Generation and characterization of AB*ab*.I transgenic mice: a novel tool for the isolation of human T cell receptors selected on a distinct human HLA class I haplotype for adoptive T cell therapy of cancer

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

ARUNRAJ DHAMODARAN

from Tamil Nadu, India

Berlin, 2021

TO MY DEAREST GRANDFATHER, KANNAN. R,
WHO PASSED AWAY FROM CANCER

This Doctoral dissertation project was conducted under the supervision of Prof. Dr. Thomas Blankenstein at the Max Delbrück Center for Molecular Medicine (MDC), Berlin. The project was performed from March 2013 until October 2020.

1st Reviewer: Prof. Dr. Thomas Blankenstein

2nd Reviewer: Prof. Dr. Gerald Willimsky

Date of Defense: 31.08.2021

Table of contents

Abstract

For efficacious adoptive T cell therapy (ATT), the appropriate selection of tumor-specific antigens (TSAs) is crucial. Recurrent somatic driver mutations yield ideal TSAs for ATT to achieve complete tumor elimination, possibly without relapse. *In silico* algorithms predict epitopes as TSAs for ATT, however, with low probability. Thus, the accuracy of prediction algorithms still needs validation in describing processed neoepitopes.

Simultaneous validation of proteasomal processing of predicted TSAs and identification of neoantigen-specific T cell receptors (TCRs) is possible from transgenic mice. One such model developed by Li *et al*. in 2010 was the AB*ab*DII mice having the complete human TCR gene loci, knocked out for murine TCR and MHC class I gene loci, and it is transgenic for human HLA class I allele, HLA-A*02:01 (HLA-A2). Though isolation of human TCRs is possible from AB*ab*DII mice, the model is limited as a source of obtaining only HLA-A2-restricted TCRs, thereby restricting epitope identification to HLA-A2 antigens.

In this doctoral study, a novel mouse model with six human HLA alleles as a single haplotype, named AB*ab*.I, was developed on AB*ab*DII background to diversify HLA genes in the existing AB*ab*DII mice for epitope discovery and broaden TCR repertoire for non-HLA-A2 restricted TCR isolation. *In silico* screening for putative neoepitopes predicted to bind high-ranked HLA alleles selected the six alleles in AB*ab*.I mice; novel class I haplotype like in humans as every individual bears six HLA class I alleles.

Initially, the *in silico* screening predicted 23 and 152 high-affinity TSAs ($IC_{50} < 50$ nM) from recurrent point mutations (n=266) that bound HLA-A2, and other frequent class I alleles (n=18), respectively.

In the first phase of this doctoral thesis, epitope immunogenicity was tested in 4 out of 23 HLA-A2-binders in AB*ab*DII mice. Only 1 out of 4 immunized epitopes elicited $CD8^+$ T cell (CTL) responses. However, this was confirmed to be not endogenously processed when tested with TCRs raised from AB*ab*DII mice. This form of reverse immunology is laborious and still a concern with open questions. A previous study on identifying epitopes from the vaccinia virus by Assarsson *et al*. in 2007 aligns with the *in silico* screen predicted data examined in this thesis. Thus, prediction algorithms offer a low probability to select immunogenic epitopes.

In the second phase, the AB*ab*.I mouse model with a novel set of chimeric class I fusion alleles, HLA-A*03:01, A*11:01, B*07:02, B*15:01, C*04:01, and C*07:02, was developed using PiggyBac transposon strategy. The introduction of six HLA alleles as a single genotype resembling a natural human situation broadened the TCR repertoire four-fold by enriching the peripheral pool with five times more unique and rarer V(D)J-TCRß clonotypes than AB*ab*DII mice. Ultimately, AB*ab*.I mice would serve as a versatile *in vivo* model system for epitope discovery, besides being a valuable tool to identify novel TCRs against high HLA-affinity tumor-specific antigens ($IC_{50} < 50$ nM) derived from recurrent somatic point mutations.

Zusammenfassung

Für eine wirksame adoptive T-Zell-Therapie (ATT) ist die Wahl geeigneter tumorspezifischer Antigene (TSAs) entscheidend. Sogenannte „Treiber-Mutationen" - somatische Mutationen, die an der Tumorentstehung beteiligt sind und gehäuft in Tumoren auftreten, stellen ideale TSAs für eine ATT mit dem Ziel einer vollständigen Tumoreliminierung, und bestenfalls ohne Rückfall, dar. *In silico*-Algorithmen können Epitope als TSAs für die ATT vorhersagen, jedoch nur mit geringer Wahrscheinlichkeit. Eine experimentelle Validierung dieser Vorhersagen und Bestätigung tatsächlich prozessierter Neoepitope ist somit weiterhin notwendig.

Die gleichzeitige Validierung der Prozessierung vorhergesagter TSAs und die Isolierung neoantigenspezifischer T-Zell-Rezeptoren (TCRs) ist in transgenen Mäusen möglich. Ein solches transgenes Model ist die AB*ab*DII-Maus, die von Li *et al*. entwickelt wurde. Diese verfügt über die vollständigen humanen TCR-Genloci sowie eines knock-out der murinen TCR- und MHC-Klasse-I-Loci und ist transgen für das humane HLA-Allel A*02:01 (HLA-A2). Auch wenn die Isolierung von humanen TCRs aus den AB*ab*DII-Mäusen möglich ist, ist das Modell auf HLA-A2-restringierte TCRs und die Identifikation HLA-A2-bindender Epitope begrenzt.

In dieser Arbeit wurde ein neues Modell (AB*ab*.I) mit einem einzelnen Haplotyp bestehend aus sechs humanen HLA-Allelen auf Basis der AB*ab*DII-Maus entwickelt, um die Identifikation von neuen nicht-HLA-A2-restringierten Epitopen und TCRs zu ermöglichen. Durch in silico-Screening nach mutmaßlichen Neoepitopen, denen eine Bindung an häufig vorkommende HLA-Allele vorhergesagt wurde, wurden die sechs Allele der AB*ab*.I-Maus ausgewählt; ein neuartiger Klasse-I-Haplotyp mit, wie im Menschen vorkommend, sechs HLA-Klasse-I-Allelen. Zunächst wurden durch das *in silico*-Screening 23 und 152 hochaffine TSAs (IC$_{50}$ < 50 nM) aus gehäuft in Tumoren auftretenden Punktmutationen (n=266) vorhergesagt, die HLA-A2 oder andere häufige Klasse-I-Allele (n=18) binden.

Im ersten Teil dieser Arbeit wurde die Immunogenität von vier der 23 HLA-A2-Binder in AB*ab*DII-Mäusen getestet. Nur eines von vier immunisierten Peptiden löste eine CD8$^+$ T-Zell-Antwort aus. Eine Testung mit aus AB*ab*DII-Mäusen isolierten TCRs ergab jedoch, dass dieses Epitop nicht endogen prozessiert wird. Diese Form der reversen Immunologie ist aufwendig und hinterlässt häufig offene Fragen. Die Vorhersagealgorithmen bieten daher eine zu geringe Wahrscheinlichkeit, immunogene Epitope zu identifizieren.

Im zweiten Teil wurde das AB*ab*.I-Mausmodell mit einem neuartigen Satz von chimären Klasse-I-Fusionsallelen – HLA-A*03:01, A*11:01, B*07:02, B*15:01, C*04:01, und C*07:02 – mit Hilfe der PiggyBac-Transposon-Strategie entwickelt. Die Einführung von sechs HLA-Allelen erweiterte das TCR-Repertoire um das Vierfache, der periphere T-Zell-Pool enthielt fünfmal mehr einzigartige und seltenere V(D)J-TCRß-Klonotypen als der in AB*ab*DII-Mäusen. Die AB*ab*.I-Mäuse können als vielseitiges *in vivo*-Modellsystem für die Entdeckung von

Epitopen dienen und darüber hinaus ein wertvolles Werkzeug zur Isolierung neuartiger TCRs gegen tumorspezifische Antigene mit hoher HLA-Affinität ($IC_{50} < 50$ nM) sein.

List of Figures

List of Tables

List of Abbreviations

Abbreviation/Symbol	Stands for/Meaning
TCR	**T C**ell **R**eceptor
pMHC	**p**eptide bound **M**ajor **H**istocompatibility **C**omplex
pHLA	**p**eptide bound **H**uman **L**eukocyte **A**ntigen
CDR	**C**omplementarity **D**etermining **R**egion
ATT	**A**doptive **T**ransfer **T**herapy
TSA	**T**umor-**S**pecific **A**ntigen
TAA	**T**umor-**A**ssociated **A**ntigen
BCR	**B C**ell **R**eceptor
PD-1	**P**rogrammed cell **D**eath protein-**1**
TCR	**T C**ell **R**eceptor
MAGE-1	**M**elanoma-associated **A**ntigen-**1**
NY-ESO-1	**N**ew **Y**ork **ESO**phageal Squamous cell carcinoma-**1**
APC	**A**ntigen **P**resenting **C**ell
aa	**a**mino **a**cid

Abbreviation/Symbol	Stands for/Meaning
mIL-2/15	**m**urine **I**nter**L**eukin-2/15
mIFN-γ	**m**urine **I**nter**F**ero**N**-gamma
KO	**K**nock-**O**ut
CTL/CD8$^+$ T cell	**C**ytotoxic **T** **L**ymphocyte
APC	**A**ntigen **P**resenting **C**ell
T$_H$	**H**elper CD4$^+$ T cell
ER	**E**ndoplasmic **R**eticulum
DMEM	**D**ulbecco's **M**odified **E**agle **M**edium
RPMI	**R**oswell **P**ark **M**emorial **I**nstitute
DC	**D**endritic **C**ell
PLC	**P**eptide **L**oading **C**omplex
V (D) J	**V**ariable **D**iversity **J**oining
DMSO	**Di**M**ethyl **S**ulf**O**xide
FBS	**F**etal **B**ovine **S**erum
HEPES	4-(2-**H**ydroxy-**E**thyl)-1-**P**iperazine-**E**thane-**S**ulfonic Acid

Abbreviation/Symbol	Stands for/Meaning
DPBS	**D**ulbecco's **P**hosphate **B**uffered **S**aline
DTT	**Di**Thio**T**hreitol
$MgCl_2$	**M**agnesium **C**hloride
NH_4Cl	**A**mmonium **C**hloride
$KHCO_3$	**P**otassium **B**icarbonate
EDTA	**Et**hylene**D**iamine**T**etra**A**cetic acid
IFA	**I**ncomplete **F**reund's **A**djuvant
TAP	**T**ransport associated with **A**ntigen **P**rocessing
ERAP	**E**ndoplasmic **R**eticulum **A**mino**P**eptidase 1
FACS	**F**luorescence-**A**ctivated **C**ell **S**orting
ACK	**A**mmonium-**C**hloride-**P**otassium
CRT	**C**al**R**e**T**iculin
BM	**B**one **M**arrow
TGF-ß	**T**ransforming **G**rowth **F**actor-**ß**
TNF- α	**T**umor **N**ecrosis **F**actor

Abbreviation/Symbol	Stands for/Meaning
PE	**P**hyco**E**rythrin
APC	**A**llo**P**hyco**C**yanin
TLR	**T**oll-**L**ike **R**eceptor
IFN	Inter**F**ero**N**
PB	**P**acific **B**lue
FITC	**F**luorescein **I**so**T**hio**C**yanate
MCA	**M**ethyl**C**hol**A**nthrene
α/ß TCR	**A**lpha/**B**eta TCR
EBV	**E**pstein **B**arr **V**irus
CMV	**C**yto**M**egalo**V**irus
10^{-6} M or 10^{-9} M, M is Molarity	1 **µ**M or 1 **n**M
IMGT	**I**m**M**uno**G**ene**T**ics
SV40	**S**imian **V**irus family **40**
LN/s	**L**ymph **N**ode/s
SPL/s	**SPL**een/s

Abbreviation/Symbol	Stands for/Meaning
eGFP	**e**nhanced **G**reen **F**luorescent **P**rotein
RT	**R**oom **T**emperature
CAR T cells	**C**himeric **A**ntigen **R**eceptor-engaged **T** cells
scFv	**S**ingle-**c**hain **V**ariable **F**ragment
S722F	mutation of **Serine** at 722 into **Phenylalanine**
shRNA	**s**hort **h**airpin RNA
IC$_{50}$	**I**nhibitory **C**oncentration, a concentration of a drug or peptide that gives a half-maximal response
MFI	**M**ean **F**luorescence **I**ntensity
P value	**P**robability value
± SD	**Plus/Minus S**tandard **E**rror of the **M**ean
ANN	**A**rtificial **N**eural **N**etwork
mTCM	**m**ouse **T C**ell RPMI **M**edium
FACS	**F**luoresence **A**ctivated **C**ell **S**orting
MACS	**M**aagnetic **A**ctivated **C**ell **S**orting
CRISPR	**C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats

Abbreviation/Symbol	Stands for/Meaning
IL	**I**nter**L**eukin
LCL	**L**ymphoblastoid **C**ell **L**ine
HuTCR	**Hu**manized **TCR,** referring AB*ab*DII mice
SOE	**S**plicing by **O**verhang **E**xtension
qRT-PCR	**q**uantitative **R**eal-**T**ime **PCR**
PB	**P**iggy**B**ac
ITRs	**I**nverted **T**erminal **R**epeat**s**
ivt-mRNA	**i**nvitro **t**ranscribed **m**essenger **RNA**
Plat-E	**P**latinum-**E**cotropic cell line
ps	**p**rotamine **s**ulfate
ELISA	**E**nzyme-**L**inked **I**mmuno**S**orbent **A**ssay
TdT	**T**erminal **d**eoxynucleotidyl **T**ransferase
TRRAP	**TR**ansformation/t**R**anscription domain-**A**ssociated **P**rotein
RACE	**R**apid **A**mplification of **c**DNA **E**nds
pA	**p**oly**A**denylation

Abbreviation/Symbol	Stands for/Meaning
bGH	**b**ovine **G**rowth **H**ormone
ß2m	**ß**eta**2 m**icroglobulin protein
TRBV	**TCR B**eta **V**ariable chain
mCALR	**m**utated **CALR**eticulin
HITI	**H**omology-**I**ndependent **T**argeted **I**nsertion

1. Introduction

1.1 Adaptive arm of the immune system

The immune system provides the body a protective barrier by fighting against foreign invasions like bacteria, viruses, and parasites using innate and adaptive defense mechanisms[1]. Innate immunity is the body's non-specific first line of defense in the detection and destruction of pathogens. Toll-like receptors (TLRs) sense pathogenic agents and recruit immune cells to the site for elimination by creating an inflammatory milieu together with macrophage-induced phagocytosis[2]. Dendritic cells (DCs) are professional antigen-presenting cells (APCs) which present endogenous and exogenous antigens to major histocompatibility molecules (MHC) I and II and act as a bridge in the recruitment of the adaptive immune system[3–5]. Several proinflammatory cues with the activation of multiple pathways like nuclear factor-κB (NF-κB), interferon (IFN)-regulatory factor (IRF), and activator protein-1 (AP-1) recruit the cells of the adaptive immune system[3,6]. Even though the innate immune mechanisms act promptly with proinflammatory cytokine release such as IFN-α, -ß (type I), and -γ (type II), cell-mediated adaptive immunity is the most effective arm of the immune system in eliminating pathogens[7,8]. The adaptive arm of the immune system consists of two cell-mediated responses, T and B cell-mediated immunity[1]. T and B lymphocytes differ in their specificity and mechanisms of action. B cells, precursors to antibody-secreting cells, recognize pathogenic antigens through B cell receptor (BCR) and support the recruitment of helper CD4[+] T cells through MHC class II proteins to the site of infections[9,10]. Naive CD8[+] T cell activation happens through the classical MHC class I presentation pathway by antigen-presenting DCs[11,12]. CD4[+] T cells independently play a crucial role in cytotoxic CD8[+] T cell activation by providing a potent signal to antigen-presenting DCs via CD40L-CD40 interaction[13–16]. Furthermore, the CD4[+] T cells help memory formation of the cytotoxic CD8[+] T cells by direct interaction through the CD40 receptor[17].

Both CD8[+] and CD4[+] conventional naive T cells interact with its cognate antigens presented on polymorphic MHC class I and II proteins through αß T cell receptors (TCRs)[18,19]. The naive TCR repertoire formed in the thymus primarily determines the precursor frequencies and distribution of distinct T cell clones[20–22]. Quantification of TCR repertoire might provide insights into the diversity of T cell-mediated responses that an organism can mount[20,23].

1.1.1 T cell receptor (TCR) and its diverse repertoire shapes T cell-mediated immunity

The αß TCRs potentially produce up to 10^{20} (theoretical) combinations by random yet abundant variable (V), diversity (D), and joining (J) recombination events in the thymus[23]. TCR αß[+] T cells undergo positive and negative selection, thymocytes are positively selected for adequate self-MHC interaction, whereas negatively selected for too high affinity towards any self-peptide MHC (pMHC)[24]. ~ 7.5% of thymocytes produce thymic signaling, from which ~ 3-5% survive

selection to populate the periphery as naive T cells[25]. Such T cells with a diverse array of αß TCRs interact intrinsically to pMHC (with co-stimulation via CD3-γ, -δ, -ε, -ζ, CD4 and CD8 chains) shape the naive CD8[+]/CD4[+] repertoire to encounter foreign cognate antigens for T cell-mediated responses[20,26,27]. MHC reactivity to ~ 5-20% of the preselection pool and ~ 10% of the peripheral T cell pool explains its intrinsic affinity towards TCRs[28,29]. Additional factors such as the complementarity determining region (CDR) 1 and CDR2 encoded by TCR V_{α} and $V_{ß}$ (germline-encoded) further instruct the germline bias of TCRs' bias towards MHC[30–33].

Some exceptions exist in which CDR1 and 2 are not the major contact domains to MHCs[34]. Crystallographic structures show unconventional diagonal topology onto MHC, in which CDR1 and 2 of the TCR Vα and $V_{ß}$ dock MHC class I's α2 and α1 helix or MHC II's ß and α helix[35]. In such interactions, hypervariable CDR3 loops of V_{α} and $V_{ß}$ are the main docking points to peptide-bound MHC (pMHC)[36,37]. Although existing pieces of evidence have shown that TCR and MHC genes inherently coevolve to interact with each other[21], distinct antigenic peptides processed and presented through class I/II antigen presentation pathways stringently define potential T cell effector functions through their CDR3α and CDR3ß interaction[38,39].

Thus, abundantly recombined CDR3-V_{α} and -$V_{ß}$ clonotypes form a diverse TCR repertoire that enables unique pMHC interactions driving T cell-mediated immunity.

1.1.2 pMHC-I/HLA-I-restricted antigen presentation pathway

The MHC glycoproteins, known to be human leukocyte antigen (HLA) in humans, processes antigens and presents 8-15 amino acid (aa) in case of class I or 12-25 aa (8-10 aa core motif) in case of class II peptides on its surface[1,40,41]. Two antigen presentation pathways assist the surface presentation of processed epitopes, peptide-loaded HLA-I (pHLA-I) and pHLA-II. HLA class I present endogenously processed epitopes like somatic point mutant, virus-derived, and aberrant frameshift proteins on its α2 and α1 domains to cytotoxic CD8[+] T cells (CTLs)[42,43]. HLA class II alleles present exogenously engulfed and processed epitopes like extracellular matrix proteins from the stroma and shredded viral proteins on its ß1 and α1 helix to helper CD4 T cells (T_H cells)[44]. Professional APCs can take up exogenous antigens into the endogenous HLA I pathway by cross-presentation[13]. Co-receptor interaction of CD8 to α3 domain of HLA I and CD4 to ß2 helix of HLA II is critical for T cell function[45].

Cellular, viral, aberrantly mutated, and misfolded post-translational proteins in the cytoplasm undergo proteasomal degradation[46]. After proteasomal processing, the fragmented peptides enter into the endoplasmic reticulum (ER) through peptide-loading complex (PLC) proteins such as transporter-associated protein (TAP) and the chaperone tapasin. ER aminopeptidase associated with antigen processing (ERAAP) protein trims the amino terminus of the ER-resident peptides before loading onto MHCs. Calreticulin (CRT), Erp57, and PLC load the processed peptides onto MHC/HLAs[47–49]. pMHC-I/HLA-I complexes are released into early

endosomes and then transit to the cell surface through late endosomal compartments for T cell recognition by CD8[+] T cells (CTLs)[50].
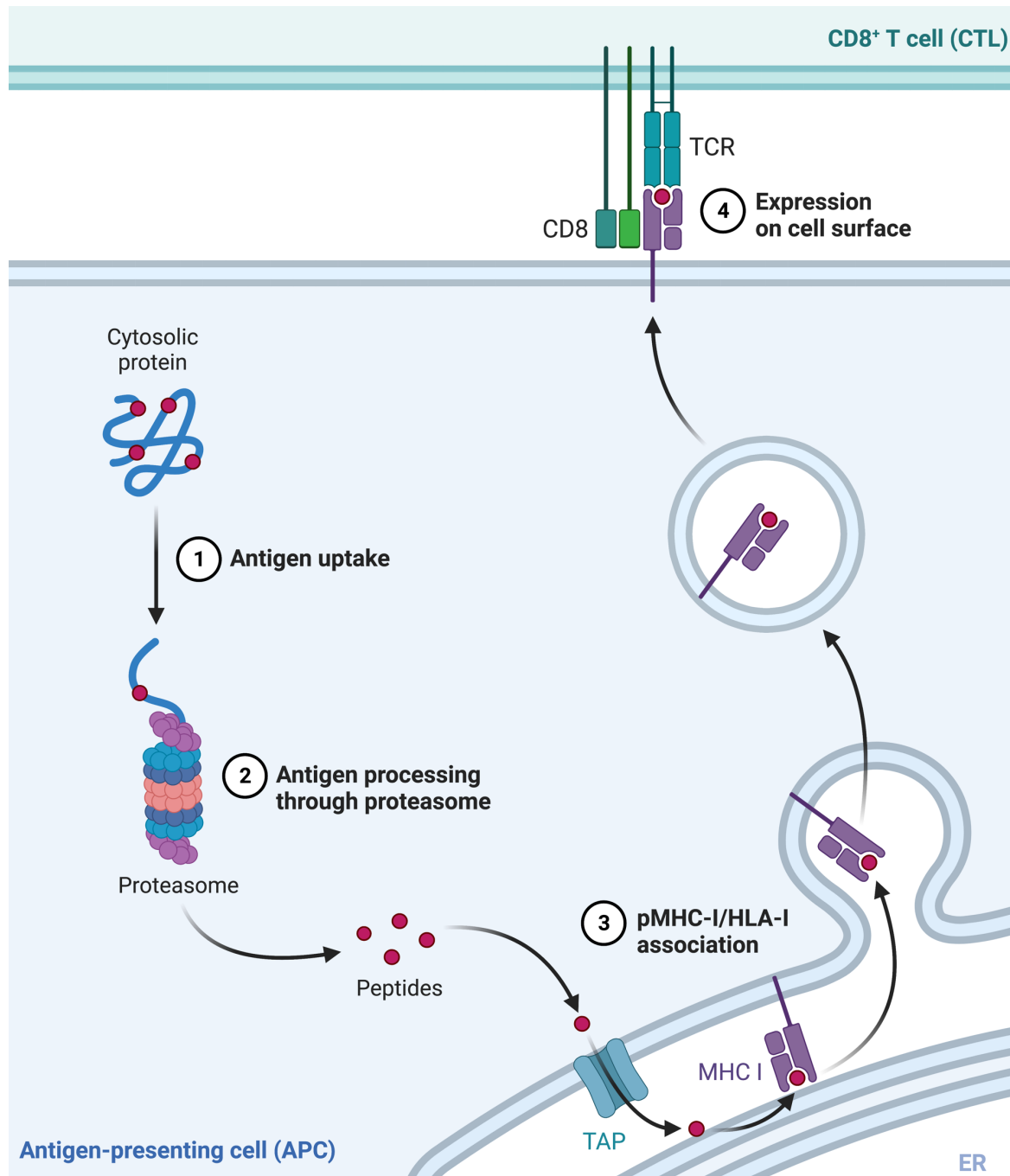
Figure 1: Schematic representation of the pMHC-I/HLA-I antigen processing pathway*.

1. Antigen uptake in the cytoplasm by the proteasome. 2. Proteins are processed by protease activity inside the proteasomal core. 3. TAP protein complex facilitate the transport of peptides into ER. PLC, Erp57, and CRT proteins instruct the peptide loading onto MHC/HLAs. 4. Peptide transport through the secretory pathway to the cell surface for presentation to CTLs.

* This simplified version does not represent all codomains and adapter proteins involved as chaperones in the antigen processing pathway. Graphic created using the BioRender software.

1.2 Tumor Immunology

1.2.1 Immune system and cancer

Cancer is a widely recognized genetic disease since oncogenic mutations, chromosomal rearrangements, and aberrations in gene expression are the most common causes detected in nearly one million cases[51,52]. Whole sequencing studies of primary lung tumor samples identified more than 50,000 point mutations (adjacent non-tumor tissue as controls), including some mutational hot-spots, for example, Gly[12] site in the RAS gene[53]. Emerging evidence proved that all cancers evolve by accumulating a small fraction of driver mutations that provide fitness and large numbers of passenger mutations that do not offer growth advantage[54,55]. For decades, research into understanding the nature of cancer focused primarily on tumor cells alone[56]. Oncogenic driver mutations make cancers thrive but not independently without additional factors supporting growth. Studies have confirmed the heterogeneous nature of tumor tissue, which makes dynamic interactions with the surrounding microenvironment, often regulate cancer progression[57,58]. Tumor microenvironment with fibroblasts, inflammatory cells, vascular endothelial cells, and the extracellular matrix comprises the so-called 'stroma'[59,60]. It is well described in the literature that a chronic inflammatory state in the stroma promotes cancer development accumulated with inflammatory cells[61,62]. One example being the stromal fibroblasts, in which the abrogation of transforming growth factor-ß (TGF-ß) signaling induced stomach and prostate tumors[63]. Under inflammatory states of progressing tumors, similar to wound healing, fibroblasts, macrophages, and endothelial cells are involved actively in assisting new blood vessel formation (as a part of stroma) to nourish tumors; a process known as 'neo-angiogenesis'[64–66].

$CD8^+$ and $CD4^+$ T cells require tumor-restricted peptide-loaded HLA-I (pHLA-I) and pHLA-II recognition, respectively for tumor rejection[67]. Though $CD8^+$ T cells (CTLs) can elicit cytotoxic activity solely in most cases, in some models, $CD4^+$ T cells reject tumors in the absence of CTLs[68]. CTLs release cytokines such as interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α) to reject tumors by attacking tumor stroma[67] and in some cases release cytotoxic molecules like perforin for tumor elimination[69–71], especially in virus-transformed tumors[72]. Studies from many analyzed models, T cells use IFN-γ as the critical effector molecule for tumor rejection[67]. In 2017, a report by Kammertoens *et al*. deduced the mechanism of IFN-γ on the tumor stroma, specifically on endothelial cells. The T cells through IFN-γ induction cause tumor ischemia showcasing a physiological state of blood vessel regression, thereby inhibiting neo-angiogenesis[73].

Another topic that links the immune system and cancer is the 'immune surveillance' concept[74]; still, a controversial theme that requires more evidence. In 2002, a study by Dunn *et al*., based on methylcholanthrene (MCA)-induced spontaneous tumor model claimed to have proved

cancer immune surveillance of tumor rejection and growth control by T cells[75]. However, follow-up studies disproved the surveillance concept of tumor rejection as T cell-mediated[76,77]. The first report proposed several other reasons for tumor rejection, such as tissue repair at the MCA site, the difference in steady-state IFN-γ concentrations between control versus experimental groups, biased IFN-γ protection due to different housing conditions of animal facilities. Further, tumor development in such models was called into question because either chemical-induced chronic inflammation or opportunistic infection could have been the reason for spontaneous growth, which did not resemble a naturally growing tumor in humans[76]. In 2005, Willimsky *et al*. developed a sporadic tumor model induced by a viral transgenic dormant oncogene (reflecting a physiological tumor) to study the influence of T cells recognizing tumor cells in the context of immune surveillance. This study proved that immunogenic sporadic tumors do not escape T cell recognition but rather avoid T cell destruction by inducing tolerance, thus proposing the tumor immune surveillance effect towards virus-associated cancers is possible[77]. Besides SV40 model[77], EBV[78]- and HPV[79]-associated tumors could be monitored by T cells, whereas immune surveillance of non-viral tumors needs more solid experimental evidence. Moreover, cautious data interpretation is necessary for reports[80,81] that show the concept of immune surveillance in non-viral-related cancers.

In summary, the effect of T cells in attacking tumor cells is evident[67,68,73]. Irrespective of T cell immune surveillance and its control over growing tumors, treatment option based on TCR-engineered CD8[+] and CD4[+] T cells using adoptive T cell therapy could give us the best possible chance for tumor rejection.

1.2.2 Adoptive T cell therapy (ATT) for cancer

In the past, many immunologists have been focusing on targeting self-antigens that were not tumor-specific using therapeutic vaccinations[82], however, with notable exceptions[83–85]. Therapeutic vaccines against tumor-associated shared antigens by active and passive cancer immunotherapy did not yield expected success due to established tolerance mechanisms, suppression, and tumor micro-environmental factors[86–89]. In contrast to cancer vaccinations, adoptive T cell therapy (ATT) has the possibility of rejecting large established tumors[90–93], specifically TCR-engineered CTLs (FasL[+]) releasing effector molecules such as IFN-γ[94], but not perforin, have the potential to destroy surrounding stroma[73,95] with complete eradication efficiency. Other than TCR-modified CTLs, the use of chimeric antigen receptor (CAR) engaged T cells (CAR T cells) in non-solid tumor clinical studies have shown success[96–98]. Early phase clinical studies with CAR T cells targeted CD19 surface antigen in hematological malignancies[97,99–101]. Though initial results targeting surface antigens in liquid tumors invited attention, challenges still exist in achieving similar success by targeting solid tumors.

CAR T cells have several obstacles to overcome to show efficacy without causing toxicity in solid cancers[102,103]. These include single-chain variable fragments (scFvs) on CAR T cells to target unique tumor-associated antigens (TAAs) on the tumor cell surface without antigen loss[104]. Next, CAR-engaged T cells should migrate to distant tumor sites for efficient homing. Besides this, circumvention of the immunosuppressive tumor microenvironment by T cells poses additional difficulties[105]. A clinical study targeting folate receptor-α in ovarian cancers reported limited T cell homing to solid tumor sites[106], which can be enhanced by co-expressing the chemokine receptor 2 (CXCR2) in CAR T cells towards Gro-α chemokine[107]. Blockade of TGF-ß signaling in prostate-specific membrane antigen targeting CAR T cells by co-expression of dominant-negative RII receptor increased their ability to infiltrate tumor stroma[108]. In another clinically relevant pleural mesothelioma study in mice, CD28 CAR T cells restored their effector function through co-transduction of a dominant-negative programmed death-1 (PD-1) receptor or > 60% knockdown of PD-1 expression by short hairpin RNAs (shRNAs)[109].

Although lessons learned from clinical studies on improving antigen selection[110], T cell migration, homing, and efficacy[102–109,111] could apply to both CAR-based and TCR-modified adoptive T cell therapy, the most crucial concern is always 'safety'[90]. TCR-modified T cells stand out in terms of safety because such T cells can precisely recognize endogenously processed antigens on tumor cells or stroma through HLA-I or HLA-II-restricted manner. In ATT, to avoid adverse autoimmune effects by reducing the chance of cross-reactivity against unintended self-targets[112], the use of TCR-modified T cells gives us a superior advantage of targeting tumor-specific antigens[90]. A unique class of such truly tumor-specific antigens could arise from non-synonymous somatic point mutations.

1.2.3 Somatic point mutations can yield tumor-specific neoantigens

Carcinogenesis is an evolving multistage process that develops through a series of genomic instability events. Instability in the genome could cause uncontrolled cell growth, invasion, and metastasis[113]. Such genomic changes include gene amplifications[114,115], chromosomal translocations[116], deletions[117], and collectively any somatic mutation[55,118–120].

Most tumor cells produce and process antigenic peptides, and these are expressed either on the cell surface or released into the extracellular matrix, the stroma. Studies have reported that a wide variety of peptide antigens are associated or shared among several cancer types, such as colon and prostate cancer, malignant melanoma, renal, breast, and lung carcinoma[121]. In all cancer entities, tumor antigens mostly fit into two categories, tumor-specific antigens (TSAs) and tumor-associated antigens (TAAs). TSAs are a class of antigens exclusively expressed on tumor cells that include all abnormal protein structures derived from somatic mutations that can be the primary cause for malignant transformation[90]. On the other hand, TAAs are

overexpressed aberrant protein-derived antigens expressed not only on tumor cells but also a subpopulation of normal cells[82].

With the latest advancements in the next-generation sequencing of whole-genome and exome (all protein-coding regions), mutation identification and screening for TSAs have become routine and are highly important[122]. Such TSAs are considered ideal and truly tumor-specific (referred to as neoantigens) targets for adoptive T cell therapy (ATT)[90]. The nature of TSAs derived from mutations varies from cancer to cancer and from individual to individual due to inter and intratumoral heterogeneity, respectively. An adult solid tumor entity bears around 50 mutations during primary diagnosis, whereas pediatric cancers and leukemia acquire fewer mutations (10 on average)[123]. UV-associated melanomas of the skin carry 100-200 mutations. However, genomically unstable colorectal cancers can contain more than 1000 mutations[123]. Above all, most mutations that frequently alter a coding protein are non-synonymous somatic point mutations caused by a single amino acid substitution such as R175H and R248Q in the TP53 gene and G12V and G13D in the RAS gene[51,53,124,125].

Regardless of the total number of somatic mutations in a tumor or its recurrent occurrence of specific mutations (hot-spots), abnormal proteins produced from a driver mutation may or may not create an epitope and induce immunogenicity. Many factors contribute to creating an epitope, such as antigen expression level, proteasomal peptide processing, transport [Figure 1], and its binding affinity to MHC/HLA determine if a mutant peptide might turn out to be a suitable epitope. The probability for a non-synonymous somatic mutation to pass through these stages to generate novel tumor-specific antigens is low. Probability speculation proved to be true in a vaccinia virus study that reported only 1 out of 14 predicted epitopes are indeed immunogenic[126]. In detail, of ~ 100 vaccinia-encoded epitopes predicted to bind HLA-A*02:01 (HLA-A2) with IC_{50} < 100 nM, only 50% of epitopes elicited $CD8^+$ T cell responses, and from these, only 15% proven for endogenously processing. 11% elicited CTL responses out of the 15% of processed epitopes, which explains that 1 out of every 14 predicted epitopes creates an immunogenic epitope[126]. Comparing the vaccinia virus to physiological cancer with somatic mutations, i.e., with similar probability, ~ 4-5 mutant epitopes per cancer could be an estimate that carries 50 mutations. Since every human individual bears six HLA alleles, one can expect 20-25 epitopes per person[90].

Immunodominance also plays a role among dominant versus recessive mutant epitopes. The recessive epitopes may be targetable only if the dominant epitope is lost[127,128]. One factor to the lack of immunodominance of recessive epitopes could be its affinity towards HLA; the dominant epitope could show a strong affinity to HLA alleles. On the other hand, some HLA alleles show prediction to bind more epitopes than others making the estimation difficult[129]. In addition to point mutations, frameshift mutations causing novel reading frames might increase the chance of potential epitopes generated per individual[130].

Supportively, in cancer patients, various forms of TSAs have been distinguished as targets for CD8[+] T cells and were also examined in pre-clinical trials[131–139]. Ample evidence available reporting that the cytotoxic T lymphocytes can effectively recognize tumor antigens resulting from mutations in various forms of cancers such as colorectal, head, and neck cancers (CASP-8), melanoma (ß-catenin), squamous cell lung carcinoma (NFYC), and leukemia[140–144].

With all the above collective understanding, somatic mutations can yield potential neoepitopes for ATT of cancer; if proven to be naturally processed and presented.

1.2.4 Recurrent somatic driver mutations as targets for ATT of cancer

The adoptive transfer of tumor antigen-specific T cells is one of the most promising and advanced medicinal products (ATMPs) in the area of cancer immunotherapy. However, problems still exist with the appropriate selection of antigens for a powerful ATT[69].

Over the past two decades, several preclinical studies represented to be tumor-specific have led to the development of diverse biological agents that block or inhibit various molecular targets derived from genetic abnormalities in cancers[145]. Targeted cancer therapies such as monoclonal antibodies (mAbs), small-molecule drugs (SMDs), and cancer vaccines used tumor-associated and self-antigens as their targets[146].

Due to the on-target attack of cells expressing self-antigens, these therapies exhibited high toxicity[147], frequent drug resistance, and relapse[148], thereby dampening the overall therapeutic benefit[149]. Hence, a careful selection of suitable tumor antigens still seems to pose a crucial challenge for an efficacious ATT. Therefore, targeting recurrent somatic driver mutations (TSAs), using TCR-engaged T cells might become a more precise strategy when compared to SMDs[150] and mAbs[151].

Anders *et al*. demonstrated that monospecific CD8[+] T cells restricted to a viral epitope (SV40-T antigen) could reject large established tumors completely (≥ 500 mm^3). Such complete tumor rejection might be possible probably by the elimination of escape variants[150]. Targeting driver mutations (ancestor in cancer cell evolution) that generate homogenously expressed epitopes could also avoid resulting antigen loss variants[128].

Not every mutation is a driver mutation and can create an epitope. Driver mutations give tumors a growth advantage. Concerning tumor heterogeneity, driver mutations can occur recurrently across (intertumoral) and between (intratumoral) tumors[152]. Recurrent neoepitopes are often derivatives of driver mutations shared (ancestral mutations) by all cancer cells and are ideal target antigens that could not be lost[90,125,153]. However, selection and validation of the most promising neoepitope candidates for ATT is a tedious task[154].

In recent times, *in silico*-based identification of neoantigens from tumor whole-genome or exome sequencing datasets seems to assist neoepitope selection. In 2019, our bioinformatics study published a prediction screen identifying and ranking the binding affinities of putative

neoepitopes to the HLA complex using TCGA datasets. We used affinity prediction algorithms based on machine learning as an opportunity to devise *off-the-shelf* T cell therapies for subgroups of cancer patients who share recurrent neoepitopes[154]. Immunogenicity validation, concerning efficient processing and presentation, is inevitable for all predicted recurrent mutant epitopes[90]. For this, humanized TCR-HLA mouse models could be utilized as a valuable tool not only to verify the immunogenicity of predicted epitopes but to discover novel epitopes by full-length protein-encoding gene immunization.

1.3 Humanized mouse models

The fact that recurrent neoepitopes are ideal targets with their homogeneous expression on all cancer cells contributed to emerging ways to isolate TCRs for ATT of cancers. Identification of neoantigen-specific TCRs is possible from humans or transgenic mice[155–157].

In humans, TCR α and ß chains can be isolated by *in vitro* priming from HLA-matched healthy individuals or immunized candidates with mutant peptides. Anyway, in the former case, the precursor TCR repertoire against neoepitopes in healthy donors is unaffected by central tolerance mechanisms; deletion of T cells that bind only the wild-type self-counterparts with high affinity. So, TCR isolation in principle is possible by autologous priming, although the percentage of neoantigen-specific T cells as shown in melanoma, gastrointestinal, lung, and ovarian cancer patients is low; only about 1.2% of mutations are spontaneously recognized[158]. Further, the expression level of neoantigens in the periphery is a decisive factor in whether the T cells are deleted or became anergic[90,159].

