Gene Editing in T-cells and T-cell Targets

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

School of Medicine

Cardiff University

Angharad Lloyd

2016

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (Candidate) Date ……

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD

Signed …… (Candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (Candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (Candidate) Date ……

Acknowledgements

I would like to thank the MRC and Cardiff University for funding the work contained in this thesis.

To my supervisors, Prof Andrew Sewell and Dr Bruno Laugel, thank you for giving me the opportunity to undertake this interesting project and to develop as a scientist. Thank you for your advice, support and for proofreading my thesis.

A special thanks to Dr Anna Bulek, Dr Garry Dolton, Dr Mai Ping Tan, Dr Meriem Attaf and Dr John Bridgeman for their help and support, Dr Claudia Consoli (CBS) for QPCR advice, Dr Catherine Naseriyan (CBS) for cell sorting, Dr Amanda Redfern (CBS) for running the bioanalyser, Dr Nicholas Kent and Angela Marchbank (Bioscience) for running the HiSeq, Dr Thomas Connor (Bioscience) for analysing the whole genome sequencing data and Prof John Phillips and Dr Colin Farrell (UTAH University) for running and analysing the RNAseq data.

Thank you to all the past and present members of the 'T-cell group' for making my time here at Cardiff University so memorable. In particular Val, Katie, Mike and Sarah, thank you for the support and friendship in and out of the lab.

To Steve, thank you for showing a real interest and for caring about what I do. Thank you for all your support, encouragement and for putting up with my mood swings, my long nights in the lab and weekends spent reading and writing.

This journey would not have been possible if not for my parents who have been incredibly patient and supportive throughout the entire process, thank you for all the words of encouragement and all the support you have given me over the years.

Publications

Publications which have contributed to chapters within this thesis:

Angharad Lloyd*, Bruno Laugel*, Erin Meermeir, Michael D. Crowther, Thomas Connor, Garry Dolton, Scott R. Burrows, Marielle Gold, David M. Lewinsohn and Andrew K. Sewell. Engineering of isogenic cells deficient for MR-1 with a novel "all-in-one" CRISPR/Cas9 lentiviral system: tools to study microbial antigen processing and presentation to human MAIT cells. ***Journal of Immunology***, **197**, 971-82. (*equal contribution)

Other publications which I have contributed to during the course of my PhD:

Angharad Lloyd, Owen Vickery and Bruno Laugel. (2013). Beyond the antigen receptor: editing the genome of T-cells for cancer adoptive cellular therapies. ***Frontiers in T cell biology***. doi: 10.3389/fimmu.2013.00221

Garry Dolton, Katie Tungatt, **Angharad Lloyd**, Valentina Bianchi, Sarah M. Theaker, Andrew Trimby, Chris Holland, Ania Skowera, Marco Donia, Andrew Godkin, David K. Cole, Per Thor Straten, Mark Peakman, Inge Marie Svane and Andrew K. Sewell. (2015). More Tricks with Tetramers: A Practical Guide to Staining T cells with peptide-MHC Multimers. ***Immunology***. doi: 10.1111/imm.12499

Sarah M. Theaker, Cristina Rius, Alexander Greenshields-Watson, **Angharad Lloyd**, Andrew Trimby, Anna Fuller, John J. Miles, David K. Cole, Mark Peakman, Andrew K. Sewell and Garry Dolton. (2016). T-cell libraries allow simple parallel procurement of multiple peptide-specific human T-cell clones. ***Journal of immunological methods***. doi:10.1016/j.jim.2016.01.014

Valentina Bianchi, Anna Bulek, Anna Fuller, **Angharad Lloyd**, Pierre J. Rizkallah, Garry Dolton, Andrew K. Sewell and David K. Cole. (2016). A molecular switch abrogates gp100 TCR-targeting of a human melanoma antigen. ***Journal of Biological chemistry*** doi:10.1074/jbc.M115.707414

David K. Cole, Hugo van den Berg, **Angharad Lloyd**, Michael D Crowther, Konrad Beck, Julia Ekeruche-Makinde, John J. Miles, Anna M. Bulek, Garry Dolton, Andrea A.J. Schauenburg, Aaron Wall, Anna Fuller, Mathew Clement, Bruno Laugel, Pierre J. Rizkallah, Linda Wooldridge, and Andrew K. Sewell. (2016). Structural mechanism underpinning cross-reactivity of a CD8+ T-cell clone that recognises a peptide derived from human telomerase reverse transcriptase. ***Journal of Biological chemistry***. doi:10.1074/jbc.M116.741603

Table of Contents

Abbreviations

ACT: Adoptive cell therapy
APC: Antigen presenting cell
β2M: Beta 2 microglobulin
bp: Base pairs
C: Celsius
CD (number): Cluster of differentiation (number)
cDNA: Complementary deoxyribonucleic acid
cGy: Centi-grey
CMV: Cytomegalovirus
CRISPR: clustered regularly interspaced short palindromic repeats
D: Diversity TCR gene fragment
DMEM: Dulbecco modified eagle's minimal essential media
DMSO: Dimethyl sulphoxide
DNA: Deoxyribonucleic acid
DTT: Dithiothreitol
E. coli: *Escherichia coli*
ELISA: Enzyme-linked immunosorbent assay
FACS: Fluorescence activated cell sorting
FBS: Foetal bovine serum
FMO: Fluorescence minus one
g: gram
GFP: Green fluorescence protein
HEK: Human embryonic kidney
HIV: Human immunodeficiency virus
HLA: Human leukocyte antigen
HLA A2: HLA A*0201
h: hour
HRP: Horse radish peroxidise
IBs: Inclusion bodies
IL: Interleukin
IFN: Interferon
IPTG: Isopropyl β-D-1-thiogalactopyranoside
IRES: Internal ribosome entry site
IU: International units
J: Joining TCR gene fragment
kDa: Kilodalton
L: Litre
LB: Luria-Bertani
MAIT: Mucosal-associated invariant T-cells
m: Milli
M: Molar
MFI: Mean fluorescence intensity
mg: Milligram
MHCI: Class I major histocompatibility complex

min: minutes
mRNA: Messenger ribonucleic acid
n: Nano
ng: Nanogram
OD: Optical density
PBMC: Peripheral blood mononuclear cell
PBS: Phosphate Buffer Saline
PCR: Polymerase chain reaction
PD1: Programmed death receptor 1
PHA: Phytohaemagglutinin
pMHC: Peptide-MHC complex
rCD2: Rat CD2
RNA: Ribonucleic acid
RNAi: Ribonucleic acid interference
rpm: Revolutions per minute
RPMI: Roswell Park Memorial Institute medium
RT: Room temperature
SDS-PAGE: Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
shRNA: Short hairpin RNA
siRNA: Short interfering RNA
TALEN: Transcription activator like effector nuclease
TCR: T-cell receptor
TNF-α: Tumour necrosis factor α
TRAJ: T-cell receptor alpha joining gene
TRAV: T-cell receptor alpha variable gene
TRBD: T-cell receptor beta diversity gene
TRBJ: T-cell receptor beta joining gene
TRBV: T-cell receptor beta variable gene
µ: micro
µL: micro litre
V: Variable TCR gene fragment
V: Volts
VSV-G: Vesicular stomatitis virus glycoprotein
WPRE: Woodchuck Hepatitis Post-transcriptional Regulatory Element
WT: Wild-type
x g: G force or relative centrifugal force
ZF: Zinc finger
ZFN: Zinc finger nuclease
ZFP: Zinc finger protein

Fluorochromes

APCY: Allophycocyanin
PB: Pacific blue
PE: R-phycoerythrin
PE Cy7: R-phycoerythrin CyChrome 7
PerCP: Peridinin chlorophyll protein

List of Figures

List of tables

Abstract

Recent years have witnessed a rapid proliferation of gene editing in mammalian cells due to the increasing ease and reduced cost of targeted gene knockout. There has been much excitement about the prospect of engineering T-cells by gene editing in order to provide these cells with optimal attributes prior to adoptive cell therapy for cancer and autoimmune disease. I began by attempting to compare short hairpin RNA (shRNA) and zinc finger nuclease (ZFN) approaches using the CD8A gene as a target for proof of concept of gene editing in Molt3 cells. During the course of my studies the clustered regularly interspaced short palindromic repeats (CRISPR) mechanism for gene editing was discovered so I also included CRISPR/Cas9 in my studies. A direct comparison of the three gene editing tools indicated that the CRISPR/Cas9 system was superior in terms of ease, efficiency of knockout and cost.

As the use of gene editing tools increases there are concerns about the inherent risks associated with the use of nuclease based gene editing tools prior to cellular therapy. Expression of nucleases can lead to off target mutagenesis and malignancy. To circumvent this problem I generated a non-nuclease based gene silencing system using the CD8A zinc finger (ZF) fused to a Krüppel associated box (KRAB) repressor domain. The ZF-KRAB fusion resulted in effective silencing of the CD8A gene in both the Molt3 cell line and in primary CD8+ T-cells. Importantly, unlike CRISPR/Cas9 based gene editing, the ZF-KRAB fusion was small enough to be transferred in a single lentiviral vector with a TCR allowing simultaneous redirection of patient T-cell specificity and alteration of T-cell function in a single construct.

To improve the efficiency of gene editing with CRISPR/Cas9 I developed an ‘all in one’ CRISPR/Cas9 system which incorporated all elements of the CRISPR/Cas9 gene editing system in a single plasmid. The ‘all in one’ system was utilised to derive MHC-related protein 1 (MR1) deficient clones from the A549 lung carcinoma and THP-1 monocytic cell lines in order to study MR1 biology. Mucosal-associated invariant T-cell (MAIT) clones were not activated by MR1 deficient A549 or THP-1 clones infected with bacteria.

1 Introduction

1.1 The immune system

The immune system is fundamental for the clearance of the many pathogens to which humans are exposed on a daily basis. The immune response can be divided into two phases - the innate and adaptive immune response. Some pathogens are cleared by the innate immune system alone, as the innate immune system provides the initial rapid defence against a broad range of pathogens. In contrast, the adaptive immune system mounts a response which is specific to the pathogen. The adaptive immune system can take a few days to initiate a response. After a particular infectious agent has been encountered by the adaptive immune system a small portion of the pathogen-specific cells persist in secondary lymphoid organs as memory cells and are able to rapidly proliferate and respond to subsequent infections from the same pathogen. This process is known as immunological memory (Panum, 1847). This study is focused on the T-cell component of the adaptive immune system and therefore the innate immune system and the other arms of the adaptive immune system will not be discussed any further here.

1.1.1 T-cells

T-cells are derived from haematopoietic stem cells in the bone marrow which differentiate and mature in the thymus. During T-cell development, thymocytes rearrange germline gene segments termed variable (V), diversity (D) and joining (J) which are fused to a constant (C) domain by alternative splicing to form an antigen receptor molecule known as a T-cell receptor (TCR). The process of gene rearrangement is called VDJ recombination. Prior to T-cells entering the circulation they must survive the process of negative and positive selection in the thymus which eliminates self and non-reactive T-cells. A diverse T-cell receptor repertoire is fundamental for the successful clearance of pathogen and abnormal cells (Messaoudi et al. 2002). Deficiencies in these T-cell subsets lead to severe combined immunodeficiency (SCID). Broadly speaking, T-cells can be subdivided into conventional T-cells and unconventional T-cells and each of these subtypes will be described in turn.

1.1.1.1 **Conventional T-cells**

Conventional T-cells recognise peptides presented by classical major histocompatibility complex (MHC) molecules (Doherty & Zinkernagel 1975). Typically, conventional T-cells express a heterodimeric TCR consisting of an α and a β chain although some unconventional T-cells also express an αβ TCR. The TCR chain usage in conventional T-cells is broad and results in a diverse TCR repertoire (Reviewed in Nikolich-Žugich et al. 2004; Laydon et al. 2015). Conventional T-cells are located in peripheral blood, tissues and lymph nodes (Gaspari & Tyring 2008). Conventional T-cells can be divided into two subsets which are identifiable by the surface expression of either CD4 or CD8.

CD4+ T-cells

CD4+ T-cells recognise peptides 12-20 aa in length presented by MHC class II molecules (MHCII) (Murphy 2012). The CD4+ subset can further divided into two main subtypes - helper cells (Murphy 2012; Kara et al. 2014) and regulatory cells subsets (Reviewed by Sakaguchi et al. 2010). In addition, cytotoxic CD4+ T-cells have been identified (Marshall & Swain 2011). The helper subset can be further divided into several subsets, T_H1, T_H2, T_H17 and follicular helper cell (T_{FH}) (Kara et al. 2014). The main role of helper T-cells is to aid B cells in antibody production, affinity maturation and class switching and activate a wide range of cell types including CD8+ T-cells, neutrophils, macrophages, basophils and eosinophils to aid in pathogen clearance (Zhu & Paul 2010). T_{reg} cells are essential for inducing self-tolerance and preventing autoimmunity by suppressing the function of other immune cells (Sakaguchi et al. 2010). The exact mechanism of suppression is currently unknown in humans, however, in vitro experiments have shown that human T_{reg} cells may kill effector T-cells through granzyme secretion (Grossman et al. 2004), secretion of immunosuppressive cytokines such as TGF-β (Powrie et al. 1996) and IL-10 (Asseman et al. 1999), through consumption of IL-2 (Höfer et al. 2012), through interaction of Fas : FasL (Strauss et al. 2009) and by interaction of immune checkpoint receptors (Tai et al. 2012).

CD8+ T-cells

CD8+ T-cells express the CD8 co-receptor formed by a heterodimer of α and β CD8 chains. CD8+ T-cells recognise MHC class I (MHCI) molecules presenting peptides 8-13 amino acids (aa) in length (Wooldridge et al. 2010). Recognition of cognate antigen leads to activation of

naive CD8+ T-cells and differentiation into cytotoxic T-cells. CD8+ cytotoxic T-cells can be divided into two subsets – type 1 (Tc1) and type (Tc2) dependent on the cytokines produced and secreted by the cells. Both Tc1 and Tc2 subsets have high cytotoxic activity (Kemp et al. 2005). Tc1 produce and secrete IFN-γ and IL-2 (Kelso & Groves 1997). Tc2 produce and secrete IL-4, IL-5, IL-6, IL-10 (Kelso & Groves 1997) and moderate levels of IFNγ (Kemp et al. 2005). Cytotoxic T-cells are so called as they synthesise cytotoxic granules such as perforin and granzyme leading to lysis of target cells (de Saint Basile et al., 2010). The release of cytotoxic granules is highly regulated as they are secreted at the immunological synapse to ensure precise killing of the target cell (Griffiths et al., 2010). Perforin released by cytotoxic T-cells leads to the formation of pores on the cell surface of the target cell allowing the entry of other cytolytic granules into the cell cytosol (Pipkin & Lieberman, 2007) such as granzyme (Elemans et al. 2012). Cytotoxic T-cells also utilise the FAS : FASL pathway to eradicate cells (Elemans et al., 2012). Activated cytotoxic cells express the FAS ligand which induces apoptosis when it binds to FAS (CD95) on target cells leading to initiation of the caspase cascade (Magnusson & Vaux 1999; Waring & Mullbacher 1999; Murphy 2012).

1.1.1.2 **Unconventional T-cells**

Unconventional T-cells are so named because they are not MHC restricted. Some unconventional T-cells may not express TCRs formed from the conventional α and β chains, and alternatively these T-cells may express a TCR formed by γ and δ chains. Unconventional T-cells do not recognise peptides presented by MHCI or II but instead can recognise lipids, phosphoantigens, metabolites, MHC like stress ligands and altered peptides presented by MHC-like molecules (Reviewed by Godfrey et al. 2015). In contrast to conventional T-cells which express a diverse TCR repertoire due to a diverse range of gene segment usage, unconventional T-cells typically express TCRs with a limited diversity (either germline encoded, invariant or semi invariant) (Reviewed by Liuzzi et al., 2015) however, there are exceptions such as γδ T-cells which are extremely diverse. MHC-like molecules are largely invariant in the population and thus unconventional T-cells are not restricted to an individual donor whereas conventional T-cells are restricted to specific allelic variants of MHC molecules. Some unconventional T-cells are described as innate-like as they do not require antigen stimulation to develop effector functions (Ribot et al. 2009; Gold et al. 2013). Several unconventional T-cell subsets have been described in the literature, these

include γδ T-cells, germline encoded mycolyl reactive (GEM) T-cells, invariant NKT (iNKT) cells and mucosal-associated invariant T-cells (MAITS) (Reviewed by Godfrey et al. 2015). Table 1.1 describes the main unconventional T-cell subsets. The MAIT subset of unconventional T-cells will be discussed in greater detail in chapter 5.

Table 1.1. Types of unconventional T-cells

Cell type	Reactive against	Presentation molecule	Chain usage
GEMs	Glucose monomycolate	CD1b	TRAV 1-2 TRAJ9
MAITs	Riboflavin derivatives	MR1	TRAV 1-2 TRAJ12, 33 or 20 TRBV 6 or 20
γδ T-cells	IPP and HMBPP	Butyrophilin	TRDV2 and TRGV9
	EPCR, ULBP and MICA	MHC like stress ligands	TRDV1 and 5
	Undefined self lipids	CD1c	TRDV1
	α-GalCer, sulfatide and other lipids	CD1d	TRDV1 and 3
	Phosphatidylethanolamine from pollen	CD1a	TRDV1
HLA-E restricted T-cells	MHCI leader and CMV derived peptides	HLA-E	
NK T-cells **Type I**	A-GalCer and other lipids	CD1d	TRBV25, TRAV10 and TRAJ18
Type II	Lysophosphatidylcholine and other lipids	CD1d	
Other CD1 restricted T-cells	Lysophosphatidylcholine, Dideoxymycobactin and other lipids	CD1a	
	Glucose monomycolate, Sulfoglycolipid and other lipids	CD1b	
	Phosphomycoketides, Methyl-lysophosphatidic acid and other lipids	CD1c	

Table adapted from Godfrey et al. 2015.

1.2 TCR ligands

1.2.1 Classical MHC molecules

Conventional T-cells recognise antigens in the form of peptides presented by classical MHC molecules known as human leukocyte antigens (HLA) in humans. The human HLA locus is found on chromosome 6 at position 6p21.3 and encompasses several genes including the

classical MHCI molecules (HLA-A, B and C) and MHCII (HLA-DR, -DP and –DQ) (Traherne 2008; Shiina et al. 2009; Murphy 2012).

MHCI

MHCI is a heterodimeric protein formed from a 44 kDa heavy α chain and a 12 kDa β2-microglobulin (β2m) chain. The heavy α chain is polymorphic with many allelic variants across the population (Neefjes et al. 2011). Folding of the heavy α chain forms three distinct domains α1, α2 and α3. The α1 and α2 domains form the peptide binding groove and the α3 domain covalently bonds to the β2m (Bjorkman et al. 1987). MHCI molecules are expressed by almost all nucleated cells, they present short peptide fragments of 8-13aa derived from antigen present in the cytosol and nucleus (Wooldridge et al. 2010; Neefjes et al. 2011; Rist et al. 2013).

MHCII

The MHCII molecule is formed from a polymorphic α and β chain. Folding of the α and β chain leads to the formation of 2 distinct domains in each chain. The α1 and β1 domains form the peptide binding groove (Murphy 2012) which houses peptides 12 to 20aa in length (Godkin et al. 2001). MHCII molecules have a more limited cellular expression pattern than MHCI molecules and are only expressed on professional APCs (Neefjes et al. 2011) such as dendritic cells (Shin et al. 2006), macrophages and B-cells (Reviewed by Reith et al. 2005). MHCII present antigens derived from exogenous peptides (Roche & Furuta 2015). As MHCII does not form an essential part of thesis it will not be discussed further.

1.2.2 Non-classical MHC molecules

In contrast to the conventional T-cells, unconventional T-cells do not recognise peptide-MHC molecules. Instead, unconventional T-cells often recognise antigens presented by non-classical MHC like molecules or MHC like stress ligands. For completeness, two of the main non-classical MHC like molecules (HLA-E and CD1) will be discussed next in turn. The non-classical MHC class I-related molecule (MR1) will be discussed in greater detail in chapter 5.

HLA-E

The HLA-E molecule was identified in 1988 and shown to have a sequence with 63-65% homology to that of HLA-A, B and C (Koller et al. 1988). Like the classical MHCI molecules, HLA-E is formed form a heavy chain which consists of three distinct domains upon folding, covalent bonded to a β2m molecule (Kaiser et al. 2008). Unlike the classical MHCI molecules, HLA-E is invariant as only two alleles have been identified in the human population HLA-E HLA-E∗0101 (HLA-E^{107R}) and HLA-E∗0103 (HLA-E^{107G}) (Grimsley et al. 2002; Pietra et al. 2010). Low expression of HLA-E has been observed on a variety of cells (Pietra et al. 2010). HLA-E has been shown to present signal sequences from classical MHCI molecules (Lemberg et al. 2001), peptides from the leader sequence from CMV and glycoprotein UL40 (Tomasec et al. 2000) which are identical in sequence to sequences from various HLA-A and –Cw alleles, peptides derived from HIV (Nattermann et al. 2005) and *Salmonella enterica* (Salerno-Gonçalves et al. 2004) which bear a high homology to leader sequences from classical MHCI molecules. Presentation of the MHCI leader sequences is important for natural killer cells immunosurveillance through NKG2A/CD94 interaction with HLA-E (Kaiser et al. 2008).

CD1 molecules

The CD1 molecules are unlike the classical MHC molecules as they do not present peptides, instead they present lipids. The structure of the CD1 molecules is similar to that of the classical MHCI molecules as they are composed of a heavy chain with three distinct extracellular domains covalently bonded to a β2m. In humans, the CD1 molecules can be divided into either group 1 (CD1a, CD1b and CD1c) and group 2 (CD1d) (Reviewed by Barral & Brenner 2007). CD1a is expressed on skin resident dendritic cells (DCs) (Wollenberg et al. 1996), CD1b is expressed on migrating DCs (Olivier et al. 2013). CD1c is expressed on Langerhans cells, monocyte derived DCs (Sugita et al. 2000; Milne et al. 2015). CD1d is expressed by intestinal epithelial cells (Landau et al. 1991), B-cells (Lang et al. 2008), dermal and monocyte derived dendritic cells (Gerlini et al. 2001). CD1 group 1 restricted T-cells were first identified in the CD4-CD8- T-cell subset (Porcelli et al. 1989) and then more recently have been identified in the CD8+ (Rosat et al. 1999) and CD4+ (Sieling et al. 2000) T-cell populations. It is thought that these cells are clonally diverse that is they express a broad TCR chain usage (Reviewed by Barral & Brenner 2007). There appear to be two

subsets of cells that that recognise CD1d - CD1d-restricted iNKT-cells and CD1d-restricted diverse NKT-cells. The CD1d restricted iNKT-cells have a invariant TCR chain usage and the CD1d restricted diverse NKT-cells are less characterised but are thought to express a diverse TCR repertoire (Reviewed by Barral & Brenner 2007).

1.3 Developing gene editing tools

The interest in developing and utilising gene editing tools has vastly increased in recent years following the ease of creating genetically modified cells for experimental purposes and the success of genetic modified cells for therapy in recent publications. The discovery of zinc finger nucleases (ZFN), transcription activator like effectors nucleases (TALENs) and more recently the clustered regularly interspaced short palindromic repeats (CRISPR) fused to CRISPR associated proteins (Cas) system have been major developments in the field of genetic engineering; these methods have extended the capabilities of gene editing on from RNAi.

1.3.1 RNAi

RNAi is the term used to describe the inhibition of gene expression using RNA molecules. Typically, this process induces gene silencing by inducing the degradation of specific messenger (m)RNA sequences. Initial studies involving RNAi suggested that the gene silencing observed was due to co-suppression (Napoli et al. 1990), quelling (Romano & Macino 1992) and sense mRNA (Guo & Kemphues 1995). An explanation for the gene silencing was published in 1998 which elicited that the silencing observed in the previous studies was due to degradation of mRNA (Fire et al. 1998). Andrew Fire and Craig Mello received a Noble prize in Physiology or Medicine for their work on RNAi in *Caenorhabditis elegans* in 2006. Of particular interest in this thesis is short hairpin RNA (shRNA). shRNA will be discussed further in chapter 3.

1.3.2 Nuclease gene editing

The alternative to RNAi interference is nuclease gene editing. The three nuclease gene editing tools which will be discussed in this thesis can be characterised as either monomeric or dimeric nucleases. The two examples of dimeric nuclease gene editing tools are ZFN and

TALEN which consist of two binding domains and two corresponding copies of the Fok1 nuclease which require dimerisation to be activated and functional. More recently a monomeric nuclease system has been identified known as the CRISPR/Cas9 system. The CRISPR/Cas9 nuclease consists of a single Cas9 protein. The ZFN, TALEN and CRISPRCas9 system will be discussed in greater detail in chapter 3.

1.3.3 Features and principles of nuclease gene editing

The priniciple of gene editing is the same for ZFN, TALEN and CRISPR/Cas9. In each case a protein or RNA binds to the specific DNA sequence at the target site. Once bound the DNA is cleaved by the nuclease protein leading to a double strand break (DSB). The DSBs in DNA are repaired by two mechanisms either by homogous recombination (HR) or non-homologous end joining (NHEJ) (Sanjana et al. 2012). HR can correctly repair the DSBs in DNA by utilising the sister chromatid as a template. In contrast, NHEJ is error prone and can result in insertions or deletions at the DNA cleavage site (Mao et al. 2008; Kim et al. 2013). The principle of the nuclease gene editing tools is shown in Figure 1.1. The medical importance of genome editing with engineered nucleases was highlighted by *Nature Methods* in 2011 when it won method of the year (Nature methods, 2012).

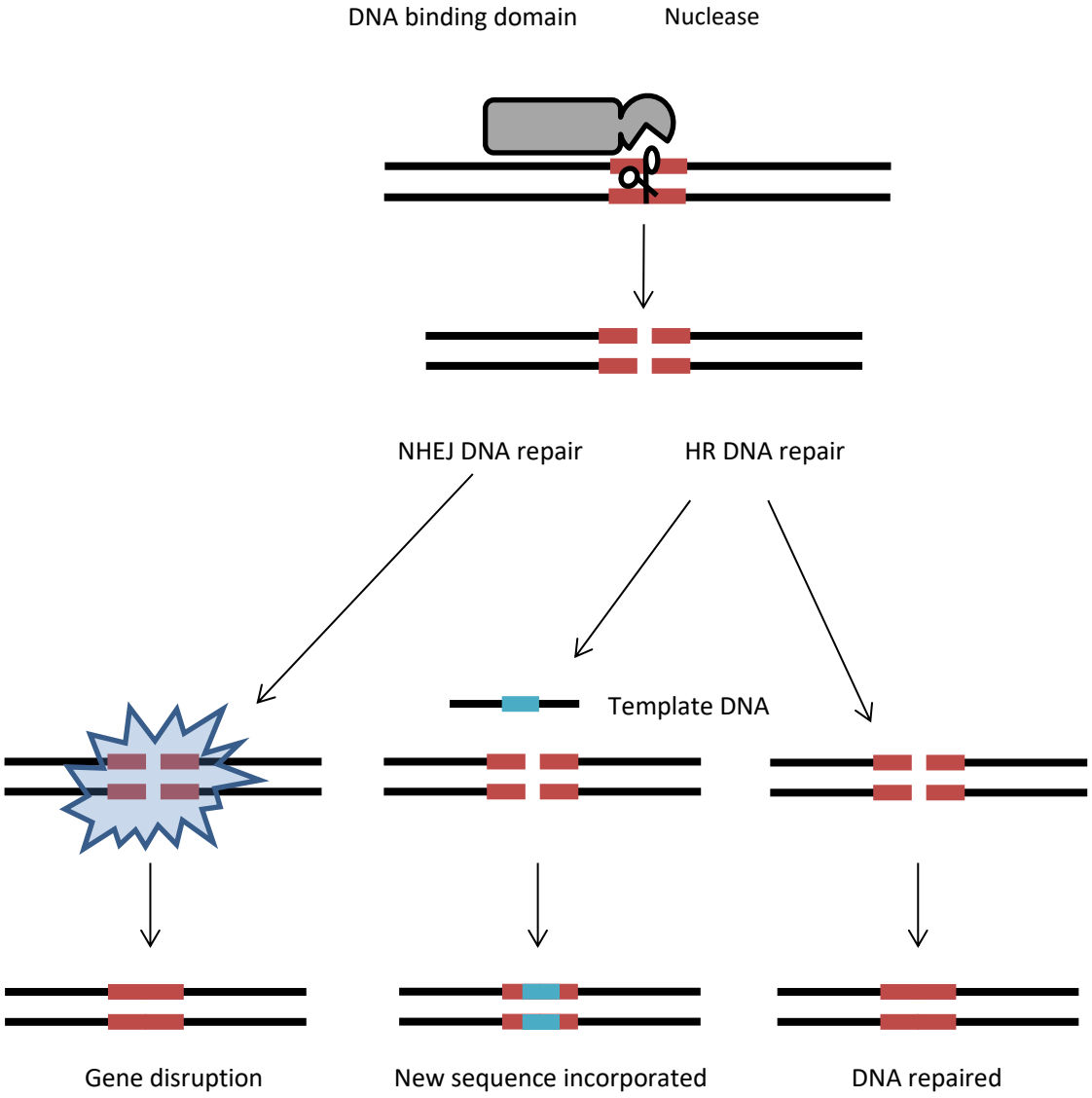


Figure 1.1. The main principles of nuclease gene editing. The DNA binding domain targets the DNA sequence of interest and the nuclease domain induces a double strand break. The DNA can be repaired by NHEJ which is error prone and commonly introduces small indels into the DNA sequence. Alternatively, DNA can be repaired by HR with or without introducing new template DNA which can lead to complete repair of the DSB.

1.3.4 Uses of gene editing tools

Several types of genetic modifications have been reported - gene disruption, gene insertion, point mutations, gene correction and structural chromosomal rearrangements. Gene disruption/ knockout rely on the introduction of small insertions and deletions (indels) by the NHEJ DNA repair pathway which can lead to frameshifts in the open reading frame of the protein. Frameshifts can lead to truncation or inactivation of the protein, which can result in a complete loss of gene expression if it occurs on both alleles. Gene insertion – A

gene editing nuclease can be co-transfected with a DNA sequence flanked with arms which are identical to the DNA near the target sequence this allows for this new sequence to be inserted into targeted areas of the genome (Li et al. 2011). Point mutations and gene correction – In a similar manner to gene insertion, gene editing tools can be co-transfected alongside targeting vectors or single stranded oligodeoxynucleotides to lead to the introduction of point mutations to repair mutated/ faulty genes (Soldner et al. 2011; Bialk et al. 2015). Finally, gene editing tools can be used to induce structural chromosomal rearrangements, such as duplications, deletions, inversions and translocations by introducing two DSBs at particular chromosomal locations (Brunet et al. 2009; Lee et al. 2010; Söllü et al. 2010; Lee et al. 2012).

1.3.5 Delivery of gene editing tools into target cells

The frequency of biallelic knockout of genes is affected by the transduction and transfection efficiency. Inefficient delivery decreases the frequency of biallelic inactivation. There are several methods used by researchers to induce expression of an exogenous protein into the cells such as mRNA (Kuhn et al. 2012), plasmid DNA (Luo & Saltzman 2000) or viral delivery (Bouard et al. 2009). In humans, lentiviral delivery has been utilised for the transfer of genes into T-cells such as TCR (Yang et al. 2008; Circosta et al. 2009) and ZFNs (Lombardo et al. 2007; Provasi et al. 2012). An integrase deficient lentiviral system has been developed which has a mutant integrase enzyme preventing uncontrolled DNA integration and thereby limiting cellular toxicity (Yanez-Munoz et al. 2006; Wanisch & Yáñez-Muñoz 2009). However, the efficiency of transduction is typically lower with transient methods therefore the integrating lentivirus system was used to deliver endogenous DNA into target cells for this project. In previous studies, genetically modified cells became immunogenic (Riddell et al. 1996), however, a more recent clinical trial with ZFNs targeting the CCR5 gene concluded that the treatment was considered safe and the genetically modified cells persisted in the patients suggesting that editing the cells in this manner did not lead to them becoming immunogenic (Tebas et al. 2014).

1.4 Aims

The work in this thesis aimed to expand the existing knowledge of gene editing tools and develop new tools that were safer and effective for gene silencing in primary T-cells. I set out to compare shRNA, ZFN, TALEN and CRISPR/Cas9 reagents that target CD8A as proof of principle. As there are currently concerns about the safety of nuclease based gene editing tools I also aimed to generate a non-nuclease based gene silencing system. Prior to my studies, the CRISPR/Cas9 system had relied upon several plasmids for the delivery of the CRISPR/Cas9 elements into the target cell for gene knockout. To improve the effectiveness of the CRISPR/Cas9 system, I also aimed to generate an 'all in one' CRISPR/Cas9 system that incorporated all of the CRISPR/Cas9 elements into a single plasmid, allowing for more efficient gene editing. As proof of principle this system was used to generate MR1 deficient APCs, in order to study MAIT / MR1 restricted T-cell biology in greater detail and to gain insight into antigen processing and presentation though MR1.

2 Materials and Methods

2.1 Bacterial cell culture

2.1.1 Culture medium and agars

All bacterial mediums and agars were supplemented with either 50 mg/ L carbenicillin (Carbenicillin Direct), 50 mg/ L Streptomycin (Sigma Aldrich) or 5 mg/ L Tetracycline HCl (Sigma Aldrich) for the selection of transformants**.**

Luria Bertani (LB) medium

10 g/ L NaCl (Fisher Scientific), 10 g/ L tryptone (Fisher Scientific) and 5 g/ L yeast extract (Fisher Scientific).

TYP medium

16 g/ L tryptone, 16 g/ L yeast extract and 5 g/ L potassium phosphate dibasic (Acros organics).

PSI broth

5 g/ L yeast extract, 20 g/ L tryptone and 5 g/ L Magnesium sulphate, (pH to 7.6 with potassium hydroxide).

Lemco culture media

5 g/ L Lemco powder (Oxoid), 10 g/ L tryptone, 5 g/ L NaCl and 0.5 mL/ L Tween 80 (Sigma-Aldrich).

LB agar

10 g/ L NaCl, 10g/ L tryptone, 5 g/ L yeast extract, 15 g/ L agar bacteriological (Oxoid).

TYP agar

16 g/ L tryptone, 16 g/ L yeast extract, 5 g/ L potassium phosphate dibasic and 15 g/ L agar bacteriological.

Lemco agar

5 g/ L Lemco powder, 10 g/ L tryptone, 5 g/ L NaCl and 15 g/ L agar bacteriological

Freezing buffer

75% LB medium and 25% glycerol.

2.1.2 Chemically competent cells

Esherichia coli XL10-Gold (Agilent) competent cells were used as host for large plasmid vectors and cultured in LB medium or on LB medium agar. BL21 Star (DE3) pLysS® (Thermo Fisher) competent cells were used for protein expression and cultured in TYP medium or on TYP agar.

2.1.2.1 **Buffers**

TfbI

30 mM CH_3COOK (Sigma Aldrich), 100 mM RbCl (Acros organics), 10 mM $CaCl_2$ (Sigma Aldrich), 50 mM $MnCl2 \cdot 4H_2O$ (Fisher scientific) and 15% w/v glycerol (Fisher Scientific). pH was adjusted to 5.8 with dilute acetic acid (Fisher Scientific).

TfbII

10 mM MOPS (Sigma Aldrich), 75 mM $CaCl_2$, 10 mM RbCl and 15% w/v glycerol. pH was adjusted to 6.5 with dilute sodium hydroxide (Fisher Scientific).

2.1.2.2 **Production of chemically competent cells**

E. coli XL10-Gold and BL21 Star (DE3) pLysS were produced in house by inoculating 1 mL of a 30 mL PSI broth starter culture in 100 mL of PSI broth and monitoring the optical density (OD) until it reached 0.45 at 600 nm (OD_{600}). At this stage the cells were placed on ice for 15 minutes (min), before being pelleted by centrifuging at 5465 x g for 10 min. The cells were resuspended in 40 mL of TfbI buffer and incubated on ice for a further 15 min. The cells were pelleted once more by centrifuging at 5465 x g for 10 min and finally suspended in 4 mL of TfbII, incubated on ice for 15 min and then snap frozen in 100 µL aliquots.

2.1.3 Transforming chemically competent cells

Transformation was carried out in accordance with the manufacturer's guidelines, briefly, 20 µL of chemically competent cells were thawed on ice, up to 100 ng (<5 µL) of DNA was added to the cells, after a 5 min incubation on ice the cells were heat shocked at 42 °C for 90 seconds (sec), then rested on ice for a further 5 min. Dependent on the initial concentration of DNA added and the antibiotic resistance the cells were immediately streaked onto agar plates or incubated in an orbital shaker for one hour (h) at 37 °C with 100 µL of SOC medium (Invitrogen) before streaking onto agar plates.

2.1.4 Bacterial cultures

Bacterial cultures of either 5 mL, 30 mL or 250 mL were produced by incubating a single colony from an agar plate into LB medium, TYP or PSI medium (with appropriate antibiotic) and incubating the culture in an orbital shaker at 37 °C and 220 rpm for 16 h.

2.1.5 *Mycobacterium smegmatis (M. smegmatis)* strain

The mc2155 strain of *M. smegmatis,* which grows quickly in synthetic media, was used in this thesis; it was kindly donated by Dr Matthias Eberl.

2.1.5.1 **Culture of *M. smegmatis***

To culture *M. smegmatis*, a loop of the frozen glycerol stock was streaked onto a Lemco agar plate and placed in an incubator at 37 °C for 48 h. Following this, a single colony was inoculated into 20 mL of Lemco media. The bacteria culture was incubated in an orbital shaker at 37 °C and 150 rpm for a minimum of 48 h. A growth curve was used to determine the concentration and phase of growth of bacteria at a given time point.

2.1.5.2 **Frozen stocks of *M. smegmatis***

M. smegmatis was cultured as described in section 2.1.5.1. After a minimum of 48 h the OD was determined and the relative number of CFU/ mL was calculated. *M. smegmatis* was then frozen in glycerol at a concentration of 10,000 CFU per vial.

2.2 Protein chemistry

2.2.1 Generation of expression plasmids

To produce the HLA-A*0201 protein (referred to as HLA A2 from now on) bound to peptide for tetramerisation and dextramerisation, the cDNA sequences encoding for the HLA A2 α chain (α1, α2 and α3 chain domains) and the β2m were inserted into separate pGMT7 plasmids (Banham & Smith 1993) as described by Boulter et al. 2003. The HLA A2 α chain (α1, α2 and α3 chain domains) and β2m were cloned into the pGMT7 plasmid downstream of an isopropyl β-D-1- thiogalactopyranoside (IPTG) inducible T7 promoter enabling maximal protein production by the introduction of IPTG during the log phase of *E.coli* growth. The HLA A2 a chain was tagged with a 15 amino acid (aa) biotinylation sequence (GLNDIFEAQKIEWHE) which allowed for the peptide-MHC (pMHC) monomer to be bound to a tetramer streptavidin or dextran streptavidin backbone for use as tetramers or dextramers. The full HLA A2 and β2m sequence is listed in Appendix Figure 1.

2.2.2 Peptides

Peptides with purity above 90% were used to produce monomers for tetramer and dextramer production. For protein refolding and T-cell activation assays crude peptides with a purity of between 50-60% were used. All peptides were synthesised by GL Biochem Ltd. (Shanghai, China) and Peptide Protein Research Ltd. (Hampshire, UK) respectively. Peptides were reconstituted in DMSO (Sigma Aldrich) at a concentration of 20 mg/ mL. For T-cell activation assays, working concentrations of peptide were prepared on the day of an experiment using R0 medium. The peptides used in this thesis were: The HLA-A2 restricted ELAGIGILTV peptide (designated as ELA) derived from Melan-A_{26-35} and recognised by the MEL5 TCR. The HLA-A2 restricted RQFGPDFPTI peptide (designated as RQF) which is a high affinity version of the ALWGPDPAAA preproinsulin$_{15-24}$ epitope recognised by the T-cells bearing the PPI TCR. The HLA- B*1801 restricted peptide SELEIKRY derived from EBV BZLF1$_{173-180}$

2.2.3 Protein expression as inclusion bodies

2.2.3.1 **Buffers**

Lysis buffer

10 mM TRIS (pH 8.1), 10 mM $MgCl_2$, 150 mM NaCl, 10% Glycerol.

Triton wash buffer

0.5% Triton X (Sigma Aldrich), 50 mM TRIS (pH 8.1), 100 mM NaCl and 10 mM EDTA.

Resuspension buffer

50 mM TRIS (pH 8.1), 100 mM NaCl and 10 mM EDTA.

Guanidine buffer

6 M guanidine, 50 mM TRIS (pH 8.1), 100 mM NaCl and 10 mM EDTA.

2.2.3.2 **Production of inclusion bodies (IB)**

The pGMT7 plasmid vectors containing the β2m and HLA-A2 α chain were transformed in BL21 Star (DE3) pLysS bacteria and cultured overnight at 37 °C on a TYP agar plate. A single colony starter culture was set up for the β2m and HLA-A*0201 α chain constructs and grown at 37 °C at 220 rpm in 30 mL of TYP medium overnight. All of the overnight starter culture was transferred into 1 L of TYP and incubated at 37 °C at 220 rpm and the OD was monitored until an OD of 0.5 at 600 nm (OD_{600}) was observed. The expression of the protein was induced by the addition of 0.5 mM IPTG (Fisher Scientific) and incubating the culture for three h in an orbital shaker at 37 °C at 220 rpm. The cells were harvested by centrifugation for 20 min at 3452 x g in a Legend RT centrifuge (Sorvall) with a Heraeus 6445 rotor, the cell pellet was resuspended in 40 mL of lysis buffer and sonicated using a MS73 probe (Bandelin) for 20 min using a 2 sec on 2 sec off programme at 60% power. The sample was kept on ice during sonication. The sample was treated with 200 µL of 20 mg/ mL DNAase (Sigma) and incubated for 60 min at room temperature (RT). The pellet was resuspended in 100 mL of Triton wash buffer and ultracentrifugation at 15,180 x g for 20 min at 4 °C in an Evolution RC

centrifuge with a SLA-1500 rotor (Sorvall). The pellet was resuspended in 100 mL of resuspension buffer and ultracentrifuged at 15,180 x g for 20 min at 4 °C in an Evolution RC centrifuge with a SLA-1500 rotor (Sorvall). The supernatant was discarded and the pellet was resuspended in 5 mL of guanidine buffer. The concentration of protein was determined using a NanoDrop ND1000 (Thermo Scientific) and calculated taking into account the reading at 280 nm wavelength and the extinction co-efficient. A 20 µL fraction of the suspension was taken at several time points (before IPTG treatment and 3 h after IPTG induction). These fractions were run on a Sodium dodecyl sulphate-polyacrylamide (SDS-PAGE) gel (section 2.2.6.) to assess the quality of the IB production.

2.2.4 Refolding of pMHCI

2.2.4.1 **Buffers**

MHC redox buffer

50 mM Tris (pH 8.1) (Fisher Scientific), 400 mM L-arginine (Sigma Aldrich), 2 mM EDTA (Fisher Scientific), 6 mM L-cysteamine (Sigma Aldrich) and 4 mM L-cystamine (Sigma Aldrich) in dH_2O.

Dialysis Buffer

10 mM Tris in dH_2O (pH 8.1).

2.2.4.2 **Refolding of pMHCI procedure**

30 mg of HLA-A2 α chain with a biotin tag inclusion bodies, 30 mg of β2m inclusion bodies and 4 mg of synthetic peptide were made up to 6.4 mL with guanidine buffer supplemented with dithiothreitol (DTT) (final concentration 10 mM) (Sigma Aldrich) and denatured at 37 °C for 15 min. The solution was then added to 1 L pre-chilled MHC redox buffer and incubated for 3 h at 4 °C on a magnetic stirrer. After which the refold solution was transferred to a 12 KD cut-off dialysis tube (Sigma Aldrich). The tube was placed in 20 L of dialysis buffer for 24 h and then placed into 20 L of dialysis buffer for a further 24 h until the conductivity of the

refold was below 2 mS/ cm. All buffers were filtered refolded protein samples were filtered through a 0.45 µm and 0.22 µm filter (Sartorius) before downstream purification steps.

2.2.5 Fast Protein Liquid Chromatography (FPLC) purification

2.2.5.1 **Buffers**

Buffer A

10 mM Tris in dH_2O (pH 8.1).

Buffer B

10 mM Tris (pH 8.1) and 1 M NaCl in dH_2O.

Monomer buffer

1x PBS in dH_2O.

2.2.5.2 **Fast Protein Liquid Chromatography (FPLC) purification procedure**

The refolded protein sample was purified using an anion exchange column (POROS® 50HQ, Life Technologies). Briefly, the column was equilibrated with buffer A and the refolded protein sample was loaded onto the column at a flow rate of 20 mL/ min with the alarm pressure set to 5 MPa. The sample was then eluted from the column by increasing the salt gradient and collecting the fractions in FPLC tubes (Greiner Bio-one). The fractions corresponding to the peak were collected and run on an SDS-PAGE gel (section 2.2.6) to assess the purity. The fractions of the refolded protein sample which were shown to contain MHC from the SDS-PAGE gel (section 2.2.6) were pooled and concentrated to a volume of 100 µL by centrifuging the sample at 1900 x g in a Vivaspin centrifugal concentrator (10 kDa molecular weight cut-off, Sartorius). The sample was made up to 700 µL with buffer A and the monomers were biotinylated by the addition of 100 µL Biomix A (Avadin), 100 µL Biomix B (Avadin), 100 µL d-Biotin 500 µM (Avadin) and 2 µL BirA enzyme (Avadin) and incubating the sample at RT overnight. To remove aggregates and biotin gel filtration was carried out using the Superdex HR 200 column (Amersham Pharmacia). The column was equilibrated in

monomer buffer by loading 40 mL of monomer buffer at a flow rate of 0.5 mL/ min with the alarm pressure set to 2 MPa. Once equilibrated the concentrated refold sample was loaded into the 2 mL loop and the purified sample was eluted into FPLC tubes. The fractions corresponding to the peak were run on a SDS-PAGE gel (section 2.2.6.) to assess purity and determine whether they contained a refold pMHCI protein. The monomers were stored as 5 µg aliquots at -80 °C until required.

2.2.6 SDS-PAGE electrophoresis

2.2.6.1 **Buffers**

5x non reducing loading buffer

125 mM Tris (pH 6.8), 4% SDS, 20% glycerol and 20 µg/ mL bromophenol blue.

5x reducing loading buffer

125 mM Tris (pH 6.8), 4% SDS, 20% glycerol, 20 µg/ mL bromophenol blue and 10% DTT.

2.2.6.2 **Running of an SDS-PAGE gel**

The quality and quantity of protein was determined by running a diluted sample of the protein on a SDS-PAGE gel. Each sample was prepared to be run on the SDS-PAGE gel by diluting in 5x non reducing loading buffer and 5x reducing loading buffer. The samples were incubated at 90 °C for 5 min and run on a 10% Bis/Tris gel in 1X NuPAGE running buffer (Invitrogen). A BLUeye, 10-245 kDa protein ladder (Geneflow) was run alongside the samples to allow for the size of the proteins to be determined. 10% Bis/Tris gels were run at 200 V for 25 min. The gel was removed from the plastic cast and stained with Quick Coomassie stain (Generon) heated in the microwave for 30 sec and then destained in dH_2O overnight.

2.3 Mammalian cell culture

2.3.1 Mammalian cell culture medium

All Mammalian cell culture medias were filtered through a 0.2 µm Stericup and Steritop Vacuum Driven Sterile Filter (Merck Millipore) using a Divac 1.2 L pump. Cells were cultured in 6, 24, 48, 96 multi well plates (Greiner Bio-One) or T25, T75 and T175 flasks (Greiner Bio-One). To ensure that cell lines were not contaminated with Mycoplasma, regular screening was carried out using the Mycoplasma kit (Lonza) manufacturer's guidelines.

R0

Roswell Park Memorial Institute medium 1640 (RPMI-1640) (Invitrogen) supplemented with 2 mM L-glutamine (Invitrogen), 100 units/ mL penicillin (Invitrogen) and 100 µg/ mL streptomycin (Invitrogen).

R5

RPMI-1640 medium supplemented with 5% FBS (fetal bovine serum) (Invitrogen), 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin.

R10

RPMI-1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin.

Antibiotic-free R10

RPMI-1640 medium supplemented with 10% FBS and 2 mM L-glutamine.

D10

Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS, 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin.

200 IU IL-2 T-cell medium

RPMI-1640 medium with 10% FBS, 2mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin, IL-15 (25 ng/ mL) (Peprotech Inc), IL-2 (200 IU/ mL) (Proleukin), non-essential amino acids (1x) (Life Technologies), Sodium pyruvate (1 mM) (Life Technologies) and HEPES buffer (1 mM) (Life Technologies).

200 IU IL-2 R20 T-cell media

RPMI-1640 medium with 20% FBS, 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin, IL-15 (25 ng/ mL), IL-2 (200 IU/ mL), non-essential amino acids (1x), Sodium pyruvate (1 mM) and HEPES buffer (1 mM).

20 IU IL-2 T-cell medium

RPMI-1640 medium with 10% FBS, 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin, IL-15 (25 ng/ mL), IL-2 (20 IU/ mL), non-essential amino acids (1x), Sodium pyruvate (1 mM) and HEPES buffer (1 mM).

20 IU IL-2 Priming T-cell medium

RPMI-1640 medium with 10% FBS, 2 mM L-glutamine, 100 units/ mL penicillin and 100 µg/ mL streptomycin, IL-2 (20 IU/ mL), non-essential amino acids (1x), Sodium pyruvate (1 mM) and HEPES buffer (1 mM).

2.3.2 Mammalian cell culture buffers

All mammalian cell culture buffers were sterilised using a 0.22 µm nitrocellulose filter (Millipore) before use.

Freezing medium

90% FBS and 10% DMSO (dimethyl sulfoxide).

Red cell lysis buffer

155 mM NH_4Cl (Fisher Scientific), 10 mM $KHCO_3$ (BDH), 0.1 mM Ethylenediaminetetraacetic acid (EDTA) (Sigma Aldrich) in dH_2O (pH adjusted to 7.2 to 7.4 with HCl (Fisher Scientific)).

Dissociation buffer

2 mM EDTA in PBS.

MACS buffer

DPBS pH 7.2 $CaCl_2^-$ and $MgCl_2^-$ (life technologies), 0.5% bovine serum albumin (BSA) (Sigma Aldrich) and 2 mM EDTA. The buffer was sterilised using a 0.2 µm Stericup before use.

2.3.3 Cell lines and culture conditions

Cells were split regularly before reaching a cell density of <90% or before the medium became yellow in colour. Suspension cell lines were split by suspending cells in fresh R10 medium at the optimal density. Adherent cell lines were split by removing the old medium and replacing the medium with either Hank's Cell Dissociation Buffer (Life Technologies) or dissociation buffer and incubating at 37 °C for 5 min. After which the cells were transferred to a 50 mL falcon tube and centrifuged at 558 x g, the cell pellet was resuspended in either R10 or D10 cell culture medium and seeded into fresh flasks at the optimal seeding density.

HEK 293T clone 17 cell line (ATCC product code CRL- 11268)

The 293T cell line is an adherent and highly transfectable cell line. 293Ts are derived from the 293 embryonic kidney cells which were transduced to stably express the SV40 large T antigen. This cell line was cultured in D10 media and maintained at 37 °C with 5% CO_2.

Molt3 (ATCC product code CRL- 1152)

The Molt3 cell line is a highly transfectable suspension hypertetraploid T lymphoblast cell line derived from a patient with relapsing acute lymphoblastic leukaemia. This cell line is cultured in R10 media and maintained at 37 °C with 5% CO_2

THP-1 (ATCC TIB 202)

The THP-1 cell line is a transfectable suspension cell line derived from a patient with acute monocytic leukaemia. This cell line is cultured in R10 media and maintained at 37 °C with 5% CO_2

A549 (ATCC CCL 185)

The A549 cell line is a transfectable adherent cell line derived from lung carcinoma. This cell line is cultured in R10 media and maintained at 37 °C with 5% CO_2. The A549 cell line has been HLA typed: A*2501/3001, B*18/4403 and C*1203/ 1601

HeLa (ATCC CCL-2)

The HeLa cell line is a transfectable adherent cell line derived from a patient with adenocarcinoma of the cervix. This cell line is cultured in R10 media and maintained at 37 °C with 5% CO_2.

MM909.24

The MM909.24 cell line is a tumour cell line derived from a tumour excised from a patient with metastatic melanoma which was kindly donated by I.M. Svane (Copenhagen University Hospital). This cell line is cultured in R10 media and maintained at 37 ˚C with 5% CO_2.

2.3.4 Culture of primary CD8+ T-cells

2.3.4.1 **Processing the peripheral blood mononuclear cells (PBMCs) from blood**

Buffy coats were ordered from the Welsh blood service in accordance with the human tissue act and appropriate ethical approval. The Welsh blood service tested and confirmed that each of the bags of buffy coats were seronegative for HIV-1, HBV and HCV before delivery. The buffy coats were separated by density gradient centrifugation by layering 25 mL the buffy coats onto 25 mL of Ficoll Hypaque (Lymphoprep) in sterile 50 mL falcon tubes and centrifugation for 20 min at 872 x g without brakes. The PBMC layer was aspirated from the density gradient using a pasteur pipette and the PBMC layer was transferred into a sterile 50 mL falcon tube. The volume was made up to 50 mL with R0 and the cells were centrifuged at 706 x g for 10 mins with the brake on. The supernatant was poured off and the cells were resuspended in 25 mL of red blood cell lysis buffer (section 2.3.2) and incubated in the water bath at 37 ˚C for 10 min. 25 mL of R10 was added to each tube to make the volume up to 50 mL and the cells were centrifuged at 314 x g for 6 min and the supernatant was poured off. The cells were resuspended in R10 media and counted.

Isolated PBMCs were either used as feeders to expand T-cells or to isolate CD8+ T-cells for lentiviral transduction.

2.3.4.2 **Irradiation of PBMCs**

An equal amount of PBMC from three donors were irradiated with Cesium-137 for 3100 centigrays (cGy) after irradiation the cells were recounted and then an equal number of cells from each donor were pooled together.

2.3.4.3 **T-cell expansion**

On day 0, 0.5 - 1 x 10^6 T-cells were transferred to a T25 flask in 15 mL of 20 IU IL-2 T-cell media containing 15 µg phytohaemagglutinin (PHA) (Alere) (1 µg/ mL final) and 1.5 x 10^7 irradiated PBMCs (section 2.3.4.2), 5 days post stimulation 7.5 mL of the medium was removed and replaced with 7.5 mL of fresh 20 IU T-cell media and incubated for a further 48 h, after which the cells were counted and plated in 24 well plates at a cell density of 3 - 4 x 10^6 cells per well in 2 mL of 200 IU IL-2 T-cell media.

2.3.4.4 **MACs separation of CD8+ T-cells from PBMCs**

PBMCs were isolated as described in section 2.3.4.1. CD8+ T-cells were purified from the PBMCs using anti-CD8 microbeads (Miltenyi Biotec) following the manufacturer's guidelines. Briefly, the cells were counted and centrifuged at 300 x g for 10 min, the supernatant was aspirated the pellet was resuspended in 80 µL of cold MACS buffer (section 2.3.2) per 1 x 10^7 cells. 20 µL of anti-CD8 microbeads (Miltenyi Biotec) was added per 1 x 10^7 cells and the cell suspension was incubated for 15 min in the refrigerator after which the cells were washed and resuspended in 500 µL of cold MACs buffer. Up to 2 x 10^8 cells were applied to a pre-rinsed MS column and allowed to drain by gravity, the column was washed in 500 µL of cold MACs buffer 3 times and finally the cells were eluted in 1 mL of cold MACs buffer. The cells were centrifuged at 300 x g for 10 mins and resuspended in 200 IU IL-2 T-cell medium and plated into 48 well plates at a density of 1 - 2 x 10^6 cells per well.

2.3.5 Cell counting by Trypan exclusion

Cells were diluted to an appropriate density in R0 medium. An equal volume of diluted cells and trypan blue (STEMCELL Technologies) were combined and mixed. The cells were loaded onto a haemocytometer (Weber Scientific International Ltd) and viable cells counted using a microscope to visualise the cells and a cell counter to record the number. The number of viable cells per mL was calculated as follows:

(Number of viable cells * Dilution factor) *10,000 = number of viable cells per mL

2.3.6 Cryopreservation of cell lines

2.3.6.1 **Cryopreservation of cell lines procedure**

Cells were centrifuged at 558 x g for 5 min, the supernatant was removed and the cell pellet was resuspended in 1 mL of freezing medium described in section 2.3.2. Cells were transferred to cryovials (Greiner bio-one) which were stored in Mr Frosty freezing pots (Thermo Scientific) at -80 ˚C for 48 h, for long term storage the cryovials were transferred to liquid nitrogen storage. Cell lines were typically frozen at a cell concentration of 1 - 2 x 10^6 cells/ mL and primary T-cells and T-cell clones were frozen at 2 - 6 x 10^6 cells/ mL.

2.3.6.2 **Thawing of cryopreserved cell lines**

Cryovials containing the frozen cells were thawed briefly in a water bath at 37 ˚C, once thawed the cells were transferred into a falcon tube containing 1 mL of R10 medium, the cells were pelleted by centrifugation (558 x g for 5 min), the supernatant was removed and the cell pellet was resuspended in appropriate medium before being transferred to a flask or plate for culture.

2.3.7 Generation of transduced and transfected cells

Cells were transduced by integrase proficient lentivirus or transfected transiently by $CaCl_2$ precipitation or lipofectamine.

2.3.7.1 **Buffers**

All of the buffers used were sterilised through a 0.22 µm nitrocellulose filter before use.

Buffered water

2.5 mM HEPES (Fisher Scientific) in dH_2O. Final pH 7.3.

$CaCl_2$

2.5 M $CaCl_2 2H_2O$ (Fluka Biochemika) in dH_2O.

HEPES-buffered saline 2×

0.28 M NaCl, 0.05 M HEPES and 1.5 mM anhydrous Na_2HPO_4 (BDH) in dH_2O. The pH was adjusted to 7 with 1 M NaOH (Fisher Scientific).

TE buffer

10 mM Tris (pH 8.1) (Fisher Scientific) and 1 mM EDTA with dH_2O. Final pH 8.0 with HCl.

2.3.7.2 **Lentivirus production**

2^{nd} and 3^{rd} generation lentivirus were used within this thesis, for the production of 2^{nd} generation lentiviruses a three plasmid system (pMD2.G, pCMV-dR8.74 and a pRRLSIN.cPPT.PGK-GFP.WPRE transfer vector containing hPGK internal promoter and target genes with or without a GFP reporter) was used. For the production of 3^{rd} generation lentiviruses a four plasmid system (pMD2.G, pMDLg/pRRE, pRSV-Rev and pELNSxv transfer vector containing EF-1 alpha promoter and target genes with or without a rCD2 reporter) was used. Prior to commencing lentiviral production 293T cells were seeded in a T175 flask, at a cell density of 2×10^6 cells in 20 mL D10 medium, the cells were at 60-80% confluency on the day of transfection. 4 h prior to transfection cell confluency was assessed; the medium was carefully removed and replaced with 20 mL of fresh D10 medium. Lentiviruses were produced as described previously (Mikkola et al. 2000). Briefly 1.1 mL TE, 550 µL buffered water, were combined with either 13 µg pMD2.G and 24 µg pCMV-dR8.74 or 15 µg pMD2.G, 15 µg pMDLg/pRRE and 15 µg pRSV-Rev per sample dependent on the generation of lentivirus being produced. The solutions were vortexed briefly to mixed, before the addition of 18.75 µg of the transfer vector containing the DNA of interest and 190 µL $CaCl_2$ mix. 1.9 mL of HEPES buffered saline was added in dropwise fashion whilst vortexing the solution. The solution was left at RT for 15 - 25 min to allow precipitates to form. After which the solution was slowly added to the media of the 293Ts in a dropwise fashion whilst

the flask was gently agitated to mix the solution into the medium. The cells were incubated overnight at 37 °C and 5% CO_2. 16 h after the transfection the medium was carefully removed and replaced with 17 mL of D10 medium, the cells were incubated for a further 48 h at 37 °C and 5% CO_2, the supernatant was collected at 24 h intervals (24 and 48 h), the supernatants were filtered through a 0.45 µm nitrocellulose filter (Millipore) and the 24 hr and 48 h supernatants were pooled and concentrated by ultracentrifugation at 150,146 x g for 2 h at 4 °C in sterilised ultra-clear ultracentrifuge tubes (Beckman Coulter). After centrifugation, the medium was discarded and the lentivirus pellet was resuspended in either 300 µL T-cell media for infecting T-cells or 1 mL R10 or D10 if infecting a cell line. The lentivirus preparations were aliquoted into cryovials and snap frozen on dry ice before transferring to -80 °C where they were stored until required.

2.3.7.3 **Transduction of cell lines**

Cells were plated at a concentration of 1 - 5 x 10^6 cells per well in a 24 well plate in 900 µL of appropriate medium and the cells were transduced with 100 µL of concentrated lentivirus. The cells were incubated overnight at 37 ˚C 5% CO_2, 16 h post transduction the media was removed from the cells and the cells were resuspended in cell appropriate cell culture medium.

2.3.7.4 **Transduction of CD8+ T-cells**

CD8+ T-cells were isolated from PBMCs as described in section 2.3.4.4, the CD8+ T-cells were activated with anti CD3 anti CD28 Dynabeads (Life technologies) at a ratio of 1 : 1 for 24 h and cultured at a concentration of 2 x 10^6 cells/ well in a 48 well plate (2 mL final volume). The following day 1.1 mL of medium was removed and 100 µL of concentrated lentivirus added to the cells in addition to Polybrene (Santa Cruz Biotechnology, CA) at a concentration of 500 µg/ mL. The cells were incubated overnight at 37 ˚C 5% CO_2, the following day 1 mL of 200 IU IL-2 R20 T-cell media was added to the cells. The beads were removed using a MACS magnet (Miltenyi Biotec) 4 days post CD8+ T-cell isolation.

2.3.8 Calcium chloride transfection of 293T cells

293T cells were plated in 6 well plates 24 h prior to $CaCl_2$ transfection, on the day of transfection the cells were 60 - 80% confluent. 5 µg of the plasmid of interest was made up to a volume of 250 µL with TE buffer, then 180 µL of HEPES buffered H_2O and 75 µL $CaCl_2$ was added to the solution. 500 µL of HEPES buffer saline was added in dropwise fashion whilst vortexing. The solution was left for 15 - 25 min for precipitates to form and then slowly added to the cells. The cells were incubated overnight at 37 °C and 5% CO_2. 16 h after the transfection the medium was carefully removed and replaced with fresh D10 medium.

2.3.9 Lipofectamine transfection

Cells were transfected using the lipofectamine 2000 reagent kits (Invitrogen). Briefly, the cells of interest were plated in 96 well plates 24 h prior to lipofectamine transfection, on the day of transfection the cells were 70 - 90% confluent. 2.5 µL of lipofectamine 2000 reagent was made up to 25 µL with opti-MEM medium (Thermo Fisher). In a separate tube 2.5 µg of plasmid DNA was made up to a volume of 125 µL with opti-MEM medium. The reagent mix and DNA mix were vortex briefly to mix and then combined at a 1 : 1 ratio. The solution was incubated at room temperature for 5 min and then 10 µL of the solution was added to the appropriate wells containing the cells of interest.

2.3.10 Flow cytometry analysis

The expression of cell surface markers and the antigen specificity of the TCR were monitored by antibody and pMHC tetramer/ dextramer staining. Flow cytometry data was collected on the FACSCalibur (BD Bioscience) or the FACSCanto II (BD Bioscience) flow cytometry machines. Cell sorting was carried out on the FACSAria (BD Bioscience) flow cytometry machine. All FACS data was analysed using FlowJo software (Treestar Inc).

2.3.10.1 **Buffers**

All buffers were filtered through a 0.22 µm nitrocellulose filter to sterilize.

FACS buffer

2% FBS and PBS (Life technologies).

Wash buffer

0.4% EDTA and PBS.

Cell fixing buffer

2% paraformaldehyde (PFA) and PBS.

2.3.10.2 **Surface marker antibodies**

For analysis by flow cytometry cells were stained with a combination of the antibodies listed in Table 2.1 and live dead markers in Table 2.2. To compensate for spectral overlap anti-mouse Igκ antibody capture beads (BD Biosciences) were used to prepare individual compensation tubes for each fluorochrome used in an experiment.

2.3.10.3 **Staining protocol**

Prior to commencing staining the cells were counted and an appropriate number of cells were washed with wash buffer in FACS tubes (Elkay) by centrifugation at 706 x g for 3 min at RT. The supernatant was discarded and the cells were stained with a dead stain dye (Vivid) diluted 1 : 40 in PBS for 5 min in the dark at RT. A cocktail of the cell surface markers with or without 7AAD was added to the cells and the samples were incubated for 20 min on ice in the dark. The cells were washed in 3 mL FACS buffer by centrifugation at 706 x g for 3 min at RT. The supernatant was discarded and the cells were resuspended in 200 µL of FACS buffer. Cells which were transduced with lentivirus and cells which were stained more than 1 hr prior to flow cytometry analyse required fixing (section 2.3.10.4).

Table 2.1.1. Cell surface marker antibodies

Marker	Colour	Brand	Clone	Isotype
CD8	APCY	Miltenyi Biotec	BW 135/80	IgG2a
	PE Vio770	Miltenyi Biotec	BW 135/80	IgG2a
CD3	PerCp	Miltenyi Biotec	BW 264/56	IgG2a
CD19	PB	Biolegend	HIB19	IgG1κ
CD14	PB	Biolegend	M5E2	IgG2a
Pan αβ TCR	APCY	Biolegend	IP26	IgG1κ
	PE	eBioscience	IP26	IgG1K
rat CD2	PE	Biolegend	OX-34	IgG2a
MR1	PE	Biolegend	26.5	IgG2a
Mouse IgG2a isotype control	PE	eBioscience	eBM2a	
HLA-A, B and C	APCY	Biolegend	W6/32	IgG2a

Table 2.2. Live / dead stain antibodies

Live dead stain	Colour	Brand
7AAD		eBioscience
Live dead vivid	Violet	Life Technologies

2.3.10.4 **Fixing cells**

After cell surface staining the cells were resuspended in cell fixing buffer (section 2.3.10.1) and incubated for 30 min on ice in the dark. The cells were washed in FACS buffer and resuspended in 200 µL of FACS buffer. The samples were stored in the dark at 4 °C for up to 48 h before being analysed on the flow cytometer.

2.3.11 Tetramer and dextramer staining

The cells of interest were counted and 5 x 10^5 – 1 x 10^6 cells were washed in wash buffer by centrifugation at 706 x g for 5 min at RT in a FACS tube.

2.3.11.1 **Treatment with Protein kinase inhibitor (PKI)**

Prior to tetramer or dextramer staining cells were treated with PKI. Briefly, a working concentration of 100 nM of PKI (Dasatinib, Axon Medchem) was prepared. The cells were resuspended in a volume of 50 µL wash buffer after washing and 50 µL of 100 mM PKI was added so that cells were treated with a final concentration of 50 nM PKI. The cells were incubated at 37 °C for 30 min before continuing to tetramer/ dextramer staining.

2.3.11.2 **Tetramer preparation**

Tetramer with a streptavidin (strept) APCY backbone was prepared by adding 3 µL of strept APCY backbone (Life Technologies) to 5 µg of biotinylated monomer (section 2.2.5) in 5 equal amounts of 0.6 µL at 20 min intervals whilst incubating on ice. Finally, 1 µL of protease inhibitor was added to the tetramer and the concentration was adjusted to a final concentration of 0.1 µg/ µL with PBS.

2.3.11.3 **Dextramer preparation**

Dextramers with a APCY backbone were prepared by adding 12.12 µL APCY dextran backbone (Immundex) per 1 µg of biotinylated monomer (section 2.2.5) and incubating the sample at RT in the dark for 30 min. 1 µL of protease inhibitor was added to the dextramer and the concentration was adjusted to a final concentration of 0.1 µg/ µL with dextramer buffer (Immundex).

2.3.11.4 **Tetramer/ dextramer staining protocol**

Prior to tetramer/ dextramer staining the cells were counted, washed in wash buffer and treated with PKI (optional) (section 2.3.11.1). Without a further wash step the cells were stained with 3 µL tetramer or 3 µL dextramer, the tubes were vortexed briefly and incubated in the dark on ice for 30 min. The cells were washed with 3 mL of wash buffer and centrifuged at 706 x g for 3 min at RT after which the supernatant was discarded. The cells were stained with live/dead and cells surface marker antibodies as described in section 2.3.10.4.

2.3.12 CD8+ T-cell effector function assays

2.3.12.1 **Peptide titration assay**

T cells were washed by the addition of excess R0 medium and centrifugation at 654 x g for 5 min. The supernatant was discarded and the cells were resuspended in R5 medium and incubated for a minimum of 18 h at 37 °C with 5% CO_2. The APC cell line was pulsed with the appropriate concentration of peptide for 2 h incubated at 37 °C with 5% CO_2 after which the APCs and T-cells were washed in R5 medium 3 times and resuspended in R5. T- cells and APCs were counted and plated at a ratio of 2 : 1 T-cells : APCs in a U bottomed 96 well plate

and incubated overnight at 37 °C with 5% CO_2. Cell supernatants were collected after 18 h and the production of cytokines was analysed by Enzyme-linked immunosorbent assay (ELISA) (section 2.3.13).

2.3.13 ELISA

2.3.13.1 **Buffers**

Wash buffer

0.05% Tween 20 (Sigma Aldrich), PBS tablets (1x) (Oxoid ^{TM}) to dH_2O.

Reagent diluent

10% BSA and PBS tablets (Oxoid) (10x) in dH_2O. A working 1x solution was made by diluting the 10x stock in the appropriate volume of dH_2O.

2.3.13.2 **ELISA procedure**

ELISAs were used for the detection of IFN-γ, TNFα and MIP-1β cytokine production using the IFN-γ, TNFα and CCL4 Duoset ELISA kits (R&D Systems). The kit was used in combination with the accessory reagents - wash buffer, reagent diluent (section 2.3.13.1) and strepavidin-HRP, reagent A (tetramethylbenzidine) (R&D Systems), reagent B (hydrogen peroxide) (R&D Systems) and stop solution (sulphuric acid) (R&D Systems) according to the manufacturer's guidelines. The plates were washed 3 times with 190 µL/ well of wash buffer using a plate washer III (tricontinent) plate washer between every step of the protocol, after washing the plates were carefully blotted to remove excess liquid before commencing with the next step of the protocol. Briefly, 96-well half well flat bottom ELISA microplates (Corning Costar) were coated with 50 µL of mouse anti-human IFN-γ, TNFα or MIP-1β capture antibody diluted in PBS (1.5 µg/ mL) and incubated overnight at RT. The plates were blocked with 150 µL of reagent diluent for a minimum of 1 h. The plate was washed 3 times and 50 µL of supernatant collected from an activation assay was added to the appropriate wells. Alongside the supernatant, 50 µL of reagent diluent or standard solution titrated from 1000 pg/ mL to 15.6 pg/ mL in reagent diluent were aliquoted into each well (in duplicate) in order to produce a standard curve allowing for the concentration of cytokine to be determined. The plate was incubated at RT for a further 75 min. The plate was washed 3

times and 50 µL of biotinylated goat anti-human IFN-γ, TNFα and MIP-1β detection antibody diluted in reagent diluent to a concentration of 50 ng/ mL was aliquoted into each well. The plates were incubated at room temperature for a further 75 mins. The plate was washed 3 times and 50 µL of strepavidin-horseradish peroxidase pre-diluted in reagent diluent according to manufacturer guidelines was aliquoted into each well. The plates were incubated for a further 20 min. Following 3 washes 50 µL of a 1 : 1 solution of reagent A and reagent B was added per well. The plates were incubated in the dark for up to 20 min, until there was an observable colour change, at which point 25 µL of stop solution (2 N sulphuric acid) was added per well. OD readings of plates were taken at 450 nm (Bio-rad iMark microplate reader) with correction set to 570 nm. The results were analysed with Excel (Microsoft) and GraphPad Prism 5 (GraphPad) software.

2.3.14 Enrichment of MR1 deficient THP-1 cells

To generate MR1 deficient THP-1 cells, THP-1 cells were first transduced with CRISPR/Cas9 MR1 crRNA A lentivirus. The bulk population was incubated with the D481A9 MAIT clone at an E : T ratio of 1 : 2 (T-cell : APCs) using 50,000 THP-1 cells to 25,000 D481A9 MAIT cells. The cells were incubated for 1 week in priming media (section 2.3.1) after which the media was replaced with R10 which allowed for the D481 A9 MAIT clone to die off whilst the THP-1 cell number expanded.

2.3.15 Enrichment of MR1 deficient MM909.24 cells

MR1 deficient MM909.24 tumour cells were generated by transducing the cells with CRISPR/Cas9 MR1 crRNA A lentivirus and incubating the bulk MM909.24 tumour cells with the MC.7.G5 MAIT clone which was identified and characterised by Michael Crowther (unpublished data). The MM909.24 deficient cells were enriched from the bulk population by incubation with the MC.7.G5 clone at an E : T ratio of 1 : 2 (T-cell : APCs) using 50,000 MM909.24 cells to 25,000 MC.7.G5 cells. The cells were incubated for 1 week in priming media after which the media was replace with R10 which allowed for the MM909.24 cell number to expand and the MC.7.G5 clone numbers to decline.

2.3.16 Assessing the function of MR1 deficient A549 and THP-1 cells by infection with *M. smegmatis*

The MAIT cell clones (D426B1 and D481A9) were washed with R0 medium and then rested in R5 medium overnight. The A549 and THP-1 cells (WT and MR1 knockouts) were washed with R0 and cultured in antibiotic-free R10 overnight. On the day of the experiment, bacteria density of the *M. smegmatis* culture was determined and the A549 and THP-1 cells were exposed to *M. smegmatis* at a multiplicity of infection (MOI) of 100 : 1 (bacteria to cells) in antibiotic-free R10 and incubated for 2 h in an incubator at 37 °C. The A549s and THP-1s were washed in R10 and the supernatant was removed and the cells were resuspended in R10. The cells were incubated for a further 2 h. To ensure all bacteria were killed the A549s and THP-1 cells were washed in R10 medium three times and finally resuspended in R5 medium. Control cells were mock-treated with R10 instead of *M. smegmatis*. The A549 and THP-1 cells were plated into U bottom 96 well plates at a density of 6 x 10^4 cells per well. The MAIT cells were plated into the appropriate wells at a density of 3 x 10^4 cells per well. The cells were cultured overnight and the supernatant was harvested after 16 h and assayed for MIP-1β, TNFα and IFN-γ ELISA (section 2.3.13).

2.3.17 Assessing the function of MR1 deficient MM909.24 cells with the MC.7.G5 clone

The MC.7.G5 T-cell clone was washed with R0 medium and then rested in R5 medium for 16 h. The WT and MR1-/- MM909.24 tumour cells were washed in R0, resuspended in R5 media and plated into U bottom 96 well plates at a density of 6 x 10^4 cells per well. The MC.7.G5 MAIT cells were washed in R0 media, resuspended in R5 media and plated into the appropriate wells at a density of 3 x 10^4 cells per well. The cells were cultured overnight and the supernatant was harvested after 16 h and assayed for MIP-1β and TNFα ELISA (section 2.3.13).

2.4 Molecular biology

2.4.1 Cloning by enzymatic digestion

During the course of these projects a variety of DNA constructs were generated. The constructs were generated by the digestion and ligation of inserts and vectors by restriction enzymes (Fast digest AgeI, BamHI, BsmBI, EcoRI, NdeI, NheI, NsiI, SalI, SpeI, XbaI and XhoI (Thermo Scientific)) (section 2.4.5). The restriction sites were introduced at DNA synthesis or by polymerase chain reaction (PCR). Plasmid maps for each of the constructs used in this thesis is shown in Figure 2.1. The cloning process for each of the constructs is described in Appendix Table 1. The inserts and vectors were ligated using T4 ligase (Thermo Scientific) (section 2.4.6). DNA was transformed into XL10-Gold *E. coli* bacteria (section 2.1.3). To confirm that the colonies contained plasmids with the correct DNA sequence, colony PCR was performed to check whether the inserts were the correct size before confirming the nucleotide sequence by Sanger sequencing and analysing data using APE software (A plasmid Editor, M. Wayne Davis).

