

Development and Application of Non-Integrating Lentiviral Vectors for Gene Therapy

Luis Filipe Silvestre Apolonia

*Submitted to the University College London for the degree of
Doctor of Philosophy*

April 2009

Division of Infection and Immunity and Institute of Child Health

University College London

Declaration

I, Luis Filipe Silvestre Apolonia, 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

Lentiviruses stably integrate their genome into the host genome. Although this feature can be advantageous for long term transgene expression, it also has the potential to cause mutagenesis and cell transformation. To address this problem, thereby improving the safety of lentiviral gene therapy, non-integrating lentiviral vectors (NILVs) were developed.

NILVs were generated by mutating the *cis*-acting sequences that interact with integrase (att sites) or by mutating specific residues in integrase in different domains (catalysis - D64V, strand transfer – Q148A, K264R, K266R, K273R, DNA or chromatin binding - N120L, W235E). Relevant mutations were then combined in order to improve the safety of these vectors.

It was shown that all mutant vectors were efficiently produced and mutations did not affect infectivity. In contrast to dividing cells, differentiated muscle cells infected with NILVs show stable transgene expression over time without degradation of episomal viral DNA. The vectors were also tested *in vivo* by intramuscular injection in neonate mice. Transgene expression from muscle cells was maintained for 8 months using both integrating and NILVs.

The vectors were then tested in a haemophilia B disease model. It was shown that plasma levels of FIX produced by muscle cells infected with integrating lentiviral vectors were above the therapeutic threshold. However, expression from NILVs was lower. This was studied in detail and it was found that integrating lentiviral vectors are transcriptionally more active than NILVs. A comparison of expression levels revealed that integrated lentivectors express more transgene protein per vector copy than NILVs and AAV vectors, but both episomal vectors display similar levels of transgene expression per vector copy.

In conclusion, NILVs have the potential to be used as tools for prolonged transgene expression in non-dividing muscle cells or transient expression in dividing cells. However, vectors may need to be optimised if high expression levels are required.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Mary Collins and Adrian Thrasher for their patience, guidance, encouragement and for nurturing my growth as a scientist. It was an invaluable experience working with them. Nicola Philpott was also very helpful during the two years that she has supervised me and I would like to thank her.

It was a pleasure to work at the Molecular Immunology Unit (ICH) while being part of the Collins lab. I am grateful for the wonderful work environment created by all of the lab members. Many people have contributed towards this work. In particular, I would like to thank Conrad for all the help, discussion, advice and for being such a good friend. Also, I would like to thank Nat, Dip, Carolina, Mike, Gerben, Ste, Jo, Joao, Austen and Sue from the MIU crowd, to Ranbir, Marieke, Sofia, Helen and Kasia from the Collins lab, to Simon and Suzy from the Imperial College and to Matthew Weitzman, Dan, Caroline and Inigo from the Salk Institute. Without their help, this journey would have been much more difficult.

I would like to acknowledge the financial funding from Fundação para a Ciência e Tecnologia.

Finally, a few words to the people that I love, although no words can possibly describe how grateful I am to them. To Paula for all the support and encouragement throughout the good and bad times and for so much more. To my family for trusting my decisions, that sometimes take me away from them. To my mother and brother for always believing in me and giving me the strength to pursue my dreams. To my grandparents, Gertrudes and Jose, for teaching me the need for knowledge and persevere until I reach all my goals. To them I dedicate this thesis

Table of contents

Abstract	3
Acknowledgements	4
List of Figures	9
List of Tables	11
Abbreviations	12
Chapter One	15
1 – Introduction	16
1.1 – Retroviruses	17
1.1.1 – Biology of Retroviruses	17
1.1.2 – Taxonomy	17
1.1.3 – Retroviral pathogenesis.....	18
1.1.4 – Genome Structure.....	19
1.1.5 – Virion structure	20
1.1.6 – Virus Life cycle.....	22
1.1.6.1 – Entry	22
1.1.6.2 – Uncoating	23
1.1.6.3 – Reverse transcription.....	23
1.1.6.4 – Nuclear entry	26
1.1.6.5 – Integration	27
1.1.6.6 – Retroviral transcription and nuclear export.....	30
1.1.6.7 – Translation of viral proteins	33
1.1.6.8 – Assembly, budding and maturation.....	34
1.2 – Retroviral vectors	36
1.2.1 – Gammaretroviral vectors.....	36
1.2.2 – Lentiviral vectors	37
1.2.3 – Self-inactivating (SIN) vectors	39
1.2.4 – Expression cassette.....	40
1.2.5 – Pseudotyping	41
1.2.6 – Vector production.....	42
1.3 – Adenoviruses.....	44

1.3.1 – Structure and Biology of Adenoviruses	44
1.3.2 – Adenoviral vectors	45
1.4 – Adeno-Associated Viruses (AAV).....	48
1.4.1 – Structure and Biology of AAV	48
1.4.2 – AAV vectors.....	50
1.5 – Vectors in gene therapy.....	52
1.5.1 – Retroviral vectors in gene therapy: Success with consequences	53
1.6 – Project aims.....	56
Chapter Two	57
2 – Material and Methods.....	58
2.1 – Materials.....	58
2.1.1 - Reagents and Enzymes	58
2.1.2 - Antibodies	59
2.1.3 - Kits.....	59
2.1.4 - Centrifuges.....	60
2.2 – Molecular Biology Techniques	65
2.2.1 – Growth and maintenance of <i>E.coli</i>	65
2.2.2 – Production and transformation of electro-competent <i>E.coli</i>	65
2.2.3 – Production and transformation of chemical-competent <i>E.coli</i>	66
2.2.4 – Plasmid DNA preparation.....	66
2.2.5 – Cloning.....	67
2.2.5.1 – Restriction enzyme digestion	67
2.2.5.2 – Polymerase Chain Reaction (PCR)	67
2.2.5.3 – Isolation of DNA fragments.....	67
2.2.5.4 – Ligation of DNA fragments	68
2.2.6 – Genomic DNA Isolation	68
2.2.7 – Total RNA Isolation.....	69
2.2.8 – Site directed mutagenesis	69
2.3 – Tissue Culture Techniques.....	70
2.3.1 – Cell lines propagation	70
2.3.2 – Cell lines storage.....	70

2.3.3 – Lentiviral vector production.....	71
2.3.4 – Adeno-Associated vector production.....	71
2.3.5 – Vector transduction <i>in vitro</i>	72
2.3.6 – Vector infection <i>in vivo</i>	72
2.3.7 – Integration efficiency	73
2.4 – Analyses of <i>in vitro</i> and <i>in vivo</i> samples.....	73
2.4.1 – Flow cytometry	73
2.4.2 – Titration of p24 in lentiviral preparations	74
2.4.3 – Quantitative real-time PCR.....	74
2.4.4 – Enzyme-linked immunosorbent assay (ELISA).....	74
2.4.4.1 – Factor IX (FIX) ELISA	74
2.4.4.2 – Enhanced green fluorescent protein (eGFP) ELISA.....	75
2.4.5 – Immunohistochemistry.....	76
2.4.6 – Western Blot.....	76
2.4.7 – Southern Blot	77
2.4.8 – Confocal microscopy	78
2.4.9 – Statistical Analysis	78
Chapter Three.....	79
3 – Development of Non-Integrating Lentiviral Vectors	80
3.1 – Introduction	80
3.2 – Production of non-integrating lentiviral vectors	82
3.3 – Assessment of integration from NILVs <i>in vitro</i>	83
3.4 – Mutant vector expression characterisation.....	85
3.5 – Quantification of background integration	90
3.6 – Multiple mutant vectors	91
3.7 – Discussion	95
Chapter Four.....	98
4 – Application of Non-Integrating Lentiviral Vectors.....	99
4.1 – Introduction	99
4.2 – Test of NILVs in non-dividing cells	100
4.3 – Viral DNA forms in non-dividing cells	102

4.4 – Expression of NILVs <i>in vivo</i> in muscle	104
4.5 – Discussion	109
Chapter Five.....	111
5 –Correction of a Disease Model with Non-Integrating Lentiviral Vectors	112
5.1 – Introduction	112
5.2 – Production of vectors expressing FIX.....	113
5.3 –Vector expression of hFIX <i>in vivo</i>	115
5.4 – Optimisation of hFIX expression from muscle cells	118
5.5 – Discussion	121
Chapter Six.....	125
6 – Comparison of Expression from Episomal and Integrated Vectors	126
6.1 – Introduction	126
6.2 – Comparison of expression from integrated and non-integrated lentiviral vectors	127
6.3 – Comparison of expression from episomal and integrated viral vectors.....	131
6.3 – Comparison of expression from lentiviral vectors and transfected DNA	137
6.4 – Discussion	139
Chapter Seven	143
7 – Protein Delivery with Lentiviral Vectors	144
7.1 – Introduction	144
7.2 – Production of vectors for delivery of proteins	146
7.2 – Protein delivery to cells by lentiviral vectors.....	151
7.3 – Discussion	155
Chapter Eight.....	157
8 – Final Discussion	158
References.....	170

List of Figures

Chapter One

Figure 1.1: Representation of the genomic organisation of the retroviral genomes	20
Figure 1.2: Model of a retrovirus structure: the mature HIV virion	21
Figure 1.3: Overview of the reverse transcription process	25
Figure 1.4: Overview of the integration process.....	28

Chapter Three

Figure 3.1: Schematic illustration of the mutations created in the integrase and in the viral genome.....	82
Figure 3.2: Transgene expression profile over time of cells infected with mutant vectors.	84
Figure 3.3: Southern Blot of DNA harvested from cells infected with integrating and non-integrating vectors.....	87
Figure 3.4: Viral DNA quantification in infected cells with lentiviral vectors.....	89
Figure 3.5: Quantification of background integration of mutant vectors.	90
Figure 3.6: Transgene expression profile over time of cells infected with multiple mutant vectors.	93
Figure 3.7: Quantification of background integration of multiple mutant vectors.	94

Chapter Four

Figure 4.1: Expression profile in dividing and non-dividing muscle cells over time.	101
Figure 4.2: Viral DNA quantification in non dividing cells.	103
Figure 4.3: Transgene expression from muscle <i>in vivo</i>	105
Figure 4.4: Identification of eGFP expressing muscle cells <i>in vivo</i> by immunohistochemistry.....	106
Figure 4.5: Quantification of eGFP protein from <i>in vivo</i> samples.....	107
Figure 4.6: Quantification of viral 2LTR circles in <i>in vivo</i> samples.....	108

Chapter Five

Figure 5.1: Expression of hFIX <i>in vitro</i>	114
Figure 5.2: hFIX expression <i>in vivo</i>	115
Figure 5.3: Comparison of hFIX expression from integrating and non-integrating lentiviral vectors <i>in vivo</i>	116
Figure 5.4: Expression of hFIX in haemophiliac mice	117
Figure 5.5: Comparison of expression of different codon optimised hFIX versions <i>in vitro</i>	119
Figure 5.6: Comparison of expression of different hFIX transgenes <i>in vivo</i>	120

Chapter Six

Figure 6.1: Comparison of expression from different promoters in integrating and non-integrating lentiviral vectors <i>in vitro</i>	128
Figure 6.2: Analysis of DNA, RNA and Protein in cells infected with lentiviral vectors.	130
Figure 6.3: Expression from different integrating and episomal viral vectors <i>in vitro</i>	133
Figure 6.4: Expression comparison between integrated and episomal viral vectors	134
Figure 6.5: Relation between expression and vector DNA concentration <i>in vitro</i>	136
Figure 6.6: Expression comparison between integrating lentivectors, NILVs and transfected DNA in 293T cells.....	138

Chapter Seven

Figure 7.1: Schematic representation of the strategy adopted to insert a foreign protein in the P2 region of gag.....	147
Figure 7.2: Infectivity of vectors harbouring mutations in the P2 region of gag.....	148
Figure 7.3: Western blot of lentiviral vectors carrying eGFP protein	150
Figure 7.4: Protein delivery to cells from vectors carrying eGFP	152
Figure 7.5: Visualisation of cells infected with a lentiviral vector carrying eGFP protein.....	154

List of Tables

Chapter One

Table 1.1: Genus of the Retroviridae family.....	17
Table 1.2: Vectors used in gene therapy	53

Chapter Two

Table 2.1: Buffers and solutions	61
Table 2.2: Primers and probes.....	62

Chapter Three

Table 3.1: Quantification of viral particles and infectivity	83
Table 3.2: Expression levels (MFI) of eGFP positive cells infected with vectors.....	86
Table 3.3: Quantification of viral particles, infectivity and expression levels (MFI) of multiple mutant vectors	92

Abbreviations

AAT	α 1-antitrypsin
AAV	Adeno-associated virus
Ad	Adenovirus
AIDS	Acquired immunodeficiency syndrome
att	Attachment site
BAF	Barrier-to-autointegration factor
CA	Capsid
CDk9	Cyclin-dependent protein kinase 9
cDNA	Complementary DNA
CGD	Chronic granulomatous disease
CIP	Calf intestinal alkaline phosphatase
CMV	Cytomegalovirus
<i>c-onc</i>	Cellular oncogene
cPPT	Central polypurine tract
CTE	Constitutive export element
DBP	DNA-binding protein
DC	Dendritic cell
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic Acid
DNW	D64V + N120L + W235E
DSB	Double stranded break
<i>E.coli</i>	<i>Escherichia coli</i>
eGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immunosorbent assay
ESCRT	Endosomal sorting complex required for transport
FIX	Factor IX
GALV	Gibbon ape leukaemia virus
hFIX	Human Factor IX
HIV	Human immunodeficiency virus