In transgenic mice, isolation of TCRs from a naive repertoire could work after multiple immunizations; to produce memory responses *in vivo*. Also, TCRs come from a tumor-free host, providing the possibility to select high-affinity T cells from a naive precursor pool that could undergo deletion in tumor-bearing human individuals due to chronic antigen stimulation[90].

1.3.1 Transgenic HLA-I mice with human HLA antigen presentation

One example of a tumor-free human HLA host is the HLA-A2/D^b transgenic mouse model, which served as a model to monitor human immune responses for TCR isolation and epitope identification[160–162]. Diverse chimeric HLA-D^b/K^b models were employed to understand heterologous immunity between viruses compared to humans[163–165]. Though chimeric MHC/HLA transgenic models existed for some time, questions remain concerning mouse TCRs education on a chimeric single human HLA allele, overall TCR repertoire, and certain V$_{αß}$ specificities gained or lost during thymic selection[21,166,167]. Furthermore, severe on-target toxicities combined with fatality were reported in cancer patients treated with an affinity matured (in CDR3α region) anti-MAGE-A3 TCR derived from HLA-A2 transgenic mice, which cross-

reacted to other shared epitopes of MAGE- A9 and -A12 antigens[168,169]. Since murine TCRs are not negatively thymus-selected on a patient's HLA alleles and the human proteome, thorough on- and off-toxicity evaluation is necessary on such mouse-derived TCRs that might cross-react with normal human proteins. Besides, partially humanized murine TCRs are indeed immunogenic when transferred to humans[170].

Hence, for scrupulous prediction and translation of murine-derived therapeutics into human systems, a comprehensive humanization of mouse models expressing diverse human HLAs with human TCR gene loci seems inevitable. While the latter, transgenic HLA-A2/K^b mice with a complete human TCR gene loci for efficient isolation of optimal affinity TCRs have been generated and established by Prof. Thomas Blankenstein and his team in the year 2010[157].

1.3.2 Human TCR transgenic mice as an ideal tool for isolating TCRs for ATT of cancer

The necessity for humanization triggered the development of a novel humanized transgenic mice named AB*ab*DII. The AB*ab*DII mice have the complete human TCR gene loci, knocked out for murine TCR and MHC I gene loci, and it is transgenic for human class I allele HLA-A2[157]. Thus, making the model a valuable tool for TCR isolation against TAAs[171], and TSAs including viral epitopes[172]. Even though its advanced humanization yielded functional CD8$^+$ T cell populations for optimal affinity TCR isolation, the AB*ab*DII model is limited as a source of obtaining only HLA-A2-restricted TCRs. On this note, the model also skews its thymic selection process towards a single human HLA-A2 allele expressed at a low level[162]. Therefore, it is indeed a bottleneck that AB*ab*DII mice are not available with other HLA alleles. So, this makes humanized TCR mice expressing other class I alleles a requirement for epitope discovery, which would benefit TCR isolation for a large number of non-HLA-A2 individuals.

1.4 Aims and objectives

One of the main goals was to broaden the allelic diversity to study the human TCR repertoire enrichment against a complete human HLA haplotype in mice. The PiggyBac transposon strategy introduced six HLA alleles into transcriptionally active sites as a single haplotype as in humans (every individual bears six HLA class I alleles). Another aim was to ensure surface expression of the newly introduced HLA alleles to analyze their presentation capacity. Ultimately, the novel mouse model could become an *in vivo* epitope discovery system for identifying a range of tumor antigens restricted towards highly frequent (allele-wise) HLA genes, thereby isolating T cell receptors for ATT of cancer.

1.4.1 The AB*ab*.I transgenic mouse model - a novel tool for isolating TCRs restricted towards a broad spectrum of HLAs

Currently, the AB*ab*DII transgenic mouse model serves as a valuable system to isolate human TCRs, as it comprises the whole human T cell receptor gene loci[157]. However, the model

exhibits a limitation of isolation of only HLA-A2-restricted TCRs. To target diverse non-HLA-A2-binding antigens, and in parallel, broaden the HLA repertoire of the existing system for epitope discovery, it is necessary to include multiple human HLA class I genes into the current AB*ab*DII mice.

In this regard, the aim was to create the 'AB*ab*.I transgenic mice' by introducing a novel set of chimeric class I fusion alleles (HLA-A*03:01, A*11:01, B*07:02, B*15:01, C*04:01, C*07:02) as a single genotype into the existing AB*ab*DII mice. Furthermore, the purpose of AB*ab*.I mice was to discover previously not described immunogenic epitopes selected for presentation on six HLA alleles post-processing.

In this doctoral study, the ultimate goal was to develop a new transgenic AB*ab*.I mouse line as an *in vivo* model system for epitope discovery, besides also being a valuable tool to identify novel TCRs against high HLA-affinity tumor-specific antigens ($IC_{50} < 50$ nM) derived from recurrent somatic point mutations.

2. Materials and Methods

2.1 Materials

2.1.1 *In silico* software algorithms

Table 1: List of web-based resources used for T cell epitope prediction-screen

Web server	Link to the open-source repository
NetMHC[173–175] Prediction of peptide-MHC class I binding using artificial neural networks (ANNs)	http://www.cbs.dtu.dk/services/NetMHC/
IEDB[176,177] Prediction of peptide-MHC class I binding using stabilized matrix methods (SMMs) and ANNs	https://www.iedb.org/home_v3.php

2.1.2 Databases and datasets

Table 2: List of datasets retrieved for HLA frequencies, ranking, and cancer incidence

Web server	Link to the open-source repository
Allele frequency net database (AFND)[178,179] Global HLA allele frequency data	http://allelefrequencies.net
National marrow donor program (NMDP)[180] bioinformatics HLA ranking datasets	https://bioinformatics.bethematchclinical.org/hla-resources/
Immunogenetics (IMGT)/HLA sequence database[181] Human major histocompatibility complex/human leukocyte antigen (HLA) sequences	https://www.ebi.ac.uk/ipd/imgt/hla/allele.html
Global cancer observatory (GLOBOCAN) dataset[182] Global cancer incidence data	http://globocan.iarc.fr

2.1.3 Cell culture media, flasks, tubes, and plates

Table 3: List of cell culture media

Media	Manufacturer
DMEM, GlutaMAX™ supplemented *	Thermo Fisher Scientific, Rockford (USA)
RPMI 1640, GlutaMAX™ supplemented *#	Thermo Fisher Scientific, Rockford (USA)

*standard media preparation: DMEM/RPMI + 10% (i.e. 50 ml per 500 ml) of heat-inactivated, sterile-filtered fetal bovine serum + 100 IU/ml penicillin and 100 µg/ml streptomycin.

#mouse T cell culture medium (mTCM): *standard RPMI media + 1 mM sodium pyruvate + 0.1 mM non-essential amino acids solution + 2 mM HEPES buffer + 50 µM ß-mercaptoethanol.

Table 4: List of consumables

Flasks, tubes and, plates	Manufacturer
Cell culture flasks: 25, 75, and 150 cm^2	TPP Products AG, Trasadingen (Switzerland)
Cell strainers: 40 and 70 µm	BD Biosciences GmbH, Heidelberg (Germany)
Syringes: 1, 5, and 10 ml	BD Biosciences GmbH, Heidelberg (Germany)
Plates: 96-well U- and flat- bottom	TPP Products AG, Trasadingen (Switzerland)
24-well plates: tissue and non-tissue treated	BD Biosciences GmbH, Heidelberg (Germany)
12-well and 48-well plates	BD Biosciences GmbH, Heidelberg (Germany)
Centrifugation tubes: 15 and 50 ml	TPP Products AG, Trasadingen (Switzerland)
FACS tubes, capped with 40 µm cell strainer	BD Biosciences GmbH, Heidelberg (Germany)
Cryogenic storage tubes and cell culture dishes	Greiner AG, Frickenhausen (Germany)
PCR reaction tubes (0.5, 1, and 2 ml)	Eppendorf AG, Hamburg (Germany)
Pipettes: 2, 5, 10, 25, and 50 ml	Starlab GmbH, Hamburg (Germany)
Pipette tips, TipOne®: 10, 200, and 1000 µl	Starlab GmbH, Hamburg (Germany)
EDTA - Mini blood collection tubes	Merck KGaA, Darmstadt (Germany)
5PRIME Phase Lock Gel tubes	Quantabio Inc., Beverly (USA)
Pasteur Pipettes	DWK Life Sciences GmbH, Mainz (Germany)

2.1.4 Chemical reagents

Table 5: List of chemical reagents

Reagents	Manufacturer
Dimethyl sulfoxide (DMSO)	Merck KgaA, Darmstadt (Germany)
Fetal bovine serum (FBS)	PAN-Biotech GmbH, Aidenbach (Germany)
Lipofectamine® 2000 transfection reagent	Thermo Fisher Scientific, Rockford (USA)
2-mercaptoethanol	Thermo Fisher Scientific, Rockford (USA)
Dulbecco's phosphate-buffered saline (DPBS)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Ammonium chloride (NH$_4$Cl)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Potassium bicarbonate (KHCO$_3$)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Ammonium acetate (NH$_4$CH$_3$CO$_2$)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Natrium chloride (NaCl)	neoLab GmbH, Heidelberg (Germany)
Protamine sulfate salt from herring	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Phorbol 12-Myristate 13-Acetate (PMA)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)

4-(2-hydroxy-ethyl)-1-piperazine-ethane-sulfonic acid (HEPES)	Thermo Fisher Scientific, Rockford (USA)
Ethylenediaminetetraacetic acid (EDTA)	neoLab GmbH, Heidelberg (Germany)
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Tris hydrochloride (Tris-HCL)	Carl Roth GmbH, Karlsruhe (Germany)
Proteinase K	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Ribonuclease A (RNase A)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Trypan blue staining 0.4% Solution	Thermo Fisher Scientific, Rockford (USA)
Phenol-chloroform-isoamyl alcohol mixture	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
Glycogen	Thermo Fisher Scientific, Rockford (USA)
Freund's Adjuvant, Incomplete (IFA)	Sigma-Aldrich Chemie GmbH, St. Louis (USA)
CpG1826 oligonucleotide (tccatgacgttcctgacgtt)	TIB Molbio GmbH, Berlin (Germany)
Ethanol (99.9%)	AppliChem GmbH, Darmstadt (Germany)
7AAD viability staining solution	BioLegend, Fell (Germany)
Protein transport inhibitor (with Brefeldin A)	BD Biosciences GmbH, Heidelberg (Germany)
CountBright™ absolute counting beads	Thermo Fisher Scientific, Rockford (USA)
Dynabeads™ mouse T-activator CD3/CD28	Thermo Fisher Scientific, Rockford (USA)

2.1.5 Buffers

Table 6: List of buffers

Buffers	Composition
Fluorescence activated cell sorting (FACS) buffer	1X PBS (pH 7.2) 1% FBS (v/v)
Magnetic-activated cell sorting (MACS) buffer	1X PBS (pH 7.2) 1% FBS (v/v) 2 mM EDTA (w/v)
Ammonium-Chloride-Potassium (ACK) lysis buffer for RBC lysis	H_2O (pH 7.2 to 7.4) NH_4Cl (0.15 M) $KHCO_3$ (10 mM) EDTA (0.1 mM)
Lysis buffer for organic phenol-chloroform DNA extraction	NaCl (0.1 M) Tris-HCl (10 mM, pH 8.0) EDTA (25 mM, pH 8.0) SDS (0.5%) Proteinase K (0.1 mg/ml)

2.1.6 Cell lines

Table 7: List of cell lines

Cell lines	Characteristics
A375 and SK-Mel-37 melanoma cell lines	Homo sapiens, skin-derived epithelial origin HLA-A*02:01$^+$ NY-ESO$^+$ TRRAP$^+$ mutant (A375) and wild-type (SK-Mel-37)
Lymphoblastoid cell line (LCL), BM14	Homo sapiens, EBV-transformed, B cell origin HLA-A*03:01$^+$, B*07:02$^+$, and C*07:02$^+$
MCA205 methylcholanthrene-induced fibrosarcoma cell line	Mus musculus (C57BL/6), mesenchymal origin H-2 D^{b+} and K^{b+}

2.1.7 Cytokines

Table 8: List of cytokines

Cytokines	Manufacturer
Recombinant murine IFN-γ	PeproTech, Hamburg (Germany)
Recombinant human/mouse IL-2	Novartis Pharma AG, Basel (Switzerland)
Recombinant murine IL-15	PeproTech, Hamburg (Germany)

2.1.8 Antibodies and Primers

Table 9: List of antibodies

Antibodies	Clone / Manufacturer
Anti-mouse CD3ε PE	145-2C11 / BioLegend
Anti-mouse CD8a APC	53-6.7 / BioLegend
Anti-mouse CD4 BV421™	GK1.5 / BioLegend
Anti-mouse IFN-γ BV421™	XMG1.2 / BioLegend
Anti-human pan HLA-ABC PE	W6/32 / BioLegend
Anti-human ß2 microglobulin PE	TÜ99 / Beckman Coulter
Anti-human HLA-A3 PE	GAP-A3 / eBioscience™
Anti-human HLA-B7 FITC	BB7.1 / Abcam
Anti-human pan HLA-C unconjugated with anti-Mouse IgG3γ PE	H-5 (unconjugated) with SB76b (secondary antibody) / Santa Cruz Biotechnology
Anti-mouse CD16/32 Fc receptor block	93 / BioLegend
Anti-mouse CD28 unconjugated for cell culture	CD28.2 / BioLegend
Anti-mouse CD3 unconjugated for cell culture	OKT-3 / BioLegend

Table 10: List of primers

Primers	Sequence
Primers used for H-2 D^b promoter activity assessment; HLAs switched for eGFP	
H-2D^{b}For-Primer	GTCT<u>GGCGCGCC</u>TTCTTCTACATA (<u>AscI-site</u>)
H-2D^{b}Rev-Primer	CTCGCCCTTGCTCACCATCTGGGATCCCGGGTGT
eGFPFor-Primer	ATGGTGAGCAAGGGCGAG
eGFPRev-Primer	GTCTTTAATTAATTACTTGTACAGCTCGTCCATGC
Primers used for CMV promoter activity assessment; H-2D^b switched for CMV promoter	
CMV-Forward Primer_AscI	GCT GGCGCGCC GTACGGGCCAGATATACGC
CMV-Reverse Primer_XmaI	GTC CCCGGG GTGGGTTCTCTAGTTAGCCAGA
Primers used for full-length H-2 D^b polyadenylation signal manipulation	
For-H-2-Full-PA	ATCTGCATCCTGTAAGCTCCATGCTAC
Rev-H-2-Full-PA	GTCTTAATTAACCATCACAGATCAAAATGGACTGAGC
Primers used for full-length H-2 D^b polyadenylation signal manipulation	
For-BGH-Full-PA	GACTGCTCTTCCAAACTGTGCCTTCTAGTTGCCAGCC
Rev-BGH-Full-PA	GTCTTAATTAACCATAGAGCCCACCGCATC
Multiple cloning site (MCS) introduced into the pMA-V vector to clone six HLA alleles one after another (5'-to-3')	
CTATCGTAGGCGCGCCTCCGATGCGGCCGCAGTTATCGCGCGCGATTGTACGCCGGCGATTGTACCTGCAGGCTTAAGATTACCGGTGCCTTGAAGCTTGTTGCAGGTACCTCGGCGGAGCTCGAAGGCGTACTAGTGCTGATACGCGTCGTAGCTCGTTAATTAACGG MCS sequence contains restriction sites (5'-to-3'): AscI, NotI, MauBI, MreI, SbfI, AflIII, AgeI, HindIII, KpnI, SacI, SpeI, MluI, PacI.	
Primers to amplify MCS	
MCS-FP	*CTATCGTAGGCGCGCCTCC*
MCS-RP	*CCGTTAATTAACGAGCTACGACGC*
Primers used for cloning six HLA alleles one after another into pMA-V (5'-to-3')	
FP-HLA-A3-AscI	CCTAGGCGCGCCTTCTTCTAC
RP-HLA-A3-NotI	AGTACGCGGCCGCCCATAGAGCCCACCGCATC
FP-HLA-A11-NotI	CGTAGCGGCCGCGCGCCTTCTTCTACATAAAACACAC
RP-HLA-A11-SbfI	CAGTTACCTGCAGGCCATAGAGCCCACCGCATC
FP-HLA-B7-SbfI	CAGCCCTGCAGGGCCTTCTTCTACATAAAACACACCC
RP-HLA-B7-AgeI	TGCAACCGGTCCATAGAGCCCACCGCATC
FP-HLA-B15-HindIII	CTAGTGAAGCTTCGCCTTCTTCTACATAAAACACACC
RP-HLA-B15-KpnI	TCGACAGGTACCCATAGAGCCCACCGCATC
FP-HLA-C4-SacI	CTAGAGCTCCGCCTTCTTCTACATAAAACACACC
RP-HLA-C4-SpeI	GATACTAGTCCATAGAGCCCACCGCATC
FP-HLA-C7-SpeI	TCAGACTAGTCGCCTTCTTCTACATAAAACACACC
RP-HLA-C7-PacI	GACATTAATTAACCATAGAGCCCACCGCATC

2.1.9 Mice

Table 11: List of transgenic and wild-type mice with relevant TCR loci and HLA genotype

Mice	Genotype		
AB*ab*DII mice[157] (in-house generated at the Max Delbrück Center mouse core facility)	HumanTCR$^{+/+}$ loci MouseTCR$^{-/-}$ loci Mouse MHC I Human HLA I	: : : :	TCR A/B$^{+/+}$ TCR a/b$^{-/-}$ ß2m$^{-/-}$, H-2 D$^{b-/-}$ HLA-A*02:01$^{+/+}$
AB*ab*.I (in-house generated at the Max Delbrück Center mouse core facility)	HumanTCR$^{+/+}$ loci MouseTCR$^{-/-}$ loci Mouse MHC I Human HLA I , and HLA-A*02:01$^{+/+}$	: : : :	TCR A/B$^{+/+}$ TCR a/b$^{-/-}$ ß2m$^{-/-}$, H-2 D$^{b-/-}$ HLA-A*03:01/A*11:01$^{+/+}$, HLA-B*07:02/B*15:01$^{+/+}$, HLA-C*04:01/C*07:02$^{+/+}$
C57BL6/N (wild-type, bred in-house)	Mouse MHC I	:	H-2 D$^{b+/+}$, H-2 D$^{b+/+}$
HHD[183]	MouseTCR$^{+/+}$ loci Mouse MHC I Human HLA I	: : :	TCR a/b$^{+/+}$ ß2m$^{-/-}$, H-2 D$^{b-/-}$ HLA-A*02:01$^{+/+}$

2.1.10 Laboratory equipment

Table 12: List of equipment

Antibodies	Clone / Manufacturer
Light microscope, primovert	Carl Zeiss Microscopy GmbH, Jena (Germany)
Axiovert 40 C microscope	Carl Zeiss, Göttingen (Germany)
Beckman GS-6 series centrifuge	Beckman Coulter Inc., California (USA)
Heraeus biofuge pico centrifuge	Heraeus Instruments, Osterode (Germany)
Centrifuge, fresco 21	Thermo Fisher Scientific, Rockford (USA)
Centrifuge, rico 21	Thermo Fisher Scientific, Rockford (USA)
Centrifuge, megafuge 8R	Thermo Fisher Scientific, Rockford (USA)
Centrifuge, megafuge 40R	Thermo Fisher Scientific, Rockford (USA)
Thermomixer comfort	Eppendorf AG, Hamburg (Germany)
Vortex mixer, Genie-2	Scientific Industries Inc., New York (USA)
Vortex mixer, Sunlab® SU1900	neoLab GmbH, Heidelberg (Germany)
Tabletop centrifuge 5415R/5415D	Eppendorf AG, Hamburg (Germany)
Finnpipette™ (2-1000 µl)	Thermo Fisher Scientific, Rockford (USA)
IBS pipetboy acu	Integra Biosciences AG, Zizers (Switzerland)
Multi pipette® M4	Eppendorf AG, Hamburg (Germany)
Nano photometer, nanovolume N50	Implen GmbH, Munich (Germany)
Cryogenic storage system, 24k	TecLab GmbH, Taunusstein (Germany)
Ice maker, FM-120KE-HC	Hoshizaki B.V., Swanley (UK)

Neubauer chamber (improved)	Paul Marienfeld GmbH, Königshofen (Germany)
Flow cytometer, FACSCelesta™	BD Biosciences GmbH, Heidelberg (Germany)
Flow cytometer, FACSCanto™ II	BD Biosciences GmbH, Heidelberg (Germany)
FACS cell sorter, FACSAria™ I	BD Biosciences GmbH, Heidelberg (Germany)
Viessmann cold chamber	Kältetechnik AG, Saale (Germany)
Mettler AE 163 analytical balance	Mettler-Waagen GmbH, Gießen (Germany)
Milli-Q® advantage A10	Merck KGaA, Darmstadt (Germany)
CO_2 incubator, ICOmed	Memmert, Schwabach (Germany)
Water bath, WNB / WNE / WPE	Memmert, Schwabach (Germany)
BVC fluid aspiration system	BrandTech Scientific, Inc., Essex (UK)
Safety workbench safe 2020	Thermo Fisher Scientific, Rockford (USA)
Freezer / refrigerator, mediline	Liebherr GmbH, Bierbach an der Riß (Germany)

2.1.11 Software

Table 13: List of commercial software

Software	Manufacturer
GraphPad Prism software (version 8.4.3)	GraphPad Software, Inc., La Jolla (USA)
FlowJo™ version 10 (version 10.7.1)	TreeStar Inc., Olten (Switzerland)
BioRender (web-based)	BioRender Inc., Toronto (Canada)
ImmunoSEQ™ analyzer (web-based)	Adaptive Biotechnologies, Seattle (USA)
Mendeley reference manager (version 1.19.8)	Elsevier group, London (UK)
SnapGene gene analyzer (version 5.1.5)	GSL Biotech LLC., Illinois (USA)

2.2 Methods

2.2.1 *In silico* design and construction of chimeric HLA class I alleles

AB*ab*.I polycistronic transgene encoding six HLA alleles segregated by 2A peptide linkers was designed *in silico* and synthesized by GeneArt™, Thermo Fisher Scientific. HLA alleles in the transgene are chimeric human-mouse fusion proteins. The α1 and α2 domains are human cDNAs from respective HLA alleles, α3 to cytoplasmic domains are murine cDNAs from the H-2 D^b gene. Human ß2m is linked to the N-terminus by a 15-amino acid glycine-serine peptide linker. Modified HLA constructs with H-2 D^b or CMV promoter, 3'-bGH polyadenylation signal variants, and HLA monochains with or without ß2m were custom-cloned in-house. Phusion® DNA polymerase combined HLA amplicons by splicing by overhang extension (SOE) PCR. Six HLA alleles were cloned as a single haplotype sequentially using unique restriction enzyme

sites into the pre-integrated multiple cloning site. PiggyBac ITRs-flanked AB*ab*.I transgene was generated using 5'-AscI and 3'-PacI sites. Stbl3™ *E. coli* (Thermo Fisher Scientific, Rockford, USA) propagated long HLA plasmids containing repetitive sequences.

2.2.2 *In vitro* characterization of HLA-AB*ab*.I transgene by flow cytometric analysis

Surface HLA class I expression was analyzed 48 hours after every transient transfection by flow cytometry. MCA205 cells were treated with or without recombinant IFN-γ (100 ng/ml) in transfection experiments. 10 µl of Lipofectamine™ 2000 delivered the primary polycistronic AB*ab*.I transgene and all modified HLA constructs. In every case, two days post-transfection, cells were stained using fluorochrome-conjugated antibodies for human pan HLA-ABC and ß2 microglobulin. MCA205/AB*ab*.I stable cell line was made by transfecting MCA205 cells using AscI-PacI digested, linearized DNA construct, followed by multiple enrichments of the HLA$^+$ cell population through flow cytometry-based cell sorting using pan HLA-ABC antibody.

2.2.3 Quantification of transgene mRNA using quantitative RT-PCR analysis

MCA205 cells were transiently transfected as described earlier with polyadenylation signal-modified constructs driven by H-2 D^b or CMV promoter. Two days later, total RNA in cell lysates was reverse transcribed using anchored oligo (dT) and random hexamers. cDNA amplification was quantified using SYBR$^®$ green dye with HLA-specific primers in the quantitative RT-PCR (qRT-PCR) reaction. As controls, 18s rRNA (housekeeping) and constitutive H-2 D^b reactions were analyzed. From Ct values, difference of expression, ΔCt = Ct (HLA/H-2 D^b) – Ct (18s rRNA) was estimated. The percentage of change in expression was measured by normalizing it to a reference gene, mouse H-2 D^b (set to 100%).

2.2.4 Construction of genome targeting vectors

Six chimeric HLA alleles, HLA-A*03:01, A*11:01, B*07:02, B*15:01, C*04:01, and C*07:02, as a natural HLA haplotype like in humans, were cloned into a minimal mammalian expression vector, the pMA plasmid. For mouse genome targeting, 5'-AscI and 3'-PacI sites were digested and blunted sequentially to insert hyperactive PiggyBac-restricted inverted terminal repeats (ITRs) to produce an oocyte targeting vector.

2.2.5 Generation of the AB*ab*.I transgenic mice using pronuclei injection technology

In vitro transcribed (*ivt*) PiggyBac (PB) transposase mRNA targeted the ITR-flanked AB*ab*.I transgene cassette into the recipient oocytes. 10-12 weeks old AB*ab*DII male mice were mating partners. 6-10 weeks old AB*ab*DII female mice were superovulated before mating to collect zygotes. NMRI mice served as pseudopregnant foster mothers. Donor HLA plasmids (AB*ab*.I transgene) were co-prepared with the PB transposase (*ivt* mRNA) at a molar ratio of 2:1. Mixed DNA with *ivt* mRNA samples were pronuclei-microinjected into fertilized AB*ab*DII oocytes as

described[184,185]. All mouse experiments were performed as per the standard guidelines approved by the Landesamt für Arbeitsschutz, Gesundheitsschutz und technische Sicherheit, Berlin.

2.2.6 Genotyping of HLA I haplotype in the AB*ab*.I Tg mice

Genomic DNA was extracted from ear biopsies of transgenic mice at 95 °C in the presence of 0.05 M NaOH. HLA-specific primers directed against all six chimeric fusion alleles amplified fragments using a routine PCR with Taq DNA polymerase. 18s rRNA and human TCR αß were control reactions. PCR amplicons were run on a 1.5% gel at 100 V for 45 minutes using an agarose gel electrophoresis system.

2.2.7 Estimation of CD3⁺ CD8⁺/CD4⁺ absolute numbers in blood and lymphoid organs

Preparation of blood: 50 µl blood per mouse (from each strain) was blocked for *Fc* receptors III/II by anti-CD16/32 antibody. Red blood cells were lysed by erythrocytes lysing buffer (ACK buffer) before staining and flow cytometry. *Preparation of lymphoid organs:* spleen and lymph nodes (axillary, brachial, mediastinal, inguinal, and mesenteric) from 8-12 weeks old mice were mashed and strained through a 70 µm cell strainer. ACK-lysed spleens and lymph nodes were pooled together by passing them through a 40 µm strainer. *Staining of blood and lymphoid organs for T cell counts:* samples were stained using fluorochrome-conjugated CD3, CD8, and CD4 antibodies. 11 µl volume (11,880 beads) of CountBright™ counting beads per sample were used for T cell quantification by flow cytometry.

Absolute T cell numbers/µl of blood or per pooled mouse lymphoid organs =

$$\frac{\text{no. of CD8/CD4 cell events}}{\text{no. of beads events}} \times \frac{\text{bead count of the lot (54,000 beads/50µl)}}{\text{volume of samples (µl)}}$$

2.2.8 Enrichment of T cell population (CD3⁺ CD8⁺) using untouched FACS sorting

Lymphoid cells and the whole blood from young AB*ab*DII or AB*ab*.I mice (8-12 weeks) were prepared and pooled using the protocol described earlier [refer 2.2.7, *preparation of blood and lymphoid organs*]. CD3⁺ cells were enriched from pooled cell suspensions through bead-based magnetic separation without columns using EasySep™ mouse T cell untouched isolation kit (Stemcell Technologies GmbH, Cologne, Germany). In brief, Fc receptors in cell suspensions (2×10^8 cells/ml) were blocked by incubating with 50 µl/ml rat serum. Pooled cells from the spleen, lymph nodes, and blood were stained using a biotinylated antibody cocktail (50 µl/ml) directed towards non-T cells (CD3⁻). Streptavidin-coated magnetic spheres (75 µl/ml) captured these CD3⁻ cells onto the EasySep™ magnet, whereas the CD3⁺ cells passed the magnetic field into the flow-through solution for collection. Untouched CD3⁺ cell populations were surface stained with fluorescent antibodies specific for CD3, CD8, and CD4 markers. Flow cytometry-

based sorting enriched CD3$^+$ CD8$^+$ T cells (≥ 97% purity) for organic DNA (genomic) extraction using the phenol-chloroform method.

2.2.9 TCRß repertoire deep sequencing of AB*ab*.I mice using ImmunoSEQ® technology

Deep sequencing of TCRß immune repertoire was performed at Adaptive Biotechnologies (Seattle, USA). ImmunoSEQ® platform precisely quantifies all V-J gene combinations in cell populations using a multiplex PCR-based assay with bias correction in the first step to minimize amplification bias[186,187]. The assay also provides quantitative abundance data with a sensitivity of 1 in 200,000 T cells. 2-4 µg of genomic DNA/sample obtained from 3×10^5 mouse (average in mice) and 1.5 µg of genomic DNA/sample obtained from ~ 1.8×10^5 human[21] CD8$^+$ T cells were deep sequenced. Three healthy human donors (age: 30, 48, and 60 years old) donated blood with informed consent. Blood collection and processing were performed[21] according to human experimental guidelines under license EA4/046/10 (Ethics committee). Computational analysis was performed on the sequencing data by Adaptive Biotechnologies. ImmunoSEQ™ analyzer portal provided access. TCR sequences were designated based on Immunogenetics (IMGT) nomenclature[181]. Statistical analysis was performed and analyzed as described for CD4$^+$ T cells using the R program[21].

2.2.10 Retroviral transduction of murine splenocytes

Ecotropic retrovirus (only infects mouse/rat cells) packaging cell line, Platinum-E[188] (Plat-E) was transfected at 80% cell confluency using Lipofectamine with pMP71 retroviral plasmids (3 µg/construct) encoding T cell receptor genes (configuration: TCRß-mCß-P2A-TCRα-mCα). 3 ml of virus supernatants were harvested twice at different time points using a 0.45 µm membrane filter for infecting murine AB*ab*DII splenocytes. Spleen and lymph nodes were prepared and pooled using the protocol described earlier [refer 2.2.7, *preparation of lymphoid organs*]. Splenocytes were pre-activated with anti-CD3 (1 µg/ml) and anti-CD28 (0.1 µg/ml) antibodies in the presence of 40 IU/ml recombinant mouse IL-2 (cell concentrations adjusted to 2×10^6 cells/ml). Transductions were performed twice at 48- and 72-hours post-transfection. On first transduction, supernatants containing virus particles were first centrifuged (3000g, 90 minutes at 4 °C) onto retronectin-coated non-tissue culture 24-well plates. 2×10^6 pre-activated splenocytes in mTCM medium supplemented with 8 µg/ml protamine sulfate (ps), 10 µl of CD3/CD28 Dynabeads™, 40 IU/ml IL-2 were added onto virus-captured plates, and centrifuged for 30 minutes, 800g at 32 °C. On second transduction, 1 ml of supernatant was removed from the top layer of the 24-well plates, replaced with a fresh batch of virus (supplemented with 8 µg/ml ps, 40 IU/ml IL-2) harvested at 72 hours' time point, and centrifuged for 90 minutes, 800g at 32 °C. Post-second transduction, transduced splenocytes were adjusted to have cell concentrations of 2×10^6 cells/ml and supplemented with 50 ng/ml murine IL-15.

2.2.11 Functional characterization of HLA alleles expressed in AB*ab*.I Tg mice *in vitro*

IFN-γ production was monitored in the supernatants of co-culture assays by enzyme-linked immunosorbent assay (ELISA). Overnight (16 hours) co-cultures were performed with $5×10^4$ TCR transduced effector T cells and $5×10^4$ target cells (MCA205/AB*ab*.I tumor cells or peripheral blood cells/lymphoid organs from AB*ab*.I mice). Target cells were pulsed with peptides (1 µM/peptide) or with phorbol myristate acetate (PMA) and ionomycin (1 µM each) as a positive control.

2.2.12 *In vivo* characterization of HLA-AB*ab*.I transgene through peptide immunization

A 200 µl injection suspension containing 100 µg of short peptides in PBS, 100 µl of incomplete Freund's adjuvant, and 50 µg CpG oligonucleotides were injected subcutaneously on both lateral sides of hind legs. 10- to 20-week-old young AB*ab*DII or AB*ab*.I mice (6-7 mice/group) were immunized with a 4-week interval between prime-boosts. Similar peptide mixtures were prepared and injected for further boosting regimens of all peptides. As standard controls, AB*ab*DII[157] and AB*ab*.I mice were immunized with 100 µg of MAGE-A1[278] KVLEYVIKV[171] peptide with a similar injection mixture. 7-10 days after every boost injection, peripheral blood cells from immunized mice were ACK-lysed and cultured with 1 µM respective specific or unspecific peptide in the presence of protein transport inhibitor for 5-6 hours. Post-cell fixation, cells were intracellularly stained for IFN-γ, surface stained for CD3, CD8 markers, and analyzed by flow cytometry.

2.2.13 Statistical analysis

GraphPad Prism generated standard curves, plotted graphs, and calculated the level of statistical significance using mean and standard deviation (P values). FlowJo™ analyzed the raw flow cytometry data (.fcs files) with statistics on mean fluorescence intensity values.

3. Results

3.1 Cancer mutanome screen yields potential neoantigens for adoptive T cell therapy

Recurrent hot-spot mutations that could generate putative tumor-specific antigens (TSAs) were screened *in silico* using open-source servers [Figure 2, Table 1, and 2]. 266 'non-synonymous somatic point mutations' previously described in 40 distinct genes were selected as TSAs using literature research[189–215] [Figure 2A]. The chosen point mutations (occurring in ≥2 patients) were detected mostly by whole-exome sequencing analysis in 21 different human cancer entities; the mutation frequencies range from 0.1% to 39%. The NetMHC program[173–175] predicted the affinity of mutant and wild-type epitopes to HLA alleles under applied selection criteria [Figure 2B]. Out of these 266 mutants, 175 epitopes showed strong binding affinities *in silico* (IC_{50}: <50 nM) for at least one of the 18 highly populated HLA class I alleles [Figure 2C]. The AB*ab*DII mouse model expressing HLA-A2 provided a platform to analyze the immunogenicity of the 23 HLA-A2-restricted binders [Figure 2D, left]. From the 152 non-HLA-A2 binders, 68 putative epitopes bound 'HLA-A*0301-A*11:01, B*07:02-B*15:01, C*04:01-C*07:02', a novel class I haplotype similar to humans. The restriction status of these six HLA alleles against multiple (>5) mutant epitopes provided the rationale to create the AB*ab*.I mice [Figure 2D, right].
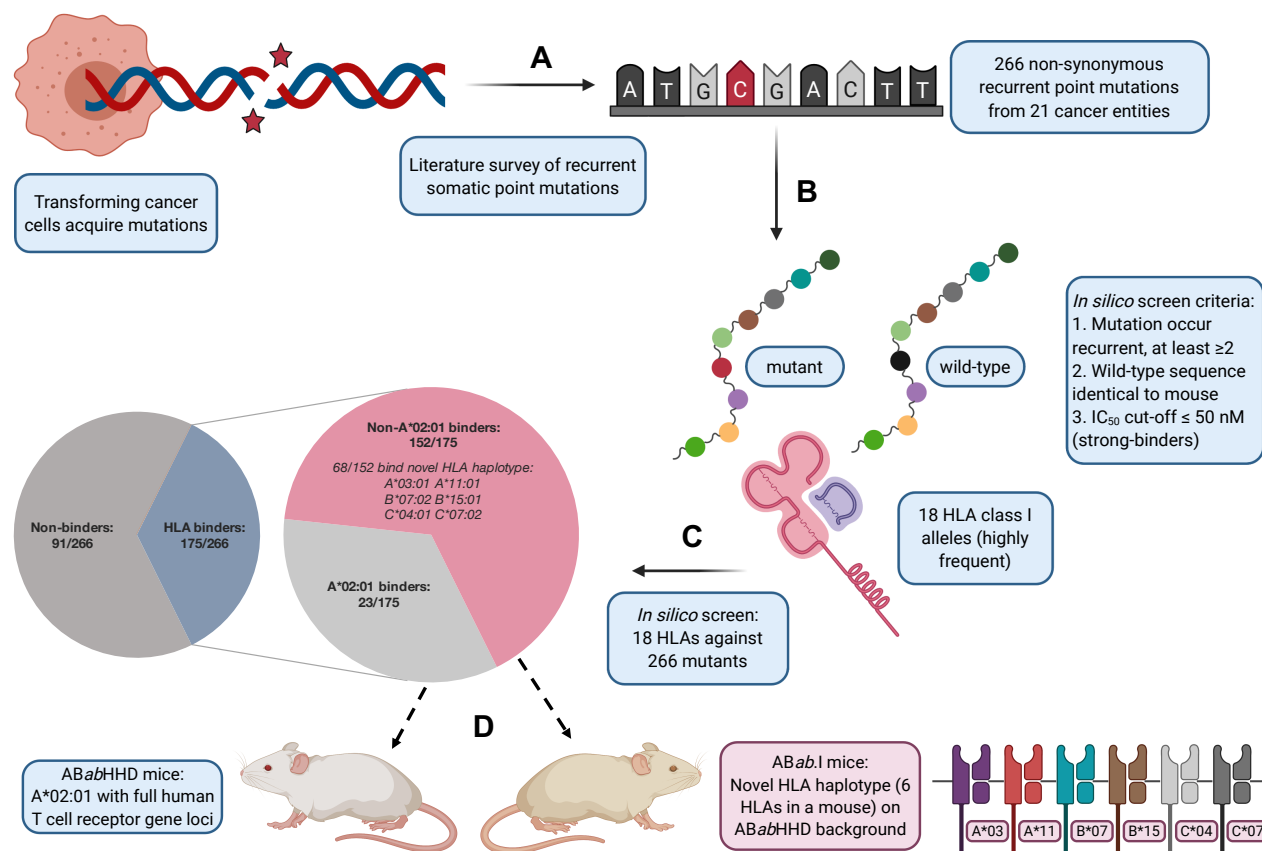
Figure 2: Tumor-specific neoantigens bind a distinct human HLA class I haplotype.

A. Schematic representation depicts recurrent neoantigens derived from cancer-driving point mutations. B. Wild-type and mutant configurations screened against 18 class I alleles using the affinity prediction software, NetMHC Server-DTU. C. HLA binding strengths of 266 mutants revealed non-HLA-A2 binders. D. In silico screen provided the rationale for the generation of ABab.I mice with six novel HLA class I human alleles.

Human T cell receptors (TCRs) isolated from AB*ab*DII and AB*ab*.I mice can be clinically applied for adoptive T cell therapy (ATT) of cancers against endogenously processed mutant antigens.