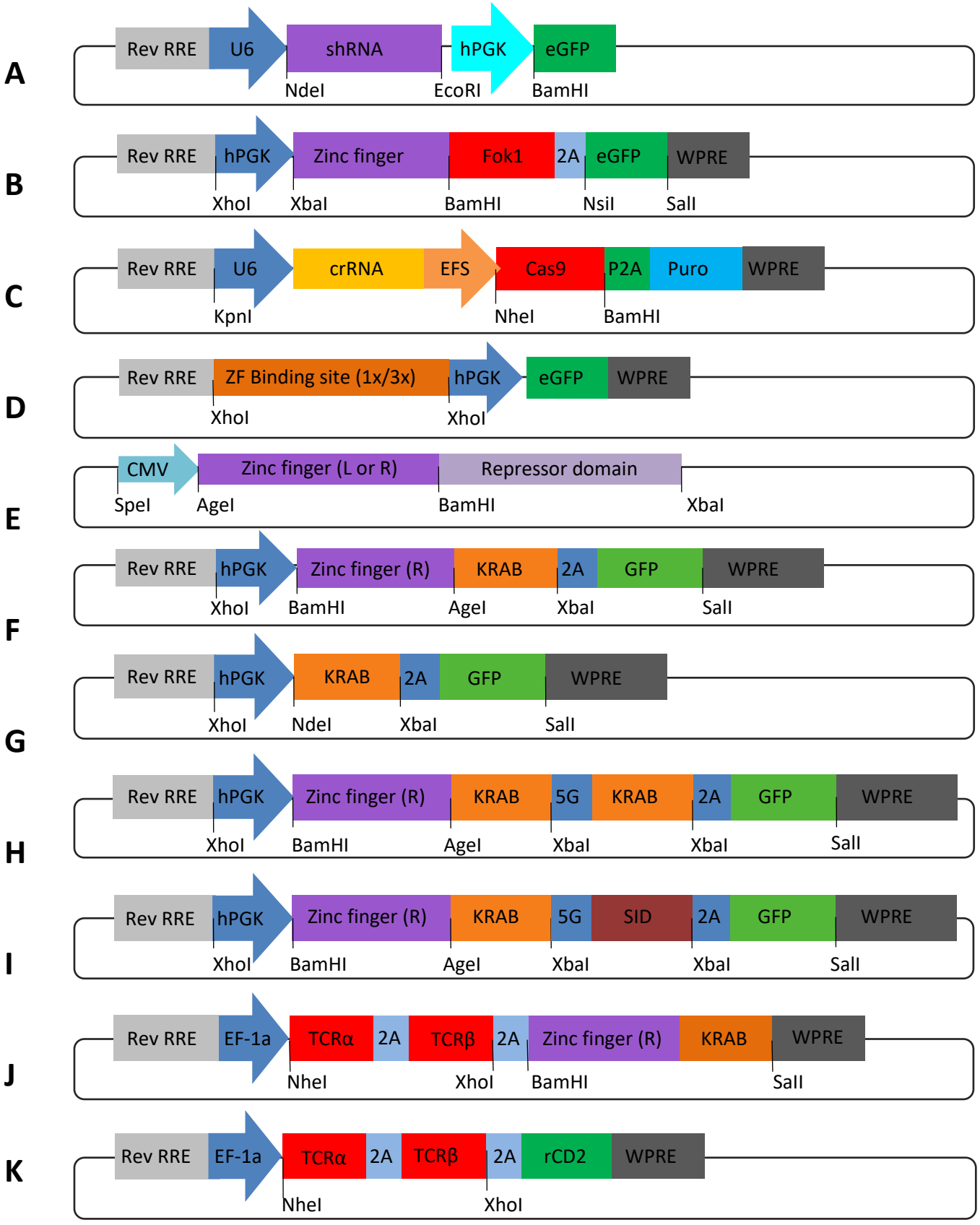


Figure 2.1. Plasmid maps of all plasmids used in this thesis. A) pLKO.1 shRNA GFP B) pRRL ZF Fok1 GFP C) crRNA CRISPRlentiV2 puro D) pRRL 1x/3x reporter system E) pV2 ZF repressor domain F) pRRL RZF-KRAB G) pRRL KRABx H) pRRL RZF-KRAB 5G KRAB I) pRRL RZF-KRAB 5G SID J) pELNSxv TCR RZF-KRAB K) pELNSxv TCR rCD2

2.4.2 Agarose gel production

2.4.2.1 **Buffers**

Electrophoresis loading buffer

A 6x loading buffer was made by combining 30% glycerol and 20 µg/ mL bromophenol blue in dH_2O.

50x Tris acetate EDTA (TAE) buffer

40 mM Tris (pH 8.1), 20 mM acetic acid and 1 mM EDTA in dH_2O.

5x Tris borate EDTA (TBE) buffer

90 mM Tris (pH 8.1), 90 mM boric acid and 2 mM EDTA in dH_2O.

2.4.3 Agarose gel electrophoresis

DNA was run on 1 - 2% agarose gel dependent on fragment size. 1 µL of 6x electrophoresis loading buffer was added to every 5 µL of DNA. Agarose (0.5 g for 1% gel or 1 g for a 2% gel) (Invitrogen) was dissolved in 50 mL Tris acetate EDTA buffer (or tris borate EDTA buffer for the Surveyor assay) by heating the solution in a microwave. Once the solution had cooled to 40 °C 2 µL of Midori green nucleic acid stain (Nippon genetics) was added and the gel was poured into a gel mold with the appropriate size combs. Agarose gels were run at 80 V for 20 - 40 min dependent on the fragment size. 5 µL of hyperladder 25 bp or hyperladder 1 kb (Bioline) was loaded alongside the DNA to determine fragment size. Agarose gels were visualised in a transilluminator (Modern Biology Inc).

2.4.4 DNA purification from agarose gel

Agarose gels were run as described in section 2.4.3, inserts and vectors were excised from the agarose gel and the DNA extracted using the Wizard SV Gel and PCR Clean-Up System (Promega) according to manufacturer's guidelines. Briefly, the excised gel was dissolved in 10 µL membrane binding solution per 10 mg of gel slice and incubated at 65 °C until the gel was dissolved. The solution was transferred to the mini column and incubated at RT for 1

min. The column was centrifuged at 16,000 x g for 1 min and the flow through was discarded. The column was washed with wash buffer twice (700 µL for 1 min at 16,000 x g, followed by a second wash 500 µL for 5 min at 16,000 x g). The flow through was discarded and the column was spun for a further 1 min at 16,000 x g to remove excess ethanol. The DNA was eluted from the column in 50 µL of dH_2O. The concentration of DNA was determined using a NanoDrop ND100 (Thermo Scientific).

2.4.5 Digestion with restriction enzymes

Restriction enzyme digestion was carried out by combining 2 µL of green fast digest buffer (Thermo Scientific), 0.5 µL of each restriction enzyme and up to 5 µg of DNA, the solution was made up to a final volume of 20 µL with nuclease free water. The reaction was briefly vortexed and centrifuged then incubated at 37 °C for 30 min and heat deactivated at 70 °C for 15 min before running the samples on an agarose gel (section 2.4.3). The appropriate DNA band was excised and the DNA extracted (section 2.4.4).

2.4.6 Ligation

Inserts were cloned into vectors by ligation with T4 ligase (New England Biolabs) following manufacturer's guidelines. Briefly, the concentration of DNA of the vector and inserts was determined using a NanoDrop ND1000. Inserts were cloned at a ratio of either 1 : 1 or 5 : 1, using 1 µL of T4 ligase, 2 µL of 10x T4 ligase buffer (New England Biolabs) and made up to a final volume of 20 µL with nuclease free water. The solution was incubated at room temperature for a minimum of 30 min before bacterial transformation (section 2.1.3).

2.4.7 Plasmid DNA extraction

Plasmid DNA was extracted from bacterial cultures with the use of miniprep kit (Sigma Aldrich) for 5 mL cultures or the PureLink HiPure plasmid filter purification kit (Invitrogen) for 250 mL cultures. DNA extractions were carried out following manufacturer's guidelines.

2.4.7.1 **Miniprep DNA extraction**

Briefly, the 5 mL culture was centrifuged at 12,000 x g for 1 min to pellet the *E.coli* cells, which were resuspended in 200 µL of resuspension buffer. The cells were lysed by the

addition of 200 µL of lysis buffer and incubated at RT for a maximum of 5 min, after which the solution was neutralised with 350 µL of precipitation buffer. The solution was centrifuged at 12,000 x g for 10 min to pellet the cell debris. The supernatant was transferred to a DNA column which had been treated with column preparation solution, the solution was loaded and centrifuged at 12,000 x g for 1 min. The flow through was discarded before washing the column 750 µL wash buffer and centrifugation at 12,000 x g for 1 min. The flow through was discarded and the column was centrifuged at 12,000 x g for 2 min to remove excess ethanol. The DNA was eluted in 100 µL of dH_2O. The concentration of DNA was determined using a NanoDrop ND1000.

2.4.7.2 **Maxiprep DNA extraction**

Briefly, the maxiprep column was prepared by the addition of 30 mL of equilibration buffer and allowing the solution to drain by gravity flow. The 250 mL *E.coli* culture was centrifuged at 4,000 x g to pellet the *E.coli*. The pellet was resuspended in 10 mL of resuspension buffer, the *E.coli* were lysed by the addition of 10 mL of lysis buffer, the falcon tube was inverted to ensure homogeneous lysis and incubated at RT for a maximum of 5 min. The solution was neutralised with 10 mL of precipitation buffer and loaded onto the maxiprep column and allowed to drain by gravity flow. The column was washed with 50 mL of wash buffer and the flow through drained by gravity flow. The DNA was eluted into a fresh 50 mL falcon tube containing 10.5 mL of isopropanol with 15 mL of elution buffer. The tube was centrifuged at 12,000 × g for 30 min at 4 °C and the supernatant was carefully removed. The pellet was resuspended in 5 mL of 70% ethanol and centrifuged at 12,000 × g for 5 min at 4 °C. The supernatant was poured off and the pellet allowed to air dry before being resuspended in an appropriate volume of dH_2O. The concentration of DNA was determined using a NanoDrop ND1000.

2.4.8 QPCR

2.4.8.1 **RNA extraction**

Up to 10 million cells were resuspended in 1 mL Trizol. 200 µL of chloroform was added and the sample was shaken vigorously for 15 seconds. The sample was incubated at RT for 3 min then centrifuged for 5 min at 12,000 x g at 4 ˚C. The aqueous phase was transferred to a new Eppendorf tube and 500 µL of 100% propanol-2-ol was added. The sample was

incubated at RT for 10 min. The sample was centrifuged for 10 min at 12,000 x g at 4 ˚C. The supernatant was discarded and 1 mL of 70% ethanol was added onto the pellet. The sample was vortexed briefly then centrifuged for 5 min at 7,500 x g at 4 ˚C. The supernatant was discarded and the pellet was allowed to air dry for 10 min before resuspension in RNAse free water and incubation at 55 ˚C for 15 min to dissolve the pellet. The concentration of RNA was determined on a NanoDrop ND1000 and the quality and integrity of RNA was determined on a bioanalyser by CBS staff.

2.4.8.2 **cDNA synthesis**

cDNA was produced utilising the Superscript 20 VILO cDNA synthesis kit following manufacturer's guidelines, briefly 4 µL of 5X VILO™ Reaction Mix was combined with 2 µL of 10X SuperScript® Enzyme Mix and up to 2.5 µg of RNA. The mixture was then made up to 20 µL using DEPC-treated water. The tube was gently mixed before being placed in a thermocycler for the following reaction programme:

Temperature	Time
25 ˚C	10 min
42 ˚C	120 min
85 ˚C	5 min

2.4.8.3 **SYBR green QPCR set up**

A master mix was prepared for each set of primers which consisted of 10 µL Power SYBR Green PCR Master Mix (Thermo Scientific), 0.3 µL forward primer (100 µM stock), 0.3 µL reverse primer (100 µM stock) and 4.4 µL H_2O per well. cDNA was synthesised as described in section 2.4.8.2, the cDNA was diluted into a final volume of 20 ng per 1 µL with dH_2O. 5 µL of diluted cDNA (100 ng) was aliquoted into each of the appropriate wells. Eppendorfs containing the master mixes and cDNA mixes were vortexed and centrifuged briefly. 15 µL of master mix was aliquoted into each of the appropriate wells. All QPCRs were set up in triplicate and run on a ViiA7 real time PCR system for the following PCR programme:

Stage	Temperature	Time	
Activation	95 ºC	10 min	
Denaturing	95 ºC	15 sec	x40 cycles
Annealing/ Extension	60 ºC	1min	

2.4.8.4 **TaqMan QPCR set up**

A master mix was prepared for each set of primers which consisted of 5 µL TaqMan Fast advance master mix (Thermo Scientific), 0.5 µL TaqMan gene expression assay and 2.5 µL dH_2O per well. cDNA was synthesised as described in section 2.4.8.2, the cDNA was diluted into a final volume of 25 ng per 1 µL with dH_2O. 2 µL of diluted cDNA (50 ng) was aliquoted into each of the appropriate wells. 8 µL of master mix was aliquoted into each of the appropriate wells. All QPCRs were set up in triplicate and then placed in a ViiA7 real time PCR system for the following PCR programme:

Stage	Temperature	Time	
Incubation	50 °C	120 sec	
Activation	90 °C	20 sec	
Denaturing	95 °C	1 sec	x40 cycles
Annealing/ Extension	60 °C	20 sec	

All QPCR data was analysed using Excel and GraphPad Prism 5 (GraphPad) software.

2.4.9 High fidelity PCR amplification

For PCR amplification of genomic DNA and plasmid cloning a high fidelity (HF) Phusion taq DNA polymerase (Thermo Scientific) was used to minimise PCR errors. The following reagents were combined in either 200 µL PCR tubes (Star labs) or 96 well PCR plates (Thermo Scientific). A master mix of all the reagents was used for large numbers of samples.

Component	Volume
DNA (5-100 ng)	variable
Phusion taq HF DNA polymerase	0.3 µL
Forward primer (100 µM)	0.1 µL
Reverse primer (100 µM)	0.1 µL
dNTPs (10 mM)	0.5 µL
HF Buffer	10 µL
H_2O	up to 50 µL
Final volume	50 µL

The PCR tubes/ plates were briefly centrifuged after set up and then placed in a thermocycler for the following PCR programme:

Stage	Temperature	Time	
Denaturation	98 °C	30 sec	
Denaturation	98 °C	10 sec	x 30 cycles
Annealing	XX °C	30 sec	
Extension	72 °C	30 sec per 1kb	
Final extension	72 °C	10 min	

2.4.10 Colony PCR

To check the size and orientation of inserts into plasmid vector was correct colony PCR was performed using a master mix of Green Dream Taq PCR master mix (Thermo Scientific) and primers, in the volumes listed below:

Component	Volume
Green Dream Taq PCR master mix	7.5 µL
Forward primer (100 µM)	0.05 µL
Reverse primer (100 µM)	0.05 µL
H_2O	up to 15 µL
Final volume	15 µL

The PCR tubes/ plates were briefly centrifuged after set up and then placed in a thermocycler for the following PCR programme:

Stage	Temperature	Time	
Denaturation	98 °C	1 min	
Denaturation	98 °C	30 sec	x 30 cycles
Annealing	XX °C	30 sec	
Extension	72 °C	1 min per 1kb	
Final extension	72 °C	5 min	

2.4.11 Genomic DNA PCR clean up

Genomic DNA was PCR amplified as described in section 2.4.9 and the size of the PCR amplification was confirmed by running 5 µL of the PCR product by gel electrophoresis (section 2.4.3). The PCR product was purified using PCR clean up beads (Axygen) following manufacturer's guidelines. Briefly, the PCR product was mixed with PCR clean up beads at a 5 : 9 ratio, the sample was mixed and incubated for 5 min at RT. The beads were placed on a magnet until a pellet was formed and the supernatant was discarded. Following this the sample was washed twice with 200 µL of 70% ethanol and the supernatant was discarded.

Finally the pellet was allowed to air dry for 10 - 15 min and the DNA was eluted from the pellet by suspending the beads in 40 µL of dH_2O.

2.4.12 Mammalian DNA extraction

Mammalian DNA was extracted from cells using a mammalian DNA extraction kit (Sigma Aldrich), briefly, cells were washed in PBS and lysed in 200 µL of Lysis Solution C. The samples were vortexed for 15 sec and then incubated at 55 °C for 10 min. To prepare the GenElute Miniprep Binding Column 500 µL of the Column Preparation Solution was added to the column which was centrifuged for 1 min at 12,000 x g. 200 µL of 100% ethanol was added to the cell lysate. The cell lysate was vortexed to mix and transferred onto the DNA binding column which was centrifuged for 1 min at 6500 x g. The flow through was discarded and 500 µL of wash Solution was added to the column, the column was centrifuged for 1 min at 6500 x g and the flow through was discarded. 500 µL of Wash Solution was added to the column which was centrifuged for a further 3 min at 16,000 x g, the through flow was discarded and the column was centrifuged for a further min at 16,000 x g to remove any existing ethanol from the column. To elute the DNA from the column 200 µL of the Elution Solution was added to the centre of the column. The column was centrifuged for 1 min at 26500 x g. The concentration was determined on the NanoDrop ND1000.

2.4.13 Surveyor assay

The Surveyor assay was carried out using the Surveyor Mutation Detection Kit (Integrated DNA Technologies) following the manufacturer's guidelines. The Surveyor assay is based on a nuclease which cleaves DNA at mismatch specific sites. Briefly, genomic DNA was extracted (section 2.4.12) from wild type (WT) cells and transfected or transduced cells. The DNA was amplified using primers which spanned the region thought to contain mutations induced by the gene editing tools. PCR amplicons derived from WT cells and transduced cells were hybridised using the thermocycler programme below.

Temperature	Time	Temperature ramp
95 °C	10 min	
95 °C to 85 °C		(–2.0 °C/ sec)
85 °C	1 min	
85 °C to 75 °C		(–0.3 °C/ sec)
75 °C	1 min	
75 °C to 65 °C		(–0.3 °C/ sec)
65 °C	1 min	
65 °C to 55 °C		(–0.3 °C/ sec)
55 °C	1 min	
55 °C to 45 °C		(–0.3 °C/ sec)
45 °C	1 min	
45 °C to 35 °C		(–0.3 °C/ sec)
35 °C	1 min	
35 °C to 25 °C		(–0.3 °C/ sec)
25 °C	1 min	
4 °C Hold	∞	

Once hybridised the samples were digested by combining the following reagents into 200 µL PCR tubes:

Reagent	Volume
Hybridised DNA (200 - 400 ng)	Variable
Surveyor enhancer S	1 µL
Surveyor nuclease S	1 µL
0.15 M $MgCl_2$	2 µL
dH_2O	Up to 20 µL
Final Volume	20 µL

The samples were incubated for 1 hr at 42 °C in a thermocycler. The reaction was stopped by the addition of 2.2 µL stop solution and the DNA was run on a 2% agarose TBE gel.

2.4.14 DNA extraction trizol method

For genomic DNA extraction from high cell numbers (>5 x 10^6) DNA was extracted using the trizol method. Briefly, up to 1 x 10^7 cells were resuspended in 1 mL Trizol and 960 µL of chloroform and 40 µL of isopropanol were added to each tube. The samples were vortexed for 60 sec and centrifuged for 5 min at 12,000 x g at 4 ˚C. The aqueous layer was transferred to a clean Eppendorf tube and sodium acetate was added (final concentration 300 mM). 2 -

2.5x volume of ice cold 100% ethanol was added and the samples were placed on dry ice for an h or in the freezer overnight. After which the samples were centrifuged for 5 min at 12,000 x g at 4 ˚C and the supernatant was discarded. 1 mL of RT 70% ethanol was added and the samples were centrifuged for 5 min at 12,000 x g at 4 ˚C, the supernatant was discarded and the pellet was air dry for 10 - 15 min. The pellet was resuspended in 100 µL H_2O and the concentration was determined by NanoDrop ND1000.

2.4.15 RNAseq

2.4.15.1 **RNAseq library preparations**

Prior to preparation of RNA libraries RNA was extracted and quality control was performed (section 2.4.8.1). Library preparation and RNAseq was performed at the University of UTAH, briefly, the RNA was purified with the Ribo-Zero Gold kit (Illumina) and RNA libraries were prepared using the TruSeq Stranded Total RNA Sample Prep kit (Illumina) following manufacturer's guidelines. The protocol is summarised in Figure 2.2. The samples were sequenced on an Illumina HiSeq using a HiSeq 125 Cycle Paired-End Sequencing v4 kit (Illumina).

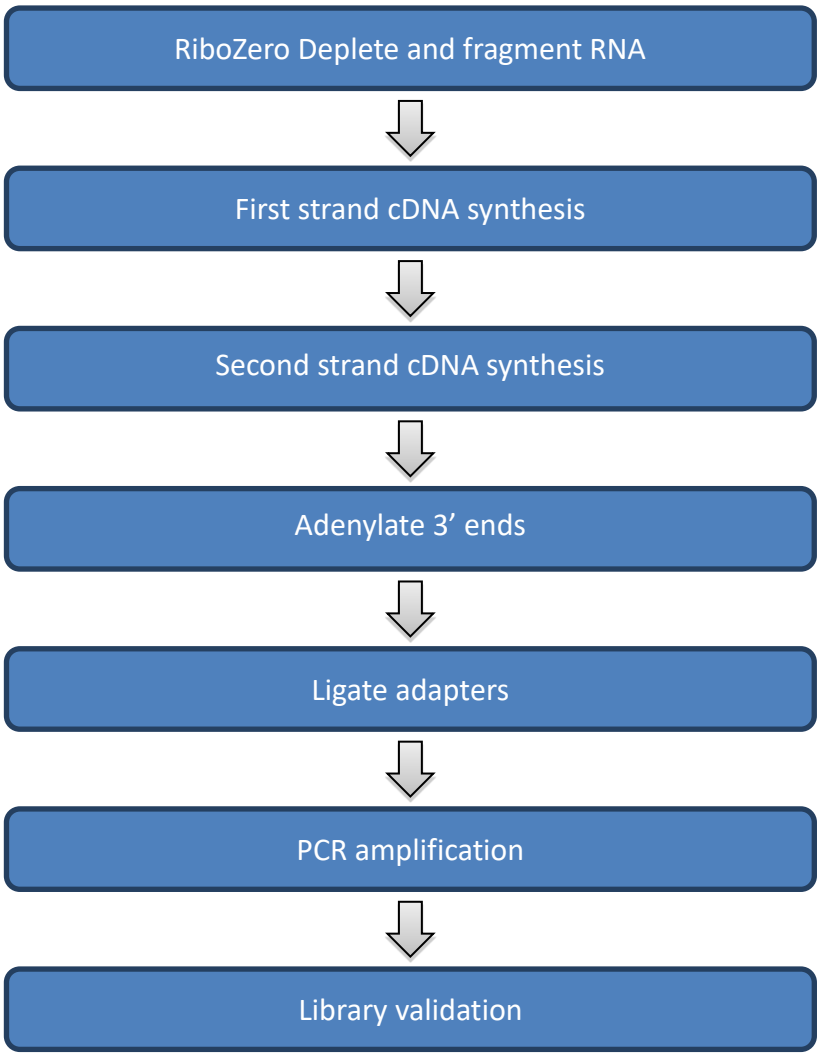


Figure 2.2. A schematic representation of the main steps in the library preparation protocol prepared prior to RNAseq

2.4.15.2 **Analysis of the RNAseq data**

The analysis of the RNAseq data was performed by Dr Colin Farrell (University of UTAH). Raw reads which did not meet quality control conditions were excluded from the analysis. The raw reads were aligned to the reference genome using Cufflinks software (Roberts et al. 2011). The differential expression of genes was determined using the Cufflinks pipeline (Roberts et al. 2011; Trapnell et al. 2012) and DeSeq2 package (Love et al. 2014) to perform pairwise comparisons between the samples.

2.4.16 Whole genome sequencing

2.4.16.1 **Next generation sequencing library preparations**

Prior to the preparation of DNA libraries DNA was extracted from cells (section 2.4.14) and the quality of DNA was determined using a bioanalyser (run by CSB staff). 1 µg of DNA was fragmented to an average size of 300 bp by a Covaris M220 focused ultrasonicator, which was confirmed by running the samples on a bioanalyser (run by CSB staff). Libraries were prepared using NEBNext Ultra library preparation kits (New England Biolabs), following manufacturer's guidelines, the protocol is summarised in Figure 2.3. A sample of the library preparations were run on a bioanalyser once more to check the fragment size of the adaptor ligated DNA fragments. To assess quality the whole genome library preps were initially run on the MiSeq (Illumina) using 250 bp paired end kit (Illumina). After the QC step the whole genome sequencing library preparations were run on a NextSeq 500, using High Output 150 bp paired end kit (Illumina) to a depth of >20x coverage for each genome.

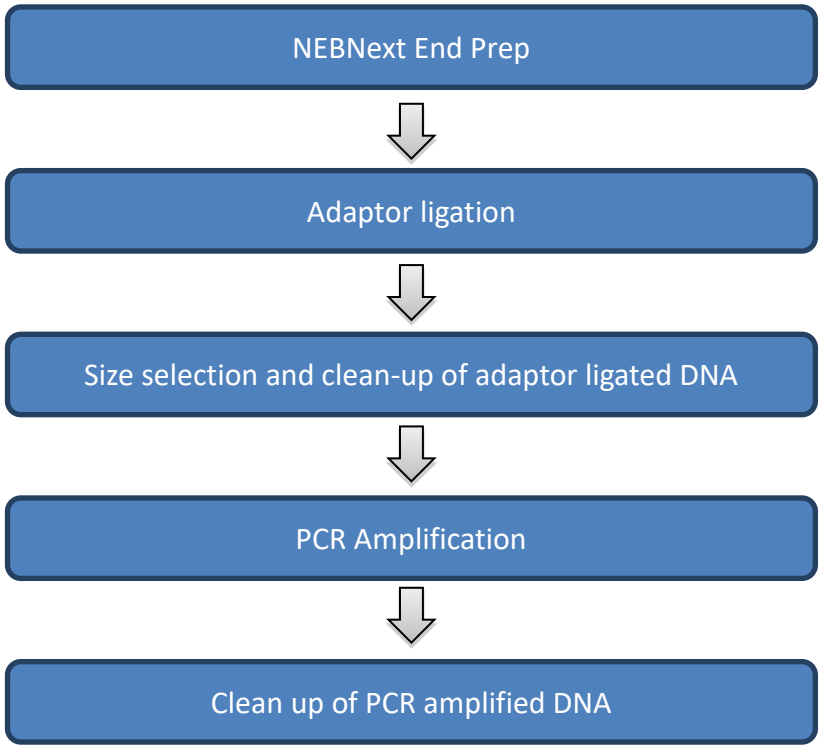


Figure 2.3. A schematic representation of the main steps in the library preparation protocol prepared prior to DNA sequencing.

2.4.16.2 **Analysis of the whole genome sequencing data**

The analysis of the whole genome sequencing data was performed by Dr Thomas Connor. Briefly, the reads were mapped against the DNA sequences using the Burrows Wheeler Alignment tool (Li & Durbin 2009) and Artemis software was used to visualise the aligned reads (Rutherford et al. 2000). The sequencing analysis for the data collected for the whole genome sequencing was completed in two phases:

1) The most likely off target sites in genes were identified using the CRISPRdesign tool (an algorithm developed at MIT) were screened.

2) Any reads containing a sequence with 15/20 bp or more matching the CRISPR RNA (crRNA) sequence were screened.

Approach 1

The most likely off target sites were identified were identified using the CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre), using these algorithms 37 potential off target sites and one unintentional on target site were identified for the crRNA A. Of the 38 sites identified, 1/38 of these predicted sites was within the RP11-46A10.6 gene, which encodes an unprocessed pseudogene, which was found to share 20/20bp homology within the crRNA. 5/38 of these predicted off target sites were within genes, the crRNA was found to share 17/20bp with the cAMP responsive element binding protein 3-like 1 (CREB3L1) gene, 16/20bp with the uncharacterized LOC728730 (LOC728730) long non-coding RNA gene, 16/20bp with tectonin beta-propeller repeat containing 1 (TECPR1) gene, 16/20bp with chromosome 16 open reading frame 59 (C16orf59) gene and 16/20bp with NLR family pyrin domain containing 1 (NLRP1) gene. Lastly, 32/38 of these predicted off target sites were not located within gene encoding DNA. The likely off target sites are shown in Table 2.3. The reads produced from the whole genome sequencing were aligned to the reference sequence GRCh38 and the 38 predicted off target sites were manually examined.

Approach 2

To analyse any off target mutations in the genome which were not identified by the CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre) algorithms, all of the sequence reads which contained a sequence matching, or complimentary to the guide RNA at 15/20 bases were extracted as a set of fastq files. The fastq files were

assembled to the reference sequence using Velvet. This analysis identified 7,574 locations in sample clone 9 and 7516 locations in clone 11 at which there were discrepancies between the reference sequence and the sequenced reads. Each query was examined manually by comparing the WT sequence to the sequenced reads.

Table 2.3. The most likely off target sites of MR1 crRNA A. The most likely off target sites of MR1 crRNA A were predicted using the CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre).Variations between the MR1 crRNA and other likely sites are highlighted in turquoise. PAM sequences are underlined.

	Sequence	Score	Mismatches	UCSC gene
MR1	GGATGGGATCCGAAACGCCCAGG	100.0	0MMs []	NM001195000
RP11-46A10.6	GGATGGGATCCGAAACGCCCAGG	100.0	0MMs []	
	CGAGGGGTTCAGAAACGCCCTGG	0.8	4MMs [1:4:8:11]	
	GGTGGGGATGCGAAACACCCTAG	0.5	4MMs [3:4:10:17]	
LOC728730	GCTTGGGACCCGAAACGCCTCGG	0.4	4MMs [2:3:9:20]	NR037875
	CTATGGGATCAGAAACGCTCAGG	0.3	4MMs [1:2:11:19]	
	AGATGGGATCCTAAAGGCCCAGG	0.3	3MMs [1:12:16]	
	GGATGGGATCAAAAACTCCCTGG	0.3	3MMs [11:12:17]	
	GGATGGGGGCGAAAACGCCCGGG	0.2	4MMs [8:9:11:12]	
	AGATGGGACCTGAAACACCCCAG	0.2	4MMs [1:9:11:17]	
	GAATGTGATCTGAAACTCCCCAG	0.2	4MMs [2:6:11:17]	
	GGAGAGGATCCTAAGGCGCCCTGG	0.2	4MMs [4:5:12:15]	
	GGCTGGGATGTGAAACGGCCAAG	0.2	4MMs [3:10:11:18]	
	GGATGGGATTAAAAACTCCCAAG	0.1	4MMs [10:11:12:17]	
TECPR1	GGTTGAGCTCCGACACGCCCAGG	0.1	4MMs [3:6:8:14]	NM015395
	GGGTGGCATCAGAAACGGCCAGG	0.1	4MMs [3:7:11:18]	
	GCATGGGTTCCAATACGCCCAGG	0.1	4MMs [2:8:12:14]	
	TGATGGGAGCAGAAATGCCCTGG	0.1	4MMs [1:9:11:16]	
	AGATGGGACCTGAAAGGCCCAAG	0.1	4MMs [1:9:11:16]	
	GGAAGGGAGCTGAAAGGCCCCAG	0.1	4MMs [4:9:11:16]	
	GGAAGGGATCTGCACCGCCCTGG	0.1	4MMs [4:11:13:15]	
	GGATGGGGCCTGAAAAGCCCAGG	0.1	4MMs [8:9:11:16]	
	GGATGGGTCCTGACACGCCCAAG	0.1	4MMs [8:9:11:14]	
	GGATGTGTTCCGAAACAGCCCGG	0.1	4MMs [6:8:17:18]	
	AGATGGGATCAGAAAAGCCATAG	0.1	4MMs [1:11:16:20]	
C16orf59	GGATGGGAGCCCGAACCCCCAGG	0.1	4MMs [9:12:13:17]	NM025108
	GGATGGGATGACAAAGGCCCTGG	0.1	4MMs [10:11:12:16]	
	GGAAGGGATCCTAAACATCCGGG	0.1	4MMs [4:12:17:18]	
	GGAAGGGATCAGCAAGGCCCGAG	0.1	4MMs [4:11:13:16]	
	GGATGGGATGAGAAAGGCCAGGG	0.1	4MMs [10:11:16:20]	
	GGATTGGATCCCAAAACCCCAAG	0.0	4MMs [5:12:16:17]	
	TGATGGGATCCCAAAGGGCCCAG	0.0	4MMs [1:12:16:18]	
	GGCTGGGATCCGAAGGCCCCAGG	0.0	4MMs [3:15:16:17]	
	GGAGGGGATCAGAGAGGCCCTGG	0.0	4MMs [4:11:14:16]	
	GGATGGGAGCCGACACCCACAAG	0.0	4MMs [9:14:17:19]	
	GGATGGGATCCAAACCCTCCCAG	0.0	4MMs [12:15:17:18]	
	GGATGAGATCCCAGAAGCCCAAG	0.0	4MMs [6:12:14:16]	
NLRP1	GGATGGGATCCGGAGGGCTTCTAG	0.0	4MMs [13:15:16:19]	NM033004
CREB3L1	GGATGGGATCAAAAACTCCCTGG	-	3MMs [11:12:17]	NM052854

2.4.17 Design of CRISPR/Cas9 systems for gene editing

Two CRISPR/Cas9 lentiviral vector plasmids were used in this thesis an 'all in one' system which was developed in house and the lentiCRISPRv2 vector which was developed by the Zhang lab. Each is described in turn:

2.4.17.1 **Development of an "all-in-one" CRISPR/Cas9 lentivector targeting the MR1 gene**

An 'all in one' CRISPR/Cas9 lentiviral vector was produced based upon the two plasmid system developed by Mali et al. 2013, which consisted of gRNA cloning vector (Addgene #41824) and hCas9 (Addgene #41815). As it was thought that a single plasmid system would improve infectivity and therefore efficiency of CRISPR/Cas9 induced gene editing. To generate the single plasmid system, the DNA sequence of the U6 RNApolIII promoter and a crRNA sequence which encompassed the MR1 specific crRNA and the trans-activating CRISPR RNA (tracrRNA) sequence were ordered and synthesised by (Eurofins MWG Operon, Ebersberg, Germany). The U6 RNApolIII promoter and a crRNA sequence were flanked by SpeI restriction sites which were used to clone the sequence into the pCDNA.3-TOPO_WT-Cas9 plasmid (a gift from George Church (Addgene#41815)) by non-directional cloning at the Spe1 restriction site upstream of CMV promoter. To produce a lentiviral vector the portion of the vector containing the U6 promoter, gRNA sequence, CMV promoter and WT Cas9 were amplified by PCR using the pCDNA.3_Fwd and pCDNA.3_Rev primers to introduce an AgeI restriction site upstream of the U6 promoter and an NsiI restriction site downstream of the Cas9 sequence. The U6 promoter, gRNA sequence, CMV promoter and WT Cas9 sequence was cloned into the 2nd generation pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone (a gift from Didier Trono's laboratory (Addgene#12252)) which had been PCR amplified using the pRRL.0_Fwd and pRRL.0_Rev primers to introduce AgeI and NsiI restriction sites and remove the human PGK promoter and GFP sequence. All of the primer sequences used in this thesis are listed in Appendix Table 2.

2.4.17.2 **Design and introduction of novel crRNA target sequences into "all-in-one" CRISPR/Cas9 lentivector**

The 'all in one' system was designed so that the 20bp crRNA sequence could be modified to target new DNA sequences. The novel crRNA target sequences were designed by identifying

potential sites with the either a 5’ 20bp NGG or 5’ CCN 20bp sequences, the sequences were run on the blastn aligorithm NCBI (Altschul et al. 1990). Novel crRNA sequences were introduced into the ‘all in one’ system by PCR cloning. The primers designed contained the new crRNA sequence flanked by a sequence complementary to the plasmid sequence. The primers were phosphate modified at the 5’ end. The primers used were in the format below (where N represents nucleotides of the target sequence). The lentivirus vector was PCR amplified (section 2.4.9) with the appropriate primers and an annealing step at 53 °C.

5’GTTTTAGAGCTAGAAATAGCAAGTTAANNNNNNNNNN 3’
3’ NNNNNNNNNNGGAAAGGACGAAACACC ’5

2.4.17.3 **Production of novel crRNA target sequences into lentiCRISPRv2 vector (Zhang lab)**

The lentiCRISPRv2 vector (Addgene plasmid #52961) was provided courtesy of the Zhang lab (Sanjana et al. 2014; Shalem et al. 2014). Novel crRNA sequences targeting genes of interest were designed with the use of the CRISPRdesign tool which ranked all of the potential crRNA sequences according to the likelihood of off target binding. Once the crRNAs were designed primers were ordered in the format (where N is crRNA sequence):

5’ CACCGNNNNNNNNNNNNNNNNNN 3’
3’ CNNNNNNNNNNNNNNNNNNNCAAA 5’

Production of the lentiviral vector was carried out following the guidelines found on the Zhang Lab GeCKO website: http://www.genome-engineering.org/gecko/. The three steps are described in sections 2.4.17.4, 2.4.17.5 and 2.4.17.6.

2.4.17.4 **Dephosphorylation of the lentiviral vector**

The lentiCRISPRv2 vector was digested and dephosphorylated by combining the reagents below in 200 µL PCR tubes:

Reagent	Volume
LentiCRISPRv2	5 µg
FastDigest BsmBI (Fermentas)	3 µL
FastAP (Fermentas)	3 µL
10X FastDigest Buffer	6 µL
100 mM DTT	0.6 µL
ddH_2O	X µL
Final volume	60 µL

The tubes were incubated at 37 °C for 30 min and run on an agarose gel (section 2.4.3) the backbone was excised from the gel and the DNA extracted as described in section 2.4.4.

2.4.17.5 **Phosphorylation of oligos**

Oligos were phosphorylated for downstream applications by combining the reagents below in 200 µL PCR tubes:

Reagent	Volume
Oligo 1 (100 µM)	1 µL
Oligo 2 (100 µM)	1 µL
10X T4 Ligation Buffer (New England Biolabs)	1 µL
ddH_2O	6.5 µL
T4 PNK (New England Biolabs)	0.5 µL
Final volume	10 µL

The samples were vortexed and centrifuged briefly and then placed in a thermocycler for the following programme:

Temperature	Time
37 °C	30 min
95 °C	5 min
95 °C to 25 °C	5 °C/ min

2.4.17.6 **Ligation of dephosphorylated backbone and phosphorylated annealed oligos**

The dephosphorylated backbone and phosphorylated annealed oligos were ligated by combining the following reagents:

Reagent	Volume
Dephosphorylated backbone (50 ng)	X µL
oligo duplex (100 nM)	1 µL
2X Quick Ligase Buffer (New England Biolabs)	5 µL
ddH_2O	X µL
Quick Ligase (New England Biolabs)	1 µL
Final volume	11 µL

A control with dH_2O replacing the oligo duplexes was also prepared. The samples were incubated at RT for 30 min and then transformed into XL10-gold bugs (section 2.1.3).

3 Comparison of tools for gene silencing in Molt3 cells

3.1 Introduction

The basic principles of gene silencing with RNAi and nuclease gene editing are described in the section 1.3. The shRNA, ZFNs, TALENs and the CRISPR/Cas9 systems will each be discussed in more depth in the following sections.

3.1.1 Small hairpin RNA (shRNA)

To induce stable silencing of the target gene shRNA is typically expressed in a viral vector. The viral vector contains a DNA sequence which consists of a sense and antisense copy of the target sequence separated by a loop sequence. After transcription the shRNA is exported from the nucleus and the loop is removed from the shRNA by a dicer, resulting in a small interfering RNA (siRNA). The double stranded siRNA is converted into a single stranded siRNA and the resulting single stranded siRNA binds to an RNA inducing silencing complex (RISC). mRNA with an identical sequence to that of the single stranded siRNA is degraded (Reviewed by Sen & Blau 2006; Moore et al. 2010). A schematic representing the process is shown in Figure 3.1.

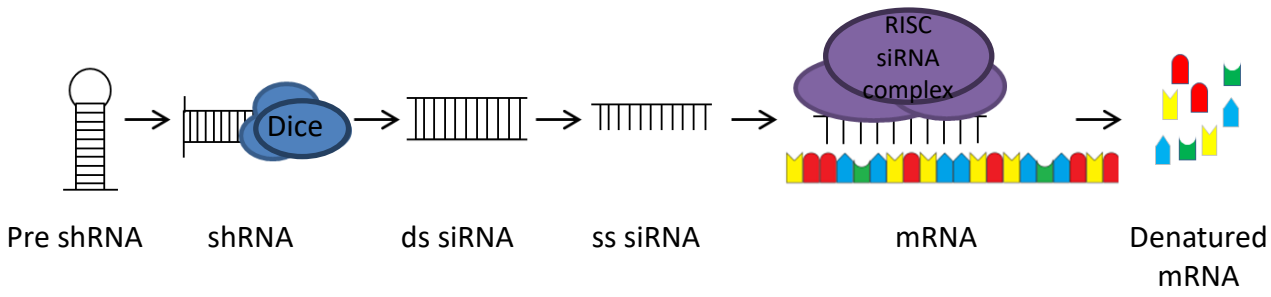


Figure 3.1. A schematic representation of shRNA processing and mRNA degradation. Pre shRNA is processed into shRNA which is subsequently converted to double stranded siRNA and then single stranded siRNA. The single stranded siRNA forms an RNA-induced silencing complex leading to denaturing of corresponding mRNA.

3.1.1.1 Design considerations

There are several design considerations when designing potential shRNA sequences. One issue is the GC content of shRNA sequence as the GC content affects the thermodynamic stability (Elbashir et al. 2002; Holen et al. 2002). A low internal stability at the 5' antisense

strand was shown to be important both for effective unwinding and entry into the RISC complex (Reynolds et al. 2004). Finally, shRNA are designed with minimal palindromic repeats within the sequence to avoid the formation of additional folds within the secondary structure (Reynolds et al. 2004). shRNA design algorithms are publically available to aid in the design of shRNA sequences (Olson et al. 2006).

3.1.1.2 **shRNA uses**

Silencing of genes with shRNAs has been a popular method since the technique was first characterised (Fire et al. 1998). Table 3.1 summarises some of the genes which have been silenced in humans with shRNA technology in recent publications.

Table 3.1. A table summarizing examples of genes in human cells which have been silenced with shRNA.

iRNA induced silencing	Genes	Reference
Gene silencing	DNMT1, DNMT3B, E-cadherin, HER2	Kawasaki & Taira 2004
	PRMT5	Zhao et al. 2009
	β-catenin	Assimakopulos 2007
	CCR5-GFP, RASSF1A	Kim et al. 2006
	EZH2	Kondo et al. 2008
	α-1-antitrypsin	Giering et al. 2008
	APK14, KIF11, IGF1R and KIF14	Siolas et al. 2005
	TF, PSKH1	Holen et al. 2002
	p53	Schomber et al. 2004
	SHP2, STAT5 and Gab2	Scherr et al. 2006
	MTA1	Jiang et al. 2011
	10839 gene library	Silva et al. 2005

3.1.2 Fok1 nuclease based technologies

Zinc finger proteins (ZFPs) (section 3.1.3) were first fused to the Fok1 restriction enzyme by Kim et al., 1996, to form ZFN. This approach was later applied to Transcription activator like effectors (TALEs) to produce TALENs. For gene editing purposes, Fok1 nucleases have been designed to be inactive in their monomeric form and only become functional once dimerised. This approach adds to the specificity of targeting and minimises potential off target effects associated with shorter DNA binding motifs (Hockemeyer et al. 2009). In order to increase specificity, all ZFNs and TALENs are designed in pairs to promote dimerisation at

the required genomic location (Hockemeyer et al. 2009). The wildtype form of the Fok1 restriction enzyme cleaves DNA when in a homodimer formation which was shown to lead to off target effects. Thus the Fok1 nuclease used in modern ZFN and TALEN design has been modified to only become activated in the heterodimeric form but retaining the cleavage efficiency shown by the wildtype form (Szczepek et al. 2007; Miller et al. 2007; Doyon et al. 2011).

3.1.3 ZFPs

A ZFP is usually comprised of several Cys2-His2 motifs (Wolfe et al. 2000; Messina et al. 2004) assembled in an array with each Cys2-His2 motif targeting three nucleotides of DNA in sequence (Pavletich & Pabo 1991). The motif also forms a hydrogen bond with one nucleotide of DNA on the opposing strand (Isalan et al. 1997). A Cys2-His2 ZF motif is composed of 2 histidine and 2 cysteine amino acid residues which interact with a central zinc ion, and a number of conserved amino acids (Phe16, Leu22 and His25) aid in the formation of a hydrophobic core (Pavletich & Pabo 1991). The first crystal structure of a ZFP bound to a DNA oligonucleotide revealed that the ZF binds in the major groove of the DNA with amino acids at residues 1, 3 and 6 determining the nucleotide specificity (Pavletich & Pabo 1991). Artificial ZFP domains can be constructed by modular assembly of individual fingers in order to target specific genomic DNA sequences, typically 9 to 18 nucleotide long (Urnov et al. 2010). This major advance allowed protein targeting of explicit sequences within the genome for the first time and was quickly utilised by fusing ZFPs to the Fok1 nuclease domain to form ZFNs. To limit the likelihood of off target effects ZFNs are designed in pairs with a spacer of between 5 - 7 nucleotides between the left and right ZFPs (Christian et al. 2010) to allow Fok1 to function (Figure 3.2).

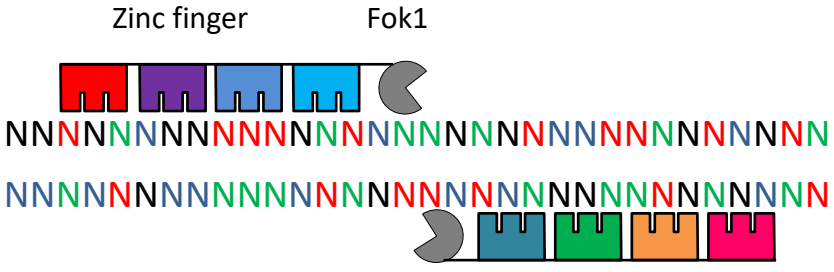


Figure 3.2. Schematic diagram of ZFN. The ZFNs are formed from two domains – the ZF arrays and the Fok1 nuclease domain. The ZF is fused to the Fok1 domain by a peptide linker.

3.1.3.1 **ZFN design considerations**

ZFPs targeting specific DNA stretches can be produced by assembly of individual ZF modules (Mandell & Barbas 2006), by oligomerized Pool Engineering (Maeder et al. 2008; Maeder et al. 2009) or through DNA synthesis using the published ZF module sequences published for 56/64 of the potential 3bp DNA stretches (Segal et al. 1999; Dreier et al. 2001; Dreier et al. 2005; Mandell & Barbas 2006). ZFPs have not been identified for all of the potential 3bp stretches, which limits the DNA sequences that can be targeted by ZFNs. There are several resources and algorithm tools available online to design gene specific ZFNs (such as toolgen/ZFNfinder and umassmed/ZFPsearch).

3.1.3.2 **ZFN uses**

ZFNs have been utilised for gene editing purposes in many organisms including virus (Cradick et al. 2010; Wayengera 2011), plants (Zhang et al. 2010) and many animals such as Zebrafish (Meng et al. 2008), Nematode (*Caenorhabditis elegans*) (Wood et al. 2011), mice and rats (Cui et al. 2011). Table 3.2 summarises some of the human genes which have been targeted by ZFNs.

Table 3.2. A table summarizing examples of genes in human cells which have been modified with ZFNs.

ZFN induced gene editing	Gene(s)	Reference
Gene disruption	CCR5	Perez et al. 2008; Kim et al. 2009
	CXCR4	Yuan et al. 2012; Didigu et al. 2014
	TP73, MAP3K14, EP300, BTK, CARM1, GNAI2 RIPK1 and KDR	Doyon et al. 2010
	TSC2	Doyon et al. 2010; Lu et al. 2013
	NR3C1	Doyon et al. 2011
	TCR	Provasi et al. 2012; Torikai et al. 2012
	VEGF, HOXB13 and CFTR	Maeder et al. 2008
Gene correction	IL2RG	Urnov et al. 2005
	A1AT	Yusa et al. 2011
	HBB	Zou et al. 2011
	SNCA	Soldner et al. 2011
	LRRK2	Sanders et al. 2014
	β globin	Sebastiano et al. 2011; Zou et al. 2011

3.1.4 TALENs

TALEs are bacterial DNA binding proteins found naturally in *Xanthomonas proteobacteria* (Joung & Sander 2013) which can be customized to target specific sequences within the genome. TALEs are formed from tandem repeat modules, which are typically 34 amino acid in length. The 12th and 13th amino acids in the module determine the binding specificity of the module (Cermak et al. 2011; Cong et al. 2012). Thus using TALE modules with alternative residues at the 12th and 13th amino acid position changes the nucleotide to which the TALE module will bind. The residues used in the literature are - NG (T), NI (A), HD (C) and NK/ NN (G) (Cermak et al. 2011; Christian et al. 2012; Schmid-Burgk et al. 2013). TALEs are most commonly fused to the Fok1 nuclease domain forming TALENs. A schematic diagram of a TALEN is shown in Figure 3.3.

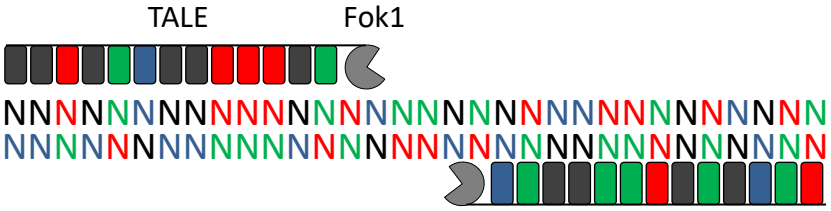


Figure 3.3. Schematic diagram of TALEN. TALENs consist of two domains – a TALE domain which is typically made up of 20 modules linked to a Fok1 domain by a peptide linker.

3.1.4.1 **TALE design considerations**

TALE proteins require a thymine nucleotide one bp upstream of the 5' of the TALE DNA binding sequence (Boch et al. 2009). There are several different TALEN backbones each with optimal spacer lengths (14 - 33 nucleotides) (Christian et al. 2012). TALENs can be synthesised or produced by modular assembly using several commercially available kits using golden gate cloning (Cermak et al. 2011; Sanjana et al. 2012).

3.1.4.2 **Applications of TALEN technology**

TALENs have been used for gene editing in many organisms including yeast (Li et al. 2011), plants (Wang et al. 2014) and many animals such as zebrafish (Solin et al. 2015), nematode (Caenorhabditis elegans) (Cheng et al. 2013), mice (Sung et al. 2013). Table 3.3 summarises some of the human genes which have been targeted by TALENs.

Table 3.3. A table summarizing examples of genes in human cells which have been modified with TALENs.

TALEN induced gene editing	Genes	Reference
Gene disruption	CCR5	Mussolino et al. 2011
	TCR	Berdien et al. 2014
	PSIP1	Fadel et al. 2014
	RAP1	Kabir et al. 2014
	NDUFA9	Stroud et al. 2013
	TERF1IP	Kabir et al. 2014
	AKT2, ANGPTL3, APOB, ATGL, C6orf106, CIITA, CELSR2, CFTR, GLUT4, LINC00116, NLRC5, PLIN1, SORT1, TRIB1, TTN	Ding, et al. 2013
	Library of TALENs targeting 17,120/18,742 (91% of protein coding genes).	Kim et al. 2013
Gene correction	HBB	Ma et al. 2013

3.1.4.3 **RNA guided endonucleases/ CRISPR/Cas system**

The most recently discovered gene editing tools are RNA guided engineered nucleases (RGEN) which are based on CRISPR prokaryotic immune system. The observation of clusters of short fragments of DNA separated by repeating DNA sequences were first identified in *E. Coli* (Ishino et al. 1987) and have been identified in a variety of bacteria and archaea since (Mojica et al. 1993; Mojica et al. 2000). The function of these DNA sequences remained uncharacterised until it was revealed that they were derived from invading bacteriophage and plasmid DNA (Bolotin et al. 2005; Mojica et al. 2005). Incorporating the DNA from invading pathogenic DNA into its own genome in CRISPR arrays allowed the host bacteria to destroy pathogenic DNA upon reinfection. The ability of CRISPR/Cas to target and cleave specific DNA sequences has been put to use for gene editing. The CRISPR/Cas II system is of particular relevance to my own work. The CRISPR/Cas II system allows for the targeting of specific DNA sequences by crRNA and the introduction of DSBs by the Cas proteins. There are several Cas proteins but typically it is the Cas9 variant which is used as a nuclease in combination with CRISPR for gene editing (van der Oost 2013). Several groups have

successfully optimised the CRISPR system to induce DSBs in DNA by co-expressing human codon optimised Cas9 with a crRNA and trans-acting crRNA (tracrRNA) in mammalian cells (Cong et al. 2013; Mali et al. 2013). A tracrRNA is needed to initiate the processing of the crRNA by ribonuclease III (Deltcheva et al. 2011) and activating the cleavage of the DNA by Cas9 (Jinek et al. 2012). The target site of the CRISPR/Cas9 protein is determined by the crRNA sequence, thus altering the crRNA sequence allows for novel genes to be targeted. A schematic diagram of a CRISPR/Cas9 is shown in Figure 3.4.

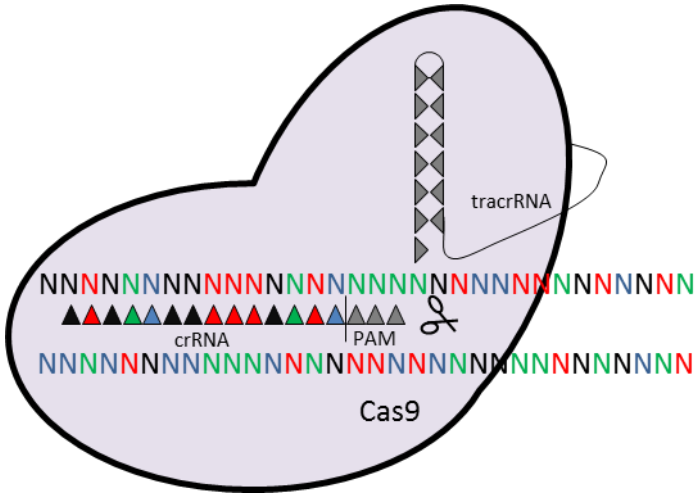


Figure 3.4. Schematic diagram of type II CRISPR/Cas9 system. The structure of *Streptococcus pyogenes* Cas9 bound to target DNA sequence. The crRNA and tracrRNA are shown.

3.1.4.3.1 CRISPR/Cas9 design considerations

The Cas9 proteins used for gene editing can be derived from several bacterial species. The protospacer adjacent motif (PAM) sequence utilised for targeted binding differs somewhat between bacterial species (Karvelis et al. 2015). The most commonly used bacterial species and the species used in this thesis is *Streptococcus pyogenes.* S. *pyogenes* Cas9 protein requires a PAM sequence of NGG where N is any nucleotide (Cong et al. 2013; Mali et al. 2013). Typically crRNAs are designed between 17-20bp in length (Mali et al. 2013; Fu et al. 2014). The NGG motif occurs regularly in the human genome (of all the possible combinations of NNN sequences 1/8 will be suitable PAMs).

3.1.4.3.2 Gene editing applications

The CRISPR/Cas9 system has been used for gene editing in many organisms including fungi (Liu et al. 2015), viruses (Yuan et al. 2015; van Diemen et al. 2016), plants (Jiang et al. 2013) and many animals such as Zebrafish (Jao et al. 2013), Nematode (Caenorhabditis elegans) (Tzur et al. 2013) and mice (Wang et al. 2013). Table 3.4 summarises some of the human genes targeted by the CRISPR/Cas9 system.

Table 3.4. A table summarizing examples of genes in human cells which have been modified with CRISPR/Cas9.

CRISPR/Cas9 induced gene editing	Genes	Reference
Gene disruption	B-globin	Cradick et al. 2013; Liang et al. 2015
	CCR5	Cho et al. 2013; Cradick et al. 2013
	AKT2, CELSR2, CIITA , GLUT4, LINC00116 and SORT1	Ding, et al. 2013
	EMX1,FANCF, RUNX1, and ZSCAN2	Kleinstiver et al. 2016
	EMX1	Cong et al. 2013; Hsu et al. 2013
	PKMYT1 and WEE1	Toledo et al. 2016
	MUC18	Chu et al. 2015
	non-coding RNAs	Ho et al. 2014
	MR1	Laugel et al. 2016
	TCR	Osborn et al. 2015
	PD1	Rupp et al. 2016
	PVALB	Cong et al. 2013
	18,080 genes (GeCKO library)	Shalem et al. 2014
Gene correction	CFTR	Schwank et al. 2013
	DMD	Li et al. 2015
	FANCC	Osborn et al. 2015
	HBB	Xie et al. 2014
	Dystrophin	Ousterout et al. 2015

3.2 Aims

My research group is interested using T-cells in therapeutic applications for a wide variety of diseases. Many studies have introduced specific T-cell receptors (TCRs) into T-cells so as to alter their antigen specificity (Morris et al. 2005; Morgan et al. 2006) and TCR transduced peripheral blood lymphocytes have been used to treat cancer in mice (Dossett et al. 2009) and humans (Morgan et al. 2006; Parkhurst et al. 2011; Morgan et al. 2014). As an extension of this approach, there is currently wide interest in gene-modifying T-cells in order to enhance treatment efficiency. Systemic blocking of T-cell checkpoint inhibitors such as CTLA-4 (Ribas et al. 2005; Callahan et al. 2010; Wolchok et al. 2010) and PD1 (Brahmer et al. 2010; Topalian et al. 2012; Hamid et al. 2013) using antibodies has shown great success in the clinic for some cancers but is associated with many autoimmune side effects. A more targeted approach would be to use gene editing to knockout checkpoint inhibitors in only the T-cells of interest. In order to examine and compare gene editing in T-cells during proof-of-concept studies, we wanted to choose a target that was easy to monitor. The CD8A gene fitted these criteria as it is highly expressed at the T-cell surface and there are a wide range of CD8 specific fluorochrome-conjugated antibodies available. This enables easy monitoring of CD8 protein expression in T-cells by flow cytometry. At the outset of my thesis, I set out to compare gene silencing of the CD8A gene by shRNA, ZFNs and TALENs. During the course of my work CRISPR/Cas9 systems became available so I also attempted to add this system to the comparison. I struggled to get good results with my TALENs so finally aimed to compare gene silencing by shRNA, ZFNs and CRISPR/Cas9.

3.3 Results

3.3.1 Design and construction of shRNA targeting the CD8A gene

A shRNA sequence targeting the CD8A gene published on the Broad Institute website (clone ID TRCN0000057583) was synthesised (Eurofins) and cloned into the PLKO.1 lentiviral backbone between NdeI and EcoRI restriction sites. The CD8A shRNA targets a 21bp stretch of exon 2 of the CD8A gene as shown in Figure 3.5.

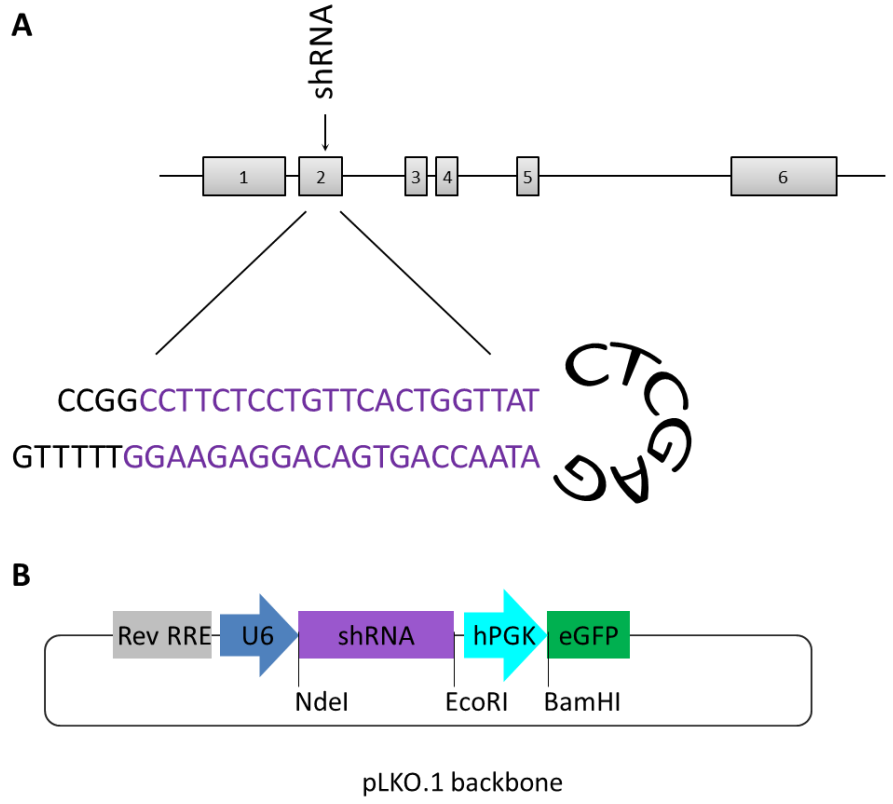


Figure 3.5. Generation of the shRNA targeting CD8A gene. A) The shRNA targets a 21bp stretch (purple lettering) of DNA located within exon 2 of the CD8A gene. B) The DNA encoding the shRNA sequence was cloned into the pLKO.1 lentiviral backbone between NdeI and EcoRI restriction sites upstream of a U6 promoter. The pLKO.1 backbone contains a GFP reporter gene.

3.3.2 Design and construction of ZFNs targeting the CD8A gene

Zinc finger nucleases targeting exon 2 of the CD8A gene were designed and validated by Sigma Aldrich. These ZFNs were cloned into the pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone. The left zinc finger (LZF) protein consists of 4 modules targeting a 16 bp stretch of DNA. The right zinc finger (RZF) protein consists of 5 modules targeting a 19 bp DNA sequence as shown in Figure 3.6. The two ZF binding sites were separated by a 4 bp spacer sequence which was necessary for Fok1 function.

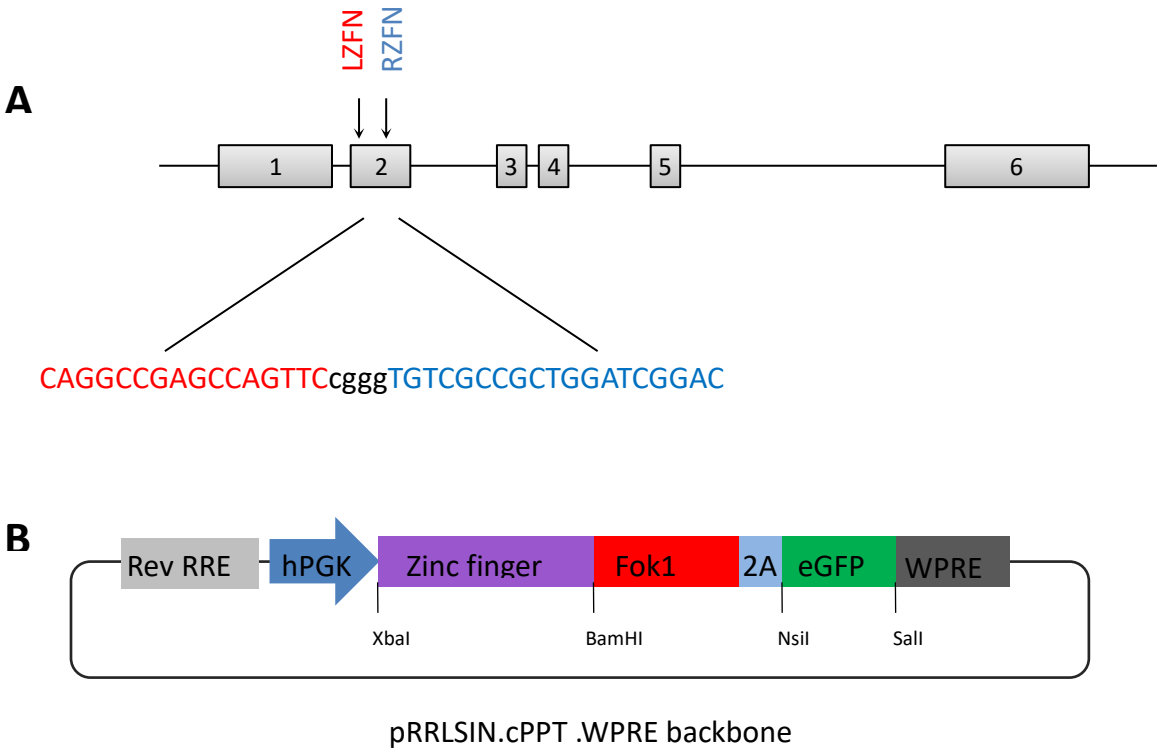
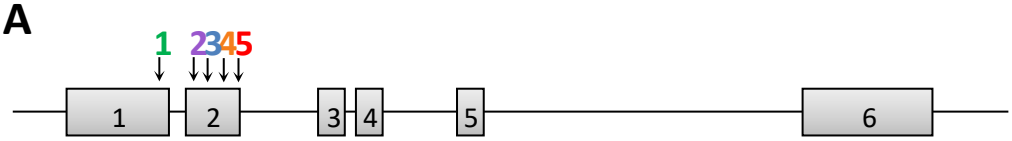


Figure 3.6. Generation of ZFN gene editing tools targeting the CD8A gene. A) The ZFNs bind within exon 2 of the CD8A gene. The target sequences of the ZFN binding sites are shown in red (left ZFN (LZFN)) and blue (right (RZFN)) with the spacer sequence shown in black. B) The lentiviral vector generated express the ZFN in the target cells.

3.3.3 Design and construction of CRISPR/Cas9 targeting the CD8A gene

5 non-overlapping crRNA sequences targeting the CD8A gene were selected from a list of potential crRNA target sequences identified by the CRISPR design tool MIT. The risk of off target mutagenesis was low as the crRNA sequences selected had a minimum of 4 mismatches from other sites in the human genome. The crRNA sequences were cloned into the lentiCRISPRv2 vector (section 2.4.17.3). crRNA 1 recognised a sequence in exon 1 of the CD8A gene, crRNAs 2, 3, 4 and 5 bound to sequences within exon 2 as shown in Figure 3.7.

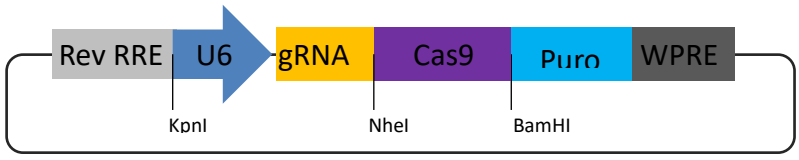
A



B

crRNA	Target sequence	PAM	Exon	Strand
1	GAGCAAGGCGGTCACTGGTA	AGG	1	+
2	GCTGCTGTCCAACCCGACGT	CGG	2	-
3	TCCGATCCAGCGGCGACACC	CGG	2	+
4	AACAAGCCCAAGGCGGCCGA	GGG	2	-
5	CTCTCGGCGGAAGTCGCTCA	GGG	2	+

C



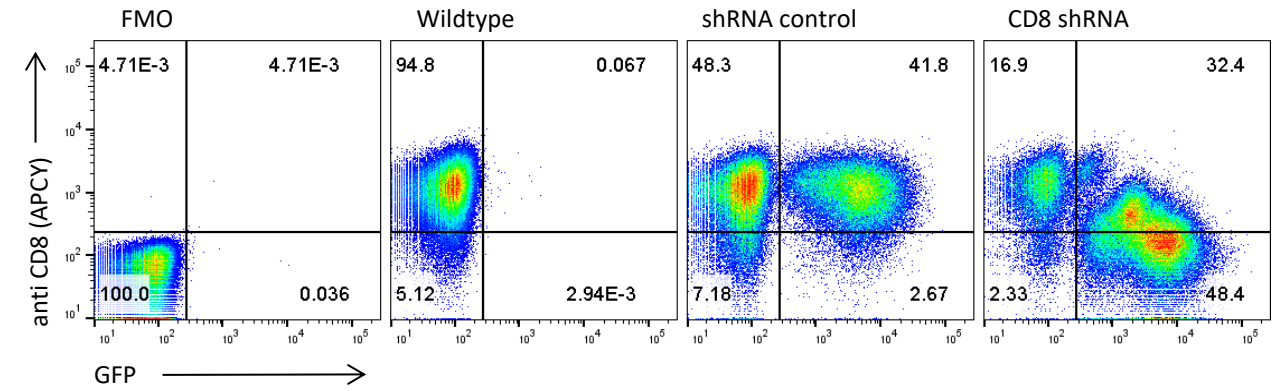
lentiCRISPRv2

Figure 3.7. Generation of CRISPR/Cas9 gene editing tools targeting the CD8A gene. A) Positioning of the 5 crRNA target sequences with respect to the intron/exon structure of the CD8A gene. crRNA 1 (green) binds within exon 1, crRNA 2 (purple), 3 (blue), 4 (orange) and 5 (red) bind within exon 2. B) A table listing the target sequences, PAM sequence and strand orientation for each of the crRNAs. C) The DNA encoding the crRNA sequence was cloned into the lentiCRISPRv2 backbone upstream of a U6 promoter. The lentiCRISPRv2 backbone contains a puromycin (puro) resistance gene which allowed for selection of transduced cells.

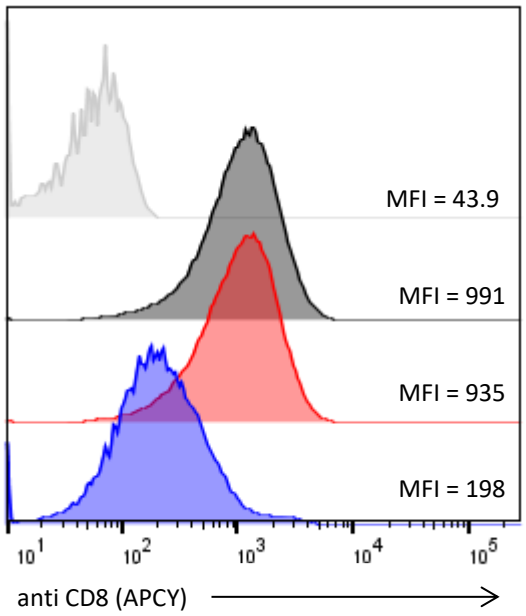
3.3.4 Validation of shRNA targeting the CD8A gene in Molt3

To assess whether the shRNA construct was functional the shRNA was expressed in Molt3 cells by lentiviral transduction. Two weeks post transduction, the cells were analysed by flow cytometry. Transduced cells were identified by expression of GFP. The MFI of CD8 was reduced by 84% in the GFP+ population compared to the WT. The MFI of the GFP+ population within the shRNA control was reduced by 6% compared to the WT (Figure 3.8).

A



B



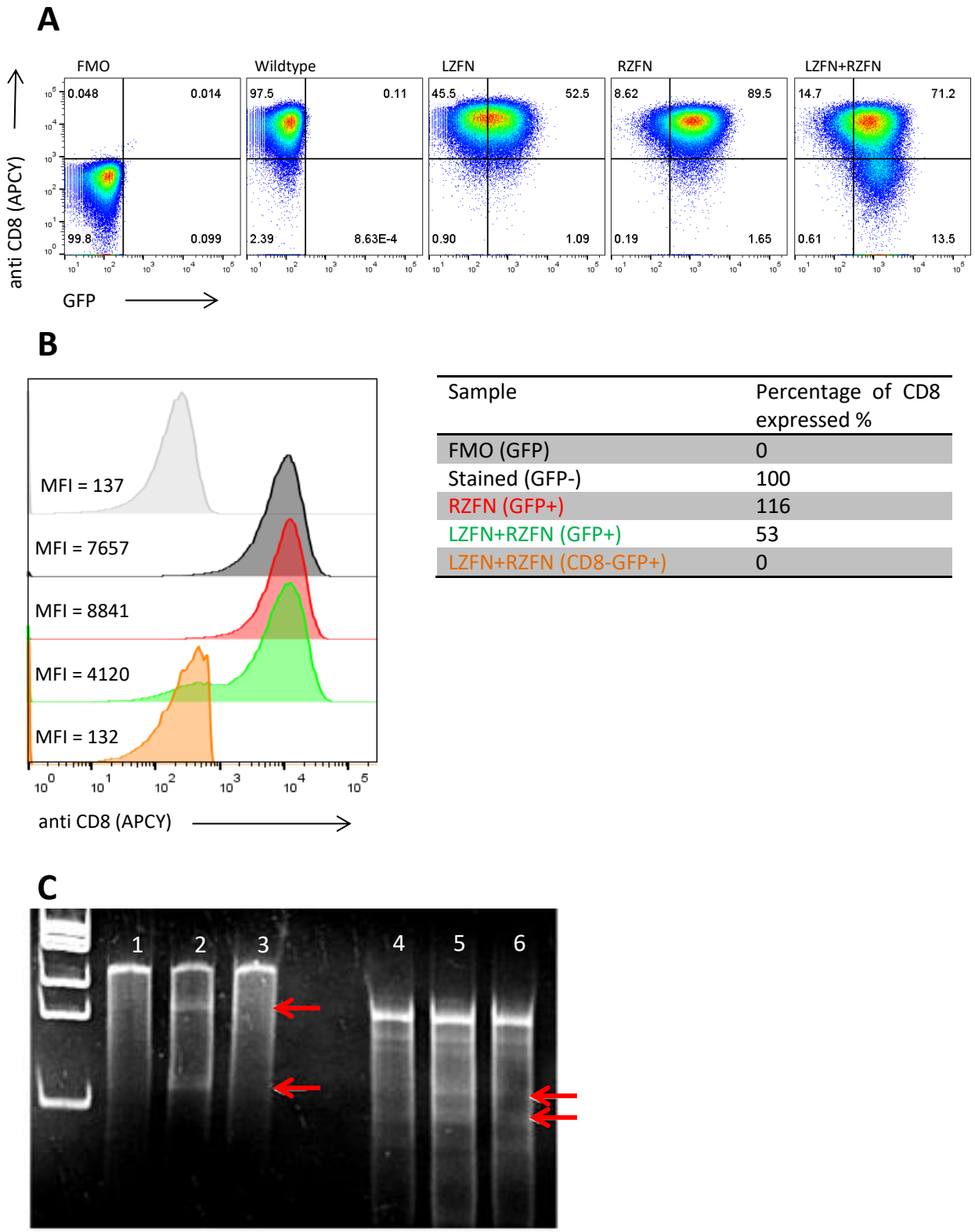
Sample	Percentage of CD8 expressed
FMO (GFP-)	0
WT (GFP-)	100
shRNA control (GFP+)	94
CD8 control (GFP+)	16

Figure 3.8. Validation of the CD8A shRNA in the Molt3 cell line. Molt3 cells were transduced with the CD8A shRNA or the MR1 shRNA as a control by lentiviral transduction. The cells were cultured for two weeks and then analysed by flow cytometry. A) Dot plots of WT and transduced Molt3 cells. Of the cells transduced with the CD8A shRNA approximately 80% were GFP, of these 60% were CD8 negative. B) Histogram plots were produced by gating on the GFP+ of transduced cells and GFP- population of untransduced cells. The percentage of CD8 expressed by each of the samples was determined relative to the WT and FMO controls. An 84% reduction in the geometric mean of CD8A shRNA was observed in the GFP+ population of transduced cells relative to WT cells. Data shown is representative of 3 experiments

3.3.5 Validation of ZFNs targeting the CD8A gene in Molt3

To assess the functionality of the ZFNs, Molt3 cells were transduced with the LZFN, RZFN or both the LZFN and RZFN by lentiviral transduction. The cells were monitored for CD8 surface expression two weeks post transduction. In order to induce knockout of CD8, the cells must have been transduced with both the LZFN and RZFN construct, as both constructs had a GFP reporter it was not possible to accurately determine the frequency of cells that were transduced with both constructs. The frequency of cells transduced with both constructs

was estimated to be 49% (multiplying the frequency of GFP+ cells in the single transduced cells ((54*91)/100). Overall, 13.5% of cells were CD8-, which equates to 28% of the estimated double transduced cells were CD8 negative (Figure 3.9A). The MFI of CD8 was reduced by 47% in the bulk GFP+ population of LZFN and RZFN transduced cells compared to the WT untransduced cells. The MFI of CD8 was increased by 15% in the RZFN transduced cells relative to the WT untransduced cells. Gating on the CD8-GFP+ population of LZFN and RZFN transduced cells shows a 100% reduction relative to the GFP+ population of the WT cells (Figure 3.9B). To confirm that the ZFN approach induced indels at a DNA level a surveyor assay was performed by extracting DNA from the wildtype and bulk LZFN and RZFN transduced cells. The DNA was amplified using CD8_surv_forward and CD8_surv_reverse primers which bound up and down stream of the ZFN binding sites. After annealing and digestion with the mismatch specific enzyme, additional bands were observed for the transduced cells on the agarose gel indicating that transducing cells with the LZFN and RZFN lead to indels at a DNA level (Figure 3.9C).



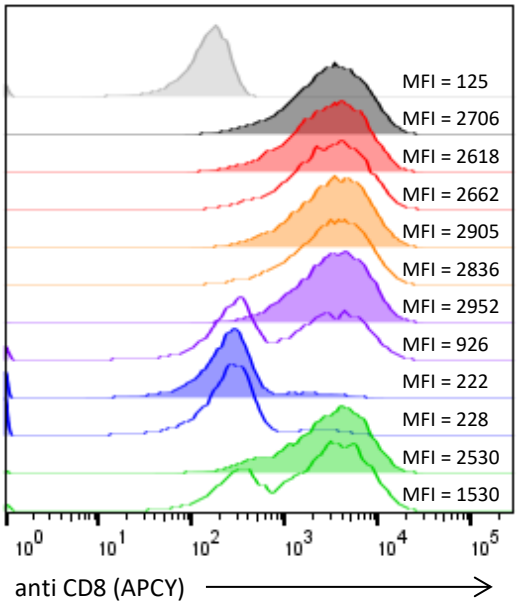
Sample	Percentage of CD8 expressed %
FMO (GFP)	0
Stained (GFP-)	100
RZFN (GFP+)	116
LZFN+RZFN (GFP+)	53
LZFN+RZFN (CD8-GFP+)	0

Figure 3.9. Validation of the ZFN gene editing approach in the Molt3 cell line. Molt3 cells were transduced with the LZFN, RZFN or both the LZFN and RZFN by lentiviral transduction the cells were cultures for 2 weeks after which the cells were monitored by flow cytometry. A) Dot plots displaying the FMO, WT and cells transduced with the LZFN, RZFN and both the LZFN and RZFN. B) Histogram plots were generated by gating on the GFP- population of the WT cells, GFP+ of the transduced cells and the CD8- population within the LZFN and RZFN transduced cells. The percentage of CD8 expressed by each of the samples was determined relative to the WT and FMO controls. C) The presence of mutations at a DNA level were identified by the surveyor assay. DNA was PCR amplified using the CD8_SURV_forward and CD8_SURV_reverse primers, which bound upstream and downstream of the LZFN and RZFN binding sites respectively. 400ng of PCR product from unmodified Molt3 (lane 4) or 200 ng of PCR product from Molt3 cells transduced with the LZFN and RZFN and 200 ng of PCR product from unmodified Molt3 cells (lane 5) or 400ng of PCR product from Molt3 cells transduced with the LZFN and RZFN (lane 6) were denatured, annealed and digested with the mismatch specific surveyor enzyme (section 2.4.12). Additional bands are observed in the DNA from the LZFN and RZFN transduced cells indicating mismatches at a DNA level were present. Denatured, annealed and digested controls are shown in lane 1 (C control), lane 2 (C+G control) and lane 3 (G control). Data shown is representative of 10 experiments.