HMGA1	High mobility group chromosome protein A1
HR	Homologous recombination
HSC	Haematopoietic stem cells
IN	Integrase
INI1	Integrase interactor 1
IRES	Internal ribosome entry site
ITR	Inverted terminal repeat
LAP2 α	lamina-associated polypeptide 2 α
LB	Luria-Bertani
LEGF/p75	Lens epithelium growth factor
LMO-2	LIM domain only-2
LTR	Long terminal repeat
MA	Matrix
MFI	Mean intensity fluorescence
MLV	Murine leukaemia virus
MOI	Multiplicity of infection
mRNA	Messenger RNA
NC	Nucleocapsid
Neo	Neomycin phosphotransferase
NF κ B	Nuclear factor-kappa B
NILV	Non-integrating lentiviral vector
NLS	Nuclear localisation signal
PBS	Primer binding site
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PIC	Pre-integration complex
PPT	Polypurine tract
PR	Protease
PRE	Post-translational regulatory element
qPCR	Quantitative real-time polymerase chain reaction
R	Repeat sequence

RNA	Ribonucleic acid
Rnase	Ribonuclease
RRE	Rev-responsive element
RSV	Rous sarcoma virus
RT	Reverse transcriptase
RTC	Reverse transcription complex
SCID	Severe combined immunodeficiency
SFFV	Spleen focus forming virus
SIN	Self-inactivating
ssDNA	Single stranded DNA
-sssDNA	Minus-strand strong stop DNA
SU	Surface glycoprotein
SV40	Simian virus 40
TAR	Trans-activation responsive region
TM	Transmembrane glycoprotein
TP	Terminal protein
tRNA	Transfer RNA
TU	Transducing units
U3	Unique in 3'
U5	Unique in 5'
UCOE	Ubiquitously acting chromatin opening element
<i>v-onc</i>	Viral oncogene
VP	Viral protein
VSVg	Vesicular stomatitis virus glycoprotein
WPRE	Woodchuck hepatitis B virus post-translational regulatory element
wt	Wild type
ZFN	Zinc finger nuclease
γc	Common cytokine receptor gamma-chain
Ψ	Packaging signal

Chapter One

Introduction

1 – Introduction

The understanding of molecular pathologies and the development of vectors that can mediate transfer of DNA into cells created a new approach in the treatment of diseases (Friedman and Roblin, 1972). Thus, gene therapy can be defined as a technique used for the delivery of genetic material into cells. Translation of this technology to clinic was initially disappointing but gene therapy became accepted as a new tool in medicine when reports of clinical trials showed that a clear clinical benefit was achieved using this technique (Aiuti *et al.*, 2002; Cavazzana-Calvo *et al.*, 2000; Gaspar *et al.*, 2004; Ott *et al.*, 2006).

The first clinical trial was initiated in 1989 in the United States. The aim was the treatment of patients with advanced cancer with autologous tumour infiltrating lymphocytes modified with a retroviral vector. This study demonstrated the successful introduction of a gene coding for the resistance to Neomycin by retroviral transduction (Rosenberg *et al.*, 1990). Over the last 20 years, more than 1400 clinical trials have been approved, with cancer being the target disease in more than 65% of the studies. Other applications where gene therapy has been used in clinical trials are cardiovascular diseases, inherited monogenic disorders and infectious diseases.

In order to improve this technology, it is clear that a profound understanding of the mechanisms employed by the vectors is necessary to achieve efficient gene delivery and expression. This introduction will then focus on the biology of viruses and the development of viral based vectors. Particular attention will be given to lentiviruses, which were the vectors under research in this project. Adeno-associated viruses (AAV) will also be reviewed as they were used in the work presented here. Adenoviruses will be covered because vectors based on this virus represent an important tool for gene therapy (one quarter of the total gene therapy trials have used these vectors).

1.1 – Retroviruses

1.1.1 – Biology of Retroviruses

Retroviruses comprise a large family of viruses that have particular features: their genome is a single-stranded RNA molecule that serves as a template for the formation of double-stranded DNA, in a process called reverse transcription, and subsequent integration into the genome of the host.

1.1.2 – Taxonomy

In the latest convention, the family Retroviridae is divided into seven genera (Index of Viruses - Retroviridae, 2006). Genetic structure and virion morphology are the main characteristics by which viruses are divided into each group (Table 1.1). Due to their importance in gene therapy, particular focus will be given to gammaretroviruses and lentiviruses in this introduction.

Table 1.1: Genus of the Retroviridae family

Genus	Species example	Virion morphology	Genome
Betaretrovirus	Mouse mammary tumour virus	eccentric, spherical core	simple
Gammaretrovirus	Murine leukaemia virus (MLV)	concentric, spherical core	simple
Alpharetrovirus	Avian leukosis virus	concentric, spherical core	simple
Deltaretrovirus	Bovine leukaemia virus	concentric, spherical core	simple
Lentivirus	Human immunodeficiency virus I (HIV 1)	concentric, cone-shaped core	complex
Epsilonretrovirus	Walleye dermal sarcoma virus	concentric, spherical core	simple
Spumavirus	Human foamy virus	central, spherical core	complex

1.1.3 – Retroviral pathogenesis

Retroviruses are associated with a wide number of diseases, such as cancer, neurological disorders and immunodeficiencies. Nowadays, the best known example of a disease induced by a retrovirus is acquired immunodeficiency syndrome (AIDS), caused by HIV. However, it was much earlier, in the beginning of the 20th century, that leukaemia and sarcoma were demonstrated to be induced in chickens within days of infection by transmissible agents, later shown to be the retroviruses avian leucosis virus and Rous sarcoma virus, respectively.

Many retroviruses, including MLV induce tumours. Acutely transforming viruses contain oncogenes (*v-onc*) in their genome, but it has been shown that the origin of *v-onc* genes is in fact, cellular (*c-onc*) (Stehelin *et al.*, 1976). This was an important find in cancer research that lead to the discovery of many cellular oncogenes. These viruses acquired at some point *c-onc* sequences by recombination and upon infection, they deliver these genes to the target cells, leading rapidly to cell transformation. However, when the oncogene is incorporated into the viral genome, some of the viruses are rendered replication defective due to loss of viral sequences.

Several retroviruses that cause tumours do not contain an oncogene in their genome and are replication competent. Transformation arises with the integration of the provirus in the host genome. This integration can activate proto-oncogenes, either by transcription from the viral promoter (Shen-Ong *et al.*, 1986) or from activation of cellular promoters by the viral enhancer (Lazo *et al.*, 1990), leading to cell transformation. The provirus could also be inserted within transcription units, resulting in coding sequence disruption or premature termination of the transcript. The resulting cellular protein can then lack important negative regulatory domains, generating a constitutively active oncoprotein (Ceci *et al.*, 1997). Another described mechanism by which the provirus integration can induce transformation is the stabilisation of RNA by the disruption of AU-rich elements that reduce the half life of mRNA. The mRNA remains in the cytoplasm for a longer period of time, leading to an increased protein translation (Selten *et al.*, 1985).

HIV was first isolated in 1983 (Barre-Sinoussi *et al.*, 1983). When this lentivirus infects a host, an acute immunosuppression may develop and, as a result, opportunistic

infections may arise. However, HIV infection can be asymptomatic for many years. During this time, there is a gradual decline of CD4⁺ T cells and the host becomes more susceptible to opportunistic infections such as pneumonia (caused by *Pneumocystis carinii* pneumonia), Kaposi sarcoma (caused by Herpes simplex virus 8), Non-Hodgkins lymphoma and anogenital cancers (reviewed in Boshoff and Weiss, 2002).

HIV replicates most efficiently in activated CD4⁺ T cells. However, these cells have a short life-span of a few days, so latent infection is established in resting/quiescent memory CD4⁺ T cells and dendritic cells (DCs). These cells do not support replication during latency, but are important reservoirs for the provirus. DCs are not infected as efficiently as CD4⁺ T cells, but they can capture virions via c-type lectins in the cell surface, storing them in intracellular compartments without degradation of the viruses. Upon migration to the lymph node, DCs can then transfer these infectious virions to CD4⁺ T cells via cell-to-cell contact (McDonald *et al.*, 2003). Afterwards, the viruses can rapidly disseminate to other lymphoid compartments. The infected CD4⁺ T cell population is destroyed by direct cytopathic effect of the virus and it is now thought that the initial drop in viremia is a consequence of the exhaustion of the pool of these cells (Phillips, 1996). This population is then replenished by naïve and central memory T cells that had not been infected with the virus, stimulating a generalised state of immune activation. These activated CD4⁺ T cells provide a source for HIV replication and it is thought that the loss of this ability to regenerate effector memory T cells leads to the eventual drop in overall CD4⁺ T cells number, increasing the susceptibility for opportunistic infections (reviewed in Picker, 2006).

1.1.4 – Genome Structure

The general genome organisation of retroviruses comprises the genes *gag*, *pol* and *env*, flanked by the unique 3' (U3) and repeat (R) sequences downstream and R and unique 5' (U5) upstream (Figure 1.1). In the provirus DNA, the genes are flanked by a long terminal repeat (LTR), which can be segmented into U3-R-U5. These sequences are formed upon reverse transcription, due to the jumps of the reverse transcriptase (RT) between templates. U3 contains most of the transcription control elements, including the

promoter and enhancer sequences. The polyadenylation site lies between the R and U5. Therefore, the promoter is only upstream of the coding sequences after reverse transcription (Hughes *et al.*, 1978; Shinnick *et al.*, 1981).

The *gag* gene encodes for the structural proteins of the virion, such as matrix (MA), capsid (CA) and nucleocapsid (NC). *pol* encodes for the viral enzymes protease (PR), reverse transcriptase (RT) and integrase (IN). The transmembrane (TM) and surface (SU) glycoproteins found in the envelope of the virus are encoded by the *env* gene.

Complex retroviruses contain additional genes. HIV 1, for example, has the regulatory genes *tat* and *rev* and the accessory genes *nef*, *vpu*, *vpr* and *vif*.

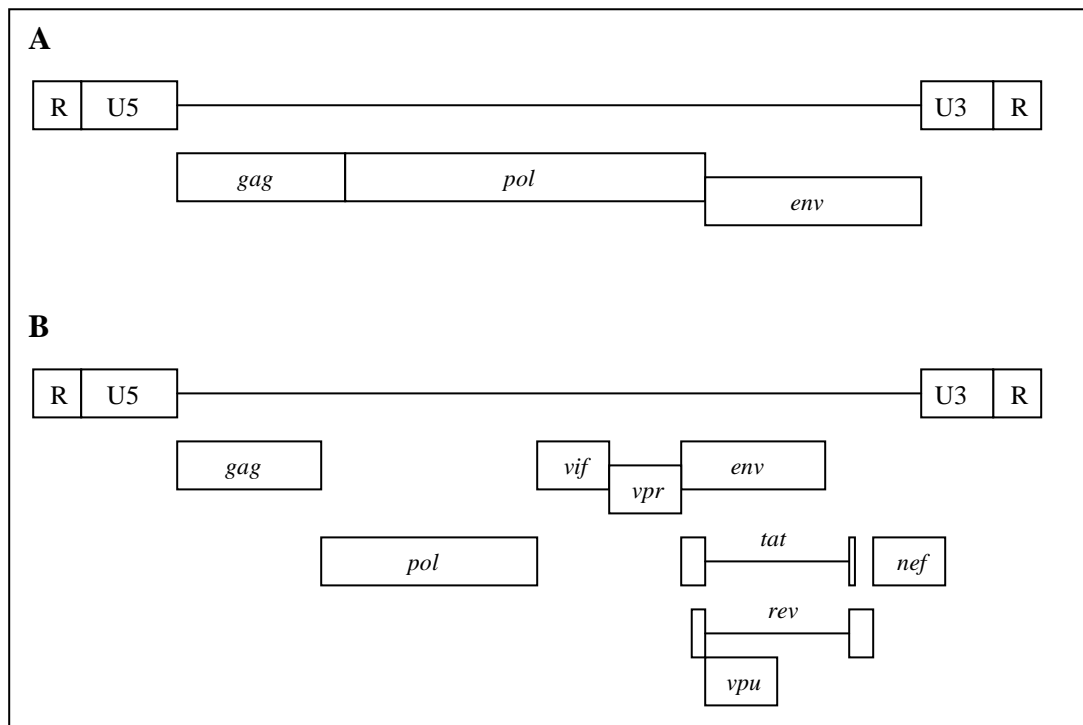


Figure 1.1: Representation of the genomic organisation of the retroviral genomes
A: MLV; **B:** HIV 1.

1.1.5 – Virion structure

The particle structure (virion) reflects the ability of retroviruses to infect new host cells, providing a safe and relatively stable environment outside cells for the accommodation of their genome. The structural proteins of the retroviruses condense the viral genome into a RNA-protein complex that is packaged into protein shells.

The particles of the RSV were described first in 1947, but only after a few years and with techniques employing electron microscopy, the ultrastructure of retroviruses was validated (Gaylord, 1955). Mature particles measure between 80-120nm, have a dense core, that can be cone-shaped (as for HIV-1) or round-shaped (as for MLV) and an outer lipidic membrane.

The envelope contains the viral glycoproteins TM anchoring SU. These multimers can form spikes projecting towards the outside of the particle. Matrix lies just underneath the envelope. Further inside, the capsid proteins form a shell, defining the outer layer of the virion core. Inside the core are the nucleocapsid-RNA complex and also other proteins such as RT and IN. Two RNA copies are packaged in each virion. A model of the HIV virion structure can be seen in Figure 1.2. The virions additionally contain proteins derived from the host cell (Arthur *et al.*, 1992).