3.2 AB*ab*DII mouse model elicits HLA-A2-restricted neoantigen-specific T cell responses

To address neoantigen-specific CD8$^+$ T cell responses restricted to HLA-A2-binders, we chose the mutation S722F frequently occurring in the N-terminus of the transformation/transcription domain-associated protein (TRRAP) gene. Serine to phenylalanine mutation at position 722 produced an HLA-A2 epitope with an IC$_{50}$ of 55 nM in our prediction-screen [Figure 3A and 3B]. TRRAP protein is a transcription co-factor involved in cell transformation through MYC activation. 4% of melanoma cases with a total of 86,000 incidences/year express S722F driver mutation. Considering this 4% with S722F express HLA-A2 with an allele frequency of 26.5% in Europe, ~ 911 individuals were estimated (theoretical) for a possible ATT per year after stratification from prior treatment lines [Figure 3B]. Likewise, a worldwide estimate of individuals/year with the TRRAP-S722F mutation using datasets from individual continents such as the U.S.A, South America, and Asia was calculated [data not shown]. Mice have a TRRAP ortholog to ensure thymic deletion of wild-type epitope-restricted T cells [Figure 3C].

A

B

TRRAP nonamer (wild-type and mutant) epitopes	NetMHC 3.2 prediction affinity score restricted to HLA-A*02:01	Mutation frequency (Melanoma)	Melanoma incidence/year (Europe)	Frequency of HLA-A*02:01 allele (Europe)	Total no. of individuals/year with the mutation (possible ATT treatment) (Europe)
KLVFGSV S→(F) L Wild-type (Mutant)	55 nM→29 nM Wild-type (Mutant)	4% (0.04)	86,000	26.5% (0.265)	~ 911 patients

C

Figure 3: HLA-A2-restricted TRRAP-S722F neoantigen as a target for ATT.

For TCR candidate discovery, young AB*ab*DII mice (n=7, 12-16 weeks) were immunized with mutant 9-mer peptides [Figure 3B] at 4-week intervals. CD8[+] T cell responses were analyzed in blood one week after every boost using intracellular staining, followed by culturing the T cells and IFN-γ capture assay to sort for peptide-specific T cells [Figure 4A]. One week post-4th injection (3rd boost) regimen, ~ 0.8% of T cells responded in blood [Figure 4B], which expanded up to ~ 17% in the spleen [Figure 4C] after T cell culturing under low peptide (10^{-9} M) and IL-2 (20 IU/ml) concentration for 10 days. Post-T cell culture, 12,000 IFN-γ[+] CD8[+] cells specific for S722F were captured after 2 hours of stimulation under 10^{-6} M peptide concentration [Figure 4D]. No wild-type specific T cell responses were observed after *in vitro* restimulations [Figure 4B, 4C, and 4D]. 5'-RACE PCR yielded one dominant TCRαß pair [Figure 4E]. 5'-RACE amplification procedure used primers annealing to constant TCRαß and the anchored 5'-adapter region as previously described[171].

A

TCR discovery pipeline of TRRAP S722F mutant

Week 1 — prime injection; Week 4 — 2nd injection; Week 8 — 3rd injection; Week 12 — 4th injection; Week 13 — T cell culture of whole splenocytes; Week 15 — IFN-γ capture of antigen-specific T cells, TCR isolation & characterization

B

Week 13 – CD8[+] T cell response in blood against S722F (7 days post 4th injection)

w/o peptide: 34,6 | 0 | 65,3 | 0,10
Anti-CD3/CD28 beads: 31,2 | 1,74 | 65,2 | 1,84
Wild-type peptide: 30,6 | 0 | 69,3 | 0,036
Mutant S722F peptide: 34,9 | 0,78 | 64,3 | 0,074

CD8 / IFN-γ

w/o peptide

4,6 | 0

65,3 | 0,10

8,0 | 0

1,5 | 0,50

w/o peptide

,10 | 0

,9 | 0,040

Anti-CD3/CD28 beads

31,2 | 1,74

65,2 | 1,84

Wild-type peptide

30,6 | 0

69,3 | 0,036

Mutant S722F peptide

34,9 | 0,78

64,3 | 0,074

C Week 15 – CD8⁺ T cells response in spleen against S722F (days post T cell culture)

w/o peptide

4865,2 | 1,84

18,0 | 0

51,5 | 0,50

Anti-CD3/CD28 beads

18,0 69,3 | 31,0,036

18,0 | 31,0

45,4 | 5,66

Wild-type peptide

49,7 64,3 | 0,074

49,7 | 0

49,7 | 0,70

Mutant S722F peptide

27,2 | 16,7

27,2 | 16,7

55,5 | 0,57

CD8 / IFN-γ

D Week 15 – IFN-γ⁺ S722F-reactive T cells sort for TCR isolation

w/o peptide

Wild-type peptide

8,10 | 0

91,8 | 0,058

Mutant S722F peptide

9,75 | 1,58

12,000 cells sorted

88,4 | 0,25

CD8 / IFN-γ

0,25

E

TRRAP TCR chains	V(D)J recombined regions	CDR3 regions	Clonally dominant TCR chains
TCR Vα chain	TRAV17-1-TRAJ58	CATDAEETSGSRLTF	11/14
TCR Vβ chain	TRBV27-1-TRBD2-2-TRBJ2-1	CASSLAGIYNEQFF	12/15
			Sub-clonally dominant TCR chains
TCR Vα chain	TRAV38-2-1-TRAJ28-1	CAYRSPYSGAGSYQ	2/20
TCR Vα chain	TRAV26-2-TRAJ29-1	CILRDDSGNTPLVF	1/23
TCR Vβ chain	TRBV27-1-TRBD2-2-TRBJ2-1	CASSLAGAYNEQFF	3/15

F

Titration of TRRAP epitopes on T2 cells

A375 (TRRAP mutant)
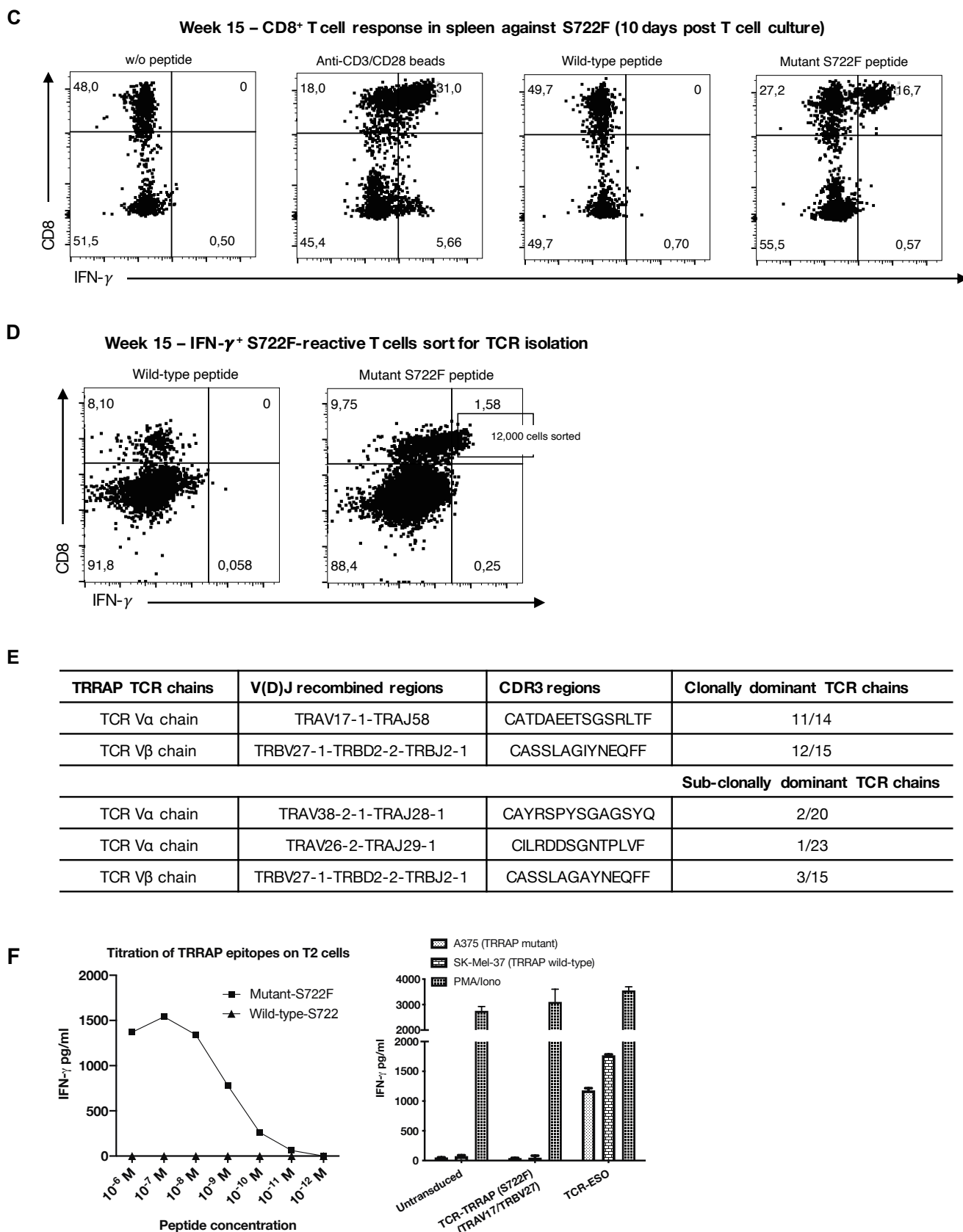SK-Mel-37 (TRRAP wild-type)
PMA/Iono

Figure 4: Generation and characterization of TRRAP S722F TCRs in ABabDII mice.

A. *TCR discovery pipeline with the immunization schedule.* *B.* *CD8⁺ T cells from ABabDII*
showed S722F peptide reactivity by IFN-γ production (gated on CD3⁺ cells, n = 3). Controls

*with wild-type S722 peptide or without (w/o) peptide. One representative flow cytometry plot out of three responder mice is given. **C.** In vitro cultured splenic CD8+ T cells (CD4-depleted) from responder mouse produced IFN-γ to 10^{-9} M peptide. **D.** IFN-γ capture assay enriched S722F-specific T cells (CD8+IFN-$γ^{high}$). **E.** 5'-RACE RT-PCR amplified clonally dominant and sub-dominant TCR variable chains. **F. Left,** Dominant $V_{αß}$ chains transduced in human T cells were co-cultured with $TAP^{-/-}$ T2 cells ($1×10^5$) loaded with titrated peptide nonamers (concentration: 10^{-6} M to 10^{-12} M) and IFN-γ release was measured by ELISA. **Right,** Endogenous processing assay with cell lines naturally expressing mutant TRRAP. IFN-γ levels measured after overnight co-culture of PBMCs ($1×10^5$) either with A375 cells (mutant-S722F+, NY-ESO+) or with SK-Mel-37 (wild-type-S722+, NY-ESO+). The TCR-ESO recognizes the NY-$ESO_{157-165}$ epitope. Bar graph represents mean ± SD (n = 2). One representation out of three co-culture experiments is given.*

Transduced T cells were co-cultured with peptide-titrated T2 ($TAP^{-/-}$) cells to ensure TCR interaction and its affinity to S722F. The dominant TCR $V_α17$-$V_ß27$ pair recognized S722F sensitive enough up to 10^{-11} M peptide concentration by producing IFN-γ with no wild-type epitope recognition [Figure 4F, left]. TRRAP+ human cell lines were co-cultured with transduced T cells to verify proteasomal processing of epitope S722F. Transduced T cells neither recognized A375 nor SK-Mel-37 cells; positive for the mutant and wild-type TRRAP, respectively. NY-$ESO_{157-165}$, the HLA-A2 epitope is expressed on melanoma cells A375 and SK-Mel-37. TCR-ESO (control TCR) recognized NY-$ESO_{157-165}$ epitope by producing IFN-γ [Figure 4F, right].

Hence, we prove that the TRRAP mutation-derived neoantigen $S722F_{715-723}$ nonamer is not endogenously processed and presented on HLA-A2.

3.3 AB*ab*.I transgene designed to reflect natural human HLA haplotype

The 23 mutated epitopes from genes, TRRAP, C-KIT, XPO1, FOXA1, RAC1, RAC2, RHOT1, TP53, MAP2K1, EGFR, TRAF7, SMO, SF3B1, FBXW7, SPOP, EZH2, FLT3, NFE2L2, TSHR, PTPN11, and NOTCH1 are predicted to bind to HLA-A2 [Figure 2C, pie chart]. Mutant peptides from the first four genes were analyzed similarly for immunogenicity like TRRAP-S772F [Figure 4A]. AB*ab*DII mice did not produce CD8+ T cell responses after immunizations with C-KIT (K642E), XPO1 (E571K), and FOXA1 (D226N) epitopes [data not shown]. The AB*ab*DII mice provide the opportunity to isolate TCRs restricted only to these 23 HLA-A2-binders.

68/152 non-HLA-A2 binders [Figure 2C, pie chart] are predicted to bind to at least one allele of a novel HLA haplotype, A*0301-A*11:01, B*07:02-B*15:01, and C*04:01-C*07:02 resembling HLA situation in humans (every individual bears six genes). To understand the biology of a whole human haplotype in mice and target the other 68 high-affinity mutants [Table

14], we designed AB*ab*.I transgene based on the *in silico* screen results to reflect a human HLA haplotype [Figure 5].

Table 14: *In silico* predicted mutant epitopes binding AB*ab*.I HLA haplotype

HLA restriction	Gene	Mutant peptide sequence (mutant AA in bold) / IC$_{50}$ nM	Wildtype peptide sequence (wild-type AA in bold) / IC$_{50}$ nM	Mutation frequency
HLA-A*03:01	EGFR	**A**SGAFGTVYK / 23 nM	**G**SGAFGTVYK / 46 nM	3%
	KLF4	KTYT**Q**SSHLK / 10 nM	KTYT**K**SSHLK / 12 nM	1%
	PIK3CA	AISTRDPLS**K** / 44 nM	AISTRDPLS**E** / 19220 nM	3.7%
	FOXA1	SLSFN**N**CFVK / 29 nM	SLSFN**D**CFVK / 77 nM	1.8%
	ABL1	**K**VYEGVWKK / 31 nM	**E**VYEGVWKK / 239 nM	2.3%
	JAK3	**G**LLKTVSYK / 37 nM	**E**LLKTVSYK / 104 nM	5.6%
	NOTCH1	RV**P**HTNVVFK / 38 nM	RV**L**HTNVVFK / 15 nM	5.5%
	PPP2R1A	**R**MVRRAAASK / 26 nM	**P**MVRRAAASK / 742 nM	9.6%
HLA-A*11:01	TP53	SSCMG**S**MNR / 35 nM	SSCMG**G**MNR / 58 nM	Hot-spot mutation (freq. unknown)
	KRAS	VVGA**V**GVGK / 40 nM	VVGA**G**GVGK / 89 nM	22%
	KRAS	VVGA**F**GVGK / 26 nM	VVGA**G**GVGK / 89 nM	1%
	KRAS	VVGA**S**GVGK / 35 nM	VVGA**G**GVGK / 89 nM	4.7%
	HRAS	VVGAG**I**VGK / 43 nM	VVGA**G**GVGK / 89 nM	Activating mutation (freq. unknown)
	EGFR	**A**SGAFGTVYK / 13 nM	**G**SGAFGTVYK / 29 nM	3%
	LRRN3	TIESLPNLK**K** / 44 nM	TIESLPNLK**E** / 32442 nM	1.5%
	KLF4	KTYT**Q**SSHLK /11 nM	KTYT**K**SSHLK /12 nM	1%
	XPO1	KTVVNKLF**K** / 24 nM	KTVVNKLF**E** / 16111 nM	1.8%
	FBXW7	TV**C**CMHLHEK /22 nM	TV**R**CMHLHEK / 22 nM	3.9%
	CTNNB1	TTAP**P**LSGK **/** 11 nM	TTAP**S**LSGK / 10 nM	4.0%
	CTNNB1	TTAP**A**LSGK / 16 nM	TTAP**S**LSGK / 10 nM	1.3%
	CTNNB1	TTAP**F**LSGK / 7 nM	TTAP**S**LSGK / 10 nM	2.7%
	CTNNB1	TTAP**Y**LSGK / 10 nM	TTAP**S**LSGK / 10 nM	1.3%
	PIK3CA	AISTRDPLS**K** / 12 nM	AISTRDPLS**E** / 27122 nM	3.7%
	ABL1	**A**MEVEEFLK / 36 nM	**T**MEVEEFLK / 38 nM	1.1%
	ABL1	SSA**T**EYLEK / 9 nM	SSA**M**EYLEK / 9 nM	2.8%
	FGFR3	**M**TTNGRLPVK / 20 nM	**K**TTNGRLPVK / 25 nM	2%
	JAK3	**G**LLKTVSYK / 19 nM	**E**LLKTVSYK / 247 nM	5.6%
	NOTCH1	RV**P**HTNVVFK / 22 nM	RV**L**HTNVVFK / 14 nM	5.5%
	PTPN11	ATLA**G**LVQYY / 40 nM	ATLA**E**LVQYY / 73 nM	0.7%
HLA-B*07:02	STK19	**N**PIFRFSSL / 14 nM	**D**PIFRFSSL / 246 nM	5%
	IDH2	SPNGTI**Q**NIL / 42 nM	SPNGTI**R**NIL / 23 nM	8.7%
	NOTCH1	F**P**RELSRVL / 11 nM	F**L**RELSRVL / 282 nM	5.5%

	NOTCH1	S**P**VLHTNVV / 26 nM	S**R**VLHTNVV / 23262 nM	5.2%
	PTPN11	MVRSQRS**A**M / 9 nM	MVRSQRS**G**M / 59 nM	0.7%
	TSHR	VPRFL**T**CNL / 30 nM	VPRFL**M**CNL / 47 nM	12%
HLA-B*15:01	RAC2	F**Q**GEYIPTVF / 35 nM	F**P**GEYIPTVF / 6216 nM	Driver mutation (freq. unknown)
	TRAF7	AQSYLYS**S**SY / 29 nM	AQSYLYS**G**SY / 33 nM	2.3%
	SMO	VLAPIGLV**F** / 40 nM	VLAPIGLV**L** / 328 nM	2.3%
	CTNNB1	HSGATTTAP**F** / 40 nM	HSGATTTAP**S** / 9015 nM	2.7%
	CTNNB1	HSGATTTAP**Y** / 43 nM	HSGATTTAP**S** / 9015 nM	1.3%
	ABL1	TQISSA**T**EY / 44 nM	TQISSA**M**EY / 48 nM	2.8%
	EZH2	VQKNEFISE**F** / 31 nM	VQKNEFISE**Y** / 37 nM	5.6%
	GATA2	YLCNACG**F**Y / 47 nM	YLCNACG**L**Y / 61 nM	2.4%
	IDH2	I**Q**NILGGTVF / 39 nM	I**R**NILGGTVF /6936 nM	8.7%
	NOTCH1	FQSATDVA**P**F / 26 nM	FQSATDVA**A**F / 28 nM	2.1%
	TSHR	RMAVLIFT**E**F /33 nM	RMAVLIFT**D**F / 45 nM	2.4%
	TSHR	RMAVLIFT**H**F / 23 nM	RMAVLIFT**D**F / 45 nM	2.4%
HLA-C*04:01 SB: strong-binder (< 1000 nM) WB: Weak binder (> 1000 nM)	TP53	YLDDRNTF**L** / SB	YLDDRNTF**R** / WB	Hot-spot mutation (freq. unknown)
	STK19	SAPE**N**PIFRF / SB	SAPE**D**PIFRF / WB	5%
	TRRAP	VFGSV**F**LF / SB	VFGSV**S**LF / WB	4%
	ZNF831	FSDAQRPS**F** / SB	FSDAQRPS**S** / WB	2%
	FOXA1	**R**FENGCYL / SB	**M**FENGCYL / WB	1.8%
	MED12	IVDGAVFAV**F** / SB	IVDGAVFAV**L** / WB	5.4%
	ABL1	FYIITE**L**M / SB	FYIITE**F**M / SB	1.8%
	DNMT3A	YTDVSNMS**H**L / SB	YTDVSNMS**R**L / SB	11%
	DNMT3A	YTDVSNMS**C**L / SB	YTDVSNMS**R**L / SB	3.2%
	FLT3	R**Y**IMSDSNYV / SB	R**D**IMSDSNYV / WB	2.7%
	MET	MYDKEY**D**SV / SB	MYDKEY**Y**SV / SB	14%
	NOTCH1	VY**P**EIDNRQCV / SB	VY**L**EIDNRQCV / WB	5.5%
	PTPN11	YYDLYGGEKF**V** / SB	YYDLYGGEKF**A** / SB	0.7%
HLA-C*07:02 SB: strong-binder (< 1000 nM) WB: Weak binder (> 1000 nM)	STK11	S**Y**LGVAEAL / SB nM	S**D**LGVAEAL / WB	1.1%
	SMO	FVLAPIGLV**F** / SB nM	FVLAPIGLV**L** / SB nM	2.3%
	MED12	FAV**F**KAVFVL / SB nM	FAV**L**KAVFVL / SB nM	5.4%
	RET	LAARNILV**F** / SB nM	LAARNILV**A** / WB nM	6%

48

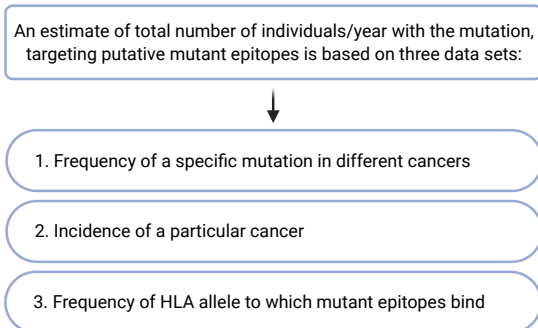|				
TSHR	RFL**T**CNLAF / SB nM	RFL**M**CNLAF / SB nM	12%
TSHR	RMAVLIFT**E**F / SB nM	RMAVLIFT**D**F / WB nM	2.4%
TSHR	YTHSEYYNHA**F** / SB nM	YTHSEYYNHA**I** / WB nM	3.5%
TSHR	YAKVS**T**CLPM / SB nM	YAKVS**I**CLPM / SB nM	4.7%

Prediction programs suggested: HLA-A*03:01 to bind a total of 8 mutant epitopes, and A*11:01 bound the maximum number of 21 epitopes. B*07:02 bound 6 and B*15:01 bound 12 epitopes, C*04:01 bound 13 epitopes, and C*07:02 showed *in silico* restriction to 8 epitopes [Figure 5A and Table 14].

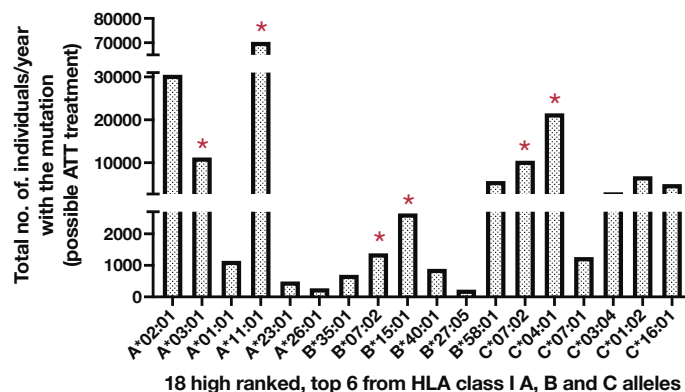**A**

In silico screen of 266 putative mutant epitopes against 18 HLA class I alleles								
Top 6 high ranked HLA-A	Mutant epitopes predicted	Allele frequency (Europe)	Top 6 high ranked HLA-B	Mutant epitopes predicted	Allele frequency (Europe)	Top 6 high ranked HLA-C	Mutant epitopes predicted	Allele frequency (Europe)
A*02:01	23	26%	B*07:02	6	10%	C*07:01	9	16%
A*03:01	8	13%	B*35:01	6	7%	C*04:01	13	12%
A*01:01	5	12%	B*15:01	12	5%	C*07:02	8	12%
A*11:01	21	7%	B*27:05	7	5%	C*03:04	6	5%
A*23:01	3	4%	B*40:01	5	4%	C*16:01	21	4%
A*26:01	3	3%	B*58:01	9	2%	C*01:02	12	3%

B

An estimate of total number of individuals/year with the mutation, targeting putative mutant epitopes is based on three data sets:

↓

1. Frequency of a specific mutation in different cancers

2. Incidence of a particular cancer

3. Frequency of HLA allele to which mutant epitopes bind

C

D

LS: Leader sequence; Linker: 15-aa polypeptide chain; TM: Transmembrane domain; CYT: Cytoplasmic domain; pA: Polyadenylation signal
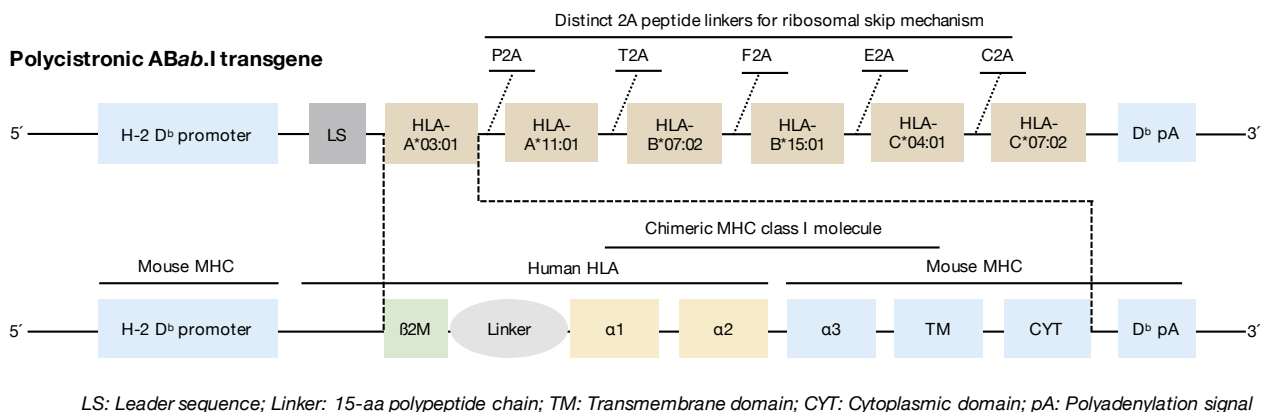
Figure 5: Design and construction of polycistronic ABab.I transgene.

A. Overview of total numbers of in silico predicted mutant epitopes across 18 HLA alleles. The top 6 HLA alleles (shaded in blue) reflect a natural human HLA haplotype (the genes present in the ABab.I transgene). B. Multiplying the three data sets estimated the no. of individuals/year with the mutation. C. Total no. of individuals/year with the mutation (possible for ATT treatment) was calculated for 175 mutant binders against 18 alleles, using data from A. and B. Red asterisks highlight top high-ranked HLA alleles. D. Top, Polycistronic ABab.I transgene designed to encode six HLA alleles segregated via five different 'self-cleaving' viral 2A peptide linkers driven by a single H-2 D^b promoter and polyadenylation signal. Bottom, A single monochain construct design is shown. Each HLA is a chimeric human-mouse fusion monochain containing α1 and α2 heavy chain cDNA regions from respective human HLA alleles, while α3, transmembrane and cytoplasmic domains from murine H-2 D^b gene fused by a glycine-serine polylinker to human ß2m.

The total number of individuals/year with the mutation based on three factors predict the possibility of ATT treatment worldwide [Figure 5B]. ~ 11,220 A*03:01$^+$ individuals/year were estimated to bear mutations that create epitopes to target. Similarly, ~ 70,389 A*11:01$^+$, ~ 1,381 B*07:02$^+$, ~ 2,651 B*15:01$^+$, ~ 21,532 C*04:01$^+$, and ~ 10,451 C*07:02$^+$ per year were theoretically calculated as possible mutation bearing individuals [Figure 5C]. In total, ~ 117,624 individuals/year were estimated worldwide carrying these 68 recurrent point mutations in their cancers; predicted epitopes as listed [Table 14] to bind at least one allele of the novel class I haplotype. Six highly represented HLA alleles in the AB*ab*.I transgene design in resemblance to human HLA architecture was encoded as a polycistronic expression cassette [Figure 5D].

3.4 Polycistronic HLA transgene is transcribed to produce mRNA but not translated to express proteins

Polycistronic primary AB*ab*.I and all modified constructs [Figure 6A] did not express the HLA alleles after staining with the pan HLA ABC antibody [Figure 6B]. The HLA expression was measured 48 hours after transient transfection. Polycistronic, H-2 D^b driven AB*ab*.I showed no surface HLA expression in MCA205 cells. 39% of cells expressed GFP. 98% of LCL-BM14 cells showed endogenous HLA expression [Figure 6B, I]. LCL-BM14 are EBV-transformed lymphoblastoid B cells.

Six HLAs were switched with enhanced GFP (eGFP) gene or CMV promoter for H-2 D^b to access promoter activity [Figure 6A, II]. 13% of MCA205 cells expressed GFP driven by H-2 D^b, and 53% of cells expressed GFP from control pMP71-eGFP retroviral plasmid [Figure 6B, II]. However, switching the H-2 D^b promoter with CMV did not rescue the expression of the six HLAs [Figure 6B, II].

Next, the 500 bp H-2 D^b polyadenylation (pA) sequence was replaced with a 1 kb full-length (FL) H-2 D^b or synthetic 350 bp bovine growth hormone (bGH) pA to analyze for defects in transcription termination. H-2 D^b or CMV promoter drove pA modified constructs, which were analyzed for surface expression using flow cytometric staining [Figure 6A, III]. None of the pA modified constructs showed significant HLA surface expression after transient transfection [data not shown].
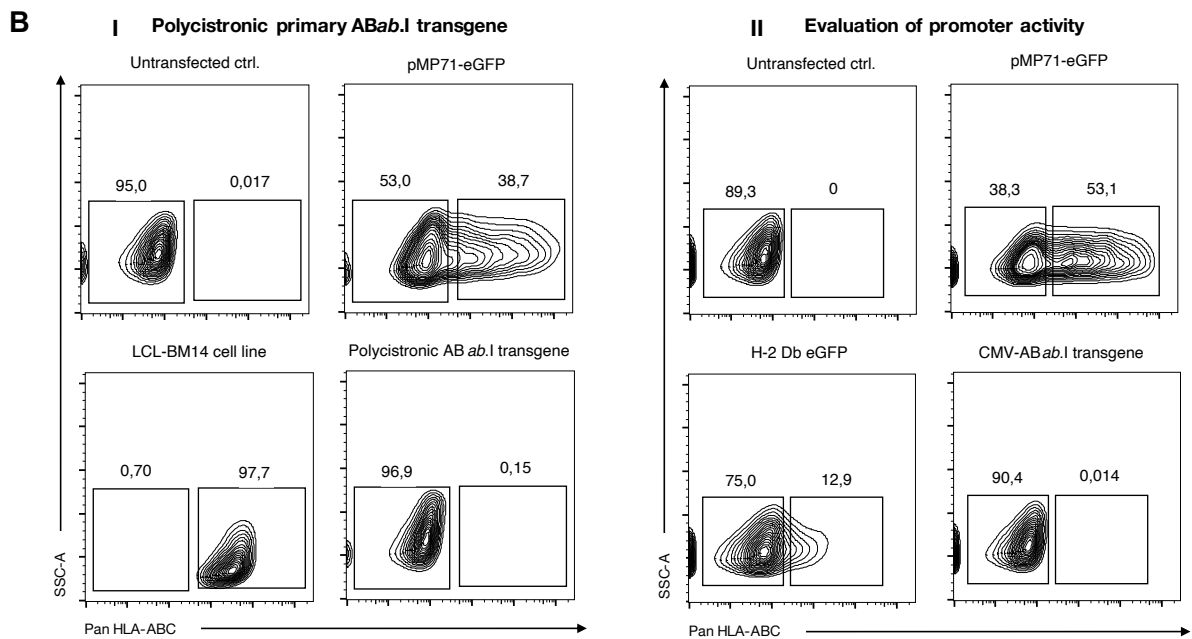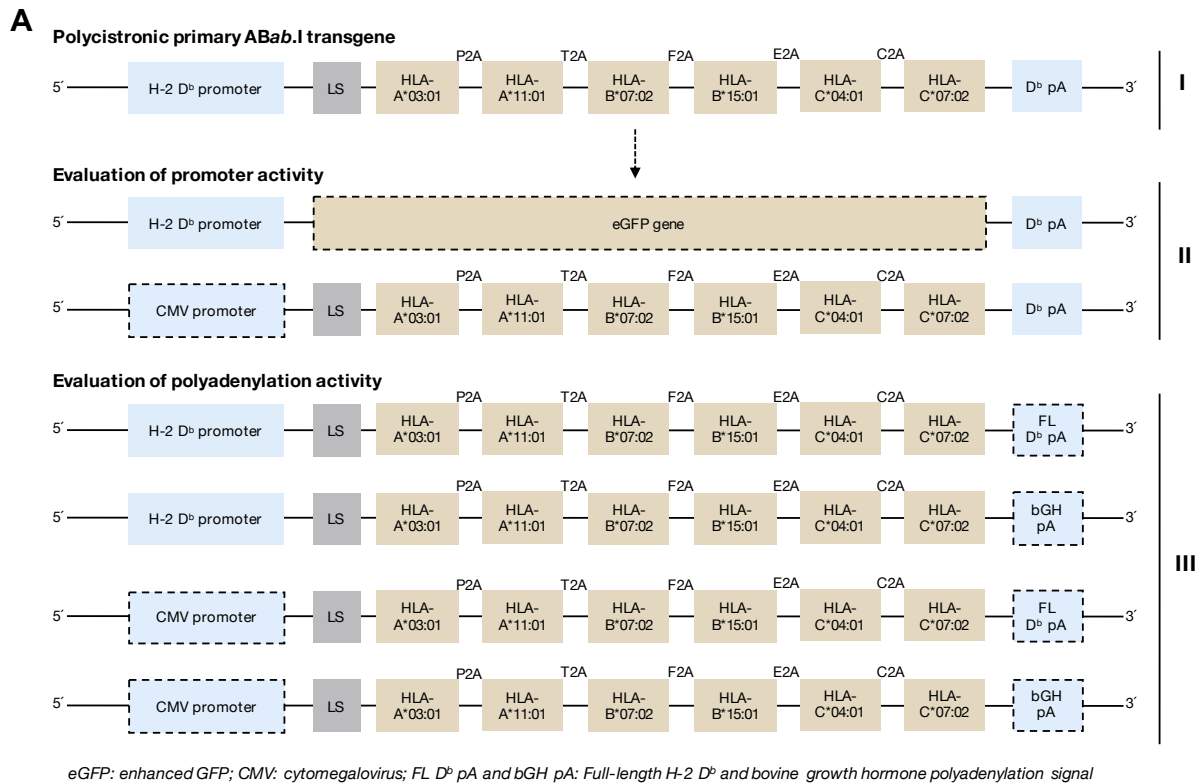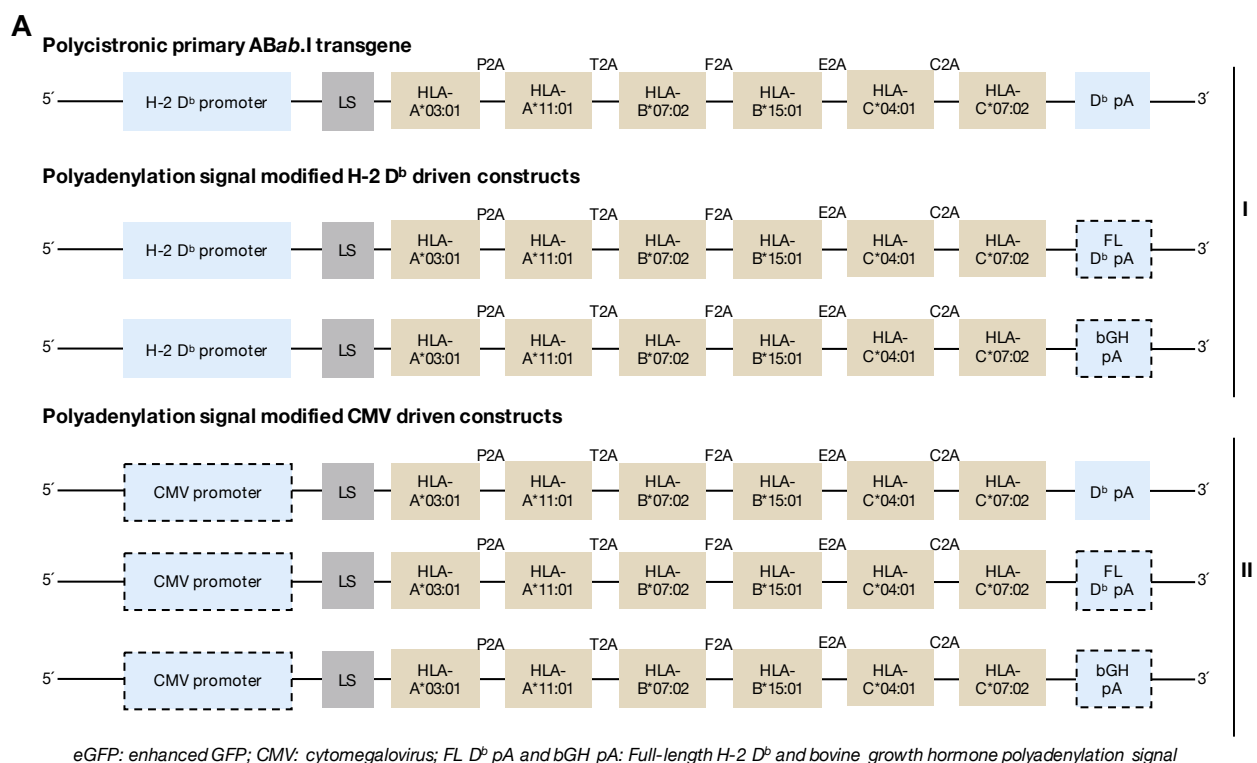

Figure 6: HLA protein expression analysis of the polycistronic ABab.I transgenes.