3.3.6 Validation of CRISPR/Cas9 targeting the CD8A gene in Molt3

Molt3 cells were transduced with the five crRNAs generated to target the CD8A gene by lentiviral transduction. 24 h post transduction; the untransduced and cells transduced with each of the CRISPR/Cas9 lentivirus were each split into two wells. One well was cultured with R10 and the other well was cultured with R10 supplemented with 2 µg/ mL puromycin as the lentiCRISPRv2 backbone contained a puromycin resistance gene to allow for selection of transduced cells. The cells were cultured for two weeks post transduction before being monitored by flow cytometry. Three patterns of CD8 expression were observed with the different crRNAs. crRNAs 1 and 2 failed to reduce CD8 expression suggesting they were non-functional. crRNAs 3 and 5 generated a minor population of CD8 negative/low cells with puromycin selection. In contrast, crRNA4 induced reduction of CD8 expression in the majority of cells in the presence or absence of puromycin selection. The MFI of CD8 was a reduced by 96% in the puromycin selected cells relative to the wildtype cells. crRNA 4 was consistently the superior crRNA and was selected for further experiments.

A



Sample	Percentage of CD8 expressed %
FMO	0
WT	100
crRNA 1	97
crRNA 1 (puro selection)	98
crRNA 2	108
crRNA 2 (puro selection)	105
crRNA 3	110
crRNA 3 (puro selection)	31
crRNA 4	4
crRNA 4 (puro selection)	4
crRNA 5	93
crRNA 5 (puro selection)	54

B

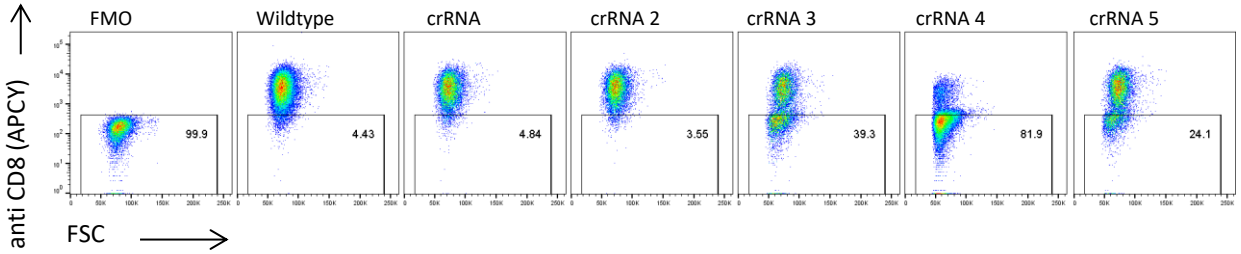


Figure 3.10. Validation of crRNAs targeting the CD8 in the Molt3 cell line. Molt3 cells were transduced with crRNAs 1, 2, 3, 4 and 5 by lentiviral transduction. Two weeks post transduction the cells were analysed by flow cytometry. A) The level of CD8 expression on the cell surface of cells was monitored by flow cytometry, the FMO control staining is shown as a solid light grey histogram, the WT is show as a dark grey histogram. The histograms for crRNAs 1, 2, 3, 4 and 5 are shown in red, orange, purple, blue and green respectively. Solid coloured histograms show bulk transduced cells without puromycin selection and coloured unfilled histograms show cells after puromycin selection. A population of CD8 negative cells is observable in cells transduced with crRNAs 3, 4 and 5. The percentage of CD8 expressed by each of the samples was determined relative to the WT and FMO controls. B) A dot plot displaying the FMO, WT and cells transduced with the 5 crRNA (after puromycin selection). The CD8 negative populations are gated within each of the samples. Data shown is representative of 3 experiments.

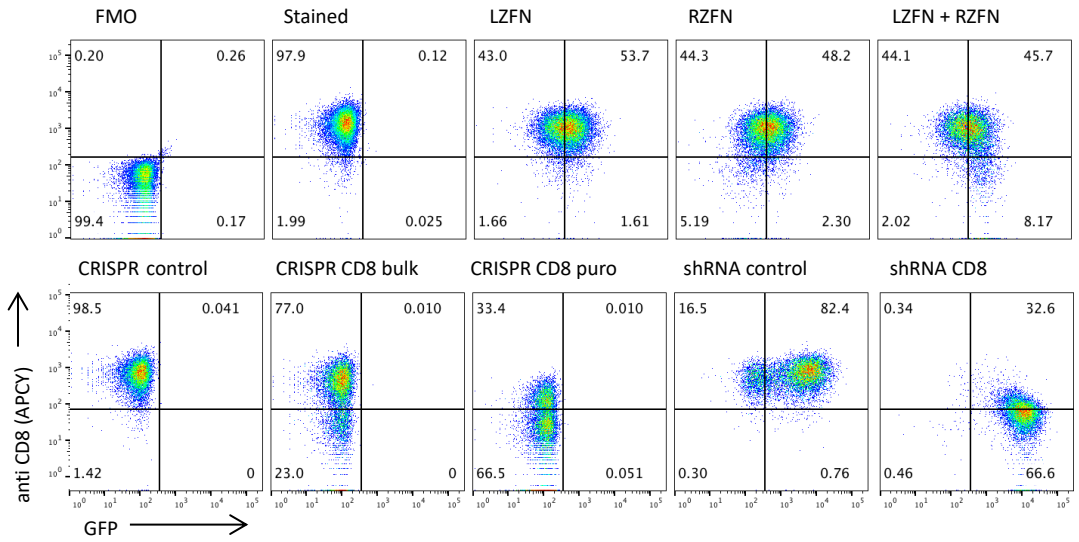
3.3.7 Comparison of shRNA, ZFNs and CRISPR/Cas9 in Molt3

In order to compare the shRNA, ZFNs and CRISPR/Cas9, Molt3 cells were transduced with gene editing tools by lentiviral transduction. The cells transduced with crRNA 4 were split 24 h post transduction and one of the two wells was treated with puromycin at 2 µg/ mL. The cells were cultured for a period of 5 weeks and monitored by flow cytometry on a weekly basis. The shRNA, ZFN and the CRISPR/Cas9 approaches all resulted in a population of CD8 negative cells (Figure 3.11A).

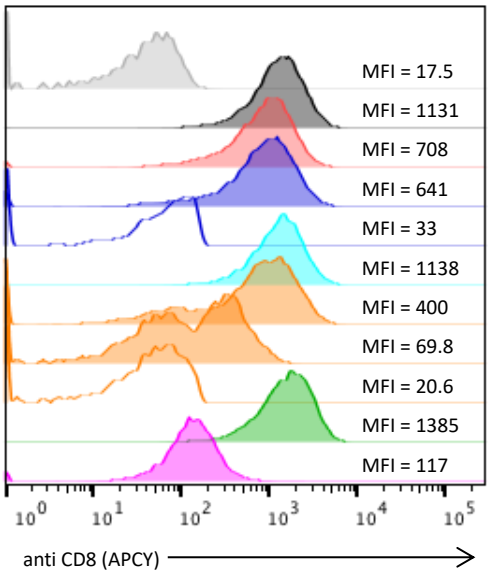
The frequency of transduced cells was determined either by the expression of GFP or by survival of puromycin selection. The frequency of cells transduced with both the LZFN and RZFN was determined by multiplying the frequency of GFP+ cells in the single transduced samples to estimate the frequency of cells transduced with both.

At two weeks post transduction, the MFI of the knock out population was reduced by 99, 100 and 91% relative to the WT in the LZFN and RZFN, crRNA4 and shRNA transduced cells respectively (Figure 3.11B). Approximately 60% of the shRNA transduced cells were CD8 negative after 1 week; over the remaining monitoring period the frequency of CD8 cells increased to a peak of 66% two weeks post transduction and then declined. At 5 weeks post transduction approximately 29% of transduced cells were CD8 negative. Transduction with the LZFN and RZFN resulted in an observable knockout efficiency of 9% of dual transduced cells 1 week post transduction, the frequency of knockout cells within the dual transduced cell population reached 52% at three weeks post transduction and remained fairly consistent over the remaining monitoring period. The CRISPR/Cas9 approach lead to knockout of CD8 in approximately 67% of transduced cells one week post transduction, the frequency of knock out cells rose slightly with 80% of transduced cells expressing CD8 at levels equivalent to the FMO control at five weeks post transduction (Figure 3.11C).

A



B



Sample	Percentage of CD8 expressed %
FMO (GFP)	0
WT	100
RZFN (GFP+)	62
LZFN+RZFN (GFP+)	56
LZFN+RZFN (CD8-GFP+)	1
crRNA control	101
crRNA 4 (bulk)	34
crRNA 4 (puro selection)	5
crRNA 4 (puro selection)(CD8-)	0
shRNA control (GFP+)	123
shRNA CD8 (GFP+)	9

C

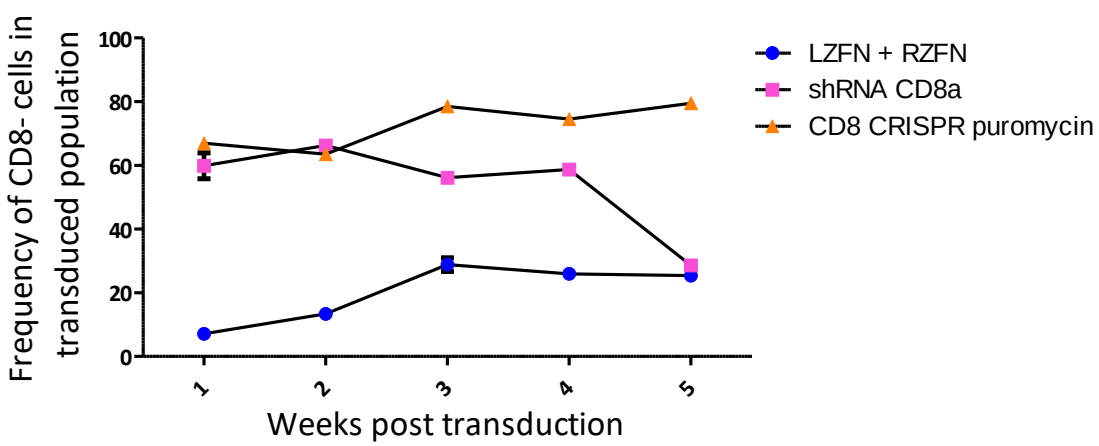


Figure 3.11. Comparison of the shRNA, ZFN and CRISPR/Cas9 gene editing tools. Molt3 cells were transduced with each of the constructs by lentiviral transduction and cultured for five weeks. The cells were analysed by flow cytometry at weekly intervals. A) Dot plots of the transduced cells two weeks post transduction. B) Histogram plots of the transduced cells at two weeks post transduction, the FMO is shown in light grey, the WT is shown in dark grey, the RZFN is shown in red, LZFN and RZFN transduced cells are shown in blue, crRNA is shown in orange, shRNA trol in green and shRNA CD8A in pink. The corresponding gating is shown in brackets. The percentage expression of each of the samples was calculated as a percentage of the WT (black) after deducting the FMO (light grey). C) A graph displaying the frequency of CD8 negative cells within the transduced cell population relative to the WT as determined by flow cytometry based on triplicate repeats of the transductions. Data shown is representative of 2 experiments.

3.4 Discussion

Several approaches have been utilised to induce targeted gene silencing and gene knockout in mammalian cells. Since shRNA was first manipulated to induce targeted gene silencing (Fire et al. 1998) it has been widely used to silence the expression of many genes. More recently, advances within the nuclease gene editing field have produced gene editing tools that induce gene knockout through the DSBs which when repaired by erroneous pathways lead to the introduction of mutations at a DNA level. Chronologically, ZFNs were identified first, followed by TALENs and most recently the CRISPR/Cas9 system. Each of these gene editing tools has been used widely since their discovery and each new technology has largely superseded its predecessors as the methodology of choice. As part of this project, plasmid systems were developed for each gene editing tool targeting the CD8A gene in order to validate and compare each of the systems. The TALEN approach (section 3.1.4) was not included in this study; the limitations of TALENs are discussed in section 3.4.2. The shRNA, ZFN and CRISPR/Cas9 approaches were validated and compared. Each of the approaches successfully induced gene silencing or gene knockout in the Molt3 cell line. The shRNA approach did not result in consistent gene silencing as the frequency of CD8 negative cells rose substantially throughout the monitoring period. The ZFN approach led to an approximate knockout efficiency of 20%, which was fairly consistent across the monitoring period. The CRISPR/Cas9 approach was found to lead to superior efficiency of gene knockout with approximately 80% of transduced cells expressing CD8A at levels equivalent to the FMO control consistently across the monitoring period when monitored by flow cytometry.

3.4.1 Use of double reporter system for identification of cells transduced with both the left and right ZFN constructs

In this thesis the ZFNs delivered by lentiviral transduction resulted in the knockout of CD8 from the Molt3 cell line as demonstrated by flow cytometry. However, the frequency of cells expressing both the LZFN and RZFN proteins was not established as a GFP marker was present in both constructs and therefore single and double transductants were indistinguishable. To improve this system for future use a second reporter gene could be

cloned into one of the ZFN constructs to allow for detection of cells transduced with both the LZFN and RZFN.

3.4.2 Generation of TALENs

Assembly of TALENs was attempted using two commercially available kits by Cermak et al. 2011 and Sanjana et al. 2012 (Methods can be found in Appendix Figure 2 and 3. The process of TALEN construction can be found in Appendix Figure 4 and 5). A TALEN targeting the CD8A gene was produced by the Cermak Kit (Appendix Figure 4); however, no TALENs were produced with the Sanjana kit (Appendix Figure 5). The TALEN approach was incompatible with lentiviral transduction due to unintentional rearrangements within the TALE sequence (Holkers et al. 2012). A pair of TALENs were synthesised (Eurofins) which were designed to minimise the repetitive nature of the TALE arrays. The TALE protein was cloned into the pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone which contained a GFP reporter. However, attempts to transduce cells with the TALE protein were unsuccessful as no GFP was expressed (Appendix Figure 6). I spent considerable time in design and testing of TALENS. Once CRISPR came along I switched my efforts to this new tool.

3.4.3 Design and construction of gene editing tools

In order to generate shRNA targeting the CD8A gene for this project the target sequence published by the Broad institute was synthesised (Eurofins) and cloned into a lentiviral backbone. This process is fast and inexpensive as shRNA constructs can be generated within 2 weeks dependent upon the synthesis of the target sequence. The ZFNs used in this project were produced by Sigma Aldrich, unless ZFNs targeting the gene of interest have previously been ordered from the company the production costs and timing are increased as due to the bespoke validation process. ZFNs for novel gene targets can take several weeks to produce and cost several thousand pounds per pair. Once the ZFN plasmid is delivered the gene must be cloned into a lentiviral backbone. The newer, CRISPR/Cas9 technology is by far the easiest and cheapest approach as altering the specificity of the simply involves design and ordering crRNA sequences. The oligos typically cost less than ten pounds per pair. The oligos are then cloned into the CRISPR/Cas9 lentiviral backbone, from start to finish this process can be completed within a week. Unless validated crRNA sequences for target genes

are published there is no guarantee that the crRNA sequences produced will be functional, however, using this system within our lab we have produced crRNAs targeting a range of genes and have found that designing and producing 5 crRNAs per target gene results in a minimum of 2 functional crRNAs. The efficiency, ease and low cost of CRISPR/Cas9 has resulted in this system now being the method of choice for gene silencing and the approach was awarded Breakthrough of the Year 2015 by Science magazine (Travis 2015).

3.4.4 Gene silencing efficiency

When directly compared against one another there was a reduction of MFI of 98, 100 and 93% for shRNA, ZFN and CRISPR/Cas9 respectively. Although the reduction in MFI was comparable between these gene editing tools the frequency of CD8 negative cells in the transduced populations varied. At two weeks post transduction, the CRISPR/Cas9 approach resulted in a knockout efficiency of 84% in transduced cells. The ZFN approach had an efficiency of 20% while 67% of cells transduced with the shRNA approach had observable CD8 gene silencing. The level of CD8 expressed by the cells transduced with the shRNA did not remain silenced and at five weeks post transduction, only 35% of the transduced cells expressed CD8 at low levels.

3.4.5 Summary

In this chapter I began by comparing shRNA, TALEN and ZFN technologies for silencing of the CD8A gene. I put considerable effort into building the TALEN tools, unfortunately, these efforts were unfruitful. While struggling with TALENs, CRISPR/Cas9 technology became available so I adopted to utilise this methodology. Overall, in direct comparisons, the CRISPR technology proved to be preferable. I went on to use CRISPR technology to study antigen presentation to unconventional T-cells in chapter 6. Some concerns have been raised about the use of nuclease based gene editing for clinical application. I next set out to find an alternative technology that did not incorporate the inherent risks of off-target nuclease activity. I will further discuss the implications of all this work in chapter 4.

4 Development of non-nuclease gene silencing technology for therapeutic manipulation of T-cells

4.1 Introduction

As described in chapter 3, recent developments in gene editing resulted in genome editing with engineered nucleases being awarded method of the year 2011 by *Nature Methods* (Nature methods, 2012). More recently the CRISPR/Cas9 gene editing technique was named breakthrough of the year 2015 by Science magazine (Travis 2015). There has been particular interest in using genome-editing technology for therapeutic applications with T-cells being at the forefront of these advances. ZFN technology has been utilised to knockout the CCR5 receptor to prevent entry of HIV into CD4+ T-cells, in order to render them resistant to viral attack (Holt et al. 2010; Tebas et al. 2014). The first clinical trial with CRISPR/Cas9 was approved in July 2016 and is due to have commenced in August 2016 (Cyranoski 2016).

To date, the most effective gene silencing technologies, ZFNs, TALENs and CRISPR/Cas9, make use of targeted nuclease activity. All these technologies therefore bear an inherent risk of off-target nuclease activity that has potential to transform cells and render them malignant. There is an urgent need for gene silencing technology that circumvents these risks during therapeutic use. In addition, the debate that rages on about who owns CRISPR has introduced uncertainly in those that wish to use CRISPR in therapeutic approaches and a new approach to gene silencing would be most welcome.

On this backdrop I decided to examine whether the precise DNA binding properties of ZFPs could be used to silence genes without the need to couple them to nucleases. In particular, I was interested in whether these domains could be successfully coupled to transcriptional repressor domains to enable precise silencing of genes in T-cells. Previously ZFPs have been fused to the KRAB repressor to silence specific promoters (SV40 promoter (Liu et al. 1997), erbB-2 (Beerli et al. 2000) and PPARγ2 (Ren et al. 2002)) however, to date no studies have used ZFPs fused to repressor domains to target and silence specific genes. Since

commencing this project there have been several advances in the CRISPR interference field which are discussed in section 4.6.4

4.1.1 Transcriptional repressors

There are two types of mammalian repressors - passive and active repressors. Passive repressors do not have intrinsic repressing activity but instead compete with transcriptional activators for DNA binding which leads to repression of mRNA synthesis. In contrast, active repressors have intrinsic repression activity. Active repressors modify the chromatin state through histone deacetylation and heterochromatin formation leading to transcriptional repression and gene silencing (Reviewed by Thiel et al. 2004). This chapter is focused on active repressors and therefore passive repressors will not be discussed further.

4.1.2 Gene transcription

In eukaryotic cells DNA is wound around histones forming the classic "beads on a string" structure. The core histone complex is formed by an octamer of two copies of H2A, H2B, H3 and H4 histone proteins (Kouzarides 2007). The gene transcription process is complex and involves the interaction of DNA promoters with the RNA polymerase II enzyme, transcription factors and genomic regulatory elements. Gene transcription is highly regulated at each stage to ensure select proteins are expressed at the correct levels (reviewed by Maston, et al. 2006). Chromatin (DNA and associated histone proteins) undergo biochemical modifications of phosphorylation, methylation and acetylation which alters the state of the chromatin and thus the levels of gene transcription. Chromatin exists as heterochromatin and euchromatin (Heintzman et al. 2007). In the heterochromatin state the nucleosomes are compacted and the promoter regions are unable to interact with transcription factors and are therefore unable to be transcribed. Alternatively the euchromatin state is a less compact state and the transription promoters are accessible to transcription machinery and can therefore be transcribed (reviewed by Gaszner & Felsenfeld 2006; Kouzarides 2007). Transcription factors are formed by a DNA binding domain (typically a ZFP as described in section 3.1.2) and an activator or repressor domain. Transcription factors play a key role in the regulation of gene transcription as they bind to specific DNA sequences and allow for targeted transcriptional repression or activation.

4.1.3 Choice of Repressor Domains

Several repressor domains have been described in the literature. I opted to test 7 previously described repressor domains for my work, namely KRAB, ARP1, GCN4-KRAB, HP1A, ScanKRAB, SID and TRIM28. Each is described below in turn.

4.1.3.1 **Krüppel-associated box (KRAB)**

In mammals ZF-KRAB family are the most common transcription factors, therefore the KRAB repressor domain has been well characterised (Bellefroid et al. 1991; Urrutia 2003). KRAB functions by binding to co-repressor proteins such as TRIM28 (Abrink et al. 2001) (section 4.1.3.7), leading to the recruitment of chromatin-modifying factors such as heterochromatin proteins (Groner et al. 2010a). Regions of genes silenced by KRAB lose histone H3 acetylation and acquire H3 lysine 9 trimethylation (H3K9me3) leading to prevention of RNA polymerase recruitment and formation of heterochromatin (Abrink et al. 2001; Huntley et al. 2006; Groner et al. 2010b), ultimately leading to the repression of DNA transcription and a loss of target protein production. KRAB proteins have been observed to mediate gene silencing several tens of kilobases away from the DNA binding sites (Groner et al. 2010b). In mice, the ZF-KRAB approach was utilised to silence a PPARγ2-specific promoter leading to a 50% reduction in PPARγ2 expression (Ren et al. 2002). Engineered ZF-KRAB proteins have been used to target erbB-2 specific promoters in the SKBR3 human breast cancer cell line resulting in a 7 fold reduction in erbB-2 expression (Beerli et al. 2000) and upstream of a SV40 promoter in a HeLa reporter cell line leading to a 93% reduction in luciferase transcription (Liu et al. 1997).

4.1.3.2 **ARP1**

ARP1, also referred to as COUP-TFII, has been described as both an activator and repressor of gene transcription. Binding of COUP-TFII to DNA leads to the recruitment of corepressor proteins such as nuclear corepressor (Bailey et al. 1997) leading to heterochromatin formation (reviewed by Litchfield & Klinge 2012).

4.1.3.3 **GCN4-KRAB**

The GCN4 transcription factor is primarily described as an activator of gene transcription and has been identified in a range of organisms (Ellenberger et al. 1992). In one study the

binding sites of 4 ZFPs were expressed in reporter plasmids upstream of a CMV promoter driving transcription of a AmCyan fluorescent protein. Reporter cell lines were transfected with plasmids containing the corresponding ZFP fused to a GCN4-KRAB repressor domain. The GCN4-KRAB repressor domain was shown to lead to a 9-16 fold reduction in the expression of the AmCyan fluorescent protein (Lohmueller et al. 2012).

4.1.3.4 **Heterochromatin protein 1a (HP1A)**

HP1a is recruited by transcription factors where it initiates silencing of gene transcription by histone 3 lysine 9 trimethylation of (H3K9me3) leading to heterochromatin formation and chromatin packaging (Fischle et al. 2005) ultimately inhibiting gene transcription.

4.1.3.5 **ScanKRAB**

The ScanKRAB domain is a variant of the KRAB repressor domain which is formed from the KRAB and the Scan box (Sander et al. 2003; Huntley et al. 2006). The KRAB domain initiates repression of gene transcription (section 4.1.3.1.) and the SCAN domain mediates protein dimerization to promote further epigenetic editing (Honer et al. 2001).

4.1.3.6 **mSin interaction domain (SID)**

SID repressor functions as a scaffold and interacts with histone deacetylases to induce compaction of the chromatin (Pang et al. 2003). The SID domain is thought to form a repressor complex which alters the structure of the chromatin blocking transcription (Ayer et al. 1996). SID has previously been shown to be an effective repression when fused to a TALE (Cong et al. 2012), suggesting it might also be effective when fused to a ZFP.

4.1.3.7 **TRIM28**

Trim28 is also known as KRAB associated protein 1 (KAP1). TRIM28 coordinates the assemble of a repressor complex composed of HP1, SETDB1, and Mi2α proteins (Ayyanathan et al. 2003; Sripathy et al. 2006) which provide a microenvironment which is favourable for heterochromatin formation by histone 3 lysine 9 trimethylation (Xiao et al. 2011).

4.2 Rational

In chapter 3, the frequency of biallelic inactivation of the CD8A gene in the Molt3 cell line was approximately 14% of bulk and 27% of cells transduced with both the LZFN and RZFN constructs, although this method of gene editing was shown to be successful the rate of biallelic inactivation was relatively low. Previous studies with ZFNs have similarly resulted in relatively low frequencies of biallelic knockout with ZFNs. In one study, modifications of the CCR5 gene were observed in 23% of clones, of which 33% of clones were shown to have homozygous mutations within the CCR5 gene (Perez et al. 2008). Holt et al. 2010, reported an estimated biallelic inactivation of between 5-7% within the CCR5 gene. ZFN induced biallelic modification of the IL2Rγ gene were observed in 7% of cells (Urnov et al. 2005). I wanted to develop a non-nuclease based gene silencing system. It was hypothesised that fusing a repressor domain to a single ZFP could be used to effectively silence gene expression at higher rates than biallelic inactivation with ZFNs. There are several advantages to the use of repressor domains over ZFNs: 1) repressor domains silence gene transcription in all cells carrying the construct; 2) potential off target effects can be monitored in an unbiased manner by gene microarray or RNAseq rather than having to sequence the whole genome; and 3) ZF repressor domains are small enough to allow co-delivery of a TCR within a single lentiviral construct thereby allowing repression of T-cell checkpoint inhibitors in redirected T-cells.

4.3 Aims

In this chapter I aimed to test whether genes could be effectively silenced in the Molt3 cell line and in primary T-cells using a repressor domain fused to a single ZFP. As the ultimate aim was to develop the technology in T-cells for cancer immunotherapy I chose to use the CD8A gene for proof-of-concept using the ZFPs utilised in chapter 3. CD8 is highly expressed at the T-cell surface (>50,000 copies) allowing ready monitoring of gene silencing by flow cytometry. I also aimed to demonstrate that a single lentiviral construct could be used to deliver a TCR while simultaneously silencing a specific gene target via a ZF repressor protein.

4.4 Results

4.4.1 Design and construction of repressor domains fused to a zinc finger protein

I first aimed to test the activity of 7 known transcriptional repressor domains described in section 4.1.3 to assess the function of each when expressed as a fusion protein with a ZF domain. The RZF targeting CD8A was cloned into the pV2 cloning backbone (Addgene #32189) between an AgeI and BamHI restriction sites. A codon-optimised cDNA sequence for each of the 7 repressor domains described (ARP1, GCN4, HP1A, KRAB, ScanKRAB, SID and TRIM28) was synthesised (Eurofins) and cloned into the pV2 cloning backbone downstream of the RZF between BamHI and XbaI restriction sites (Figure 4.1). This cassette system allows easy substitution of both the ZFP and/or the repressor domain at a later date as required.

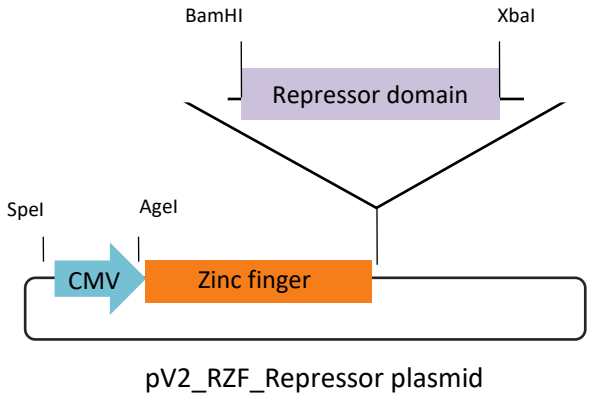


Figure 4.1. Generation of the ZF repressor domain constructs. The codon optimised DNA sequence of the ARP1, GCN4, HP1A, KRAB, ScanKRAB, SID and TRIM28 repressor domains were synthesised by Eurofins flanked by BamHI and XbaI restriction sites. The repressor domains were cloned into the pV2 backbone downstream of a CMV promoter and the RZF targeting CD8.

4.4.2 Design and construction of a reporter system to assess the functionality of the repressor domains

To test the RZF repressor domains a reporter 293T cell line was produced. First, a single copy (referred to as 1x) or a triple copy (referred to as 3x) of the binding sites for the CD8 ZFPs (separated by the 4bp linker) were synthesised (Eurofins). These binding sites were flanked by XhoI restriction sites and were cloned upstream of a hPGK promoter and a GFP

gene into the 2nd generation pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone. 293T cells were transduced with the pRRL_1x and pRRL_3x lentiviral vectors. A diagrammatic representation of the reporter construct is shown in Figure 4.2A. To produce a clone for each of the reporter cell lines the transduced 293Ts were cloned by limiting dilution. Initially, 26 1x clones and 22 3x clones were screened (Appendix Figure 7 and 8) and the best four clones were expanded for each and screened again and a GFP+ clone was identified for both as shown in Figure 4.2B.

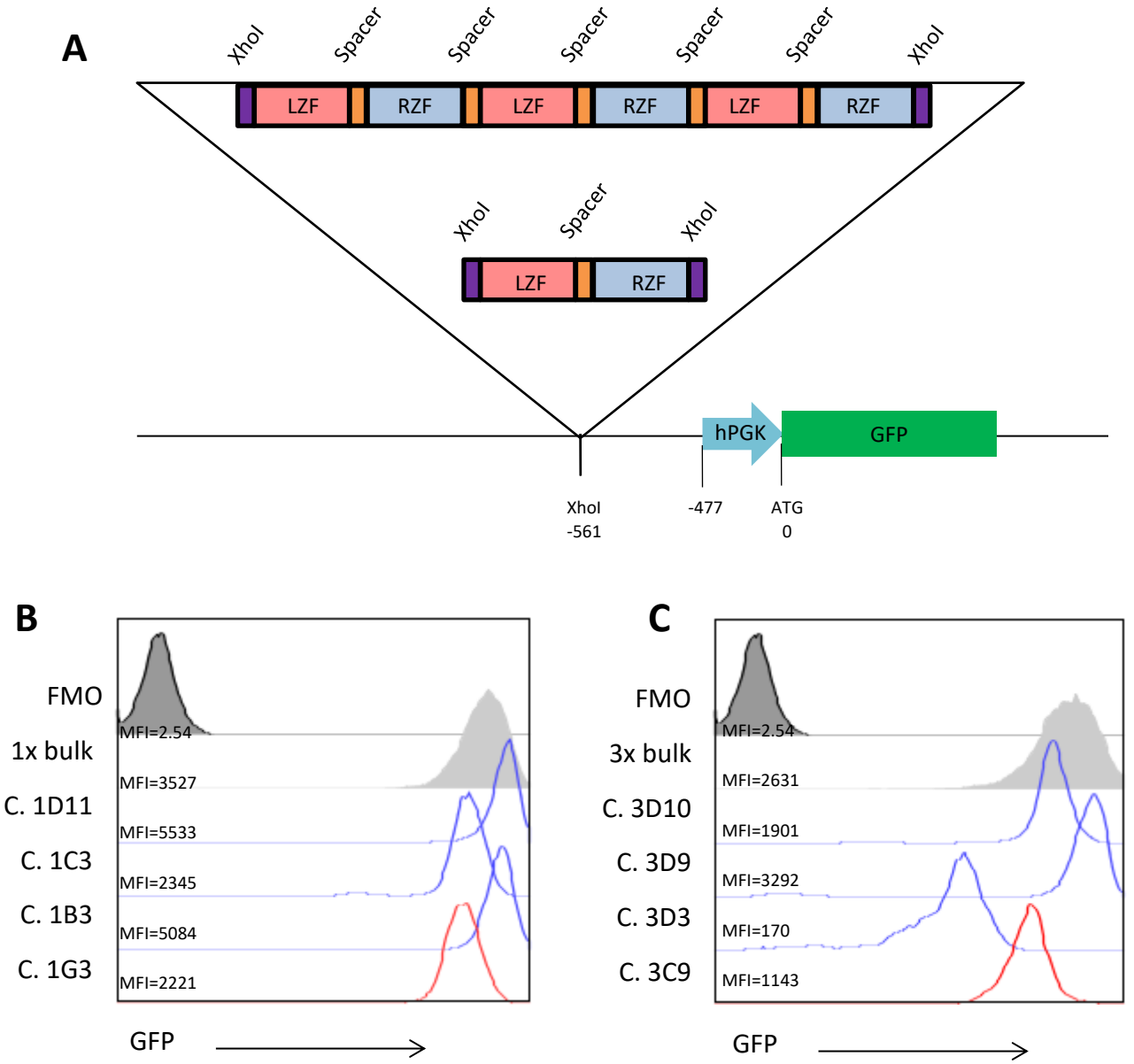


Figure 4.2. Generation of the GFP reporter system to validate the repressor domains. A) Diagrammatic representation of the 1x and 3x reporter system. The left and right zinc finger binding sites were flanked by XhoI restriction sites and the binding sites were separated by a 6bp spacer sequence. Either 1 or 3 copies of the binding sites were cloned into the pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone upstream of the hPGK promoter and GFP reporter gene. 293T cells were transduced with either the 1x (B) or 3x (C) reporter construct by lentiviral transduction. The bulk transduced cells were single cell cloned by limiting dilution. 26 1x clones and 22 3x clones were screened by flow cytometry (Appendix Figure 7 and 8) the FMO is shown as the dark grey histogram, the bulk transduced cells are shown in light grey, three clones representing the variety of GFP are shown by blue histograms and the 1x and 3x clones selected are shown by a red histogram. The MFI of each of the histograms is shown.

4.4.3 Comparison of the repressor domain constructs in the reporter system

To validate the effector domains generated in section 4.4.1 the 1x and 3x 293T reporter cells were transfected with each of the RZF repressor domain constructs by $CaCl_2$ transfection. Cells were transfected as described in section 2.3.8. The incubation time post transfection was optimised and 16 h incubation time was found to lead to optimal levels of transfection

of a GFP control whilst minimising cell death (Appendix Figure 9). The MFI of GFP was consistently reduced in three of the seven repressor domains tested – GCN4, KRAB and TRIM28 suggesting that these repressor domains were effective at repressing the hPGK promoter (Figure 4.3). To assess the degree of MFI reduction, the transfections of the GCN4, KRAB and TRIM28 repressor domains were repeated in triplicate. In the 1x reporter system the overall MFI was reduced by 30, 45 and 33% for the GCN4, KRAB and TRIM28 repressor domains respectively. In the 3x system the overall MFI was reduced by 36, 62 and 32% for the GCN4, KRAB and TRIM28 repressor domains respectively Figure 4.3C and D. The true extent of the functionality of these repressors was difficult to determine as a population of GFP+ cells remained as the transfections were not 100% efficient. The KRAB repressor domain was consistently superior to the other repressor domains tested and was utilised for downstream experiments.

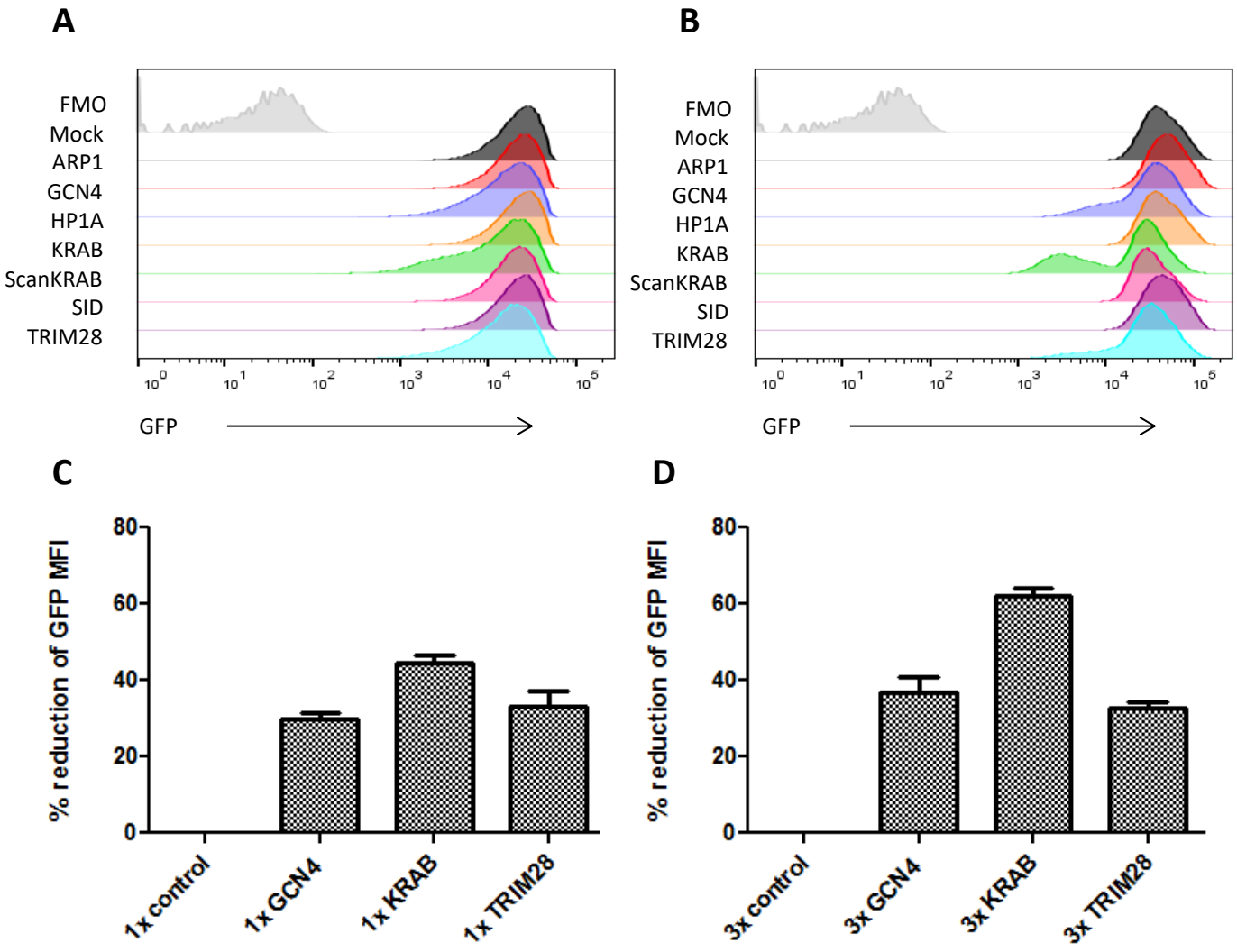


Figure 4.3. Validation of the 7 repressor domains in the 1x and 3x reporter cell lines. The 1x and 3x reporter cell lines were transduced with each of the 7 repressor domain constructs by $CaCl_2$ transfection and the level of GFP extinction was determined by flow cytometry five days post transfection. A) Histogram plots of the 1x reporter cells transfected with the repressor domain constructs. The FMO (light grey; MFI = 10.7), Mock (dark grey; MFI =19599), ARP1 (red; MFI = 19269), GCN4 (blue; MFI =13958), HP1A (orange; MFI =19965), KRAB (green; MFI =10568), ScanKRAB (pink; MFI =17456), SID (purple; MFI =18666) and TRIM28 (light blue; MFI =13094). B) Histogram plots of the 3x reporter cells transfected with the repressor domain constructs. The FMO (light grey; MFI = 10.7), Mock (dark grey; MFI =41654), ARP1 (red; MFI = 48904), GCN4 (blue; MFI =28182), HP1A (orange; MFI =41657), KRAB (green; MFI =14958), ScanKRAB (pink; MFI =33077), SID (purple; MFI =43953) and TRIM28 (light blue; MFI =27382). The data in Figure 4.3 A and B is representative of 5 experiments. The 1x (C) and 3x (D) reporter cell lines were transfected with the GCN4, KRAB and TRIM28 repressor domain constructs and the cells were analysed by flow cytometry 5 days post transfection. The graph shows the percentage reduction in GFP MFI in cells transfected with each of the three repressor domain constructs relative to the mock transfected cells. The standard error of the mean values of triplicate wells is shown.

4.4.4 Comparison of the LZFP and RZFP fused to the KRAB domain.

The LZF as described in section 3.3.2 was cloned into the pV2 KRAB repressor domain construct as described in section 4.4.1 in the place of the RZF. The 1x and 3x reporter cell lines were transduced with the pV2 RZF-KRAB and pV2 LZF-KRAB repressor construct by $CaCl_2$ transfection. The cells were monitored by flow cytometry 5 days post transfection. There was no observable GFP extinction in the cells transduced with the pV2 LZF-KRAB repressor construct. However, GFP extinction was observed in cells transduced with the pV2 RZF-KRAB repressor construct (Figure 4.4).

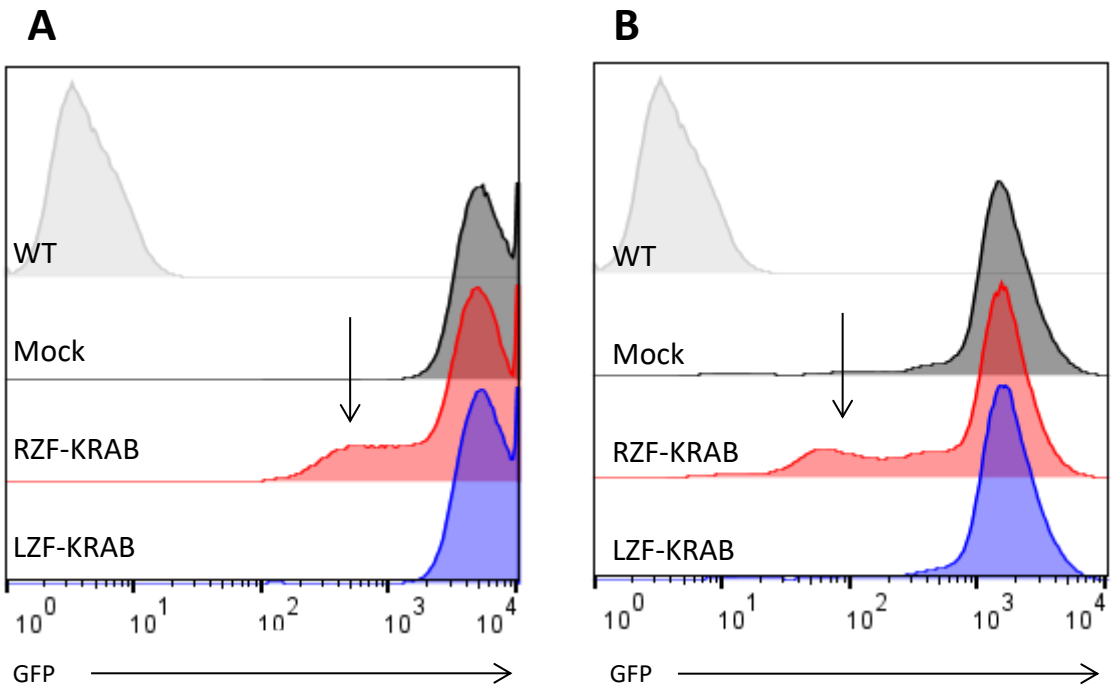
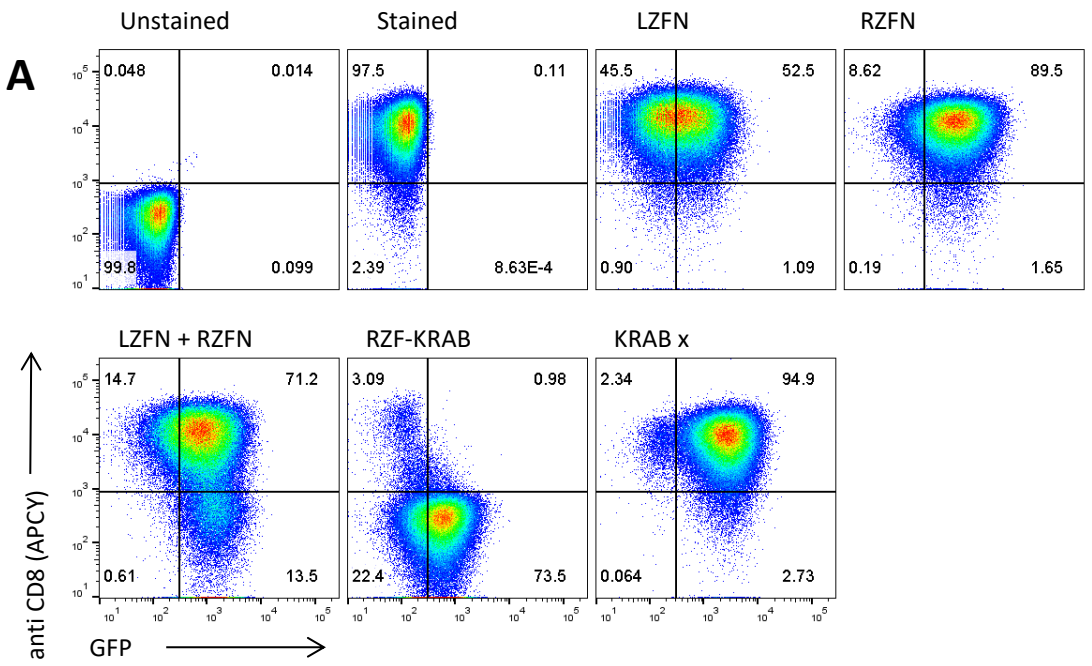
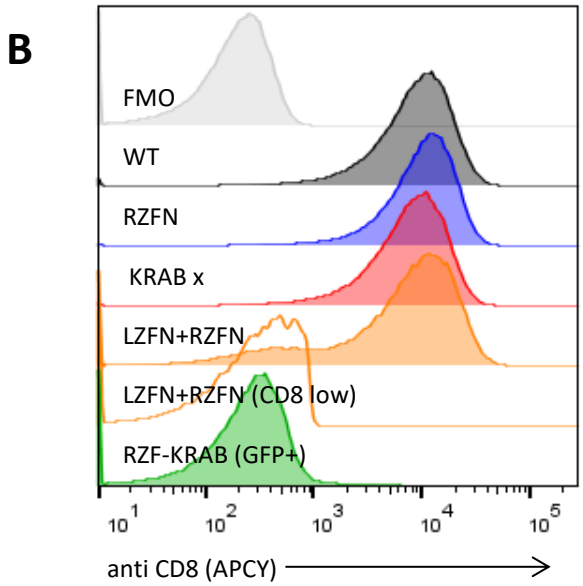


Figure 4.4. Comparison of the LZF-KRAB and RZF-KRAB constructs. The 1x (A) and 3x (B) reporter cell lines were transfected with either the LZF-KRAB or RZF-KRAB repressor domain constructs by $CaCl_2$ transfection. The expression of GFP was analysed by flow cytometry 5 days post transfection. A) WT 293T (light grey; MFI = 4.23), 1x mock (dark grey; MFI = 5182), 1x RZF-KRAB (red; MFI = 3045) and the 1x LZF-KRAB (blue; MFI = 5211). B) WT 293T (light grey; MFI = 4.23), 3x mock (dark grey; MFI = 1313), 3x RZF-KRAB (red; MFI = 759) and the 3x LZF-KRAB (blue; MFI = 1390). GFP extinction in a population of cells is indicated by an arrow. The data is representative of 3 experiments.

4.4.5 Validation of the ZF-KRAB repressor domain in Molt3

To establish whether the ZF-KRAB fusion was a valid alternative to the ZFN approach the RZF targeting CD8 fused to the KRAB repressor was cloned into the 2nd generation pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone. Molt3 cells were transduced with the LZFN, RZFN, both the LZFN and RZFN or RZF-KRAB by lentiviral transduction. The cells were cultured and analysed by flow cytometry at day 3, 7, 14, 28, 35, 49 and 77 post transduction. Three days post transduction there was no observable knock out in the LZFN and RZFN transduced cells, yet, 97.1% of the cells transduced with the RZF-KRAB were CD8 negative (after deducting background). There was an observable population of CD8low cells in the LZFN and RZFN transduced cells 7 days post transduction, this population was more clearly defined 14 days post transduction (Figure 4.5A). At two weeks post transduction, 16% of cells predicted to be transduced with both the LZFN and RZFN constructs were CD8 negative. In comparison, 99% of the cells transduced with the RZF-KRAB construct were CD8 negative. The MFI of CD8 was reduced by 43% compared to WT in the GFP+ portion of the LZFN and RZFN transduced cells (deducting background) and a reduction of 99.7% was observed when gating on the CD8- GFP+ population within the LZFN and RZFN transduced cells (deducting background). In comparison the MFI of CD8 within the GFP+ population of RZF-KRAB transduced cells was reduced by 100% compared to WT cells (deducting background) (Figure 4.5B).





Sample	Percentage reduction of CD8 MFI (%)
	100
WT	0
RZFN (GFP+)	-15
KRAB x (GFP+)	12
LZFN + RZFN (GFP+)	43
LZFN + RZFN (CD8 low)	100
RZF-KRAB (GFP+)	100

Figure 4.5. Comparison of ZFNs and ZF-KRAB gene silencing in Molt3 cells. Molt3 cells were transduced with the LZFN, RZFN, LZFN+RZFN, RZF-KRAB and KRAB x constructs by lentiviral transduction and monitored by flow cytometry at various time points. A) 2 weeks post transduction, cells transduced with the LZFN, RZFN and KRAB x constructs had an observable GFP+ population. The cells transduced with the RZF-KRAB had an observable GFP+CD8- population. There was a GFP+CD8- population present in the cells transduced with the LZFN + RZFN lentiviruses. B) The percentage reduction of CD8 MFI was determined relative to the WT untransduced cells (GFP- population), The FMO is shown in light grey (gating on GFP-), the WT is shown in dark grey (gating on GFP-), the RZFN is shown in blue (gating on GFP+), the KRAB x is shown in red (gating on GFP+), the LZFN+RZFN is shown in orange (gating on GFP+), the GFP+CD8- population within the LZFN+RZFN transduced cells is shown as an orange histogram with no filled colouring and the green histogram represents the GFP+ population of cells transduced with the RZF-KRAB lentivirus.

4.4.6 Long-term monitoring of the RZF-KRAB induced silencing in Molt3

The cells were monitored by flow cytometry 3, 7, 14, 28, 35, 49 and 77 days post transduction, at each time point the percentage of transduced cells that were CD8 negative was calculated. The proportion of CD8 negative cells within the cells transduced with both the LZFN and RZFN was approximately 20% consistently after day 7 monitoring. By comparison, 97% of cells transduced with the RZF-KRAB lentivirus were CD8 negative at 3 days post transduction, the frequency declined slightly over time with 87.8% of transduced cells being CD8 negative by day 77 (Figure 4.6). Cells transduced with the RZF-KRAB construct lost expression of GFP over time, at day 3 95% of cells transduced with the RZF-KRAB by lentiviral transduction were GFP+. By day 77, 11% of the cell population were GFP+, however, 88% of cells transduced with the RZF-KRAB construct were CD8 negative.

To determine whether repression of CD8 could be improved tandem repeats of the KRAB repressor fused to the RZF were generated and tested, however there was no benefit to tandem repeats of the repressor domain and this approach was not pursued further (Appendix Figure 10).

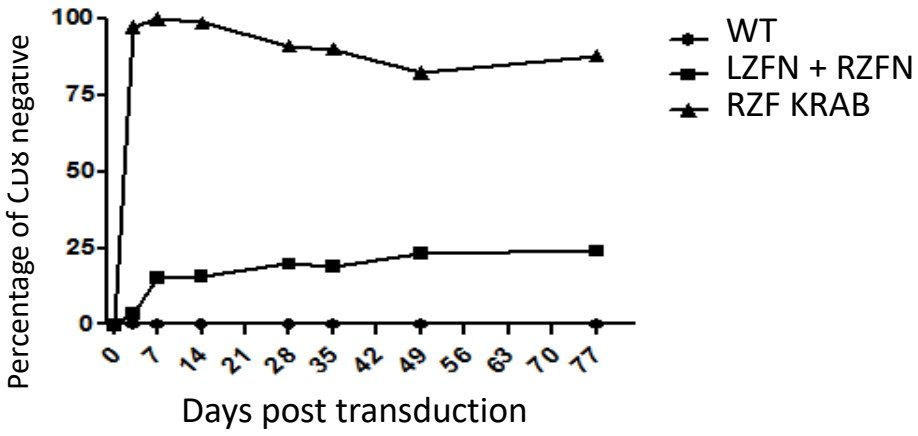


Figure 4.6. Long-term monitoring of Molt3 cells transduced with both the LZFN and RZFN and RZF-KRAB constructs. A graph displaying the percentage of CD8- cells within the transduced cells population over a 77 day time course. The data is representative of 3 experiments.

4.4.7 Confirmation of CD8a repression by QPCR

Cells were transduced with the RZFN and RZF-KRAB constructs by lentiviral transduction, the cells were analysed by flow cytometry 72 h post transduction by flow cytometry (Figure 4.7). At the 72 h time point RNA was extracted from the cells and the quality of the RNA was checked by running the samples on the bioanalyser. cDNA was synthesised from the RNA and used to perform QPCR. A 94% reduction in CD8A mRNA was observed in the RZF-KRAB transduced cells relative to the WT (Figure 4.7). The expression of CD8A in the RZFN was reduced by 41% relative to the WT despite the MFI of CD8 as measured by flow cytometry having no observable difference between RZFN and WT.

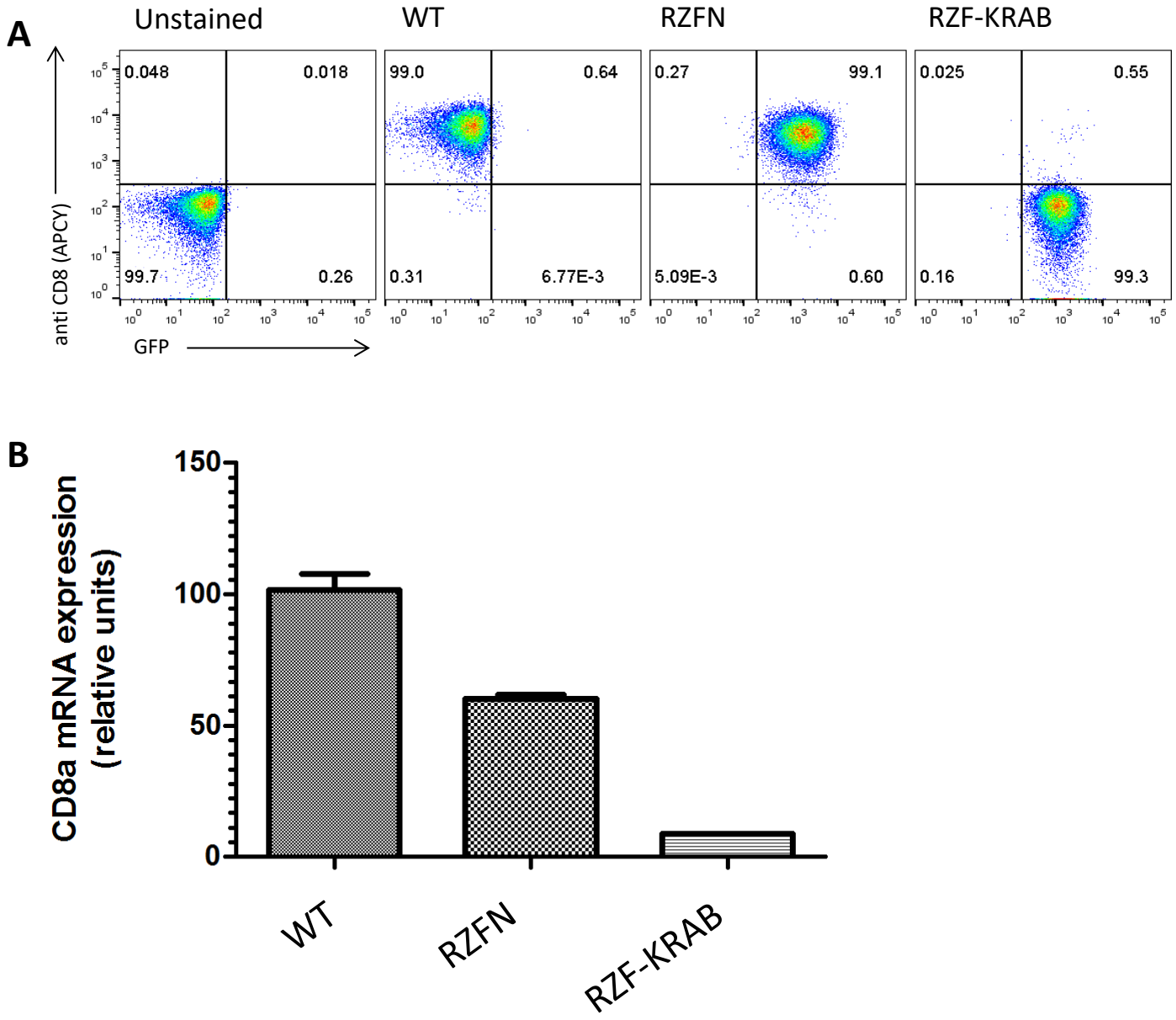


Figure 4.7. Assessment of the level of CD8 repression in Molt3 cells by QPCR. Molt3 cells were transduced with the RZFN and RZF-KRAB constructs by lentiviral transduction. Cells were monitored by flow cytometry 3 days post transduction. A) The dot plots of the cell samples 3 days post transduction. B) The relative levels of CD8A mRNA expression in the cells was determined by QPCR using 18S and β-actin as endogenous controls.. The primers used for the QPCR are listed in the Table 2 of the Appendix. The CD8A mRNA expression is shown as a percentage of the WT. Standard error of the mean of 3 replicates is shown. The data is representative of 2 experiments.

4.4.8 Long-term monitoring CD8 expression in Molt3 clones transduced with RZF-KRAB

Cells transduced with the RZF-KRAB construct lost expression of GFP over time and the frequency of CD8+ cells rose over time. It was hypothesised that either the CD8- transduced cells were outgrown by the untransduced population or the transduced cells were able to overcome the suppression. Molt3 cells were transduced with the RZFN and RZF-KRAB by

lentiviral transduction and monitored at 72 h post transduction and clones were generated by single cell cloning. The clones were screened 1 month post transduction by flow cytometry and 5 CD8- clones were selected. The 5 clones were cultured for a further 2 months before being analysed by flow cytometry to assess whether these clones harboured a CD8+ population after continuous culture. The CD8 expression by the 5 clones remained constant over the monitoring period suggesting that the cells were not overcoming the KRAB repression (Figure 4.8A). To assess the level of CD8A mRNA expressed by the clones, RNA was extracted from 5 of the clones in addition to cells transduced with the RZFN and RZF-KRAB and WT untransduced cells. cDNA synthesised from the RNA was analysed by QPCR. Expression of CD8A mRNA was reduced by 90% in the bulk RZF-KRAB transduced cells. Expression of CD8A mRNA was reduced by 89, 87, 85, 90 and 93% in clones 4, 6, 7, 10 and 13 respectively (Figure 4.8B). Therefore, there was no improvement of CD8A repression in the clonal derivatives compared to the bulk RZF-KRAB transduced cells. There was a 41% reduction of CD8A expression in the RZFN transduced cells compared to the WT untransduced cells.

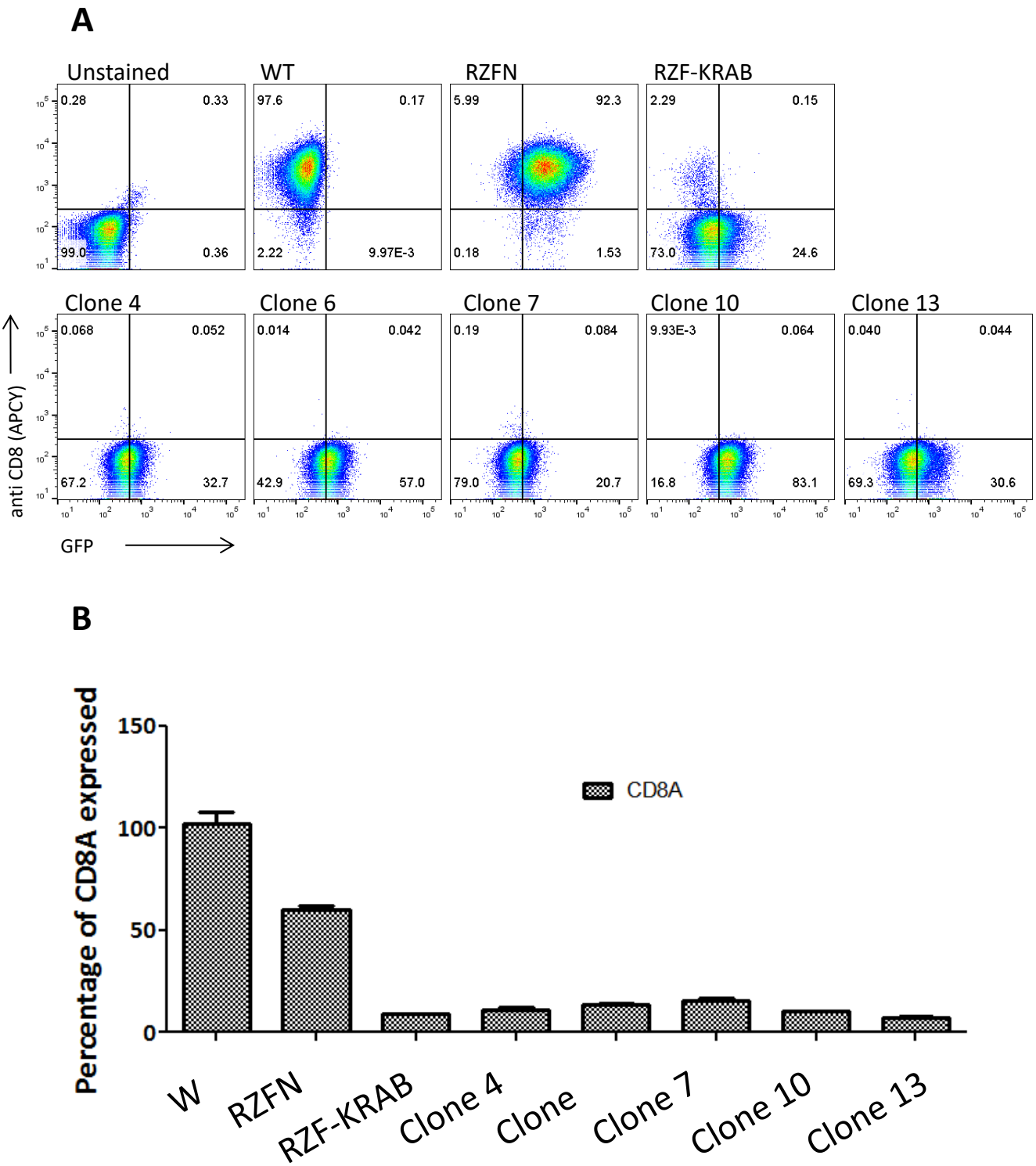


Figure 4.8. Monitoring of the clonal derivatives of RZF-KRAB transduced Molt3 cells. Clonal derivatives of the bulk RZF-KRAB transduced Molt3 cells were monitored by flow cytometry and QPCR to assess whether the population of CD8 negative cells remained constant over time and whether the level of CD8A mRNA expressed was reduced in the clonal derivatives compared to the bulk RZF-KRAB transduced cells. A) Dot plots of the WT, RZFN transduced, RZF-KRAB transduced and RZF-KRAB clones 4, 6, 7, 10 and 13 three months post transduction. B) A graph displaying the percentage of CD8A mRNA expressed in comparison to the WT untransduced cells of the bulk RZFN transduced, bulk RZF-KRAB transduced, RZF-KRAB clones 4, 6, 7, 10 and 13. QPCR was performed using 18S and β-actin as endogenous controls. The mean and standard error from the mean of 3 replicates is shown; the data is representative of 2 experiments.

4.4.9 Comparison of ZF-KRAB and ZFN in human primary T-cells

In order to be useful therapeutically, a gene silencing system would need to be effective in primary T-cells. I next examined this possibility using CD8+ T cells isolated from the buffy coats of three donors (section 2.3.4). The CD8+ T cells were transduced with the LZFN, RZFN, LZFN and RZFN, RZF-KRAB and KRAB x constructs by lentiviral transduction and monitored by flow cytometry. The frequency of cells transduced with both the LZFN and RZFN was estimated by multiplying the frequency of cells transduced with the single transductants (LZFN or RZFN). Two weeks post transduction 10, 8 and 10% of GFP+ cells within the LZFN and RZFN transduced cells were CD8low in donors A, B and C, respectively after deducting background. This was estimated to represent 29, 28 and 24% of cells transduced with both the LZFN and RZFNs. In comparison, 82, 88 and 88% of cells transduced with the RZF-KRAB lentivirus were CD8low in donors A, B and C respectively after deducting background (Figure 4.9) suggesting that the ZF-KRAB approach is far more effective at inducing loss of expression of CD8 from the cell surface than the ZFN approach in primary CD8 T-cells.

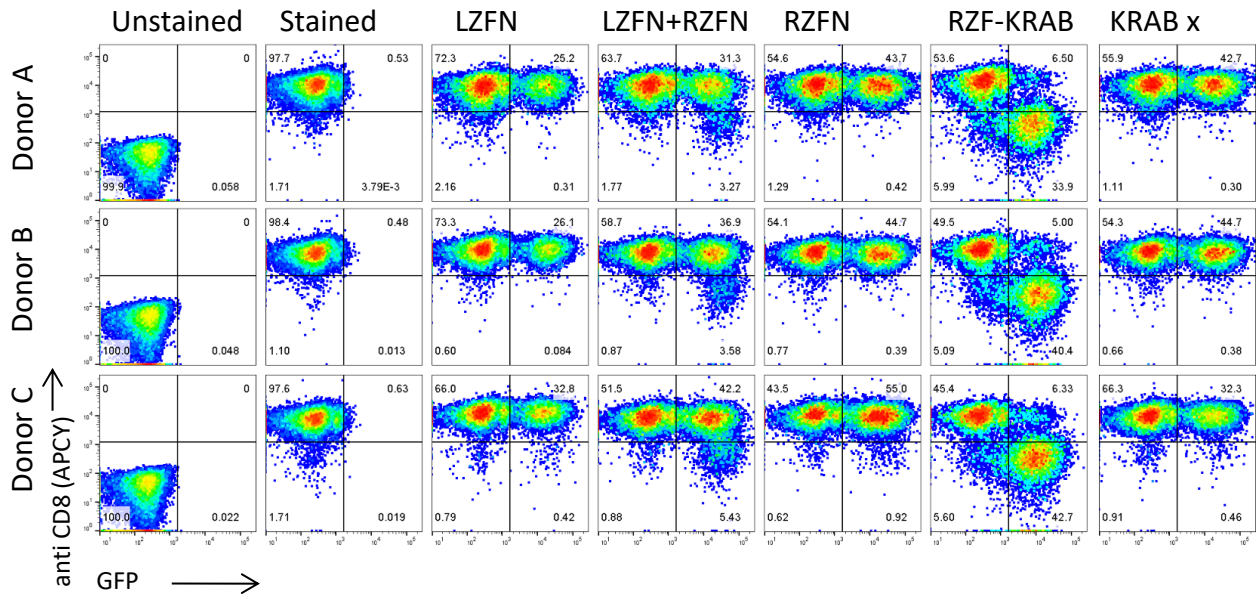
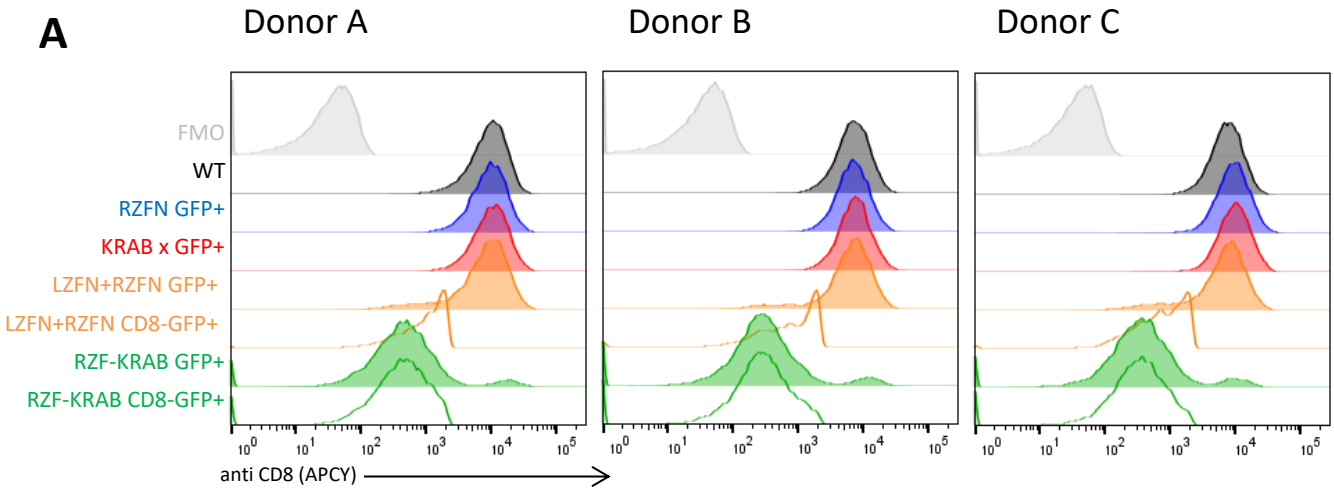


Figure 4.9. Comparison of the ZF-KRAB approach in human CD8+ T-cells. CD8 T-cells from three donors were transduced with LZFN, RZFN, LZFN and RZFN, RZF-KRAB and KRAB x constructs by lentiviral transduction. Dot plots of the cells 2 weeks post transduction, a CD8- population is observed in cells transduced with the LZFN and RZFN and RZF-KRAB lentivirus in donors A, B and C.

The level of CD8 expressed by the CD8low population of cells transduced with the LZFN and RZFN and RZF-KRAB lentivirus was determined by gating on the cells and calculating the percentage reduction in MFI of the different populations. For Donor A, there was a 91% and

96% reduction in CD8 MFI of the CD8low cells transduced with LZFN and RZFN and RZF-KRAB respectively. For Donor B, there was a 91% and 97% reduction in CD8 MFI of the CD8low cells transduced with LZFN and RZFN and RZF-KRAB respectively. For Donor C, there was a 91% and 96% reduction in CD8 MFI of the $CD8^{low}$ cells transduced with LZFN and RZFN and RZF-KRAB respectively (Figure 4.10). This suggests that the ZF-KRAB approach is a viable and more effective alternative to ZFNs in this system.

A



B

Donor A

Sample	Percentage reduction of CD8 MFI (%)
Unstained	100
Stained	0
RZFN	-3
KRAB x	-10
LZFN + RZFN (GFP+)	25
LZFN + RZFN (CD8-GFP+)	91
RZF-KRAB (GFP+)	95
RZF-KRAB (CD8-GFP+)	96

Donor B

Sample	Percentage reduction of CD8 MFI (%)
Unstained	100
Stained	0
RZFN	1
KRAB x	-6
LZFN + RZFN (GFP+)	22
LZFN + RZFN (CD8-GFP+)	91
RZF-KRAB (GFP+)	96
RZF-KRAB (CD8-GFP+)	97

Donor C

Sample	Percentage reduction of CD8 MFI (%)
Unstained	100
Stained	0
RZFN	-26
KRAB x	-32
LZFN + RZFN (GFP+)	24
LZFN + RZFN (CD8-GFP+)	91
RZF-KRAB (GFP+)	95
RZF-KRAB (CD8-GFP+)	96

Figure 4.10. Comparison of the ZFN and ZF-KRAB approach in human CD8+ T-cells. CD8 T-cells from three donors were transduced with LZFN, RZFN, LZFN and RZFN, RZF-KRAB and KRAB x constructs by lentiviral transduction. A) Histogram plots of the CD8 MFI of donors A, B and C are shown. FMO shown in light grey, WT in dark grey, RZFN gating on GFP+ (blue), KRAB x gating on GFP+ (red), LZFN +RZFN gating on GFP+ (orange), LZFN +RZFN gating on GFP+CD8- (orange histogram with no coloured fill), RZF-KRAB gating on GFP+ (green) and RZF-KRAB gating on GFP+CD8- (green histogram with no coloured fill). B) The percentage reduction in CD8 MFI of each of the samples relative to the WT is shown in the table below each of the histogram plots after deducting the background fluorescence.

4.4.10 Monitoring the longevity of gene silencing in T-cells

T-cells transduced with the LZFN, RZFN, LZFN and RZFN, RZF-KRAB and KRAB x by lentiviral transduction were monitored by flow cytometry for 9 weeks. The CD8low cells within the LZFN + RZFN transduced and RZF-KRAB cell populations declined over the monitoring course and were barely detectable at the end of the monitoring period (Figure 4.11). The decline in CD8- cells suggests that the untransduced cells within these populations proliferated at a greater level than CD8- cells. However, this finding was not unexpected as CD8 is thought to be involved in basal signalling and T-cell turnover.

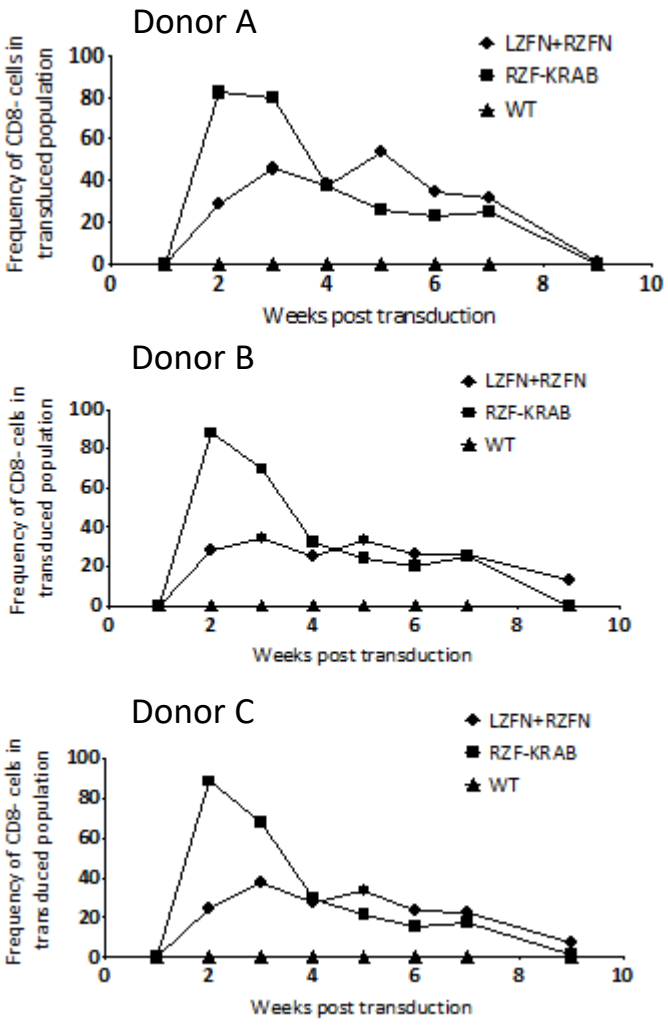


Figure 4.11. Monitoring the CD8low population within the ZFN and ZF-KRAB transduced T-cells. CD8 T-cells from three donors were transduced with LZFN, RZFN, LZFN and RZFN, RZF-KRAB and KRAB x constructs by lentiviral transduction. The cells were analysed by flow cytometry for nine weeks post transduction. A time course of the frequency of CD8- cells within the transduced cell population of cells transduced with the LZFN+RZFN and RZF-KRAB by lentiviral transduction for Donor A, B and C. The frequency was estimated using the frequency of GFP+ cells 2 weeks post transduction and the frequency of CD8- cells at each time point. The data is representative of 2 experiments.

4.4.11 Confirming the reduction of CD8A using RZF-KRAB by QPCR in primary T-cells

To determine whether the level of reduction of CD8 MFI correlated with the mRNA levels, the cells were sorted by flow cytometry gating on GFP+CD8- population of cells within the RZF-KRAB transduced sample, the GFP+ population of RZFN cells and KRAB x cells. RNA was extracted from the untransduced and sorted cells. QPCR confirmed that there was 100% reduction in CD8A mRNA in the CD8- cells transduced with the RZF-KRAB compared to WT untransduced cells in donor A, B and C (Figure 4.12). Expression of CD8A was consistent between the RZFN and KRAB x controls despite differing from the WT.

