Construction and characterisation of models for X-linked severe combined immunodeficiency for targeted gene correction by zinc finger nucleases.

Celeste Elizabeth Pallant

INSTITUTE OF CHILD HEALTH

UNIVERSITY COLLEGE LONDON

A thesis submitted for the degree of Doctor of Philosophy

Declaration

I, Celeste Pallant confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

X-linked Severe Combined Immunodeficiency (SCID-X1) is an immunopathy caused by a mutation of the common gamma chain (γc) gene, *IL2RG*, which results in a lack of T cells, NK cells and with dysfunctional B cells. Current gene therapy methods involve the addition of a correct γc gene via integrating viral vectors. However, these current non-targeting gene addition strategies can result in transformation of the cell. A novel solution to this problem is met by targeted gene correction via homologous recombination stimulated by a site specific cleavage event caused by zinc finger nucleases (ZFN) within the disease gene.

A γc deficient mouse has been created by replacing the murine *il2rg* locus with a mutated human *IL2RG* containing a point mutation frequently seen in SCID-X1 patients. The mutant human *IL2RG* is transcribed and initial analysis of this new SCID-X1 model has revealed a phenotype mirroring γc gene knockout mice. Lineage negative bone marrow cells from these mice, transduced with integrating lentiviral vector encoding functional *IL2RG* can reconstitute the immune cells in the $Rag2^{-/-}\gamma c^{-/-}$ double knockout SCID mouse model. Therefore the humanised mouse model of SCID-X1 can be corrected and is an appropriate platform to assess the efficiency of various gene targeting and correction strategies for the human mutation including ZFN induced homologous recombination.

We have successfully achieved targeted homologous recombination in both a human T cell SCID-X1 cell line model and the humanised mouse embryonic stem cells with *IL2RG* specific ZFN.

Acknowledgments

I thank my primary supervisor Adrian Thrasher for allowing me the opportunity to do this well considered and interesting project and I thank Steven Howe for being a caring supervisor, showing me the ropes and giving me so much of his time and patience.

I also thank the group of Toni Cathomen who taught me about ZFN and looked after me in Hannover.

And for the MIU, a poem.

To the beautiful people at the MIU,

My PhD was completed all thanks to you.

I witnessed experimental finesse,

And received unexpected acts of kindness.

Stopping for chats in the corridor,

Letting me sneak treats from the food draw.

Friends who helped me through the sad moods,

And accepted all kinds of my bad rudes.

I met so many wonderful folk,

Round the corner having a smoke,

Fantastic postdocs who are learned thinkers,

And witty and kind and quite often drinkers.

The specials ones who can talk with the mice,

Thank you all so much for your laughs, time and advice.

I thank my family for giving me the support and space to write up and my sister and friends for their love and making sure we had regular chats which helped me keep some sanity.

Finally, I want to regard the Mice, Cattle, Rats, Rabbits and Hamsters that were sacrificed for the direct benefit and enablement of this project.

For Nadine

Table of contents

List of figures

List of tables

List of abbreviations

A	Adenosine
AAV	Adeno-associated virus
ADA	Adenosine deaminase
AML	Acute myeloid leukemia
AMP^R	Ampicillin resistance
AAVS1	Adeno-associated virus site 1

B cell	Bone marrow derived lympocyte
BM	Bone Marrow
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
CCR5	Chemokine receptor 5
cDNA	Complementary DNA
CGD	Chronic granulomatous disease
CMV	Cytomegalovirus
CO_2	Carbon Dioxide
Con A	Concanavalin A
COS cell	Simian Cercopithecus aethiops monkey kidney (CV-1) origin cells Immortalised with simian vacuolating virus 40
CpG	Cytosine nucleotide adjacent to guanine nucleotide
cPPT	Central polypurine tract
CRISPR	Clustered regularly interspaced short pallendromic repeats
DAPI	4'-6-diamidino-2-phenylindole
DBD	DNA-binding domain
H_20	Water
DMEM	Dulbecco modified Eagle's minimal essential medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DSB	Double strand DNA break
EDTA	Ethylenediaminetetraacetic acid
EF1α	Elongation Factor 1α
eGFP	Enhanced green fluorescent protein
EJ cell	Endometrioid adenocarcinoma cells
ELISA	Enzyme-linked immunosorbent assay
Env	Envelope
ES cell	Embryonic stem cell
FACS	Fluorescence-activated cell sorter
FCS	Foetal calf serum
FITC	Fluorescein isothiocyanate
G	Guanine
Gag	Group specific antigens
GFP	Green fluorescent protein
Gy	Gray
H	Histidine
h	Human
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hγcmut	Humanised mutant gamma chain
HIV-1	Human immunodeficiency virus type 1

HLA	Human leukocyte antigen
HR	Homologous recombination
HRP	Horseradish peroxidase
HSC	Human stem cell
HSCT	Human stem cell therapy
hTPO	Human thrombopoietin
IDLV	Integrase deficient lentiviral vector
IFN	Interferon
Ig	Immunoglobulin
IL-	Interleukin
IL2RG	Interleukin 2 receptor common gamma chain
LV	(integrating proficient) Lentiviral vector
JAK	Janus-activated kinase
kb	Kilobase
kDa	Kilodalton
LB	Luria-Bertani
LCL	Lymphoblastoid cell line
Lin^-	Lineage negative
LMO2	LIM domain only 2
LTR	Long terminal repeat
M	Molar
m	Milli (10^{-3}) or murine
MFI	Mean flourescence intensity
mFlt-3	Murine foetal liver tyrosine kinase 3 ligand
MOI	Multiplicity of infection
MoMLV	Moloney murine leukaemia virus
mRNA	Messenger RNA
n	Nano (10^{-9})
Neo	Neomycin phosphotransferase
NHEJ	Non-homologous end joining
NK	Natural killer cell
p24	HIV-1 p24 capsid protein
PBMC	Peripheral blood mononuclear cell
PBS	Primer binding site or phosphate buffered saline
PBST	Phosphate buffered saline with Tween
PCR	Polymerase chain reaction
PE	Phycoerythrin
PEI	Polyethylenimine
PFA	Paraformaldehyde
Pol	Polymerase
PPT	Polypurine tract
PRE	Rev-response element

PVDF	polyvinylidene fluoride
qPCR	quantitative PCR
R	Repeat region or Arginine
rag	Recombinase activating gene
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
RPMI	Roswell Park Memorial Institute medium
RRE	Rev response element
RT	Reverse transcriptase
RT-qPCR	Reverse transcriptase polymerase chain reaction
SCF	Stem cell factor
SCID	Severe combined immunodeficiency
SCID-X1	X-linked Severe combined immunodeficiency
SFFV	spleen focus-forming virus
SH2	Src homology domain
SIN	Self-inactivating
ss	Single stranded
ssODN	Single stranded oligodeoxynucleotides
STAT	Signal transducer and activator
T	Thymidine
TALEN	Transcription Activator-Like Effector Nucleases
T cell	Thymus derived lymphocyte
TAE	Tris-acetate-EDTA
T-ALL	T-cell acute lymphoblastic leukaemia
TCR	T cell receptor
TFO	Triplex-forming oligonucleotide
tRNA	Transfer RNA
U3	Unique at 3’ region
U5	Unique at 5’ region
VSV-G	Vesicular stomatitis virus glycoprotein
WPRE	Woodchuck hepatitis virus posttranscriptional regulatory element
WT	Wild type
ZFN	Zinc finger nuclease
α	Alpha
β	Beta
δ or Δ	Delta
γ	Gamma
γc	common cytokine receptor gamma chain
μ	Micro (10^{-6})
Ψ	RNA encapsidation signal

1 Introduction

This introduction outlines the cause of the genetic disorder X-linked severe combined immunodeficiency and discusses treatment for the disease in the context of gene therapy. This will lead to the potential for targeted gene correction strategies, in particular, homologous recombination stimulated by zinc finger nucleases.

1.2 The common gamma chain and Severe Combined Immunodeficiency

X-linked severe combined immunodeficiency (SCID-X1) occurs as a result of the loss of signalling by interleukins vital for immune system development. The affected individuals are severely immuno-compromised with no T cells, no NK cells and defective B cells. Consequently, the individual is not protected from infection and prone to disease caused by common opportunistic pathogens (Buckley 2004). The disease incidence could be up to 1 in 60000 births (Cavazzana-Calvo, Lagresle et al. 2005) and if left untreated, the patient will die at infancy (Fischer, Le Deist et al. 2005). SCID-X1 is caused by dysfunctional common cytokine receptor gamma chain (γc) caused by mutations in its gene, *IL2RG,* resulting in an inherited disorder which follows an X-linked recessive pattern occurring in males.

1.2.1 IL2RG

Linkage analysis revealed the gene for the transmembrane protein, γc, to be located on chromosome Xq13.1 (de Saint Basile, Arveiler et al. 1987; Noguchi, Yi et al. 1993; Puck, Deschenes et al. 1993) and consist of 8 exons. The extracellular domain contains two pairs of conserved cysteine residues shared between exon 2 and 3 and a conserved WSXWS motif located in exon 5. Both motifs are typically found in the cytokine receptor super family. Exon 6 encodes the transmembrane domain and the cytoplasmic tail is encoded by exons 7 and 8 (Noguchi, Adelstein et al. 1993).

γc was first discovered for its association with the interleukin-2 Receptor (IL2R) subunit IL2Rβ (Takeshita, Asao et al. 1990; Takeshita, Ohtani et al. 1992) which achieve cytoplasmic signalling on heterodimerisation (Nakamura, Russell et al. 1994). Together with IL2Rα the high affinity receptor complex, IL2Rαβγ, binds IL-2 which in turn induces tyrosine phosphorylation of Jak1 and Jak3 (Miyazaki, Kawahara et al. 1994) located on the cytoplasmic tails of IL2Rβ and γc respectively (Russell, Johnston et al. 1994). This recruits STAT proteins via their Src homology (SH2) domains, which are in turn

phosphorylated giving rise to STAT dimerisation and subsequent translocation to the nucleus where the dimer is recruited by an IFNγ activation sequence (GAS) motif and in the presence of other transcription factors, regulates transcription (Leonard and O'Shea 1998; Leonard 2001) (Figure 1.1a). Jak3 is only activated by the presence of γc (Leonard and O'Shea 1998) and mutations in its gene also give rise to immunodeficiencies with the same immune cell profile as SCID-X1 (Candotti, Oakes et al. 1997). The immune cell profile of SCID-X1 patients could not be attributed to disruptions in IL-2 signalling alone as normal absolute numbers and percentages of circulating CD4+ and CD8+ T cells are observed in humans with IL-2 deficiency (Chatila, Castigli et al. 1990). This observation was supported by the normal development of thymocytes in IL-2 deficient mice (Schorle, Holtschke et al. 1991).

The IL2R receptor subunit γc was subsequently found be common to the receptor complexes for IL-4 (Kondo, Takeshita et al. 1993; Russell, Keegan et al. 1993), IL-7 (Noguchi, Adelstein et al. 1993; Kondo, Takeshita et al. 1994), IL-9 (Kimura, Takeshita et al. 1995), IL-15 (Giri, Ahdieh et al. 1994) and IL-21 (Asao, Okuyama et al. 2001) each contributing to immune cell function. IL-2 was originally identified as a T cell growth factor (Morgan, Ruscetti et al. 1976), is involved in further regulation of T cell populations (Liao, Lin et al. 2011) and is in-disposable for inducing T regulatory cells (Malek and Bayer 2004). IL-4 was first identified as a B cell growth factor (Howard, Farrar et al. 1982), is involved in B cell development including immunoglobulin class switching to IgE in humans and is implied in thymic development (Paul 1991). IL-7 was also identified for its B cell proliferation properties in mice (Namen, Lupton et al. 1988) however patients with IL7Rα deficiency lack T cells but their B cell and NK cell populations remain intact (Puel, Ziegler et al. 1998). IL-9 was first observed in mice and originally characterized as a T cell growth factor (Van Snick, Goethals et al. 1989) and subsequently shown to prolong human T cell activation *in vitro* (Renauld, Goethals et al. 1990). IL-15 was first observed to stimulate peripheral blood mononuclear cells in humans (PBMC) (Grabstein, Eisenman et al. 1994) and studies in IL-15 knockout mice have revealed its potential in NK cell and CD8+ T cell development (Lodolce, Boone et al. 1998). IL-21 was investigated due to its structural similarities with IL-2, IL-4 and IL-15 (Asao, Okuyama et al. 2001) and was found to induce proliferation in human B cells lines (Habib, Senadheera et al. 2002) and IL12R was later found responsible for human

B cell differentiation *in vivo* (Recher, Berglund et al. 2011). It is the wide application of the γc in these cytokine receptors and the fact γc is constitutively expressed on all peripheral blood leukocytes (Takeshita, Asao et al. 1992) that the immune system is severely compromised in SCID-X1 patients.

SCID-X1-causing mutations including point mutations, insertions and deletions have been found throughout the 8 exons of *IL2RG* (Figure 1.1b). Mutations were identified in 23 of 87 unrelated patients in the WSXWS containing exon 5. This exon is host to mutational hotspots including a CpG dinucleotide mutated in 11 of the patients studied. 5 were point mutation C690T causing amino acid code change R226C and 6 were G691A causing R226H (Puck, Pepper et al. 1997). The amino acid at position R226 is conserved between human, dog (Henthorn, Somberg et al. 1994) and mouse (Cao, Kozak et al. 1993) *IL2RG*. Interestingly, a larger T cell population is observed as a result of the R226C conversion than R226H (Pepper, Buckley et al. 1995; Kumaki, Ishii et al. 2000). The mutation G691A is the common SCID-X1 mutation considered for targeted correction in this project.

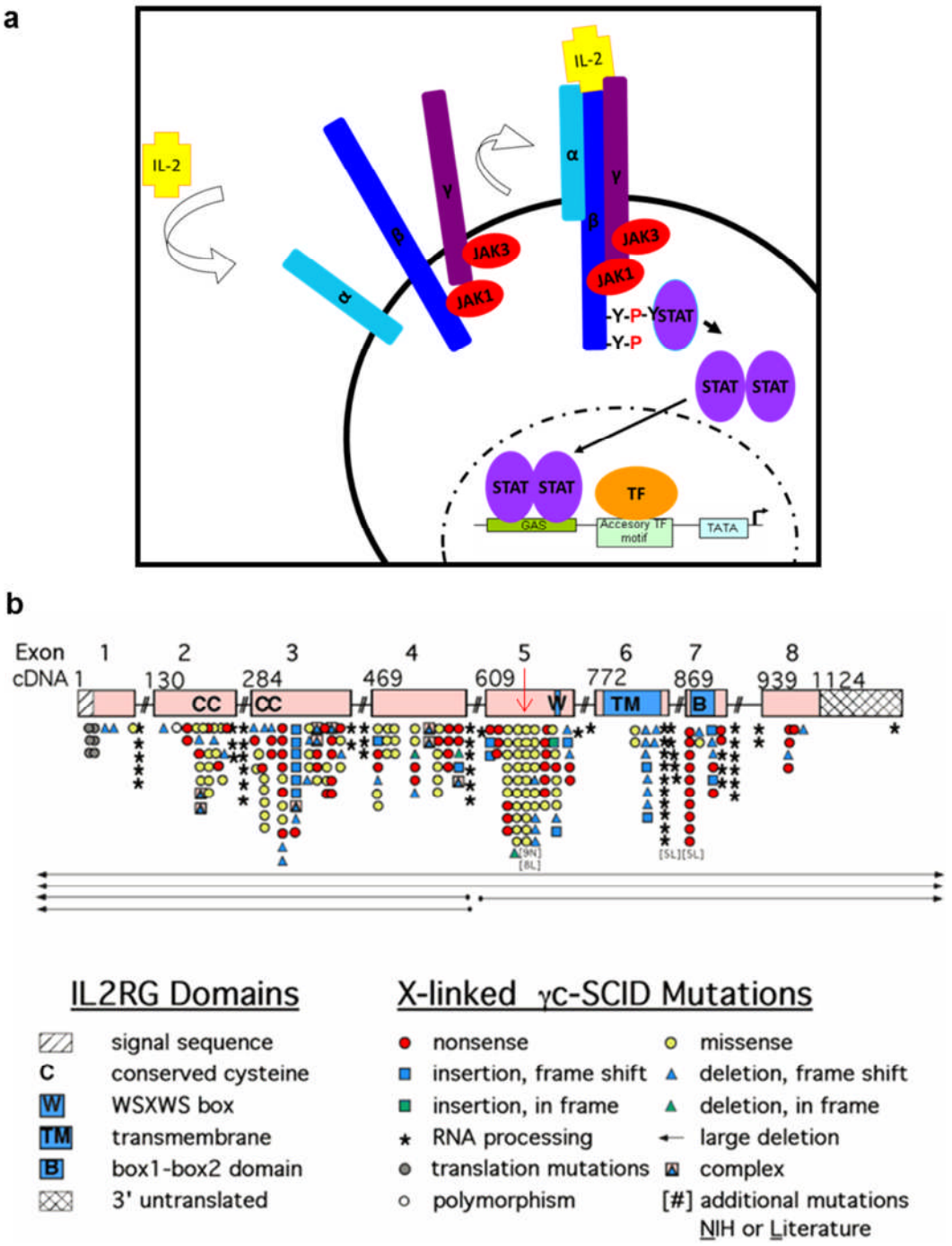


Figure 1.1 The common gamma chain (γc) at the cell surface and SCID-X1 associated mutations.

(a) A schematic of the downstream signalling events on IL-2 binding at the receptor resulting in the signalling pathway whereby Jak1 and Jak3 associate with and phosphorylate STAT which dimerises, enters the nucleus and is recruited by a GAS motif in proximity to a transcriptional start site. Figure adapted from Leonard, O'Shea 1998 (Leonard and O'Shea 1998; Maeder, Thibodeau-Beganny et al. 2008). (b) A map of mutations observed in SCID-X1 patients www.*research.nhgri.nih.gov/scid/IL2RGbase* with the location of the G691A mutation (red arrow).

1.3 Treatment of SCID-X1

The most successful current treatment for SCID-X1 is a bone marrow transplant (BMT) from which donor haematopoietic stem cells (HSC) can provide a constant source of immune cells and due to the absence of immune cells in the recipient, transplant rejection is avoided. Whilst the patient waits for a transplant, prophylaxis is administered including antimicrobial drugs and immunoglobulin therapy. However, obtaining an identical HLA match is not always possible as there is not always a sibling who is a fully matched donor. Haploidentical (half a tissue match) bone marrow from a parent or an unrelated donor can be used and the incidence of graft versus host disease (GVHD) reduced by depleting the donor cells of the T cell compartment before transplantation (Reisner, Kapoor et al. 1983). In a report of European SCID (all severe combined immunodeficiency) cases treated with haematopoietic stem cell therapy (HSCT), identical donor matches were found to be significantly more successful at restoring a SCID patient's immune system than haploidentical, T cell depleted donors (Antoine, Muller et al. 2003). In a later comparison, survival rates were seen to be >90% for fully matched HSCT and 75% for haploidentical T cell depleted donor bone marrow (Cavazzana-Calvo, Lagresle et al. 2005). In another study, transplants in the surviving recipients were able to restore T cell numbers to normal levels and increase NK cell populations however 62% of patients still required immunoglobulin replacement therapy (Buckley, Schiff et al. 1999). Due to limited availability of matched donors for HSCT and ineffective engraftment of donor derived B cells, therapeutic strategies progressed to gene therapy by inserting a functional copy of *IL2RG* into the genome of the SCID-X1 patient's haematopoietic stem cells by retroviral gene transfer.

1.3.1 HSCT

The multipotent quality of HSC has been demonstrated by the repopulation of all the blood cell lineages in animal models. By using the HSC cell surface marker CD34, CD34+ cells isolated from bone marrow have returned bone marrow cellularity to normal in lethally irradiated baboons (Berenson, Andrews et al. 1988) and human bone marrow CD34+ cells achieved chimeric immune cell populations when transplanted into sheep *in utero* (Srour, Zanjani et al. 1993). Humans receiving autologous (Civin, Trischmann et

al. 1996) or haploidentical (Link, Arseniev et al. 1996) CD34+ cells also achieved successful immune cell reconstitution.

CD34+ cells make up approximately 1% of human bone marrow and <10% of these are lineage negative (lin^-) precursor cells, capable of differentiation into various blood cell types. Such cells from human sources were capable of restoring long term, multilineage engraftment in sheep (Civin, Almeida-Porada et al. 1996) and NOD/SCID mice (Bhatia, Wang et al. 1997). Although these are not the only cells capable of haematopoietic repopulation, the small fraction of bone marrow capable of maintaining the blood can be exploited in HSCT.

1.3.2 Retroviral vectors

Vectors designed to modify a patient's stem cells by the incorporation of a therapeutic gene are based on members of the retrovirus family due to their ability to integrate their genetic payload into the host's genome (Provirus). The main candidates for retroviral vector design to deliver *IL2RG* have been gammaretroviruses and lentiviruses. The single stranded RNA genomes of these two viruses are very similar and the basic processes involved from cell entry to provirus integration in brief are described in Figure 1.2a.

Retroviral vectors are designed to keep the essential proteins required for packaging and delivery of the viral genome, but any non-essential accessory proteins including those responsible for their disease profile are removed. Importantly, the vectors have been made replication defective by splitting the genome into separate parts, which can then be individually expressed by plasmid vectors or in specifically engineered packaging cell lines. The *gag* polyprotein is responsible for the structural proteins for the protein matrix, capsid and nucleocapsid. The *Gag-Pol* polyprotein contains reverse transcriptase, necessary for viral ssRNA processing into dsDNA, integrase, for viral DNA integration into the host genome and protease. The env protein is involved in viral-cell membrane fusion. By isolating the *gag, pol* and *env* genes and permitting their expression with accompanying promoters, the viral proteins generated can package the separate, LTR flanked ssRNA carrying the therapeutic gene, but renders the resulting particle unable to generate further virions (Figure 1.2b). Once the viral genome has been reverse transcribed into dsDNA and has entered the nucleus of the cell, this nucleic acid is integrated into the host genome aided by the viral enzyme,

integrase. The therapeutic gene is now expressed due to the promoter/enhancer elements of the viral LTR, additional modifications can remove this feature from the LTR, resulting in a self-inactivating (SIN) vector where therapeutic gene expression is mediated by a provided promoter. This multipart genome also avoids the risk of recombining with other wild type retroviruses that may be present in the target cell.

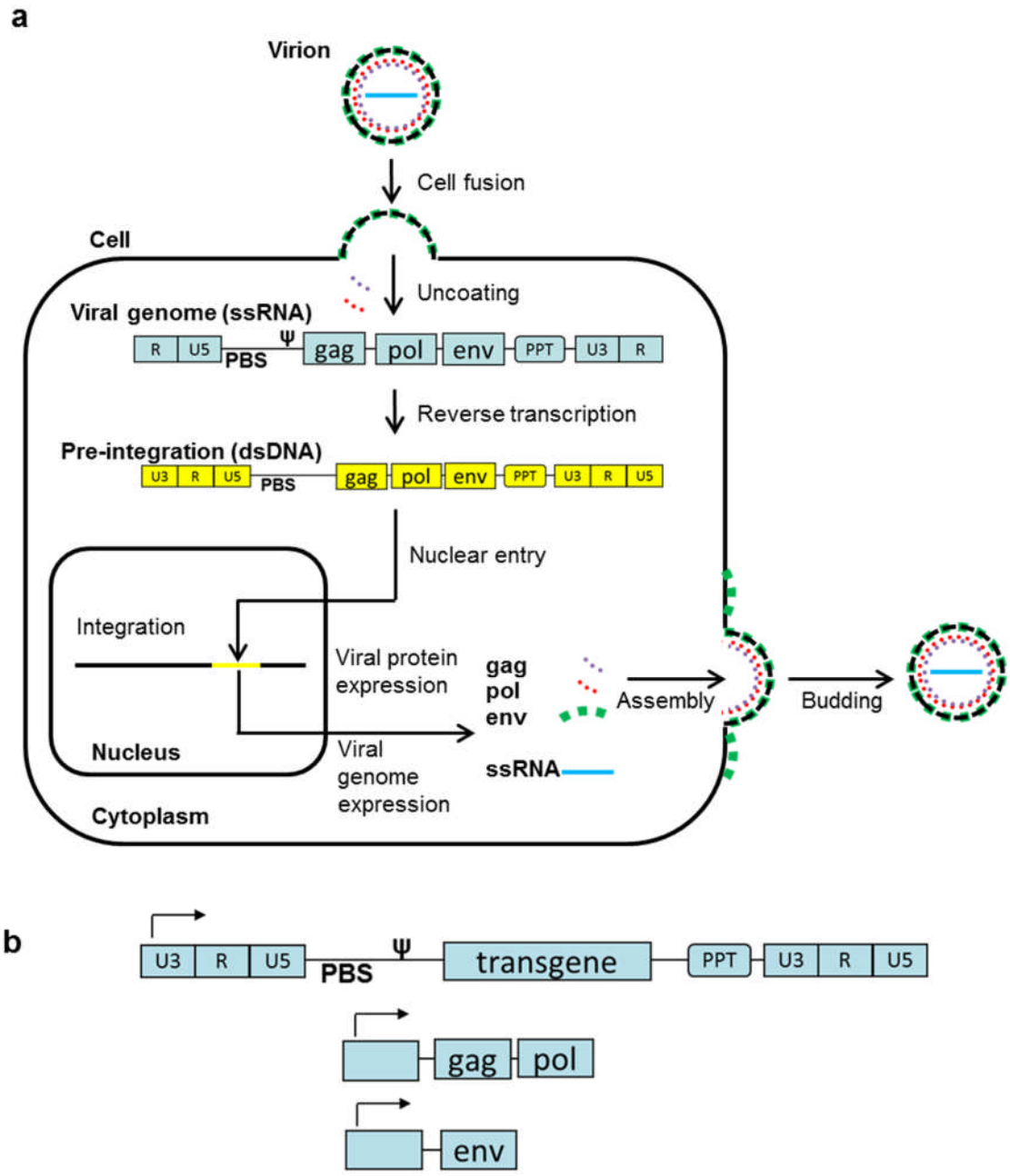


Figure 1.2 The life cycle of a retrovirus and the essential proteins for lentiviral production.

(a) The env protein makes contact with the host cell and the membrane fuses with the host membrane delivering the capsid into the cell's cytosol. The viral core is uncoated and reverse transcription is initiated in order to convert the single stranded (ss)RNA genome into double stranded (ds)DNA. This now enters the nucleus where it is integrated into a host chromosome by the viral integrase. Now the LTR initiates transcription of the *gag*, *pol* and *env* genes which are required for packaging and viral processes. The proteins assemble at the host membrane and the ssRNA contains a packaging signal necessary for virion assembly. The virion leaves the cell by budding off. (b) Viral vector plasmid constructs containing the LTR flanked transgene with packaging signal or the *gag* and *pol* genes or the *env* gene.

1.3.3 Retroviral gene therapy of SCID-X1

By introducing a transgene via retroviral vectors into the genome of multipotent stem cells from a patient, it is possible to repopulate tissue with corrected cells without the complications from mismatched tissue. SCID-X1 is an ideal candidate for such *ex vivo* treatment as it is a monogenic disorder and there is no immune system to recognise the newly incorporated protein as foreign thus avoiding challenge of the adaptive immune system towards the corrected cells. Furthermore, the incorporated γc should provide a selective advantage, due to the growth signals mediated by IL-2 binding at the cell surface, as demonstrated by the functional T cells seen in a patient who experienced a spontaneous mutation reversion in a single T cell precursor (Stephan, Wahn et al. 1996).

Three individual groups demonstrated that they were able to reconstitute the T, B and NK cells lacking in γc deficient mice by transducing bone marrow with retroviral vectors to deliver human *IL2RG* (Lo, Bloom et al. 1999) or murine *il2rg* (Otsu, Anderson et al. 2000; Soudais, Shiho et al. 2000). All 3 studies demonstrated long term reconstitution with no toxic effects.

Gene therapy trials have been carried out on infants with no HLA-identical donor using gammaretroviral vector mediated gene transfer of *IL2RG* to autologous CD34+ HSC for subsequent re-infusion. 10 patients were recruited in Paris (Cavazzana-Calvo 2000; Hacein-Bey-Abina, Le Deist et al. 2002; Cavazzana-Calvo, Lagresle et al. 2005), 10 in London (Gaspar, Parsley et al. 2004) and 1 in Sydney who had a T-B+NK+ phenotype (Ginn, Curtin et al. 2005). Of the total 21 SCID-X1 patients, 20 developed T and NK cells plus functional B cells. Furthermore, 15 patients had polyclonal T cell populations reaching normal levels (Gaspar, Parsley et al. 2004) from a limited number of transduced progenitor cells (Schmidt, Hacein-Bey-Abina et al. 2005). The level of reconstitution reflected the number of transduced CD34+ cells the patient received. The French trial observed clearance of Varicella zoster virus, responses to vaccination, immunoglobulin production and the development of a normal size thymus in their patients. Their patient who did not develop immune cells as a consequence of gene therapy had accumulated the transduced cells in their spleen and subsequently underwent a BMT from an unmatched donor (Cavazzana-Calvo, Lagresle et al. 2005). The British trial observed responses to vaccination and *Candida* and immunoglobulin

production (Gaspar, Parsley et al. 2004). The Australian patient initially responded well but after an upper respiratory tract infection and neurological complications was subsequently treated by BMT from an unrelated donor (Ginn, Curtin et al. 2005).

The overall observations of the clinical trials for retroviral gene therapy of SCID-X1 holds much promise for permanent disease treatment with a better success rate than haploidentical HSCT. Unfortunately, due to the uncontrolled integration of these viral vectors, 5 patients went on to develop a lymphoproliferative disease 2-6 years after gene therapy.

1.3.4 Severe adverse effects of the SCID-X1 clinical trials

The key events leading to T cell acute lymphoblastic leukaemia (T-ALL) in these patients resulted from insertional mutagenesis due to the viral vector transformation of the treated cells. 4 patients were reported from the French trial (Hacein-Bey-Abina, von Kalle et al. 2003; Marshall 2003; Check 2005; Baum 2007) and 1 from the British trial (Howe, Mansour et al. 2008). Longitudinal sample analysis revealed that the copy number did not increase over time (Hacein-Bey-Abina, Garrigue et al. 2008) and Southern blots confirmed there was no presence of env, reverse transcriptase or transposons (Hacein-Bey-Abina, Von Kalle et al. 2003) negating the possibility that the viral vector had become capable of replication.

In 4 of these patients, the clonal expansion correlated with vector integration sites near to or within *LMO2,* a gene implicated in T-ALL (Boehm, Foroni et al. 1991), resulting in its expression (Hacein-Bey-Abina, Garrigue et al. 2008; Howe, Mansour et al. 2008). A second insertion site was also found in one patient located near the T-ALL associated gene *BMI1* resulting in up-regulated gene expression. The remaining patient who developed cancer in the absence of LMO2 interference had a vector integration site upstream of the transcription start site of the oncogene *CCND2* and quantitative mRNA analysis revealed high expression levels in the patient's leukemic cells, consistent with T-ALL patients.

As these oncogenes are not expressed in mature T cells, their expression was attributed to the LTR viral promoter/enhancer which also appeared responsible for the up-regulation of nearby genes (Howe, Mansour et al. 2008). This widespread enhancer effect of the LTR was therefore considered a major determinant in the development of

the cancer. In addition, the LTR of Moloney murine leukemia virus (MoMLV) on which this therapy vector was based has been shown to activate genes at long distances in rats (Lazo, Lee et al. 1990). Furthermore, MoMLV has a preference for actively transcribed genes with an insertional profile near transcriptional start sites (Wu, Li et al. 2003). Further study into the relationship of the LMO2 gene and T cell development revealed that LMO2 is actively transcribed in CD34+ HSC and down regulated during T cell development, this may have been overridden as a consequence of the proximal LTR compounded by the pressure for T cell proliferation (Pike-Overzet, de Ridder et al. 2007).

Both groups observed that the leukemic clones had acquired oncogenic rearrangements associated with T-ALL, indicative of a multistep pathway leading to tumorigenisis and a reasonable explanation for cancer progression considering the late onset of presentation. With the suggestion that these cancers were not due to a single transformation event, it was speculated that *IL2RG* itself aided cancer progression as it has been shown to be dysregulated alongside *Lmo2* in a proportion of murine leukaemias (Dave, Akagi et al. 2009) and thymic lymphomas develop in mice overexpressing γc (Woods, Bottero et al. 2006). This theory was subsequently confounded with the observation that an excessive amount of retroviral vector was used to achieve high expression, increasing the likelihood of insertional mutagenesis, *IL2RG* had not been shown to be oncogenic in previous studies and Jak3 kinase was not constitutively active in the patient leukaemias (Pike-Overzet, de Ridder et al. 2006).

It can be concluded from the SCID-X1 gene therapy clinical trials that although the efficacy of the treatment to reconstitute the immune system was high, the long term production of immune cells conferred by transduced multipotent HSC also included a risk of leukaemia in approximately 1 in 4 patients as a direct result of viral vector insertion. Consequently, all but one of these patients have responded to chemotherapy and survived.

1.3.5 Improved retroviral vectors

Since the adverse events resulting from the retroviral vector treated cells, the LTR has been identified as a likely candidate for pathogenicity due to its strong enhancer/promoter properties. This has been improved with the incorporation of Self-Inactivating (SIN) LTR which lose function after reverse transcription in the host cell

(Figure 1.3a). By deleting the 3' U3 sequence in the ssRNA viral genome, the duplication event of the 3' end in order to initiate 5' cDNA synthesis means the dsDNA integrated into the cell has inactive LTR ends (Yu, von Ruden et al. 1986) (Figure 1.3b). In a comparative study SIN retroviral vectors were shown to be less likely to cause bone marrow cell immortality *in vitro* than the original LTR retroviral vector (Modlich, Bohne et al. 2006). In the absence of a viral promoter, the promoters of housekeeping genes can be introduced such as PGK or EF1α which do not have far reaching enhancer activity (Zychlinski, Schambach et al. 2008) and can include regulatory elements that confer tissue specific expression.

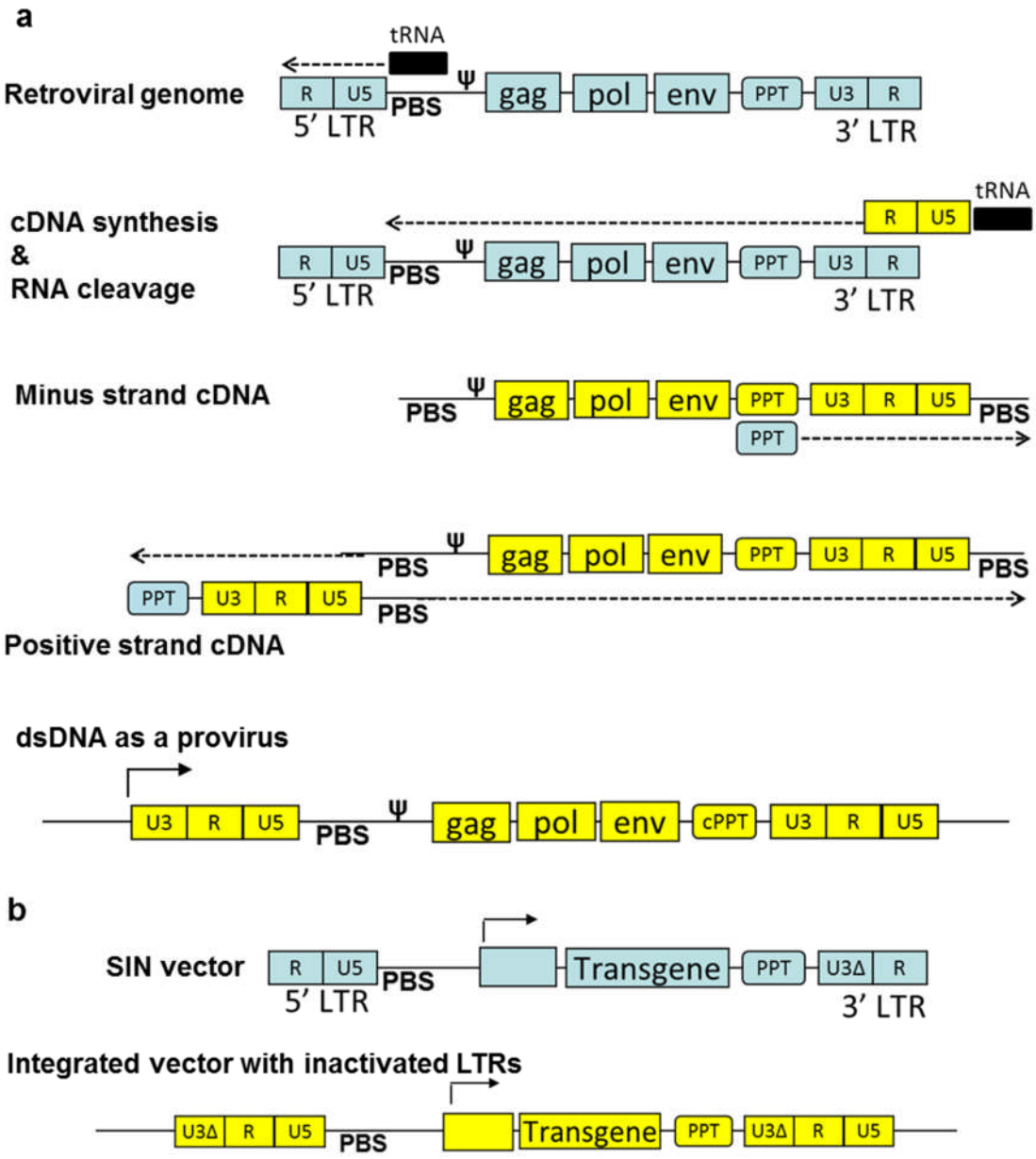


Figure 1.3 Reverse transcription of HIV and the generation of the SIN lentiviral vector.

(a) Reverse transcription of the retroviral genome. Host tRNA primer anneals to the viral genome's (blue) primer binding site (PBS). Viral reverse transcriptase (RT) synthesises minus strand cDNA (yellow) from the 5' end of the genome and copies the U5 and repeat (R) sequence of the viral genome. This minus strand disassociates and transfers to the 3' R sequence. Minus strand cDNA synthesis continues to copy the viral genome up to and including the PBS. Meanwhile the RNase H activity of RT cleaves the RNA leaving the 3' polypurine tract (PPT) intact which then acts as a primer to synthesis the positive strand cDNA from the 3' end of the minus strand cDNA. The result is a dsDNA with the U3, R and U5 (LTR) sequences originating from the 5' end flanking the genome on either-side. (b) The self-inactivating (SIN) lentiviral provirus. By deleting the U3 sequence in the 3' LTR, the resulting provirus does not have functional LTRs and allows for the incorporation of different promoters.

Other modifications to viral vector design have been made to aid production and transduction. In order to alter the vector's tissue tropism, different envelope proteins may be incorporated into the packaging vector, a process known as pseudotyping. Where the clinical trials used the envelope of moMLV (Hacein-Bey-Abina, Le Deist et al. 2002; Ginn, Curtin et al. 2005) or the gibbon ape leukemia virus (GALV) (Gaspar, Parsley et al. 2004)a wider tissue tropism is asserted by the vesicular stomatitis virus glycoprotein (VSV-G) (Naldini, Blomer et al. 1996) and is more commonly used. Another addition was the woodchuck post-transcriptional regulatory element (WPRE) which aids nuclear export and the accumulation of viral RNA in the cytoplasm (Donello, Loeb et al. 1998; Popa, Harris et al. 2002) and was found to enhance viral vector titre (Schambach, Bohne et al. 2006). All of these modifications can be applied to all retroviruses.

HIV based lentiviral vectors are created with the same design as the Gammaretroviral vectors but with the inclusion of their *rev* protein in the packaging constructs as *Rev* binds the rev response element (RRE) and is thought to aid nuclear export of viral RNA and viral encapsulation (Colleaux, D'Auriol et al. 1988) and therefore the RRE sequence is included alongside the therapeutic gene.

Unlike Gammaretrovirues, Lentiviruses are able to transduce quiescent cells. This makes them a more attractive candidate for gene therapy as they avoid the necessity for cell cycle progression during *ex vivo* transduction which can cause the loss of stem cell qualities and will cause cell cycle progression genes to be openly transcribed and more accessible for vector integration. The central polypurine tract (cPPT) of HIV-1 aids nuclear import of the viral genome (Zennou, Petit et al. 2000) and lentiviral transduction was enhanced by the inclusion of the cPPT in the viral vector design (Follenzi, Ailles et al. 2000). Another difference between these vectors is the integration profile. Where MoMLV has a preference for actively transcribed genes with an insertional profile near the transcriptional start site, HIV is more likely to insert within the gene (Schroder, Shinn et al. 2002; Wu, Li et al. 2003). This pattern has also been shown in MoMLV vectors and HIV based SIN lentiviral vectors (Montini, Cesana et al. 2006). Although this integration pattern does not defy insertional mutagenesis, SIN-Lentiviral vectors did not accelerate tumour progression in tumour prone mice compared to the MoMLV vector. However, MoMLV is a murine leukaemia virus

whereas HIV is not and the lack of viral promoter/enhancer in the SIN LTR is more likely to confer the reduced incidence of cancer than it's integration profile as HIV has been implied in causing B-cell lymphoma by integrating upstream of the *c-fes/fps* proto-oncogene (Shiramizu, Herndier et al. 1994) however this is achieved by WT virus.

Viral vector mediated gene addition strategies have proven to be beneficial due to the long term gene expression as a result of gene insertion and although safer vector constructs have been engineered, it cannot truly be pre-determined whether treatment with SIN-Lentiviral vectors modified to control gene expression will prevent pathogenic transforming events. Such events may include bystander enhancer effects leading to proto-oncogene activation, intergenic insertion which could disrupt a tumour suppressor gene or due to the splice donor and acceptor sites within the integrated sequence, aberrant splicing can lead to mRNA species consisting of host cellular mRNA fused to viral mRNA with unknown pathogenic potential as seen in a recipient of an integrating therapeutic vector for β-thalassaemia (Cavazzana-Calvo, Payen et al. 2010). Due to the relatively unpredictable location of integration, gene dysregulation may still be unavoidable with the viral vector gene insertion strategy.

Another complication arising from viral vector gene insertion is the host defence mechanism of viral DNA silencing. DNA methylation is a normal process in gene expression regulation and attained by the transfer of methyl group to CpG dinucleotides by methyl transferases. Methylation of CpG islands in viral promoters can occur and silence viral gene expression which has been observed with MoMLV vectors in mHSC (Challita and Kohn 1994). Progress in vector design has incorporated viral promoters without methylation sensitive CpG elements in order to maintain long-term gene expression, for example the ubiquitously actin chromatin opening element (Zhang, Thornhill et al. 2007).

1.3.6 Non integrating viral vectors

The gene delivery efficiency mediated by viral transduction can still be utilised without active integration with the use of integrase deficient lentiviral vectors (IDLV) (Leavitt, Robles et al. 1996). By mutating the integrase gene the lentiviral vector will not actively integrate into the host genome and remain episomal either as linear dsDNA or the LTR ends can be joined by end to end or can recombine resulting in circular DNA (Wanisch and Yanez-Munoz 2009). IDLV can either transduce dividing cells and

meditate transient expression as the episomal vector cDNA is lost with cell division or transduce non-dividing cells for long term gene expression (Yanez-Munoz, Balaggan et al. 2006). However background integration can still occur via integrase independent recombination mechanisms between vector and chromosomal DNA due to regions of microhomology and homologous recombination.

1.4 Targeted gene correction

Although retroviral gene therapy is an effective method of treating genetic disease, due to random integration leading to insertional mutagenesis, the necessity to develop a safer genome modification strategy is a priority for gene medicine. An alternative strategy for gene therapy is to correct the genetic mutation *in situ*. Genomic modifications can be executed without the need for insertion vectors by harnessing the cell's HR machinery. It has been estimated that HR events occur at a frequency of 1 in 10^6 mammalian cells (Vasquez, Marburger et al. 2001) however this rate of recombination can be increased by introducing a double strand break (DSB) and thereby activating cellular DNA repair mechanisms. This resolution will permit endogenous gene expression and will avoid any viral vector "footprint" left in the genome. A DSB can be introduced in a specific sequence of DNA by enzymatically active molecules such as Zinc finger nucleases (ZFN), Transcription Activator-Like Effector Nucleases (TALEN) and Meganucleases.

1.4.1 Homologous recombination (HR)

HR is an endogenous gene editing process which brings about the exchange of genetic material between DNA sequences with high sequence similarity. During meiosis, HR is employed by the cell to exchange genetic material between homologous regions of parental chromosomes, creating new DNA combinations. HR is also essential for genome maintenance and capable of repairing DNA lesions in the form of DSB which can lead to deleterious mutations or cell death. Repair in somatic cells by HR is possible due to the diploid karyotype as the undamaged chromosome can act as the repair template. In addition, HR usually occurs after DNA replication, in the S/G2 phase of the cell cycle, so the homogenous sister chromatid can act as the repair template. Using the correct DNA sequence to repair a DSB is more beneficial to the cell than the alternative mechanism, non-homologous end joining (NHEJ). NHEJ occurs throughout the cell cycle and acts only to re-ligate the broken ends of the DNA, an error prone

system which frequently involves deletions or insertions of nucleotides resulting in sequence frame shift and/or mutation.

1.4.2 Models of HR.

As pathways of HR are varied and coupled to specific cell fate, this section will focus on the DNA repair pathways that can be engaged as a result of targeted gene DSB with and without provision of repair template.

When DNA replication is halted due to a DSB, a signal is generated to activate the transcription of DNA repair genes, for example those used in HR. Models of HR and the proteins which regulate the process of eukaryotic HR are reviewed by Sung and Klein (Sung and Klein 2006) and are depicted in Figure 1.4. In short, the 5' ends of the DSB are resected leaving behind 3' overhangs. This single stranded DNA is bound by the repair protein Rad51. Rad51 is a DNA dependent ATPase able to support a DNA synapse of both single stranded and double stranded DNA and catalyse directional branch migration with ATP. In humans the presynaptic filament can also be formed by Dmc1. The presynaptic filament can search for, invade and pair with undamaged homologous DNA, forming a Holliday structure and, in so doing, is referred to as the invading strand. As a result of this strand invasion, the unpaired strand on the recipient DNA is displaced forming a D-loop. Having synthesised genetic material to accurately repair the DSB, the synapse can be resolved in a number of ways. Firstly, the double strand break repair (DSBR) model is possible whereby a subsequent strand exchange results in a second Holliday junction. This junction can be resolved to generate repaired DNA molecules with or without homologous DNA crossover (Szostak, Orr-Weaver et al. 1983). Either the double Holiday junction is cut in the orientation it was formed resulting in the exchange of the joining molecule only, this is known as non-crossover. Alternatively, the junction is resolved as the isomer and the non-invading strands are cut, resulting in crossover. The latter resolution is regarded as meiotic crossing-over between the parental chromosomes. Alternatively, a process distinct from DSBR and independent of a second Holliday junction is the synthesis dependent strand annealing (SDSA) model of HR (Allers and Lichten 2001). In this case the newly synthesised DNA is displaced from the D-loop and anneals with the other DSB end. This

model always results in non-crossover DNA resolution and is more likely to be utilised in mitotic cells.

Non-crossover resolution of homologous DNA is followed by a multistep repair pathway resulting in the exchange of genetic material between homologous strands whereby the information on the DNA with the DSB is replaced by the homologous information of the repair template and *vice versa*. Homologous regions between chromosomes are likely to be the same allele of a gene, however, if the alleles differ, the recombination event may result in gene conversion. Gene conversion following a DSB is exploited by certain organisms as discovered by the identification of meganucleases in *Saccharomyces cerevisiae*.

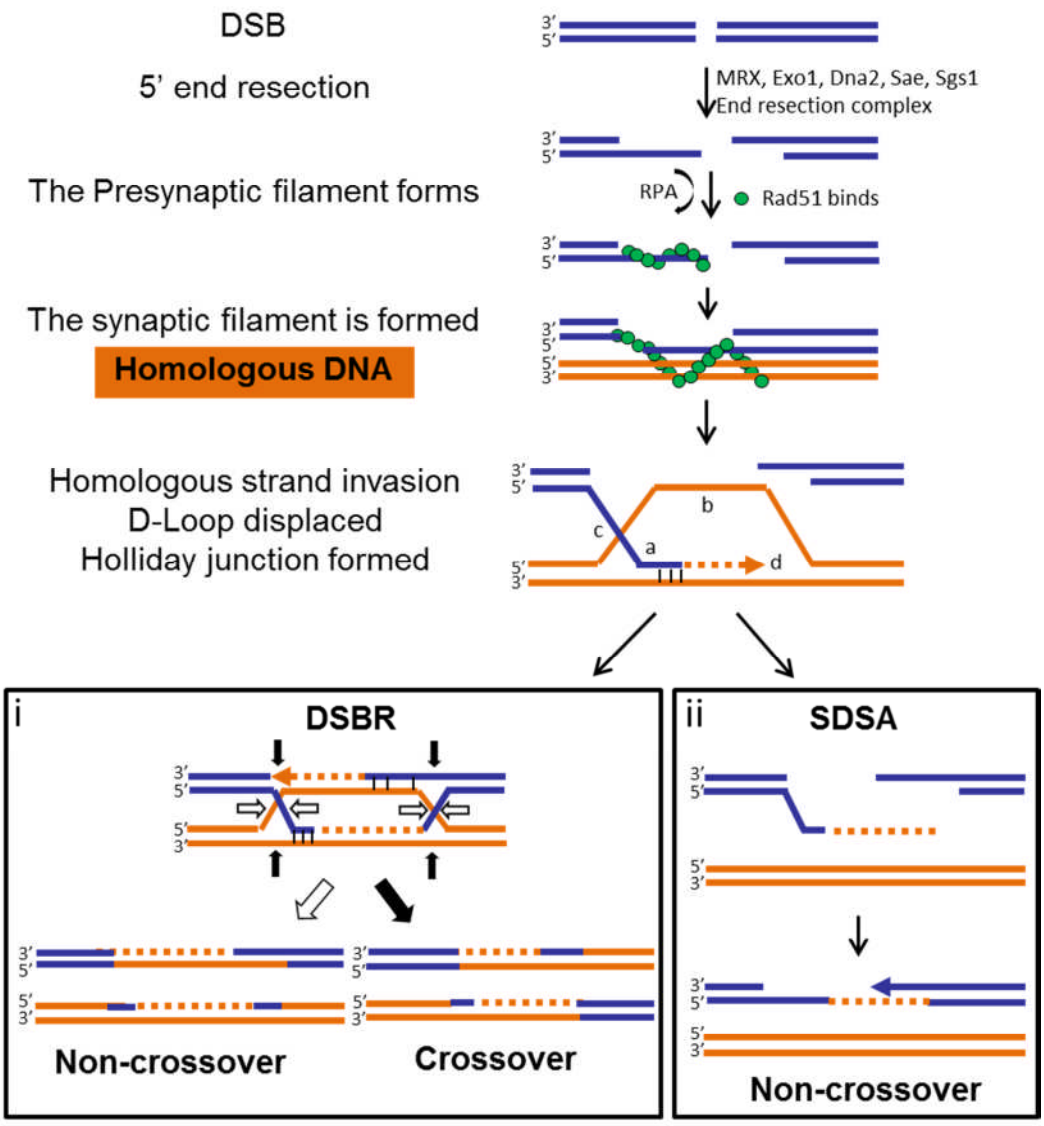


Figure 1.4 A schematic of the events that follow a double strand break (DSB) and resolved by homologous recombination (HR) in S. cerevisiae.

Limited degradation occurs at the 5' ends by a nuclease complex containing MRX, Exo1, Dna2 and Sae plus the helicase Sgs1 in. After 5' end resection a 3' single stranded DNA (ssDNA) overhang remains. Rad51 polymerises on the single strand to form the presynaptic filament, overcoming competition for ssDNA with replication protein A (RPA) by other proteins from the Rad52 gene group such as Rad52 and the Rad55-Rad57 complex. The presynaptic filament (a) invades and displaces homologous DNA resulting in a D-loop (b) and forming a Holliday junction (c). DNA is synthesised from the 3' end of the invading strand (d). Resolution of the D-loop. (i) Double strand break repair (DSBR) resolution results in Non-crossover (White arrows) or crossover (Black arrows). (ii) Synthesis dependent strand annealing (SDSA) non- crossover resolution of a single Holliday junction.

In addition to the DSBR and SDSA models of HR there are further, non-conservative types of DNA repair. In the event that a DSB occurs in the middle of 2 repeat sequences, the 5' degradation at the cut ends can lead to the exposure of single stranded complementary homologous DNA which anneal allowing DNA synthesis to repair the DSB. This form of recombination is known as single-strand annealing (SSA) and results in the loss of DNA (Figure 1.5a). Another non-conservative DNA repair model is NHEJ which does not depend on the availability of homologous DNA and acts to re-ligate the broken DNA ends. This process frequently incorporates or deletes nucleotides but can sometimes faithfully restore DNA to its WT sequence (Figure 1.5b).

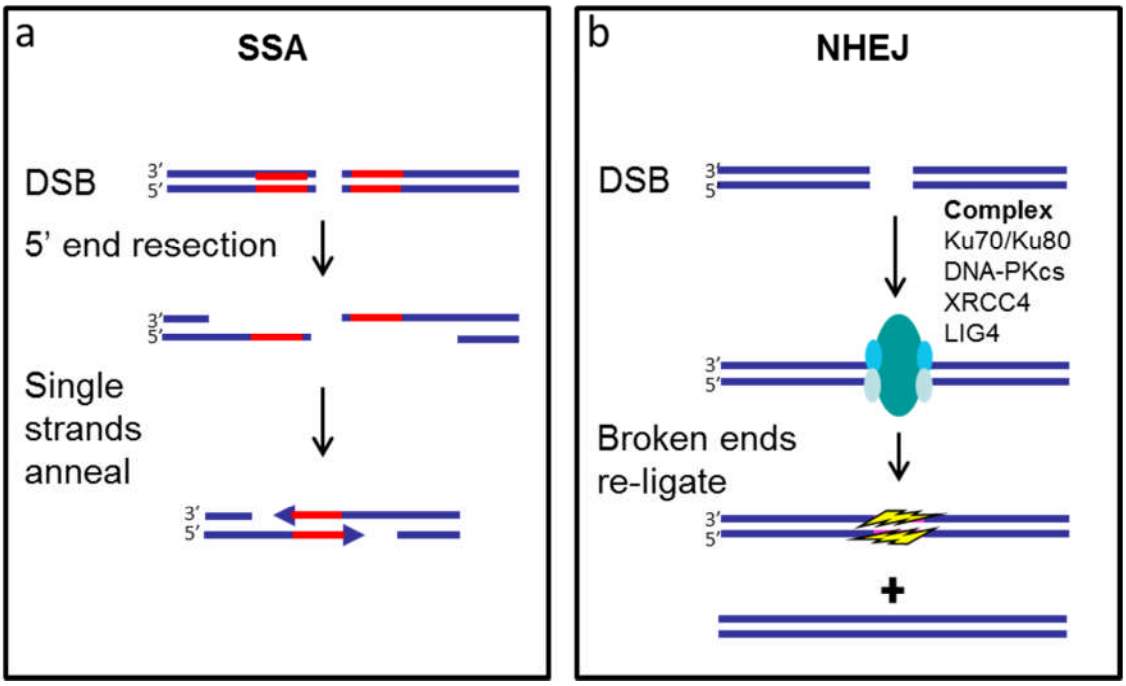


Figure 1.5 DSB stimulate non-conservative repair pathways.

(a) SSA, repetitive DNA (Red) either side of a DSB undergo 5' end degradation and the single strands anneal followed by DNA synthesis. (b) NHEJ in mammals, the broken end of the DSB protected by Ku proteins (Ku70 & Ku80) in association with DNA-PK_{cs} which activates XRCC4 which interacts with DNA ligase IV (LIG4) resulting in DNA with or without mutations.

1.4.3 Introducing novel DNA into eukaryotic genomes by HR

In order to modify the endogenous genome by HR, a molecule of extragenous DNA can be introduced into a known locus as part of a donor DNA repair matrix. It is possible to integrate genes carried on a bacterial plasmid into the yeast genome via transformation. Hinnen *et al*. transformed *leu2*- mutant yeast with donor DNA comprising of a segment of the yeast chromosome III with the LEU2+ gene carried by the *E.coli* plasmid ColE1 (Hinnen, Hicks et al. 1978). Transformants able to grow in leucine negative media occurred at a rate of 10^{-7} where gene targeting frequency = targeting events/total cells transfected. 71% of cells had transformation events in chromosome III that showed close linkage to the original *leu2*- gene. 12% had an unchanged chromosome III and linkage analysis of the LEU2+ and *leu2*- gene confirmed that the transformation event had occurred on a different chromosome. No evidence of donor DNA was detected in the remaining transformants as in the untreated control. As revertants had not been seen in this model before, it was suggested that LEU2+ was incorporated by conversion or double crossover. This study explored the opportunity of harnessing HR to modify endogenous eukaryotic genome with prokaryotic plasmid DNA. It also revealed the possibility that recombination can occur off site.

Engineering of mouse strains to investigate genetic disease has relied on HR events to incorporate or excise genetic material. Capecchi et al. found the rate of recombination in mES when microinjecting donor DNA to be, on average 10^{-3} per treated cell (Thomas, Folger et al. 1986). Due to the labour intensive process of microinjection, the system was developed to include double selection to engineer knockout (KO) strains where extragenous donor DNA containing the gene of interest (GOI), disrupted by a neomycin gene (*neo*) and with a tyrosine kinase gene (*tk*) outside the region of homology. Murine embryonic stem (mES) cells were electoporated and the cells that had incorporated the donor DNA by random integration were negatively selected by gancyclovir treatment, and HR events were positively selected by G418 treatment (Mansour, Thomas et al. 1988). Out of the cells that had survived electroporation, up to 5.4×10^{-6} survived positive and negative selection.

1.4.4 DSB increase the rate of HR

Early studies in HR revealed that cleaving the DNA increases the frequency of recombination. Orr-Weaver *et al*. 1961 showed that transforming yeast with linearised plasmid vector cut in the region of homology increased plasmid integration up to 1000 fold compared to circular vector (Orr-Weaver, Szostak et al. 1981). The ability of transformed yeast to integrate linear plasmid DNA was greatly reduced in *Rad52* mutant strains however this was not observed for circular plasmids indicating different processes for these events. It was also shown that introducing the cut site outside of the homologous region reduced the efficiency of recombination. To add weight to the advantage of introducing a DSB into the vector, plasmids containing multiple homologous regions to distinct chromosomal sites only integrated at the chromosomal site corresponding to the cut homologous region. Furthermore plasmids made linear with a 300-1200bp gap were repaired by homologous chromosomal DNA.

It was later identified that 3 genetic modifications could occur on transforming yeast with a plasmid made linear by removing 60bp from its region of yeast homology (Orr-Weaver and Szostak 1983). (i) The gapped plasmid was repaired with yeast chromosomal DNA, the rational for this event was non-crossover HR (ii) the plasmid was integrated into the yeast chromosome in the region of homology. In this instance, it was reasoned that the D-loop created by the plasmid's 3' invading strand migrated by DNA synthesis until the plasmids 5' break was met and the free end ligated to the yeast chromosomal DNA and therefore the plasmid had integrated (iii) the broken ends of the plasmid were re-ligated by NHEJ.

Lin, Sperle and Sternberg continued the assessment of DSB stimulated HR in mammalian cells (Lin, Sperle et al. 1984). Phage constructs containing defective tyrosine kinase (*tk*) genes were delivered to mouse L cells (mouse fibroblast cells, strain L (Miyamoto, Zeuthen et al. 1973)), where recombination between 2 defective copies would result in a functional *tk* gene (*tk*+ transformants). Firstly, a construct containing 2 defective *tk* genes in the same orientation and sharing 984bp of homology, were separated by 3kb of non *tk* DNA. Cleavage by a restriction enzyme in the 3kb separating region resulted in intramolecular recombination leading to *tk*+ transformants at a rate 30 fold higher than un-cut control. Intermolecular recombination between the defective *tk* genes on separate plasmids was as efficient as the intramolecular recombination however no significant increase in *tk*+ transformants

was observed when cleaving only one plasmid compared to uncut. Therefore the type of HR observed here was SSA due to the obligation of 2 exposed homologous genes to anneal. It was also noted that a region of 320bp homology shared between the 2 plasmids resulted in a transformation efficiency of 7.2%. When the homologous region was reduced to 46bp the transformation efficiency dropped to 0.2%. These findings show that the rate of recombination declines as the region of homology is reduced.

Further studies of intermolecular recombination between plasmids in mammalian cells verified the relationship between the length of homology and rate of recombination. Gradually decreasing the homologous region from 500bp to 25bp caused a drop in recombination in both in monkey COS cells and EJ human bladder carcinoma cells (Ayares, Chekuri et al. 1986). Although a homologous region of 25bp is sufficient for recombination between plasmids, a greater length of 295bp is necessary to achieve HR between chromosomes (Liskay, Letsou et al. 1987). An explanation for the greater extent of homology required between chromosomes is to avoid any fortuitous recombination between repeat sequences.

These observations provided evidence that the rate of HR is increased when DNA damage in the form of a DSB is experienced in mammalian cells and that the type of DNA repair varies. Furthermore, HR depends on the extent of homology provided by the donor DNA.

1.4.5 Meganucleases

Considering a DSB stimulates HR, it is an attractive prospect to control the site of DNA repair with reagents that can increase the rate of gene targeting at a pre-determined location in the genome and thus avoiding the unpredictable integration profile of retroviral vectors that may lead to insertional mutagenesis.

Meganucleases are sequence specific endonucleases which can recognise target DNA of ≥12bp (Thierry and Dujon 1992), greater than the 6-8bp recognition site necessary for the rare cutter type II restriction endonucleases. The meganuclease I-SceI recognises an 18bp sequence (Colleaux, D'Auriol et al. 1988) and was shown to stimulate SSA in extrachromosomal DNA in mouse cells (Rouet, Smih et al. 1994). Furthermore, there is no naturally occurring I-SceI recognition site in the mouse genome and the highly specific endonuclease activity was confirmed by the lack of

genotoxicity in the cells. Early gene targeting studies in mammalian cells using the meganuclease I-SceI, are listed in the overview by (Paques and Duchateau 2007) and although gene targeting does not represent targeted HR it is of note that gene targeting frequencies of 10^{-1} can be achieved in human 293T cells (Alwin, Gere et al. 2005). However, these experimental models require the pre-requisite of introducing the naturally occurring I-SceI target site into cell types which do not normally harbour this sequence. When the LAGLIDADG protein motif was identified in the so called family of meganucleases, shuffling of these residues revealed that the recognition sequence could be modified. Further identifying 2 subdomains and engineering 4 mutants gave rise to a practical approach to modify the DNA binding domain of this protein as shown by a meganuclease engineered to cleave the human Rag1 gene (Smith, Grizot et al. 2006).

Given the development of engineered proteins to cut predetermined sequences within the genome, it is possible to introduce novel DNA at these sites via HR. For this technology to progress certain factors must be considered in order to permit HR with extragenous DNA. The length of the homologous tract on exogenous donor DNA for endogenous gene targeting is vital to the efficiency of HR mediated repair. However, an ensuing question is the necessary extent of homology in this homologous tract. In order for a mutant gene to be corrected, some change in sequence must occur and therefore some disparity in gene sequence will be expected between the chromosomal and donor DNA.