*A. The ABab.I transgenes: primary construct (I) and modified versions (II, III) from the initial design. **I.** Primary transgene in which viral peptide 2A linkers segregate the HLAs. **II.** Modified transgenes: H-2D^b promoter-driven eGFP gene and CMV promoter-driven HLAs. **III.** Modified transgenes with different polyadenylation signal sequences. Dotted boxes represent modifications to the initial primary construct. **B.** Surface staining of transiently transfected MCA205 cells with constructs from **A. I. and II.** is shown (**III.** is not shown). Cells were stained 48 hours post-transfection using pan HLA-ABC antibody (clone: W6/32). Controls: pMP71-eGFP (GFP vector, measured in same channel) and LCL-BM14 (lymphoblastoid cells), positive for HLA ABC. One representation out of multiple flow cytometry experiments is given (n > 5).*

H-2 D^b promoter- and CMV promoter-driven primary ABab.I construct with all respective pA modifications [Figure 7A] produced mRNA in MCA205 cells [Figure 7B]. HLA mRNA was detected using cDNA amplification with specific primers against three viral peptide linkers, P2A (front region), F2A (middle), and C2A (end region). The H-2 D^b promoter-driven bGH pA construct produced up to 7% of HLA expression compared to that of mouse MHC expression (endogenous D^b gene in MCA205 cells) after transient transfection. The H-2 D^b promoter-driven FL pA and the primary ABab.I construct respectively produced 6% and 2% of HLA expression [Figure 7B, I]. A 193% of relative change in the mRNA production to that of the endogenous D^b expression in CMV-driven bGH pA constructs was significantly measured. The FL pA and CMV-primary ABab.I constructs did show a relative change, respectively, of 138% and 75% from basal H-2 D^b [Figure 7B, II].
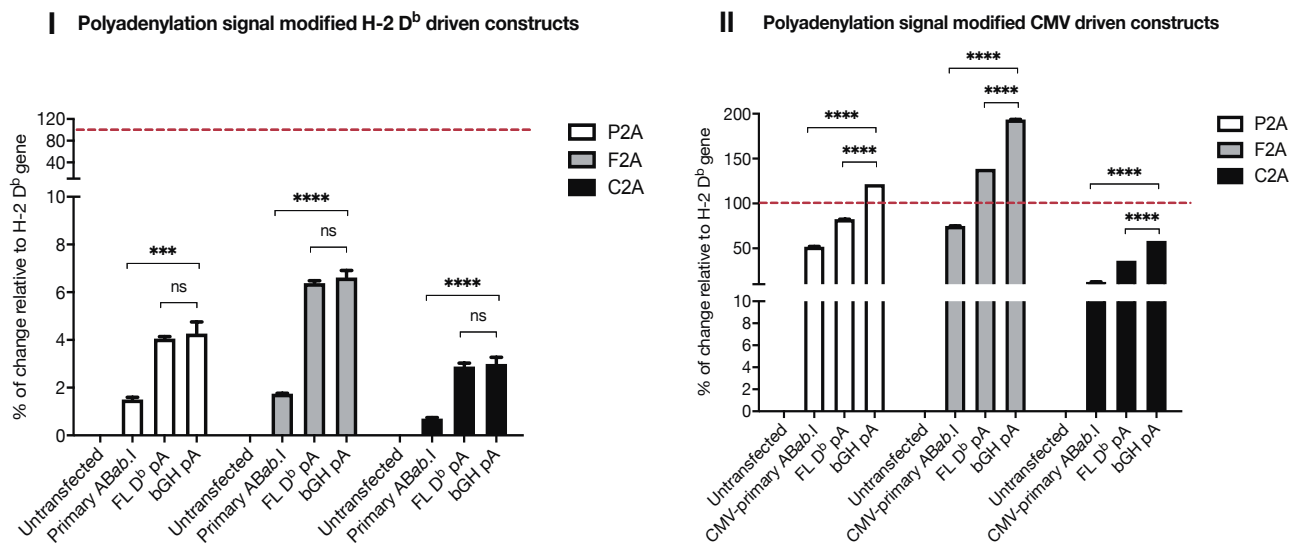
A

eGFP: enhanced GFP; CMV: cytomegalovirus; FL D^b pA and bGH pA: Full-length H-2 D^b and bovine growth hormone polyadenylation signal

B

Figure 7: mRNA expression profile of polycistronic ABab.I transgenes.

*A. I., Top, The ABab.I primary transgene. I., Bottom, Polyadenylation (pA) signal-modified constructs driven by H-2 D^b or II. CMV viral promoter. B. Quantitative RT-PCR analysis of mRNA transcripts from constructs A., I. and II. is shown. cDNA regions of ABab.I transgenes were amplified from MCA205 cell lysates 48 hours after transfection using primers flanking 2A peptide linkers (P2A, F2A, and C2A). Percentage change relative to constitutive H-2 D^b, mouse MHC class I gene (red dotted lines) is compared. Bar graphs represent mean ± SD (n = 3). P values: ****, P ≤ 0.0001; ***, P ≤ 0.001; ns, not significant (unpaired two-tailed t-test).*

3.5 HLA alleles are stably expressed when driven by its own 5' and 3' regulatory elements

H-2 D^b promoter driving six HLA alleles in polycistronic configuration with bGH pA did produce HLA mRNA, but not proteins [Figure 6 and 7].

Six HLA alleles were cloned with or without ß2 microglobulin (ß2m) as monocistrons [Figure 8A, I] to test their expression before they were placed together into a single class I haplotype, each with its regulatory elements [Figure 8A, II]. Monochains without ß2m (linked) were co-transfected exogenously with ß2m plasmid. Respective leader sequences led HLAs to translocate in the ß2m non-linked monochains, e.g., A*03 leader peptide was cloned instead of ß2m leader sequence in the HLA-A*03:01 construct.
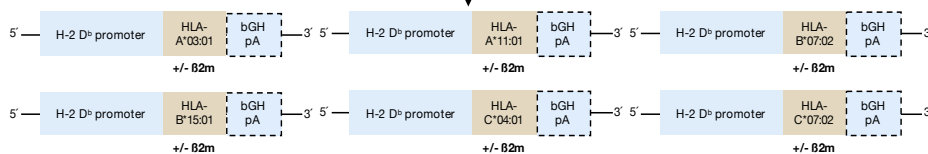
MCA205 cells were transfected with or without ß2m-linked monochains [Figure 8A, I] in the presence or absence of IFN-γ. HLA monochains transfected without ß2m (ß2m co-transfected) did not show any significant difference in the surface expression of HLA alleles [data not shown]. Here, all six monocistrons (as ß2m-linked monochains) were stained using pan HLA ABC and anti-human ß2m antibody and measured by flow cytometry in transiently transfected MCA205 cells [Figure 8B].
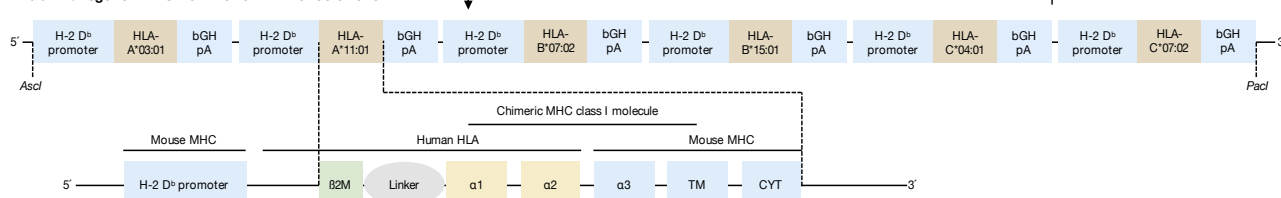
A

Polycistronic primary AB*ab*.I transgene

Single HLA monochains tested for expression with or without ß2m

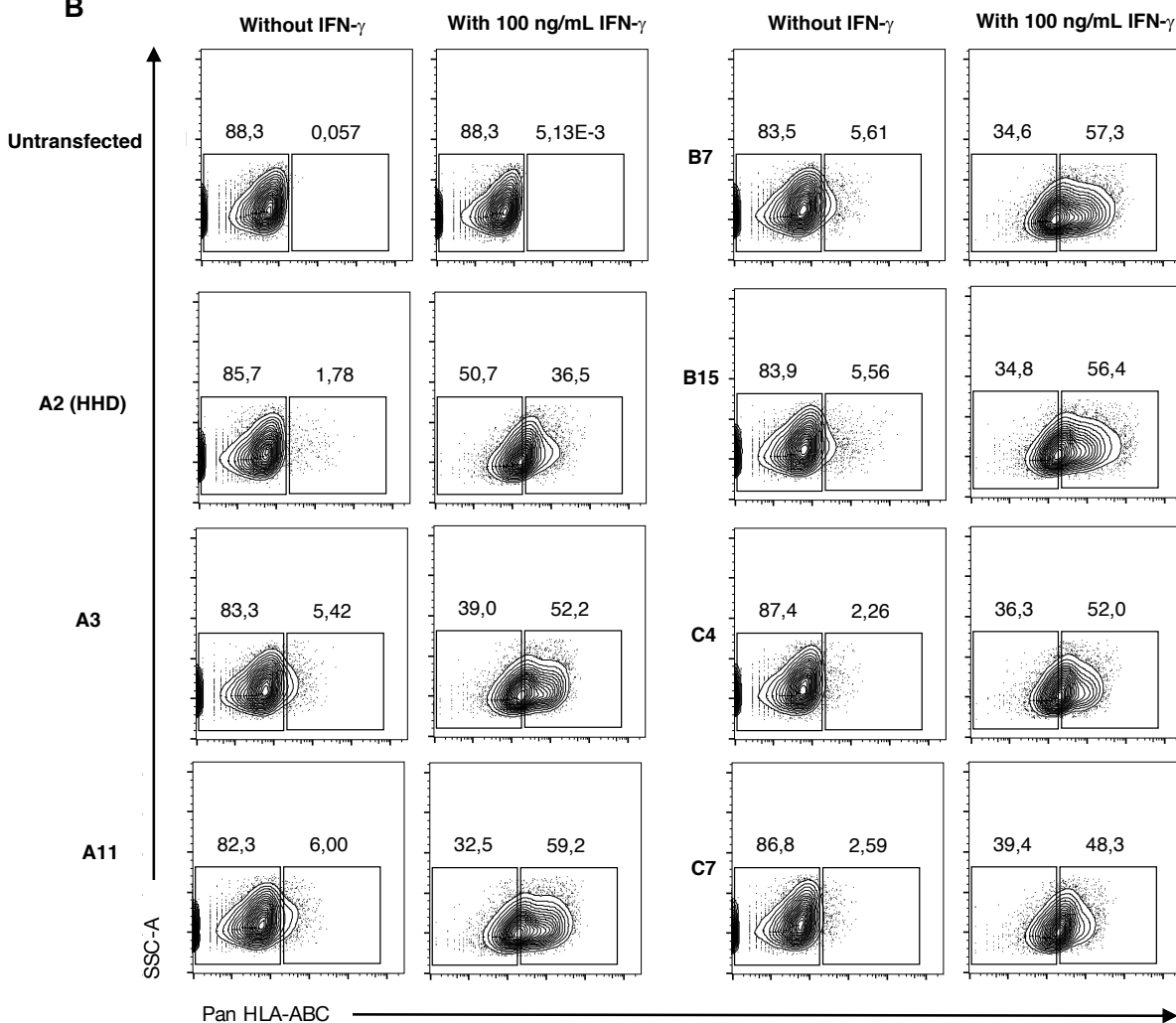Transgene construct with each HLA monochains driven by its own regulatory elements

AB*ab*.I transgene with six chimeric HLA monocistrons

LS: Leader sequence; Linker: 15-aa polypeptide chain; TM: Transmembrane domain; CYT: Cytoplasmic domain; bGH pA: bovine growth hormone polyadenylation signal

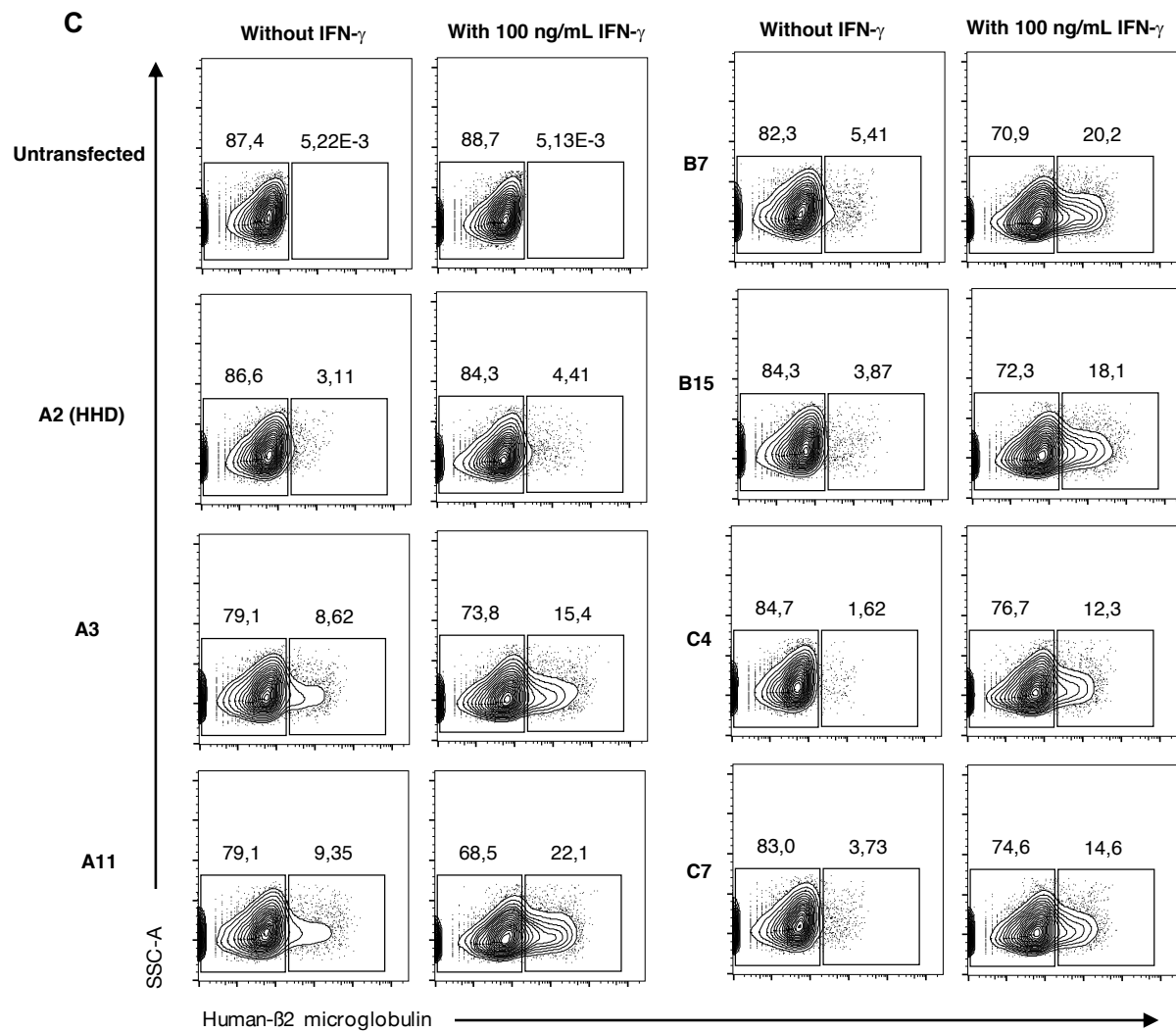## I Surface expression profile of HLA monochains

B

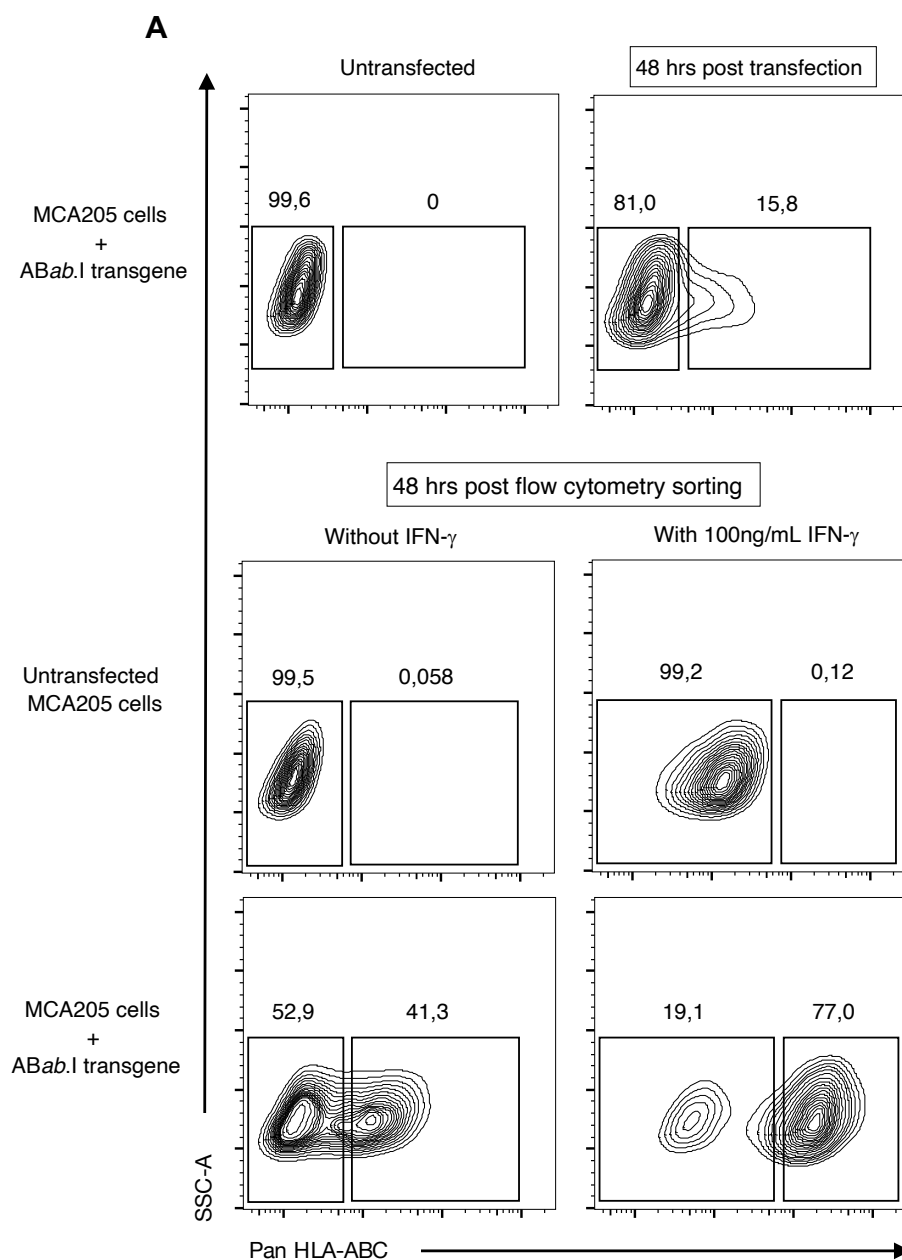**Figure 8: HLA expression analysis of single monocistrons from ABab.I transgene.**

A. I. Single monochains, with or without ß2m, were cloned from ABab.I primary transgene. Every HLA allele as a monocistron is driven by its own H-2D^b^ promoter and pA signal. Dotted boxes represent additional modifications from the primary construct. II. The ABab.I transgene expresses six monocistrons (as shown in A. I.) as a single haplotype (5'-to-3') with its own regulatory elements. B. and C. MCA205 cells expressing six monochains were surface stained with or without IFN-γ pre-treatment. The HLA-A2 protein was stained for reference. Cells were stained 48 hours post-transfection using, B. pan HLA-ABC antibody (clone: W6/32), C. human-ß2 microglobulin antibody (clone: TÜ99). One representative flow cytometry plot out of three experiments is given.

Pan HLA ABC antibody stained transiently transfected MCA205 cells and the increase in HLA expression was in the range of 9- to 23-fold in A*03 (5.4% to 52%), A*11 (6% to 59%), B*07 (5.6% to 57%), B*15 (5.6% to 56%), C*04 (2.3% to 52%) and C*07 (2.6% to 48%) monochains with 100 ng/ml IFN-γ treatment for 48 hours during transfection [Figure 8B]. Similarly, anti-
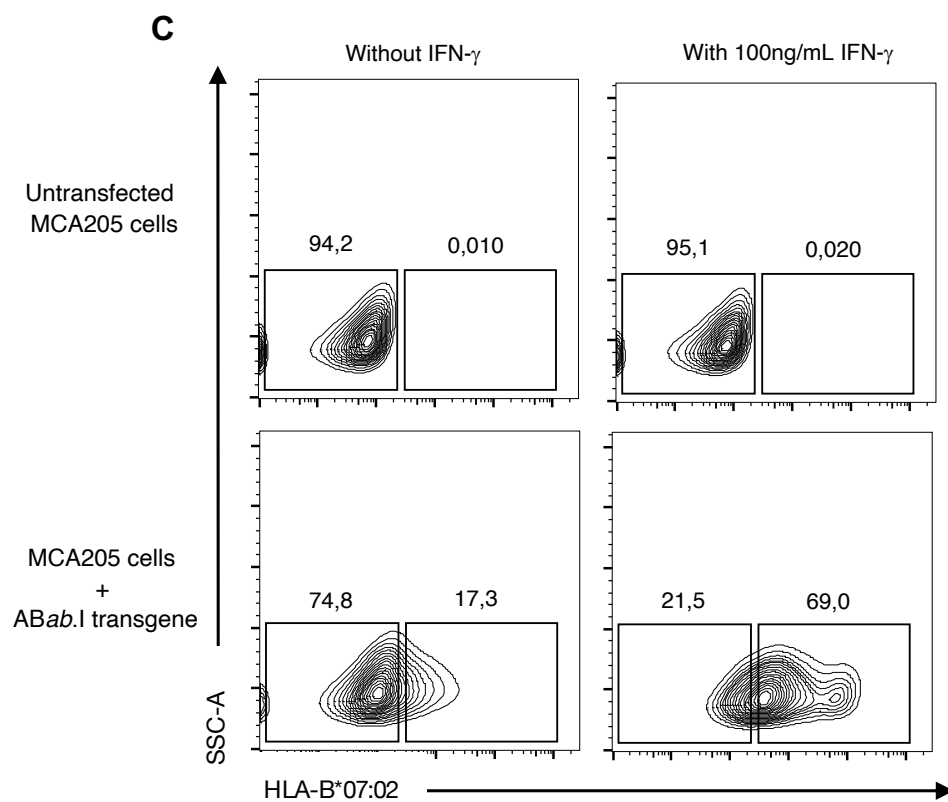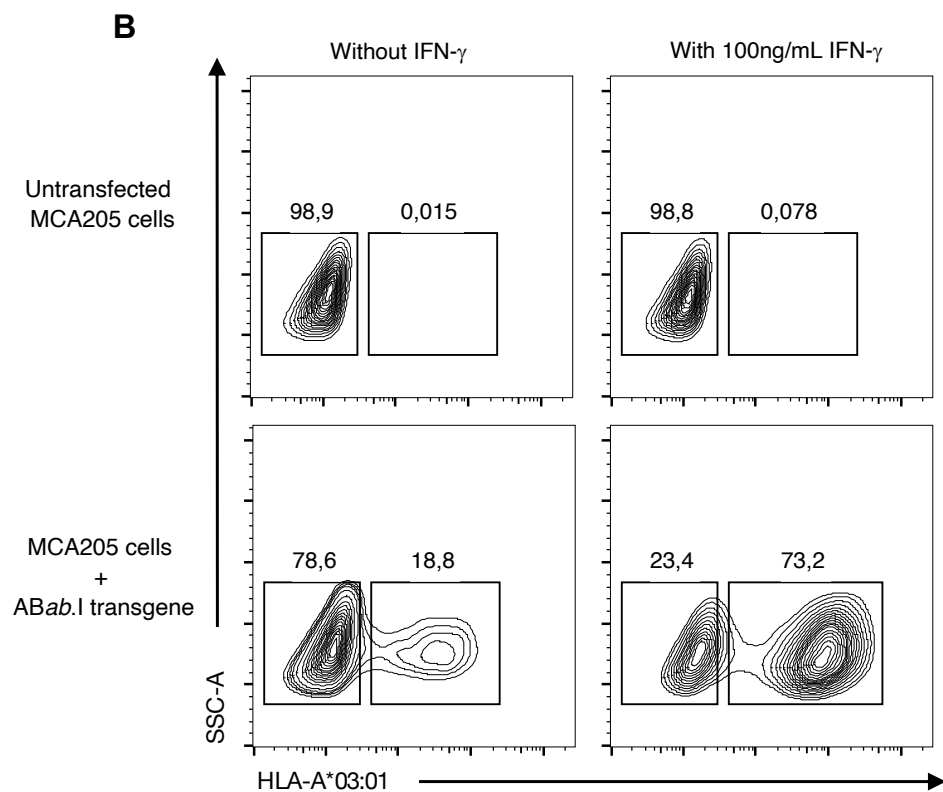
human ß2m stained transiently transfected MCA205 cells, the increase in HLA expression was in the range of 1.5- to 8-fold in A*03 (8.6% to 15%), A*11 (9.3% to 22%), B*07 (5.4% to 20%), B*15 (3.9% to 18%), C*04 (1.6% to 12%) and C*07 (3.7% to 14.6%) monochains with 100 ng/ml IFN-γ treatment for 48 hours during transfection [Figure 8C].

The AB*ab*.I transgene, encoding six chimeric monocistrons, was transfected into the MCA205 cell line for HLA expression analysis. AB*ab*.I HLA haplotype encodes six alleles back-to-back in a head-to-tail configuration. All six genes comprise a promoter and a polyadenylation signal of their own [Figure 8A, II].

II **ABab.I transgene with six chimeric HLA monocistrons as a single class I haplotype**

A

Transient

MCA205 cells
+
AB*ab*.I transgene

Untransfected

99,6 | 0

48 hrs post transfection

81,0 | 15,8

Transfection with AscI-PacI linearized ABab.I transgene
Staining after multiple flow sort enrichments

48 hrs post flow cytometry sorting

12 weeks post transfection

48 hrs post flow cytometry sorting, stained for HLA-specific antibodies

Without IFN-γ | With 100ng/mL IFN-γ | Without IFN-γ

Untransfected
MCA205 cells

99,5 | 0,058 — 99,2 | 0,12 — 98,9 | 0,015

MCA205 cells
+
AB*ab*.I transgene

52,9 | 41,3 — 19,1 | 77,0 — 78,6 | 18,8

SSC-A

Pan HLA-ABC

48 hours post flow cytometry sorting, stained for HLA-specific antibodies

Transient
12 weeks post transfection

12 weeks post transfection

With 100ng/mL IFN-γ

B

Without IFN-γ
Without IFN-γ

With 100ng/mL IFN-γ
With 100ng/mL IFN-γ

Without IFN-γ
Without IFN-γ

99,2   0,12

Untransfected
Untransfected
MCA205 cells

Transient

| 99,5 | 0,058 |
| 98,9 | 0,015 |

Untransfected

| 99,2 | 0,12 |
| 98,8 | 0,078 |

48 hrs post transfection

| 98,9 | 0,015 |

MCA205 cells
+
ABab.I transgene
MCA205 cells
MCA205+ cells
ABab.I transgene
ABab.I transgene

19,1   77,0

99,6	0
52,9	41,3
78,6	18,8

81,0	15,8
19,1	77,0
23,4	73,2

| 78,6 | 18,8 |
| 78,6 | 18,8 |

SSC-A

HLA-A*03:01

Transient

12 weeks post transfection

C

Without IFN-γ
Without IFN-γ

With 100ng/mL IFN-γ
With 100ng/mL IFN-γ

Without IFN-γ
Without IFN-γ

With 100ng/mL IFN

Without IFN-γ

Untransfected
MCA205 cells
Untransfected
MCA205 cells

5,1   0,020

| 94,2 | 0,010 |

| 95,1 | 0,020 |

| 98,9 | 0,015 |
| 94,2 | 0,045 |
| 0,045 |

MCA205 cells
+
ABab.I transgene
MCA205 cells
+
ABab.I transgene

1,5   69,0

| 74,8 | 17,3 |

| 21,5 | 69,0 |

| 78,6 | 18,8 |
| 58,9 | 31,6 |

SSC-A

HLA-B*07:02

52 weeks post transfection

52 weeks post transfection
Stable

Without IFN-γ
Untransfected

With 100ng/mL IFN-γ
ABab.I transgene expressing
MCA205 stable cell line

52 weeks post transfection

Without IFN-γ

.I transgene expressing
CA205 stable cell line
Stable

ABab.I transgene expressing
MCA205 stable cell line

| 98,3 | 0,017 |
Untransfected

| 7,39 | 91,5 |
ABab.I transgene expressing
MCA205 stable cell line

20   91,5

Figure 9 (continued): Panel D shows flow cytometry contour plots.

D — Without IFN-γ / With 100ng/mL IFN-γ

Untransfected MCA205 cells:
- Without IFN-γ: 94,2 / 0,045
- With 100ng/mL IFN-γ: 93,5 / 0,081

MCA205 cells + ABab.I transgene:
- Without IFN-γ: 58,9 / 31,6
- With 100ng/mL IFN-γ: 25,2 / 69,0

Axes: SSC-A (y), Pan HLA-C (x)

E Stable

Transfection with AscI-PacI linearized ABab.I transgene
Staining after multiple flow sort enrichments

Untransfected:
- 98,3 / 0,017

ABab.I transgene expressing MCA205 stable cell line
48 hrs post flow cytometry sorting:
- 7,39 / 91,5

48 hrs post flow cytometry sorting, stained for HLA-specific antibodies

Axes: SSC-A (y), Pan HLA-ABC (x)

Figure 9: HLA haplotype expression profile of ABab.I transgene in MCA205 cell line.

*A. MCA205 cells were stained using pan HLA-ABC antibody. **Top,** HLA class I haplotype expressing six alleles (depicted in **Figure 8 A. II.**) was stained 48 hours post-transfection. **Bottom,** Transfected MCA205 cells were enriched for HLA+ cell population and stained for HLA alleles with or without IFN-γ pre-treatment (48 hours prior staining). **B., C., and D.** Post-enrichment, MCA205 cells were stained using HLA-specific antibodies, pre-treated with or without IFN-γ. **B.** HLA-A*03:01 specific antibody (clone: GAP.A3) stained HLA-A3 allele. **C.***

Two days after transfection, 16% of cells were positive for HLA alleles [Figure 9A, top]. Flow cytometry sorting enriched HLA$^+$ population in the transfected cells. Two days post-sorting and culturing *in vitro*, cells were pre-treated with or without IFN-γ and stained using pan HLA ABC antibody. 41% of cells stained for HLA without IFN-γ, whereas 77% of MCA205 cells stained for HLA alleles after enrichment in the presence of IFN-γ for 48 hours in culture [Figure 9A, bottom]. Simultaneously, the sorted cells (HLA$^+$) were stained using HLA-specific antibodies. The increase in HLA expression was in the range of 2- to 4-fold after staining with A*03-specific (19% to 73%), B*07-specific (17% to 69%), and pan HLA-C (32% to 69%) antibodies after 48 hours of 100 ng/ml IFN-γ treatment [Figure 9B, 9C, and 9D]. AscI-PacI digested and linearized AB*ab*.I DNA was used for MCA205 transfection to produce a stable cell line, MCA205/AB*ab*.I. After multiple enrichments of HLA$^+$ cells by flow cytometry sorting at different time points, 92% of cells expressed the AB*ab*.I transgene [Figure 9E].

3.6 PiggyBac transposon-mediated targeting produced AB*ab*.I founder mice with six HLAs as a single haplotype
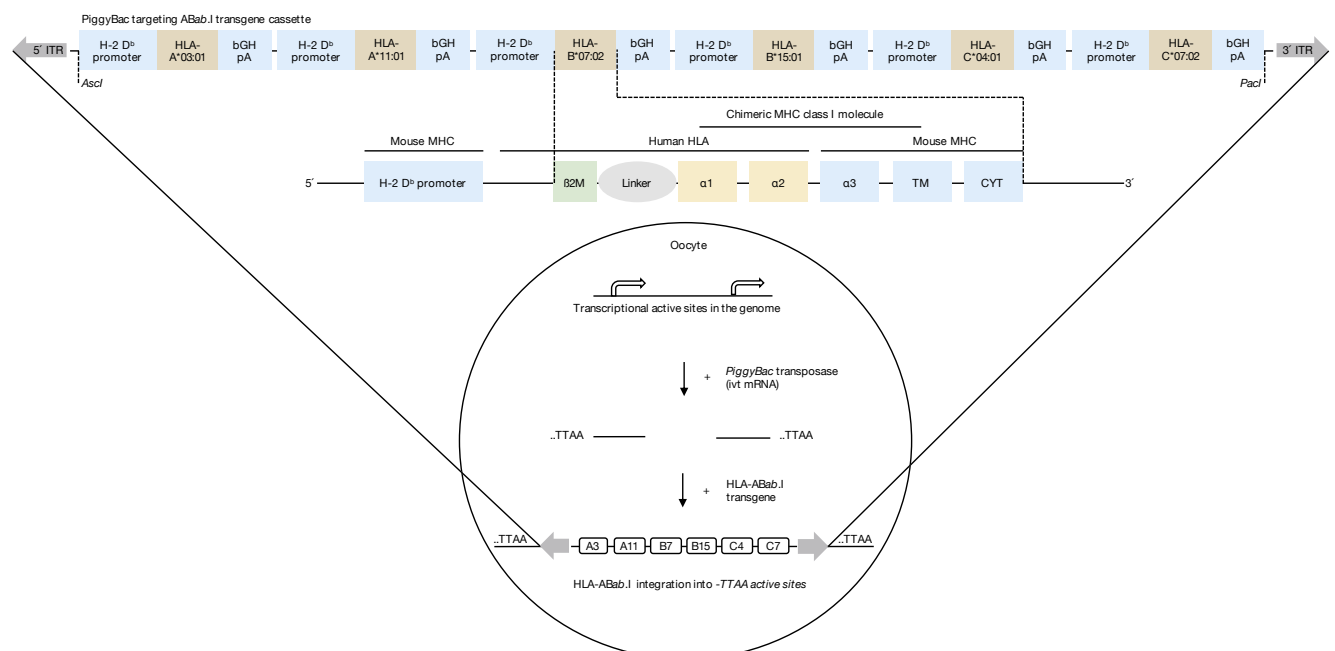

***Figure 10: PiggyBac transposons target ABab.I transgene into oocytes.***

Schematic representation of PiggyBac (PB) transposon strategy for the generation of ABab.I mice using pronuclei microinjection technology. The PB targeting ABab.I targeting transgene

*(flanked by ITRs for PB transposase catalysis) has repeating elements of six chimeric HLA genes constructed one after another (5'-to-3') as depicted in **Figure 7 A. II.** each with its own promoter and 3'-UTR components.*

AB*ab*.I HLA cassette was shown to be both transiently [Figure 9A and 9B] and stably [Figure 8C] expressed on the surface of murine MCA205 cells *in vitro*.

In vivo integration of the HLA haplotype was achieved using the hyperactive PiggyBac (PB) transposon system. The PB transposase acts on the ITRs-flanked targeting vector (cassette bearing HLAs) by a cut-paste mechanism. The enzyme was co-transfected as mRNA for the excision (cut) of 5' and 3' ITR regions to release the insert (HLAs) for integration (paste) into transcriptionally active sites of the mouse genome [Figure 10].

After successful rounds of pronuclei injections, allele-specific PCRs confirmed the presence of six HLA alleles in two founder animals [Figure 11A]. Ear biopsy DNA confirmed the genomic integration of alleles in F0_Q4115 and F0_Q4118 using genotyping PCRs [Figure 11A, left]. As controls, 18s rRNA housekeeping gene and human TCR αß genes were checked [Figure 11A, right]. Three additional founders: F0_Q6844, F0_Q8552, and F0_Q8556 were positive genotyped [data not shown]. Stable homozygous AB*ab*.I mice strain from the Q6844 line was established according to the standard breeding pattern [Figure 11B].
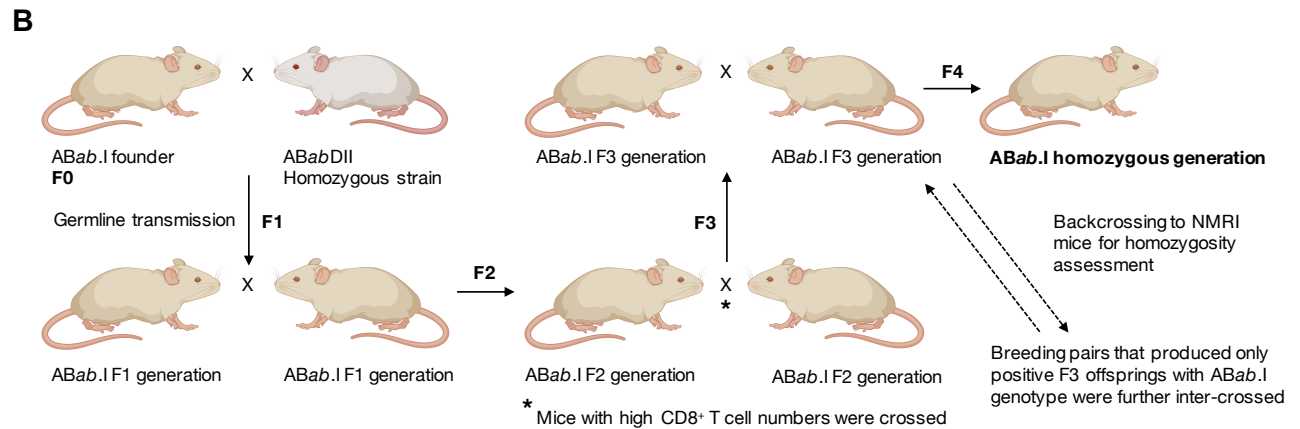
B
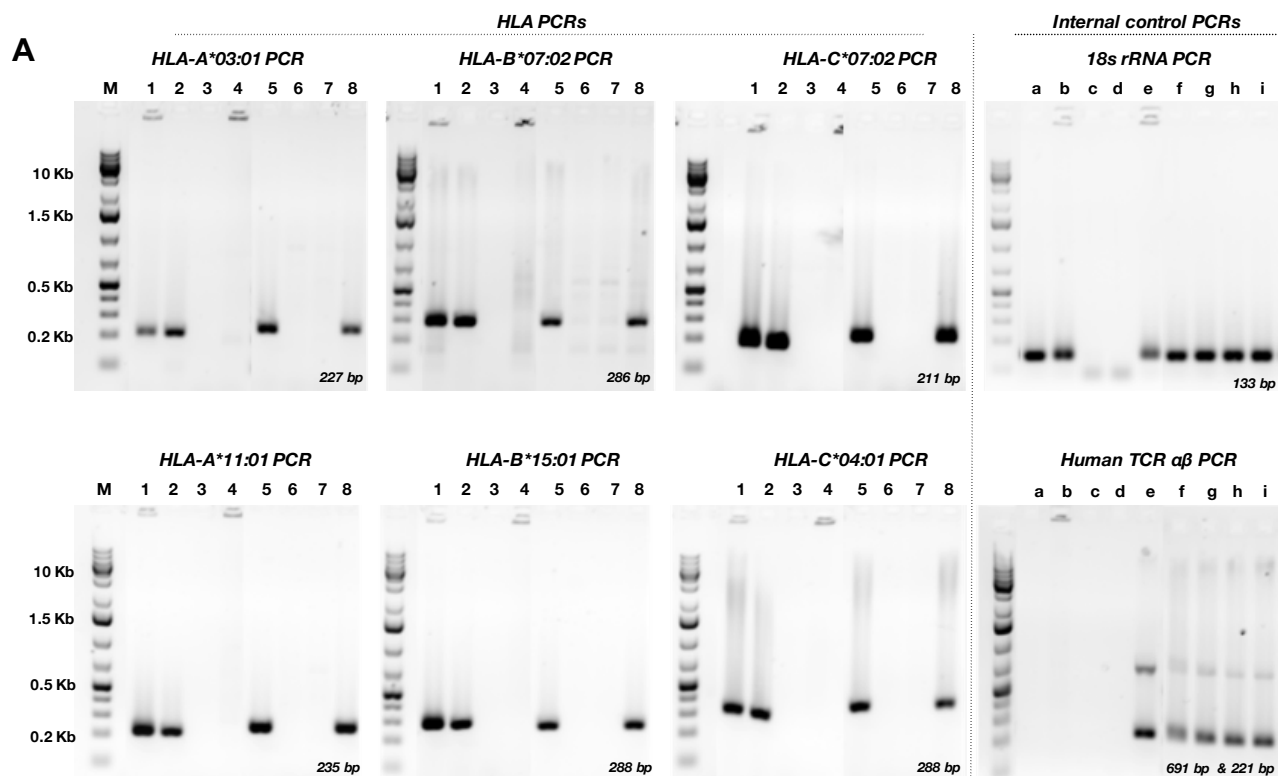
Figure 11: Genotyping profile and breeding scheme of ABab.I transgenic founders.