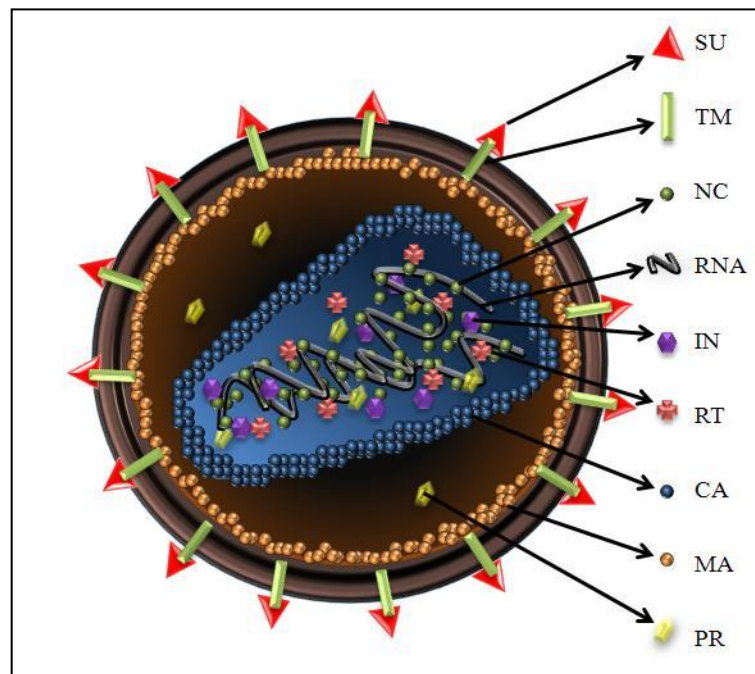


Figure 1.2: Model of a retrovirus structure: the mature HIV virion

Env proteins are shown jutting out from the lipid membrane. The cone-shaped core is also represented. Not all proteins incorporated in virion particles are shown in this model. Viral proteins in this model: surface glycoproteins (SU), transmembrane glycoproteins (TM), matrix (MA), capsid (CA), nucleocapsid (NC), retroviral protease (PR), reverse transcriptase (RT) and integrase (IN).

1.1.6 – Virus Life cycle

1.1.6.1 – Entry

The first step in the virus life cycle is the entry of the virion in the target cell. The SU protein protrudes from the particle and makes a contact with specific receptors in the target cell. The receptors define the target cells and host of each virus. Virions lacking the envelope glycoproteins are non-infectious and cells lacking the receptor are non-permissive towards the virus.

The disease that HIV induces in humans is characterised by the depletion of a subset of T-cells. This specificity was correlated with the ability of infection of CD4⁺ lymphocytes. The primary receptor of HIV-1 was thus identified as CD4, a cell surface marker of helper T cells that is also expressed on monocytes, macrophages and dendritic cells (Dalglish *et al.*, 1984; Klatzmann *et al.*, 1984). Direct evidence was shown by Maddon and colleagues, upon expression of this receptor in HeLa cells, changing the permissivity of these cells towards HIV-1 (Maddon *et al.*, 1986). Nonetheless, since not all human cells expressing the CD4 are infected with HIV-1, this virus requires an additional receptor for efficient fusion. The chemokine receptors CXCR4 (Feng *et al.*, 1996) or CCR5 (Deng *et al.*, 1996; Dragic *et al.*, 1996) were shown to be required as co-receptors to mediate membrane fusion.

After the initial binding between SU and the receptor, fusion is mediated by the TM protein, as a result of conformational changes. Interestingly, the glycoprotein of the Vesicular stomatitis virus (rhabdoviridae family), a commonly used glycoprotein for the production of viral vectors, requires the acidic environment of the endosome to activate fusion. For that virus, after initial contact with the cell receptor, it is taken into coated pits, undergo endocytosis and fusion happens between the virus and the endosome membranes.

1.1.6.2 – Uncoating

The fusion of viral and cellular membranes delivers the viral core into the cytoplasm, where uncoating and reverse transcription takes place. Disassembling of virus capsid appears to be essential for progress of reverse transcription but is one of the least known processes of retroviruses.

Cyclophilin A, a host protein, is thought to be involved in the entry or uncoating process of HIV. This protein binds to CA (Luban *et al.*, 1993) and the inhibition of this interaction reduces virion infectivity in a step prior to reverse transcription but after receptor binding and membrane fusion (Braaten *et al.*, 1996; Franke *et al.*, 1994).

The uncoating of HIV is rapid, in contrast to MLV where the CA protein is associated with the reverse transcription complex (RTC) throughout reverse transcription (Fassati and Goff, 1999; Fassati and Goff, 2001). Nevertheless, the RTC in both viruses contain at least IN and RT and also Vpr, MA and NC for HIV (Bukrinsky *et al.*, 1993; Nermut and Fassati, 2003).

1.1.6.3 – Reverse transcription

Reverse transcription is the hallmark of the Retroviridae and is performed by the viral enzyme reverse transcriptase (RT). RT possesses RNA- and DNA-dependent DNA polymerase activity (Baltimore, 1970; Temin and Mizutani, 1970) and also an ribonuclease (RNase H) domain that degrades RNA in RNA:DNA duplexes (Molling *et al.*, 1971).

Reverse transcription initiates with the annealing of a tRNA primer to the viral genomic RNA. All retroviruses encapsidate host tRNA and virions contain a large excess of 50-100 free different tRNA molecules. Nevertheless, the use of a specific tRNA for primer is different between genera. HIV-1 uses tRNA^{Lys3} (Jiang *et al.*, 1993) for reverse transcription initiation whereas MLV uses tRNA^{Pro} (Peters *et al.*, 1977).

The primer binds to a region called the primer binding site (PBS) and the viral enzyme initiates the synthesis of a negative single stranded DNA. Although initiation can occur inside the virion (Trono, 1992), extension and completion of the process is accomplished in the cytoplasm of the host cell.

After initiation, synthesis proceeds to the 5' end of the viral genome, so the unique in 5' (U5) and repeat (R) regions are encoded on this minus-strand strong stop DNA (-sssDNA). Reverse transcriptase degrades the RNA template afterwards with its RNase H activity. The R region is redundant in the viral genome (hence its name), allowing for a jump of the -sssDNA from the R in 5' to the R in 3' of the viral genome (Mitra *et al.*, 1979) once the RNA that served as template has been degraded (Tanese *et al.*, 1991). After this strand transfer, the -sssDNA is extended up to the 5' of the template RNA. RT degrades the template RNA while it is forming the -DNA strand except in the polypurine tract (PPT). This RNA sequence is then used for initiation of the +DNA strand (Sorge and Hughes, 1982; Finston and Champoux, 1984; Smith *et al.*, 1984). Additionally, other RNA sequences, such as the central polypurine tract (cPPT) in HIV, are not degraded and are also used as primers for +DNA strand synthesis, creating discontinuous DNA products (Charneau and Clavel, 1991; Varmus *et al.*, 1978). The +DNA strand is then synthesised using the -sssDNA as a template, stopping at the tRNA primer. The tRNA primer is degraded by the RNase H activity of RT, facilitating an intramolecular second jump. Under the current accepted model, the PBS region at the 3' end of the -DNA strand anneals to the complementary PBS of the +DNA strand, allowing extension of the -DNA strand using the +DNA strand as a template. In this way, the U3, R and U5 regions are duplicated, forming the LTR of the viral DNA. The +DNA strand is also extended, using the -DNA strand as a template, completing the reverse transcription process. The final product is a blunt-ended double stranded DNA that serves as a template for integration (Brown *et al.*, 1989). A schematic representation of the process can be seen in Figure 1.3.

In HIV, the synthesis of the +DNA strand initiated at the PPT displaces the plus strand initiated at the cPPT. The displacement creates a discontinuous DNA (DNA flap), that appears to be important in steps of the virus life cycle posterior to reverse transcription (Charneau *et al.*, 1994). The discontinuous DNA is efficient for integration. Therefore it is possible that cellular host proteins complete the DNA synthesis (Miller *et al.*, 1995).

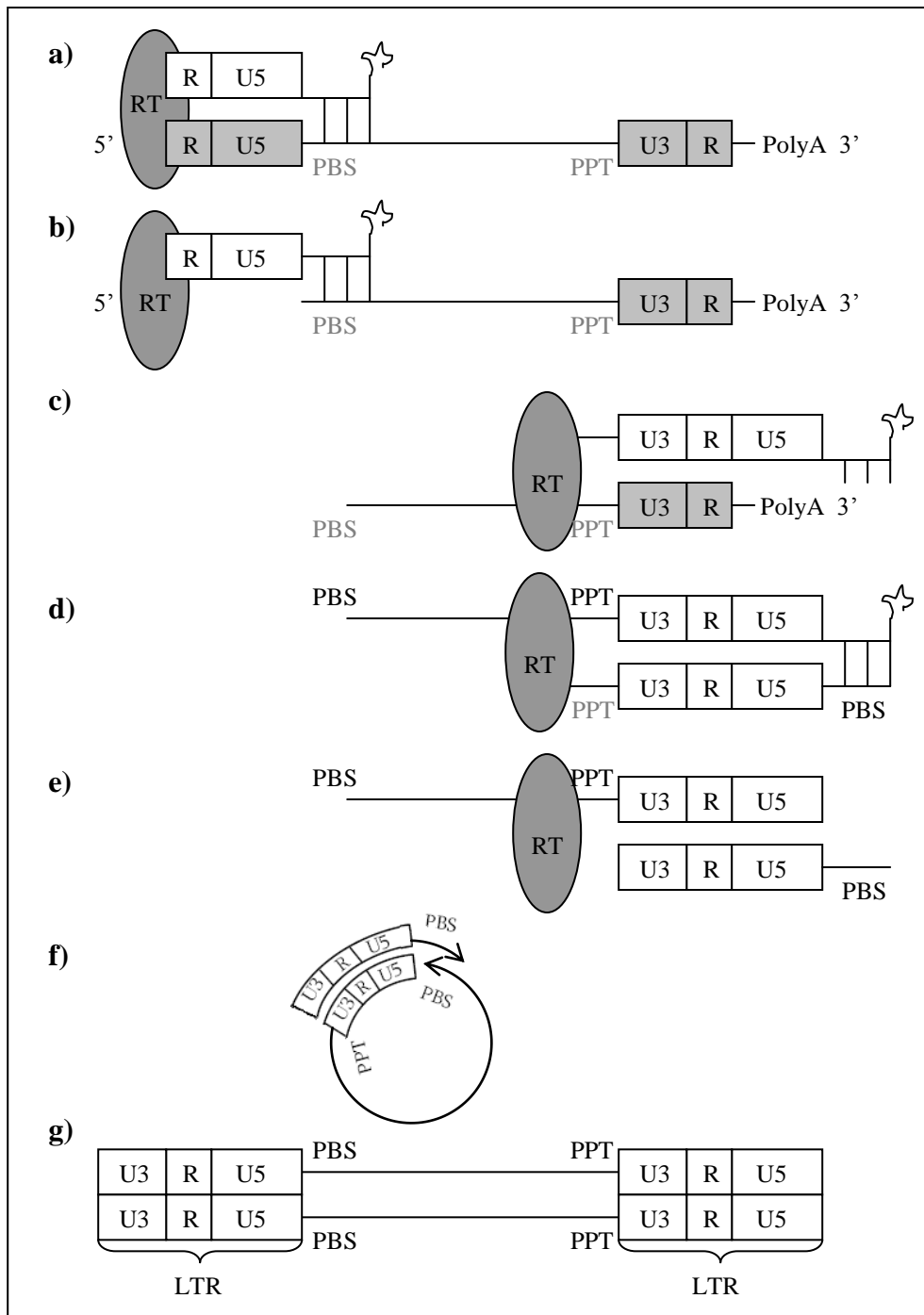


Figure 1.3: Overview of the reverse transcription process

a) minus strand DNA synthesis initiated from the annealed tRNA primer to the primer binding site (PBS) in the RNA genome by the reverse transcriptase (RT); **b)** RNase H degradation of the R and U5 sequences in the genomic RNA; **c)** first strand transfer, with annealing of the -ssDNA R sequence to the complementary R in the 3'RNA, -DNA strand synthesis with further digestion of RNA template, except for the polypurine tract (PPT); **d)** tRNA removal by RNase H activity, synthesis of +DNA with the PPT as primer; **e)** second strand transfer through the annealing of the PBS region; **f)** reverse transcription with strand displacement. Grey boxes and letters represent regions of the genomic RNA, white boxes represent provirus DNA.

1.1.6.4 – Nuclear entry

While the reverse transcription takes place, the RTC migrates to the nucleus, using the cellular cytoskeleton. Initial movements of the virus occur in association with the actin cytoskeleton, but subsequent translocation to the nucleus occurs along the microtubule network (Arhel *et al.*, 2006). It is likely that when the RTC reaches the nucleus, reverse transcription is finished and RT disassociates from the complex, reducing the size of the complex, allowing the translocation through the nuclear pore (Bukrinsky, 2004).

The mechanism of nuclear import of the viral DNA is still unknown. Gammaretroviruses, such as MLV require mitosis to access the nucleus and thus, these viruses can only infect dividing cells (Roe *et al.*, 1993). However, lentiviruses can infect both dividing and non-dividing cells with similar efficiency (Weinberg *et al.*, 1991; Lewis *et al.*, 1992).