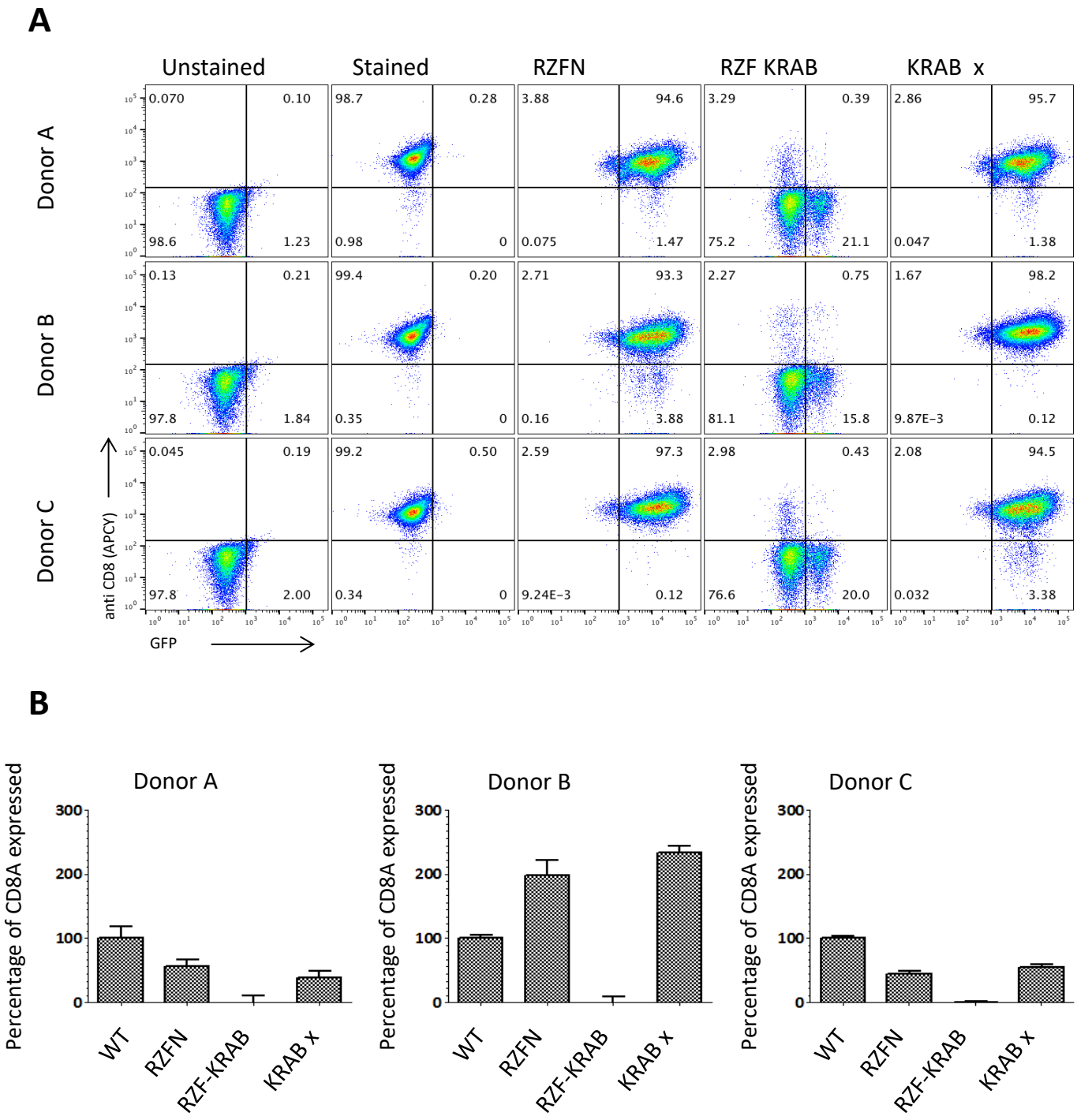


Figure 4.12. Confirmation of CD8A silencing by QPCR. CD8+ T-cells from three donors were transduced with RZFN, RZFKRAB and KRAB x by lentiviral transduction. The transduced cells were sorted from the bulk population by FACS, the purity of the sorted transduced T-cells was established by flow cytometry. A) Dot plots are shown for the three donors (A, B and C). B) QPCR was performed using RNA extracted from the sorted RZFN, KRAB x and RZF-KRAB cells in addition to the WT cells from donors A, B and C. β-actin and 18S were used as endogenous controls. All of the primer sequences can be found in Appendix Table 2.

4.4.12 Assessment of off target effects by RNAseq

The CD8A-specific ZF-KRAB domain was composed of 6 modules which targeted a 19bp stretch of DNA. RNAseq was performed to assess whether there were any off target effects as a result of using the ZF-KRAB approach to silence expression of the CD8A gene using RNA

extracted from the purified GFP+ RZFN, GFP+CD8- RZF-KRAB and WT T-cells from donor B in section 4.4.11. The RNAseq was performed by the University of Utah sequencing core as described in section 2.4.15

The 19bp target sequence was run through the blastn aligorithm NCBI (Altschul et al. 1990) to identify any genes that were likely to be off target sites. The 5 most likely off targets sites were identified as the DDR1, PPP1R16B, ZNF844, WTIP and SLCO3A1 genes. Alignment of the ZF binding site to each of these genes is shown in Figure 4.13A. However, no significant alteration of the expression of these genes was observed in the RNAseq data. The analysis revealed that there was significant change ($p=<0.05$) of expression in 4 genes in the RZF-KRAB transduced cells compared to the RZFN and WT cells. Of the 4 genes identified as being differentially expressed the CD8A showed the most significant change. The other 3 genes identified were Adrenergic, Beta, Receptor Kinase 2 (ADRBK2), gephyrin (GPHN) and solute carrier family 17 member 9 (SLC17A9) (Figure 4.13B). There was a 98.84, 97.35, 91.87 and 97.11% reduction in the number of reads detected for CD8A, ADRBK2, GPHN and SLC17A9 for the RZF-KRAB sample in comparison to the WT and RZFN samples (Figure 4.13B).

A

Sequence	Gene
TGTCGCCGCTGGATCGGAC	CD8A
-GTCGCCGCTGGAT-----	DDR1
TGTCGCCGCTGGA------	PPP1R16B
TGTCGCCGCTGGA------	ZNF844
----GCCGCTGGATCGG--	WTIP
-----CCGCTGGATCGGA-	SLCO3A1

B

	Sample		
Gene	WT	RZFN	RZF-KRAB
CD8A	100	94.23	1.16
GPHN	100	74.94	2.65
ADRBK2	100	100.95	8.13
SLC17A9	100	156.99	2.89

Figure 4.13. Analysis of RNAseq data. A) An alignment of the 5 most likely off target sites as determined by BLAST, the genes in which the sequences are found is listed. Mismatches are shown as '-'. There was no significant difference between the expression pattern of these genes in the WT, RZFN and RZF-KRAB samples. B) A table displaying the percentage expression of the genes found to have significantly different expression patterns between the WT, RZFN and RZF-KRAB samples.

The 19bp ZFP target sequence was aligned to the exonic sequence of the 4 genes identified by the RNAseq data (Figure 4.14). This alignment revealed that the ADRBK2, GPHN and SLC17A9 genes do not align with the 19bp target sequence. The longest sequence homology between the 19bp ZFP target sequence and the exonic sequences of ADRBK2, GPHN and SLC17A9 are 7, 6 and 6 bp respectively suggesting that the ZFP does not directly bind to these genes and induce gene silencing. It seems likely that the reductions observed by RNAseq for ADRBK2, GPHN and SLC17A9 are not genuine. QPCR experiments aimed at determining whether these potential off target effect are real are currently in progress. It will be important to establish whether silencing of CD8A could have an indirect effect on these genes.

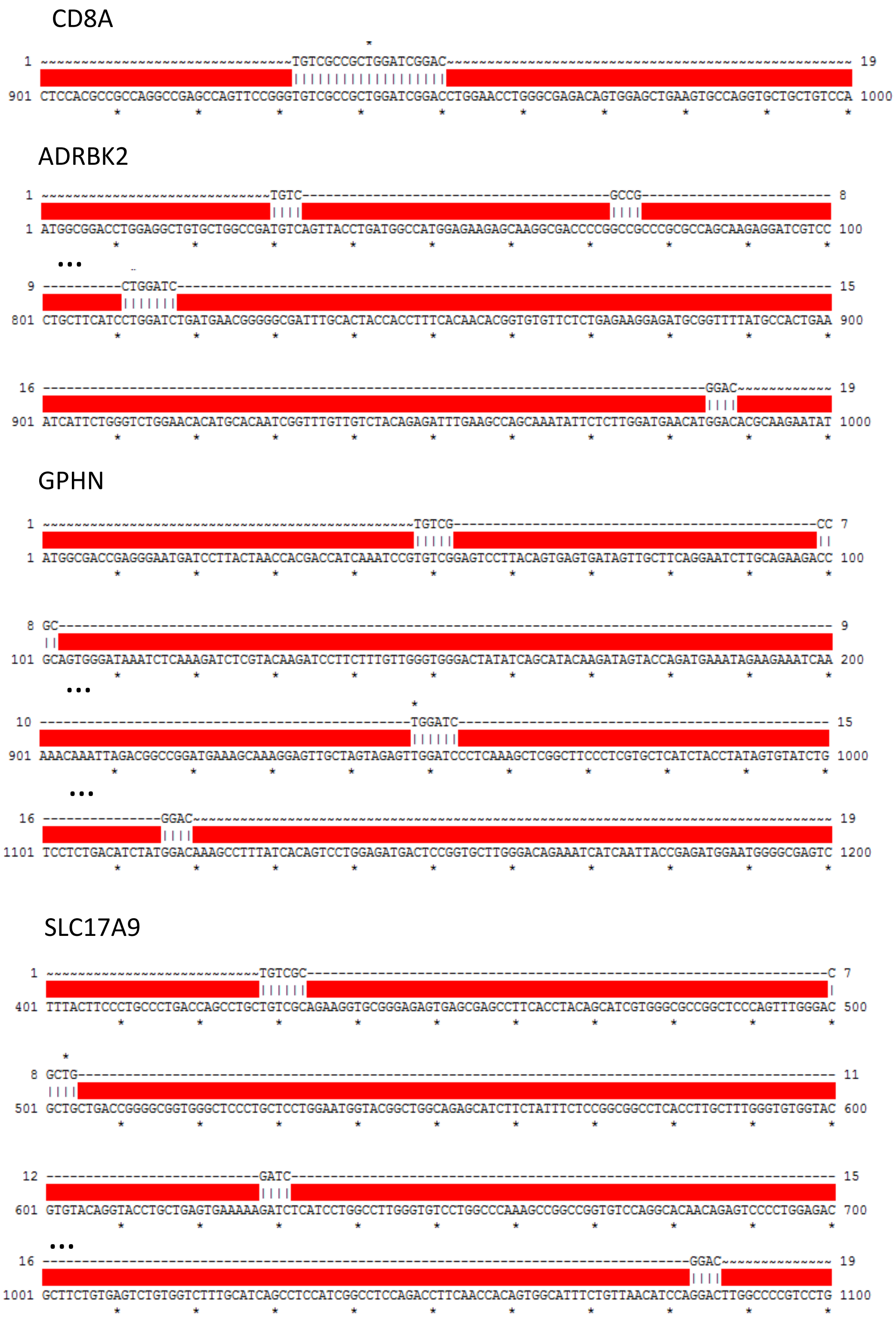


Figure 4.14. Alignment of 19bp ZFP target site and exonic sequences of the genes identified by RNAseq. The alignments were performed using APE software. Mismatches are shown in red. The nucleotide position in the gene is shown by the numerical values on the left of the alignment.

4.4.13Assessment of level of gene expression of CD8A, GPHN, ADRBK2 and SLC17A9 by QPCR

The data from the RNAseq experiments (Figure 4.13) revealed that the expression of 4 genes (CD8A, GPHN, ADRBK2 and SLC17A9) were significantly reduced in the RZF-KRAB transduced cells compared to the WT and RZFN control samples. In order to confirm whether the expression of these genes was altered, QPCR was performed using the RNA from donor B on which the RNAseq was performed and on two additional donors (A and C). QPCR was performed on the RNA using the Taqman approach as described in section 2.4.8.4. There was a 99.9% reduction in the expression of GPHN, ADRBK2 and SLC17A9 mRNA compared to CD8A in the WT samples from all 3 donors, suggesting that GPHN, ADRBK2 and SLC17A9 are expressed at substantially lower levels relative to CD8A (Figure 4.15A). The QPCR indicated that the levels of mRNA for CD8A, GPHN, ADRBK2 and SLC17A9 were significantly reduced in the RZF-KRAB samples compared to the WT and RZFN samples in all three donors (Figure 4.15B) (p value <0.05).

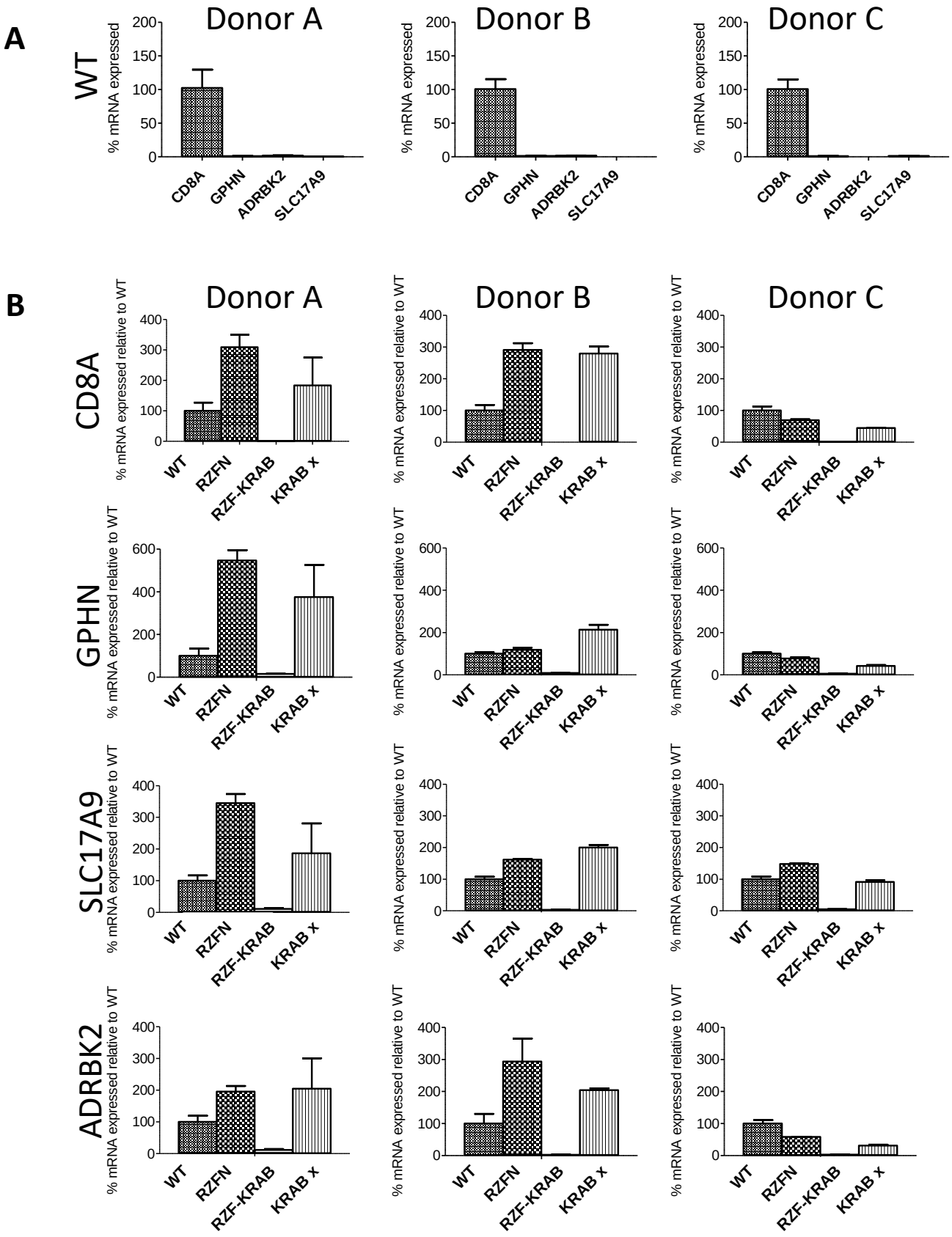


Figure 4.15. Validation of gene expression by QPCR. CD8+ T-cells from three donors were transduced with RZFN, RZFKRAB and KRAB x by lentiviral transduction. The transduced cells were sorted and QPCR was performed on the extracted RNA. A) The relative mRNA expression of the GPHN, ADRBK2 and SLC17A9 compared to CD8A. B) The percentage of mRNA expressed for CD8A, GPHN, ADRBK2 and SLC17A9 relative to the WT sample for each of the three donors. All data was normalised using β2m as an endogenous control. The data shown is from a single repetition of the experiment

4.4.14Combined delivery of ZF-KRAB and TCR in a single vector

The use of T-cell checkpoint inhibitor antibodies for cancer immunotherapies has shown great promise. However, these systemic approaches 'take the brakes off' all T-cells and come with some severe side effects. Future approaches should aim to delimit only cancer-specific T-cells. One way of doing this would be to knockdown genes like PD-1 and CTLA-4 in tumour-reactive TIL or to deliver a tumour-specific TCR to T-cells while simultaneously silencing genes that regulate T-cell responses. Combined delivery of ZF-KRAB with an antigen-specific TCR might enable such an approach. I next set out to test whether a ZF-KRAB construct targeting CD8 could be delivered to Molt3 within a lentivirus that delivered a TCR to change their specificity.

In order to generate a combined delivery vector, the MEL5 or PPI TCRα and TCRβ chains were cloned between NheI and XhoI restriction sites downstream of the U6 promoter in the pELNSxv lentiviral backbone (Dull et al. 1998). This cassette system allowed for alterative TCR chains to be substituted in future experiments. The CD8 ZF-KRAB sequences were inserted downstream from the TCR chains between BamHI and SalI restriction sites to allow for substitution of these domains in future experiments. As a control, a construct was generated which was composed of the MEL5 or PPI TCRα and TCRβ chains in combination with a rCD2 marker. A diagram of these constructs is shown in Figure 4.16A. Molt3 cells were transduced with the rCD2 TCR or CD8 RZF-KRAB TCR by lentiviral transduction. The cells were cultured and monitored by flow cytometry 1 week post transduction. A population of TCR+ cells was observed within the Molt3 cells transduced with the rCD2 TCR and CD8 RZF-KRAB TCR constructs. The TCR+ population of CD8 RZF-KRAB TCR transduced cells expressed CD8 at a lower level than the untransduced cells indicating that the CD8 RZF-KRAB protein was functional (Figure 4.16B).

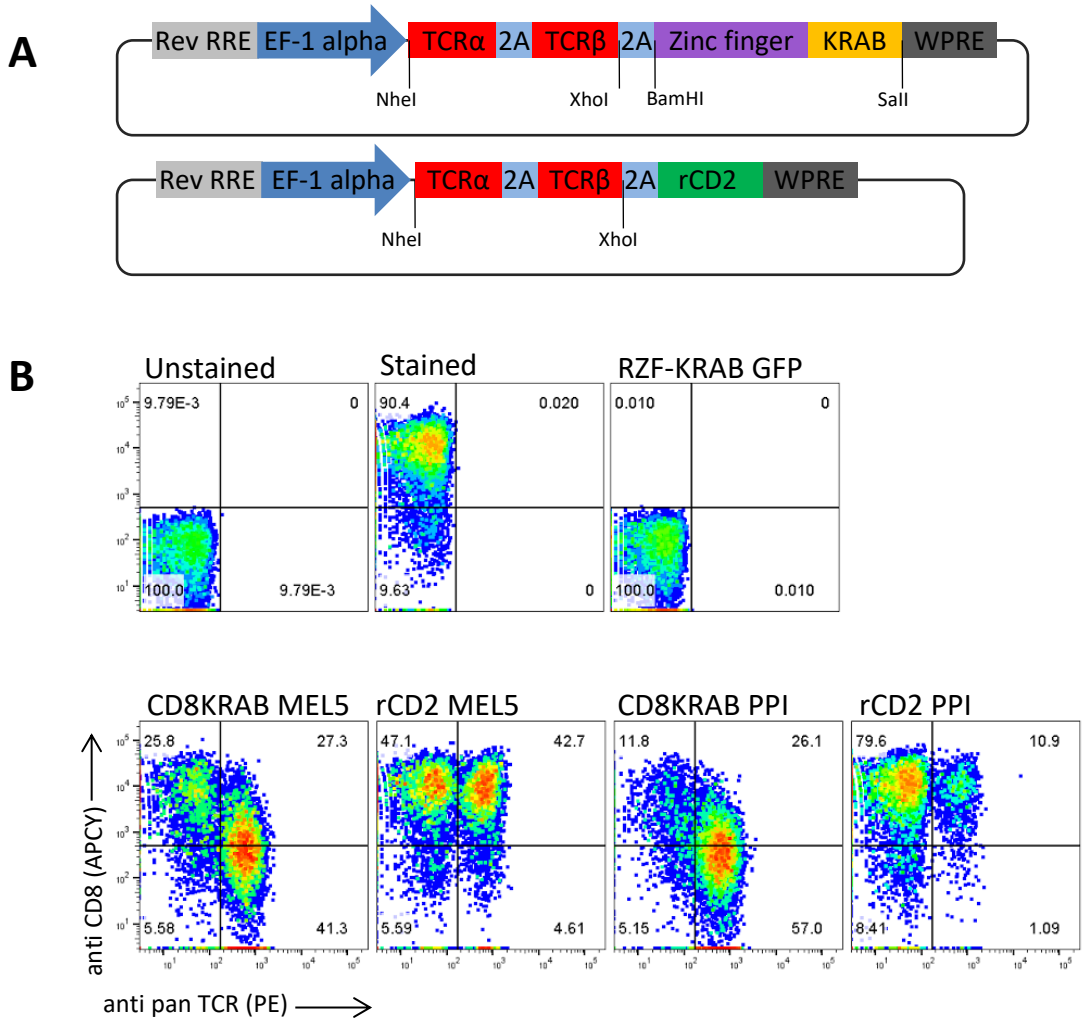


Figure 4.16. Validation of the TCR constructs. A) A diagrammatic representation of the ZF-KRAB TCR and rCD2 TCR constructs. B) Molt3 cells transduced with the RZF-KRAB TCR and rCD2 TCR constructs by lentiviral transduction were analysed by flow cytometry 1 week post transduction.

4.4.15 Dextramer staining in Molt3

To determine whether the RZF-KRAB TCR construct could silence the CD8A gene and direct Molt3 cells to recognise the cognate antigen. Molt3 cells were transduced with the RZF-KRAB, rCD2 PPI TCR and RZF-KRAB PPI TCR constructs by lentiviral transduction and stained with dextramers two weeks post transduction. The multimer staining process was optimised to maximise the frequency of cells which stained. Optimisation with PKI and tetramers / dextramers can be found in Appendix Figure 11. Dextramers were titrated using concentrations – 0.3, 0.03 and 0.003 µg/ µL, the cells were then analysed by flow cytometry.

The RZF-KRAB TCR construct functioned well as CD8A silencing and expression of the PPI TCR was observed. Of the TCR+ population of cells 40, 26 and 5% stained with 0.3, 0.03 and

0.003 µg/ µL concentrations of RQF dextramer. In comparison 52, 27 and 5% of TCR+ rCD2 PPI TCR transduced cells stained with the 0.3, 0.03 and 0.003 µg/ µL RQF dextramer (Figure 4.17). The increased frequency of dextramer positive cells in the rCD2 PPI TCR transduced cells compared to the RZF-KRAB PPI TCR transduced cells was expected as CD8 plays an important role in stabilising the TCR/pMHC interaction (Wooldridge et al. 2005).

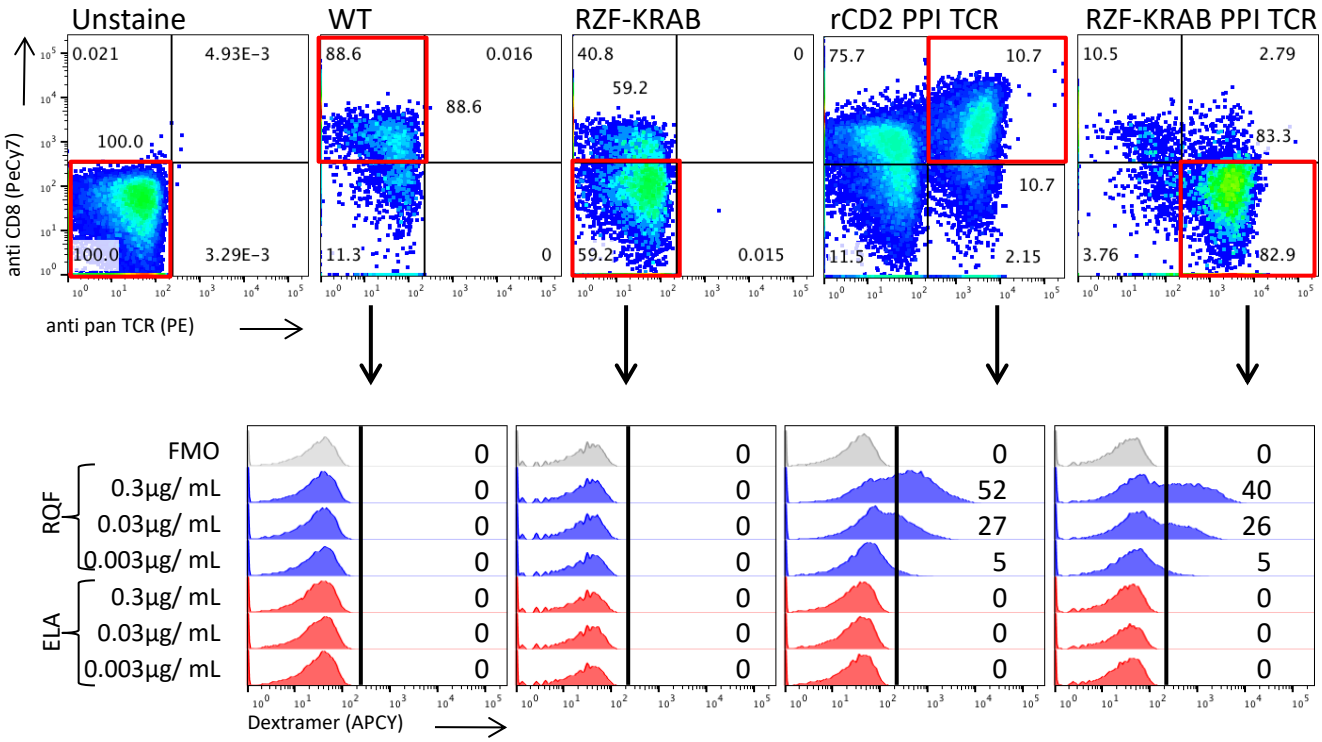


Figure 4.17. Dextramer staining of Molt3 cells transduced with rCD2 TCR and RZF-KRAB TCR constructs. Molt3 cells were transduced with the RZF-KRAB, rCD2 PPI TCR and RZF-KRAB PPI TCR constructs by lentiviral transduction. Cells were incubated with PKI and stained with dextramer two weeks post transduction. A) Dot plots of the transduced cells and appropriate gate used to monitor dextramer staining. B) The histogram plots show the FMO (grey), cells stained with the RQF dextramer (blue), ELA irrelevant dextramer (red). The frequency of dextramer positive cells within the TCR positive population of transduced cells is shown. The data is representative of two experiments.

4.5 Discussion

Nuclease based gene editing has risen in popularity over the last few years due to the ease and efficiency of gene disruption. However, there are risks associated with the expression of a nuclease, namely off target mutagenesis which can lead to malignancy. There is a need for non-nuclease based gene silencing technology as this would be safer in the clinic. Here I developed a non-nuclease based gene silencing system utilising the ZFPs platform. Of the 7 repressor domains tested the KRAB repressor domain was shown to be superior in the reporter cell line. The RZF-KRAB construct was validated in both the Molt3 cell line and in human primary T-cells and led to a reduction of CD8 expression on the cell surface as confirmed by flow cytometry and QPCR. RNAseq performed on RNA extracted from WT and T-cells transduced with the RZF-KRAB and RZFN constructs revealed three genes in addition to CD8A which showed reduced expression in the RZF-KRAB transduced cells relative to the RZFN transduced and WT untransduced cells. QPCR confirmed that expression of these genes was significantly altered in RZF-KRAB compared to WT and RZFN cells. An explanation for the alteration in expression of these genes has not yet been identified. A construct was generated which incorporated a ZF-KRAB and a TCR as a single entity. The TCR ZF-KRAB construct was validated in Molt3 cells which were found to express the TCR on the cell surface and expressed CD8A at a reduced level relative to untransduced and control cells. Overall, the RZF-KRAB construct is an effective and nuclease free use of ZFPs allowing for silencing of genes without the risk of potential off target nuclease activity.

4.5.1 Length of ZF and affinity for target site

The ZF pair used in this chapter consisted of a RZF formed by 6 ZF modules (targeting a 19bp stretch of DNA) and a LZF formed from 5 modules (targeting a 16bp stretch of DNA). The LZF and RZF were fused to the KRAB repressor domain and validated in the 1x and 3x reporter cell lines. Of the two constructs only the RZF-KRAB was functional. It is suspected that the RZF has an increased affinity for the DNA binding site due to the additional ZF module. Increased repression of a AmCyan fluorescent protein was observed when repressor domains were fused to a 6 module ZFP in comparison ZFPs with less modules (Lohmueller et al. 2012). These results indicate that the length of the ZFP affects the affinity that the ZFP

has for its corresponding target site. In addition, longer ZFPs may increase the specificity of the ZFP and reduce potential off target binding.

4.5.2 RNAseq

RNAseq identified 3 additional genes which had significantly altered expression in the RZF-KRAB transduced cells compared to both the WT and RZFN transduced cells these were GPHN, SLC17A9 and ADRBK2. The GPHN membrane plays a key role in anchoring glycine receptors to cytoskeleton and subsynaptic microtubules (David-Watine 2001). The SLC17A9 gene encodes a lysosomal ATP transporter which is expressed by chromaffin cells, pancreatic and T-cells (Cao et al. 2014). ADRBK2 encodes a β- adrenergic receptor kinase which phosphorylates GPCRs allowing β-arrestin binding and initiation of receptor endocytosis (Luttrell & Lefkowitz 2002; Feng et al. 2005). The function of these genes would suggest that repression of these genes is unlikely to be a result of the proteins being involved in the downstream pathways of CD8. The likelihood of direct off target binding was assessed by aligning the 19bp target sequence with the genomic sequences of the identified genes. There is minimal alignment of the 19bp target sequence and the genomic sequence suggesting that unlikely that the altered expression is due to a direct silencing. To assess the expression of CD8A, GPHN, SLC17A9 and ADRBK2 in the RZF-KRAB transduced cells relative to the WT and RZFN samples QPCR was performed using RNA extracted from three donors, the data from a single QPCR experiment suggested that expression of CD8A, GPHN, SLC17A9 and ADRBK2 was significantly reduced in in the RZF-KRAB samples of three donors. Unfortunately, due to time constraints, the genes identified by the RNAseq data were not fully confirmed and further experiments are required to confirm the expression of these genes and to investigate the reasoning

4.5.3 Immunogenicity and persistence

In this study the frequency of CD8 negative cells within the bulk transduced population initially declined in Molt3 then remained stable over the monitoring period. The frequency of CD8 negative cells within the bulk transduced T-cell population declined steadily over time and the cells were undetectable at 11 weeks post transduction. This decline of CD8 negative T cells over time is thought to be due to the untransduced T-cells outgrowing the

CD8 negative T-cells. The CD8 molecule is thought to play an important role in basal signalling in T-cells, therefore the CD8 negative cells generated in this chapter were at a survival disadvantage to the untransduced T-cells. Previous studies have highlighted the importance of 'co-receptor tuning' for homeostasis, suggesting that CD8 expression plays an important role in the long-term survival of T-cells (Park et al. 2007; Zuniga-Pflucker 2007), as basal signalling is thought to occur though interactions between TCR and self pMHC interactions (Reviewed by Hogquist & Jameson 2014). In hindsight, CD8A was not an ideal target as proof of principle in T-cells.

A previous clinical trial evaluating the adoptive transfer of T-cells genetically modified to express a HIV specific TCR transfer via retroviral transduction revealed that genetic modification of T-cells led to cells becoming immunogenic whereby 5/6 patients rejected the genetically modified cells (Riddell et al. 1996). However, more recent clinical trials assessing the safety and persistence of genetically engineered T-cells have been promising. A clinical trial to evaluate the safety and feasibility of genetic modification of autologous T-cells using ZFNs to modify the CCR5 gene concluded that the use of ZFNs did not lead to the cells becoming immunogenic as they were found to persist long-term after transfer and the frequency of biallelic knockout correlated to the level of viral load with high viral load being associated with lower frequency of CCR5 KO (Tebas et al. 2014).

4.5.4 Advances in CRISPR interference technology

Since the start of my PhD, CRISPR/Cas9 technology has been identified and utilised for gene editing. Recent developments have focused on gene silencing with CRISPR based technologies. An inactive form of Cas9 termed dCas9 was generated by two nucleotide substitutions D10A and H840A (Qi et al. 2013). Fusion of the nuclease inactive dCas9 and the KRAB repressor has led to epigenetic silencing of the TP53 gene (Lawhorn et al. 2014) and the HS2 enhancer (Thakore et al. 2015). The ZF-KRAB approach developed in this chapter has clear advantages over the CRISPR/dCas9-KRAB approach as the small size of the ZF-KRAB construct allows for co-transfer of TCR chains within a single viral vector.

4.5.5 Summary

There is a significant need for the generation of novel non-nuclease based gene silencing technologies as nuclease based gene editing technologies have an inherent risk of malignancy resulting from off target nuclease activity. To address this issue I have developed a system which incorporates ZF DNA binding proteins and the KRAB repressor domain. The ZF-KRAB sequence was incorporated into a single plasmid with TCRα and TCRβ which allowed for specific gene knockout and redirection of Molt3 cells. Delivery of non-nuclease based gene silencing tools and an antigen specific TCR within a single plasmid has great potential for therapeutic use.

5 Development of an 'all in one' CRISPR/Cas9 lentiviral system to engineer isogenic cells deficient in MR1

5.1 Introduction

5.1.1 MAITS and MR1 restricted T-cells

MAITs are the most frequently occurring unconventional T-cell subset (Sandberg et al. 2013) and are found at high frequencies in the gut lamina propria (Huang et al. 2009), liver and mucosal surfaces (Kurioka et al. 2015). Typically, MAIT cells are identified by their semi invariant TCRs (TRAV1-2, TRAJ33/12/20) (Tilloy et al. 1999). A subset of MR1 restricted T-cells have recently been identified which do not express the canonical TRAV1-2 chains (Meermeier et al. 2016) and it may yet prove that there is a wide family of MR1-restricted T-cells in addition to MAITs. Unlike conventional T-cells, MAITs and MR1 restricted T-cells do not recognise peptide antigens restricted by classical MHCI or MHCII molecules. Instead, these unconventional T-cells are activated by pathogen-derived metabolites from a range of microbes (Le Bourhis et al. 2010) presented in the binding groove of MR1. MAITs are described as innate like as their effector functions are acquired during thymic selection and therefore MAITs have a rapid response time and do not require prior antigen priming (Gold et al. 2013). MAITs play a key role in controlling yeast and bacterial infections and have been shown to be activated by metabolites produced by a range of yeast species *Candida albicans*, *Candida glabrata* and *Saccharomyces cerevisiae* (Le Bourhis et al. 2010) and bacterial species such as *Escherichia coli* (Gold et al. 2010), *Klebsiella pneumoniae*, *Lactobacillus acidophilus* (Le Bourhis et al. 2010), *mycobacteria* (Gold et al. 2010), *Pseudomonas aeruginosa*, (Le Bourhis et al. 2010), *Staphylococcus aureus*, (Gold et al. 2010; Le Bourhis et al. 2010), *Shigella dysenteriae, Staphylococcus epidermidis* and *Shigella flexneri*, (Le Bourhis et al. 2013). However, some bacterial species are unable to activate MAIT cells such as *Enterococcus faecalis* (Le Bourhis et al. 2010), *Listeria monocytogenes* (Kjer-Nielsen et al. 2012) and group A *Streptococcus* (Le Bourhis et al. 2010) and this is thought to be due to an inability of these bacteria to produce riboflavin metabolites,

potentially due to defects in the metabolic pathway (Le Bourhis et al. 2010). Activation of MAITs leads to the release of granzyme and cytokines (IFNγ and TNF) (Gold et al. 2008; Gold & Lewinsohn 2012) which result in target cell death. MAITs have been shown to be CD4- CD8- orCD8αβ^{int} (Dusseaux et al. 2011) and express a pattern of chemokine receptors which are associated with homing to tissues in the liver and intestines (Martin et al. 2009; Dusseaux et al. 2011). MAIT cells in the thymus and in the cord blood have a naïve phenotype, in contrast peripheral blood MAITs have a memory T cell phenotype (Dusseaux et al. 2011; Gold et al. 2013). Selection and expansion of MAITs is dependent upon the presence of B cells and commensal microbiota (Martin et al. 2009). The MAIT cell population is absent in patients and mice deficient in B cells and germfree mice (Treiner et al. 2003) suggesting that the presence of commensal microbiota is essential for their survival. It has been shown that MAITs with distinct TCR usage are activated by different pathogens (Gold et al. 2014). There is strong evidence that both the TCR receptors and MR1 molecule have been conserved throughout evolution due to their abilities in protecting the host from a wide range of microbes (Gold & Lewinsohn 2012).

5.1.2 MR1 molecule

The MR1 molecule has a high sequence homology with and similar structure to the classical MHCI molecules. The homology between the α1 and α2 domains of MR1 and the other MHCI molecules is approximately 90% in humans and mice (Yamaguchi et al. 1997; Tsukamoto et al. 2013). In a similar manner to the other classical MHCI molecules, the MR1 molecule is formed from a heavy α chain and a light β2m chain. Within the α chain, the α1 and α2 domains forming the antigen binding groove and the α3 domain covalently binding to the β2m molecule (Kjer-Nielsen et al. 2012) (Figure 5.1). The MR1 molecule presents small metabolites from vitamin B biosynthesis such as metabolites of folic acid (vitamin B9). However, only three examples of vitamin B2 derivatives have been shown to activate MAIT cells (rRL-6-CH_2OH, RL-6-Me-7-OH and RL-6,7-diMe) (Kjer-Nielsen et al. 2012).

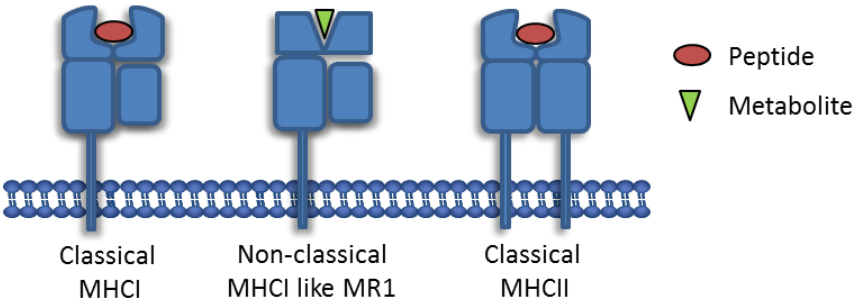


Figure 5.1. Schematic representation of MHC molecules (MHCI, MHCI like MR1 and MHCII). Classical MHCI and MHCII molecules present peptides, whilst the non-classical MHCI like MR1 molecules present metabolites.

5.1.3 *M. smegmatis* as a model for studying T-cell responses to bacteria presented through MR1

M. smegmatis was used as a model in this study as it is a non-pathogenic, fast growing class I bacterial strain, which can be handled in the biosafety level 1 facility available to me (Brown-Elliott & Wallace 2002). *M. smegmatis* has been shown to be phagocytosed and processed by APCs, which present the *M. smegmatis* antigens on MR1 in addition to the classical MHC molecules. Although *M. smegmatis* is non-pathogenic it has a high homology with *Mycobacterium tuberculosis,* including 12/19 virulence factors, (Reyrat & Kahn 2001) and therefore serves as a good laboratory model for this human pathogen.

5.2 Rationale

The invariant nature of MR1 across the human population and its established role in the presentation of pathogen-derived antigen are of outstanding interest for the potential development of universal therapeutic and diagnostic tools in infectious diseases. MR1 expression also appears to be ubiquitous among different cells and tissues, which may indicate that MR1-driven antigen responses are relevant to the pathogenesis of a broad number of immune mediated diseases. However, the invariance and ubiquity of MR1 also complicate basic investigations of its ligand binding and antigen presentation properties as well as in the understanding of MR1-restricted T-cell biology. The presence of MR1 on most APC lines and primary cells that express other classical and non-classical HLA molecules can make the unambiguous identification of microbe-specific MAIT cells and their distinction from conventional T-cells that express the TRAV1-2 TCR chain problematic. Currently, the role that MAITs play in autoimmune, non-infectious inflammatory and highly infectious disease is largely unknown despite localisation of MAIT cells at disease site in numerous diseases such as inflammatory bowel disease (Serriari et al. 2014), autoimmune arthritis (Billerbeck et al. 2010), multiple sclerosis (Annibali et al. 2011), autoimmune hepatitis, chronic hepatitis C, alcoholic liver disease, non-alcoholic steatohepatitis, primary biliary cirrhosis (Oo et al. 2016) and *Mycobacterium tuberculosis* (Karp et al. 2015). Development of MR1 deficient cell lines will allow for studying the role of MAIT cells in autoimmune, non-infectious inflammatory diseases and highly infectious diseases.

5.3 Aims

When the CRISPR/Cas9 technology was first published, researchers required three individual plasmids for efficient targeted knockout that expressed tracrRNA, crRNA and Cas9 protein. One aim of this project was to combine all of these plasmids into a single gene knockout vector to allow easy testing of crRNA efficiency. I then aimed to show that CRISPR/Cas9 technology could be useful for the study of antigen presentation to unconventional T-cell subsets. There has been much recent interest in the presentation of bacterial metabolites to MR1-restricted T-cells. I thus aimed to use the 'all in one' CRISPR/Cas9 system as a tool to produce MR1 deficient cell lines as a proof-of-concept that this approach could be used to study microbial antigen processing and presentation to human MAIT cells.

5.4 Results

5.4.1 Generation of an 'all in one' lentiviral CRISPR/Cas9 system

The existing CRISPR/Cas9 system described by Mali et al. 2013 consisted of two plasmids, one containing the Cas9 protein and the other containing the gRNA sequences. To improve the efficiency of knockout using the CRISPR/Cas9 system a single plasmid system was produced. The DNA sequence of the U6 RNApolIII promoter and a gRNA sequence which encompassed the MR1 specific crRNA A and the tracrRNA sequence were synthesised (Eurofins). The sequence was cloned into the pCDNA.3-TOPO_Cas9 upstream of the CMV promoter. To produce a lentiviral vector the pCDNA.3-TOPO_Cas9+gRNA plasmid was PCR amplified using pCDNA.3_Fwd and pCDNA.3_Rev primers to introduce an AgeI restriction site upstream of the U6 RNApolIII promoter and an NsiI restriction site downstream of the Cas9 sequence. Corresponding restriction sites were introduced into the 2nd generation pRRL.SIN.cPPT.pgk-gfp.WPRE lentivector backbone by PCR amplification with pRRL.0_Fwd and pRRL.0_Rev primers. The CRISPR elements were cloned into the pRRL.SIN.cPPT.pgk-gfp.WPRE lentivector backbone (Figure 5.2). All primer sequences are found in Appendix Table 2.

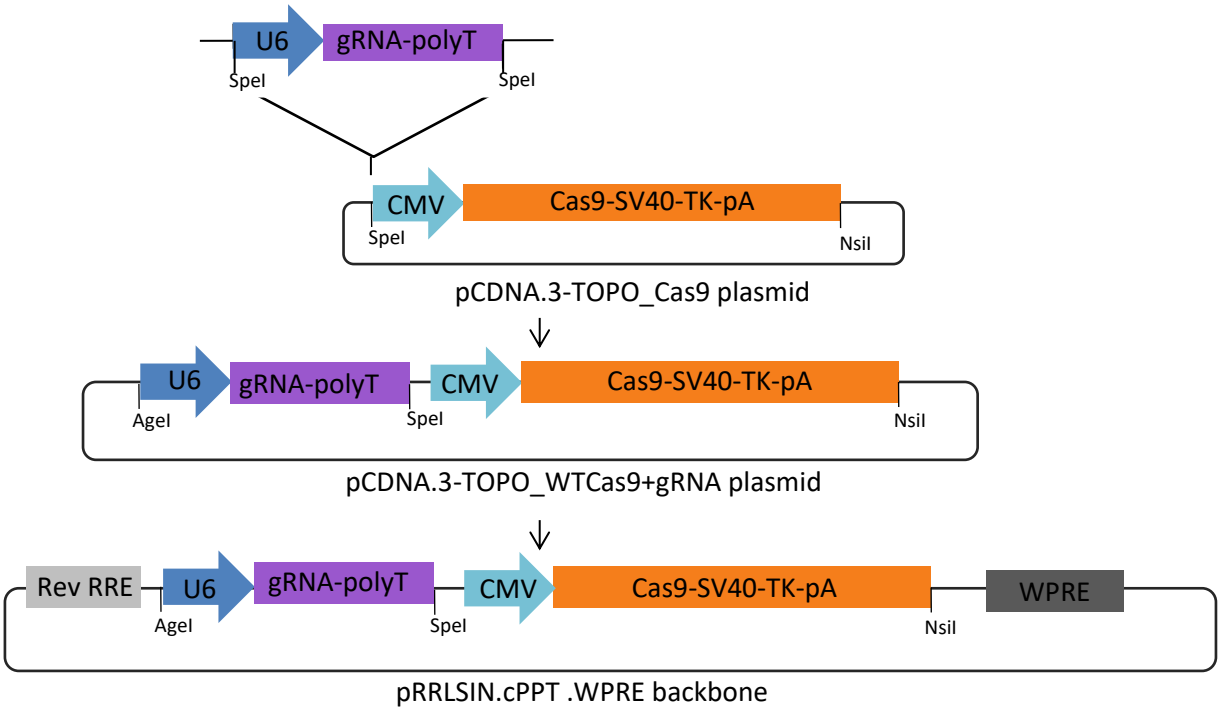
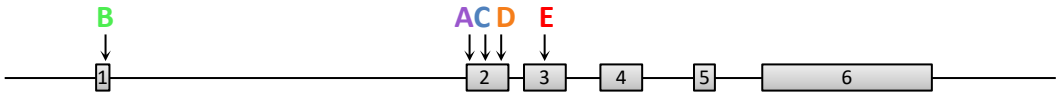


Figure 5.2. Development of the CRISPR/Cas9 "all-in-one" lentiviral system. Graphic representation of the "all-in-one" lentiviral system generated for this study. The synthesised gRNA sequence was cloned into pCDNA.3-TOPO_WT-Cas9 plasmid. The U6 promoter, gRNA sequence, CMV promoter and Cas9 elements were cloned into the 2nd generation pRRL.SINpRRLSIN.cPPT.WPRE lentivector backbone.

5.4.2 Production of 5 novel crRNA sequences targeting the MR1 gene

5 crRNA target sequences were designed by identifying $GN_{19}GG$ DNA motifs present on either strand of the MR1 genomic DNA sequence. The crRNA sequences were run through the basic local alignment tool - nucleotide (blastn) aligorithm to identify likely off target sites (Altschul et al. 1990) and crRNA sequences with potential off target sites were excluded. 5 non-overlapping target sequences were selected located in the first 3 exons of the MR1 coding sequence. crRNA B targeting exon 1, crRNA A, C and D targeting exon 2 and crRNA E targeting exon 3. The novel crRNA sequences were cloned into the 'all in one' CRISPR/Cas9 system for validation. The location and sequence of each crRNA is shown in Figure 5.3.

A



B

crRNA	Target sequence	PAM	Exon	Strand
A	**GGATGGGATCCGAAACGCCC**	**AGG**	**2**	**+**
B	**GGTGAAGCACAGCGATTCC**	**CGG**	**1**	**-**
C	**GTCCCTGAATTTATTTCGGT**	**TGG**	**2**	**-**
D	**GAACCTCGCGCCTGATCACT**	**GGG**	**2**	**-**
E	**GCAGTATGCATATGACGGGC**	**AGG**	**3**	**-**

Figure 5.3. Design of novel crRNA sequences targeting the MR1 gene. A) Positioning of 5 active crRNA target sequences with respect to the intron/exon structure of the MR1 gene. B) Table listing the CRISPR/Cas9 crRNA target sequences within the MR1 gene generated during this study.

5.4.3 Comparison of the novel crRNA sequences targeting the MR1 gene

The level of MR1 expression on the cell surface of normal non-bacterially infected cells is low, so in order to test the 5 crRNAs a reporter cell line was generated by overexpressing the native MR1 sequence in HeLa cells (referred to as HeLa-MR1 cells). Lentivirus for each of the 5 crRNAs was produced as described in section 2.3.7.2 and used to transduce the HeLa-MR1 cell line. 3/5 of the crRNAs, A, B and D, resulted in MR1 knock down/knock out as evident by flow cytometry. To assess whether mutations were present at a DNA level, DNA was extracted from WT HeLa-MR1 and HeLa-MR1 cells transduced with crRNAs A, B and D. The surveyor assay was performed as described in section 2.4.13. crRNA A was found to

induce the greatest levels of knockout on a protein level (flow cytometry) and DNA level (Surveyor assay) (Figure 5.4). Therefore, crRNA A was chosen for downstream applications.

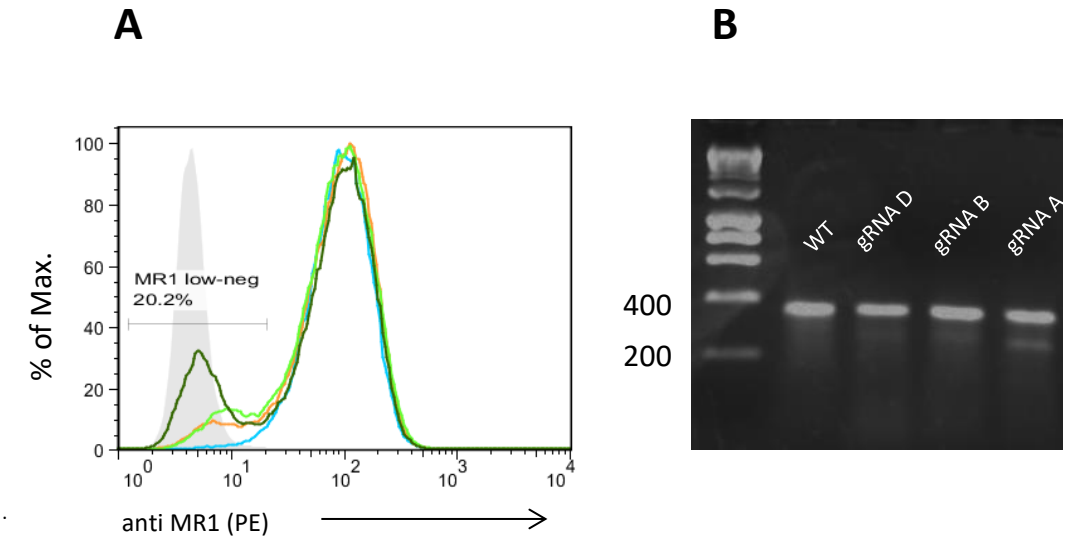


Figure 5.4. Validation of crRNAs targeting the MR1 in the HeLa-MR1 cell reporter system. HeLa-MR1 cells were transduced with crRNAs A, B and D by lentivirus. A) The level of MR1 expression on the cell surface of cells was monitored by flow cytometry, the isotype control (solid light grey), crRNA A (dark green), crRNA B (light green), crRNA D (orange) and WT unmodified HeLa-MR1 cells (light blue). A population of MR1 negative cells is observable in cells transduced with crRNAs A, B and D. B) The presence of mutations at a DNA level was monitored by surveyor assay. DNA was PCR amplified using the MR1_SURV1_Fwd and MR1_SURV1_Rev primers, which bound immediately upstream of the MR1 cDNA start and downstream of the target site respectively. 400ng of PCR product from unmodified Hela-MR1 or 200 ng of PCR product from Hela-MR1 cells transduced with each crRNA target (crRNA A, B and D) and 200 ng of PCR product from unmodified Hela-MR1 cells were denatured, annealed and digested with the mismatch specific surveyor enzyme as described in section 2.4.13. Additional bands indicating mismatches at a DNA level were observed weakly in crRNA B and D wells and more strongly in the crRNA A well.

5.4.4 Efficiency of MR1 knockout using lipofectamine transfection and lentiviral transduction

The crRNA A was expressed in A549 cells either transiently by lipofectamine transfection or stably by lentiviral transduction. The cells were monitored by flow cytometry which showed that both transient and stable expression of the crRNA A led to an observable knock down/ knock out population and a reduction in the MFI of MR1 in A549 cells transfected with lipofectamine and transduced with lentivirus in comparison to the WT A549 cells (Figure 5.5A). A surveyor assay confirmed the presence of mutations at a DNA level in A549 cells transfected with lipofectamine and transduced with lentiviral in comparison to the WT A549 cells (Figure 5.5B).

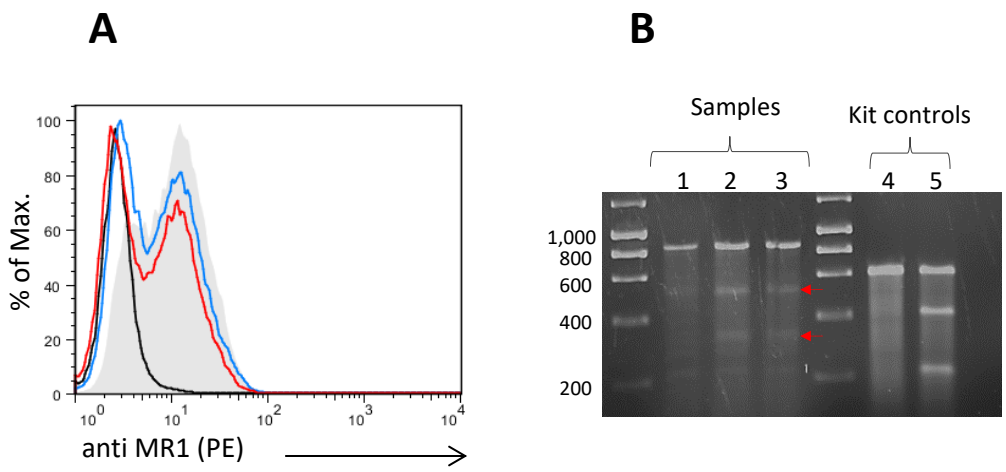


Figure 5.5. Comparison of crRNA A lipofectamine transfection and lentiviral transduction. MR1 deficient A549 cells were generated by lipofectamine transfection and lentiviral transduction. The cells were cultured for 8 days before being analysed by flow cytometry and surveyor assay. A) The cells were treated with acetyl-6-formylpterine (6-f-p) (50 mg/ mL) for 16 h and then were stained with the MR1 antibody or with an isotype control. Cells were stained with the isotype control antibody (black; MFI = 2.78), mock lipofected cells (solid grey; MFI = 9.53), cells transduced with lentivirus (red; MFI = 5.84) and cells transfected with CRISPR/Cas9 plasmid DNA by lipofectamine (blue; MFI = 7.07). The lipofectamine and lentiviral transduction approaches both lead to a substantial MR1 negative population. B) The surveyor assay was performed with 500 ng of DNA extracted from WT unmodified cells and cells transduced or transfected with the crRNA lentivirus or lipofectamine. The DNA was amplified using MR1_SURV2_Fwd and MR1_SURV2_Rev primers. Homoduplex of WT DNA amplicons (lane 1) or heteroduplexes of annealed WT and amplicons from transduced cells (lane 2) or lipofected cells (lane 3) were digested with the Cel1 DNA-mismatch specific enzyme. The red arrows indicate the two bands expected after digestion with the mismatch specific Cel1 enzyme (521 and 331 bp). The surveyor assay kit controls - undigested (lane 4) and digested (lane 5) assay controls are shown.

5.4.5 Generation of clonal A549 cell populations bearing CRISPR/Cas9 induced mutations in the MR1 locus

The lentiviral transduction and lipofectamine transfection approaches both effectively knocked out the expression of MR1 from the cell surface of a substantial proportion of the A549 cells. Despite the lentiviral transduction approach leading to a higher frequency of knockout than the lipofectamine transfection approach, it was thought long-term expression of the CRISPR/Cas9 elements by integrated lentiviral transduction may increase the risk of off target mutations. Therefore, the bulk lipofectamine transfected cells were used to produce MR1 deficient A549 clones. $MR1^{low}$ cells were sorted by flow cytometry and cloned by limiting dilution to generate clonal populations. Eight clones were screened by flow cytometry, of these clones 1, 9 and 11 expressed the lowest levels of MR1 consistently (Figure 5.6A). The surveyor assay was performed on these clones to establish whether there were mismatches at a DNA level. The DNA flanking the crRNA sequence was amplified by PCR in each of the clones and the WT A549 cells using the MR1_SURV2_Fwd and MR1_SURV2_Rev primers which bound upstream and downstream of the crRNA respectively. Interestingly two bands were observed for clone 9 (one at the expected size of

852 bp and other 100-150 bp shorter), suggesting that there was a deletion on one allele. A surveyor assay was performed as described in section 2.4.13. Mismatches at the crRNA locus were expected to produce two products (approximately 520 and 330 bp) after digestion with the Cel1 mismatch specific enzyme. Additional products were observed for clone 9 and 11 indicating that there were mutations at a DNA level (Figure 5.6B). Clone 1 did not produce additional bands and was not taken forward.

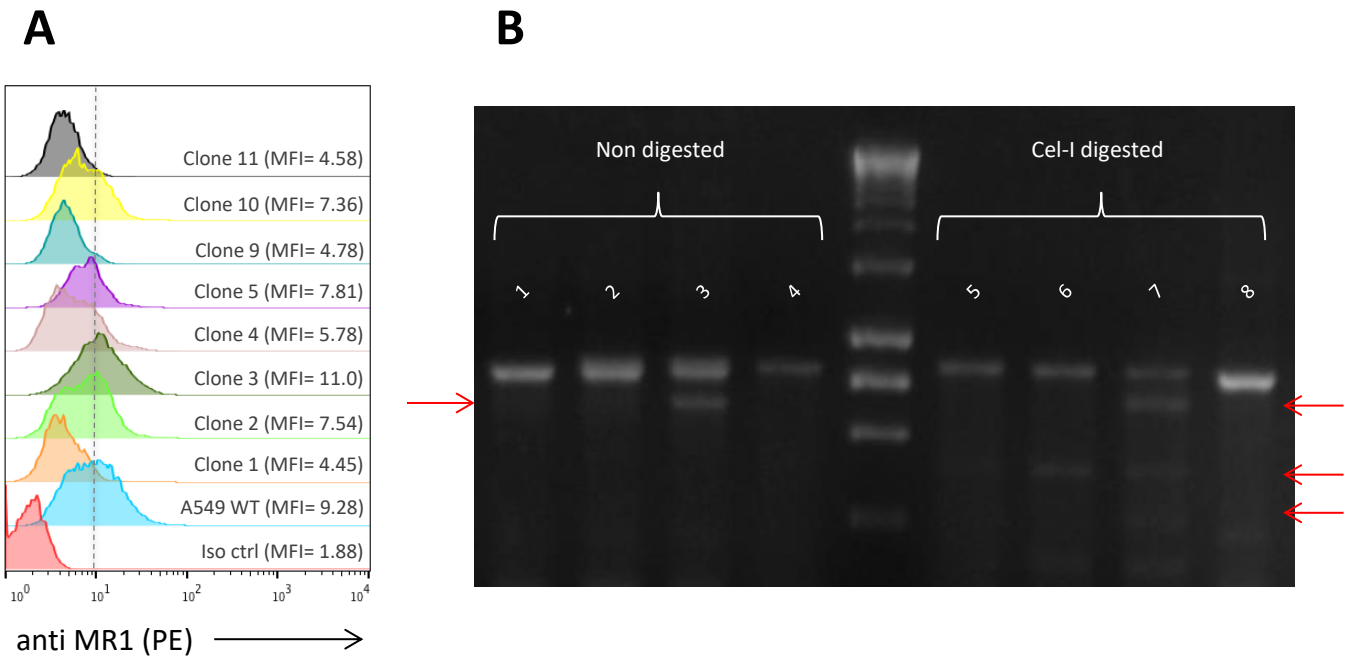


Figure 5.6. Generation of the MR1 negative A549 clones. Clonal derivatives were generated by single cell sorting of the bulk population of A549 cell population transfected with the crRNA CRISPR/Cas9 elements by lipofectamine. A) Eight clones and the WT A549 cells were treated with 6-f-p (50 mg/mL) for 16 h; cells were then stained with the MR1 antibody or with an isotype control antibody. Isotype control (red; MF1= 1.88), WT (blue; MFI=9.28) and clones shown as a variety of colours with MFIs ranging from 4.45 to 11. Clone 1, 9 and 11 were found to have to lowest MFI consistently and were chosen for downstream applications. B) DNA was extracted from clone 1, 9, 11 and WT A549 cells. The DNA was amplified by PCR using MR1_SURV2_Fwd and MR1_SURV2_Rev primers and the amplicons were run on an agarose gel in the absence of hybridisation and digestion WT A549 cells (lane 1) and A549 clone 11 (lane 2), clone 9 (lane 3) and clone 1 (lane 4). The PCR amplicons were hybridised with WT DNA and digested with the Cel1 enzyme and run on an agarose gel homoduplex WT DNA (lane 5) or heteroduplexes of annealed WT and modified amplicons from clone 11 (lane 6), clone 9 (lane 7) or clone 1 (lane 8). Prior to digestion a second smaller band was observed in clone 9 suggesting a 100-150bp deletion on one allele. After digestion additional bands were observed in clone 9 and clone 11.

5.4.6 Characterization of disruptive mutations in the MR1 gene of MR1-/- A549 clonal derivatives

For further characterisation, the region of DNA flanking the crRNA binding site was amplified using the MR1_SURV_Fwd_BsaI and MR1_SURV_Rev_BsaI primers which bound upstream and downstream of the crRNA respectively and introduced BsaI restriction sites

onto the PCR product and the PCR products were cloned into a cloning vector (Addgene#32189). Ten colonies were selected for each clone and minipreps of the DNA were sequenced by Sanger sequencing using MR1_Seq_Fwd and MR1_Seq_Rev primers. Clone 9 was found to have a 126 bp deletion on allele 1. On allele 2 of clone 9 there was a single nucleotide deletion and two nucleotide substitutions. Clone 11 had an identical single nucleotide deletion on both alleles. The single nucleotide deletion in clone 9 and 11 were identical (Figure 5.7). All mutations induced a frameshift in the reading frame on both alleles of clone 9 and clone 11. Clone 9 allele 1 is disrupted from amino acid 23, the second allele of clone 9 and both alleles of clone 11 are disrupted from amino acid 34 onwards (Figure 5.7C). Taken in combination with the lack on MR1 staining on the cell surface, my results indicate that the CRISPR/Cas9 system effectively abrogated MR1 expression from the cell surface.

A

Sample	DNA sequence
WT	GGAGCACTCGTGTGGGGCTAAATGAATGCAGTTGAAGGATCTGGATCATCTGGGACCCTACATGTCTTCC TTCTTTGCCTCCTTTCCAGGGACGCACTCTCTGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATG
C9 a1	GGA~~GCACTCGTGTGGGGCTAAATGAATGCAGTTGAAGGATCTGGATCATCTGGGACCCTACATGTCTTCC~~ ~~TTCTTTGCCTCCTTTCCAGGGACGCACTCTCTGAGATATTTTCGCCTGGGCGTTTCGGA~~TCCCATCCATG

B

Sample	DNA sequence
WT	CTGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAAT
C9 a2	CTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAAT
C11 a1	CTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAAT
C11 a2	CTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAAT

C

Sample	Amino acid sequence
WT	SDSRTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLRG
C9 a1 **C9 a2**	SDS? SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C11 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C11 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!

Figure 5.7. Analysis of the mutations induced by CRISPR/Cas9 at a DNA level in A549 clone 9 and 11. The crRNA binding sequence is highlighted in yellow and the PAM sequence is highlighted in blue. Nucleotides in exon 2 are blue, red nucleotides with a strikethrough represent deleted nucleotides. A) The DNA sequence of A549 clone 9 allele 1. The reference WT sequence is shown above. A 126bp deletion is identifiable which encompasses a portion of the intron and the first 33bp of exon 2. B) The DNA sequence of A549 clone 9 allele 2 and both alleles of clone 11, the WT sequence is shown above. All alleles share a single nucleotide deletion in the crRNA sequence 4 bp upstream of the PAM sequence. C) The predicted primary structures of the mutant MR1 proteins were determined by translating the DNA sequences. MR1 amino acids 20 - 90 are shown. The WT MR1 protein sequence is shown above. Out of frame reads are highlighted in grey. (C9 a1 = clone 9 allele 1, C9 a2 = clone 9 allele 2, C11 a1 = clone 11 allele 1 and C11 a2 = clone 11 allele 2).

5.4.7 Identification of CRISPR/Cas9 mutagenesis off-target effects using genomics.

To assess whether the CRISPR/Cas9 system targeting the MR1 gene was specific for MR1 whole genome sequencing was performed using DNA extracted from clone 9 and clone 11. The data was analysed with the help of Dr. Thomas Connor using the methodologies described in section 2.4.15. The most likely off target sites were identified using the CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre), using these algorithms 38 potential off target sites were identified for the crRNA A sequence. Table 2.3 shows the 37 potential off target sites. A further, unanticipated, on-target site was predicted within the RP11-46A10.6 gene, which encodes an unexpressed partial MR1 pseudogene. This site shares 23/23bp homology within the crRNA and the PAM sequence. 5/38 of these predicted off target sites were within genes and varied from the crRNA site by

3 or more nucleotides. 32/38 of these predicted off target sites were not located within gene encoding DNA. To ensure that no off target sites were excluded from the analysis, in excess of 7000 additional sites that varied from the crRNA sequence by 5bp or less were identified. All reads were aligned to the GRCh38 reference sequence and assembled using Artemis software. All of the potential off target sites were screened by eye. Only one case of an unintended mutation was identified in the RP11-46A10.6 pseudogene (as discussed in section 5.3.8.). The whole genome sequencing data analysis confirmed the mutations identified in the MR1 gene of clone 9 and clone 11 by Sanger sequencing (Figure 5.8).

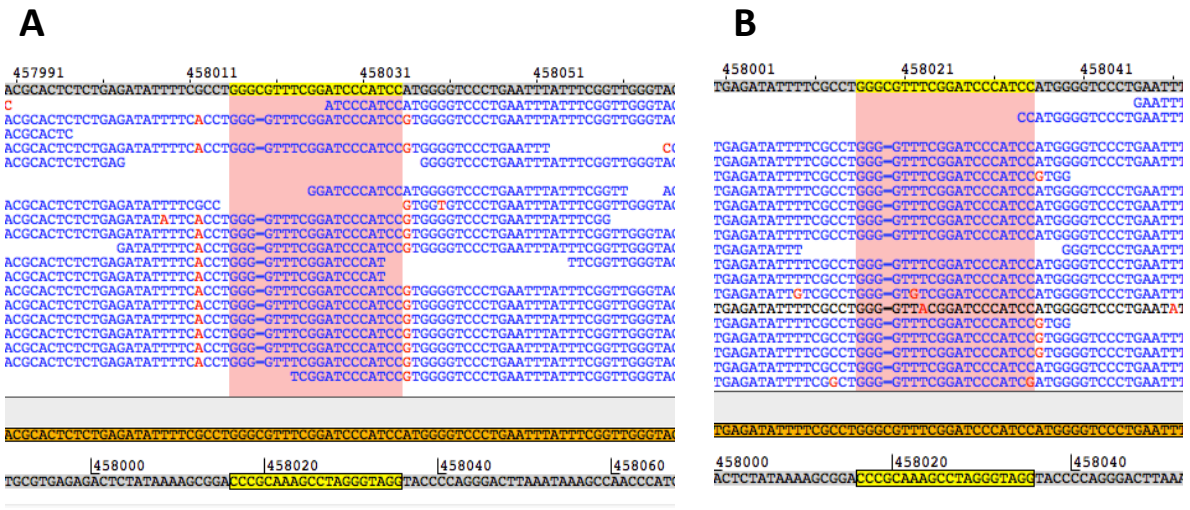


Figure 5.8. Confirmation of the mutations detected by PCR amplification of the MR1 gene by whole genome sequencing in clone 9 and clone 11. The 'on target' effects observed in the MR1 gene of A) clone 9 and B) clone 11.

5.4.8 MR1 CRISPR/Cas9 mutagenesis resulted in an unintended DNA mutation in the RP11-46A10.6 pseudogene.

The RP11-46A10.6 pseudogene exon 1 had a 94% sequence homology with MR1 exon 2 (Figure 5.9) and 100% homology at the crRNA binding site. Whole genome sequencing identified a single base pair deletion in the A549 clones C9 and C11 which was identical to that observed in the MR1 gene in C9 allele 2 and clone 11 allele 1 and 2 (Figure 5.9). The RP11-46A10.6 gene is located in the intron region between exon 1 and 2 of the STX6 protein-coding sequence. The exons of STX6 are a considerable distance from this mutation in the RP11-46A10.6 target sequence and therefore it is unlikely that this mutation will have any adverse effects on the STX6 gene.

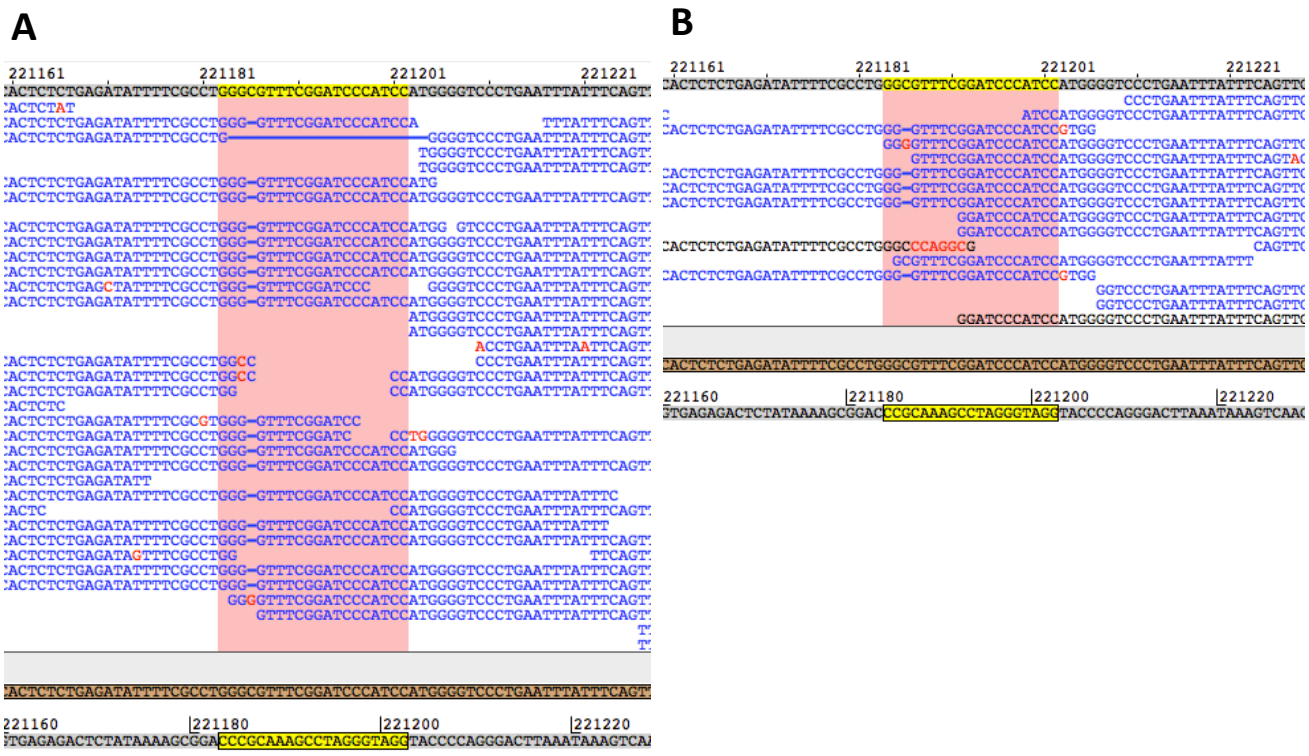


Figure 5.9. A screenshot of the mutations identified in the RP11-46A10 gene by whole genome sequencing in clone 9 and clone 11 using the Artemis software. Both alleles of clone 9 and clone 11 show a single nucleotide deletion at the same position as the deletion observed in the MR1 gene of clone 9 and clone 11 (A and B respectively).

In order to confirm the mutations induced in the RP11-46A10.6 which were identified by whole genome sequencing were present the DNA from the A549 clones 9 and 11 and WT cells was amplified using the RP11_FWD and RP11_REV primers which bind upstream and downstream of the crRNA sequence respectively. The amplified DNA was sequenced by Sanger sequencing which confirmed a single base pair deletion at the same site as observed in C9 allele 2 and C11 allele 1 and 2 MR1 crRNA binding site (Figure 5.10).

A

Sample	DNA sequence
WT	TGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C9 a1	TGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C9 a2	TGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C11 a1	TGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C11 a2	TGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT

B

```
 11 GATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTT     60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 21 GATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTT     70

 61 ATTTCGGTTGGGTACGTGGACTCGCACCCTATCACCACATATGACAGTGT    110
    |||||.|||||||||.||||||||||||||||||||||||||||||||||
 71 ATTTCAGTTGGGTATGTGGACTCGCACCCTATCACCACATATGACAGTGT    120

111 CACTCGGCAGAAGGAGCCACGGGCCCCATGGATGGCAGAGAACCTCGCGC    160
    |||||.||||||||||||||||||||||||||||||||||||||||.||
121 CACTCAGCAGAAGGAGCCACGGGCCCCATGGATGGCAGAGAACCTCGTGC    170

161 CTGATCACTGGGAGAGGTACACTCAGCTGCTGAGGGGCTGGCAGCAGATG    210
    |||||||||||||||||||||||||||||||.||||||||||||||||
171 CTGATCACTGGGAGAGGTACACTCAGCTGCTGAAGGGCTGGCAGCAGATG    220

211 TTCAAGGTGGAACTGAAGCGCCTACAGAGGCACTACAATCACTCAG     256
    ||||.||||||||||||||..|||||.|||||||||||||||
221 TTCAGGGTGGAACTGAAGCGCCAGCAGAGACACTACAATCACTCAG     266
```

Figure 5.10. Analysis of the mutations induced by CRISPR/Cas9 at a DNA level in A549 clone 9 and 11 within the RP11-46A10.6 gene. A) The DNA sequence of A549 clone 9 allele 1, 2 and clone 11 allele 1 and 2, the WT sequence is shown above. Each of the alleles of C9 and C11 share a single nucleotide deletion in the crRNA sequence 4 bp upstream of the PAM sequence. (C9 a1 = clone 9 allele 1, C9 a2 = clone 9 allele 2, C11 a1 = clone 11 allele 1 and C11 a2 = clone 11 allele 2). B) Exon 2 of MR1 and exon 1 of the RP11-46A10.6 gene were aligned using the EMBOSS MATCHER software. Exon 2 of MR1 is the top sequence and exon 1 of RP11-46A10.6 is shown as the bottom sequences. Mismatches between MR1 and the RP11-46A10.6 gene are shown as dots. The crRNA sequence is highlighted in yellow and the PAM sequence is highlighted in blue.

5.4.9 A549 mutant clones infected with bacteria selectively fail to activate MAIT cell clones infected with *M. smegmatis.*

To confirm that the MR1 deficiency on the cell surface led to a functional deficiency, WT and MR1 deficient A549 clones were infected with *M. smegmatis* bacteria and incubated with one of two MAIT clones (D426B1 and D481A9). Both clones expressed the TRAV1-2_TRAJ33 MAIT TCR α-chain rearrangement, however, the MAIT clones expressed different CDR3

sequences and different β chain pairings (D426B1 expresses TRBV6-4 and D481A9 expresses TRBV20-1). An ELISA was performed on the supernatant after 16 h incubation with infected target cells. The infected WT A549 cells stimulated the MAIT cell clones to produce MIP1β and IFN-γ, in contrast, the MR1 deficient clones were unable to activate the MAIT clones and unable to drive cytokine release (Figure 5.11). These findings suggest that the MR1 deficient A549 clones were unable to present bacterial antigens through MR1 on the cell surface.

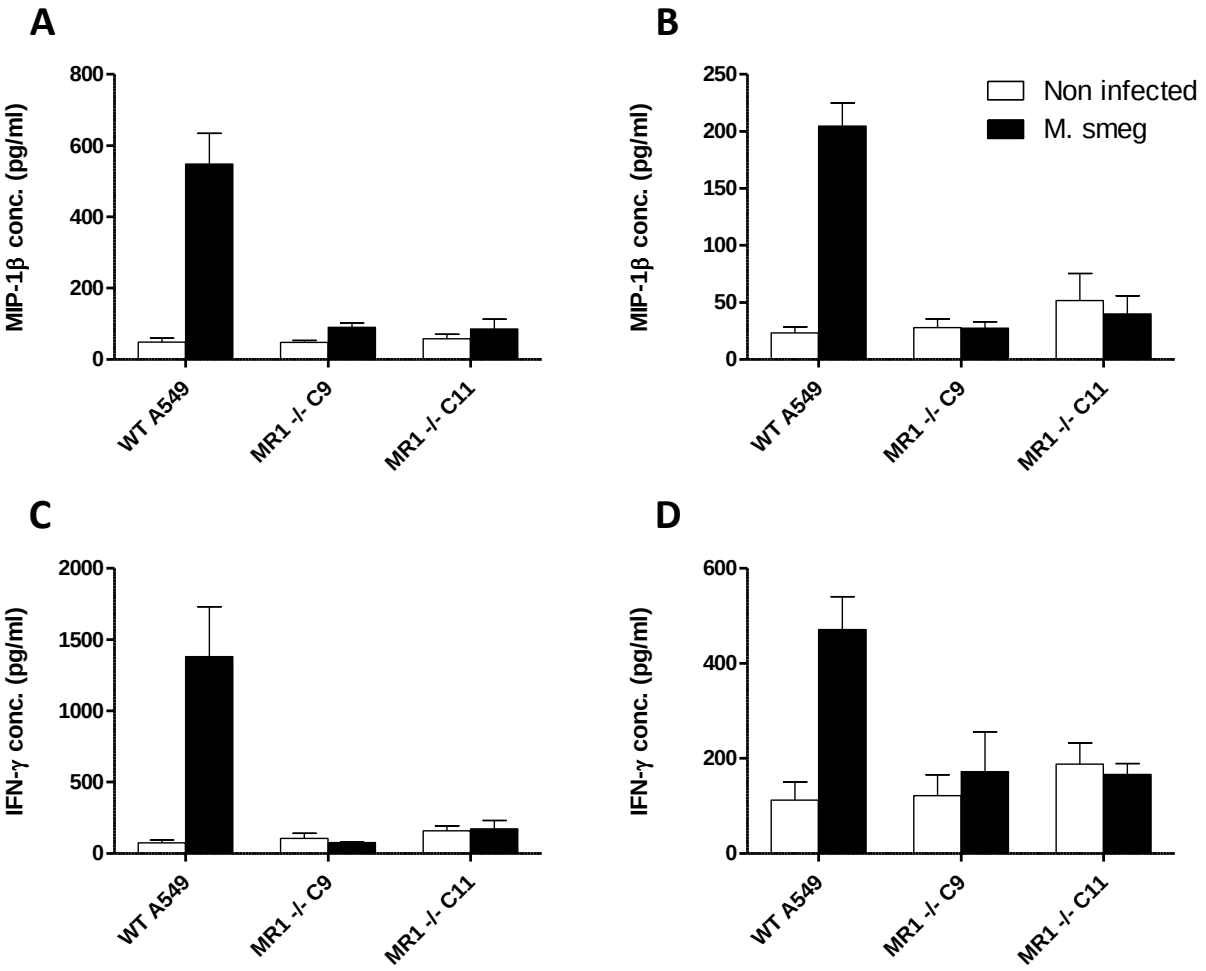


Figure 5.11. MR1 deficient A549 clones C9 and C11 infected with *M. smegmatis* fail to activate The D481A9 and 426B1 MAIT clones. The concentration of MIP-1β produced by the D481A9 MAIT cell clone (A) and D426B1 MAIT cell clone (B), were quantified by ELISA. The concentration of IFN-γ was also measured for the D481A9 MAIT cell clone (C) and D426B1 MAIT cell clone (D). Standard error of the mean values for 3 replicates are shown; data are representative of 3 experiments.