A. Left, HLA-PCR reactions: *PCR genotyping data of 2 founder mice (F0), Q4115 and Q4118.*
HLA- A*03, A*11, B*07, B*15, C*04, and C*07 PCRs: *Genomic (g) DNA from MCA205/ABab.I stable cell line (lane 1), ABab.I transgene as plasmid DNA (lane 2), control H_2O (lane 3), gDNA from C57BL6/N mice (lane 4), positive ABab.I, 1st founder: Q4115 (lane 5), two negative mice (lane 6 and 7), positive ABab.I, 2nd founder: Q4118 (lane 8).*

Right, Internal control PCRs: 18s rRNA and human TCR αß PCRs: *Genomic (g) DNA from C57BL6/N mice (lane a), gDNA from MCA205/ABab.I stable cell line (lane b), ABab.I transgene as plasmid DNA (lane c), control H_2O (lane d), gDNA from ABabDII mice (lane e), 1st founder: Q4115 (lane f), two negative mice (lane g and h), positive ABab.I, 2nd founder: Q4118 (lane i).*

B. *Breeding scheme representation from F0 to F4 homozygous mice. F0 backcrossed with ABabDII for germline transmission. NMRI backcrossing was used for homozygosity testing.*

HLA expression was stained and quantified in the peripheral blood cells of C57BL6/N, AB*ab*DII, and AB*ab*.I mice using anti-human ß2m antibody with LCL-BM14 cells (99% ß2m$^+$) as control [Figure 12A and 12B]. 3% and 7% of lymphocytes from AB*ab*DII mice, whereas 35% and 40% of lymphocytes from AB*ab*.I mice stained for HLA expression with the ß2m antibody [Figure 12A]. Shown here are two AB*ab*.I mice from founder line Q6844. Other founders tested for HLA expression included Q4115, Q4118, Q8552, and Q8556 [data not shown]. Mouse to mouse variations was analyzed by staining multiple mice per strain. Mean fluorescence intensity (MFI) comparison of ß2m revealed 4-times more HLA expression in AB*ab*.I compared to that in AB*ab*DII. 120-times higher MFI was observed in human HLA (non-chimeric) expressing LCL-BM14 cells [Figure 12C and 12D]. One possible explanation could be that the folding characteristics of human HLA proteins differ from that of a chimeric fusion molecule, which could hinder epitope accessibility to ß2m antibody, which needs to be further analyzed.
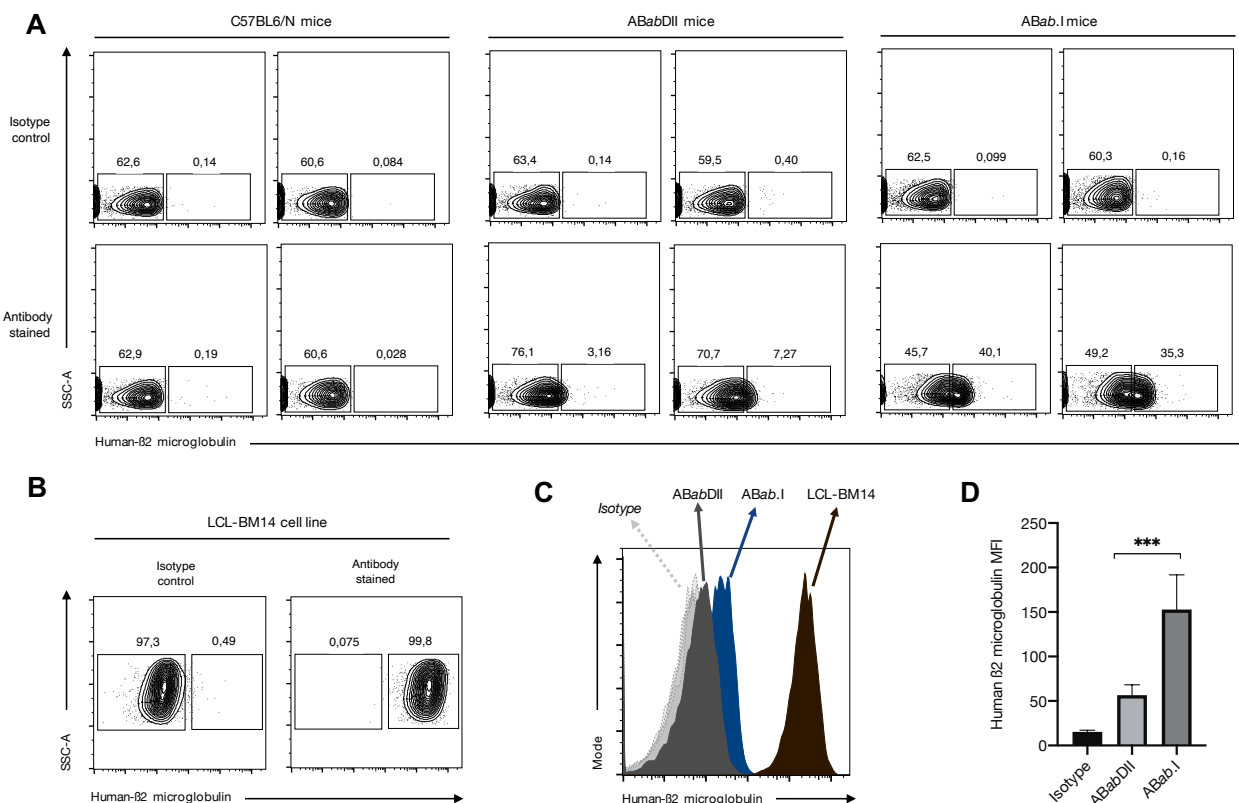
Figure 12: Human HLA class I haplotype expression in ABab.I mice.

*A. HLA staining of peripheral blood cells from C57BL6/N, ABab.I and ABabDII mice was performed using ß2 microglobulin (ß2m) antibody. Two representative plots (two mice/strain) out of multiple experiments are given (n > 10). B. LCL-BM14 (lymphoblastoid cells), positive for HLA ABC, stained with ß2m antibody. C. Histogram shows mean fluorescence intensity (MFI) of ß2m staining among mouse strains: ABabDII and ABab.I, and cell line, LCL-BM14. One representative staining out of multiple mice is given (n > 20). D. ß2m MFI compared between ABabDII and ABab.I mice. Bar graph represents mean ± SD (n = 5). P value: ***, P ≤ 0.001 (unpaired two-tailed t-test).*

3.7 ABab.I mice have higher CD8⁺ T cell counts compared with the single HLA-bearing ABabDII mice

Phenotypic T cell characterization depicted increased thymic output/peripheral survival in ABab.I mice (founder F0_Q6844 lineage) compared with ABabDII and HHD mice [Figure 13]. HHD mice have a mouse TCR repertoire selected on a human HLA-A2[183]. CD8⁺ CD4⁺ profiling of peripheral blood cells (on CD3⁺ cells) from four mouse strains by flow cytometry showed 13% and 14% of CD8⁺ T cells in ABabDII, around 6% in both HHD mice, 36% and 45% in ABab.I, and 36% and 41% in C57BL6/N mice [Figure 13A]. Shown here are two mice/strain. CD8⁺ T cell fractions (%) in ABab.I mice were comparable to that in C57BL6/N mice. On average, CD3⁺ T cells were 90% CD4⁺ in HHD mice, 64% in ABabDII, 45% in ABab.I, and 55%

in C57BL6/N mice [Figure 13A]. A significant number of AB*ab*.I mice accumulated 4-times more CD8$^+$ T cell numbers in the periphery than AB*ab*DII mice [Figure 13B]. AB*ab*.I mice with high CD8$^+$ T cell counts were inter-crossed at F-2 to produce a stable mouse line [Figure 11B]. No significant difference in the CD4$^+$ T cell numbers was observed [Figure 13C]. CD8/CD4 ratio was higher in AB*ab*.I mice, similar to that in C57BL6/N mice, 3-times more than that in AB*ab*DII mice [Figure 13D].
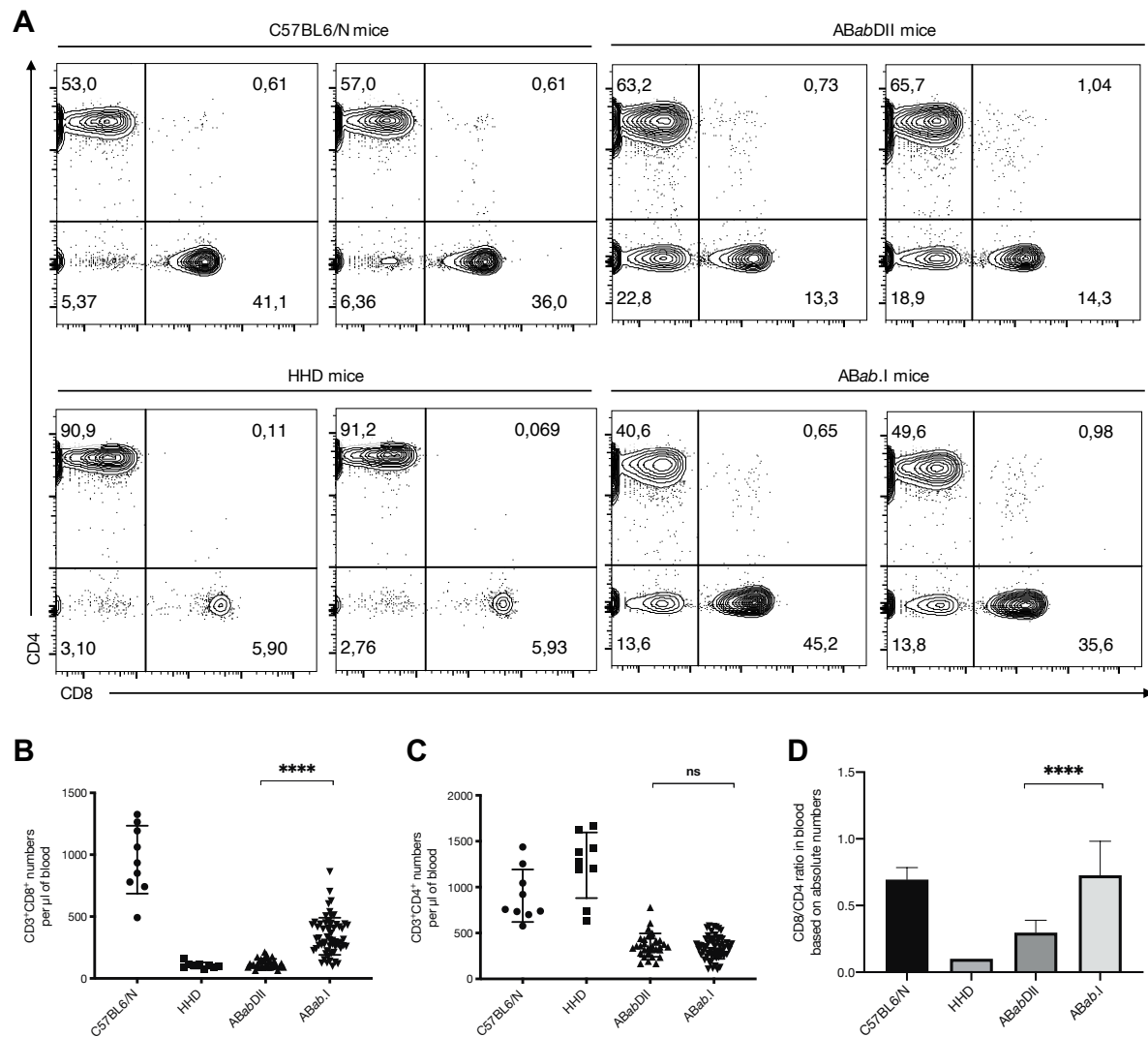

Figure 13: Phenotypic characterization of peripheral blood T cells in ABab.I mice.

*A. Peripheral blood cells from young (8-12 weeks) indicated mouse strains were stained with antibodies specific for CD3, CD8, and CD4 chains and analyzed by flow cytometry (gated on CD3$^+$ lymphocytes). Two representative plots (two mice/strain) out of multiple experiments is given (n > 30). **B. and C.** Absolute T cell numbers/µl blood, **B.** CD3$^+$CD8$^+$ and **C.** CD3$^+$CD4$^+$ were quantified using CountBright™ beads by flow cytometry. Each data point represents a single young mouse from indicated strains. Horizontal intervals on scattered charts represent mean ± SD. Summarized data from, C57BL6/N (n = 9), HHD (n = 9), ABabDII (n = 32), and ABab.I (n = 60) mice. **D.** CD8/CD4 ratio based on absolute numbers in peripheral blood among*

Comparison of CD8$^+$ and CD4$^+$ T cell populations in pooled lymphoid organs (on CD3$^+$ cells) showed 14% and 12% of CD8$^+$ T cells in AB*ab*DII with 62% and 67% of CD4$^+$ T cells, and 47% and 44% of CD8$^+$ T cells in AB*ab*.I with 39% of CD4$^+$ T cells [Figure 14A]. On average, AB*ab*.I mice estimated to have 3.5-times more absolute CD8$^+$ T cell numbers in the major secondary lymphoid organs [Figure 14B] with no significant difference in absolute CD4$^+$ T cell numbers [Figure 14C]. Similar to peripheral blood, CD8/CD4 ratio was higher in AB*ab*.I mice, 2.5-times more than that in AB*ab*DII mice [Figure 14D].
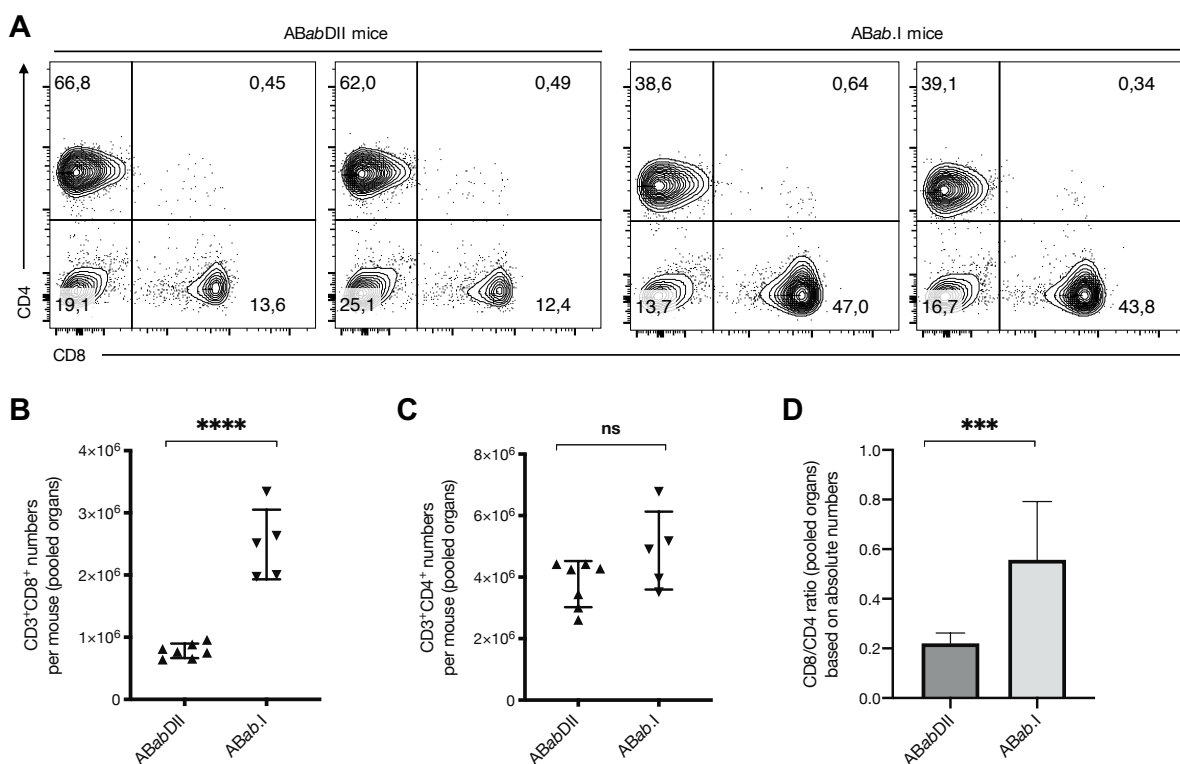

Figure 14: Phenotypic characterization of T cells from lymphoid organs in ABab.I mice.
*A. Spleen and lymph nodes including whole blood from young (8-12 weeks) indicated mice strains were stained with antibodies specific for CD3, CD8, and CD4 chains and analyzed by flow cytometry (gated on CD3$^+$ lymphocytes). Two representative plots (two mice/strain) out of multiple experiments is given (n > 10). **B. and C.** Absolute T cell numbers/mouse, **B.** CD3$^+$CD8$^+$ and **C.** CD3$^+$CD4$^+$ were quantified using CountBright™ beads by flow cytometry. Each data point represents a single young mouse from indicated strains. Horizontal intervals on scattered charts represent mean ± SD. Summarized data from, ABabDII (n = 7), and ABab.I (n = 5) mice. **D.** CD8/CD4 ratio based on absolute numbers in pooled organs among different*

*mice strains. Bar graphs represent mean ± SD (n = 5). P values in **B., C., and D.** indicate: ****,*
P ≤ 0.0001; ***, P ≤ 0.001; ns, not significant (unpaired two-tailed t-test).

3.8 Broader CD8⁺ T cell repertoire in AB*ab*.I mice compared with AB*ab*DII mice

TCRß repertoire deep sequencing quantified $11.2 \pm 2.7 \times 10^4$ and $4.8 \pm 0.80 \times 10^4$ numbers of unique in-frame amino acid (aa) clonotypes in the CD8⁺ T cells of AB*ab*.I and AB*ab*DII mice, respectively [Figure 15A]. From the three age-unmatched humans, the 30-year old donor had similar numbers of clonotypes, $11.4 \pm 4.1 \times 10^4$, compared to AB*ab*.I mice [Figure 15A].

The total number of unique aa clonotypes in human donors had a high variance. Considering widespread differences in age (30, 50, and 65 years old), we excluded human data from the clone-size analysis. AB*ab*.I mice had 8.5% of rare clones, whereas AB*ab*DII had 5.7% rare clones [Figure 15B]. However, AB*ab*DII mice had three times more small clones of 4.1% than AB*ab*.I mice with 1.4% of small clones [Figure 15B].

Sequencing of the complete repertoire with full coverage was not possible using ImmunoSEQ™ deep sequencing. So, the estimation of the total TCR repertoire in CD8⁺ T cells of mice and humans was determined using computational statistics.

The iChao1 estimator calculated the total number of clonotypes per mouse (observed and undetected) using information from clones that occurred only once or twice[216]. iChao1 estimates true species richness using lower bound rarely occurring ones. The first human donor, a 30-year old individual had the most diverse repertoire with up to 2×10^6 clonotypes with 1×10^6 on average (n=3). After humans, AB*ab*.I mice had 0.66×10^6 clonotypes on average (n=5) with a diverse repertoire. AB*ab*DII had the lowest estimation with 1.8×10^5 clonotypes [Figure 15C].
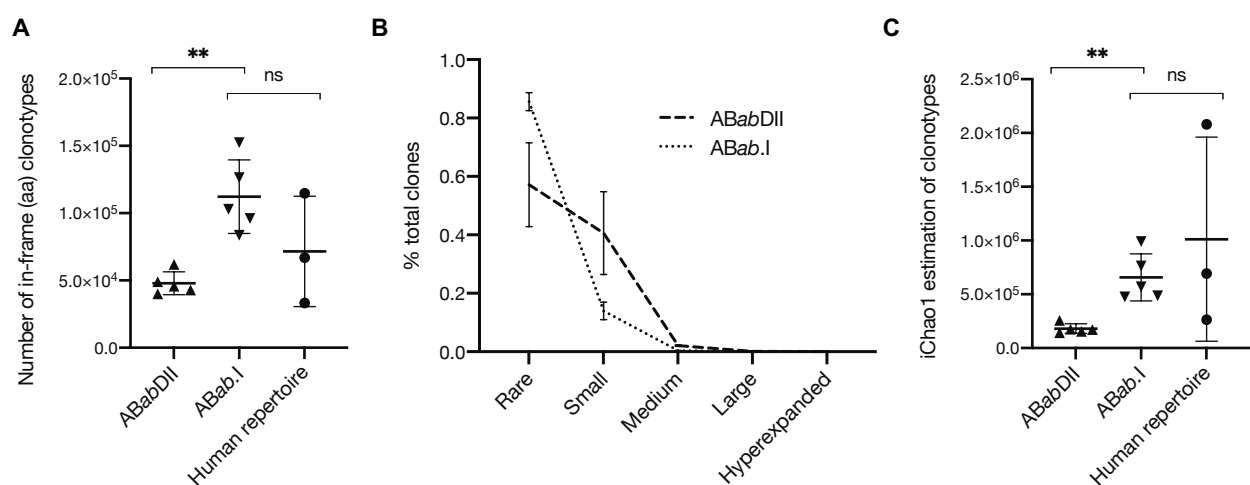

Figure 15: Comparison of TCRß repertoire among ABabDII, ABab.I, and human donors.
A. Absolute numbers of unique TCRß amino acid (aa) clonotypes within 3×10^5 (average in mice strains) or 1.8×10^5 human CD8⁺ T cells. B. Clonal distribution of TCRß aa clonotypes of

*different sizes: rare, 0 < x ≤ 0.001%; small, 0.001% < x ≤ 0.01%; medium 0.01% < x ≤ 0.1%; large, 0.1% < x < 1%; hyperexpanded, 1% < x < 10%. Human donors were excluded from this analysis. **C.** TCRß diversity was calculated using the iChao1 estimator (lower bound richness of total numbers of unique templates within an individual repertoire). Horizontal intervals on scattered charts represent mean ± SD. Data are from, ABabDII (n = 5); ABab.I (n = 5) mice, and humans (n = 2 for **B.**; n = 3 for **A. and C.**). P values in **A. and C.** indicate: **, P ≤ 0.01; ns, not significant (unpaired two-tailed t-test).*
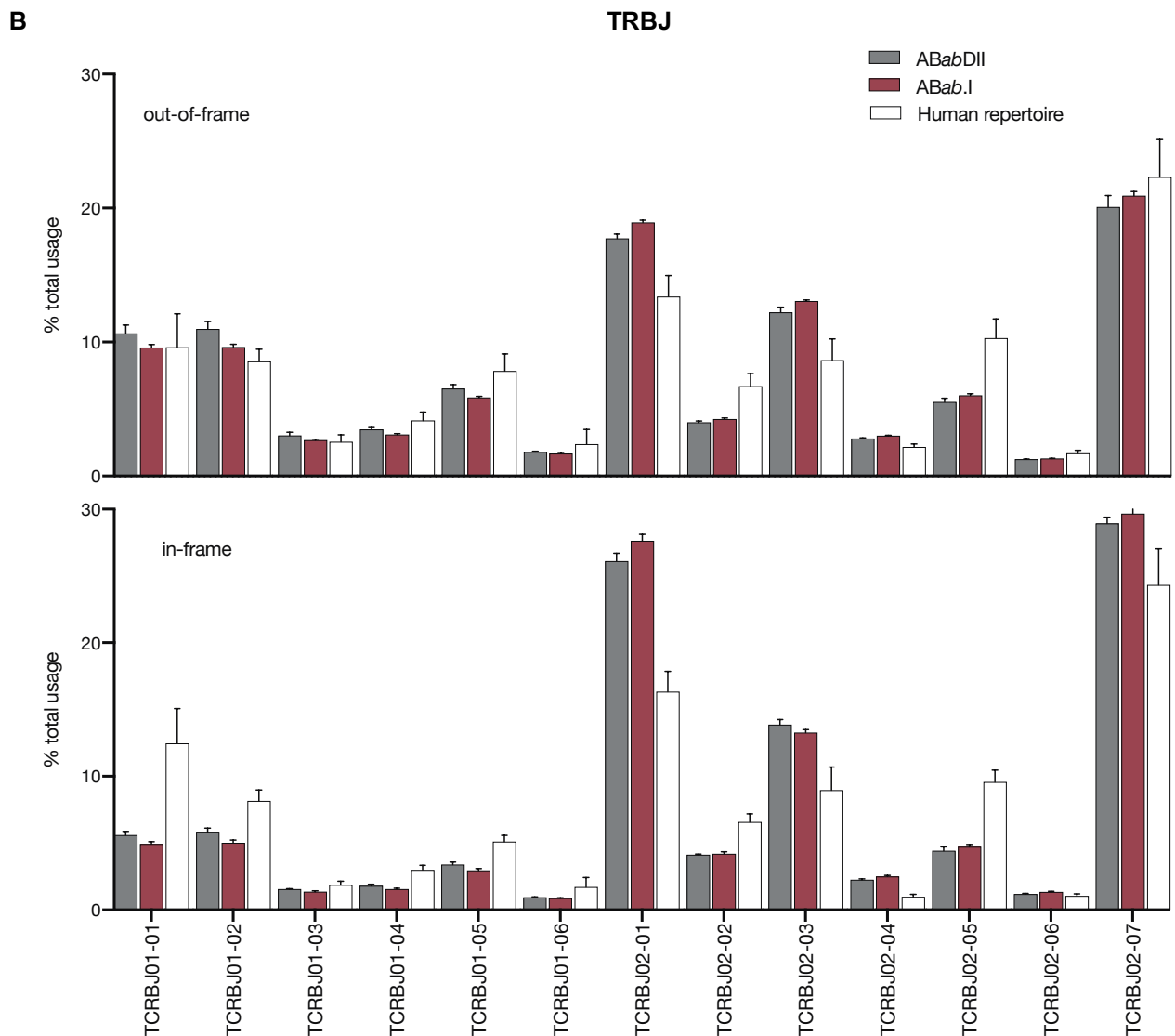
B

TRBJ
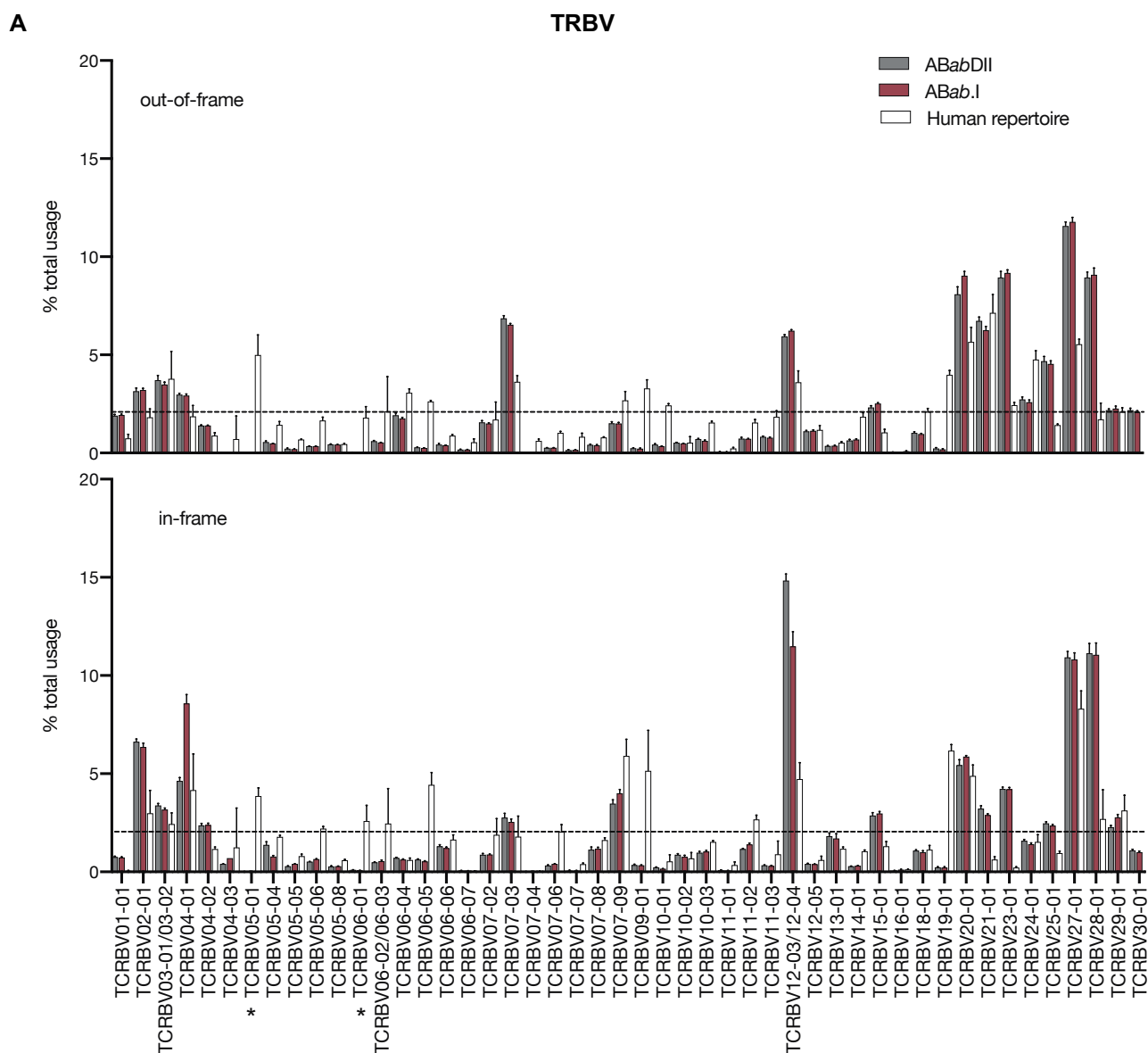
Figure 16: V_ß and J_ß gene usage frequencies of in-frame and out-of-frame TCRß clonotypes.

*A. and B. Frequencies of V_ß and J_ß gene usages of unique TCRß clonotypes in CD8⁺ T cells of ABabDII and ABab.I mice, and humans. **A. and B. Top,** Out-of-frame and **Bottom,** In-frame frequencies. Arrangement of V_ß gene (**A.**) or J_ß (**B.**) segments mentioned on the x-axis according to their position on the human chromosome from 5' to 3'. The dotted line in **A.** represents the frequency of random V_ß gene usage (TRBV, 2.1%). Bar graphs represent mean ± SD. Data are from, ABabDII (n = 5); ABab.I (n = 5) mice, and humans (n = 3). * Expression of V_ß genes is missing in both ABabDII and ABab.I mice.*

V_ß [Figure 16A] and J_ß [Figure 16B] gene usages were analyzed present in both out-of-frame and in-frame TCRß clonotypes. The former represents the preselection pool, which is an unbiased measure of the diversity of all possible V(D)J recombination events that could have occurred before positive selection. In AB*ab*DII mice, and so, thereby in AB*ab*.I mice, most of

the detected V$_\text{ß}$ segments were found to be rearranged, except for TRBV5-1 and TRBV6-1, which were previously reported for non-existence or lack of expression [Figure 16A]. AB*ab*DII and AB*ab*.I mice showed no difference in any preferred V$_\text{ß}$ usage in the preselection pool. Similarly, in humans, the out-of-frame V$_\text{ß}$ gene usage pattern was comparable without any selection preference.

Both in the pre- and post-selection pool, some V$_\text{ß}$ genes were either over or underrepresented in mice. Overrepresented V$_\text{ß}$ genes include TRBV20, TRBV21, TRBV23, TRBV27, and TRBV28. Underrepresented V$_\text{ß}$ genes include TRBV6-2/6-3, TRBV6-4, TRBV7-9, TRBV9-1, and TRBV10-1, mostly closer to the 5' end [Figure 16A, top]. Single nucleotide polymorphisms might be a reason for this difference in representation. Preferential overrepresentation of two V$_\text{ß}$ genes such as TRBV7-3 and TRBV12-3/12-4 in AB*ab*DII and AB*ab*.I mice was observed compared to humans [Figure 16A, top].

In-frame V$_\text{ß}$ genes represent the post-selection pool with similar frequencies, except some genes in the 5' region such as TRBV2-1 and TRBV4-1, and TRBV7-9 were highly selected-in by positive selection, whereas, V$_\text{ß}$ genes such as TRBV7-3, TRBV20/21/23/24, and TRBV25 were selected-out in the thymus [Figure 16A, bottom]. Only one V$_\text{ß}$ gene in each strain, V$_\text{ß}$4-1 in AB*ab*.I and V$_\text{ß}$12-3/12-4 in AB*ab*DII mice, seem to be predominantly represented in the post-selection pool in one or the other mice [Figure 16A, bottom].

J$_\text{ß}$ gene usages were also non-random and similar between the transgenic mice and humans [Figure 16B], although some J$_\text{ß}$ segments were used more frequently in mice than in humans: TRBJ1-2, TRBJ2-1, and TRBJ2-3 [Figure 16B, top]. Notably, J$_\text{ß}$2-07 in the 3' end was represented high with $\sim$ 30 % total usage in the post-selection pool in mice than in humans [Figure 16B, bottom].

3.9 AB*ab*.I mice enrich longer CDR3-containing T cell receptors

The V(D)J recombination event in the TCRß chain generates CDR3 diversity and is responsible for antigen recognition in conjugation with the VJ events in the TCRα chain. AB*ab*.I mice had enriched significantly longer CDR3ß regions than AB*ab*DII mice.

Considering shorter CDR3 lengths, in AB*ab*DII, 5.5% of clonotypes expressed TCRs with 33 bp, 11.5% with 36 bp, and 21% with 39 bp CDR3 sequences. In AB*ab*.I, lengths corresponded to 4.9% with 33 bp, 10% with 36 bp, and 19% with 39 bp CDR3 sequences. In AB*ab*.I mice, a higher percentage of clones expressed longer CDR3 sequences than AB*ab*DII mice, 23% versus 21% with 48 bp, 10% versus 8.3% with 51 bp, and 3.7% versus 3.4% with 54 bp [Figure 17A]. However, humans produced the lengthiest CDR3ß sequences compared to mice [Figure 17A]. On average, AB*ab*.I mice had TCRs with 42 ± 0.1 bp CDR3 lengths compared with 41.5 ± 0.06 bp CDR3 lengths in AB*ab*DII mice, but not longer than humans with an average CDR3 of 43.5 ± 0.18 bp [Figure 17B].
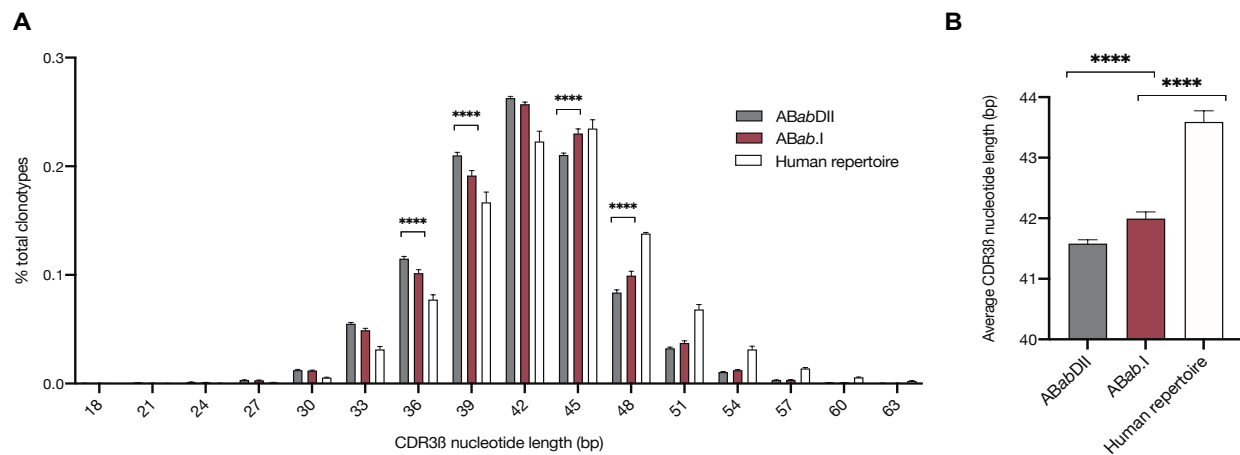
Figure 17: CDR3ß region analysis among ABabDII, ABab.I, and human donors.

*A. CDR3 length distribution of TCRß clonotypes. In-frame frequencies of different CDR3ß lengths for ABabDII and ABab.I mice, and humans are shown in the bar graph. **B.** Average CDR3ß lengths were compared between ABabDII and ABab.I mice, and humans. Bar graphs represent mean ± SD. Data are from, ABabDII (n = 5); ABab.I (n = 5) mice, and humans (n = 3). P values in **A. and B.** indicate: ****, P ≤ 0.0001 (unpaired two-tailed t-test).*

3.10 Natural human HLA haplotype in AB*ab*.I mice educates a wide array of unique clones

Shared clones were compared between mouse strains and among humans using Jaccard similarity index estimation. Jaccard index (J) evaluates the shared immune repertoire between samples. J index ranges from 0 to 1, and the score is calculated based on the formula: the total number of shared clones divided by the total number of unique clones across two samples, $J(A, B) = A \cap B / A \cup B$. High shared clonality within strains (AB*ab*.I with AB*ab*.I or AB*ab*DII with AB*ab*DII) were detected than across strains (AB*ab*.I with AB*ab*DII or vice versa). AB*ab*.I mice shared ∼ 6.4% of clones among each other, Jaccard index: 0.0636 ± 0.003, and AB*ab*DII mice shared ∼ 5% of clones within their group, Jaccard index: 0.0498 ± 0.004 [Figure 18A]. Humans shared more clones with AB*ab*.I mice (Jaccard index: 0.006 ± 0.001) than with AB*ab*DII mice (Jaccard index: 0.005 ± 0.001) or among each other with no statistically significant difference [Figure 18A].

Both AB*ab*DII and AB*ab*.I produced more shared TCRß clones among each other than they shared with humans, likely the possibility of similar genetic background with human TCR gene loci. Notably, AB*ab*.I mice shared fewer clones with AB*ab*DII, Jaccard index: 0.04 ± 0.003 than within their group, 0.0636 ± 0.003 [Figure 18A].