Important factors concerning the disparity of donor DNA repair template were investigated by Elliot *et al*. (Elliott, Richardson et al. 1998). Analysis of gene conversion between plasmid and chromosomal DNA was studied in mES cells modified to contain the I-SceI target site in their genome. Using I-SceI to induce a DSB in the chromosome, plasmid donor DNA was provided containing increasing numbers of silent mutations. It was observed that as the sequence disparity increased by 5, 7 and 8bp the rate of targeted HR decreased by 2.5, 4 and 6 fold respectively. Furthermore, silent mutations close to the DSB were more likely to be incorporated than the silent mutations further upstream. This indicated that the extent of gene tract conversion is finite. 68% of the analysed clones had gene conversion tracts of 58bp and the longest was 511bp seen in just 2.5%. The combined results gave an estimated average of a 200bp conversion

tract. Another interesting observation made from these recombinants was the direction of repair from the DSB. The transformants that displayed a short tract of gene conversion also displayed unidirectional conversion. Conversely, the transformants that had longer gene conversion tracts showed bidirectional conversion. This could be interpreted as evidence that 2 different HR pathways can be employed to repair the DSB in this model. For example, SDSA could account for the unidirectional gene conversion due to the nature of the single strand synthesis. Alternatively non-crossover resolution of DSBR could account for the bidirectional repair observed with more extensive gene conversion due to the gene synthesis that occurs off of both strands of the donor DNA. Similar observations were made where the length of homologous region dictates the nature of HR and longer homologous regions in the donor DNA can retain homology after 5' end resection allowing for a second Holliday junction to form and be resolved by DSBR. In contrast, short homology regions, 25-35bp can lose homology by 5'end resection and are more likely to recombine via SDSA (Ayares, Chekuri et al. 1986).

So far it has been established that the cell's endogenous repair mechanism, HR, can be adopted to integrate novel DNA from plasmid DNA into a mammalian genome as long as the prerequisites for HR are available. It is also possible to greatly increase the rate of HR by the introduction of a DSB which can be administered to a target site by nucleases thus resulting in gene targeting.

1.5 Zinc Finger Nucleases

Like meganucleases, ZFN can be manufactured to confer specificity for and cleave a target site in the genome increasing the rate of HR (Porteus and Baltimore 2003). ZFN are chimeric proteins consisting of a DNA binding domain made up of an array of zinc fingers (ZF) joined to a FokI nuclease domain. The FokI nuclease domain is active as a dimer so 2 ZFN monomers pair at the DNA target site in order to achieve cleavage, the structure and DNA recognition of a ZFN pair is shown in figure 1.6a. Due to the modular nature of ZF, there are fewer constraints than meganucleases for modifying the DNA binding domain.

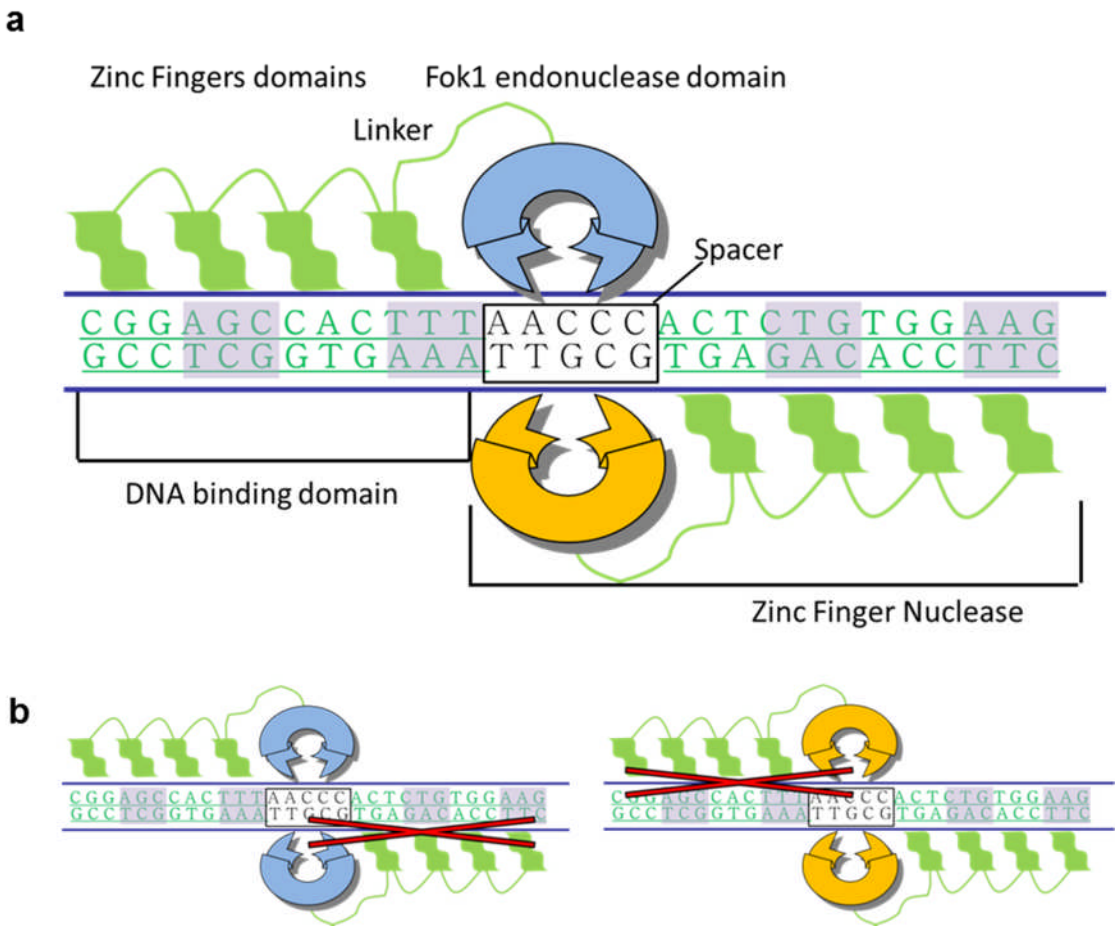


Figure 1.6 Binding of the ZFN heterodimer.

(a) A diagram of the Zinc finger nuclease (ZFN) chimera as shown as a dimer. Multiple zinc fingers are joined to heterodimeric FokI nuclease subunits via a linker domain and the ZF binding domains are separated by a spacer region. (b) Homodimers are avoided by the repelling forces of the Fok1 domains and the lack of DNA recognition by the ZF.

Once the chromosomal target site has been activated by a DSB, subsequent modification of the gene via the HR pathway can include correction *in situ* by the exchange of base pairs, targeted gene addition by the introduction of a new DNA sequence and gene excision or gene disruption by NHEJ (Figure 1.7). ZFN mediated gene targeting is an attractive alternative to retroviral gene therapy as it avoids the introduction of viral elements into the host genome and the target site is pre-determined.

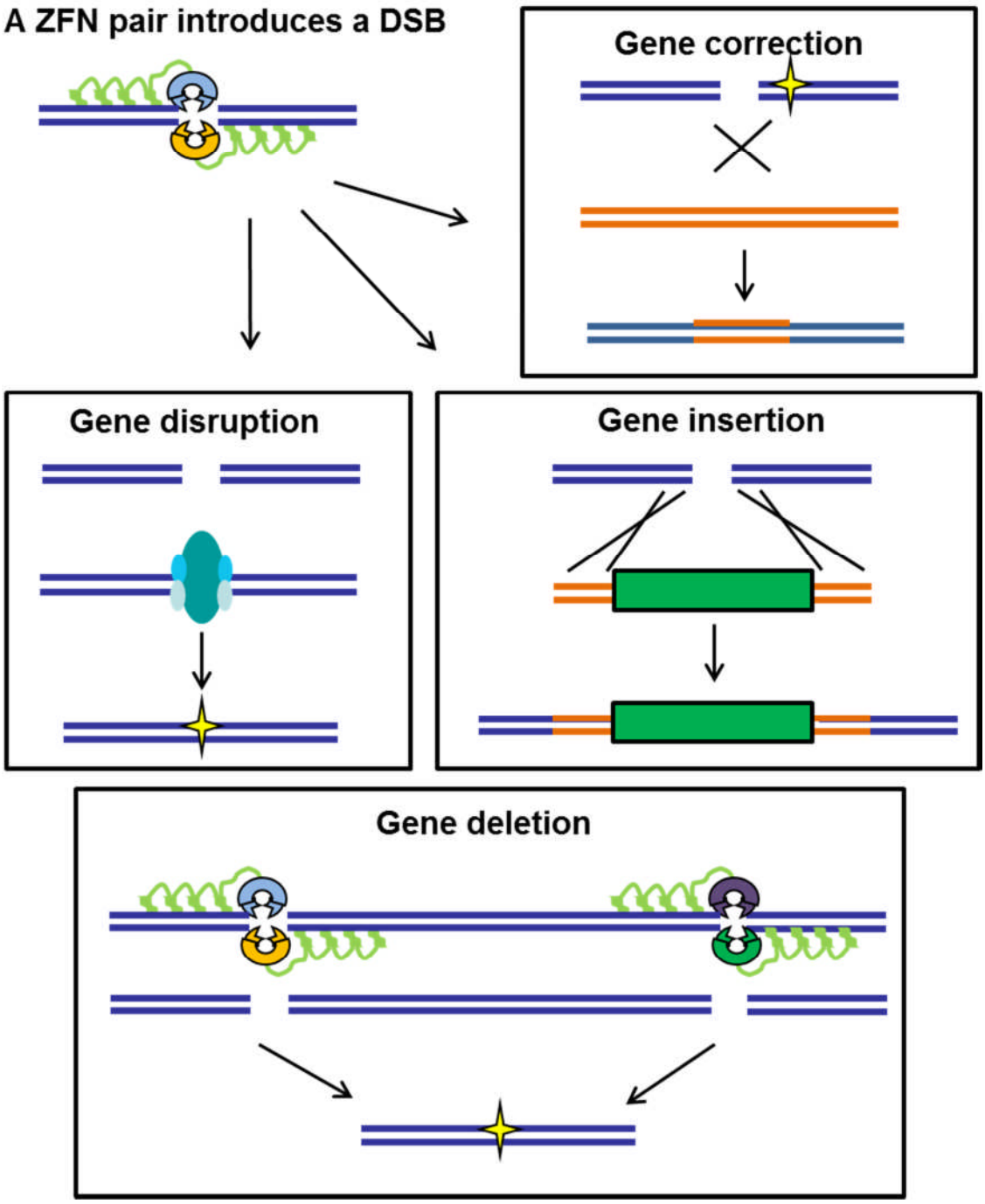


Figure 1.7 The preferential outcomes of a DSB

Gene correction of a mutation (yellow star) proximal to the DSB can be mediated on the provision of homologous DNA containing restorative sequence. Gene insertion can be targeted to a DSB on provision of the GOI flanked by sequence homologous to the target locus. Gene disruption results when no homologous DNA is used to correct the DSB frequently resulting in mutation. Gene excision is mediated by two ZFN pairs which are directed to 2 loci resulting in the removal of a stretch of DNA.

1.5.1 The Zinc Finger

Zinc fingers (ZF) were first discovered as part of the eukaryotic transcription factor TFIIIA (Miller, McLachlan et al. 1985). The common Cys_2-His_2 zinc finger comprises of a ββα-fold made up of ~30 amino acids that co-ordinate a Zn+ atom. The helix of one ZF binds the major groove of DNA and typically confers affinity to a triplicate of DNA nucleotides. It is with this DNA specificity of the α-helix protein motif that transcription factors are able to assemble at the promoter sites of specific genes or groups of genes and control their expression. Multiple variants of ZF can be linked to form polydactyl zinc finger proteins (ZFP) (Liu, Segal et al. 1997) designed to recognise DNA targets in multiples of 3bp. Polydactyl ZF are capable of re-directing DNA effector molecules (Blancafort, Segal et al. 2004) and ZF can also be used to bind an unwanted gene and serve to inhibit transcription, such blocking ZF were generated against the BCR-ABL carcinogenic translocation (Choo, Sanchez-Garcia et al. 1994).

1.5.2 Modification of ZF DNA recognition

A complete set of ZF capable of recognising all the possible DNA triplicates would amount to 64 molecules (Beerli and Barbas 2002). The length of DNA sequence to confer exclusivity in the human genome is calculated to be at least 16bp. The meganuclease, I-SceI has an 18bp recognition site which is emulated by fusing 6 ZF domains (Moore, Klug et al. 2001).

Based on 6 residues in the conserved α-helix, ZF with different DNA binding specificities can be rationally designed and selected according to their dissociation constant by using phage display (Segal, Dreier et al. 1999; Dreier, Segal et al. 2000; Pabo, Peisach et al. 2001; Dreier, Fuller et al. 2005). Identifying families of ZF formed libraries from which modular assembly of polydactyl ZF can be constructed. These ZF combinations can also be selected in accordance with the effect of target site overlap seen when neighbouring ZF bind 4bp instead of 3bp. This modular method of designing ZFP has provided effective reagents for gene targeting studies (Bibikova, Golic et al. 2002) however at an estimated success rate of 24% for the assembly of a 3 finger domain (Cathomen and Joung 2008) which is further reduced for a 6 finger domain, has limited the use of this method.

An alternative approach to engineering ZF domains is by bacterial cell based two-hybrid system (B2H). In this method candidate ZF are individually selected for low

stringency for the target site and supported by known ZF. These potential ZF are randomly recombined and selected for high stringency for the target site. This second round accounts for both target site specificity and neighbouring ZF interference (Joung, Ramm et al. 2000; Hurt, Thibodeau et al. 2003). The reagents and method for this strategy were made available to the public as "Oligomerized Pool Engineering" or OPEN (Maeder, Thibodeau-Beganny et al. 2008) and custom ZFN for zebrafish were promptly developed (Foley, Maeder et al. 2009; Foley, Yeh et al. 2009).

An archive of 2-finger modules that are randomly combined and selected by target sequence specificity by a B1H-based selection is also utilised (Gupta, Christensen et al. 2012) and is thought to be a similar strategy to the system employed by Sangamo Biosciences.

Modification of the ZF binding domain has provided novel DNA binding proteins with a variety of applications for gene targeting, one of which is the ZFN which mediates targeted DSB.

1.5.3 The ZF-FokI chimera

Work by the Chandrasegaran laboratory established that the type II restriction endonuclease, FokI, had a non-specific DNA cleavage domain which could be separated from its DNA recognition domain and fused to the Ultrabithorax (*Ubx*) homeobox domain of *Drosophila* with a $(glycine_3Serine)_2$ linker and demonstrated that this chimera could cleave the *Ubx* target site on a pUC13 plasmid (Kim and Chandrasegaran 1994). Subsequently, they combined ZF molecules with the FokI cleavage domain to produce the chimeric ZFN protein (Kim, Cha et al. 1996). Having two $(ZF)_3$ constructs at their disposal, they fused both with FokI to produce 2 individual hybrids. In their analysis of λ DNA digestion by these hybrids, they showed preferential cleavage by one construct and specific, yet unpredicted cleavage by the other. Their findings emphasised the importance of rigorous ZF screening and also highlighted the possible interference by FokI itself and the necessary optimisation of the linker domain between the FokI and ZF.

1.5.4 The ZFN pair

Subsequent to the ZF-FokI hybrid now given the more general term ZFN, FokI was discovered to be active as a dimer (Bitinaite, Wah et al. 1998). Importantly it was

shown that, so long as DNA recognition is provided by one WT FokI monomer, a catalytic dimer can be formed with a second FokI molecule with a disrupted DNA binding domain. Due to this character of the FokI dimer, DNA binding at off target sites would result in cleavage at unwanted sites. This was proven by an early study in the pairing of ZFN monomers by the Carroll-Chandrasagaran collaboration (Smith, Bibikova et al. 2000). Here they used ZFN monomers and target DNA with either a single ZFN target site or inverted repeats of the target site. It was shown *in vitro* that the target DNA with a single ZFN recognition site which would bind one ZFN monomer could be cleaved at high concentrations of the ZFN. Further observations made with the target DNA with inverted repeats revealed that the FokI dimer cleaved in the middle of the 2 ZF binding domains, now known as the spacer region. These findings had implications for the imminent ZFN pair design where the total number of ZF selected for specific DNA binding are distributed equally between 2 FokI monomers. With this design, the ZFN will recognise 2 adjacent DNA binding half sites and the FokI will dimerise and cleave the at spacer region.

Using a similar experimental model but this time targeting extrachomosomal DNA in *Xenopus* oocyte nuclei, the group identified the relationship between linker length and spacer domain. It was seen that longer linkers (up to 18 amino acids) could target a broad range of spacer regions from 6-35bp, conversely, short linkers (0 amino acids to diminishing the FokI domain) were more restricted in the range of spacer length they could target and able to cut short spacers of 4-6bp. This study also attempted the first ZFN pair using ZFN that recognised different sequence which yielded 19% recombination in a plasmid designed to correspond to the respective ZFN (Bibikova, Carroll et al. 2001). The linker to spacer relationship was also reported by Porteus *et al*. (Porteus and Baltimore 2003) and was confirmed by Händel *et al*., who also noted that longer linker domains which conveyed broader activity were also more toxic (Handel, Alwin et al. 2009).

The Carroll group moved their ZFN analysis *in vivo* by constructing two 3-finger ZFN monomers by modular assembly that, when paired, would bind 2 adjacent 9bp “half sites” separated by a 6bp spacer in the *yellow* gene of *Drosophila*. Targeted DSB was measured by the rate of mutation in *yellow* resulting from the NHEJ pathway of repair. The rate of germline mutation was calculated to be 1/250 confirming *in vivo* targeted

DSB (Bibikova, Golic et al. 2002). A subsequent study observed HR events in this model. Circular or linear donor DNA, homologous to *yellow,* were provided alongside the ZFN. The best rate of HR was found using linear donor and occurred in mutant male and female germlines at 63% and 73% respectively (Bibikova, Beumer et al. 2003). Lethality was associated with one of the ZFN used in these studies and explained by its potential to recognise "half sites" throughout the genome. As seen previously, the DNA binding event of one ZFN allowed a second FokI catalytic domain to dimerise and introduce a DSB off target. To circumvent this, a mutation was introduced to inactivate the FokI domain and no lethality was observed in flies as a result (Beumer, Bhattacharyya et al. 2006). Studies carried out by the Carroll group made important observations regarding the DNA repair pathways NHEJ and HR stimulated by ZFN and the toxic effects of off target DSB.

1.5.5 ZFN associated toxicity

In order to make ZFN a viable therapeutic reagent, the degree of damage they exert on the cell must be understood and minimised. DSB are genotoxic and if a second cut is administered, chromosomal translocations can occur >2000 fold more than in the presence of a single cut due to the non-conservative DNA repair pathways NHEJ and SSA (Richardson and Jasin 2000). Toxcitiy can be measured by cell viability and quantification of DNA repair foci components, phosphorylated histone H2AX (ϒ-H2AX) and 53BP1.

By attributing the majority of toxic incidence caused by ZFN to the frequency of DSB they exert, it can be conceived that by constructing ZFN that are only capable of cleaving a single site in the genome cell injury will be reduced. The potential areas for improving the ZFN architecture are: resolving the promiscuity of FokI dimerization, improving the sequence specificity of the DNA binding domain and using a linker with a narrow range of spacer region affiliation.

A solution to the problem of off target DSB due to nonspecific FokI dimerization was developed whereby charged residues introduced at the dimer interface resulted in polar FokI domains which maintained the ability to cleave DNA. This absolved the WT homodimer and forced the FokI monomers to heterodimerize (Figure 1.7b). Obligate heterodimers permit cleavage of the target site by the ZFN pair and greatly reduced the toxicity associated with off target DSB. Miller *et al*. exchanged 2 residues in the

dimer interface of WT FokI which reduced the number of γ-H2AX positive cells from 18.9% to 1.7% with no loss of target site cleavage (Miller, Holmes et al. 2007). Similarly, Szczepek *et al*., demonstrated that 4 ZFN pairs with different catalytically active FokI heterodimers each resulted in a significant reduction of γ–H2AX+ cells compared to WT (Szczepek, Brondani et al. 2007). Further reduction in DNA damage as measured by 53BP1 was reported when testing a FokI heterodimer that combined 3 of the polar residues used in the Miller *et al.* and Szczepek *et al*. studies (Ramalingam, Kandavelou et al. 2011).

Another approach to control the cleavage domain of ZFN is to fuse the ZF to a sequence specific nuclease rather than the catalytic FokI domain. This was done with the type II endonuclease PvuII and no off target cleavage activity was observed (Schierling, Dannemann et al. 2011). This design of ZFN is constrained to cleaving sequence that includes both the PvuII and ZF recognition sites and will limit their application for endogenous genes, however, progress in protein engineering may provide a library of site specific cleavage domains which could be applied to targeted gene therapy.

An inverse correlation exists between ZF specificity and toxicity. Using the B2H system of ZF selection, Cornu *et al*. assembled 3 ZFN for 3 different targets and compared their binding activity and toxicity. They noted that in the majority of cases the ZFN that activated the greatest frequency of targeted HR incurred the lowest incidence of γ-H2AX+ cells (Cornu and Cathomen 2007). A direct comparison of modular and B2H based assembly of ZFN revealed that not only does the B2H method result in ZFN with greater gene targeting efficiency but also achieved 100% cell survival in HEK293 cells whereas ZFN created through modular assembly resulted in a 50% cell survival rate (Pruett-Miller, Connelly et al. 2008). It was speculated that increasing the number of ZF in a ZFN could increase specificity thereby decreasing off target DSB. However, changing the number of ZF from 3 to 6 per ZFN considerably decreased the target activity and showed the optimal finger number to be 3 or 4 per ZFN monomer (Shimizu, Sollu et al. 2011).

In addition to optimising the structure of the ZFN, further adjustments can be made in administering ZFN to reduce genotoxicity. The off target events of ZFN appear to have

a positive correlation with dosage (Pruett-Miller, Connelly et al. 2008) as observed with heterodimeric ZFN designed for zebrafish using the B2H system (Meng, Noyes et al. 2008). Analysis of these off target events incurred at high ZFN doses revealed that sites with 5bp mismatch from the intended target were cleaved at a statistically significant frequency whereas sites with 1 or 2bp differences remained uncut (Gupta, Meng et al. 2011). The unexpected observation that the relationship between target homology and ZFN binding was not necessarily linear prompted more investigation.

In order to minimise and monitor off target ZFN activity, software such ZFN-Site (Cradick, Ambrosini et al. 2011) and Systemic Evolution of Ligands by Exponential Enrichment (SELEX) are available to predict potential ZFN binding sites based on target site homology which account for spacer length. Potential off target sites of the ZFN designed to target the CCR5 gene (*CCR5* ZFN) were predicted *in silico* by SELEX and of the top 15 matches only the closest match, in *CCR2*, had been disrupted *in vitro* at a tenfold lesser extent than *CCR5* (Perez, Wang et al. 2008). Because DSB can capture fragments of linear DNA by NHEJ, the accuracy of *in silico* predictions, based on target site homology was evaluated by mapping the sites of linear IDLV capture at DSB sites incurred by the *CCR5* ZFN (Gabriel, Lombardo et al. 2011). Clustered integration site (CLIS) analysis of the captured IDLV revealed that unintended target sites could differ by as many as 7-8bp from the target site and only represented a portion of the *in silico* predicted sites. This capture assay is thought to give a more accurate representation of off target activity by ZFN. Studying *in silico* predicted sites for *CCR5* ZFN also revealed that increasing the concentration of ZFN leads to more cleavage sites *in vitro* and mutations can be tolerated in a DNA binding “half site” if a ZFN monomer has bound a corresponding non-mutated half site. The threshold of base pair dissimilarity was calculated to be up to 3bp hence ZFN should be designed to recognise a target site with at least 3bp difference than the next closest possible target (Pattanayak, Ramirez et al. 2011). These observations reveal that ZFN targets can be unpredictable and probably not entirely conducted by target site homology. One possible explanation is that targets can be made inaccessible to endonucleases by CpG methylation (Valton, Daboussi et al. 2012).

The positive correlation between ZFN dose and toxicity indicates the necessity to ensure transient ZFN expression in the cell. Transient expression can be deployed in a

number of ways including using an expression vector that does not integrate into the genome, incorporating the ZFN target site in the ZFN expression vector, as demonstrated with I-SceI (Takashima, Sakuraba et al. 2009), adding a ubiquitin tag to the ZFN and therefore marking it for proteasome degradation (Pruett-Miller, Reading et al. 2009) and treating the cell with ZFN protein rather than expression vectors has shown gene targeted frequencies of up to 27% (Gaj, Guo et al. 2012).

As optimisation of the ZFN architecture develops and the associated toxicity is monitored, ZFN gene targeting strategies can be assessed for safer therapeutic intervention.

1.5.6 NHEJ vs HR

Different DNA repair pathways result in different DSB repair outcomes so depending on the gene targeting strategy it is of interest to monitor the type of DNA repair employed by cells treated with ZFN and to determine a strategy that will utilise the desired pathway.

DNA repair pathways are dependent on the type of DNA damaging agent and the phase of the cell cycle. HR is more likely to occur during the S and G2 phases of the cell cycle as indicated by the assembly of Rad51 at the site of DNA damage in these phases (Gasior, Olivares et al. 2001). NHEJ is the main repair mechanism in G1 as shown by increased γ–H2AX foci in this phase when part of the NHEJ repair complex, DNA-PKcs is inhibited (Rothkamm, Kruger et al. 2003). However NHEJ and HR repair pathways are not always mutually exclusive as seen in intrachromosomal gene exchange stimulated by I-SceI where DNA synthesis initiated by HR goes beyond the region of homology and ends with NHEJ resulting in a junction containing random nucleotide deletions (Richardson and Jasin 2000).

The studies in *Drosophila* demonstrated that NHEJ and HR are both stimulated by ZFN induced DSB when a donor DNA is provided. Of the total gene targeted events mediated by *rosy* specific ZFN in female and male parents, HR was observed at frequencies of 59% and 41% respectively and the remaining modifications were a result of NHEJ (Beumer, Bhattacharyya et al. 2006). Inhibition of the *Drosophila* Rad51 homologue significantly reduces HR and disallowing SSA to occur by using circular plasmid inhibits HR completely. Furthermore, inhibition of the *Drosophila* Ligase IV,

part of the NHEJ repair complex, results in a higher frequency of HR events than in Ligase IV+ controls (Bozas, Beumer et al. 2009).

The choice of DNA repair pathway can be moderated by manipulating levels of repair pathway proteins. For example targeted HR will increase 2.1 fold in HT1080 cells overexpressing Rad51 without any increase in random integration and in the absence of targeting endonucleases (Yanez and Porter 1999). However, when considering other members of the *RAD52* epistasis group of genes, chromosomal targeted HR is inhibited by Rad52 overexpression and this environment significantly increases cell morbidity (Yanez and Porter 2002).

The availability of sister chromatids, hence a homologous copy of DNA, is the preferred template for DNA correction (Kadyk and Hartwell 1992) and their availability after replication in S phase reflects the increased incidence of HR at the S/G2 phase. However the possibility that donor DNA template could compete with NHEJ to initiate repair by HR at G1 is unlikely due to blocking of end resection and therefore the presynaptic filament in this phase (Shrivastav, De Haro et al. 2008).

It can be inferred from these observations that the incidence of NHEJ is unavoidable when incurring a DSB and that inhibition or support of a particular DNA repair pathway may be an opportunity to promote the desired repair however this may result in cell damage.

1.5.7 ZFN as therapeutic reagents

Up to this point general considerations of ZFN mediated gene targeting have been discussed. Some of the remaining factors such as delivery of ZFN, design of donor DNA repair matrix and the rate of gene targeting may depend on cellular factors and the manner of genome editing. This section will discuss these considerations in the context of targeting disease genes with ZFN in the interest of gene therapy.

The outcomes of the targeting of various human disease associated genes are summarised in Table 1.1. (Urnov, Miller et al. 2005; Lombardo, Genovese et al. 2007; Moehle, Rock et al. 2007; Maeder, Thibodeau-Beganny et al. 2008; Perez, Wang et al. 2008; Benabdallah, Allard et al. 2010; Cradick, Keck et al. 2010; DeKelver, Choi et al. 2010; Greenwald, Cashman et al. 2010; Holt, Wang et al. 2010; Chen, Pruett-Miller et al. 2011; Herrmann, Garriga-Canut et al. 2011; Li, Haurigot et al. 2011; Lombardo, Cesana et al. 2011;

Sebastiano, Maeder et al. 2011; Soldner, Laganiere et al. 2011; Yusa, Rashid et al. 2011; Zou, Sweeney et al. 2011; Gaj, Guo et al. 2012; Overlack, Goldmann et al. 2012; Provasi, Genovese et al. 2012; Sakkhachornphop, Barbas et al. 2012; Yuan, Wang et al. 2012)

From the observations made in these studies it can be seen that, regardless of strategy, a greater rate of gene targeting is seen in cell lines compared to primary cells. This could be due to altered DNA repair in transformed cells or differences in chromatic structure for example, less efficient gene targeting in iPS cells is probably due to inactive chromatin blockage of the ZFN target site (van Rensburg, Beyer et al. 2013). Various methods of gene delivery have been utilised to introduce ZFN expression vectors and/or donor DNA to target cells. In the majority of studies, integrating deficient vectors were used with the exception of the Adeno-Associated viral vector (AAV) which has a low propensity for integration into the host genome (Nathwani, Davidoff et al. 2005). The different methods of gene targeting quantification prohibit comparison of ZFN and donor delivery however it is of note that nucleofection or other methods of plasmid transfection strategies are the most popular, probably chosen to avoid any background integration of viral sequence, and that direct treatment with ZFN protein is promising. The most frequently used donor DNA construct was circular plasmid however, a high rate of HR was achieved with a 2.9Kb linear plasmid (Greenwald, Cashman et al. 2010) but as these results were extrapolated from qPCR values, the method of quantification is not comparable with the more common Southern and RFLP methods. Interestingly, when quantification was done by flow cytometry and PCR or flow cytometry and the Cel-1 surveyor assay the methods consistently gave similar values demonstrating the robustness of these procedures. The frequency of NHEJ concurring with HR is extremely variable and not routinely assessed. Also, due to the infrequent analysis of off target DSB, no insight into the delivery strategies or ZFN design was made.

.

				Delivery		Rate of HR	Rate of NHEJ
	Author et al(year)	Target gene	Cell or animal model	ZFN	Donor	(method)	(method)
Gene Correction	Urnov (2005)	*IL2RG*	k-562	T	T	13.4% (RFLP)	n.q.
	Urnov (2005)	*IL2RG*	Primary CD4+ T cells	T	T	5.3% (RFLP)	n.q.
	Lombardo (2007)#	*IL2RG*	k-562	IDLV	IDLV	16 ± 9% (RFLP)	10% (RFLP)
	Lombardo (2007)#	*IL2RG*	Lymphoblastoid	IDLV	IDLV	5 ± 1.5% (RFLP)	5% (RFLP)
	Maeder (2008)	*IL2RG*	k-562	T	T	4.1% (RFLP)	1.8% (S)
	Maeder (2008)	*VEGF-A*	k-562	T	T	7.7% (RFLP)	9% (S)
	Greenwald (2010)	*rhodopsin*	Her cells	L	L	17% (qPCR)	n.q
	Herrmann (2011)	*p53*	Human BT549	T	T	0.1% (S)	n.q
	Soldner (2011)	α-Synuclein	Patient iPS	T	T	0.004% (S)	2% (SB)
	Overlack (2012)	m*Ush1c*	Murine Ush1c-mut	L	L	PCR+ Protein+ (n.q.)	n.q
	Chen (2011)	*RPS6KA3*	k-562	N	N :ssODN	22-32% (RFLP)	n.q
Gene insertion	Moehle (2007)#	*IL2RG*	K-562	T	T: 12bp	15% (RFLP)	n.q.
	Moehle (2007)#	*IL2RG*	K-562	T	T 1.5Kb	6% (FACS+PCR)	n.q.
	Moehle (2007)#	*IL2RG*	K-562	T	T 7.6Kb	5-8% (SB)	n.q.
	Lombardo (2007)#	*IL2RG*	K-562	IDLV	IDLV	3.4 ± 1% (FACS)	n.q.
	Lombardo (2007)#	*IL2RG*	Lymphoblastoid	IDLV	IDLV	1.8 ± 0.7% (FACS)	n.q.
	Zou (2011)	*PIG-A*	human ESC	N	N	83%* (FACS)	n.q.
	Zou (2011)	*PIG-A*	human iPS	N	N	45 ± 20%* (FACS)	n.q.
	Yusa (2011)	*A1AT*	human iPS	E	E	58%* (PCR)	n.q.
	Sebastiano (2011)	*β-Globulin*	patient iPS	N	N	9.8%* (PCR)	0% (S)
	Li (2011)#	F9	Mouse (*hF9*)	AAV I.P.	AAV I.P.	2.2% (PCR)	13.8 (Cel-1)
	***AAVS1* (Safe Harbour)**						
	DeKelver (2010)#	(GFP)	K562	N	N	10% (FACS)	n.q.
	DeKelver (2010)#	(GFP)	HEK293/Hep3B	N	N	3% (FACS)	n.q.
	DeKelver (2010)#	(GFP)	hTERT Fibroblasts	IDLV	IDLV	3% (FACS)	n.q.
	Lombardo (2011)#	(GFP)	Primary h T cells	Ad5/35	IDLV	5% (FACS)	n.q.
	Lombardo (2011)#	(GFP)	NSC	IDLV	IDLV	6% (FACS)	n.q.
	Lombardo (2011)#	(GFP)	iPS	IDLV	IDLV	1% (FACS_	n.q.
	Zou (2011)	(*CYBB*)	human iPS	N:mRNA	N	75%* (PCR)	50%~ (S)
	***CCR5* (Safe Harbour)**						
	Lombardo (2007)#	(GFP)	CD34+ Progenitor cells	IDLV	IDLV	0.06 ± 0.02% (FACS)	n.q.
	Lombardo (2007)#	(GFP)	hES	IDLV	IDLV	3.5 ± 1.1% (FACS)	12.7 ± 6.4%~ (Cel1)
	Benabdallah (2010)#	(mEpo)	hMSC	Ad5/35	IDLV	PCR+ (n.q.)	n.q.
	Lombardo (2011)#	(GFP)	Primary h T cells	Ad5/35	IDLV	5% (FACS)	n.q.
Gene Disruption	Perez (2008)#	*CCR5*	Primary CD4+ T cells	Ad5/35	n.a.	n.a.	2.4% (Cel-1)
	Holt (2010)#	*CCR5*	CD34+ Progenitor cells	N	n.a.	n.a.	17 ± 10% (Cel-1)
	Cradick (2010)	HBV genome	Huh7 Hepatoma cells	T	n.a.	n.a.	36% (RFLP)
	Sakkhachornphop (2012)	HIV 2-LTR	293T cells	N	n.a.	n.a.	26-49% (FACS)
	Li (2011)#	F9	Mouse (*hF9*)	AAV8 I.V.	n.a.	n.a.	41% (Cel-1)
	Yuan (2012)#	CXCR4	SupT1 T cells	Ad5/35	n.a.	n.a.	20-35% (Cel-1)
	Provasi (2012)#	*TRBC*	Primary T cells	IDLV	n.a.	n.a.	6-7% (FACS, Cel-1)
	Provasi (2012)#	*TRBC*	Jurkat	IDLV	n.a.	n.a.	20% (FACS) 26% (Cel-1)
	Gaj (2012)	*CCR5*	HEK293	SP	n.a.	n.a.	27% (Cel-1)
	Gaj (2012)	*CCR5*	THP1	SP	n.a.	n.a.	14% (Cel-1)
	Gaj (2012)	*CCR5*	HDF	SP	n.a.	n.a.	23% (Cel-1)
	Gaj (2012)	*CCR5*	Primary CD4+ T cells	SP	n.a.	n.a.	8% (Cel-1)

Table 1.1 Publications assessing ZFN targeting of genes with therapeutic potential.

Transfection (T), nucleofection (N), lipofection (L) and electroporation (E) were carried out with plasmid unless otherwise noted. Abbreviations: *Sharkey* protein (SP), intraperitoneal injection (IP), intravenous injection (IV), Integrase deficient lentivirus (IDLV), Adenoassociated virus (AAV), Adenovirus (Ad), Restriction Fragment Length Polymorphism (RFLP), Sequencing (S), Southern blot (SB), not quantified (n.q.), Not applicable (n.a.). Key to symbols: (#) in collaboration with or directly from Sangamo Biosciences, (*) Selection of resistance marker in the transgene, (~) quantification carried out on a sub-population. (Method) = (Method of quantification). Rates of targeting are given as the range (-), mean or mean ± standard deviation when provided. Values for NHEJ are observed when in conjunction with HR unless excluding gene disruption. Nucleofection is with plasmid unless otherwise stated.

Although these examples of ZFN mediated gene therapy are predominated by genetic disease, ZFN strategies are also being developed to combat globally significant infectious diseases. This challenge can be approached in two ways, the disruption of the infectious agent's genome or the modification of host factors that confer resistance. In the instance of malaria, ZFN are being developed to target the *Plasmodium* genome itself (Nain, Sahi et al. 2010) or disrupt the DARC and glycophorin-α genes in erythrocytes, seen to confer resistance to *Plasmodia vivax* and *falciparum* infection respectively (Kajumbula, Byarugaba et al. 2012). HIV is a major candidate for ZFN mediated gene therapy. ZFN have been designed to target the LTR (Sakkhachornphop, Barbas et al. 2012) and *Pol* (Wayengera 2011) regions of the HIV genome however, the most promising development is by ZFN mediated disruption of the host's HIV entry molecule, CCR5, which in turn provides disease resistance (Perez, Wang et al. 2008) and is the first ZFN mediated gene therapy to start phase I clinical trials.

There is much impetus on correcting patient stem cells as their regeneration properties can repopulate autologous tissue with as little as a single cell thereby avoiding heterologous BMT. Due to the limited availability of stem cells, a current area of research is the generation of induced pluripotent stem (iPS) cells which are re-programmed differentiated cells that have the pluripotent quality of stem cells (Takahashi, Okita et al. 2007). These cells can be corrected by ZFN, as ZFN treatment does not disrupt stem cell properties (Soldner, Laganiere et al. 2011; Hoher, Wallace et al. 2012), and reintroduced into the original patient. With this in mind *CCR5* disrupted by ZFN in hES and iPS cells are capable of multi-lineage differentiation *in vitro* and *in vivo* thus suggesting that pluripotency is maintained regardless of *CCR5* expression (Holt, Wang et al. 2010; Yao, Nashun et al. 2012). As a result, the *CCR5* locus is considered a suitable target site for transgene insertion. Another target locus deemed a "safe harbour" is the AAV integration site 1, *AAVS1*, within the *PPP1R12C* gene. Insertions at this site have no known pathophysiological side effects thus avoiding insertional mutagenesis and the *PPP1R12C* promoter can be used for expression of a transgene. Furthermore, integration at this site is stable (DeKelver, Choi et al. 2010)

and does not affect the pluripotency of ES or iPS cells (Hockemeyer, Soldner et al. 2009).

The *CCR5* and *AAVS1* loci were compared by inserting GFP at the respective locus and subsequent expression was analysed. (Lombardo, Cesana et al. 2011). Separate integration of 3 GFP constructs with SFFV, PGK or EF1α promoters showed equal target recombination efficiency at the 2 loci, however, the expression of GFP at the *AAVS1* locus was consistently greater than at the *CCR5* locus. Importantly, introducing a promoter at the *CCR5* locus up-regulated *CCR5* and the proximal genes *CCR3* and *CCR1*. Such an effect was not observed in *AAVS1* likely due, in part, to the chromatin insulator in the *PPP1R12C* promoter. Epigenetic markers RNA polymerase II and histone H3 revealed *AAVS1* to be a more open and transcribed region than *CCR5* (van Rensburg, Beyer et al. 2013) yet transgene insertion at both loci enhanced active transcription.

Due to the low rate of HR induced by ZFN it is beneficial to have a selection advantage within the treatment strategy. This can be achieved by including a resistance marker in the donor DNA and therefore selecting recombinants or the selective advantage may be conferred by the genomic modification itself, as seen when CXCR4- or CCR5- cells proliferate in favour of CXCR4+ or CCR5+ cells respectively during HIV-1 infection (Perez, Wang et al. 2008; Yuan, Wang et al. 2012).

Gene deletion is another genome editing outcome made possible by sequence specific endonucleases. Up to 15,000kb can be specifically deleted from a chromosome (Lee, Kim et al. 2010). Such a strategy was employed to mimic the *F8* inversion associated with Haemophilia A in HEK293 cells with the hypothesis that such a manipulation could be carried out to correct the disease's genetic abnormality (Lee, Kweon et al. 2012). It is also possible that 2 pairs of ZFN can be administered with 2 different FokI obligate heterodimers that do not cross react (Sollu, Pars et al. 2010) thus minimising the possibility of off target DSB when 4 ZFN monomers are administered.

In addition to gene therapy strategies mentioned in Table 1.1, ZFN can also be used to add fluorescent markers to endogenous genes of interest in order to visualise intracellular protein distribution including endocytosis (Doyon, Zeitler et al. 2011) and iPS cell reprogramming (Hockemeyer, Soldner et al. 2009) and differentiation (Wang,

Zhang et al. 2012). ZFN mediated gene targeting in the well-defined insertion sites, *AAVS1* and *CCR5* can also be used to harbour the reprogramming factors necessary to generate iPS cells (Ramalingam, London et al. 2013).

1.5.8 ZFN targeting of *IL2RG*

SCID-X1 is an ideal candidate for gene targeted correction. Firstly, the low rate of targeted correction would be superseded by the selective advantage of corrected stem cells. Secondly, due to the lack of T cells, introduction of the γc protein will not provoke an autoimmune response and thirdly, as SCID-X1 is an X linked recessive genetic disorder, correction of only one allele is necessary to restore protein expression.

A ZFN pair was designed by Urnov et al. (Urnov, Miller et al. 2005) at Sangamo Biosciences to recognise exon 5 of *IL2RG* (IL2RG ZFN) at the mutation hotspot G691A. The pair consists of 2 chimeric proteins that each have 4 zinc finger (ZF) domains and a FokI domain. Considering each zinc finger recognises 3bp of DNA, the ZFN pair will recognise 24bp of DNA which is a suitable length to target a single site in the genome. The gene targeting capability of these ZFNs was examined in the same study in K-562 cells (an immortalised chronic myelogenous leukemia cell line) (Lozzio and Lozzio 1975) by using a donor DNA to introduce the novel restriction enzyme recognition site BsrBI with a single base pair exchange at the WT *IL2RG* locus. RFLP analysis of the novel restriction enzyme site was carried out on PCR product 96 hours after transfection with both ZFN and donor plasmids. Up to 13.4% of target loci had undergone recombination and in cells arrested at the G2/M phase of the cell cycle HR had increased to 18.3%. No background integration was observed in cells transfected with donor alone and the recombination events were stable over time as shown by a Southern blot 1 month later. Clonal cell lines of the K-562 cells transfected with both ZFN and donor plasmid were analysed for biallelic modification and it was observed that 13.2% of arrested cells had the BsrBI site in one allele and 6.6% in both alleles. In asynchronous cells, 12.2% were mono-allelic and 2.4% were bi-allelic. The *IL2RG* ZFN were also used to introduce a 1bp frameshift from a suitably designed donor DNA which resulted in decreased mRNA expression, this frameshift was subsequently corrected by using the BsrBI containing donor which restored mRNA expression. The

ZFN and BsrBI donor also achieved HR in primary CD4+ T cells at a frequency of 5.3% where the transfection efficiency was 30%.

Further non-selection based assays with this ZFN pair examined the rate of targeted integration at the *IL2RG* locus (Moehle, Rock et al. 2007). A donor DNA was constructed to contain two 750bp homology arms flanking a 12bp sequence encoding a 4 amino acid tag and a novel Stu1 restriction enzyme site. K-562 cells transfected with IL2RG ZFN and this donor DNA yielded Stu1 sites in 15% of *IL2RG* sites and a precise copy of the 12bp tag sequence. Next, a 1.5kb GFP expression cassette provided by a donor DNA was integrated into 6% of K-562 cells when co-transfected with the ZFN pair, 8.5 fold above random integration. Random integration was not increased by this ZFN pair as compared with etoposide, a reagent capable of random DSB, providing evidence that this pair has low toxicity. Finally a donor consisting of a 7.6Kb gene destined for insertion was flanked by 750bp homology arms was recombined with 5-8% of the target loci. Subsequent analyses showed that this was homologous integration and did not involve NHEJ events.

Lombardo *et al*., then assessed the efficiency of IL2RG ZFN mediated repair by delivering the constructs using IDLV (Lombardo, Genovese et al. 2007). 2 constructs were made to combine the donor region containing the novel BsrBI site and either ZFN monomer with an SFFV promoter in the opposite orientation to the *IL2RG* homology region, hence on delivery of both ZFN monomers, 2 donor homology regions were delivered in unison. K-562 cells transduced with increasing ug/mL of these constructs achieved up to 39% gene targeting events at the *IL2RG* locus, 74% of which were BsrBI sensitive indicating HR and the remaining were a result of NHEJ. A lymphoblastoid cell line (LCL) was created by transforming male B-lymphocytes with Epstein Barr Virus and on transduction with the IDLV constructs yielded 13% gene targeting, 61% of which was repaired by HR.

Subsequently, a donor DNA with a GFP gene flanked by *IL2RG* homology arms was constructed which would disrupt exon 5 of *IL2RG* on recombination resulting in γc knockdown. LCLs were transfected with ZFN and GFP expression vectors and interestingly, insertion of the GFP gene was achieved with lower doses of ZFN. The GFP+ and - cells were sorted by flow cytometry and γc expression knockdown was

observed from 73% in GFP- cells to 4% in GFP+ cells. A further donor DNA consisting of promoter-less partial *IL2RG* cDNA flanked by *IL2RG* homology arms was delivered by IDLV and integrated in 6% of LCLs (note the single *IL2RG* allele in male cells) when the Donor:ZFN ratio was 3:1 and 3:2 and the donor was 0.75ug/mL. RT-PCR showed evidence of the integrated *IL2RG* cDNA by mRNA expression however protein expression wasn't scrutinised.

As gene targeting was performed in K-562 cells using the *IL2RG* ZFN with WT FokI by both Urnov *et al*.,(Urnov, Miller et al. 2005) and Lombardo *et al*. (Lombardo, Genovese et al. 2007) the greater rate of HR achieved using IDLV may lead to the conclusion that such a delivery is more efficient. However, neither the promoter used to achieve ZFN expression nor the ratio of donor to ZFN was disclosed by Urnov *et al*.(Urnov, Miller et al. 2005). Furthermore the combined IDLV design of donor and ZFN restricts the ratio to 2:1 thereby limiting this potential area for optimisation and lastly, the transduction and transfection efficiencies are not provided in either study.

The development of obligate FokI heterodimers led to a succeeding *IL2RG* ZFN pair which, when using the Sangamo-designed BsrBI donor plasmid, achieved HR at frequencies up to 12.5% compared to 10.6% observed for the WT FokI. Analysis of DSB events in K-562 cells transfected with IL2RG ZFN alone revealed an 11 fold drop in γ-H2AX when using the heterodimer pair as compared to the WT FokI with no loss of *IL2RG* targeting (Miller, Holmes et al. 2007). This genome wide analysis of DNA damage revealed that these ZFN were obligate heterodimers and therefore capable of targeted gene modification. It is of interest that the same protocol and cell line was used by Maeder *et al.,* (Maeder, Thibodeau-Beganny et al. 2008) however in their hands the frequency of HR was 4.1% with concurrent NHEJ at 9%. Pruett-Miller *et al*. determined the maximal targeting efficiency of episomal DNA in HEK 293 cells by each ZFN to be 10-20ng per 10^6 cells (Pruett-Miller, Connelly et al. 2008).

The IL2RG ZFN pair bind 5bp apart, leave 5' 5bp overhangs at the site of the DSB. Such overhangs are capable of capturing short DNA fragments by NHEJ (Orlando, Santiago et al. 2010). DNA capture analysis by CLIS was carried out using 2 IL2RG ZFN pairs which differed in DNA binding domain (Gabriel, Lombardo et al. 2011). 14 and 15 genomic loci were identified as targets for these ZFN pairs, 3 of which, *FAM133B*,

SLC31A1 and *SEC16A*, were common for both pairs and the *KIAA0528* locus was also predicted *in silico* which has 88% homology with the *IL2RG* target. One IL2RG ZFN pair was noted to have more on target and less off target activity however whether this was the ZFN pair used in previous Sangamo studies was undisclosed.

Assessment of the potential application of the IL2RG ZFN is promising. The rates of gene targeting would be suitable to restore immune cells in SCID-X1 patients and the low toxicity associated with this particular ZFN pair may pass for clinical use like CCR5. However in order to make this progression these ZFN must be analysed in a suitable model of SCID-X1 containing human *IL2RG* that on correction will restore the model's immune system and reveal any toxicity that may have been overlooked in *in vitro* studies.

1.5.9 TALENs

ZFN utilise the dimerization of the endonuclease domain of FokI to unite 2 ZF regions and complete the DNA binding region. A more recent advance in this type of gene targeting strategy has been found in Transcription Activator-Like Effector Nucleases or TALENs which unite DNA recognition molecules from pathogenic plant bacteria and the FokI nuclease dimer. Unlike ZF, the TALE DNA recognition molecule is defined by its two adjacent amino acids also known as the repeat variable di-residue (RVD) each recognising 1 base pair (Boch, Scholze et al. 2009) which appear to be context independent (Moscou and Bogdanove 2009), granting simple assembly of the complete recognition domain. Another difference is where ZFN have been optimised to bind DNA half sites with a 5-6bp spacer, TALENs require a 12-15bp spacer. The structure of meganucleases, ZFN and TALENs are compared in figure 1.9.

A TALEN pair has been designed to target the same region of *IL2RG* as the aforementioned ZFN (Mussolino, Morbitzer et al. 2011). Each TALEN monomer recognises an 18bp sequence separated by a 15bp spacer. The TALEN pair actively cleave exogenous DNA at the same rate as the IL2RG ZFN obligate heterodimer however ZFN gene targeting in HEK293 cells is twice as efficient as the TALENs as measured by NHEJ. However, these TALENs appear to be less toxic than the ZFN as

seen in the statistically significant drop in cell survival from nearly 100% in TALENs to 50% in ZFN treated cells.

Due to the low toxicity shown by TALENs they may supersede ZFN for gene therapy, but yet again, suitable animal models are a necessary step to clinical trials.

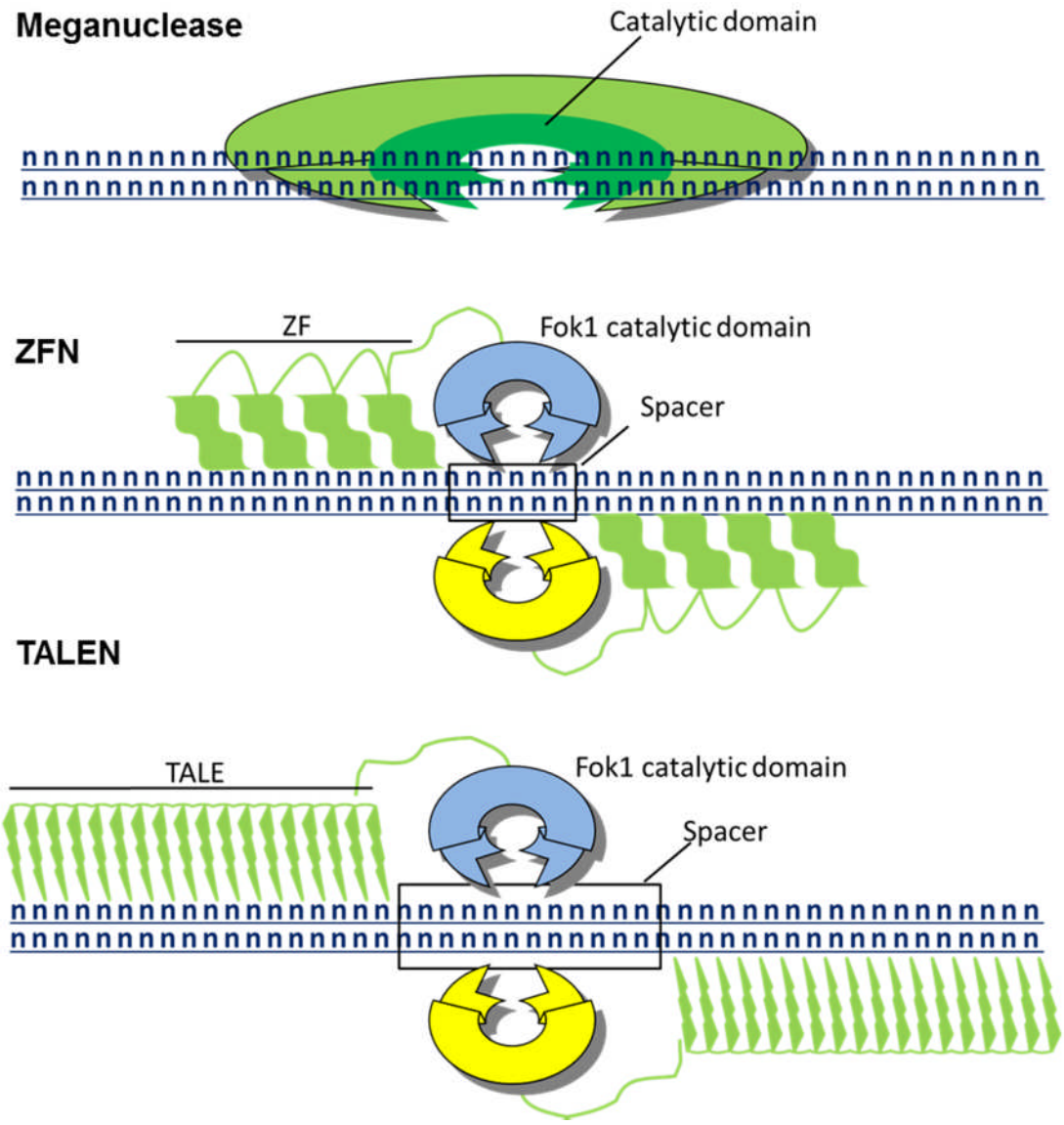


Figure 1.8 Comparison of a meganuclease, ZFN and TALEN.

A schematic representation, adapted from Curtin et al. (Curtin, Voytas et al. 2012) comparing the binding of engineered genome modifying enzymes to DNA. DNA is represented by individual nucleotides (n). Top, a monomeric meganuclease, for example I-SceI bound to DNA with the catalytic site in dark green. The DNA binding LAGLIDADG motif is not shown. Middle, a zinc finger nuclease (ZFN) binding DNA via the zinc fingers (ZF) in green. Bottom, a transcription activator like effector nuclease (TALEN) binding DNA via the transcription activator like effectors (TALE) in green. In both cases, the ZF and TALE are fused to the catalytic domain of the FokI heterodimer which cuts the spacer region between the DNA binding domains.

1.6 Models of SCID-X1 for gene targeting

Animal models of SCID-X1 assessed for the efficacy of gene transfer from viral vectors were murine *il2rg* knockouts and although these animals do have an immunodeficient phenotype suitable for such gene transfer analysis, these knockouts are unsuitable for *IL2RG* targeted correction strategies due to the lack of human *IL2RG* target. In order to assess the efficacy of targeted gene correction of SCID-X1 it is important to establish mammalian *in vivo* models with a genome suitable for the analysis of targeting nucleases intended for human genes.

The progress made in gene targeting has been applied to animal model engineering. A rat γc knockout was achieved by targeting exon 2 of the rat *il2rg* with ZFN (Mashimo, Takizawa et al. 2010). The homozygous founder animals showed an immunodeficient phenotype similar to that of the $\gamma c^{-/-}$ mouse, with depleted CD8+ T cells, NK cells, B cells and reduced IgG and diminished IgA and with a hypoplastic thymus. Although this model of SCID-X1 has a suitable phenotype, it does not contain the target for human *IL2RG* and cannot be used to ascertain the effects of ZFN designed for human SCID-X1.

Murine models are a very useful resource for human disease as they are mammalian, their genome has been sequenced and they have already been extensively studied in regards SCID-X1. Furthermore there is evidence of successful ZFN mediated gene targeting in *in vitro* and *in vivo* murine studies.

By micro-injection of murine eggs with ZFN mRNA the surviving foetuses showed NHEJ in 24-75% (Carbery, Ji et al. 2010) and HR in up to 25% on provision of plasmid donor DNA (Cui, Ji et al. 2011). In the interest of therapeutic strategies, gene targeting of pluripotent cells from mice is relevant to the treatment of human disease. mES cells with a GFP transgene treated with GFP ZFN delivered by lipofection achieved knockdown in 8% of cells (Osiak, Radecke et al. 2011) and HR in 0.21% (Connelly, Barker et al. 2010). These ZFN treated mES cells were capable of forming teratomas (a tumour with various tissue types) *in vivo* thus providing evidence that ZFN treatment does not abrogate the pluripotent quality of ES cells. However, chromosomal abnormalities were observed in a clonal cell line from the ZFN treated mES cells which may have been due to chromosome activation by DSB or long term cell culture. Although it has been observed that high doses of ZFN can increase cytotoxicity in

murine stem cells, this does not appear to effect the expression of stem cell markers or differentiation *in vitro* (Hoher, Wallace et al. 2012).

Gene targeting in primary cells has also been observed using nucleofection. Targeted NHEJ of the murine "safe harbour" locus, ROSA26, occurred at a frequency of 11% in primary skeletal myoblasts (Perez-Pinera, Ousterout et al. 2012) and targeted correction of a GFP transgene in 1.8% of fibroblasts and 0.17% of astrocytes, furthermore, ZFN treated fibroblasts were successfully transplanted into an immunocompetent recipient (Connelly, Barker et al. 2010).

A very promising development in the *in vivo* ZFN treatment of a humanised murine model of Haemophilia B was carried out by Li *et al.,* (Li, Haurigot et al. 2011). This mouse model has a mutated human gene for factor IX (*hF9*) knocked into the ROSA26 locus and integration of a cDNA containing exons 2-8 at the ZFN binding site in intron 1 will restore Factor IX expression. Delivery of ZFN and donor DNA by intraperotineal injection of hepatotophic AAV vector (AAV8) resulted in targeted HR ranging from 1-3% in liver tissue analysed after 10 weeks. NHEJ also occurred at a frequency ranging from 11-17% and no integration of donor alone was detected by PCR. The corrected loci resulted in circulating hFactor.IX at 3-7% of the normal levels and reduced clotting time to almost normal levels. Analysis of 20 off target sites predicted by SELEX showed one site that had been targeted at a tenth of the rate of *hF9* site.

It can be seen that ZFN gene editing is applicable in mice. The gene targeting frequencies in murine primary cells are similar to the rates seen in human (Table 1.1) and a range of DNA delivery can be utilised and optimised. The benefit of using humanised mouse models is the analysis of designed for humans in an environment that contains physiological epigenetic features and a mammalian DNA repair response.

1.7 Summary and Project Aims

Due to the efficacy of *IL2RG* modified HSCT for the treatment of SCID-X1 and the potential to correct this gene *in situ*, the purpose of this project is to continue the assessment of the suitability of IL2RG ZFN induced HR for the treatment of SCID-X1. Using a novel cell line model of SCID-X1 it is hypothesised that ZFN plus a correction donor plasmid will restore γc signalling *in vitro.* This project also investigates a new humanised model of SCID-X1 to confirm a loss of γc function that can subsequently be

restored by gene targeted strategies and it is postulated that optimised ZFN mediated correction of HSC could restore a complete and functional immune system *in vivo*.

Materials and methods

2.1 Reagents and abbreviations

Reagent	Supplier
Dimethyl sulfoxide (DMSO)	Sigma
Restriction endonucleases and buffer	Promega or New England Biolabs (NEB)
Agarose	Life Technologies
1kb Plus DNA Ladder	Life Technologies
SeeBlue Plus2	Life Technologies
Agar	MERCK
Ampicillin	Stratagene
Dulbecco's modified eagle medium (DMEM)	Life Technologies
Roswell Park Memorial Institute (RPMI) medium	Life Technologies
Fetal Calf serum (FCS)	Life Technologies
Antibiotic-Antimitotic (Anti-Anti)	Life Technologies
OPTI-MEM	Life Technologies
Trypan Blue	Sigma
Phosphate Buffered Saline (PBS)	Sigma
SOC medium	Life Technologies
Stem Span	StemCellTechnologies
ESGRO	Millipore
Accutase	Millipore
0.1% Gelatine	Millipore
Alkaline phosphatase detection kit	Millipore
10x Red Blood Cell Lysis buffer	Biolegend
High Fidelity (HF) Taq	New England Biolabs
Go Taq Green 2x mastermix	Promega
Platinum Quantitative PCR supermix–UDG with ROX	Life Technologies
Cells Direct One-step qRT-PCR kit	Life Technologies
T4 ligase	Promega
DNeasy	Qiagen
Sodium Carbonate (Na_2CO_3)	Sigma – Aldrich
Sodium Hydrogen Carbonate ($NaHCO_3$)	AnalaR
Bovine Serum Albumin (BSA)	Sigma
Sucrose	AnalaR
Tween 20	Sigma – Aldrich
Sulphuric Acid	Aldrich
Sodium Dodecyl Sulfate (SDS)	Sigma – Aldrich
Sodium hydroxide (NaOH)	Sigma – Aldrich
Sodium Chloride (NaCl)	Fluka
Tri-Sodium Citrate	AnalaR
Tris base	Fluka
Ethylenediaminetetraacetic acid (EDTA)	AlfaAesar
Sodium phosphate dibasic (Na_2HPO_4)	Sigma

Sodium dihydrogen phosphate (NaH_2PO_4)	Sigma
Potassium Chloride	Sigma
KaryoMAX (Colcemid)	Life Technologies
Methanol	Sigma
Acetic Acid, 100% (analytical grade)	Sigma
Tetramethylbenzidine (TMB) Solution	BD biosciences
Streptavidin Horse Radish Peroxidase (Strep – HRP)	BD Pharmingen
Streptavidin-Cy5 conjugate	Life Technologies
AccuCheck Counting Beads	Life Technologies
Kits	
TnT quick coupled transcription/translation system	Promega
QuikChange Site Directed mutagenesis kit	Stratagene
Endofree maxi prep kit	Qiagen
Gel extraction kit	Qiagen
Mini/maxi prep kit	Qiagen
PCR purification kit	Qiagen
RNA extraction kit	Qiagen
DNA labelling system	GE Healthcare
Murine lineage negative Magnetic activated cell selection	Miltenyi biotech
Amaxa Cell line nucleofector kit V	Lonza
Amaxa mES cell kit	Lonza
Amaxa human CD34+ cell nucleofector kit	Lonza
Amaxa mouse macrophage nucleofector kit	Lonza
Cytokines	
Stem Cell Factor (SCF)	Peprotech
Murine Flt3	Peprotech
Human Thrombopoietin (hTPO)	Peprotech
Human Interleukin 2 (hIL-2)	Peprotech
Human Interleukin 7 (hIL-7)	Peprotech
Concanavalin A (ConA)	MP Biomedicals

Table 2.1 Reagents and abbreviations

2.2 Antibodies

Antibody	**Clone**	**Supplier**
Western		
Anti-HA (rabbit)	Polyclonal Rabbit anti HA	Novus Biologicals
Anti-HA-HRP (Goat)	A190-107P	Bethyl Laboratories
Anti-GFP (Mouse)	MAB3580	Chemicon International
Anti-Rabbit - 680LT		LI-COR
Anti-mouse - 680LT		LI-COR
Flow Cytometry		
Anti CD3 – PE	Hamster anti mouse 145-2C11	BD Pharmingen
Anti CD4 – PE	Rat anti Mouse GK1.5	eBiosciences
Anti CD4 – FITC	Rat anti mouse GK1.5	BD Pharmingen
Anti CD8 – APC	Rat anti mouse 53 – 6.7	BD Pharmingen

Anti B220 – APC	Rat anti mouse RA3 – 6B2	BD Pharmingen
Anti IgM – PE	Rat anti mouse R6-60.2	BD Pharmingen
Anti IgM – FITC	Rat anti mouse R6-60.2	BD Pharmingen
Anti NK1.1 – APC	Mouse anti mouse PK136	BD Pharmingen
Anti γc – PE	AG184	BD Pharmingen
Anti γc – APC	TUG-H4	BioLegend
Anti pSTAT-5 – PE	Mouse anti pSTAT5 pY694	BD Pharmingen
ELISA		
Capture antibody		
IgA	C10-3	BD Pharmingen
IgE	R35-72	BD Pharmingen
IgM	II/41	BD Pharmingen
IgG	Polyclonal rabbit anti mouse IgG	AdD Serotec
IgG1	A85-3	BD Pharmingen
IgG2a	R11-89	BD Pharmingen
IgG2b	R9-91	BD Pharmingen
IgG3	R9-91	BD Pharmingen
Detection Antibody		
IgA - Biotin	C10-1	BD Pharmingen
IgE - Biotin	R35-118	BD Pharmingen
IgM - Biotin	II/41	eBioscience
IgG - Biotin	Polyclonal rabbit anti mouse IgG	Dako
IgG1 - Biotin	X56	BD Pharmingen
IgG2a - HRP	R19-15	BD Pharmingen
IgG2b - HRP	R12-3	BD Pharmingen
IgG3 - Biotin	R40-82	BD Pharmingen

Table 2.2 Antibodies

2.3 Cell lines

Escherichia coli (E. coli)	XL10-Gold Chemically competent	Stratagene
Escherichia coli (E. coli)	DH5α Chemically Competent Cells	Invitrogen
ED-7R	Human T cell line	Gift: I. Alexander
ED-γc-7R	Human T cell line	Gift: I. Alexander
293T	Human embryonic kidney cell line (HEK293T)	ATCC: CRL-11268
hγcmut mES	Murine embryonic Stem cells (mESC)	InGenious

Table 2.3 Cell lines

2.4 Animal maintenance

All mice (wild type and genetically modified) came from the C57BL/6 genetic background and were maintained adhering strictly to the Home Office licencing laws.