In lentiviruses, it is probable that nuclear import requires the use of the nuclear pore complex. Nonetheless, a mechanism of disruption of the nuclear membrane by Vpr has also been proposed (De Noronha *et al.*, 2001). Nuclear cellular proteins contain a nuclear localisation signal (NLS). MA, IN and Vpr are associated with the viral DNA, in the pre-integration complex (PIC) and are candidates for nuclear import as they have putative NLS. Mutations in the putative NLSs have shown reduced nuclear import of viral DNA, although replication was not impaired (Yamashita and Emerman, 2005). The DNA flap has been also implicated in nuclear import (Zennou *et al.*, 2000). Nevertheless, absence of the cPPT and hence the lack of the DNA flap, was shown to only reduce viral infectivity (Dvorin *et al.*, 2002). Some experiments have not been replicated by other labs and some experiments are still debated, so consensus has not been reached yet about the mechanism. Possible models proposed imply that nuclear import involves more than one karyophilic protein or a still unknown NLS for binding to cellular nuclear importins and that the DNA flap gives the best conformation for nuclear import (Bukrinsky, 2004). Other hypothesis rely on the velocity of the uncoating process of the lentiviruses and also the cell cycle of the target cell (Yamashita and Emerman, 2006). Additionally, cellular factors have been shown to be involved in the nuclear import of HIV-1, including importin 7 (Fassati *et al.*, 2003) and a 3' CCA deficient tRNA species that is incorporated into HIV-1 virions (Zaitseva *et al.*, 2006).

1.1.6.5 – Integration

Integration of the provirus into the host DNA is an important step of the retroviruses life cycle. It ensures that the provirus DNA is replicated along with the host DNA and transmitted to the progeny as an element of the host chromosomes (Hughes *et al.*, 1978; Temin, 1964).

The 32kDa viral integrase is the key protein in integration. This protein is a cleavage product of the gag-pol polyprotein and contains three major domains. The N-terminal domain contains a zinc binding motif (HHCC) (Johnson *et al.*, 1986). IN can indeed bind a zinc ion (Burke *et al.*, 1992) and zinc promotes multimerisation of IN and increases the enzymatic activity (Zheng *et al.*, 1996). In the central domain lies the catalytic core, the conserved DD35E motif (Engelman and Craigie, 1992). The C-terminal domain has DNA binding properties (Mumm and Grandgenett, 1991). IN binds the viral DNA in the attachment sites (att), which are conserved CA dinucleotides situated at the end of the LTR (Sherman and Fyfe, 1990).

Integration occurs in two catalytic steps. In the first, known as 3'-processing, IN cleaves and removes a dinucleotide adjacent to the att sites and at the end of the viral DNA (Katzman *et al.*, 1989). This reaction occurs in the cytoplasm within the pre-integration complex. The second catalytic step, called strand transfer, happens in the nucleus. Here, IN mediates a concerted nucleophilic attack by the 3'-hydroxyl residues of the viral DNA on phosphodiester bridges located on either side of the major groove in the target DNA and a transesterification reaction where the viral 3' ends are ligated to the 5'-O-phosphate residues of the target DNA (Bushman and Craigie, 1991). In the final step of integration, which is probably mediated by cellular enzymes, the unpaired dinucleotides from the 5' ends of the viral DNA are removed and the single-stranded gaps created between the viral 5' end and target DNA are filled. Figure 1.4 represents a schematic overview of the integration process. In lentiviruses, the sites of integration in the target DNA are separated by five bases. Therefore, there is a duplication of these bases in the final product, so the 5' and 3' ends of the integrated DNA are flanked by the same 5 bases (Craigie, 2001). For MLV, the repeats flanking the integrated proviruses have 4 bases (Van Beveren *et al.*, 1980).

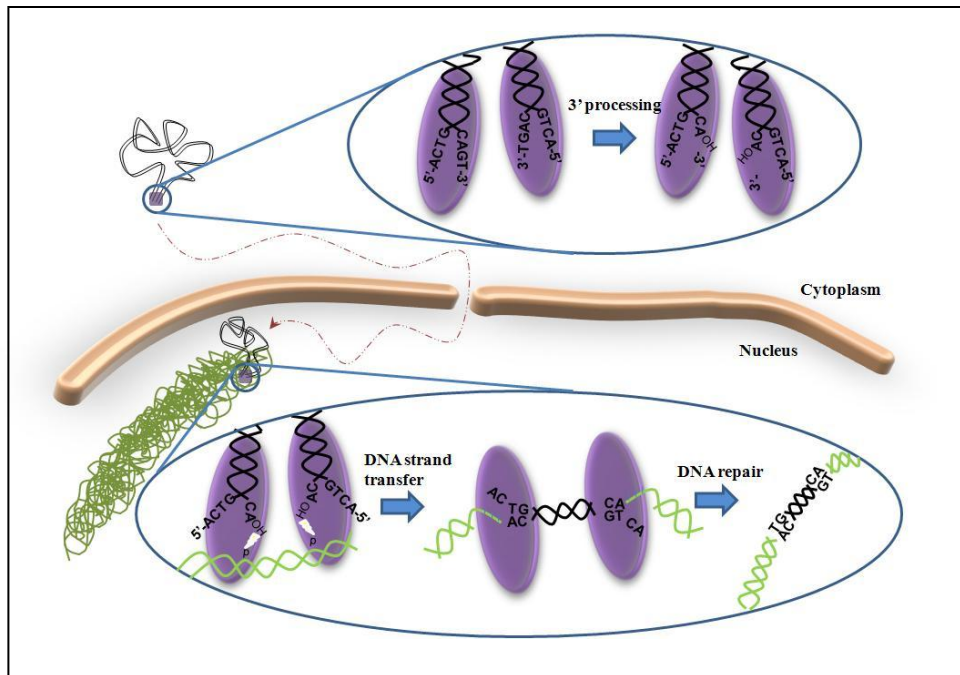


Figure 1.4: Overview of the integration process

Viral DNA is represented in black, host DNA in green and viral integrase in purple. In the cytoplasm, two dinucleotides on the 3' of each viral DNA strand are removed by integrase (3' processing). In the nucleus, integrase mediates a nucleophilic attack by the hydroxyl residues in the viral DNA strand on the phosphodiester bridges of the host DNA and ligation of the two DNA molecules (DNA strand transfer). The removal of the unpaired viral dinucleotides on the 5' viral DNA strand and filling of the single stranded gaps in the host genome (DNA repair) is thought to be mediated by host proteins.

In addition to the integrated provirus, there are other extrachromosomal viral DNA species (Gianni *et al.*, 1975; Varmus *et al.*, 1974). These molecules can either be linear or circular. The later can have one LTR (formed by homologous recombination between the two LTRs) or two LTRs produced by non-homologous end-joining of the double stranded blunt viral DNA produced upon reverse transcription (Ju and Skalka, 1980). Additionally, there are other viral DNA circles produced by an intramolecular integration when the viral ends of a DNA molecule uses itself as a target (Shoemaker *et al.*, 1980). The formation of circular genomes is thought to occur in the nucleus. As a consequence, these molecules have been used to quantify nuclear viral DNA. It was thought that these molecules were by-products of integration, but recent finding suggested that this non-integrated viral DNA

could support transcription (Brussel and Sonigo, 2004; Engelman *et al.*, 1995; Poon and Chen, 2003; Wu and Marsh, 2003).

The integration site potential of the provirus is very big, spanning virtually the whole of the host genome (Withers-Ward *et al.*, 1994). Nonetheless, it appears that each virus has specific preferential sites. HIV provirus has been shown to integrate preferentially anywhere within transcriptional units, particularly in highly active genes (Mitchell *et al.*, 2003; Schroder *et al.*, 2002). MLV has a different integration pattern. This virus strongly favours CpG islands (whereas HIV does not) and integrates the provirus preferentially near transcription start sites (either upstream or downstream) of actively transcribed genes (Wu *et al.*, 2003). The differences in the selection of integration target sites may be due to the viral integrase, cellular proteins or regional features in the target DNA.

It has been shown that integration in nucleosomes occurs more efficiently compared with nucleosome-free regions and that preferential targets occur where the major groove is exposed. Also, integration is prevented in regions occupied by site-specific DNA-binding proteins (Pryciak and Varmus, 1992).

Several cellular proteins are known to bind the pre-integration complex and therefore may regulate the preferential target site for integration. Lee and Craigie have shown that there is a cellular protein bound to the PIC that prevents the “suicidal” auto-integration, called barrier-to-autointegration factor (BAF) (Lee and Craigie, 1994; Lee and Craigie, 1998). Also, the lamina-associated polypeptide 2 α (LAP2 α) interacts with BAF and was recently shown to be a component of the MLV PIC. The depletion of LAP2 α in cells decreased MLV replication, demonstrating a contribution to the nucleoprotein organisation of the PIC (Suzuki *et al.*, 2004). However, LAP2 α was shown to be required for infection of cells by HIV with the wild-type envelope but not with a VSVg envelope pseudotype. In the same study, it was also demonstrated that emerin, an integral inner-nuclear-envelope protein and a binding partner of BAF, is required for efficient integration of the HIV provirus in cells but not MLV DNA. When cells are depleted of emerin or BAF, HIV episomal DNA circles are accumulated in the nucleus but not integrated in the host chromosome. Therefore, HIV and MLV use different inner-nuclear-envelope proteins to co-localise with chromatin (Jacque and Stevenson, 2006). Another protein localised in

the PIC is the high mobility group chromosome protein A1 (HMGA1), a non-histone DNA-binding protein that can modulate transcriptional regulation and chromatin structure (Farnet and Bushman, 1997). HMGA1 may not be required for integration, but it has been implicated in transcription of the viral genome (Beitzel and Bushman, 2003; Henderson *et al.*, 2000). A cellular protein (integrase interactor – INI 1) was found to bind tightly to IN and stimulates its DNA-joining activity (Kalpana *et al.*, 1994). Nonetheless, up until now, there is no evidence for a possible role of the INI 1 in integration in cells. Lens epithelium growth factor (LEGF/p75) was also shown to associate tightly with the HIV IN (Cherepanov *et al.*, 2003). A integrase mutant (Q168A) defective for LEGF/p75 binding failed to integrate the provirus but not nuclear import of viral DNA, demonstrating the importance of this cellular factor in the integration process (Emiliani *et al.*, 2005). The analysis of integration sites in LEGF/p75 depleted cells, showed that integration sites are less frequent in transcription units and in LEGF/p75 regulated genes, compared with integrants in cells that have normal amounts of this factor (Ciuffi *et al.*, 2005). These results revealed the first cellular protein that influences the integration site selection.

The viral integrase is also an important factor for site integration selection. This was well demonstrated by the Bushman group. A chimeric HIV virus containing a MLV integrase caused the hybrid to integrate the viral DNA with a specificity close to the MLV virus (Lewinski *et al.*, 2006). Also, an HIV integrase fused to a sequence-specific DNA binding protein was found to direct integration selectively to targets where the fused protein binds (Bushman, 1994; Bushman and Miller, 1997).

1.1.6.6 – Retroviral transcription and nuclear export

Once integrated in the host cell genome, the provirus behaves like a normal host gene, thus relying on the cell machinery for transcription and expression. Therefore, the retroviral genome needs to have all the *cis*-acting elements necessary for interaction with the host machinery. These are contained in the viral LTR.

Viral RNA molecules are transcribed by the host RNA polymerase II. The full length genomic RNA transcriptional start is located at the 5' end of R at the 5'LTR, defining the upstream R boundary, and a 3'-end processing event is located at the 3' end of

R in the 3'LTR, which defines the downstream R boundary. Therefore, the genomic RNA organisation is R-U5-*gag-pol-env*-U3-R.

The production levels of viral RNA are important. In simple retroviruses, such as MLV, transcription is regulated by the host and so, depends on the cell cycle and differentiation stage of cell. In cells actively producing virus, the viral RNA levels increase. As an example, cells infected with RSV may contain viral RNA reaching up to 10% of all of the cellular mRNA (Lee *et al.*, 1979). However, in complex virus, such as HIV, the transcription is also controlled by viral proteins.

The LTRs provides the *cis* elements where transcription factors bind to repress or augment transcription. In MLV, the U3 region of the LTR can be divided into an enhancer and a promoter. The promoter contains the TATA box, necessary for the binding of the RNA polymerase II and transcription initiation. The enhancer contains numerous binding sites for cellular transcription factors and many of them are closely packed or even overlapping. The nature of these elements contributes for the tissue specific regulation of transcription. The enhancer, apart from activating the LTR promoter, is also able to activate cellular promoters over large distances, accounting for one of the mechanisms of insertional mutagenesis (Bartholomew and Ihle, 1991; Lazo *et al.*, 1990).

Apart from cellular transcription factors, HIV also encodes a *trans*-activator protein (Tat) that confers tight control over transcription. This additional control may explain the ability of the virus to produce latent infection. During the latent stage of lentiviruses, only low levels of viral transcripts encoding Tat are found. In the presence of other cellular transcription activators, Tat is produced, resulting in high levels of viral expression. This may confer an advantage as continuous production of high levels of potentially toxic proteins could induce cell death before productive replication. Additionally, low viral protein levels may suppress an effective host immune response.

The HIV LTR can be divided into 4 functional domains: the *trans*-activation responsive region (TAR), the core promoter, the enhancer region and the regulatory region. Several host transcription factor can bind the LTR. The enhancer contains two NFkB binding sites (Nabel and Baltimore, 1987), important for efficient gene expression in activated T-cells. In unstimulated T-cells, NFkB p50 homodimers occupy the binding sites in the enhancer (Montano *et al.*, 1996). Histone deacetylase 1 is bound to p50 and

deacetylates the histones nucleosome 0 and 1, inhibiting the binding of RNA polymerase II (Williams *et al.*, 2006). After T-cell stimulation, NF κ Bp65/p50 translocates to the nucleus due to phosphorylation and subsequent degradation of I κ B that was restricting p65/p50 heterodimers to the cytoplasm (Whiteside and Israel, 1997). The heterodimers displace p50 in the viral enhancer. Additionally, they bind to acetyl transferase p300 that can acetylate the nucleosome 1, allowing the initiation of transcription by the RNA polymerase II (Gerritsen *et al.*, 1997). In the absence of Tat, although transcription can be initiated, elongation is inefficient (Kao *et al.*, 1987). However, NF κ B p65/p50 can recruit TFIIF and p-TEFb that phosphorylates RNA polymerase II stimulating elongation (Barboric *et al.*, 2001; Kim *et al.*, 2006). Upon Tat production, an increase of the levels of expression can be observed (Sodroski *et al.*, 1985). Tat binds to a stable stem loop located in the 5' of nascent viral transcripts (TAR). It recruits the host protein p-TEFb complex, including the subunit cyclin-dependent protein kinase 9 (Cdk9). Cdk9 hyperphosphorylates the RNA polymerase II transcription complex, stimulating transcription elongation (Herrmann and Rice, 1995).