5.4.10 MR1 mutations do not affect the expression of classical HLA molecules in A549s.

To assess whether the expression of classical HLA molecules was altered in MR1 deficient A549 cells the A549 WT and MR1 deficient cells were stained with a pan class I antibody and the expression was analysed by flow cytometry, an unpaired T test revealed there was no significant difference between the WT or MR1 deficient cells. WT and MR1 deficient cells were pulsed with the HLA- B*1801 restricted SELEIKRY peptide derived from EBV $BZLF1_{173-180}$ at concentrations ranging from 10^{-6} to 10^{-10} M. The SELEIKRY-specific T-cell clone was incubated with the peptide pulsed A549 for 16 h after which an IFN-γ and MIP-1β ELISA were performed with the supernatants (Figure 5.12). The MR1 deficient A549 clones C9 and C11 were able to present the SELEIKRY peptide derived from the EBV protein by HLA-B*1801 and activate the SELEIKRY-specific T-cell clone in a similar manner to the WT A549s as evidenced by equivalent cytokine release in peptide titration experiments.

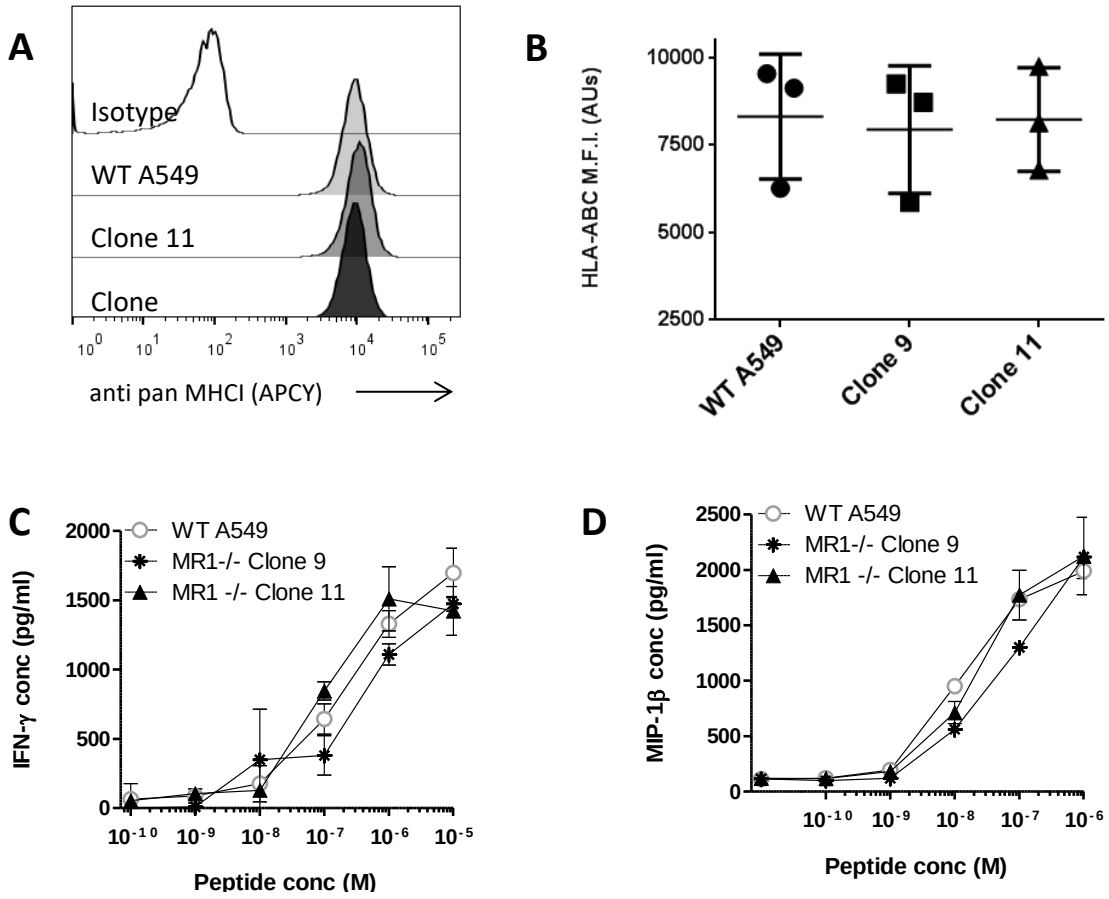


Figure 5.12. MR1 disruption does not affect HLA class I expression. A) A549 WT and the MR1 deficient clones were stained with an isotype control or the pan class 1 antibody. The MFI values determined by flow cytometry were 56.4 (isotype control), 8,733 (WT A549), 9,880 (clone 11) and 8,682 (clone 9). B) The MFI from 3 surface stainings with the pan class I antibody. The MFI data points are plotted in addition to mean values and the standard error bars. There is no significant difference in expression of classical MHC expressed by the WT or MR1 deficient clones C9 and C11, the P values were 0.81 and 0.94 respectively. C + D) The WT and MR1 deficient A549 cells were pulsed with the HLA-B*1801 restricted SEL peptide and incubated with the B9 T cell clone for 16 h after which the supernatants were used to perform IFN-γ and MIP-1β ELISA (C and D respectively).

5.4.11 Generation of MR1-/- THP-1 clones.

The 'all-in-one' system was used to knockout MR1 in the THP-1 cell line to further demonstrate the versatility of this CRISPR/Cas9 system for targeted gene knockout. The crRNA A was expressed in THP-1 cells either transiently by lipofectamine transfection or stably by lentiviral transduction alongside a GFP control to establish the efficiency of transfection/transduction. The lipofectamine transfection was not a viable method for introducing the plasmid encoding the crRNA sequence into THP-1 as the efficiency was extremely low. However, the lentiviral transduction approach led to sufficient knockout to generate MR1 deficient THP-1 cells.

MR1 deficient THP-1 clones were generated by two methods from the bulk transduced THP-1 cells. Either the bulk transduced THP-1 cells were infected with *M. smegmatis* and incubated with the D481A9 MAIT clone as described in section 2.3.16 which allowed for an enrichment of MR1 deficient THP-1 cells. Clones were then derived from the surviving cells by limiting dilution. Alternatively, the $MR1^{low}$ cells were sorted from the bulk transduced THP-1 cells by flow cytometry. Clonal derivatives were produced by limiting dilution. 20 clones were generated by limiting dilution of the THP-1 cells after selection with the D481A9 MAIT clone and 10 clones were generated by limiting dilution of the THP-1 after cell sorting.

The bulk population of cells derived by the enrichment process had a substantially reduced MFI compared to the WT cells (Figure 5.13A). The level of MR1 expressed by the 20 clones produced by MAIT enrichment and 10 clones produced by flow cytometry sorting was greatly reduced relative to the WT THP-1 cells (Figure 5.13B and D). The 5 clones expressing the lowest levels of MR1 on the cell surface for both the MAIT enrichment and flow cytometry sorting methods were subjected to a Surveyor assay. Mismatches were observed in the 10 clones produced indicating that mutations had been induced at a DNA level relative to the WT sequence (Figure 5.13C and E).

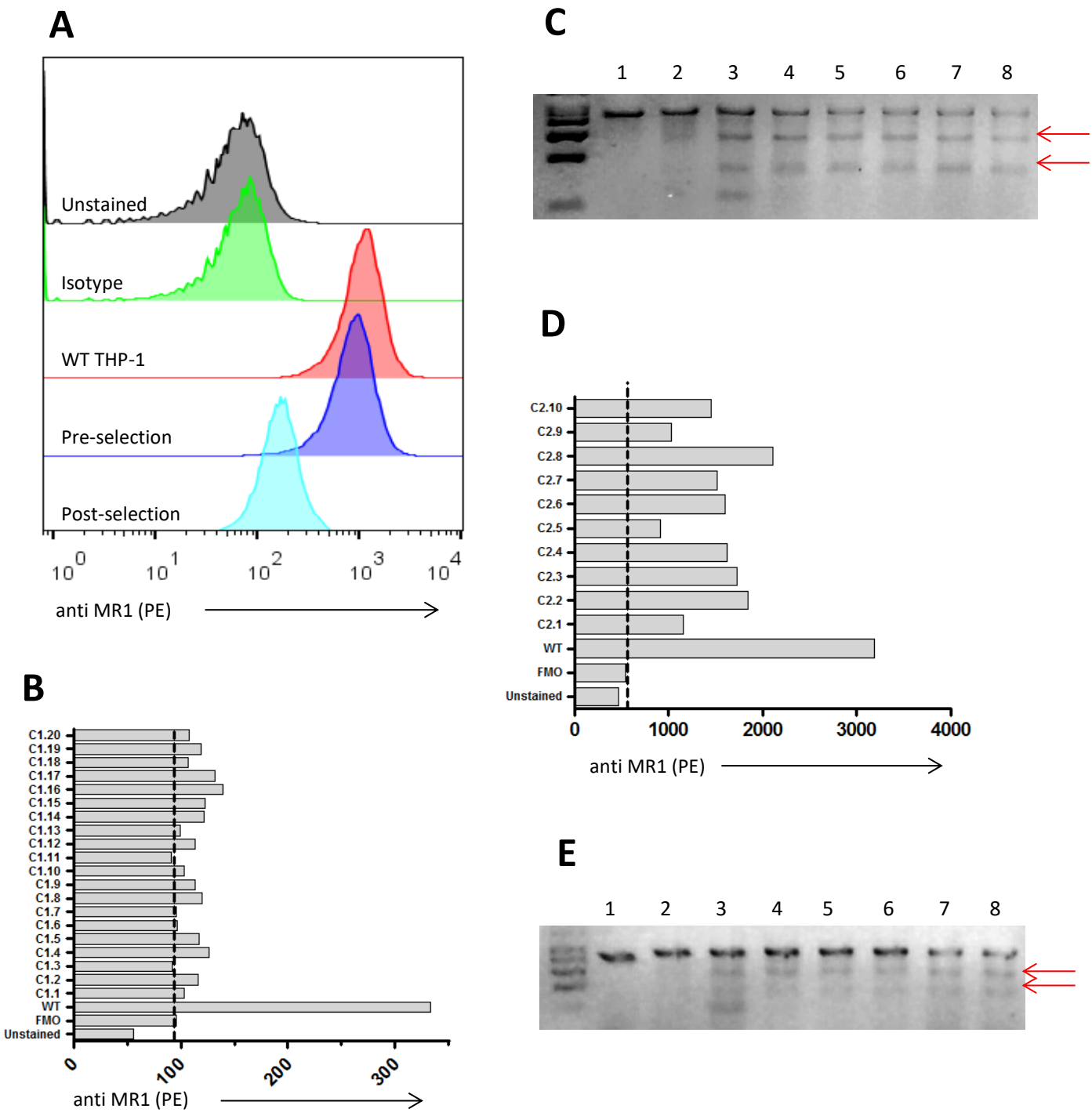


Figure 5.13. Generation of MR1 deficient THP-1 cells. THP-1 cells were transduced with CRISPR/Cas9 MR1 crRNA A by lentiviral transduction and the MR1 deficient THP-1 cells were enriched by MAIT cell selection with the 426B1 cell clone to derive MR1 deficient clones. A) Flow cytometry analysis of MR1 expression in THP-1 cells transduced with MR1 CRISPR/Cas9 lentiviral particles before and after enrichment of MR1 deficient cells by selection with the 426B1 MAIT clone MFI values were: Unstained 45.5, isotype control 53.9, WT THP-1 1033, pre-selection edited cells 820 and post-selection edited cells 161. B) The 20 clones produced after single cell cloning of the THP-1 cells after MAIT enrichment were stained with the MR1 antibody or an isotype control alongside the wildtype THP-1 cells. C) Surveyor assay performed on the WT and 5 clones (1.3, 1.6, 1.7, 1.11 and 1.13) which expressed the lowest levels of MR1 on the cell surface. Cel-I digested PCR amplicon homoduplex of WT THP-1 cells (lane 2), heteroduplexes of WT THP-1 cells DNA hybridized with DNA from bulk-transduced THP-1 pre- (lane 3) and post-selection (lane 4), as well as clones 1.3 (lane 5), 1.6 (lane 6), 1.7 (lane 7), 1.11 (lane 8) and 1.13 (lane 9). The Surveyor assay was performed with a total of 500 ng of PCR amplicon obtained from genomic DNA of unmodified THP-1 cells or from the different bulk and clonal populations.

5.4.12 THP-1 mutant clones infected with *M. smegmatis* bacteria selectively fail to activate MAIT cell clone D481.

In order to determine whether the MR1 deficient THP-1 clones were able to function as APCs, the 10 THP-1 clones (C1.3, C1.6, C1.7, C1.11, C1.13, C2.1, C2.5, C2.7, C2.9 and C2.10) were infected with *M. smegmatis* and then incubated with the D481 A9 MAIT clone for 16 h. The supernatants were collected and used to perform an IFN-γ ELISA. The 10 THP-1 clones failed to activate the D481 MAIT clone relative to the WT THP-1 cells which were able to activate the D481 A9 clone as determined by the production and secretion of IFN-γ (Figure 5.14). The results indicated that the mutations in the DNA of the MR1 gene lead to a loss of a functional MR1 protein.

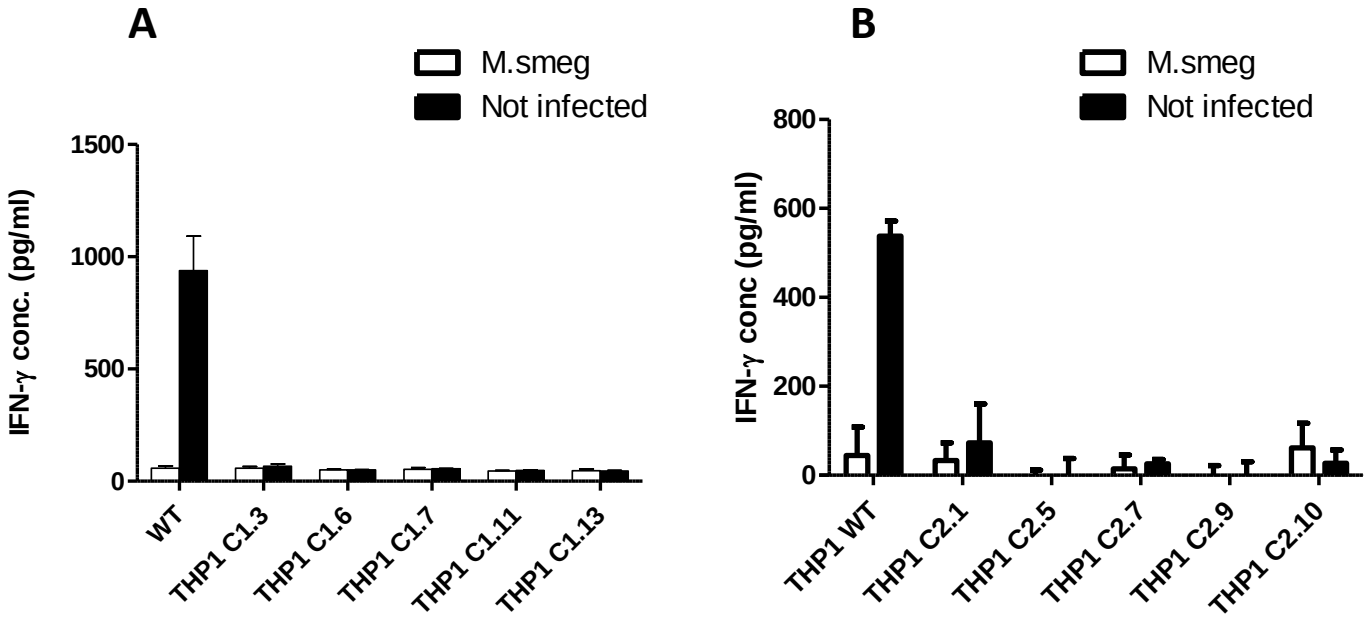


Figure 5.14. MR1 deficient THP-1 clones infected with *M. smegmatis* bacteria fail to activate the D481A9 MAIT clone. The concentration of IFN-γ produced by the D481A9 MAIT cell clone after a 16 h incubation with (A) the 5 MR1 deficient THP-1 clones derived by MAIT-cell enrichment and B) the 5 MR1 deficient THP-1 clones derived by cell sorting. ELISAs were carried out in triplicate wells. The mean and standard error of the mean values are shown.

5.4.13 Genotype characterization of disruptive mutations in the MR1 gene of THP-1 MR1-/- clones

5 clones were selected for downstream analysis (clone 2.1, 2.5, 2.7, 2.9 and 2.10) as they were generated without infecting the clones with *M. smegmatis.* To establish the mutations present in the 5 THP-1 MR1-/- clones the DNA was amplified, cloned and sequenced as described for the A549s in section 5.4.6. Clones 1, 5 and 10 harboured a biallelic single nucleotide deletion in the crRNA binding sequence 4 bp downstream of the PAM sequence.

Clone 7 harboured a biallelic 33bp deletion encompassing the entire crRNA. Clone 9 harboured a single nucleotide deletion on allele 1 crRNA binding sequence 4 bp downstream of the PAM sequence and a 45bp deletion on allele 2 which encompassed the entire crRNA. All of the single nucleotide deletions identified in the THP-1 clones were identical to that observed in the A549 clone 9 allele 2 and both alleles of clone 11 (Figure 5.15A). The amino acid sequence was interrupted from amino acid position 34 onwards in the case of clone 2.1 allele 1 and 2, clone 2.5 allele 1 and 2, clone 2.9 allele 2 and clone 2.10 allele 1 and 2. Clone 2.7 allele 1 and 2 had an 11 amino acid deletion from amino acid position 32 to 43 and clone 9 allele 1 had an 15 amino acid deletion from amino acid position 27 to 42 (Figure 5.15B).

A

Sample	DNA sequence
WT	GGACGCACTCTCTGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
1 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
1 a2	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
5 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
5 a2	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
7 a1	GGACGCACTCTCTGAGATATT~~TTCG~~~~CCT~~~~GGGCGTTTCGGATCCCATCC~~~~ATGGGG~~TCCCTGAATTTATTT
7 a2	GGACGCACTCTCTGAGATATT~~TTCG~~~~CCT~~~~GGGCGTTTCGGATCCCATCC~~~~ATGGGG~~TCCCTGAATTTATTT
9 a1	GGACGCACT~~CTCTGAGATATTTTCG~~~~CCT~~~~GGGCGTTTCGGATCCCATCC~~~~ATGGGG~~TCCCTGAATTTATTT
9 a2	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
10 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
10 a2	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT

B

Sample	Amino acid sequence
WT	SDSRTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLR
1 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
1 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
5 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
5 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
7 a1	SDSRTHSLRYF- - - - - - - - - - PEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLR
7 a2	SDSRTHSLRYF- - - - - - - - - - PEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLR
9 a1	SDSRTH - - - - - - - - - - - - - FPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLR
9 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
10 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
10 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!

Figure 5.15. Analysis of the mutations induced in the MR1 gene of MR1 deficient THP-1 clones A) The DNA sequence of the WT THP-1 cells is shown on top and each allele of the 5 MR1 deficient clones is shown. The PAM sequence is highlighted in blue and the crRNA sequence is highlighted in yellow. Deletions are shown as red letters with a strikethrough. Blue letters represent nucleotides present in exon 2 of the MR1 gene. B) The amino acid sequences were determined by translating the DNA sequence into the corresponding amino acid. Amino acids altered by a frameshift are highlighted in grey and deleted amino acids are shown by '–'.

5.4.14 MR1 mutations do not affect the expression of classical HLA I molecules in THP-1.

To establish whether the expression of classical MHCI was altered in the 5 THP-1 MR1-/- clones relative to the WT THP-1 cells the MR1 deficient and WT THP-1 cells were stained with pan MHC class I antibody. There was no significant difference in the levels of MHCI on the surface of the WT or the 5 THP-1 MR1-/- clones suggesting that targeting the MR1 for gene knockout had no effect on the expression of MHCI molecules as shown by Figure 5.16. An unpaired T test revealed that there was no significant difference between expression of classical MHC on the surface of the WT and THP-1 clones.

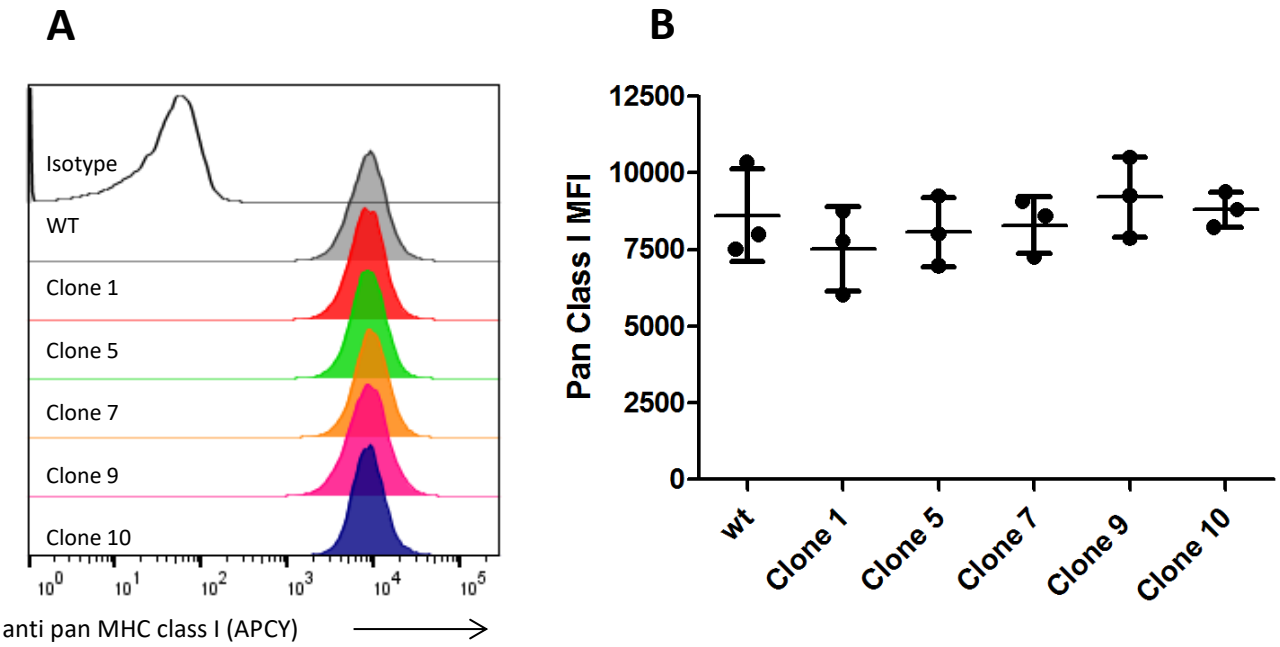


Figure 5.16. Expression of Pan MHCI in MR1 deficient THP-1 cells. A) The WT and the 5 THP-1 clones were stained with either an isotype control or with the pan MHC I antibody and the cells were analysed by flow cytometry. In MR1 deficient cells the expression of pan MHCI on the cell surface was not affected. The MFIs were isotype (22.5), WT (8051), C1 (7785), C5 (8025), C7 (8696), C9 (7914) and C10 (8244). B) The cells were stained in triplicate and the expression of pan MHCI was determined by flow cytometry. Individual MFI data points as well as mean values and standard error of the mean are shown. There was no significant difference between the expression of classical MHC on the WT and MR1 deficient clones. The P values were 0.4041, 0.6419, 0.7775, 0.6307 and 0.8515 for clone 1, 5, 7, 9 and 10 respectively.

5.4.15 Generation of MR1-/- MM909.24 clones

To further validate that the 'all in one' CRISPR/Cas9 system developed during this project was an effective system for producing MR1-/- cells, the MM909.24 tumour cell line was also targeted. The crRNA A was expressed either by lipofectamine transfection or by lentiviral transduction alongside a GFP control to establish the efficiency of transfection/transduction. Similarly to the THP-1 cells, the lipofectamine transfection was not a suitable method for

introducing the plasmid encoding the crRNA sequence into MM909.24 cells as the efficiency was extremely low. However, MR1 deficient MM909.24 cells were generated using the lentiviral transduction approach. Unlike the A549 and THP-1 cells the MM909.24 stained poorly for MR1 even after incubation with 6-f-p as analysed by flow cytometry. MR1 deficient clones were produced by incubating the bulk transduced MM909.24 cells with the MC7G5 MR1-restricted T-cell clone which enriched the $MR1^{neg}$ MM909.24 cells. 21 clones were isolated from the enriched cells by single cell cloning. As the MM909.24 cells expressed low levels of MR1 on the cell surface, the MR1 deficient cells were identified by incubating the WT MM909.24 and the 21 clones with the MC7G5 MR1-restricted T-cell clone for 16 h and performing a MIP-1β and a TNF-α ELISA with the supernatants to assess whether the clones could activate the MAITs at a similar level to the WT MM909.24 cells. The 21 clones appeared to be MR1 deficient as they were unable to activate the MC7G5 MAIT clone unlike the WT MM909.24 which was able to activate the MC7G5 MR1-resticted T-cell clone and induce the production of MIP-1β and TNF-α (Figure 5.17A and B). A surveyor assay was performed as described in section 2.4.12 using DNA extracted from the WT, pre selection, post selection and clones 1, 4, 8, 9 and 11. Mismatches at the crRNA locus were expected to produce two products (approximately 520 and 330 bp) after digestion with the Cel1 mismatch specific enzyme. Mismatches were observed in all 5 of the clones tested (Figure 5.17C)

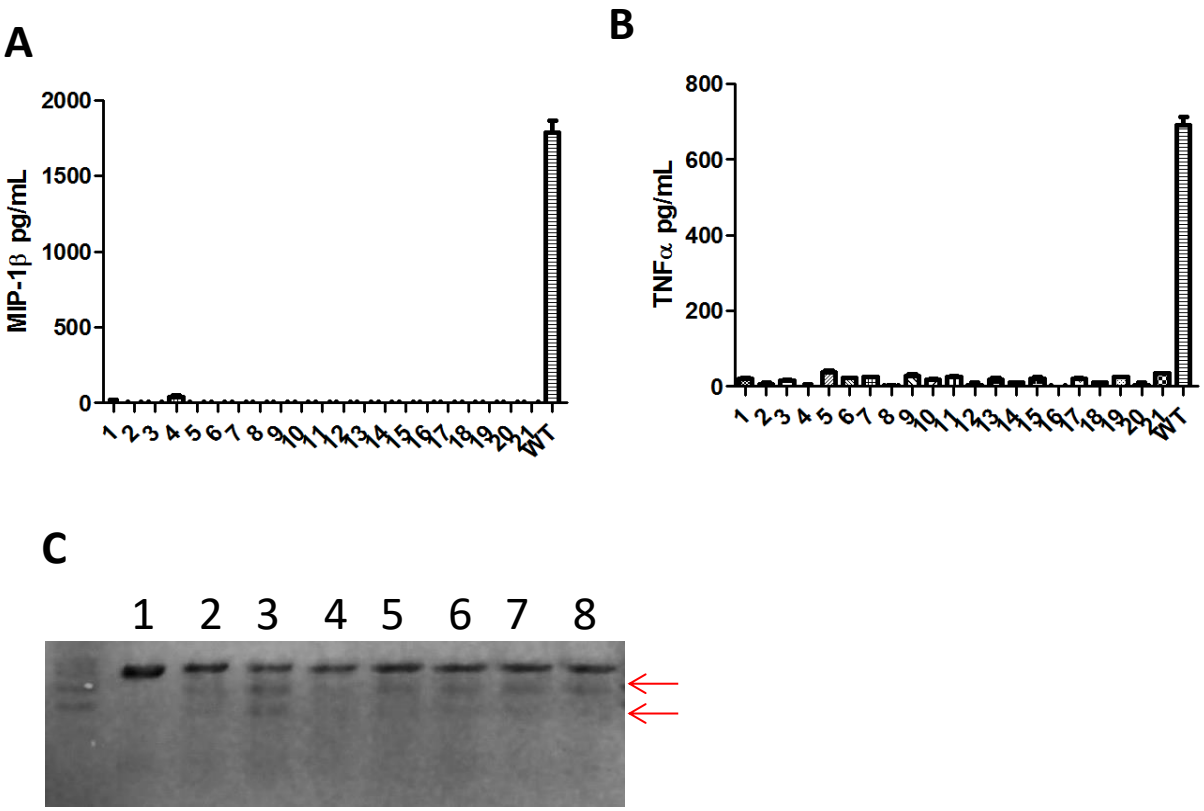


Figure 5.17. Phenotypic characterisation of MM909.24 clones generated by selection with the MC7G5 MR1 restricted T-cell clone. The 21 clones and WT MM909.24 cells were incubated it the MC7G5 MR1-restricted clone for 16 h after which the supernatants were collected and used to perform ELISAs to determine the concentration of A) MIP-1β and B) TNF-α produced respectively. C) The surveyor assay was performed on the WT (lane 1), pre-selection (lane 2), post-selection (lane 3) and clones 1 (lane 4), 4 (lane 5), 8 (lane 6), 9 (lane 7) and 11 (lane 8). The surveyor assay was performed using 500ng of DNA which was amplified using the MR1_SURV2_Fwd and MR1_SURV2_Rev primers amplified and annealed DNA from the WT and each of the clones. The red arrows indicate the presence of additional products for each of the clones.

5.4.16 Genotype characterization of disruptive mutations in the MR1 gene of MM909.24 MR1-/- clones

DNA was extracted from 5 of the MR1 deficient MM909.24 clones (clones 1, 4, 8, 9 and 11) in order to characterise the mutations present in the MR1 gene. The DNA was sequenced as described for the A549 and THP-1 cells by amplifying the DNA flanking the crRNA sequencing with MR1_SURV_Fwd_BsaI and MR1_SURV_Rev_BsaI primers which bound upstream and downstream of the crRNA respectively and introduced BsaI restriction site to the PCR product. The amplified DNA was cloned into a cloning vector (Addgene#32189). 10 colonies were selected for each MM909.24 clone and a miniprep of DNA was prepared and sequenced by Sanger sequencing.

On allele 1 of clone 1 there was a 91bp deletion and a nucleotide substitution (A>G) on the 5' end of the crRNA sequence. On allele 2 of clone 1 there was a 23bp deletion (Figure 5.18A). There was a 7bp deletion on allele 1 of clone 4 that encompassed the PAM sequence and 4bp of the crRNA, on allele 2 there was a 2bp deletion within the crRNA sequence. Clone 8 contained a biallelic single nucleotide deletion 4 bp from the PAM sequence. There was a biallelic single nucleotide deletion 4 bp from the PAM sequence on both alleles of clone 9, with the second allele containing an addition single base pair deletion 3bp from the PAM sequence. Both alleles of clone 11 encompassed a single bp deletion, on allele one the deletion was 4 bp from the PAM sequence and on allele 2 the deletion was 3 bp from the PAM sequence (Figure 5.18B).

Each mutation resulted in a frameshift the MR1 reading frame (Figure 5.18A and B). Thus, if these mutated alleles do express any protein it will not be functional MR1 (Figure 5.18C). The amino acid sequence of clone 1 is disrupted from amino acid 23 on allele 1 and amino acid 32 on allele 2 (Figure 5.18C). The amino acid sequence of clone 4 is disrupted from amino acid 32 on allele 1 and amino acid 33 on allele 2. The amino acid sequence of clone 8 and 9 is disrupted from amino acid 34 on both alleles. The amino acid sequence of clone 11 is disrupted from amino acid 34 on allele 1 and amino acid 33 on allele 2.

The lack of MAIT cell activation and the mutations presence within the MR1 gene suggest that the CRISPR/Cas9 system effectively abrogated MR1 expression from the cell surface of these MM909.24 clones.

A

Sample	DNA sequence
WT	GCAGTTGAAGGATCTGGATCATCTGGGACCCTACATGTCTTCCTTCTTTGCCTCCTTTCCAGGGACGCACTC TCTGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C1 a1	GCAGTT~~GAAGGATCTGGATCATCTGGGACCCTACATGTCTTCCTTCTTTGCCTCCTTTCCAGGGAC~~ ~~GCACTCTCTGAGATATTTTCGCCTGGGCGTT~~TCGGATCCCATCCGTGGGGTCCCTGAATTTATTT
C1 a2	GCAGTTGAAGGATCTGGATCATCTGGGACCCTACATGTCTTCCTTCTTTGCCTCCTTTCCAGGGACGCACTC TCTGAGATATTTTCGCCT~~GGGCGTTTCGGATCCCATCC~~ATGGGGTCCCTGAATTTATTT

B

Sample	DNA sequence
WT	GGACGCACTCTCTGAGATATTTTCGCCTGGGCGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C4 a1	GGACGCACTCTCTGAGATATTTTCGCCT~~GGGC~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C4 a2	GGACGCACTCTCTGAGATATTTTCGCCTG~~G~~CGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C8 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C8 a2	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C9 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C9 a2	GGACGCACTCTCTGAGATATTTTCGCCT~~G~~GG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C11 a1	GGACGCACTCTCTGAGATATTTTCGCCTGGG~~C~~GTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT
C11 a2	GGACGCACTCTCTGAGATATTTTCGCCTGG~~G~~CGTTTCGGATCCCATCCATGGGGTCCCTGAATTTATTT

C

Sample	Amino acid sequence
WT	SDSRTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWERYTQLLR
C1 a1	SDSLGSHPWGP!
C1 a2	SDSRTHSLRYFRWGP!
C4 a1	SDSRTHSLRYFRFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C4 a2	SDSRTHSLRYFRLRFGSHPWGP!
C8 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C8 a2	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C9 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C9 a2	SDSRTHSLRYFRLGFGSHPWGP!
C11 a1	SDSRTHSLRYFRLGFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!
C11 a2	SDSRTHSLRYFRLAFRIPSMGSLNLFRLGTWTRTLSPHMTVSLGRRSHGPHGWQRTSRLITGRGTLSC!

Figure 5.18. Characterisation of the MM909.24 clonal derivatives in the MR1 gene. A) The DNA sequence of the MR1 gene was determined by Sanger sequencing. The WT sequence is shown above and both alleles of clone 1 are shown below. B) The DNA sequence of the region flanking the crRNA binding sequence of MR1 in clones 4, 8, 9 and 11 was determined by Sanger sequencing. The WT sequence is shown above and all alleles of the clones are shown below. Nucleotide substitutions are shown as green highlighted letters. Deletions are shown as red strikethrough letters. Nucleotides in exon 2 of the MR1 gene are blue in colour and nucleotides in the intron upstream of exon 2 are shown as black letters. C) The predicted amino acid sequence was determined by translating the DNA sequences. The WT amino acid sequence from position 20 to 90 is shown above. Out of frame reads are highlighted in grey. Stop codons appear as '!'. (C1 = clone 1, C4 = clone 4, C8 = clone 8, C9 = clone 9, C11 = clone 11, a1 = allele 1 and a2 = allele 2.

5.4.17 Mutations do not affect the expression of classical HLA molecules in MM909.24s.

The WT and 5 MR1 deficient MM909.24 clones characterised in section 5.4.16 were stained with pan MHC class I antibody and analysed by flow cytometry to assess whether knocking out MR1 affected the expression of the classical MHC on the cell surface. There was no significance difference in the expression of the classical MHC molecules on the cell surface the WT MM909.24 and MR1-/- clones (Figure 5.19).

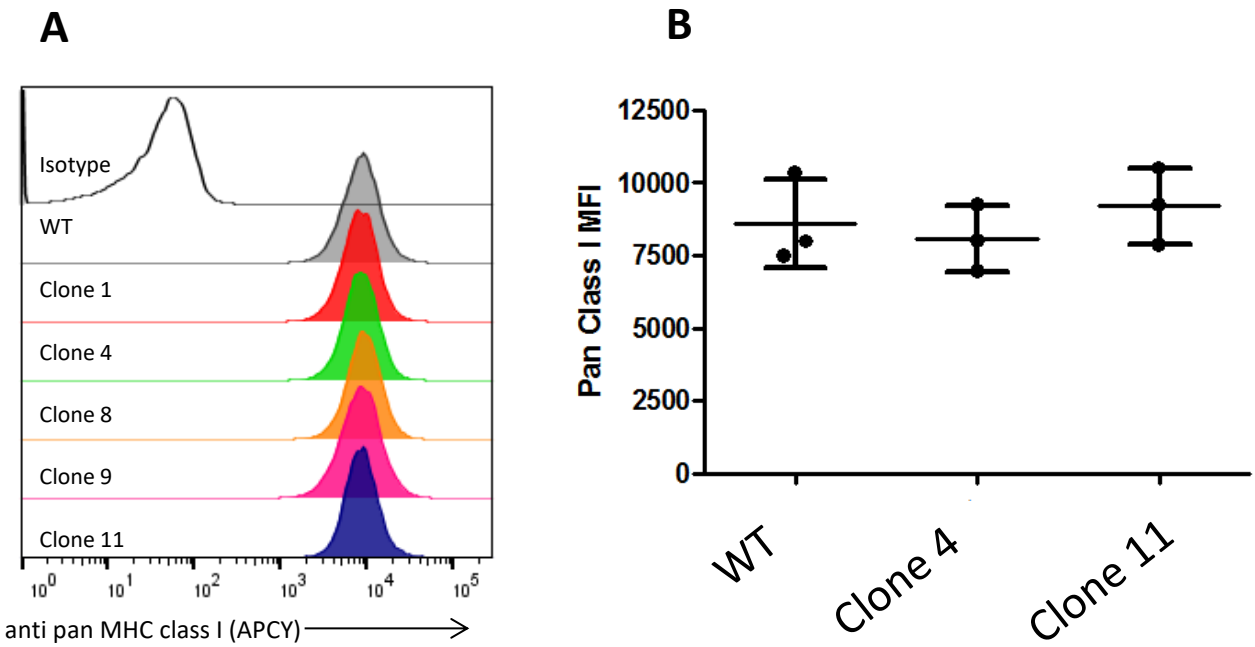


Figure 5.19. Expression of Pan MHCI in MR1 deficient MM909.24 cells. A) The WT and the 5 MM909.24 clones were stained with either an isotype control or with the pan MHCI antibody and the cells were analysed by flow cytometry. In MR1 deficient cells the expression of pan MHCI on the cell surface was not affected. The MFIs were isotype (19.5), WT (815), C1 21), C4 (613), C8 (785), C9 (610) and C11 (764). B) The WT, clone 4 and clone 11 cells were stained in triplicate and the expression of pan MHC I was determined by flow cytometry. Individual MFI data points as well as mean values and standard error of the mean are shown. There was no significance difference in pan MHCI expression between WT and the MR1 deficient MM909.24 clones. The p values were 0.3195 and 0.1109 for clones 4 and 11 respectively.

5.5 Discussion

Since the CRISPRCas9 system was first used for gene editing in 2013 (Mali et al. 2013; van der Oost 2013; Cong et al. 2013) it was been shown to be an efficient and reliable tool for knocking out a broad spectrum of human genes (Sanjana et al. 2014) in a range of cell lines and primary cells. As part of this project an "all-in-one" lentiviral vector single plasmid system was produced which allowed for simultaneous delivery of the Cas9 endonuclease and an associated gRNA sequences by modifying the two plasmid system produced by Mali et al. 2013. Five crRNA target sequences were designed to target the MR1 gene, after an initial screen in a reporter system a superior crRNA sequence (crRNA A) was identified. The CRISPR/Cas9 system with the verified crRNA sequence was efficient at inducing targeting knockout of the MR1 gene and was used to produce MR1 deficient A549, THP-1 and MM909.24 clones. MR1 deficient A549 clones were derived by cell sorting $MR1^{low}$ cells from the bulk transfected A549 and deriving clones by limiting dilution, 16 clones were screened by flow cytometry and 8 clones were observed to consistently express MR1 MFIs lower than the parental cells. A surveyor assay was performed on the 3 clones expressing the lowest levels of MR1, which revealed that mismatches were present at a DNA level in 2 of these clones. MR1 deficient THP-1 clones were derived by either positive enrichment with MAIT cells or by cell sorting the $MR1^{low}$ cells from the bulk transduced cells, in each case clones were derived through limiting dilution. The 20 clones derived from positive enrichment with MAIT cells and 10 clones derived from cell sorting were stained and the surface expression of MR1 was assessed. All of the clones appeared to be MR1 deficient. A surveyor assay was performed on 5 clones derived from positive enrichment with MAIT cells and 5 clones derived by cell sorting, leading to the confirmation of mismatches in all 10 clones. MR1 deficient MM909.24 clones were derived by selection with an MR1-restricted T-cell clone and MM909.24 clones were derived by limiting dilution. The 21 MM909.24 clones were determined to be MR1 deficient as they were unable to activate the MR1-restricted clone (MC7GS). A surveyor assay was performed on 5 of the clones, which revealed that mismatches were present in each of the 5 clones. As all of the THP-1 and MM909.24 clones derived from positive selection were unable to activate MAIT cells confirms that the positive selection approach is an effective method of selecting a potentially low frequency of MR1 deficient cells from a bulk population of cells. The positive

selection approach would be particularly beneficial in the context of primary cell lines. Disruption of the MR1 gene completely abrogated the level of MR1 expressed at the cell surface and the cells ability to present and activate MAIT cells when the clones are infected with *M. smegmatis* in the case of the A549 and THP-1 clones. In the case of the MM909.24 MR1 deficient clones the mutant forms of MR1 completely abrogated the cells ability to present to an MR1-restricted T cell. The MR1 deficient A549, THP-1 and MM909.24 clones produced were found to harbour mutations at a DNA level (as identified by Surveyor assay and Sanger sequencing of the region flanking the crRNA). Processing and presentation of antigen through the classical MHCI pathway was shown to be unaffected indicating that the CRISPR/Cas9 induced gene editing was specific for the MR1 gene.

5.5.1 Lipofectamine transfection vs lentiviral transduction approaches

The A549, THP-1 and MM909.24 cell lines were transfected transiently with lipofectamine and transduced stably with lentivirus. An MR1 deficient population was observable in the A549 cells using both approaches (Figure 6.4). In contrast, lipofectamine transfection in THP-1 and MM909.24 was very poor and MR1 deficient clones were generated by lentiviral transduction. There were no apparent adverse effects observed in cells transduced with the lentivirus. However, the lipofectamine approach is preferred as continuous expression of the CRISPR/Cas9 elements may have potential off target effects at later time points as Cas9 is continually transcribed.

5.5.2 Monitoring of MR1 expression on the surface of different cell lines

To confirm that CRISPR/Cas9 was a valid approach to knockout MR1 from cell lines the A549 were chosen to initially test the approach and once validated I extended the approach to THP-1 and MM909.24 cell lines. In the case of the A549s the cells required a 16 h incubation with 6-f-p to enhance the expression of MR1 on the cell surface after which a population of MR1 deficient cells was identifiable from the bulk transfected population. The THP-1 cell line expressed sufficient MR1 on the cell surface that the cells did not need to be pre-treated with 6-f-p prior to staining. Even after incubation with 6-f-p there were not sufficient levels of MR1 on the cell surface of MM909.24 cells to be detectable using a fluorochrome conjugated MR1 antibody.

5.5.3 Characterisation of mutations within the MR1 gene of the A549, THP-1 and MM909.24 clonal derivatives

Sanger sequencing revealed that mutations had been introduced into the MR1 loci in all of the A549, THP-1 and MM909.24 clones sequenced. Of particular note are A549 clone 9 allele 1 (126bp deletion), THP-1 Clone 7 allele 1 and 2 (33bp deletion), THP-1 clone 9 allele 1 (45bp deletion), MM909.24 clone 1 allele 1 (91bp deletion) and MM909.24 clone 1 allele 2 (23bp deletion) as each of these alleles encompasses a fairly large deletion. Large deletions are not typically observed as a result of NHEJ DNA repair. The mutations observed in the alleles of the remaining clones are more typically, of what is normally observed as a result of NHEJ DNA repair in the literature.

5.5.4 Monitoring the expression of the classical MHCI molecules on the surface of the MR1 deficient clonal derivatives

Despite there being little homology between the MR1 target sequence and sequences present within the other classical HLA genes, it was important to ensure that targeting MR1 did not affect the expression of the other classical HLA molecules or interfere with the antigen presentation pathways. Although this method was not extensive, the analysis of the surface expression of classical HLA class I molecules in A549, THP-1 and MM909.24 clonal derivatives was not significantly different to the WT cell line (Figure 5.11 and 5.15). In addition, the presentation of cognate exogenous peptide to an HLA-B*1801-restricted $CD8^+$ T-cell clone by the A549 derivatives led to production of cytokines at the same level in MR1 deficient and unmodified A549 cells. Collectively, these results indicate that the expression and processing of HLA alleles and peptide presentation are intact within these cells.

5.5.5 Monitoring of off target mutations

The morphology and growth rate of the MR1 knockout cells were unaffected compared to the WT. However, this does not exclude other potential off target sites in other unrelated genes. To assess the frequency and occurrence of mutations in undesired off target sites whole genome sequencing was performed in the A549 clone 9 and 11. Several off target sites were predicted by the CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre) (Table 1 and 2). In addition, analysis of the reads identified approximately

1400 sites where reads harboured potential off target mutations. Analysis of these sites manually did not identify any off target mutations. On target mutations were confirmed at the MR1 locus and at an additional unintentional target site within the RP11-46A10.6 pseudogene which shared the crRNA target sequence with the MR1 gene. Interestingly, the mutations observed in both A549 clones (clone 9 and clone 11) within the RP11-46A10.6 gene were identical to the most common mutation observed within the MR1 gene (a single base deletion at position 17 of the crRNA sequence). Although unintended, this mutation is considered on target as the crRNA sequence was present in both MR1 and the RP11-46A10.6 gene. The initial assessment of potential off target sites using the BLAST tool did not identify the RP11-46A10.6 gene, as the BLAST tool was limited to coding genomic regions. Given that this RP11-46A10.6 gene is noncoding there should be no consequences to this off target mutation. The introduction of mutations within both the MR1 and RP11-46A10.6 gene highlights both the high efficiency and specificity of the CRISPR/Cas9 system. Since the initial screening of the crRNA sequences was performed several algorithms have become available online (CRISPRdesign tool (MIT) and eCRISP software (German cancer research centre)) which allow for rapid identification of all potential off target sites within coding and non-coding DNA. These algorithms take into account data from previous studies which have shown that the 13 nucleotides proximal to the PAM sequence (3' end) are crucial for crRNA binding. Differences between the target sequence and crRNA sequence at the 3' end abolish binding and vastly reduce the frequency of gene editing. In contrast up to 6 mismatches at the 5' end of the crRNA sequence can be tolerated (Jinek et al. 2012). These algorithms can be used to identify crRNA sequences with low potential for off target binding. These programmes can be used to identify the most likely off target sites which can then by screened to assess the occurrence and frequency of off target mutations by amplicon sequencing to identify the presence of off target mutations at a lower cost than whole genome sequencing.

5.5.6 Summary

In this chapter I produced an 'all in one' CRISPR/Cas9 plasmid which incorporated all of the CRISPR/Cas9 elements in a single plasmid as initial studies with CRISPR/Cas9 required multiple plasmids for gene disruption. Using this 'all in one' CRISPR/Cas9 system I generated

A549, THP-1 and MM909.24 clonal derivatives which were MR1 deficient. The clonal derivatives did not express MR1 on the cell surface and were unable to activate MAIT cells indicating that the cells were MR1 deficient. It is hoped that the cells produced in chapter of the thesis will be used to understand the biology of MR1 and MR1 restricted T-cells. We have already had many requests for these cells since publication in June this year (Laugel et al. 2016) and it seems likely that they will go on to make a substantial contribution to the field. Thus far, my laboratory has already collaborated to use these cells to confirm the existence of a new family of MR1-restricted T-cells that do not express the canonical TRAV1-2 chains or recognise canonical MAIT antigens (Meermeier et al. 2016).

6 General discussion

The discovery and development of gene editing tools that can induce targeted knockout of specific genes has been revolutionary by enabling researchers to study the function of proteins and pathways in ways not previously possible. Genome editing was awarded method of the year in 2011 by Nature (http://www.nature.com/nmeth/journal/v9/n1/full/nmeth.1852.html), and genome engineering was runner up to breakthrough of the year awarded by Science in both 2012 and 2013. More recently, the CRISPR/Cas9 system was awarded breakthrough of the year by Science in 2015 (Travis 2015). These awards serve to highlight the importance of the topic.

6.1 Summary of work

In this thesis, shRNA, ZFN and CRISPR/Cas9 gene editing tools were generated that targeted the CD8A gene as proof of concept. The shRNA, ZFN and CRISPR/Cas9 approaches all functioned, resulting in a CD8 negative population within the transduced cells. Of the 3 approaches the CRISPR/Cas9 system was superior in terms of cost of production, ease of development and efficiency of gene knockout (chapter 3).

A novel non-nuclease based gene silencing system was generated as an alternative to the current nuclease systems by fusing the ZFP targeting the CD8A gene to the KRAB repressor domain as proof of concept. This system was shown to induce silencing of CD8 in both the Molt3 cell line and primary CD8+ T-cells. The ZFP-KRAB fusion was combined in a single plasmid with a TCR which allowed for silencing of CD8 in combination with delivery of a TCR (chapter 4).

A single plasmid system incorporating all of the CRISPR/Cas9 elements was developed to improve knockout efficiency of the two plasmid system described by Mali et al. 2013. The 'all in one' CRISPR/Cas9 system was utilised to knock out MR1 from a range of APCs to generate MR1-deficient cells. The MR1 deficient APC clones derived during this project were infected with *M. smegmatis* were unable to activate MAIT cells clones indicating that the mutations induced by the CRISPR/Cas9 system abolished MR1 expression from the cell surface (Laugel et al. 2016) (chapter 5).

6.2 Implications of findings

6.2.1 CRISPR/Cas9 system

Of the three gene editing approaches utilised in this thesis the CRISPR/Cas9 approach was the most cost effective, most efficient and easiest system to modify to incorporate new targets, this corroborates with the published literature and thus CRISPR/Cas9 would be the approach of choice for inducing targeted knockout within cell lines. However, the safety of transducing cells modified with CRISPR/Cas9 into human patients has yet to be tested, with the first clinical trial for CRISPR/Cas9 being approved in July 2016 and commencing in August 2016 (Cyranoski 2016). There is currently an ongoing debate regarding the ownership of the CRISPR/Cas9 technology which once resolved could lead to licensing costs and future limitations to the use of CRISPR/Cas9 technology. This could have severe consequences for translational exploitation of CRISPR/Cas9 by researchers worldwide (Sherkow 2016).

6.2.2 The ZF-KRAB approach

The ZF-KRAB gene silencing approach which was developed during this thesis (chapter 4) was shown to be a valid alternative to the nucleases approaches. The ZF-KRAB construct induced gene silencing in the majority of transduced cells. The absence of nuclease activity is advantageous as it eliminates the risk of off-target DNA damaging DSBs within the DNA. Repair of DSBs can cause translocations and genomic instability which can lead to tumorigenesis (reviewed by Khanna & Jackson 2001; Aparicio et al. 2014). This ZF-KRAB gene silencing system could benefit T-cell based immunotherapies as the ZF-KRAB DNA sequence can be incorporated into a single plasmid with a TCR which would allow for targeted gene knockout and re-direction of a T-cell in a single plasmid. The knockout of PD1 or CTLA-4 checkpoint inhibitors in primary T-cells while simultaneously introducing a tumour-specific TCR would be of particular interest as it would allow checkpoint control of only tumour-specific responses as opposed to the systemic blocking antibody approaches currently in use and that are known to induce autoimmune side effects.

6.2.3 Generation of MR1 deficient cells

The APCs generated in this thesis will enable fundamental investigations into the biology of MR1 and MR1-restricted T-cells. 1) The use of MR1 deficient cells will help identify MR1 restricted T-cells better than current techniques such as antibody blockade which can lead to incomplete blockade. The MR1 deficient A549 clones produced during this work have been used to identify a population of MR1 restricted T-cells which do not express the canonical TCR chains (Meermeier et al. 2016). 2) Using MR1 deficient APCs, in combination with WT APCs in *in vitro* activation experiments will allow for the characterisation and quantification of MR1-restricted T-cells within the bulk pathogen-specific T-cell response. 3) MR1 deficient cells may proof useful for delineating the mechanisms behind the processing and loading of antigen on MR1 molecules. Currently, whether T-cells can discriminate between distinct MR1 antigen complexes and the chemical and structural variability of bacterial antigens presented by MR1 is unknown. However, pulsing MR1 sufficient and deficient cells with chemically pure compounds could be utilised to assess the response of T-cell clones with distinct TCR usage. 4) The production of MR1 deficient cell lines will enable researchers to experiment with altered versions of MR1 proteins to establish whether improved or chimeric versions can improve function and allow characterisation of the cellular biology of MR1 such as distribution in cellular compartments and trafficking. The invariant nature of MR1 across the population makes MR1 an attractive target for the development of therapeutic and diagnostic tools and thus further investigations into MR1 biology and MR1-restricted T-cells may prove valuable for universal therapies.

6.3 Implications of adoptive cell transfer with genetic engineered cells

The process of adoptive cell therapy (ACT) of autologous tumour infiltrating lymphocytes to treat cancer was first described in 1988 (Rosenberg et al. 1988). Since this initial publication, ACT protocols have improved through lymphodepletion prior to ACT which was shown to improve the level of cancer regression (Dudley et al. 2002; Dudley et al. 2005) and administration of IL-2 which improved the therapeutic potency of T-cells (Rosenberg et al. 1986). ACT has proven to be effective at inducing clearance of cancers in a variety of studies

(Yee et al. 2002; Mackensen et al. 2006). The transfer of T-cells genetically engineered to express cancer specific TCRs has resulted in successfully tumour clearance (Morgan et al. 2006). The ZFN and CRISPR/Cas9 systems have great potential in the ACT setting as T-cells can be further engineered to improve function while simultaneously re-programming to target a specific antigen. The ZF-KRAB TCR system described in this thesis could potentially be utilised to genetically engineer cells for ACT by silencing genes such as those encoding immune checkpoint inhibitors and re-directing T-cells to target tumour associated antigens through the transfer of a tumour specific TCR.

6.3.1 Long-term effects of genetic engineered cells

There are several potential issues with transferring genetic engineered T-cells into patients, such as TCR chain mispairing, lentiviral integration into unknown sites in the genome and immunogenicity.

6.3.1.1 **TCR mispairing**

When transducing T-cells with a TCR there is potential for the TCR α or β chains to mispair with the endogenous TCR chains, resulting in the expression of a TCR with an unknown specificity. Such TCR mispairing in mice led to autoimmunity (Bendle et al. 2010), however this has not been recorded in human clinical trials (Rosenberg 2010). TCR mispairing has been shown to be minimised by the use of murinised TCRs which are expressed at higher levels in human T-cells than human TCRs (Cohen et al. 2006). Fusion of human TCR variable chains to mouse constant domains promotes pairing of the exogenous chains (Sommermeyer & Uckert 2010). Furthermore, an additional disulphide bond between the TCR α and β chains through the addition of an cysteine residues on each chain additionally promotes pairing of the exogenous chains (Kuball et al. 2007). Knocking out the expression of the existing endogenous TCR chains using gene editing tools (Provasi et al. 2012) can also be used to minimise TCR chain mispairing. Finally transfer of CD3 components in addition to TCR α and β chains promotes pairing and surface expression of the exogenous TCR chains (Sebestyen et al. 2008; Ahmadi et al. 2011). While TCR chain mispairing might cause some problems, there are therefore several novel ways to minimise this potential problem.

6.3.1.2 Integrating versus non-integrating approaches for expression of gene editing tools

Long-term expression of the gene editing tools and TCR chains in cells in this thesis was achieved by integrating lentiviral transduction. One of the major risks of integrating lentivirus as a delivery system is integration mutagenesis. Integration near a promoter may alter gene expression (Reviewed by Sinn et al. 2005) and integration near oncogenes can drive cellular proliferation and lead to the development of cancers (Hacein-Bey-Abina et al. 2003). In addition, the transgene is continually expressed within the cell and in the case of TCR transduction this is desirable. In contrast, in the case of gene editing tools this could be problematic as continual expression of the nuclease protein may result in off target mutagenesis after transduction. Nevertheless, comparison of plasmid transfection and lentiviral transduction of ZFN resulted in similar levels of off target disruption (Cai et al. 2014). There are alternative methods of transgene transfer including non-integrating lentivirus and plasmid transfection. Non-integrating lentiviral transduction has been successful used to express ZFNs transiently to induce knockout of the CCR5 gene (Holt, et al. 2010). Cationic lipid based transfection methods have been used to transiently express gene editing tools in human cells (Miller et al. 2007; Zuris et al. 2014). The presence of mutations at undesired locations induced as an off target effect of gene editing tools can easily be detected by whole genome sequencing as performed in chapter 5 (Laugel et al. 2016) and other studies to detect off target effects (Smith et al.2014; Veres et al. 2014).

6.3.1.3 Suicide genes

To improve the safety of genetically engineered cells for ACT it would be advantageous to incorporate a suicide gene into the construct containing the gene editing tool, which would allow for eradication of transduced cells in the event of adverse effects post adoptive transfer. One way of inducing suicide is with the use of a prodrug. The first prodrug used for transduced T-cells which has been through clinical testing was Herpes simplex virus thymidine kinase/ ganciclovir (HSV-TK/ GCV). GCV was introduced into cells with an adeno associated virus. When active, HSV-TK phosphorylates GCV which interferes with DNA synthesis killing the transduced cells (Pan et al. 2012), however, this approach was slow as interference with DNA can take several days to induce cell death. As an alternative approach, a iCasp9 transgene was expressed in transduced cells, administration of the

AP1903 drug causes dimerization and activation of the iCasp9 gene resulting in the induction of the apoptotic pathway and cell death (Di Stasi et al. 2011).

6.4 Future work

The generation of CD8 deficient T-cells allows for investigation of the CD8 molecule in greater detail as previous studies have worked with antibody blocking (MacDonald et al. 1982) or with mutated MHC (Laugel et al. 2007), whereas this approach would allow for studying T-cells with CD8 abolished from the cell surface. Comparison of T-cell activation in the presence and absence of the CD8 molecule ± HLA with and without the ability to engage CD8 will allow dissection of which CD8-mediated events occur via CD8-HLA interaction and those that are independent of this interaction and occur, for instance via the TCR-CD8 interaction (Arcaro et al. 2000; Arcaro et al. 2001). There is potential to target many other genes encoding proteins expressed by T-cells, of particular interest are checkpoint inhibitors such as CTLA-4 and PD1 for which blocking antibodies have shown to be effective for tumour clearance in mice (Fecci et al. 2007; Hirano et al. 2005) and humans (Hodi et al. 2010; Topalian et al. 2012). It is thought that cell specific knockout of checkpoint inhibitors will minimise the adverse effects that have been observed with systemic delivery of blocking monoclonal antibodies. There may also be the potential in future to modify and improve the function of TCR-modified Treg via gene knockout once the mechanisms by which these cells operate are delineated.

There is potential to substitute the ZFP and TCR chains contained within the ZF-KRAB TCR construct to produce a toolkit of constructs that would allow for the generation of T-cells deficient in numerous checkpoint inhibitors with different specificities. These T-cells could be useful in determining whether this approach would have any benefit to the clinic.

A loss of function was observed in the A549, THP-1 and MM909.24 as a result of an MR1 deficiency, however, to validate this I plan to re-express the MR1 protein in these cells and repeat the *M. smegmatis* infection and activation experiments to show that function is restored upon re-expression of MR1. Another line of work which I wish to pursue is the generation of APCs deficient in members of the CD1 family, HLA-E or other MHCIb genes

that may serve as ligands for the TCR (Attaf et al. 2015). This would generate additional cell lines for the study of these other invariable T-cell ligands.

6.4.1 Concluding remarks

Gene editing tool technology has developed rapidly over recent years. To date there have been no studies which have compared all of the shRNA, ZFN and CRISPR/Cas9 systems in parallel. As part of this thesis I compared these systems. The CRISPR/Cas9 system was found to be the most efficient and cost effective system. Therefore, the CRISPR/Cas9 system would be the method of choice for generating knockout cells in the future.

There are risks associated with the expression of nuclease proteins in cells to generate cells deficient in a target protein, particularly if these cells are to be transfused into patients for therapy. Therefore, a non-nuclease based gene silencing system was generated by the fusion of a ZFP targeting the CD8A gene to the KRAB repressor domain. This system was shown to induce silencing of the CD8 protein at levels comparable to the ZFN approach. The ZF-KRAB was cloned into a lentiviral vector alongside a TCR, enabling expression of a specific TCR and silencing of the CD8 as proof of concept. There is great potential for this system in the immunotherapy field as the ZF-KRAB approach would lead to specific gene silencing whilst eradicating DNA damage and breakages. Although not tested in this thesis there is potential to utilise ZFPs targeting PD-1, CTLA-4 or other checkpoint inhibitors in this system which could improve the functionality of transduced T-cells whilst simultaneously expressing an alternate TCR to redirect T-cells.

Finally, an ‘all in one’ CRISPR/Cas9 system was developed which was effectively utilised to generate MR1 deficient clonal derivatives of the A549, THP-1 and MM909.24 cell lines. The generation of MR1 deficient cells will be useful for fundamental investigations into the biology of MR1 and MR1-restricted T-cells and the development of therapeutic and diagnostic tools which aim to utilise this invariant non-classical HLA molecule. It is hoped that the generation cell lines deficient in HLA-E, members of CD1 family and other MHCIb molecules will enable a greater understanding into the biology of the HLA-E, CD1 and other MHCIb molecules and the T-cells that recognise antigen presented by these molecules.

7 References

Abrink, M., Ortiz, J.A., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P. & Losson, R., 2001. Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1 beta. *Proc Natl Acad Sci U S A*, 98(4), pp.1422–1426.

Ahmadi, M., King, J.W., Xue, S.-A.A., Voisine, C., Holler, A., Wright, G.P., Waxman, J., Morris, E. & Stauss, H.J., 2011. CD3 limits the efficacy of TCR gene therapy in vivo. *Blood*, 118(13), pp.3528–37.

Ahmed, R. & Rouse, B.T., 2006. Immunological Memory. *Immunological reviews*, (211), pp.5–7.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J., 1990. Basic local alignment search tool. *Journal of molecular biology*, 215(3), pp.403–410.

Annibali, V., Ristori, G., Angelini, D.F., Serafini, B., Mechelli, R., Cannoni, S., Romano, S., Paolillo, A., Abderrahim, H., Diamantini, A., Borsellino, G., Aloisi, F., Battistini, L. & Salvetti, M., 2011. CD161highCD8+T cells bear pathogenetic potential in multiple sclerosis. *Brain*, 134(2), pp.542–554.

Aparicio, T., Baer, R. & Gautier, J., 2014. DNA double-strand break repair pathway choice and cancer. *DNA Repair*, 19, pp.169–175.

Arcaro, A., Grégoire, C., Bakker, T.R., Baldi, L., Jordan, M., Goffin, L., Boucheron, N., Wurm, F., van der Merwe, P.A., Malissen, B. & Luescher, I.F., 2001. CD8β Endows CD8 with Efficient Coreceptor Function by Coupling T Cell Receptor/CD3 to Raft-associated CD8/p56 lck Complexes. *The Journal of Experimental Medicine*, 194(10), pp.1485–1495.

Arcaro, A., Gregoire, C., Boucheron, N., Stotz, S., Palmer, E., Malissen, B. & Luescher, I.F., 2000. Essential Role of CD8 Palmitoylation in CD8 Coreceptor Function. *The Journal of Immunology*, 165(4), pp.2068–2076.

Asseman, C., Mauze, S., Leach, M.W., Coffman, R.L. & Powrie, F., 1999. An Essential Role for Interleukin 10 in the Function of Regulatory T Cells That Inhibit Intestinal Inflammation. *The Journal of Experimental Medicine*, 190(7), pp.995–1004.

Assimakopulos, SF; Scopa, ChD; Vagianos, C., 2007. Pathophysiology of increased permeability in obstructive jaundice. *World Journal of Gastroenterology*, 13(48), pp.6455–6618.

Attaf, M., Legut, M., Cole, D.K. & Sewell, A.K., 2015. The T cell antigen receptor: the Swiss army knife of the immune system. *Clinical & Experimental Immunology*, 181(1), pp.1–18.

Ayer, D.E., Laherty, C.D., Lawrence, Q.A., Armstrong, A.P. & Eisenman, R.N., 1996. Mad proteins contain a dominant transcription repression domain. *Molecular and cellular biology*, 16(10), pp.5772–81.

Ayyanathan, K., Ayyanathan, K., Lechner, M.S., Lechner, M.S., Bell, P., Bell, P., Maul, G.G., Maul, G.G., Schultz, D.C. & Schultz, D.C., 2003. Heterochromatin protein 1 (HP1) is a key component of constitutive heterochromatin in. *Genes & Development*, pp.1855–1869.

Bailey, P.J., Dowhan, D.H., Franke, K., Burke, L.J., Downes, M. & Muscat, G.E., 1997. Transcriptional repression by COUP-TF II is dependent on the C-terminal domain and involves the N-CoR variant, RIP13delta1. *The Journal of steroid biochemistry and molecular biology*, 63(4–6), pp.165–174.

Banham, A.H. & Smith, G.L., 1993. Characterization of vaccinia virus gene B12R. *The Journal of general virology*, 74. Pt 12, pp.2807–2812.

Barral, D.C. & Brenner, M.B., 2007. CD1 antigen presentation: how it works. *Nature reviews. Immunology*, 7(12), pp.929–941.

Beerli, R.R., Dreier, B. & Barbas, C.F., 2000. Positive and negative regulation of endogenous genes by designed transcription factors. *Proceedings of the National Academy of Sciences* , 97(4), pp.1495–1500.

Bellefroid, E.J., Poncelet, D. a, Lecocq, P.J., Revelant, O. & Martial, J.A., 1991. The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 88(9), pp.3608–12.

Bendle, G.M., Linnemann, C., Hooijkaas, A.I., Bies, L., de Witte, M.A., Jorritsma, A., Kaiser, A.D.M., Pouw, N., Debets, R., Kieback, E., Uckert, W., Song, J.-Y., Haanen, J.B. a G. & Schumacher, T.N.M., 2010. Lethal graft-versus-host disease in mouse models of T cell receptor gene therapy. *Nature medicine*, 16(5), p.565–70, 1p following 570.

Berdien, B., Mock, U., Atanackovic, D. & Fehse, B., 2014. TALEN-mediated editing of endogenous T-cell receptors facilitates efficient reprogramming of T lymphocytes by lentiviral gene transfer. *Gene Therapy*, 21(6), pp.539–548.

Bialk, P., Rivera-Torres, N., Strouse, B. & Kmiec, E.B., 2015. Regulation of Gene Editing Activity Directed by Single-Stranded Oligonucleotides and CRISPR/Cas9 Systems. *PloS one*, 10(6), p.e0129308.

Billerbeck, E., Kang, Y.-H., Walker, L., Lockstone, H., Grafmueller, S., Fleming, V., Flint, J., Willberg, C.B., Bengsch, B., Seigel, B., Ramamurthy, N., Zitzmann, N., Barnes, E.J., Thevanayagam, J., Bhagwanani, A., Leslie, A., Oo, Y.H., Kollnberger, S., Bowness, P., Drognitz, O., Adams, D.H., Blum, H.E., Thimme, R. & Klenerman, P., 2010. Analysis of CD161 expression on human CD8+ T cells defines a distinct functional subset with tissue-homing properties. *Proceedings of the National Academy of Sciences* , 107(7), pp.3006–3011.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. & Wiley, D.C., 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*, 329(6139), pp.506–12.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. & Bonas, U., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science (New York, N.Y.)*, 326(5959), pp.1509–12.

Bolotin, A., Quinquis, B., Sorokin, A. & Dusko Ehrlich, S., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 151(8), pp.2551–2561.

Bouard, D., Alazard-Dany, D. & Cosset, F.-L., 2009. Viral vectors: from virology to transgene expression. *British journal of pharmacology*, 157(2), pp.153–65.

Boulter, J.M., Glick, M., Todorov, P.T., Baston, E., Sami, M., Rizkallah, P. & Jakobsen, B.K., 2003. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein engineering*, 16(9), pp.707–711.

Le Bourhis, L., Dusseaux, M., Bohineust, A., Bessoles, S., Martin, E., Premel, V., Cora, M., Sleurs, D., Serriari, N.-E., Treiner, E., Hivroz, C., Sansonetti, P., Gougeon, M.-L., Soudais, C. & Lantz, O., 2013. MAIT Cells Detect and Efficiently Lyse Bacterially-Infected Epithelial Cells. *PLoS Pathogens*, 9(10), p.e1003681.

Le Bourhis, L., Martin, E., Péguillet, I., Guihot, A., Froux, N., Coré, M., Lévy, E., Dusseaux, M., Meyssonnier, V., Premel, V., Ngo, C., Riteau, B., Duban, L., Robert, D., Huang, S., Rottman, M., Soudais, C. & Lantz, O., 2010. Antimicrobial activity of mucosal-associated invariant T cells. *Nature immunology*, 11(8), pp.701–708.

Brahmer, J.R., Drake, C.G., Wollner, I., Powderly, J.D., Picus, J., Sharfman, W.H., Stankevich, E., Pons, A., Salay, T.M., McMiller, T.L., Gilson, M.M., Wang, C., Selby, M., Taube, J.M., Anders, R., Chen, L., Korman, A.J., Pardoll, D.M., Lowy, I. & Topalian, S.L., 2010. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: Safety, clinical activity, pharmacodynamics, and immunologic correlates. *Journal of Clinical Oncology*, 28(19), pp.3167–3175.

Brown-Elliott, B.A. & Wallace, R.J., 2002. Clinical and Taxonomic Status of Pathogenic Nonpigmented or Late-Pigmenting Rapidly Growing Mycobacteria. *Clinical Microbiology Reviews* , 15(4), pp.716–746.

Brunet, E., Simsek, D., Tomishima, M., DeKelver, R., Choi, V.M., Gregory, P., Urnov, F., Weinstock, D.M. & Jasin, M., 2009. Chromosomal translocations induced at specified loci in human stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26),

pp.10620–5.

Cai, Y., Bak, R.O. & Mikkelsen, J.G. iehm, 2014. Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. *eLife*, 3, p.e01911.

Callahan, M.K., Wolchok, J.D. & Allison, J.P., 2010. Anti-CTLA-4 Antibody Therapy: Immune Monitoring During Clinical Development of a Novel Immunotherapy. *Seminars in oncology*, 37(5), pp.473–484.

Cao, Q., Zhao, K., Zhong, X.Z., Zou, Y., Yu, H., Huang, P., Xu, T.-L. & Dong, X.-P., 2014. SLC17A9 protein functions as a lysosomal ATP transporter and regulates cell viability. *The Journal of biological chemistry*, 289(33), pp.23189–23199.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. a, Somia, N. V, Bogdanove, A.J. & Voytas, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*, 39(12), p.e82.

Cheng, Z., Yi, P., Wang, X., Chai, Y., Feng, G., Yang, Y., Liang, X., Zhu, Z., Li, W. & Ou, G., 2013. Conditional targeted genome editing using somatically expressed TALENs in C. elegans. *Nat Biotech*, 31(10), pp.934–937.

Cho, S.W., Kim, S., Kim, J.M. & Kim, J.-S., 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology*, 31(3), pp.230–2.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. & Voytas, D.F., 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 186(2), pp.757–61.

Christian, M.L., Demorest, Z.L., Starker, C.G., Osborn, M.J., Nyquist, M.D., Zhang, Y., Carlson, D.F., Bradley, P., Bogdanove, A.J. & Voytas, D.F., 2012. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PloS one*, 7(9), p.e45383.

Chu, H.W., Rios, C., Huang, C., Wesolowska-Andersen, A., Burchard, E.G., O'Connor, B.P., Fingerlin, T.E., Nichols, D., Reynolds, S.D. & Seibold, M.A., 2015. CRISPR-Cas9-mediated gene knockout in primary human airway epithelial cells reveals a proinflammatory role for MUC18. *Gene therapy*, 22(10), pp.822–829.

Circosta, P., Granziero, L., Follenzi, A., Vigna, E., Stella, S., Vallario, A., Elia, A.R., Gammaitoni, L., Vitaggio, K., Orso, F., Geuna, M., Sangiolo, D., Todorovic, M., Giachino, C. & Cignetti, A., 2009. T cell receptor (TCR) gene transfer with lentiviral vectors allows efficient redirection of tumor specificity in naive and memory T cells without prior stimulation of endogenous TCR. *Human gene therapy*, 20(12), pp.1576–1588.

Cohen, C.J., Zhao, Y., Zheng, Z., Rosenberg, S.A. & Morgan, R.A., 2006. Enhanced antitumor activity

of murine-human hybrid T-cell receptor (TCR) in human lymphocytes is associated with improved pairing and TCR/CD3 stability. *Cancer Research*, 66(17), pp.8878–8886.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L. a & Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339(6121), pp.819–23.

Cong, L., Zhou, R., Kuo, Y.-C., Cunniff, M. & Zhang, F., 2012. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature communications*, 3, p.968.

Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G., 2013. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic acids research*, 41(20), pp.9584–9592.

Cradick, T.J., Keck, K., Bradshaw, S., Jamieson, A.C. & McCaffrey, A.P., 2010. Zinc-finger Nucleases as a Novel Therapeutic Strategy for Targeting Hepatitis B Virus DNAs. *Mol Ther*, 18(5), pp.947–954.

Cui, X., Ji, D., Fisher, D. a, Wu, Y., Briner, D.M. & Weinstein, E.J., 2011. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nature biotechnology*, 29(1), pp.64–7.

Cyranoski D., 2016. First trial of CRISPR in people. *Nature*, 535, pp.476–477.

David-Watine, B., 2001. The human gephyrin (GPHN) gene: structure, chromosome localization and expression in non-neuronal cells. *Gene*, 271(2), pp.239–245.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. & Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), pp.602–607.

Didigu, C. a., Wilen, C.B., Wang, J., Duong, J., Secreto, A.J., Danet-Desnoyers, G. a., Riley, J.L., Gregory, P.D., June, C.H., Holmes, M.C. & Doms, R.W., 2014. Simultaneous zinc-finger nuclease editing of the HIV coreceptors ccr5 and cxcr4 protects CD4+ T cells from HIV-1 infection. *Blood*, 123(1), pp.61–69.

van Diemen, F.R., Kruse, E.M., Hooykaas, M.J.G., Bruggeling, C.E., Schürch, A.C., van Ham, P.M., Imhof, S.M., Nijhuis, M., Wiertz, E.J.H.J. & Lebbink, R.J., 2016. CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections. *PLoS Pathog*, 12(6), p.e1005701.

Ding, Q., Lee, Y.-K., Schaefer, E.A.K., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T., Trevisan, M., Gupta, R.M., Moisan, A., Banks, E., Friesen, M., Schinzel, R.T., Xia, F., Tang, A., Xia, Y., Figueroa, E., Wann, A., Ahfeldt, T., Daheron, L., Zhang, F., Rubin, L.L., Peng, L.F., Chung, R.T., Musunuru, K. & Cowan, C.A., 2013. A TALEN Genome-Editing System for Generating Human Stem Cell-Based Disease Models. *Cell Stem Cell*, 12(2), pp.238–

251.

Ding, Q., Regan, S.N., Xia, Y., Oostrom, L. a, Cowan, C. a & Musunuru, K., 2013. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell stem cell*, 12(4), pp.393–4.

Doherty, P.C. & Zinkernagel, R.M., 1975. A BIOLOGICAL ROLE FOR THE MAJOR HISTOCOMPATIBILITY ANTIGENS. *The Lancet*, 305(7922), pp.1406–1409.

Dossett, M.L., Teague, R.M., Schmitt, T.M., Tan, X., Cooper, L.J.N., Pinzon, C. & Greenberg, P.D., 2009. Adoptive Immunotherapy of Disseminated Leukemia With TCR-transduced, CD8+ T Cells Expressing a Known Endogenous TCR. *Mol Ther*, 17(4), pp.742–749.

Doyon, Y., Choi, V.M., Xia, D.F., Vo, T.D., Gregory, P.D. & Holmes, M.C., 2010. Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. *Nature methods*, 7(6), pp.459–60.

Doyon, Y., Vo, T.D., Mendel, M.C., Greenberg, S.G., Wang, J., Xia, D.F., Miller, J.C., Urnov, F.D., Gregory, P.D. & Holmes, M.C., 2011. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Meth*, 8(1), pp.74–79.

Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D. & Barbas, C.F., 2001. Development of Zinc Finger Domains for Recognition of the 5'-ANN-3' Family of DNA Sequences and Their Use in the Construction of Artificial Transcription Factors. *Journal of Biological Chemistry*, 276(31), pp.29466–29478.