All experimental practises were performed under both a personal licence (Celeste Pallant unless otherwise indicated) and a project licence.

2.4.1 Engineering the humanised mouse model

The outline of the generation of the humanised mouse model was carried out by InGenious as follows. A 5518bp region of the human γc gene was subcloned from a human BAC clone into a shuttle vector (pSP72, Promega). *Mlu*I sites were engineered at the 5' and 3' ends by Red/ET recombineering. The G691A mutation was incorporated into exon 5 of the human γc gene by overlap extension PCR before transfer into the mouse sequence. A 14.4kb region of the mouse γc gene was subcloned from a C57BL/6 BAC (RP23: 263O9) clone; this contained the murine homology arms and the target region to be replaced by the human gene, flanked by *Mlu*I. The 5' murine homology arm was 7.1kb and the 3' homology arm was 2.3kb. The murine sequence replaced was 5106bps long and included the entire coding sequence of the γc mouse gene. The resulting vector also included a neomycin resistance gene. The target vector was delivered to murine embryonic cells by electroporation. Successfully recombined cells were selected for by neomycin treatment. The neomycin cassette was flanked with LoxP-FRT allowing the neomycin resistance gene to be excised from the genome via Cre-Lox recombination. Successful generations resulting from the injection of these embryonic stem cells into pseudo-pregnant mice were sequenced to confirm the presence of the mutant human γc allele and deletion of the neomycin resistance cassette.

2.4.2 Isolation of primary murine cells

Primary cells were harvested from the following tissues: Spleen, lymph nodes, thymus, bone marrow (muscle were removed and thebone marrow was flushed from the femurs and tibias using a 23G needle) and peripheral blood. Peripheral blood was harvested either at sacrifice by cardiac puncture with a 23G needle or was harvested from the living mouse by warming the animal with an infrared lamp to locate the tail vein for blood extraction with a 23G needle. All tissues except peripheral blood were macerated through 7um pore gauze to achieve a single cell suspension. The single cell suspension was then centrifuged at 400g for 5 minutes, re-suspended in 10mL of red blood cell lysis buffer and incubated at room temperature for 10 minutes. The cell

suspension was then centrifuged at 400g for 5 minutes and re-suspended in the medium suitable for the subsequent step.

2.4.3 Isolation and culture of Murine Lin$^-$ cells

Bone marrow isolate passed through 7um pore gauze was negatively selected for Lin$^-$ cells as per the MACS Lin$^-$ selection kit instructions. 10^8 cells were prepared per LS sized MACS column and values follow per 10^8 cells. Cells were centrifuged at 300g for 10 minutes and re-suspended in 400µl of MACS buffer (0.5% BSA, 2mM EDTA), 100µl Biotin-antibody cytokine cocktail was added and the cells incubated at 4ºC for 20 minutes. Cells were washed in 5ml of MACS buffer and centrifuged at 300g for 10 minutes. The cells were re-suspended in 800µl of MACS buffer and 200µl of Anti-Biotin microbeads added. Cells were incubated at 4ºC for 30 minutes. The cells were washed in 5ml of MACS buffer and centrifuged at 300g for 10 minutes. Cells were re-suspended in 500µl of MACS buffer and the cells were applied to the LS columns, pre-prepared by rinsing with 3ml of MACS buffer. Cell effluent was allowed to drip through by gravity. For the reconstitution experiment the effluent was applied to a second LS column pre-prepared by rinsing with 3ml of MACS buffer. The effluent from the MACS columns was centrifuged at 300g for 10 minutes and re-suspended at 10^6 cells/mL in complete StemSpan 1% penicillin/streptomycin plus cytokines at the following concentrations: SCF (100ng/mL), mFlt3 (100µg/mL) and hTPO (25ng/µL). A 24 well plate was then seeded at 10^5 cells per well. The cells were left to recover at 37°C for 4 hours.

2.4.4 Transduction of Murine Lin$^-$ cells

To prepare the murine Lin$^-$ cells for tail vein injection into mice, the recovered Lin$^-$ cells were transduced mL in complete StemSpan, 1% penicillin/streptomycin plus cytokines at the following concentrations: SCF (100ng/mL), mFlt3 (100µg/mL) and hTPO (25ng/µL), with integration proficient lentiviral vector at various multiplicity of infection (MOI) overnight. The cells were then centrifuged at 300g for 10 minutes and re-suspended in 200µl of RPMI.

2.4.5 Irradiation and tail vein injection of mice

Mice were irradiated with total of 10 Gray by a split dose of 6 Gray followed by 4 Gray 24 hours later. $2x10^5$ transduced murine lin$^-$ cells were administered via tail vein. Both

the irradiation and tail vein injection were performed by Michael Blundell, holder of a personal licence and relevant project licence.

2.4.6 Splenocyte proliferation assay

Total splenocytes suspended in splenocyte medium (DMEM, 10%FCS, 0.05 Beta-2-mercaptoethanol, 1x antibiotic-antimitotic) were plated at $2x10^5$ cells per well in a round bottom 96 well plate, in triplicate. The cells were incubated with and without Concanavalin A (Con A) (1µg/ml). Cells from either splenocyte medium alone or with Con A were also stimulated with either IL-2 (20ng/ml) or IL-7 (20ng/ml) or without. Following incubation at 37°C, 5% CO_2 for 48 hours, the cells were then pulsed with 1µl of methyl-3H-thymidine (0.037 MBq) for 12 hours. Uptake of methyl-3H-thymidine was determined by counts per minute detected by a beta-counter (MicroBeta TAILUX, Wallac).

2.5 Tissue Culture

2.5.1 ED7R

The ED7R cell line is a non-adherent cell line (Table 2.1) (Maeda, Shimizu et al. 1985). The cells were grown in ED7R growth medium (RPMI1640 with Glutamax, 20% FCS and 1% Antibiotic-Antimitotic) and incubated in 175cm^2 tissue flasks (NUNC) at 37°C with 5% CO_2. Following medium colour change, the cells were passaged to achieve a $1x10^5$ cells/ml concentration.

2.5.2 293T

The 293T cell line is an adherent, packaging cell line. The cells were grown in 293T growth medium (DMEM with Glutamax, 10% FCS and 1% Antibiotic-Antimitotic) and incubated in 175cm^2 tissue culture flasks (NUNC) at 37°C with 5% CO_2. When the cells reached 80-90% confluence, the monolayer was washed in PBS and detached from the flask with trypsin/EDTA then incubated at room temperature for 5 minutes. The cells were then passaged at a 1:10 dilution with growth medium.

2.5.3 murine ES cells (mES)

An adherent murine embryonic stem (mES) cell line from the hγcmut mouse were cultured in a feeder free system whereby 3ml of 0.1% gelatine is incubated in a 25 cm^2 flask for 30 minutes and the excess removed. mES cells are then thawed, washed in PBS and seeded in ESGRO. Once the cells formed embryoid bodies but before they became completely differentiated, accutase (1x) was used to manufacturers

instruction to detach the cells from the flask and from each other. The cells were then split 1:3.

2.5.4 Long term storage

The cells were stored long term by centrifuging the cells at 400g for 5 minutes, re-suspended in freezing medium (90%FCS, 10%DMSO), aliquoted into cryovials (NUNC) and slowly frozen in a freezing chamber to -80°C. The cells were then transferred to liquid nitrogen.

Cells were thawed rapidly in a 37°C water bath and the DMSO was diluted with growth medium. The cells were centrifuged at 500g for 5 minutes, re-suspended in growth medium and propagated as described.

2.6 Bacterial Manipulation

2.6.1 Gel electrophoresis

DNA fragments were visualised by loading on to a 0.8% - 1.5% agarose gel containing 0.5µg/mL ethidium bromide and separated by electrophoresis at 20 volts – 100 volts in 1x TAE buffer (40mM Tris-acetate, 5mM EDTA) as indicated. Expected fragment length of <1kb underwent electrophoresis for 30 minutes or 60 minutes for >1kb. The DNA was visualised using ultra violet (UV) light.

2.6.2 Bacterial transformation

A frozen aliquot of 30-50µl of chemically competent *E. coli* was kept on ice, 10-50ng of DNA was added and the cells were incubated on ice for 30 minutes. They then underwent heat shock at 42°C for 30 seconds and put directly on ice. SOC medium was added and the cells were incubated at 37°C for 1 hour, agitated at 250rpm. 50-500µl of the transformed *E.coli* were then streaked onto LB agar plates containing the appropriate antibiotic and incubated at 37°C overnight. Individual *E.coli* colonies were then selected and grown in LB broth at 37°C, agitated at 250rpm. *E.coli* was stored long term in 16% glycerol at -80°C.

2.6.3 Plasmid DNA preparation

E.coli expressing the plasmid of choice was grown for either small scale or large scale plasmid DNA preparation. DNA was synthesised by *E.coli* and extracted using a Mini-Prep kit (small scale) or a Maxi-prep kit (Large scale) carried out to the manufacturer's instructions.

DNA concentration was measured using a Nano-Drop spectrophotometer. The measurements were taken with a 0.2mm pathlength at a wavelength of 269nm (A_{260}).

2.6.4 Restriction enzyme digests

For a 10µl reaction, miniprep DNA (1µg), restriction enzyme (10 units), corresponding restriction enzyme buffer (1µl) and water to the volume of 10µl were incubated at 37°C for 1 hour. The products were then loaded alongside a 1Kb+ DNA ladder on a 1% agarose gel containing ethidium bromide. After electrophoresis the gel was visualised using UV light.

2.7 DNA cloning and analysis

2.7.1 Karyotype analysis

Cells were prepared for karyotype analysis as follows. 10^6 ED7R cells in 10ml of ED7R growth medium (RPMI1640 with Glutamax, 20% FCS and 1% Antibiotic-Antimitotic) were treated with 50µl Colcemid (KaryoMAX 10µg/ml) over night. Cells were centrifuged at 300g for 10 minutes. Cells were re-suspended in Potassium Chloride (0.075M). Cells were centrifuged at 300g for 10 minutes and as much potassium chloride was removed as possible. Freshly made (within 30 minutes) ice cold fix (30ml methanol, 10ml acetic acid) was added drop wise for 10 drops with intermittent flicking followed by 5ml. Cells were centrifuged at 300g for 10 minutes and re-suspended in 5ml of fresh fix without slow re-suspension. Karyotype analysis was carried out by Nikki Austin at Great Ormond Street Hospital.

2.7.2 PCR

DNA template (100ng) was used per reaction. Annealing temperatures and extension times varied according to the primers and fragment length as indicated.

When 2x Go-Taq green mastermix was used the reaction was prepared to manufacturer's instruction and performed with an initial denaturation at 95°C for 2min followed by 35 amplification cycles consisting of denaturation at 95°C for 30s, primer annealing for 30s, extension at 72°C for 1min/kb, and a final extension step at 72°C.

When 2x HF Taq mastermix was used the reaction was prepared to manufacturer's instruction and performed with an initial denaturation at 98°C for 2min followed by 35 amplification cycles consisting of denaturation at 98°C for 30s, primer annealing for 30s, extension at 72°C for 30 seconds/kb, and a final extension step at 72°C.

2.7.3 Genotyping the humanised mouse strain

Mice were ear notched at 3 weeks of age for identification and genotyping. DNA was released from the tissue using DNAreleasy to manufacturers instruction. The genomic DNA was then amplified with the GoTaqGreen 2x mastermix containing forward (F) primers to either human *IL2RG*:

5'TCTGGAATTTCTGGGCTCAC

or murine *il2rg*:

5'TTCATTCAACCCACCTGCGTCTC

and the reverse (R) primer to the region downstream of murine *il2rg*:

5' CCACCCCTATAGATGCTGACAAC

Polymerase chain reaction (PCR) was performed with an annealing temperature of 54°C. Products were then run on a 1% agarose gel for 30 minutes at 100 volts and visualised using UV light.

2.7.4 Site Directed mutagenesis

Complementary primers were designed using Stratagene's primer design software.

F: 5'- GTTTCGTGTTCGGAGCCACTTTAACCCACTCTGTG and

R: 5'- CACAGAGTGGGTTAAAGTGGCTCCGAACACGAAAC

Site directed mutagenesis was then carried out to manufacturer's instruction.

2.7.5 Modifying the ZFN binding site in the donor

Linker PCR was conducted in order to introduce the modified binding site MB17. HF-Taq was used to amplify 2 fragments with the following primer pairs:

F:5'-GGAGCCGCACAATGGATTGAATCTTGATCTAACACGAAACGTGTAGCGTTTCTG

R:5'-TCACCCTTCTCCCAGTTGTC and

R5'-GGAGCCGCACAATGGATTGAATCTTGATCTAACACGAAACGTGTAGCGTTTCTG

F5'- GACAGAGCCTCACTCTGTTGC

respectively generated the 2 arms of the homology domain with a complimentary overlapping section. These products were purified on Qiagen PCR purification columns

(to manufacturer's instructions). A second HF-Taq PCR using these two fragments plus the flanking primers

F: 5'-TCACCCTTCTCCCAGTTGTC and R: 5'- GACAGAGCCTCACTCTGTTGC

generated the homology domain for the modified binding site donor. This blunt PCR product was then cloned into pJet using the Fermentas clone jet system (performed to manufacturer's instruction).

2.7.6 Cloning the modified binding site (MBS) into the complete γc gene

The MBS donor vector and the Wild type γc (WT γc) expression vector were digested with restriction enzymes NcoI and Bsu361 to create sticky ends. The correct bands were excised from the gel and purified using the gel purification kit (Qiagen, to manufacturer's instruction). The fragments were then ligated at a 1:3 ratio of γc backbone:MBS donor insert with Promega T4 ligase to manufacturer's instruction. The modified γc expression vector was transfected into 293T cells, 48 hours later γc expression analysed by Flow cytometry.

2.7.7 DNA preparation

To assess the transgene copy number in the ED7RγcΔ and ED7RγcWT clonal cell lines, Chromosomal DNA was isolated by the salting out method of DNA extraction. 10^6 cells were resuspended in 3mL of nuclear lysis buffer (10mM Tris-HCl pH8.2, 0.4M NaCl, 2mM Na_2EDTA), 200µl 10% SDS and 537µl proteinase K solution (2mg/ml Proteinase K, 1% SDS, 2mM Na_2EDTA) and incubated at 37 °C overnight. 1mL of 6M NaCl was added and shaken to precipitate out the cellular debris which was then pelleted by centrifugation at 1309g for 15 minutes. The DNA in the supernatant was isolated from the debris and precipitated out with 10mL of ethanol. The chromosomal DNA was then pelleted by centrifugation at 3351g for 30 minutes. The pellet was washed with 70% ethanol, centrifuged again at 3351g for 15 minutes and dissolved in 500µl TE buffer (10mM Tris, 1mM EDTA, pH8) at 37°C overnight.

10µg of chromosomal DNA was digested in 10 units of restriction enzyme to a maximum volume of 20µl at 37°C overnight and the DNA products were separated on a 0.8% agarose gel at 30 volts overnight.

In a separate use of the Southern blot, utilised to quantify the rate of homologous recombination in ZFN and repair donor DNA treated samples, template for the

Southern blot was prepared by amplifying a 1,901bp product from the samples to be quantified for targeted HR with primer pair 3

F: 5'-TGAACCACTGTTTGGAGCAC and R: 5'-AGGTTCTTCAGGGTGGGAAT

Following gel purification (Qiagen, to manufacturer's instruction) 100ng of the PCR product was digested with BsrBI with or without DraIII, to a total of 1µl per 10µl digestion reaction, at 37°C for 1 hour. The digestion products were separated by electrophoresis on a 1% agarose gel alongside a 1Kb plus ladder.

2.7.8 Southern blot

The DNA products were blotted onto a membrane (Amersham Hybond-N membrane) overnight by capillary transfer in alkaline denaturing buffer (0.4M NaOH). The membrane was washed in 2 x SSC (3M NaCl, 0.3M tri – sodium citrate dihydrate), dried and stored at 4°C.

The membrane was pre-hybridised in 10ml of Church buffer (1% BSA, 7% SDS, 0.4M Na_2HPO_4, 0.1M NaH_2PO_4) at 65°C for 1 hour.

To quantify the rate of HR, a radiolabelled probe was designed to bind the 5' half of the digested products and was produced by amplifying a 1002bp product from a γc expression plasmid with the primer pair 1:

F: 5'-CAAGATTCAATCCATTGTGCG and R: 5'-AGGTTCTTCAGGGTGGGAAT

Two probes were generated to detect the copy number of transgene by restriction enzyme digest of the lentiviral construction plasmid.

Following gel purification (Qiagen, to manufacturer's instruction), 25ng of the PCR product was labelled with [α-P^{32}]dCTP (Megaprime DNA Labelling System) to manufacturer's instructions. Unincorporated nucleotides were removed from the probe by centrifugation through a microSpin –s-300 HR column (Amersham). Having denatured the probe at 98°C for 5 minutes and cooling on ice, it was added to the hybridisation tube containing the membrane and hybridisation was performed with rotation at 68°C overnight.

The membrane underwent sequential washes with SDS (0.5%) and decreasing concentrations of SSC (2x, 1x 0.5x, and 0.2x) at 65°C for 30 minutes/wash until no

significant background signal (as determined by a Geiger counter) could be determined. The membrane was then dried and sealed in cellophane and exposed for a minimum of 10 minutes to a maximum of 2 weeks to a phosphor screen and was processed using a Molecular Dynamics Typoon 9410 phosphorimager (GE Life Sciences, Piscataway, NJ) and analysed using the ImageJ software.

2.7.9 Quantitative PCR (qPCR)

Genomic DNA (100ng) extracted using the Qiagen DNA extraction kit was used as a template for each reaction. Each reaction consisted of primer (0.9 μM), fluorescent probe (0.2 μM) and the Platinum qPCR SuperMix-UDG with ROX mastermix and was performed in triplicate and real time PCR was carried out to manufacturers instruction and quantified using an ABI Prism 7000 (Applied Biosystems, Foster City, CA). Primers and probe designed to detect *β-actin* were designed in house and the primers and probes to detect *titin* and WPRE were a kind gift from Anne Galy, Genathon. A plasmid encoding the WPRE, *β-actin* and murine *titin* sequences, designed by Conrad Vink, was titrated and used as a standard for DNA copy number

Primers to human *β-actin*

F: 5'-TCACCCACACTGTGCCCATCTACGA and

R: 5'-CAGCGGAACCGCTCATTGCCAATGG

Probe to detect human *β-actin*

F: 5'-FAM-ATGCCCTCCCCCATGCCATCCTGCGT-TAMRA

Primers to WPRE

5'-TTCTCCTCCTTGTATAAATCCTGGTT and R: 5'-CGCCACGTTGCCTGACA

Probe to detect WPRE

5'-FAM-CTGTCTCTTTATGAGGAGTTGTGGCCCG-TAMRA

Primers to detect murine *titin*

F: 5'-AAAACGAGCAGTGACGTGAGC and R: 5'-TTCAGTCATGCTGCTAGCGC

Probe to detect murine *titin*

5’- FAM-TGCACGGAAGCGTCTCGTCTCAGTC-TAMRA

2.7.10 Detecting HR after transfection with ZFN and MBS donor

48 hours after transfection with plasmid, genomic DNA was harvested and a PCR performed using the Primer pairs:

To detect the MB17 cassette:

1: F:5’-CAAGATTCAATCCATTGTGCG and R:5’-AGGTTCTTCAGGGTGGGAAT

2: F:5’-CGCACAATGGATTGAATCTTG and R:5’-TGAACCACTGTTTGGAGCAC

To detect the MB3 cassette:

1: F:5’-GGAGCCGCTTCAATCCG and R:5’-AGGTTCTTCAGGGTGGGAAT

2: F:5’-CGCACAATGGATTGAATCTTG and R:5’-CACTTCCACAGAGCGGATTG

To amplify the region outside of donor homology:

3: F:5’-TGAACCACTGTTTGGAGCAC and R:5’-AGGTTCTTCAGGGTGGGAAT

GoTaq 2x master mix was used to manufacturer’s instruction with an annealing temperature of 60°C and an extension time of 90 seconds. The products then underwent gel electrophoresis on a 1% agarose gel at 100V for 30 minutes.

2.7.11 Detecting insertions and deletions (Indels)

Genomic DNA that has undergone targeted NHEJ can be denatured and re-annealed to give rise to areas of mismatched DNA which can subsequently be recognised and digested by T7 endonuclease or Cel-1. 96 hours after 293T cells or ED7R cells were transfected or transduced with ZFN encoding vectors, genomic DNA was extracted (Qiagen DNeasy kit to manufacturer’s instruction). Genomic PCR was carried out with High Fidelity (HF) Taq and primers which flank the ZFN target sequence on either side

F: 5’-TCAGTGAAGGGAGCAGTGTG and R: 5’-AACAACACGCTAACCCAACC

to amplify a 500bp template for the T7 endonuclease assay. To achieve a DNA template with an area of mismatch 100ng of purified PCR product was incubated with enzyme buffer 2 (NEB) at 98°C for 5minutes and allowed to cool. T7 endonuclease (NEB) was then added and incubated at 37°C for 20 minutes. Alternatively the Cel-1

surveyor assay was performed to manufacturer's instruction. The cleavage product was separated by gel electrophoresis on a 1.5% agarose gel at 90V for 30 minutes.

2.7.12 TnT lysate assay

500bp products were amplified from Wild type, mutant γc or the MBS donor with primers

F: 5'-TCAGTGAAGGGAGCAGTGTG and R: 5'-AACAACACGCTAACCCAACC

ZFN were translated in a microtube using the TnT rabbit lysate kit (to manufacturer's instruction). The product of translation was then incubated with 100ng of purified PCR product NEB buffer 4, 0.1M NaCl, 1% BSA at 37°C for 1 hour 30 minutes. The digest then underwent electrophoresis on a 1.5% agarose gel at 90V for 30 minutes.

2.8 Lentivirus production

The packaging cell line 293T cells were grown to 70-80% confluence in a 175cm^2 flask and washed with optimem before transfection. The Lentiviral vector DNA (40µg), VSV-G plasmids pMDG.2 (10µg), packaging plasmid for integrating proficient lentivirus pCMVRΔ8.74 (gag-pol) or integrating deficient lentivirus pCMVRΔ8.74 D64V (30µg) and OPTI-MEM (5mL) were filtered through a 0.22µl filter (Millipore). The transfection reagent polyethylenimine (PEI) (1µl) and OPTI-MEM (5mL) was filtered separately through a 0.22µm pore filter. The filtered vector and transfection reagent were mixed and incubated together at room temperature for 20 minutes and added to the 293T cells before incubating for 4 hours at 37°C and 5% CO_2. At this point the supernatant was replaced with fresh 293T growth medium and incubated at 37°C and 5%CO_2 for 48 hours. The viral supernatant was then harvested, the cell debris was spun down at 3351g for 10 minutes in a table top centrifuge and the supernatant filtered through a 0.22µm pore filter. The virus was then concentrated by centrifugation at 98,000g for 2 hours. The pellets were re-suspended in OPTI-MEM (200µl) and stored at -80°C. Fresh 293T growth medium replaced the viral supernatant on the same transfected 239T cells which were again incubated at 37°C 5% CO_2 and the viral supernatant was harvested again as described 24 hours later.

2.8.1 Quantification of viral vector stocks

The titres of lentiviral vector stocks were quantified by either qPCR or flow cytometry. 5x10^5 ED7R cells in ED7R growth medium were plated in round bottomed 96 well

plates (NUNC). A titration of virus (0.5, 1, 2.5, 5 or 7.5µl) was then added to the wells with a final volume of 200µl. The cells were then incubated at 37°C and 5% CO_2 for 2 days. On day 3 the cells were harvested and analysed. ED7R cells transduced with ZFN and donor DNA constructs were analysed for integrated WPRE by qPCR. ED7R cells transduced with *IL2RG* of GFP constructs were analysed for γc and GFP expression respectively by flow cytometry. Samples less than 30% positive for the viral construct were used for the calculation of viral titre as positive cells should only have 1 copy of vector. To calculate the amount of viral vector per mL, the following equation was used: (percentage+/100) x number of ED7R cells transduced x (1000/µl of vector used). In subsequent transductions a multiplicity of infection (MOI) was calculated from these values.

2.9 Nucleofection

Cells were centrifuged at 300g for 10 minutes, re-suspended in nucleofection solution and pulsed per manufacturer's instruction. Cells were then transferred to medium pre-warmed to 37°C and incubated at 37°C and 5% CO_2.

2.9.1 ED7R cells

$2x10^6$ ED7R cells were re-suspended in Lonza nucleofection solution V and DNA (2µg) then pulsed with programme X-001. The cells were immediately transferred to ED7R growth medium.

2.9.2 mES cells

$2x10^6$ mES cells were then re-suspended in Lonza nucleofection solution for murine ES cells and 2µg of DNA and pulsed with programme A-30 as directed. The cells were immediately transferred to ESGRO.

2.9.3 Lineage negative (lin^-) cells

1x106 lin^- cells were then re-suspended in Lonza nucleofection solution. 3 nucleofection kits were used (i) for macrophages and pulsed with the suggested programmes for nucleofection (ii) for human CD34+ cells and pulsed with the programme U-008 (iii) with solution V and pulsed with programme X-001. The cells were immediately transferred to complete Stem-Span.

2.10 Flow cytometry

Cells were incubated with fluorophore conjugated antibodies, diluted to manufacturer's instruction in a volume of 100µl of FACS buffer (PBS with 1% FCS),

incubated for 30 minutes at 4°C and washed twice with FACS buffer. The cells were acquired on a Calibur (BD Biosciences) flow cytometer and the cell populations analysed using FLOWJO software.

2.10.1 Episomal knockdown assay

$1.3x10^5$ 293T cells were and seeded in to a 24 well plate and incubated at 37°C for 4-17 hours. The adhered cells were then transfected with a combination of plasmids encoding ZFN EA, KV, γc-GFP fusion construct, mCherry, ΔSce1 or mock (Appendix 3). A total of 2.5µg of DNA was incubated with 50µl PEI (0.1g PEI/L, 150nM NaCl, HCL to pH5.5), incubated for 10 minutes at room temperature and transferred to the cells. 96 hours later the cells were harvested and acquired by a flow cytometer.

2.11 Western blot analysis

96 hours after 293T cells were transfected with ZFN encoding plasmids, the protein lysate was harvested with Lysis buffer (20mM Tris pH 8.0, 100mM NaCl, 0.2% NP-40, 0.2% Triton X-100, 0.2% deoxycholate) plus protease inhibitor. Protein concentration was determined with the BIORAD kit (carried out to manufacturer's instruction) and the absorbance read at 750nm. Equal amounts of protein lysate was loaded on to a separating gel (1M Tris pH8.8, 30% AA/BAA, 20%SDS, 10%APS, TEMED) with stacking gel (1M Tris pH 6.8, 30% AA/BAA, 20%SDS, 10%APS, TEMED) and underwent electrophoresis at 175V for 70 minutes. The separated protein bands were transferred on to a polyvinylidene fluoride (PVDF) membrane in transfer buffer (25mM Tris, 190 mM glycine, 20% methanol, 12g Tris, 57.6g glycine, H_2O and MeOH) between 2 sponges. Transfer was performed at 100V for 1 hour at 4°C. The membrane was removed from the transfer set up and incubated in blocking buffer (5% milk powder in PBS, 0.5% tween) under gentle agitation overnight. The membrane was incubated with primary antibody diluted to manufacturer's instruction in PBS, 0.5% Tween, 2% BSA and incubated for 1 hour at 25°C under agitation. The membrane was washed in PBS 0.5% tween 2x 5 minutes. The membrane was incubated with the appropriate secondary antibody diluted to manufacturer's instruction for 1 hour under gentle agitation. When the membrane bound protein bands were detected with fluorescence conjugated antibodies, the bands were detected by their fluorescence. Alternatively, when protein bands were detected by HRP conjugated antibodies, they were visualised

using enhanced chemiluminescence (ECL) using an UVIchemi (UVItec) chemiluminescence documentation system and UVIsoft (UVItec) software.

2.12 The pSTAT-5 assay

$1x10^5$ nucleofected or non-nucleofected, ED7Rγc or ED7R γcm cells were incubated overnight in complete RPMI medium. The following day the cells were washed once in PBS, pelleted by centrifugation in a table-top centrifuge at 400g for 5 minutes and re-suspended in 200 µl of RPMI only. The cells were split between two FACS tubes and human IL-2 was added to one of the tubes to a final concentration of 100 ng/µl. The cells were incubated at 37°C with 5% CO_2 for 10 minutes, pre-warmed (37°C) FACS lyse/fix buffer (1:5 dilution) was added to each tube. The cells were then incubated for a further 10 minutes at 37°C with 5% CO_2. The cells were pelleted by centrifuging at 400g for 5 minutes, the supernatant was decanted and the cell pellets re-suspended in 1ml of cold Perm Buffer III. The cells were then incubated on ice for 30 minutes. The cells were washed once with 1ml of FACS buffer. The cells were then stained with anti-pSTAT5-PE and incubated on ice for 30 minutes. The cells were subsequently washed twice with 1ml of FACS buffer and finally re-suspended in FACS buffer supplemented with 1% (volume for volume) paraformaldehyde solution and analysed by flow cytometry.

2.13 Immunoglobulin ELISA assay

Capture Antibody	FINAL CONC	Sample Serum dilution	Detection Antibody Dilution	HRP
IgA	2µg/ml	1/100	1/250	yes
IgE	2µg/ml	1/10	1/250	yes
IgM	2µg/ml	1/250	1/250	yes
IgG	2µg/ml	1/81000	1/5000	yes
IgG1	2µg/ml	1/1000	1/250	no
IgG2a	2µg/ml	1/1000	1/500	no
IgG2b	2µg/ml	1/1000	1/250	yes
IgG3	2µg/ml	1/50	1/250	yes

Table 2.4 Antibody and Sample dilution for ELISA

Greiner flat bottom plates were coated with 50µl coating antibody in coating buffer (0.5M carbonate-bicarbonate, pH9.6) and incubated at 4°C for 16 hours. The plates were washed 3 times with 200µl wash buffer (1x PBS, 0.05% Tween). The wells were then incubated for 2 hours at room temperature with 100µl of block buffer (1x PBS, 1% BSA, 5% sucrose, 0.05% tween). Samples and standards were diluted in sample diluent (1x PBS, 1% BSA, 5% sucrose), and 50µl/well incubated in duplicate at room temperature for 2 hours followed by 3 washes in 200µl wash buffer. The diluted detector antibody was added at 50µl per well and incubated at room temperature for 1 hour followed by 3 washes in 200µl wash buffer. 50µl of diluted Streptomycin-HRP conjugated antibody was applied to samples detected with biotinytlated antibody and incubated in the dark at room temperature for 30 minutes followed by 3 washes with 200µl of wash buffer. The TMB substrate (OptEIA BD) was applied per manufacturer's instruction and after incubation in the dark for 5-10 minutes the absorbance was read using an Optima plate reader.

3 Targeting the common gamma chain gene with *IL2RG* specific ZFN for correction with a donor repair matrix designed to be resistant to ZFN cleavage

3.1 Aims

- Confirm ZFN expression and cleavage of the desired target
- Design and construct a donor repair matrix resistant to ZFN cleavage
- Demonstrate targeted HR as a result of specific *IL2RG* cleavage by ZFN

3.2 Introduction

As gene therapy for SCID-X1 has shown, restoration of a functional immune system in patients is possible with the introduction of functional *IL2RG* via integrating retroviruses. Unfortunately, due to the insertional mutagenesis that can occur as a result of vector insertion near a proto-oncogene, the incidence of cancer associated has highlighted the necessity for safer gene therapy strategies to be developed.

A DSB can be repaired by NHEJ and, due to the error prone nature of this pathway, can result in the disruption of the gene and therefore loss of protein expression. Alternatively, the DSB may be repaired by HR which relies on homologous gene sequence to be present and can be utilised to faithfully introduce predetermined gene sequence. Both of these gene repair pathways have their uses in gene therapy In order to harness the cell's response to DNA damage for a therapeutic strategy, the DSB stimulus can be directed to an exact point in the genome by ZFN.

The ZF designed by Sangamo Biosciences and described to confer specificity for *IL2RG* were subsequently fused to an obligate Fok I heterodimer by T. Cathomen et al., which has been modified from WT and installed with opposite charges at the dimer interface. The Fok I subunit KV has amino acids K and V at positions 490 and 538 respectively and confers a positive charge (Sollu, Pars et al. 2010). The partner subunit, EA has increased negative charge at amino acids 486 and 499. The two variants EA and KV are fused to the Left and Right ZF DNA recognition domain respectively (Figure 3.1). This pair was shown to be a faithful heterodimer and capable of stimulating HR in an episomal GFP targeting assay (Szczepek, Brondani et al. 2007).

Correcting mutant *IL2RG* with ZFN remains to be seen. The strategy in this project is to deliver the *IL2RG* ZFN pair alongside a donor DNA repair matrix and induce HR to correct the common SCID-X1 mutation, G691A, *in situ*. To develop this strategy DSBs and evidence of NHEJ and HR will be assayed to verify the possibility of gene targeted correction.

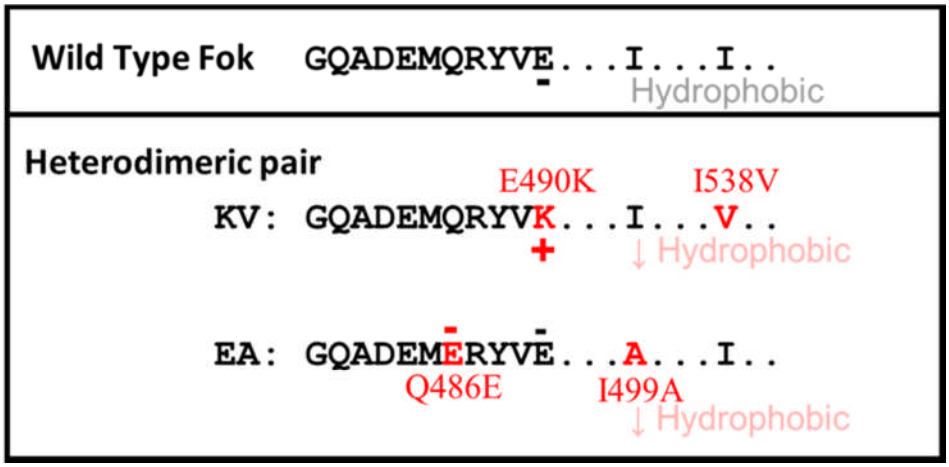


Figure 3.1 The FokI obligate heterodimer

The charges of the amino acids at the Wild Type Fok1 dimer interface compared to the charges engineered for the heterodimeric pair.

3.3 ZFN expression analysis

Due to the novel ZFN chimeric protein assembly of the Sangamo designed *IL2RG* specific zinc finger sequences and the alternative FokI heterodimeric pair: EA-KV, it was necessary to confirm whether this construct was capable of ZFN expression *in vitro*. Furthermore, it was essential to confirm that such expression was capable of cleaving the *IL2RG* target site.

3.3.1 Expression of the individual ZFN: EA and KV

In order to confirm that the Left and Right ZFN expression plasmids with EA or KV FokI nuclease domain respectively were capable of the individual ZFN expression a Western blot was carried out (Figure 3.2). 293T cells were transfected with the individual HA tagged ZFN expression plasmids and a GFP expression plasmid was delivered alongside as a transfection control. Protein lysate from the ZFN chimera treated cells was isolated 48 hours post transfection and quantified in order to assess expression in equal measure. The 40kDa Left or Right ZFN protein stained positive for HA. Likewise, GFP was also present as confirmed by antibodies directed against GFP.

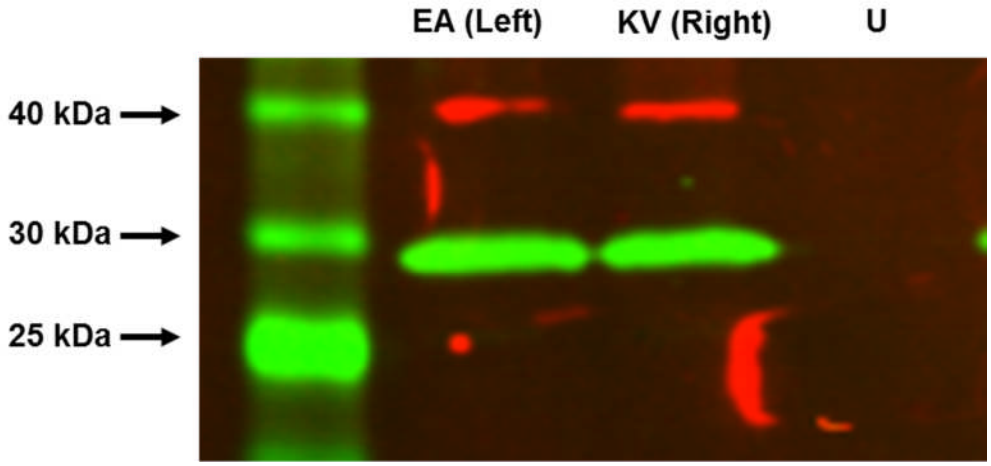


Figure 3.2 A Western blot to demonstrate ZFN expression.

293T cells were transfected with plasmids encoding either the Left or Right ZFN and the protein lysate harvested 48 hours later. A western blot of the lysate was carried out. Red; the HA tag. Green, the GFP positive control. Samples were run alongside an untransfected 293T control (U).

3.3.2 Detecting NHEJ in DNA as a result of ZFN directed cleavage

The principal function of ZFNs is to induce a DSB at the target site for which the zinc fingers were designed and assembled. Such a genetic modification can result in the DNA repair process of error-prone NHEJ to re-ligate the DNA strands together, frequently incorporating insertions or deletions (indels) in the process. Indels incorporated at the ZFN target site provide evidence of NHEJ which in turn confirms the functional expression of ZFN by their induction of a DSB.

Indels created as a result of ZFN treatment invariably result in a variety of different DNA sequences which can be recognised and digested by T7 endonuclease when these heterogenous DNA fragments are annealed at random. A schematic for the indel detection assay is shown in Figure 3.3a. 293T cells were transfected with expression plasmids for either the individual ZFN Left or Right or as a ZFN pair and incubated for 96 hours. The *IL2RG* target sequence was analysed for indel incorporation as follows: the genomic DNA was isolated from the cells under each condition and PCR carried out to achieve 500bp amplicons with the ZFN target site at the centre of the amplicon. After denaturing and re-annealing of the PCR product, the DNA was incubated with T7 endonuclease. The 250bp product indicative of T7 endonuclease digestion was present when the ZFN pair were administered together thus confirming *IL2RG* targeting (Figure 3.3b). Furthermore, the Left or Right ZFN alone did not result in the incorporation of indels, at the ZFN target site in *IL2RG*. The presence of indels at the *IL2RG* target site in the cells transfected with the ZFN pair but not in cells transfected with individual ZFN expression plasmid alone confirms that this pair is capable of both target sequence cleavage and appropriate heterodimerisation at the target site.

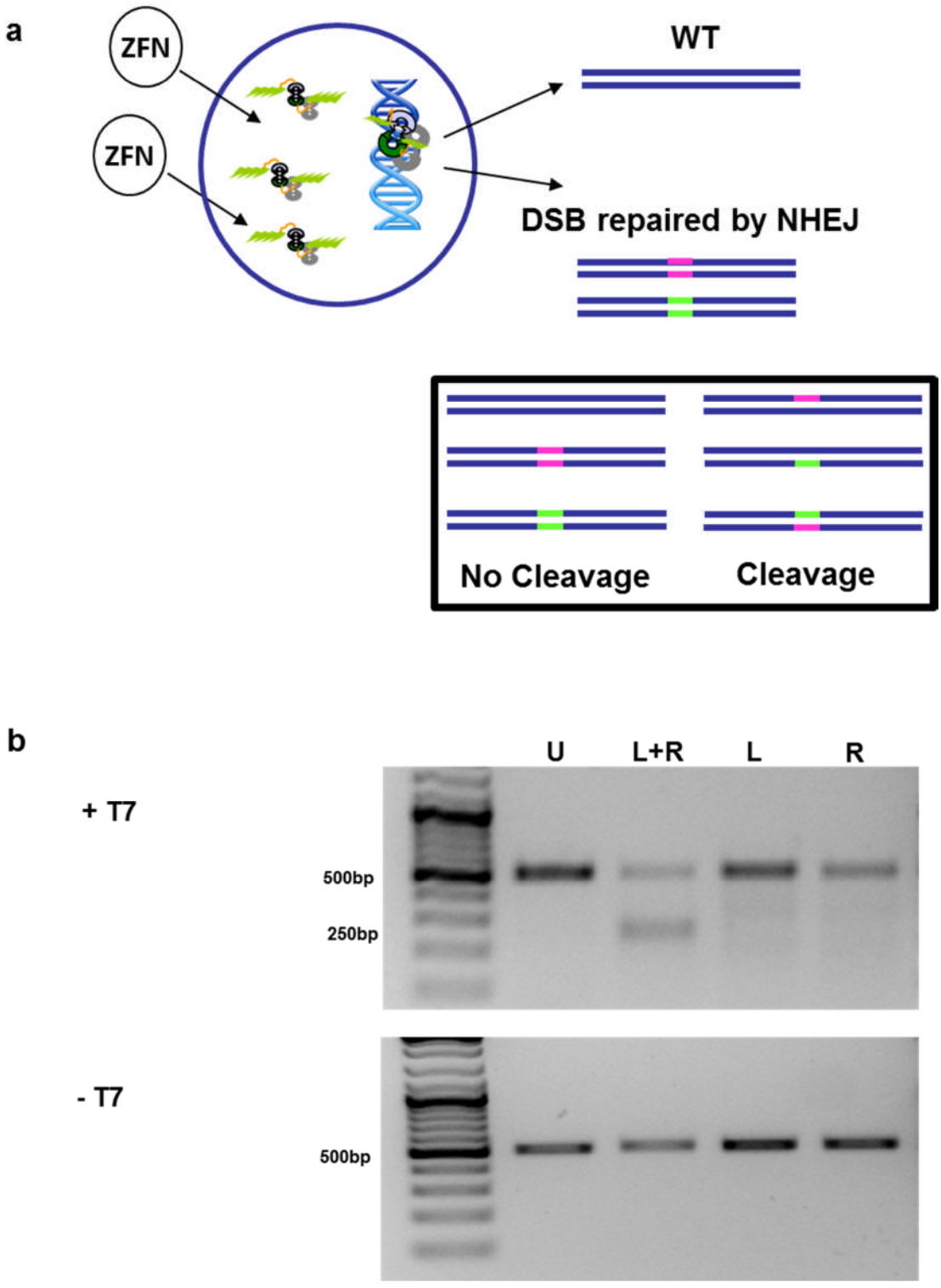


Figure 3.3 Detecting insertions and deletions (indels) as a result of ZFN induced double strand breaks (DSB).

(a) A schematic of the T7 endonuclease assay. 293T cells are transfected with Left (L), Right (R) or both ZFN expression plasmids. Genomic DNA isolated after 96 hours and the ZFN target site is amplified by PCR. Indels incorporated as a result of ZFN induced NHEJ are detected and cleaved by T7 endonuclease. (b) Results of the T7 endonuclease assay. 500bp target templates produced by PCR from genomic DNA were incubated with T7 endonuclease. The 250bp digest products are separated by electrophoresis as shown.

3.3.3 ZFN mediated cleavage assessed by an episomal GFP knockdown assay

To help quantify the rate of ZFN cleavage at the *IL2RG* locus, a GFP knockdown assay was performed. An expression plasmid containing *IL2RG* cDNA upstream of the GFP gene is transcribed into a γc-GFP fusion protein, on cleaving the *IL2RG* target site by ZFN the plasmid will be disrupted and the expression of GFP will be prevented (Figure 3.4a). Considering both the γc-GFP and ZFN rely on the CMV promoter for expression, to account for any loss of GFP expression due the incorporation of the transcription machinery at the ZFN's promoter, plasmids containing CMV upstream of SceI were transfected in place of ZFN in the GFP control. An mCherry expression plasmid was included in all conditions as a transfection control. 293T cells were transfected with γc-GFP expression plasmids and either ZFN or CMV control plasmids and analysed for GFP expression 48 hours later by flow cytometry (Figure 3.4b). Transfection with the ZFN pair reduced the mean fluorescence intensity (MFI) of GFP signal more than 4 fold compared to the CMV control, and although the signal was not completely abolished, the notable reduction in the fluorescence of GFP shows that the ZFN pair is capable of episomal *IL2RG* disruption.

It was also considered that the mCherry positive cells may not have represented all of the transfected GFP and control 293T cells. Further analysis of GFP expression in the total live cells confirmed the same pattern of GFP knockdown as mCherry gated cells (Appendix 3).

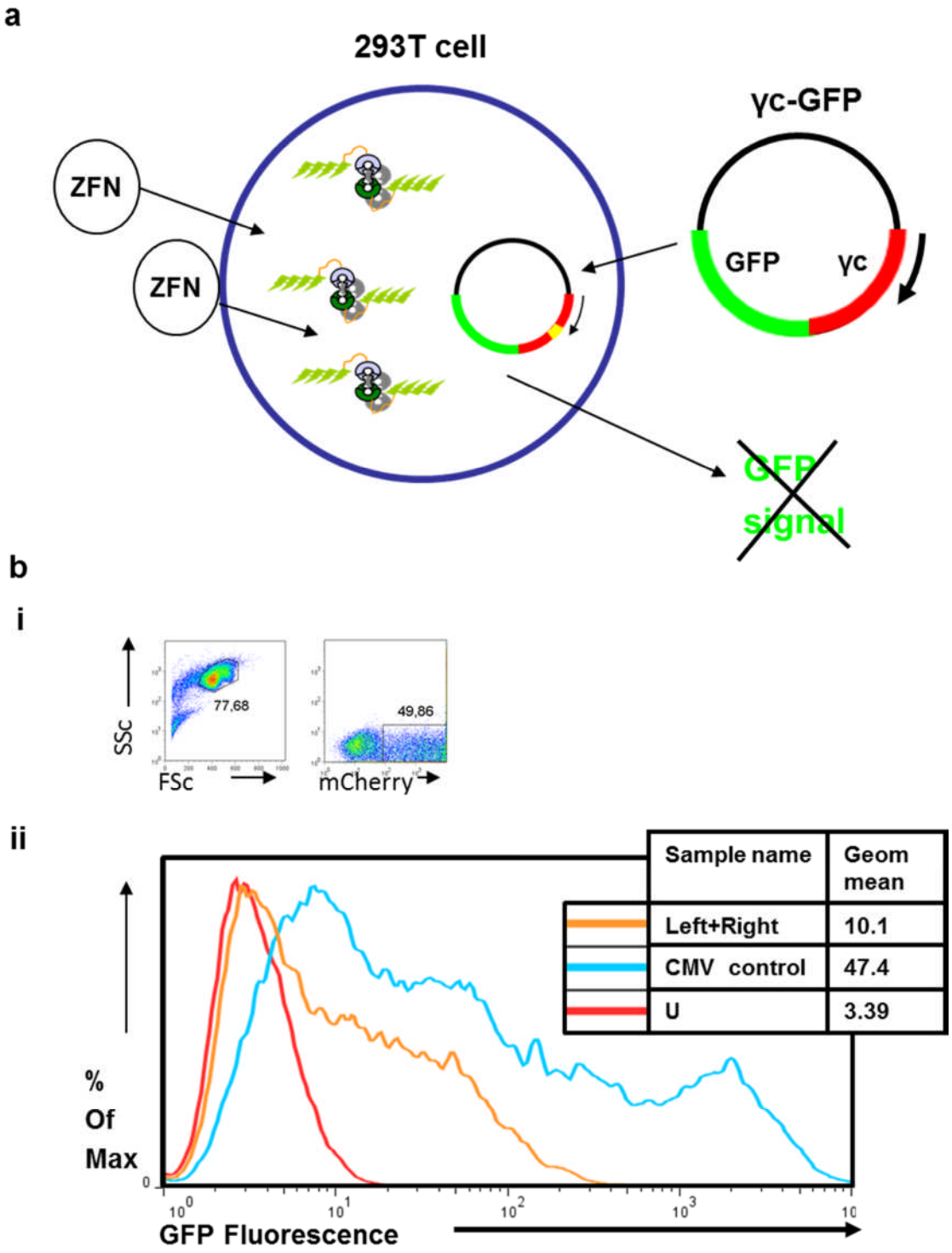


Figure 3.4 GFP knockdown assay.

(a) A schematic for the process for episomal GFP knockdown. 293T cells are transfected with γc-GFP expression plasmids alongside ZFN expression or CMV control plasmids. GFP expression is analysed by flow cytometry. (b) The results of the GFP knockdown assay. (i) Live cells were gated positive for the transfection control, mCherry. (ii) The MFI and Geometric mean (Geom mean) of GFP on mCherry+ cells is shown by a histogram for ZFN pair treated (Left+Right), CMV promoter control and untransfected control (U).

3.4 ZFN mediated cleavage of WT, Mutant and repair *IL2RG* DNA

It has been established so far that the IL2RG ZFN are capable of targeted cleavage of the *IL2RG* locus as shown by the detection of DSB and the indels incorporated as a result of NHEJ. Although techniques to detect NHEJ are useful to assess the performance of the ZFN they cannot provide information regarding HR events which are necessary to achieve gene correction.

In order to achieve targeted HR, together with delivering ZFN expression vectors to a cell to attain a DSB, a donor molecule containing a sequence capable of correcting the gene must also be provided. Due to the faithful DNA binding of sequence-specific ZFN constructs to their target, certain considerations are necessary for the design of donor DNA constructs for the correction of one base pair. There are three major concerns concerning target sequence similarity. Firstly, cleavage of the donor fragment by ZFN could hinder the efficiency of HR, secondly if the repair sequence is a target for the ZFN, this would permit further cutting after recombination (Pruett-Miller, Connelly et al. 2008) and finally, conversion of the point mutation G691A to WT will not introduce a new restriction enzyme site and therefore would be difficult to detect during genomic analysis. These concerns can be overcome by designing a donor DNA containing ZFN blocking mutations which can subsequently be used to detect targeted recombination.

The IL2RG ZFN pair has been shown to cut WT sequence at the site of a common SCID-X1 associated point mutation: G691A located in exon 5 of the gene (Miller, Holmes et al. 2007). As this missense mutation occurs in the zinc finger recognition site (Figure 3.5ai) it was necessary to confirm that the ZFN pair was capable of performing site directed cleavage of the mutated target site. Furthermore the extent of base pair diversity that is tolerated by the binding of this particular ZFN pair should be known to construct suitable donor DNA repair constructs.

3.4.1 Design of donor DNA constructs with ZFN blocking mutations are capable of restoring γc expression

Using the degenerate nature of the codon code, the ZFN binding sequence of the donor plasmid was modified to incorporate sequence disparity from WT by 1, 3 and 17bp whilst maintaining the same amino acid code. These proposed constructs were

named MB1, MB3 and MB17 respectively (Figure 3.5aii). In all three designs, the sequence modification resulted in a loss of a restriction site for DraIII and a gain of a BsrBI site in all three plus an additional DpnII restriction site in the MB17 construct.

These modified binding cassettes were constructed as follows: A WT donor construct was achieved by amplifying a 1,532bp fragment from WT *IL2RG* sequence from the full length gene located in a γc expression plasmid (Appendix 1) with the ZFN target sequence located in the middle, flanked by left and right homology arms of 612bp and 889bp. The A-G modification in MB1 was introduced into the WT donor construct consisting of the same homology arms flanking the ZFN target sequence by site directed mutagenesis. The T-C, C-T and A-G modifications in MB3 were manufactured by GeneArt and the synthesised cassette was subsequently cloned in the WT donor construct. The MB17 cassette was constructed by linker mediated PCR whereby primers were designed to incorporate the 17bp changes whilst maintaining the ability to bind WT sequence (Figure 3.5b). The first round of PCR involved two PCR reactions each generating a PCR product with the MB17 sequence overhangs at the 3' or 5' end. The second round of PCR then relied on the modified overhangs to anneal and the PCR was completed by the flanking primers used in the first round, achieving an amplicon containing the 17bp modifications which replaced the WT ZFN target site. This amplicon was purified from the PCR gel and cloned into the pJet backbone (Appendix 5).

To confirm that the modified region, when incorporated into *IL2RG,* would still be capable of γc expression, the modified cassette from MB17, containing the maximum number of DNA base modifications used in this study, was cloned into a γc expression plasmid containing the complete endogenous *IL2RG* gene and transfected into 293T cells. An increase in γc expression was observed 48 hours post transfection in cells treated with both WT *IL2RG* and *IL2RG* containing the MB17 cassette compared to untransfected 293T cells (Figure 3.5c).

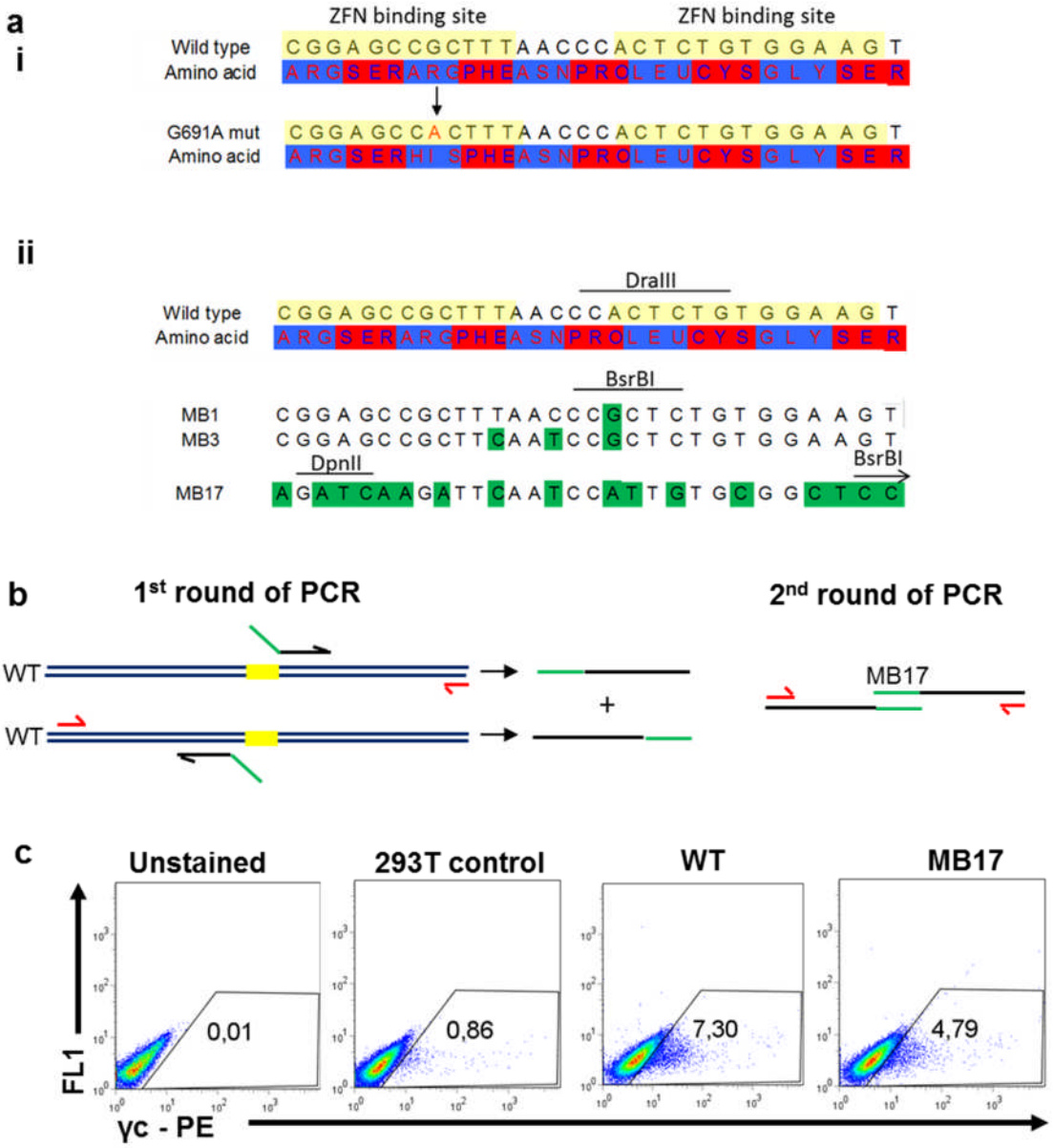


Figure 3.5 Modifying the ZFN binding domain.

(a) The ZFN binding domain (yellow boxes) of *IL2RG.* (i) The G691A SCID-X1 associated mutation is located within the ZFN binding site (ii) Designing a ZFN *IL2RG* donor with a modified binding site. 3 different Donor DNA constructs, MB1, MB3 and MB17 were designed based on the flexibility of the redundant amino acid code, changes in DNA sequence from WT are shown in green and the corresponding restriction enzyme site for each construct is labelled. (b) Construction of MB17 by linker mediated PCR. Primers designed to incorporate the 17bp change (green) at the ZFN binding site (Yellow) undergo 1 round of PCR with flanking primers (Red) and generate an amplicon from WT template (Blue) followed by a second round where the PCR amplicon becomes the template and the cassette is completed by the flanking primers. (c) 293T cells were transfected with either plasmid encoding the complete *IL2RG* gene with either the WT or MB17 modified ZFN target sequence, untransfected 293T cells were included as a control. After 48 hours the cells were stained with PE conjugated antibodies directed against γc and acquired by flow cytometry. The numbers shown indicate the percentage of γc positive cells.

3.4.2 ZFN translated *in vitro* is capable of targeted DNA cleavage of both WT and SCID-X1 associated *IL2RG* mutation G691A but not donor constructs with ZFN blocking mutations

Having established that the 17 DNA base modifications did not disrupt γc expression, it was necessary to confirm that these modifications abolished ZFN targeted cleavage. ZFN protein was translated *in vitro* using rabbit reticulate lysate and incubated with 500bp target fragments generated by PCR from the WT, MB1, MB3 and MB17 donor constructs. The digestion products were then separated by electrophoresis. In order to confirm the correct product size and inclusion of a new restriction enzyme site, the WT was also digested with DraIII and the modified donors were digested with BsrBI and separated alongside the ZFN digestion products (Figure 3.6a). The WT template digested with DraIII resulted in 250bp products. Likewise, all three modified donors were digested by BsrBI to achieve nearly 100% digestion. A faint 250bp band indicates the translated ZFN pair was able to cleave about 13% of WT template. There was evidence of digestion of the MB3 template by the ZFN pair however at a lower rate of 3.5%. No other donor template showed any digestion product when treated with ZFN. As the WT control rendered a weak digestion product I am unable to confirm that these donors are resistant to ZFN cleavage from this assay. Furthermore, successful ZFN translation was confirmed by Western blot (Appendix 6).

Considering the missense SCID-X1 mutation occurs in the ZFN binding domain of *IL2RG*, the *in vitro* ZFN digestion assay was carried out to assess the ability of the ZFN pair to target the SCID-X1 mutation. 500bp ZFN target templates from WT, G691A *IL2RG* mutant and the MB17 donor construct were amplified by PCR and incubated with ZFN translated by rabbit reticulate lysate *in vitro* (Figure 3.6b). Digestion with DraIII or DpnII was included as a marker for cleavage product size. The results of the digestion show that the ZFN pair was able to cleave both the WT and SCID-X1 target sequence at 67% and 61% respectively, however, the MB17 target remained uncut. Furthermore, when the target was incubated with either ZFN individually, no cleavage occurred, thereby confirming that the ZFN must be present as a heterologous pair in order to introduce a DSB at their target site. Due to the clear result that the translated ZFN can digest the WT and G691A mutant, the complete lack of MB17 digestion implies that the MB17 modified binding site is resistant to cleavage by *IL2RG* ZFN.