The viral RNA molecules are modified with a 5' cap of m⁷G5'ppp5' (Keith and Fraenkel-Conrat, 1975) and by polyadenylation at the 3' end (Lai and Duesberg, 1972). The provirus does not control the transcriptional stop, and transcription can carry on into flanking host sequences. If the processing of the 3' viral end of transcripts containing host flanking sequences is not correct, those sequences can be translated and also packaged into the virion. This is another mechanism of insertional mutagenesis described for retroviruses, where a cellular oncogene is expressed at high levels due to insertion of the viral promoter in an adjacent site (Hayward *et al.*, 1981; Shen-Ong *et al.*, 1986).

Apart from the genomic RNA, shorter RNA molecules are also produced (Fan and Baltimore, 1973). These RNA molecules are the product of RNA splicing. In simple retroviruses, apart from the genomic RNA, a single spliced RNA transcript is also produced but in complex viruses, multiple spliced transcripts are found (Schwartz *et al.*, 1990).

Cellular mRNAs containing introns and unspliced sites are normally retained in the nucleus. This poses a problem for retroviruses, as for efficient genomic RNA package and translation, the full length RNA need to be exported from the nucleus, especially during

the replication stage of the virus. Simple retroviruses may contain *cis*-elements, termed constitutive export element (CTE) that allow the transport of such RNAs (Bray *et al.*, 1994). Complex viruses have evolved a different mechanism for efficient unspliced RNA export. HIV encodes the Rev protein that regulates this process. This viral protein binds to a complex structure of 5 stem loops (the Rev-responsive element, RRE) in the viral RNA (Mann *et al.*, 1994). Rev then recruits the cellular nuclear proteins Eukaryotic initiation factor 5A (Ruhl *et al.*, 1993) and importin 1 (Neville *et al.*, 1997). Importin 1 can bind the nuclear export factor Ran guanosine triphosphatase (Fornerod *et al.*, 1997). Once assembled it is thought that this complex is competent to be exported from the nucleus into the cytoplasm.

1.1.6.7 – Translation of viral proteins

The RNA molecules transcribed from the provirus resembles host mRNA. Therefore, translation of the viral proteins uses the host machinery. The ribosomes bind to the RNA and scan it towards the 3' end until finding the initiation codon AUG in the kozak consensus sequence (Kozak, 1987). In some retroviruses, this sequence lays in a stable secondary structure of the RNA and an internal ribosomal entry site has been described for MLV (Berlioz and Darlix, 1995).

All viral RNA transcripts are template for translation, including the genomic RNA. The translation of the full length genomic RNA produces a polyprotein containing the *gag* and *pol* genes. This gag-pol protein is produced by a bypass of the stop codon of *gag*. The mechanism used by MLV is read-through, where the ribosome recognises the stop codon as a sense codon (Yoshinaka *et al.*, 1985). Another mechanism, used for the majority of retrovirus, including HIV, is ribosomal frameshift, where the ribosome goes back one nucleotide in this region and can therefore proceed with protein synthesis (Jacks and Varmus, 1985). Using this strategy, only 5% of the total viral protein is gag-pol (Jamjoom *et al.*, 1977).

After protein synthesis, gag and gag-pol proteins of most retroviruses are modified in the amino termini by the addition of a myristyl group to glycine (Henderson *et al.*,

1983). This post-translational modification has been shown to be required for the binding of gag to the cellular membrane (Rein *et al.*, 1986).

The envelope proteins are produced from a spliced RNA transcript. The nascent polyprotein is glycosylated (Rifkin and Compans, 1971). This process occurs in the endoplasmic reticulum, as with typical host glycoproteins. Afterwards, the env protein is assembled in oligomeric structure and then transported to the golgi compartment, where it is cleaved by the cellular protease furin to form the TM and SU proteins (Bradac and Hunter, 1986; Wills *et al.*, 1984; Stein and Engleman, 1990; Hallenberger *et al.*, 1992; Einfeld and Hunter, 1988).

1.1.6.8 – Assembly, budding and maturation

After synthesis, gag, gag-pol and env proteins need to come together at the cell membrane to assemble and incorporate the genomic RNA to produce infectious virions.

Typically, around 2000 molecules of gag are needed for particle assembly (Stromberg *et al.*, 1974) and gag alone is able to form virus-like particles, not containing any RNA, env or pol proteins (Gheysen *et al.*, 1989). In MLV and HIV, this process occurs at the plasma membrane. The M domain, located in MA, is necessary to direct the gag proteins to the membrane (Yuan *et al.*, 1993). Gag proteins must then join with each other to form the assembly site. This interaction is made through the I domain located in NC (Gheysen *et al.*, 1989). However, for both MLV and HIV, the carboxy-terminal of CA may be required for additional interactions (Schwartzberg *et al.*, 1984; Jowett *et al.*, 1992).

Gag-pol polyprotein is also directed to the assembly site through the signals in gag. Interestingly, it has been shown that gag-pol alone is not capable of forming virions (Felsenstein and Goff, 1988).

The RNA contains a sequence that forms a secondary structure necessary for the interaction with gag and incorporation to the virus particle. This sequence is called the packaging signal (Ψ) and it interacts with NC. In HIV, it is localised between the major donor and acceptor splicing sites. Therefore, only genomic unspliced RNA is packaged (Darlix *et al.*, 1990).

Env proteins are not required for gammaretrovirus or lentivirus budding, but virions without env are not efficiently infectious. In HIV, it is thought that env proteins interact with MA in order to actively localise to the assembly site (Egan *et al.*, 1996; Yu *et al.*, 1992).

Vpr is packed in the virion through interaction with the p6 region in gag (Lu *et al.*, 1993). On the other hand, Nef and Vif are thought to be incorporated non-specifically in the virion, although Nef requires myristoylation for efficient incorporation (Bukovsky *et al.*, 1997; Camaur and Trono, 1996).

After all the components have been localised and assembled, buds emerge and the virion is then released from the cell. The late domains in gag mediate the process. In HIV, the late domain, a PTAP motif in p6 (Gottlinger *et al.*, 1991), interacts with Tsg101, a sub unit of the endosomal sorting complex required for transport I (ESCRT I). It is thought that this protein interacts then with ESCRT III, a key host protein involved in this process (Garrus *et al.*, 2001). Another secondary late domain (YPXnL) is located downstream the PTAP sequence has been shown to be also involved in budding. ALIX binds to YPXnL and then interacts with ESCRT III (Fisher *et al.*, 2007), promoting budding. Additionally, Vpu is necessary for virion release in some human cells, including CD4+ T Cells (Terwilliger *et al.*, 1989). These cells encode tetherin, a host protein that causes retention of fully formed virions in the cell surface. Vpu colocalises with tetherin, inhibiting this effect (Neil *et al.*, 2008).

After release from the cell, the virions are immature and not infectious. The polyproteins are cleaved by protease (PR) to produce the enzymes IN and RT, as well as the MA, CA and NC structural proteins. This cleavage is an ordered and sequential process (Ledbetter, 1979). Subsequently, the structural proteins rearrange via a process called maturation to form the infectious virus particle (Vogt, 1996).

1.2 – Retroviral vectors

Retroviruses insert their genome into the host DNA. This feature is highly desirable for gene therapy as the viral DNA introduced in a progenitor cell will be passed on to all the progeny enabling stable gene transfer to a tissue.

The two main types of viral vectors were derived from gammaretrovirus and lentivirus. The later have the advantage that they infect non-dividing cells.

The basic principle behind vector development is to replace the coding region (*gag-pol-env*) with the transgene of interest. The *cis* elements required for the RNA encapsidation, reverse transcription and transcription need to be retained in the transfer vector. These include the packaging signal (Ψ), the viral LTR, primer binding site (PBS) and the polypurine tract (PPT). Additional proteins necessary for vector production are supplemented in *trans*, in packaging constructs.

1.2.1 – Gammaretroviral vectors

Gammaretroviral vectors are most commonly based on the Moloney murine leukaemia virus (MLV). In the simplest transfer vectors, the transgene is placed between the packaging signal and the PPT, transcription is regulated by the enhancer and promoter in the 5'LTR and the polyadenylation signal is provided in the 3'LTR.

The splicing signals are elements that overlap with viral genes. The splice donor is located upstream of the packaging signal and the splice acceptor is downstream, in the *pol* gene. In the N2 vector, the transgene is placed in the *gag* position. Therefore, this vector does not contain the splice acceptor sequence (Gilboa *et al.*, 1982; Eglitis *et al.*, 1985). Other transfer vectors were made where *env* is replaced by the transgene, thus containing the splice donor and acceptor sequences, such as the MFG vector (Danos and Mulligan, 1988; Dranoff *et al.*, 1993). Increased levels of spliced RNA in target cells explains the higher levels of transgene expression from the MFG vector compared with the N2 vector (Krall *et al.*, 1996). The MFG vector was therefore chosen as the transfer vector in two

gene therapy trials for severe combined immunodeficiency (SCID)-X1 disease (Cavazzana-Calvo *et al.*, 2000; Gaspar *et al.*, 2004).

The viral proteins encoded in the viral *gag*, *pol* and *env* genes are provided in *trans*, in retroviral packaging constructs. The first generation of these constructs (pMOV-psi-) was accomplished using a retroviral genome missing the packaging signal (Mann *et al.*, 1983). However, there was a high potential for generating replication competent virus due a recombination event between the vector and the packaging construct. In a second generation of retroviral packaging constructs, pPAM3, apart from the Ψ signal, all of the 3'LTR, 352bp on the 5' of U3 of the 5'LTR and the PPT were also removed (Miller and Buttimore, 1986). A further improvement of packaging constructs was made by Danos and Mulligan, where the helper functions were split into different constructs. Here, *env* was separated from the *gag* and *pol* genes to create the pCRIPenv- (containing the *gag* and *pol* genes) and the envelope constructs pCRIPAMgag- (expressing the amphotropic envelope) and pCRIPgag-2 (expressing the ecotropic envelope). With these constructs, further recombination events were needed to produce replication competent virus, thus increasing the safety of vector production. These constructs also allowed the development of pseudotyping, the generation of virion with heterologous *env* proteins (Danos and Mulligan, 1988).

1.2.2 – Lentiviral vectors

The development of lentiviral based transfer vectors followed the gammaretrovirus system. In the vector pHR, in addition to the viral LTRs, Ψ signal, PBS and PPT, other viral elements were left. These sequences are part of the *gag* gene which had been shown to augment packaging efficiency, and the Rev response element (RRE), that is encompassed in part of the *env* gene. In the transfer vector, the Ψ signal and RRE are flanked by splice acceptor and donor sequences. The *gag* gene reading frame was blocked by a frameshift mutation. The *trans*-activation responsive region (TAR) augments the transcription efficiency and RRE allows efficient nuclear export of unspliced RNA in the presence of Tat and Rev, respectively. These two viral proteins would only be present in the producer cells, supplemented in *trans*. The lentivirus LTR requires Tat for efficient

expression, so, in this vector design, transcription of the transgene in the target cells had to be controlled by an internal promoter (Naldini *et al.*, 1996a; Naldini *et al.*, 1996b).

Additional modifications to this transfer vector included the incorporation of the central polypurine tract (cPPT) to enable the formation of a DNA flap during reverse transcription and the post-translational regulatory element of the Woodchuck hepatitis B virus (WPRE). This DNA flap has been shown to increase nuclear import of the viral DNA, thus increasing the transduction efficiency in both dividing and non-dividing cells (Follenzi *et al.*, 2000; Zennou *et al.*, 2000). The WPRE improves transgene expression levels, by facilitating the nuclear export of transcripts containing this element (Zufferey *et al.*, 1999).

The packaging constructs necessary to supplement the viral sequences necessary for the production of vectors were designed based on the split retroviral vectors. In the first generation pCMV Δ R9, expressing the *gag* and *pol* genes, the viral LTRs were substituted with the human cytomegalovirus (CMV) immediate early promoter and the polyadenylation signal from the insulin gene at the 5' and 3', respectively. The Ψ signal and adjacent sequences were also removed, but the splice donor site was preserved. Additionally, the envelope proteins and the accessory protein Vpu were also removed. The envelope glycoprotein was supplied in another packaging construct.

A second generation packaging construct was made by deleting the accessory genes, as these proteins were shown to be dispensable for efficient vector production, transduction and integration (Zufferey *et al.*, 1997). These constructs would confer a higher safety as replication competent vectors that may arise during production would lack essential viral proteins for HIV-virulence *in vivo*.

In the third generation constructs, the LTR of the transfer vector was replaced by constitutively active promoter sequences (either from RSV or MLV) controlling the RNA production in producer cells, thus allowing the removal of the *tat* gene from the packaging vector. Furthermore, the *rev* gene was shown to be necessary for efficient nuclear export of the transfer vector, so it was separated from the *gag* and *pol* genes and supplemented in a separate vector (Dull *et al.*, 1998).

1.2.3 – Self-inactivating (SIN) vectors

The U3 region containing most of the enhancer region can be deleted from the 3'LTR in the transfer vector, thus creating a self-inactivating vector. Upon reverse transcription, the U3 will be deleted in both LTRs, so expression of the transgene will have to be controlled from an internal promoter.

