39

B2m_RG5            51 TCCTCAACTGCTTCGCGGCCGGCTTCCACCCACCCAAGATCTCCATCACG
100
                      ||||||||||||||||||||||||||||||.|||||||||||||||.||
CB_PGC_B2m.7_      40 TCCTCAACTGCTTCGCGGCCGGCTTCCACCCGCCCAAGATCTCCATCGCG
89

B2m_RG5           101 CTGATGAAGGACGGCGTGCCCATGGAGGGTGCGCAGTACTCCGACATGTC
150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.7_      90 CTGATGAAGGACGGCGTGCCCATGGAGGGTGCGCAGTACTCCGACATGTC
139

B2m_RG5           151 CTTCAACGACGACTGGACGTTCCAGCGCCTGGTGCACGCCGACTTCACGC
200
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.7_     140 CTTCAACGACGACTGGACGTTCCAGCGCCTGGTGCACGCCGACTTCACGC
189

B2m_RG5           201 CCAGCAGCGGTTCCACCTACGCGTGCAAGGTGGAGCACGAGACCCTGAAG
250
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
CB_PGC_B2m.7_     190 CCAGCAGCGGTTCCACCTACGCGTGCAAGGTGGAGCACGAGACCCTGAAG
239

B2m_RG5           251 GAGCCGCAGGTCTACAAGT    269
                      ||||||||||||||||||
CB_PGC_B2m.7_     240 GAGCCGCAGGTCTACAAGT    258
```

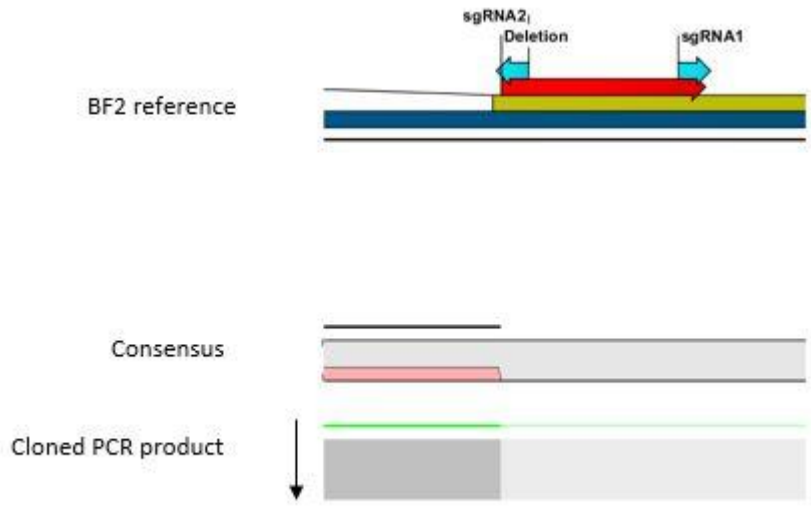
Appendix iib DNA sequence alignment of B12 PGC clonal line B2m.1 cl-7 shows editing.

The editing events revealed by sequencing exon 2 of B2m were mapped onto the exon 2 sequence of B2m and compared to the WT, the editing events can be seen in at position 19. The two sequences are labelled on the left-hand margin; WT sequence (B2m_RG5) and edited alleles of clonal line b2m.1 cl-1 (CB_PGC_B2m.7).



Appendix iic DNA sequence alignment of B12 PGC polyclonal line BF2.1/2 shows a large deletion.

Two BF2 targeting sgRNAs were co-transfected into B12 PGC cells and sorted for GFP expression, the GFP positive polyclonal population was expanded and genomic DNA was extracted. BF2 was amplified, cloned and sequenced. Editing events revealed by sequencing exon 2 of BF2 were mapped onto the whole sequence of BF2 and compared to the WT. The combination of both sgRNAs created a 120 bp deletion between the two targeted sites of the sgRNAs. sgRNA targeting sites (blue arrows) and the deletion generated (red arrow) are shown.



Appendix iid DNA sequence alignment of B12 PGC polyclonal line BF2.1/2 shows a large deletion.

Two BF2 targeting sgRNAs were co-transfected into B12 PGC cells and sorted for GFP expression, the GFP positive polyclonal population was expanded and genomic DNA was extracted. BF2 was amplified, cloned and sequenced. Editing events revealed by sequencing exon 2 of BF2 were mapped onto the whole sequence of BF2 and compared to the WT. The combination of both sgRNAs created a 121 bp deletion between the two targeted sites of the sgRNAs. sgRNA targeting sites (blue arrows) and the deletion generated (red arrow) are shown.