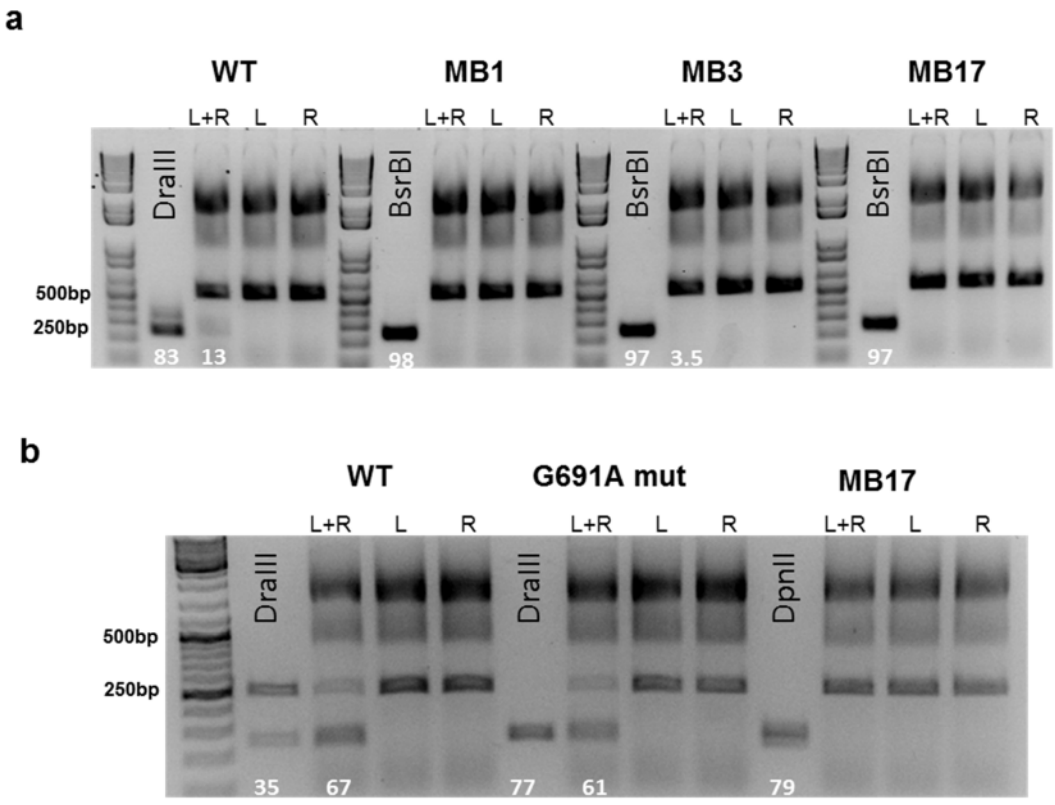


Figure 3.6 ZFN mediated cleavage of the WT, SCID-X1 and modified donor DNA sequences.

Both the Left (L) and Right (R) ZFN were translated *in vitro* by rabbit reticulate lysate. The translation products were then incubated with target template produced by PCR. L or R ZFN were incubated as a control to confirm ZFN pair were obligate heterodimers. Restriction enzyme digest of the target template confirmed inclusion of the new restriction enzyme site and cleavage product size. (a) The results of the *in vitro* ZFN digestion assay. Target template from WT, MB1, MB3 and MB17 were digested with L, R or L+R alongside either BsrBI or DraIII. The products were separated by electrophoresis. (b) 500bp DNA template from wild type, SCID-X1 mutation or Modified binding site (MB17) were incubated with L, R or L+R ZFN. DNA digested with DraIII and DpnII were used as a positive cleavage control. The numbers in white denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

3.5 *IL2RG* ZFN are capable of targeted HR as shown by HR specific PCR

It has been established that a donor DNA molecule containing the WT *IL2RG* sequence and therefore the ZFN target, can be cut resulting in an incomplete repair matrix and therefore theoretically unsuitable for HR. Modification of the ZFN target site in the donor DNA, MB17, produced DNA resistant to ZFN cleavage and resulted in the incorporation of a new restriction enzyme site. Furthermore, the disparity of 17bp from WT, if incorporated into the endogenous *IL2RG*, will provide a new sequence at the target site which may be exclusive for specific oligonucleotides. Utilising this in a PCR reaction may offer the ability to rapidly detect targeted HR with less expense compared to high throughput screening of recombination of the target site.

Another consideration for the donor DNA repair matrix is the size of the homology domain (HD) between the donor and target gene. Donor DNA containing two sizes of HD were compared, one of 978bp (short) and the other 1531bp (long) with the target site for correction and detection, located in the middle.

The PCR strategy to detect targeted HR events is outlined in Figure 3.7 Primer pair 1 consists of a forward primer which anneals to the modified binding site (MBS) and a reverse primer binding in exon 7 of the genomic DNA, outside of the homology domain corresponding in the long donor DNA. Primer pair 2 is designed with the same principle but with the forward primer targeted to exon 4, upstream of the long donor's homology domain and the reverse primer to the MBS. This strategy is employed to detect HR in 293T cells, ED7R cells and finally mES cells over the course of this project. Concatamers that may have been incorporated as a result of multiple HR events at the target site can be detected by using the exon 4 forward primer and the exon 7 reverse primer and a PCR elongation time suitable to amplify DNA regions in excess of the endogenous size.

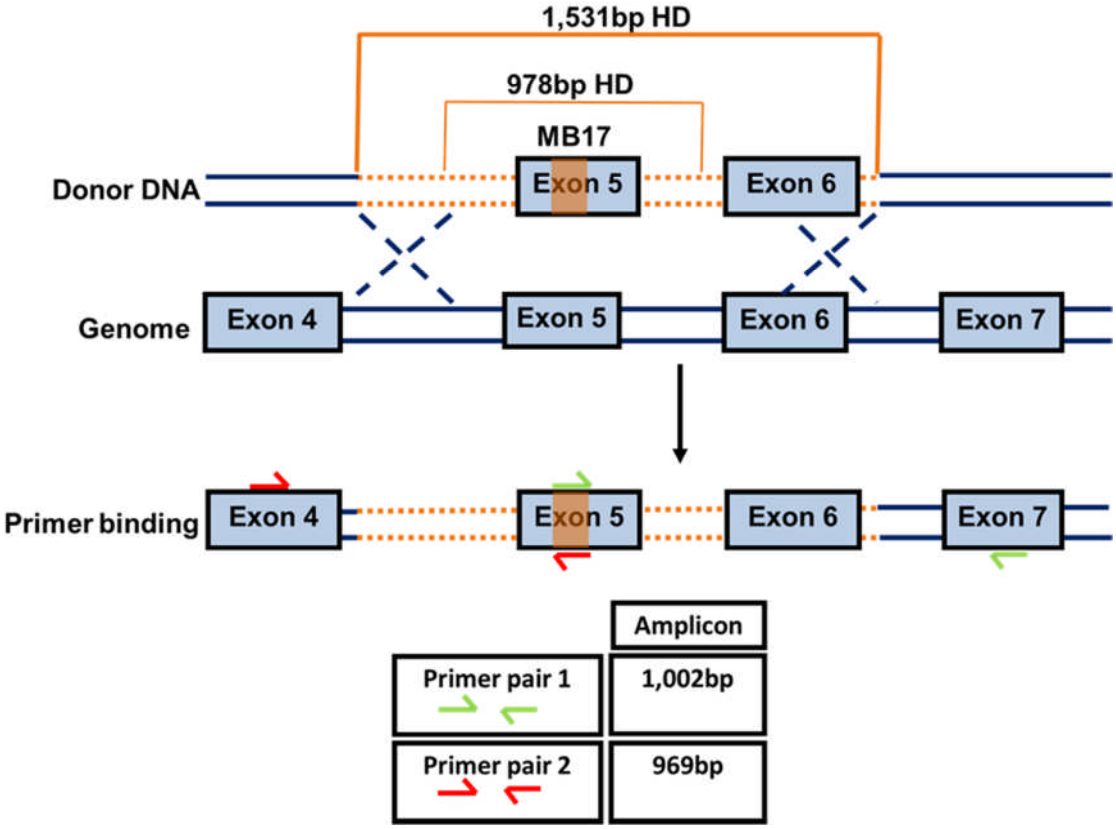


Figure 3.7 A schematic of the primer strategy to amplify targeted HR.

A schematic of the primer strategy to amplify targeted HR. The long (1,531bp) or short (978bp) homology domain (HD) of the donor DNA (Broken orange line) is incorporated into the genomic DNA by HR. the modified binding site MB17 in exon 5 (orange rectangle) is detected by primer pair 1 (green arrows) and results in an amplicon length of 1,002bp. Primer pair 2 (red arrows) results in an amplicon length of 969bp.

3.5.1 Detecting ZFN induced HR in 293T cells

Two sets of transfections were carried out on 293T cells, with ZFN plasmids and donor DNA constructs containing the MB17 cassette at ZFN:Donor ratios of 1:3, 1:1 and 3:1. One set of transfections was carried out with the short (978bp) HD, the other with the long (1,531bp) HD at a total of 2µg per transfection alongside an un-transfected control. Transfected cells were incubated at 37°C for 96 hours at which point the genomic DNA was isolated and analysed for targeted HR by PCR (Figure 3.8a). The results show that on transfection with the short donor, the PCR product amplified by primer pair 2 was present at a transfection ratio of 1:3 and 3:1 ZFN:Donor. However there were no amplicons produced using primer pair 1. Cells transfected with the ZFN plasmids and the long donor DNA construct resulted in positive PCR amplicons in all three ZFN:Donor ratios detected using both primer pair 1 and 2. Regarding the consistency of the long donor plasmid to achieve HR, this length of HD at 1,531bp was chosen for all further donor DNA construct analysis. The positive control used in this PCR is the *IL2RG* expression plasmid containing the MB17 modified binding cassette in place of the WT ZFN target. Using this template, it appears that the band produced by primer pair 1 is of a greater intensity than that produced by primer pair 2. It is worth noting that there is no detectable background integration of long or short donor alone in 293T cells at the *IL2RG* target site.

In order to verify that the targeted HR detected by PCR had resulted in a single exchange of donor DNA with the endogenous locus, a PCR was performed to detect any concatamers, similar to the head to tail repeats seen by Lombardo *et al*. (Lombardo, Genovese et al. 2007) that may have resulted during the DNA repair process (Figure 3.8b). The primers for exon 4 and exon 7 were utilised to detect any genomic DNA in excess of the 1,901bp endogenous sequence by using a PCR elongation time of 4 minutes. Using this technique, no evidence of concatamers was seen.

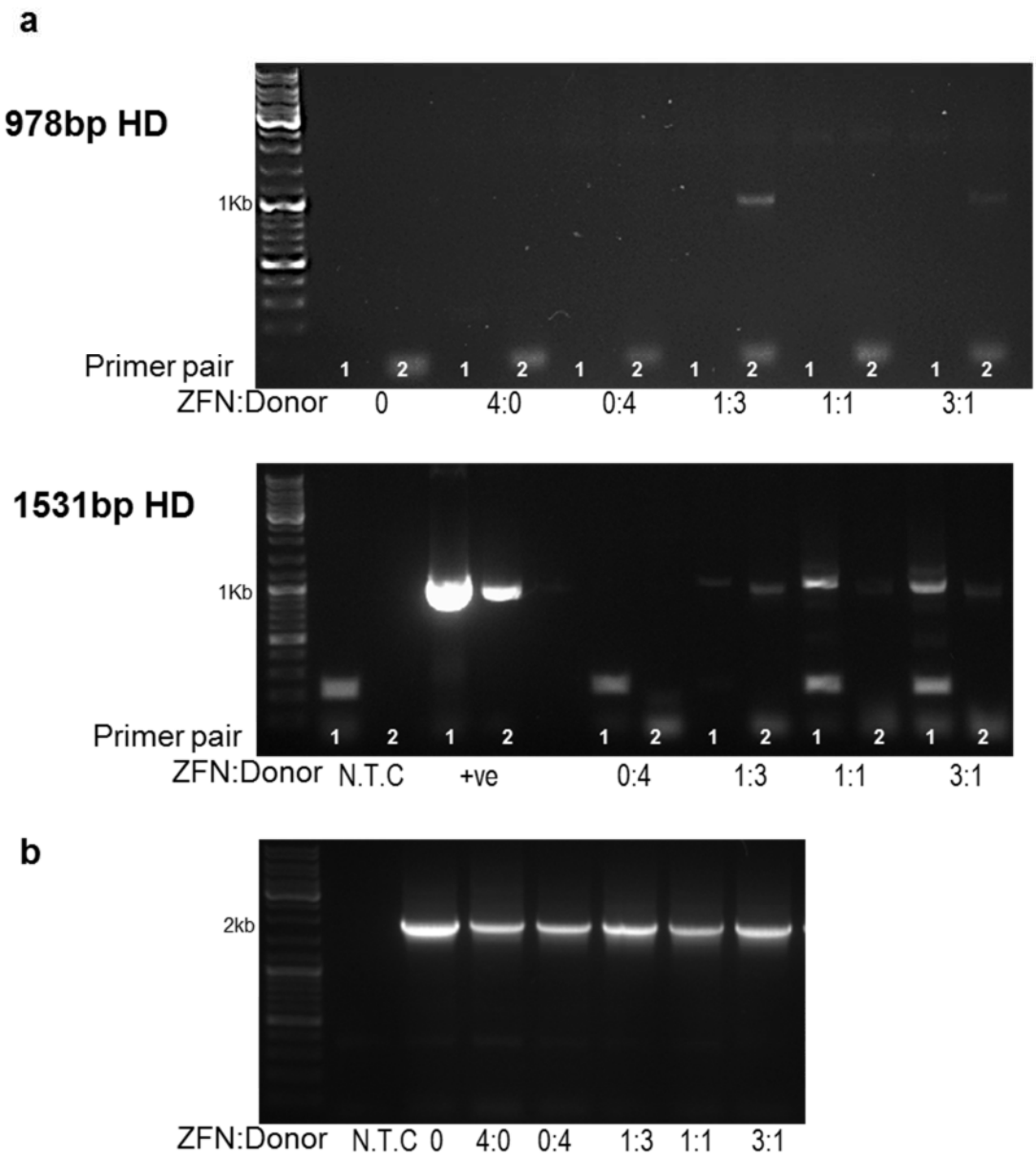


Figure 3.8 Targeted HR in 293T cells.

293T cells were transfected with ZFN expression plasmids or donor DNA constructs at various ZFN:donor ratios as indicated. Genomic DNA was isolated 96 hours later. (a) Results of the PCR to detect HR. Top, cells transfected with the short (978bp) HD domain. Bottom, cells transfected with the long (1531bp) HD domain. (c) Results of the PCR using primers to exon 4 and exon 7 to detect concatamers.

3.6 Discussion

This section shows that the IL2RG-specific zinc finger fused to heterodimeric FokI domains can be expressed in 293T cells. Furthermore, the ZFN pair is functional as they are capable of inducing a DSB at the *IL2RG* target locus in 293T cells as shown by the signature indels introduced as result of NHEJ. It was also shown that the IL2RG ZFN can cleave both the WT sequence and a common SCID-X1 associated mutation (G691A). Although ZFN exhibit excellent sequence specificity, there is a degree of flexibility in bp recognition as observed previously between A and G in the middle finger of the murine transcription factor Zif268 (Choo and Klug 1994).

Due to the extent of homology necessary between the donor DNA and the endogenous sequence, the donor molecule was modified to render it resistant to ZFN mediated cleavage. The ZFN recognition sequence was modified in 3 donor DNA constructs by 1, 3 and 17bp. On assessing the cleavage resistant modifications, the faint cleavage product from the MB3 template detected by band intensity analysis maybe an artefact of unusually low background or the ZFN have a low affinity for the new sequence. Furthermore, an inferior rate of digestion in the WT control at 13% compared with 67% in the subsequent assay with the same indicates the cleavage resistance should be re-assessed. The latter experiment did however provide evidence that the MB17 donor construct is resistant to ZFN targeting. Furthermore, the 17bp modifications from WT in this construct were shown not to disrupt γc expression making it a suitable candidate for *IL2RG* correction.

The advantages of such designs are that the donor molecule will remain intact during ZFN treatment and will prevent subsequent cutting once incorporated into the endogenous *IL2RG*. In addition, such a change in sequence is a marker for the detection of HR events. Firstly, the introduction of the novel BsrBI restriction site can be detected exclusively of the endogenous sequence as seen by Millar *et al.* (Miller, Holmes et al. 2007). Furthermore, the incorporation of a new sequence with enough sequence disparity from endogenous sequence can provide a target detectable by PCR. This was shown in 293T cells transfected with ZFN expression plasmids and the MB17 donor. A possible disadvantage is that the decrease in homology between the donor

and endogenous sequence could decrease the likelihood of HR events. Direct comparison of the 3 modified donor DNA would be useful to investigate the effect these modifications have on HR stimulated by ZFN induced DSB.

The length of HD between the donor DNA and endogenous target locus was briefly investigated. Targeted HR in 293T cells was more consistent when an HD of 1531bp was present in the donor DNA when compared to an HD of 978bp. Although confounding evidence that homology arms >50bp are not necessary to exchange a short region of sequence (Orlando, Santiago et al. 2010) this extent of homology was also successful for the insertion of a 7,762bp sequence at the same *IL2RG* locus by Moehle *et al.* (Moehle, Rock et al. 2007). Another observation when comparing these 2 HD lengths is the inconsistency in modified binding site detection. This is likely to be due to differences in primer pair efficiency as observed in the positive control. Finally, the ratio of ZFN to donor DNA was compared in transfected 293T cells. The ZFN:Donor DNA ratios of 1:1 and 3:1 when using the Donor DNA with the long HD, resulted in a PCR product of visibly more intensity.

To overcome the limitations of PCR to quantify targeted HR, it is possible to design a fluorescent probe to bind the modified site for use in qPCR. Accompanying the quantification of the modified binding site with a house keeping gene such as actin, different conditions such ZFN to donor DNA ratios could be compared. This method has been successful for Greenwald *et al.*(Greenwald, Cashman et al. 2010).

Having established functional expression of the ZFN pair, observing faithful heterodimerisation at the *IL2RG* locus and having developed donor DNA constructs which can undergo targeted HR, it remained to be assessed whether targeted gene correction of *IL2RG* can result in functional γc expression which is explored in the following chapter.

4 Developing a SCID-X1 cell line model for ZFN targeting analysis.

4.1 Aims

- To characterise the ED7R cell line
- To modify this cell line to contain 1 copy of either WT or G691A *IL2RG*
- To demonstrate ZFN mediated gene targeted knockdown of WT *IL2RG*
- To demonstrate ZFN mediated correction of the G691A mutation resulting in γc restoration and signalling
- To assess different ZFN and donor DNA delivery methods
- To assess the effect of sequence dissimilarity at the ZFN binding site on the donor DNA

4.2 Introduction

As previously described, the *IL2RG* ZFN pair is capable of targeting *IL2RG* in an episomal γc-GFP fusion construct and the endogenous locus in 293T cells, furthermore, this targeted DSB can be repaired by HR on the provision of a donor DNA construct in these cells. In order to further examine this gene correction strategy, a cell line model of SCID-X1 has been constructed that when corrected will restore γc expression. The common SCID-X1 mutation G691A was chosen to be included in the target locus for ZFN due to its frequency in SCID-X1 cases and its location in the mutation hotspot of exon 5. Due to the nature of base pair exchange during HR, extensive stretches of DNA can be recombined >100bp from the cut site (Elliott, Richardson et al. 1998; Herrmann, Garriga-Canut et al. 2011) which would suggest that a host of SCID-X1 mutations could be corrected via HR stimulated by this ZFN pair.

The ED7R cell line is a model of SCID-X1 based on the characteristic that it does not express γc resulting in incomplete IL2R, which doesn't respond to IL-2 and is unable to phosphorylate the subsequent signalling steps of the Jak-Stat pathway. The cell line was originally cultured from a patient with HTLV-1 associated T cell leukaemia and lacked both γc and IL7Rα. Kumaki *et al*., re-introduced IL7Rα cDNA to the genome, hence the name ED7R, and on subsequent addition of the *IL2RG* cDNA, termed EDγc7R, signalling via the receptors for IL-2, IL-4 and IL-7 was restored as shown by the tyrosine phosphorylation of Jak3 (Kumaki, Ishii et al. 1999). Lentiviral vectors (LV) have been used to reintroduce *IL2RG* into ED7R cells with the promoters SFFV, UCOE, (Zhang, Thornhill et al. 2007), PGK, LTR, WASP, Jak3 (Smyth, Ginn et al. 2007) and EF1α

(Smyth, Ginn et al. 2007; Throm, Ouma et al. 2009) and the subsequent γc expression proven functional by cytokine signalling resulting in the phosphorylation of Stat5. Stat5a/b have been shown to be IL2RG specific as mice deficient in these proteins are unable to undergo IL-2 induced cell cycle progression (Moriggl, Topham et al. 1999).

An LV vector was designed (W.Qasim) incorporating full length *IL2RG* in reverse orientation to avoid intron splicing during lentiviral vector production and transgene integration (Appendix 1). The *IL2RG* endogenous promoter, designed by S. Thornhill to include binding sites for members of the Ets family of transcription factors, (Markiewicz, Bosselut et al. 1996; Ohbo, Suda et al. 1996) and regulatory elements conserved between human and mouse. This promoter, in conjunction with *IL2RG* cDNA successfully restored γc expression in SC-1 fibroblasts and a SCID-X1 LCLs and *IL2RG* downstream signalling in ED7R cells. This WT *IL2RG* LV is used to construct ED7R based cell lines for ZFN gene targeting in this chapter.

Integrase deficient lentiviral vectors (IDLV) have been used to deliver targeting endonucleases and donor DNA (Cornu and Cathomen 2007; Lombardo, Genovese et al. 2007) from which the transient endonuclease expression resulted in successful targeted gene modification. ZFN delivery by LV, IDLV and nucleofection is attempted in this chapter.

4.3 Characterising and constructing the ED7R cell line SCID-X1 model

Due to the unknown nature of the *IL2RG* locus in ED7R cells and why γc is not expressed, the cell line must be characterised to establish whether it is suitable model for targeted for gene correction. The aim of this chapter is to characterise the cell line and to incorporate a single copy of mutant *IL2RG* or WT *IL2RG* for control. These modified cell lines will be used to optimise ZFN mediated gene targeting and correction and to assess whether this gene correction strategy results in functional IL2R signalling.

4.3.1 The ED7R Karyotype

The *IL2RG* locus is situated on the X-chromosome at Xq13.1. Karyotype analysis of the male derived ED7R cell (Figure 4.1a) revealed gross chromosomal aberrations, most importantly a duplication of the X-chromosome. Sequencing of the *IL2RG* locus from 207bp upstream of the promoter and regulatory elements to 350bp downstream of

the polyA signal sequence revealed the deletion of a thymine (T) in the polypyrimidine tract in intron 4 (Figure 4.1bi). As this deletion occurs in an intron it will not result in frameshift of translated sequence, furthermore the affected polyT tract is not located near the 3' splice acceptor site which negates aberrant spicing. However, it is not refuted that intron 4 may be involved in the regulation of γc expression.

The sequence analysis also confirmed the presence of intact *IL2RG* ZFN target sequence and clinical sequencing confirmed this sequence to be homologous which, considering the duplication of the X-chromosome equates to two ZFN target sequences per ED7R cell with the assumption that no recombination events have occurred to incorporate the target external to the normal loci (Figure 4.1bii).

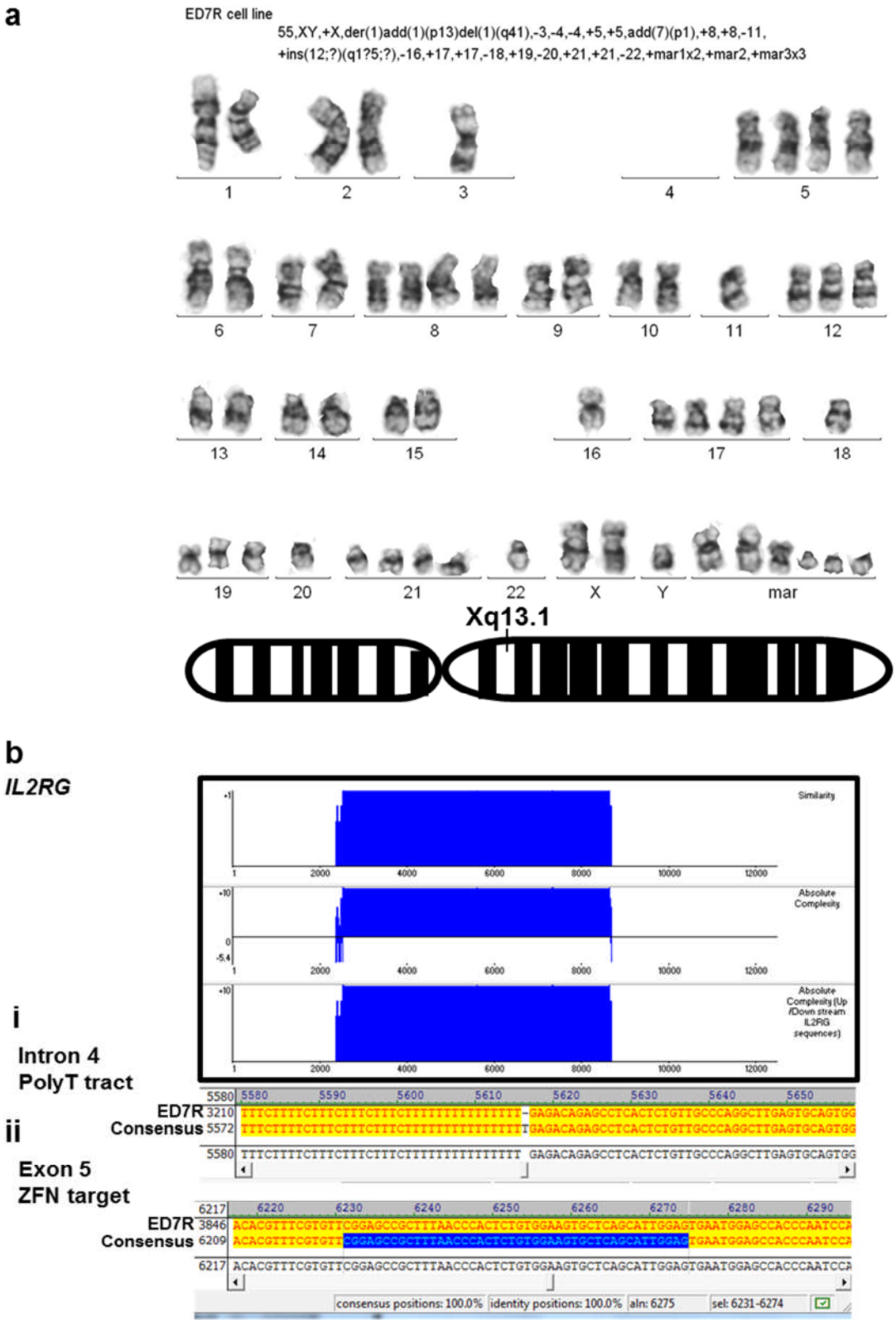


Figure 4.1 *IL2RG* in ED7R cells.

(a) The ED7R karyotype. Top, chromosomes identified by G-banding. Bottom, the location of the *IL2RG* locus in the X-chromosome. (b) Sequence analysis of ED7R compared to the consensus sequence available from NCBI. (i) The ED7R *IL2RG* locus has a deletion of a thymine (T) in the polyT tract of intron 3. (ii) The ED7R genome has an intact *IL2RG* ZFN target sequence.

4.3.2 WT or G691A *IL2RG* is incorporated as a transgene into the ED7R genome

Although the ED7R cell line contains the ZFN target site, subsequent genomic modification at this site will not show translated evidence of *IL2RG* correction by ZFN. The ED7R cells were transduced with LV encoding either WT or G691A *IL2RG* with the endogenous *IL2RG* promoter. To accompany the SCID-X1 cell line, ED7R cells were also modified with WT *IL2RG* to confirm that the lentiviral vector was capable of inserting a transgene capable of functional γc expression and secondly, the ED7R cell line with WT sequence can be used to measure ZFN targeting as NHEJ of the DSB would result in the knockdown of γc protein expression.

The mutant *IL2RG* LV construct was engineered by incorporating the G691A mutation via site directed mutagenesis of the plasmid vector containing the lentiviral backbone and WT *IL2RG* by oligonucleotides containing the point mutation (Figure 4.2a). Following a DpnI digest of un-mutated, methylated parental plasmid and amplification in *E. Coli*, sequencing confirmed successful incorporation of this missense mutation (Appendix 7). ED7R cells were transduced with a titration of either lentiviral vector followed by limiting dilution of the transduced cells to create clonal cell lines. ED7RγcWT clones were analysed for γc expression by flow cytometry in order to confirm successful insertion of the *IL2RG* transgene. As no γc expression was expected on the ED7RγcΔ clones, genomic DNA was isolated and analysed by PCR for WPRE to confirm insertion of the transgene. γc expression of a pooled population of ED7RγcΔ clones and a representative ED7RγcWT clone as compared to ED7R was analysed by flow cytometry (Figure 4.2b). As expected, γc expression was not detected on the cell surface of ED7RγcΔ cells as compared to the positive shift seen in the γc stained ED7RγcWT clone. To establish whether the γc mutation leads to a protein incapable of cell surface expression, intracellular γc staining was carried out, again the ED7RγcWT clone stained positive and the ED7RγcΔ were negative for γc confirming that the G691A mutation does not result in protein expression.

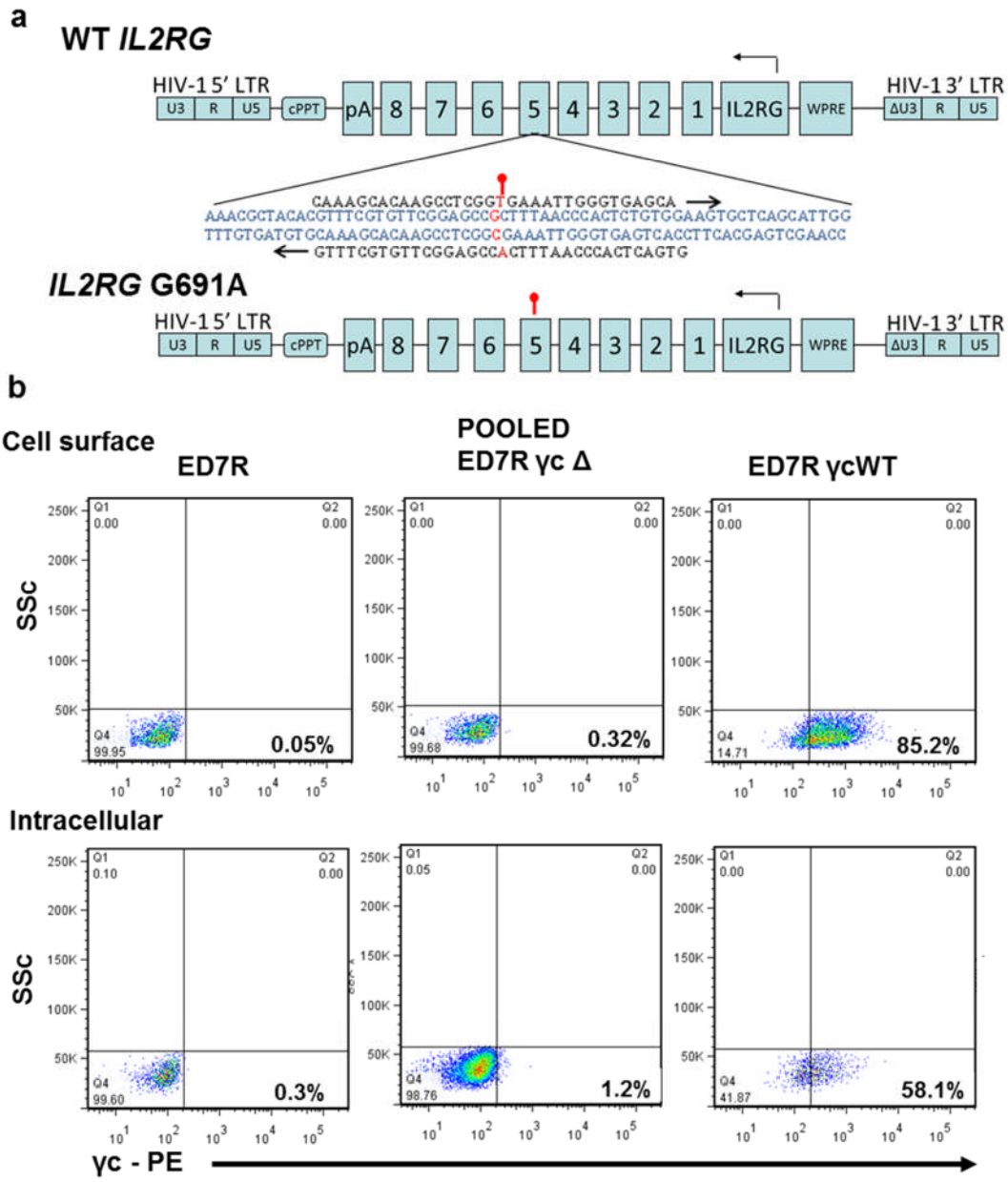


Figure .4.2 Making the ED7RγcΔ and ED7RγcWT cell lines.

ED7R cells were transduced with lentiviral vectors encoding WT or mutant *IL2RG*. (a) Self inactivating (SIN) lentiviral constructs used to introduce the G691A and WT *IL2RG* into ED7R cells. Top, the WT vector sequence was mutated via site directed mutagenesis by complementary primers (black) containing the G691A mutation (Red lollipop) resulting in a lentivirus construct with mutated *IL2RG*. (b) Cell surface (top) and intracellular (bottom) γc expression by pooled ED7RγcΔ clones, an ED7RγcWT clone and untransduced ED7R cells as detected by flow cytometry analysis of anti-γc antibody conjugated to PE.

4.3.3 Transcriptional and functional analysis of WT and mutant *IL2RG*

Having established that, like ED7R cells the ED7RγcΔ cells do not express γc and ED7RγcWT clones have γc expression restored, further analysis of these cell lines was carried out to investigate whether the ED7R or ED7RγcΔ cells express *IL2RG* mRNA and also, whether the γc expression on ED7RγcWT is functional.

Total RNA was extracted from ED7R, ED7RγcΔ and ED7RγcWT cells. The mRNA was reverse transcribed to cDNA by reverse transcriptase PCR (RT-PCR) and primers specific for *IL2RG* mRNA were used to amplify a 78bp product and primers specific for *GAPDH* mRNA was used as a template control (Figure 4.3a). A faint band represents the *IL2RG* mRNA PCR product from un-transduced ED7R cells and a band of greater intensity can be seen for both the ED7RγcΔ and ED7RγcWT clones. The intensity of *GAPDH* PCR product suggests a suitable concentration of template was used for each sample. The >300bp product seen in the NTC and ED7R was assumed to be an artefact of the *IL2RG* mRNA specific primers and not mRNA contamination as no *IL2RG* mRNA product is present.

In order to confirm that the WT IL2RG in the ED7RγcWT cells is functional and to confirm that no functional protein is translated in the ED7RγcΔ cells, the signalling molecule, pSTAT5 was analysed. On the binding of IL-2 at the cell surface by IL2R, the intracellular cytoplasmic tails of the IL2Rαβγ complex recruit JAK kinases which phosphorylate STAT5 which dimerise and enter the nucleus to activate transcription (Figure 4.3b). In order to prevent background phosphorylation of STAT5 by cross reactivity with bovine IL-2 in the foetal calf serum (FCS), the cells were starved in medium lacking FCS overnight. ED7R, ED7RγcΔ, ED7RγcWT and the previously published EDγc7R cells containing functional *IL2RG* cDNA were then pulsed with IL-2 and intracellular pSTAT5 was stained with antibodies and analysed by flow cytometry (Figure 4.3c). The results of this assay clearly show that on pulsing the cells with IL-2, ~52% of ED7RγcWT and ED7Rγc7R positive control cells demonstrated STAT5 phosphorylation, whereas the ED7R and ED7RγcΔ cells remained pSTAT5 negative. This confirms that the WT *IL2RG* transgene is functional and the G691A mutation renders the *IL2RG* gene product non-functional.

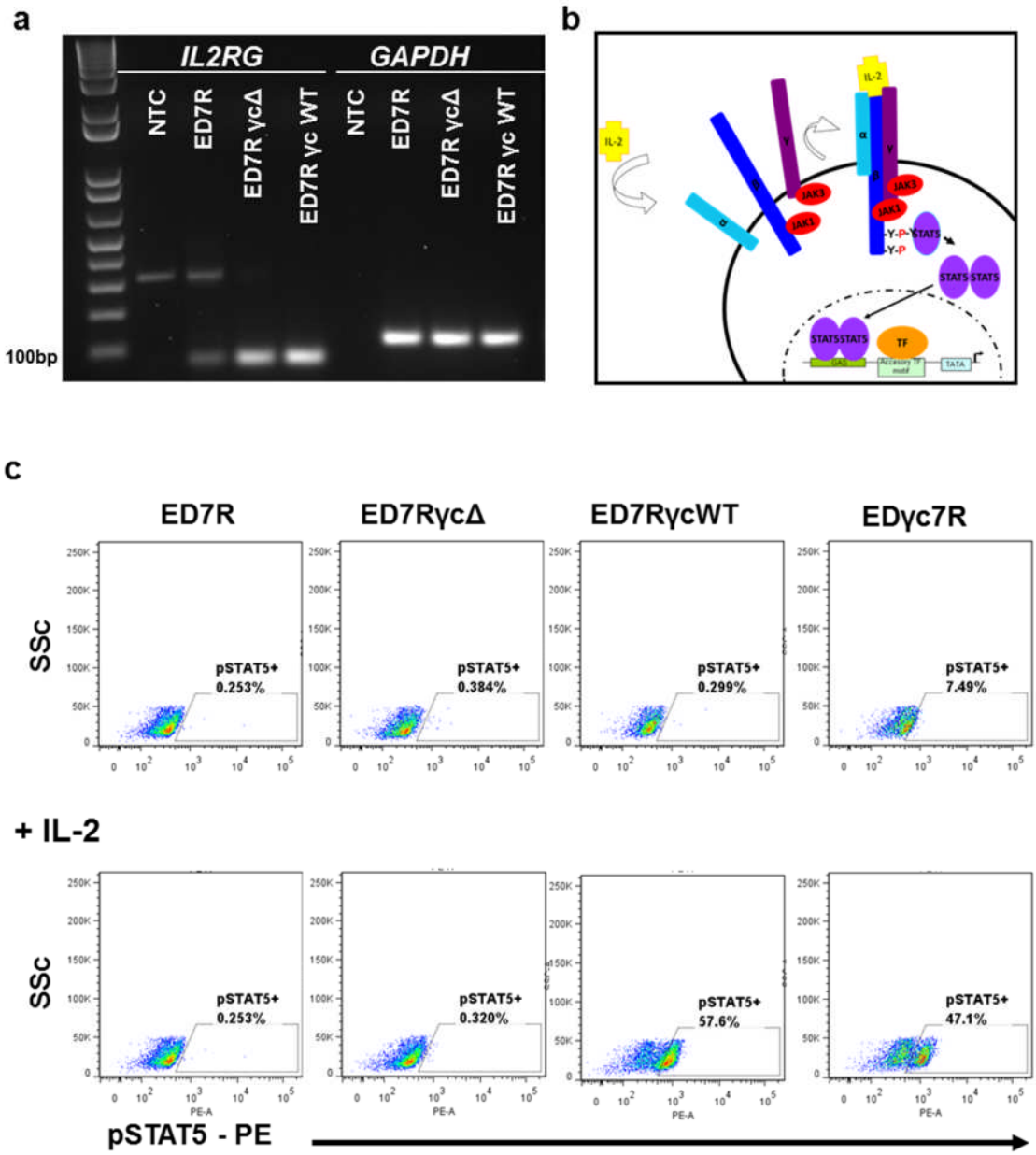


Figure 4.3 Further characterisation of *IL2RG* in ED7R cells.

(a) Reverse transcription PCR (RT-PCR) carried out on total RNA extract from ED7R, ED7RγcΔ and ED7RγcWT and amplified for *IL2RG* (78bp product) and *GAPDH* (107bp product) mRNA. PCR products were separated by gel electrophoresis and visualised. (b+c) The pSTAT5 assay of ED7R, ED7RγcΔ, ED7RγcWT and EDγc7R cells (b) a schematic of the downstream signalling events on IL-2 binding at the receptor resulting in the phosphorylated STAT5 (pSTAT5). (c) The results of IL-2 pulsing (bottom) or media alone (top) on starved cells as detected by flow cytometry analysis of pSTAT5.

4.3.4 qPCR of the transgene in ED7RγcΔ and ED7RγcWTclones

In order for the cell lines to be suitable for ZFN targeting analysis, a single copy of the transgene should be present in the genome. This will represent a more physiological target and reduce the toxic effect of multiple DSB.

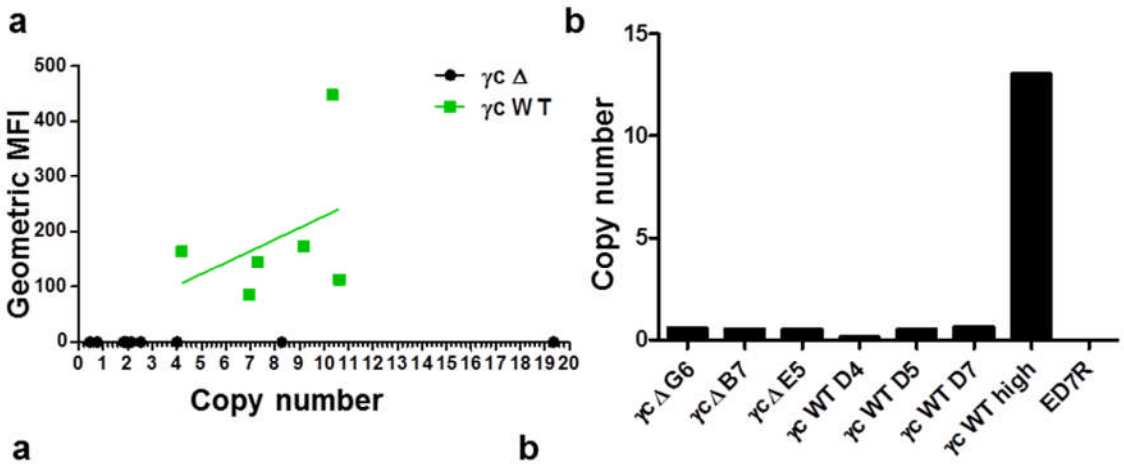
To determine the copy number of the transgene in the genomic DNA of each of the cell lines, quantitative PCR (qPCR) was performed using fluorogenic FAM-TAMRA probes to detect and quantify either the transgene's WPRE or the housekeeping gene, β-Actin. The quantitative values from the qPCR was compared to the MFI of γc expression on transduced cell clones as shown in Figure 4.4a where each point on the graph represents a clone. A positive linear association can be seen for WPRE copy number and the MFI of γc measured flow cytometry for the ED7RγcWT clones demonstrating that multiple *IL2RG* transgenes result in greater γc expression with the greatest MFI expression seen in the ED7RγcWT clone with a transgene number quantified at 10, now referred to as ED7RγcWT high. Regression analysis did not confirm this correlation to be significant and no such pattern exists for the ED7RγcΔ cells as there is no γc expression on these cells. From these data it can be seen that no ED7RγcWT clones have a copy number of 1 whereas a WPRE copy number of less than 1 is seen in 2 ED7RγcΔ clones.

In order to generate ED7RγcWT clones with a single copy number, the previous data were used to quantify the multiplicity of infection (MOI) of the *IL2RG* lentiviral vector stock. ED7R cells were transduced with an MOI titration of 0.4, 0.8 and 1.2 viral particles (vp) per cell and γc+ clones were isolated. The ED7RγcΔ clones with a WPRE copy number of <1 were also re-cloned by limited dilution to confirm there was no contaminating un-transduced ED7R cells which would bring the ratio of WPRE to β-Actin down.

A second round of qPCR was performed using TAMRA probes for WPRE and Albumin on the genomic DNA from ED7RγcWT generated by low MOI (Figure 4.4b). The data from this qPCR analysis valued the WPRE copies per cell for the re-cloned ED7RγcΔ and WT cell lines at less than 1. The ED7Rγchigh cells consistently had a high copy number and as expected the un-transduced ED7R control had no transgene. The Albumin housekeeping gene is normally located on chromosome 4 (Chr4) and the ED7R karyotype analysis revealed that Chr4 is absent in these cells. The unknown quantity of

this gene in ED7R cells may explain the low copy number in the freshly isolated ED7R clones.

To verify the transgene copy number using the WPRE qPCR data, a further calculation was possible with the karyotype analysis at hand, and calculated using the mass of the ED7R genome. By using a database of normal human chromosome length in base pairs the ED7R genome is valued at an estimated $6.71x10^9$ bp (Table 4.1a). With the assumption that a mole of nucleotides is 650g, the copy number of ED7R genome DNA used in the qPCR reaction can be calculated using the amount of template in nanograms and Avogadro's number, $6.022x10^{23}$ (Table 4.1b). From these calculations it can be seen that the copy number of transgene was close to 1 in 2/3rds of the clonal cell lines for both ED7RγcΔ and γcWT and that the ED7RγcWThigh remained consistent at 10 copies.



a

Chr.	Mass (bp)	No	Subtotal mass (bp)
1	2.E+08	2	498501242
2	2.E+08	2	486398746
3	2.E+08	1	198022430
4	2.E+08	0	0
5	2.E+08	4	723661040
6	2.E+08	2	342230134
7	2.E+08	2	318277326
8	1.E+08	4	585456088
9	1.E+08	2	282426862
10	1.E+08	2	271069494
11	1.E+08	1	135006516
12	1.E+08	3	401555685
13	1.E+08	2	230339756
14	1.E+08	2	214699080
15	1.E+08	2	205062784
16	9.E+07	1	90354753
17	8.E+07	4	324780840
18	8.E+07	1	78077248
19	6.E+07	3	177386949
20	6.E+07	1	63025520
21	5.E+07	4	192519580
22	5.E+07	1	51304566
X	2.E+08	2	310541120
Y	6.E+07	1	59373566
Other	**Est. mass**		
M1	9.E+07	2	180709506
M2	6.E+07	1	59128983
M3	9.E+07	2.5	225886883
		total	6.71E+09

b

	ng	WPRE	ED7R*	Copy No~
Δ G6	46.2	4155	7110	0.6
Δ B7	53	2626	8154	0.3
Δ E5	55.4	5630	8523	0.7
WT D4	60.2	1530	9262	0.2
WT D5	71.2	5730	10954	0.5
WT D7	51.6	4785	7938	0.6
WT high	0.5	748	77	9.7

Figure 4.4 Quantitative analysis of the transgene copy number in transduced ED7RγcWT and ED7RγcΔ clonal cell lines.

The transgene copy number is calculated as (WPRE value) / (Housekeeping gene value) x copies of housekeeping gene (a) Geometric MFI of γc expression quantified by flow cytometry compared to transgene copy number determined by qPCR of WPRE and βActin. Each point represents clonal cell line of ED7RγcWT (green square) or ED7RγcΔ (black circle). (b) qPCR analysis of WPRE in ED7RγcΔ or γcWT clones compared to the controls ED7RγcWThigh and un-transduced.

Table 4.1 Quantitative analysis of the transgene copy number in transduced ED7RγcWT and ED7RγcΔ clonal cell lines.

(a) Estimate of the total number of base pairs in a single ED7R genome based on karyotype analysis (b) Determining the transgene copy number based on the mass and amount of template ED7R DNA in ng and Avogadro's number. (*) = (ng template x 6.022×10^{23})/(6.71×10^{9} bp x 1×10^{9} x 650) and (~) = WPRE /ED7R*

4.3.5 Southern blot of the transgene in ED7RγcΔ and ED7RγcWTclones

Southern blot analysis of the clones was carried out to verify the copy number values generated by qPCR.

10µg of genomic DNA from the clonal cell lines was incubated with BamHI, EcoRI or XbaI alongside 100ng of *IL2RG* lentivirus plasmid vector for blotting and hybridisation control. These restriction enzymes are estimated to cut $6x10^5$ sites in the human genome which results in an average of fragment size of 5kb. The digests (Appendix 8) were blotted onto positively charged nitrocellulose membrane and hybridised with a radioactive probe. Probes were made to detect either the rev response element (RRE) or WPRE gene of the lentiviral backbone by digesting the *IL2RG* lentiviral construction plasmid with either restriction enzyme NotI or ClaI respectively, separating the digestion fragments by gel electrophoresis and isolating the correct size DNA of 1079bp (RRE) or 728bp (WPRE) with a gel extraction kit (Qiagen) (Figure 4.5a). The plasmid control hybridised at the expected fragment length of 13351bp, 14131bp, and 1093bp for BamHI, EcoRI and XbaI respectively confirming that both the DNA to membrane transfer and the radioactive probe were prepared successfully, this part of the membrane was removed after an exposure time of 30 minutes to allow the more extensive exposure time of 5 days for the genomic DNA. The RRE probe gave rise to bands in the region of 11kb but as these bands appeared for all of the clones including untransduced ED7R, this hybridisation is expected to be an artefact of background recognition. The WPRE probe did not achieve any quantitative bands for any of the DNA digests.

From the total data acquired to determine the transgene copy number in the clonal cell lines it cannot be confirmed that a copy number of one has been achieved.

By using the clonal cell lines generated in this section, ZFN mediated gene knockdown or correction can be assessed using the ED7RγcWT or ED7RγcΔ cell lines respectively and the ED7RγcWT and ED7RγcΔ clones which gave a copy number of <1 by qPCR are used in experiments where a cell line with a single transgene is required. Due to the functional gene in ED7RγcWT and the mutation designed for ZFN mediated correction in ED7RγcΔ cells gene targeting can be observed both at the genomic and the proteomic level of expression. With these tools, different methods of delivery can be assessed and compared.

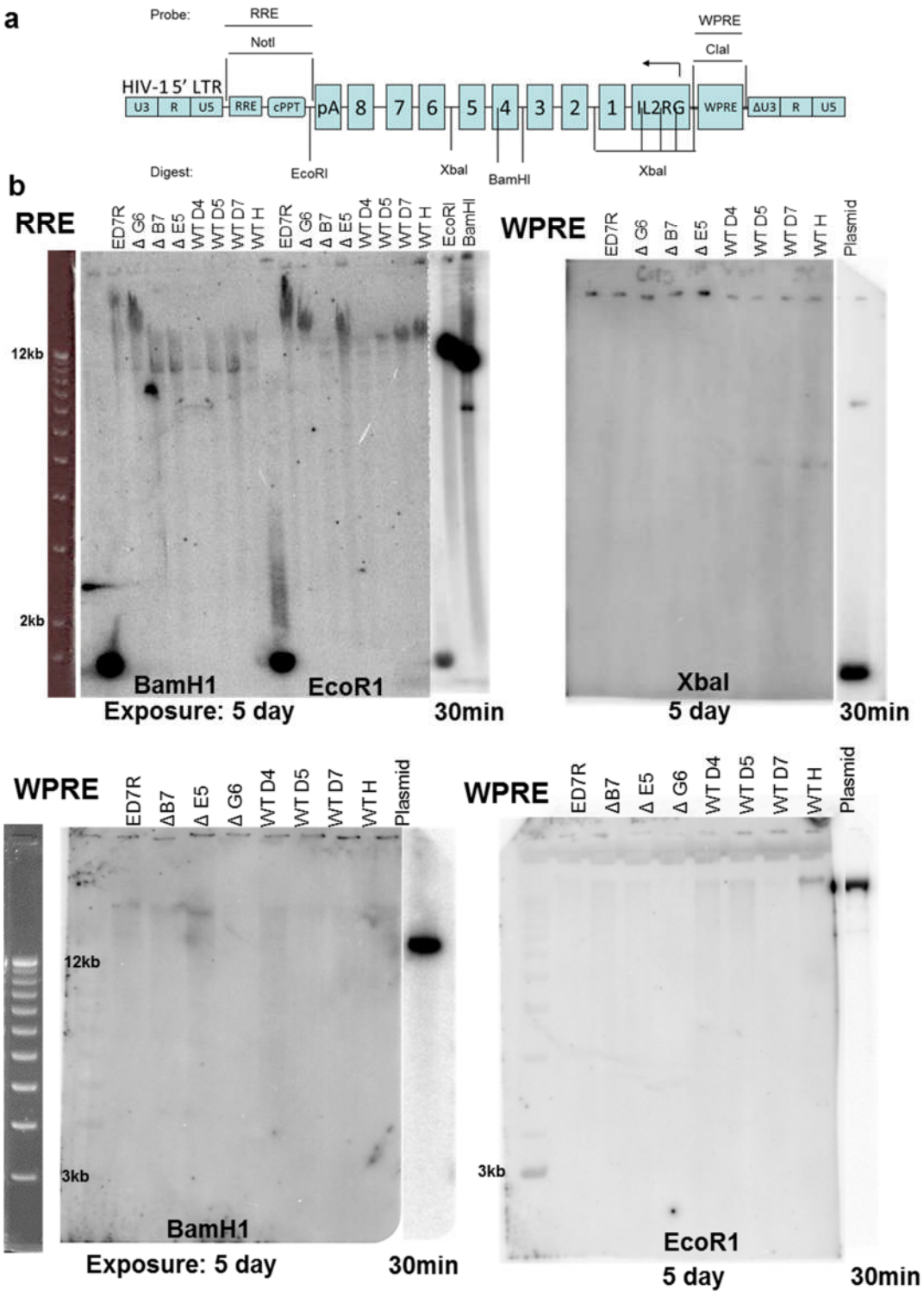


Figure 4.5 Southern blot analysis of the transgene copy number in ED7R cell lines.

Genomic and plasmid DNA digested with restriction enzyme for Southern blot and hybridised with radioactive probe. (a) The restriction enzyme and probe positions relative to the *IL2RG* transgene. (b) Southern blots of clonal cell lines digested by restriction enzyme and hybridised with radioactive probe as indicated.

4.4 Knockdown of γc by ZFN induced NHEJ

DBS which occur in protein coding sequence and are repaired by non-conservative NHEJ will frequently result in a non-functional gene and protein knockdown can be observed. In order to demonstrate *IL2RG* ZFN gene targeting, the ED7RγcWT clonal cell lines are used to ascertain the rate of knockdown by lentiviral delivery or nucleofection of ZFN.

4.4.1 Knockdown by Lentivirus

The Left and Right ZFN genes were individually cloned into lentiviral constructs which were subsequently used to produce either integrating proficient (LV) or integrating deficient (IDLV) lentiviral vectors, the MOI for each vector stock was valued by qPCR of the WPRE gene. It is important to note that constitutive ZFN expression achieved by LV ZFN vectors is only useful to confirm ZFN activity as such expression is likely to be toxic to the cell and gene targeting by transient ZFN expression provided by IDLV is far more relevant to therapeutic strategies.

In order to confirm that these lentiviral vectors were functional the effects of constitutive ZFN expression by the transgene integrated by LV was assessed. ED7RγcWTlow clonal cells with a copy number of 3 as quantified by qPCR were transduced in triplicate with a total MOI of 0.5, 1, 2 and 4vp per cell of the ZFN pair and for control the Left or Right ZFN at an MOI of 4. The γc expression was analysed by flow cytometry, 4 and 15 days post transduction (Figure 4.6a). It was seen 15 days post transduction that an MOI of 4 appeared to knockdown γc expression and deplete γc+ cells, however the statistical power of this decrease in expression cannot be distinguished when compared the cells treated with the L or R ZFN alone as there are only two values for this negative control. Evidence of gene targeting at the genomic level was assessed by analysis of indels detected by T7 endonuclease (Figure 4.6b). Using the same primers that generate the target template from genomic DNA, an indel positive control was generated by amplifying the 500bp target from a WT donor and the MB17 donor and combined at a ratio of 1:1. On denaturing and reannealing, the 17bp dissimilarity from WT in the MB17 will be recognised by mismatch sensitive endonuclease and digested, resulting in 250bp fragments. This 500bp target template is hence forth known as the indel positive control. Unlike the flow cytometry analysis, indels were detected 15 days post transduction in cells treated with an MOI of 4 or 2.

There was also evidence of indel incorporation at an MOI of 0.5 but as no indels are detected at an MOI of 1 this may be a false positive due to a DNA shearing.

The ZFN targeting which was observed by genomic analysis but not by change in γc expression at the cell surface could be explained by the multiple copies of *IL2RG* in the cell line used where knockdown of a single copy would not be reflected by γc expression at the cell surface. To overcome this, cell lines with a single copy of the transgene are used in subsequent experiments and denoted as ED7RγcWT.

To achieve both viral transduction and transient expression, ZFN expression constructs were delivered by IDLV to target *IL2RG* in ED7RγcWT cells. Transductions were carried out with a titration of the IDLV ZFN pair at 0, 100, 200 and 300vp per cell in total and ZFN left only at an MOI of 300 for control. Flow cytometry analysis of γc expression 14 days post transduction revealed a slight reduction in γc expression when cells were transduced with the ZFN pair at an MOI of 300 (Figure 4.6c) however this was not confirmed by the Cel-1 surveyor assay as no indels were detected in any of the transduced samples (Figure 4.6d).

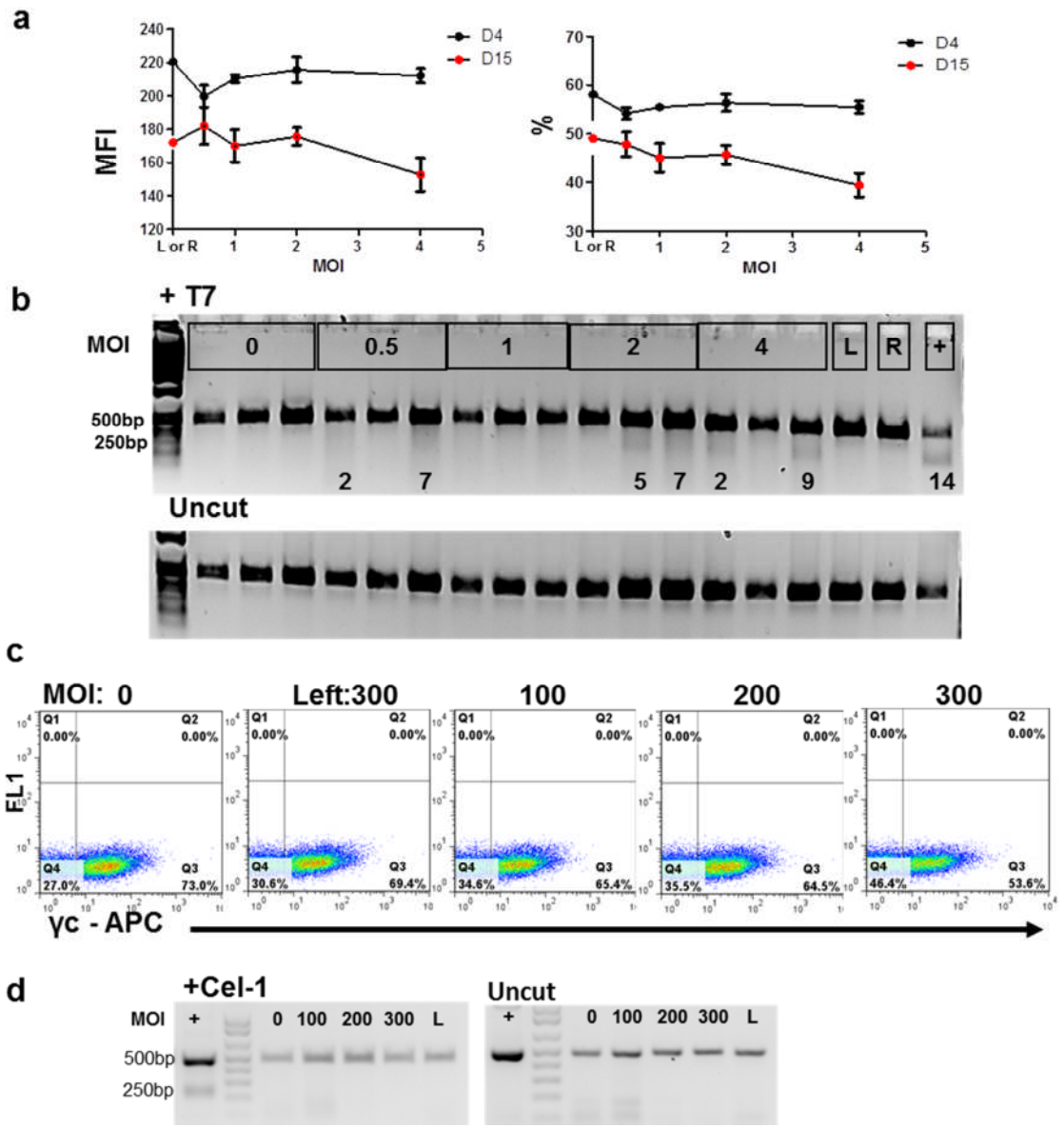


Figure 4.6 *IL2RG* targeting in ED7RγcWT cell lines by ZFN delivered by lentiviral vectors.

Knockdown of γc expression by ZFN expressing LV or IDLV was by analysed flow cytometry or indel incorporation detected by T7 endonuclease or Cel-1 alongside an uncut control. A 500bp indel positive control (+) which will be digested to 250bp fragments on exposure to the endonucleases was also included. (a+b) $5x10^4$ ED7RγcWTlow cells were transduced with LV at increasing MOIs of the ZFN pair or ZFN Left (L) or Right (R) alone at an MOI of 4. (a) Expression analysed by flow cytometry at 4 and 15 days post transduction. Graphs show the geometric mean of γc expression (MFI) and the percentage of γc+ cells (%). (b) Indel incorporation in the treated cells, detected by the T7 endonuclease 15 days post transduction. The numbers in black denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ. (c+d) $5x10^4$ ED7RγcWT cells transduced with IDLV at increasing MOI of the ZFN pair or ZFN left only and analysed after 14 days. (c) Flow cytometry analysis of % γc expression, 14 days post transduction (d) Indel incorporation in the treated cells detected by the Cel-1 surveyor assay 14 days post transduction.

4.4.2 Knockdown by nucleofection of ZFN

To avoid the use of viral vectors altogether, transfection with plasmid DNA is a popular method to introduce novel genetic material for transient expression. Nucleofection has been developed to transfect virtually all cell types with various types of nucleic acid including endotoxin free plasmid vector.

To assess the knockdown of γc expression by the *IL2RG* ZFN pair using this method, $2x10^6$ ED7RγcWT cells were nucleofected with a titration of Left and Right ZFN expression plasmid together with a GFP control (pMax GFP) in order to consistently use 2μg total DNA in each sample. The combined ZFN plasmid DNA increased from 0μg with 0.5μg intervals to 2μg and samples were run in triplicate. Flow cytometry analysis of γc expression was carried out 96hours post nucleofection (Figure 4.7a) and knockdown of expression was clearly demonstrated, from 60% γc+ cells in the GFP only control to 1.4% in cells treated with 2μg of total ZFN expression plasmid. Further analysis of these triplicate data (Figure 4.7b) revealed that nucleofection $2x10^6$ cells with 1.5μg of ZFN statistically reduces the percentage of γc+ cells however the reduction in the total level of γc expression as measured by MFI did not reach significance. Propidium Iodide (PI) uptake of dead cells was used to analyse cell viability and a linear negative association observed between ZFN treatment and cell viability where ≥1.5μg ZFN expression plasmid caused a significant drop in live cells. Interestingly, no indels were detected by the Cel-1 surveyor assay in the DNA pooled from each triplicate (Figure 4.7c).