SIN design is believed to increase the biosafety of these vectors by decreasing the probability of generating replication competent vectors in both produced and target cells, reducing therefore the risk of mobilisation by wild-type virus in an infected host (Zufferey *et al.*, 1998). Most importantly, the deletion of an enhancer reduces the risk of insertional mutagenesis. Additionally, it may increase transgene expression by reducing transcription interference between the viral LTR and the internal promoter.

The first SIN retroviral vectors were made in an MLV-based vector (Yu *et al.*, 1986; Yee *et al.*, 1987). The first vector had a deletion of 299bp in the U3 region, containing the enhancer and the CAAT box, the second had the additional deletion of the TATA box. However, retroviral vectors display weak polyadenylation sites and deletion of U3 to generate SIN vectors increased the likelihood of read-through thus augmenting the potential for insertional mutagenesis (Furger *et al.*, 2001; Zaiss *et al.*, 2002). Schambach and colleagues have described the insertion of upstream polyadenylation elements to improve 3' processing to increase the biosafety of these vectors for gene therapy (Schambach *et al.*, 2007).

The first lentiviral SIN vector was described by Zufferey and colleagues (Zufferey *et al.*, 1998). In HIV, the upstream polyadenylation element is located between the TATA box and the transcriptional initiation site and deletion of a region further upstream had minimal effect on 3' processing (Valsamakis *et al.*, 1991). This characteristic allowed the deletion of an extensive part of the U3 region without affecting considerably 3' processing. The lentiviral SIN vector was generated by deletion of 400bp in U3, including the TATA box and leaving only 38bp from the 5' end (containing the attachment sites for integrase recognition of the viral DNA) and 18bp from the 3' end of U3 (Zufferey *et al.*, 1998). However, even this large deletion did not completely inhibit transcription from the viral

SIN LTR, due to regions containing binding sites for transcriptional activators (SBF1 and SP1) located downstream the transcriptional initiation site (Logan *et al.*, 2004).

1.2.4 – Expression cassette

In the simplest approach for gammaretroviral vectors, the expression of a cDNA, encoding the protein of interest, is controlled by the viral LTR. However, in lentivectors and SIN vectors, transcription is driven by an internal promoter. This strategy allows a more flexible choice of a promoter.

The promoter can be constitutively active and derived from virus (such as the LTR from the spleen focus-forming virus (SFFV), or the immediate early promoter from the cytomegalovirus) or cellular genes (such as the phosphoglycerate kinase or the elongation factor 1 α). These promoters are active in a range of different cell types. However, transcription can be restricted to a given cell type by the use of tissue-specific promoters, such as the α 1-antitrypsin promoter in liver (Ciliberto *et al.*, 1985) or the myogenin promoter in skeletal muscle (Yee and Rigby, 1993). Alternatively, the incorporation of microRNA target sequences in the transcript can also prevent expression in cells where those microRNAs are expressed (Brown *et al.*, 2006). The restriction of expression to the target cells may avoid an immune response by preventing expression in antigen presenting cells. In certain applications, temporal control of transgene expression may be required. This can be achieved by the incorporation of an inducible promoter, such as a tetracycline-responsive promoter (Gossen and Bujard, 1992).

Several strategies can be taken when multiple transgenes are required to be expressed from one vector. Proteins can be expressed using alternative splicing, although this strategy presents several drawbacks, such as the ratio of spliced transcripts or the generation of alternative splicing with the incorporation of cellular sequences (Cepko *et al.*, 1984; Korman *et al.*, 1987). Multiple promoters can be included, each driving expression of a different transgene. However, this strategy has the potential problem of promoter suppression, where one of the transcripts is expressed more than the other (Emerman and Temin, 1984). Another strategy relies on the introduction of an internal ribosome entry site (IRES) between transgenes that are being translated from the same

transcript (Adam *et al.*, 1991). Two transgene may also be expressed as a fusion protein combining both functions in a chimeric protein or separated during translation by a ribosomal skip mechanism or after translation by a protease cleavage site (Germann *et al.*, 1990; Morrison *et al.*, 1991; Szymczak *et al.*, 2004).

1.2.5 – Pseudotyping

A process that commonly occurs during viral assembly of viruses in cells infected with two or more virus is the formation of phenotypically mixed particles. These particles have envelope proteins from both viruses and the tropism of the parental virus is therefore altered (reviewed in Zavada, 1982).

In the case of HIV, several studies have demonstrated that the wild-type virus produced in cells infected with xenotropic MLV (Lusso *et al.*, 1990), amphotropic MLV (Spector *et al.*, 1990) or herpes simplex virus gives (Zhu *et al.*, 1990) gives rise to viruses with an expanded host range, suggesting that heterologous glycoproteins were incorporated in the virus envelope.

A new strategy to alter the tropism of vectors has then emerged with the design of packaging constructs where *env* was split from the *gag* and *pol* genes thus allowing the expression of different envelope glycoproteins in the producer cell. Using this strategy, the tropism of gammaretrovirus vectors based on ecotropic MLV has been broadened to include human cells by using the MLV amphotropic envelope proteins (Danos and Mulligan, 1988).

The most commonly used glycoprotein to pseudotype lentiviral vectors in order to expand vector tropism is the glycoprotein from the vesicular stomatitis virus (VSVg). The receptor for VSVg, although still undetermined, appears to be ubiquitous in all cell types, explaining the broad host range of VSV (Schlegel *et al.*, 1982; Coil and Miller, 2004). Furthermore, VSVg pseudotyped vectors are highly stable and can be efficiently concentrated by ultracentrifugation, enabling the production of serum-free, high-titre vector particles (Burns *et al.*, 1993; Naldini *et al.*, 1996b). However, VSVg is associated with cytotoxicity, hindering the creation of packaging cell lines for production of VSVg pseudotyped vectors. Nevertheless, progress has been reported using inducible expression

of this envelope protein. Another drawback of this pseudotype is the inactivation by human serum complement, preventing its use *in vivo* (DePolo *et al.*, 2000). Furthermore, it has been reported that lentiviral vector preparations pseudotyped with VSVg are contaminated with tubulovesicular structures carrying nucleic acids or proteins that may elicit an undesirable immune response (Pichlmair *et al.*, 2007).

Heterologous glycoproteins have also been used to restrict the tropism of the vectors. The neurotropic properties of Lyssavirus, including the rabies and Mokola virus, have early been studied. Vectors pseudotyped with those glycoproteins infected neurons preferentially (Mochizuki *et al.*, 1998; Desmaris *et al.*, 2001). Retrograde transport has been achieved by the use of Rabies glycoprotein (Mazarakis *et al.*, 2001).

Another strategy to target specific cell types is the engineering of envelope glycoproteins. Viral envelope proteins have been modified to contain ligands or single chain antibodies (Hatzioannou *et al.*, 1999). Using the second approach, by fusing single chain antibodies to the sindbis glycoprotein, lentiviral vectors have efficiently been targeted to melanoma cells (Morizono *et al.*, 2005).

1.2.6 – Vector production

Two strategies have been adopted to produce gammaretroviral and lentiviral vectors. The production can be accomplished by transient transfection of plasmids coding for the transfer vector, packaging and envelope into highly transfectable cells, usually 293T (Naldini *et al.*, 1996b; Soneoka *et al.*, 1995). Stable packaging cell lines have also been developed to produce these vectors. For clinical applications, the use of packaging cell lines is more appropriate. It allows the production of large and reproducible batches enabling control of the quality of the vector.

Gammaretrovirus and lentivirus present some differences for the development of stable cell lines, especially in the complexity of the viral genome. Furthermore, contrary to gammaretroviral proteins, the lentiviral protease and Vpr proteins have been reported to be cytotoxic (Kaplan and Swanstrom, 1991; Planelles *et al.*, 1995; Rogel *et al.*, 1995). Therefore, the development of packaging cells for the production of lentiviral vectors has

relied on the use of inducible expression of the viral proteins (Dull *et al.*, 1998; Kafri *et al.*, 1999; Yu *et al.*, 1996).

The development of a stable cell line producing lentiviral vectors has been recently reported. The STAR cells are derived from 293T or HT1080 human cell lines that were transduced an MLV-based vector expressing a codon optimised version of *gag* and *pol* (lacking RRE) and then co-transduced with Tat and Rev being expressed from separate gammaretroviral vectors. The STAR cells were then stably transfected with plasmids expressing either MLV-A, GALV+, or RD114pro glycoproteins to allow for high plasticity regarding vector pseudotyping. The production of SIN and non-SIN versions of lentivectors in these cells is similar, with a high titre and stable for up to 3 months (Ikeda *et al.*, 2003).

A variety of gammaretroviral cell lines have been developed and some have already been used in clinical trials, including the Ψ CRIP cells producing gammaretroviral vectors pseudotyped with the MLV amphotropic envelope (Danos and Mulligan, 1988) and the PG13 cells that produces gibbon ape leukaemia virus pseudotyped vectors (Miller *et al.*, 1991). However, these cell lines were developed from mouse cell lines and vectors produced by these cells are rapidly inactivated by human serum (Takeuchi *et al.*, 1994; Takeuchi *et al.*, 1996) and may cross-pack contaminating endogenous viral sequences (Patience *et al.*, 1998). Therefore these vectors would not be suitable for *in vivo* human gene therapy. Packaging cell lines derived from human cells have been developed to solve these problems (Cosset *et al.*, 1995).

1.3 – Adenoviruses

1.3.1 – Structure and Biology of Adenoviruses

Adenoviruses (Ad) have been associated with mild diseases, such as colds or acute respiratory diseases. The virions are non-enveloped, icosahedral structures with long fibres extending from each of the 12 vertices, with a double-stranded molecule of DNA as genome. This complex structure is composed of hexons (forming each of the 20 triangular facets of the capsid), penton bases and fibres, and other “minor” components. There are six other structural proteins situated in the virion core (Green *et al.*, 1967; Horne *et al.*, 1959; Valentine and Pereira, 2003).

The fibre is composed of a rod and a knob attached to the penton base. The length and flexibility of the rod varies among the virus serotypes and the knob function as the cellular attachment site. The major receptor for most adenoviruses is the Coxsackie adenovirus receptor (Coyne and Bergelson, 2005; Philipson and Pettersson, 2004). After this initial interaction, the RGD peptide on the penton base binds to cellular $\alpha_v\beta_3/\alpha_v\beta_5$ integrins (Mathias *et al.*, 1994), facilitating virus internalisation via clatherin coated vesicles and into endosomes for further processing (Patterson and Russell, 1983). After uncoating within the early endosome, the nucleoprotein particle is released to the cytoplasm and migrates to the nucleus. The viral terminal protein (TP) remains attached to the DNA and drives it to specific sites in the nucleus where viral transcription is initiated (Schaack *et al.*, 1990).

DNA replication of these viruses is a complex process. In Ad2 virus, the 36kb DNA contains a long inverted terminal repeat (ITR), where the origin of replication is situated. Gene expression can be divided in early and late transcription units, based on the kinetics of this process. Six early transcripts (E1A, E1B, E2A, E2B, E3 and E4) encode proteins that orchestrate the virus expression. After infection, the first units to be transcribed are the E1A and E1B, whose products regulate cellular genes and transactivate other transcription units. The products from the E3 unit modulate the host immune response (Fessler *et al.*, 2004; Flomenberg *et al.*, 1987). E2A generates the DNA-binding protein (DBP) and E2B produces the TP and a DNA polymerase. These three proteins,

together with the cellular nuclear factors I, II and III are involved in the replication of the virus DNA. After a TP-primed DNA replication initiation, the viral polymerase elongates the nascent strand. After completion, the process is repeated, generating high concentrations of viral DNA. The major late promoter is then activated, producing the late transcripts that encode the structural proteins. Virion assembly occurs in the nucleus. This process is thought to arise from a series of concerted maturation steps involving the viral protease and the assembly of scaffolding proteins and precursor structural polypeptides. Early experiments suggested that the viral genome is inserted in these preformed capsids (Edvardsson *et al.*, 1976; D'Halluin *et al.*, 1978a; D'Halluin *et al.*, 1978b) but recent findings indicate that the genome initiates the assembly of the capsid (Zhang and Imperiale, 2000). Regardless of the mechanism, it is well accepted that selective encapsidation of DNA molecules occur by the recognition of *cis*-acting sequences, located close to the left ITR (Grable and Hearing, 1990). Once assembly and DNA encapsidation have been completed, the viral protease cleaves the structural proteins, forming fully mature particles (reviewed in Mangel *et al.*, 2003). Finally, the adenovirus death protein is involved in the lyses of the cells and release of the virus progeny (Tollefson *et al.*, 1996).

1.3.2 – Adenoviral vectors

Due to the better understanding of the genome of Ad 2 and 5, and because these viruses never induced tumours in mice models, they were used to develop the first generation vectors. The principle was to delete essential sequences of the viral genome in order to impair replication and to facilitate the accommodation of the transgenes of interest. The E1 genes are required to initiate viral replication and thus, are essential for viral replication. Furthermore, it is known that E1 coded proteins are involved in cellular transformation. Therefore, the first strategy to produce adenoviral vectors was to delete the E1 transcription units. Vector production would require supplementing the E1 genes in *trans*. A cell line was then produced for this purpose. The 293 cell line (Graham *et al.*, 1977) was transformed with the E1 coding region. Then, a plasmid containing the virus ITR, packaging signal and the E1 genes substituted by foreign DNA could be transfected in such cells. This technique allowed the production of high concentration of vectors.