10. References

Anders, A.-K., Call, M. J., Schulze, M.-S. E. D., Fowler, K. D., Schubert, D. A., Seth, N. P., … Wucherpfennig, K. W. (2011). HLA-DM captures partially empty HLA-DR molecules for catalyzed removal of peptide. *Nature Immunology*, *12*(1), 54–61. http://doi.org/10.1038/ni.1967

Bacon, L. D., Witter, R. L., Crittenden, L. B., Fadly, A., & Motta, J. (1981). B-Haplotype Influence on Marek's Disease, Rous Sarcoma, and Lymphoid Leukosis Virus-Induced Tumors in Chickens. *Poultry Science*, *60*(6), 1132–1139. http://doi.org/10.3382/ps.0601132

Barker, J. N., Navsaria, H. A., Leigh, I. M., & MacDonald, D. M. (1988). Gamma-interferon induced human keratinocyte HLA-DR synthesis: the role of dermal activated T lymphocytes. *The British Journal of Dermatology*, *119*(5), 567–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2462899

Beck, S., & Trowsdale, J. (1999). Sequence organisation of the class II region of the human MHC. *Immunological Reviews*, *167*, 201–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10319262

Bénaroch, P., Yilla, M., Raposo, G., Ito, K., Miwa, K., Geuze, H. J., & Ploegh, H. L. (1995). How MHC class II molecules reach the endocytic pathway. *The EMBO Journal*, *14*(1), 37–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7530198

Benoist, C., & Mathis, D. (1990). Regulation of Major Histocompatibility Complex Class-II Genes: X, Y and Other Letters of the Alphabet. *Annual Review of Immunology*, *8*(1), 681–715. http://doi.org/10.1146/annurev.iy.08.040190.003341

Bertolino, P., & Rabourdin-Combe, C. (1996). The MHC class II-associated invariant chain: a molecule with multiple roles in MHC class II biosynthesis and antigen presentation to CD4+ T cells. *Critical Reviews in Immunology*, *16*(4), 359–79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8954255

Bhat, K. P., Truax, A. D., & Greer, S. F. (2010). Phosphorylation and ubiquitination of degron proximal residues are essential for class II transactivator (CIITA) transactivation and major histocompatibility class II expression. *The Journal of Biological Chemistry*, *285*(34), 25893–903. http://doi.org/10.1074/jbc.M110.127746

Bikoff, E. K., Huang, L. Y., Episkopou, V., van Meerwijk, J., Germain, R. N., & Robertson, E. J. (1993). Defective major histocompatibility complex class II assembly, transport, peptide acquisition, and CD4+ T cell selection in mice lacking invariant chain expression. *The Journal of Experimental Medicine*, *177*(6), 1699–712. http://doi.org/10.1084/JEM.177.6.1699

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*, *329*(6139), 506–512. http://doi.org/10.1038/329506a0

Blees, A., Januliene, D., Hofmann, T., Koller, N., Schmidt, C., Trowitzsch, S., … Tampé, R. (2017). Structure of the human MHC-I peptide-loading complex. *Nature*, *551*(7681), 525–528. http://doi.org/10.1038/nature24627

Bodmer, H., Viville, S., Benoist, C., & Mathis, D. (1994). Diversity of endogenous epitopes bound to MHC class II molecules limited by invariant chain. *Science (New York, N.Y.)*, *263*(5151), 1284–6. http://doi.org/10.1126/SCIENCE.7510069

BOONYANUWAT, K., THUMMABUTRA, S., SOOKMANEE, N., VATCHAVALKHU, V., & SIRIPHOLVAT, V. (2006). Influences of major histocompatibility complex class I haplotypes on avian influenza virus disease traits in Thai indigenous chickens. *Animal Science Journal*, *77*(3), 285–289. http://doi.org/10.1111/j.1740-0929.2006.00350.x

Bourlet, Y., Béhar, G., Guillemot, F., Fréchin, N., Billault, A., Chaussé, A. M., … Auffray, C. (1988). Isolation of chicken major histocompatibility complex class II (B-L) beta chain sequences: comparison with mammalian beta chains and expression in lymphoid organs. *The EMBO Journal*, *7*(4), 1031–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2841107

Bouvier, M., & Wiley, D. (1994). Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science*, *265*(5170), 398–402. http://doi.org/10.1126/science.8023162

Boyle, L. H., Hermann, C., Boname, J. M., Porter, K. M., Patel, P. A., Burr, M. L., … Trowsdale, J. (2013). Tapasin-related protein TAPBPR is an additional component of the MHC class I presentation pathway. *Proceedings of the National Academy of Sciences*, *110*(9), 3465–3470. http://doi.org/10.1073/pnas.1222342110

Bridgen, J., Snary, D., & Crumpton, M. J. (1976). Isolation and N-terminal amino acid sequence of membrane-bound human HLA-A and HLA-B antigens. *Nature*, *261*(5557), 200–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1272392

Briles, W. E., Goto, R. M., Auffray, C., & Miller, M. M. (1993). A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. *Immunogenetics*, *37*(6), 408–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8436415

Briles, W. E., Stone, H. A., & Cole, R. K. (1977). Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science (New York, N.Y.)*, *195*(4274), 193–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/831269

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, *364*(6432), 33–39. http://doi.org/10.1038/364033a0

Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J., & Wiley, D. C. (1988). A hypothetical model of the foreign antigen binding site of Class II histocompatibility molecules. *Nature*, *332*(6167), 845–850. http://doi.org/10.1038/332845a0

Busch, R., Doebele, R. C., Patil, N. S., Pashine, A., & Mellins, E. D. (2000). Accessory molecules for MHC class II peptide loading. *Current Opinion in Immunology*, *12*(1), 99–106. http://doi.org/10.1016/S0952-7915(99)00057-6

Carrington, M., Nelson, G. W., Martin, M. P., Kissner, T., Vlahov, D., Goedert, J. J., … O'Brien, S. J. (1999). HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*

(New York, N.Y.), *283*(5408), 1748–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10073943

Carrington, M., & O'Brien, S. J. (2003). The influence of HLA genotype on AIDS. *Annual Review of Medicine*, *54*(1), 535–51. http://doi.org/10.1146/annurev.med.54.101601.152346