The consistently contradictory evidence of *IL2RG* targeting from the flow cytometry analysis and indel detection prompted investigation into the sensitivity of the Cel-1 assay. In addition this enzyme was compared to the T7 nuclease which has also been used for the same purpose in this study. As for the indel positive control, template was prepared from plasmid containing the WT and MB17 donor plasmids and a total of 500ng per reaction was used with a decreasing percentage of template prepared from MB17 thereby decreasing the potential template containing mismatches (Figure 4.7d). The detection of indels using Cel-1 in this instance was more accurate than the T7 however the detection level of both the Cel-1 and T7 endonuclease could not be seen above background when mismatch potential dropped to 2% and the drop off point could be anywhere between 8% and 2%. This comparative analysis suggests that the

Cel-1 surveyor assay is more sensitive than using the T7 endonuclease and from this assessment it can be speculated that gene targeting events must reach a rate of ≥8% to be detected with this assay.

Flow cytometry analysis of the γc knockdown experiments in ED7RγcWT cells imply there is ZFN activity resulting in DSB. However the absence of indels in these samples do not confirm that the DSB are repaired by NHEJ. This could be due to the sensitivity of the indel surveyor assay or due to a deregulated DNA repair mechanism in ED7R cells.

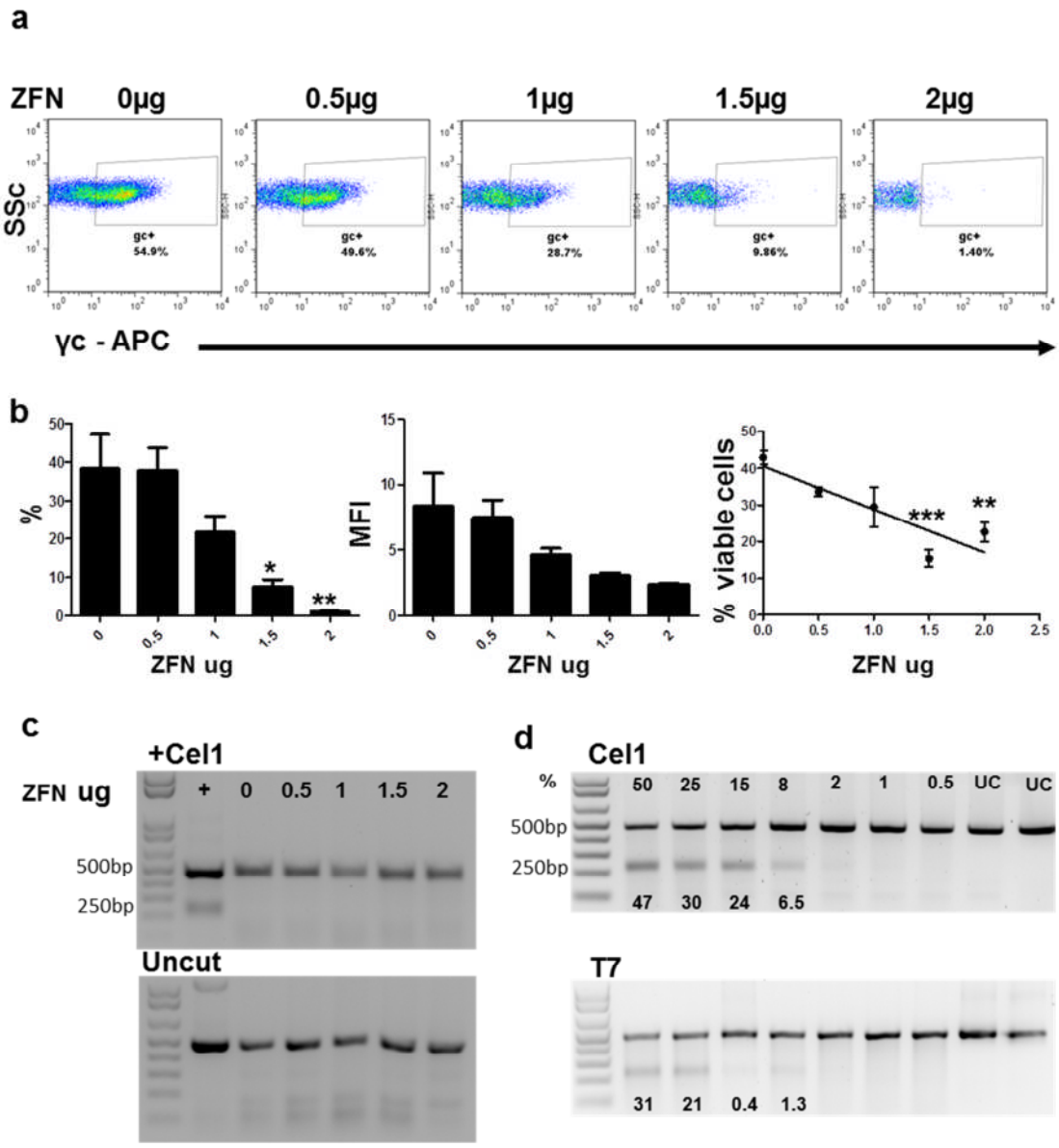


Figure 4.7 *IL2RG* targeting in ED7RγcWT cell lines by ZFN delivered by nucleofection.

(a-c) ED7RγcWT cells nucleofected with increasing amounts of ZFN expression plasmid. (a) Flow cytometry analysis of γc expression detected on PI-ve cells. (b) The triplicate values of γc+ cells, the MFI of γc and the percentage of PI-ve (viable) cells as determined by flow cytometry, plotted against the ZFN titration and the significance calculated by a one way analysis of variance (*= P-value<0.05) (**=P-value<0.01)(***=P<0.001). (c) Indels detected by the Cel-1 surveyor assay in template prepared from nucleofected cells alongside uncut template control and an indel positive control. (d) A decreasing percentage of DNA capable of forming mismatched sequence (%) were digested by Cel-1 and T7 endonuclease and uncut template (UC) was included for control. The numbers in black below the bands denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

4.5 Gene targeted correction of *IL2RG* by ZFN induced HR

The ED7RγcΔ cell lines were made to assess the possibility that the DSB induced by ZFN can be repaired by HR from an endogenous donor fragment and in so doing, correct the SCID-X1 associated G691A mutation. Again lentiviral vector delivery and nucleofection of ZFN expression plasmids was performed but this time in unison with donor DNA delivery. In addition to establishing an efficient method of delivery, the most effective ratio of ZFN to donor DNA constructs was also considered in this section.

When the MB17 donor is used for HR the MB17 sequence is incorporated and can be distinguished from WT by an HR specific PCR as established in chapter 3. In addition, the novel restriction enzyme site, BsrBI is introduced and the endogenous DraIII site is lost. This exchange can be detected by RFLP analysis of 1,901bp template DNA amplified from the treated cells using primers that anneal outside of the region of donor homology and are therefore specific for unaltered genomic DNA (Figure 4.8a). The digestion products can then be separated by gel electrophoresis and the band intensity of cut and uncut bands can be determined (Figure 4.8bi). Subsequent blotting of the DNA onto a membrane and hybridisation to radioactive probe can increase the sensitivity of RFLP analysis as seen by the Southern blot analysis of a titration of BsrBI containing fragments in total template DNA where the percentage of digested product more accurately represented the percentage of BsrBI sites in the starting template (Figure 4.8bii). As the HR specific PCR of the MB17 donor cassette is not quantitative, this RFLP Sothern blot can be used to quantify the extent of HR in cells treated with ZFN and Donor DNA.

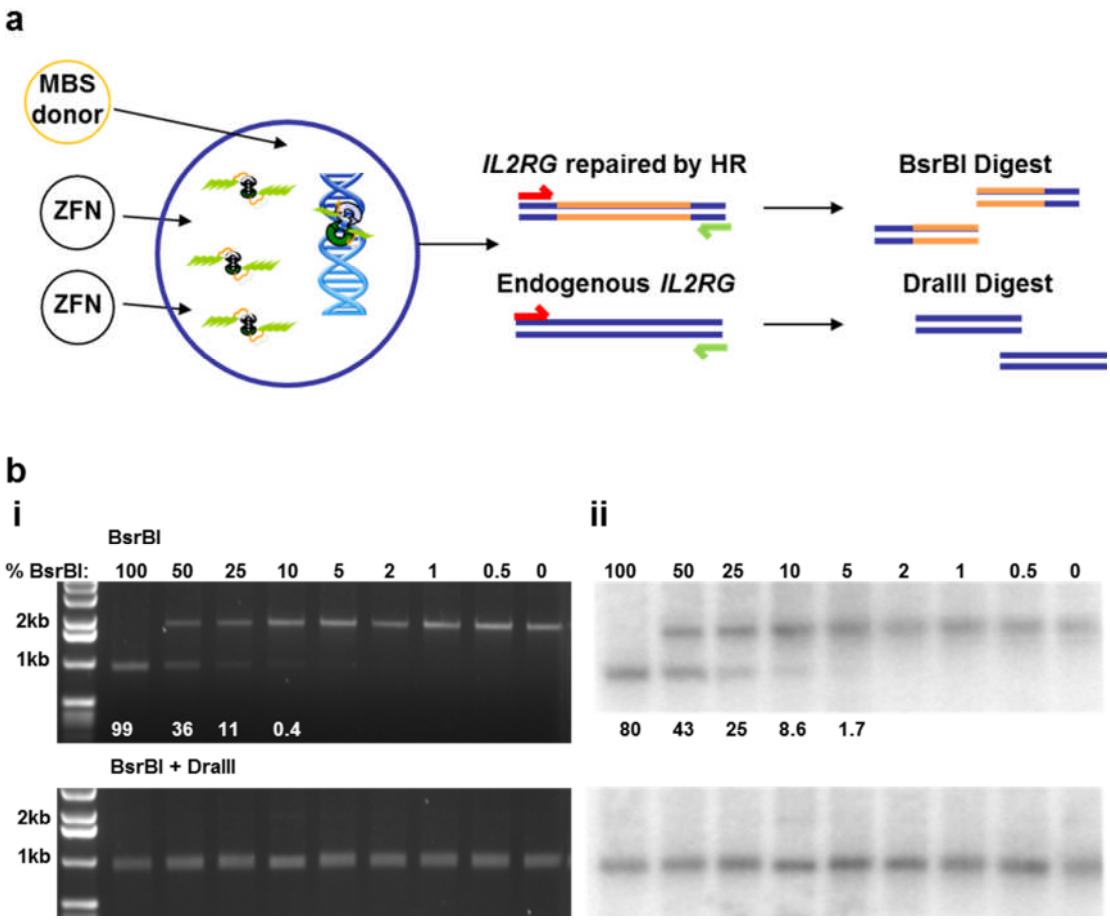


Figure 4.8 RFLP analysis of cells treated with ZFN expression vectors and Donor repair DNA.

(a) A schematic of the RFLP analysis: genomic DNA is extracted from treated cells and the target loci are amplified by primers (red and green) that anneal outside the region of donor homology (Orange). The 2kb amplicons are purified and digested with BsrBI or DraIII to achieve 1kb digestion products. (b) RFLP analysis of template DNA containing a decreasing percentage of BsrBI (% BsrBI) containing fragments made up to 100ng with WT template, digested with BsrBI or BsrBI plus DraIII. (i) Gel electrophoresis of the cut and uncut fragments are visualised by UV. (ii) The separated fragments are blotted onto a negatively charged membrane, hybridised with radioactive probe and visualised by autoradiography. The numbers below the bands denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

4.5.1 Gene targeted correction of *IL2RG* by lentiviral vector delivery of ZFN and donor DNA

To accompany the ZFN delivered by lentiviral vectors, the MB17 donor DNA was cloned into a lentiviral backbone and IDLV was prepared. Donor DNA is always delivered by integrase deficient vectors as randomly integrated donor constructs are not readily available for HR.

To confirm the proof of principle of ZFN targeted correction and to establish an effective ratio with donor DNA, $5x10^4$ ED7RγcΔ cells were transduced with integrase proficient ZFN LV vectors and IDLV donor DNA at different ratios (Figure 4.9a). The total MOI of ZFN LV pair was increased starting at an MOI of 5 in accordance with the knockdown experiments where the most activity was observed at an MOI of 4. HR specific PCR reveals that a ZFN:Donor DNA ratio of 3:1 consistently resulted in HR positive amplicons. However, the smear from the PCR product, amplified by primer pair 1, from cells treated with a 3:1 ratio and ZFN LV at an MOI of 10 (Figure 4.9a lower gel) may or may not contain an HR positive product. HR occurred in 2/3rds of samples treated with a ratio of 1:3 and the least effective ratio is 1:1 which only achieved HR one in four times. In this assay the HR was frequently detected by primer pair 1 and infrequently by primer pair 2 and the positive PCR control detected by both. It is also important to note that the donor DNA integrated at the target locus without ZFN activity when an MOI of 150 was used.

Subsequent analysis of ED7RγcΔ cells transduced with lentiviral vectors was performed with a titration of IDLV donor construct without ZFN vectors, and a ratio of 3:1, ZFN:Donor DNA using either LV or IDLV ZFN vectors (Figure 4.9b). The HR specific PCR performed on these transduced cells revealed that treatment with IDLV donor DNA with MOI ranging from 1 to 100 viral particles per cell did not result in HR as no HR specific band was detected at 1kb for primer pair 1 or 970bp for primer pair 2. When donor DNA was combined with ZFN expression vector, targeted integration was only observed when an MOI of 90 ZFN LV was combined with Donor DNA at an MOI of 30. No HR was observed when ZFN IDLV constructs were used.

Cells treated with an MOI of 30:90 ZFN LV:Donor IDLV and the corresponding control samples treated with either an MOI of 30 LV ZFN or 90 IDLV donor DNA and untransduced control were analysed by RFLP (Figure 4.9c). Digestion product analysis

revealed that 17% of target IL2RG had undergone HR and gene exchange was observed in cells treated with donor alone. The positive control, prepared from full length *IL2RG* containing the MB17 cassette was quantified at 150% digestion, this artefact may be due to band saturation of the cut product or the background exceeding the remaining uncut product. Double digest with both BsrBI and DraIII is incomplete so band intensity was not calculated to determine whether there was a loss of the DraIII site by NHEJ in ZFN treated samples.

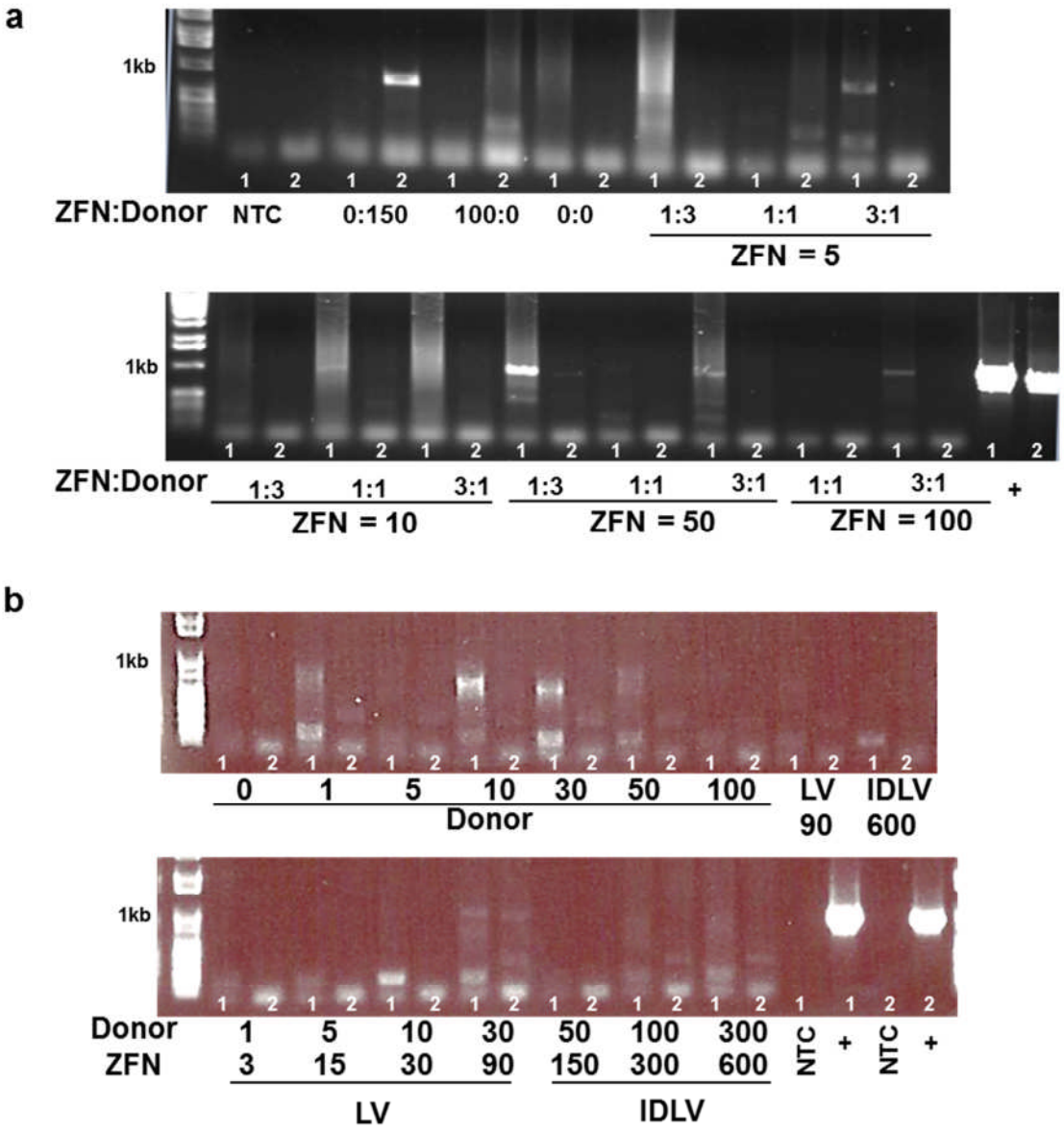


Figure 4.9 Targeted HR in ED7RγcΔ cells by ZFN and donor DNA delivered by lentiviral vectors.

$5x10^4$ ED7RγcΔ cells treated with ZFN delivered by increasing MOI of LV or IDLV ZFN construct with or without IDLV donor at different ratios as indicated. (a+b) HR specific PCR performed on genomic DNA extracted from cells 15 days post transduction alongside a plasmid containing full length *IL2RG* with the MB17 cassette for control (+). (c) RFLP analysis by Southern blot. BsrBI or BsrBI plus DraIII digest of template from cells treated with either ZFN LV (Z) or Donor IDLV DNA (D) at an MOI of 90 or 30 respectively, or both (Z+D) alongside untransduced control (C) and the BsrBI positive control (+). The numbers below the bands denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

4.5.2 Gene targeted correction of *IL2RG* by nucleofection with ZFN and donor DNA plasmid vectors

Because the transient expression of ZFN by IDLV delivery did not result in any detectable target HR, nucleofection was used to assess different ZFN to Donor DNA ratios. Furthermore, different donor DNA constructs containing the WT, MB1, MB3 and MB17 cassette in the *IL2RG* homology domains were compared to establish whether increasing the sequence disparity between the target and donor affects HR and if WT homologous donor containing intact ZFN target hindered HR.

$2x10^6$ ED7RγcΔ cells were nucleofected with 1:3, 1:1 or 3:1 ZFN:Donor DNA constructs to a total of 2µg per sample. A titration of 0.5, 1 and 1.5µg of ZFN or donor construct were also nucleofected for control along with nucleofection solution alone. Flow cytometry analysis 96 hours after nucleofection did not reveal correction resulting in γc expression (Appendix 9) and HR specific PCR was carried out to determine whether any genomic modification had taken place. A set of primers specific for the MB3 cassette with the same binding position as the MB17 primers were used to detect HR in MB3 donor treated cells (Figure 4.10a). HR was detected in cells treated with the MB3 donor at all ratios tested and in the donor control with both primer pairs 1 and 2. The absence of an HR specific product in the nucleofection solution alone and ZFN treated cells confirmed that the primers can distinguish from WT. HR specific bands were also detected in all the cells nucleofected with the MB17 donor. However no BsrBI digestion was detected by RFLP analysis in cells treated with donor capable of introducing the BsrBI restriction site (MB1, MB3 and MB17). Therefore it cannot be determined whether cells treated with ZFN and donor DNA constructs had a greater rate of HR than cells treated with donor alone (Figure 4.10b).

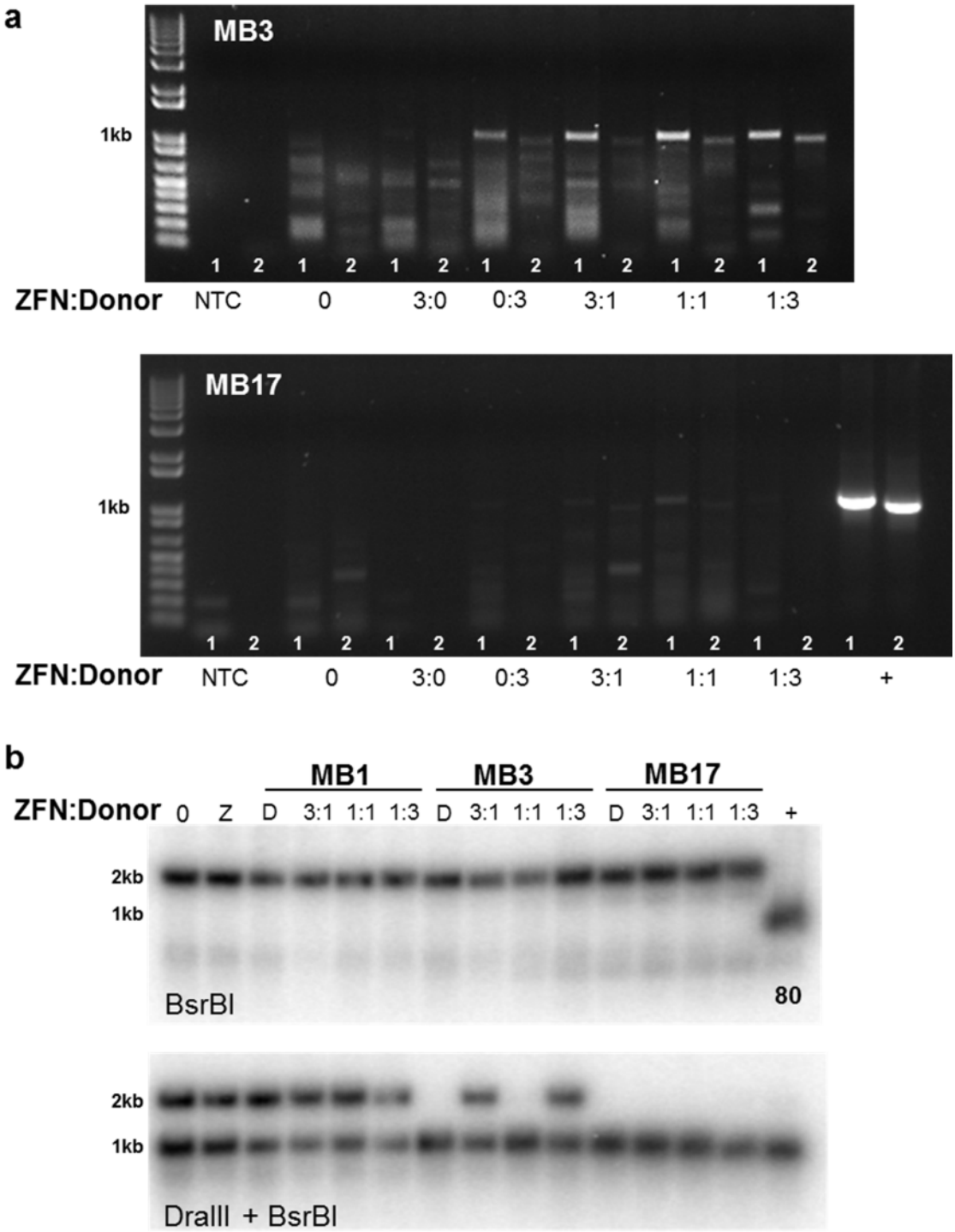


Figure 4.10 Targeted HR in ED7RγcΔ cells treated with ZFN expression vectors and donor DNA delivered by nucleofection.

Genomic DNA analysis of targeted HR, 96 hours after nucleofection of $2x10^6$ ED7RγcΔ with 2µg of plasmid at different ZFN:Donor DNA or 1.5µg of ZFN or donor DNA constructs. (a) HR specific PCR to detect the MB3 or MB17 cassette at the *IL2RG* locus. (b) RFLP analysis of the cells nucleofected with different ratios of ZFN to donor DNA capable of introducing a BsrBI restriction site at the IL2RG locus alongside BsrBI positive control (+). The positive values of DNA digested by BsrBI above background are noted as a percentage below the digested bands. The number below the + band denotes the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

The nucleofection of ED7RγcΔ cells was repeated in order to monitor whether γc expression could be detected after a longer incubation period. The amount and type of ZFN and Donor used remained the same however the Lonza GFP control was used to bring the total DNA to 2µg for control samples. Flow cytometry analysis at day 12 showed γc expression in 2.4% of cells receiving the WT donor at a ZFN to Donor DNA ratio of 3:1. Treatment with the MB1 and MB3 donor constructs at the same ZFN:Donor ratio also achieved γc expression where no γc expression was detected in the other ratios or in the donor controls (Figure 4.11a). These cells were kept in culture for a further 20 days and subsequent flow cytometry analysis revealed a >15 fold increase of expression (Figure 4.11b). In addition to successful gene targeted correction stimulated by ZFN, these results also suggest that γc expression conveys a selective advantage *in vitro*.

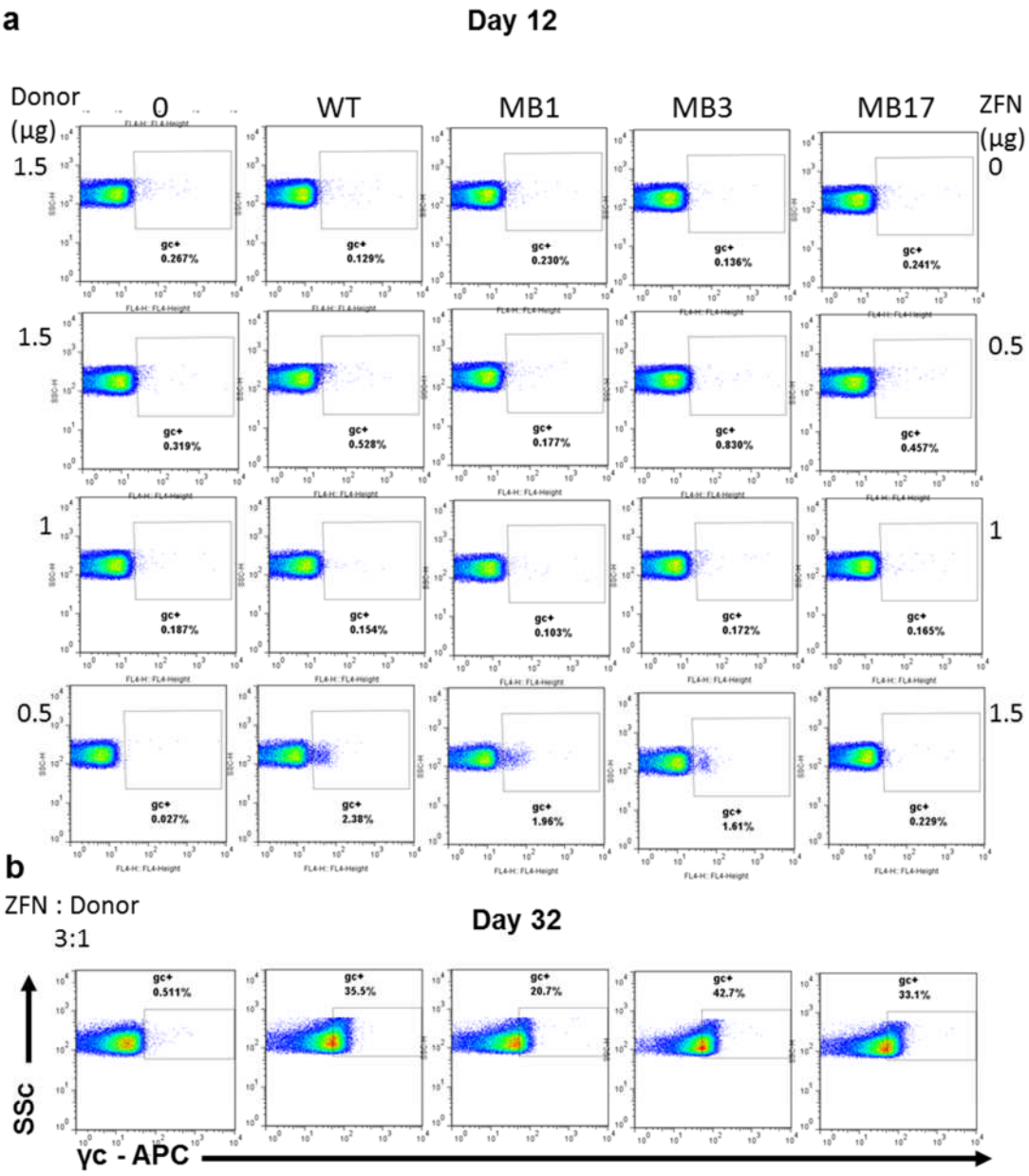


Figure 4.11 Cell surface expression of γc in ZFN and donor DNA treated ED7RγcΔ cells.

$2x10^6$ ED7RγcΔ nucleofected with 2µg of plasmid at different ZFN:Donor DNA. Or 1.5µg of ZFN or donor DNA constructs plus 0.5µg of GFP expression plasmid. 2µg GFP expression plasmid was used where no ZFN or donor DNA plasmid was used (0). Cells were stained for γc expression, acquired by flow cytometry and gated on PI-ve live cells. (a) Flow cytometry analysis 12 days post nucleofection. (b) Flow cytometry analysis of ED7RγcΔ receiving a ZFN:Donor ratio of 3:1 kept in culture for 32 days compared to untreated ED7RγcΔ control cells.

4.5.3 IL2 Receptor signalling in corrected cells

In order to determine whether the γc expression in the ED7RγcΔ cells treated with a ZFN:Donor ratio of 3:1 detected by flow cytometry restored IL2R signalling. The γc positive or negative cells were sorted by Fluorescence activated cell sorting (FACS) and pSTAT5 was monitored following IL-2 stimulation (Figure 4.12). Medium alone did not stimulate STAT5 phosphorylation in any of the sorted cells or the ED7RγcWT control (Appendix 10). The pronounced shift of pSTAT5 positive cells seen in the ED7RγcWT cells was not seen in any of the IL-2 stimulated cells. However a slight pSTAT5+ shift was seen in the γc+ cells sorted after treatment with ZFN plus the WT, MB3 and MB17 donor compared to the γc- sorted cells. Conversely, this pattern was reversed in the cells treated with the MB1 donor.

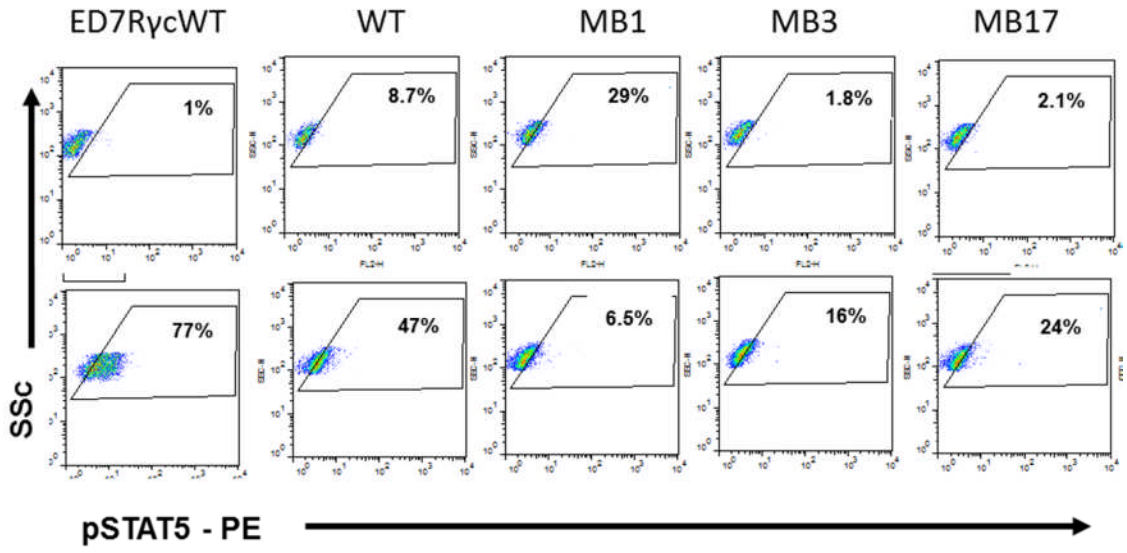


Figure 4.12 pSTAT-5 Signalling in treated ED7RγcΔ cells.

ED7RγcΔ cells treated by ZFN and donor ratio of 3:1 were cell sorted by FACS according to γc expression. Top, γc- . bottom γc+. After ON cell starvation of FCS, the cells were stimulated with IL2 and were analysed for pSTAT5 by flow cytometry.

4.6 Discussion

In this chapter 2 cell lines, ED7RγcWT and ED7RγcΔ have been constructed, characterised and successfully used for ZFN mediated γc knockdown and correction respectively.

The data show that the ED7R cell line is capable of *IL2RG* transcription which follows from the intact *IL2RG* locus. However, the lack of intracellular and cell surface γc expression suggests that there is a problem between translation and post-translational processing of γc. The disproportionate, low intensity of the ED7R mRNA PCR product compared to the transduced controls could be speculated as (i) RNA instability, possibly due to the loss of the T from the polyT tract in intron 4. (ii) The greater mRNA product in transduced controls could be due to the transgene located in more active chromatin leading to higher levels of expression. (iii) unequal concentrations of mRNA in the sample. In order to resolve the latter 2 issues, a repeat RT-PCR of mRNA expression should be compared between freshly isolated mRNA from the ED7R cells and primary cells with endogenous levels of *IL2RG* and quantified with corresponding RT-PCR amplicons generated from the housekeeping gene, *GAPDH* mRNA.

A point of consideration for the lack of γc expression in the ED7R cell line is the possible interference of IL-2 signalling by the original HTLV-1 infection. ED7R cells have already been shown to be HTLV-1 protein *tax* negative (Hironaka, Mochida et al. 2004) and the HTLV-1 protein $p12^{I}$, capable of binding and stabilising γc in the pre-Golgi compartment (Mulloy, Crownley et al. 1996), is not suspected as ED7RγcWT cells are capable of cell surface γc expression. It would, however, be useful to investigate whether the HTLV-1 genome is located in proximity to the *IL2RG* locus and whether cis-acting elements interfere with *IL2RG* expression.

The mRNA transcribed by ED7RγcΔ cells is a reflection of the nature of the G691A mutation which does not result in a stop codon or a splice site but the conversion of arginine at codon position 226 to histidine (R226H) and mRNA was detected in a SCID-X1 patient with this mutation by Northen blot (Puck, Pepper et al. 1997). This study also observed a trace of cell surface expression of the mutant γc on the patient's B cells however, this was not seen in a subsequent study by a different group on the same mutation using the same monoclonal antibody, TugH4 (Kumaki, Ishii et al. 2000). This residue is located in domain-2 of the extracellular region of γc (Stauber, Debler et

al. 2006) and is involved in binding IL-2. The location of this mutation could have consequences in the recognition by antibodies raised to recognise WT γc, however over the course of this chapter, 2 different antibodies raised to recognise two different epitopes of γc were used. Both the PE conjugated AG184 clone and the APC conjugated TuGH4 clone successfully discriminated between the γc-ve ED7R cells and the γc+ve ED7RγcWT and neither detected γc on the ED7RγcΔ cells, Furthermore, the AG184 clone recognises an epitope distinct from the IL-2 binding site of γc suggesting that this antibody would recognise the G691A mutation if the protein was expressed on the cell surface. Justification for the lack of protein expression includes (i) the protein is missfolded and promptly degraded, (ii) the protein is unable to traverse the cell membrane and is degraded as a consequence. The latter is unlikely as no intracellular protein was detected.

In trying to establish a cell line with a single copy of the WT or G691A *IL2RG* transgene, 3 qPCR analysis strategies gave consistent copy numbers which in turn justified the copy number values used in subsequent experiments. The lack of supporting evidence for these values by Southern blot analysis can neither confirm nor deny the qPCR quantification. Limitations of the Southern blot analysis included the presence of large >12kb fragments that were not separated by electrophoresis, this could be improved by using a more extensive array of restriction enzymes used individually or in combination to generate shorter fragments. It would also be of benefit to include a probe that annealed to a known digested fragment length in the genomic DNA as this would confirm successful digestion and DNA of suitable quality.

A limitation of introducing the ZFN target as a randomly integrated transgene means it is unlikely that the chromatin infrastructure in and around the target site accurately represent that of physiological conditions. Another drawback is due to the ED7R cell line, despite the fact they were originally isolated from a human male, the karyotype of the ED7R cell line contains two X chromosomes and therefore 2 copies of non-functional *IL2RG*. These precursory, non-functioning copies of the *IL2RG* could “mop up” the ZFN provided and the multiple DSB incurred could increase the toxicity experienced by the cell. A solution to this latter problem would be to treat ED7R cells with ZFN before introducing the transgene. By sequencing NHEJ events at the *IL2RG* loci in ZFN treated cells, those which had both alleles disrupted by non-conservative

repair would no longer have ZFN targets. This would ensure that any *IL2RG* target sequence added to the cells would be the only target for ZFN and therefore decrease toxicity and possibly increase the sensitivity of gene modification detection.

Gene targeting by flow cytometry analysis showed that γc gene knockdown was achieved by DSB mediated by ZFN however these data was not consistently supported by the detection of indels which would evaluate DNA repair by NHEJ. As the level of indel detection was determined to be sensitive enough for a dissimilarity of 8% of DNA molecules, it is unlikely that the surveyor assay is not sensitive enough to detect *IL2RG* modification where ZFN treatment resulted in a 59% drop of γc+ cells. It is well documented that over-expression of ZFN leads to toxicity due to off target cleavage (Beumer, Bhattacharyya et al. 2006; Gupta, Meng et al. 2011) and ED7R cells may be sensitive to the extent that the treated cells had undergone apoptosis and DNA degradation before indel analysis was carried out. Contrary to this is the evidence of indels in ED7RγcWTlow cells, containing an estimated 5 targets, undergoing constitutive ZFN expression. However the multiple copies may have resulted in enough gene targets to be detected in surviving cells.

Genomic analysis detected targeted HR in the ED7RγcΔ cells 96 hours after ZFN and donor treatment, this is in accordance with the expected ZFN translation and activity observed by Porteus *et al.* (Porteus and Baltimore 2003). The γc expression observed by flow cytometry after a longer incubation time of 12 days is corroborated by the 15 day incubations used by Lombardo *et al*. (Lombardo, Genovese et al. 2007). RFLP analysis only detected targeted HR 15 days after transduction with ZFN LV and donor IDLV. The 15 day time point was not analysed after nucleofection as γc expression was confirmed by flow cytometry but this analysis would be useful to corroborate the flow cytometry data. The absence of targeted HR with the use of ZFN and donor DNA delivered by IDLV is contradictory to the successful use of IDLV used by Lombardo *et al*. It is not reasonable to compare the amount of IDLV vector used in either study as the IDLV quantified by Lombardo *et al.* was in µg of p24/mL. This method of quantification is not in accordance with MOI calculations for vector stocks and does not exclude empty vectors which may overestimate the quantity of complete viral vector.

On transient expression of ZFN, HR was only successful by nucleofection where a 3:1 ZFN:Donor ratio was performed using 2μg of plasmid per $2x10^6$ cells. Using Avogadro's number, this equates to $1.6x10^5$ plasmids per cell. This is 180 fold greater than the highest MOI of IDLV transfected and therefore a direct comparison cannot be made with nucleofection. What is of interest is that by increasing the MOI of donor IDLV alone led to HR at the *IL2RG* target when 150vp/cell was used but not ≤100vp/cell. It is of no surprise that ZFN independent HR occurred when nucleofected with $1.5x10^5$ donor plasmids/cell.

Background HR of donor DNA at the target site in the absence of ZFN was frequently observed in HR specific PCR but not in flow cytometry or RFLP analysis. As the HR specific PCR cannot quantify the rate of HR but may be more sensitive than RFLP and flow cytometry, it can only be assumed from the flow cytometry analysis which only showed γc expression in cells that received both the ZFN pair and donor DNA, that the rate of recombination is increased when ZFN are administered to activate the chromosomal target.

The comparison of different donor DNA constructs with increasing sequence dissimilarity shed light on various aspects of correction in this model. Firstly, the PCR method of detection is only possible in the donors with 3 or 17 base pair changes from WT. Both of these donors were used for HR when DNA was delivered by nucleofection. Using the quantifiable method of flow cytometry, after 12 days of culture post nucleofection, it would seem that by increasing the base pair dissimilarity from the mutant target, the rate of HR declines, from 2.38% with WT donor (1bp difference) to none in MB17 (18bp difference). However, this pattern is not seen after a month of culture where corrected γc was observed following treatment by all of the donors. Firstly this demonstrates that the PCR method of detection is more sensitive. Secondly, the hypothesis that a donor containing the ZFN target site would inhibit HR is not shown in this model. Thirdly, γc expression likely provides a growth advantage in ED7R cells *in vitro*, as the percentage of corrected cells increased over time.

The γc corrected cells did not convincingly demonstrate IL2R signalling by IL2 stimulation. This may in part be due to the strenuous effects of cell sorting. It will be necessary to repeat this assay to incorporate an extended recovery period for the

sorted cells and to ensure that the inconsistent result observed after treatment with the donor construct MB1, was not a result of sample exchange. It will also be optimal to carry out RFLP and high throughput sequencing on the γc positive and negative sorted cells to ensure that the γc expression was the result of targeted *IL2RG* correction.

5 Characterising the humanised mouse model of SCID-X1

5.1 Aims

- To characterise the nature of human *IL2RG* in a new humanised SCID-X1 mouse model
- To confirm the immune cell profile of the humanised mouse is deficient by comparing it to WT and γc knockout mice
- To reconstitute the immune system of the humanised mouse with *IL2RG* expressing integrating lentivirus

5.2 Introduction

Studying ZFN in murine models is a prerequisite of human clinical trials. Variations of γc knockout mice have so far been used as murine models of SCID-X1 and although gene transfer by retroviral vectors is successful in restoring the immune system in these models, due to the lack of endogenous human *IL2RG* (h*IL2RG*), there is no platform on which to perform *IL2RG* targeting with ZFN designed for h*IL2RG*.

Murine *il2rg* (m*il2rg*) like human, is expressed in blood cells rather than other tissue and is located on the X chromosome where it shares 69% sequence homology with h*IL2RG* at the cDNA level. It contains the conserved cysteine residues and WSXWS motif and the greatest region of homology is in the cytoplasmic tail however not in the SH2 subdomian (Cao, Kozak et al. 1993). In order to investigate whether a mouse model devoid of *il2rg* expression would result in the same phenotype as SCID-X1 patients DiSanto *et al*. engineered a $\gamma c^{-/-}$ mouse strain by excising exons 2-6 of murine *il2rg* via cre-*lox*P recombination of intronic *lox*P (DiSanto, Muller et al. 1995) in parallel Cao *et al*. developed another $\gamma c^{-/-}$ mouse with a gene truncated from exon 3 (Cao, Shores et al. 1995). These models have thymi that are markedly reduced in size and cellularity with the same architecture as normal mice but their thymocytes are unable to respond to mitogens *in vitro*. NK cells, intraepithelial lymphocytes and TCRγδ T cells are all absent and the population of B cells is decreased in the bone marrow with no IgE in the sera and no immunoglobulin isotype switching in response to IL-4 *in vitro*. The mice have smaller spleens than normal at age 3-4 weeks however further observation by Cao *et al*. showed that at age 4-9 weeks the mice displayed splenomegaly. Other observations by Cao *et al*. showed the lymph nodes lacked germinal centres and that the CD4+:CD8+ T cell ratio in the periphery was increased

due to the decreased population of CD8+ T cells. DiSanto *et al*. showed normal T cell populations in the thymus, a 12-fold reduction of mature IgM+ IgD+ B cells and a 3 fold reduction in IgM in the sera. In contrast Cao *et al*. observed no difference in IgM compared to normal. No mutant γc mRNA was detected by Cao *et al*. and DiSanto *et al*. confirmed the absence of γc by antibody detection but did not confirm whether any malformed protein was expressed. A third SCID-X1 mouse model was generated by deleting the cytoplasmic domain of m*il2rg* (Ohbo, Suda et al. 1996). The observations made in this mouse corroborated those made by both Cao *et al*. and DiSanto *et al*. and the IgM+ B cells were 15-20 fold less than control littermates, however IgM in the sera was increased in mice >8 weeks but was reduced at 10-11 weeks and the IgG levels were lower than WT control mice. The conclusions that can be made from these SCID-X1 animals are that they are hypothymic, have no NK cells, have reduced levels of mature B cells and CD8+ T cells yet their CD4+ T cell numbers remain comparable with WT.

The reduced B cell population in the $\gamma c^{-/-}$ mice is unlike the non-functional B cell population in human SCID-X1 patients. This is thought to be due to IL-7 in mice as B cell development is blocked before transition to the Pro B cell phase in IL-7 Knockout (von Freeden-Jeffry, Vieira et al. 1995) and IL7Rα knockout (Peschon, Morrissey et al. 1994) mice. In contrast, IL7Rα mutations in humans do not affect B cell development (Puel, Ziegler et al. 1998). However, T cell development is disrupted in both mouse and human on the loss of IL-7 signalling.

The T cell population present in $\gamma c^{-/-}$ mice does not reflect the T cell absence in SCID-X1 patients. When challenged with *Toxoplasma Gondii*, the depleted NK cells and CD8+ T cells which are vital for protection against this opportunistic pathogen in humans, where shown to be redundant in the $\gamma c^{-/-}$ model and the CD4+ population produced IFNγ in response to the infection (Scharton-Kersten, Nakajima et al. 1998). The enduring CD4+ T cells are also implicated in the occurrence of anaemia, splenomegaly and colitis in $\gamma c^{-/-}$ mice (Sharara, Andersson et al. 1997) and are likely to be due to the loss of the immunomodulatory function of IL-2 (Sadlack, Merz et al. 1993). Colitis can be resolved by keeping the animals in a pathogen free controlled environment which is essential for the maintenance of SCID models.

The outcome of positive and negative selection of T cells in the thymus of $\gamma c^{-/-}$ mice crossed with mice expressing a TCR specific for male antigen demonstrated that autoreactive CD8+ cells were negatively selected in the male mice negating the role of γc in this process. Likewise positive selection of male antigen specific CD8+ cells in female mice was unaltered (DiSanto, Guy-Grand et al. 1996). The thymus of SCID-X1 patients is hypoplastic, lacks the cortico-medullary distinction and Hassall bodies seen in normal thymus architecture and completely lacks CD3+ T cells (Hale, Buckley et al. 2004). The difference in T cell population between SCID-X1 humans and mouse models can be explained by the role of the tyrosine kinase. Where Jak3 kinase is used in humans, c-kit is used in murine thymocyte development as shown by the alymphoid thymus in c-$kit^{-/-}$ $\gamma c^{-/-}$ double knockouts (Rodewald, Ogawa et al. 1997).

Another alymphoid mouse model was generated by crossing the $Rag2^{-/-}$ mouse which is incapable of V(D)J rearrangement and therefore unable to develop B and T cells (Shinkai, Rathbun et al. 1992) with the $\gamma c^{-/-}$ mouse of DiSanto *et al*. (Goldman, Blundell et al. 1998). This genetic cross resolved the issue of splenomegaly and colitis observed in the $\gamma c^{-/-}$ mouse and the absent innate immunity in these mice makes them suitable recipients for HSC transplants from different animal origins.

In light of the requisite for a new animal model of SCID-X1, a humanised mouse model was engineered by InGenious (figure 5.1a). Briefly, the full length m*il2rg* with its promoter and enhancer elements was replaced by full length h*IL2RG* with the G691A point mutation (figure 5.1b). This humanised mouse model of SCID-X1 not only provides the mutant *IL2RG* target for site specific endonucleases, but due to the location of the novel human loci, it also takes into account the chromatin structure and epigenetic factors that can affect gene targeting which may be applicable to the human location of *IL2RG*.

This chapter determines whether this SCID-X1 mouse model, hγcmut, results in a SCID phenotype by comparing it to a $\gamma c^{-/-}$ mouse. Subsequently, lineage negative (lin^-) HSC from this mouse are enriched and transduced with lentiviral constructs in order to introduce h*IL2RG* as a transgene and assess whether these transformed multipotent cells are capable of restoring a fully functional immune system. The 2 constructs used to demonstrate immune cell reconstitution by restoring γc expression were, the

IL2RGp-*IL2RG* lentiviral vector used in the previous chapter or an SFFVp-*IL2RG* lentiviral vector construct, comprising the U3 region of the spleen focus forming virus as a promoter (SFFV) and *IL2RG* cDNA. Both the IL2RG and SFFV promoters contain an Ets binding site (EBS) (Markiewicz, Bosselut et al. 1996), (Baum, Itoh et al. 1997). The Ets family of transcription factors regulate hematopoietic cell development including T cell, B cell and NK cell differentiation (Anderson, Hernandez-Hoyos et al. 1999) and the EBS in the *IL2RG* promoter is involved in the regulation of γc expression in various cell types (Markiewicz, Bosselut et al. 1996).

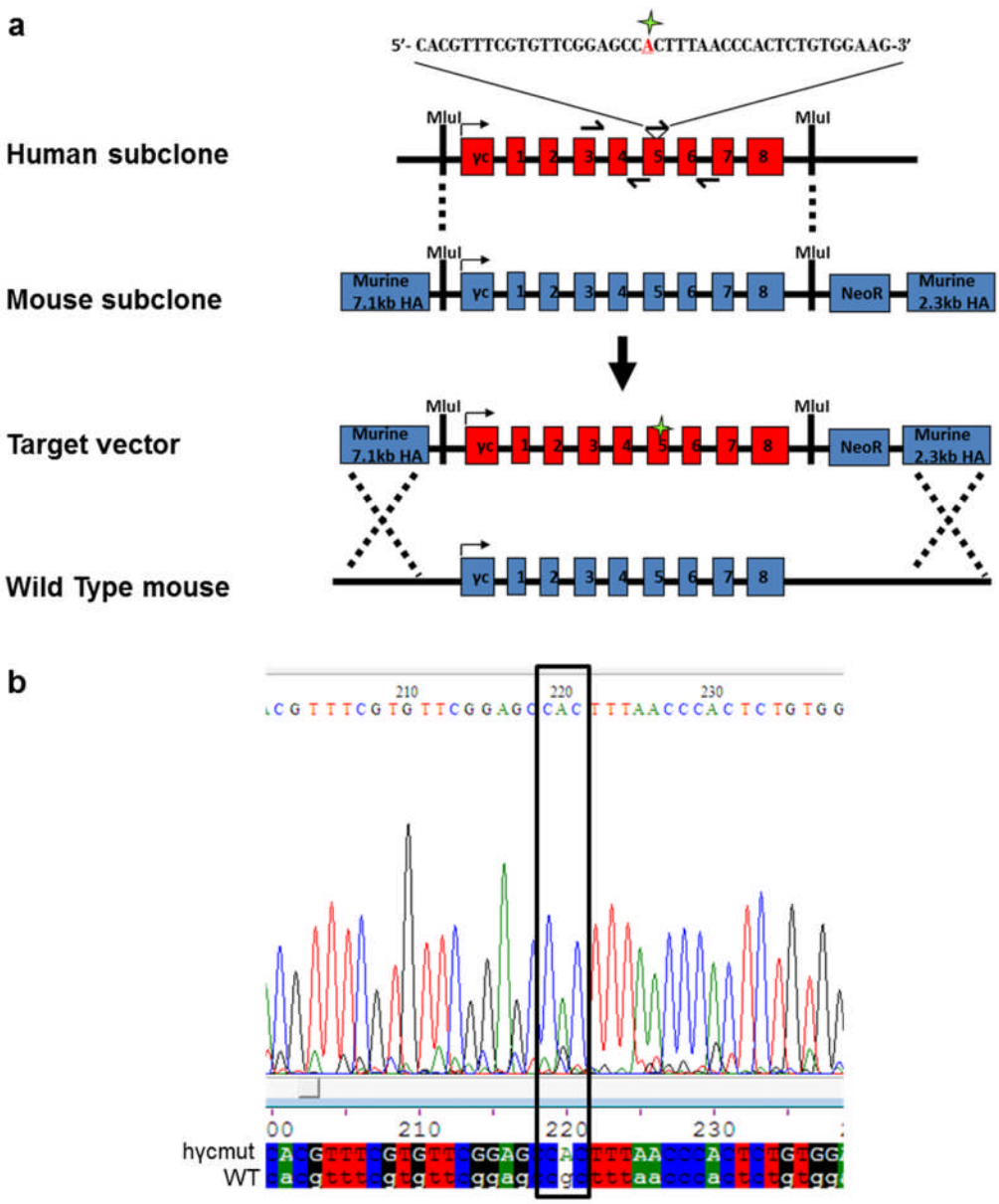


Figure 5.1 Engineering the humanised mouse model of SCID-X1.

(a) The human *IL2RG* was subcloned into a shuttle vector and the G691A point mutation introduced by site directed mutagenesis. The murine *il2rg* was also subcloned and both the human and murine *il2rg* loci were flanked by MluI sites which allowed for gene exchange. Now the mutated human gene was flanked by murine homology arms, the shuttle vector was electroporated into murine ES cells and successful homologous recombination was selected for by Neomycin resistance (NeoR). Following selection the LoxP flanked NeoR was excised by Cre resulting in humanised mouse (hγcmut) ES cell lines ready for implantation. (b) The hγcmut loci containing the G691A mutation in the selected ES cell line was amplified by PCR which underwent high throughput sequencing. The read out is aligned with the known WT sequence to confirm the mutation as present seen at base pair 220 of the read out.

5.3 Characterisation of the humanised γc mutant mouse

InGenious provided 2 $\gamma c^{\Delta/Y}$ homozygous males and one $\gamma c^{\Delta/+}$ heterozygous female. A pure strain of hγcmut mouse was bred whereby the females were homozygous for the mutated human γc gene. To genotype the mice, a PCR strategy was designed to generate a 623bp or 206bp product if the human or mouse gene were present, respectively (Figure 5.2a). Figure 5.2b shows the PCR products that relate to the genotype of the mice received from InGenious as compared to a BL6 WT mouse. Only PCR products generated from murine specific primers were seen in the BL6 wild type mouse and the heterozygous female. The PCR products generated by human specific primers were only seen in both of the knock in males and the heterozygous female but not in the WT sample. A combination of mouse and human specific primers were successful in amplifying both m*il2rg* and h*IL2RG* targets in the heterozygous female, this combination was therefore used as the method of screening for litters up to and including the founding litter as seen in Figure 5.2c. The founding litter contained only human mutant *IL2RG* (h$IL2RG^{\Delta}$) and there was no evidence of m*il2rg* as compared to a heterozygous control, confirming a pure strain had been bred.

Having confirmed the presence of h$IL2RG^{\Delta}$ in the hγcmut mouse, it was then investigated whether the missense mutation resulted in lack of transcript expression. Total mRNA was isolated from hγcmut and WT mice and human peripheral blood and cDNA generated by RT-PCR (Figure 5.2d). The 76bp amplicon generated by the primers recognising the exon 6 and 7 boundary of the cDNA template was seen exclusively in the hγcmut mouse and human control samples. An amplicon was detected in the WT sample, but as its size was at least 3 fold greater than the 76bp specific product, it was more likely to be due to non-specific binding. Amplicons generated by human *GAPDH* and murine *gapdh* cDNA detecting primers verified the presence of human and murine RNA respectively. It can be confirmed that the hγcmut mice are capable of transcribing the h$IL2RG^{\Delta}$ gene mRNA.

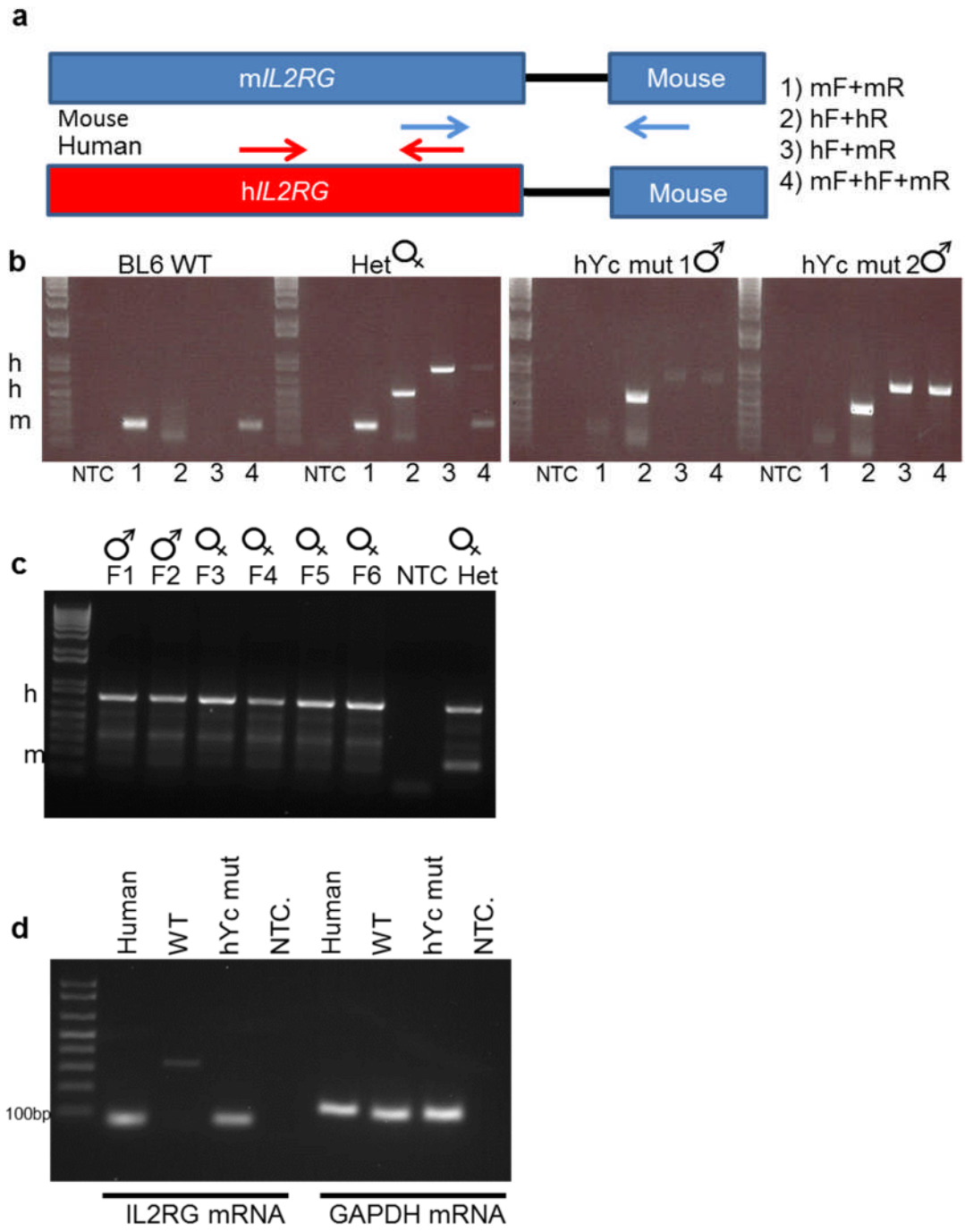


Figure 5.2 Genotyping the hγcmut mouse model.

(a)The PCR strategy for detecting the human *IL2RG* or murine *il2rg* in the mouse model, blue arrows represent murine specific primers and the red arrow is human specific. The PCR products amplified from transgenic mice by primer pair combinations 1-5. Human forward (hF) and mouse reverse (mR) or mouse forward (mF) and mouse reverse (mR) confirm the heterozygous female and 2 knock-in males in (b). (c) The homozygous founding litter genotype as amplified by the mouse and human forward with mouse reverse is shown alongside a heterozygous control. (d) cDNA was generated from total mRNA from human, hγcmut and WT blood cells. The 76bp *IL2RG* amplicon was seen in human and hγcmut samples. *GAPDH* was amplified for control with primers specific for human *GAPDH* or murine *gapdh*.

5.3.1 Immune cells in the peripheral blood

It is not possible to differentiate hγc in murine immune cells using the current antibodies available. Further analysis of the hγcmut mouse was carried out to ratify this model as immunodeficient by comparing the immune cell profile to WT BL6 mice and the pre-existing γc knockout ($\gamma c^{-/-}$) mouse model (DiSanto, Muller et al. 1995). Immune cell populations were acquired from the following tissues: Blood, spleen, thymus, lymph node and bone marrow. Specific immune cell subsets were determined using fluorophore conjugated antibodies specific for lymphocyte markers and quantified using flow cytometry. The lymphocyte subsets analysed included T cells, defined by their $CD4^+$ and $CD8^+$ cell surface markers, $NK1.1^+$ $CD3^-$ NK cells and IgM^+ $B220^+$ B cells. Representative flow cytometry plots of the proportion of immune cell types in the peripheral blood are shown as a percentage of the total live gate, determined by forward and side scatter (FSc and SSc respectively) (Figure 5.3a) and the total data accumulated from analysed mice are shown in figure 5.3b. Here it is clearly demonstrated that the peripheral blood of the $\gamma c^{-/-}$ and the hγcmut mouse model have significantly fewer CD8+ cells, NK cells and IgM+ B cells than WT ($P< 0.001$). A well-established CD4+ cell population was seen in all three mouse models. Because the immune cell populations were calculated as a proportion of total cells, this method of analysis does not consider the total number of immune cells in circulation. Due to the nature of pathogenic *IL2RG* deficiency, it is possible that there are fewer immune cells circulating the peripheral blood and tissues as compared to wild type. As a result, it is possible that the proportions of immune cells found in each tissue don't accurately represent the insufficient immune repertoire of the SCID-X1 models. In order to address this observation, the absolute number of cells analysed by flow cytometry were calculated using a known number of absolute counting beads (Figure 5.3c). Again, there is a CD4+ cell population in the hγcmut mouse similar to WT, and in this instance, the decrease in CD8+ cells does not reach significance, possibly due to a smaller sample number. In support of the proportional data, the IgM+ B cells and NK cells are consistently reduced, as compared to WT. Taking both the proportional data from the various tissues examined and the absolute cell counts from the blood, it can be concluded that the hγcmut mice, like the $\gamma c^{-/-}$ mice, do not have the full repertoire of immune cells and are therefore immunodeficient.