This first generation vector permitted the incorporation of 5kb foreign cDNA. Also, some vectors were generated with a further deletion on the E3 genes, as these are not required for viral replication, increasing the capacity to insert expression cassettes of 8kb. Although these vectors showed great potential for gene delivery, problems arose from the fact that replicating virus were also produced. The E1 coding region in the complementing cell line could recombine with the vector, giving rise to viral DNA with functional E1 generating replication competent viruses. Other complementing cell lines, such as PER.C6 (Fallaux *et al.*, 1998), were constructed. In these, the E1 coding region was trimmed to minimise the overlap with the vector DNA, reducing homologous recombination. However, experiments in animal models demonstrated that, although the vector could efficiently infect cells, expression was transient. This happened due to the activation of the host immune responses.

A new strategy was then generated to minimise the immune response triggered by the vectors. In these so-called second generation vectors, further regions of the viral genome, such as E2 and E4, were deleted. Cell lines with the complementary deleted sequences were constructed (Schaack *et al.*, 1995; Amalfitano *et al.*, 1996; Amalfitano and Chamberlain, 1997), but the vector production in these cells yielded lower titres, compared with first generation vectors. The E2 deleted vectors showed longer transgene expression in immune competent mice (Hu *et al.*, 1999), but results from the E4 deleted vectors were less conclusive (Grave *et al.*, 2000; Armentano *et al.*, 1997). The best strategy then seemed to be the complete excision of all transcription units of the adenovirus DNA. These third generation vectors, also called gutless adenoviral vectors, only have the ITRs and the packaging signal essential for encapsidation of the vector DNA. Therefore, the packaging capacity of the vector increased to 36kb.

In order to produce gutless vectors, several efforts have been made to produce a cell line to provide in *trans* all the necessary viral genes. However, due to the complexity of the Ad genome and its regulation, such task has not yet been reported. The most common strategy relies on the infection of the producer cell with a helper Ad. However, separation of the produced vectors from the helper virus had to be addressed. The strategies reported so far, rely on the reduction of the packaging efficiency of the helper genome compared to the gutless vector. This has been accomplished by mutating the

helper virus packaging signal, a different genome size (smaller or bigger genomes than the optimal do not pack efficiently) or by specific removal of the packaging signal during vector production using a recombinase such as Cre, Flp or the unidirectional recombinase Φ C31. Nevertheless, contaminating helper virus still remains in the final preparation, although at low levels (approximately 0.2%).

The immunogenicity and transient expression of adenoviral vectors may be a hurdle for the application to the treatment of inherited genetic disorders. However, these same characteristics provide an excellent tool for vaccination and cancer treatment. In fact, the first gene therapy product is an E1-deleted adenoviral vector (GencidineTM, SiBiono GeneTech), that is now commercially available in China (Garber, 2006). The E1B-55k Ad protein inactivates the p53 protein (Dobner *et al.*, 1996) and thus, vectors lacking this gene can only replicate in cells that lack p53 function (Bischoff *et al.*, 1996). Since a high percentage of tumour cells have mutations in the p53 gene, the oncolytic vectors selectively replicate in these cells, lysing tumour cells but not normal cells (Chang *et al.*, 1995; Heise *et al.*, 1997).

1.4 – Adeno-Associated Viruses (AAV)

1.4.1 – Structure and Biology of AAV

The AAV virus was first discovered as a contaminant of Adenoviruses preparations (Atchison *et al.*, 1965). Although most humans have been infected, this virus has not been associated with any disease.

This is one of the smallest viruses known, with an icosahedral capsid of 20nm in diameter and a small linear single-stranded DNA genome of 4.7kb. The genome is very compact, with only two coding regions flanked by ITRs (reviewed in Berns, 1990). The rep gene encodes four regulatory proteins. These proteins are produced through the use of two promoters (p5 and p19). Splicing of the transcript originated from the p5 promoter gives rise to the Rep78 and Rep68 proteins that are important in all stages of the virus life cycle. The p19 promoter transcript can be spliced to produce the smaller proteins Rep52 and Rep40 that are involved in the accumulation of the genome during replication. The cap gene is transcribed from one promoter (p40), but alternative splicing gives rise to two transcripts. The larger produces the VP1 capsid protein and the smaller encodes the VP2 and VP3 structural proteins through the use of a noncanonical start codon. These three proteins assemble to form the capsid (Hermonat *et al.*, 1984; Tratschin *et al.*, 1984; Chejanovsky and Carter, 1989).

AAV2 particles use the ubiquitous heparan sulphate proteoglycans as docking sites, explaining the broad tropism of these viruses (Summerford and Samulski, 1998). Internalisation is aided by the co-receptors $\alpha_v\beta_5$ integrin (Summerford *et al.*, 1999), fibroblast co-receptor type 1 (Qing *et al.*, 1999) and the hepatocyte growth factor receptor c-met (Kashiwakura *et al.*, 2005).

There are two stages in the AAV life cycle, depending if the cell is superinfected with a helper virus, such as Adenoviruses or Herpes Simplex viruses. When AAV infects human cells alone, expression is auto-repressed and latency is established by the integration of the virus genome in the chromosome 19, in the AAVS1 site (Kotin *et al.*, 1990). This site-specific integration is unique among viruses and therefore, very attractive in the gene therapy context. The process is not completely understood, but the virus

elements necessary for site-specific integration have been elucidated. These include one of the Rep 68 or Rep78 proteins in *trans* and the ITRs or the integration efficiency element in the p5 promoter in *cis* (Surosky *et al.*, 1997; Philpott *et al.*, 2002). In the host genome, there are multiple Rep binding elements but only the AAVS1 site has the required Rep specific nicking site (known by terminal resolution site) at the proper distance (Weitzman *et al.*, 1994).

When the cell is infected with a helper virus, the AAV starts a productive cycle. Rep expression is repressed by the binding of the host factor YY1 to a sequence in the p5 promoter. When the cell is infected with Adenovirus, the E1A protein binds to YY1 relieving the Rep expression inhibition (Shi *et al.*, 1991). Rep78 and Rep68 are then produced and act as a transactivator of the p19 and p40 promoters to produce the smaller Rep proteins and cap proteins. Also, they excise the integrated provirus and start replication. The ITR of the genome can fold on itself due to base complementarity and form a characteristic T-shaped hairpin structure. This secondary structure provides the initiation primer for DNA replication (Hauswirth and Berns, 1977). Upon completion of DNA polymerisation of the leading strand, the double-stranded molecule, closed at one end, is resolved by site-specific nicking at the terminal resolution site on the ITR of the parental strand (Snyder *et al.*, 1990). Then, the ITR is copied using the free 3'hydroxyl group in the parental strand. Finally, the palindromic linear duplex termini can renature into terminal hairpins, the strands separate and the process can be re-initiated. Concatemers can also be formed when the nicking does not occur, but resolution of these structures may still occur afterwards (Straus *et al.*, 1976).

The other adenovirus proteins necessary for AAV replication are the E1B 55K and E4orf6 proteins, that have been shown to facilitate the export of the mRNA transcripts to the cytoplasm (Samulski and Shenk, 1988; Huang and Hearing, 1989) and also the E2A protein and the VA RNAs which are required for translation of the capsid mRNA (Janik *et al.*, 1989). Finally, the genome packaged into assembled capsids and released from the cells by lyses induced by the helper virus.

1.4.2 – AAV vectors

When Samulski and colleagues cloned the whole AAV genome into a transfer plasmid, transfected it in cells and infected them with Adenoviruses, they could rescue infective AAV virions (Samulski *et al.*, 1982). This observation provided the basis for AAV vector development. The only *cis* sequences necessary for encapsidation of the AAV genome are the ITRs. Therefore, the AAV vectors can be striped out of the cap and rep genes. Nevertheless, these genes need to be supplemented in *trans* for replication of the vector genome and production of the viral capsid. Also, AAV is a virus whose replication is dependent on the infection of a helper virus. Therefore, in order to produce a vector based on AAV, it was necessary to supplement in *trans* the necessary virus helper proteins to accomplish effective production.

Initially, producer cells were transfected with the AAV vector plasmid, containing the expression cassette of interest flanked by the ITRs, and a plasmid coding for the cap and rep genes (Samulski *et al.*, 1989). These cells would then be infected with helper adenovirus to produce AAV vectors. Nevertheless, it was necessary to subsequently remove the Adenovirus from the preparation. The elimination of Adenovirus relied on the difference in physical properties of the virions, such as thermostability and density. Typically, the Adenovirus could be eliminated by heat inactivation or the AAV vectors could be rescued after separation in a caesium chloride density ultracentrifugation. Later, the discovery that AAV had a receptor for heparan sulphate proteoglycan permitted the development of a purification strategy based on affinity chromatography (Zolotukhin *et al.*, 1999).

The finding of which helper proteins were involved in the AAV replication permitted the construction of a plasmid expressing all of the necessary proteins and enabled the production of AAV vectors without the need for Adenovirus infection (Xiao *et al.*, 1998).

The efficiency of these vectors is hindered by the slow process of second strand synthesis necessary for transgene expression. AAV vectors can package plus and minus single-stranded DNA genomes. Therefore, when multiple vectors infect one cells, the plus and minus DNA strands can anneal by base pairing. Although the probability of

occurrence of this mechanism increases with vector dose, host recombination factor may play an important role in promoting strand annealing so, it is difficult to predict the efficiency in certain cell types. However, this step can be optimised by the use of self complementary vectors (reviewed in McCarty, 2008). These vectors carry an inverted repeat genome that can fold into a double-stranded DNA molecule, without the need of DNA synthesis or strand annealing, thus increasing the efficiency of the vector expression. However, in order to package such genomes, the expression cassette can only contain a maximum of 2.2Kb, which limits the application of these vector to the expression of small proteins, but there are still several applications for these vectors, such as haemophilia B (Nathwani *et al.*, 2006).

The common AAV vector has been derived from the AAV-2 serotype. Gene transfer from these vectors can be hampered by neutralising anti-AAV-2 antibodies, which are highly prevalent in human population. However, this drawback can be circumvented by the use of different serotypes (reviewed in Wu *et al.*, 2006). The development of new AAV vectors based on different serotypes can help to evade the humoral immune response towards AAV-2 and also enhance the infection in target cells. The AAV-2 vector genome can be cross-packaged into the capsids of other serotypes and, it has been shown that these “pseudotyped” AAV vectors display different tissue tropism. Therefore, this technique can be exploited in order to broaden the targets for AAV vectors or to develop tissue-targeted vectors while escaping the anti-AAV-2 immune response.

1.5 – Vectors in gene therapy

In order to ameliorate a genetic disease, the defective gene can be repaired or substituted or a corrected gene can be introduced in the target cell. This can be accomplished *ex-vivo* or directly in patients and there are different methods to deliver nucleic acids to target cells.

Physical delivery of naked DNA has been achieved by microinjection, electroporation, ultrasound, hydrodynamic pressure and particle bombardment. Although the concept of direct delivery of the DNA without disruption of the cells or incorporation of any other foreign molecules would be ideal, this technology has been hindered by the low transfer of DNA and the fact that these techniques are invasive *in vivo* (Reviewed in Wells, 2004).

Chemical approaches have been developed to enhance DNA transfer by mimicking viral properties. Normally, these non-viral vectors consist in a carrier molecule (lipid, peptide or polymer-based) complexed with DNA. Although success has been achieved in animal models, the transfer of this technology into a clinical setting remains elusive due to factors controlling stability and pharmacokinetics of these vectors and because the results obtained may not be similar in humans (reviewed in Niidome and Huang, 2002).

Viruses have evolved to efficiently transfer their genome to the host cell and many have then been exploited to develop vectors for gene therapy. Viral based vectors are currently the most effective means of gene delivery. In fact, they represent ~70% of the vectors used in clinical trials. The vectors used to achieve gene delivery have been a wide range, with adenovirus and retrovirus comprising the majority used in clinical trials (Journal of Gene Medicine, clinical trials database).

Each viral vector system is characterised by a set of properties, some of which inherited from the parental virus. Hence, given the diversity of disease targets, it is clear that no single vector is suitable for all applications. The choice of vector depends largely on the cell type to be targeted and duration of expression required, but other properties such as low cytotoxicity and low immune response are highly desirable to ensure safety in a clinical setting. Table 1.2 summarises the properties of vectors.

Table 1.2: Vectors used in gene therapy

Vector	Packaging capacity	Host range	Features
AAV	~4 Kb	Dividing and non-dividing cells	Slow expression onset, long term expression
Adenovirus	Up to 30 Kb	Broad (low in neurons)	Transient expression, strong immunogenicity
Alphaviruses	<7.5 Kb	Broad (neuron and glial cell-specific strains)	Transient high expression, low immunogenicity
Herpes simplex virus	40 Kb 150Kb in amplicons	Broad, (strong for neurons)	Latent infection, long-term expression, low toxicity
Lentivirus	8-10 Kb	Dividing and non-dividing cells	Genome integration, long-term expression, inefficient production
Retrovirus	8 Kb	Restricted to dividing cells	Genome integration, long-term expression

1.5.1 – Retroviral vectors in gene therapy: Success with consequences

Retroviral vectors can integrate their genome into the host genome. This particular characteristic makes them useful for target diseases that require prolonged transgene expression in dividing cells. Indeed, the first successful clinical trials used these vectors for the treatment of immunodeficiencies.