Chan, J. L., Tang, K. C., Patel, A. P., Bonilla, L. M., Pierobon, N., Ponzio, N. M., & Rameshwar, P. (2006). Antigen-presenting property of mesenchymal stem cells occurs during a narrow window at low levels of interferon-gamma. *Blood*, *107*(12), 4817–24. http://doi.org/10.1182/blood-2006-01-0057

Chappell, P. E., Meziane, E. K., Harrison, M., Magiera, Ł., Hermann, C., Mears, L., … Kaufman, J. (2015). Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. *ELife*, *4*. http://doi.org/10.7554/eLife.05345

Choi, N. M., Majumder, P., & Boss, J. M. (2011). Regulation of major histocompatibility complex class II genes. *Current Opinion in Immunology*, *23*(1), 81–87. http://doi.org/10.1016/J.COI.2010.09.007

Choi, N. M., Majumder, P., & Boss, J. M. (2010). Regulation of major histocompatibility complex class II genes. http://doi.org/10.1016/j.coi.2010.09.007

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, *339*(6121), 819–23. http://doi.org/10.1126/science.1231143

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, *339*(6121), 819–23. http://doi.org/10.1126/science.1231143

Costantino, C. M., Hang, H. C., Kent, S. C., Hafler, D. A., & Ploegh, H. L. (2008). Lysosomal cysteine and aspartic proteases are heterogeneously expressed and act redundantly to initiate human invariant chain degradation. *Journal of Immunology (Baltimore, Md. : 1950)*, *180*(5), 2876–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18292509

Cresswell, P. (1996). Invariant chain structure and MHC class II function. *Cell*, *84*(4), 505–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8598037

Criscitiello, M. F., Ohta, Y., Graham, M. D., Eubanks, J. O., Chen, P. L., & Flajnik, M. F. (2012). Shark class II invariant chain reveals ancient conserved relationships with cathepsins and MHC class II. *Developmental and Comparative Immunology*, *36*(3), 521–33. http://doi.org/10.1016/j.dci.2011.09.008

Czernielewski, J. M., & Bagot, M. (1986). Class II MHC antigen expression by human keratinocytes results from lympho-epidermal interactions and gamma-interferon production. *Clinical and Experimental Immunology*, *66*(2), 295–302. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2434272

Dausset, J. (1958). Iso-leuco-anticorps. *Acta Haematologica*, *20*(1–4), 156–166. http://doi.org/10.1159/000205478

de la Salle, H., Zimmer, J., Fricker, D., Angenieux, C., Cazenave, J.-P., Okubo, M., … Hanau, D. (1999). HLA class I deficiencies due to mutations in subunit 1 of the peptide transporter TAP1. *Journal of Clinical Investigation*, *103*(5), R9–R13. http://doi.org/10.1172/JCI5687

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., … Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–7. http://doi.org/10.1038/nature09886

Denzin, L. K., Fallas, J. L., Prendes, M., & Yi, W. (2005). Right place, right time, right peptide: DO keeps DM focused. *Immunological Reviews*, *207*(1), 279–292. http://doi.org/10.1111/j.0105-2896.2005.00302.x

Deverson, E. V., Gow, I. R., Coadwell, W. J., Monaco, J. J., Butcher, G. W., & Howard, J. C. (1990). MHC class II region encoding proteins related to the muKidrug resistance family of transmembrane transporters. *Nature*, *348*(6303), 738–741. http://doi.org/10.1038/348738a0

Doebele, R. C., Busch, R., Scott, H. M., Pashine, A., & Mellins, E. D. (2000). Determination of the HLA-DM interaction site on HLA-DR molecules. *Immunity*, *13*(4), 517–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11070170

Elliott, E. A., Drake, J. R., Amigorena, S., Elsemore, J., Webster, P., Mellman, I., & Flavell, R. A. (1994). The invariant chain is required for intracellular transport and function of major histocompatibility complex class II molecules. *The Journal of Experimental Medicine*, *179*(2), 681–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8294875

Falk, K., Rötzschke, O., Stevanović, S., Jung, G., & Rammensee, H.-G. (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, *351*(6324), 290–296. http://doi.org/10.1038/351290a0

Ferrante, A. (2013). Thermodynamics of Peptide-MHC Class II Interactions: Not all Complexes are Created Equal. *Frontiers in Immunology*, *4*, 308. http://doi.org/10.3389/fimmu.2013.00308

Fling, S. P., Arp, B., & Pious, D. (1994). HLA-DMA and -DMB genes are both required for MHC class II/peptide complex formation in antigen-presenting cells. *Nature*, *368*(6471), 554–558. http://doi.org/10.1038/368554a0

Fremont, D. H., Crawford, F., Marrack, P., Hendrickson, W. A., & Kappler, J. (1998). Crystal structure of mouse H2-M. *Immunity*, *9*(3), 385–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9768758

Frickel, E.-M., Riek, R., Jelesarov, I., Helenius, A., Wuthrich, K., & Ellgaard, L. (2002). TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(4), 1954–9. http://doi.org/10.1073/pnas.042699099

Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, *31*(7), 397–405. http://doi.org/10.1016/j.tibtech.2013.04.004

Geppert, T. D., & Lipsky, P. E. (1985). Antigen presentation by interferon-gamma-treated endothelial cells and fibroblasts: differential ability to function as antigen-presenting cells despite comparable Ia expression. *Journal of Immunology (Baltimore, Md. : 1950)*, *135*(6), 3750–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3934267

Germain, R. N. (1994). MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. *Cell*, *76*(2), 287–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8293464

Glazier, K. S., Hake, S. B., Tobin, H. M., Chadburn, A., Schattner, E. J., & Denzin, L. K. (2002). Germinal center B cells regulate their capability to present antigen by modulation of HLA-DO. *The Journal of Experimental Medicine*, *195*(8), 1063–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11956297

Glimcher, L. H., & Kara, C. J. (1992). Sequences and Factors: A Guide to MHC Class-II Transcription. *Annual Review of Immunology*, *10*(1), 13–49. http://doi.org/10.1146/annurev.iy.10.040192.000305

Godkin, A. J., Smith, K. J., Willis, A., Tejada-Simon, M. V, Zhang, J., Elliott, T., & Hill, A. V. (2001). Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions. *Journal of Immunology (Baltimore, Md. : 1950)*, *166*(11), 6720–7. http://doi.org/10.4049/JIMMUNOL.166.11.6720