Dreier, B., Fuller, R.P., Segal, D.J., Lund, C. V., Blancafort, P., Huber, A., Koksch, B. & Barbas, C.F., 2005. Development of zinc finger domains for recognition of the 5???-CNN-3??? family DNA sequences and their use in the construction of artificial transcription factors. *Journal of Biological Chemistry*, 280(42), pp.35588–35597.

Dudley, M.E., Wunderlich, J.R., Robbins, P.F., Yang, J.C., Hwu, P., Schwartzentruber, D.J., Topalian, S.L., Sherry, R., Restifo, N.P., Hubicki, A.M., Robinson, M.R., Raffeld, M., Duray, P., Seipp, C. a, Rogers-Freezer, L., Morton, K.E., Mavroukakis, S. a, White, D.E. & Rosenberg, S. a, 2002. Cancer regression and autoimmunity in patients after clonal repopulation with antitumor lymphocytes. *Science (New York, N.Y.)*, 298(5594), pp.850–854.

Dudley, M.E., Wunderlich, J.R., Yang, J.C., Sherry, R.M., Suzanne, L., Restifo, N.P., Royal, R.E., Kammula, U., White, D.E., Sharon, A., Rogers, L.J., Gracia, G.J., Jones, S.A., Mangiameli, D.P., Pelletier, M.M., Gea-banacloche, J., Robinson, M.R., Berman, D.M., Filie, A.C., Abati, A. & Rosenberg, S.A., 2005. Adoptive Cell Transfer Therapy Following Non-Myeloablative but Lymphodepleting Chemotherapy for the Treatment of Patients With Refractory Metastatic Melanoma. *journal of clinical oncology*, 23(10), pp.2346–2357.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D. & Naldini, L., 1998. A third-

generation lentivirus vector with a conditional packaging system. *Journal of virology*, 72(11), pp.8463–8471.

Dusseaux, M., Martin, E., Serriari, N., Pe, I., Premel, V., Louis, D., Milder, M., Bourhis, L. Le, Soudais, C., Treiner, E. & Lantz, O., 2011. Human MAIT cells are xenobiotic-resistant , tissue-targeted , CD161 hi IL-17 – secreting T cells. *Blood*, 117(4), pp.1250–1260.

Elbashir, S.M., Harborth, J., Weber, K. & Tuschl, T., 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, 26(2), pp.199–213.

Elemans, M., Seich Al Basatena, N.-K. & Asquith, B., 2012. The efficiency of the human CD8+ T cell response: how should we quantify it, what determines it, and does it matter? *PLoS computational biology*, 8(2), p.e1002381.

Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C., 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α Helices: Crystal structure of the protein-DNA complex. *Cell*, 71(7), pp.1223–1237.

Fadel, H.J., Morrison, J.H., Saenz, D.T., Fuchs, J.R., Kvaratskhelia, M., Ekker, S.C. & Poeschla, E.M., 2014. TALEN Knockout of the PSIP1 Gene in Human Cells: Analyses of HIV-1 Replication and Allosteric Integrase Inhibitor Mechanism. *Journal of Virology* , 88(17), pp.9704–9717.

Fecci, P.E., Ochiai, H., Mitchell, D.A., Grossi, P.M., Sweeney, A.E., Archer, G.E., Cummings, T., Allison, J.P., Bigner, D.D. & Sampson, J.H., 2007. Systemic CTLA-4 blockade ameliorates glioma-induced changes to the CD4 + T cell compartment without affecting regulatory T-cell function. *Clinical Cancer Research*, 13(7), pp.2158–2167.

Feng, Y.-H., Wang, L., Wang, Q., Li, X., Zeng, R. & Gorodeski, G.I., 2005. ATP stimulates GRK-3 phosphorylation and β-arrestin-2-dependent internalization of P2X7 receptor. *American Journal of Physiology - Cell Physiology*, 288(6), p.C1342 LP-C1356.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. & Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669), pp.806–811.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B. a, Shabanowitz, J., Hunt, D.F., Funabiki, H. & Allis, C.D., 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature*, 438, pp.1116–1122.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. & Joung, J.K., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature biotechnology*, 32(3), pp.279–84.

Gaspari, A. & Tyring, S.K.Gaspari, A. & Tyring, S.K., 2008. *Chapter 6: Conventional and unconventional T cells. In Clinical and Basic Immunodermatology*,

Gaszner, M. & Felsenfeld, G., 2006. Insulators: exploiting transcriptional and epigenetic mechanisms.

Nature reviews. Genetics, 7, pp.703–713.

Gerlini, G., Hefti, H.P., Kleinhans, M., Nickoloff, B.J., Burg, G. & Nestle, F.O., 2001. Cd1d is expressed on dermal dendritic cells and monocyte-derived dendritic cells. *The Journal of investigative dermatology*, 117(3), pp.576–582.

Giering, J.C., Grimm, D., Storm, T. a & Kay, M. a, 2008. Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic. *Molecular therapy : the journal of the American Society of Gene Therapy*, 16(9), pp.1630–1636.

Godfrey, D.I., Uldrich, A.P., Mccluskey, J., Rossjohn, J. & Moody, D.B., 2015. The burgeoning family of non- conventional T cells . *Nature Publishing Group*, 16(11), pp.1114–1123.

Godkin, A.J., Smith, K.J., Willis, A, Tejada-Simon, M. V, Zhang, J., Elliott, T. & Hill, A.V, 2001. Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions. *Journal of immunology (Baltimore, Md. : 1950)*, 166(11), pp.6720–6727.

Gold, M.C., Cerri, S., Smyk-Pearson, S., Cansler, M.E., Vogt, T.M., Delepine, J., Winata, E., Swarbrick, G.M., Chua, W.-J., Yu, Y.Y.L., Lantz, O., Cook, M.S., Null, M.D., Jacoby, D.B., Harriff, M.J., Lewinsohn, D. a. D.M., Hansen, T.H. & Lewinsohn, D. A. D.M., 2010. Human Mucosal Associated Invariant T Cells Detect Bacterially Infected Cells. *PLoS Biol*, 8(6), p.e1000407.

Gold, M.C., Ehlinger, H.D., Cook, M.S., Smyk-pearson, S.K., Wille, P.T., Ungerleider, R.M., Lewinsohn, D. a & Lewinsohn, D.M., 2008. Human Innate Mycobacterium tuberculosis– Reactive abTCR. *PLoS Pathogens*, 4(2).

Gold, M.C., Eid, T., Smyk-Pearson, S., Eberling, Y., Swarbrick, G.M., Langley, S.M., Streeter, P.R., Lewinsohn, D. a & Lewinsohn, D.M., 2013. Human thymic MR1-restricted MAIT cells are innate pathogen-reactive effectors that adapt following thymic egress. *Mucosal Immunology*, 6(1), pp.35–44.

Gold, M.C. & Lewinsohn, D.M., 2012. Co-dependents: MR1-restricted MAIT cells and their antimicrobial function. *Nature Reviews Microbiology*, 11(1), pp.14–19.

Gold, M.C., McLaren, J.E., Reistetter, J. A., Smyk-Pearson, S., Ladell, K., Swarbrick, G.M., Yu, Y.Y.L., Hansen, T.H., Lund, O., Nielsen, M., Gerritsen, B., Kesmir, C., Miles, J.J., Lewinsohn, D.M. A., Price, D. A. & Lewinsohn, D.M. A., 2014. MR1-restricted MAIT cells display ligand discrimination and pathogen selectivity through distinct T cell receptor usage. *Journal of Experimental Medicine*, 211(8), pp.1601–1610.

Griffiths, G.M., Tsun, A. & Stinchcombe, J.C., 2010. The immunological synapse: A focal point for endocytosis and exocytosis. *Journal of Cell Biology*, 189(3), pp.399–406.

Grimsley, C., Kawasaki, A., Gassner, C., Sageshima, N., Nose, Y., Hatake, K., Geraghty, D.E. & Ishitani,

A., 2002. Definitive high resolution typing of HLA-E allelic polymorphisms: Identifying potential errors in existing allele data. *Tissue Antigens*, 60(3), pp.206–212.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P. & Trono, D., 2010a. KRAB–Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading. *PLoS Genet*, 6(3), p.e1000869.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P. & Trono, D., 2010b. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS genetics*, 6(3), p.e1000869.

Grossman, W.J., Verbsky, J.W., Tollefsen, B.L., Kemper, C., Atkinson, J.P., Ley, T.J. & Ctls, C.T., 2004. Differential expression of granzymes A and B in human cytotoxic lymphocyte subsets and T regulatory cells. *Analysis*, 104(9), pp.2840–2848.

Guo, S. & Kemphues, K.J., 1995. par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell*, 81(4), pp.611–620.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J.I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L.E., Wissler, M., Prinz, C., Rabbitts, T.H., Le Deist, F., Fischer, A. & Cavazzana-Calvo, M., 2003. Associated Clonal T Cell Proliferation in Two Patients after Gene Therapy for SCID-X1. *Science*, 302(5644), p.415 LP-419.

Hamid, O., Robert, C., Daud, A., Hodi, F.S., Hwu, W.-J., Kefford, R., Wolchok, J.D., Hersey, P., Joseph, R.W., Weber, J.S., Dronca, R., Gangadhar, T.C., Patnaik, A., Zarour, H., Joshua, A.M., Gergich, K., Elassaiss-Schaap, J., Algazi, A., Mateus, C., Boasberg, P., Tumeh, P.C., Chmielowski, B., Ebbinghaus, S.W., Li, X.N., Kang, S.P. & Ribas, A., 2013. Safety and Tumor Responses with Lambrolizumab (Anti–PD-1) in Melanoma. *New England Journal of Medicine*, p.130602081510007.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E. & Ren, B., 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics*, 39(3), pp.311–8.

Hirano, F., Kaneko, K., Tamura, H., Dong, H., Wang, S., Ichikawa, M., Rietz, C., Flies, D.B., Lau, J.S. & Zhu, G., 2005. Blockade of B7-H1 and PD-1 by Monoclonal Antibodies Potentiates Cancer Therapeutic Immunity Blockade of B7-H1 and PD-1 by Monoclonal Antibodies. *cancer research*,

pp.1089–1096.

Ho, T.-T., Zhou, N., Huang, J., Koirala, P., Xu, M., Fung, R., Wu, F. & Mo, Y.-Y., 2014. Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. *Nucleic Acids Research* .

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E. A, Zeitler, B., Meng, X., Miller, J.C., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D. & Jaenisch, R., 2009. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature biotechnology*, 27(9), pp.851–7.

Hodi, F.S., O'Day, S.J., McDermott, D.F., Weber, R.W., Sosman, J.A., Haanen, J.B., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J.C., Akerley, W., van den Eertwegh, A.J.M., Lutzky, J., Lorigan, P., Vaubel, J.M., Linette, G.P., Hogg, D., Ottensmeier, C.H., Lebbé, C., Peschel, C., Quirt, I., Clark, J.I., Wolchok, J.D., Weber, J.S., Tian, J., Yellin, M.J., Nichol, G.M., Hoos, A. & Urba, W.J., 2010. Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *The New England journal of medicine*, 363(8), pp.711–723.

Höfer, T., Krichevsky, O. & Altan-Bonnet, G., 2012. Competition for IL-2 between regulatory and effector T cells to chisel immune responses. *Frontiers in Immunology*, 3(SEP), pp.1–9.

Hogquist, K. a & Jameson, S.C., 2014. The self-obsession of T cells: how TCR signaling thresholds affect fate "decisions" and effector function. *Nature Immunology*, 15(9), pp.815–823.

Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E. & Prydz, H., 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic acids research*, 30(8), pp.1757–1766.

Holkers, M., Maggio, I., Liu, J., Janssen, J.M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T. & Gonçalves, M. a F. V, 2012. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic acids research*, 41(5).

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C. & Cannon, P.M., 2010. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature biotechnology*, 28(8), pp.839–47.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Michael, C. & Cannon, P.M., 2010. Zinc finger nuclease-mediated CCR5 knockout hematopoietic stem cell transplantation controls HIV in vivo. *Nature biotechnology*, 28(8), pp.839–847.

Honer, C., Chen, P., Toth, M.J. & Schumacher, C., 2001. Identification of SCAN dimerization domains in four gene families. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1517(3), pp.441–448.

Hsu, P.D., Scott, D. a, Weinstein, J. a, Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L. a, Bao, G. & Zhang, F., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*, 31(9), pp.827–32.

Huang, S., Martin, E., Kim, S., Yu, L., Soudais, C., Fremont, D.H., Lantz, O. & Hansen, T.H., 2009. MR1 antigen presentation to mucosal-associated invariant T cells was highly conserved in evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 106, pp.8290–8295.

Huntley, S., Baggott, D.M., Hamilton, A.T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E. & Stubbs, L., 2006. A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome research*, 16(5), pp.669–77.

Isalan, M., Choo, Y. & Klug, A., 1997. Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 94(11), pp.5617–21.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of bacteriology*, 169(12), pp.5429–33.

Jao, L.-E., Wente, S.R. & Chen, W., 2013. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences* , 110(34), pp.13904–13909.

Jiang, Q., Zhang, H. & Zhang, P., 2011. ShRNA-mediated gene silencing of MTA1 influenced on protein expression of ER alpha, MMP-9, CyclinD1 and invasiveness, proliferation in breast cancer cell lines MDA-MB-231 and MCF-7 in vitro. *Journal of experimental & clinical cancer research : CR*, 30(1), p.60.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. & Weeks, D.P., 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research*.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. a & Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096), pp.816–21.

Joung, J.K. & Sander, J.D., 2013. NIH Public Access. *Nature Review of Molecular and Cellular Biology*, 14(1), pp.49–55.

Kabir, S., Hockemeyer, D. & de Lange, T., 2014. TALEN Gene Knockouts Reveal No Requirement for the Conserved Human Shelterin Protein Rap1 in Telomere Protection and Length Regulation.

Cell Reports, 9(4), pp.1273–1280.

Kaiser, B.K., Pizarro, J.C., Kerns, J. & Strong, R.K., 2008. Structural basis for NKG2A/CD94 recognition of HLA-E. *Proceedings of the National Academy of Sciences* , 105(18), pp.6696–6701.

Kara, E.E., Comerford, I., Fenix, K. a., Bastow, C.R., Gregor, C.E., McKenzie, D.R. & McColl, S.R., 2014. Tailored Immune Responses: Novel Effector Helper T Cell Subsets in Protective Immunity. *PLoS Pathogens*, 10(2), p.e1003905.

Karp, C.L., Wilson, C.B. & Stuart, L.M., 2015. Tuberculosis vaccines: Barriers and prospects on the quest for a transformative tool. *Immunological Reviews*, 264(1), pp.363–381.

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M. & Siksnys, V., 2015. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biology*, 16(1), pp.1–13.

Kawasaki, H. & Taira, K., 2004. Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. *Nature*, 431(7005), pp.211–217.

Kelso, A. & Groves, P., 1997. A single peripheral CD8+ T cell can give rise to progeny expressing type 1 and/or type 2 cytokine genes and can retain its multipotentiality through many cell divisions. *Proc Natl Acad Sci U S A*, 94(15), pp.8070–8075.

Kemp, R.A., Bäckström, B.T. & Ronchese, F., 2005. The phenotype of type 1 and type 2 CD8(+) T cells activated in vitro is affected by culture conditions and correlates with effector activity. *Immunology*, 115(3), pp.315–324.

Khanna, K.K. & Jackson, S.P., 2001. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature genetics*, 27(3), pp.247–54.

Kim, D.H., Villeneuve, L.M., Morris, K. V & Rossi, J.J., 2006. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol*, 13(9), pp.793–797.

Kim, H.J., Lee, H.J., Kim, H., Cho, S.W. & Kim, J.-S., 2009. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome research*, 19(7), pp.1279–88.

Kim, Y., Kweon, J., Kim, A., Chon, J.K., Yoo, J.Y., Kim, H.J.H., Kim, S.S., Lee, C., Jeong, E., Chung, E., Kim, D., Lee, M.S., Go, E.M., Song, H.J., Cho, N., Bang, D. & Kim, J.-S., 2013. A library of TAL effector nucleases spanning the human genome. *Nature biotechnology*, 31(October 2012), pp.1–9.

Kim, Y.G., Cha, J. & Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93(3), pp.1156–1160.

Kjer-Nielsen, L., Patel, O., Corbett, A.J., Le Nours, J., Meehan, B., Liu, L., Bhati, M., Chen, Z., Kostenko, L., Reantragoon, R., Williamson, N. a., Purcell, A.W., Dudek, N.L., McConville, M.J.,

O'Hair, R. a. J., Khairallah, G.N., Godfrey, D.I., Fairlie, D.P., Rossjohn, J. & McCluskey, J., 2012. MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature*. 491, 717-723

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z. & Keith Joung, J., 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, 529(7587), pp.490–495.

Koller, B.H., Geraghty, D.E., Shimizu, Y., DeMars, R. & Orr, H.T., 1988. HLA-E. A novel HLA class I gene expressed in resting T lymphocytes. *The Journal of Immunology* , 141(3), pp.897–904.

Kondo, Y., Shen, L., Cheng, A.S., Ahmed, S., Boumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B., Gold, D.L., Sekido, Y., Huang, T.H.-M. & Issa, J.-P.J., 2008. Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nature genetics*, 40(6), pp.741–750.

Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, 128(4), pp.693–705.

Kuball, J., Dossett, M.L., Wolfl, M., Ho, W.Y., Voss, R.H., Fowler, C. & Greenberg, P.D., 2007. Facilitating matched pairing and expression of TCR chains introduced into human T cells. *Blood*, 109(6), pp.2331–2338.

Kuhn, A.N., Beißert, T., Simon, P., Vallazza, B., Buck, J., Davies, B.P., Türeci, Ö. & Sahin, U., 2012. mRNA as a versatile tool for exogenous protein expression. *Current Gene Therapy*, 12(5), pp.347–361.

Kurioka, A., Ussher, J.E., Cosgrove, C., Clough, C., Fergusson, J.R., Smith, K., Kang, Y.-H., Walker, L.J., Hansen, T.H., Willberg, C.B. & Klenerman, P., 2015. MAIT cells are licensed through granzyme exchange to kill bacterially sensitized targets. *Mucosal immunology*, 8(2), pp.429–40.

Landau, S.B., Trier, J.S., Balk, S.P., Blumberg, R.S., Terhorst, C.O.X., Bleicher, P., Mcdermott, F. V, Allan, C.H., Landau, S.B., Trier, J.S. & Balkz, S.P., 1991. Expression of a nonpolymorphic MHC class I-like molecule , CD1D , by human intestinal epithelial cell. *Journal of Immunology*, 147, pp.2518–2524.

Lang, G.A., Devera, T.S. & Lang, M.L., 2008. Requirement for CD1d expression by B cells to stimulate NKT cell–enhanced antibody production. *Blood*, 111(4), pp.2158–2162.

Laugel, B., van den Berg, H. a, Gostick, E., Cole, D.K., Wooldridge, L., Boulter, J., Milicic, A., Price, D. a, Sewell, A.K. & Berg, H.A. Van Den, 2007. Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties. *The Journal of biological chemistry*, 282(33), pp.23799–810.

Laugel, B., Lloyd, A., Meermeier, E.W., Crowther, M.D., Connor, T.R., Dolton, G., Miles, J.J., Burrows, S.R., Gold, M.C., Lewinsohn, D.M. & Sewell, A.K., 2016. Engineering of Isogenic Cells Deficient for MR1 with a CRISPR/Cas9 Lentiviral System: Tools To Study Microbial Antigen Processing and

Presentation to Human MR1-Restricted T Cells. *The Journal of Immunology*.

Lawhorn, I.E.B., Ferreira, J.P. & Wang, C.L., 2014. Evaluation of sgRNA target sites for CRISPR-mediated repression of TP53. *PLoS ONE*, 9(11).

Laydon, D.J., Bangham, C.R.M. & Asquith, B., 2015. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. *Phil Trans R Soc B*, 370(1675), p.20140291-.

Lee, H.J., Kim, E. & Kim, J., 2010. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome research*, pp.81–89.

Lee, H.J., Kweon, J., Kim, E., Kim, S. & Kim, J.S., 2012. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Research*, 22(3), pp.539–548.

Lemberg, M.K., Bland, F.A., Weihofen, A., Braud, V.M. & Martoglio, B., 2001. Intramembrane Proteolysis of Signal Peptides: An Essential Step in the Generation of HLA-E Epitopes. *The Journal of Immunology* , 167(11), pp.6441–6446.

Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.

Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., Murphy, S.L., Finn, J.D., Khazi, F.R., Zhou, S., Paschon, D.E., Rebar, E.J., Bushman, F.D., Gregory, P.D., Holmes, M.C. & High, K. A, 2011. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature*, 475(7355), pp.217–21.

Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., Yamamoto, T., Yamanaka, S. & Hotta, A., 2015. Precise Correction of the Dystrophin Gene in Duchenne Muscular Dystrophy Patient Induced Pluripotent Stem Cells by TALEN and CRISPR-Cas9. *Stem Cell Reports*, 4(1), pp.143–154.

Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P. & Yang, B., 2011. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic acids research*, 39(14), pp.6315–25.

Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., Lv, J., Xie, X., Chen, Y., Li, Y., Sun, Y., Bai, Y., Songyang, Z., Ma, W., Zhou, C. & Huang, J., 2015. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein and Cell*, 6(5), pp.363–372.

Litchfield, L.M. & Klinge, C.M., 2012. Multiple roles of COUP-TFII in cancer initiation and progression. *Journal of molecular endocrinology*, 49(3), pp.R135–R148.

Liu, Q., Segal, D.J., Ghiara, J.B. & Barbas, C.F., 1997. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proceedings of the National Academy of Sciences* , 94(11), pp.5525–5530.

Liu, R., Chen, L., Jiang, Y., Zhou, Z. & Zou, G., 2015. Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. *Cell Discovery*, 1, p.15007.

Liuzzi, A.R., McLaren, J.E., Price, D. a & Eberl, M., 2015. Early innate responses to pathogens: pattern recognition by unconventional human T-cells. *Current Opinion in Immunology*, 36, pp.31–37.

Lohmueller, J.J., Armel, T.Z. & Silver, P. A, 2012. A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic acids research*, 40(11), pp.5180–7.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.-L., Kim, K. a, Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., Holmes, M.C. & Naldini, L., 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature biotechnology*, 25(11), pp.1298–306.

Love, M.I., Anders, S. & Huber, W., 2014. *Differential analysis of count data - the DESeq2 package*,

Lu, J., Zavorotinskaya, T., Dai, Y., Niu, X.-H., Castillo, J., Sim, J., Yu, J., Wang, Y., Langowski, J.L., Holash, J., Shannon, K. & Garcia, P.D., 2013. Pim2 is required for maintaining multiple myeloma cell growth through modulating TSC2 phosphorylation. *Blood*, 122(9), pp.1610–1620.

Luo, D. & Saltzman, W.M., 2000. Synthetic DNA delivery systems. *Nat Biotech*, 18(1), pp.33–37.

Luttrell, L.M. & Lefkowitz, R.J., 2002. The role of β-arrestins in the termination and transduction of G-protein-coupled receptor signals. *Journal of Cell Science*, 115(3), p.455 LP-465.

Ma, N., Liao, B., Zhang, H., Wang, L., Shan, Y., Xue, Y., Huang, K., Chen, S., Zhou, X., Chen, Y., Pei, D. & Pan, G., 2013. Transcription activator-like effector nuclease (TALEN)-mediated Gene correction in integration-free β-Thalassemia induced pluripotent stem cells. *Journal of Biological Chemistry*, 288(48), pp.34671–34679.

MacDonald, H.R., Glasebrook, A.L. & Cerottini, J.C., 1982. Clonal heterogeneity in the functional requirement for Lyt-2/3 molecules on cytolytic T lymphocytes: analysis by antibody blocking and selective trypsinization. *The Journal of experimental medicine*, 156(6), pp.1711–1722.

Mackensen, A., Meidenbauer, N., Vogl, S., Laumer, M., Berger, J. & Andreesen, R., 2006. Phase I study of adoptive T-cell therapy using antigen-specific CD8 + T cells for the treatment of patients with metastatic melanoma. *Journal of Clinical Oncology*, 24(31), pp.5060–5069.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D. a, Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J. a, Unger-Wallace, E., Sander, J.D., Müller-Lerch, F., Fu, F., Pearlberg, J., Göbel, C., Dassie, J.P., Pruett-Miller, S.M., Porteus, M.H., Sgroi, D.C., Iafrate, a J., Dobbs, D., McCray, P.B., Cathomen, T., Voytas, D.F. & Joung, J.K., 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Molecular cell*, 31(2), pp.294–301.

Maeder, M.L., Thibodeau-Beganny, S., Sander, J.D., Voytas, D.F. & Joung, J.K., 2009. Oligomerized

pool engineering (OPEN): an “open-source” protocol for making customized zinc-finger arrays. *Nat. Protocols*, 4(10), pp.1471–1501.

Magnusson, C. & Vaux, D.L., 1999. Signalling by CD95 and TNF receptors: Not only life and death. *Immunol Cell Biol*, 77(1), pp.41–46.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. & Church, G.M., 2013. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339(6121), pp.823–6.

Mandell, J.G. & Barbas, C.F., 2006. Zinc Finger Tools: Custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Research*, 34, pp.516–523.

Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V., 2008. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair*, 7(10), pp.1765–1771.

Marshall, N.B. & Swain, S. l, 2011. Cytotoxic CD4 T cells in antiviral immunity. *Journal of biom*.

Martin, E., Treiner, E., Duban, L., Guerri, L., Laude, H., Toly, C., Premel, V., Devys, A., Moura, I.C., Tilloy, F., Cherif, S., Vera, G., Latour, S., Soudais, C. & Lantz, O., 2009. Stepwise development of MAIT cells in mouse and human. *PLoS biology*, 7(3), p.e54.

Maston, G. A. et al., 2006. Transcriptional Regulatory Elements in the Human Genome. *Annual Review of Genomics and Human Genetics*, 7(1), pp.29–59.

Meermeier, E.W., Laugel, B.F., Sewell, A.K., Corbett, A.J., Rossjohn, J., McCluskey, J., Harriff, M.J., Franks, T., Gold, M.C. & Lewinsohn, D.M., 2016. Human TRAV1-2-negative MR1-restricted T cells detect S. pyogenes and alternatives to MAIT riboflavin-based antigens. *Nature Communications*, 7, p.12506.

Meng, X., Noyes, M.B., Zhu, L.J., Lawson, N.D. & Wolfe, S. A, 2008. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotech*, 26(6), pp.695–701.

Messaoudi, I., Guevara Patino, J.A., Dyall, R., LeMaoult, J. & Nikolich-Zugich, J., 2002. Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense. *Science (New York, N.Y.)*, 298(5599), pp.1797–1800.

Messina, D.N., Glasscock, J., Gish, W. & Lovett, M., 2004. An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome research*, 14(10B), pp.2041–7.

Mikkola, H., Woods, N., Sjögren, M., Helgadottir, H., Hamaguchi, I. & Jacobsen, S., 2000. Lentivirus Gene Transfer in Murine Hematopoietic Progenitor Cells Is Compromised by a Delay in Proviral Integration and Results in Transduction Mosaicism and Heterogeneous Gene Expression in Progeny Cells Lentivirus Gene Transfer in Murine Hematopoietic P. *journal of virology*, 74(24), pp.11911–11918.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.-L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K. a, Gregory, P.D., Pabo, C.O. & Rebar, E.J., 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature biotechnology*, 25(7), pp.778–85.

Milne, P., Bigley, V., Gunawan, M., Haniffa, M. & Collin, M., 2015. CD1c(+) blood dendritic cells have Langerhans cell potential. *Blood*, 125(3), pp.470–473.

Mojica, F.J.M., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E., 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution*, 60(2), pp.174–182.

Mojica, F.J.M., Díez-Villaseñor, C., Soria, E. & Juez, G., 2000. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, 36(1), pp.244–246.

Mojica, F.J.M., Juez, G. & Rodriguez-Valera, F., 1993. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Molecular Microbiology*, 9(3), pp.613–621.

Moore, C.B., Guthrie, E.H., Huang, M.T.H. & Taxman, D.J., 2010. Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *Methods in molecular biology (Clifton, N.J.)*, 629(2), pp.141–158.

Morgan, R.A., Chinnasamy, N., Abate-daga, D.D., Gros, A., Robbins, F.P., Zheng, Z., Feldman, S.A., Yang, J.C., Sherry, R.M., Phan, Q.G., Hughes, M.S., Kammula, U.S., Miller, A.D., Hessman, C.J., Stewart, A.A., Restifo, N.P., Quezado, M.M., Alimchandani, M., Rosenberg, Z.A., Nath, A., Wang, T., Bielekova, B., Wuest, S.C., Nirmala, A., Mcmahon, F.J., Wilde, S., Mosetter, B., Schendel, D.J., Laurencot, C.M. & Rosenberg, S.A., 2014. Cancer regression and neurologic toxicity following anti-MAGE- A3 TCR gene therapy. *J Immunother*, 36(2), pp.133–151.

Morgan, R.A., Dudley, M.E., Wunderlich, J.R., Hughes, M.S., Yang, J.C., Sherry, R.M., Royal, R.E., Topalian, S.L., Kammula, U.S., Restifo, N.P., Zheng, Z., Nahvi, A., Vries, C.R. De, Rogers-Freezer, L.J., Mavroukakis, A.S. & Rosenberg, S. A, 2006. Cancer Regression in Patients After Transfer of Genetically Engineered Lymphocytes. *Science*, 314(5796), pp.126–129.

Morris, E.C., Tsallios, A., Bendle, G.M., Xue, S. & Stauss, H.J., 2005. A critical role of T cell antigen receptor-transduced MHC class I-restricted helper T cells in tumor protection. *Proceedings of the National Academy of Sciences* , 102(22), pp.7934–7939.

Murphy, K.M.Murphy, K.M., 2012. *Janeway's Immunobiology*.

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. & Cathomen, T., 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity.

Nucleic acids research, 39(21), pp.9283–93.

Napoli, C., Lemieux, C. & Jorgensen, R., 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant cell*, 2(4), pp.279–289.

Narni-Mancinelli, E. & Vivier, E., 2014. Delivering Three Punches to Knockout Intracellular Bacteria. *Cell*, 157(6), pp.1251–1252.

Nattermann, J., Nischalke, H.D., Hofmeister, V., Kupfer, B., Ahlenstiel, G., Feldmann, G., Rockstroh, J., Weiss, E.H., Sauerbruch, T. & Spengler, U., 2005. HIV-1 infection leads to increased HLA-E expression resulting in impaired function of natural killer cells. *Antiviral Therapy*, 10(1), pp.95–107.

Nature Methods, 2011. Method of the Year 2011. *Nature Methods*, 9(1), p.1.

Neefjes, J., Jongsma, M.L.M., Paul, P. & Bakke, O., 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol*, 11(12), pp.823–836.

Nikolich-Žugich, J., Slifka, M.K. & Messaoudi, I., 2004. The many important facets of T-cell repertoire diversity. *Nature Reviews Immunology*, 4(2), pp.123–132.

Olivier, M., Foret, B., Le Vern, Y., Kerboeuf, D. & Guilloteau, L.A., 2013. Plasticity of Migrating $CD1b^{+}$ and $CD1b^{-}$ Lymph Dendritic Cells in the Promotion of Th1, Th2 and Th17 in Response to Salmonella and Helminth Secretions. *PLoS ONE*, 8(11), p.e79537.

Olson, a, Sheth, N., Lee, J.S., Hannon, G. & Sachidanandam, R., 2006. RNAi Codex: a portal/database for short-hairpin RNA (shRNA) gene-silencing constructs. *Nucleic acids research*, 34, pp.D153–D157.

Oo, Y.H., Banz, V., Kavanagh, D., Liaskou, E., Withers, D.R., Humphreys, E., Reynolds, G.M., Lee-Turner, L., Kalia, N., Hubscher, S.G., Klenerman, P., Eksteen, B. & Adams, D.H., 2016. CXCR3-dependent recruitment and CCR6-mediated positioning of Th-17 cells in the inflamed liver. *Journal of Hepatology*, 57(5), pp.1044–1051.

van der Oost, J., 2013. New tool for genome surgery. *Science (New York, N.Y.)*, 339(6121), pp.768–70.

Osborn, M.J., Gabriel, R., Webber, B.R., DeFeo, A.P., McElroy, A.N., Jarjour, J., Starker, C.G., Wagner, J.E., Joung, J.K., Voytas, D.F., von Kalle, C., Schmidt, M., Blazar, B.R. & Tolar, J., 2015. Fanconi Anemia Gene Editing by the CRISPR/Cas9 System. *Human Gene Therapy*, 26(2), pp.114–126.

Osborn, M.J., Webber, B.R., Knipping, F., Lonetree, C., Tennis, N., DeFeo, A.P., McElroy, A.N., Starker, C.G., Lee, C., Gabriel, R., Merkel, S., Lund, T.C., Kelly-Spratt, K.S., Jensen, M.C., Voytas, D.F., von Kalle, C., Schmidt, M., Hippen, K.L., Miller, J.S., Scharenberg, A.M., Tolar, J. & Blazar, B.R., 2015. Evaluation of TCR Gene Editing achieved by TALENs, CRISPR/Cas9 and megaTAL nucleases.

Molecular therapy : the journal of the American Society of Gene Therapy, 24(October), pp.570–581.

Ousterout, D.G., Kabadi, A.M., Thakore, P.I., Majoros, W.H., Reddy, T.E. & Gersbach, C. A., 2015. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nature communications*, 6, p.6244.

Pan, J.G., Zhou, X., Luo, R. & Han, R.F., 2012. The adeno-associated virus-mediated HSV-TK/GCV suicide system: A potential strategy for the treatment of bladder carcinoma. *Medical Oncology*, 29(3), pp.1938–1947.

Pang, Y.-P., Kumar, G. a, Zhang, J.-S. & Urrutia, R., 2003. Differential binding of Sin3 interacting repressor domains to the PAH2 domain of Sin3A. *FEBS Letters*, 548(1–3), pp.108–112.

Panum PL. Observations Made During The Epidemic Of Measles On The Faroe Islands In The Year 1846. Copenhagen: Bibiliothek for Laeger; 1847.

Park, J.-H., Adoro, S., Lucas, P.J., Sarafova, S.D., Alag, A.S., Doan, L.L., Erman, B., Liu, X., Ellmeier, W., Bosselut, R., Feigenbaum, L. & Singer, A., 2007. "Coreceptor tuning": cytokine signals transcriptionally tailor CD8 coreceptor expression to the self-specificity of the TCR. *Nat Immunol*, 8(10), pp.1049–1059.

Parkhurst, M.R., Yang, J.C., Langan, R.C., Dudley, M.E., Nathan, D.-A.N., Feldman, S.A., Davis, J.L., Morgan, R.A., Merino, M.J., Sherry, R.M., Hughes, M.S., Kammula, U.S., Phan, G.Q., Lim, R.M., Wank, S.A., Restifo, N.P., Robbins, P.F., Laurencot, C.M. & Rosenberg, S.A., 2011. T Cells Targeting Carcinoembryonic Antigen Can Mediate Regression of Metastatic Colorectal Cancer but Induce Severe Transient Colitis. *Mol Ther*, 19(3), pp.620–626.

Pavletich, N.P. & Pabo, C.O., 1991. Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 A. *Science*, 252(May), pp.809–817.

Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K. a, Liu, O., Wang, N., Lee, G., Bartsevich, V. V, Lee, Y.-L., Guschin, D.Y., Rupniewski, I., Waite, A.J., Carpenito, C., Carroll, R.G., Orange, J.S., Urnov, F.D., Rebar, E.J., Ando, D., Gregory, P.D., Riley, J.L., Holmes, M.C. & June, C.H., 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature biotechnology*, 26(7), pp.808–16.

Pietra, G., Romagnani, C., Manzini, C., lorenzo moretta & Mingari, M.C., 2010. The emerging role of HLA-E-restricted CD8+ T lymphocytes in the adaptive immune response to pathogens and tumors. *Journal of Biomedicine and Biotechnology*, 2010.

Pipkin, M.E. & Lieberman, J., 2007. Delivering the kiss of death: progress on understanding how perforin works. *Current opinion in immunology*, 19(3), pp.301–8.

Porcelli, S., Brenner, M.B., Greenstein, J.L., Balk, S.P., Terhorst, C. & Bleicher, P. A, 1989. Recognition

of cluster of differentiation 1 antigens by human CD4-CD8-cytolytic T lymphocytes. *Nature*, 341(6241), pp.447–450.

Powrie, F., Carlino, J., Leach, M.W., Mauze, S. & Coffman, R.L., 1996. A critical role for transforming growth factor-beta but not interleukin 4 in the suppression of T helper type 1-mediated colitis by CD45RB(low) CD4+ T cells. *The Journal of experimental medicine*, 183(6), pp.2669–74.

Provasi, E., Genovese, P., Lombardo, A., Magnani, Z., Liu, P.-Q., Reik, A., Chu, V., Paschon, D.E., Zhang, L., Kuball, J., Camisa, B., Bondanza, A., Casorati, G., Ponzoni, M., Ciceri, F., Bordignon, C., Greenberg, P.D., Holmes, M.C., Gregory, P.D., Naldini, L. & Bonini, C., 2012. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nature medicine*, 18(5), pp.807–15.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. & Lim, W.A., 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, 152(5), pp.1173–1183.

Reith, W., Leibund Gut-Landmann, S. & Waldburger, J.-M., 2005. Regulation of MHC class II gene expression by the class II transactivator. *Nature reviews. Immunology*, 5, pp.793–806.

Ren, D., Collingwood, T.N., Rebar, E.J., Wolffe, A.P. & Camp, H.S., 2002. PPAR gamma knockdown by engineered transcription factors : exogenous PPAR gamma 2 but not PPAR gamma 1 reactivates adipogenesis. *Genes & Development*, 16(734), pp.27–32.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S. & Khvorova, A., 2004. Rational siRNA design for RNA interference. *Nature biotechnology*, 22(3), pp.326–30.

Reyrat, J.-M. & Kahn, D., 2001. Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends in Microbiology*, 9(10), pp.472–473.

Ribas, A., Camacho, L.H., Lopez-Berestein, G., Pavlov, D., Bulanhagui, C.A., Millham, R., Comin-Anduix, B., Reuben, J.M., Seja, E., Parker, C.A., Sharma, A., Glaspy, J.A. & Gomez-Navarro, J., 2005. Antitumor Activity in Melanoma and Anti-Self Responses in a Phase I Trial With the Anti-Cytotoxic T Lymphocyte–Associated Antigen 4 Monoclonal Antibody CP-675,206. *Journal of Clinical Oncology* , 23(35), pp.8968–8977..

Ribot, J.C., deBarros, A., Pang, D.J., Neves, J.F., Peperzak, V., Roberts, S.J., Girardi, M., Borst, J., Hayday, A.C., Pennington, D.J. & Silva-Santos, B., 2009. CD27 is a thymic determinant of the balance between interferon-[gamma]- and interleukin 17-producing [gamma][delta] T cell subsets. *Nat Immunol*, 10(4), pp.427–436.

Riddell, S., Elliott, M., Lewinsohn, D., Gilbert, M., Wilson, L., Manley, S., Lupton, S., Overell, R., Reynolds, T., Corey, L. & Greenberg, P., 1996. T-cell mediated rejection of gene-modified HIV-specific cytotoxic T lymphocytes in HIV-infected patients. *Nat Med*, 2(2), p.216–23.

Rist, M.J., Theodossis, A., Croft, N.P., Neller, M. a., Welland, A., Chen, Z., Sullivan, L.C., Burrows, J.M., Miles, J.J., Brennan, R.M., Gras, S., Khanna, R., Brooks, A.G., McCluskey, J., Purcell, A.W., Rossjohn, J. & Burrows, S.R., 2013. HLA peptide length preferences control CD8+ T cell responses. *Journal of immunology (Baltimore, Md. : 1950)*, 191(2), pp.561–71.

Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L., 2011. Identification of novel transcripts in annotated genome using RNA-Seq. *Bioinformatics*, 27(17), pp.2325–2329.

Roche, P.A. & Furuta, K., 2015. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol*, 15(4), pp.203–216.

Romano, N. & Macino, G., 1992. Quelling: Transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. *Molecular Microbiology*, 6(22), pp.3343–3353.

Rosat, J.P., Grant, E.P., Beckman, E.M., Dascher, C.C., Sieling, P. a, Frederique, D., Modlin, R.L., Porcelli, S. a, Furlong, S.T. & Brenner, M.B., 1999. CD1-restricted microbial lipid antigen-specific recognition found in the CD8+ alpha beta T cell pool. *Journal of immunology (Baltimore, Md. : 1950)*, 162(1), pp.366–371.

Rosenberg, S.A., Packard, B.S., Aebersold, P.M., Solomon, D., Topalian, S.L., Toy, S.T., Simon, P., Lotze, M.T., Yang, J.C., Seipp, C.A., Simpson, C., Carter, C., Bock, S., Schwartzentruber, D., Wei, J.P. & White, D.E., 1988. Use of Tumor-Infiltrating Lymphocytes and Interleukin-2 in the Immunotherapy of Patients with Metastatic Melanoma. *New England Journal of Medicine*, 319(25), pp.1676–1680.

Rosenberg, S.A., Spiess, P. & Lafreniere, R., 1986. A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. *Science*, 233(4770), pp.1318–1321.

Rosenberg, S.A., 2010. Of Mice , Not Men : No Evidence for Graft-versus-Host Disease in Humans Receiving T-Cell Receptor – Transduced Autologous T Cells Prophylactic Antiparasitic Transgenesis. *Molecular Therapy*, 18(10), pp.1744–1745.

Rupp, L., Schumann, K., Roybal, K.T., Marson, A. & Lim, W.A., 2016. CRISPR/Cas9-mediated PD-1 disruption enhances anti-tumor efficacy of human chimeric antigen receptor T cells. *The Journal of Immunology* , 196(1 Supplement), p.214.24-214.24.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. a & Barrell, B., 2000. Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)*, 16(10), pp.944–945.

de Saint Basile, G., Ménasché, G. & Fischer, A., 2010. Molecular mechanisms of biogenesis and exocytosis of cytotoxic granules. *Nature reviews. Immunology*, 10(8), pp.568–79.

Sakaguchi, S., Miyara, M., Costantino, C.M. & Hafler, D. A., 2010. FOXP3+ regulatory T cells in the

human immune system. *Nature Reviews Immunology*, 10(7), pp.490–500.

Salerno-Gonçalves, R., Fernandez-Viña, M., Lewinsohn, D.M. & Sztein, M.B., 2004. Identification of a Human HLA-E-Restricted CD8+ T Cell Subset in Volunteers Immunized with Salmonella enterica Serovar Typhi Strain Ty21a Typhoid Vaccine. *The Journal of Immunology* , 173(9), pp.5852–5862.

Sandberg, J., Dias, J., Barbara l shacklett & Leeansyah, E., 2013. will loss of your MAITs weaken your HAART? *AIDS*, 27(16), pp.2901–2544.

Sander, T.L., Stringer, K.F., Maki, J.L., Szauter, P., Stone, J.R. & Collins, T., 2003. The SCAN domain defines a large family of zinc finger transcription factors. *Gene*, 310, pp.29–38.

Sanders, L.H., Laganière, J., Cooper, O., Mak, S.K., Vu, B.J., Huang, Y.A., Paschon, D.E., Vangipuram, M., Sundararajan, R., Urnov, F.D., Langston, J.W., Gregory, P.D., Zhang, H.S., Greenamyre, J.T., Isacson, O. & Schüle, B., 2014. LRRK2 mutations cause mitochondrial DNA damage in iPSC-derived neural cells from Parkinson's disease patients: Reversal by gene correction. *Neurobiology of Disease*, 62, pp.381–386.

Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G. & Zhang, F., 2012. A transcription activator-like effector toolbox for genome engineering. *Nature protocols*, 7(1), pp.171–92.

Sanjana, N.E., Shalem, O. & Zhang, F., 2014. Improved vectors and genome-wide libraries for CRISPR screening. *nature methods*, 11(8), pp.783–784.

Scherr, M., Chaturvedi, A., Battmer, K., Dallmann, I., Schultheis, B., Ganser, A. & Eder, M., 2006. Enhanced sensitivity to inhibition of SHP2, STAT5, and Gab2 expression in chronic myeloid leukemia (CML). *Blood*, 107(8), pp.3279–3287.

Schmid-Burgk, J.L., Schmidt, T., Kaiser, V., Höning, K. & Hornung, V., 2013. A ligation-independent cloning technique for high-throughput assembly of transcription activator–like effector genes. *Nature biotechnology*, 31(1), pp.76–81.

Schomber, T., Kalberer, C.P., Wodnar-filipowicz, A. & Skoda, R.C., 2004. Brief report Gene silencing by lentivirus-mediated delivery of siRNA in human CD34 ƨ cells. *Gene*, 103(12), pp.4511–4513.

Schwank, G., Koo, B.-K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., Nieuwenhuis, E.E.S., Beekman, J.M. & Clevers, H., 2013. Functional Repair of CFTR by CRISPR/Cas9 in Intestinal Stem Cell Organoids of Cystic Fibrosis Patients. *Cell Stem Cell*, 13(6), pp.653–658.

Sebastiano, V., Maeder, M.L., Angstman, J.F., Haddad, B., Khayter, C., Yeo, D.T., Goodwin, M.J., Hawkins, J.S., Ramirez, C.L., Batista, L.F.Z., Artandi, S.E., Wernig, M. & Joung, J.K., 2011. In Situ Genetic Correction of the Sickle Cell Anemia Mutation in Human Induced Pluripotent Stem Cells Using Engineered Zinc Finger Nucleases. *STEM CELLS*, 29(11), pp.1717–1726.

Sebestyen, Z., Schooten, E., Sals, T., Zaldivar, I., San Jose, E., Alarcon, B., Bobisse, S., Rosato, A., Szollosi, J., Gratama, J., Willemsen, R. & Debets, R., 2008. Human TCR that incorporate CD3zeta induce highly preferred pairing between TCRalpha and beta chains following gene transfer. *Journal of immunology (Baltimore, Md : 1950)*, 180(11), pp.7736–7746.

Segal, D.J., Dreier, B., Beerli, R.R. & Barbas, C.F., 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), pp.2758–2763.

Sen, G.L. & Blau, H.M., 2006. A brief history of RNAi : the silence of the genes ELUCIDATION OF THE SILENCING TRIGGER. *FASEB*, 20, pp.1293–1299.

Serriari, N.-E., Eoche, M., Lamotte, L., Lion, J., Fumery, M., Marcelo, P., Chatelain, D., Barre, A., Nguyen-Khac, E., Lantz, O., Dupas, J.-L. & Treiner, E., 2014. Innate mucosal-associated invariant T (MAIT) cells are activated in inflammatory bowel diseases. *Clinical & Experimental Immunology*, 176(2), pp.266–274.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F. & Zhan, F., 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science (New York, N.Y.)*, 343(6166), pp.84–7.

Sherkow, J.S., 2016. Pursuit of Profit Poisons Collaboration. *Nature*, 532(April 2014), p.172.

Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J.K., 2009. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of human genetics*, 54(1), pp.15–39.

Shin, J.-S., Ebersold, M., Pypaert, M., Delamarre, L., Hartley, A. & Mellman, I., 2006. Surface expression of MHC class II in dendritic cells is controlled by regulated ubiquitination. *Nature*, 444(7115), pp.115–118.

Sieling, P. a, Ochoa, M.T., Jullien, D., Leslie, D.S., Sabet, S., Rosat, J.P., Burdick, a E., Rea, T.H., Brenner, M.B., Porcelli, S. a & Modlin, R.L., 2000. Evidence for human CD4+ T cells in the CD1-restricted repertoire: derivation of mycobacteria-reactive T cells from leprosy lesions. *Journal of immunology (Baltimore, Md. : 1950)*, 164(9), pp.4790–6.

Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R., Sheth, N., Bradshaw, J., Burchard, J., Kulkarni, A., Cavet, G., Sachidanandam, R., McCombie, W.R., Cleary, M. a, Elledge, S.J. & Hannon, G.J., 2005. Second-generation shRNA libraries covering the mouse and human genomes. *Nature genetics*, 37(11), pp.1281–8.

Sinn, P.L., Sauter, S.L. & McCray, P.B., 2005. Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. *Gene therapy*, 12(14), pp.1089–98.

Siolas, D., Lerner, C., Burchard, J., Ge, W., Linsley, P.S., Paddison, P.J., Hannon, G.J. & Cleary, M. A, 2005. Synthetic shRNAs as potent RNAi triggers. *Nature biotechnology*, 23(2), pp.227–231.

Smith, C; Gore, A; Yan, W; Abalde-Atristain, L; Li, Z; He, C; Wang, Y; Brodsky, R; Zhang, K; Cheng, L; Ye, Z., 2014. Whole-Genome Sequencing Analysis Reveals High Specificity of CRISPR/Cas9 and TALEN-Based Genome Editing in Human iPSCs. *Stem cell Cell*, 15(1), pp.520–529.

Soldner, F., Laganière, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H., Lindquist, S., Zhang, L., Guschin, D., Fong, L.K., Vu, B.J., Meng, X., Urnov, F.D., Rebar, E.J., Gregory, P.D., Zhang, H.S. & Jaenisch, R., 2011. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, 146(2), pp.318–31.

Solin, S.L., Shive, H.R., Woolard, K.D., Essner, J.J. & McGrail, M., 2015. Rapid tumor induction in zebrafish by TALEN-mediated somatic inactivation of the retinoblastoma1 tumor suppressor rb1. *Scientific Reports*, 5(April), p.13745.

Söllü, C., Pars, K., Cornu, T.I., Thibodeau-Beganny, S., Maeder, M.L., Joung, J.K., Heilbronn, R., Cathomen, T., Öllü, C., Pars, K., Cornu, T.I., Thibodeau-Beganny, S., Maeder, M.L., Joung, J.K., Heilbronn, R. & Cathomen, T., 2010. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic acids research*, 38(22), pp.8269–76.

Sommermeyer, D. & Uckert, W., 2010. Minimal amino acid exchange in human TCR constant regions fosters improved function of TCR gene-modified T cells. *Journal of immunology (Baltimore, Md. : 1950)*, 184(11), pp.6223–31.

Sripathy, S.P., Stevens, J. & Schultz, D.C., 2006. The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Molecular and cellular biology*, 26(22), pp.8623–38.

Di Stasi, A., Tey, S.-K., Dotti, G., Fujita, Y., Kennedy-Nasser, A., Martinez, C., Straathof, K., Liu, E., Durett, A.G., Grilley, B., Liu, H., Cruz, C.R., Savoldo, B., Gee, A.P., Schindler, J., Krance, R.A., Heslop, H.E., Spencer, D.M., Rooney, C.M., Brenner, M.K. & S-k, T., 2011. Inducible apoptosis as a safety switch for adoptive cell therapy. *The New England journal of medicine*, 365(18), pp.1673–83.

Strauss, L., Bergmann, C. & Whiteside, T.L., 2009. Human Circulating CD4+ CD25high FOXP3+ regulatory T cells kill autologous CD8+ but not CD4+ responder cells by Fas medated apoptosis. *Journal of Immunology*.

Stroud, D.A., Formosa, L.E., Wijeyeratne, X.W., Nguyen, T.N. & Ryan, M.T., 2013. Gene Knockout Using Transcription Activator-like Effector Nucleases (TALENs) Reveals That Human NDUFA9

Protein Is Essential for Stabilizing the Junction between Membrane and Matrix Arms of Complex I. *Journal of Biological Chemistry* , 288(3), pp.1685–1690.

Sugita, M., van der Wel, N., Rogers, R.A., Peters, P.J. & Brenner, M.B., 2000. CD1c molecules broadly survey the endocytic system. *Proceedings of the National Academy of Sciences of the United States of America*, 97(15), pp.8445–8450.

Sung, Y.H., Baek, I.-J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.-S. & Lee, H.-W., 2013. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotech*, 31(1), pp.23–24.

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D.J. & Cathomen, T., 2007. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature biotechnology*, 25(7), pp.786–93.

Tai, X., Van Laethem, F., Pobezinsky, L., Guinter, T., Sharrow, S.O., Adams, A., Granger, L., Kruhlak, M., Lindsten, T., Thompson, C.B., Feigenbaum, L. & Singer, A., 2012. Basis of CTLA-4 function in regulatory and conventional CD4(+) T cells. *Blood*, 119(22), pp.5155–5163.

Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G., Holmes, M.C., Gregory, P.D., Ando, D.G., Kalos, M., Collman, R.G., Binder-Scholl, G., Plesa, G., Hwang, W.-T., Levine, B.L. & June, C.H., 2014. Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *New England Journal of Medicine*, 370(10), pp.901–910.

Thakore, P.I., M, D.A., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E. & Gersbach, C.A., 2015. Highly specific epigenome editing by {CRISPR-Cas9} repressors for silencing of distal regulatory elements. *Nat Methods*, 12(12), pp.1143–1149.

Thiel, G., Lietz, M. & Hohl, M., 2004. How mammalian transcriptional repressors work. *European Journal of Biochemistry*, 271(14), pp.2855–2862.

Tilloy, F., Treiner, E., Park, S.H., Garcia, C., Lemonnier, F., de la Salle, H., Bendelac, a, Bonneville, M. & Lantz, O., 1999. An invariant T cell receptor alpha chain defines a novel TAP-independent major histocompatibility complex class Ib-restricted alpha/beta T cell subpopulation in mammals. *The Journal of experimental medicine*, 189(12), pp.1907–21.

Toledo, C., Yu, D., Hoellerbauer, P., Davis, R., Basom, R., Girard, E., Corrin, P., Bolouri, H., Davison, J., Zhang, Q., Nam, D.-H., Lee, J., Pollard, S., Delrow, J., Clurman, B., Olson, J. & Paddison, P., 2016. Abstract 4370: Genome-wide CRISPR-Cas9 screens reveal loss of redundancy between PKMYT1 and WEE1 in patient-derived glioblastoma stem-like cells. *Cancer Research*, 76(14 Supplement), p.4370 LP-4370.

Tomasec, P., Braud, V.M., Rickards, C., Powell, M.B., McSharry, B.P., Gadola, S., Cerundolo, V., Borysiewicz, L.K., McMichael, A.J. & Wilkinson, G.W.G., 2000. Surface Expression of HLA-E, an

Inhibitor of Natural Killer Cells, Enhanced by Human Cytomegalovirus gpUL40. *Science*, 287(5455), pp.1031–1033.

Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., Mcdermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., Leming, P.D., Spigel, D.R., Antonia, S.J., Horn, L., Drake, C.G., Pardoll, D.M., Chen, L., Sharfman, W.H., Anders, R.A., Taube, J.M., Mcmiller, T.L., Xu, H., Korman, A.J., Jure-kunkel, M., Agrawal, S., McDonald, D., Kollia, G.D., Gupta, A., Wigginton, J.M. & Sznol, M., 2012. Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer. *New England Journal of Medicine*, 366(26), pp.2443–2454.

Torikai, H., Reik, A., Liu, P.-Q., Zhou, Y., Zhang, L., Maiti, S., Huls, H., Miller, J.C., Kebriaei, P., Rabinovitch, B., Lee, D.A., Champlin, R.E., Bonini, C., Naldini, L., Rebar, E.J., Gregory, P.D., Holmes, M.C. & Cooper, L.J.N., 2012. A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. *blood*, 119(24), pp.5697–5705.

Traherne, J. A, 2008. Human MHC architecture and evolution: implications for disease association studies. *International journal of immunogenetics*, 35(3), pp.179–92.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. & Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, 7(3), pp.562–78.

Travis, J., 2015. Making the Cut. *Science*, 39(6), pp.23–25.

Treiner, E., Duban, L., Bahram, S., Radosavljevic, M., Wanner, V., Tilloy, F., Affaticati, P., Gilfillan, S. & Lantz, O., 2003. Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature*, 422(6928), pp.164–9.

Tsukamoto, K., Deakin, J.E., Graves, J.A.M. & Hashimoto, K., 2013. Exceptionally high conservation of the MHC class I-related gene, MR1, among mammals. *Immunogenetics*, 65(2), pp.115–124.

Tzur, Y.B., Friedland, A.E., Nadarajan, S., Church, G.M., Calarco, J.A. & Colaiácovo, M.P., 2013. Heritable Custom Genomic Modifications in Caenorhabditis elegans via a CRISPR–Cas9 System. *Genetics*, 195(3), pp.1181–1185.

Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D. & Holmes, M.C., 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435(7042), pp.646–51.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D., 2010. Genome editing with engineered zinc finger nucleases. *Nature reviews. Genetics*, 11(9), pp.636–46.

Urrutia, R., 2003. KRAB-containing zinc-finger repressor proteins. *Genome biology*.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Talkowski, M.E. &

Musunuru, K., 2014. Low incidence of Off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell*, 15(1), pp.27–30.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. & Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 153(4), pp.910–8.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. & Qiu, J.-L., 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotech*, 32(9), pp.947–951.

Wanisch, K. & Yáñez-Muñoz, R.J., 2009. Integration-deficient lentiviral vectors: a slow coming of age. *Molecular therapy : the journal of the American Society of Gene Therapy*, 17(8), pp.1316–32.

Waring, P. & Mullbacher, A., 1999. Cell death induced by the Fas/Fas ligand pathway and its role in pathology. *Immunol Cell Biol*, 77(4), pp.312–317.

Wayengera, M., 2011. Identity of zinc finger nucleases with specificity to herpes simplex virus type II genomic DNA: novel HSV-2 vaccine/therapy precursors. *Theoretical Biology & Medical Modelling*, 8, p.23.

Wolchok, J.D., Neyns, B., Linette, G., Negrier, S., Lutzky, J., Thomas, L., Waterfield, W., Schadendorf, D., Smylie, M., Guthrie Jr, T., Grob, J.-J., Chesney, J., Chin, K., Chen, K., Hoos, A., O'Day, S.J. & Lebbé, C., 2010. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *The Lancet Oncology*, 11(2), pp.155–164.

Wolfe, S.A., Nekludova, L. & Pabo, C.O., 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struc*, pp.183–212.

Wollenberg, A., Kraft, S., Hanau, D. & Bieber, T., 1996. Immunomorphological and Ultrastructural Characterization of Langerhans Cells and a Novel, Inflammatory Dendritic Epidermal Cell (IDEC) Population in Lesional Skin of Atopic Eczema. *Journal of Investigative Dermatology*, 106(3), pp.446–453.

Wood, A.J., Lo, T.-W., Zeitler, B., Pickle, C.S., Ralston, E.J., Lee, A.H., Amora, R., Miller, J.C., Leung, E., Meng, X., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D. & Meyer, B.J., 2011. Targeted genome editing across species using ZFNs and TALENs. *Science (New York, N.Y.)*, 333(6040), p.307.

Wooldridge, L., van den Berg, H. a, Glick, M., Gostick, E., Laugel, B., Hutchinson, S.L., Milicic, A., Brenchley, J.M., Douek, D.C., Price, D. a, Sewell, A.K., Berg, H.A. Van Den & Jason, M., 2005. Interaction between the CD8 coreceptor and major histocompatibility complex class I stabilizes

T cell receptor-antigen complexes at the cell surface. *The Journal of biological chemistry*, 280(30), pp.27491–501.

Wooldridge, L., Laugel, B., Ekeruche, J., Clement, M., van den Berg, H. A, Price, D. A & Sewell, A.K., 2010. CD8 controls T cell cross-reactivity. *Journal of immunology (Baltimore, Md. : 1950)*, 185(8), pp.4625–32.

Xiao, T.Z., Bhatia, N., Urrutia, R., Lomberk, G. a, Simpson, A. & Longley, B.J., 2011. MAGE I transcription factors regulate KAP1 and KRAB domain zinc finger transcription factor mediated gene repression. *PloS one*, 6(8), p.e23747.

Xie, F., Ye, L., Chang, J.C., Beyer, A.I., Wang, J., Muench, M.O. & Kan, Y.W., 2014. Seamless gene correction of ??-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Research*, 24(9), pp.1526–1533.

Yamaguchi, H., Hirai, M., Kurosawa, Y. & Hashimoto, K., 1997. A Highly Conserved Major Histocompatibility Complex Class I-Related Gene in Mammals. *Biochemical and Biophysical Research Communications*, 238(3), pp.697–702.

Yanez-Munoz, R.J., Balaggan, K.S., MacNeil, A., Howe, S.J., Schmidt, M., Smith, A.J., Buch, P., MacLaren, R.E., Anderson, P.N., Barker, S.E., Duran, Y., Bartholomae, C., von Kalle, C., Heckenlively, J.R., Kinnon, C., Ali, R.R. & Thrasher, A.J., 2006. Effective gene therapy with nonintegrating lentiviral vectors. *Nat Med*, 12(3), pp.348–353.

Yang, S., Rosenberg, S.A. & Morgan, R.A., 2008. Clinical-scale lentiviral vector transduction of PBL for TCR gene therapy and potential for expression in less differentiated cells. *Journal of Immunotherapy*, 31(9), pp.830–839.

Yee, C., Thompson, J.A., Byrd, D., Riddell, S.R., Roche, P., Celis, E. & Greenberg, P.D., 2002. Adoptive T cell therapy using antigen-specific CD8+ T cell clones for the treatment of patients with metastatic melanoma: in vivo persistence, migration, and antitumor effect of transferred T cells. *Proceedings of the National Academy of Sciences of the United States of America*, 99(25), pp.16168–73.

Yuan, J., Wang, J., Crain, K., Fearns, C., Kim, K. a, Hua, K.L., Gregory, P.D., Holmes, M.C. & Torbett, B.E., 2012. Zinc-finger nuclease editing of human cxcr4 promotes HIV-1 CD4(+) T cell resistance and enrichment. *Molecular therapy : the journal of the American Society of Gene Therapy*, 20(4), pp.849–59.

Yuan, M., Zhang, W., Wang, J., Al Yaghchi, C., Ahmed, J., Chard, L., Lemoine, N.R. & Wang, Y., 2015. Efficiently Editing the Vaccinia Virus Genome by Using the CRISPR-Cas9 System. *Journal of Virology* , 89(9), pp.5176–5179.

Yusa, K., Rashid, S.T., Strick-Marchand, H., Varela, I., Liu, P.-Q., Paschon, D.E., Miranda, E., Ordóñez,

A., Hannan, N.R.F., Rouhani, F.J., Darche, S., Alexander, G., Marciniak, S.J., Fusaki, N., Hasegawa, M., Holmes, M.C., Di Santo, J.P., Lomas, D. a, Bradley, A. & Vallier, L., 2011. Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. *Nature*, 478(7369), pp.391–4.

Zhang, F., Maeder, M.L., Unger-Wallace, E., Hoshaw, J.P., Reyon, D., Christian, M., Li, X., Pierick, C.J., Dobbs, D., Peterson, T., Joung, J.K. & Voytas, D.F., 2010. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America*, 107(26), pp.12028–12033.

Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D., Cunningham, J.M. & Jane, S.M., 2009. PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nature structural & molecular biology*, 16(3), pp.304–311.

Zhu, J. & Paul, W.E., 2010. Heterogeneity and plasticity of T helper cells. *Cell research*, 20(1), pp.4–12..

Zou, J., prashant mali, Huang, X., Dowey, S.N., Cheng, L., Mali, P., Huang, X., Dowey, S.N. & Cheng, L., 2011. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*, 118(17), pp.4599–4608.

Zuniga-Pflucker, J.C., 2007. CD8+ T cells are kept in tune by modulating IL-7 responsiveness. *Nat Immunol*, 8(10), pp.1027–1028.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y. & Liu, D.R., 2014. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature biotechnology*, 33(1), pp.73–80.

8 Appendix

HLA-A2 heavy chain sequence:

MGSHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDGETRK
VKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDGKDYIALKEDLRSW
TAADMAAQTTKHKWEAAHVAEQLRAYLEGTCVEWLRRYLENGKETLQRTDAPKTHMTHHAVSDHEA
TLRCWALSFYPAEITLTWQRDGEDQTQDTELVETRPAGDGTFQKWAAVVVPSGQEQRYTCHVQHEGL
PKPLTLRWEPGLNDIFEAQKIEWHE

β2-microglobulin sequence:

MIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFT
PTEKDEYACRVNHVTLSQPKIVKWDRDM

Appendix Figure 1. The amino acid sequence of the HLA-A2 heavy chain and β2-microglobulin. The biotin tag is underlined in red.

Appendix Table 1. Cloning methods used to generate each of the constructs used in this thesis.

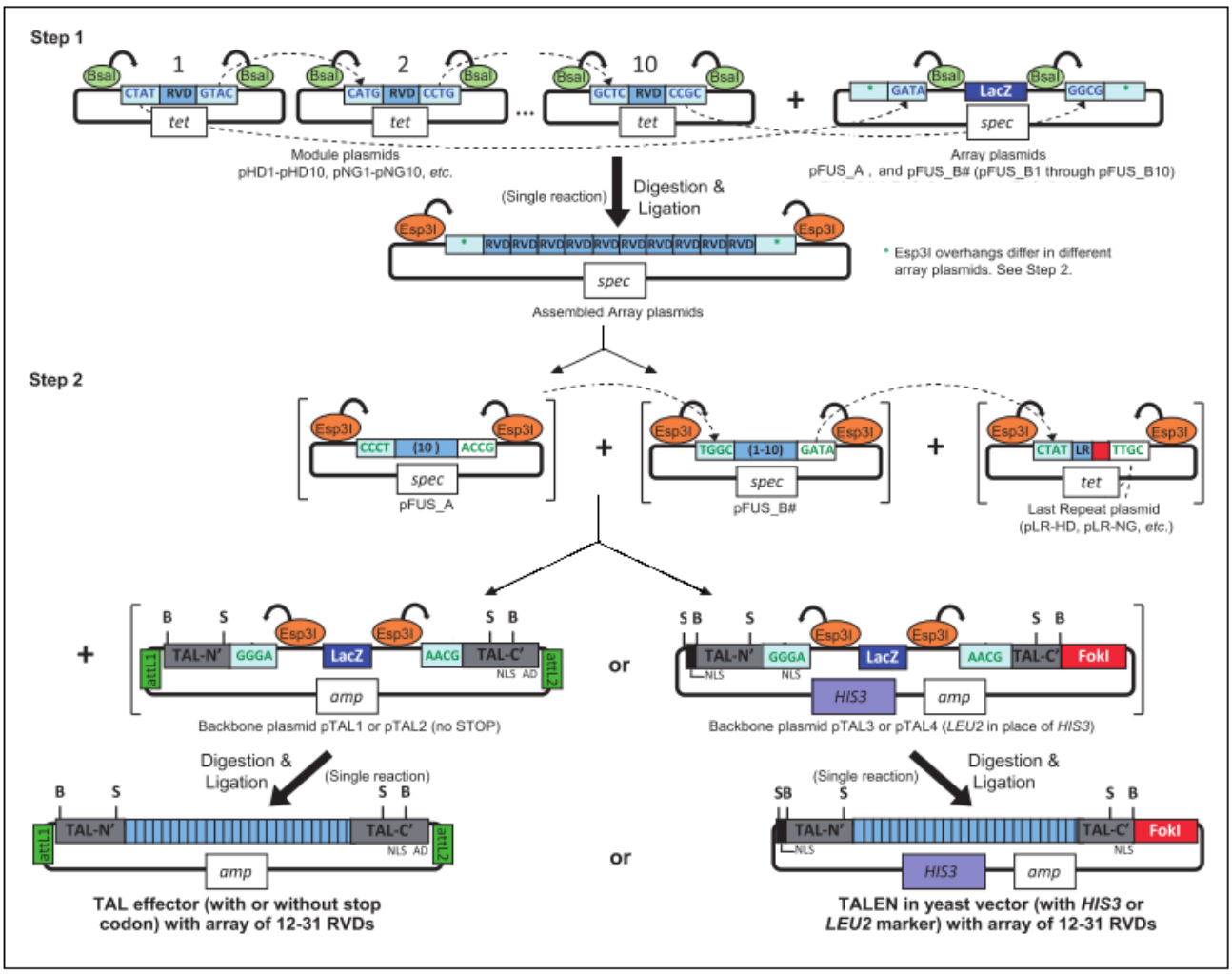
Plasmid Backbone	Insert	Cloning method	Chapter
pLKO.1	shRNA CD8 shRNA MR1	Inserts ordered were flanked by NdeI and EcoRI restriction sites. Inserts and pLKO.1 backbone were digested with NdeI and EcoRI enzymes. DNA was purified and ligated.	3
pRRL	LZFN RZFN	The LZFN and RZFN DNA sequences were PCR amplified from the transfer plasmid using L_ZFN_XbaI_F or R_ZFN_XbaI_F and L_Fok1_NsiI_R or R_Fok1_NsiI_R to introduce XbaI and NsiI restriction sites, the PCR product was treated with DpnI. The treated PCR product and pRRLSIN.cPPT.PGK-GFP.WPRE backbone were digested with XbaI and NsiI. DNA was purified and ligated.	3
GECKO	CD8 crRNA 1 CD8 crRNA 2 CD8 crRNA 3 CD8 crRNA 4 CD8 crRNA 5	Cloning is described in section 2.4.17.3	3
pRRL	ZF binding site (1x) ZF binding site (3x)	The inserts ordered were flanked by XhoI restriction sites. The inserts and pRRLSIN.cPPT.PGK-GFP.WPRE backbone were digested with XhoI enzymes. DNA was purified and ligated.	4
pV2	HP1A KRAB ScanKRAB TRIM28 ARP1 GCN4 SID	pV2 backbone was amplified with pV2_ backbone_XbaI_ F and pV2_Age1prox_Rev to introduce AgeI and XbaI restriction sites, the PCR product was treated with DpnI. The treated PCR product and ZFP sequences (left or right) flanked with AgeI and XbaI were digested with AgeI and XbaI restriction enzymes. The DNA was purified was ligated. The pV2 backbone containing the LZF or RZF was PCR amplified with pV2_ZF_R_Codop_R or pV2_ZF_L_Codop_R and pV2_ backbone_XbaI_ F to introduce BamHI and XbaI restriction sites. The PCR products were treated with dpnI. The treated PCR product and each of the repressor domains was digested with BamHI and XbaI. The DNA was purified was ligated.	4
pRRL	RZF-KRAB	The pV2 RZF-KRAB construct was PCR amplified with pV2_ZF_R_codop_BamHI_Forward and KRAB_Xba1_Reverse, the PCR product was treated with DpnI. The pRRL 2AGFP backbone and PCR product was digested with BamHI and XbaI. DNA was purified and ligated.	4
pRRL	KRAB x	The pRRL RZF-KRAB construct was amplified with ZFless_fwd and ZFless_rev primers to insert a NdeI restriction site upstream and downstream of the ZF sequence. The PCR product was treated with dpnI and digested with NdeI. DNA was purified and ligated.	4

pELNSxv	rCD2 PPI	The PPI TCRα and TCRβ chains separated by a 2A sequences were synthesised flanked by NheI and XhoI restriction sites. The pELNSxv backbone and TCR sequences were digested with NheI and XhoI. The DNA was purified and ligated.	4
pELNSxv	CD8 KRAB PPI	The pRRL R ZF KRAB construct was PCR amplified with pV2_ZF_R_codop_BamHI_Forward and KRAB_Sal1_Reverse, the PCR product was treated with DpnI. The PCR product and pELNsxv PPI CD8 KRAB backbone were digested with BamHI and SalI. DNA was purified and ligated.	4
pRRL	MR1 crRNA 1 MR1 crRNA 2 MR1 crRNA 3 MR1 crRNA 4 MR1 crRNA 5	Cloning is described in section 2.4.17.2	5

Appendix Table 2. Name and sequence of all primers used in thesis

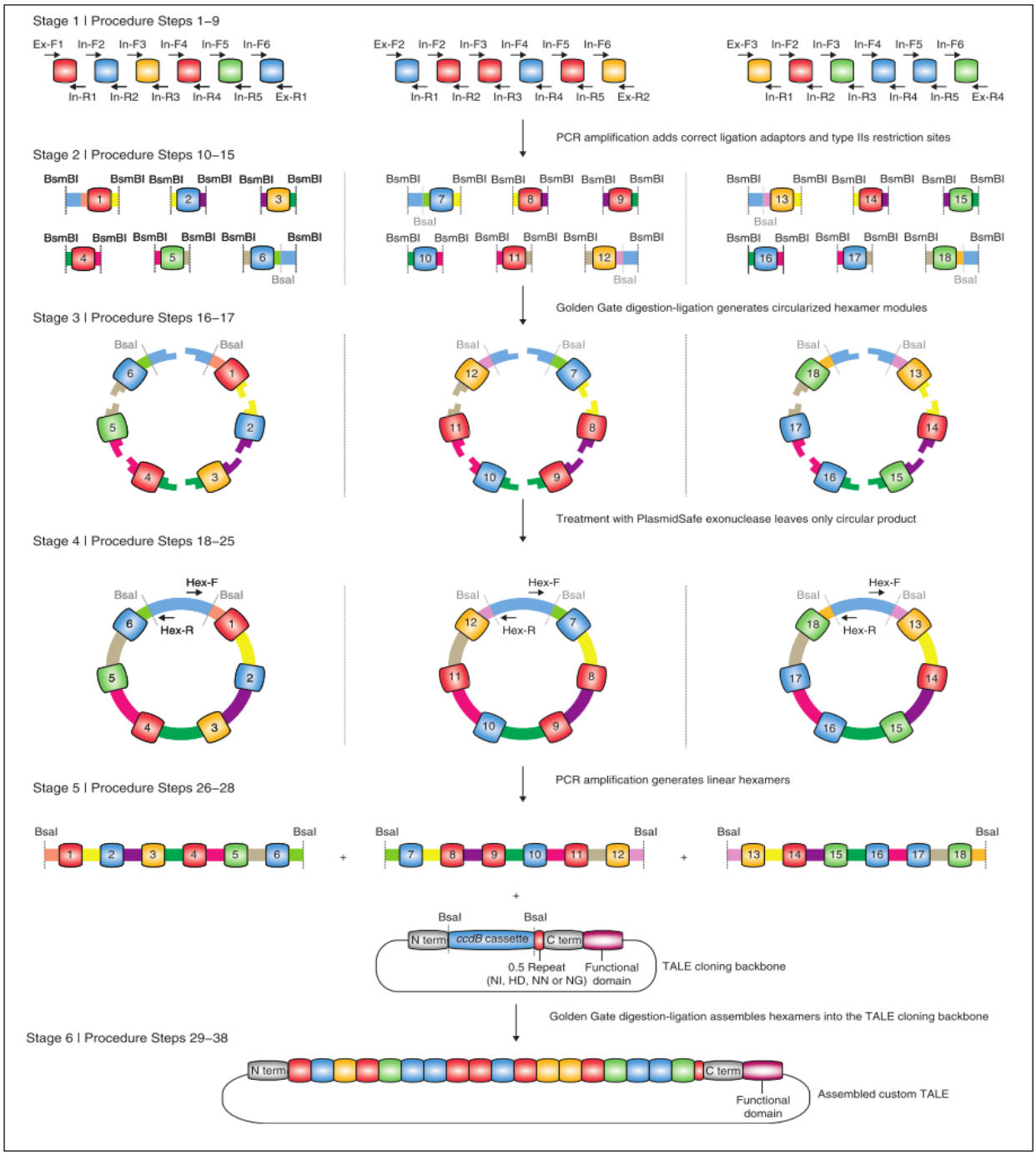
Comments	Primer name	Sequence
Primers to insert XbaI and NsiI restriction sites of pRRL LZFN cloning	L_ZFN_XbaI_F	TCTAGACAGTGTGGTGGAATTCGCC
	L_Fok1_NsiI_R	GTGACCGAGTTCAAGTTCCTGTTATGCAT
Primers to insert XbaI and NsiI restriction sites of pRRL RZFN cloning	R_ZFN_XbaI_F	TCTAGAcagtgtggtggaattcgcc
	R_Fok1_NsiI_R	AACAACGGCGAGATCAACTTCATGCAT
Primers to insert XbaI and AgeI restriction sites in pV2 backbone for cloning	pV2_ backbone_XbaI_ F	tctagaaaaatcagcctcgactgtgccttc
	pV2_Age1prox_Rev	Gcgcaccggtggtggccgtacgcc
Primers to insert XbaI and BamHI restriction sites flanking the LZF and RZF for cloning	pV2_ backbone_XbaI_ F	tctagaaaaatcagcctcgactgtgccttc
	pV2_ZF_L_Codop_R	ggatccTGCGCAGATGGATTTT
	pV2_ZF_R_Codop_R	ggatccGCGCAGATGGATCTT
Primers to insert XbaI and BamHI restriction sites flanking the RZFKRAB for cloning	pV2_ZF_R_codop_Ba mHI_Forward	ggatccATGCGGTCAGACTACAAAG
	KRAB_Xba1_Reverse	tctagattaAACCAGCCAAGGTTCTTCTCC
Primers used to insert NdeI restriction sites flanking the RZF of 'pRRL_RZF_KRAB' allowing for removal of ZF	ZFless_fwd	GGCATATGCGGACTCTGGTGACCT
	ZFless_rev	GGCATATGCCCTGGGGAGAGAGGTCG
Primers used to insert BamHI and SalI restriction sites to clone RZFKRAB into pELNSxv backbone	pV2_ZF_R_codop_Ba mHI_Forward	ggatccATGCGGTCAGACTACAAAG
	KRAB_Sal1_Reverse	GTCGACttaAACCAGCCAAGGTTCTTCTCC
CD8 crRNA target sequence PCR cloning primers	CD8_crRNA1_F	CACCGGAGCAAGGCGGTCACTGGTA
	CD8_crRNA1_R	AAACTACCAGTGACCGCCTTGCTCC
	CD8_crRNA2_F	CACCGGCTGCTGTCCAACCCGACGT
	CD8_crRNA2_R	AAACACGTCGGGTTGGACAGCAGCC
	CD8_crRNA3_F	CACCGTCCGATCCAGCGGCGACACC
	CD8_crRNA3_R	AAACGGTGTCGCCGCTGGATCGGAC
	CD8_crRNA4_F	CACCGAACAAGCCCAAGGCGGCCGA
	CD8_crRNA4_R	AAACTCGGCCGCCTTGGGCTTGTTC
	CD8_crRNA5_F	CACCGCTCTCGGCGGAAGTCGCTCA
	CD8_crRNA5_R	AAACTGAGCGACTTCCGCCGAGAGC
Primers for genomic CD8 PCR and SURVEYOR assay	CD8_surv_forward	GAAGGGCGCAACTTTCCC
	CD8_surv_reverse	GTAGCCCTCGTTCTCTCGG
Primers for genomic MR1 PCR and SURVEYOR assay	MR1_SURV1_Fwd	GCATGTGTTTGTGTGCCTGT
	MR1_SURV1_Rev	GGTGCAATTCAGCATCCGC
Primers for MR1 cDNA PCR and SURVEYOR assay	MR1_SURV2_Fwd	GGTCTTACTGACATCCACTTTGC
	MR1_SURV2_Rev	CAGTGATCAGGCGCGAG
MR1 amplicon cloning primers	MR1_SURV_Fwd_BsaI	gcgcGGTCTCcGCATGTGTTTGTGTGCCTGT
	MR1_SURV_Rev_BsaI	gcgcGGTCTCcTGCCGGTGCAATTCAGCATCCGC
Primers used to sequence MR1 amplicons	MR1_Seq_Fwd	CCAGTTGCTGAAGATCGCGAAGC
	MR1_Seq_Rev	TGCCACTCGATGTGATGTCCTC
gRNA/pCMV-Cas9 PCR amplification in pCDNA.3	pCDNA.3_Fwd	GCACCGGTTGTACAAAAAAGCAGGCTTTA
	pCDNA.3_Rev	GCATGCATTCACACCTTCCTCTTCTTCT
"Empty" pRRL vector PCR	pRRL.0_Fwd	GCATGCATAATCAACCTCTGGATTACAAAATTTG

amplification	pRRL.0_Rev	ACCGGTGCTAGTCTCGTGATCGATAAAAT
MR1 crRNA target sequence PCR cloning primers	MR1crRNA_A_Fwd	GAAACGCCCGTTTTAGAGCTAGAAATAGCAAGTTAA
	MR1crRNA_A_Rev	GGATCCCATCCGGTGTTTCGTCCTTTCC
	MR1crRNA_B_Fwd	CAGCGATTCCGTTTTAGAGCTAGAAATAGCAAGTTAA
	MR1crRNA_B_Rev	TGCTTCACCGGTGTTTCGTCCTTTCC
	MR1crRNA_C_Fwd	AATTTATTTCGGTGTTTTAGAGCTAGAAATAGCAAGTTAA
	MR1crRNA_C_Rev	CAGGGACGGTGTTTCGTCCTTTCC
	MR1crRNA_D_Fwd	GCCTGATCACTGTTTTAGAGCTAGAAATAGCAAGTTAA
	MR1crRNA_D_Rev	GCGAGGTTCGGTGTTTCGTCCTTTCC
	MR1crRNA_E_Fwd	TATGACGGGCGTTTTAGAGCTAGAAATAGCAAGTTAA
	MR1crRNA_E_Rev	TGCATACTGCGGTGTTTCGTCCTTTCC
Primers for sequencing the region of the RP11 gene which contains the crRNA A target site	RP11_FWD	gctaaatgaatgcagttgaaggacctg
	RP11_REV	CAGTGATCAGGCACGAGGTTCTC
Colony PCR primers flanking gRNA target sites	pLKO.1-A	GACTATCATATGCTTACCGT
	gRNAcolPCR_R	CACTTGATGTACTGCCAAGT
For sequencing	CMV_fwd	CGCAAATGGGCGGTAGGCGTG
	hPGK_fwd	GTGTTCCGCATTCTGCAAG
	eGFP_rev	CGTCGCCGTCCAGCTCGACCAG
QPCR primers	RTPCR_CD8a_F	GACGTGTTTGCAAATGTCCC
	RTPCR_CD8a_R	AAAATGAAAGGGAAGGACTTGCT
	RTPCR_18S_F	GTAACCCGTTGAACCCCATT
	RTPCR_18S_R	CCATCCAATCGGTAGTAGCG
	RTPCR_ACTB_F	GACCCAGATCATGTTTGAGACCTT
	RTPCR_ACTB_R	CAGAGGCGTACAGGGATAGCA

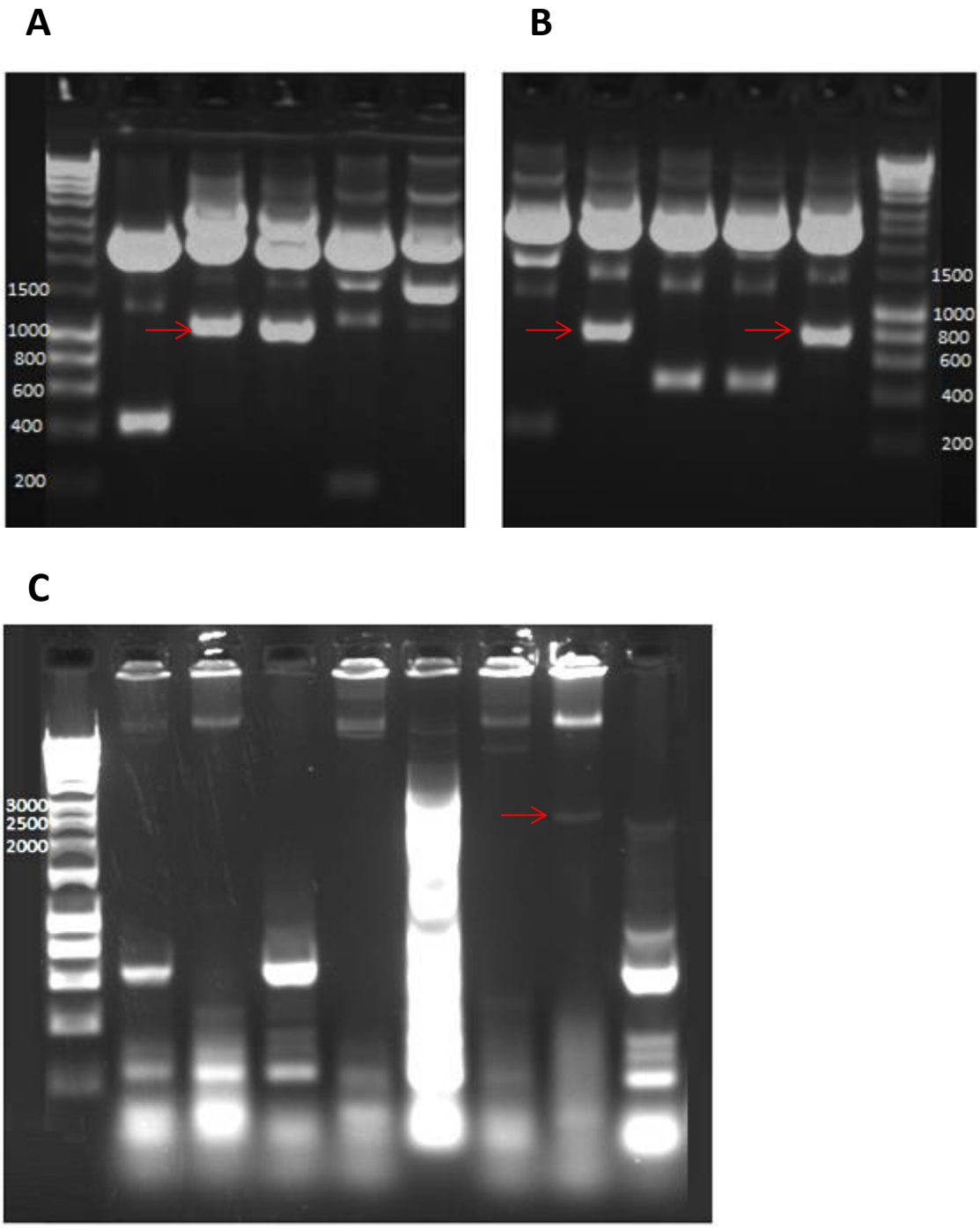


Appendix Figure 2. TALEN construction with the Cermak kit (Cermak et al., 2011). TALENs were assembled in two steps. Step 1 involved the modular assembly of the monomers 1-11 of the TALE in a pFUS_A backbone and the assembly of monomers 12-19 of the TALE in a pFUS_B backbone in a golden gate reaction using the restriction endonucleases BsaI and Esp3I. The golden gate reaction was treated with endonuclease and transformed onto Xgal IPTG spectinomycin plates. DNA from the colonies was digested with BsmBI and run on an agarose gel. Plasmids containing the correct size fragments were selected and used for stage 2 of the assembly. The second stage involved ligating the 10 monomer plasmid from pFUS_A and 8 monomer plasmid from pFUS_B with the final half repeat and the destination backbone to produce complete constructs by a golden gate reaction. The golden gate reaction was treated with endonuclease and transformed onto a Xgal IPTG spectinomycin plate. DNA was extracted from the colonies, digested with BsmBI and run on an agarose gel. Samples containing the correct size fragment were sequenced by Sanger sequencing.