AB*ab*.I mice generated more unique clones compared with AB*ab*DII mice. Pooled analysis from 5 mice per strain determined the total number of shared and unique clones. AB*ab*DII and AB*ab*.I mice shared less than 100,000 clones (0.8×10^5) between each other. AB*ab*.I mice

produced almost 4-times more unique and rare clones (AB*ab*.I: 7.5×10^5) than ABabDII mice: 2×10^5 TCRß clones [Figure 18B].
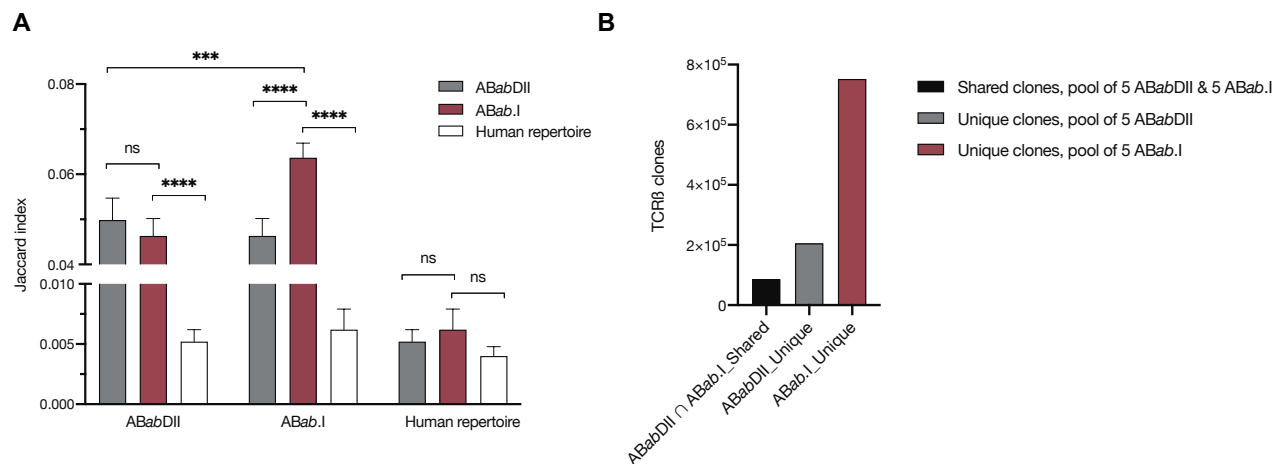

Figure 18: Shared repertoire analysis among ABabDII, ABab.I and human donors.
*A. Jaccard index represents a score based on the number of shared TCRß clones within and between groups of ABabDII and ABab.I mice, and human repertoire. Bar graph represents mean ± SD. Data are from, ABabDII (n = 5); ABab.I (n = 5) mice, and humans (n = 3). P values indicate: ****, P ≤ 0.0001; ***, P ≤ 0.001; ns, not significant (unpaired two-tailed t-test). B. Total numbers of unique TCRß clones in ABabDII and ABab.I from pooled data. Five mice per strain were pooled for analysis.*

3.11 HLA alleles in AB*ab*.I mice can efficiently present epitopes and trigger T cell responses *ex vivo*

To determine epitope presentation by HLA alleles in the newly generated AB*ab*.I mice, peripheral blood or lymphoid organ cells isolated from the mice *ex vivo* were co-cultured with TCR-transduced T cells recognizing model epitopes. Previously described (A3-, A11-, B7-, B15-, and C7-restricted), or in-house generated (C4-restricted) TCRs against a panel of model antigens acquired from literature resources[217–221] were used in co-culture assays [Figure 19A]. Co-culture supernatants were analyzed for murine IFN-γ to assess effector T cells for HLA allele recognition bound with its respective model epitope. As positive control, MCA205/AB*ab*.I cells expressing six HLA alleles as in AB*ab*.I mice were used as target cells to present model epitopes to transduced T cells. After 16 hours of co-culture, IFN-γ was measured in the supernatant to confirm the functional activity. All five model TCRs recognized the respective peptide HLA combination (pHLA-TCR) and produced IFN-γ in the range of 9437 ± 863 pg/ml - 11,877 ± 299 pg/ml [Figure 19B].

Peripheral blood and lymphoid organ cells used as target cells to measure pHLA-TCR interaction in AB*ab*.I mice *ex vivo*. Post-recognition of blood pHLAs, all TCR-transduced T cells produced IFN-γ, and the lowest release was 4168 ± 308 pg/ml, and the highest was 7873 ±

212 pg/ml [Figure 19C], proving functional activity. Similarly, pHLAs from secondary lymphoid organs were recognized (spleen, lymph nodes as target cells), all model TCR-transduced T cells released IFN-γ in the range of 5540 ± 216 pg/ml - 8054 ± 270 pg/ml [Figure 19C]. PMA/Ionomycin activated protein kinase C and opened calcium (Ca^{2+}) channels in transduced T cells for maximum IFN-γ release. Whereas, in untransduced T cells, no IFN-γ release (only low background) was measured [Figure 19A and 19B]. HLA-A2 (present in AB*ab*.I mice) restricted TCR[171], T1367-transduced T cells were used in co-cultures as an internal reference control in all presentation assays. A model TCR against CMV pp65 epitope[220] was isolated in-house from immunized HuTCR mice expressing HLA-C*04:01 as monochain [refer to discussion section 4.5]. All model epitopes [Figure 19A] used in co-culture assays were taken forward for *in vivo* immunizations into AB*ab*.I mice to analyze CD8[+] T cell responses by IFN-γ release using intracellular staining.

A

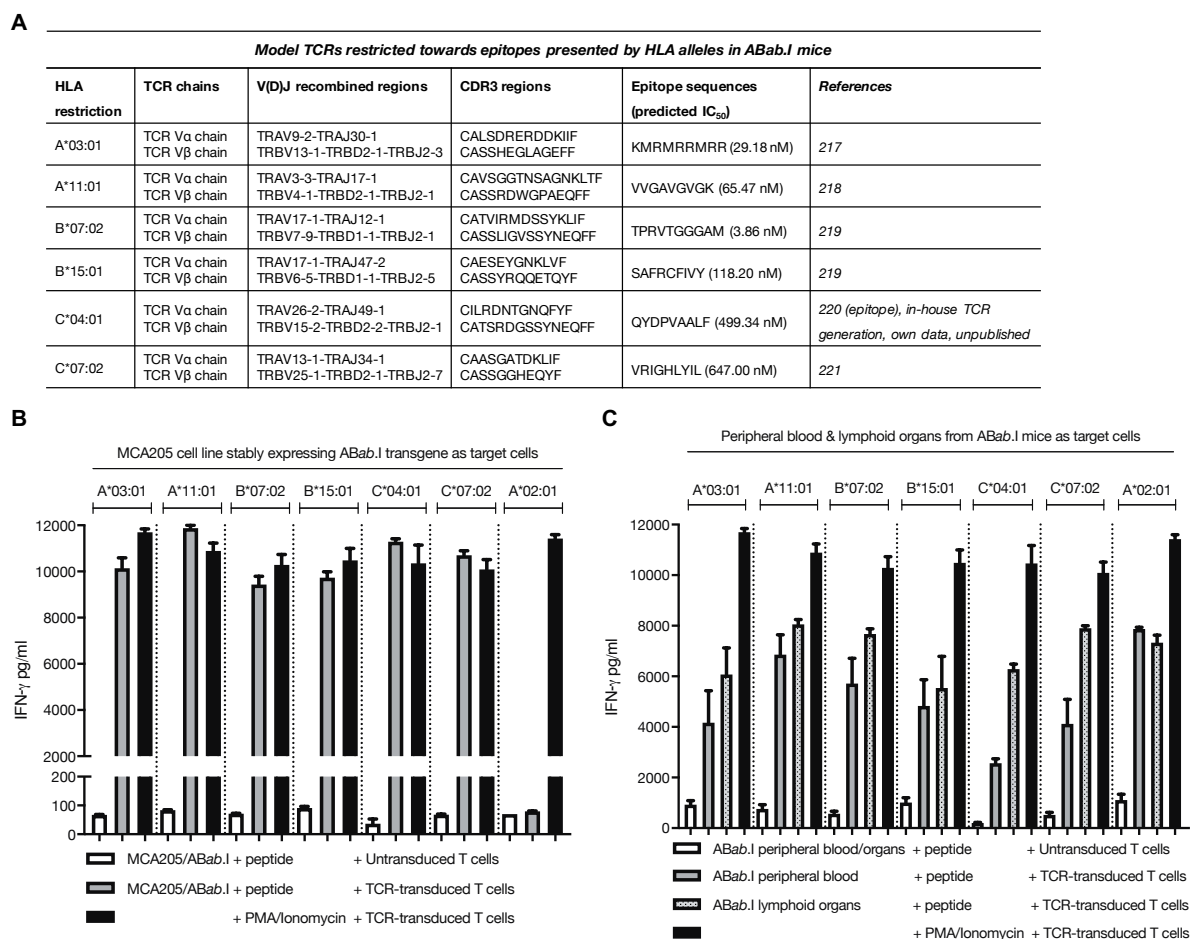|					
Model TCRs restricted towards epitopes presented by HLA alleles in ABab.I mice					
HLA restriction	**TCR chains**	**V(D)J recombined regions**	**CDR3 regions**	**Epitope sequences (predicted IC$_{50}$)**	*References*
A*03:01	TCR Vα chain TCR Vβ chain	TRAV9-2-TRAJ30-1 TRBV13-1-TRBD2-1-TRBJ2-3	CALSDRERDDKIIF CASSHEGLAGEFF	KMRMRRMRR (29.18 nM)	*217*
A*11:01	TCR Vα chain TCR Vβ chain	TRAV3-3-TRAJ17-1 TRBV4-1-TRBD2-1-TRBJ2-1	CAVSGGTNSAGNKLTF CASSRDWGPAEQFF	VVGAVGVGK (65.47 nM)	*218*
B*07:02	TCR Vα chain TCR Vβ chain	TRAV17-1-TRAJ12-1 TRBV7-9-TRBD1-1-TRBJ2-1	CATVIRMDSSYKLIF CASSLIGVSSYNEQFF	TPRVTGGGAM (3.86 nM)	*219*
B*15:01	TCR Vα chain TCR Vβ chain	TRAV17-1-TRAJ47-2 TRBV6-5-TRBD1-1-TRBJ2-5	CAESEYGNKLVF CASSYRQQETQYF	SAFRCFIVY (118.20 nM)	*219*
C*04:01	TCR Vα chain TCR Vβ chain	TRAV26-2-TRAJ49-1 TRBV15-2-TRBD2-2-TRBJ2-1	CILRDNTGNQFYF CATSRDGSSYNEQFF	QYDPVAALF (499.34 nM)	*220 (epitope), in-house TCR generation, own data, unpublished*
C*07:02	TCR Vα chain TCR Vβ chain	TRAV13-1-TRAJ34-1 TRBV25-1-TRBD2-1-TRBJ2-7	CAASGATDKLIF CASSGGHEQYF	VRIGHLYIL (647.00 nM)	*221*

Figure 19: Functional characterization of HLA alleles present in ABab.I mice.
*A. List of model TCRs restricted towards HLA alleles in ABab.I with known peptide binders. In addition to this list, HLA-A2-restricted T1367 MAGE-A1 TCR[171] was used as an internal reference in all presentation assays. **B. and C.** Functional activity of ABab.I HLA alleles through*

*surface expression and model peptide presentation on, **B.** MCA205 cell line stably expressing six HLA alleles as in ABab.I mice, except HLA-A2 or **C.** Ex vivo-isolated peripheral blood or cells from lymphoid organs, co-cultured with model TCR-transduced or untransduced T cells. Cytokine levels were measured after overnight co-culture by mouse IFN-γ ELISA assay. pHLA-TCR pairs are mentioned in **A.** Bar graphs represent mean of intra-assay duplicates ± SD. One representation out of multiple co-culture experiments is given (n > 5).*

3.12 AB*ab*.I mice mount immune responses against peptide antigens *in vivo*
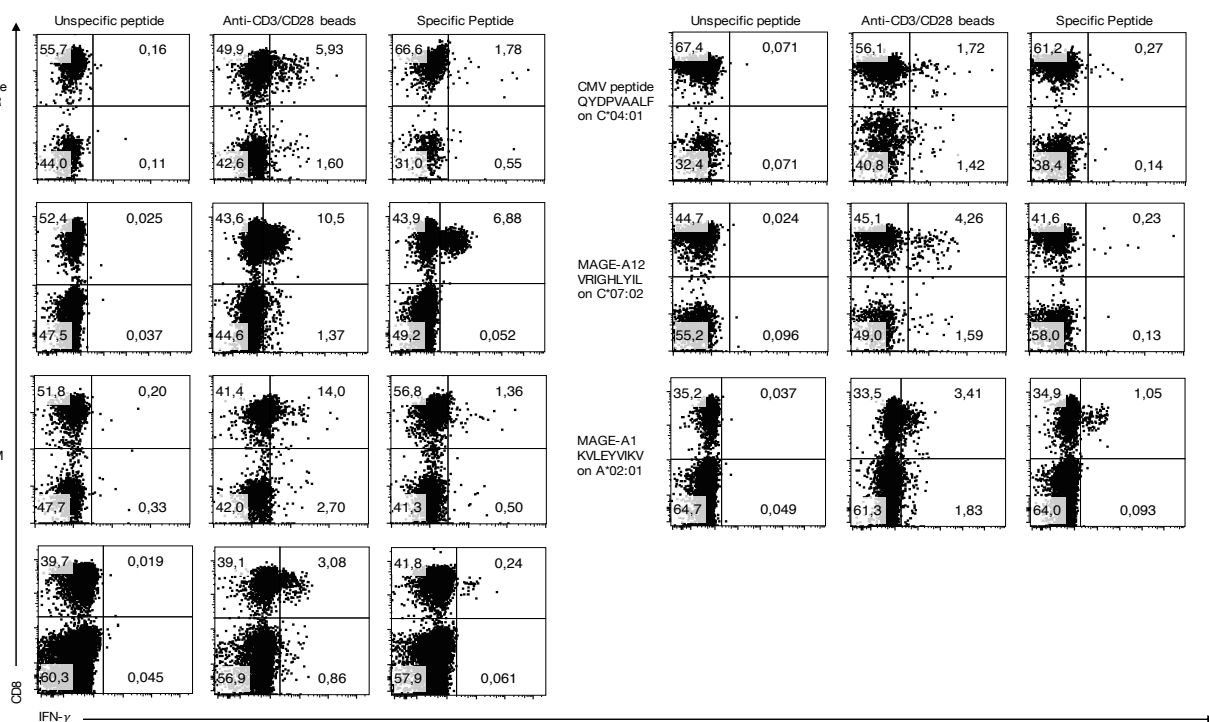
Figure 20: CD8⁺ T cell responses against a panel of human antigens in ABab.I mice.

*Specific CD8⁺ T cell responses in ABab.I mice was observed after immunization with a panel of model antigens (mentioned along x-axis). Young ABab.I mice were immunized (minimum twice) with an interval of at least 4 weeks between peptide injections. Seven days post-injection, peripheral blood cells were stimulated in vitro with the same specific peptides used for immunization (**right**), or unspecific peptides (**left**), or CD3/CD28 beads (**middle**, positive control). Cells were analyzed for CD3, CD8, and intracellular IFN-γ expression by flow cytometry. Dot plots depict CD8⁺ IFN-γ⁺ T cells (gated on CD3⁺ lymphocytes). Unspecific NY-ESO_{APRGPHGGAASGL} peptide was used only for in vitro stimulation. One representative flow cytometry plot out of multiple responders is given (out of 6-7 mice/antigen, 6 mice responded against KRAS, 3 mice against MAGE-A1 and HPV, 1 mouse responded against mCALR, CMV, and MAGE-A12 peptides).*

We analyzed AB*ab*.I mice for antigen-specific CD8[+] T cell responses against described model antigens [Figure 19A] using peptide immunization [Figure 20]. All six model antigens were *in silico* predicted to be strong-binders to their respective HLA alleles. Intracellular staining measured ~ 1.8% of CD8[+] T cells produced IFN-γ against the mutated calreticulin (mCALR) protein-derived epitope, KMRMRRMRR. ~ 6.9% of CD8[+] T cells released IFN-γ against the mutated KRAS protein-derived G12V epitope, VVGAVGVGK. ~ 1.4% against B*07:02 bound CMV epitope, TPRVTGGGAM, ~ 0.2% against B*15:01 bound HPV epitope, SAFRCFIVY, ~ 0.3% against C*04:01 bound CMV epitope, QYDPVAALF, and ~ 0.23% against MAGE-A12 epitope, VRIGHLYIL [Figure 20]. The MAGE-A1 epitope, KVLEYVIKV bound HLA-A2 elicited ~ 1% of CD8[+] T cell response. The unspecific peptide used in *in vitro* restimulations was a NY-ESO epitope, APRGPHGGAASGL. No T cell responses were observed in any case against the unspecific peptide proving a specific response towards the immunized peptide. CD8[+] T cell responses against anti-CD3/CD28 beads were in the range of 1.7% - 14% [Figure 20]. Thus, CD8[+] T cell responses towards a panel of described epitopes in the AB*ab*.I mouse model depicts *in vivo* functionality of the human HLA haplotype.

4. Discussion

In this doctoral study, we generated the 'AB*ab*.I transgenic mouse model' using the PiggyBac transposon strategy with six HLA class I alleles on the 'humanized TCR' mouse background[157]. We confirmed the human HLA haplotype in mice by genotyping the six HLAs using allele-specific PCRs. We confirmed the phenotype by staining with pan- and HLA-specific antibodies, as much as available, using flow cytometry analysis. We provided evidence supporting the increased thymic output in AB*ab*.I mice by profiling CD8$^+$ and CD4$^+$ T cells in the periphery. We characterized the immune repertoire using ImmunoSEQ™ deep sequencing and observed that AB*ab*.I mice have a four-fold broader T cell receptor repertoire than the AB*ab*DII mice. We proved the antigen presentation efficiency of all six HLA alleles present in AB*ab*.I mice using IFN-γ release assays by the effector CD8$^+$ T cells *ex vivo* and *in vivo*.

4.1 Computational neoantigen prediction and the need for epitope immunogenicity validation - mutant neoepitope S722F is not endogenously processed

We predicted 175 class I mutant epitopes out of 266 mutations to bind 18 highly ranked (population-wise) HLA alleles. All 266 mutations were somatic point mutations that occurred recurrently more than twice in different cancer entities. We focused only on recurrent driver mutations as they are present in every cancer cell throughout the disease[123,125] and yield truly tumor-specific antigens[90,154]. Earlier evidence compared chemotherapy by a drug inactivating the cancer-driving oncogene with tumor-specific CD8$^+$ ATT treatment and proved relapse-free tumor elimination happens by eradicating escape variants with tumor-stroma destruction[150]. Based on this, we reasoned that in our study, by raising TCRs against recurrent neoantigens derived from driver mutations, complete eradication of tumors without antigen loss and relapse might be possible.

In our *in silico* screen, we marked epitopes as strong-binders with an IC$_{50}$ < 50 nM criterion. pMHC affinity ranges between IC$_{50}$ <1 nM to >20,000 nM. Engels *et al*. investigated pMHC affinities with therapy outcome on tumor rejection versus relapse. The study reported that strong pMHC affinities with IC$_{50}$ <10 nM caused tumor rejection in every case and relapsed if IC$_{50}$ > 100 nM[222]. Following, two reviews highlighted the importance of selecting high-affinity (IC$_{50}$ <50 nM) class I epitopes for CD8$^+$ ATT treatments in the context of a tripartite interaction between peptide-bound-MHC and TCR[90,223]. Based on these studies, in our prediction screen, we selected mutant neoepitopes with an IC$_{50}$ of less than 50 nM to raise TCRs against these recurrent antigens using AB*ab*DII[157] and AB*ab*.I (generated as a part of this doctoral thesis).

Not all predicted neoantigens yield immunogenic epitopes. Epitope immunogenicity combines natural endogenous processing for pMHC presentation with the ability to trigger CD8$^+$ T cell responses[224]. *In silico* defined epitopes must be evaluated with caution as they could generate false positives[154]. In our study, we predicted TRRAP-S722F driver mutation to contain a

putative neoantigen, a 9-mer epitope. The S722F mutant epitope triggered an immune response in AB*ab*DII mice. TCRs isolated from mice were specific for S722F and recognized low amounts of the exogenously loaded peptide. However, they did not react to the melanoma cell line that naturally expressed the mutated gene confirming that the 9-mer epitope is not endogenously processed and presented. By this reverse immunology approach, we proved that the S722F-predicted epitope is not endogenously processed. In addition to the TRRAP mutant (S722F), predicted neoepitopes from genes, C-KIT (K642E), XPO1 (E571K), and FOXA1 (D226N) did not elicit CD8[+] T cell responses in AB*ab*DII mice [own data not shown, unpublished].

Several other findings confirmed our observation that reverse immunology (first predicting epitopes and at last investigating endogenous processing) is not the best way to choose target antigens for adoptive T cell therapy. In an epitope immunogenicity study of the vaccinia virus, $\sim$ 100 epitopes showed good predicted binding affinities to HLA-A2 with an IC_{50} of < 100 nM. Only 50% of those elicited CD8[+] T cell responses after HLA-A2 transgenic mice immunization. 15% of response-eliciting peptides further got processed by the proteasome. Out of this, 11% elicited spontaneous CD8[+] responses upon vaccinia virus infection, proving that only 1/14 of the predicted fraction created an immunogenic epitope[126]. In another melanoma RNA vaccination trial, CD8[+] T cell responses were relatively low compared with CD4[+] T cell responses, which might be due to more promiscuous class II peptide presentation by the MHC[225]. In yet another long peptide melanoma vaccination trial, only 16% of the predicted neoepitopes elicited CD8[+] T cell responses[226].

Software programs do not properly predict endogenous epitope modifications in the antigen presentation pathway using their algorithms. These alterations inside proteasomal-, post-proteasomal-, and peptide transport-compartments are decisive factors to determine epitopes as immunogenic and targetable. One study showed proteasomal epitope destruction of a proposed epitope[227]. Another example reported that ERAP-trimmed epitopes managed easier escape than not-trimmed epitopes[228]. In a third study, although the human individual showed CD8[+] T cell responses against the CDK4 mutant antigen, it was later verified in an experimental cancer model to be a relatively poor target for T cell therapy[229].

Thus, artificial network algorithms[175–177] cannot accurately predict proteasomal processing, peptide transport, N-terminal trimming by ERAP enzymes, or splicing mechanisms to define human neoantigens for ATT of cancers.

4.2 Polycistronic transgene design requires mRNA stability and protein misfolding considerations - lessons learned

Polycistronic primary AB*ab*.I transgene segregated by viral 2A linker elements showed no surface protein expression from any of the six HLA alleles. So, we manipulated the primary

construct driven by the H-2 D^b promoter with three polyadenylation signals (short D^b, full-length D^b, and bGH pA). Simultaneously, we replaced the promoter sequence in these three constructs with the viral CMV promoter. In all of these six constructs, though we quantified significant amounts of HLA mRNA by qRT-PCR in H-2 D^b primary constructs, 30x more in CMV-primary constructs, we did not detect cell surface protein expression by flow cytometry. Multiple reasons could have caused this untranslatable situation of HLA mRNAs affecting its cell surface presentation. Such considerations could be mRNA instability of the long HLA transcript, inefficient excision of viral 2A linkers, and unfavorable post-translational modifications causing misfolding of nascent polypeptide chains.

We speculate that mRNA stability was affected by a single common 3'-UTR region regulating all six HLA alleles[230]. Moreover, recent evidence confirmed that native 2A linkers functioned inefficiently by not self-cleaving themselves to release proteins[231]; of note, our primary transgene contained native 2A peptide linkers. A study on sequence divergence versus homology reported that transient misfolding does occur in proteins derived from polycistronic mRNAs, whereas they disappear in the case of low sequence identity[232]; HLA class I chimeric chains maintain 90-98% homology[233]. In line with this, a recent review postulated that open questions exist on how nascent polypeptides from polycistrons fold without forming aggregates for cellular clearance[234].

Based on these factors, we decided to clone and express all six HLA alleles as monocistrons driven by their own 5'- and 3' regulatory elements.

4.3 Protein characterization of chimeric HLAs in AB*ab*.I mice and the need for monochain-specific antibodies - human lymphoblastoid cell lines express high levels of HLA proteins

Single HLA alleles were cloned from the primary polycistronic transgene and recloned as monocistrons for a successful surface expression. To confirm cell surface expression of HLAs, always a pan antibody or an antibody against the human ß2m chain present in all chimeric alleles were stained. There are no specific antibodies that exist against the unique HLA alleles in the AB*ab*.I mice, which can be used for flow cytometry analysis, except for HLA-A*03:01 and HLA-B*07:02, which makes validation of HLA proteins difficult.

Notably, both pan HLA and ß2m antibodies stained different HLA chains with different intensities, especially the ß2m antibody stained HLA genes in lymphoblastoid cell lines (LCL) with high fluorescence intensity profile (MFI) than the chimeric alleles in AB*ab*.I mice. Notably, the HLA genes in LCL lines are fully human[235]; pHLA stabilization happens by human proteome-derived antigens. HLA alleles in AB*ab*.I mice are chimeric; pHLA stabilization occurs here by mouse proteosome-derived antigens. This difference in the expression pattern raised the question, whether it was the consequence of peptides on the HLA stability that make the

ß2m antibody stain the fully human HLAs better with an increased shift in the mean fluorescence intensity.

Interestingly, a study using HLA antibodies from multiparous women confirmed the impact of peptides on HLA reactivity by antibodies[236]. Based on this, we speculated, the more stable the pHLAs are on the surface, the more they can be detectable through antibody staining. Thus, we could relate alloantibody staining differences to our ß2m staining difference between LCLs and AB*ab*.I mice. Another possible explanation could be that the chimeric HLAs in AB*ab*.I mice fold differently in an unnatural fashion, which could obstruct epitope accessibility by the ß2m antibody, which requires further analysis.

On a different note, a fully human HLA expression in transgenic mice without ß2m linked as a fusion protein could enhance surface stability, thereby improving detection. However, the maintenance of mouse CD8α interaction with the α3 domain of the chimeric mouse-human HLA heavy chain is crucial for CD8+ T cell co-activation[237–239].

4.4 Generation of AB*ab*.I transgenic mice using pronuclei injection technology

We produced two versions of AB*ab*.I transgene-targeting vectors for injections into AB*ab*DII donor oocytes. The first vector flanked the Rosa26 guide RNA sequences for CRISPR-Cas9 recognition, and the second vector flanked inverted terminal repeat sequences for hyperactive PiggyBac recognition on both 5' and 3' ends, respectively.

Our exploratory attempts of using CRISPR-Cas9 technology through homology-independent targeted integration (HITI)[240] failed even after several trials. Injections for random integration with vector backbone-free naked DNA transgene were not successful as well. For HITI with CRISPR-Cas9, we reasoned that a better strategy could have been to use homologous arms flanking the transgene cassette for efficient integration into the Rosa26 locus. Until now, the HITI method showed high integration efficiency only in the post-mitotic neurons but not in oocytes[240], so it requires in-depth optimization to make a successful attempt in mouse embryos. DNA strand breaks in the mouse genome during pronuclei microinjection procedures are the primary cause of random transgene integration (multiple copies) in a head-to-tail fashion. During chromosomal end-joining, the break-point ends take up the available DNA transgene as donor template to ligate the junctions[241]. We speculate that the large size of the AB*ab*.I transgene with homologous sequences (~ 17 kb) could have been the reason for failed naked transgene integration as concatemers. The use of the PiggyBac transposon strategy might enable us to target one copy per locus[184].

Hence, simultaneously we utilized hyperactive PiggyBac transposase to catalyze AB*ab*.I transgene cassette with ITRs for integration into the mouse genome[184,185]. We successfully produced five founders, Q4115, Q4118, Q6844, Q8552, and Q8556. The first founder, Q4115, did not germline transmit the HLA alleles in its genome to F1 generation in nearly 250

genotyped mice. We believe that PiggyBac catalyzed the transgene for integration into the genome after multiple cell divisions of the embryo or late integration, which resulted in a mosaic founder mouse not being able to germline transmit to the F1 generation as per Mendelian genetics[242]. Nonetheless, we managed to achieve germline transmission in all the other founders, Q4118, Q6844, Q8552, and Q8556.

We investigated the phenotype of different founder mice based on the absolute numbers of CD8[+] and CD4[+] T cells in the blood and the secondary lymphoid organs [data not shown for other founders except Q6844 line]. The Q6844 lineage showed the highest and consistent CD8[+] T cell numbers across generations from F1 until F4 providing indirect evidence for thymic and likely peripheral HLA expression. Significant HLA expression is necessary for positive selection in the thymus and T cell maintenance in the periphery. Hence, we hypothesized that the high CD8[+] T cell numbers (Q6844 lineage) reflect efficient positive thymic selection and increased TCR repertoire diversity. Besides, the impact on the homeostatic proliferation of T cells is another interesting question to address. For this, it is crucial to consider CD44 staining as an activation marker[243] to understand homeostasis in the periphery based on thymic output and the number of precursor T cells that come into circulation[244].

However, using next-generation ImmunoSEQ repertoire sequencing, we managed to deep sequence the TCRß immune repertoire of the CD8[+] T cells of AB*ab*.I in comparison with AB*ab*DII mice.

4.5 Next-generation deep sequencing of AB*ab*.I and AB*ab*DII mice - CD8[+] TCRß immunosequencing

Having observed the differences in the absolute CD8[+] T cell numbers between AB*ab*.I and AB*ab*DII mice, we deep sequenced the TCRß repertoire in these strains to understand whether the increase in CD8[+] T cell numbers resulted in a broader TCR repertoire. Being this the case, we could exclude the possibility that the increase in CD8[+] T cell numbers in AB*ab*.I compared to AB*ab*DII mice is a result of increased homeostatic expansion[245]. Of note, AB*ab*DII mice bear a single HLA-A2[157], whereas AB*ab*.I mice have additionally a complete human HLA haplotype of six alleles. Thus, we investigated whether or not the introduction of multiple class I alleles into the humanized TCR mice made the repertoire broader and also how broad it was compared with AB*ab*DII mice.

We proved using iChao1 estimation that AB*ab*.I mice have a diversified repertoire and on average, four times broader than that of AB*ab*DII mice. While we deep-sequenced approximately two-fold more numbers of in-frame amino acid clonotypes in AB*ab*.I mice versus AB*ab*DII mice ($11.2 \pm 2.7 \times 10^4$ versus $4.8 \pm 0.80 \times 10^4$), the iChao1 estimator theoretically extrapolated the repertoire to be four times broader; back-calculated to a complete mouse repertoire by including clones that occurred only once or twice (lower bound events)[216]. By

theoretical extrapolations, we showed that AB*ab*.I mice used diverse V and J genes and produced close to 1×10^6 clonotypes compared with AB*ab*DII, which only generated a maximum of 2.57×10^5 clonotypes (~ four times lower). More diversity in AB*ab*.I mice reflected higher percentages of rare and unique clones and that the thymic selection indeed selected novel V(D)J TCRß combinations on multiple HLA alleles. Thus, the expression of six HLA alleles in AB*ab*.I mice selected a highly diverse human TCR repertoire, likely because of enhanced positive selection and overall thymic output, which is reduced in AB*ab*DII mice[157].

The out-of-frame TCR repertoire represents the preselection pool before positive selection and gives a precise estimate, how frequently individual V and J segments are rearranged[246] in AB*ab*DII and AB*ab*.I mice. As expected, the preselection (out-of-frame) TCR repertoire was nearly identical between AB*ab*DII and AB*ab*.I mice, because the human TCR gene loci and the recombination enzymes were the same. The post-positive selection (in-frame) repertoire reflected, on one hand, the frequency by which individual V and J segments were rearranged, and on the other hand, revealed that some $V_ß$ gene segments were overrepresented in AB*ab*.I mice, for example, TRBV4-1, whereas those were underrepresented in AB*ab*DII mice. Similarly, TRBV12-3/12-4 is overrepresented in AB*ab*DII, but not in AB*ab*.I mice.

We have made an apparent connection between multiple HLA expression and their impact on repertoire diversity. Looking from the TCR side, we speculate that some $V_ß$ genes could have a higher intrinsic affinity towards some but not other HLA alleles. A critical requirement for positive selection is the interaction between the CDR1 and CDR2 regions with the selecting MHC class I molecule[30,32]. The 40-50 $V_ß$ genes differ in their CDR1 and CDR2 regions and the MHC class I genes are highly polymorphic. Therefore, one can assume that the intrinsic affinity for some $V_ß$ gene segments is higher for some MHC I alleles but lower for others and that the MHC and TCR gene loci co-evolved to ensure more efficient positive selection[32]. This hypothesis was supported by Chen *et al*. They sequenced the human CD4$^+$ TCR repertoire in mice selected on either a single mouse or a single human MHC class II molecule. They found that the species-compatible human MHC class II selected a broader CD4$^+$ TCR repertoire compared to the species-incompatible mouse MHC molecule I-A^b. This was explained by, on average, the increased inherent affinity of the human TCRs for the human MHC class II[21].

In analogy, one might assume that some $V_ß$ gene segments have a higher inherent affinity to HLA-A2, whereas other $V_ß$ gene segments have a lower intrinsic affinity to HLA-A2, but in turn a higher inherent affinity for any of the other six MHC class I alleles in AB*ab*.I mice. This then would mean, that certain TCRs can only poorly be selected by HLA-A2, but are then more efficiently selected by the other MHC I alleles in AB*ab*.I mice.

Chen *et al*. also observed that TCRs selected on mouse MHC II had, on average, shorter CDR3 sequences compared to human TCRs that were selected on human MHC II molecules[21]. The hypothesis is that different inherent affinities of $V_ß$ gene segments for MHC were adjusted

by CDR3 lengths, i.e. a particular V$_\beta$ gene segment with a relatively low inherent affinity for MHC gets better selected if it carries a shorter CDR3. We extend these findings, because a TCR repertoire selected on multiple class I alleles contained, on average, slightly larger CDR3 regions compared to the TCR repertoire selected only on a single MHC class I allele, in our case HLA-A2. The best explanation is that some V$_\beta$ segments have an increased inherent affinity for any of the six novel MHC class I alleles in ABab.I mice, which allows for the selection of TCRs with a slightly larger CDR3 region. This not only explained the diversity in AB*ab*.I compared with AB*ab*DII mice, but could also result in TCRs with higher specificity. This assumption is based on studies in mice that were deficient in terminal deoxynucleotidyl transferase (TdT) expression, which resulted in TCRs with short CDR3s. These TCRs tended to be more promiscuous and cross-reactive[247]. Despite the fact that the TCR repertoire in AB*ab*DII is relatively diverse, we speculate that the TCR repertoire in AB*ab*DII mice is not optimal because the CDR3 regions are relatively short[21]. In contrast, thymocytes in AB*ab*.I mice can choose one out of six additional class I alleles for getting selected. Therefore, we assume that different MHC alleles in the thymus compete with each other for selecting a given T cell clone. If this is correct, T cell clones that are not being able to be selected by HLA-A2 could be selected by any of the other six MHC class I alleles for which they have an increased inherent affinity and then also allowing longer CDR3s being accepted by not impeding positive selection.

It should be noted, however, that the hypothesis that the T cells in AB*ab*.I mice find its HLA fit with optimal intrinsic affinity at a price of reduced selection on HLA-A2 remains speculative so far. This hypothesis is difficult to prove because it is also not known whether different MHC class I alleles can select similar diverse TCR repertoire.

Therefore, we are in the process of generating and characterizing mice with single HLA alleles of those that are contained in AB*ab*.I mice. By TCR repertoire sequencing of mice expressing single HLA alleles on the same background, we can verify whether or not certain V$_\beta$ segments have a different inherent affinity for different HLA alleles[21,248]. For this, single AB*ab*-A*03, AB*ab*-A*11, AB*ab*-B*07, AB*ab*-B*15, AB*ab*-C*04, and AB*ab*-C*07 founder lines were generated and require characterization before deep sequencing [own data not shown, unpublished]. In these mice, we can prove or disprove whether certain V$_\beta$ gene segments are preferentially selected or preferentially ignored by individual MHC I alleles.

However, mice, in general, select shorter CDR3 TCRs than humans[21,249]. The reason for this could be higher TdT and exonuclease activity in humans compared to mice or a longer time window of expression. The TdT-exonuclease machinery with the recombination activating genes (RAG) evolved to magnify the addition or deletion of nucleotides (more editing); such increased activity generates higher numbers of unique TCRs with lengthier CDR3s impacting the overall repertoire[250].

Because AB*ab*.I mice generated four times more unique clones than AB*ab*DII, they also shared more TCRß clonotypes (around 1000 compared to around 500) with the repertoire of human donors. The numbers of shared clones between AB*ab*.I mice are surprisingly high, as they were not HLA-matched. The high clonal overlap is likely to be explained by the non-random V and J usage. In humans, the theoretical number of TCRß sequences is 5×10^{11}. If V-J rearrangement was random, less than five shared TCRß clones would be expected between any two individuals, but in fact, around 10,000 shared clones have been detected, consistent with the detection of only 0.1% of the theoretical 5×10^{11} in human peripheral blood[251]. Our data in AB*ab*.I mice support the data in humans and suggest TCRα/TCRß combinatorial diversity is larger than previously anticipated[252]. We also deep-sequenced the TCRα repertoire from these two mouse groups. Preliminary data analysis showed that the TCRα repertoire is also more diverse in AB*ab*.I compared to AB*ab*DII mice [own data not shown, unpublished]. Detailed analysis is still required to reveal combinatorial diversity between these mice.

In summary, a more diverse repertoire, closer to humans, might be achieved through extended humanization of the human TCR gene loci mouse model.

4.6 Functional characterization of AB*ab*.I mice *ex vivo and in vivo*

We showed that *ex vivo* derived APCs from AB*ab*.I mice can efficiently present model epitopes on all six HLA alleles and triggered effector T cells to produce IFN-γ. Due to the lack of HLA-specific antibodies, we devised this *in vitro* co-culture assay to estimate both HLA expression and its functionality by activating model-epitope-restricted TCR-modified T cells. APCs from cells of blood, spleen, and lymph nodes were used as targets to effectors, as they recapitulate the physiological circulating immune system when it comes to immunization[1].

The purpose of generating AB*ab*.I mice was to isolate T cell receptors to treat cancer patients targeting tumor-restricted antigens. Here, we show not only *ex vivo* but *in vivo* by immunizing AB*ab*.I mice with a wide range of model antigens that each were known to bind to one of the six HLA alleles, represented in the mice. These epitopes were selected to cover different antigen categories such as a frame-shift derived mutant (mCALR on A*03), point mutant (KRAS-G12V on A*11), viral (CMV-pp65 on B*07 and C*04, HPV-E5 on B*15), and tumor-associated (MAGE-A12 on HLA-C*07:02) epitopes.

Although most epitopes selected for the panel were proven to be endogenously processed, mice responded stronger to mutant KRAS and viral epitopes than others, irrespective of the predicted IC_{50} nM values. It is not a concern for processed epitopes, but for epitopes not proven to be processed, DNA or Adenovirus (encoding antigens) immunization expressing complete protein sequence can leverage the AB*ab*.I mouse model as a platform to discover novel epitopes.

In conclusion, the AB*ab*.I mouse model expressing six HLA alleles as like in a natural human situation can become an incomparable novel tool to isolate novel T cell receptors to be used clinically in ATT of cancer.

5. Outlook

5.1 Understanding the biology behind a full human HLA haplotype in a mouse model

The AB*ab*.I model with its diverse TCRß repertoire has proven to be a valuable tool to isolate unique T cell receptors for therapy. However, understanding the TCRα locus repertoire might determine the combinatorial diversity of the overall combination of TCRs that AB*ab*.I mice could produce. By this, we can also study how comparable is the TCR diversity in AB*ab*.I mice to the theoretically possible combinations of TCR diversity ($\sim 10^{15}$ clonotypes).

Different intrinsic $V_ß$ affinity towards some but not other human HLA alleles could be the next possible step to understand skewness during positive selection in the thymus. For this, comparison of deep sequencing data from AB*ab*.I mice with single AB*ab*- animals on humanized TCR background, AB*ab*-A*03, AB*ab*-A*11, AB*ab*-B*07, AB*ab*-B*15, AB*ab*-C*04, and AB*ab*-C*07 mice will allow us to clearly understand if there is a bias in TCR affinity towards HLAs and whether different HLA alleles can select similar or dissimilar TCR repertoires.

Knocking-out HLA-A2 should be the next event from single AB*ab*- and AB*ab*.I mice to generate pure lines devoid of the A2 gene; the ideal representation of a normal HLA class I loci in humans.

5.2 Epitope discovery of non-HLA-A*02:01-restricted antigens using AB*ab*.I mice

Reverse immunology based on epitope prediction is a concern. Thus, AB*ab*.I mice provide an opportunity to identify immunogenic epitopes and isolate T cell receptors simultaneously. Direct immunology by immunizing AB*ab*.I mice with DNA, adenovirus, or mRNA encoding the complete protein for epitope discovery would yield TCRs restricted to such newly discovered epitopes for ATT.