Gorer, P. A. (1936). The Detection of Antigenic Differences in Mouse Erythrocytes by the Employment of Immune Sera. *British Journal of Experimental Pathology*, *17*(1), 42. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2065193/

Greer, S. F., Harton, J. A., Linhoff, M. W., Janczak, C. A., Ting, J. P.-Y., & Cressman, D. E. (2004). Serine residues 286, 288, and 293 within the CIITA: a mechanism for down-regulating CIITA activity through phosphorylation. *Journal of Immunology (Baltimore, Md. : 1950)*, *173*(1), 376–83. http://doi.org/10.4049/JIMMUNOL.173.1.376

Groettrup, M., Kirk, C. J., & Basler, M. (2010). Proteasomes in immune cells: more than peptide producers? *Nature Reviews Immunology*, *10*(1), 73–78. http://doi.org/10.1038/nri2687

Guce, A. I., Mortimer, S. E., Yoon, T., Painter, C. A., Jiang, W., Mellins, E. D., & Stern, L. J. (2013). HLA-DO acts as a substrate mimic to inhibit HLA-DM by a competitive mechanism. *Nature Structural & Molecular Biology*, *20*(1), 90–98. http://doi.org/10.1038/nsmb.2460

Guillemot, F., Kaufman, J. F., Skjoedt, K., & Auffray, C. (1989). The major histocompatibility complex in the chicken. *Trends in Genetics : TIG*, *5*(9), 300–4. http://doi.org/10.1016/0168-9525(89)90112-1

Hershberg, R. M., Framson, P. E., Cho, D. H., Lee, L. Y., Kovats, S., Beitz, J., … Nepom, G. T. (1997). Intestinal epithelial cells use two distinct pathways for HLA class II antigen processing. *The Journal of Clinical Investigation*, *100*(1), 204–15. http://doi.org/10.1172/JCI119514

Hornell, T. M. C., Burster, T., Jahnsen, F. L., Pashine, A., Ochoa, M. T., Harding, J. J., … Mellins, E. D. (2006). Human dendritic cell expression of HLA-DO is subset specific and regulated by maturation. *Journal of Immunology (Baltimore, Md. : 1950)*, *176*(6), 3536–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16517722

Hosomichi, K., Miller, M. M., Goto, R. M., Wang, Y., Suzuki, S., Kulski, J. K., … Shiina, T. (2008). Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. *Journal of Immunology (Baltimore, Md. : 1950)*, *181*(5), 3393–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18714011

Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L.-F., Sontheimer, E. J., & Thomson, J. A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15644–9. http://doi.org/10.1073/pnas.1313587110

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., … Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9), 827–832. http://doi.org/10.1038/nbt.2647

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., … Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9), 827–832. http://doi.org/10.1038/nbt.2647

Hudson, J. C., Hoerr, F. J., Parker, S. H., & Ewald, S. J. (n.d.). Quantitative Measures of Disease in Broiler Breeder Chicks of Different Major Histocompatibility Complex Genotypes after Challenge with Infectious Bursal Disease Virus. *Avian Diseases*. American Association of Avian Pathologists. http://doi.org/10.2307/1592937

Hughes, A. L., & Nei, M. (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(3), 958–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2492668

Hughes, A. L., & Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, *335*(6186), 167–170. http://doi.org/10.1038/335167a0

Hughes, E. A., Hammond, C., & Cresswell, P. (1997). Misfolded major histocompatibility complex class I heavy chains are translocated into the cytoplasm and degraded by the proteasome. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(5), 1896–901. http://doi.org/10.1073/PNAS.94.5.1896

Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., … Engelhard, V. H. (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science (New York, N.Y.)*, *255*(5049), 1261–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1546328

Huse, M., Lillemeier, B. F., Kuhns, M. S., Chen, D. S., & Davis, M. M. (2006). T cells use two directionally distinct pathways for cytokine secretion. *Nature Immunology*, *7*(3), 247–55. http://doi.org/10.1038/ni1304

Hüttl, S., Helfrich, F., Mentrup, T., Held, S., Fukumori, A., Steiner, H., … Schröder, B. (2016). Substrate determinants of signal peptide peptidase-like 2a (SPPL2a)-mediated intramembrane proteolysis of the invariant chain CD74. *The Biochemical Journal*, *473*(10), 1405–22. http://doi.org/10.1042/BCJ20160156

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3316184

Jacob, J. P., Milne, S., Beck, S., & Kaufman, J. (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), 138–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10663576

Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1996). Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(2), 734–8. http://doi.org/10.1073/PNAS.93.2.734

Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Chi, Y., … Wiley, D. C. (1994). Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. *Nature*, *368*(6473), 711–718. http://doi.org/10.1038/368711a0

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. http://doi.org/10.1126/science.1225829

Jones, E. Y. (1997). MHC class I and class II structures. *Current Opinion in Immunology*, *9*(1), 75–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9039778

Kaufman, J. F., Auffray, C., Korman, A. J., Shackelford, D. A., & Strominger, J. (1984). The class II molecules of the human and murine major histocompatibility complex. *Cell*, *36*(1), 1–13. http://doi.org/10.1016/0092-8674(84)90068-0

Kaufman, J. (1999). Co-evolving genes in MHC haplotypes: the "rule" for nonmammalian vertebrates? *Immunogenetics*, *50*(3–4), 228–236. http://doi.org/10.1007/s002510050597

Kaufman, J. (2018). Unfinished Business: Evolution of the MHC and the Adaptive Immune System of Jawed Vertebrates. *Annual Review of Immunology*, *36*(1), 383–409. http://doi.org/10.1146/annurev-immunol-051116-052450

Kaufman, J. (2018). Generalists and Specialists: A New View of How MHC Class I Molecules Fight Infectious Pathogens. *Trends in Immunology*, *39*(5), 367–379. http://doi.org/10.1016/j.it.2018.01.001

Kaufman, J. (2015). Co-evolution with chicken class I genes. *Immunological Reviews*, *267*(1), 56–71. http://doi.org/10.1111/imr.12321