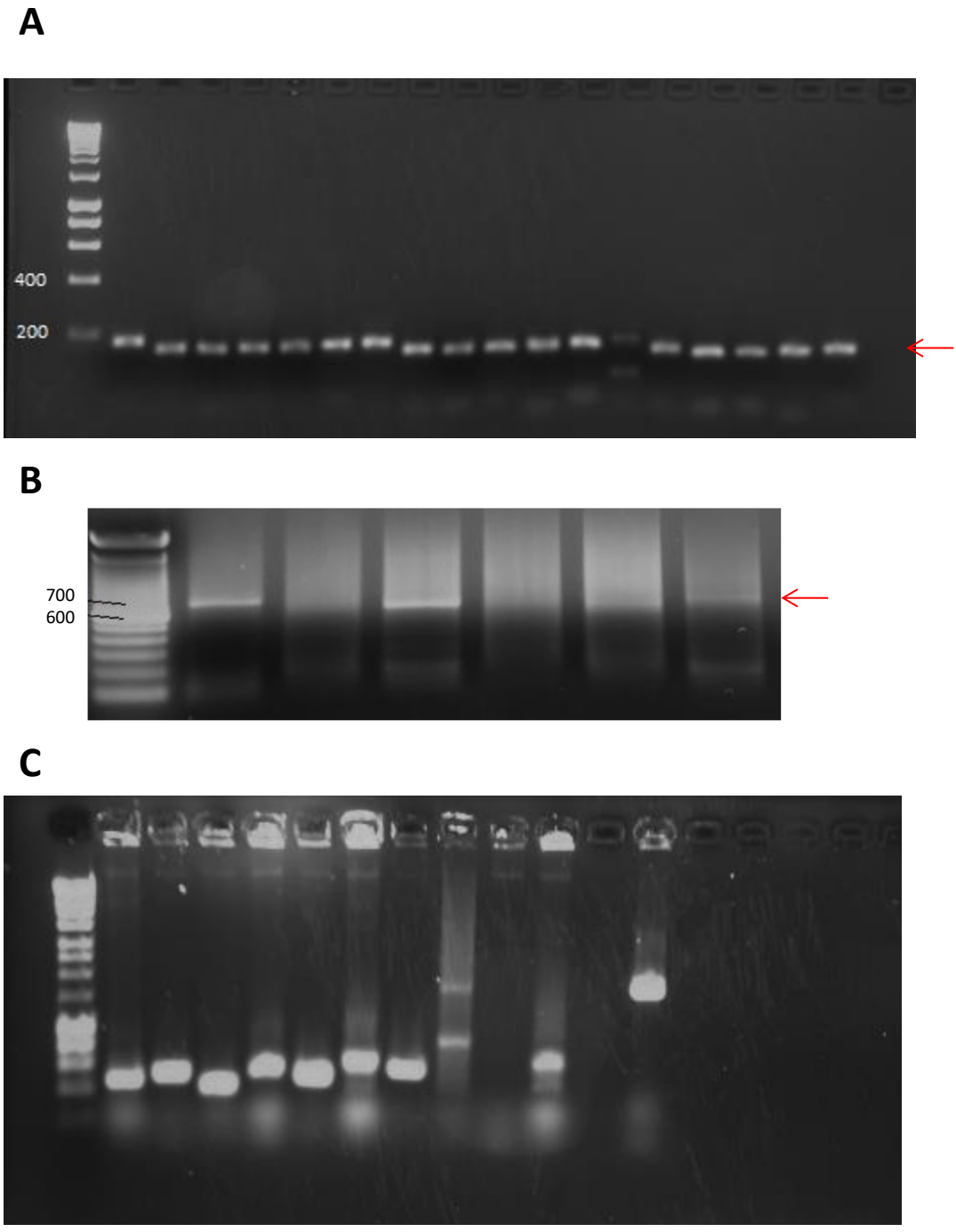
NLS - nuclear localization signal. AD - transcriptional activation domain. tet - tetracycline resistance. spec - spectinomycin resistance. amp - ampicillin resistance.



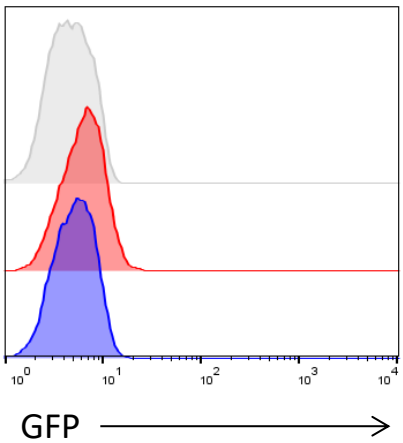
Appendix Figure 3. TALEN construction with the Sanjana kit (Sanjana et al., 2012). Assembly of the TALENs with the Sanjana kit had 6 stages. Stage 1 - Monomers were produced with specific ligation adapters by PCR amplification. Stage 2 – Hexameric tandem repeats were assembled from the purified monomers from stage 1 by a Golden Gate reaction. Stage 3 - The hexamer reactions were treated with exonuclease. Stage 4 - Hexamers were PCR amplified and purified. Stage 5 - The three adjacent hexamers were ligated into the TALEN backbone. Stage 6 - The assembled TALEN was transformed into competent cells and DNA from clones was sequenced by Sanger sequencing.



Appendix Figure 4. Generation of TALENs with the Cermak kit. A) DNA extracted from colonies transduced with the 10mer pFUS_A plasmids were digested with BsmBI and run on a 1% agarose gel. Lane 1 .Bioline hyper ladder 1. Lanes 2 – 6 minipreps from 5 different colonies. Lanes 3 and 4 contain a fragment of the correct size ~ 1000bp (shown by arrow). B) DNA extracted from colonies transduced with the 8mer pFUS_B plasmids were digested with BsmBI and run on a 1% agarose gel. Lanes 1 – 5 minipreps from 5 different colonies. Lanes 2 and 5 contain a fragment of the correct size ~ 800bp (shown by arrow). All minipreps were sequenced to ensure RVDs had been assembled in the correct order before continuing to the next stage of the assembly. C) DNA extracted from colonies transduced with the final half repeat backbone ligated to the 18mer TALE were digested with BsmBI and run on a 1% agarose gel. Lane 1 .Bioline hyper ladder 1. Lanes 2 – 9 PCR amplification from 10 different colonies. Lane 8 (shown by arrow) contains a fragment of the correct size ~ 2200bp.



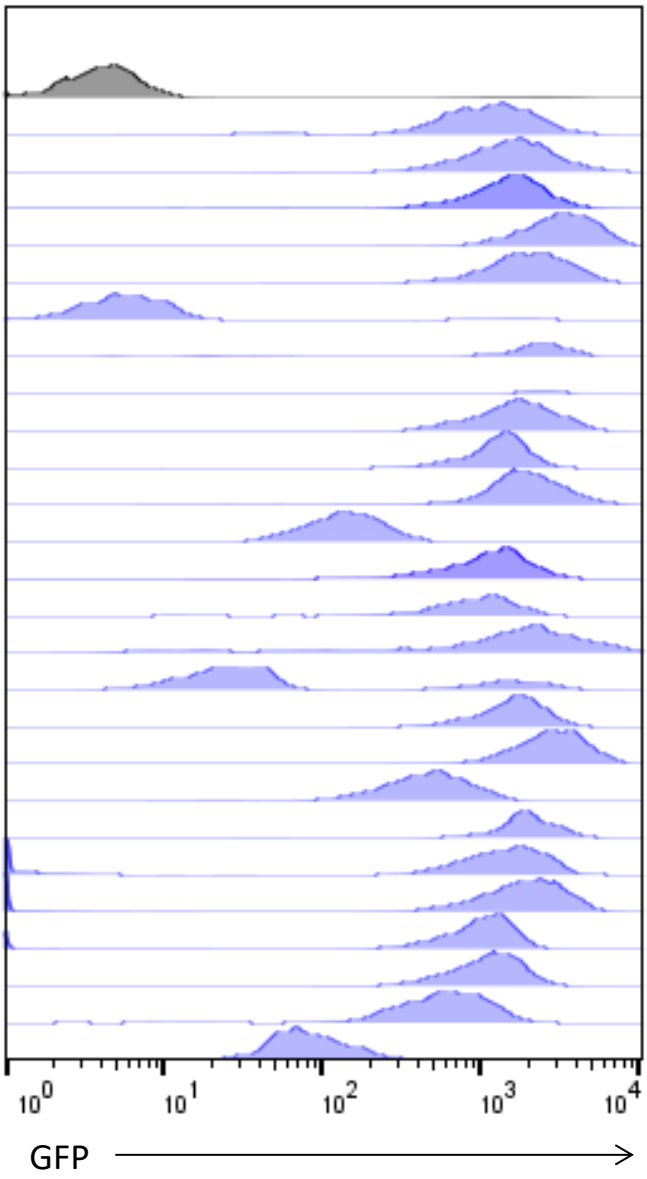
Appendix Figure 5. A) PCR amplification of NI monomer with ligation adapter primers for each of the 18 positions on a 1% agarose gel (representative of NN, NG and HD). Lane 1. Bioline hyper ladder 1. Lanes 2 – 19 monomers for each position. Monomers in positions 1, 6, 7, 12, 13 and 18 were 170bp in length, the monomers in the remaining positions were 150bp in length. B) PCR amplification of hexamer units three for each TALE on a 1% agarose gel. Lane 1. Track it 100bp DNA ladder. Lanes 2 -4 the three hexamers of the left TALE. Lanes 5 - 7 the three hexamers of the right TALE. C) The 3 hexamer units were ligated into the FokI backbone and transformed into XL10 Gold. 5 colonies were screened by colony PCR for each TALE and run on a 1% agarose gel. Lane 1 - Bioline hyper ladder 1. Lanes 2 – 6 5 colonies for the left TALE. Lanes 7 - 13 5 colonies for the right TALE. Correctly assembled TALEs should have resulted in a fragment of 2175bp, none of the colonies screened were the correct size. Sequencing confirmed that the TALEs were not assembled correctly.



Sample	Geometric mean of CD8 (MFI)
FMO	4.59
L TALE	6.22
R TALE	5.06

Appendix Figure 6. Transduction of Molt3 cells with the left or right TALE constructs by lentiviral transduction. The cells were cultured for 48 hours post transduction and analysed by flow cytometry. There is no GFP+ population observed in the cells transduced with the left TALE (red) or right TALE (blue) compared to the FMO control (grey).

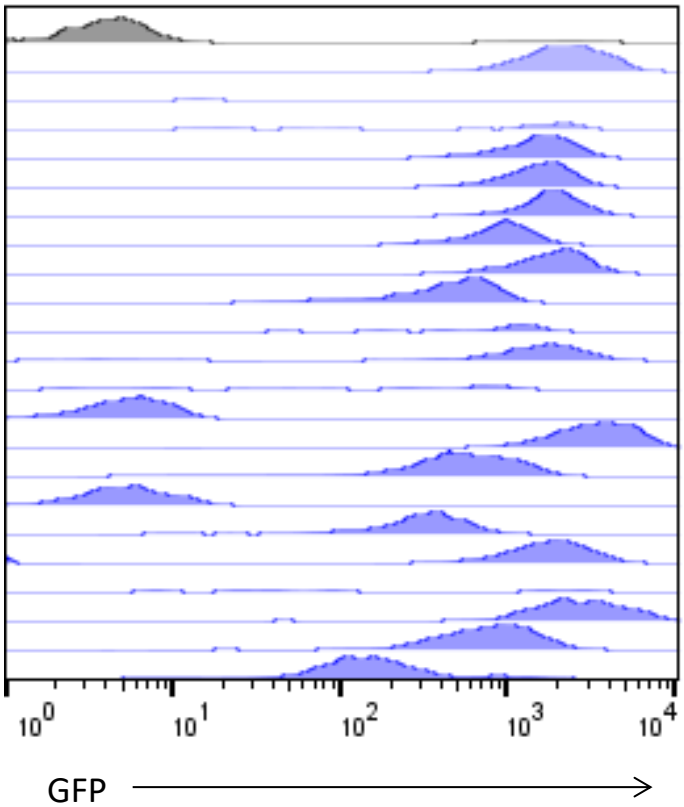
1x



Sample	Geometric mean GFP
FMO	6.12
1	868
2	1345
3	1277
4	2925
5	1830
6	7.17
7	933
8	2377
9	1543
10	1186
11	1662
12	131
13	947
14	819
15	874
16	67.4
17	1453
18	2606
19	405
20	1705
21	261
22	677
23	661
24	1031
25	539
26	77.9

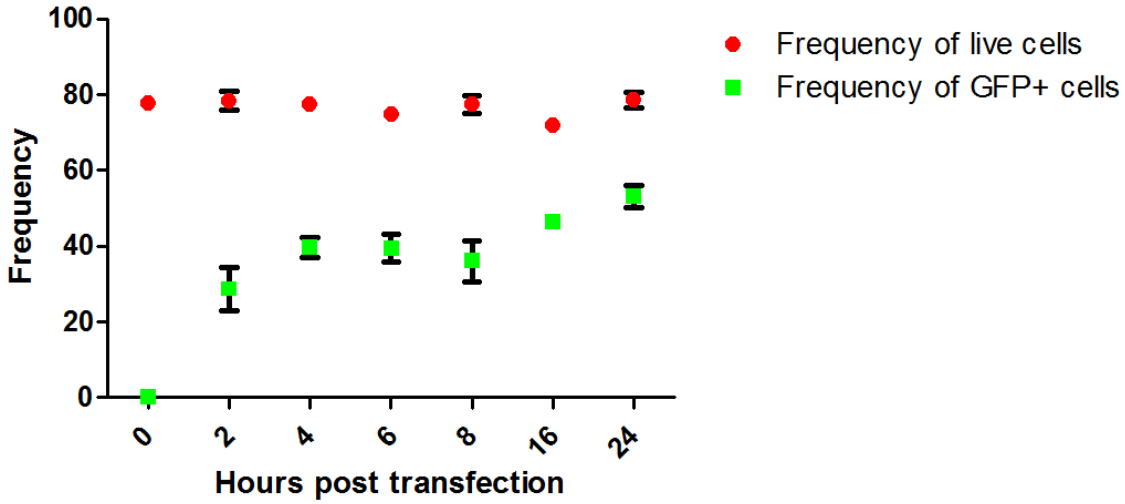
Appendix Figure 7. Screening of 26 clones produced by limiting dilution of the bulk 293T cells transduced with the pRRL_1x reporter construct by lentiviral transduction

3x

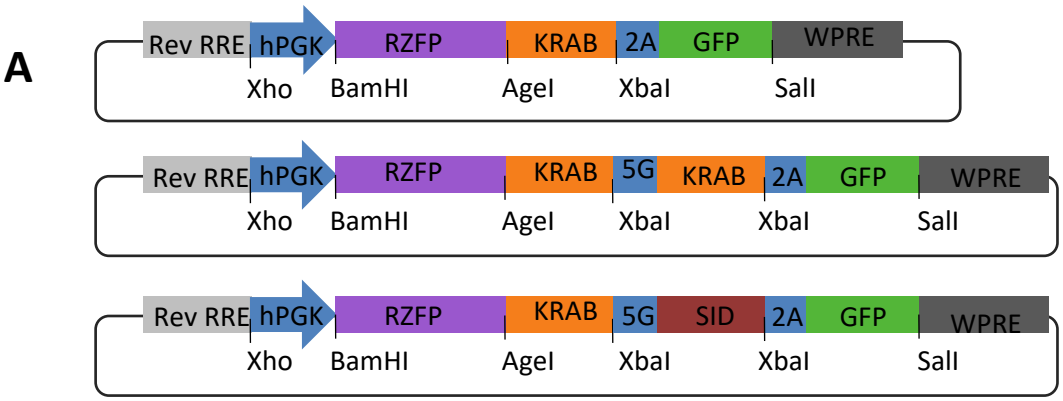


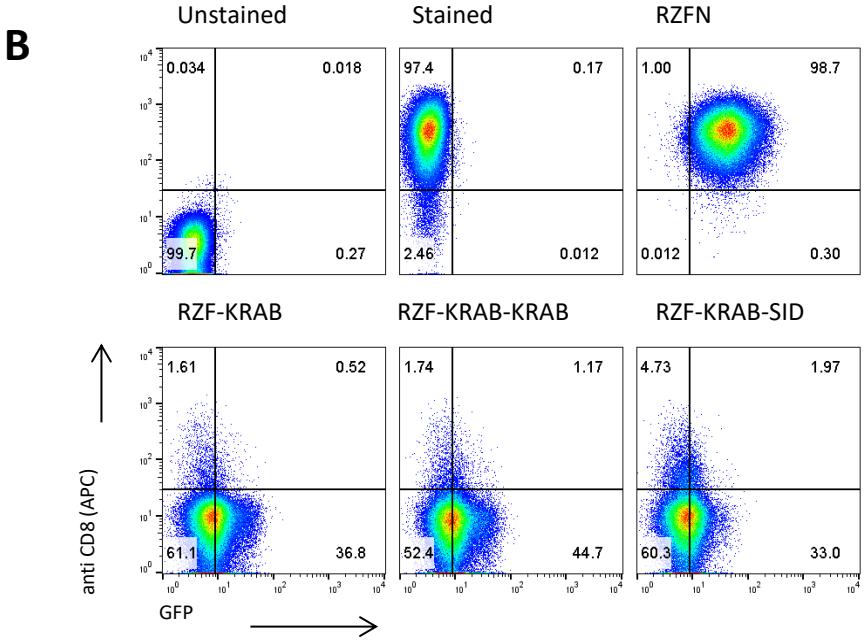
Sample	Geometric mean GFP
FMO	6.12
1	2037
2	14.2
3	480
4	1356
5	1372
6	1701
7	791
8	1614
9	366
10	650
11	772
12	75.4
13	5.17
14	2928
15	451
16	5.28
17	262
18	1252
19	122
20	2320
21	658
22	150

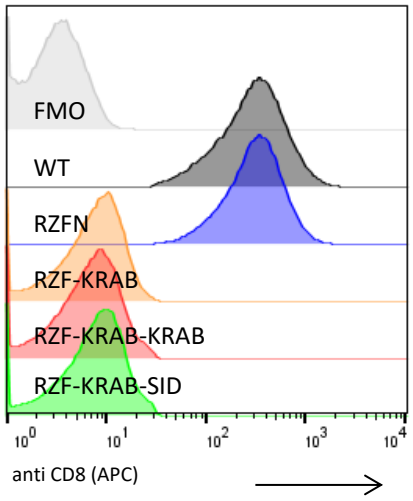
Appendix Figure 8. Screening of 22 clones produced by limiting dilution of the bulk 293T cells transduced with the pRRL_3x reporter construct by lentiviral transduction



Appendix Figure 9. Optimisation of incubation time post $CaCl_2$ transfection. WT 293T cells were transfected with a plasmid containing the GFP reporter gene by $CaCl_2$ transfection and incubated for 2, 4, 6, 8, 16 or 24 hours after which the supernatant was removed and replaced with 4 mL of fresh medium. The cells were cultured for a further 48 h and then stained with a live / dead stain and analysed by flow cytometry. All transfections were carried out in triplicate wells. Standard error of the mean of 3 replicates is shown. The data is representative of 2 experiments.

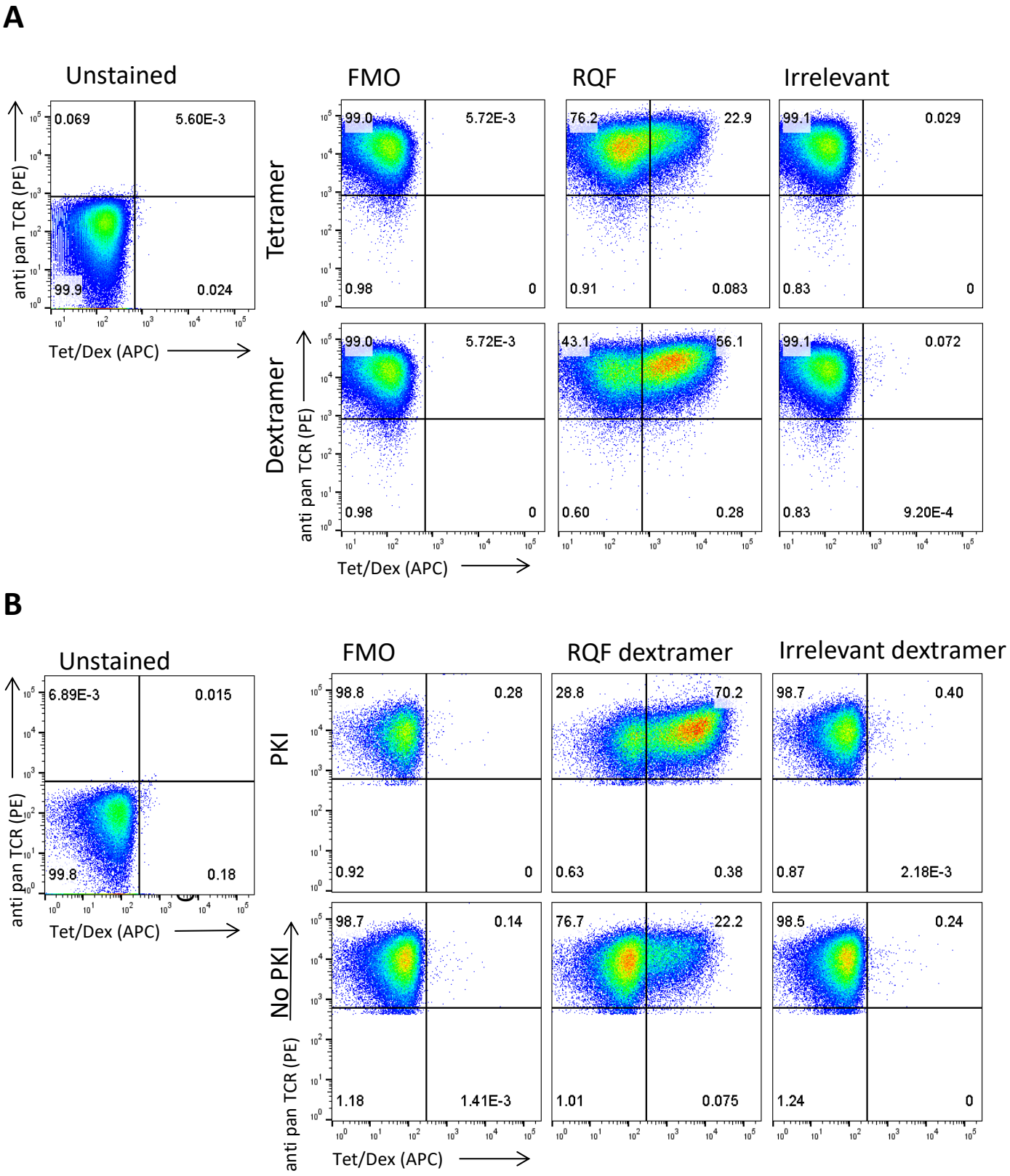






Sample	Percentage reduction of CD8 MFI (%)
FMO	100
WT	0
RZFN (GFP+)	-4
RZF-KRAB (GFP+)	99
RZF-KRAB-KRAB (GFP+)	99
RZF-KRAB-SID (GFP+)	98

Appendix Figure 10. Generation and validation of tandem repressor module constructs A) A diagrammatic representation of the RZF-KRAB, RZF-KRAB-KRAB and RZF-KRAB-SID repressor domain constructs. B) Molt3 cells were transduced with the RZFN, the RZF-KRAB, RZF-KRAB-KRAB and RZF-KRAB-SID constructs by lentiviral transduction. The cells were cultured and monitored 2 weeks post transduction by flow cytometry, a CD8- population was observable in the samples transduced with each of the repressor domain constructs. C) The CD8 histograms of FMO is shown in light grey, WT is shown in dark grey, RZFN (GFP+) is shown in blue, RZF-KRAB (GFP+) is shown in orange, RZF-KRAB-KRAB (GFP+) is shown in red and RZF-KRAB-SID (GFP+) is shown in green. The table displays the percentage reduction of CD8 MFI relative to the WT after deducting background is shown. The data is representative of 2 experiments.



Appendix Figure 11. Optimisation of multimer staining. Molt3 cells were transduced with the rCD2 PPI construct by lentiviral transduction. At two weeks post transduction the Molt3 cells were stained with multimers. A) Dot plot showing Molt3 stained with RQF tetramers and dextramers, ELA tetramers and dextramers as irrelevant controls or without tetramers and dextramers as FMO controls. B) Dot plot showing Molt3 incubated with and without PKI prior to staining with RQF and ELA dextramers. The dot plots were gated on the TCR+ population within the transduced cells. This data is representative of 2 experiments.

The Journal of Immunology

Engineering of Isogenic Cells Deficient for MR1 with a CRISPR/Cas9 Lentiviral System: Tools To Study Microbial Antigen Processing and Presentation to Human MR1-Restricted T Cells

Bruno Laugel, Angharad Lloyd, Erin W. Meermeier, Michael D. Crowther, Thomas R. Connor, Garry Dolton, John J. Miles, Scott R. Burrows, Marielle C. Gold, David M. Lewinsohn and Andrew K. Sewell

J Immunol 2016; 197:971-982; Prepublished online 15 June 2016;
doi: 10.4049/jimmunol.1501402
http://www.jimmunol.org/content/197/3/971

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
9650 Rockville Pike, Bethesda, MD 20814-3994.

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Engineering of Isogenic Cells Deficient for MR1 with a CRISPR/Cas9 Lentiviral System: Tools To Study Microbial Antigen Processing and Presentation to Human MR1-Restricted T Cells

Bruno Laugel,*[,1,2] Angharad Lloyd,*[,2] Erin W. Meermeier,[†] Michael D. Crowther,* Thomas R. Connor,[‡] Garry Dolton,* John J. Miles,[§] Scott R. Burrows,[§] Marielle C. Gold,[†] David M. Lewinsohn,[‡] and Andrew K. Sewell*

The nonclassical HLA molecule MHC-related protein 1 (MR1) presents metabolites of the vitamin B synthesis pathways to mucosal-associated invariant T (MAIT) cells and other MR1-restricted T cells. This new class of Ags represents a variation on the classical paradigm of self/non-self discrimination because these T cells are activated through their TCR by small organic compounds generated during microbial vitamin B_2 synthesis. Beyond the fundamental significance, the invariant nature of MR1 across the human population is a tantalizing feature for the potential development of universal immune therapeutic and diagnostic tools. However, many aspects of MR1 Ag presentation and MR1-restricted T cell biology remain unknown, and the ubiquitous expression of MR1 across tissues and cell lines can be a confounding factor for experimental purposes. In this study, we report the development of a novel CRISPR/Cas9 genome editing lentiviral system and its use to efficiently disrupt MR1 expression in A459, THP-1, and K562 cell lines. We generated isogenic $MR1^{-/-}$ clonal derivatives of the A549 lung carcinoma and THP-1 monocytic cell lines and used these to study T cell responses to intracellular pathogens. We confirmed that MAIT cell clones were unable to respond to $MR1^{-/-}$ clones infected with bacteria whereas Ag presentation by classical and other nonclassical HLAs was unaffected. This system represents a robust and efficient method to disrupt the expression of MR1 and should facilitate investigations into the processing and presentation of MR1 Ags as well as into the biology of MAIT cells. ***The Journal of Immunology*****, 2016, 197: 971–982.**

Mucosal-associated invariant T (MAIT) cells are the most abundant nonconventional T cell subset, accounting for up to 5% of all T cells in humans, and are thought to be important for the control of a number of bacterial, fungal, and yeast infections (1–5). These so-called innate-like T cells, which are mostly found in the blood, the liver, and at mucosal surfaces, express a semi-invariant TCR consisting of an α-chain using the canonical TRAV1-2–TRAJ33/12/20 (Vα7.2-Jα33/12/20) rearrangements (6). MAIT cells acquire effector functions during thymic selection and readily respond to Ags derived from many (but not all) bacteria such as *Escherichia coli*, *Klebsiella pneumoniae*, *Mycobacterium tuberculosis*, or *Staphylococcus epidermis* as well as several yeast species in the periphery without prior priming (3, 7). MAIT cell activation is mediated by the interaction between the TCR and microbe-derived Ags presented by the nonclassical MHC-related protein 1 (MR1) and results in the secretion of cytokines as well as in granzyme- and perforin-dependent cytotoxicity (2, 8). The nature of these Ags has been recently discovered by Kjer-Nielsen et al. (9) who showed that MR1 binds and presents small organic metabolite compounds derived from the vitamin B synthesis pathways (10). A number of intermediates of the folic acid (vitamin B_9) and riboflavin (vitamin B_2)

*Division of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, United Kingdom; [†]Department of Pulmonary and Critical Care Medicine, Oregon Health and Science University, Portland, OR 97239; [‡]Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom; and [§]QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4006, Australia

[1]Current address: Adaptimmune Ltd., Oxon, U.K.

[2]B.L. and A.L. contributed equally to this work.

ORCIDs: 0000-0002-9709-7624 (E.W.M.); 0000-0003-2505-3291 (M.C.G.); 0000-0003-3194-3135 (A.K.S.).

Received for publication June 22, 2015. Accepted for publication May 18, 2016.

This work was supported in part by United States Department of Veterans Affairs Biomedical Laboratory Research and Development Merit Review Awards I01 BX001231 (to M.C.G.) and I01 BX000533 (to D.M.L.) with resources and the use of facilities at the Veterans Affairs Portland Health Care System and by National Institutes of Health Grants R01 AI078965 (to M.C.G.), R01 AI048090 (D.M.L.), and T32 AI078903-05 and T32 HL83808-05 (to E.W.M.). A.K.S. is a Wellcome Trust Senior Investigator. B.L. was supported by Biotechnology and Biological Sciences Research Council Grant BB/H001085/1. A.L. and M.D.C. are Ph.D. students supported by the Medical Research Council and Health and Care Research Wales, respectively.

The sequencing data presented in this article have been submitted to the European Molecular Biology Laboratory/European Bioinformatics Institute European Nucleotide Archive (http://www.ebi.ac.uk/ena) under accession numbers PRJEB12991, ERS1078785, SAMEA3891651, ERS1078786, SAMEA3891652, ERX1378672, ERR1307049, ERX1378673, and ERR1307050.

Address correspondence and reprint requests to Dr. Bruno Laugel and Prof. Andrew K. Sewell, Division of Infection and Immunity, Cardiff University School of Medicine, Henry Wellcome Building, Heath Park, Cardiff CF14 4XN, U.K. E-mail addresses: LaugelBF@cardiff.ac.uk (B.L.) and SewellAK@cardiff.ac.uk (A.K.S.)

The online version of this article contains supplemental material.

Abbreviations used in this article: Cas9, clustered regularly interspaced palindromic repeat–associated protein 9; CRISPR, clustered regularly interspaced palindromic repeat; gRNA, guide RNA; MAIT, mucosal-associated invariant T; MFI, mean fluorescence intensity; MR1, MHC-related protein 1; PAM, proto-spacer–associated motif; sgRNA, single-guide RNA; WT, wild-type.

pathways act as ligands for MR1 (10, 11). However, only compounds derived from the riboflavin pathway, which is absent in mammals but present in microbes, were found to activate MAIT cells, therefore providing a molecular basis for the specific recognition of microbially infected cells (9). Our recent study showed that human MAIT cells isolated from a single individual use distinct TCR repertoires to recognize cells infected with different bacteria in an MR1-specific manner (12). Moreover, Gherardin et al. (13) have recently characterized the crystal structure and biophysical properties of TCRs from T cells with discrete Ag specificity for folate- or riboflavin-derived compounds presented by MR1. Remarkably, several of these MR1-restricted T cell clonotypes did not express the canonical MAIT TRAV1-2 TCR α-chain (13), indicating that non-MAIT αβ T cells are also able to recognize MR1 Ags. This TCR usage heterogeneity may provide a degree of specificity in MAIT- and MR1-restricted T cell activation and hints that different pathogens could generate MR1-restricted Ags of varied structure and chemical composition. In addition to MR1-restricted activation, MAIT cells respond to proinflammatory innate cytokines such as IL-12 and IL-18 (1, 14), which can act as autonomous stimuli or combine with TCR signals to potentiate MAIT cell activation (15). This Ag-independent activation process may be relevant to the pathogenesis of a number of inflammatory conditions in which the number, distribution, phenotype, and functions of MAIT cells were found to be altered (1, 16–18).

The biology of MR1-restricted T cells is a rapidly emerging field in immunology. The invariant nature of MR1 across the human population and its established role in the presentation of pathogen-derived Ags are of outstanding interest for the potential development of universal therapeutic and diagnostic tools in infectious diseases. MR1 expression also appears to be ubiquitous among different cells and tissues (19, 20), which may indicate that MR1-driven Ag responses are relevant to the pathogenesis of a broad number of immune-mediated diseases. However, the invariance and ubiquity of MR1 also complicate basic investigations of its ligand-binding and Ag presentation properties as well as in the understanding of MR1-restricted T cell biology. Indeed, the presence of MR1 on most APC lines and primary cells that also express other classical and nonclassical HLA molecules can make the unambiguous identification of microbe-specific MAIT cells and their distinction from conventional T cells that express the TRAV1-2 TCR chain problematic. Besides, solely relying on known MAIT cell phenotypic markers could result in the exclusion of previously undescribed bona fide MR1-restricted T cells. So far, confirmation of MR1 restriction has exclusively relied on the use of an anti-MR1 blocking mAb to abrogate activation (21), yet this type of functional validation can sometimes produce inconclusive experimental data because of incomplete blockade. In this context, the availability of isogenic model cell lines bearing MR1 knockout mutations would be a valuable tool for reverse genetics and loss of function studies.

The clustered regularly interspaced palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) genome editing platform has recently been developed as a powerful tool to induce targeted disruptive mutations in, or edit the DNA sequence of, a specific gene (22–28). Current CRISPR/Cas9 genome editing approaches rely on targeting the endonuclease activity of *Streptococcus pyogenes* or *Staphylococcus aureus* (29) Cas9 to a 17- to 23-nt-long genomic DNA sequence using a single-guide RNA (sgRNA) polynucleotide that acts both as a scaffold for Cas9 and as a DNA tethering agent. Cas9 induces DNA strand breaks near a so-called proto-spacer–associated motif (PAM) that are repaired either via nonhomologous end joining, often introducing insertions or deletions that disrupt the translational reading frame, or via homologous recombination using a repair template (25). CRISPR/Cas9 has been used to rapidly generate transgenic animals, correct deleterious mutations causing congenital diseases in cell lines and pluripotent stem cells, or identify and validate therapeutic targets. In this study, we report the generation of a novel versatile all-in-one CRISPR/Cas9 lentiviral vector and its use to derive isogenic clonal variants of carcinomic human alveolar basal epithelial A549 cells and of leukemic monocytic THP-1 cells commonly used to study the immune response to several intracellular bacteria, including *M. tuberculosis* and *Mycobacterium smegmatis* (2).

Materials and Methods

Generation of an all-in-one CRISPR/Cas9 lentivector

A synthetic polynucleotide sequence containing the canonical U6 RNA polymerase III promoter and an sgRNA sequence containing the *trans*-activating CRISPR RNA and the MR1-specific target sequence, based on the design of Mali et al. (24), was purchased from Eurofins MWG Operon (Ebersberg, Germany) and introduced in pCDNA.3-TOPO_wt-Cas9 (Addgene no. 41815) by nondirectional cloning at the unique Spe1 restriction site immediately upstream of the CMV promoter. To generate the all-in-one lentivector construct, the plasmid region containing the U6 promoter–sgRNA as well as the CMV promoter and Cas9 sequences was PCR amplified with primers containing Age1 and NsiI (pCDNA.3_Fwd and pCDNA.3_Rev) restriction sites from the pCDNA.3-TOPO_U6-sgRNA_wt-Cas9. This fragment was fused to a PCR amplicon consisting of the second generation pRRL.sin.cppt.pgk-gfp.wpre lentivector backbone developed by Didier Trono's laboratory (Addgene no. 12252) devoid of the human PGK promoter and GFP cDNA. This amplicon was obtained by PCR amplification with primers containing matching restriction sites (pRRL.0_Fwd and pRRL.0_Rev). The final sequence of the resulting pRRL.sin.CRISPR/Cas9 plasmid and all the primer sequences are provided in Supplemental Fig. 1 and Supplemental Table I, respectively.

Introduction of novel sgRNA target sequences

Repurposing of the CRISPR/Cas9 system was performed by introducing new 19-nt target sequences using a PCR cloning approach whereby the new sequence is fused to 5′-phosphate–modified primers complementary to plasmid sequences immediately flanking the target sequence (forward primer. 5′-GTTTTAGAGCTAGAAATAGCAAGTTAAx_9-3′, reverse primer, 5′-GGTGTTTCGTCCTTTCCx_{10}-3′, where x represents nucleotides of the target sequence). PCR amplification conditions were as follows: an initial 2-min denaturation step at 95˚C followed by 30 three-step cycles (30 s denaturation at 95˚C, 30 s annealing at 53˚C, 5 min elongation at 72˚C) and a final elongation step of 5 min at 72˚C using the Phusion high-fidelity DNA polymerase (Thermo Scientific, Cambridge, U.K.). The PCR product corresponding to the full-size plasmid (~11 kbp) was gel extracted, purified, and circularized by ligation with standard T4 DNA ligase (Promega, Southampton, U.K.). Ligation products were then transformed into *Escherichia coli* XL10 Gold (Stratagene, La Jolla, CA). Bacteria were plated onto Luria–Bertani agar containing 100 μg/ml carbenicillin (Biochemical Direct) following a 1-h recovery step at 37˚C in SOC medium (Invitrogen, Paisley, U.K.). After an overnight incubation at 37˚C, 10 colonies per new target sequence were tested by colony PCR to assess plasmid integrity and confirm the presence of the sgRNA region of interest (forward primer pLKO1.A; reverse primer gRNAcolPCR.R). PCR products of the expected size (450 bp) were purified, cleaned on DNA Clean and Concentrator columns (Zymo Research, Irvine, CA), and sent for sequencing. A total of five sgRNA target sequences were introduced for testing: one targeting exon 1, three targeting exon 2, and one specific for a site on exon 3 (Table I).

Cell culture and clonal expansion of $MR1^{-/-}$ cells

Adenocarcinoma HeLa (ATCC CCL-2), human acute monocytic leukemia THP-1 (ATCC TIB-202), lymphoblastic chronic myelogenous leukemic K-562 (ATCC CCL-243), and carcinomic alveolar basal epithelial A549 (ATCC CCL-185) cell lines were maintained in RPMI 1640 growth medium (Life Technologies, Paisley, U.K.) supplemented with 2 mM glutamine, 50 U/ml penicillin, and 50 μg/ml streptomycin (Life Technologies) and 10% FBS (Life Technologies), referred to as R10 hereon. Cells were passaged when reaching 80% confluence using enzyme-free cell-dissociation buffer HBSS (Life Technologies). Bulk A549 cells transfected with the CRISPR/

Table I. CRISPR/Cas9 sgRNA target sequences within the MR1 gene tested in this study

gRNA	Target Sequence (5′→3′)	PAM (5′→3′)	Exon	Strand
A	GGATGGGATCCGAAACGCCC	AGG	2	+
B	GGTGAAGCACAGCGATTCC	CGG	1	−
C	GTCCCTGAATTTATTTCGGT	TGG	2	−
D	GAACCTCGCGCCTGATCACT	GGG	2	−
E	GCAGTATGCATATGACGGGC	AGG	3	−

Target sequences were identified by searching for $GN_{19}GG$ DNA motifs on either strand of the MR1 cDNA sequence. We selected five nonoverlapping target sequences located in the first half of the MR1 coding sequence.

Cas9 plasmid were sorted by flow cytometry by gating on the $MR1^-$ population. $MR1^{-/-}$ cells were isolated from bulk THP-1 cells transduced with the CRISPR/Cas9 lentivector and subjected to MAIT cell positive selection. The selection was performed by using THP-1 cells infected with *M. smegmatis* at a 20:1 multiplicity of infection. Target cells were cocultured for 7 d with the MAIT cell clone D481A9 at an E:T ratio of 1:2. After the selection period, cells were transferred to flasks for a larger cell culture and further expanded. Subsequently, the resulting cells were clonally expanded by isolating single cells using a limiting dilution approach. An average of 0.3 cell was plated in individual wells of four flat-bottom 96-well plates containing 200 μl medium. A total of 16 clonal A549 and 20 THP-1 populations were then screened by flow cytometry. Three A549 and five THP-1 clones were ultimately selected and analyzed using a mismatch-specific endonuclease assay as described in *Results*. T cell clones were cultured in RPMI 1640 media supplemented with 2 mM glutamine, 50 U/ml penicillin, and 50 μg/ml streptomycin (Life Technologies), 10% FBS, 0.01 M HEPES buffer, nonessential amino acids, sodium pyruvate (Life Technologies), 25 ng/ml IL-15 (PeproTech, Rocky Hill, NJ), and either 20 or 200 IU/ml IL-2 (aldesleukin [Proleukin]; Prometheus, San Diego, CA), depending on the stage of culture.

MAIT and T cell clones activation

IFN-γ ELISPOT assays with the HLA-E–restricted T cell clone D160 1–23 were performed as described in Lewinsohn et al. (30). Briefly, A549 cells or $MR1^{-/-}$ A549 cells (10^4/well) were incubated with 5 mg/ml *M. tuberculosis* pronase-digested cell wall (31) in RPMI 1640 medium supplemented with 10% human serum for 2 h. D160 1–23 (10^4/well) was then added to each well. The plate was incubated overnight at 37°C and then developed using an AEC Vectastain kit (Vector Laboratories). PHA stimulation was used as a positive control for T cell signaling viability. Data shown are representative of three independent experiments. The D426B1 and D481 MAIT cell clones were washed with RPMI 1640 medium and then rested in RPMI 1640 supplemented with 50 U/ml penicillin and 50 μg/ml streptomycin, 2 mM L-glutamine, and 5% FBS overnight. A549 cells (wild-type [WT] and MR1 knockouts) were cultured in antibiotic-free R10 overnight. A549 and THP-1 cells were exposed to *M. smegmatis* at a multiplicity of infection of 100:1 (bacteria to cells) for 2 h in antibiotic-free R10, followed by 2 h incubation with R10 containing 50 U/ml penicillin and 50 μg/ml streptomycin (Life Technologies). A549 and THP-1 cells were washed to remove extracellular *M. smegmatis*. Control cells were mock treated as if they had been coincubated with bacteria. A549 and THP-1 cells were then plated into 96U-well plates at a density of 6×10^4 cells per well. MAIT cells (3×10^4) were then added to each well and cultured overnight. Supernatant (60 μl) was then harvested, diluted to 120 μl with R5, and assayed for MIP-1β, TNF-α, and IFN-γ ELISA (R&D Systems) according to the manufacturer's instructions. Data were plotted and analyzed with GraphPad Prism software version 5.03. Activation of the $CD8^+$ T cell clone B9 specific for an octameric peptide (SELEIKRY) derived from the EBV BZLF1 protein and presented by HLA-B*1801 (32) was done by coculturing T cells with A549 WT or $MR1^{-/-}$ cells overnight in 200 μl R5. A549 cells (1×10^6) were pulsed with each peptide concentration in 100 μl R5 for 2 h at 37°C. The cells were washed with 10 ml R5 five times, aspirating all excess supernatant between washes. The A549 cells were counted and resuspended at 6×10^5 cells/ml in a final volume of 700 μl. Then, 100 μl cell suspension (i.e., 60,000 cells) was plated into 96-well plates. T cells (3×10^4) in 100 μl R5 were added into each well. Then, the remaining APCs were centrifuged and the supernatant was plated with T cells as a control to ensure that the T cell response was mediated through APC presentation and not T:T presentation. Each assay condition was performed in triplicate. For ELISAs, 50 μl supernatant per well was used for TNF-α and IFN-γ ELISA and 25 μl supernatant diluted with 25 μl of R5 per well was used for MIP-1β. Supernatants were frozen and assayed for MIP-1β and IFN-γ in a sandwich ELISA assay (R&D Systems) according to the manufacturer's instructions.

Flow cytometry

The PE-conjugated anti-MR1 Ab clone 26.5 (BioLegend, London, U.K.) was used to stain HeLa-MR1, A549, and THP-1 cells at the dilution recommended by the manufacturer. To stabilize the cell-surface levels of MR1, A549 cells were incubated overnight in the presence of 50 μg/ml acetyl-6-formylpterine (Schircks Laboratories, Jona, Switzerland) prior to staining (11). Where indicated, the W6/32 monoclonal anti–HLA-ABC Ab conjugated to APC (eBioscience, Hatfield, U.K.) was also used to monitor classical HLA-I levels on the cell surface. 7-Aminoactinomycin D (BD Biosciences, Oxford, U.K.) or the viability Vivid Dye (Molecular Probes, Life Technologies) was added to all the staining before data acquisition on a FACSCalibur instrument (Becton-Dickinson, UK) or a FACSCanto II (Becton Dickinson, Oxford, U.K.). Data analysis was performed with FlowJo software (Tree Star, Ashland, OR) by drawing a population gate on the forward scatter area/side scatter area plot and on viable cells (7-aminoactinomycin D^- or $Vivid^-$).

Monitoring of mutations at the MR1 locus

Genomic DNA from A549 and THP-1 cells was isolated with the GenElute mammalian genomic DNA miniprep kit (Sigma-Aldrich, Gillingham, U.K.). Mutations at the target site were detected using the CEL-I enzyme, as part of the Surveyor assay (Transgenomic, Glasgow, U.K.), which cleaves DNA duplexes bearing base pair mismatches, caused by insertions or deletions at proximity of the PAM sequence, within the PCR amplicons generated with primers flanking the genomic target site. The PCR forward primer (SURV1_Fwd) is located in the intron region upstream of the target site, and the reverse primer (SURV1_Rev) is located downstream on exon 2. The predicted size of the full-length PCR product was 852 bp, and the expected position of Cas9 cleavage (immediately downstream of the PAM sequence) is located 521 bp downstream of the start of the forward primer and 331 bp from the reverse primer. In the case of HeLa-MR1 cells, which overexpress MR1 cDNA, a PCR amplicon flanking the sgRNA target sequences was obtained by amplifying a 360-bp region from template DNA isolated as described above using a forward primer complementary to a plasmid region immediately upstream of the MR1 cDNA (SURV2_Fwd) and a reverse primer downstream of the target site (SURV2_Rev). To sequence the modified MR1 locus in the A549 clonal derivatives clone 9 and clone 11, the same primers fused to the type IIs endonuclease BsaI were used to amplify the genomic DNA and clone the amplicons into a cloning vector (Addgene no. 32189). Plasmid minipreps from 10 colonies obtained from the resulting transformation for each clone were sent for Sanger sequencing at Eurofins MWG Operon.

Identification of CRISPR/Cas9 mutagenesis off-target effects using genomics

Undertaking a BLAST search against the human genome refseq database using the guide RNA (gRNA) sequence, we identified six sites (excluding MR1) located within genes that contained sequence that differed at four or fewer sites compared with the gRNA. Having identified these genes, we undertook whole-genome sequence of two samples, clone 9 and clone 11, to assess whether these genes contained off-target effects. To sequence the sample, we extracted the genomic DNA from the two A549 clones using the GenElute mammalian genomic DNA miniprep kit (Sigma-Aldrich). We fragmented 1 μg DNA to an average of 300 bp fragments by sonication and prepared libraries using NEBNext ultra library preparation kits (New England Biolabs, Herts, U.K.). The libraries were sequenced using a NextSeq 500, running high output 150-bp paired end sequencing to a depth of >20× coverage for each genome. Taking the sequence reads generated, we mapped these against the gene sequences of the target sites, and in five of the six cases using the Burrows–Wheeler alignment tool (33), visualizing the mapping results using Artemis (34). To identify inserts that are larger than those that could be detected by mapping, we screened all of the

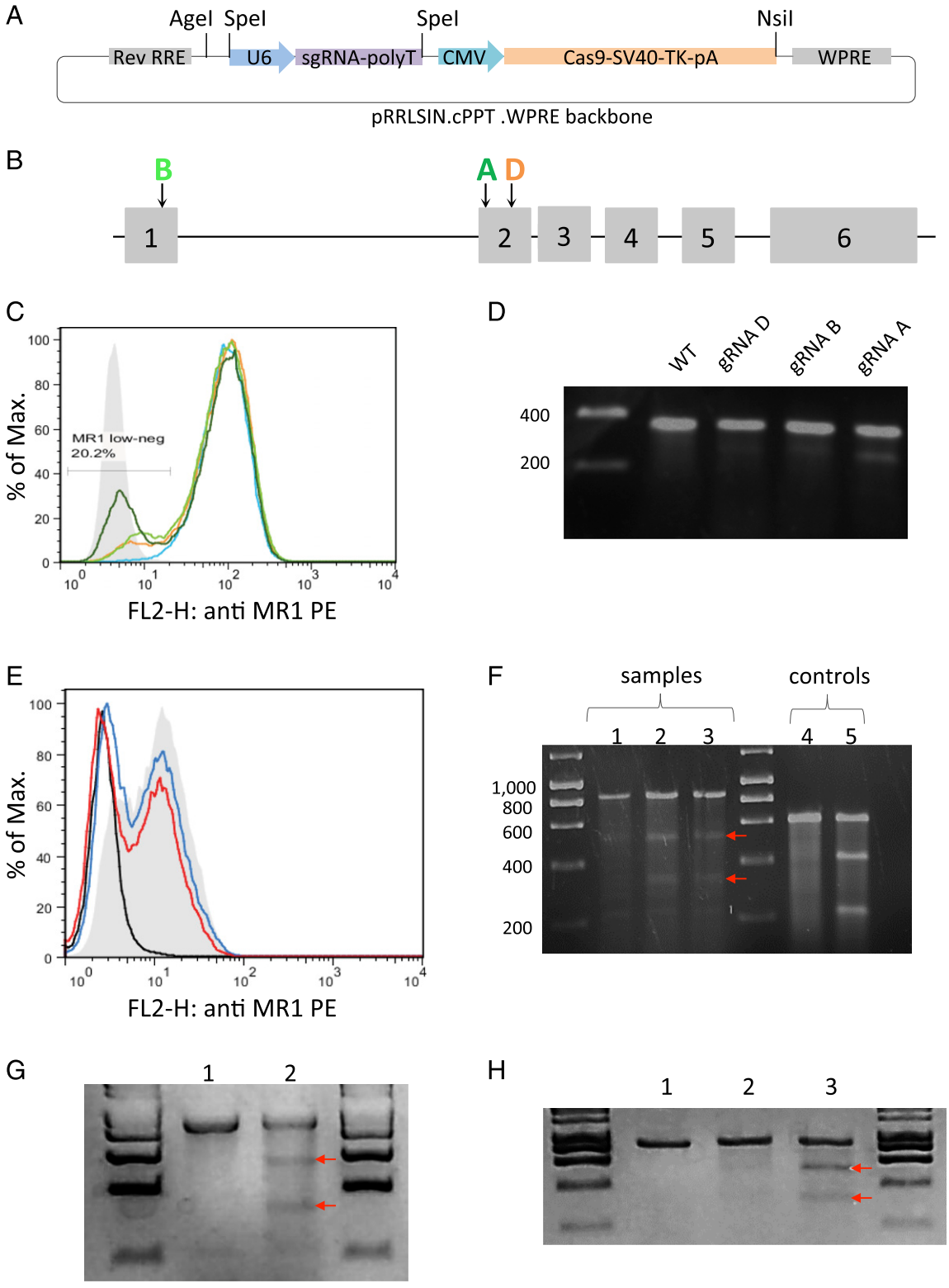


FIGURE 1. Description of the lentiviral CRISPR/Cas9 system and identification of gRNA target sequences efficiently disrupting MR1 gene expression. (**A**) Graphic representation of the all-in-one lentiviral system generated for this study. The CRISPR/Cas9 elements derived from the pCDNA.3-TOPO_wt-Cas9 plasmid were incorporated within the second generation pRRLSIN.cPPT.WPRE lentivector backbone between the Age1 and Nsi1 unique restriction sites. (**B**) Positioning of three active sgRNA target sequences with respect to the intron/exon structure of the MR1 gene. (**C**) Disruption of MR1 protein surface expression on HeLa reporter cells overexpressing MR1 transduced with lentiviruses expressing Cas9 in concert with a control gRNA (light blue histogram), MR1 gRNAs A (dark green histogram), B (light green histogram), or D (orange histogram). Isotype control staining is shown as a solid light gray histogram. The histogram plot shown is representative of four independent staining experiments. (**D**) Monitoring of mutations within the genomic DNA of transduced HeLa cells. MR1 cDNA PCR amplicons generated with the genomic DNA of WT cells or of cells transduced with lentiviruses expressing each sgRNA target were denatured, annealed, and digested with the CEL-I enzyme as described in *Materials and Methods*. Four hundred (WT lane) or 200 ng of PCR products from unmodified HeLa-MR1 cells mixed with 200 ng amplicon generated with DNA obtained from cells transduced with each sgRNA target (gRNA D, B, and A lanes) were used. Digestion of PCR amplicons was only obvious for the WT/gRNA A (*Figure legend continues*)

reads generated using the sequence 20 bp on the 5′ side of each of the off-target sites identified above. Extracting these reads, we identified any cases where larger deletions were present within these reads by manually examining the sequence of the reads by eye against the expected reference sequence. Following this we also undertook an analysis to identify other potential off-target sites that may not have been in genes, or detected by our screen of the human genome. To do this, we identified all of the sequence reads that contained a sequence matching, or complementary to, the gRNA at 15 of 20 bases, and extracted these into a set of fastq files. We then assembled these sequences using Velvet and extracted all contigs with a sequence length of >200 bp. In total, this analysis identified 7574 locations in sample clone 9 and 7516 locations in clone 11 for further investigation. Extracting the BLAST alignments for each of the sequences queried, we examined, by eye, all of the alignments containing insertions within the contig for evidence of insertions within 100 bp of the match to the gRNA. In total, other than the known deletion within MR1 and the detected RP11-46A10.6 site, we found no other cases where there appeared to have been an off-target effect in either of our samples. Sequencing data were submitted to the EMBL-EBI European Nucleotide Archive (http://www.ebi.ac.uk/ena). The study accession number is PRJEB12991. Sample accession and secondary accession numbers for A549 clone 9 are ERS1078785 and SAMEA3891651. Sample accession and secondary accession numbers for A549 clone 11 are ERS1078786 and SAMEA3891652. Experiment accession and run accession numbers for A549 clone 9 are ERX1378672 and ERR1307049. Experiment accession and run accession numbers for A549 clone 11 are ERX1378673 and ERR1307050.

Production of lentiviral particles and cell transduction

The pRRL.sin.CRISPR/Cas9 lentivector was cotransfected with the second generation VSV.G envelope pMD2.G (Addgene no. 12259) and packaging pCMV-dR8.74 (Addgene no. 22036) plasmids in HEK-293T cells by $CaCl_2$ transfection in six-well plates. Lentiviral supernatants (~10 ml total volume) were collected at 24 and 48 h, filtered using a 0.45-μm cellulose acetate syringe filter, concentrated by centrifugation at 28,000 rpm for 2 h, and resuspended in 0.5 ml R10. HeLa-MR1, THP-1, K562, or A549 cells grown to ~50% confluence in six-well plates were transduced with the concentrated viral supernatant diluted 3-fold in a final volume of 3 ml.

Results

Development of a versatile all-in-one lentiviral CRISPR/Cas9 system

We chose to adapt the two plasmid delivery platform described by Mali et al. (24), whereby the Cas9 and sgRNA elements are expressed on separate plasmids, to generate a single delivery system embedded within a second generation lentivector backbone. We reasoned that such a tool would combine the ease of an all-in-one delivery method with the high infectivity and broad cellular tropism of lentiviruses. For this purpose, we introduced a codon-optimized version of *S. pyogenes* Cas9 under the control of the CMV promoter as well as a U6 RNA polymerase III promoter sgRNA complex within the self-inactivating second generation lentivirus pRRL.sin.cppt.pgk-gfp.wpre (Fig. 1A), a well-characterized lentiviral system that produces high-titer viral particles. We designed the sgRNA so that the 19-nt target sequence could be easily swapped through a cloning PCR approach, a feature that distinguishes our lentivector from other available CRISPR/Cas9 systems.

Generation of cell populations bearing CRISPR/Cas9-induced mutations in the MR1 locus

A total of five lentivectors, each containing a different CRISPR target sequence specific for the MR1 gene (Fig. 1B, Table I), were generated and tested in a reporter system consisting of HeLa cells overexpressing native human MR1 cDNA. $MR1^{low/-}$ cells could be identified by flow cytometry within HeLa-MR1 cell populations transduced with three of the five lentiviruses (Fig. 1C). gRNA.A showed the highest disruption efficiency both at the protein and DNA levels (Fig. 1C, 1D). Based on these results, we selected gRNA.A to edit the genome of the A549 cell line. We compared MR1 gene modifications induced in A549 cells in which Cas9 and the MR1 gRNA.A were either stably or transiently expressed using lentiviral transduction or plasmid DNA transfection, respectively. Flow cytometry analysis showed that for both transient and stable CRISPR/Cas9 expression, the overall mean fluorescence intensities (MFIs) with anti-MR1 Ab were decreased and the proportion of cells appearing MR1 low or negative was increased compared with the control WT A549s (Fig. 1E). At the DNA level, the Surveyor assay revealed that DNA mismatches were introduced at the expected positions, as digestion of heteroduplex PCR amplicons containing the CRISPR/Cas9 target sites from WT and modified cells yielded two digestion products of the expected sizes (Fig. 1F). We also used our CRISPR/Cas9 system to disrupt the endogenous MR1 genes of the THP-1 and K562 cell lines. Transient transfection of K562 and THP-1 cells with CRISPR/Cas9 plasmid DNA did not allow MR1 gene editing with high enough frequencies for monitoring with the Surveyor assay. However, CRISPR/Cas9 lentivirus transduction successfully generated MR1 indels in both cell lines (Fig. 1G, 1H). In the case of THP-1, we used an enrichment strategy based on the positive selection of MR1 mutants able to escape killing by a MAIT cell clone. Following 7 d of coculture with MAIT cells at an E:T ratio of 1:2, we observed an increase in the frequency of mutations in the MR1 gene within the expanding THP-1 population, suggesting that the process allowed for the selection of disruptive MR1 mutations (Fig. 1H).

Generation of $MR1^{-/-}$ A549 clones: phenotypic and genotypic characterization

Although CRISPR/Cas9 gene disruption by lentiviral transduction appeared more efficient than transient expression, we reasoned that continuous expression of the gene-editing elements from a stable, integrated vector within the cells was more likely to generate detrimental off-target DNA cleavage, and we therefore elected to

heteroduplexes (lane gRNA A). Data shown are representative of two independent experiments. (**E**) A549 cells were lipofected with plasmid DNA containing the CRISPR/Cas9 elements (including the gRNA A target sequence) or transduced with lentiviral particles made from the same plasmid. Eight days after transduction/lipofection, cells were treated overnight with acetyl-6-formylpterine (50 μg/ml) prior to flow cytometry analysis. A549 cells stained with an isotype control Ab are shown as a black histogram (MFI = 2.78), mock lipofected cells as a solid gray histogram (MFI = 9.53), cells transduced with lentivirus as a red histogram (MFI = 5.84), and those lipofected with the CRISPR/Cas9 plasmid DNA as a blue histogram (MFI = 7.07). The data shown are representative of three stainings. (**F**) Surveyor assay performed with a total of 500 ng PCR amplicon obtained from genomic DNA of unmodified A549 cells or from cells lipofected or transduced with gRNA.A CRISPR/Cas9. Homoduplex WT DNA (*lane 1*) or heteroduplexes of annealed WT and modified amplicons from transduced cells (*lane 2*) or lipofected cells (*lane 3*) were digested with the CEL-I DNA-mismatch–specific enzyme. Two bands matching the sizes of the expected digestion products (521 and 331 bp) can be identified (indicated with red arrows). Undigested (*lane 4*) and digested (*lane 5*) assay controls are shown. The gel shown is representative of at least five different experiments. (**G**) Surveyor assay performed with a total of 500 ng PCR amplicon obtained from genomic DNA of unmodified K562 cells or from cells transduced with gRNA.A CRISPR/Cas9. Homoduplex WT DNA (*lane 1*) or heteroduplexes of annealed WT and modified amplicons from transduced cells (*lane 2*) were digested with the CEL-I DNA-mismatch–specific enzyme. Two bands matching the sizes of the expected digestion products (521 and 331 bp) can be identified (indicated with red arrows). (**H**) Surveyor assay performed with a total of 500 ng PCR amplicon obtained from genomic DNA of unmodified THP-1 cells or from cells transduced with gRNA.A CRISPR/Cas9. Homoduplex WT DNA (*lane 1*) or heteroduplexes of annealed WT and modified amplicons from transduced cells before and after MAIT cell selection (*lanes 2* and *3*) were digested with the CEL-I DNA-mismatch–specific enzyme. Two bands matching the sizes of the expected digestion products (521 and 331 bp) can be identified (indicated with red arrows). The data shown in Fig. 1G and 1H are representative of two independent experiments.

use bulk-transfected A549 cells to derive MR1$^{-/-}$ clonal populations by limiting dilution. Following a first screen by flow cytometry of 16 clonal A549 derivatives (not shown), we selected 8 candidates showing homogeneous cell surface MR1 protein expression. The putative clones 1, 9, and 11 consistently displayed the lowest MFIs (Fig. 2A) and were taken forward for further molecular analysis of their genomic DNA. Interestingly, PCR amplification of the region flanking the sgRNA.A target sequence yielded two products in the case of clone 9 (Fig. 2B). In addition to a PCR amplicon of the predicted size (852 bp) common to all samples, a smaller fragment, shorter by 100–150 bp, was amplified, likely indicating a deletion near the site of Cas9 cleavage. Digestion of DNA heteroduplexes with the CEL-I enzyme generated additional bands matching digestion product sizes predicted by the positioning of the gRNA.A target and PAM sequences within the amplicon (~520 and 330 bp) for clones 9 and 11 but not clone 1 (Fig. 2B). In the case of clone 9, further characterization of PCR products by Sanger sequencing revealed a 126-nt deletion spanning the intron/exon junction and resulting in the partial deletion of the CRISPR/Cas9 target sequence on one allele (Fig. 2C). The other allele bore a single nucleotide deletion as well as two base substitutions compared with the MR1 reference sequence (Fig. 2C), in line with the PCR analysis. Although it is unclear what the exact consequence

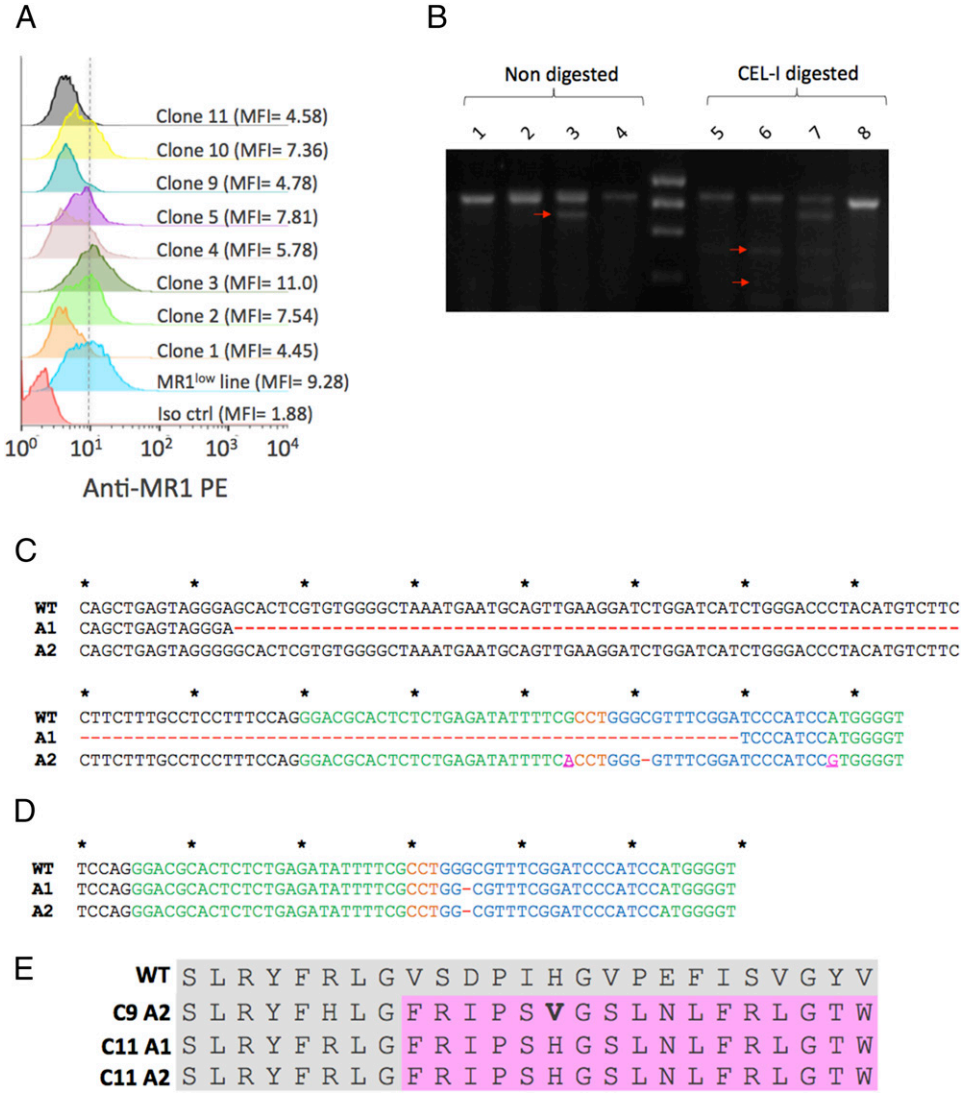


FIGURE 2. Characterization of A549 clonal derivatives bearing biallelic disruptive mutations in the MR1 gene. (**A**) Flow cytometry analysis of MR1 expression in eight different A549 clones obtained by limiting dilution and comparison with the bulk-transfected A549 parental cell line. Cells were treated overnight with acetyl-6-formylpterin (50 μg/ml) prior to staining with a PE-conjugated anti-MR1 Ab and data acquisition. The histogram plot shown is representative of three experiments. (**B**) Molecular characterization of mutations at the MR1 locus target site. PCR amplicons of WT A549 cells (*lane 1*) and A549 clones 11 (*lane 2*), 9 (*lane 3*), and 1 (*lane 4*) are shown in the absence of hybridization and CEL-I digestion. The Surveyor assay was performed with a total of 500 ng PCR amplicon obtained from genomic DNA of unmodified A549 cells or from the three different clones. Homoduplex WT DNA (*lane 5*) or heteroduplexes of annealed WT and modified amplicons from clone 11 (*lane 6*), clone 9 (*lane 7*), or clone 1 (*lane 8*) were digested with the CEL-I DNA-mismatch–specific enzyme. The gel shown is representative of at least four independent experiments. (**C**) Analysis of the MR1 genomic sequences recovered from A549 clone 9. The reference WT sequence is shown on top and the sequences recovered from clone corresponding to each allele are shown underneath. (**D**) WT MR1 reference sequence (*top*) and sequences corresponding to each allele of A549 clone 11 are shown. Nucleotide deletions are shown as red hyphens, and substitutions are shown as purple characters. Nucleotides corresponding to the sequence of the exon 2 MR1 are shown as green characters, the 19-nt gRNA.A target sequence is shown in blue characters, and the PAM sequence is shown in orange. (**E**) Predicted primary structures of the mutant MR1 proteins as determined by translating PCR amplicon DNA sequences. MR1 amino acids 26–50 are shown as alignments. The WT MR1 protein sequence is indicated as gray shadows and out-of-frame reads are shown as pink shadows. C9A2, clone 9 allele 2; C11A1, clone 11 allele 1; C11A2, clone 11 allele 2.

the large deletion has on MR1 protein structure, the latter mutation likely disrupts the protein reading frame from amino acid 34 onward (Fig. 2E). In the case of clone 11, an identical single nucleotide deletion within the sgRNA target sequence could be detected on each allele (Fig. 2D). The reading frame on both alleles of clone 11's coding sequence therefore shows the same frame shift as allele 2 of clone 9, where the native MR1 protein sequence is disrupted from amino acid 34 onward (Fig. 2E). Collectively, the lack of MR1 staining by flow cytometry and the presence of disruptive mutations on each allele for both clones indicated that MR1 expression was almost certainly abrogated as a result of CRISPR/Cas9 mutagenesis.

MR1 CRISPR/Cas9 mutagenesis resulted in an unintended nucleotide deletion in the RP11-46A10.6 pseudogene

To establish whether there were off-target effects resulting from the CRISPR/Cas9 modification, we undertook whole-genome sequencing for A549 clones 9 and 11. We identified likely off-target sites that occurred within genes (defined as sites varying at up to five bases compared with the MR1gRNA target sequence), totaling six possible genes where off-target effects could have been observed. Mapping the whole-genome sequence data back to the reference sequences for these genes, and for MR1, we identified only one case of unintended mutation. This was in a pseudogene bearing a high degree of sequence homology (94%) with the targeted MR1 region (Fig. 3A) and containing the exact same gRNA target sequence (Fig. 3B). In both samples the same nucleotide deletion at position 16 of the gRNA sequence was observed at this locus as on both clone 11 MR1 loci and on one MR1 locus of clone 9 (Figs. 2C, 2D, 3C). Another coding gene (STX6) sits within the RP11-46A10.6 region. However, the exons making up the STX6 protein–coding sequence are a considerable distance from the RP11-46A10.6 target sequence (Fig. 3A). We also examined our sequence reads to identify any other off-target effects that had longer deletions. Other than the one identified in MR1 for clone 9, we observed no evidence for other long deletions in the reads generated as part of the whole-genome sequencing (not shown).