5.3 AB*ab*.I mice as a novel tool to isolate human TCRs for ATT of cancer

The prediction screen pointed out the number of individuals per year with the mutation. In principle, AB*ab*.I mice immunized with such mutant epitopes will yield a range of high- to optimal affinity TCRs to be taken forward to clinical applications, however, information on endogenous is necessary.

Hence, AB*ab*.I mice are not entitled only to understand human repertoire shaped on a natural class I HLA environment but also a novel model to isolate T cell receptors against a wide variety of antigens for adoptive T cell therapy.

6. References

1.   Paul, W. E. *Fundamental immunology*. Wolters Kluwer : Lippincott Williams and Wilkins (2013).

2.   Trinchieri, G. & Sher, A. Cooperation of Toll-like receptor signals in innate immune defence. *Nat. Rev. Immunol.* **7**, 179–190 (2007).

3.   Akira, S., Uematsu, S. & Takeuchi, O. Pathogen Recognition and Innate Immunity. *Cell* **124**, 783–801 (2006).

4.   Blander, J. M. & Medzhitov, R. Toll-dependent selection of microbial antigens for presentation by dendritic cells. *Nature* **440**, 808–812 (2006).

5.   Marshall, J. S., Warrington, R., Watson, W. & Kim, H. L. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* **14**, 49 (2018).

6.   Zhong, B., Tien, P. & Shu, H.-B. Innate immune responses: Crosstalk of signaling and regulation of gene transcription. *Virology* **352**, 14–21 (2006).

7.   Lacy, P. & Stow, J. L. Cytokine release from innate immune cells: association with diverse membrane trafficking pathways. *Blood* **118**, 9–18 (2011).

8.   Iwasaki, A. & Medzhitov, R. Regulation of Adaptive Immunity by the Innate Immune System. *Science* **327**, 291–295 (2010).

9.   Adler, L. N. *et al.* The Other Function: Class II-Restricted Antigen Presentation by B Cells. *Front. Immunol.* **8**, 319 (2017).

10.   Jiang, W., Adler, L. N., Macmillan, H. & Mellins, E. D. Synergy between B cell receptor/antigen uptake and MHCII peptide editing relies on HLA-DO tuning. *Sci. Rep.* **9**, 13877 (2019).

11.   Inaba, K., Young, J. W. & Steinman, R. M. Direct activation of CD8[+] cytotoxic T lymphocytes by dendritic cells. *J. Exp. Med.* **166**, 182–194 (1987).

12.   Kast, W. M., Boog, C. J., Roep, B. O., Voordouw, A. C. & Melief, C. J. Failure or success in the restoration of virus-specific cytotoxic T lymphocyte response defects by dendritic cells. *J. Immunol.* **140**, 3186–3193 (1988).

13.   Bennett, S. R. M., Carbone, F. R., Karamalis, F., Miller, J. F. A. P. & Heath, W. R. Induction of a CD8[+] Cytotoxic T Lymphocyte Response by Cross-priming Requires Cognate CD4[+] T Cell Help. *J. Exp. Med.* **186**, 65–70 (1997).

14.   Bennett, S. R. M. *et al.* Help for cytotoxic-T-cell responses is mediated by CD40 signalling. *Nature* **393**, 478–480 (1998).

15.   Smith, C. M. *et al.* Cognate CD4[+] T cell licensing of dendritic cells in CD8[+] T cell immunity. *Nat. Immunol.* **5**, 1143–1148 (2004).

16.   Schoenberger, S. P., Toes, R. E. M., van der Voort, E. I. H., Offringa, R. & Melief, C. J. M. T-cell help for cytotoxic T lymphocytes is mediated by CD40–CD40L interactions.

Nature **393**, 480–483 (1998).

17. Bourgeois, C., Rocha, B. & Tanchot, C. A Role for CD40 Expression on CD8+ T Cells in the Generation of CD8+ T Cell Memory. *Science* **297**, 2060–2063 (2002).

18. Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).

19. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature* **335**, 744 (1988).

20. Blattman, J. N. *et al.* Estimating the precursor frequency of naive antigen-specific CD8 T cells. *J. Exp. Med.* **195**, 657–664 (2002).

21. Chen, X., Poncette, L. & Blankenstein, T. Human TCR-MHC coevolution after divergence from mice includes increased nontemplate-encoded CDR3 diversity. *J. Exp. Med.* **214**, 3417–3433 (2017).

22. de Greef, P. C. *et al.* The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes. *Elife* **9**, e49900 (2020).

23. Zarnitsyna, V. I., Evavold, B. D., Schoettle, L. N., Blattman, J. N. & Antia, R. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front. Immunol.* **4**, 485 (2013).

24. McDonald, B. D., Bunker, J. J., Erickson, S. A., Oh-Hora, M. & Bendelac, A. Crossreactive αβ T Cell Receptors Are the Predominant Targets of Thymocyte Negative Selection. *Immunity* **43**, 859–869 (2015).

25. Merkenschlager, M. *et al.* How many thymocytes audition for selection? *J. Exp. Med.* **186**, 1149–1158 (1997).

26. Blackman, M. *et al.* The T cell repertoire may be biased in favor of MHC recognition. *Cell* **47**, 349–357 (1986).

27. Gras, S. *et al.* Allelic polymorphism in the T cell receptor and its impact on immune responses. *J. Exp. Med.* **207**, 1555–1567 (2010).

28. Zerrahn, J., Held, W. & Raulet, D. H. The MHC Reactivity of the T Cell Repertoire Prior to Positive and Negative Selection. *Cell* **88**, 627–636 (1997).

29. Suchin, E. J. *et al.* Quantifying the Frequency of Alloreactive T Cells In Vivo: New Answers to an Old Question. *J. Immunol.* **166**, 973–981 (2001).

30. Garcia, K. C., Adams, J. J., Feng, D. & Ely, L. K. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nat. Immunol.* **10**, 143–147 (2009).

31. Sim, B.-C., Zerva, L., Greene, M. I. & Gascoigne, N. R. J. Control of MHC Restriction by TCR VαCDR1 and CDR2. *Science* **273**, 963–966 (1996).

32. Marrack, P., Scott-Browne, J. P., Dai, S., Gapin, L. & Kappler, J. W. Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu. Rev. Immunol.* **26**, 171–203 (2008).

33. Huseby, E. S. *et al.* How the T Cell Repertoire Becomes Peptide and MHC Specific. *Cell* **122**, 247–260 (2005).

34. Beringer, D. X. *et al.* T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat. Immunol.* **16**, 1153–1161 (2015).

35. Hahn, M., Nicholson, M. J., Pyrdol, J. & Wucherpfennig, K. W. Unconventional topology of self peptide–major histocompatibility complex binding by a human autoimmune T cell receptor. *Nat. Immunol.* **6**, 490–496 (2005).

36. Adams, J. J. *et al.* Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. *Nat. Immunol.* **17**, 87–94 (2016).

37. Rossjohn, J. *et al.* T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).

38. Garcia, K. C. & Adams, E. J. How the T Cell Receptor Sees Antigen—A Structural View. *Cell* **122**, 333–336 (2005).

39. Borg, N. A. *et al.* The CDR3 regions of an immunodominant T cell receptor dictate the 'energetic landscape' of peptide-MHC recognition. *Nat. Immunol.* **6**, 171–180 (2005).

40. Rammensee, H.-G., Friede, T. & Stevanović, S. MHC ligands and peptide motifs: first listing. *Immunogenetics* **41**, 178–228 (1995).

41. Falk, K., Rötzschke, O., Stevanovíc, S., Jung, G. & Rammensee, H.-G. Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules. *Immunogenetics* **39**, 230–242 (1994).

42. Germain, R. N. & Margulies, D. H. The Biochemistry and Cell Biology of Antigen Processing and Presentation. *Annu. Rev. Immunol.* **11**, 403–450 (1993).

43. York, I. A. & Rock, K. L. Antigen processing and presentation by the class I major histocompatibility complex. *Annu. Rev. Immunol.* **14**, 369–396 (1996).

44. Tewari, M. K., Sinnathamby, G., Rajagopal, D. & Eisenlohr, L. C. A cytosolic pathway for MHC class II–restricted antigen processing that is proteasome and TAP dependent. *Nat. Immunol.* **6**, 287–294 (2005).

45. Salter, R. D. *et al.* A binding site for the T-cell co-receptor CD8 on the α3 domain of HLA-A2. *Nature* **345**, 41–46 (1990).

46. Schubert, U. *et al.* Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404**, 770–774 (2000).

47. Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T. & Cresswell, P. Roles for Calreticulin and a Novel Glycoprotein, Tapasin, in the Interaction of MHC Class I Molecules with TAP. *Immunity* **5**, 103–114 (1996).

48. Blum, J. S., Wearsch, P. A. & Cresswell, P. Pathways of antigen processing. *Annu. Rev. Immunol.* **31**, 443–473 (2013).

49. Nguyen, T. T. *et al.* Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. *Nat. Struct. Mol. Biol.* **18**, 604–613 (2011).

50. Peters, P. J., Neefjes, J. J., Oorschot, V., Ploegh, H. L. & Geuze, H. J. Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature* **349**, 669–676 (1991).

51. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).

52. Hudson, T. J. *et al.* International network of cancer genome projects. *Nature* **464**, 993–998 (2010).

53. Lee, W. *et al.* The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* **465**, 473–477 (2010).

54. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–158 (2007).

55. Haber, D. A. & Settleman, J. Drivers and passengers. *Nature* **446**, 145–146 (2007).

56. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100**, 57–70 (2000).

57. Junttila, M. R. & de Sauvage, F. J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**, 346–354 (2013).

58. Lu, P., Weaver, V. M. & Werb, Z. The extracellular matrix: A dynamic niche in cancer progression. *J. Cell Biol.* **196**, 395–406 (2012).

59. Tlsty, T. D. & Hein, P. W. Know thy neighbor: stromal cells can contribute oncogenic signals. *Curr. Opin. Genet. Dev.* **11**, 54–59 (2001).

60. Bissell, M. J. & Radisky, D. Putting tumours in context. *Nat. Rev. Cancer* **1**, 46–54 (2001).

61. Coussens, L. M. & Werb, Z. Inflammation and cancer. *Nature* **420**, 860–867 (2002).

62. Philip, M., Rowley, D. A. & Schreiber, H. Inflammation as a tumor promoter in cancer induction. *Semin. Cancer Biol.* **14**, 433–439 (2004).

63. Bhowmick, N. A. *et al.* TGF-ß Signaling in Fibroblasts Modulates the Oncogenic Potential of Adjacent Epithelia. *Science* **303**, 848–851 (2004).

64. Folkman, J. Fundamental concepts of the angiogenic process. *Curr. Mol. Med.* **3**, 643—651 (2003).

65. Mantovani, A., Sozzani, S., Locati, M., Allavena, P. & Sica, A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol.* **23**, 549–555 (2002).

66. Silzle, T., Randolph, G. J., Kreutz, M. & Kunz-Schughart, L. A. The fibroblast: Sentinel cell and local immune modulator in tumor tissue. *Int. J. Cancer* **108**, 173–180 (2004).

67. Blankenstein, T. The role of tumor stroma in the interaction between tumor and immune

system. *Curr. Opin. Immunol.* **17**, 180–186 (2005).

68. Greenberg, P. D. Adoptive T Cell Therapy of Tumors: Mechanisms Operative in the Recognition and Elimination of Tumor Cells. *Advances in Immunology* **49**, 281–355 (1991).

69. Spiotto, M. T., Rowley, D. A. & Schreiber, H. Bystander elimination of antigen loss variants in established tumors. *Nat. Med.* **10**, 294–298 (2004).

70. Poehlein, C. H. *et al.* TNF Plays an Essential Role in Tumor Regression after Adoptive Transfer of Perforin/IFN-γ Double Knockout Effector T Cells. *J. Immunol.* **170**, 2004–2013 (2003).

71. Hollenbaugh, J. A., Reome, J., Dobrzanski, M. & Dutton, R. W. The Rate of the CD8-Dependent Initial Reduction in Tumor Volume Is Not Limited by Contact-Dependent Perforin, Fas Ligand, or TNF-Mediated Cytolysis. *J. Immunol.* **173**, 1738–1743 (2004).

72. Trapani, J. A. & Smyth, M. J. Functional significance of the perforin/granzyme cell death pathway. *Nat. Rev. Immunol.* **2**, 735–747 (2002).

73. Kammertoens, T. *et al.* Tumour ischaemia by interferon-γ resembles physiological blood vessel regression. *Nature* **545**, 98–102 (2017).

74. Burnet, M. Immunological factors in the process of carcinogenesis. *Br. Med. Bull.* **20**, 154–158 (1964).

75. Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002).

76. Qin, Z. & Blankenstein, T. A cancer immunosurveillance controversy. *Nat. Immunol.* **5**, 3–4 (2004).

77. Willimsky, G. & Blankenstein, T. Sporadic immunogenic tumours avoid destruction by inducing T-cell tolerance. *Nature* **437**, 141–146 (2005).

78. Straathof, K. C. M., Bollard, C. M., Rooney, C. M. & Heslop, H. E. Immunotherapy for Epstein-Barr Virus-Associated Cancers in Children. *Oncologist* **8**, 83–98 (2003).

79. de Jong, A. *et al.* Human Papillomavirus Type 16-Positive Cervical Cancer Is Associated with Impaired CD4[+] T-Cell Immunity against Early Antigens E2 and E6. *Cancer Res.* **64**, 5449–5455 (2004).

80. Street, S. E. A., Trapani, J. A., MacGregor, D. & Smyth, M. J. Suppression of Lymphoma and Epithelial Malignancies Effected by Interferon γ . *J. Exp. Med.* **196**, 129–134 (2002).

81. Shankaran, V. *et al.* IFNγ and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* **410**, 1107–1111 (2001).

82. Schietinger, A., Philip, M. & Schreiber, H. Specificity in cancer immunotherapy. *Semin. Immunol.* **20**, 276–285 (2008).

83. Sampson, J. H., Archer, G. E., Mitchell, D. A., Heimberger, A. B. & Bigner, D. D. Tumor-specific immunotherapy targeting the EGFRvIII mutation in patients with malignant

glioma. *Semin. Immunol.* **20**, 267–275 (2008).

84. Neller, M. A., López, J. A. & Schmidt, C. W. Antigens for cancer immunotherapy. *Semin. Immunol.* **20**, 286–295 (2008).

85. Lucas, S. & Coulie, P. G. About human tumor antigens to be used in immunotherapy. *Semin. Immunol.* **20**, 301–307 (2008).

86. Buckwalter, M. R. & Srivastava, P. K. "It is the antigen(s), stupid" and other lessons from over a decade of vaccitherapy of human cancer. *Semin. Immunol.* **20**, 296–300 (2008).

87. Rosenberg, S. A. *et al.* Tumor Progression Can Occur despite the Induction of Very High Levels of Self/Tumor Antigen-Specific CD8$^+$ T Cells in Patients with Melanoma. *J. Immunol.* **175**, 6169–6176 (2005).

88. Srivastava, P. K. Immunotherapy of human cancer: lessons from mice. *Nat. Immunol.* **1**, 363–366 (2000).

89. Schietinger, A., Delrow, J. J., Basom, R. S., Blattman, J. N. & Greenberg, P. D. Rescued tolerant CD8 T cells are preprogrammed to reestablish the tolerant state. *Science* **335**, 723–727 (2012).

90. Blankenstein, T., Leisegang, M., Uckert, W. & Schreiber, H. Targeting cancer-specific mutations by T cell receptor gene therapy. *Curr. Opin. Immunol.* **33**, 112–119 (2015).

91. Koeppen, H., Singh, S. & Schreiber, H. Genetically Engineered Vaccines Comparison of Active versus Passive Immunotherapy against Solid Tumorsa. *Ann. N. Y. Acad. Sci.* **690**, 244–255 (1993).

92. Schreiber, K. *et al.* Spleen cells from young but not old immunized mice eradicate large established cancers. *Clin. Cancer Res.* **18**, 2526–2533 (2012).

93. Schietinger, A. *et al.* Longitudinal confocal microscopy imaging of solid tumor destruction following adoptive T cell transfer. *Oncoimmunology* **2**, e26677–e26677 (2013).

94. Listopad, J. J. *et al.* Fas expression by tumor stroma is required for cancer eradication. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2276–2281 (2013).

95. Kammertoens, T., Schüler, T. & Blankenstein, T. Immunotherapy: target the stroma to hit the tumor. *Trends Mol. Med.* **11**, 225–231 (2005).

96. Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N. Engl. J. Med.* **365**, 725–733 (2011).

97. Maude, S. L. *et al.* Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).

98. Grupp, S. A. *et al.* Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *N. Engl. J. Med.* **368**, 1509–1518 (2013).

99. Brentjens, R. J. *et al.* Safety and persistence of adoptively transferred autologous CD19-targeted T cells in patients with relapsed or chemotherapy refractory B-cell leukemias.

Blood **118**, 4817–4828 (2011).

100. Kochenderfer, J. N. *et al.* B-cell depletion and remissions of malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor-transduced T cells. *Blood* **119**, 2709–2720 (2012).

101. Hirayama, A. V *et al.* High rate of durable complete remission in follicular lymphoma after CD19 CAR-T cell immunotherapy. *Blood* **134**, 636–640 (2019).

102. Srivastava, S. & Riddell, S. R. Chimeric Antigen Receptor T Cell Therapy: Challenges to Bench-to-Bedside Efficacy. *J. Immunol.* **200**, 459–468 (2018).

103. Castellarin, M., Watanabe, K., June, C. H., Kloss, C. C. & Posey, A. D. Driving cars to the clinic for solid tumors. *Gene Ther.* **25**, 165–175 (2018).

104. O'Rourke, D. M. *et al.* A single dose of peripherally infused EGFRvIII-directed CAR T cells mediates antigen loss and induces adaptive resistance in patients with recurrent glioblastoma. *Sci. Transl. Med.* **9**, 399 (2017).

105. Martinez, M. & Moon, E. K. CAR T Cells for Solid Tumors: New Strategies for Finding, Infiltrating, and Surviving in the Tumor Microenvironment. *Front. Immunol.* **10,** 128 (2019).

106. Kershaw, M. H. *et al.* A phase I study on adoptive immunotherapy using gene-modified T cells for ovarian cancer. *Clin. Cancer Res.* **12**, 6106–6115 (2006).

107. Kershaw, M. H. *et al.* Redirecting Migration of T Cells to Chemokine Secreted from Tumors by Genetic Modification with CXCR2. *Hum. Gene Ther.* **13**, 1971–1980 (2002).

108. Kloss, C. C. *et al.* Dominant-Negative TGF-β Receptor Enhances PSMA-Targeted Human CAR T Cell Proliferation And Augments Prostate Cancer Eradication. *Mol. Ther.* **26**, 1855–1866 (2018).

109. Cherkassky, L. *et al.* Human CAR T cells with cell-intrinsic PD-1 checkpoint blockade resist tumor-mediated inhibition. *J. Clin. Invest.* **126**, 3130–3144 (2016).

110. Hinrichs, C. S. & Restifo, N. P. Reassessing target antigens for adoptive T-cell therapy. *Nat. Biotechnol.* **31**, 999–1008 (2013).

111. Geiger, R. *et al.* L-Arginine Modulates T Cell Metabolism and Enhances Survival and Anti-tumor Activity. *Cell* **167**, 829–842.e13 (2016).

112. Kieback, E., Charo, J., Sommermeyer, D., Blankenstein, T. & Uckert, W. A safeguard eliminates T cell receptor gene-modified autoreactive T cells after adoptive transfer. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 623–628 (2008).

113. Shen, Z. Genomic instability and cancer: an introduction. *J. Mol. Cell Biol.* **3**, 1–3 (2011).

114. Albertson, D. G. Gene amplification in cancer. *Trends Genet.* **22**, 447–455 (2006).

115. Santarius, T., Shipley, J., Brewer, D., Stratton, M. R. & Cooper, C. S. A census of amplified and overexpressed human cancer genes. *Nat. Rev. Cancer* **10**, 59–64 (2010).

116. Rabbitts, T. H. Chromosomal translocations in human cancer. *Nature* **372**, 143–149

(1994).

117. Storchova, Z. & Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nat. Rev. Mol. Cell Biol.* **5**, 45–54 (2004).

118. Bozic, I. *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 18545–18550 (2010).

119. Loeb, K. R. & Loeb, L. A. Significance of multiple mutations in cancer. *Carcinogenesis* **21**, 379–385 (2000).

120. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).

121. Urban, J. L. & Schreiber, H. Tumor Antigens. *Annu. Rev. Immunol.* **10**, 617–644 (1992).

122. Liu, X. S. & Mardis, E. R. Applications of Immunogenomics to Cancer. *Cell* **168**, 600–612 (2017).

123. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).

124. Monach, P. A., Meredith, S. C., T.Siegel, C. & Schreiber, H. A unique tumor antigen produced by a single amino acid substitution. *Immunity* **2**, 45–59 (1995).

125. Beck-Engeser, G. B. *et al.* Point mutation in essential genes with loss or mutation of the second allele: relevance to the retention of tumor-specific antigens. *J. Exp. Med.* **194**, 285–300 (2001).

126. Assarsson, E. *et al.* A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. *J. Immunol.* **178**, 7890–7901 (2007).

127. Wettstein, P. J. & Bailey, D. W. Immunodominance in the immune response to "multiple" histocompatibility antigens. *Immunogenetics* **16**, 47–58 (1982).

128. Wortzel, R. D., Philipps, C. & Schreiber, H. Multiple tumour-specific antigens expressed on a single tumour cell. *Nature* **304**, 165–167 (1983).

129. Paul, S. *et al.* HLA Class I Alleles Are Associated with Peptide-Binding Repertoires of Different Size, Affinity, and Immunogenicity. *J. Immunol.* **191**, 5831–5839 (2013).

130. Bauer, K. *et al.* T cell responses against microsatellite instability-induced frameshift peptides and influence of regulatory T cells in colorectal cancer. *Cancer Immunol. Immunother.* **62**, 27–37 (2013).

131. Baurain, J.-F. *et al.* High Frequency of Autologous Anti-Melanoma CTL Directed Against an Antigen Generated by a Point Mutation in a New Helicase Gene. *J. Immunol.* **164**, 6057–6066 (2000).

132. Chiari, R. *et al.* Two Antigens Recognized by Autologous Cytolytic T Lymphocytes on a Melanoma Result from a Single Point Mutation in an Essential Housekeeping Gene. *Cancer Res.* **59**, 5785–5792 (1999).

133. Echchakir, H. *et al.* A Point Mutation in the α-Actinin-4 Gene Generates an Antigenic

Peptide Recognized by Autologous Cytolytic T Lymphocytes on a Human Lung Carcinoma. *Cancer Res.* **61**, 4078–4083 (2001).

134. Coulie, P. G. *et al.* A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7976–7980 (1995).

135. Brändle, D., Brasseur, F., Weynants, P., Boon, T. & Van den Eynde, B. A mutated HLA-A2 molecule recognized by autologous cytotoxic T lymphocytes on a human renal cell carcinoma. *J. Exp. Med.* **183**, 2501–2508 (1996).

136. Guéguen, M. *et al.* An Antigen Recognized by Autologous CTLs on a Human Bladder Carcinoma. *J. Immunol.* **160**, 6188–6194 (1998).

137. Hogan, K. T. *et al.* The Peptide Recognized by HLA-A68.2-restricted, Squamous Cell Carcinoma of the Lung-specific Cytotoxic T Lymphocytes Is Derived from a Mutated Elongation Factor 2 Gene. *Cancer Res.* **58**, 5144–5150 (1998).

138. Gaudin, C., Kremer, F., Angevin, E., Scott, V. & Triebel, F. A hsp70-2 Mutation Recognized by CTL on a Human Renal Cell Carcinoma. *J. Immunol.* **162**, 1730–1738 (1999).

139. Graf, C. *et al.* A neoepitope generated by an FLT3 internal tandem duplication (FLT3-ITD) is recognized by leukemia-reactive autologous CD8[+] T cells. *Blood* **109**, 2985–2988 (2006).

140. Mandruzzato, S., Brasseur, F., Andry, G., Boon, T. & van der Bruggen, P. A CASP-8 mutation recognized by cytolytic T lymphocytes on a human head and neck carcinoma. *J. Exp. Med.* **186**, 785–793 (1997).

141. Robbins, P. F. *et al.* A mutated beta-catenin gene encodes a melanoma-specific antigen recognized by tumor infiltrating lymphocytes. *J. Exp. Med.* **183**, 1185–1192 (1996).

142. Takenoyama, M. *et al.* A point mutation in the NFYC gene generates an antigenic peptide recognized by autologous cytolytic T lymphocytes on a human squamous cell lung carcinoma. *Int. J. Cancer* **118**, 1992–1997 (2006).

143. De Smet, C. *et al.* Genes coding for melanoma antigens recognised by cytolytic T lymphocytes. *Eye* **11**, 243–248 (1997).

144. Renkvist, N., Castelli, C., Robbins, P. F. & Parmiani, G. A listing of human tumor antigens recognized by T cells. *Cancer Immunol. Immunother.* **50**, 3–15 (2001).

145. Ma, W. W. & Adjei, A. A. Novel Agents on the Horizon for Cancer Therapy. *CA. Cancer J. Clin.* **59**, 111–137 (2009).

146. Kammertoens, T. & Blankenstein, T. Making and circumventing tolerance to cancer. *Eur. J. Immunol.* **39**, 2345–2353 (2009).

147. Keefe, D. M. K. & Bateman, E. H. Tumor control versus adverse events with targeted anticancer therapies. *Nat. Rev. Clin. Oncol.* **9**, 98–109 (2012).

148. Vanneman, M. & Dranoff, G. Combining immunotherapy and targeted therapies in cancer treatment. *Nat Rev Cancer* **12**, 237–251 (2012).

149. Amos, S. M. *et al.* Autoimmunity associated with immunotherapy of cancer. *Blood* **118**, 499–509 (2011).

150. Anders, K. *et al.* Oncogene-targeting T cells reject large tumors while oncogene inactivation selects escape variants in mouse models of cancer. *Cancer Cell* **20**, 755–767 (2011).

151. Modjtahedi, H., Ali, S. & Essapen, S. Therapeutic application of monoclonal antibodies in cancer: Advances and challenges. *British Medical Bulletin* **104**, 41–59 (2012).

152. Schreiber, K., Wu, T. H., Kast, W. M. & Schreiber, H. Tracking the Common Ancestry of Antigenically Distinct Cancer Variants. *Clin. Cancer Res.* **7**, 871-875 (2001).

153. Tao, Y. *et al.* Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12042–12047 (2011).

154. Blanc, E. *et al.* Identification and ranking of recurrent neo-epitopes in cancer. *BMC Med. Genomics* **12**, (2019).

155. Linnemann, C. *et al.* High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nat. Med.* **19**, 1534–1541 (2013).

156. Ali, M. *et al.* Induction of neoantigen-reactive T cells from healthy donors. *Nat. Protoc.* **14**, 1926–1943 (2019).

157. Li, L.-P. *et al.* Transgenic mice with a diverse human T cell antigen receptor repertoire. *Nat. Med.* **16**, 1029 (2010).

158. Karpanen, T. & Olweus, J. The Potential of Donor T-Cell Repertoires in Neoantigen-Targeted Cancer Immunotherapy. *Front. Immunol.* **8**, 1718 (2017).

159. Redmond, W. L., Marincek, B. C. & Sherman, L. A. Distinct Requirements for Deletion versus Anergy during CD8 T Cell Peripheral Tolerance In Vivo. *J. Immunol.* **174**, 2046–2053 (2005).

160. Man, S. *et al.* Definition of a human T cell epitope from influenza A non-structural protein 1 using HLA-A2.1 transgenic mice. *Int. Immunol.* **7**, 597–605 (1995).

161. Shirai, M. *et al.* CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL of humans carrying HLA-A2.1. *J. Immunol.* **154**, 2733–2742 (1995).

162. Pascolo, S. *et al.* HLA-A2.1-restricted education and cytolytic activity of CD8(+) T lymphocytes from beta2 microglobulin (beta2m) HLA-A2.1 monochain transgenic H-2Db beta2m double knockout mice. *J. Exp. Med.* **185**, 2043–2051 (1997).

163. Terajima, M. *et al.* Identification of Vaccinia CD8[+] T-Cell Epitopes Conserved among Vaccinia and Variola Viruses Restricted by Common MHC Class I Molecules, HLA-A2

or HLA-B7. *Hum. Immunol.* **67**, 512–520 (2006).

164. Terajima, M. *et al.* Quantitation of CD8⁺ T Cell Responses to Newly Identified HLA-A*0201–restricted T Cell Epitopes Conserved Among Vaccinia and Variola (Smallpox) Viruses. *J. Exp. Med.* **197**, 927–932 (2003).

165. Pasquetto, V. *et al.* HLA-A*0201, HLA-A*1101, and HLA-B*0702 Transgenic Mice Recognize Numerous Poxvirus Determinants from a Wide Variety of Viral Gene Products. *J. Immunol.* **175**, 5504–5515 (2005).

166. Kotturi, M. F. *et al.* Of mice and humans: how good are HLA transgenic mice as a model of human immune responses? *Immunome Res.* **5**, 3 (2009).

167. Wentworth, P. A. *et al.* Differences and similarities in the A2.1-restricted cytotoxic T cell repertoire in humans and human leukocyte antigen-transgenic mice. *Eur. J. Immunol.* **26**, 97–101 (1996).

168. Morgan, R. A. *et al.* Cancer regression and neurologic toxicity following anti-MAGE-A3 TCR gene therapy. *J. Immunother.* **36**, 133–151 (2013).

169. Chinnasamy, N. *et al.* A TCR Targeting the HLA-A*0201–Restricted Epitope of MAGE-A3 Recognizes Multiple Epitopes of the MAGE-A Antigen Superfamily in Several Types of Cancer. *J. Immunol.* **186**, 685–696 (2011).

170. Davis, J. L. *et al.* Development of Human Anti-Murine T-Cell Receptor Antibodies in Both Responding and Nonresponding Patients Enrolled in TCR Gene Therapy Trials. *Clin. Cancer Res.* **16**, 5852–5861 (2010).

171. Obenaus, M. *et al.* Identification of human T-cell receptors with optimal affinity to cancer antigens using antigen-negative humanized mice. *Nat. Biotechnol.* **33**, 402 (2015).

172. Gavvovidis, I. *et al.* Targeting Merkel Cell Carcinoma by Engineered T Cells Specific to T-Antigens of Merkel Cell Polyomavirus. *Clin. Cancer Res.* **24**, 3644–3655 (2018).

173. Lundegaard, C. *et al.* NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res.* **36**, W509–W512 (2008).

174. Nielsen, M. *et al.* Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017 (2003).

175. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32**, 511–517 (2016).

176. Buus, S. *et al.* Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue Antigens* **62**, 378–384 (2003).

177. Peters, B. & Sette, A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* **6**, 132 (2005).

178. Gonzalez-Galarza, F. F., Christmas, S., Middleton, D. & Jones, A. R. Allele frequency

net: A database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* **39**, D913–D919 (2011).

179. Takeshita, L. Y. C., Jones, A. R., Gonzalez-Galarza, F. F. & Middleton, D. Allele frequencies database. *Transfus. Med. Hemother.* **41**, 352–355 (2014).

180. Maiers, M. *et al.* Maintaining updated DNA-based HLA assignments in the National Marrow Donor Program Bone Marrow Registry. *Rev. Immunogenet.* **2**, 449–460 (2000).

181. Robinson, J. *et al.* IPD-IMGT/HLA Database. *Nucleic Acids Res.* **48**, D948–D955 (2020).

182. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–E386 (2015).

183. Pascolo, S. *et al.* HLA-A2.1–restricted Education and Cytolytic Activity of CD8(+) T Lymphocytes from β2 Microglobulin (β2m) HLA-A2.1 Monochain Transgenic H-2D(b) β2m Double Knockout Mice. *J. Exp. Med.* **185**, 2043–2051 (1997).

184. Ding, S. *et al.* Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).

185. Ittner, L. M. & Götz, J. Pronuclear injection for the production of transgenic mice. *Nat. Protoc.* **2**, 1206–1215 (2007).

186. Robins, H. S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099–4107 (2009).

187. Carlson, C. S. *et al.* Using synthetic templates to design an unbiased multiplex PCR assay. *Nat. Commun.* **4**, 2680 (2013).

188. Morita, S., Kojima, T. & Kitamura, T. Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* **7**, 1063–1066 (2000).

189. Elliott, N. E. *et al.* FERM domain mutations induce gain of function in JAK3 in adult T-cell leukemia/lymphoma. *Blood* **118**, 3911–3921 (2011).

190. Wei, X. *et al.* Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat. Genet.* **43**, 442–448 (2011).

191. Tzoneva, G. *et al.* Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nat. Med.* **19**, 368–371 (2013).

192. Dias-Santagata, D. *et al.* Braf V600E mutations are common in pleomorphic xanthoastrocytoma: Diagnostic and therapeutic implications. *PLoS One* **6**, e17948–e17948 (2011).

193. Kompier, L. C. *et al.* FGFR3, HRAS, KRAS, NRAS AND PIK3CA mutations in bladder cancer and their potential as biomarkers for surveillance and therapy. *PLoS One* **5**, e13821–e13821 (2010).

194. Petitjean, A. *et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: Lessons from recent developments in the IARC TP53 database.

Hum. Mutat. **28**, 622–629 (2007).

195. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).

196. Khalili, J. S., Hanson, R. W. & Szallasi, Z. In silico prediction of tumor antigens derived from functional missense mutations of the cancer gene census. *Oncoimmunology* **1**, 1281–1289 (2012).

197. Di Martino, E., Tomlinson, D. C. & Knowles, M. A. A decade of FGF receptor research in bladder cancer: Past, present, and future challenges. *Adv. Urol.* **2012**, 429213 (2012).

198. Banerji, S. *et al.* Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405–409 (2012).

199. Kohlmann, A. *et al.* The Interlaboratory RObustness of Next-generation sequencing (IRON) study: A deep sequencing investigation of TET2, CBL and KRAS mutations by an international consortium involving 10 laboratories. *Leukemia* **25**, 1840–1848 (2011).

200. Kan, Z. *et al.* Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466**, 869–873 (2010).

201. Hodis, E. *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).

202. Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat. Genet.* **44**, 685–689 (2012).

203. More, H. *et al.* Identification of seven novel germline mutations in the human E-cadherin (CDH1) gene. *Hum. Mutat.* **28**, 203 (2007).

204. Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44**, 694–698 (2012).

205. Mitsudomi, T. & Yatabe, Y. Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer. *FEBS J.* **277**, 301–308 (2010).

206. Ghadjar, P. *et al.* MET Y1253D-activating point mutation and development of distant metastasis in advanced head and neck cancers. *Clin. Exp. Metastasis* **26**, 809–815 (2009).

207. Clark, V. E. *et al.* Genomic analysis of non-NF2 meningiomas reveals mutations in TRAF7, KLF4, AKT1, and SMO. *Science* **339**, 1077–1080 (2013).

208. Le Gallo, M. *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat. Genet.* **44**, 1310–1315 (2012).

209. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).

210. Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat. Genet.* **44**, 47–52 (2012).

211. Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**, 101–105 (2011).

212. Church, D. N. *et al.* DNA polymerase 1 and d exonuclease domain mutations in endometrial cancer. *Hum. Mol. Genet.* **22**, 2820–2828 (2013).

213. Shih, I. M. *et al.* Somatic mutations of PPP2R1A in ovarian and uterine carcinomas. *Am. J. Pathol.* **178**, 1442–1447 (2011).

214. Bullock, A. N., Henckel, J. & Fersht, A. R. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: Definition of mutant states for rescue in cancer therapy. *Oncogene* **19**, 1245–1256 (2000).

215. Shibata, T. *et al.* NRF2 mutation confers malignant potential and resistance to chemoradiation therapy in advanced esophageal squamous cancer. *Neoplasia* **13**, 864–873 (2011).

216. Chiu, C.-H., Wang, Y.-T., Walther, B. A. & Chao, A. An improved nonparametric lower bound of species richness via a modified good–turing frequency formula. *Biometrics* **70**, 671–682 (2014).

217. Tubb, V. M. *et al.* Isolation of T cell receptors targeting recurrent neoantigens in hematological malignancies. *J. Immunother. Cancer* **6**, 70 (2018).

218. Wang, Q. J. *et al.* Identification of T-cell Receptors Targeting KRAS-Mutated Human Tumors. *Cancer Immunol. Res.* **4**, 204–214 (2016).

219. Lorenz, F. K. M. *et al.* Unbiased Identification of T-Cell Receptors Targeting Immunodominant Peptide-MHC Complexes for T-Cell Receptor Immunotherapy. *Hum. Gene Ther.* **28**, 1158–1168 (2017).

220. Kondo, E. *et al.* Identification of novel CTL epitopes of CMV-pp65 presented by a variety of HLA alleles. *Blood* **103**, 630–638 (2004).

221. Zhu, S. *et al.* Characterization of T-cell receptors directed against HLA-A*01-restricted and C*07-restricted epitopes of MAGE-A3 and MAGE-A12. *J. Immunother.* **35**, 680–688 (2012).

222. Engels, B. *et al.* Relapse or eradication of cancer is predicted by peptide-major histocompatibility complex affinity. *Cancer Cell* **23**, 516–526 (2013).

223. Kammertoens, T. & Blankenstein, T. It's the Peptide-MHC Affinity, Stupid. *Cancer Cell* **23**, 429–431 (2013).

224. Bjerregaard, A.-M. *et al.* An Analysis of Natural T Cell Responses to Predicted Tumor Neoepitopes. *Front. Immunol.* **8**, 1566 (2017).

225. Sahin, U. *et al.* Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **547**, 222–226 (2017).

226. Ott, P. A. *et al.* An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* **547**, 217–221 (2017).

227. Popović, J. *et al.* The only proposed T-cell epitope derived from the TEL-AML1 translocation is not naturally processed. *Blood* **118**, 946–954 (2011).

228. Textor, A. *et al.* Preventing tumor escape by targeting a post-proteasomal trimming independent epitope. *J. Exp. Med.* **213**, 2333–2348 (2016).

229. Leisegang, M., Kammertoens, T., Uckert, W. & Blankenstein, T. Targeting human melanoma neoantigens by T cell receptor gene therapy. *J. Clin. Invest.* **126**, 854–858 (2016).

230. Koh, W. S., Porter, J. R. & Batchelor, E. Tuning of mRNA stability through altering 3′-UTR sequences generates distinct output expression in a synthetic circuit driven by p53 oscillations. *Sci. Rep.* **9**, 5976 (2019).

231. Wang, Y., Wang, F., Wang, R., Zhao, P. & Xia, Q. 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. *Sci. Rep.* **5**, 16273 (2015).

232. Borgia, A. *et al.* Transient misfolding dominates multidomain protein folding. *Nat. Commun.* **6**, 8861 (2015).

233. Davidson, W. F., Kress, M., Khoury, G. & Jay, G. Comparison of HLA class I gene sequences. Derivation of locus-specific oligonucleotide probes specific for HLA-A, HLA-B, and HLA-C genes. *J. Biol. Chem.* **260**, 13414–13423 (1985).

234. Kramer, G., Shiber, A. & Bukau, B. Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. *Annu. Rev. Biochem.* **88**, 337–364 (2019).

235. Rooney, C. M. *et al.* Infusion of Cytotoxic T Cells for the Prevention and Treatment of Epstein-Barr Virus–Induced Lymphoma in Allogeneic Transplant Recipients. *Blood* **92**, 1549–1555 (1998).

236. Mulder, A. *et al.* Impact of Peptides on the Recognition of HLA Class I Molecules by Human HLA Antibodies. *J. Immunol.* **175**, 5950–5957 (2005).