Kaufman, J. (2015). What chickens would tell you about the evolution of antigen processing and presentation. *Current Opinion in Immunology*, *34*, 35–42. http://doi.org/10.1016/J.COI.2015.01.001

Kaufman, J., Milne, S., Göbel, T. W. F., Walker, B. A., Jacob, J. P., Auffray, C., … Beck, S. (1999). The chicken B locus is a minimal essential major histocompatibility complex. *Nature*, *401*(6756), 923–925. http://doi.org/10.1038/44856

Kelley, J., Walter, L., & Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics*, *56*(10), 683–695. http://doi.org/10.1007/s00251-004-0717-7

Kelly, A. P., Monaco, J. J., Cho, S. G., & Trowsdale, J. (1991). A new human HLA class II-related locus, DM. *Nature*, *353*(6344), 571–3. http://doi.org/10.1038/353571a0

Kim, H., & Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. *Nature Reviews. Genetics*, *15*(5), 321–34. http://doi.org/10.1038/nrg3686

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8577732

King, G., & Dixon, A. M. (2010). Evidence for role of transmembrane helix–helix interactions in the assembly of the Class II major histocompatibility complex. *Molecular BioSystems*, *6*(9), 1650. http://doi.org/10.1039/c002241a

Koch, N., Zacharias, M., König, A., Temme, S., Neumann, J., & Springer, S. (2011). Stoichiometry of HLA Class II-Invariant Chain Oligomers. *PLoS ONE*, *6*(2), e17257. http://doi.org/10.1371/journal.pone.0017257

Lamb, C. A., Yewdell, J. W., Bennink, J. R., & Cresswell, P. (1991). Invariant chain targets HLA class II molecules to acidic endosomes containing internalized influenza virus. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(14), 5998–6002. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2068076

Lamont, S. J. (1998). The chicken major histocompatibility complex and disease. *Revue Scientifique et Technique (International Office of Epizootics)*, *17*(1), 128–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9638806

Landsverk, O. J. B., Bakke, O., & Gregers, T. F. (2009). MHC II and the Endocytic Pathway: Regulation by Invariant Chain. *Scandinavian Journal of Immunology*, *70*(3), 184–193. http://doi.org/10.1111/j.1365-3083.2009.02301.x

Leach, M. R., Cohen-Doyle, M. F., Thomas, D. Y., & Williams, D. B. (2002). Localization of the lectin, ERp57 binding, and polypeptide binding sites of calnexin and calreticulin. *The Journal of Biological Chemistry*, *277*(33), 29686–97. http://doi.org/10.1074/jbc.M202405200

Livant, E. J., Brigati, J. R., & Ewald, S. J. (2004). Diversity and locus specificity of chicken MHC B class I sequences. *Animal Genetics*, *35*(1), 18–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14731225

Long, E. O. (1985). In search of a function for the invariant chain associated with Ia antigens. *Survey of Immunologic Research*, *4*(1), 27–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3890055

Majumder, P., & Boss, J. M. (2018). Insulators through Interactions with MHC Class II Genes MHC Class II Cohesin Regulates. http://doi.org/10.4049/jimmunol.1100688

Makarova, K. S., Grishin, N. V, Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, *1*(1), 7. http://doi.org/10.1186/1745-6150-1-7

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., … Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science*, *339*(6121), 823–826. http://doi.org/10.1126/science.1232033

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nature Methods*, *10*(10), 957–63. http://doi.org/10.1038/nmeth.2649

Marks, M. S., Blum, J. S., & Cresswell, P. (1990). Invariant chain trimers are sequestered in the rough endoplasmic reticulum in the absence of association with HLA class II antigens. *The Journal of Cell Biology*, *111*(3), 839–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2391366

Marsh, S. G., Bodmer, J. G., Albert, E. D., Bodmer, W. F., Bontrop, R. E., Dupont, B., … Terasaki, P. I. (2001). Nomenclature for factors of the HLA system, 2000. *European Journal of Immunogenetics : Official Journal of the British Society for Histocompatibility and Immunogenetics*, *28*(3), 377–424. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11422419

Matsumura, M., Fremont, D. H., Peterson, P. A., & Wilson, I. A. (1992). Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science (New York, N.Y.)*, *257*(5072), 927–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1323878

Mellins, E. D., & Stern, L. J. (2014). HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. *Current Opinion in Immunology*, *26*, 115–22. http://doi.org/10.1016/j.coi.2013.11.005

Miller, M. M., Goto, R. M., Taylor, R. L., Zoorob, R., Auffray, C., Briles, R. W., … Bloom, S. E. (1996). Assignment of Rfp-Y to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(9), 3958–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8632997

Miller, M. M., Goto, R., Bernot, A., Zoorob, R., Auffray, C., Bumstead, N., & Briles, W. E. (1994). Two Mhc class I and two Mhc class II genes map to the chicken Rfp-Y system outside the B complex. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(10), 4397–401. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7910407

Miller, M. M., & Taylor, R. L. (2016). Brief review of the chicken Major Histocompatibility Complex: the genes, their distribution on chromosome 16, and their contributions to disease resistance. *Poultry Science*, *95*(2), 375–392. http://doi.org/10.3382/ps/pev379

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733–740. http://doi.org/10.1099/mic.0.023960-0

Mosyak, L., Zaller, D. M., & Wiley, D. C. (1998). The structure of HLA-DM, the peptide exchange catalyst that loads antigen onto class II MHC molecules during antigen presentation. *Immunity*, *9*(3), 377–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9768757

Mulder, D. J., Pooni, A., Mak, N., Hurlbut, D. J., Basta, S., & Justinich, C. J. (2011). Antigen Presentation and MHC Class II Expression by Human Esophageal Epithelial Cells: Role in Eosinophilic Esophagitis. *The American Journal of Pathology*, *178*(2), 744–753. http://doi.org/10.1016/J.AJPATH.2010.10.027