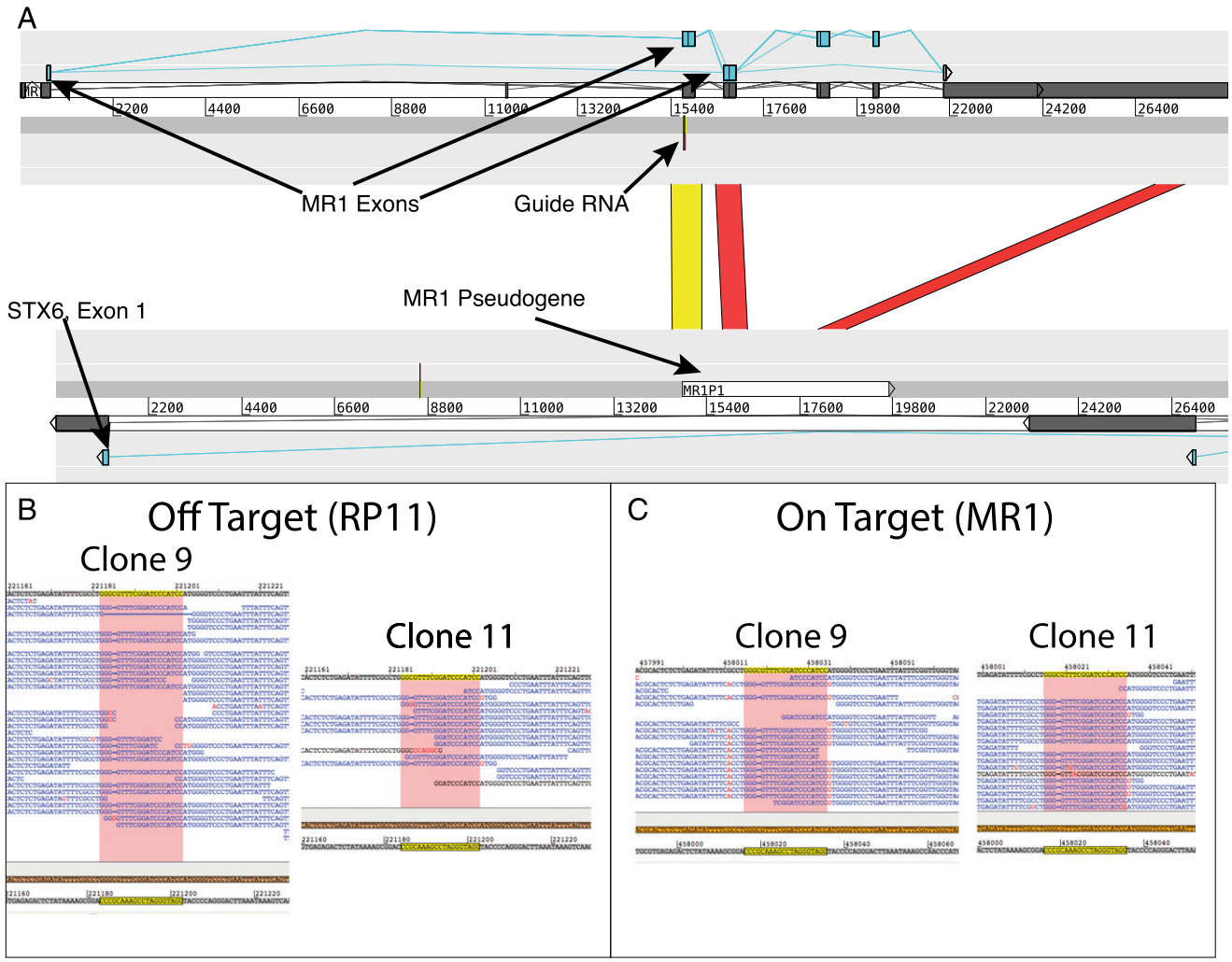


FIGURE 3. Results of whole-genome sequencing to identify off-target effects. (**A**) Figure generated from the Artemis comparison tool comparing the DNA sequence of MR1 to the RP11-46A10.6 pseudogene (located between exon 1 and 2 of STX6 gene). The red ribbons indicate large contiguous regions of 90%+ homology between the two genes, with the yellow ribbon being the region that contains the sequence targeted by the gRNA. The exons of MR1, STX6, and the RP11-46A10.6 pseudogene are annotated in this region and are shown in turquoise, dark gray, and white, respectively. (**B**) Unintended on-target effects within RP11-46A10.6 in samples clone 9 and clone 11. Two potential effects appear to be evident: one single base deletion seen in both clone 9 and clone 11, and one longer deletion in clone 9 (third read down from the top). The gRNA sequence is highlighted in yellow, and SNPs relative to the reference are shown in red characters. (**C**) On-target effects on MR1 in samples clone 9 and clone 11. The single base deletion observed in both clones, which is identical to that in RP11-46A10.6 relative to the gRNA target sequence, is evident for the vast majority of reads. The gRNA sequence is highlighted in yellow, and single nucleotide polymorphisms relative to the reference are shown in red characters.

A549 mutant clones infected with bacteria selectively fail to activate MAIT cell clones

Next we sought to confirm MR1 deficiency at the functional level. WT A549 cells or clones 9 and 11 infected with *M. smegmatis* were cocultured with the MAIT cell clones D426B1 and D481A9 known to express canonical TRAV1-2_TRAJ33 MAIT TCR α-chain rearrangement differing in their CDR3 sequences and two distinct β-chains (TRBV6-4 for D426B1 and TRBV20-1 for D481A9) (2, 12). In contrast to infected WT A549s, which triggered efficient MIP-1β and IFN-γ release, MR1-deficient clones 9 and 11 failed to activate either MAIT cell clone (Fig. 4). These results show that both A549 clonal derivatives are unable to present bacterial Ags in the context of cell-surface MR1. Such a distinct phenotype implies that CRISPR/Cas9 mutagenesis induced biallelic disruptive mutations, resulting in loss of MR1 protein function in both A549 clones.

MR1 mutations do not affect the expression of nonclassical or classical HLA molecules

To further assess the genetic integrity of both mutant A549 clones, we performed additional phenotypic and functional experiments. These established that cell-surface expression levels of classical HLA-I molecules were not affected in MR1-deficient cells (Fig. 5A, 5B). Moreover, both MR1$^{-/-}$ A549 clones were able to present HLA-B*1801–restricted peptides derived from the EBV protein BZLF1 to a CD8$^+$ T cell clone with efficiencies similar to the WT A549 cells, as evidenced by equivalent cytokine release in peptide titration experiments (Fig. 5C, 5D). Finally, when cocultured with the HLA-E–restricted CD8$^+$ T cell clone D160 1–23, the A549 clones 9 and 11 presenting *M. tuberculosis* cell wall fractions triggered similar IFN-γ secretion levels compared with the WT A549 cells (Fig. 5E, 5F). Collectively, these data suggest that classical and nonclassical HLA-I Ag presentation is intact in the MR1-deficient cells.

Generation and functional characterization of MR1$^{-/-}$ THP-1 clones

The monocytic THP-1 cell line is a broadly used and well-characterized tool to study bacterial phagocytosis and Ag presentation (35). The availability of MR1$^{-/-}$ THP-1 cells would likely be of interest to study MAIT cell responses to a range of pathogens. Following the selection of bulk-transduced THP-1 cells bearing increased MR1 mutations that conferred resistance to MAIT cell killing (Fig. 1H), we sought to derive clonal populations by single-cell dilution and assessed their MR1 phenotype. MR1 protein expression, as measured by flow cytometry, was markedly reduced in bulk THP-1 cells that underwent the positive selection process compared with the initial transduced preselection population (Fig. 6A), confirming the mismatch nuclease assay (Figs. 1H, 6C). All 20 tested THP-1 derivatives isolated by single-cell dilution of the selected population showed MR1 expression levels similar to those of the negative control used in the assay (Fig. 6B). Moreover, CEL-I digestion of five of these cell populations pointed to the presence of mismatch mutations at high frequencies within MR1 (Fig. 6C). Taken together, these data suggested that many, if not most, THP-1 derivatives bore biallelic disruptive MR1 mutations. Next we tested the ability of THP-1 cells to activate MAIT cells following *M. smegmatis* infection. Similar to the A549 clone 9 and clone 11, all the tested THP-1 derivatives failed to activate both D426B1 and D481A9 MAIT cell clones, as measured in cytokine release assays (Fig. 7), whereas the infected parental THP-1 cells induced the secretion of high levels of both TNF-α and IFN-γ by

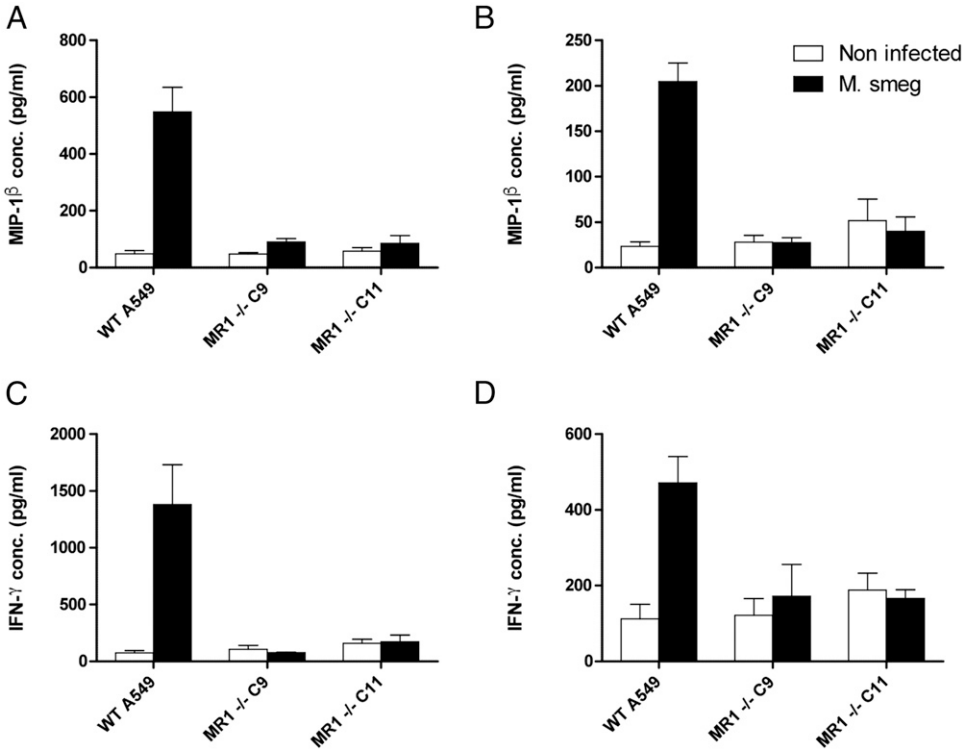


FIGURE 4. MAIT cells fail to recognize A549 derivatives clone 9 and clone 11 infected with *M. smegmatis*. The D481A9 and 426B1 MAIT cell clones were cocultured with the indicated A549 APCs infected with *M. smegmatis* or not as described in *Materials and Methods*. Supernatants were collected following overnight incubation. MIP-1β produced by MAIT cell clones D481A9 (**A**) and D426B1 (**B**), respectively, were quantified by ELISA. IFN-γ levels were also measured for both D481A9 (**C**) and D426B1 (**D**). Assays were carried out in triplicate wells. Means ± SEM are shown on the graph using data representative of three experiments.

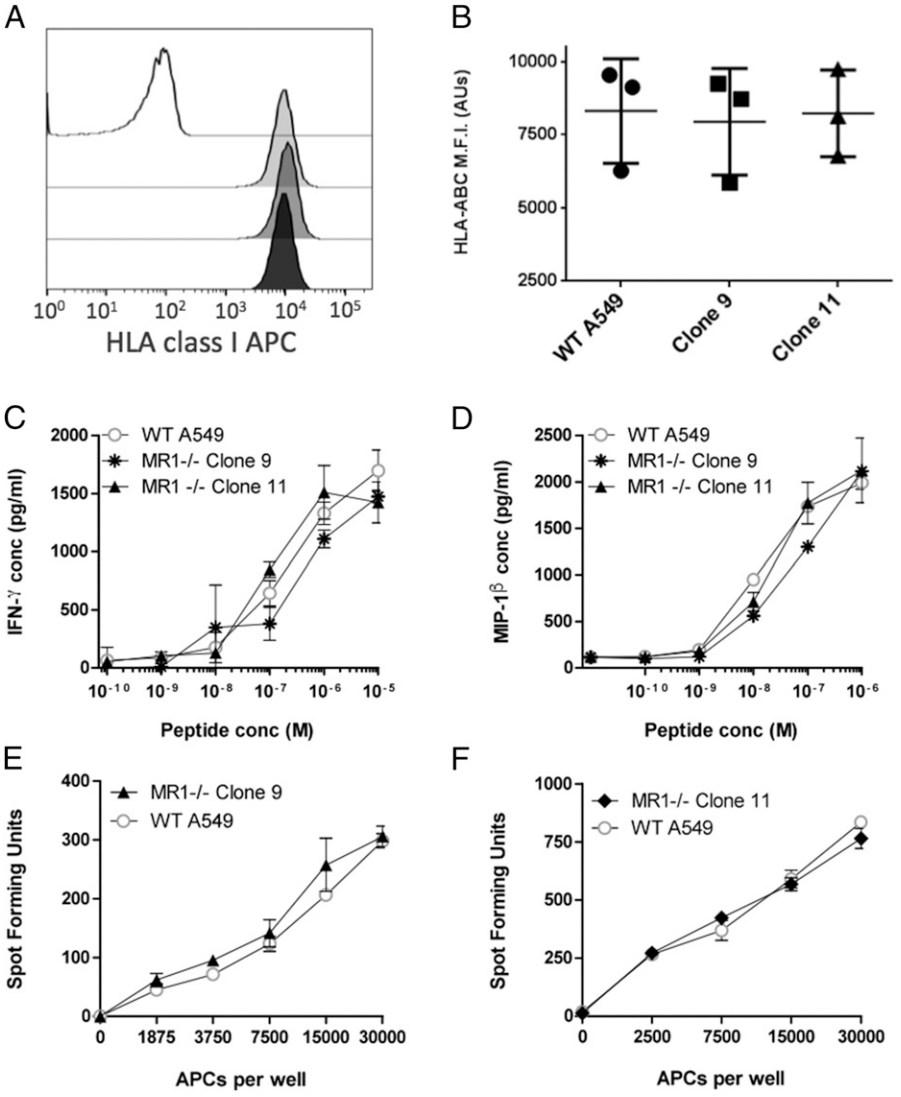


FIGURE 5. MR1 disruption does not affect HLA class Ia and HLA-E Ag presentation. (**A**) Representative cell surface HLA-I expression by A549 cells. MFI values were 56.4 (isotype control), 8733 (WT A549), 9880 (clone 11), and 8682 (clone 9). (**B**) MFI from three surface HLA-I staining of A549 WT and MR1$^{-/-}$ cells. Individual MFI datum points as well as mean values ± SEM are shown. The data shown are representative of two independent staining experiments. (**C** and **D**) Secretion of IFN-γ and MIP-1β cytokines by CD8$^+$ T cell clones B9 in response to increasing concentrations of an agonist peptide (SELEIKRY) derived from the EBV BZLF1 protein and presented by HLA-B*1801 expressed on the surface of WT A549, clone 9, or clone 11 as indicated in the key. (**E**) Activation of the D160 1–23 HLA-E–restricted CD8$^+$ T cell clone in response to increasing numbers of WT A549 or A549 MR1$^{-/-}$ clone 9 cells exposed to pronase-digested *M. tuberculosis* cell wall extracts. (**F**) Activation of the D160 1–23 HLA-E–restricted CD8$^+$ T cell clone measured as spot-forming units in response to increasing numbers of WT A549 or A549 MR1$^{-/-}$ clone 11 exposed to pronase-digested *M. tuberculosis* cell wall extracts in ELISPOT assays. The data shown are representative of two independent experiments.

D481A9 (Fig. 7A, 7C) by D426B1 (Fig. 7B, 7D), thereby demonstrating that these cells are unable to present MR1-restricted Ags and confirming their knockout status.

Discussion

RNA-guided genome engineering with CRISPR/Cas9 is a versatile and efficient tool to manipulate the genome of cell lines, including somatic and germinal cells. We have adapted this technology in a novel all-in-one lentiviral vector for the simultaneous delivery of the Cas9 endonuclease and an associated sgRNA. We designed sgRNA target sequences specific for the human MR1 locus and, following an initial screen in a reporter system, selected a target sequence we used to efficiently disrupt the expression of endogenous MR1 in three cell lines. We subsequently expanded clonal A549 and THP-1 populations bearing disruptive biallelic MR1 mutations. T cell assays where MAIT cells were activated by the parental A549 and THP-1 cell lines but failed to recognize cells obtained by limiting dilution upon infection with *M. smegmatis* confirmed MR1 protein loss of function. Importantly, CRISPR/Cas9-directed mutagenesis in these cells did not alter Ag processing and presentation by classical and nonclassical HLAs, highlighting the specificity of the genome editing approach and validating the use of these two clones to accurately characterize MR1-restricted Ags and their recognition by MAIT cells.

The two MR1-deficient A549 clones characterized in this study were generated from transient expression of an sgRNA and Cas9 elements in the parental cell line. The expression level of MR1 on the surface of A549 cells is relatively weak, and monitoring by flow cytometry, even after stabilization with high acetyl-6-formylpterine concentrations, did not allow us to unambiguously identify a population of MR1$^-$ cells in the bulk-transfected population (Fig. 1E). However, the CEL-I mismatch-specific nuclease assay clearly showed that mutations were introduced at the intended site within the MR1 gene (Fig. 1F). The level of MR1 disruption was increased following flow cytometry cell sorting of an MR1$^{-/low}$ population (not shown) from which we isolated and expanded individual clones by limiting dilution. A screen of 16 clonal populations identified eight derivatives that exhibited MR1 MFIs consistently lower than those of the parental cells (Fig. 2A). Three of these clones appeared MR1$^-$ by flow cytometry and were taken forward for molecular analysis of genomic DNA. Disruptions at the MR1 locus could be identified in two clones using the CEL-I assay, and Sanger sequencing confirmed the presence of frameshift mutations on both alleles for each clone (Fig. 2C). Overall 2 of 16 of the clones (12.5%) displayed the genotype of interest. This suggested that extending the induction of disruptive MR1 mutations in other cell lines and, possibly, in primary cells using this system should be possible. We confirmed this by applying

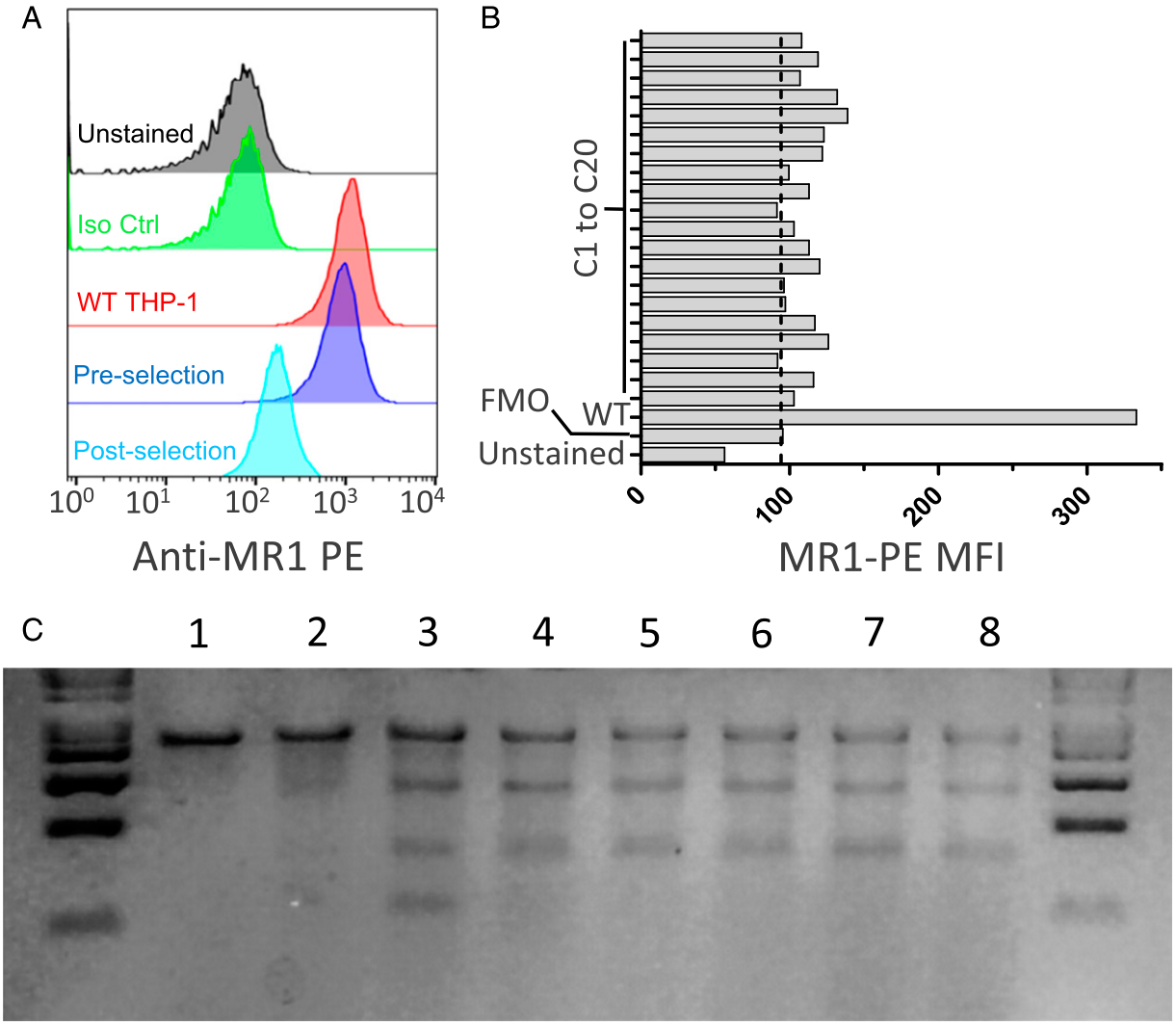


FIGURE 6. Isolation and characterization of THP-1 clonal derivatives bearing disruptive mutations in the MR1 gene. (**A**) Flow cytometry analysis of MR1 expression in THP-1 cells transduced with MR1 CRISPR/Cas9 lentiviral particles before and after enrichment of MR1-deficient cells by selection with the 426B1 MAIT cell clone, as indicated in the key to the figure. MFI values were: unstained, 45.5; isotype control, 53.9; WT THP-1, 1033; preselection edited cells, 820; and postselection edited cells, 161. (**B**) Staining of 20 different THP-1 clones obtained by limiting dilution from the postselection THP-1 parental cell line. MFI values of the indicated populations are represented. The data shown are representative of two independent experiments. (**C**) Molecular characterization of mutations at the MR1 locus target site. CEL-I–digested PCR amplicon homoduplexes of WT THP-1 cells (*lane 1*), heteroduplexes of WT THP-1 cells DNA hybridized with DNA from bulk-transduced THP-1 preselection (*lane 2*) and postselection (*lane 3*), as well as clones 3 (*lane 4*), 6 (*lane 5*), 7 (*lane 6*), 11 (*lane 7*), and 13 (*lane 8*). The Surveyor assay was performed with a total of 500 ng PCR amplicon obtained from genomic DNA of unmodified THP-1 cells or from the different bulk and clonal populations. The gel shown is representative of two independent experiments.

CRISPR/Cas9 mutagenesis to K562 and THP-1 cells. In both lines we were able to show the induction of MR1 mutations at the intended target site with the mismatch nuclease assay (Fig. 1G, 1H). In the case of the highly phagocytic THP-1 cell line, we set up a mutation enrichment strategy relying on the positive selection of THP-1 mutants unable to present MR1 Ags following infection with *M. smegmatis* and on MAIT cell cytotoxicity as a selecting force. The frequency of MR1 mutations was markedly increased by this selection process (Figs. 1H, 6A, 6C), and we were able to isolate monoclonal THP-1 cells by limiting dilution and expansion. All these derivatives failed to stain with anti-MR1 Ab (Fig. 6A, 6B), and the five we tested functionally did not activate MAIT cells in cytokine release assays (Fig. 7). Although we did not establish the clonal nature of these populations, it seems likely that at least some of them are $MR1^{-/-}$ clones that can be used to study intracellular bacteria of different species than those infecting the A549 cells. The fact that all the THP-1 derivatives failed to activate MAIT cells (Fig. 7) is a clear indication of the efficiency of the positive selection approach we undertook, which could prove a useful and powerful way to establish the MHC restriction or Ag specificity of T cells without having to generate clonal target cell derivatives bearing biallelic mutations. Single-cell cloning is a cumbersome process and can be impossible to implement in the context of primary cells with limited proliferative capacity. Moreover, it is conceivable that the efficiency of this approach can be accurately quantified through a deep amplicon sequencing approach.

Concerning the nature of the A549 genomic mutations, it is notable that, on one allele of clone 9, we identified a 126-bp deletion spanning the junction between MR1 exon 2 and the flanking upstream intronic region and ending at the sgRNA target sequence (Fig. 2C). Such large deletions are not commonly observed following nonhomologous end joining genome editing with nuclease systems (CRISPR/Cas9 or others). The genomic DNA disruptions observed on the three other alleles were more in line with what is described in the literature, as we identified a single nucleotide

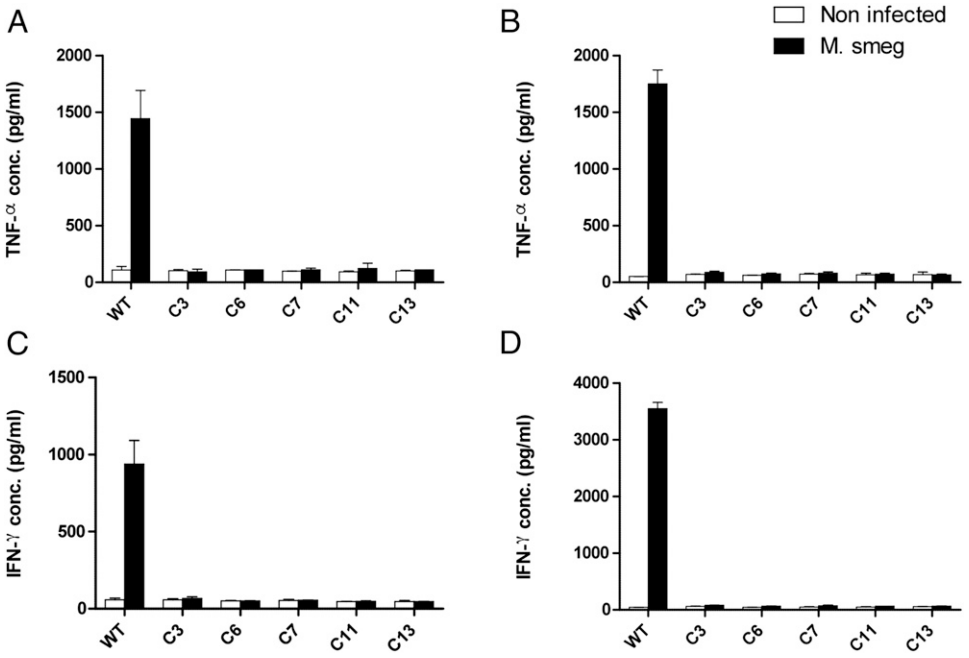


FIGURE 7. THP-1 single-cell derivatives infected with *M. smegmatis* are unable to activate MAIT cells. The D481A9 and 426B1 MAIT cell clones were cocultured with the indicated THP-1 derivatives infected with *M. smegmatis* or not as described in *Materials and Methods*. Supernatants were collected following overnight incubation. TNF-α produced by MAIT cell clones D481A9 (**A**) and D426B1 (**B**), respectively, was quantified by ELISA. IFN-γ levels were also measured for both D481 (**C**) and D426B1 (**D**). Assays were carried out in triplicate wells. Means ± SEM are shown on the graph using data representative of two experiments. Assays were carried out in triplicate wells.

deletion common to all three other alleles that disrupted the protein reading frame from amino acid 34 onward (Fig. 2C–E).

Although the MR1 target sequence we used bears little homology with other classical and nonclassical HLA genes, it was important to test whether other Ag presentation pathways had been disrupted as a result of the genome editing process. Albeit not exhaustive, our control experiments showed that classical HLA class I expression (Fig. 5A, 5B) as well as presentation of cognate exogenous peptide to an HLA-B*1801–restricted $CD8^+$ T cell clone (Fig. 5C, 5D) by both A549 derivatives were equivalent to the unmodified A549 cells. Moreover, both clones activated an HLA-E–restricted $CD8^+$ T cell clone with efficiencies similar to WT A549 (Fig. 5E, 5F). Collectively, these results imply that both the expression and processing of HLA alleles as well as peptide presentation are intact in these cells. Additionally, the basic properties of A549 clones 9 and 11 such as growth rate and morphology appeared unaffected (not shown); however, this does not exclude the possibility of other off-target mutations in unrelated genes. To address this possibility, we performed whole-genome sequencing on both A549 clones and searched for mutations in genes containing sequences with homology to our MR1 gRNA target. In total we examined 7580 locations with off-target potential across the genome. We found no mutations in these except in the RP11-46A10.6 pseudogene (Fig. 3B) locus that contains DNA regions of high homology with some MR1 introns and exons (Fig. 3A). Most notably, the 23 nucleotides composing the MR1 gRNA target sequence and adjacent PAM are 100% conserved in RP11-46A10.6 (Fig. 3B, 3C). Interestingly, the RP11-46A10.6 A549 loci contained the same mutation we identified as being the most common in MR1, that is, a single base deletion at position 17 of the sgRNA seed region. Even though this mutation was clearly unintended, it is on-target in the sense that it occurred at a locus bearing the full intact sgRNA target and PAM sequences. The potential for such an unintended effect was missed during the initial assessment of our MR1 gRNA candidates because we only searched for gRNA target and PAM sequence homology against coding genomic regions using the BLAST tool. Both A549 clones bear this pseudogene mutation, most likely in a homozygous manner, yet we anticipate no functional consequences given the noncoding nature of this locus. The fact that our sgRNA induced mutation in both MR1 and the RP11-46A10.6 pseudogene, where there is a 100% conservation of sequence, serves to highlight both the high efficiency and on-target specificity of CRISPR/Cas9.

The availability of MR1-deficient cells should be useful for fundamental investigations into the biology of MR1 and MR1-restricted T cells as well as for the potential development of therapeutic and diagnostic tools seeking to harness this invariant nonclassical HLA molecule. First, $MR1^{-/-}$ cells will help with the unambiguous identification of MR1-restricted T cells because, unlike Ab blockade, the absence of MR1 completely obliterates MAIT cell responses to bacterially infected cells (Fig. 4). Second, in in vitro activation experiments, MR1-deficient APCs should allow characterizing and quantifying MR1-restricted T cells within the bulk pathogen-specific T cell response. Third, it is unclear at present whether bacterial Ags presented by MR1 show a degree of chemical and structural variability (36) and whether responding T cells can discriminate between distinct cognate MR1/Ag complexes via their clonal TCR. $MR1^{-/-}$ cells should be a useful tool to test the ability of individual putative MR1 ligands to activate T cells. This would require obtaining chemically pure compounds used to pulse MR1-sufficient or -deficient APCs to test the response of T cell clones with distinct TCR usage and assess MR1 restriction. Fourth, the availability of cell lines in which endogenous MR1 expression is abrogated could also prove useful to knock-in mutated or chimeric reporter MR1 proteins and characterize the cellular biology of MR1, such as its trafficking and distribution in subcellular compartments. Such an approach may help shed light on the mechanisms governing the processing and loading of vitamin B derivatives without interference from endogenous MR1 molecules. Finally, $MR1^{-/-}$ cells are likely to prove valuable for the generation and/or validation of biologicals and compounds targeting MR1.

In summary, we have generated THP-1 and A549 isogenic cells deficient for MR1 and a CRISPR/Cas9 lentivector that should allow applying the same process to other cell types. These reagents should facilitate investigations seeking to further understand the mechanisms underpinning the unique biology of MR1 and to harness its therapeutic and diagnostic potential. In the future, similar genome editing approaches could also be extended to fa-

cilitate basic and applied research on other invariant MHC-related proteins such as the CD1 family, HLA-E, or other MHC class Ib gene products that may serve as ligands for the TCR (37).

Disclosures

The authors have no financial conflicts of interest.

References

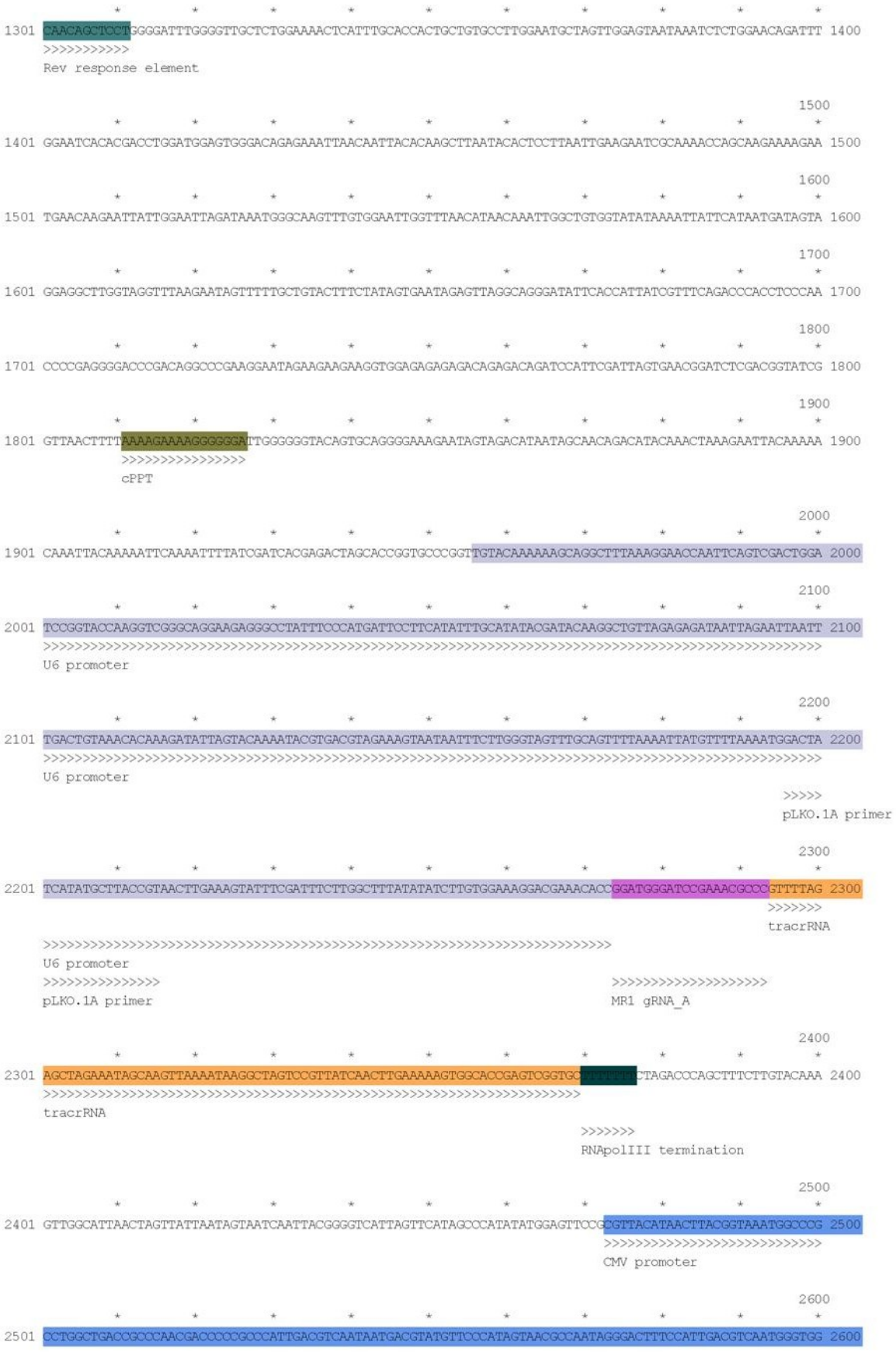
1. Ussher, J. E., P. Klenerman, and C. B. Willberg. 2014. Mucosal-associated invariant T-cells: new players in anti-bacterial immunity. *Front. Immunol.* 5: 450.
2. Gold, M. C., S. Cerri, S. Smyk-Pearson, M. E. Cansler, T. M. Vogt, J. Delepine, E. Winata, G. M. Swarbrick, W.-J. Chua, Y. Y. L. Yu, et al. 2010. Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biol.* 8: e1000407.
3. Gold, M. C., and D. M. Lewinsohn. 2013. Co-dependents: MR1-restricted MAIT cells and their antimicrobial function. *Nat. Rev. Microbiol.* 11: 14–19.
4. Le Bourhis, L., M. Dusseaux, A. Bohineust, S. Bessoles, E. Martin, V. Premel, M. Coré, D. Sleurs, N.-E. Serriari, E. Treiner, et al. 2013. MAIT cells detect and efficiently lyse bacterially-infected epithelial cells. *PLoS Pathog.* 9: e1003681.
5. Meierovics, A., W.-J. C. Yankelevich, and S. C. Cowley. 2013. MAIT cells are critical for optimal mucosal immune responses during in vivo pulmonary bacterial infection. *Proc. Natl. Acad. Sci. USA* 110: E3119–E3128.
6. Tilloy, F., E. Treiner, S. H. Park, C. Garcia, F. Lemonnier, H. de la Salle, A. Bendelac, M. Bonneville, and O. Lantz. 1999. An invariant T cell receptor alpha chain defines a novel TAP-independent major histocompatibility complex class Ib-restricted α/β T cell subpopulation in mammals. *J. Exp. Med.* 189: 1907–1921.
7. Gold, M. C., T. Eid, S. Smyk-Pearson, Y. Eberling, G. M. Swarbrick, S. M. Langley, P. R. Streeter, D. A. Lewinsohn, and D. M. Lewinsohn. 2013. Human thymic MR1-restricted MAIT cells are innate pathogen-reactive effectors that adapt following thymic egress. *Mucosal Immunol.* 6: 35–44.
8. Le Bourhis, L., E. Martin, I. Péguillet, A. Guihot, N. Froux, M. Coré, E. Lévy, M. Dusseaux, V. Meyssonnier, V. Premel, et al. 2010. Antimicrobial activity of mucosal-associated invariant T cells. *Nat. Immunol.* 11: 701–708.
9. Kjer-Nielsen, L., O. Patel, A. J. Corbett, J. Le Nours, B. Meehan, L. Liu, M. Bhati, Z. Chen, L. Kostenko, R. Reantragoon, et al. 2012. MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* 491: 717–723.
10. McWilliam, H. E., R. W. Birkinshaw, J. A. Villadangos, J. McCluskey, and J. Rossjohn. 2015. MR1 presentation of vitamin B-based metabolite ligands. *Curr. Opin. Immunol.* 34: 28–34.
11. Corbett, A. J., S. B. G. Eckle, R. W. Birkinshaw, L. Liu, O. Patel, J. Mahony, Z. Chen, R. Reantragoon, B. Meehan, H. Cao, et al. 2014. T-cell activation by transitory neo-antigens derived from distinct microbial pathways. *Nature* 509: 361–365.
12. Gold, M. C., J. E. McLaren, J. A. Reistetter, S. Smyk-Pearson, K. Ladell, G. M. Swarbrick, Y. Y. Yu, T. H. Hansen, O. Lund, M. Nielsen, et al. 2014. MR1-restricted MAIT cells display ligand discrimination and pathogen selectivity through distinct T cell receptor usage. *J. Exp. Med.* 211: 1601–1610.
13. Gherardin, N. A., A. N. Keller, R. E. Woolley, J. Le Nours, D. S. Ritchie, P. J. Neeson, R. W. Birkinshaw, S. B. Eckle, J. N. Waddington, L. Liu, et al. 2016. Diversity of T cells restricted by the MHC class I-related molecule MR1 facilitates differential antigen recognition. *Immunity* 44: 32–45.
14. Chua, W.-J., S. M. Truscott, C. S. Eickhoff, A. Blazevic, D. F. Hoft, and T. H. Hansen. 2012. Polyclonal mucosa-associated invariant T cells have unique innate functions in bacterial infection. *Infect. Immun.* 80: 3256–3267.
15. Turtle, C. J., J. Delrow, R. C. Joslyn, H. M. Swanson, R. Basom, L. Tabellini, C. Delaney, S. Heimfeld, J. A. Hansen, and S. R. Riddell. 2011. Innate signals overcome acquired TCR signaling pathway regulation and govern the fate of human CD161^{hi} CD8α^{+} semi-invariant T cells. *Blood* 118: 2752–2762.
16. Teunissen, M. B. M., N. G. Yeremenko, D. L. P. Baeten, S. Chielie, P. I. Spuls, M. A. de Rie, O. Lantz, and P. C. Res. 2014. The IL-17A-producing CD8^{+} T-cell population in psoriatic lesional skin comprises mucosa-associated invariant T cells and conventional T cells. *J. Invest. Dermatol.* 134: 2898–2907.
17. Miyazaki, Y., S. Miyake, A. Chiba, O. Lantz, and T. Yamamura. 2011. Mucosal-associated invariant T cells regulate Th1 response in multiple sclerosis. *Int. Immunol.* 23: 529–535.
18. Magalhaes, I., K. Pingris, C. Poitou, S. Bessoles, N. Venteclef, B. Kiaf, L. Beaudoin, J. Da Silva, O. Allatif, J. Rossjohn, et al. 2015. Mucosal-associated invariant T cell alterations in obese and type 2 diabetic patients. *J. Clin. Invest.* 125: 1752–1762.
19. Yamaguchi, H., M. Hirai, Y. Kurosawa, and K. Hashimoto. 1997. A highly conserved major histocompatibility complex class I-related gene in mammals. *Biochem. Biophys. Res. Commun.* 238: 697–702.
20. Riegert, P., V. Wanner, and S. Bahram. 1998. Genomics, isoforms, expression, and phylogeny of the MHC class I-related MR1 gene. *J. Immunol.* 161: 4066–4077.
21. Huang, S., S. Gilfillan, M. Cella, M. J. Miley, O. Lantz, L. Lybarger, D. H. Fremont, and T. H. Hansen. 2005. Evidence for MR1 antigen presentation to mucosal-associated invariant T cells. *J. Biol. Chem.* 280: 21183–21193.
22. Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819–823.
23. Mali, P., K. M. Esvelt, and G. M. Church. 2013. Cas9 as a versatile tool for engineering biology. *Nat. Methods* 10: 957–963.
24. Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. 2013. RNA-guided human genome engineering via Cas9. *Science* 339: 823–826.
25. Sander, J. D., and J. K. Joung. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32: 347–355.
26. Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816–821.
27. Stromnes, I. M., C. Fowler, C. C. Casamina, C. M. Georgopolos, M. S. McAfee, T. M. Schmitt, X. Tan, T.-D. Kim, I. Choi, J. N. Blattman, and P. D. Greenberg. 2012. Abrogation of SRC homology region 2 domain-containing phosphatase 1 in tumor-specific T cells improves efficacy of adoptive immunotherapy by enhancing the effector function and accumulation of short-lived effector T cells in vivo. *J. Immunol.* 189: 1812–1825.
28. Fu, Y., J. D. Sander, D. Reyon, V. M. Cascio, and J. K. Joung. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32: 279–284.
29. Ran, F. A., L. Cong, W. X. Yan, D. A. Scott, J. S. Gootenberg, A. J. Kriz, B. Zetsche, O. Shalem, X. Wu, K. S. Makarova, et al. 2015. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520: 186–191.
30. Lewinsohn, D. M., A. L. Briden, S. G. Reed, K. H. Grabstein, and M. R. Alderson. 2000. *Mycobacterium tuberculosis*-reactive CD8^{+} T lymphocytes: the relative contribution of classical versus nonclassical HLA restriction. *J. Immunol.* 165: 925–930.
31. Heinzel, A. S., J. E. Grotzke, R. A. Lines, D. A. Lewinsohn, A. L. McNabb, D. N. Streblow, V. M. Braud, H. J. Grieser, J. T. Belisle, and D. M. Lewinsohn. 2002. HLA-E-dependent presentation of Mtb-derived antigen to human CD8^{+} T cells. *J. Exp. Med.* 196: 1473–1481.
32. Rist, M. J., K. M. Hibbert, N. P. Croft, C. Smith, M. A. Neller, J. M. Burrows, J. J. Miles, A. W. Purcell, J. Rossjohn, S. Gras, and S. R. Burrows. 2015. T cell cross-reactivity between a highly immunogenic EBV epitope and a self-peptide naturally presented by HLA-B*18:01^{+} cells. *J. Immunol.* 194: 4668–4675.
33. Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.
34. Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16: 944–945.
35. Ackerman, M. E., B. Moldt, R. T. Wyatt, A. S. Dugast, E. McAndrew, S. Tsoukas, S. Jost, C. T. Berger, G. Sciaranghella, Q. Liu, et al. 2011. A robust, high-throughput assay to determine the phagocytic activity of clinical antibody samples. *J. Immunol. Methods* 366: 8–19.
36. Soudais, C., F. Samassa, M. Sarkis, L. Le Bourhis, S. Bessoles, D. Blanot, M. Hervé, F. Schmidt, D. Mengin-Lecreulx, and O. Lantz. 2015. In vitro and in vivo analysis of the Gram-negative bacteria-derived riboflavin precursor derivatives activating mouse MAIT cells. *J. Immunol.* 194: 4641–4649.
37. Attaf, M., M. Legut, D. K. Cole, and A. K. Sewell. 2015. The T cell antigen receptor: the Swiss army knife of the immune system. *Clin. Exp. Immunol.* 181: 1–18.

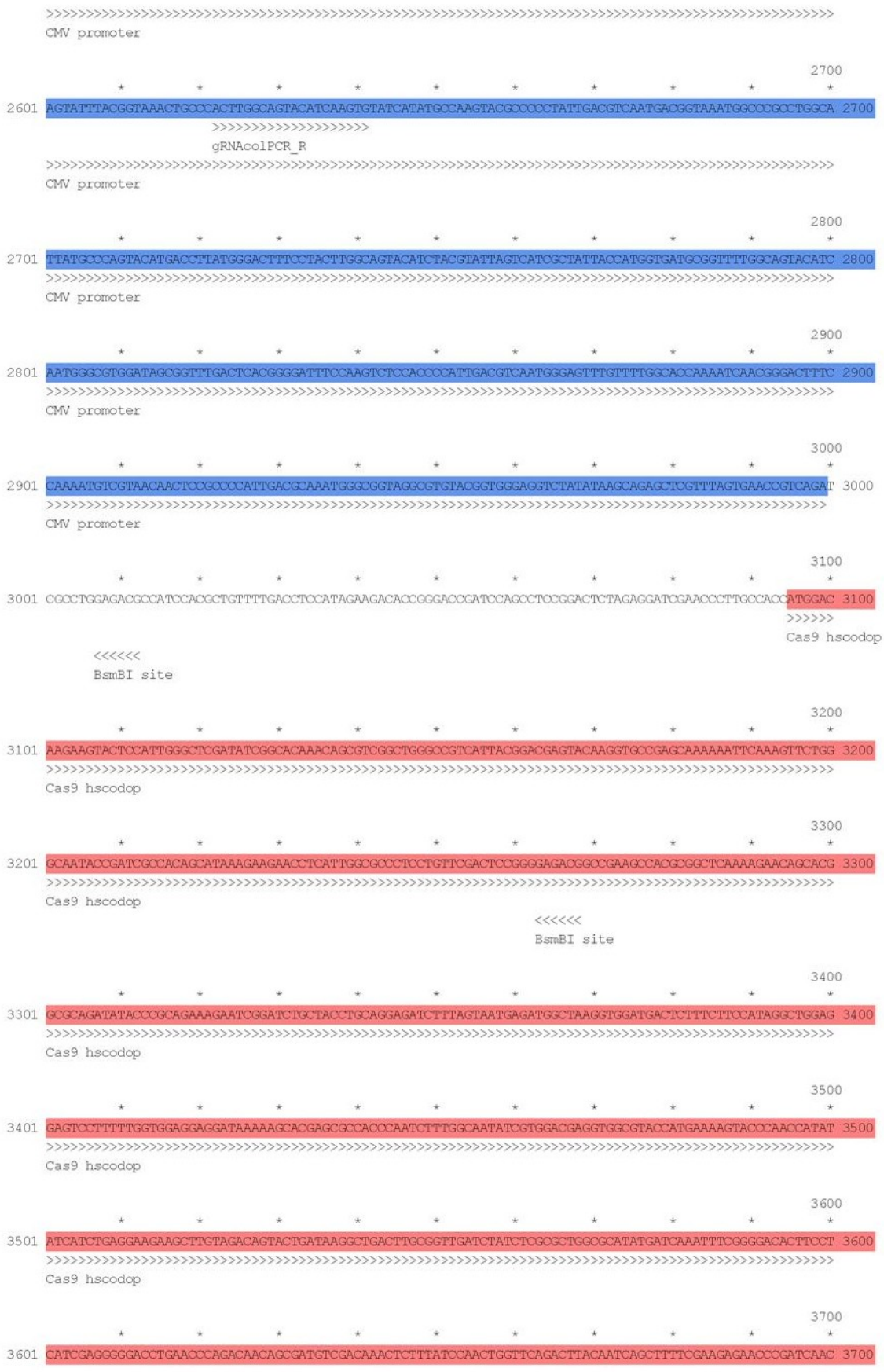
gRNA target sequence PCR cloning primers	**gRNA_A_Fwd**	GAAACGCCC**GTTTTAGAGCTAGAAATAGCAAGTTAA**
	gRNA_A_Rev	GGATCCCATC**CGGTGTTTCGTCCTTTCC**
	gRNA_B_Fwd	CAGCGATTCC**GTTTTAGAGCTAGAAATAGCAAGTTAA**
	gRNA_B_Rev	TGCTTCAC**CGGTGTTTCGTCCTTTCC**
	gRNA_C_Fwd	AATTTATTTCGGT**GTTTTAGAGCTAGAAATAGCAAGTTAA**
	gRNA_C_Rev	CAGGGA**CGGTGTTTCGTCCTTTCC**
	gRNA_D_Fwd	GCCTGATCACT**GTTTTAGAGCTAGAAATAGCAAGTTAA**
	gRNA_D_Rev	GCGAGGTT**CGGTGTTTCGTCCTTTCC**
	gRNA_E_Fwd	TATGACGGGC**GTTTTAGAGCTAGAAATAGCAAGTTAA**
	gRNA_E_Rev	TGCATACTG**CGGTGTTTCGTCCTTTCC**
Primers for genomic MR1 PCR and SURVEYOR assay	**SURV1_Fwd**	GCATGTGTTTGTGTGCCTGT
	SURV1_Rev	GGTGCAATTCAGCATCCGC
Primers for MR1 cDNA PCR and SURVEYOR assay	**SURV2_Fwd**	GGTCTTACTGACATCCACTTTGC
	SURV2_Rev	CAGTGATCAGGCGCGAG
MR1 amplicon cloning primers	**SURV_Fwd_BsaI**	gcgcGGTCTCcGCATGTGTTTGTGTGCCTGT
	SURV_Rev_BsaI	gcgcGGTCTCcTGCCGGTGCAATTCAGCATCCGC
Primers used to sequence MR1 amplicons	**Seq_Fwd**	CCAGTTGCTGAAGATCGCGAAGC
	Seq_Rev	TGCCACTCGATGTGATGTCCTC
Colony PCR primers flanking gRNA target sites	**pLKO.1-A**	GACTATCATATGCTTACCGT
	gRNAcolPCR_R	CACTTGATGTACTGCCAAGT
gRNA/pCMV-Cas9 PCR amplification in pCDNA.3	**pCDNA.3_Fwd**	GCACCGGTTGTACAAAAAAGCAGGCTTTA
	pCDNA.3_Rev	GCATGCATTCACACCTTCCTCTTCTTCT
"Empty" pRRL vector PCR amplification	**pRRL.0_Fwd**	GCATGCATAATCAACCTCTGGATTACAAAATTTG
	pRRL.0_Rev	ACCGGTGCTAGTCTCGTGATCGATAAAAT

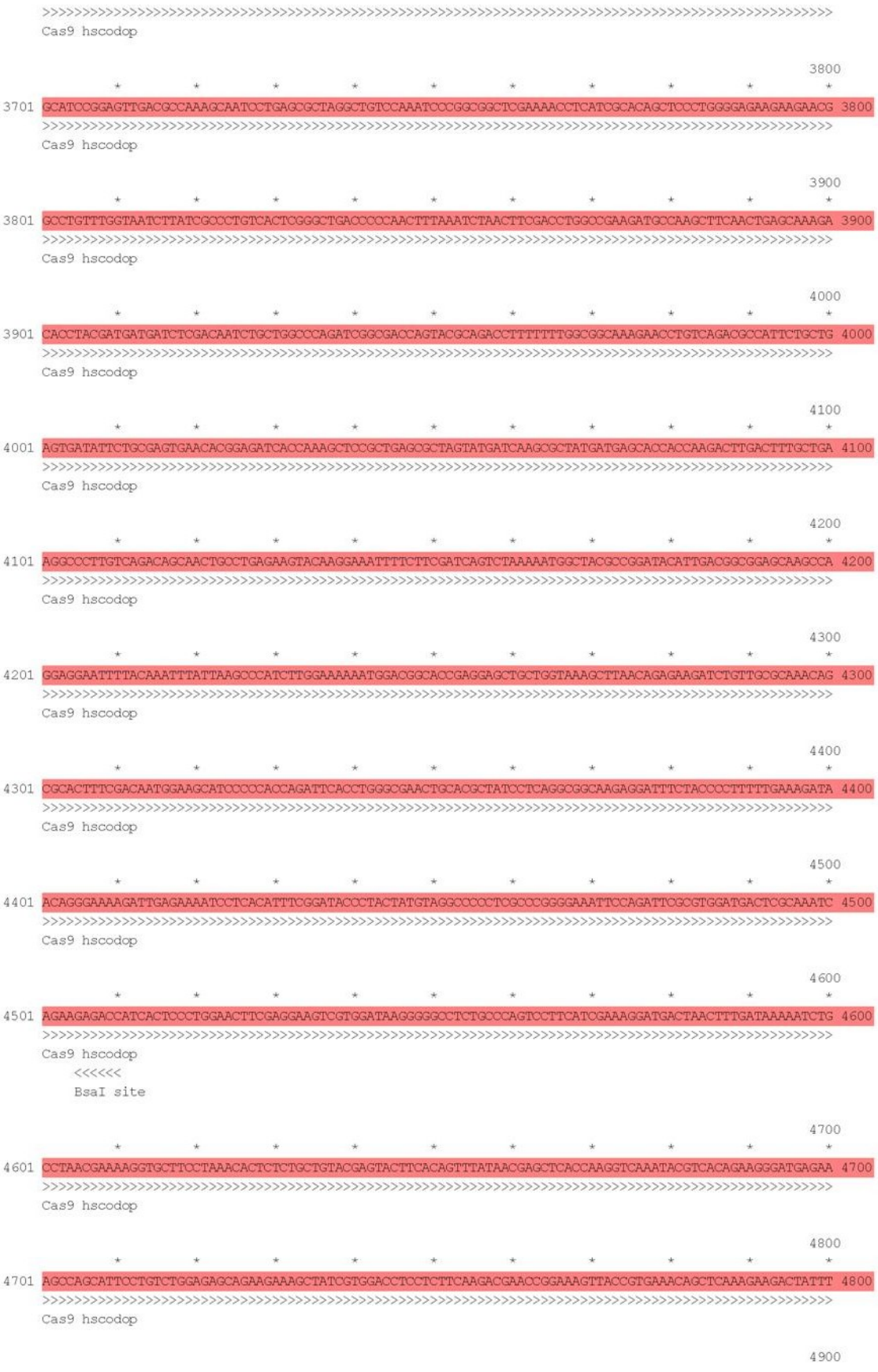
Supplementary Table 1. List of primers used in project.

Thu Jan 22, 2015 23:17 -0000
pRRL_Cas9_MR1gRNA_A.str
Text Map

```
                                                                                                        100
               *         *         *         *         *         *         *         *         *         *
     1 AGCTTAATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCATGCC 100

                                                                                                        200
               *         *         *         *         *         *         *         *         *         *
   101 GATTGGTGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGCCGCATTGCAGAGA 200

                                                                                                        300
               *         *         *         *         *         *         *         *         *         *
   201 TATTGTATTTAAGTGCCTAGCTCGATACAATAAACGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCT 300
                                          >>>>>>
                                          BsaI site
                                         >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
                                         HIV-1 5'LTR

                                                                                                        400
               *         *         *         *         *         *         *         *         *         *
   301 TAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTG 400
       >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
       HIV-1 5'LTR

                                                                                                        500
               *         *         *         *         *         *         *         *         *         *
   401 TGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGC 500
       >>>>>>>>>>>>>>>
       HIV-1 5'LTR

                                                                                                        600
               *         *         *         *         *         *         *         *         *         *
   501 AAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGG 600
                               >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
                               HIV-1_psI_pack

                                                                                                        700
               *         *         *         *         *         *         *         *         *         *
   601 AGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGAT 700

                                                                                                        800
               *         *         *         *         *         *         *         *         *         *
   701 TCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATC 800

                                                                                                        900
               *         *         *         *         *         *         *         *         *         *
   801 ATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAA 900

                                                                                                       1000
               *         *         *         *         *         *         *         *         *         *
   901 AGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTA 1000

                                                                                                       1100
               *         *         *         *         *         *         *         *         *         *
  1001 AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCT 1100
                                                                                    >>>>>>>>>>>>>>>>>>>>>>>
                                                                                    Rev response element

                                                                                                       1200
               *         *         *         *         *         *         *         *         *         *
  1101 TGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCT 1200
       >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
       Rev response element

                                                                                                       1300
               *         *         *         *         *         *         *         *         *         *
  1201 GAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGAT 1300
       >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
       Rev response element

                                                                                                       1400
```

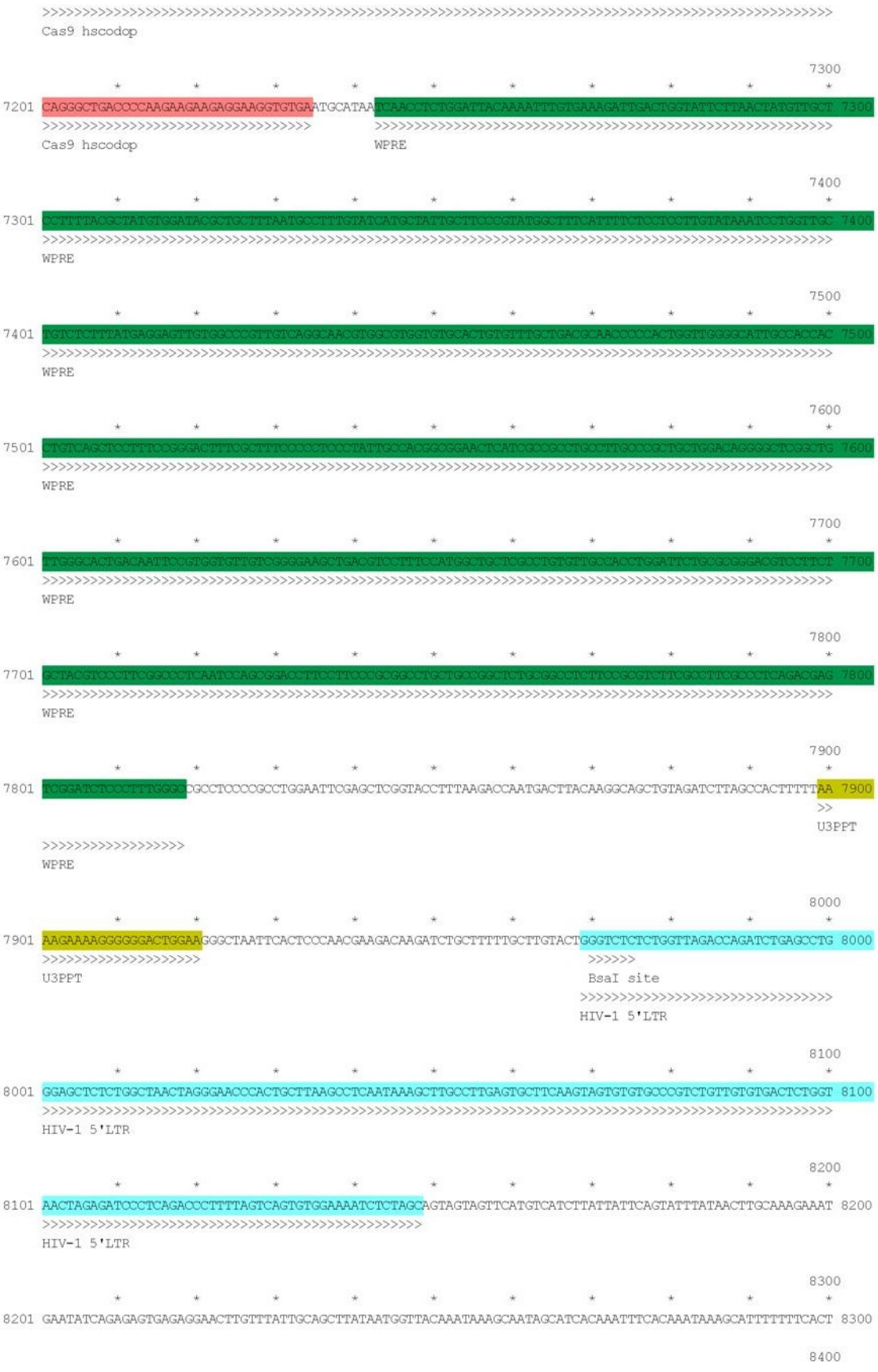

1301 CAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTT 1400
Rev response element
1401 GGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAA 1500
1501 TGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTA 1600
1601 GGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAA 1700
1701 CCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCG 1800
1801 GTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAA 1900
cPPT
1901 CAAATTACAAAAATTCAAAATTTTATCGATCACGAGACTAGCACCGGTGCCCGGTTGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGA 2000
2001 TCCGGTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATT 2100
U6 promoter
2101 TGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTA 2200
U6 promoter
pLKO.1A primer
2201 TCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGATGGGATCCGAAACGCCCGTTTTAG 2300
tracrRNA
U6 promoter
pLKO.1A primer
MR1 gRNA_A
2301 AGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCTAGACCCAGCTTTCTTGTACAAA 2400
tracrRNA
RNApolIII termination
2401 GTTGGCATTAACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCG 2500
CMV promoter
2501 CCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG 2600

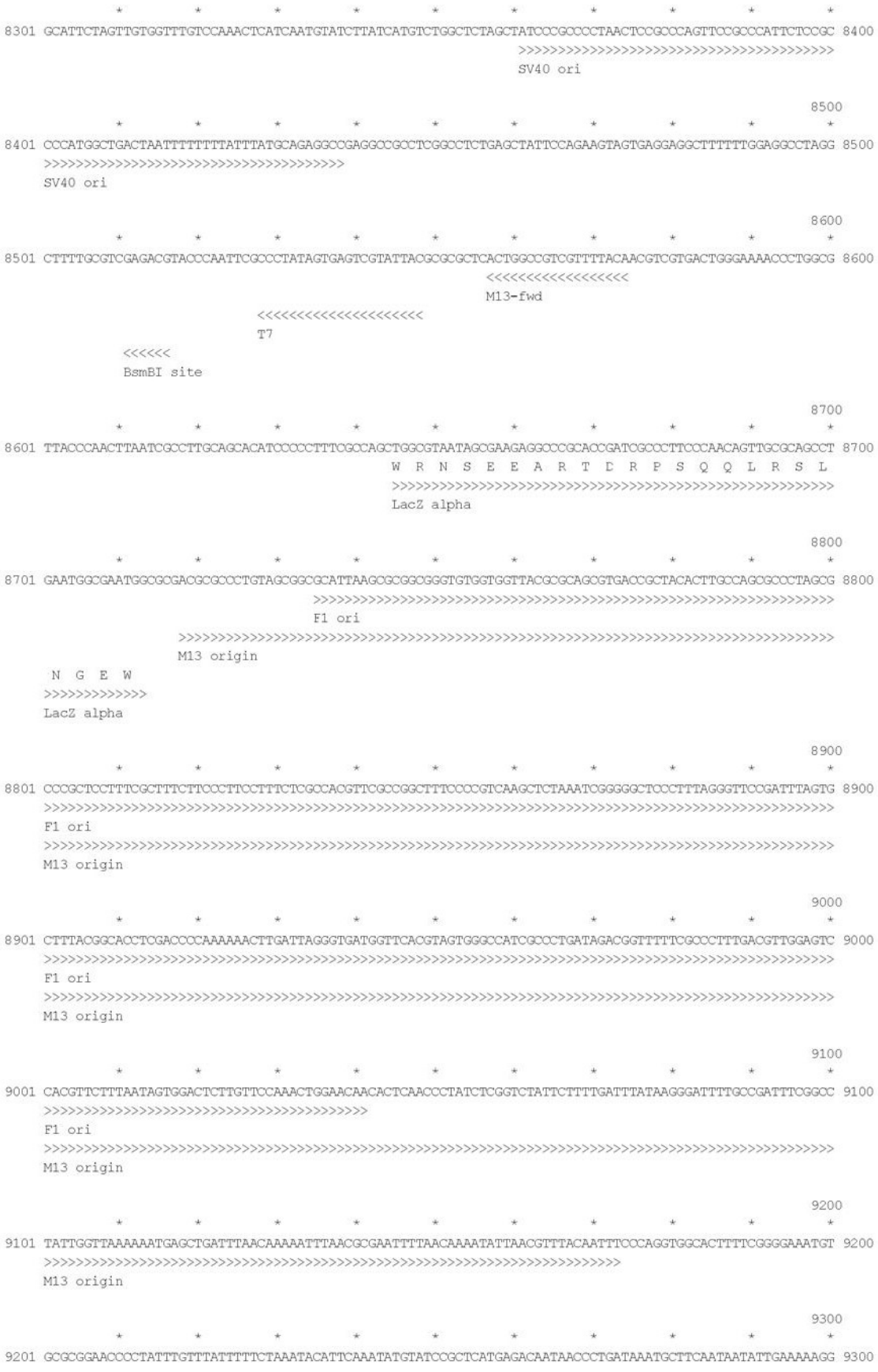

CMV promoter
2700
2601 AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCA 2700
gRNAcolPCR_R
CMV promoter
2800
2701 TTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATC 2800
CMV promoter
2900
2801 AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC 2900
CMV promoter
3000
2901 CAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAT 3000
CMV promoter
3100
3001 CGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGAGGATCGAACCCTTGCCACCATGGAC 3100
Cas9 hscodop
BsmBI site
3200
3101 AAGAAGTACTCCATTGGGCTCGATATCGGCACAAACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGG 3200
Cas9 hscodop
3300
3201 GCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACG 3300
Cas9 hscodop
BsmBI site
3400
3301 GCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAG 3400
Cas9 hscodop
3500
3401 GAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCATAT 3500
Cas9 hscodop
3600
3501 ATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATTTCGGGGACACTTCCT 3600
Cas9 hscodop
3700
3601 CATCGAGGGGGACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAGAGAACCCGATCAAC 3700


Cas9 hscodop
3800
3701 GCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACG 3800
Cas9 hscodop
3900
3801 GCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGA 3900
Cas9 hscodop
4000
3901 CACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTG 4000
Cas9 hscodop
4100
4001 AGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACTTGACTTTGCTGA 4100
Cas9 hscodop
4200
4101 AGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGGCGGAGCAAGCCA 4200
Cas9 hscodop
4300
4201 GGAGGAATTTTACAAATTTATTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAG 4300
Cas9 hscodop
4400
4301 CGCACTTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTGAAAGATA 4400
Cas9 hscodop
4500
4401 ACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTGGATGACTCGCAAATC 4500
Cas9 hscodop
4600
4501 AGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTG 4600
Cas9 hscodop
<<<<<<
BsaI site
4700
4601 CCTAACGAAAAGGTGCTTCCTAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAA 4700
Cas9 hscodop
4800
4701 AGCCAGCATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGACTATTT 4800
Cas9 hscodop
4900

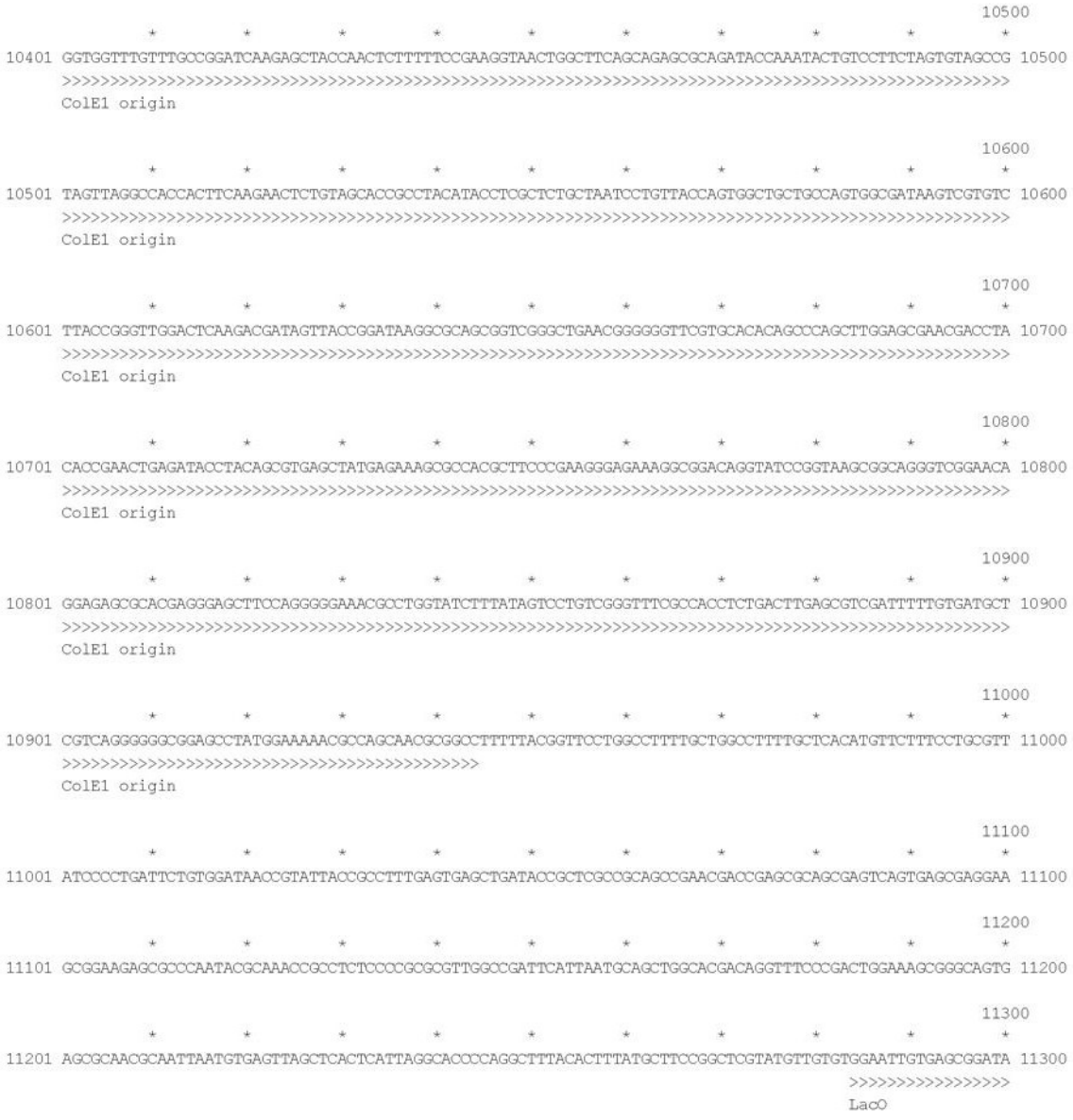

4801 CAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTAAA 4900
Cas9 hscodop
4901 GACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCT 5000
Cas9 hscodop
5001 TGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCAATGG 5100
Cas9 hscodop
5101 GATCCGAGACAAGCAGAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGATGACTCTCTC 5200
Cas9 hscodop
5201 ACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGG 5300
Cas9 hscodop
5301 GAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAAC 5400
Cas9 hscodop
5401 TACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAA 5500
Cas9 hscodop
5501 AACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATCGGCTCTCCGACT 5600
Cas9 hscodop
5601 ACGACGTGGATCATATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAATAGAGGGAAGAGTGATAA 5700
Cas9 hscodop
5701 CGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAG 5800
Cas9 hscodop
5801 GCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCG 5900
Cas9 hscodop
5901 ATTCACGCATGAACACCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGA 6000
Cas9 hscodop
BsaI site


6100
6001 CTTTCAGTTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATATCCC 6100
Cas9 hscodop
6200
6101 AAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGT 6200
Cas9 hscodop
6300
6201 ACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAAC 6300
Cas9 hscodop
6400
6301 AGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACC 6400
Cas9 hscodop
6500
6401 GGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCGGATTCGATT 6500
Cas9 hscodop
6600
6501 CTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCAT 6600
Cas9 hscodop
6700
6601 GGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCT 6700
Cas9 hscodop
6800
6701 CTCTTTGAGCTTGAAAACGGCCGGAAACGAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTAATTTCT 6800
Cas9 hscodop
6900
6801 TGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAACACTACCTTGATGAGAT 6900
Cas9 hscodop
>>>>>>
BsaI site
7000
6901 CATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATC 7000
Cas9 hscodop
7100
7001 AGGGAGCAGGCAGAAAACATTATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACCACCATAGACAGAAAGCGGT 7100
Cas9 hscodop
7200
7101 ACACCTCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGAGACAG 7200


Cas9 hscodop
7300
7201 CAGGGCTGACCCCAAGAAGAAGAGGAAGGTGTGAATGCATAAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCT 7300
Cas9 hscodop
WPRE
7400
7301 CCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGC 7400
WPRE
7500
7401 TGTCTCTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCAC 7500
WPRE
7600
7501 CTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTG 7600
WPRE
7700
7601 TTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCT 7700
WPRE
7800
7701 GCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAG 7800
WPRE
7900
7801 TCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGGAATTCGAGCTCGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAA 7900
U3PPT
WPRE
8000
7901 AAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTG 8000
U3PPT
BsaI site
HIV-1 5'LTR
8100
8001 GGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT 8100
HIV-1 5'LTR
8200
8101 AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGAAAT 8200
HIV-1 5'LTR
8300
8201 GAATATCAGAGAGTGAGAGGAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACT 8300
8400


SV40 ori
M13-fwd
T7
BsmBI site
W R N S E E A R T D R P S Q Q L R S L
LacZ alpha
F1 ori
M13 origin
N G E W
LacZ alpha
F1 ori
M13 origin
F1 ori
M13 origin
M13 origin

```
                                                                                                      9400
                *         *         *         *         *         *         *         *         *         *
 9301 AAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA 9400

                                                                                                      9500
                *         *         *         *         *         *         *         *         *         *
 9401 AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCC 9500

                                                                                                      9600
                *         *         *         *         *         *         *         *         *         *
 9501 AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAT 9600
          M  S  T  F  K  V  L  L  C  G  A  V  L  S  R  I  D  A  G  Q  E  Q  L  G  R  R  I  H  Y  S  Q  N
          >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
          AmpR

                                                                                                      9700
                *         *         *         *         *         *         *         *         *         *
 9601 GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG 9700
      D  L  V  E  Y  S  P  V  T  E  K  H  L  T  D  G  M  T  V  R  E  L  C  S  A  A  I  T  M  S  D  N  T  A
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR

                                                                                                      9800
                *         *         *         *         *         *         *         *         *         *
 9701 CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACC 9800
        A  N  L  L  L  T  T  I  G  G  P  K  E  L  T  A  F  L  H  N  M  G  D  H  V  T  R  L  D  R  W  E  P
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR

                                                                                                      9900
                *         *         *         *         *         *         *         *         *         *
 9801 GGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACT 9900
       E  L  N  E  A  I  P  N  D  E  R  D  T  T  M  P  V  A  M  A  T  T  L  R  K  L  L  T  G  E  L  L  T
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR

                                                                                                      10000
                *         *         *         *         *         *         *         *         *         *
 9901 CTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATA 10000
      L  A  S  R  Q  Q  L  I  D  W  M  E  A  D  K  V  A  G  P  L  L  R  S  A  L  P  A  G  W  F  I  A  D  K
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR

                                                                                                      10100
                *         *         *         *         *         *         *         *         *         *
10001 AATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA 10100
        S  G  A  G  E  R  G  S  R  G  I  I  A  A  L  G  P  D  G  K  P  S  R  I  V  V  I  Y  T  T  G  S  Q
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR
                           >>>>>>
                           BsaI site

                                                                                                      10200
                *         *         *         *         *         *         *         *         *         *
10101 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAG 10200
        A  T  M  D  E  R  N  R  Q  I  A  E  I  G  A  S  L  I  K  H  W
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      AmpR

                                                                                                      10300
                *         *         *         *         *         *         *         *         *         *
10201 ATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT 10300
                                                                   >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
                                                                   ColE1 origin

                                                                                                      10400
                *         *         *         *         *         *         *         *         *         *
10301 GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGC 10400
      >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
      ColE1 origin
```



Supplementary Figure 1. Annotated full DNA sequence of the pRRL.sin.CRISPR/Cas9 plasmid..