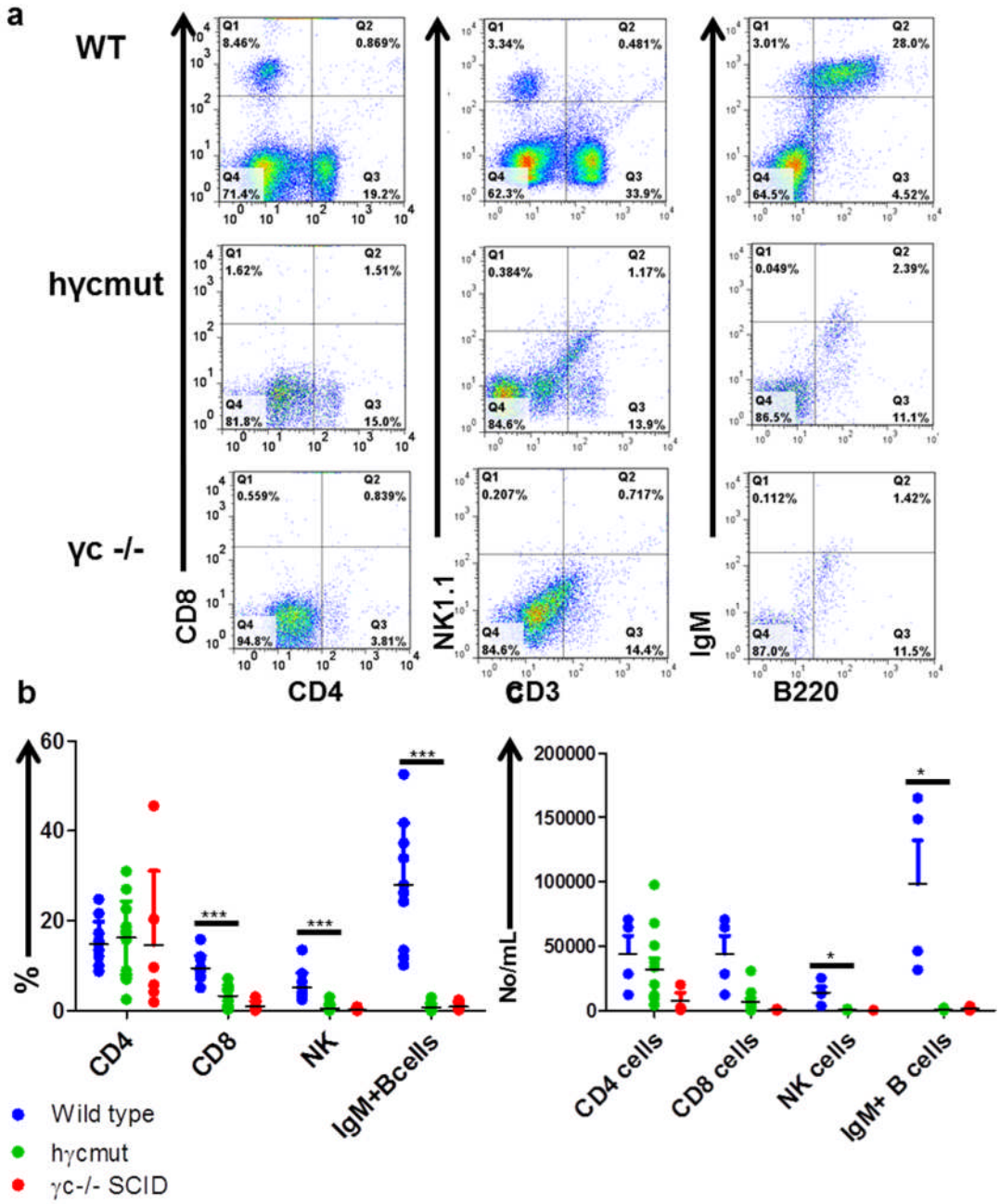


Figure 5.3 Immune cells in the blood of hγcmut, WT and γc$^{-/-}$ mice.

(a) Representative flow cytometry plots from the analysis of immune cell isolation from the peripheral blood of WT, hγcmut and γc$^{-/-}$ mouse model. The live cells were established by the FSc, SSc plot. T cells are established by CD4 or CD8 expression, NK cells determined by CD3- NK1.1+ expression and Mature B cells by IgM and B220 expression. (b) Vertical scatter plots of the immune cell percentage in the peripheral blood, of WT (blue), hγcmut (green) and γc$^{-/-}$ (red) mice, as analysed by flow cytometry. (c) Scatter plots of the absolute cells count per mL of immune cells in a known volume of peripheral blood as analysed by flow cytometry. The comparison of CD4+ or CD8+ T cells, NK cells and IgM+ B cells did not have a normal distribution and were analysed using the Kruskal Wallis extension of the Mann-Whitney test and the P values generated by comparing the hγcmut mouse to WT or γc$^{-/-}$ mice are indicated by star ratings: *** : P=<0.001; ** : P=<0.01, * : P=<0.05.

5.3.2 Thymic development in SCID-X1 mice

Characterising the phenotype of the hγcmut mouse continued with a more in-depth analysis of particular tissues that may be affected by aberrant γc expression. As it has been observed that mice lacking γc cannot develop a normal bi-lobular thymus and only a rudimentary thymus is seen, thymi of hγcmut and $\gamma c^{-/-}$ mice were compared to WT mice (Figure 5.4a). Example thymi are shown *in situ* and isolated. The WT thymus example was clearly bi-lobular with a width of at least 8mm. Conversely, thymi dissected from mature hγcmut and $\gamma c^{-/-}$ mice were smaller and a bi-lobular structure was not present. Although it can be seen that a thymus is unable to fully develop in the hγcmut mouse, T cells can develop in this rudimentary thymus. The flow cytometry plots (Figure 5.4b). demonstrated that the CD4+ and CD8+ single positive populations and the CD4+ CD8+ double positive population were found in the hγcmut mouse and the $\gamma c^{-/-}$ mouse. This corroborates with the CD4+ cell populations found in the periphery of the hγcmut and $\gamma c^{-/-}$ mouse. However, the CD8+ cell population was greater in the WT mouse compared to the hγcmut and $\gamma c^{-/-}$ mice. On analysing the immune cell populations of a greater sample size (Figure 5.4c), there was no difference in the CD4+ or CD8+ cells between the WT and hγcmut thymi. The contrasting decrease of T cells seen in the $\gamma c^{-/-}$ mouse may be due to the larger number of hγcmut mice analysed compared to $\gamma c^{-/-}$ mice. The populations of NK cells and IgM+ B cells in WT were not seen in the hγcmut and $\gamma c^{-/-}$ counterparts. To summarise, when the mutated human *IL2RG* is present in place of the functional murine *il2rg* locus, only rudimentary thymus develops.

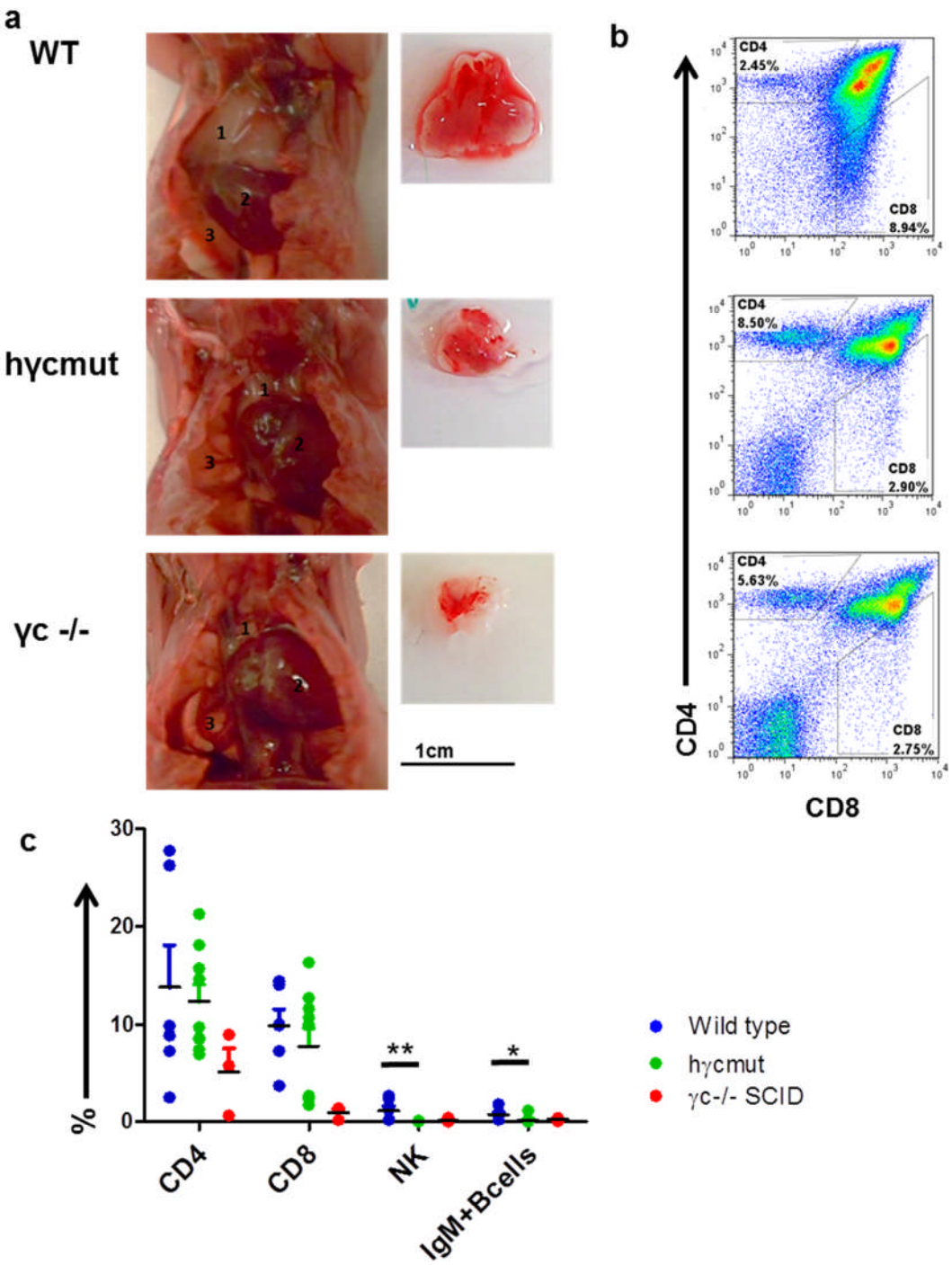


Figure 5.4 Immune cells in the hγcmut, WT and γc$^{-/-}$ thymi.

(a) examples of mature murine thymi from hγcmut WT and γc$^{-/-}$ aged over 10 weeks. Left, thymi (1) heart (2) and lung (3) *in situ*. Right, isolated thymi with a corresponding scale bar. (b) Representative flow cytometry plots of of CD4+ and CD8+ single and positive T cells analysed by FACS, 100,000 events were acquired with a minimum of 40,000 events in the live gate. (c) Vertical scatter plots of the collective data of all the immune cells analysed in the thymus did not have a normal distribution and was analysed using the Kruskal Wallis extension of the Mann-Whitney test and the P values generated by comparing the hγcmut mouse to WT or γc$^{-/-}$ mice are indicated by star ratings: ** : P=<0.01; * : P=<0.05.

5.3.3 Immune cells in the spleen, bone marrow and lymph node

It has also been observed in $\gamma c^{-/-}$ mice that there is a propensity for splenomegaly. Spleens from WT, hγcmut and $\gamma c^{-/-}$ mice over 9 weeks of age were compared. Representative spleens isolated from each mouse type demonstrated that the spleen size from the SCID-X1 mouse is greater than WT (Figure 5.5a). The scatter graph of splenic mass from the different mouse groups shows an increased spleen mass in the SCID-X1 mouse models (Figure 5.5b). Analysis of the immune cell populations reveals immune cell distribution similar to the observations made from peripheral blood (Figure 4.5c) and taking into account the splenic mass and size, it can be concluded that like the $\gamma c^{-/-}$ mouse, the hγcmut mouse also exhibits signs of splenomegaly.

To complete the comparison of the immune cells of the mouse models, the immune cells of the bone marrow and lymph node were also analysed by flow cytometry (Figure 4.5d). The bone marrow showed a similar immune cell distribution to the blood with a decrease in NK cells and IgM+ B cells however a significant decrease in CD8+ cells was not well defined. The immune cell populations of the lymph nodes were not significantly different between the mouse models which is possibly due to the low cell numbers retrieved from the isolation of the tissue.

Overall, the immune cell populations from the hγcmut mouse reflect that of the $\gamma c^{-/-}$ mouse and also have the same rudimentary thymus associated with a lack of γc expression. Furthermore, the existing CD4+ cell population observed in the SCID mouse models can be implicated in the occurrence of splenomegaly observed in these mouse models.

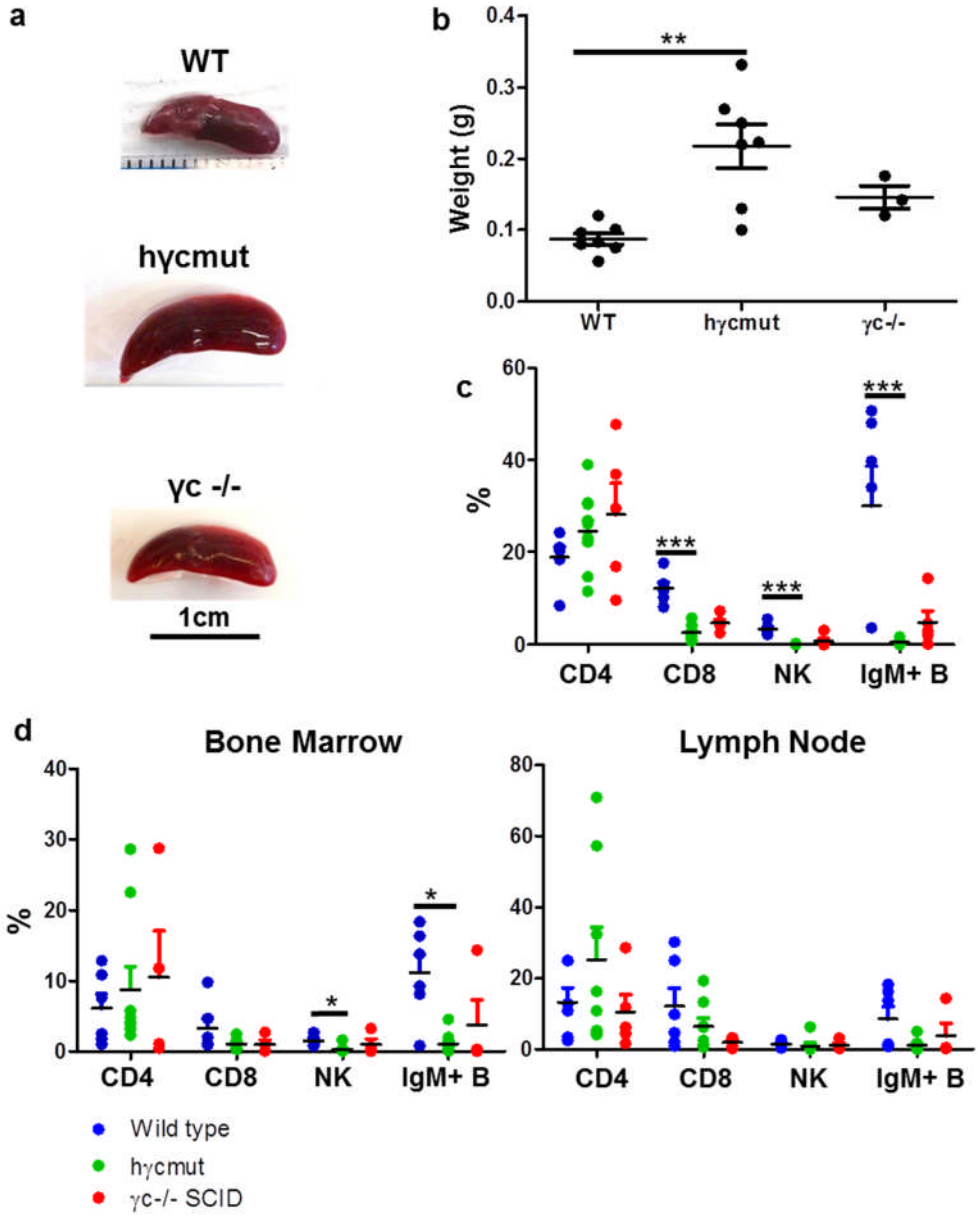


Figure 5.5 Immune cells and mass of the hγcmut, WT and γc$^{-/-}$ spleen.

(a) representative age matched spleens from >9 week old WT, hγcmut mice and γc$^{-/-}$ with scale bar. (b) scatter plots of the splenic mass of mice >9 weeks of age. (c+d) vertical scatter plot of the percentage of immune cell populations in WT (blue), hγcmut (green) and γc$^{-/-}$ (red) mice. (c) spleens (d) bone marrow and lymph node. Data were analysed using the Kruskal Wallis extension of the Mann-Whitney test and the P values generated by comparing the hγcmut mouse to WT or γc$^{-/-}$ mice indicated by star ratings: *** : P=<0.001; ** : P=<0.01; * : P=<0.05.

5.4 Reconstitution of the immune system by introducing human *IL2RG*

In order for the hγcmut mouse to be a suitable model for gene targeting of the h*IL2RG* locus, it must be possible to reverse the disease phenotype in this model of SCID-X1 by introducing h*IL2RG*. The ability of functional human γc to complete interleukin receptors missing the murine γc and develop a functional immune cell repertoire was assessed in this model by transducing progenitor cells from the hγcmut model with a h*IL2RG* gene and observing the possibility of restoring a functional immune system when transplanted into a Rag2$^{-/-}$γc$^{-/-}$ immunodeficient mouse of the same, C57BL/6 genetic background. This strategy has proven successful as Rag2$^{-/-}$γc$^{-/-}$c5$^{-/-}$ are suitable recipients as they have no immune cells to compete with, or reject the transplant (Thornhill, Schambach et al. 2008). As haematopoietic progenitors capable of long term engraftment are resting in the G0 phase of the cell cycle immediately after isolation from the donor (Reems and Torok-Storb 1995), the lin^{-} cells from the donor mice were stimulated with Stem cell factor (SCF), murine Flt3 ligand (mFlt3) and human thrombopoietin (hTPO) as these factors aid self-renewal of lin^{-} cells *in vitro* with limited multipotency (Hennemann, Conneally et al. 1999; Oostendorp, Audet et al. 2000) and progression from the G0 phase of the cell cycle enhances lentiviral vector transduction (Sutton, Reitsma et al. 1999).

5.4.1 The reconstitution assay

A reconstitution assay was carried out whereby lin^{-} cells were enriched from the bone marrow of hγcmut mice, transduced with lentiviral vector overnight in StemSpan supplemented with SCF, mFlt3 and hTPO. The cells were then re-suspended in RPMI and injected intravenously into lethally irradiated Rag2$^{-/-}$γc$^{-/-}$ mice. (Figure 5.6a), Three integrating lentiviral vectors were each used at an MOI of 10 as calculated by flow cytometry analysis of either GFP or cell surface γc in the case of GFP or h*IL2RG* constructs respectively. 3 mice were treated with cells transduced with the full length h*IL2RG* with its promoter and regulatory elements (IL2RGp-*IL2RG*). 3 with h*IL2RG* cDNA regulated by an SFFV promoter (SFFVp-*IL2RG*) and 2 with an SFFVp-GFP construct for control (Figure 5.6b). Untransduced lin^{-} cells from WT mice were also injected into 2 additional lethally irradiated mice at this point for positive control. Flow cytometry analysis of the lin^{-} cells selected by magnetic activated cell sorting (MACS) from either

WT or hγcmut bone marrow confirms that the NK cells and IgM+ B cells were successfully depleted however some T cells were still present (Figure 5.6c). The extent of lentiviral transduction was assessed at the pre-injection stage by incubating the cells in complete StemSpan at 37°C for 72 hours, extracting the genomic DNA and quantifying the copy number of transgene by qPCR (Figure 5.6d). Primers and probe specific for the WPRE detected the integrated viral backbone and primers and probe for the murine housekeeping gene, *titin*, quantified the cell number. The proviral copy number was calculated by dividing the WPRE copy number by the *titin* copy number and multiplied by two, to account for the 2 copies of *titin*. From the copy number observed in these cells it was seen that the IL2RGp-*IL2RG* and the SFFVp-GFP constructs achieved a greater copy number than the SFFVp-*IL2RG*. The transplanted animals were maintained for 16 weeks in an environment suitable for immunodeficient animals.

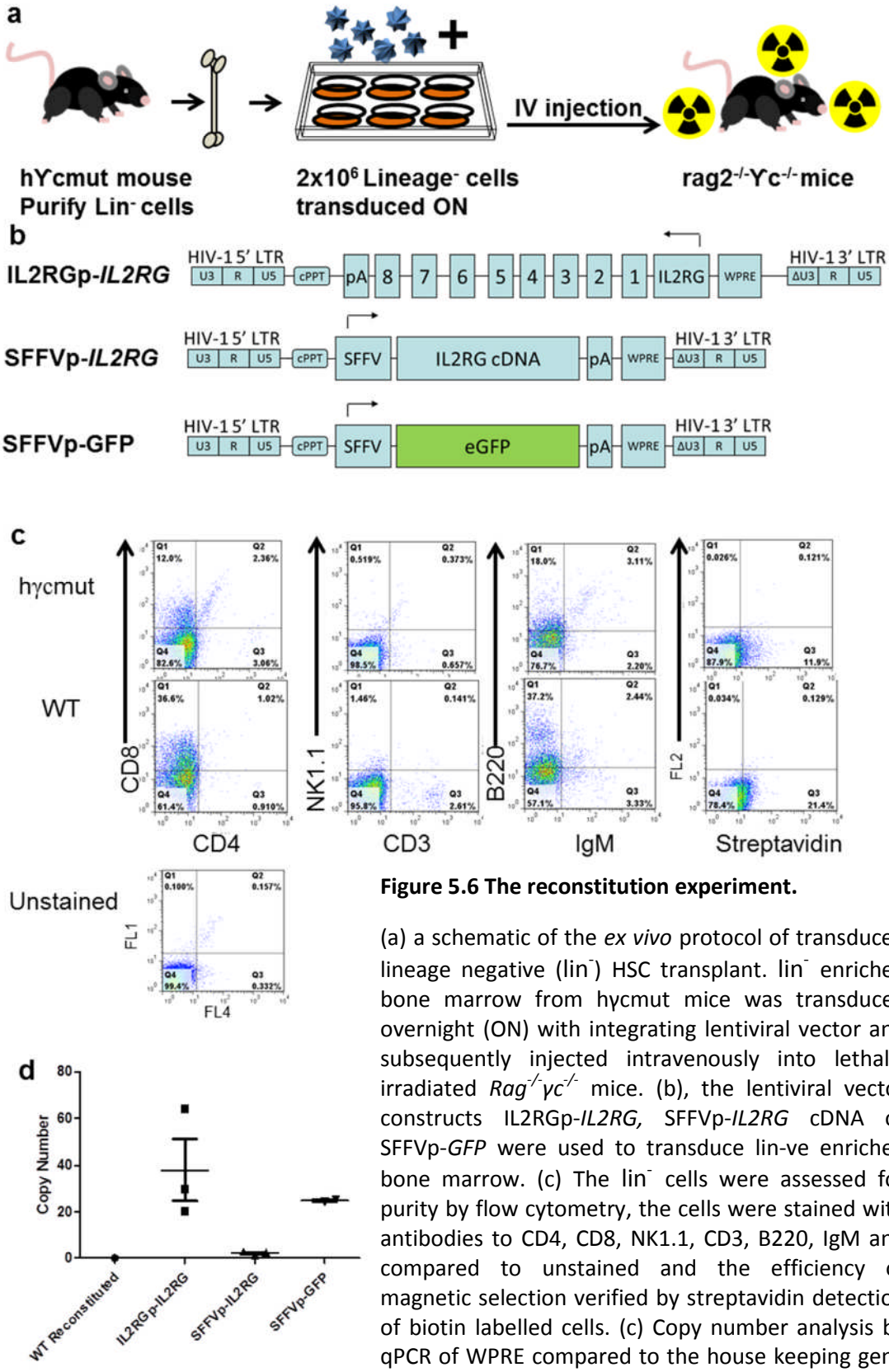


Figure 5.6 The reconstitution experiment.

(a) a schematic of the *ex vivo* protocol of transduced lineage negative (lin$^-$) HSC transplant. lin$^-$ enriched bone marrow from hγcmut mice was transduced overnight (ON) with integrating lentiviral vector and subsequently injected intravenously into lethally irradiated *Rag*$^{-/-}$*γc*$^{-/-}$ mice. (b), the lentiviral vector constructs IL2RGp-*IL2RG,* SFFVp-*IL2RG* cDNA or SFFVp-*GFP* were used to transduce lin-ve enriched bone marrow. (c) The lin$^-$ cells were assessed for purity by flow cytometry, the cells were stained with antibodies to CD4, CD8, NK1.1, CD3, B220, IgM and compared to unstained and the efficiency of magnetic selection verified by streptavidin detection of biotin labelled cells. (c) Copy number analysis by qPCR of WPRE compared to the house keeping gene *titin* in lin$^-$ cells, 72 hours post transduction

5.4.2 Reconstitution at 12 weeks

The presence of immune cells reconstituted by the transduced donor hγcmut bone marrow cells was analysed by flow cytometry. Peripheral blood was extracted from the mice 12 weeks post-transplant and stained with antibodies to detect mature cell lineages (Figure 5.7a). Example flow cytometry plots reveal CD4+ cells, CD8+ cells. NK cells and IgM+ B cells in all the mice treated with cells transduced with functional *IL2RG* constructs and in mice receiving WT lin^- cells. The mice transplanted with lin^- cells transduced with SFFVp-GFP did not show any reconstitution of the immune cells analysed. The ratio of CD4+ to CD8+ cells was 1.9 in mice transplanted with WT and 2.3 and 3.3 in mice treated with cells transduced with IL2RGp-*IL2RG* and SFFVp-*IL2R*G respectively. In the previous characterisation data this ratio in an untreated WT mouse was seen to be 2.3 (Figure 5.3a). The greatest reconstitution of NK cells was observed in the WT reconstitution control and second greatest in the SFFVp-IL2RG treated example. Interestingly, the IgM+ B cells were more frequent in the recipients of *IL2RG* transduced donor cells than the WT control. The total immune cell data retrieved from all 10 of the transplanted mice (Figure 5.7 b) showed an increase in CD4+, CD8+ and IgM B cells in the mice treated with either WT or IL2RG transduced lin^- cells compared to the GFP control. No significant conclusions could be drawn as only 2 GFP control mice were analysed.

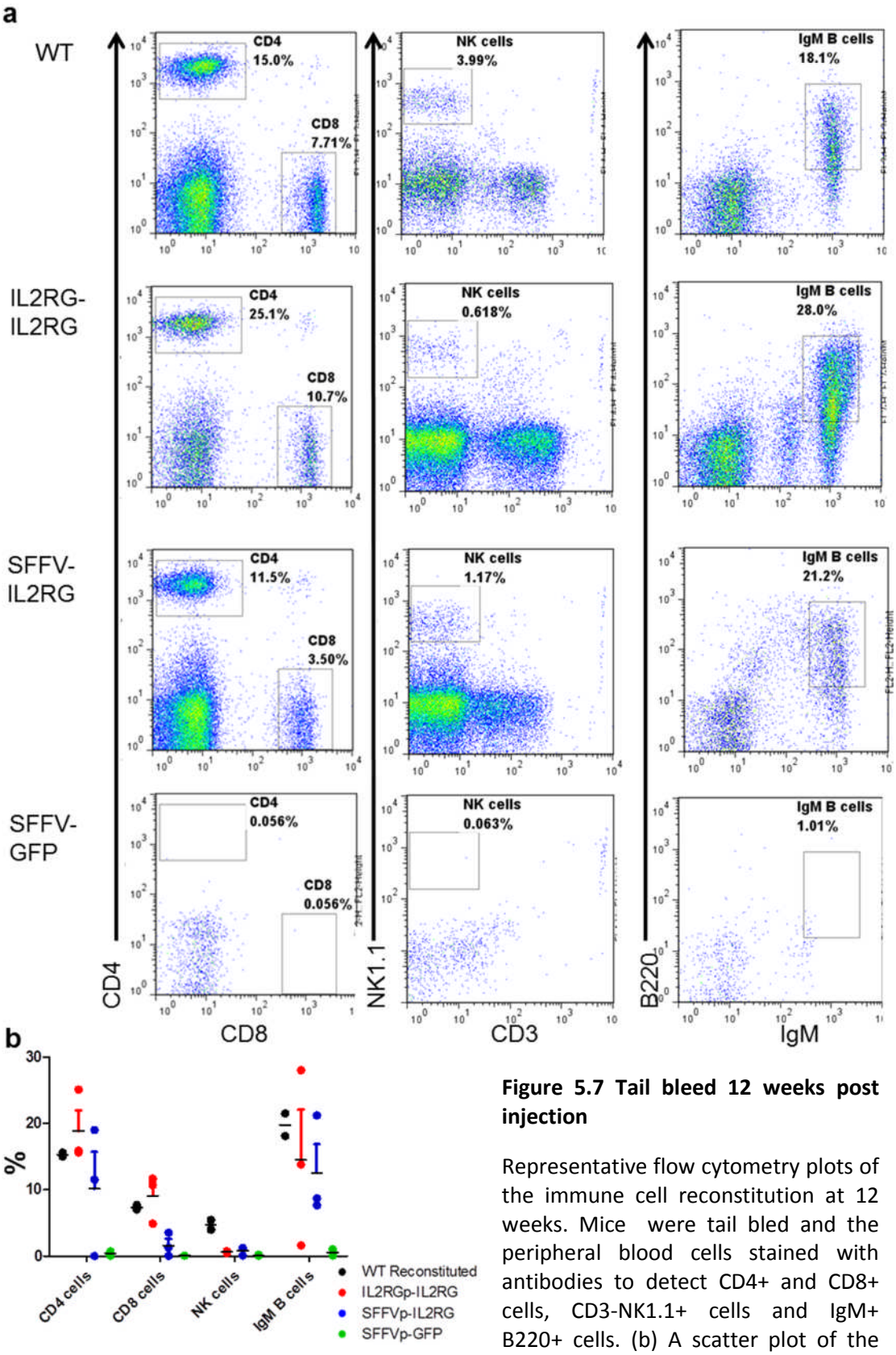


Figure 5.7 Tail bleed 12 weeks post injection

Representative flow cytometry plots of the immune cell reconstitution at 12 weeks. Mice were tail bled and the peripheral blood cells stained with antibodies to detect CD4+ and CD8+ cells, CD3-NK1.1+ cells and IgM+ B220+ cells. (b) A scatter plot of the percentages of the reconstituted immune cells 12 weeks after transplant.

5.4.3 Reconstitution at sacrifice

Due to the successful reconstitution of immune cells analysed in the periphery at 12 weeks post-transplant, further analysis of the tissue in the recipient mice was carried out. Apart from one SFFVp-*IL2RG* treated mouse that died at week 12 due to an unrelated circumstance, the remaining mice were sacrificed at week 16 plus 3 untreated Rag2$^{-/-}$γc$^{-/-}$ mice for comparison. Cells from all the mice were isolated from the blood, bone marrow, spleen, and thymus. Immune cells were stained with lymphoid specific antibodies and acquired by flow cytometry. The populations, as a percentage of the total live cells, (Figure 5.8a) confirmed that all of the immune cell types analysed were increased in both the *IL2RG* treated hγcmut lin^{-} cells and the WT positive control mice. However, due to small sample sizes, the clear trend is not supported by statistical analysis. The increase is most notable in the percentage of T cells in the blood, bone marrow and spleen when the IL2RGp-*IL2RG* lentiviral vector was used and the increase of IgM+ B cells in the blood and spleen when the SFFVp-*IL2RG* vector was used. The SFFVp-*IL2RG* lentiviral vector did appear to reconstitute CD4+, NK and IgM+ B cells but the difference was less striking concerning the CD8+ cells. The GFP control recipients showed a moderate presence of T cells in the spleen yet fewer than untreated hγcmut mice (Figure 5.5c) and a remarkable disparity of these cells in the thymus between the 2 animals analysed. Conversely there is no evidence of NK or IgM+ B cells in the treated negative control. There are no values for immune cell populations in Rag2$^{-/-}$γc$^{-/-}$ thymi due to the absence of this tissue in these animals.

The overall pattern of immune cell repertoire in the treated mice is that IL2RG treated hγcmut lin^{-} cells are capable of restoring immune cell populations similar to the WT positive control than the GFP negative control which displays immune cell populations with a similar pattern to untreated Rag2$^{-/-}$γc$^{-/-}$ mice.

To determine the engraftment capacity of the transduced donor cells, the proviral copy number was assessed in the various tissues of the treated mice by quantitative PCR (qPCR) (Figure 5.8b). Genomic DNA was extracted from the peripheral blood, bone marrow, spleen and thymus of mice sacrificed at week 16. Primers and probe specific for the WPRE detected integrated viral backbone. Likewise, primers and probe for the murine housekeeping gene, *titin*, quantified the cell number. The proviral copy number

was calculated by dividing the WPRE copy number by the *titin* copy number. Unlike the transduction efficiency at 72 hours where the quantity of transgene achieved by the IL2RGp-*IL2RG* construct was greater than the SFFVp-*IL2RG* construct after transplantation, the SFFV promoter constructs now exceed the IL2RGp-*IL2RG* construct's copy number. This could be in part due to the higher transduction efficiency resulting in cell morbidity or it may be due to regulatory elements which allow physiological control of the haematopoietic stem cell differentiation.

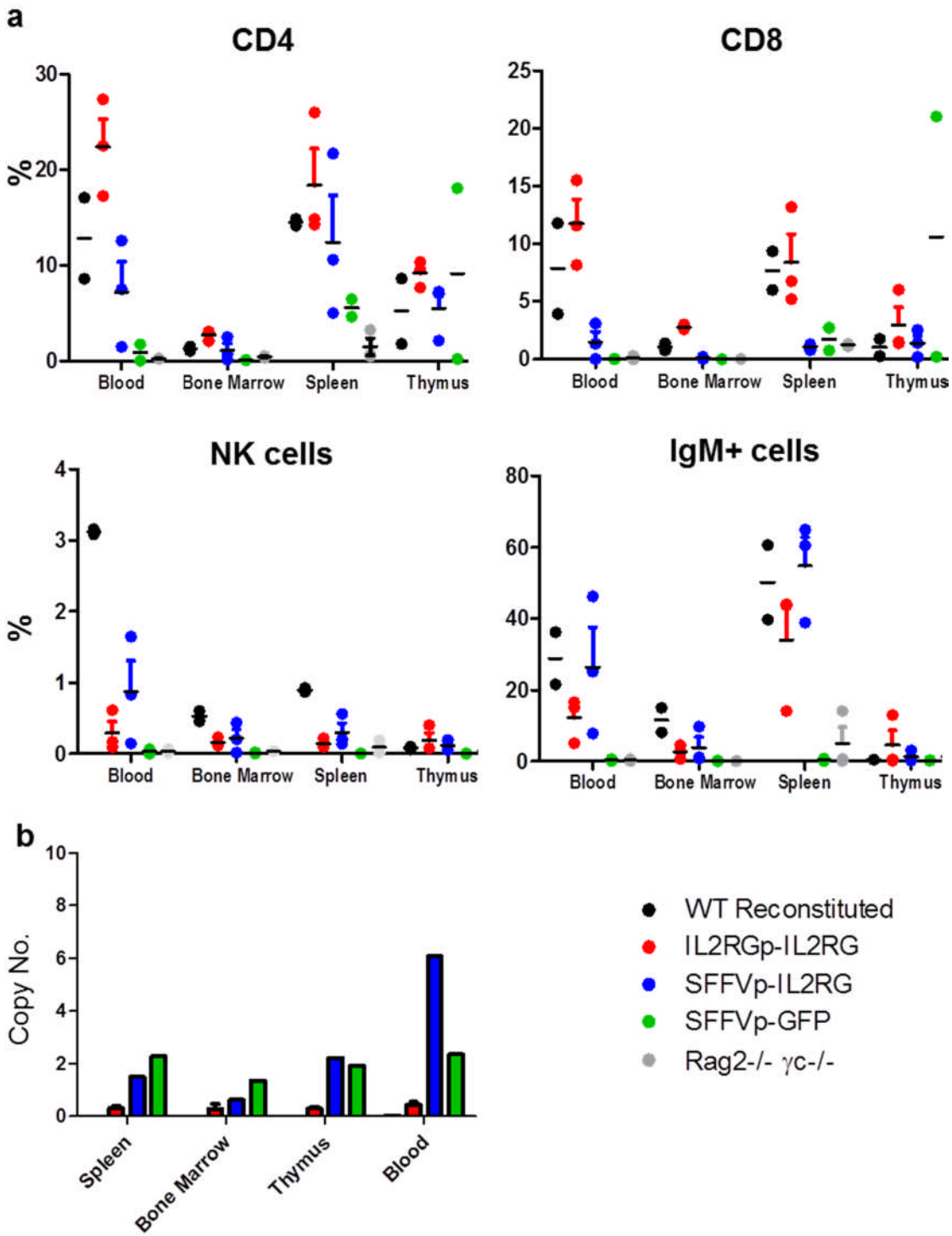


Figure 5.8 Immune cell reconstitution at sacrifice and transgene copy number at 16 weeks

(a) Scatter plots of the flow cytometry analysis of the immune cells from the spleen, bone marrow, thymus and peripheral blood in mice transplanted with lin$^-$ cells from either WT mice or hγcmut mice transduced with IL2RGp-*IL2RG,* SFFVp-*IL2RG* or SFFVp-GFP and Rag2$^{-/-}$γc$^{-/-}$ non-transplanted control. The small sample sizes do not warrant a one way analysis of variance. (b) Proviral copy number at 16 weeks detected by qPCR in the spleen, bone marrow, thymus and peripheral blood mice transplanted with lin$^-$ from WT (n=2), or hγcmut transduced with IL2RGp-*IL2RG* (n=3), SFFVp-*IL2RG* (n=2), SFFVp-GFP (n=2) and calculated by the WPRE copy divide by the copy number of *titin* housekeeping gene.

5.4.4 Thymic development in reconstituted mice

The Rag2$^{-/-}$γc$^{-/-}$ mouse is incapable of developing a thymus and a sign of successful immune system reconstitution is the development of thymic tissue after treatment. In all of treated mice thymic tissue was present at sacrifice, including the GFP control (Figure 5.9a). From the examples of thymi *in situ* and isolated from the treated mice and untreated WT and Rag2$^{-/-}$γc$^{-/-}$ controls it can be seen that the thymi are smaller than those dissected from WT mice. Example thymi from mice treated with lin^{-} cells from hγcmut mice incorporating an *IL2RG* transgene and the WT reconstitution control more closely resembled the bi-lobular structure of the untreated WT positive control than the SFFVp-GFP treated control. Furthermore the CD4+/CD8+ single positive and double positive T cells of the mice reconstituted with functional *IL2RG* containing lin^{-} cells had a consistent pattern of distribution whereas the thymic tissue isolated as a result of SFFVp-GFP treatment was uni-lobular like the thymi seen in hγcmut mice and had an aberrant distribution of CD4+/CD8+ single and double positive T cells (Figure 5.9b). The second SFFVp-GFP treated mouse did not have any CD4 or CD8 staining (Appendix 12) accounting for the large variation of the CD4+ and CD8+ cells observed in the 4 organ comparison flow cytometry data in the thymi of SFFVp-GFP treated mice. The untreated WT positive control displayed a unique CD4+/CD8+ distribution, most likely due to a greater amount of cells isolated from the tissue resulting in fewer antibodies per cell and therefore a dimmer stain. The tissue extracted from the Rag2$^{-/-}$ γc$^{-/-}$ mouse for control contains no cells with either the CD4 or CD8 T cell marker is unlikely to be thymic tissue. The immune cells of the thymi were analysed by flow cytometry and it was clearly demonstrated that the CD4+, CD8+ single and double positive populations of T cells are present the mice which received lin^{-} cells capable of γc expression.

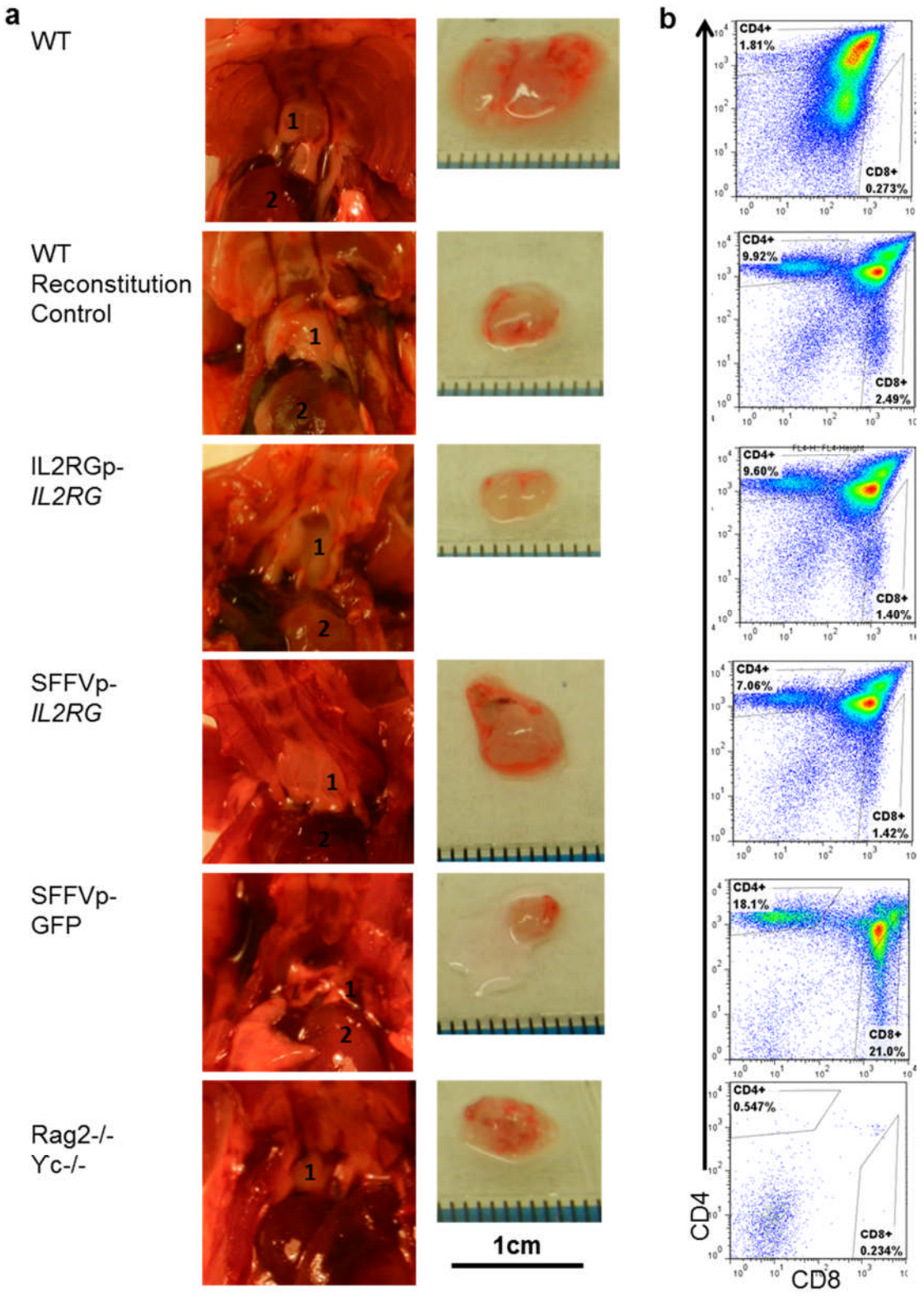


Figure 5.9 Thymic reconstitution.

(a) Examples of thymi dissected from Rag2$^{-/-}$γc$^{-/-}$ mice treated with IL2RGp-*IL2RG* , SFFpV-*IL2RG* and SFFVp-GFP transduced hγcmut lin^{-} cells or WT lin^{-} cells for reconstitution control compared to untreated WT and Rag2$^{-/-}$γc$^{-/-}$ control mice, *in situ* (Left) thymi (1) heart (2), and isolated with a scale bar (Right). (b) The flow cytometry analysis of the CD4+ and CD8+ single and double positive T cells from these thymi.

5.4.5 Splenic observations in the reconstituted mice

The Rag2$^{-/-}$γc$^{-/-}$ mouse has a hypocellular spleen due to a lack of lymphocytic infiltrate, therefore an increase in splenic mass and cellularity after treatment will be indicative of a recovered immune system. Spleens were removed from the treated and control groups of mice. Examples of dissected spleens from the test and control mice demonstrated an increase in splenic size (Figure 5.10a) and mass (Figure 5.10b) in all the mice that received donor cells compared to the Rag2$^{-/-}$γc$^{-/-}$ control. However, the splenic mass of the IL2RGp-*IL2RG* treated mice is not significantly greater than the Rag2$^{-/-}$γc$^{-/-}$ control.

To assess the proliferative capacity of the engrafted immune cells in response to stimuli, splenocytes were isolated from the spleens of Rag2$^{-/-}$γc$^{-/-}$ mice treated with lin- cells from either WT mice or hγcmut mice, transduced with IL2RGp-*IL2RG* or SFFVp-GFP. Untreated WT and Rag2$^{-/-}$γc$^{-/-}$ control murine splenocytes were also included. The splenocytes were incubated with cytokines IL-2 or IL-7 with or without Concanavalin A (ConA) which is a T cell mitogen (Figure 5.10c). The proliferation index, a measure of proliferation in response to cytokines divided by medium alone, demonstrated that providing functional *IL2RG* results in a response to ConA which was increased with the further addition of IL-2 or IL-7. No proliferative cytokine response was observed in the SFFVp-GFP or Rag2$^{-/-}$γc$^{-/-}$ control splenocytes.

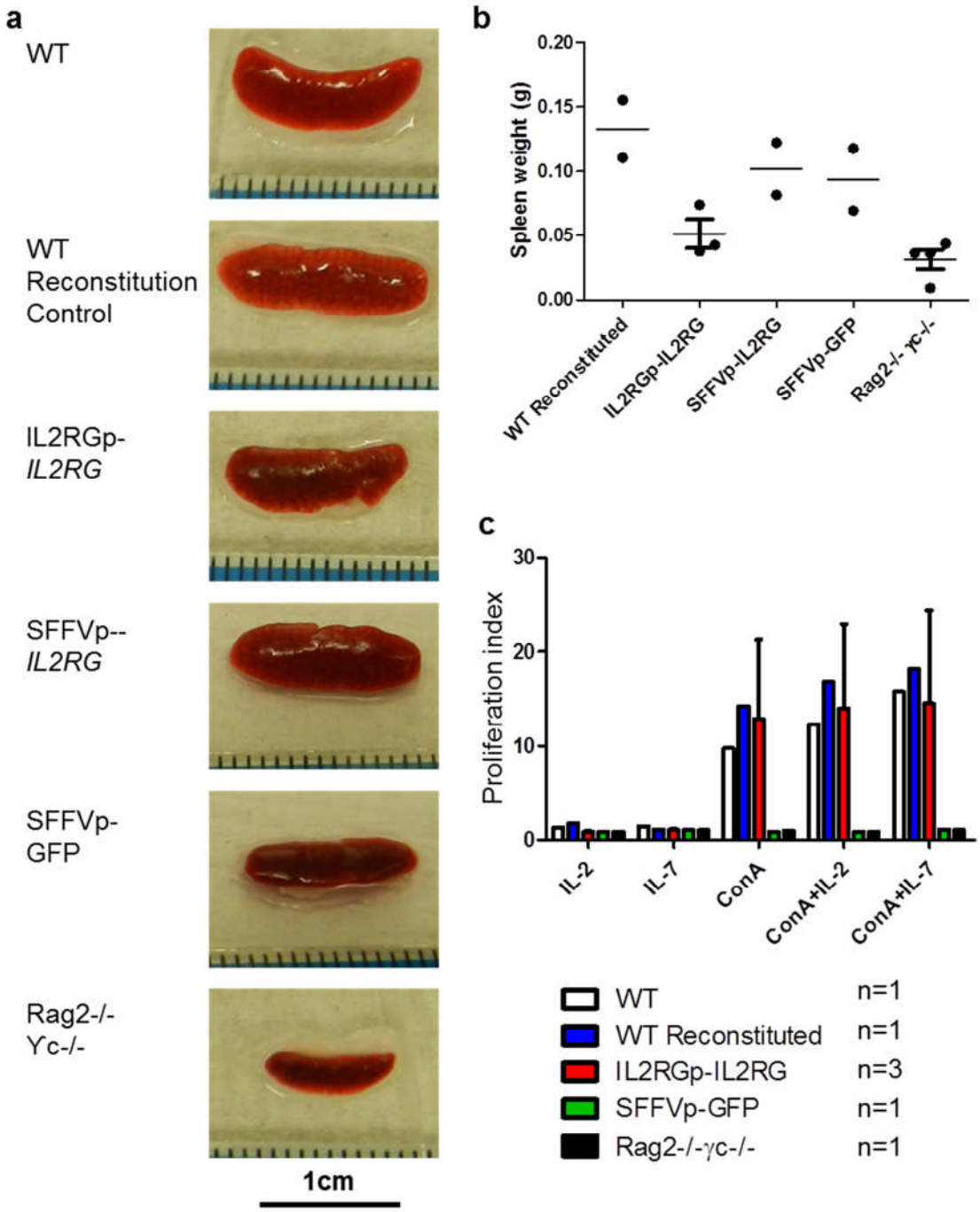


Figure 5.10 Splenic reconstitution.

(a) Examples of spleens dissected from $Rag2^{-/-}\gamma c^{-/-}$ mice treated with WT or IL2RGp-*IL2RG* , SFFVp-*IL2RG* and SFFVp-GFP transduced hγcmut lin^- cells and untreated $Rag2^{-/-}\gamma c^{-/-}$ mice with a scale bar. (b) The splenic weight of all the treated and control mice are compared in a scatter plot. (c) Splenocytes from $Rag2^{-/-}\gamma c^{-/-}$ mice treated with WT or IL2RGp-*IL2RG* and SFFVp-GFP transduced hγcmut lin^- cells and untreated $Rag2^{-/-}\gamma c^{-/-}$ controls at the sample number indicated (n) were incubated with mitogens IL-2 (20µg/ml) or IL-7 (20µg/ml) with or without Concanavalin A (1µg/ml). Proliferation index values above 1 denote splenocyte proliferation above that of media alone measured as counts per minute.

5.4.6 Serum immunoglobulin levels in reconstituted mice

The final analysis performed on the all of the reconstituted mice at sacrifice measured the immunoglobulin (Ig) levels in the serum as a measure of the functionality of the B cells (Figure 5.11). It has already been described that mature IgM+ B cells developed in the mice reconstituted with lin$^-$ cells containing a functional *IL2RG* gene whereas no IgM+ B cells developed in SFFVp-GFP control mice. To assess whether there was an increase in Ig types compared to untreated hγcmut mice, serum from these animals were also included as a control. From the data it can clearly be seen that all Ig types were detected at normal WT levels in mice treated with WT lin$^-$ cells and hγcmut lin$^-$ cells treated with either IL2RGp-IL2RG or SFFVp-IL2RG. With the exception of IgG2a, there was an increase in Ig levels in IL2RG treated mice compared to control and the increase of the total Ig analysed across the sample conditions revealed that this increase is significant. Not only does this imply successful B cell development, it also reveals the presence of functional T helper cells.

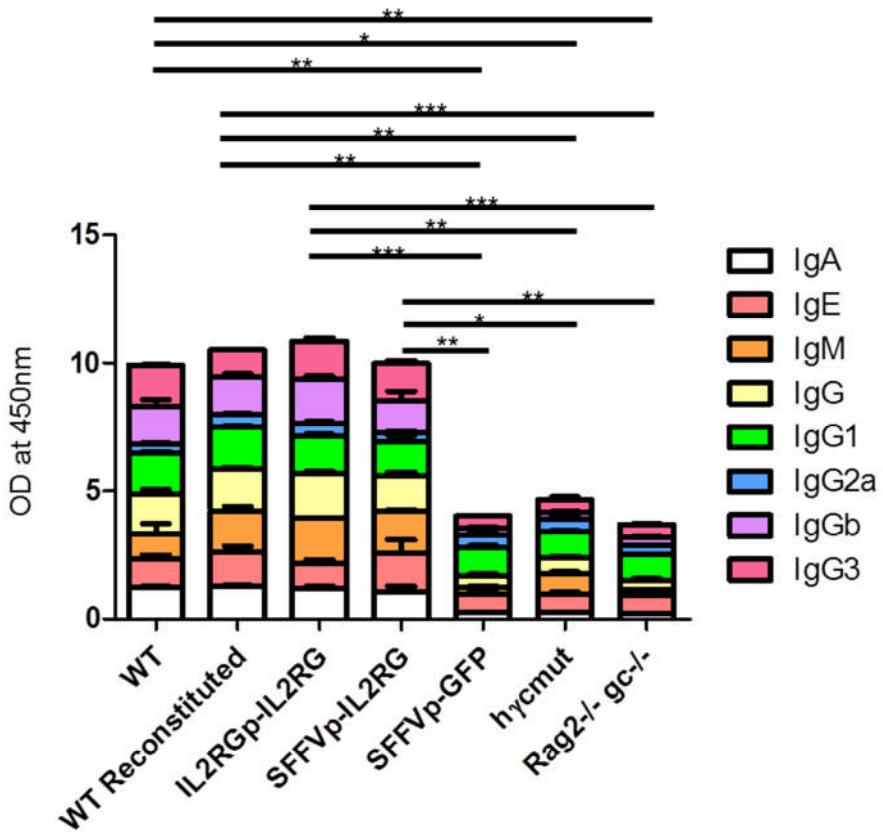


Figure 5.11 Serum immunoglobulin levels in reconstituted mice.

Serum immunoglobulin levels in Rag2$^{-/-}$γc$^{-/-}$ mice treated with WT or IL2RGp-*IL2RG*, SFFVp-*IL2RG* or SFFVp-GFP transduced hγcmut lin$^-$ cells and untreated WT, hγcmut and Rag$^{-/-}$γc$^{-/-}$ controls were run in triplicate and analysed by ELISA. The data have a normal distribution according to the Kolmogorov-Smirnov test and analysed by one way analysis of variance with a Bonferroni post-test (*= P-value<0.05) (**=P-value<0.01).(*** : P=<0.001).

5.5 Discussion

The hγcmut mouse has the same distribution of CD4+, CD8+, NK and IgM+ B cells as the $\gamma c^{-/-}$ mouse. The CD4+ cells are of normal levels as seen in the previous SCID-X1 mouse models (Cao, Shores et al. 1995; DiSanto, Muller et al. 1995; Ohbo, Suda et al. 1996) and are likely to cause the splenomegaly observed in these mice. The rudimentary thymi and the associated T cell maturation are also common to both the hγcmut mouse and the $\gamma c^{-/-}$ SCID-X1 mouse.

Although a direct comparison of Ig was not performed with the $\gamma c^{-/-}$ mouse, the similar distribution of Ig in the hγcmut mouse and the $Rag2^{-/-}\gamma c^{-/-}$ mouse highlights the B cell deficiency in the new model, contributing to its SCID phenotype.

Concerning the ability to reconstitute a $Rag2^{-/-}\gamma c^{-/-}$ mouse, these data imply that stem cells from the hγcmut mouse are not able to generate a functional immune system without the addition of *IL2RG* to the genome. However, of the 2 thymi extracted from the GFP treated control group, one did demonstrate that the stem cells of the hγcmut mouse can support the development of a rudimentary thymus as seen in the SCID-X1 models however, the pattern of T cell distribution was unlike that seen in the mice receiving *IL2RG* expressing cells. In the other control, the thymus did not have any CD4+ or CD8+ cells present. Furthermore, neither GFP control mice were able to generate CD4+ or CD8+ cells to reach normal levels in the peripheral blood, bone marrow or the spleen. Together with the Ig data, demonstrating no significant difference between untreated $Rag2^{-/-}\gamma c^{-/-}$ mice and those receiving GFP treated hγcmut stem cells, it can be speculated that hγcmut stem cells alone are not able to reconstitute a functional immune system. However a greater number of GFP negative control mice should be included to confirm this finding as significant.

The overall observations from mice receiving hγcmut cells treated with *IL2RG* expressing lentiviral vectors clearly demonstrate that this reconstitution strategy was successful when using both constructs. It is however interesting to note that the IL2RGp-*IL2RG* vector appears to be more effective in generating CD8+ cells and the SFFVp-*IL2RG* vector better at generating NK cells although normal levels of NK cells were not achieved in any of the tissues analysed, this is in accordance with a study

using SIN gammaretroviral vectors expressing *IL2RG* treated γc$^{-/-}$ progenitor cells to reconstitute Rag2$^{-/-}$γc$^{-/-}$c5$^{-/-}$ mice (Thornhill, Schambach et al. 2008) where NK cells constituted 1.8% of bone marrow in mice treated with an SFFVp-*IL2RG* vector, almost 2 fold less than that of normal. It is also important to note that a greater sample number would be required to determine a significant different between these vectors.

It is of interest that the level of viral transduction was at least 10 fold greater in the cells treated with the IL2RGp-*IL2RG* construct compared to SFFVp-*IL2RG* at 72 hours, and after 16 weeks of animal maintenance, the copy number of IL2RGp-*IL2RG* in all tissues analysed was both lower than the SFFVp-*IL2RG* copy number and consistent in each tissue. It is possible that this reduction in IL2RGp-*IL2RG* transgene was due to a high degree of mortality as a result of the high transduction rate. Another explanation could be that the IL2RG promoter was methylated more readily than the SFFV promoter, leading to the loss of *IL2RG* gene expression and moderating the survival advantage provided by γc. The SFFVp-*IL2RG* transgene however was found at variable levels in each tissue, the greatest in the blood and the lowest in the bone marrow. This may be a result of the regulatory elements additional to the EBS included within the promoter or intronic regions of *IL2RG* not included in the SFFVp-*IL2RG* construct. Alternatively, the divergent copy numbers in different tissues seen in cells treated with SFFV containing constructs may be due to clonal dominance as a result of activating growth promoting genes by the SFFV promoter which is capable of far reaching promoter activity (Ott, Schmidt et al. 2006).

The results from this chapter clearly demonstrate that the hγcmut mouse mirrors the phenotype of the γc$^{-/-}$ mouse and like the γc$^{-/-}$ mouse, hγc expression can be restored with integrating h*IL2RG* encoding lentiviral vectors. Importantly, it is also demonstrated in this chapter that the h*IL2RG* promoter and regulatory elements are effective at providing the transcription control of *IL2RG* in mice which is vital for the use of this model in gene targeting studies as the mutant human gene, if corrected *in situ*, will provide γc capable of reconstituting this SCID-X1 mouse model's immune system.

6 Correcting the humanised mouse model of SCID-X1 by ZFN mediated HR

6.1 Aims

- To optimise the efficiency of ZFN gene delivery to lineage negative cells by nucleofection
- To attempt ZFN mediated targeted HR in stem cells

6.2 Introduction

Having established that the *IL2RG* ZFN pair is capable of targeting the G691A mutation and thereby stimulating HR at this target site with extragenous donor, it is possible to apply this strategy to correct the novel humanised mouse model, hγcmut. This mouse model has been confirmed to contain the h*IL2RG* with the common SCID-X1 mutation G691A in replace of m*il2rg* and is immunodeficient. Furthermore, full length h*IL2RG* with its corresponding promoter and regulatory elements are capable of restoring γc expression as seen by the reconstitution of the immune cells on treatment of lin^- bone marrow cells from these animals with h*IL2RG* expressing lentiviral vectors.

IL2RG targeting in published work has been limited to the WT gene in human cell lines K-562 and LCLs (Urnov, Miller et al. 2005; Lombardo, Genovese et al. 2007; Moehle, Rock et al. 2007). The hγcmut mouse model provides an opportunity to correct mutant h*IL2RG in situ* and to assess the efficiency of ZFN mediated corrected HSC to repopulate the immune system with fully functional immune cells.

For such a therapeutic strategy to develop, it is necessary to correct lin^- HSC in order to provide the multipotent properties for haematological reconstitution. ZFN mediated gene targeting has been shown not to disrupt the pluripotent quality of stem cells *in vitro* (Soldner, Laganiere et al. 2011; Hoher, Wallace et al. 2012), and ZFN mediated HR in one cell rat and mouse embryos has resulted in heritable transgene (Cui, Ji et al. 2011), further providing evidence for the efficacy of this method regarding the correction of stem cells. It is also an important consideration that haematopoietic progenitors are resting in the G0 phase of the cell cycle (Reems and Torok-Storb 1995) and HR occurs in the S/G2 phase. To aid cell cycle progression and the conditions for HR in the this set of experiments, the cytokines used in the reconstitution experiment,

SCF, mFlt3 and hTPO were included in the medium for the lin$^-$ bone marrow cells after the ZFN and donor constructs were introduced.

An important factor to consider at this juncture is the method of ZFN gene and donor DNA delivery. As has been shown in this work, integrating lentiviral vector delivery has proven successful in both the delivery of a transgene to lin$^-$ cells and of ZFN genes to cell lines. However, as the purpose of ZFN mediated gene correction is to avoid the risks associated with integrating viral vectors and because constitutive ZFN expression could result in unwanted DSB events, this method of delivery is not included in this section. Furthermore, the alternative, IDLV ZFN vector did not provide detectable targeted HR in the ED7R cell line. Considering nucleofection meets the integration deficient criteria and was proven successful in the delivery of donor DNA and IL2RG ZFN expression plasmids to the ED7RγcΔ cell line as seen by targeted HR events at the *IL2RG* locus, this chapter explores the use of this method to achieve the correction of murine stem cells.

6.3 Nucleofection optimisation of lin$^-$ bone marrow cells.

To remain consistent with the reconstitution strategy of SCID-X1 mouse models, nucleofection was optimised for lin$^-$ bone marrow cells. However, murine lin$^-$ cells are not included in the primary cell repertoire optimised for nucleofection by Lonza. Lonza's technical support advised testing the murine macrophage nucleofection kit, in addition to this, the human CD34+ cell kit and the murine T cell kit were also included for optimisation.

Lin$^-$ cells were enriched from the bone marrow of the hγcmut mice by MACS, any biotin labelled differentiated cells remaining in the selection were analysed by flow cytometry by staining with Cy5 conjugated streptavidin (Appendix 13). 1x10^6 of the lin$^-$ enriched bone marrow cells were nucleofected according to the pulse programme recommended for each kit, with 2µg of the GFP expression vector, pMax GFP or were resuspended in the nucleofection solution without a pulse to establish the toxicity of the solution. The cells were incubated for 72 hours then analysed by flow cytometry for cell viability, as measured by the resistance to uptake of PI, and GFP expression (Figure 6.1). Nucleofection with the murine macrophage kit was performed using 3 different pulse programmes and the viability of the cells decreased in all of the pulsed cells compared to the no pulse control (Figure 6.1a). The greatest cell survival was seen

using the pulse programme A-023 resulting in GFP expression of 5.65%, however, this programme resulted in a >2 fold decrease in cell survival compared to no pulse. The greatest GFP expression observed for this kit was with the pulse programme Y-001 however, this programme also resulted in the most cell death.

The human CD34+ cell nucleofection kit had a greater cell survival compared to the murine macrophage kit and achieved 4.2% GFP positive cells (Figure 6.1b). The lin^- enriched cells used for this nucleofection included 43% biotin labelled cells (Appendix 13), indicating a poor purity of the MACS selection. The murine T cell nucleofection kit appeared to have the least toxic nucleofection solution, where 62.1% of the cells survived compared to 23.2% for the murine macrophage kit and 57.5% for the human CD34+ cell kit when no pulse was applied (Figure 6.1c). The pulse resulted in a 2.4 fold reduction in cell survival however the kit did result in the greatest GFP expression at 14% GFP positive cells. However, the lin^- cells prepared for this assay contained 53% biotin labelled differentiated cells.