Neefjes, J., Jongsma, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*, *11*(12), 823–836. http://doi.org/10.1038/nri3084

Nelson, C. A., Viner, N., Young, S., Petzold, S., Benoist, C., Mathis, D., & Unanue, E. R. (1996). Amino acid residues on the I-Ak alpha-chain required for the binding and stability of two antigenic peptides. *Journal of Immunology (Baltimore, Md. : 1950)*, *156*(1), 176–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8598459

Nemudryi, A. A., Valetdinova, K. R., Medvedev, S. P., & Zakian, S. M. (2014). TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta Naturae*, *6*(3), 19–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25349712

O'Brien, C., Flower, D. R., & Feighery, C. (2008). Peptide length significantly influences in vitro affinity for MHC class II molecules. *Immunome Research*, *4*, 6. http://doi.org/10.1186/1745-7580-4-6

Olias, P., Adam, I., Meyer, A., Scharff, C., & Gruber, A. D. (2014). Reference Genes for Quantitative Gene Expression Studies in Multiple Avian Species. *PLoS ONE*, *9*(6), e99678. http://doi.org/10.1371/journal.pone.0099678

Ortmann, B., Copeman, J., Lehner, P. J., Sadasivan, B., Herberg, J. A., Grandea, A. G., … Cresswell, P. (1997). A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science (New York, N.Y.)*, *277*(5330), 1306–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9271576

Painter, C. A., Negroni, M. P., Kellersberger, K. A., Zavala-Ruiz, Z., Evans, J. E., & Stern, L. J. (2011). Conformational lability in the class II MHC 310 helix and adjacent extended strand dictate HLA-DM susceptibility and peptide exchange. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(48), 19329–34. http://doi.org/10.1073/pnas.1108074108

Park, S. J., Sadegh-Nasseri, S., & Wiley, D. C. (1995). Invariant chain made in Escherichia coli has an exposed N-terminal segment that blocks antigen binding to HLA-DR1 and a trimeric C-terminal segment that binds empty HLA-DR1. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(24), 11289–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7479981

Parker, A., & Kaufman, J. (2017). What chickens might tell us about the MHC class II system. *Current Opinion in Immunology*, *46*, 23–29. http://doi.org/10.1016/J.COI.2017.03.013

Patil, N. S., Hall, F. C., Drover, S., Spurrell, D. R., Bos, E., Cope, A. P., … Mellins, E. D. (2001). Autoantigenic HCgp39 epitopes are presented by the HLA-DM-dependent presentation pathway in human B cells. *Journal of Immunology (Baltimore, Md. : 1950)*, *166*(1), 33–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11123274

Pazderka, F., Longenecker, B. M., Law, G. R. J., & Ruth, R. F. (1975). The major histocompatibility complex of the chicken. *Immunogenetics*, *2*(1), 101–130. http://doi.org/10.1007/BF01572280

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., … Gersbach, C. A. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods*, *10*(10), 973–6. http://doi.org/10.1038/nmeth.2600

Peters, P. J., Neefjes, J. J., Oorschot, V., Ploegh, H. L., & Geuze, H. J. (1991). Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature*, *349*(6311), 669–76. http://doi.org/10.1038/349669a0

Peters, P. J., Neefjes, J. J., Oorschot, V., Ploegh, H. L., & Geuze, H. J. (1991). Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature*, *349*(6311), 669–676. http://doi.org/10.1038/349669a0

Petersen, J. L., Hickman-Miller, H. D., McIlhaney, M. M., Vargas, S. E., Purcell, A. W., Hildebrand, W. H., & Solheim, J. C. (2005). A charged amino acid residue in the transmembrane/cytoplasmic region of tapasin influences MHC class I assembly and maturation. *Journal of Immunology (Baltimore, Md. : 1950)*, *174*(2), 962–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15634919

Pink, J. R. L., Droege, W., Hála, K., Miggiano, V. C., & Ziegler, A. (1977). A three-locus model for the chicken major histocompatibility complex. *Immunogenetics*, *5*(1), 203–216. http://doi.org/10.1007/BF01570477

Plachý, J., Hála, K., Hejnar, J., Geryk, J., & Svoboda, J. (1994). src-specific immunity in inbred chickens bearing v-src DNA- and RSV-induced tumors. *Immunogenetics*, *40*(4), 257–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8082891

Ploegh, H. L., Orr, H. T., & Strominger, J. L. (1981). Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell*, *24*(2), 287–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7016338

Pos, W., Sethi, D. K., Call, M. J., Schulze, M.-S. E. D., Anders, A.-K., Pyrdol, J., & Wucherpfennig, K. W. (2012). Crystal structure of the HLA-DM-HLA-DR1 complex defines

mechanisms for rapid peptide selection. *Cell*, *151*(7), 1557–68. http://doi.org/10.1016/j.cell.2012.11.025

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–308. http://doi.org/10.1038/nprot.2013.143

Reith, W., LeibundGut-Landmann, S., & Waldburger, J.-M. (2005). Regulation of MHC class II gene expression by the class II transactivator. *Nature Reviews Immunology*, *5*(10), 793–806. http://doi.org/10.1038/nri1708

Riberdy, J. M., Newcomb, J. R., Surman, M. J., Barbosa, J. A., & Cresswell, P. (1992). HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature*, *360*(6403), 474–7. http://doi.org/10.1038/360474a0

Roche, P. A., & Cresswell, P. (2011). Proteolysis of the class II-associated invariant chain generates a peptide binding site in intracellular HLA-DR molecules. Proc. Natl. Acad. Sci. USA. 1991. 88: 3150-3154. *Journal of Immunology (Baltimore, Md. : 1950)*, *187*(3), 1076–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21772034

Rock, K. L., Reits, E., & Neefjes, J. (2016). Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends in Immunology*, *37*(11), 724–737. http://doi.org/10.1016/j.it.2016.08.010

Rogers, S., Shaw, I., Ross, N., Nair, V., Rothwell, L., Kaufman, J., & Kaiser, P. (2003). Analysis of part of the chicken Rfp-Y region reveals two novel lectin genes, the first complete genomic sequence of a class I alpha-chain gene, a truncated class II beta-chain gene, and a large CR1 repeat. *Immunogenetics*, *55*(2), 100–8. http://doi.org/10.1007/s00251-003-0553-1