237. Irwin, M. J., Heath, W. R. & Sherman, L. A. Species-restricted interactions between CD8 and the alpha 3 domain of class I influence the magnitude of the xenogeneic response. *J. Exp. Med.* **170**, 1091–1101 (1989).

238. Newberg, M. H., Ridge, J. P., Vining, D. R., Salter, R. D. & Engelhard, V. H. Species specificity in the interaction of CD8 with the alpha 3 domain of MHC class I molecules. *J. Immunol.* **149**, 136–142 (1992).

239. Samberg, N. L., Scarlett, E. C. & Stauss, H. J. The α3 domain of major histocompatibility complex class I molecules plays a critical role in cytotoxic T lymphocyte stimulation. *Eur. J. Immunol.* **19**, 2349–2354 (1989).

240. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144–149 (2016).

241. Rülicke, T. & Hübscher, U. Germ Line Transformation of Mammals by Pronuclear Microinjection. *Exp. Physiol.* **85**, 589–601 (2000).

242. Haruyama, N., Cho, A. & Kulkarni, A. B. Overview: engineering transgenic constructs and mice. *Curr. Protoc. cell Biol.* **Chapter 19**, Unit-19.10 (2009).

243. Rocha, B. & von Boehmer, H. Peripheral selection of the T cell repertoire. *Science* **251**, 1225–1228 (1991).

244. Almeida, A. R. M., Borghans, J. A. M. & Freitas, A. A. T Cell Homeostasis: Thymus Regeneration and Peripheral T Cell Restoration in Mice with a Reduced Fraction of Competent Precursors. *J. Exp. Med.* **194**, 591–600 (2001).

245. Marrack, P. *et al.* Homeostasis of αβ TCR+ T cells. *Nat. Immunol.* **1**, 107–111 (2000).

246. Zvyagin, I. V *et al.* Distinctive properties of identical twins' TCR repertoires revealed by high-throughput sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5980–5985 (2014).

247. Gavin, M. A. & Bevan, M. J. Increased peptide promiscuity provides a rationale for the lack of N regions in the neonatal T cell repertoire. *Immunity* **3**, 793–800 (1995).

248. Sharon, E. *et al.* Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat. Genet.* **48**, 995–1002 (2016).

249. Casrouge, A. *et al.* Size Estimate of the αβ TCR Repertoire of Naive Mouse Splenocytes. *J. Immunol.* **164**, 5782– 5787 (2000).

250. Nikolich-Žugich, J., Slifka, M. K. & Messaoudi, I. The many important facets of T-cell repertoire diversity. *Nat. Rev. Immunol.* **4**, 123–132 (2004).

251. Robins, H. S. *et al.* Overlap and effective size of the human CD8[+] T cell receptor repertoire. *Sci. Transl. Med.* **2**, 47ra64 (2010).

252. Arstila, T. P. *et al.* A Direct Estimate of the Human αβ T Cell Receptor Diversity. *Science* **286**, 958–961 (1999).

7. Acknowledgements

all my experimental mice. Next, I would like to express my gratitude to Elisa Kieback, CTO of T-knife GmbH, who offered me a job position with great trust even before finishing this Doctoral thesis. Thanks to all your scientific inputs and the push you gave me to get this done! Many thanks to my current colleagues Niklas, Lorenz, Sandra, and Jennifer. Many thanks to Niklas for all the great support he provided in purifying T cells for deep sequencing, which is also a part of his Master's thesis, and for writing the 'Zusammenfassung' of this thesis. Thanks to Lorenz and Sandra for their help hands in functional characterization assays! Thanks, Jennifer, for correcting the 'Zusammenfassung' and for all the pleasant science conversations!

I have no words to thank my family members. I am who I am because of them! Without them, I couldn't have made it to Germany. Thank you, Appa (Dhamodaran) and Amma (Kalaiselvi). They know nothing about Ph.D. life, and how it works in Germany, but always ask me if I am well, does my mouse strains are alive and working. The only thing they had was to trust me and stand by me always, and they know one day the research I did in Thomas's lab will help save at least one life. Thanks, Pa and Ma! My little sister, no matter what, has always wanted to bring out the best in me. From childhood fights, chit-chats to discussions on family values, she has always tried to trick me to bring me back to India to live with her. I will be back soon, Divi ma! Also, thank you, Shobi, Kishore, Mano, Murali, Kutti ma (Dhatchayani), Amuthamma (Amutha), Chithappas (Balu and Ravi), and Periya Athai. A close-to-heart special thanks to Natalie and Bernd Hofmann for all the love they showered on me and for their moral support. You are like my family to me, Natalie! Thank you, Seenu Uncle, Kumar Maama, Kamaraj and Shanthi (father and mother-in-law), Venkat, Sindhu, Naresh Thambi, and my grad school friends! (Amba, Karthi, Gokul, and Kadhambari) Because you showed such great interest in my research whenever I fly to India! Thanks to all my Berlin friends for their encouragement!

In my mother tongue, Tamil language, there is a proverb saying, you save the best for the last. My last thanks, but the most important of all the above acknowledgments, will be for my dearest wife, Abinaya Kamaraj, and my cute little son, Iaan (who became our family in 2019). I am deeply grateful to my loving life partner, my advisor, my life coach, my dearest wife, Abi. She has always stood by me at a more personal level than anyone in the world. She understood my project in and out and sacrificed a lot to achieve what I am doing. Without her, I can say that this dissertation wouldn't have made it this far. Thanks, Abi, for everything right from accompanying me to the lab for late-night experiments to long walks in depressing moments of our life up to sudden fitness regimen to veganism and now to sustainability. A salute to you! Thanks to my little-lucky charm, Iaan paapa! The month he was born, mice start expressing the phenotype we wanted. Let it be a coincidence (says my scientist's mind voice) or his precious positive vibe, whatever it may be! I am super happy to thank our little chellam, too! Thank you, Thatha (my late Grandfather)! This path of life started from the spark you gave. Though you are not here, I can feel your presence around me always!

8. Publications

Blanc, E., Holtgrewe, M., **Dhamodaran, A.** *et al*. Identification and ranking of recurrent neo-epitopes in cancer. *BMC Med Genomics* **12, 171** (2019). https://doi.org/10.1186/s12920-019-0611-7.

9. Curriculum Vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version.

10. Invention disclosures & IP

Prof. Dr. Thomas Blankenstein and I (Arunraj Dhamodaran) had filed an invention disclosure in 2019 on seven mouse strains at the Max Delbrück Center for Molecular Medicine (MDC), Berlin. As part of this Doctoral thesis (2013-2021), seven humanized TCR mice were generated and maintained at the MDC animal facilities. These include AB*ab*.I mice (six HLA alleles: A*03-A*11-B*07-B*15-C*04-C*07 as one haplotype) and six single transgenic mouse models: AB*ab*-A*03, AB*ab*-A*11, AB*ab*-B*07, AB*ab*-B*15, AB*ab*-C*04, and AB*ab*-C*07.

Patent application filing representing the mouse models as *in vivo* novel epitope discovery platform is pending. A publication is planned in a peer-reviewed journal for this dissertation after successful IP filing.

11. Declaration statement

Hiermit erkläre ich gemäß der Promotionsordnung der Freie Universität Berlin, dass ich

* die vorliegende Arbeit eigenständig unter Anleitung und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

* die Arbeit in gleicher oder ähnlicher Form nicht in anderen Promotionsverfahren vorgelegt wurde.

Berlin, den 27.04.2021

Arunraj Dhamodaran

12. Appendix

12.1 DNA sequence of the AB*ab*.I transgene construct in the founder mice

5'-
GCAGGCTGAATAATAAAAAAATTAGAAACTATTATTTAACCCTAGAAAGATAATCATATTGTGACGT
ACGTTAAAGATAATCATGCGTAAAATTGACGCATGTGTTTTATCGGTCTGTATATCGAGGTTTATTT
ATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAA
TTTATTTATGTTTATTTATTTATTAAAAAAAAACAAAAACTCAAAATTTCTTCTATAAAGTAACAAAAC
TTTTAAACATTCTCTCTTTTACAAAAATAAACGGCGCGCCTTCTTCTACATAAAACACACCCATCTG
GAGCTACAGAGGCTTCATATGGGAAGAAACAGGAGGAGTCTGAGACTAAGCCCGAGGCTGAGAA
GACAGATCCTGAGGAAATAGGCAAAGTCTCCCCTTTACAGATGAGAGTCCTGCACTCAGGCTTGG
CAGTGTGAGCCGCCCATTGCAGGTGAACAGAGCCTGGTCTCTGTGGGATCCCTGTGGGGCTTGC
AGGCCAGCGCCTCTGCTTTAAAGAGAAGCCTCTCTCCACTGCATCCCTAAGCGCTTGTGTCGCCA
TTGTATTCCCGGAAGTGACCTTTCTTCTAGAAGACTCTAGGGTGTGACTTCTGAAGAGAAGAAGG
AAGAGGAAGGGTGGAGGTTAGGAAACAGTGAGTCGGGCTTGTGGGTCTCTCCTGGTGATCTGAC
AGCTTCTGGGTCAGAACTCGGAGTCACCACGACAAACTGCTCTCTGTCCGCAGTACAGGGTTCA
GGCAAAGTCTTGGTTGCCAGGCGGTGAGGTCAGGGGTGGGGAAGCCCAGGGCTGGGGATTCCC
CATCTCCTCAGTTTCACTTCTGCACCTAACCTGGGTCAGGTCCTTCTGCCGGGACACTGATGACG
CGCTGGCAGGTCTCACTATCATTGGGTGGCGAGATTCCAGGAGCCAATCAGCGTCGCCGCGGAC
GCTGGTTATAAAGTCCACGCAACCCGCGGGACTCAGACACCCGGGATCCCAGATGGGCGCCAT
GGCCCCTAGAACCCTGCTGCTGCTCCTGGCCGCTGCCCTGGCCCCTACCCAGACCAGAGCCAT
CCAGAGAACCCCCAAGATCCAGGTGTACAGCAGACACCCCGCCGAGAACGGCAAGAGCAACTTC
CTGAACTGCTACGTGTCCGGCTTCCACCCCAGCGACATCGAGGTGGACCTGCTGAAGAACGGCG
AGAGAATCGAGAAGGTGGAACACAGCGACCTGAGCTTCAGCAAGGACTGGTCCTTCTACCTGCT
GTACTACACCGAGTTCACCCCCACCGAGAAGGACGAGTACGCCTGCAGAGTGAACCACGTGACC
CTGAGCCAGCCCAAGATCGTGAAGTGGGACAGAGACATGGGCGGAGGCGGCTCTGGTGGCGGA
GGAAGCGGAGGCGGAGGCAGCGGCAGCCACAGCATGAGATACTTTTTCACCAGCGTGTCCAGA
CCCGGCAGAGGCGAGCCCAGATTCATTGCCGTGGGCTACGTGGACGACACCCAGTTCGTCAGAT
TCGACAGCGACGCCGCCAGCCAGAGAATGGAACCCAGAGCCCCCTGGATCGAGCAGGAAGGCC
CCGAGTACTGGGACCAGGAAACCAGAAACGTCAAGGCCCAGAGCCAGACCGACAGAGTGGACC
TGGGCACCCTGAGAGGCTACTACAACCAGAGCGAGGCCGGCTCCCACACCATCCAGATTATGTA
CGGCTGCGACGTGGGCAGCGACGGCAGATTCCTGAGGGGCTACAGACAGGACGCCTACGACGG
CAAGGACTATATCGCCCTGAACGAGGACCTGAGAAGCTGGACAGCCGCCGACATGGCCGCCCA
GATCACCAAGAGAAAGTGGGAGGCCGCCCACGAGGCCGAGCAGCTGAGAGCCTACCTGGACGG
CACCTGTGTGGAATGGCTGCGGAGATACCTGGAAAACGGCAAAGAGACACTGCAGAGAACCGAC
AGCCCCAAGGCCCACGTGACACACCACCCTAGAAGCAAGGGCGAAGTGACCCTGCGGTGCTGG
GCTCTGGGCTTCTACCCCGCCGACATCACCCTGACCTGGCAGCTGAACGGCGAGGAACTGACCC
AGGACATGGAACTGGTGGAAACCAGACCTGCCGGCGACGGCACCTTCCAGAAATGGGCCAGCG
TGGTGGTGCCCCTGGGCAAAGAGCAGAACTACACCTGTAGAGTGTACCACGAGGGCCTGCCCG
AGCCCCTGACCCTGAGATGGGAGCCTCCTCCCAGCACCGACTCCTACATGGTCATCGTGGCCGT
GCTGGGAGTGCTGGGCGCTATGGCCATCATCGGCGCCGTGGTGGCCTTCGTGATGAAGAGGCG
GAGAAACACCGGCGGCAAGGGCGGCGATTACGCCCTGGCTCCTGGCAGCCAGAGCAGCGAGAT
GAGCCTGAGAGACTGCAAGGCCTGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTC
CCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAA
TTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA
GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGGCGGCC
GCGCGCCTTCTTCTACATAAAACACACCCATCTGGAGCTACAGAGGCTTCATATGGGAAGAAACA
GGAGGAGTCTGAGACTAAGCCCGAGGCTGAGAAGACAGATCCTGAGGAAATAGGCAAAGTCTCC
CCTTTACAGATGAGAGTCCTGCACTCAGGCTTGGCAGTGTGAGCCGCCCATTGCAGGTGAACAG
AGCCTGGTCTCTGTGGGATCCCTGTGGGGCTTGCAGGCCAGCGCCTCTGCTTTAAAGAGAAGCC
TCTCTCCACTGCATCCCTAAGCGCTTGTGTCGCCATTGTATTCCCGGAAGTGACCTTTCTTCTAGA
AGACTCTAGGGTGTGACTTCTGAAGAGAAGAAGGAAGAGGAAGGGTGGAGGTTAGGAAACAGTG
AGTCGGGCTTGTGGGTCTCTCCTGGTGATCTGACAGCTTCTGGGTCAGAACTCGGAGTCACCAC
GACAAACTGCTCTCTGTCCGCAGTACAGGGTTCAGGCAAAGTCTTGGTTGCCAGGCGGTGAGGT
CAGGGGTGGGGAAGCCCAGGGCTGGGGATTCCCCATCTCCTCAGTTTCACTTCTGCACCTAACC
TGGGTCAGGTCCTTCTGCCGGGACACTGATGACGCGCTGGCAGGTCTCACTATCATTGGGTGGC
GAGATTCCAGGAGCCAATCAGCGTCGCCGCGGACGCTGGTTATAAAGTCCACGCAACCCGCGG
GACTCAGACACCCGGGATCCCAGATGGGCGCCATGGCCCCTAGAACCCTGCTGCTGCTCCTGG
CCGCTGCCCTGGCCCCTACCCAGACCAGAGCCATCCAGAGAACCCCCAAGATCCAGGTGTACAG

CAGACACCCCGCCGAGAACGGCAAGAGCAACTTCCTGAACTGCTACGTGTCCGGCTTCCACCCC
AGCGACATCGAGGTGGACCTGCTGAAGAACGGCGAGAGAATCGAGAAGGTGGAACACAGCGAC
CTGAGCTTCAGCAAGGACTGGTCCTTCTACCTGCTGTACTACACCGAGTTCACCCCCACCGAGAA
GGACGAGTACGCCTGCAGAGTGAACCACGTGACCCTGAGCCAGCCCAAGATCGTGAAGTGGGA
CAGAGACATGGGCGGAGGCGGCTCTGGTGGCGGAGGAAGCGGAGGCGGAGGCAGCGGCAGC
CACAGCATGAGATACTTTTACACCTCCGTGTCCCGGCCTGGCAGGGGAGAGCCTCGGTTTATCG
CCGTGGGATATGTGGATGATACACAGTTTGTCCGCTTCGACTCCGACGCCGCCTCTCAGCGGAT
GGAACCTCGGGCTCCCTGGATTGAACAGGAAGGACCTGAATATTGGGATCAGGAAACACGGAAC
GTCAAAGCTCAGTCCCAGACAGACCGGGTCGACCTGGGAACACTGCGGGGATATTACAACCAGT
CCGAGGATGGCAGCCATACAATTCAGATTATGTATGGATGTGATGTGGGCCCTGACGGCCGGTT
CCTGAGAGGATACCGGCAGGATGCTTACGATGGAAAGGATTACATTGCCCTCAATGAGGACCTG
CGGTCCTGGACCGCCGCTGATATGGCTGCTCAGATTACAAAGCGGAAGTGGGAAGCTGCTCACG
CCGCTGAGCAGCAGCGGGCTTACCTGGAAGGCAGATGCGTCGAGTGGCTGAGGCGCTACCTCG
AGAACGGAAAAGAAACCCTGCAGCGGACCGACTCTCCTAAGGCTCACGTGACCCATCATCCCAG
GTCCAAGGGGGAAGTCACACTGAGATGTTGGGCCCTGGGCTTTTATCCTGCTGATATTACCCTCA
CATGGCAGCTCAATGGGGAAGAACTCACACAGGATATGGAACTCGTCGAGACAAGGCCCGCTGG
CGACGGAACATTTCAGAAGTGGGCTTCCGTCGTCGTGCCTCTCGGAAAAGAACAGAATTACACAT
GCCGGGTGTACCATGAAGGACTGCCTGAACCTCTCACACTCCGCTGGGAGCCCCCACCCTCCAC
AGACAGCTATATGGTCATTGTCGCTGTGCTCGGCGTCCTGGGAGCCATGGCTATCATTGGAGCT
GTGGTCGCTTTTGTCATGAAGCGCAGAAGAAACACAGGGGGAAAAGGCGGAGACTACGCTCTGG
CCCCAGGCTCCCAGTCCAGCGAGATGTCTCTGCGGGATTGCAAGGCTTGACTGTGCCTTCTAGT
TGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCAC
TGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
GGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG
ATGCGGTGGGCTCTATGGCCTGCAGGGCCTTCTTCTACATAAAACACACCCATCTGGAGCTACAG
AGGCTTCATATGGGAAGAAACAGGAGGAGTCTGAGACTAAGCCCGAGGCTGAGAAGACAGATCC
TGAGGAAATAGGCAAAGTCTCCCCTTTACAGATGAGAGTCCTGCACTCAGGCTTGGCAGTGTGAG
CCGCCCATTGCAGGTGAACAGAGCCTGGTCTCTGTGGGATCCCTGTGGGGCTTGCAGGCCAGC
GCCTCTGCTTTAAAGAGAAGCCTCTCTCCACTGCATCCCTAAGCGCTTGTGTCGCCATTGTATTC
CCGGAAGTGACCTTTCTTCTAGAAGACTCTAGGGTGTGACTTCTGAAGAGAAGAAGGAAGAGGAA
GGGTGGAGGTTAGGAAACAGTGAGTCGGGCTTGTGGGTCTCTCCTGGTGATCTGACAGCTTCTG
GGTCAGAACTCGGAGTCACCACGACAAACTGCTCTCTGTCCGCAGTACAGGGTTCAGGCAAAGT
CTTGGTTGCCAGGCGGTGAGGTCAGGGGTGGGGAAGCCCAGGGCTGGGGATTCCCCATCTCCT
CAGTTTCACTTCTGCACCTAACCTGGGTCAGGTCCTTCTGCCGGGACACTGATGACGCGCTGGC
AGGTCTCACTATCATTGGGTGGCGAGATTCCAGGAGCCAATCAGCGTCGCCGCGGACGCTGGTT
ATAAAGTCCACGCAACCCGCGGGACTCAGACACCCGGGATCCCAGATGGGCGCCATGGCCCCT
AGAACCCTGCTGCTGCTCCTGGCCGCTGCCCTGGCCCCTACCCAGACCAGAGCCATCCAGCGC
ACTCCTAAGATTCAGGTCTACAGCCGCCACCCAGCTGAAAACGGAAAGTCCAATTTCCTCAACTG
CTATGTCTCTGGATTTCACCCCTCCGACATTGAAGTGGACCTGCTCAAAAATGGCGAACGCATTG
AGAAAGTCGAGCATAGCGATCTCAGCTTCTCCAAGGATTGGTCTTTTTATCTGCTCTATTACACAG
AATTCACCCCTACAGAAAAGGATGAATACGCATGTCGAGTGAATCACGTCACACTCTCCCAGCCA
AAAATTGTGAAGTGGGACCGCGATATGGGAGGCGGAGGATCTGGGGGAGGTGGTAGCGGAGGG
GGCGGATCCGGCTCTCACTCCATGAGATATTTCTACACATCTGTGTCTAGGCCCGGACGGGGCG
AGCCCCGGTTCATCTCTGTCGGATACGTCGACGATACTCAGTTTGTCAGATTTGACTCTGATGCC
GCTAGCCCCAGAGAGGAACCACGCGCCCCTTGGATCGAACAGGAAGGGCCAGAGTATTGGGAC
CGGAACACCCAGATCTACAAGGCTCAGGCCCAGACCGATCGCGAGAGCCTGAGAAACCTGAGG
GGGTACTACAATCAGTCCGAAGCCGGAAGCCACACCCTGCAGTCTATGTACGGTTGTGACGTCG
GCCCCGATGGCAGACTGCTGAGAGGCCACGACCAGTACGCTTATGACGGCAAAGACTACATTGC
TCTCAACGAAGATCTCCGCAGCTGGACCGCTGCCGATACCGCTGCACAGATCACCCAGCGCAAA
TGGGAAGCCGCTAGAGAGGCTGAGCAGAGAAGGGCCTACCTCGAGGGCGAGTGTGTCGAATGG
CTCAGACGGTATCTGGAAAATGGCAAGGACAAGCTGGAAAGGGCCGACTCCCCCAAAGCACATG
TGACCCACCATCCACGCAGCAAAGGCGAAGTCACTCTCCGCTGTTGGGCACTCGGATTCTACCC
AGCTGATATTACACTGACTTGGCAGCTGAATGGGGAGGAACTGACTCAGGATATGGAACTGGTC
GAGACTCGCCCAGCCGGGGATGGAACTTTTCAGAAATGGGCCTCTGTGGTCGTCCCACTGGGAA
AAGAGCAGAATTATACCTGTCGCGTCTACCATGAGGGACTCCCAGAGCCACTCACTCTCAGATGG
GAACCCCCCCCATCTACCGATAGTTACATGGTCATTGTGGCAGTCCTCGGGGTGCTCGGTGCCA
TGGCCATTATCGGGGCTGTCGTGGCATTTGTCATGAAGAGAAGGCGCAACACAGGCGGGAAGG
GGGGGGGACTATGCACTGGCCCCTGGAAGCCAGTCCTCCGAGATGAGTCTCCGGGACTGTAAAG
CCTGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCC
TGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTA
GGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACA

ATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGACCGGTGCCTTGAAGCTTCGCCTTCTTC
TACATAAAACACACCCATCTGGAGCTACAGAGGCTTCATATGGGAAGAAACAGGAGGAGTCTGAG
ACTAAGCCCGAGGCTGAGAAGACAGATCCTGAGGAAATAGGCAAAGTCTCCCCTTTACAGATGA
GAGTCCTGCACTCAGGCTTGGCAGTGTGAGCCGCCCATTGCAGGTGAACAGAGCCTGGTCTCTG
TGGGATCCCTGTGGGGCTTGCAGGCCAGCGCCTCTGCTTTAAAGAGAAGCCTCTCTCCACTGCA
TCCCTAAGCGCTTGTGTCGCCATTGTATTCCCGGAAGTGACCTTTCTTCTAGAAGACTCTAGGGT
GTGACTTCTGAAGAGAAGAAGGAAGAGGAAGGGTGGAGGTTAGGAAACAGTGAGTCGGGCTTGT
GGGTCTCTCCTGGTGATCTGACAGCTTCTGGGTCAGAACTCGGAGTCACCACGACAAACTGCTCT
CTGTCCGCAGTACAGGGTTCAGGCAAAGTCTTGGTTGCCAGGCGGTGAGGTCAGGGGTGGGGA
AGCCCAGGGCTGGGGATTCCCCATCTCCTCAGTTTCACTTCTGCACCTAACCTGGGTCAGGTCCT
TCTGCCGGGACACTGATGACGCGCTGGCAGGTCTCACTATCATTGGGTGGCGAGATTCCAGGAG
CCAATCAGCGTCGCCGCGGACGCTGGTTATAAAGTCCACGCAACCCGCGGGACTCAGACACCC
GGGATCCCAGATGGGCGCCATGGCCCCTAGAACCCTGCTGCTGCTCCTGGCCGCTGCCCTGGC
CCCTACCCAGACCAGAGCCATTCAGAGGACTCCAAAAATCCAGGTCTACTCTAGACATCCTGCCG
AAAATGGGAAAAGCAATTTTCTGAATTGCTACGTCAGCGGGTTCCACCCATCTGACATTGAGGTC
GACCTGCTCAAGAACGGGGAACGGATTGAAAAGGTGGAACATTCTGACCTGAGCTTTTCTAAAGA
TTGGTCCTTCTATCTCCTGTATTATACTGAATTCACTCCAACCGAAAAAGACGAATATGCATGCCG
CGTGAACCATGTCACTCTGTCTCAGCCCAAAATCGTCAAGTGGGATCGGGATATGGGCGGAGGG
GGTTCTGGCGGTGGTGGATCTGGCGGGGGAGGTTCCGGCAGCCACTCCATGCGCTACTTTTACA
CAGCCATGAGCAGGCCCGGCAGGGGGGAACCACGGTTCATTGCAGTCGGCTATGTCGACGATA
CACAGTTCGTCCGCTTTGATAGCGACGCCGCTTCCCCCAGAATGGCCCCCAGGGCACCTTGGAT
TGAGCAGGAAGGTCCAGAATACTGGGATCGCGAGACACAGATCAGCAAGACCAACACCCAGACA
TACAGAGAGTCCCTGCGCAACCTGCGCGGCTACTATAATCAGTCTGAGGCTGGCTCCCATACCC
TGCAGAGGATGTATGGATGCGACGTCGGACCAGACGGCCGGCTGCTGAGGGGACACGATCAGA
GCGCCTATGATGGAAAAGATTATATCGCTCTGAATGAAGATCTGTCCTCTTGGACCGCAGCCGAC
ACAGCAGCCCAGATTACTCAGCGCAAGTGGGAGGCAGCCAGGGAAGCCGAGCAGTGGCGGGCC
TATCTGGAAGGACTGTGCGTGGAATGGCTGCGCCGGTATCTCGAGAATGGGAAAGAGACTCTCC
AGCGCGCCGATTCCCCTAAGGCCCATGTCACACATCATCCAAGATCTAAAGGCGAAGTGACACTC
AGATGCTGGGCACTCGGTTTTTATCCAGCAGACATCACTCTCACTTGGCAGCTGAATGGCGAAGA
ACTCACTCAGGACATGGAACTGGTCGAAACACGGCCAGCTGGGGACGGGACATTCCAGAAATGG
GCATCTGTCGTGGTGCCTCTGGGGAAGAACAGAACTATACTTGCCGAGTCTATCACGAAGGGC
TGCCAGAACCTCTGACTCTCAGATGGGAGCCACCACCTAGCACAGATAGTTATATGGTCATCGTC
GCCGTCCTGGGTGTCCTCGGGGCAATGGCAATTATTGGCGCCGTGGTCGCATTTGTGATGAAGC
GGAGGCGGAATACCGGCGGAAAGGGTGGCGACTATGCTCTCGCTCCCGGATCTCAGAGTTCCG
AGATGTCCCTCAGGGACTGCAAAGCATGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAG
GAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA
GCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGGTA
CCTCGGCGGAGCTCCGCCTTCTTCTACATAAAACACACCCATCTGGAGCTACAGAGGCTTCATAT
GGGAAGAAACAGGAGGAGTCTGAGACTAAGCCCGAGGCTGAGAAGACAGATCCTGAGGAAATAG
GCAAAGTCTCCCCTTTACAGATGAGAGTCCTGCACTCAGGCTTGGCAGTGTGAGCCGCCCATTG
CAGGTGAACAGAGCCTGGTCTCTGTGGGATCCCTGTGGGGCTTGCAGGCCAGCGCCTCTGCTTT
AAAGAGAAGCCTCTCTCCACTGCATCCCTAAGCGCTTGTGTCGCCATTGTATTCCCGGAAGTGAC
CTTTCTTCTAGAAGACTCTAGGGTGTGACTTCTGAAGAGAAGAAGGAAGAGGAAGGGTGGAGGTT
AGGAAACAGTGAGTCGGGCTTGTGGGTCTCTCCTGGTGATCTGACAGCTTCTGGGTCAGAACTC
GGAGTCACCACGACAAACTGCTCTCTGTCCGCAGTACAGGGTTCAGGCAAAGTCTTGGTTGCCA
GGCGGTGAGGTCAGGGGTGGGGAAGCCCAGGGCTGGGGATTCCCCATCTCCTCAGTTTCACTT
CTGCACCTAACCTGGGTCAGGTCCTTCTGCCGGGACACTGATGACGCGCTGGCAGGTCTCACTA
TCATTGGGTGGCGAGATTCCAGGAGCCAATCAGCGTCGCCGCGGACGCTGGTTATAAAGTCCAC
GCAACCCGCGGGACTCAGACACCCGGGATCCCAGATGGGCGCCATGGCCCCTAGAACCCTGCT
GCTGCTCCTGGCCGCTGCCCTGGCCCCTACCCAGACCAGAGCCATTCAGCGCACACCTAAGATC
CAGGTGTACTCTCGGCATCCCGCCGAGAACGGGAAGTCTAATTTCCTCAATTGCTACGTGTCAGG
GTTCCATCCATCCGACATCGAAGTGGATCTGCTCAAGAACGGGGAAAGGATCGAGAAAGTCGAA
CACAGTGATCTGTCATTCTCTAAAGACTGGTCTTTCTACCTCCTCTATTACACTGAGTTCACACCTA
CTGAGAAAGATGAATACGCCTGCCGGGTCAACCACGTCACACTCAGTCAGCCAAAGATCGTGAA
ATGGGATAGGGATATGGGTGGCGGAGGCTCCGGCGGAGGCGGATCAGGGGGAGGCGGATCTG
GCAGCCATTCTATGAGATACTTCAGCACAAGCGTGTCCTGGCCAGGGCGAGGCGAGCCACGCTT
TATTGCAGTGGGTTATGTCGATGACACCCAGTTTGTCCGATTTGATTCAGACGCTGCCAGCCCCA
GGGGCGAGCCTAGAGAACCTTGGGTGGAACAGGAAGGTCCTGAGTACTGGGATAGAGAAACCC
AGAAGTACAAGAGACAGGCCCAGGCTGATAGAGTGAACCTGAGAAAGCTGCGCGGGTATTATAA
CCAGAGTGAAGATGGCTCTCACACCCTCCAGCGGATGTTCGGCTGTGACCTGGGCCCAGACGGT

AGACTGCTGCGGGGCTACAACCAGTTCGCCTATGACGGGAAGGATTACATTGCACTGAACGAGG
ATCTGAGATCTTGGACCGCCGCAGACACCGCCGCTCAGATTACCCAGAGGAAATGGGAGGCTGC
TCGCGAAGCCGAACAGCGCAGAGCCTATCTCGAGGGAACCTGCGTCGAATGGCTGAGAAGATAT
CTCGAGAACGGCAAAGAAACACTCCAGAGGGCTGACTCTCCAAAAGCTCATGTCACTCACCACC
CCCGCTCTAAGGGGGAAGTGACTCTCAGGTGTTGGGCTCTGGGGTTTTACCCAGCAGATATCAC
TCTGACCTGGCAGCTCAACGGGGAAGAACTCACCCAGGACATGGAACTCGTGGAAACTCGGCCC
GCAGGCGACGGGACTTTTCAGAAATGGGCAAGCGTCGTCGTGCCACTCGGCAAAGAGCAGAACT
ATACTTGCAGGGTCTACCATGAGGGCCTGCCTGAACCACTGACACTCCGATGGGAACCACCTCC
TTCAACCGACTCTTATATGGTCATTGTGGCTGTCCTGGGCGTGCTGGGGGCTATGGCTATTATTG
GTGCAGTCGTGGCTTTCGTCATGAAGCGGCGACGGAATACTGGGGGCAAGGGGGGAGACTACG
CCCTCGCACCTGGTTCTCAGTCCTCTGAAATGTCTCTCCGCGACTGCAAAGCTTGACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACT
CCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATT
CTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGC
TGGGGATGCGGTGGGCTCTATGGACTAGTCGCCTTCTTCTACATAAAACACACCCATCTGGAGCT
ACAGAGGCTTCATATGGGAAGAAACAGGAGGAGTCTGAGACTAAGCCCGAGGCTGAGAAGACAG
ATCCTGAGGAAATAGGCAAAGTCTCCCCTTTACAGATGAGAGTCCTGCACTCAGGCTTGGCAGTG
TGAGCCGCCCATTGCAGGTGAACAGAGCCTGGTCTCTGTGGGATCCCTGTGGGGCTTGCAGGC
CAGCGCCTCTGCTTTAAAGAGAAGCCTCTCTCCACTGCATCCCTAAGCGCTTGTGTCGCCATTGT
ATTCCCGGAAGTGACCTTTCTTCTAGAAGACTCTAGGGTGTGACTTCTGAAGAGAAGAAGGAAGA
GGAAGGGTGGAGGTTAGGAAACAGTGAGTCGGGCTTGTGGGTCTCTCCTGGTGATCTGACAGCT
TCTGGGTCAGAACTCGGAGTCACCACGACAAACTGCTCTCTGTCCGCAGTACAGGGTTCAGGCA
AAGTCTTGGTTGCCAGGCGGTGAGGTCAGGGGTGGGGAAGCCCAGGGCTGGGGATTCCCCATC
TCCTCAGTTTCACTTCTGCACCTAACCTGGGTCAGGTCCTTCTGCCGGGACACTGATGACGCGCT
GGCAGGTCTCACTATCATTGGGTGGCGAGATTCCAGGAGCCAATCAGCGTCGCCGCGGACGCT
GGTTATAAAGTCCACGCAACCCGCGGGACTCAGACACCCGGGATCCCAGATGGGCGCCATGGC
CCCTAGAACCCTGCTGCTGCTCCTGGCCGCTGCCCTGGCCCCTACCCAGACCAGAGCCATTCAG
AGAACACCAAAGATTCAGGTGTACTCAAGGCACCCAGCCGAAAACGGCAAGTCAAACTTTCTGAA
TTGCTATGTGTCTGGGTTTCATCCTAGCGATATCGAGGTGGACCTCCTCAAAAATGGGGAGAGGA
TTGAGAAAGTGGAACACTCCGATCTCAGTTTCAGTAAAGATTGGAGTTTCTATCTGCTCTACTATA
CCGAATTCACACCAACTGAAAAGGACGAGTACGCCTGTAGGGTCAACCATGTCACTCTGAGCCA
GCCAAAGATTGTGAAGTGGGATAGGGACATGGGGGGTGGTGGCAGTGGTGGTGGTGGTTCAGG
CGGCGGTGGCAGCTGCTCCCACTCTATGCGGTATTTCGATACCGCCGTGTCTAGACCAGGCCGC
GGAGAACCTAGATTCATCAGCGTCGGCTATGTGGATGACACTCAGTTCGTCCGCTTTGACTCCGA
TGCTGCCTCCCCTAGAGGGGAGCCTAGGGCCCCTTGGGTCGAGCAGGAAGGCCCTGAATATTG
GGACAGAGAGACACAGAAGTATAAGCGGCAGGCTCAGGCAGACAGAGTGTCTCTGAGGAACCTC
AGAGGCTACTATAACCAGTCCGAGGACGGATCTCACACCCTGCAGAGAATGAGCGGCTGCGATC
TGGGCCCTGATGGACGGCTGCTCCGCGGCTACGACCAGAGCGCATACGATGGGAAGGACTATA
TCGCACTCAATGAGGATCTCCGCTCATGGACAGCTGCAGACACTGCTGCCCAGATCACACAGAG
GAAGCTCGAGGCCGCCAGAGCCGCTGAACAGCTGAGGGCATATCTCGAAGGCACTTGTGTCGA
GTGGCTCCGCCGATACCTCGAAAATGGCAAAGAAACTCTGCAGCGAGCTGATAGCCCAAAGGCA
CACGTCACCCATCACCCTCGCTCAAAGGGCGAAGTCACTCTGAGGTGCTGGGCCCTCGGGTTCT
ATCCTGCCGATATCACACTCACCTGGCAGCTCAATGGCGAGGAACTGACACAGGACATGGAACT
CGTCGAAACTAGACCCGCTGGGGATGGCACCTTTCAGAAGTGGGCCAGCGTCGTCGTCCCCCTG
GGGAAAGAACAGAATTACACTTGTCGCGTGTACCACGAAGGGCTCCCTGAACCACTCACACTGA
GATGGGAGCCACCCCCTTCAACTGACAGCTACATGGTCATCGTGGCTGTCCTCGGAGTCCTGGG
CGCCATGGCAATTATCGGAGCAGTCGTCGCCTTCGTGATGAAGCGACGCCGGAACACAGGTGGC
AAAGGTGGCGATTATGCACTGGCACCAGGCTCTCAGAGTAGCGAAATGTCACTGAGAGATTGCA
AAGCCTGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGA
CCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG
AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAA
GACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGTTAATTAAGCAGAGAGGATATGCT
CATCGTCTAAAGAACTACCCATTTTATTATATATTAGTCACGATATCTATAACAAGAAAATATATATA
TAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAA
GTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTATAGTTCAAAATCAGTGACA
CTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCA
CAGCGACGGATTCGCGCTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTAC
GCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACAC - 3'

12.2 PCR primers used to genotype AB*ab*.I mice

Primers	Sequence
Primers used for HLA-A*03:01 PCR reaction (product length - 227 bp)	
GP-A3-1	ACCGACAGAGTGGACCTGG
GP-A3-2	TCCCACTTTCTCTTGGTGATCTGG
Primers used for HLA-A*11:01 PCR reaction (product length - 235 bp)	
GP-A11-3	CAAAGCTCAGTCCCAGACAGAC
GP-A11-4	CCGCTTTGTAATCTGAGCAGC
Primers used for HLA-B*07:02 PCR reaction (product length - 286 bp)	
GP-B7-5	CCCAGATCTACAAGGCTCAGGCC
GP-B7-6	TAGGCCCTTCTCTGCTCAGC
Primers used for HLA-B*15:01 PCR reaction (product length - 288 bp)	
GP-B15-7	CAACACCCAGACATACAGAGAGTCC
GP-B15-8	ACGCACAGTCCTTCCAGATAGG
Primers used for HLA-C*04:01 PCR reaction (product length - 288 bp)	
GP-C4-9	CCCAGAAGTACAAGAGACAGGCC
GP-C4-10	GATAGGCTCTGCGCTGTTCG
Primers used for HLA-C*07:02 PCR reaction (product length - 211 bp)	
GP-C7-11	GAACCTAGATTCATCAGCGTCGG
GP-C7-12	CGGACTGGTTATAGTAGCCTCTGAG
Primers used for 18s rRNA - control PCR reaction (product length - 133 bp)	
18s rRNA For	GGCCGTTCTTAGTTGGTGGAGCG
18s rRNA Rev	CTGAACGCCACTTGTCCCTC
Primers (mixed) used for human TCR αß - control PCR reaction (product length - 221+691 bp)	
TCR Alpha forward + Beta forward	ATGGCAAAAACCAAGTGGAG + CACATTAGGCCAGGAGAAGC
TCR Alpha reverse + Beta reverse	TTTGCTTTGTGTCTGCATCC + CCTGCTTAGTGGCTGAGTGG