It is important to consider the purity of the lin^- cells used in the nucleofection optimisation and whether the differentiated cells, such as T cells, were nucleofected at a greater rate than the lin^- cells. To overcome the low level of purity seen in these experiments it is possible to apply the MACS effluent to a second magnetic column. This can decrease the yield of lin^- cells yet increase the purity.

Considering the overall cell survival and GFP expression across the nucleofection kits tested, the murine T cell nucleofection kit was used for subsequent gene delivery to lin^- enriched bone marrow. The low rate of GFP+ cells seen in the nucleofected cells may have been due to loss of expression plasmids as the cells divided over the extensive incubation time of 72 hours. This incubation time was chosen due to the high cell morbidity of nucleofection and the personal observation that cell recovery is enhanced if left to rest for this time in replenished medium.

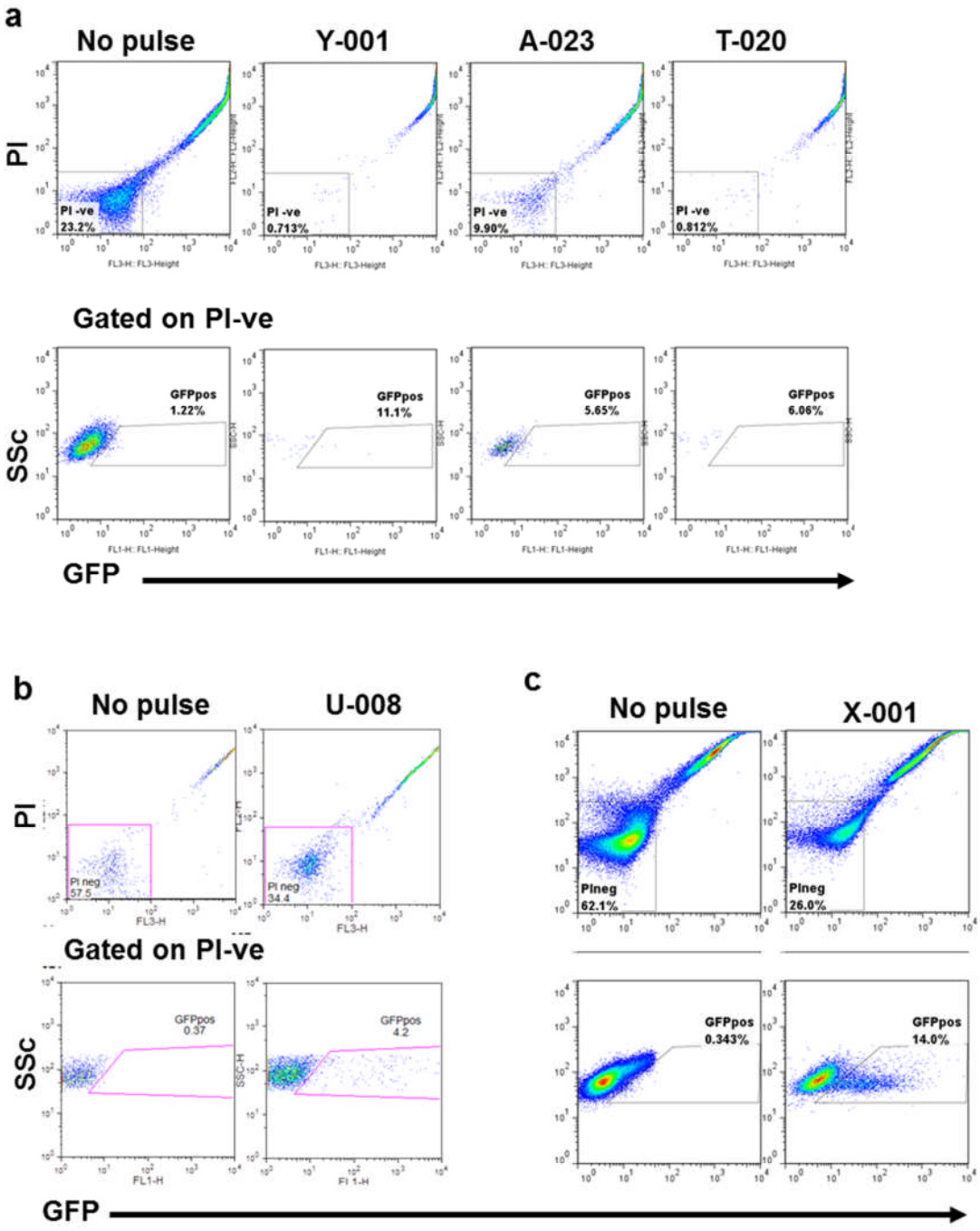


Figure 6.1 Optimising nucleofection of murine bone marrow cells.

$1x10^{6}$ lin^{-} enriched bone marrow from hgcmut mice was nucleofected with 2µg pMAX GFP expression plasmid and analysed by flow cytometry 72 hours post nucleofection for the viability stain Propidium Iodide (PI) and GFP expression in PI-ve cells. (a) Optimisation of the macrophage nucleofection kit with the voltage programme indicated. (b) Nucleofection with the human CD34+ cell kit. (c) Nucleofection with the T cell line kit.

6.4 ZFN mediated targeted HR in murine stem cells

Having established a nucleofection kit capable of delivering expression plasmids to lin^- enriched bone marrow cells, gene targeting by ZFN was assessed in these cells. A second source of stem cells was available from the hγcmut mice in the form of a murine embryonic stem (mES) cell line established during the development of the mouse (Appendix 14).

$1x10^6$ lin^- enriched bone marrow cells, obtained by MACS or $2x10^6$ mES cells were nucleofected per sample, with the murine T cell and mES nucleofection kit respectively. Both cell types were nucleofected with a total of 2µg of plasmid of the ZFN pair in equal measure, the donor DNA MB17 or the ZFN pair plus the donor DNA at a ratio of 3:1, a no pulse control was also included for both cell types. The ZFN to donor DNA ratio of 3:1 was chosen for this assay due to the observation that this ratio resulted in the most consistent target HR in ED7R cells, observed in chapter 4. The lin^- cells were incubated in StemSpan supplemented with SCF, mFlt3 and hTPO and the mES cells were incubated in ESGRO.

96 hours post nucleofection the genomic DNA was extracted and analysed by HR specific PCR (Figure 3.7) for the presence of the MB17 sequence by using primer pairs 1 or 2. A template control was also included whereby the forward and reverse primers which recognise exons 4 and 7 respectively and bind outside the region of donor DNA homology were used as a primer pair (primer pair 3) in a further PCR reaction (Figure 6.2a). The PCR did not detect targeted HR in lin^- enriched bone marrow cells nucleofected with the ZFN expression plasmids and donor DNA. The full length *IL2RG* expression plasmid containing the MB17 cassette, included as a positive control, did confirm correct PCR conditions as a faint 1,002bp band was seen for primer pair 1 and a more intense 969bp band present for primer pair 2. The DNA template for all conditions was also shown to be suitable by the 1.9kb band amplified by primer pair 3.

Targetted HR was observed in the mES cells when the cells were nucleofected with both the ZFN pair and donor DNA plasmids as demonstrated by the 1,002bp and 969bp bands for primer pair 1 and 2 respectively (Figure 6.2b). Again, primer pair 3 confirmed the presence of DNA template for all the nucleofection conditions, although the faint band for donor DNA may indicate a lower concentration of this template or the presence PCR inhibitors.

Having established the targeted HR events in the mES cells, RFLP analysis by Southern blot was carried out to determine the rate of HR in the cells treated with ZFN expression plasmids and donor DNA alongside the nucleofection controls and a positive control (Figure 6.2c). 100ng of purified template DNA, amplified from the genomic DNA by primer pair 3 was incubated with either BsrBI to detect the inclusion of this novel restriction enzyme site by target recombination with the donor plasmid MB17, or BsrBI and DraIII as a template control. Template prepared from the full length *IL2RG* expression plasmid containing the MB17 cassette was included as a positive control. No inclusion of the BsrBI restriction site was detected by RFLP analysis in any of the nucleofected samples and the assay conditions were shown to be suitable by the 78% digestion detected for the BsrBI positive control.

The RFLP analysis has been shown to be sensitive enough to detect HR events when BsrBI is present in >5% of total DNA template. Therefore it maybe speculated that the HR events detected by HR specific PCR did not reach the rate necessary for quantification by RFLP.

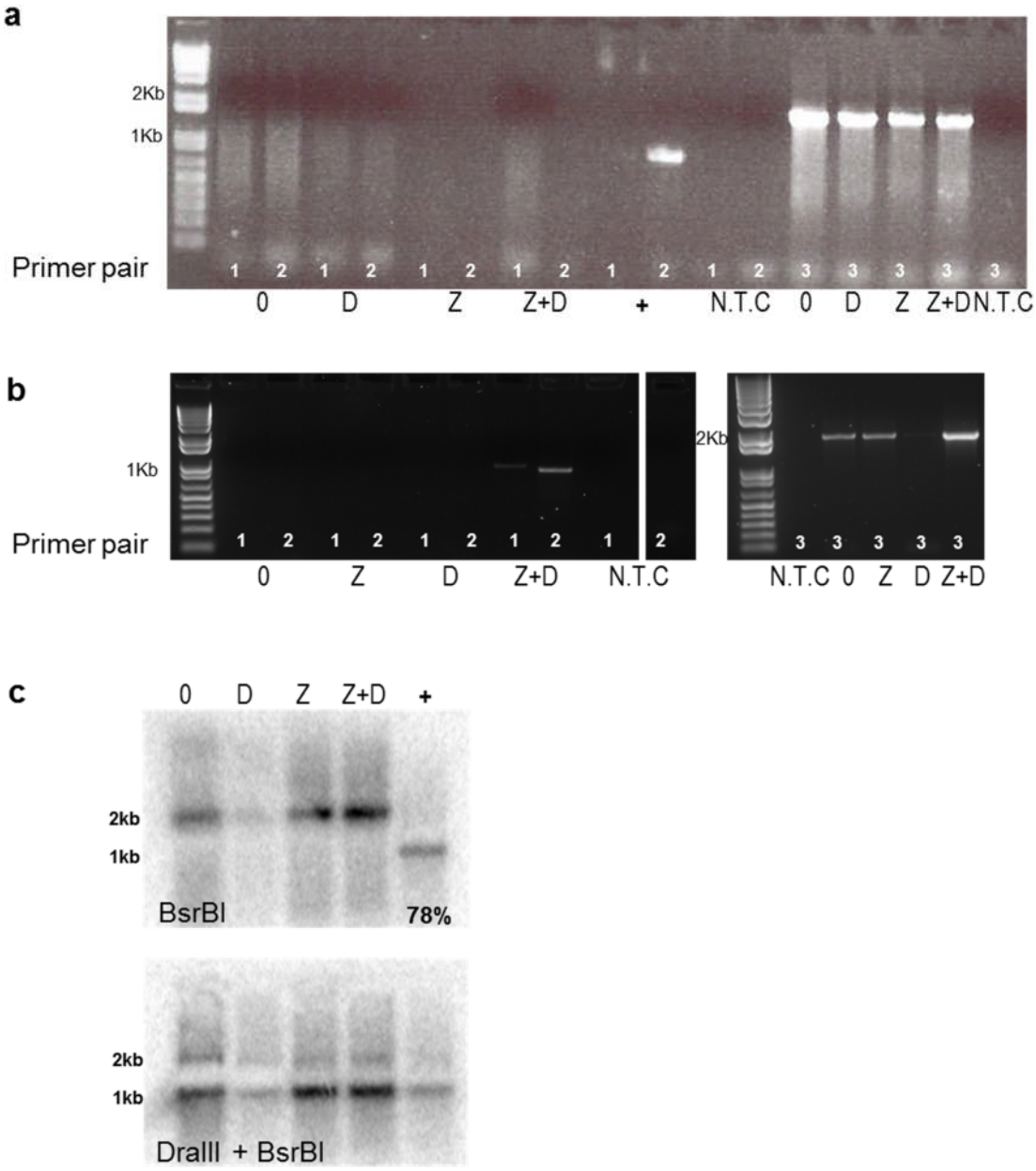


Figure 6.2 Targeted correction of human *IL2RG* in murine pluripotent cells.

96 hours after nucleofection with 2µg of ZFN and Donor constructs at a ratio of 3:1 and untreated, ZFN or donor DNA constructs alone for control. (a,b) HR specific PCR to detect the MB17 cassette at the genomic *IL2RG* locus was carried out with primer pairs 1 and 2. A full length IL2RG expression construct containing MB17 was included as a PCR control, A PCR with primer pair 3, designed to recognise exon 4 and 7 was also carried out to verify the template suitability. (a) $1x10^6$ lin^- enriched bone marrow. (b) $2x10^6$ ES. (c) RFLP analysis of the cells nucleofected with different ratios of ZFN to donor DNA capable of introducing a BsrBI restriction site at the IL2RG locus alongside BsrBI positive control (+) and BsrBI+DraIII template control. The numbers below the bands denote the percentage of the 250bp digestion product of total 500bp DNA template above background, calculated from the band intensity by ImageJ.

6.5 Discussion

The observations in this chapter have demonstrated the requirement for further optimisation of the delivery of ZFN expression vectors and donor DNA to stem cells. It has been demonstrated that targeted HR is possible in hγcmut mES cells however, in order for this correction strategy to progress it is necessary to demonstrate the correction of primary HSC *ex vivo* for the subsequent re-infusion of autologous cells.

More stringent methods must be employed to yield a pure fraction of lin^- cells. This will ensure that the gene targeting events do occur in multipotent cells. Another consideration for gene delivery to lin^- cells is the potential loss of mulipotency. It has been demonstrated that mES cells (Capecchi 2001) and ZFN treated human primary CD34+ HSC (Holt, Wang et al. 2010) retain pluripotency after nucleofection. However, it will be advantageous to culture colonies from treated lin^- cells to determine their stem cell properties and to assess the rate of target HR by high throughput sequencing.

Nucleofection was chosen as the mechanism for gene delivery to murine ES cells in this chapter and has proven successful for ZFN mediated targeted gene insertion in human ES cells, human iPS cells (Sebastiano, Maeder et al. 2011; Zou, Sweeney et al. 2011) and gene disruption in primary human CD34+ cells (Holt, Wang et al. 2010), this technique is not established for murine lin^- cells. In order to achieve an appropriate rate of gene delivery to these cells, other methods must be identified and optimised. The potential of other gene delivery methods is overviewed in the final discussion.

The ZFN:Donor DNA ratio that provided the most consistent evidence of targeted HR in 293T cells and ED7R cells was 3:1, as compared to 1:3 and 1:1. The most pronounced evidence for this observed in the nucleofection of ED7RγcΔ cells where γc+ cells were detected 12 days post nucleofection at this ratio in 3 of 4 of the treatment conditions whereas no evidence of correction was detected in the other ratios. Whether this method of gene delivery or ratio is optimal for murine lin^- cells is yet to be discerned. Future attempts at gene targeting in primary mouse cells would also benefit from a titration of ZFN vs Donor vectors as increasing dosage of ZFN can inhibit HR (Connelly,

Barker et al. 2010; Handel, Gellhaus et al. 2012) and observations from the ED7R cell line may not be applicable in the murine lin^- cells.

The data in this chapter confirms that the IL2RG ZFN are capable of targeted HR in hγcmut mES cells, thereby confirming the utility of these animals in the future of gene targeting therapeutic strategies for SCID-X1. And in addition to the optimisation of the ZFN:Donor ratio and delivery strategy, future experiments should also compare the WT, MB1 and MB3 donor constructs designed for *IL2RG* correction. It is possible to regenerate multilineage hematopoietic engraftment in mice with mES cells modified to ectopically express the Homeodomain gene *HoxB4* and the Homeobox domain transcription factor *Cdx4* which modulates *Hox* expression (Wang, Yates et al. 2005). Therefore it could be possible to utilise this method of regeneration to investigate the principal that the corrected hγcmut mES cells are capable of reconstituting the immune system of an immunodeficient mouse.

7 Final discussion

In the case of SCID-X1, when a suitably matched donor is not available for a bone marrow transplant, it is possible to cure the disease by the addition of *IL2RG* via a retroviral vector into the genome of the patient's haematopoietic stem cells. This gene addition strategy results in integration of *IL2RG* and its promoter with limited predictability. This may result in insertional mutagenesis due to off target gene regulation and gene disruption. An alternative strategy for SCID-X1 treatment is to correct the gene at the site of the mutation. *In situ* resolution will permit endogenous expression and reduce the risk associated with retroviral vector therapy. Gene correction can be instigated by stimulating HR with a DSB introduced by a ZFN pair.

Previous reports have demonstrated that the ZFN designed to target *IL2RG* can mediate targeting of WT *IL2RG* in K-562 cells and a lymphoblastic cell line. This project has observed IL2RG ZFN sequence specific activity in additional cell lines, 293T cells, ED7R cells and mES cells. Furthermore, the IL2RG ZFN pair has been shown to target a common SCID-X1 mutation. This mutation is included in a humanised mouse model designed for *in situ* correction studies and this work has confirmed the SCID-X1 mouse model to have an immunodeficient phenotype which can be resolved by the addition of full length human *IL2RG* with its endogenous promoter to their lin- bone marrow cells.

This study has shown that the ZFN pair designed to target *IL2RG* is capable of cleaving both WT *IL2RG* and the SCID-X1 mutation hotspot, G691A. Due to the nature of HR and the potential to convert DNA sequence up to 511bp from the initiation site (Elliott, Richardson et al. 1998), this IL2RG ZFN pair could, in theory, correct a host of mutations within this distance from the target providing the mutation does not occur within the ZFN binding site. Such a mutation is the G691A neighbouring mutation, C690T, which is also common (Puck, Pepper et al. 1997). It has not yet been demonstrated whether this ZFN pair can target this mutation and it would be of interest to investigate whether the IL2RG ZFN is capable of binding and cutting sequence containing this mutation. A further genome editing strategy could be employed, regardless of SCID-X1 associated mutation, which would be to introduce the *IL2RG* gene at the safe harbour sites *CCR5* or *AAVS1* using the corresponding ZFN pairs.

However, the potential to correct the SCID-X1 mutation *in situ* would have the benefits of physiological expression and no unforeseen consequences of disruption at the *CCR5* and *AAVS1* locus would be challenged.

During this study, the detection of HR was limited in the majority of cases to targeted HR specific PCR which could detect recombination events that remained undetectable by RFLP and flow cytometry. Unfortunately this method does not provide a means to quantify the rate of recombination. Furthermore, the PCR product generated by primer pair 1 and 2 did not result in consistent band intensity. This could be due to the efficiency of the individual primer pair however, the primer pair generating the band of greater intensity alternated from the more intense band seen in the 293T cells and ED7RγcΔ cells to the less intense band observed in mES cells treated with ZFN and or Donor. A possible explanation for this is the presence of contaminants in the PCR reaction that inhibited the efficiency of the reaction. It would be beneficial to overcome the limitations for sensitive quantification by analysing clones from the treated cells by high throughput sequencing as not only would this quantify the rate of correction, it would also identify the mechanism of correction, including unilateral SDSA or bilateral DSBR mechanism of HR and allow for the comparison with NHEJ events at the DSB. This may also serve an explanation for the variable primer binding efficiency for if the unilateral SDSA mechanism of HR was employed, and half of the primer binding site was recombined, the target site may have been recognised, but at a decreased efficiency.

The comparison of 4 donor DNA constructs designed to correct the G691A mutation which differed in sequence from the target site by an increasing number of base pairs did not confirm the hypothesis that divergent sequence at the site of HR initiation would inhibit strand invasion. Nor was it shown that the WT donor, with an intact ZFN target site, would be cleaved by ZFN and therefore unsuitable, as this donor DNA generated the greatest percentage of corrected ED7RγcΔ cells as observed by flow cytometry. However, these data have not been repeated or verified by another form of quantifiable analysis.

Over the course of this study, targeted HR was demonstrated in 293T cells, ED7RγcΔ cells and mES cells. All of these examples are cell lines, which are liable to accrue

transformation events during extensive periods of cell culture *in vitro* including mES cells which can accumulate copy number variations (Liang, Conte et al. 2008) and other chromosomal abnormalities reviewed by Rebuzzini et al. (Rebuzzini, Neri et al. 2008). In order for the optimal assessment of gene targeting strategies, the target locus should be representative of the physiological genome. In this study, the ED7R genome was shown to contain at least 2 copies of the IL2RG locus which does not represent the male SCID-X1 patient. Furthermore, the necessary addition of the h*IL2RG*Δ transgene increased the number of ZFN target sites at unknown locations in the genome. These conditions are not ideal for accurate assessment of ZFN targeting efficiency as multiple target sites could cause multiple DSB events thereby unnecessarily increasing these toxic events experienced by the cell. Furthermore, the analysis of gene correction is hindered by the presence of ZFN targets that will not result in functional protein as a result of targeted HR. Such an obstacle has been overcome with the development of the hγcmut mouse which provides a single copy of the ZFN target in male cells and also provides a source of mulipotent stem cells.

The long term goal of this project is to correct haematopoietic stem cells from disease patients *ex vivo* and return the corrected progenitor cells to re-populate the immune-system with fully functioning cells. The humanised hγcmut mouse has been confirmed to contain the h*IL2RG*Δ gene in place of m*il2rg* and has been characterised as immunodeficient in this study. Furthermore, an immunodeficient phenotype can be reversed in lin^- cells from these mice following transduction with vectors carrying the full length h*IL2RG* with its endogenous promoter. This novel SCID-X1 mouse model will be a major source of information on whether targeted gene correction strategies for *IL2RG* can restore healthy phenotypes. Initial observations pertaining to the IL2RG ZFN stimulated HR events at the target locus in hγcmut mES cells further support the advances that can be made with this model concerning genome editing of *IL2RG* in murine stem cells.

An outstanding problem still remains in that no targeted HR has been observed in the lin^- bone marrow cells which model the HSC used in autologous BMT in SCID-X1 patients. This study has begun initial gene delivery strategies in the form of nucleofection, but poor lin^- purity, high cell mortality, low transduction efficiency and the absence of targeted HR highlight the need for optimisation of this strategy.

Firstly, the most suitable gene delivery mechanism for lin$^-$ cells must be established. As previously discussed in chapter 6, it is not justified to use integrating proficient lentiviral vectors. Although nucleofection is a popular method of DNA delivery to mES cells (Capecchi 2001) a comparative study demonstrated that lipofection achieved a greater rate of transfection in these cells (Osiak, Radecke et al. 2011). Integrase deficient viral vectors can also be utilised to deliver ZFN and donor DNA to stem cells as observed by the targeted gene integration of 6% of human neural stem cells (Lombardo, Cesana et al. 2011). Adenoviral vector (Ad), Ad5/35, which is naturally non-integrating, has successfully delivered ZFN to primary T lymphocytes (Perez, Wang et al. 2008; Lombardo, Cesana et al. 2011) and the CCR5 safe harbour in human Mesenchymal stromal cells, achieving stable expression of the transgene in 40% of cells without affecting the plasticity of the cell (Benabdallah, Allard et al. 2010). However this delivery method has not been directly compared to IDLV.

AAV vectors have been developed by deleting the gene for the *rep* protein which mediates HR at the AAVS1 site. An AAV vector was developed to target hepatocytes and successfully delivered both ZFN and donor DNA to the liver of *f9* deficient mice via I.P. injection resulting in long term targeted correction (Li, Haurigot et al. 2011). An AAV integration profile was then carried out, revealing genome wide integration events had occurred in genes claimed not to be oncogenic. This is in contrast to hepatocellular carcinoma that has been observed in mice due to insertional mutagenesis caused by the integration of AAV vectors via regions of microhomology (Li, Haurigot et al. 2011). However, an AAV associated risk of insertional mutagenesis has not been observed in human cells.

An advantage to using AAV vectors is their ability to recombine with host chromatin. Two independent studies of the hypoxanthine-guanine phosphoribosyltransferase locus in the HT1080 cell line estimated targeted HR at a rate of 10^{-6} for Ad viral vectors, similar to plasmid DNA (Stephen, Sivanandam et al. 2008) and 10^{-3} for AAV vectors (Russell and Hirata 1998). However both vectors have shown a rate of heterologous integration up to 10 fold more than HR. It has been observed that ZFN mediated DSB can increase the rate of targeted HR by AAV donor DNA in hES cells up to 15 fold and it was estimated that background integration occurred at 1 AAV genome per 100 cells (Asuri, Bartel et al. 2012). Background integration has also been reported in

approximately 0.04% human CD34+ cells by IDLV (Lombardo, Genovese et al. 2007) and remains a concern regarding the safety issues surrounding genome modification.

Other limitations of AAV delivery include the inefficiency of multipotent cell transduction (Liu, Yan et al. 2004) and the high MOI needed to achieve transduction which can inhibit HR events (Handel, Gellhaus et al. 2012). This can, in part, be overcome by combining the ZFN pair in one vector (Sollu, Pars et al. 2010) or by combining the a ZFN and the donor DNA as previously described for IDLV (Lombardo, Genovese et al. 2007). The benefit of this is the reduced transduction events experienced by a cell. Possible drawbacks include the possibility of unequal ZFN expression and in the case of ZFN-Donor constructs, less flexibility in optimising the most appropriate ratio of ZFN to donor. However, a viral vector could be designed for clinical use once these conditions are proved optimal.

The problem of background integration of episomal DNA could prove deleterious if either the ZFN vector is randomly integrated resulting in constitutive ZFN expression or whether elements of either the ZFN or donor DNA vector were introduced to sites that could cause insertional mutagenesis. To avoid the risk of ZFN vector integration, it is possible to deliver ZFN mRNA via nucleofection which has proven effective in gene insertion at the AAVS1 safe harbour site in human iPS cells (Zou, Sweeney et al. 2011), this transient expression also avoids the risk of background integration of the ZFN expression vector in to the host genome which would be deleterious as over expression of ZFN increases the risk of off target DSB. Another promising method of transient ZFN delivery which demonstrates a decrease in off-site ZFN activity is by cellular uptake of the ZFN protein itself which has proven successful in targeted gene disruption in primary human dermal fibroblasts and CD4+ T cells. This method of delivery was shown to be comparable to plasmid transfection of ZFN at the CCR5 loci in HEK293 cells with targeting frequencies of 13% for direct protein and 15% for plasmid transfection but resulted in fewer events at predicted off target sites (Gaj, Guo et al. 2012).

Although direct application of ZFN protein to cells can result in gene targeting, in order to achieve HR, donor DNA must be provided. Furthermore, this strategy is optimal following repeated treatment over 3 days and at hypothermic, 30°C conditions, where

ZFN activity is increased. This extensive “cold shock” could prove deleterious to primary cells treated *ex vivo*. The activity of ZFN at 37°C has been increased by selecting an obligate heterodimer FokI nuclese pair with the greatest performance at 37°C (Doyon, Vo et al. 2011) and by the development of the more enzymatically active *Sharkey* FokI catalytic domain (Guo, Gaj et al. 2010).

To limit the the possible integration of unwanted elements from the donor DNA construct, it is possible to create plasmid DNA with no bacterial backbone by generating minicircles whereby the bacterial features of a plasmid are excised by intramolecular recombination (Chen 2003). These minicircles can be transfected into cells and could provide the correction template for HR. Another donor DNA design which avoids the presence of superfluose DNA are single stranded oligodeoxynucleotides (ssODN). Gene correction via ZFN induced DSB and ssODN donor DNA has achieved gene conversion in 22-23% of K-562 cells (Chen, Pruett-Miller et al. 2011). However another study observed unfaithful DNA repair whereby the ssODN was added between the DSB termini rather than acting as the template for HR (Radecke, Radecke et al. 2010) thus limiting their use for gene correction strategies. HR can be enhanced by tethering a dsDNA molecule to a triplex-forming oligonucleotide (TFO) (Chan, Lin et al. 1999). TFOs form a triplex with dsDNA by forming Hoogsteen hydrogen bonds with polypurine or polypyramidine tracts in the genome in a sequence specific manner. These could potentially serve to aid the localisation of donor DNA to a targeted DSB. Although the TFO site would be outside of the donor DNA homology region, a potential drawback to this approach could be interference of the TFO with the formation of the Holliday junction at the DSB site.

Although the focus of this study has been on gene targeting by ZFN, these are not the only reagent capable of mediating a sequence specific DSB. As breifly mentioned in the introduction, TALENs are a promising alternative to ZFN and due to their simple assembly, researchers can construct their own custom TALENs which has aided their populatrity. TALENs have successfully achieved gene targeting in human ES cell with a similar rate of efficiency to their ZFN counterparts and a high degree of specificity yet they are also capable of off target activity (Hockemeyer, Wang et al. 2011). TALENs have not, as yet, superseded ZFN in *IL2RG* targeting efficiency (Mussolino, Morbitzer et

al. 2011). A more recent development for mediating a sequence specific DSB is by clustered regularly interspaced short pallendromic repeats (CRISPR). CRISPR technology is based on an RNA directed Cas9 nuclease from *Streptococcus pyogenes*. The nuclease guiding sequences in the non-coding RNA can be modified to correspond to a target gene in eukaryotes with the same specificity as TALENs, as observed for the *EMX1* locus (Cong, Ran et al. 2013). However, constraints due to RNA processing and stability limit the range of possible targets. The Cas9 nuclease can also be modified to cut a single strand rather than introduce a DBS, and so called a nickase.

The single stranded nick achieved by nickases can be repaired seamlessly and rarely results in NHEJ however the with a lower rate of HR (Certo, Ryu et al. 2011). The FokI nuclease domain of a ZFN or TALEN can also be engineered to become a nickase. In a direct comparison ZFN ahieved targeted HR at a rate 3-10 fold higher than its ZFNickase counterpart (Ramirez, Certo et al. 2012).

Considering the progress made in genome engineering, the IL2RG ZFN is the most efficient reagent for *IL2RG* targeting so far. This leaves and structure of donor DNA and the method of delivery currently remaining for optimisation. The safest method of delivery would, in theory, avoid the end use of any promoter/ enhancer elements, viral or otherwise, that could randomly integrate at unpredictable sites in the genome. Circular DNA ensures the donor DNA will not be wholly integrated by DSB end capture which can occur with linear DNA. The ZFN expression must be well tolerated by the host cell to allow DSB repair to progress rather than cell death in response to the toxic nature of DSB. The greatest way to achieve this is by the administration of ZFN protein itself, thus ensuring transient expression with no risk of ZFN gene integration.

In developing an *IL2RG* correction strategy the method and dosage of reagent delivery must be optimised. In the face of the challenges involved in gene targeting it is important to note that different cells respond differently to DSB, resulting in variable rates of DNA repair and it would be advisable to use the most relevant source of cell for optimisation. As the purpose of developing such a multifaceted gene therapy strategy is to improve the safety of gene manipulation it is also necessary to monitor off target integration of the delivery vectors and employ a method that minimises these events. Off target IL2RG ZFN activity has been observed in K-562 cells (Gabriel, Lombardo et al. 2011). However, it is important to consider the state of the chromatin profile and therefore the accessibility of sequence specific endonucleases to their

target sequences which may differ between a human derived cell line and primary human multipotent cells. A limitation of the hγcmut murine model, due to their murine genome, is the prediction of off target integration and DSB activity that may occur as result of targeted gene correction treatment in humans. Therefore it would be beneficial to monitor off site activity in human haematopoietic progenitor cells. An additional measure of the safety of this strategy could include transformation assays adapted from those used to assess retroviral gene insertion (Modlich, Bohne et al. 2006) to help predict whether ZFN/Donor constructs are potentially oncogenic. Proof of safety combined with efficient targeted HR will make ZFN mediated gene correction a viable therapy strategy for the future.

In addition to the efficiency of ZFN mediated HR and the safety profile of this reagent, additional features learnt from the retroviral treatment of SCID-X1 must be considered.

The clinical trials of SCID-X1 with retroviral vectors inluded the *ex vivo* treatment of HSC over 4 days to stimulate the progression of the cell cycle in order to achieve gammaretroviral transduction (Cavazzana-Calvo 2000; Gaspar, Parsley et al. 2004). Due to the nature of HR which occures in the S/G2 phase of the cell cycle, it will be necessary to stimulate cell cycle progression during the *ex vivo* treatment of the cells for targeted correction by HR. Although detectable HR events did not occure in the lin^- cells stimulated with SCF, mFlt3 and hTPO, this may have been due to the gene delivery technique yet optimisation of cell cycle progression for HR should be considered in future HSC correction strategies. A drawback in promoting stem cell progression is the relationship seen between lineage committed cells and the rate of cycling where cells in the S/G2 and M phase of the cell cycle are often committed progenitors (Gothot, Pyatt et al. 1997).

Another observation made in a clinical trial for the treatment of SCID-X1 is the age at which gene therapy will prove effective at regenerating a functional immune system. 2 patients aged 15 and 20 years old underwent retroviral gene therapy with the same protocol for the infants in the SCID-X1 therapy clinical trials however the T cell compartment remained lymphopenic (Thrasher, Hacein-Bey-Abina et al. 2005). This was likely to be due to the inability to regenerate a functional thymus as has been seen

in adult mice which were unable to reverse the disturbed architecture of the thymic tissue that had developed without the appropriate signalling by thymocytes (van Ewijk, Hollander et al. 2000). It is therefore of great importance that long-term SCID-X1 therapy is administered as soon as viably possible after diagnosis and this includes gene correction strategies.

Even though the rate of HR stimulated by ZFN remains lower than gene addition mediated by retroviral transduction, with advances in stem cell therapy, a small number of stem cells can theoretically repopulate the missing popultaions on immune cells as seen in the T cell progenitor revertent which repopulated the T cells of a SCID-X1 patient (Stephan, Wahn et al. 1996). Considering the survival advantage confered by γc expression by corrected cells in a SCID-X1 patient, ZFN mediated correction of stem cells holds great potential for this disease. Once the correction of *IL2RG* in human multipotent haematopoetic stem cells has been optimised with a diminished associated risk confirmed, the gene correction strategy for SCID-X1 will be a suitable candidate for therapy.

This project has continued the assessment of IL2RG ZFN and has confirmed their cleavage of a SCID-X1 associated mutation. A SCID-X1 mouse model has also been described which is suitable for ZFN mediated correction of *IL2RG* and will provide a platform for the assessment of ZFN treatment of primary stem cells and subsequent *in vivo* studies necessary to confirm the suitability of *in situ* correction of *IL2RG* for use in clinical trials.

8 Reference List

Allers, T. and M. Lichten (2001). "Differential timing and control of noncrossover and crossover recombination during meiosis." Cell **106**(1): 47-57.

Alwin, S., M. B. Gere, et al. (2005). "Custom zinc-finger nucleases for use in human cells." Molecular therapy : the journal of the American Society of Gene Therapy **12**(4): 610-617.

Anderson, M. K., G. Hernandez-Hoyos, et al. (1999). "Precise developmental regulation of Ets family transcription factors during specification and commitment to the T cell lineage." Development **126**(14): 3131-3148.

Antoine, C., S. Muller, et al. (2003). "Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99." **361**(9357): 553-560.

Asao, H., C. Okuyama, et al. (2001). "Cutting edge: the common gamma-chain is an indispensable subunit of the IL-21 receptor complex." Journal of immunology **167**(1): 1-5.

Asuri, P., M. A. Bartel, et al. (2012). "Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells." Molecular therapy : the journal of the American Society of Gene Therapy **20**(2): 329-338.

Ayares, D., L. Chekuri, et al. (1986). "Sequence homology requirements for intermolecular recombination in mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **83**(14): 5199-5203.

Baum, C. (2007). "What are the consequences of the fourth case?" Molecular therapy : the journal of the American Society of Gene Therapy **15**(8): 1401-1402.

Baum, C., K. Itoh, et al. (1997). "The potent enhancer activity of the polycythemic strain of spleen focus-forming virus in hematopoietic cells is governed by a binding site for Sp1 in the upstream control region and by a unique enhancer core motif, creating an exclusive target for PEBP/CBF." Journal of Virology **71**(9): 6323-6331.

Beerli, R. R. and C. F. Barbas, 3rd (2002). "Engineering polydactyl zinc-finger transcription factors." Nature biotechnology **20**(2): 135-141.

Benabdallah, B. F., E. Allard, et al. (2010). "Targeted gene addition to human mesenchymal stromal cells as a cell-based plasma-soluble protein delivery platform." Cytotherapy **12**(3): 394-399.

Berenson, R. J., R. G. Andrews, et al. (1988). "Antigen CD34+ marrow cells engraft lethally irradiated baboons." The Journal of clinical investigation **81**(3): 951-955.

Beumer, K., G. Bhattacharyya, et al. (2006). "Efficient gene targeting in Drosophila with zinc-finger nucleases." Genetics **172**(4): 2391-2403.

Bhatia, M., J. C. Wang, et al. (1997). "Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice." Proceedings of the National Academy of Sciences of the United States of America **94**(10): 5320-5325.

Bibikova, M., K. Beumer, et al. (2003). "Enhancing gene targeting with designed zinc finger nucleases." Science **300**(5620): 764.

Bibikova, M., D. Carroll, et al. (2001). "Stimulation of homologous recombination through targeted cleavage by chimeric nucleases." Molecular and cellular biology **21**(1): 289-297.

Bibikova, M., M. Golic, et al. (2002). "Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases." Genetics **161**(3): 1169-1175.

Bitinaite, J., D. A. Wah, et al. (1998). "FokI dimerization is required for DNA cleavage." Proceedings of the National Academy of Sciences of the United States of America **95**(18): 10570-10575.

Blancafort, P., D. J. Segal, et al. (2004). "Designing transcription factor architectures for drug discovery." Molecular pharmacology **66**(6): 1361-1371.

Boch, J., H. Scholze, et al. (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors." Science **326**(5959): 1509-1512.

Boehm, T., L. Foroni, et al. (1991). "The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13." Proceedings of the National Academy of Sciences of the United States of America **88**(10): 4367-4371.

Bozas, A., K. J. Beumer, et al. (2009). "Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila." Genetics **182**(3): 641-651.

Buckley, R. H. (2004). "Molecular defects in human severe combined immunodeficiency and approaches to immune reconstitution." Annual review of immunology **22**: 625-655.

Buckley, R. H., S. E. Schiff, et al. (1999). "Hematopoietic stem-cell transplantation for the treatment of severe combined immunodeficiency." The New England journal of medicine **340**(7): 508-516.

Candotti, F., S. A. Oakes, et al. (1997). "Structural and functional basis for JAK3-deficient severe combined immunodeficiency." Blood **90**(10): 3996-4003.

Cao, X., C. A. Kozak, et al. (1993). "Characterization of cDNAs encoding the murine interleukin 2 receptor (IL-2R) gamma chain: chromosomal mapping and tissue specificity of IL-2R gamma chain expression." Proceedings of the National Academy of Sciences of the United States of America **90**(18): 8464-8468.

Cao, X., E. W. Shores, et al. (1995). "Defective lymphoid development in mice lacking expression of the common cytokine receptor gamma chain." Immunity **2**(3): 223-238.

Capecchi, M. R. (2001). "Generating mice with targeted mutations." Nature medicine **7**(10): 1086-1090.

Carbery, I. D., D. Ji, et al. (2010). "Targeted genome modification in mice using zinc-finger nucleases." Genetics **186**(2): 451-459.

Cathomen, T. and J. K. Joung (2008). "Zinc-finger nucleases: the next generation emerges." Molecular therapy : the journal of the American Society of Gene Therapy **16**(7): 1200-1207.

Cavazzana-Calvo, M. (2000). "Gene Therapy of Human Severe Combined Immunodeficiency (SCID)-X1 Disease." Science **288**(5466): 669-672.

Cavazzana-Calvo, M., C. Lagresle, et al. (2005). "Gene therapy for severe combined immunodeficiency." Annual review of medicine **56**: 585-602.

Cavazzana-Calvo, M., E. Payen, et al. (2010). "Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia." Nature **467**(7313): 318-322.

Certo, M. T., B. Y. Ryu, et al. (2011). "Tracking genome engineering outcome at individual DNA breakpoints." Nature methods **8**(8): 671-676.

Challita, P. M. and D. B. Kohn (1994). "Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo." Proceedings of the National Academy of Sciences of the United States of America **91**(7): 2567-2571.

Chan, P. P., M. Lin, et al. (1999). "Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide." The Journal of biological chemistry **274**(17): 11541-11548.

Chatila, T., E. Castigli, et al. (1990). "Primary combined immunodeficiency resulting from defective transcription of multiple T-cell lymphokine genes." Proceedings of the National Academy of Sciences of the United States of America **87**(24): 10033-10037.

Check, E. (2005). "Gene therapy put on hold as third child develops cancer." Nature **433**(7026): 561.

Chen, F., S. M. Pruett-Miller, et al. (2011). "High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases." Nature methods **8**(9): 753-755.

Chen, Z. (2003). "Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo." Molecular Therapy **8**(3): 495-500.

Choo, Y. and A. Klug (1994). "Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions." Proceedings of the National Academy of Sciences of the United States of America **91**(23): 11168-11172.

Choo, Y., I. Sanchez-Garcia, et al. (1994). "In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence." Nature **372**(6507): 642-645.

Civin, C. I., G. Almeida-Porada, et al. (1996). "Sustained, retransplantable, multilineage engraftment of highly purified adult human bone marrow stem cells in vivo." Blood **88**(11): 4102-4109.

Civin, C. I., T. Trischmann, et al. (1996). "Highly purified CD34-positive cells reconstitute hematopoiesis." Journal of clinical oncology : official journal of the American Society of Clinical Oncology **14**(8): 2224-2233.

Colleaux, L., L. D'Auriol, et al. (1988). "Recognition and cleavage site of the intron-encoded omega transposase." Proceedings of the National Academy of Sciences of the United States of America **85**(16): 6022-6026.

Cong, L., F. A. Ran, et al. (2013). "Multiplex genome engineering using CRISPR/Cas systems." Science **339**(6121): 819-823.

Connelly, J. P., J. C. Barker, et al. (2010). "Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease." Molecular therapy : the journal of the American Society of Gene Therapy **18**(6): 1103-1110.

Cornu, T. I. and T. Cathomen (2007). "Targeted genome modifications using integrase-deficient lentiviral vectors." Molecular therapy : the journal of the American Society of Gene Therapy **15**(12): 2107-2113.

Cradick, T. J., G. Ambrosini, et al. (2011). "ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites." BMC bioinformatics **12**: 152.

Cradick, T. J., K. Keck, et al. (2010). "Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs." Molecular therapy : the journal of the American Society of Gene Therapy **18**(5): 947-954.

Cui, X., D. Ji, et al. (2011). "Targeted integration in rat and mouse embryos with zinc-finger nucleases." Nature biotechnology **29**(1): 64-67.

Curtin, S. J., D. F. Voytas, et al. (2012). "Genome Engineering of Crops with Designer Nucleases." Plant Gen. **5**(2): 42-50.

Dave, U. P., K. Akagi, et al. (2009). "Murine leukemias with retroviral insertions at Lmo2 are predictive of the leukemias induced in SCID-X1 patients following retroviral gene therapy." PLoS genetics **5**(5): e1000491.

de Saint Basile, G., B. Arveiler, et al. (1987). "Close linkage of the locus for X chromosome-linked severe combined immunodeficiency to polymorphic DNA markers in Xq11-q13." Proceedings of the National Academy of Sciences of the United States of America **84**(21): 7576-7579.

DeKelver, R. C., V. M. Choi, et al. (2010). "Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome." Genome research **20**(8): 1133-1142.

DiSanto, J. P., D. Guy-Grand, et al. (1996). "Critical role for the common cytokine receptor gamma chain in intrathymic and peripheral T cell selection." The Journal of experimental medicine **183**(3): 1111-1118.

DiSanto, J. P., W. Muller, et al. (1995). "Lymphoid development in mice with a targeted deletion of the interleukin 2 receptor gamma chain." Proceedings of the National Academy of Sciences of the United States of America **92**(2): 377-381.

Donello, J. E., J. E. Loeb, et al. (1998). "Woodchuck hepatitis virus contains a tripartite posttranscriptional regulatory element." Journal of Virology **72**(6): 5085-5092.

Doyon, J. B., B. Zeitler, et al. (2011). "Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells." Nature cell biology **13**(3): 331-337.

Doyon, Y., T. D. Vo, et al. (2011). "Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures." Nature methods **8**(1): 74-79.

Dreier, B., R. P. Fuller, et al. (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." The Journal of biological chemistry **280**(42): 35588-35597.

Dreier, B., D. J. Segal, et al. (2000). "Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains." Journal of molecular biology **303**(4): 489-502.

Elliott, B., C. Richardson, et al. (1998). "Gene conversion tracts from double-strand break repair in mammalian cells." Molecular and cellular biology **18**(1): 93-101.

Fischer, A., F. Le Deist, et al. (2005). "Severe combined immunodeficiency. A model disease for molecular immunology and therapy." Immunological reviews **203**: 98-109.

Foley, J. E., M. L. Maeder, et al. (2009). "Targeted mutagenesis in zebrafish using customized zinc-finger nucleases." Nature protocols **4**(12): 1855-1867.

Foley, J. E., J. R. Yeh, et al. (2009). "Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN)." PloS one **4**(2): e4348.

Follenzi, A., L. E. Ailles, et al. (2000). "Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences." Nature genetics **25**(2): 217-222.

Gabriel, R., A. Lombardo, et al. (2011). "An unbiased genome-wide analysis of zinc-finger nuclease specificity." Nature biotechnology **29**(9): 816-823.

Gaj, T., J. Guo, et al. (2012). "Targeted gene knockout by direct delivery of zinc-finger nuclease proteins." Nature methods **9**(8): 805-807.

Gasior, S. L., H. Olivares, et al. (2001). "Assembly of RecA-like recombinases: distinct roles for mediator proteins in mitosis and meiosis." Proceedings of the National Academy of Sciences of the United States of America **98**(15): 8411-8418.

Gaspar, H. B., K. L. Parsley, et al. (2004). "Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector." The Lancet **364**(9452): 2181-2187.

Ginn, S. L., J. A. Curtin, et al. (2005). "Treatment of an infant with X-linked severe combined immunodeficiency (SCID-X1) by gene therapy in Australia." The Medical journal of Australia **182**(9): 458-463.

Giri, J. G., M. Ahdieh, et al. (1994). "Utilization of the beta and gamma chains of the IL-2 receptor by the novel cytokine IL-15." The EMBO journal **13**(12): 2822-2830.

Goldman, J. P., M. P. Blundell, et al. (1998). "Enhanced human cell engraftment in mice deficient in RAG2 and the common cytokine receptor gamma chain." British journal of haematology **103**(2): 335-342.

Gothot, A., R. Pyatt, et al. (1997). "Functional heterogeneity of human CD34(+) cells isolated in subcompartments of the G0 /G1 phase of the cell cycle." Blood **90**(11): 4384-4393.

Grabstein, K. H., J. Eisenman, et al. (1994). "Cloning of a T cell growth factor that interacts with the beta chain of the interleukin-2 receptor." Science **264**(5161): 965-968.

Greenwald, D. L., S. M. Cashman, et al. (2010). "Engineered zinc finger nuclease-mediated homologous recombination of the human rhodopsin gene." Investigative ophthalmology & visual science **51**(12): 6374-6380.

Guo, J., T. Gaj, et al. (2010). "Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases." Journal of molecular biology **400**(1): 96-107.

Gupta, A., R. G. Christensen, et al. (2012). "An optimized two-finger archive for ZFN-mediated gene targeting." Nature methods **9**(6): 588-590.

Gupta, A., X. Meng, et al. (2011). "Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases." Nucleic acids research **39**(1): 381-392.

Habib, T., S. Senadheera, et al. (2002). "The common gamma chain (gamma c) is a required signaling component of the IL-21 receptor and supports IL-21-induced cell proliferation via JAK3." Biochemistry **41**(27): 8725-8731.

Hacein-Bey-Abina, S., A. Garrigue, et al. (2008). "Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1." The Journal of clinical investigation **118**(9): 3132-3142.

Hacein-Bey-Abina, S., F. Le Deist, et al. (2002). "Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy." The New England journal of medicine **346**(16): 1185-1193.

Hacein-Bey-Abina, S., C. von Kalle, et al. (2003). "A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency." The New England journal of medicine **348**(3): 255-256.

Hacein-Bey-Abina, S., C. Von Kalle, et al. (2003). "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1." Science **302**(5644): 415-419.

Hale, L. P., R. H. Buckley, et al. (2004). "Abnormal development of thymic dendritic and epithelial cells in human X-linked severe combined immunodeficiency." Clinical immunology **110**(1): 63-70.

Handel, E. M., S. Alwin, et al. (2009). "Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity." Molecular therapy : the journal of the American Society of Gene Therapy **17**(1): 104-111.

Handel, E. M., K. Gellhaus, et al. (2012). "Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors." Human gene therapy **23**(3): 321-329.

Hennemann, B., E. Conneally, et al. (1999). "Optimization of retroviral-mediated gene transfer to human NOD/SCID mouse repopulating cord blood cells through a systematic analysis of protocol variables." Experimental hematology **27**(5): 817-825.

Henthorn, P. S., R. L. Somberg, et al. (1994). "IL-2R gamma gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease." Genomics **23**(1): 69-74.

Herrmann, F., M. Garriga-Canut, et al. (2011). "p53 Gene repair with zinc finger nucleases optimised by yeast 1-hybrid and validated by Solexa sequencing." PloS one **6**(6): e20913.

Hinnen, A., J. B. Hicks, et al. (1978). "Transformation of yeast." Proceedings of the National Academy of Sciences of the United States of America **75**(4): 1929-1933.

Hironaka, N., K. Mochida, et al. (2004). "Tax-independent constitutive IkappaB kinase activation in adult T-cell leukemia cells." Neoplasia **6**(3): 266-278.

Hockemeyer, D., F. Soldner, et al. (2009). "Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases." Nature biotechnology **27**(9): 851-857.

Hockemeyer, D., H. Wang, et al. (2011). "Genetic engineering of human pluripotent cells using TALE nucleases." Nature biotechnology **29**(8): 731-734.

Hoher, T., L. Wallace, et al. (2012). "Highly efficient zinc-finger nuclease-mediated disruption of an eGFP transgene in keratinocyte stem cells without impairment of stem cell properties." Stem cell reviews **8**(2): 426-434.

Holt, N., J. Wang, et al. (2010). "Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo." Nature biotechnology **28**(8): 839-847.

Howard, M., J. Farrar, et al. (1982). "Identification of a T cell-derived b cell growth factor distinct from interleukin 2." The Journal of experimental medicine **155**(3): 914-923.

Howe, S. J., M. R. Mansour, et al. (2008). "Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients." The Journal of clinical investigation **118**(9): 3143-3150.

Hurt, J. A., S. A. Thibodeau, et al. (2003). "Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection." Proceedings of the National Academy of Sciences of the United States of America **100**(21): 12271-12276.

Joung, J. K., E. I. Ramm, et al. (2000). "A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions." Proceedings of the National Academy of Sciences of the United States of America **97**(13): 7382-7387.

Kadyk, L. C. and L. H. Hartwell (1992). "Sister chromatids are preferred over homologs as substrates for recombinational repair in Saccharomyces cerevisiae." Genetics **132**(2): 387-402.

Kajumbula, H., W. Byarugaba, et al. (2012). "Targeting wild-type Erythrocyte receptors for Plasmodium falciparum and vivax Merozoites by Zinc Finger Nucleases In-silico: Towards a Genetic Vaccine against Malaria." Genetic vaccines and therapy **10**(1): 8.

Kim, Y. G., J. Cha, et al. (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." **93**(3): 1156-1160.

Kim, Y. G. and S. Chandrasegaran (1994). "Chimeric restriction endonuclease." Proceedings of the National Academy of Sciences of the United States of America **91**(3): 883-887.

Kimura, Y., T. Takeshita, et al. (1995). "Sharing of the IL-2 receptor gamma chain with the functional IL-9 receptor complex." International immunology **7**(1): 115-120.

Kondo, M., T. Takeshita, et al. (1994). "Functional participation of the IL-2 receptor gamma chain in IL-7 receptor complexes." Science **263**(5152): 1453-1454.

Kondo, M., T. Takeshita, et al. (1993). "Sharing of the interleukin-2 (IL-2) receptor gamma chain between receptors for IL-2 and IL-4." Science **262**(5141): 1874-1877.

Kumaki, S., N. Ishii, et al. (2000). "Characterization of the gammac chain among 27 unrelated Japanese patients with X-linked severe combined immunodeficiency (X-SCID)." Human Genetics **107**(4): 406-408.

Kumaki, S., N. Ishii, et al. (1999). "Functional role of interleukin-4 (IL-4) and IL-7 in the development of X-linked severe combined immunodeficiency." Blood **93**(2): 607-612.

Lazo, P. A., J. S. Lee, et al. (1990). "Long-distance activation of the Myc protooncogene by provirus insertion in Mlvi-1 or Mlvi-4 in rat T-cell lymphomas." Proceedings of the National Academy of Sciences of the United States of America **87**(1): 170-173.

Leavitt, A. D., G. Robles, et al. (1996). "Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection." Journal of Virology **70**(2): 721-728.

Lee, H. J., E. Kim, et al. (2010). "Targeted chromosomal deletions in human cells using zinc finger nucleases." Genome research **20**(1): 81-89.

Lee, H. J., J. Kweon, et al. (2012). "Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases." Genome research **22**(3): 539-548.

Leonard, W. J. (2001). "Cytokines and immunodeficiency diseases." Nature reviews. Immunology **1**(3): 200-208.

Leonard, W. J. and J. J. O'Shea (1998). "Jaks and STATs: biological implications." Annual review of immunology **16**: 293-322.

Li, H., V. Haurigot, et al. (2011). "In vivo genome editing restores haemostasis in a mouse model of haemophilia." Nature **475**(7355): 217-221.
Liang, Q., N. Conte, et al. (2008). "Extensive genomic copy number variation in embryonic stem cells." Proceedings of the National Academy of Sciences of the United States of America **105**(45): 17453-17456.
Liao, W., J. X. Lin, et al. (2011). "IL-2 family cytokines: new insights into the complex roles of IL-2 as a broad regulator of T helper cell differentiation." Current opinion in immunology **23**(5): 598-604.
Lin, F. L., K. Sperle, et al. (1984). "Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process." Molecular and cellular biology **4**(6): 1020-1034.
Link, H., L. Arseniev, et al. (1996). "Transplantation of allogeneic CD34+ blood cells." Blood **87**(11): 4903-4909.
Liskay, R. M., A. Letsou, et al. (1987). "Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells." Genetics **115**(1): 161-167.
Liu, Q., D. J. Segal, et al. (1997). "Design of polydactyl zinc-finger proteins for unique addressing within complex genomes." Proceedings of the National Academy of Sciences of the United States of America **94**(11): 5525-5530.
Liu, X., Z. Yan, et al. (2004). "Targeted Correction of Single-Base-Pair Mutations with Adeno-Associated Virus Vectors under Nonselective Conditions." Journal of Virology **78**(8): 4165-4175.
Lo, M., M. L. Bloom, et al. (1999). "Restoration of lymphoid populations in a murine model of X-linked severe combined immunodeficiency by a gene-therapy approach." Blood **94**(9): 3027-3036.
Lodolce, J. P., D. L. Boone, et al. (1998). "IL-15 receptor maintains lymphoid homeostasis by supporting lymphocyte homing and proliferation." Immunity **9**(5): 669-676.
Lombardo, A., D. Cesana, et al. (2011). "Site-specific integration and tailoring of cassette design for sustainable gene transfer." Nature methods **8**(10): 861-869.
Lombardo, A., P. Genovese, et al. (2007). "Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery." Nature biotechnology **25**(11): 1298-1306.
Lozzio, C. B. and B. B. Lozzio (1975). "Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome." Blood **45**(3): 321-334.
Maeda, M., A. Shimizu, et al. (1985). "Origin of human T-lymphotrophic virus I-positive T cell lines in adult T cell leukemia. Analysis of T cell receptor gene rearrangement." The Journal of experimental medicine **162**(6): 2169-2174.
Maeder, M. L., S. Thibodeau-Beganny, et al. (2008). "Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification." Molecular cell **31**(2): 294-301.
Malek, T. R. and A. L. Bayer (2004). "Tolerance, not immunity, crucially depends on IL-2." Nature reviews. Immunology **4**(9): 665-674.
Mansour, S. L., K. R. Thomas, et al. (1988). "Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes." Nature **336**(6197): 348-352.
Markiewicz, S., R. Bosselut, et al. (1996). "Tissue-specific activity of the gammac chain gene promoter depends upon an Ets binding site and is regulated by GA-binding protein." The Journal of biological chemistry **271**(25): 14849-14855.

Marshall, E. (2003). "Gene therapy. Second child in French trial is found to have leukemia." Science **299**(5605): 320.

Mashimo, T., A. Takizawa, et al. (2010). "Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases." PloS one **5**(1): e8870.

Meng, X., M. B. Noyes, et al. (2008). "Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases." Nature biotechnology **26**(6): 695-701.

Miller, J., A. D. McLachlan, et al. (1985). "Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes." The EMBO journal **4**(6): 1609-1614.

Miller, J. C., M. C. Holmes, et al. (2007). "An improved zinc-finger nuclease architecture for highly specific genome editing." Nature biotechnology **25**(7): 778-785.

Miyamoto, H., E. Zeuthen, et al. (1973). "Clonal growth of mouse cells (strain L)." Journal of cell science **13**(3): 879-888.

Miyazaki, T., A. Kawahara, et al. (1994). "Functional activation of Jak1 and Jak3 by selective association with IL-2 receptor subunits." Science **266**(5187): 1045-1047.

Modlich, U., J. Bohne, et al. (2006). "Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity." Blood **108**(8): 2545-2553.

Moehle, E. A., J. M. Rock, et al. (2007). "Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases." Proceedings of the National Academy of Sciences of the United States of America **104**(9): 3055-3060.

Montini, E., D. Cesana, et al. (2006). "Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration." Nature biotechnology **24**(6): 687-696.

Moore, M., A. Klug, et al. (2001). "Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units." Proceedings of the National Academy of Sciences of the United States of America **98**(4): 1437-1441.

Morgan, D. A., F. W. Ruscetti, et al. (1976). "Selective in vitro growth of T lymphocytes from normal human bone marrows." Science **193**(4257): 1007-1008.

Moriggl, R., D. J. Topham, et al. (1999). "Stat5 is required for IL-2-induced cell cycle progression of peripheral T cells." Immunity **10**(2): 249-259.

Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors." Science **326**(5959): 1501.

Mulloy, J. C., R. W. Crownley, et al. (1996). "The human T-cell leukemia/lymphotropic virus type 1 p12I proteins bind the interleukin-2 receptor beta and gammac chains and affects their expression on the cell surface." Journal of Virology **70**(6): 3599-3605.

Mussolino, C., R. Morbitzer, et al. (2011). "A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity." Nucleic acids research **39**(21): 9283-9293.

Nain, V., S. Sahi, et al. (2010). "CPP-ZFN: a potential DNA-targeting anti-malarial drug." Malaria journal **9**: 258.

Nakamura, Y., S. M. Russell, et al. (1994). "Heterodimerization of the IL-2 receptor beta- and gamma-chain cytoplasmic domains is required for signalling." Nature **369**(6478): 330-333.

Naldini, L., U. Blomer, et al. (1996). "Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector." Proceedings of the National Academy of Sciences of the United States of America **93**(21): 11382-11388.

Namen, A. E., S. Lupton, et al. (1988). "Stimulation of B-cell progenitors by cloned murine interleukin-7." Nature **333**(6173): 571-573.

Nathwani, A. C., A. M. Davidoff, et al. (2005). "A review of gene therapy for haematological disorders." British journal of haematology **128**(1): 3-17.

Noguchi, M., S. Adelstein, et al. (1993). "Characterization of the human interleukin-2 receptor gamma chain gene." The Journal of biological chemistry **268**(18): 13601-13608.

Noguchi, M., H. Yi, et al. (1993). "Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans." Cell **73**(1): 147-157.

Ohbo, K., T. Suda, et al. (1996). "Modulation of hematopoiesis in mice with a truncated mutant of the interleukin-2 receptor gamma chain." Blood **87**(3): 956-967.

Oostendorp, R. A., J. Audet, et al. (2000). "High-resolution tracking of cell division suggests similar cell cycle kinetics of hematopoietic stem cells stimulated in vitro and in vivo." Blood **95**(3): 855-862.

Orlando, S. J., Y. Santiago, et al. (2010). "Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology." Nucleic acids research **38**(15): e152.

Orr-Weaver, T. L. and J. W. Szostak (1983). "Yeast recombination: the association between double-strand gap repair and crossing-over." Proceedings of the National Academy of Sciences of the United States of America **80**(14): 4417-4421.

Orr-Weaver, T. L., J. W. Szostak, et al. (1981). "Yeast transformation: a model system for the study of recombination." Proceedings of the National Academy of Sciences of the United States of America **78**(10): 6354-6358.

Osiak, A., F. Radecke, et al. (2011). "Selection-independent generation of gene knockout mouse embryonic stem cells using zinc-finger nucleases." PloS one **6**(12): e28911.

Otsu, M., S. M. Anderson, et al. (2000). "Lymphoid development and function in X-linked severe combined immunodeficiency mice after stem cell gene therapy." Mol Ther **1**(2): 145-153.

Ott, M. G., M. Schmidt, et al. (2006). "Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1." Nature medicine **12**(4): 401-409.

Overlack, N., T. Goldmann, et al. (2012). "Gene repair of an Usher syndrome causing mutation by zinc-finger nuclease mediated homologous recombination." Investigative ophthalmology & visual science **53**(7): 4140-4146.

Pabo, C. O., E. Peisach, et al. (2001). "Design and selection of novel Cys2His2 zinc finger proteins." Annual review of biochemistry **70**: 313-340.

Paques, F. and P. Duchateau (2007). "Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy." Current gene therapy **7**(1): 49-66.

Pattanayak, V., C. L. Ramirez, et al. (2011). "Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection." Nature methods **8**(9): 765-770.

Paul, W. E. (1991). "Interleukin-4: a prototypic immunoregulatory lymphokine." Blood **77**(9): 1859-1870.

Pepper, A. E., R. H. Buckley, et al. (1995). "Two mutational hotspots in the interleukin-2 receptor gamma chain gene causing human X-linked severe combined immunodeficiency." American journal of human genetics **57**(3): 564-571.

Perez-Pinera, P., D. G. Ousterout, et al. (2012). "Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases." Nucleic acids research **40**(8): 3741-3752.

Perez, E. E., J. Wang, et al. (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases." Nature biotechnology **26**(7): 808-816.

Peschon, J. J., P. J. Morrissey, et al. (1994). "Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice." The Journal of experimental medicine **180**(5): 1955-1960.

Pike-Overzet, K., D. de Ridder, et al. (2006). "Gene therapy: is IL2RG oncogenic in T-cell development?" Nature **443**(7109): E5; discussion E6-7.

Pike-Overzet, K., D. de Ridder, et al. (2007). "Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34+ cells: implications for leukemogenesis in gene therapy." Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K **21**(4): 754-763.

Popa, I., M. E. Harris, et al. (2002). "CRM1-dependent function of a cis-acting RNA export element." Molecular and cellular biology **22**(7): 2057-2067.

Porteus, M. H. and D. Baltimore (2003). "Chimeric nucleases stimulate gene targeting in human cells." Science **300**(5620): 763.