Romieu-Mourez, R., François, M., Boivin, M.-N., Stagg, J., & Galipeau, J. (2007). Regulation of MHC class II expression and antigen processing in murine and human mesenchymal stromal cells by IFN-gamma, TGF-beta, and cell density. *Journal of Immunology (Baltimore, Md. : 1950)*, *179*(3), 1549–58. http://doi.org/10.4049/JIMMUNOL.179.3.1549

Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., & Cresswell, P. (1996). Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity*, *5*(2), 103–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8769474

Salomonsen, J., Chattaway, J. A., Chan, A. C. Y., Parker, A., Huguet, S., Marston, D. A., … Kaufman, J. (2014). Sequence of a Complete Chicken BG Haplotype Shows Dynamic Expansion and Contraction of Two Gene Lineages with Particular Expression Patterns. *PLoS Genetics*, *10*(6), e1004417. http://doi.org/10.1371/journal.pgen.1004417

Salomonsen, J., Marston, D., Avila, D., Bumstead, N., Johansson, B., Juul-Madsen, H., … Kaufman, J. (2003). The properties of the single chicken MHC classical class II ? chain (B-LA) gene indicate an ancient origin for the DR/E-like isotype of class II molecules. *Immunogenetics*, *55*(9), 605–614. http://doi.org/10.1007/s00251-003-0620-7

Schou, T. W., Permin, A., Juul-Madsen, H. R., Sorensen, P., Labouriau, R., Nguyen, T. L. H., … Pham, S. L. (2007). Gastrointestinal helminths in indigenous and exotic chickens in Vietnam: association of the intensity of infection with the Major Histocompatibility Complex. *Parasitology*, *134*(04), 561. http://doi.org/10.1017/S0031182006002046

Segal, D. J., & Meckler, J. F. (2013). Genome Engineering at the Dawn of the Golden Age. *Annual Review of Genomics and Human Genetics*, *14*(1), 135–158. http://doi.org/10.1146/annurev-genom-091212-153435

Sette, A., Ceman, S., Kubo, R. T., Sakaguchi, K., Appella, E., Hunt, D. F., … Rudersdorf, R. (1992). Invariant chain peptides in most HLA-DR molecules of an antigen-processing mutant. *Science (New York, N.Y.)*, *258*(5089), 1801–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1465617

Shah, S. A., Erdmann, S., Mojica, F. J. M., & Garrett, R. A. (2013). Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biology*, *10*(5), 891–9. http://doi.org/10.4161/rna.23764

Shaw, I., Powell, T. J., Marston, D. A., Baker, K., van Hateren, A., Riegert, P., … Kaufman, J. (2007). Different evolutionary histories of the two classical class I genes BF1 and BF2 illustrate drift and selection within the stable MHC haplotypes of chickens. *Journal of Immunology (Baltimore, Md. : 1950)*, *178*(9), 5744–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17442958

Sisk, T. J., Nickerson, K., Kwok, R. P. S., & Chang, C.-H. (2003). Phosphorylation of class II transactivator regulates its interaction ability and transactivation function. *International Immunology*, *15*(10), 1195–205. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13679389

Smith, M. A., Wright, G., Wu, J., Tailor, P., Ozato, K., Chen, X., … Wright, K. L. (2011). Positive Regulatory Domain I (PRDM1) and IRF8/PU.1 Counter-regulate MHC Class II Transactivator (CIITA) Expression during Dendritic Cell Maturation. *Journal of Biological Chemistry*, *286*(10), 7893–7904. http://doi.org/10.1074/jbc.M110.165431

Spies, T., Bresnahan, M., Bahrain, S., Arnold, D., Blanck, G., Mellins, E., … DeMars, R. (1990). A gene in the human major histocompatibility complex class II region controlling the class I antigen presentation pathway. *Nature*, *348*(6303), 744–747. http://doi.org/10.1038/348744a0

Springer, T. A., Kaufman, J. F., Terhorst, C., & Strominger, J. L. (1977). Purification and structural characterisation of human HLA-linked B-cell antigens. *Nature*, *268*(5617), 213–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/301990

Stebbins, C. C., Loss, G. E., Elias, C. G., Chervonsky, A., & Sant, A. J. (1995). The requirement for DM in class II-restricted antigen presentation and SDS-stable dimer formation is allele and species dependent. *The Journal of Experimental Medicine*, *181*(1), 223–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7807005

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature*, *368*(6468), 215–221. http://doi.org/10.1038/368215a0

Suri, A., Lovitch, S. B., & Unanue, E. R. (2006). The wide diversity and complexity of peptides bound to class II MHC molecules. *Current Opinion in Immunology*, *18*(1), 70–77. http://doi.org/10.1016/J.COI.2005.11.002

Tang, T.-H., Bachellerie, J.-P., Rozhdestvensky, T., Bortolin, M.-L., Huber, H., Drungowski, M., … Huttenhofer, A. (2002). Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. *Proceedings of the National Academy of Sciences*, *99*(11), 7536–7541. http://doi.org/10.1073/pnas.112047299

Tewari, M. K., Sinnathamby, G., Rajagopal, D., & Eisenlohr, L. C. (2005). A cytosolic pathway for MHC class II–restricted antigen processing that is proteasome and TAP dependent. *Nature Immunology*, *6*(3), 287–294. http://doi.org/10.1038/ni1171

Thelemann, C., Eren, R. O., Coutaz, M., Brasseit, J., Bouzourene, H., Rosa, M., … Acha-Orbea, H. (2014). Interferon-γ Induces Expression of MHC Class II on Intestinal Epithelial Cells and Protects Mice from Colitis. *PLoS ONE*, *9*(1), e86844. http://doi.org/10.1371/journal.pone.0086844

Ting, J. P.-Y., & Trowsdale, J. (2002). Genetic control of MHC class II expression. *Cell*, *109 Suppl*(2), S21-33. http://doi.org/10.1016/S0092-8674(02)00696-7

Tregaskes, C. A., Harrison, M., Sowa, A. K., van Hateren, A., Hunt, L. G., Vainio, O., & Kaufman, J. (2016). Surface expression, peptide repertoire, and thermostability of chicken class I molecules correlate with peptide transporter specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 692–7. http://doi.org/10.1073/pnas.1511859113

Trowsdale, J., Hanson, I., Mockridge, I., Beck, S., Townsendt, A., & Kelly, A. (1990). Sequences encoded in the class II region of the MHC related to the “ABC” superfamily of transporters. *Nature*, *348*(6303), 741–744. http://doi.org/10.1038/348741a0

Umetsu, D. T., Katzen, D., Jabara, H. H., & Geha, R. S. (1986). Antigen presentation by human dermal fibroblasts: activation of resting T lymphocytes. *Journal of Immunology (Baltimore, Md. : 1950)*, *136*(2), 440–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3484491

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636–646. http://doi.org/10.1038/nrg2842

Vainio, O., Veromaa, T., Eerola, E., Toivanen, P., & Ratcliffe, M. J. (1988). Antigen-presenting cell-T cell interaction in the chicken is MHC class II antigen restricted. *Journal of Immunology (Baltimore, Md. : 1950)*, *140*(9), 2864–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3258887

van der Poel, J. J., Groenen, M. A., Dijkhof, R. J., Ruyter, D., & Giphart, M. J. (1990). The nucleotide sequence of the bovine MHC class II alpha genes: DRA, DOA, and DYA. *Immunogenetics*, *31*(1), 29–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2298463

van Hateren, A., Bailey, A., Werner, J. M., & Elliott, T. (2015). Plasticity of empty major histocompatibility complex class I molecules determines peptide-selector function. *Molecular Immunology*, *68*(2 Pt A), 98–101. http://doi.org/10.1016/j.molimm.2015.03.010

Van Kaer, L., Ashton-Rickardt, P. G., Ploegh, H. L., & Tonegawa, S. (1992). TAP1 mutant mice are deficient in antigen presentation, surface class I molecules, and CD4-8+ T cells. *Cell*, *71*(7), 1205–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1473153

Vilen, B. J., Cogswell, J. P., Ting, J. P., Strominger, J. L., Merika, M., Papamatheakis, J., & Ting, J. P. (1991). Stereospecific alignment of the X and Y elements is required for major histocompatibility complex class II DRA promoter function. *Molecular and Cellular Biology*, *11*(5), 2406–15. http://doi.org/10.1128/mcb.20.17.6185-6194.2000

Viville, S., Neefjes, J., Lotteau, V., Dierich, A., Lemeur, M., Ploegh, H., … Mathis, D. (1993). Mice lacking the MHC class II-associated invariant chain. *Cell*, *72*(4), 635–648. http://doi.org/10.1016/0092-8674(93)90081-Z

Vyas, J. M., Van der Veen, A. G., & Ploegh, H. L. (2008). The known unknowns of antigen processing and presentation. *Nature Reviews Immunology*, *8*(8), 607–618. http://doi.org/10.1038/nri2368

Walker, B. A., Hunt, L. G., Sowa, A. K., Skjødt, K., Göbel, T. W., Lehner, P. J., & Kaufman, J. (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), 8396–401. http://doi.org/10.1073/pnas.1019496108

Wallny, H.-J., Avila, D., Hunt, L. G., Powell, T. J., Riegert, P., Salomonsen, J., … Kaufman, J. (2006). Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1434–9. http://doi.org/10.1073/pnas.0507386103

Williams, A., Peh, C. A., & Elliott, T. (2002). The cell biology of MHC class I antigen presentation. *Tissue Antigens*, *59*(1), 3–17. http://doi.org/10.1034/j.1399-0039.2002.590103.x

Xiu, F., Côté, M.-H., Bourgeois-Daigneault, M.-C., Brunet, A., Gauvreau, M.-É., Shaw, A., & Thibodeau, J. (2011). Cutting edge: HLA-DO impairs the incorporation of HLA-DM into exosomes. *Journal of Immunology (Baltimore, Md. : 1950)*, *187*(4), 1547–51. http://doi.org/10.4049/jimmunol.1100199

Yuhki, N., Beck, T., Stephens, R. M., Nishigaki, Y., Newmann, K., & O'Brien, S. J. (2003). Comparative genome organization of human, murine, and feline MHC class II region. *Genome Research*, *13*(6A), 1169–79. http://doi.org/10.1101/gr.976103

Zhang, Y., Baig, E., & Williams, D. B. (2006). Functions of ERp57 in the Folding and Assembly of Major Histocompatibility Complex Class I Molecules. *Journal of Biological Chemistry*, *281*(21), 14622–14631. http://doi.org/10.1074/jbc.M512073200

Zoorob, R., Béhar, G., Kroemer, G., & Auffray, C. (1990). Organization of a functional chicken class II B gene. *Immunogenetics*, *31*(3), 179–87. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1969383

10. Acknowledgements

Firstly, I would like to acknowledge my supervisor, mentor and much more, Jim Kaufman. Jim has provided me with much support, advice and guidance (in nearly every way), in and out of the lab, at anytime and regardless of how busy he is. I would like to say a big thank you to Jim for getting me to this point!

In general, I would like to say thank you to the whole Kaufman lab, members past and present. In particular, Aimee Parker, Andrew Chan, John Chattaway, Hannah Siddle, Clive Tregaskes, Samer Halabi, Flic Coulter, Tom Yip, Magdalena Migalska, Becky Martin and Lei Chen. All of whom have helped and supported me in a multitude of ways. I would also like to thank the whole department, academic and non-academic staff for facilitating my studies and creating a supportive environment. Especially, Melvin and Gerry from building maintenance for going above and beyond to resolve any issues that arose, as well as providing me with much needed distraction at various points of stress.

I would like to give a special mention to the following people. Andreas Neerincx for many useful discussions and being a great friend. Ben Trigg for providing daily comedy and running advice. My mentor John Doorbar, who's door was always open and was very engaging, helpful and supportive. Mara Sgroi for much patience and care (during what has been a particularly stressful period) and generally being one of the most wonderful people I have ever met.

Two friends deserve a big thank you, Paris Flynn and Michael Adamson. Both of which have been there for me in times of need over many many years. Providing me with nonsense whimsical conversation and real-world advice when required. Both of whom are fantastic people and are incredibly appreciated.

I would also like to thank my family who have been incredibly supportive and caring during this process. Most importantly my brother Steven Wise. Steve has been there for me no matter what I have done or where I have been, and has always gone above and beyond to help me in any way he could, regardless of how difficult it was. The level of commitment and loyalty shown by Steve is something I feel is truly rare, so thank you.

Lastly, I would like to dedicate my thesis to my son, Solomon Wise.