Provasi, E., P. Genovese, et al. (2012). "Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer." Nature medicine **18**(5): 807-815.

Pruett-Miller, S. M., J. P. Connelly, et al. (2008). "Comparison of zinc finger nucleases for use in gene targeting in mammalian cells." Molecular therapy : the journal of the American Society of Gene Therapy **16**(4): 707-717.

Pruett-Miller, S. M., D. W. Reading, et al. (2009). "Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels." PLoS genetics **5**(2): e1000376.

Puck, J. M., S. M. Deschenes, et al. (1993). "The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X-linked severe combined immunodeficiency, SCIDX1." Human molecular genetics **2**(8): 1099-1104.

Puck, J. M., A. E. Pepper, et al. (1997). "Mutation analysis of IL2RG in human X-linked severe combined immunodeficiency." Blood **89**(6): 1968-1977.

Puel, A., S. F. Ziegler, et al. (1998). "Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency." Nature genetics **20**(4): 394-397.

Radecke, S., F. Radecke, et al. (2010). "Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications." Molecular therapy : the journal of the American Society of Gene Therapy **18**(4): 743-753.

Ramalingam, S., K. Kandavelou, et al. (2011). "Creating designed zinc-finger nucleases with minimal cytotoxicity." Journal of molecular biology **405**(3): 630-641.

Ramalingam, S., V. London, et al. (2013). "Generation and genetic engineering of human induced pluripotent stem cells using designed zinc finger nucleases." Stem cells and development **22**(4): 595-610.

Ramirez, C. L., M. T. Certo, et al. (2012). "Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects." Nucleic acids research **40**(12): 5560-5568.

Rebuzzini, P., T. Neri, et al. (2008). "Karyotype analysis of the euploid cell population of a mouse embryonic stem cell line revealed a high incidence of chromosome abnormalities that varied during culture." Cytogenetic and genome research **121**(1): 18-24.

Recher, M., L. J. Berglund, et al. (2011). "IL-21 is the primary common gamma chain-binding cytokine required for human B-cell differentiation in vivo." Blood **118**(26): 6824-6835.

Reems, J. A. and B. Torok-Storb (1995). "Cell cycle and functional differences between CD34+/CD38hi and CD34+/38lo human marrow cells after in vitro cytokine exposure." Blood **85**(6): 1480-1487.

Reisner, Y., N. Kapoor, et al. (1983). "Transplantation for severe combined immunodeficiency with HLA-A,B,D,DR incompatible parental marrow cells fractionated by soybean agglutinin and sheep red blood cells." Blood **61**(2): 341-348.

Renauld, J. C., A. Goethals, et al. (1990). "Cloning and expression of a cDNA for the human homolog of mouse T cell and mast cell growth factor P40." Cytokine **2**(1): 9-12.

Richardson, C. and M. Jasin (2000). "Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells." Molecular and cellular biology **20**(23): 9068-9075.

Richardson, C. and M. Jasin (2000). "Frequent chromosomal translocations induced by DNA double-strand breaks." Nature **405**(6787): 697-700.

Rodewald, H. R., M. Ogawa, et al. (1997). "Pro-thymocyte expansion by c-kit and the common cytokine receptor gamma chain is essential for repertoire formation." Immunity **6**(3): 265-272.

Rothkamm, K., I. Kruger, et al. (2003). "Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle." Molecular and cellular biology **23**(16): 5706-5715.

Rouet, P., F. Smih, et al. (1994). "Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **91**(13): 6064-6068.

Russell, D. W. and R. K. Hirata (1998). "Human gene targeting by viral vectors." Nature genetics **18**(4): 325-330.

Russell, S. M., J. A. Johnston, et al. (1994). "Interaction of IL-2R beta and gamma c chains with Jak1 and Jak3: implications for XSCID and XCID." Science **266**(5187): 1042-1045.

Russell, S. M., A. D. Keegan, et al. (1993). "Interleukin-2 receptor gamma chain: a functional component of the interleukin-4 receptor." Science **262**(5141): 1880-1883.

Sadlack, B., H. Merz, et al. (1993). "Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene." Cell **75**(2): 253-261.

Sakkhachornphop, S., C. F. Barbas, 3rd, et al. (2012). "Zinc finger protein designed to target 2-long terminal repeat junctions interferes with human immunodeficiency virus integration." Human gene therapy **23**(9): 932-942.

Schambach, A., J. Bohne, et al. (2006). "Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression." Gene therapy **13**(7): 641-645.

Scharton-Kersten, T., H. Nakajima, et al. (1998). "Infection of mice lacking the common cytokine receptor gamma-chain (gamma(c)) reveals an unexpected role for CD4+ T lymphocytes in early IFN-gamma-dependent resistance to Toxoplasma gondii." Journal of immunology **160**(6): 2565-2569.

Schierling, B., N. Dannemann, et al. (2011). "A novel zinc-finger nuclease platform with a sequence-specific cleavage module." Nucleic acids research **40**(6): 2623-2638.

Schmidt, M., S. Hacein-Bey-Abina, et al. (2005). "Clonal evidence for the transduction of CD34+ cells with lymphomyeloid differentiation potential and self-renewal capacity in the SCID-X1 gene therapy trial." Blood **105**(7): 2699-2706.

Schorle, H., T. Holtschke, et al. (1991). "Development and function of T cells in mice rendered interleukin-2 deficient by gene targeting." Nature **352**(6336): 621-624.

Schroder, A. R., P. Shinn, et al. (2002). "HIV-1 integration in the human genome favors active genes and local hotspots." Cell **110**(4): 521-529.

Sebastiano, V., M. L. Maeder, et al. (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): 1717-1726.

Segal, D. J., B. Dreier, et al. (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): 2758-2763.

Sharara, L. I., A. Andersson, et al. (1997). "Deregulated TCR alpha beta T cell population provokes extramedullary hematopoiesis in mice deficient in the common gamma chain." European journal of immunology **27**(4): 990-998.

Shimizu, Y., C. Sollu, et al. (2011). "Adding fingers to an engineered zinc finger nuclease can reduce activity." Biochemistry **50**(22): 5033-5041.

Shinkai, Y., G. Rathbun, et al. (1992). "RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement." Cell **68**(5): 855-867.

Shiramizu, B., B. G. Herndier, et al. (1994). "Identification of a common clonal human immunodeficiency virus integration site in human immunodeficiency virus-associated lymphomas." Cancer research **54**(8): 2069-2072.

Shrivastav, M., L. P. De Haro, et al. (2008). "Regulation of DNA double-strand break repair pathway choice." Cell research **18**(1): 134-147.

Smith, J., M. Bibikova, et al. (2000). "Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains." Nucleic acids research **28**(17): 3361-3369.

Smith, J., S. Grizot, et al. (2006). "A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences." Nucleic acids research **34**(22): e149.

Smyth, C. M., S. L. Ginn, et al. (2007). "Limiting {gamma}c expression differentially affects signaling via the interleukin (IL)-7 and IL-15 receptors." Blood **110**(1): 91-98.

Soldner, F., J. Laganiere, et al. (2011). "Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations." Cell **146**(2): 318-331.

Sollu, C., K. Pars, et al. (2010). "Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion." Nucleic acids research **38**(22): 8269-8276.

Soudais, C., T. Shiho, et al. (2000). "Stable and functional lymphoid reconstitution of common cytokine receptor gamma chain deficient mice by retroviral-mediated gene transfer." Blood **95**(10): 3071-3077.

Srour, E. F., E. D. Zanjani, et al. (1993). "Persistence of human multilineage, self-renewing lymphohematopoietic stem cells in chimeric sheep." Blood **82**(11): 3333-3342.

Stauber, D. J., E. W. Debler, et al. (2006). "Crystal structure of the IL-2 signaling complex: paradigm for a heterotrimeric cytokine receptor." Proceedings of the National Academy of Sciences of the United States of America **103**(8): 2788-2793.

Stephan, V., V. Wahn, et al. (1996). "Atypical X-linked severe combined immunodeficiency due to possible spontaneous reversion of the genetic defect in T cells." The New England journal of medicine **335**(21): 1563-1567.

Stephen, S. L., V. G. Sivanandam, et al. (2008). "Homologous and heterologous recombination between adenovirus vector DNA and chromosomal DNA." The journal of gene medicine **10**(11): 1176-1189.

Sung, P. and H. Klein (2006). "Mechanism of homologous recombination: mediators and helicases take on regulatory functions." Nature reviews. Molecular cell biology **7**(10): 739-750.

Sutton, R. E., M. J. Reitsma, et al. (1999). "Transduction of human progenitor hematopoietic stem cells by human immunodeficiency virus type 1-based vectors is cell cycle dependent." Journal of Virology **73**(5): 3649-3660.

Szczepek, M., V. Brondani, et al. (2007). "Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases." Nature biotechnology **25**(7): 786-793.

Szostak, J. W., T. L. Orr-Weaver, et al. (1983). "The double-strand-break repair model for recombination." Cell **33**(1): 25-35.

Takahashi, K., K. Okita, et al. (2007). "Induction of pluripotent stem cells from fibroblast cultures." Nature protocols **2**(12): 3081-3089.

Takashima, Y., M. Sakuraba, et al. (2009). "Dependence of DNA double strand break repair pathways on cell cycle phase in human lymphoblastoid cells." Environmental and molecular mutagenesis **50**(9): 815-822.

Takeshita, T., H. Asao, et al. (1992). "Cloning of the gamma chain of the human IL-2 receptor." Science **257**(5068): 379-382.

Takeshita, T., H. Asao, et al. (1990). "An associated molecule, p64, with high-affinity interleukin 2 receptor." International immunology **2**(5): 477-480.

Takeshita, T., K. Ohtani, et al. (1992). "An associated molecule, p64, with IL-2 receptor beta chain. Its possible involvement in the formation of the functional intermediate-affinity IL-2 receptor complex." Journal of immunology **148**(7): 2154-2158.

Thierry, A. and B. Dujon (1992). "Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes." Nucleic acids research **20**(21): 5625-5631.

Thomas, K. R., K. R. Folger, et al. (1986). "High frequency targeting of genes to specific sites in the mammalian genome." Cell **44**(3): 419-428.

Thornhill, S. I., A. Schambach, et al. (2008). "Self-inactivating gammaretroviral vectors for gene therapy of X-linked severe combined immunodeficiency." Molecular therapy : the journal of the American Society of Gene Therapy **16**(3): 590-598.

Thrasher, A. J., S. Hacein-Bey-Abina, et al. (2005). "Failure of SCID-X1 gene therapy in older patients." Blood **105**(11): 4255-4257.

Throm, R. E., A. A. Ouma, et al. (2009). "Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection." Blood **113**(21): 5104-5110.

Urnov, F. D., J. C. Miller, et al. (2005). "Highly efficient endogenous human gene correction using designed zinc-finger nucleases." Nature **435**(7042): 646-651.

Valton, J., F. Daboussi, et al. (2012). "5'-Cytosine-phosphoguanine (CpG) methylation impacts the activity of natural and engineered meganucleases." The Journal of biological chemistry **287**(36): 30139-30150.

van Ewijk, W., G. Hollander, et al. (2000). "Stepwise development of thymic microenvironments in vivo is regulated by thymocyte subsets." Development **127**(8): 1583-1591.

van Rensburg, R., I. Beyer, et al. (2013). "Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells." Gene therapy **20**(2): 201-214.

Van Snick, J., A. Goethals, et al. (1989). "Cloning and characterization of a cDNA for a new mouse T cell growth factor (P40)." The Journal of experimental medicine **169**(1): 363-368.

Vasquez, K. M., K. Marburger, et al. (2001). "Manipulating the mammalian genome by homologous recombination." Proceedings of the National Academy of Sciences of the United States of America **98**(15): 8403-8410.

von Freeden-Jeffry, U., P. Vieira, et al. (1995). "Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine." The Journal of experimental medicine **181**(4): 1519-1526.

Wang, Y., F. Yates, et al. (2005). "Embryonic stem cell-derived hematopoietic stem cells." Proceedings of the National Academy of Sciences of the United States of America **102**(52): 19081-19086.

Wang, Y., W. Y. Zhang, et al. (2012). "Genome editing of human embryonic stem cells and induced pluripotent stem cells with zinc finger nucleases for cellular imaging." Circulation research **111**(12): 1494-1503.

Wanisch, K. and R. J. Yanez-Munoz (2009). "Integration-deficient lentiviral vectors: a slow coming of age." Molecular therapy : the journal of the American Society of Gene Therapy **17**(8): 1316-1332.

Wayengera, M. (2011). "Proviral HIV-genome-wide and pol-gene specific zinc finger nucleases: usability for targeted HIV gene therapy." Theoretical biology & medical modelling **8**: 26.

Woods, N. B., V. Bottero, et al. (2006). "Gene therapy: therapeutic gene causing lymphoma." Nature **440**(7088): 1123.

Wu, X., Y. Li, et al. (2003). "Transcription start regions in the human genome are favored targets for MLV integration." Science **300**(5626): 1749-1751.

Yanez-Munoz, R. J., K. S. Balaggan, et al. (2006). "Effective gene therapy with nonintegrating lentiviral vectors." Nature medicine **12**(3): 348-353.

Yanez, R. J. and A. C. Porter (1999). "Gene targeting is enhanced in human cells overexpressing hRAD51." Gene therapy **6**(7): 1282-1290.

Yanez, R. J. and A. C. Porter (2002). "Differential effects of Rad52p overexpression on gene targeting and extrachromosomal homologous recombination in a human cell line." Nucleic acids research **30**(3): 740-748.

Yao, Y., B. Nashun, et al. (2012). "Generation of CD34+ cells from CCR5-disrupted human embryonic and induced pluripotent stem cells." Human gene therapy **23**(2): 238-242.

Yu, S. F., T. von Ruden, et al. (1986). "Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **83**(10): 3194-3198.

Yuan, J., J. Wang, et al. (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): 849-859.

Yusa, K., S. T. Rashid, et al. (2011). "Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells." Nature **478**(7369): 391-394.

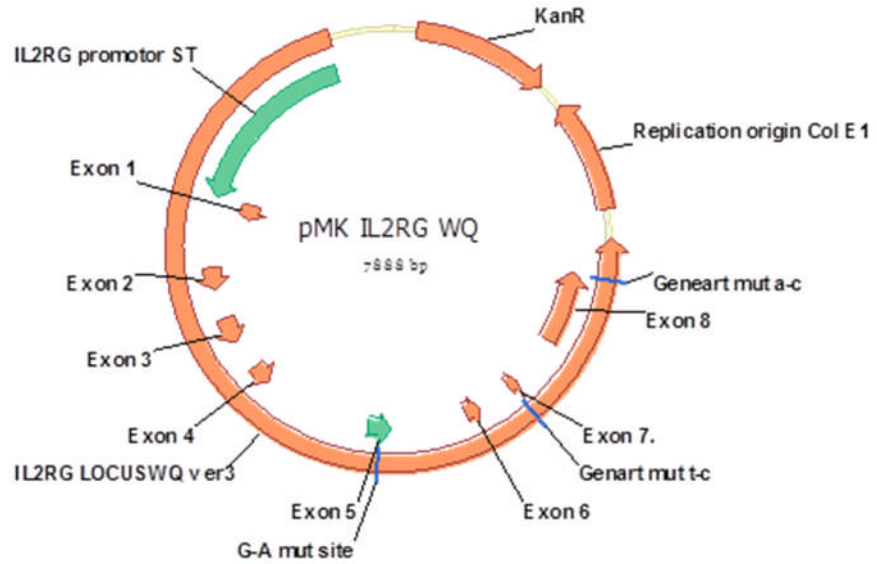
Zennou, V., C. Petit, et al. (2000). "HIV-1 genome nuclear import is mediated by a central DNA flap." Cell **101**(2): 173-185.

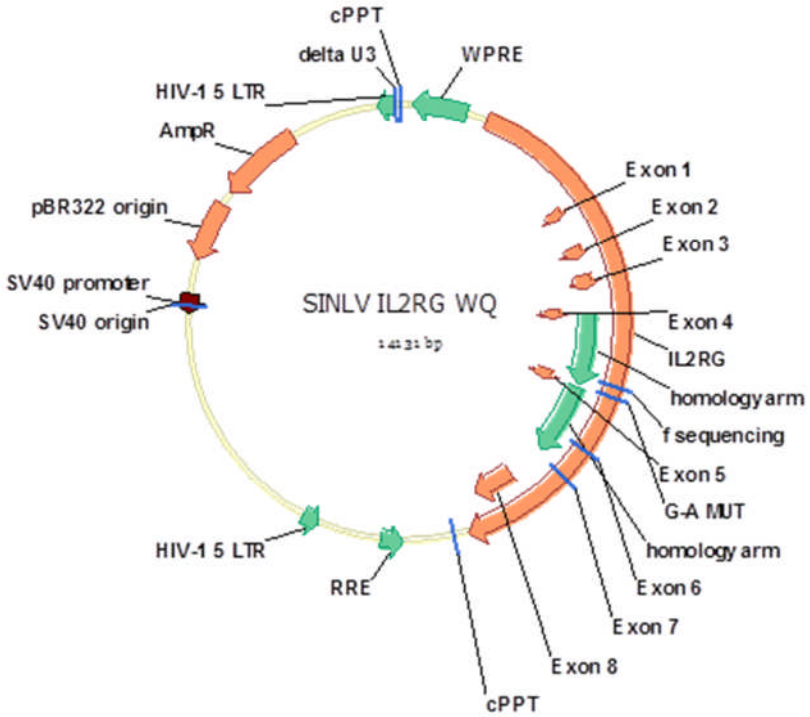
Zhang, F., S. I. Thornhill, et al. (2007). "Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells." Blood **110**(5): 1448-1457.

Zou, J., C. L. Sweeney, et al. (2011). "Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting." Blood **117**(21): 5561-5572.

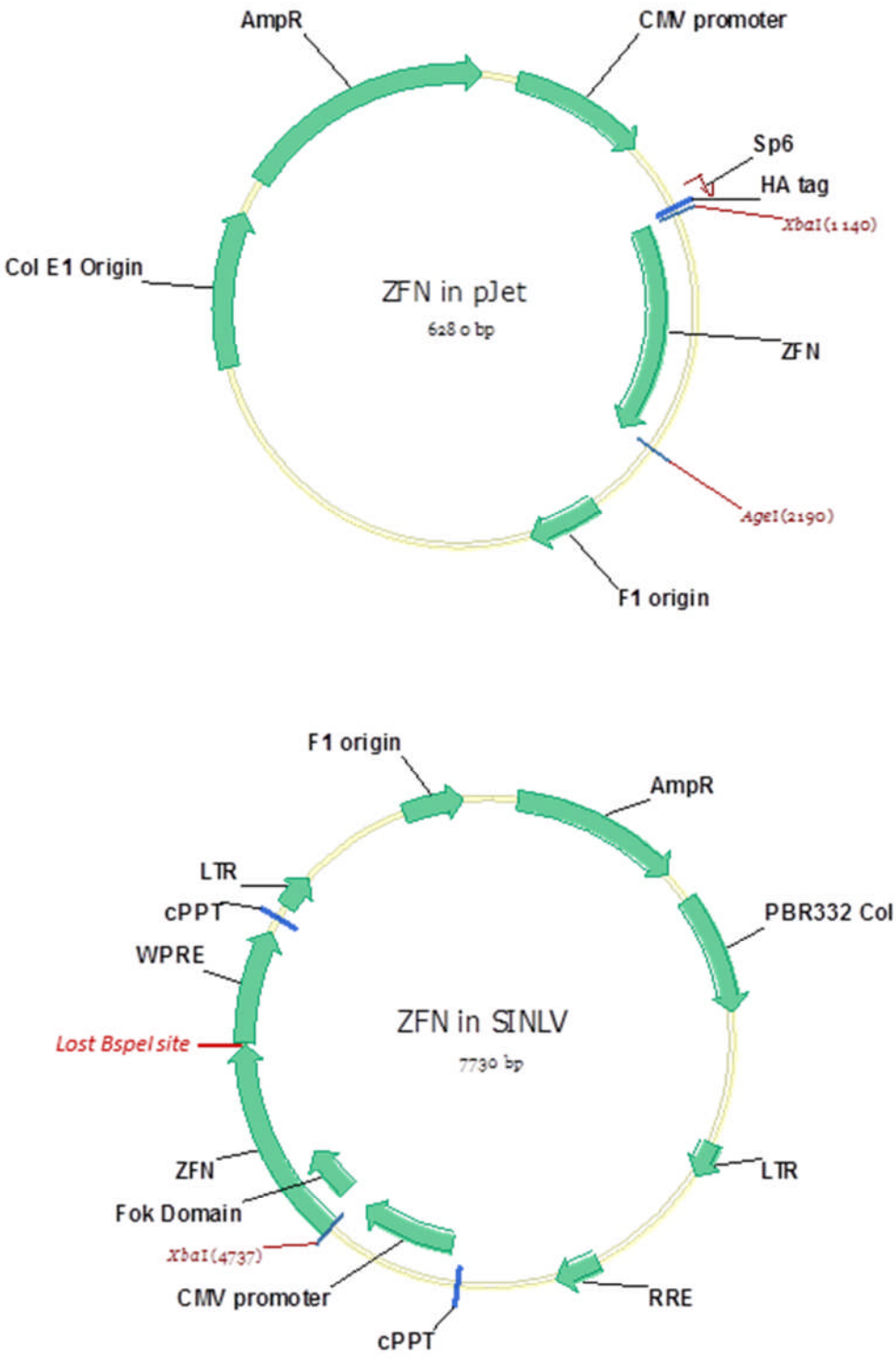
Zychlinski, D., A. Schambach, et al. (2008). "Physiological promoters reduce the genotoxic risk of integrating gene vectors." Molecular therapy : the journal of the American Society of Gene Therapy **16**(4): 718-725.

9 Appendices

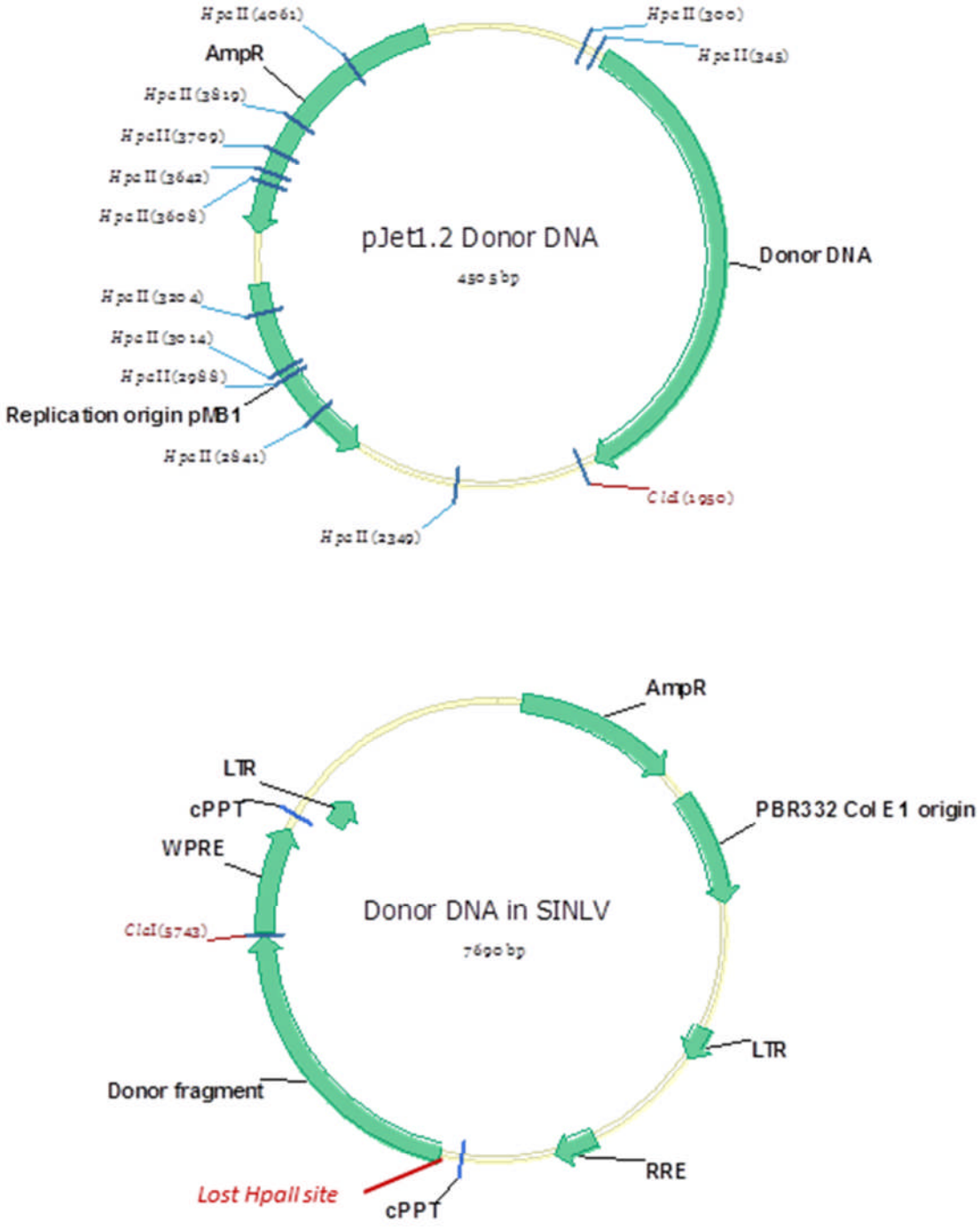




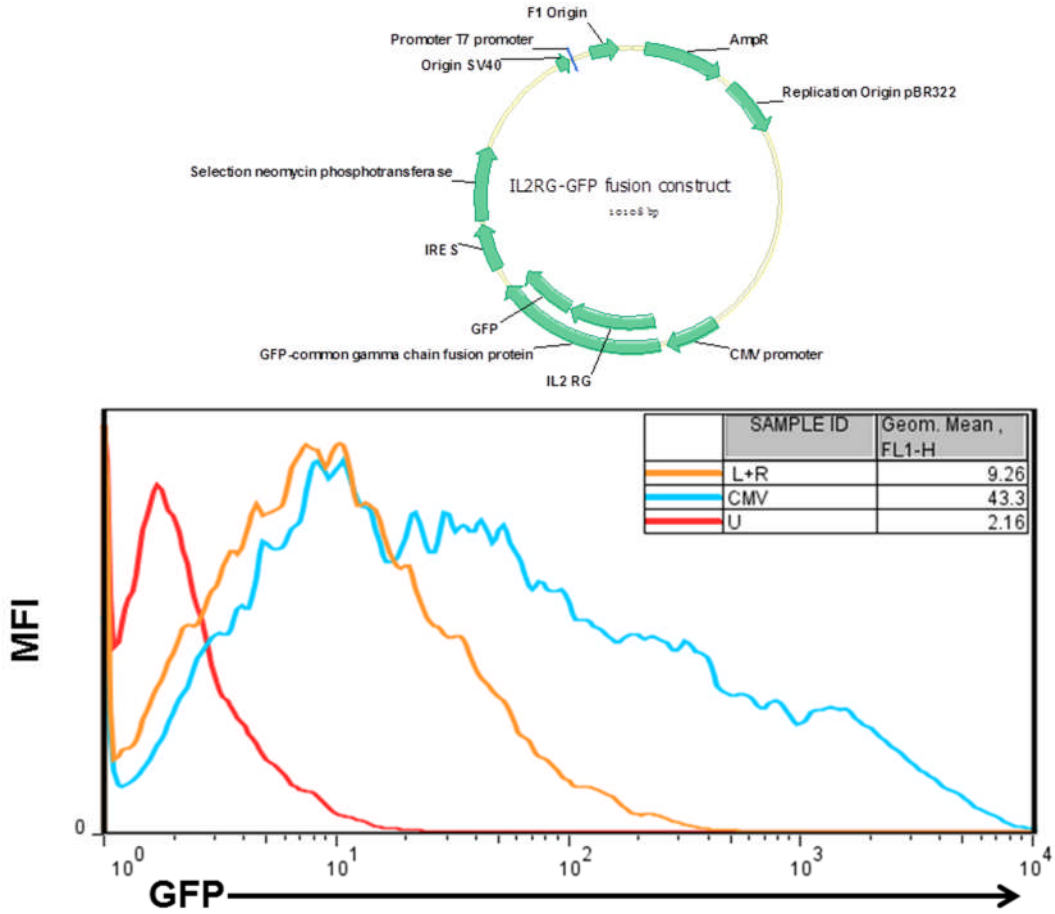
Appendix 1: Full length *IL2RG* expression vectors. Top, *IL2RG* in a plasmid expression vector deigned by W. Qasim (WQ) incorporating the IL2RG promoter region isolated by S. Thornhill (ST). Bottom, the IL2RG in reverse orientation in a lentiviral packaging vector. The donor DNA homology arms are marked on the lentiviral vector to demonstrate their position in regards to the endogenous gene.



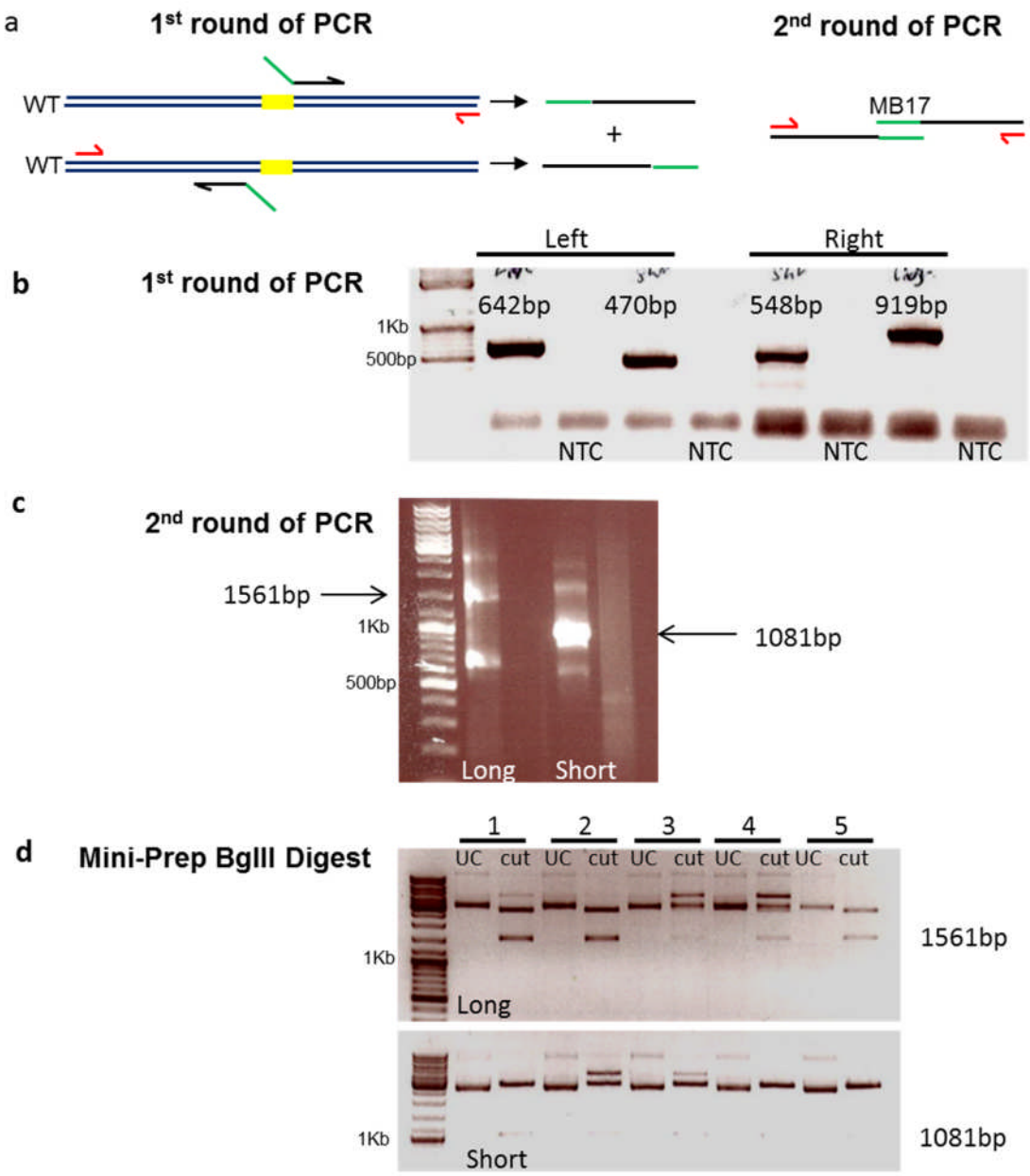
Appendix 2: Expression vectors representing either ZFN, EA or KV. The ZFN gene was cloned from a plasmid expression vector (Top) in to a Self-inactivating lentiviral packaging vector (SINLV) (Bottom). The ZFN expression plasmid was digested with restriction enzymes ApaI and XhoI and the ZFN fragment isolated and purified. Corresponding sticky ends were generated by linearising the lentiviral backbone with BspeI and XbaI. The ZFN containing insert was ligated into the lentiviral backbone and confirmed by a restriction enzyme digest and DNA sequencing. On ligation of the insert the BspeI site is lost and the XbaI site remains, as shown.



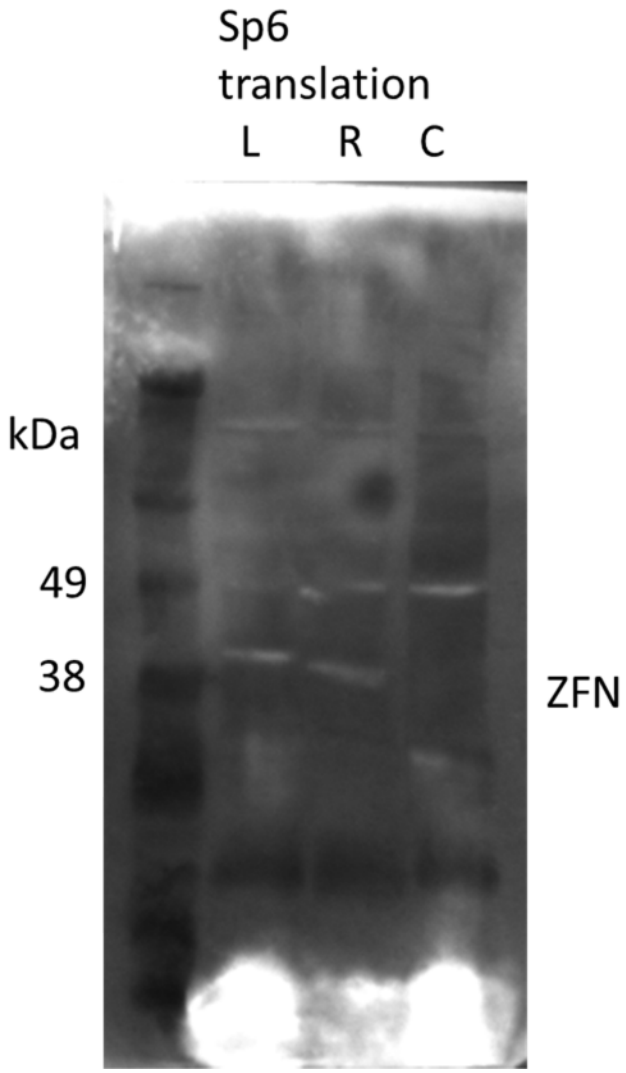
Appendix 3: Representative Donor DNA constructs. (Top) The 1531bp homology domain for WT *IL2RG* was amplified by PCR and ligated into pJET1.2 by blunt-end ligation The MB1, MB3 and MB17 donor DNA cassettes were subsequently cloned to replace the ZFN binding domain. The MB17 donor DNA variant was cloned into a self-inactivating lentiviral packaging vector (SINLV) (Bottom). The 1531bp homology domain with the MB1 cassette was digested from the pJet1.2 plasmid with restriction enzymes ClaI and HpaII and the donor DNA fragment isolated and purified. Complementary sticky ends were generated by digesting the lentiviral backbone with ClaI (ClaI has compatible sticky ends with HpaII) followed by dephosphorylation. The donor DNA insert was ligated into the lentiviral backbone and confirmed by a restriction enzyme digest and DNA sequencing. On ligation of the insert the HpaII site on the insert is lost and the ClaI site remains, as shown.



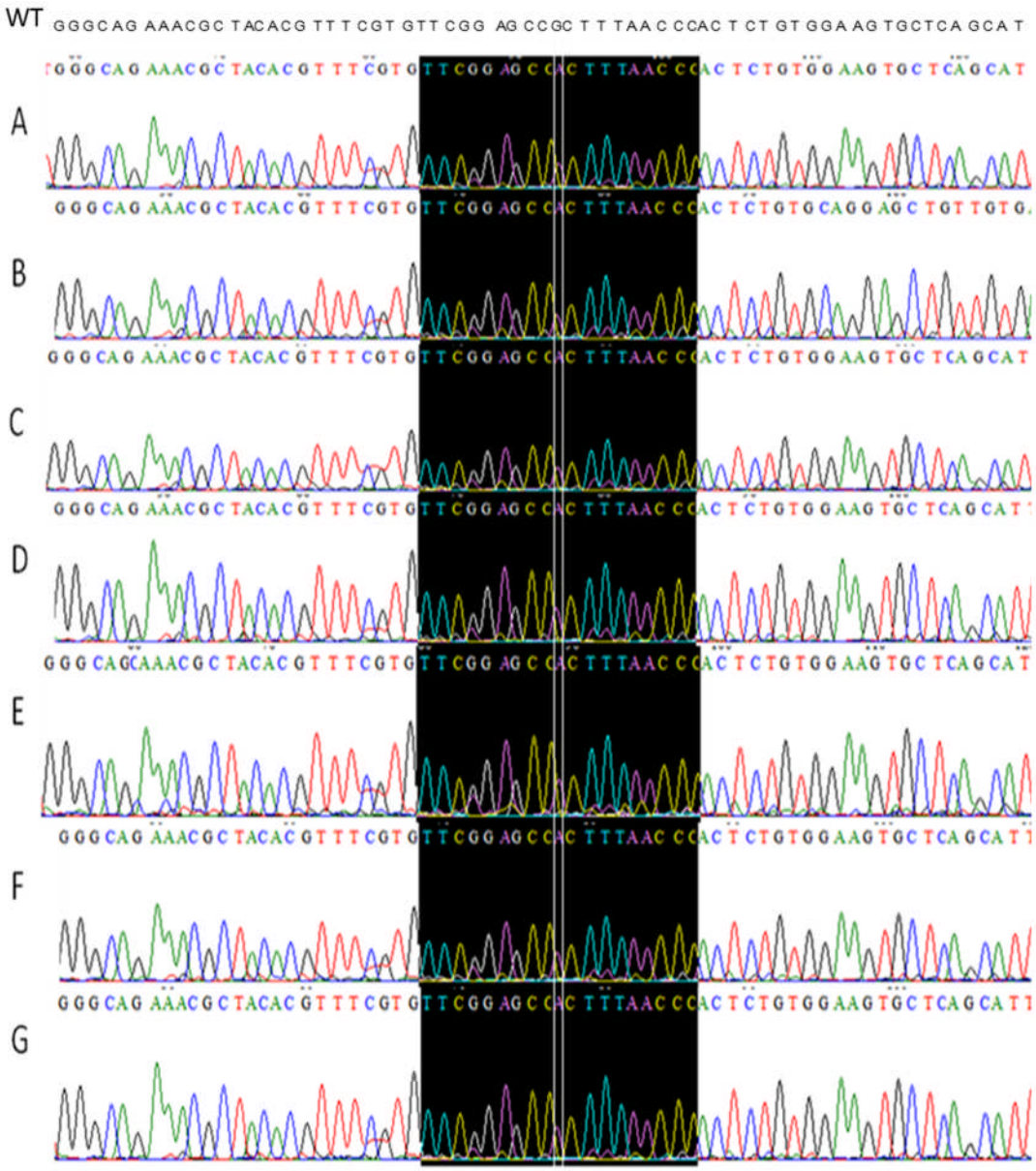
Appendix 4: The episomal GFP knockdown assay. 293T cells were transfected with *IL2RG*-GFP expression plasmids (top) alongside ZFN expression or CMV control plasmids. GFP expression was analysed by flow cytometry. (bottom) The MFI and Geometric mean (Geom mean) of GFP on live cells is shown by a histogram for ZFN pair treated (L+R), CMV promoter control and untransfected (U).



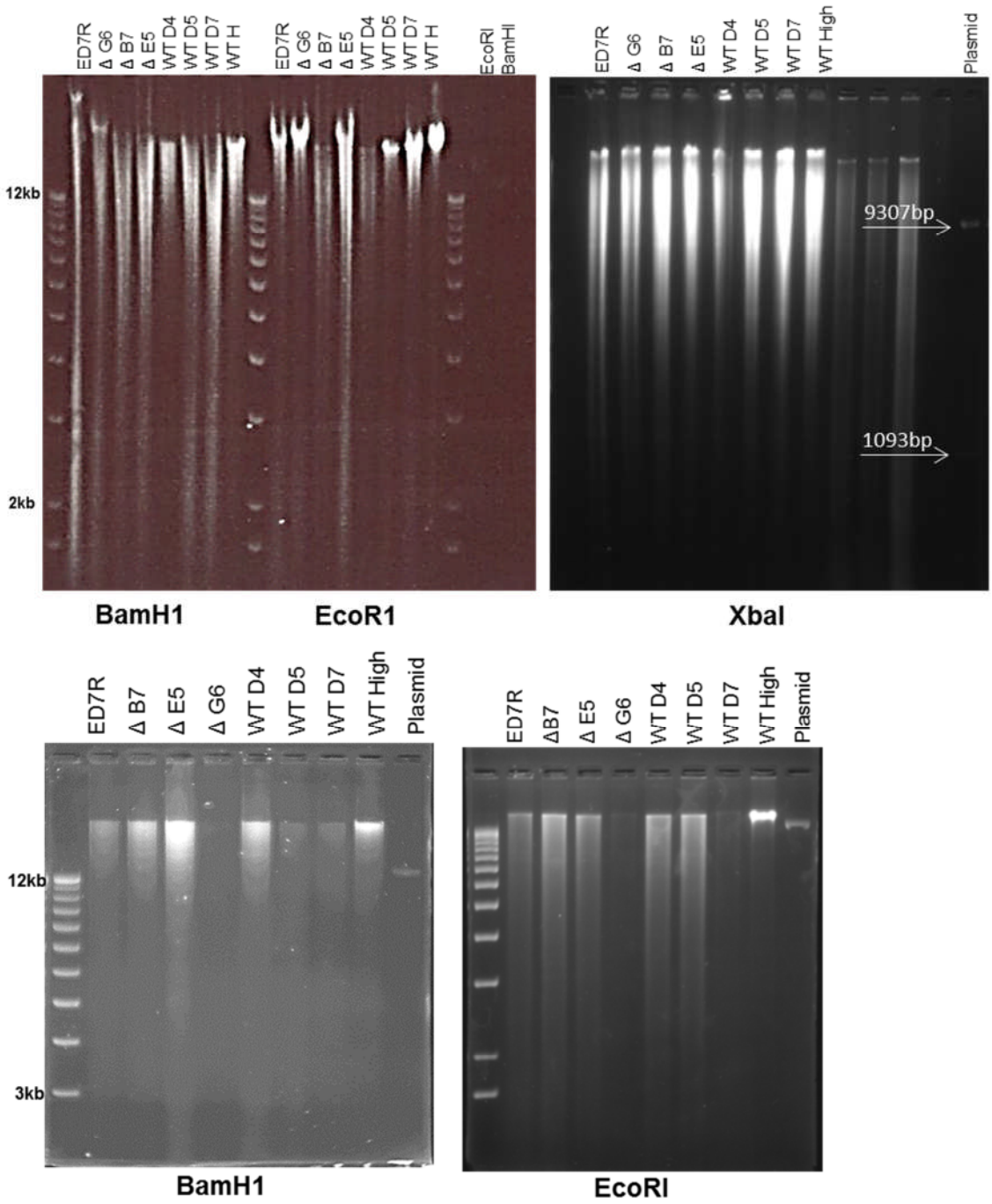
Appendix 5. Construction of the long and short MB17 donor DNA plasmids. (a) the linker PCR strategy. (b) the first round of PCR, using primers designed to make short, 470 and 548bp and long, 642 and 919bp product. pMK-*IL2RG* WQ was used for template Non template controls were included for each primer pair, as indicated. (c) the second round of PCR was carried out to combine the short and long PCR products from (b) resulting in 1018bp and long 1561bp MB17 donor DNA cassettes. (d) The 1018bp Long and 1561bp Short PCR bands were purified by gel extraction and cloned into pJet. Plasmid DNA was extracted from clones 1-5 and digested with BglII (Cut) which cleaves on either site of the pJet multiple cloning site. Uncut (UC) was included for control.



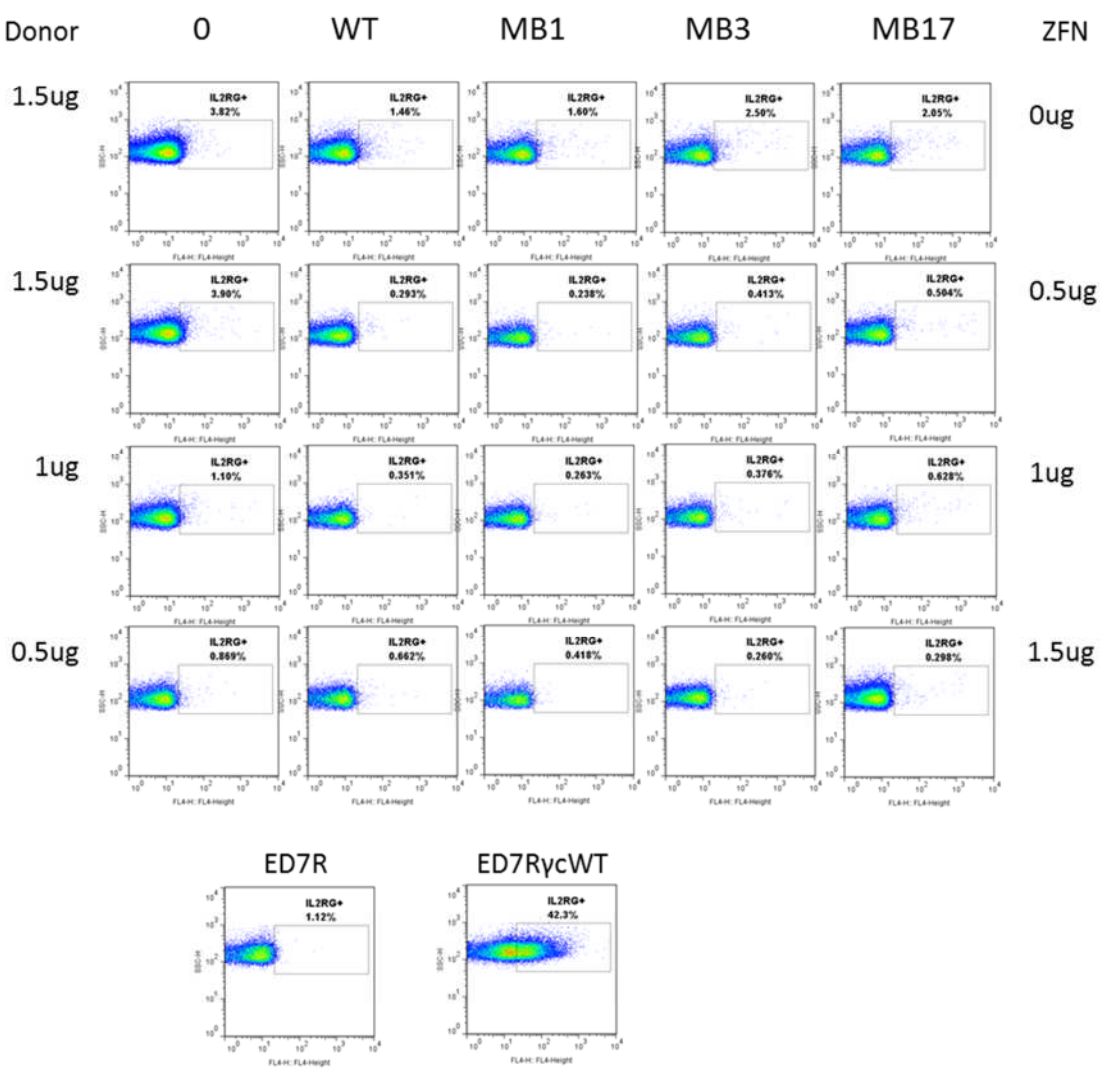
Appendix 6: Western blot to detect ZFN protein production in the Sp6 rabbit reticulate lysate assay. Left (L) and right (R) ZFN expression plasmids were incubated with Rabbit reticulate. A no plasmid control (C) was also included. The lysate was run on an SDS page gel, the protein bands were then transferred to a membrane and Anti-HA antibody, conjugated to HRP, was applied to the membrane at a 1:1000 dilution in 5% milk powder, 0.1% Tween in PBS, to detect the HA tag on the 40KDa ZFN protein.



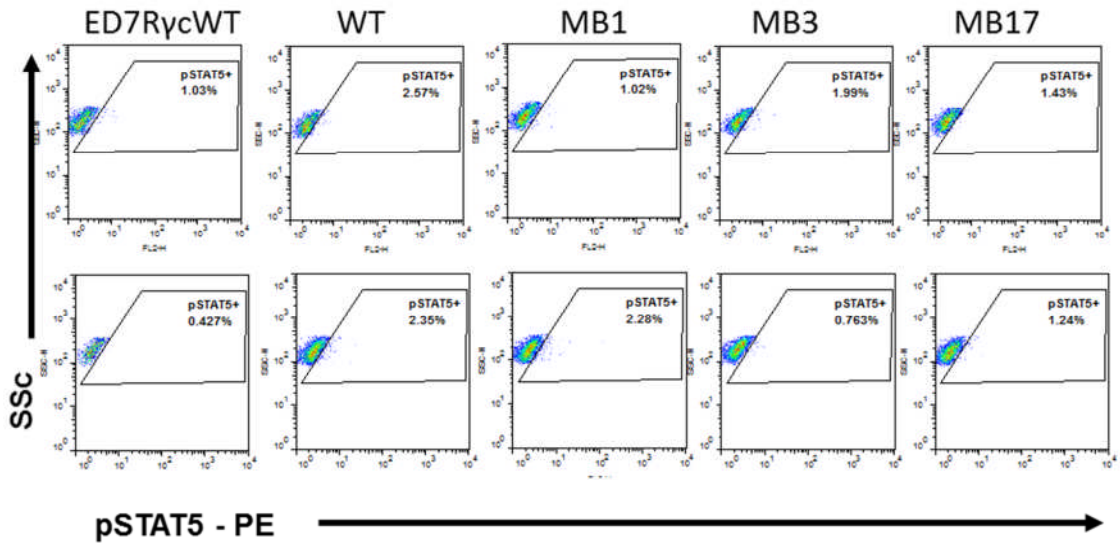
Appendix 7: Sequencing results of the site directed mutagenesis. Chromatographs (A-G) of the G691A point mutation and the surrounding DNA sequence from 7 plasmid samples isolated from 7 *E. Coli* clones as a result of site directed mutagenesis compared to the wild type (WT) sequence (Top).



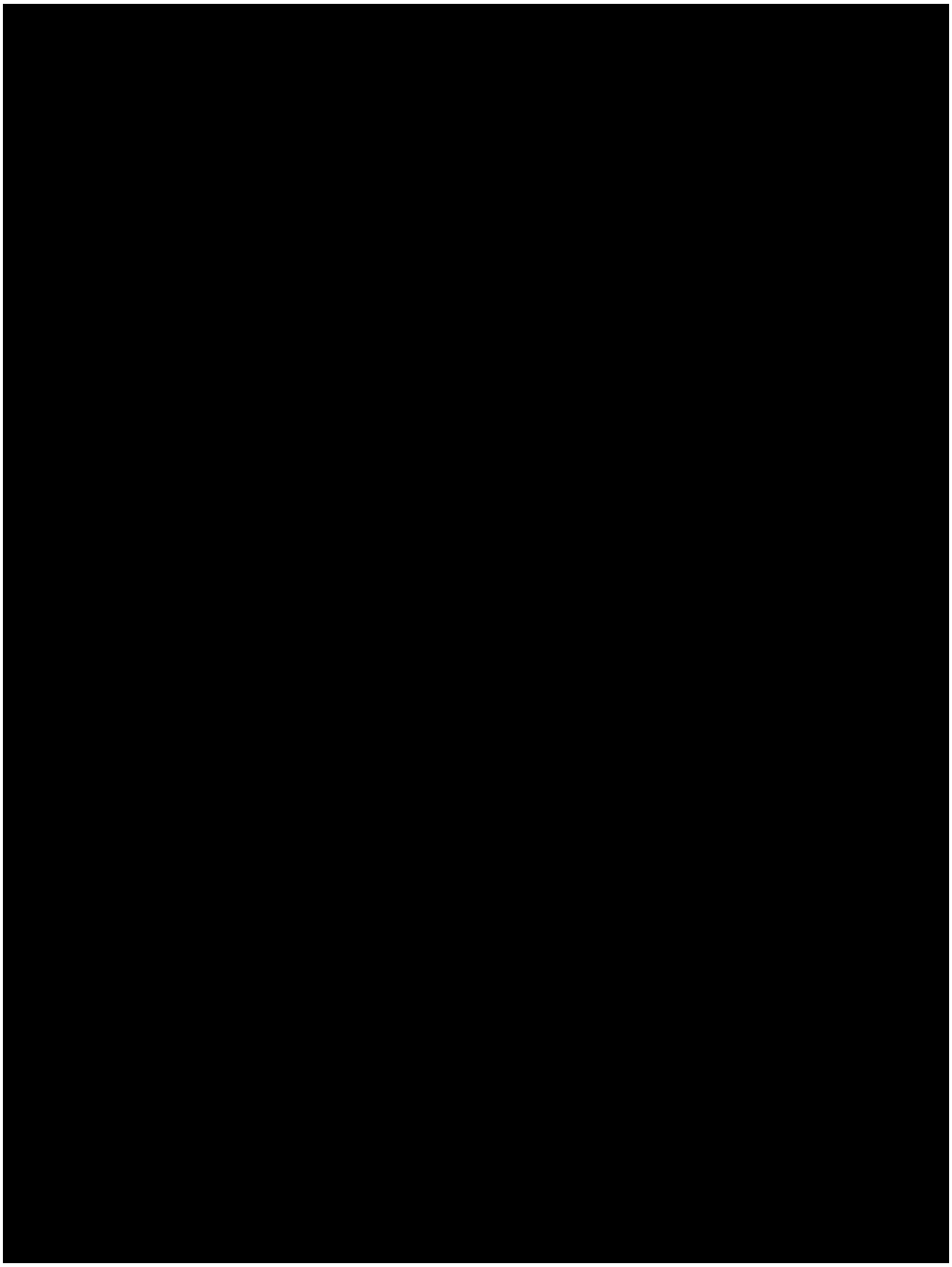
Appendix 8: Copy number Southern blot hybridisation accompanying gel. Genomic DNA was extracted from the cell lines indicated and was digested with restriction enzyme as indicated alongside plasmid control DNA. Digestion fragments were separated on a 0.8% agarose gel by gel electrophoresis at 20V for 24 hours and ethidium bromide intercalation was visualised by UV light.



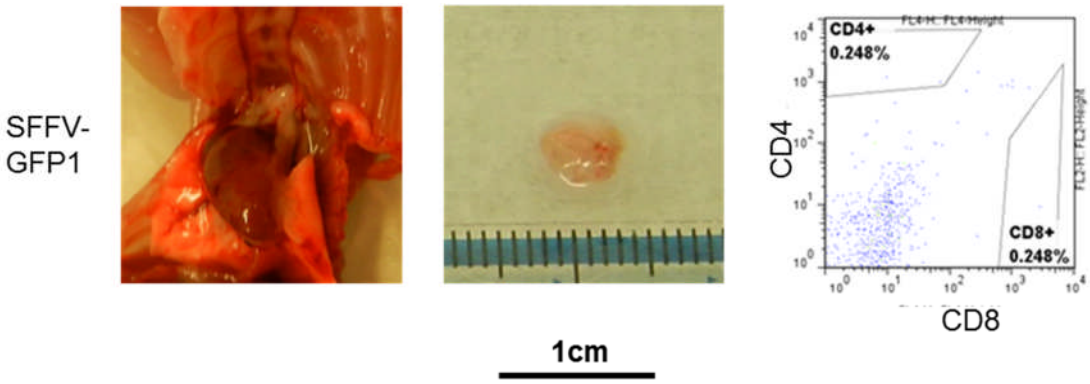
Appendix 9: Flow cytometry analysis of γc correction by nucleofection at 96 hours. $2x10^6$ ED7RγcΔ was nucleofected with 2µg of plasmid at different ZFN:Donor DNA ratios. Or 1.5µg of ZFN or donor DNA constructs. 2µg GFP expression plasmid was used instead of ZFN or donor DNA plasmid for negative control (0). Cells were stained for γc expression, acquired by flow cytometry and gated on PI-ve live cells.



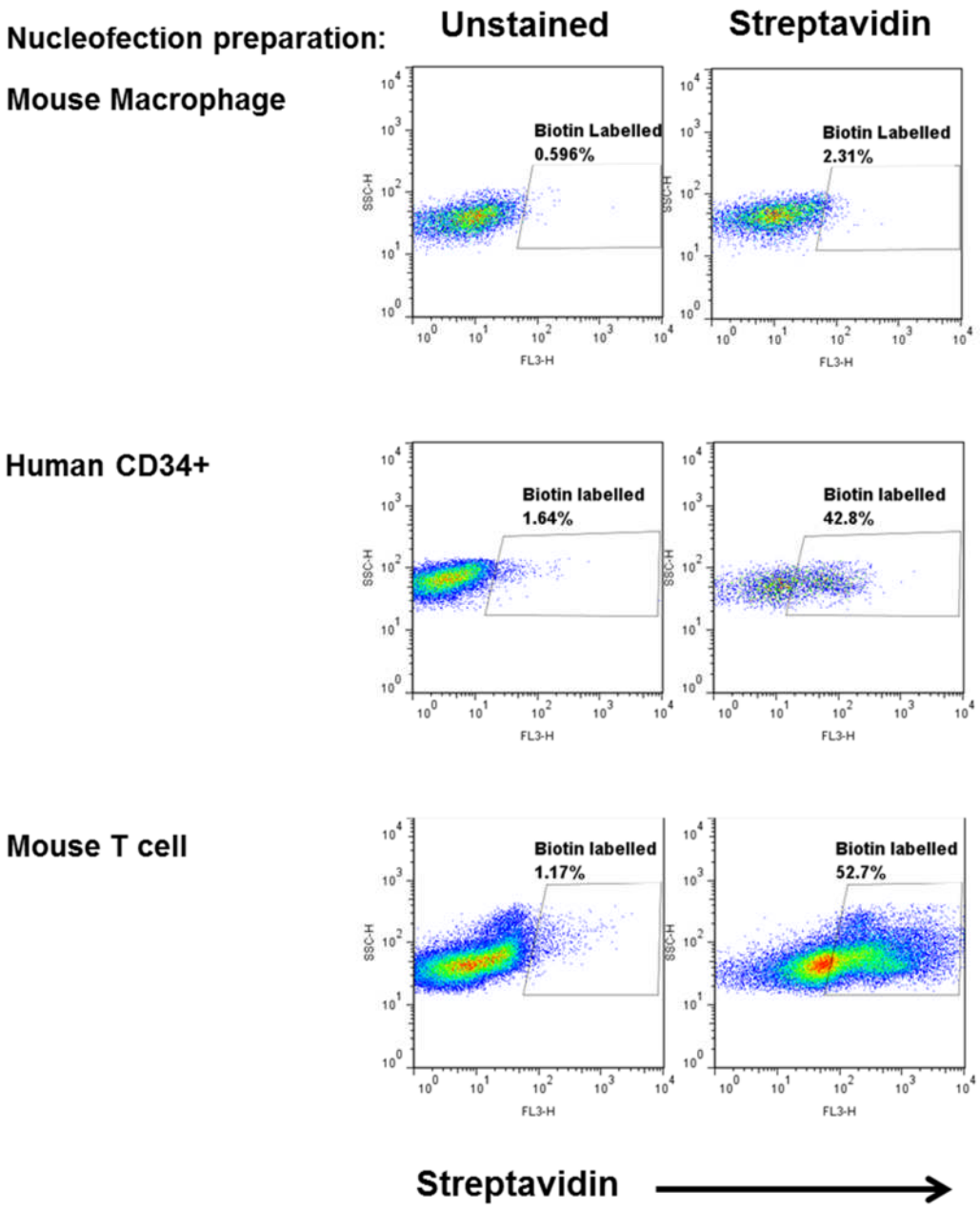
Appendix 10: Detection of pSTAT5 without IL-2 stimulation. ED7RγcΔ nucleofected with ZFN expression plasmids and donor DNA at a 3:1 ratio were cell sorted as either γc+ or γc- by FACS. Cells were starved overnight to prevent STAT5 phosphorylation and the cells were analysed for intracellular pSTAT5 without IL-2 stimulation by flow cytometry analysis.



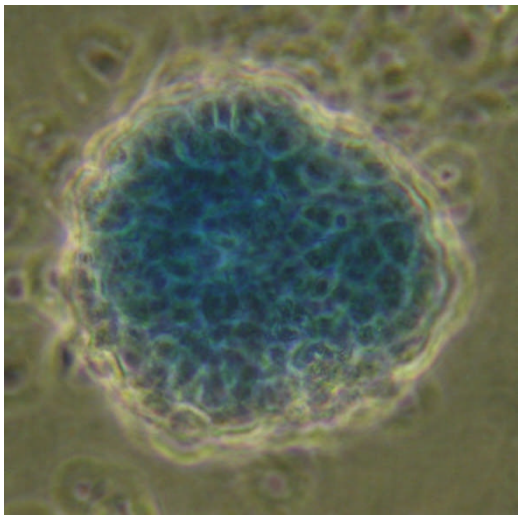
Appendix 11: Self inactivating lentiviral (SINLV) expression vector. Top, an enhanced GFP (eGFP) gene with an Spleen Focus Forming Virus (SFFV) promoter. Bottom, *IL2RG* cDNA with SFFV promoter (SF gc). Excerpt from the Thesis of Suzy Bailey.



Appendix 12: The thymus and associated T cells of a control Rag2-/-γc-/- mouse treated with hγcmut lin- cells treated with the integrating lentiviral vector SFFV-GFP. Left, the thymus *in situ*. Middle, the isolated thymus with a scale bar. Right, the CD4+ and CD8+ immune cell populations analysed by flow cytometry.



Appendix 13: The purity of lin- cells enriched from the bone marrow of hycmut mice in preparation for the nucleofection indicated. Murine bone marrow was stained with biotin labelled antibodies directed against differentiated blood cell markers and captured on a magnetic column by magnetic activated cell separation (MACS). The non-magnetic effluent was stained with Cy5-conjugated streptavidin to assess the contaminating, biotin labelled cells. Flow cytometry analysis of the stained cells compared to unstained was carried out for the cells prepared for nucleofection for the kits indicated.



Appendix 14: An mES embryoid body from the hγcmut mouse mES cell line. Alkaline phosphatase staining in blue.