X-linked severe combined immunodeficiency (SCID-X1) results from mutations in the gene encoding the common cytokine receptor gamma-chain (γ_c). The failure of γ_c signalling results in a classical phenotype characterised by the absence of T and natural killer (NK) cells, whilst B cells are present but poorly-functional (White *et al.*, 2000). Unless patients are treated with haematopoietic stem cells (HSC), they succumb to infections during the first years of life. The aim of two clinical trials using gene therapy for the treatment of this disease was the immune reconstitution with autologous HSCs manipulated to express γ_c . So, HSCs were isolated from patient and transduced *ex vivo* with a gammaretroviral vector expressing γ_c . The cells were then infused back into patients and reconstitution of the immune system was monitored over time. The reports from the two clinical trials conducted in France and United Kingdom demonstrated clearly

that the cure of this disease had been achieved by the use of gene therapy (Cavazzana-Calvo *et al.*, 2000; Gaspar *et al.*, 2004). Since then, similar encouraging results have been published for the gene therapy clinical trials to treat adenosine deaminase deficient - SCID and chronic granulomatous disease (CGD) (Aiuti *et al.*, 2002; Gaspar *et al.*, 2006; Ott *et al.*, 2006).

After the positive news, the scientific community soon became apprehensive due to the development of a lymphoproliferative disorder that arose in two patients following gene therapy treatment. Analysis of the T-cell clones revealed a retroviral vector integration in the vicinity of the LIM domain only-2 (LMO2) proto-oncogene promoter, leading to aberrant translation of the protein. Although there is already a strong advantage for the genetically modified cells to proliferate, the activation of the LMO2 gene expression probably boosted this ability of the clones to the point of malignancy (Hacein-Bey-Abina *et al.*, 2003a; Hacein-Bey-Abina *et al.*, 2003b).

The leukaemias in the French SCID-X1 gene therapy trial represent the first reports of gene therapy vector-induced insertional mutagenesis in humans. These news prompted the scientific community to endeavour all necessary efforts to understand how viral vectors can cause insertional mutagenesis. Several *in vivo* and *in vitro* assays have been developed to investigate and quantify this process. A protocol has been adapted from the observation that primary bone marrow cells can be immortalised by infection with gammaretroviruses (Du *et al.*, 2005). This *in vitro* method is able to quantify insertional mutagenesis that leads to transformation by comparison with background. It has been possible to establish that the transformation frequency detected with a SIN retroviral vector is lower compared with a non-SIN counterpart and that the cellular promoter EF1a used as internal promoter in a SIN retroviral vector does not give rise to detectable transformed cells (Zychlinski *et al.*, 2008; Modlich *et al.*, 2006). Another recent report using a different assay to quantify insertional mutagenesis also demonstrates that SIN vectors are safer compared with the non-SIN design. In this study, the rise of clones that can grow independently of interleukin-3 in BAF3 derived cells, that normally need this factor for growth, gives an estimate of the frequency of insertional mutagenesis. Retroviral and lentiviral-based vectors were compared in this study and it was found that the transformation frequency is similar, but by a different mechanism. Lentiviral vectors transform cells by activating the

growth hormone receptor (*ghr*) transcription whereas gammaretroviral vectors upregulate expression of cytokine IL-3. It was also shown that the SIN lentivector design prevents activation of the *ghr* gene (Bokhoven *et al.*, 2009).

One assay using mice models have also been used to compare insertional mutagenesis induced by retroviral and lentiviral-based vectors. In this study, bone marrow cells from the tumour prone *Cdkn2a*^{-/-} mouse were transduced with a SIN lentivector and a non-SIN retroviral vector, transplanted to lethally irradiated mice and tumourigenesis was monitored over time. The results demonstrate that SIN-lentiviral vectors have a reduced potential for insertional mutagenesis (Montini *et al.*, 2006).

1.6 – Project aims

Lentiviral vectors are very promising tools for gene delivery due to their relatively large packaging capacity and ability to infect a range of different cell types. The vector is integrated into the host genome promoting stable transgene expression in dividing cells. However, integration can also be problematic due to a variation in gene expression between cells, possible gene silencing and most importantly, insertional mutagenesis which can lead to undesirable effects such as malignant transformation. Expression of a transgene in non-dividing cells does not require integration. Also, transient expression may have useful applications. The aim of this project is therefore to study non-integrating lentiviral vectors.

In order to produce NILVs, integration has to be blocked specifically. IN cannot be deleted because it is involved in reverse transcription (Zhu *et al.*, 2004) and possibly in nuclear entry (Gallay *et al.*, 1996). Nevertheless, several groups have demonstrated that mutations in IN or in the *cis*-acting sequences interacting with IN (att sites) blocks integration. Here, a combination of selected mutations has been introduced in the lentiviral transfer vector and in the packaging construct.

The resulting vectors will then be suitable for gene therapy in post-mitotic tissues. NILVs have previously been studied in retina and brain. Hence, in order to extend the NILVs application, expression from these vectors has been studied in muscle. Expression has been compared with the integrating counterpart and with other episomal vectors.

Another possible application of NILVs is the delivery of a protein to a target cell (pseudo-transduction). A new vector, based on NILVs, has been developed with this objective. In this vector, a foreign protein has been incorporated within the virion during vector production and delivery to the target cells was studied.

Chapter Two
Material and Methods

2 – Material and Methods

2.1 – Materials

All tissue culture reagents were supplied by Gibco BRL (Invitrogen) and all general chemicals were supplied by Sigma, unless otherwise stated. Solutions and buffers stated in the text are listed in table 2.1. Primers (Invitrogen) and probes (MWG) used are listed in table 2.2.

2.1.1 - Reagents and Enzymes

1kb Plus DNA Ladder	Invitrogen
AEC substrate	Vector Laboratories
Agar	Merck
Agarose	Invitrogen
Alexa Fluor 488 Phalloidin	Invitrogen
Ampicillin	Stratagene
Calf intestinal alkaline phosphatase	New England Biolabs
DNA polymerase I large (Klenow) fragment	Promega
ECL western blotting substrate	Pierce
Formaldehyde (methanol free) 16% Ultrapure	Polysciences, Inc
HALT protease inhibitor cocktail	Thermo Scientific
Histoclear solution	R.A. Lamb
HiTrap Heparin column	Sigma
Iodixanol	Accurate Chemical
Kodak BioMax film	Kodak
Lipofectamine 2000	Invitrogen
Lysis buffer	Roche
MOPS Buffer	Invitrogen
NuPAGE NOVEX Bis-Tris gels	Invitrogen
<i>PfuTurbo</i> DNA polymerase, buffer	Stratagene

Platinum qPCR SuperMix-UDG with ROX	Invitrogen
Proteinase K, PCR-grade	Roche
Rapid-Hyb TM Buffer	GE Healthcare
Rediprime II Random Prime Labelling System	GE Healthcare
Redivue [α - ³² P] dCTP	GE Healthcare
Restriction Endonucleases	New England Biolabs
SDS PAGE running buffer	Invitrogen
T4 DNA Ligase	New England Biolabs
<i>Taq</i> DNA polymerase, dNTPs, buffer	Promega
Vectashield hard set mounting medium containing Dapi	Vector Laboratories

2.1.2 - Antibodies

Matched-pair of antibodies set for hFIX ELISA	Affinity Biologicals Inc.
Monoclonal anti-eGFP (AB1218-100)	Abcam
Biotin-conjugated polyclonal anti-eGFP (AB6658-100)	Abcam
Rabbit anti-GFP rabbit serum (A6455)	Invitrogen
Biotin-conjugated goat anti-rabbit antibody (BA-1000)	Vector Laboratories

2.1.3 - Kits

Plasmid miniprep/maxiprep/megaprep kit	Qiagen
QIAquick gel extraction kit	Qiagen
QIAquick PCR purification kit	Qiagen
Quick change II XL site-directed mutagenesis kit	Stratagene
HIV-1 p24 Antigen ELISA	Beckman Coulter
Biorad protein assay kit	Biorad
Vectastain ABC kit	Vector Laboratories

2.1.4 - Centrifuges

Microcentrifuge

Tabletop centrifuge

Superspeed centrifuge

Ultracentrifuge

Ultra centrifuge

Heraeus Biofuge Fresco

Sorvall Legend RT

Sorvall Evolution RC

Sorvall *Discovery* SE

Beckman Coulter 60 Ti

Table 2.1: Buffers and solutions

Buffers	
1x Phosphate-buffered saline (PBS)	1.9mM NaH ₂ PO ₄ , 8.1mM Na ₂ HPO ₄ and 154mM NaCl, pH 7.4
1x Hanks balanced salt solution (HBSS)	5.4mM KCl, 0.3mM Na ₂ HPO ₄ ·7H ₂ O, 0.4mM KH ₂ PO ₄ , 4.2mM NaHCO ₃ , 1.3mM CaCl ₂ , 0.5mM MgCl ₂ ·6H ₂ O, 0.6mM MgSO ₄ ·7H ₂ O, 137mM NaCl and 5.6mM D-glucose, pH 7.4
1x TAE buffer	40mM Tris (pH 7.8), 20mM sodium acetate and 1mM EDTA
6x DNA loading buffer	10mM Tris pH 7.5, 50mM EDTA, 10% Ficol 400 and 0.05% bromophenol blue
Nuclei lysis buffer	10mM Tris-HCl, 400mM NaCl, and 2mM Na ₂ EDTA, pH 8
Proteinase K buffer	2mM Na ₂ EDTA and 1% SDS
ELISA coating buffer	1.59g/L of Na ₂ CO ₃ and 2.93g/L of NaHCO ₃ , pH 9.6
ELISA wash buffer	PBS with 0.05% (v/v) Tween 20
ELISA blocking buffer	3% of BSA in PBS with 0.05% (v/v) Tween 20
ELISA sample diluent	23.8g/L HEPES, 5.84g/L NaCl, 3.72g/L Na ₂ EDTA, 10% bovine serum albumin with 0.1% Tween-20, pH 7.2
Citrate-phosphate buffer	5.2g/L citric acid and 13.8g/L Na ₂ HPO ₄
OPD substrate solution	5mg of o-Phenylenediamine.2HCl added to 12mL of citrate-phosphate buffer containing 0.03% H ₂ O ₂
Sodium-citrate buffer	2.94 g/L sodium citrate, pH 6
Bacterial media and solutions	
LB (Luria-Bertani) agar	1% bacto typtone, 0.5% bacto yeast, 0.5% NaCl (pH 7.0) with 15g/L bactoagar
LB (Luria-Bertani) broth	1% bacto typtone, 0.5% bacto yeast and 0.5% NaCl (pH 7.0)
SOC medium	2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl ₂ , 10mM MgSO ₄ and 20mM glucose

Table 2.2: Primers and probes

Primers for PCR	
eGFP Forward 1	5' – ATGGTGAGCAAGGGCGAGGAGCTG – 3'
eGFP Reverse 2	5' – TGCTGGTAGTGGTCGGCGAGCTGC – 3'
eGFP Forward 3	5' – CCTGGTCGAGCTGGACGGCC – 3'
eGFP Reverse 4	5' – CCGTCCTCGATGTTGTGGCG – 3'
eGFP BspEI Forward	5' – ATCCGGTCTCCGGAATGGTGAGCAAGG – 3'
eGFP BspEI Reverse	5' – CCTTGCTCACCATTCCGGAGACCGGAT – 3'
Primers for mutagenesis	
D64V Forward	5' – CCAGGAATATGGCAGCTAGTTTG TACACATTTAGAAGG – 3'
D64V Reverse	5' – CCTTCTAAATGTGTACAAAC TAGCTGCCATATTCCTGG – 3'
N120L Forward	5' – GTACATACAGACAATGGCA GCCTTTTCACCAGTACTACAGTTAAGG – 3'
N120L Reverse	5' – CCTTAACTGTAGTACTGGTG AAAAGGCTGCCATTGTCTGTATGTAC – 3'
W235E Forward	5' – GGACAGCAGAGATCCAGTT GAGAAAGGACCAG CAAAGC – 3'
W235E Reverse	5' – GCTTTGCTGGTCCTTTCTCAACTGG ATCTCTGCTGTCC – 3'
Q1248A Forward	5' – GGCATTCCCTACAATCCCCAA AGTGCAGGAGTAATAGAATCTATG – 3'
Q148A Reverse	5' – CATAGATTCTATTACTCCTGCA CTTTGGGGATTGTAGGGAATGCC – 3'
K264R Forward	5' – GTAGTGCCAAGAAGAA GAGCAAAGATCATCAGGG – 3'
K264R Reverse	5' – CCCTGATGATCTTTG CTCTTCTTCTTGCACTAC – 3'
K264,266,273R Forward	5' – CCAAGAAGAAGAGCAAGGATC ATCAGGGATTATGGAAGACAGATGGCAGG – 3'
K264,266,273R Reverse	5' – CCTGCCATCTGTCTTCCATAAT CCCTGATGATCCTTGCTCTTCTTCTTGG – 3'
att 3'LTR forward	5' – GCCACTTTTTAAAAGAAAA GGGGGACCAGAAGGGCTAATTC – 3'

att 3'LTR reverse	5' – GAATTAGCCCTTCTGGTCCC CCCTTTTCTTTTAAAAAGTGGC – 3'
att 5'LTR forward	5' – GGAAAATCTCTAGTGGTGGC GCCCGAACAGGGACTTGAAAGCG – 3'
att 5'LTR reverse	5' – CGCTTTC AAGTCCCTGTTCCGGG CGCCACCACTAGAGATTTTCC – 3'
P2mut Forward	5' – GGCTGAAGCAATGAGCT CCGGAACAAATCCAGCTACC – 3'
P2mut Reverse	5' – GGTAGCTGGATTTGTTCC GGAGCTCATTGCTTCAGCC – 3'
P2ins1 Forward	5' – GCAATGAGCCAAGTAACAAATGGC AATTCCGGACCAGCTACCATAATGATA C – 3'
P2ins1 Reverse	5' – GTATCATTATGGTAGCTGG TCCGGAATTGCCATTTGTTACTTGGCTCATTGC – 3'
P2ins2 Forward	5' – GCAATGAGCCAAGTAACAAA TATGATACAGAAAGGCAATTCCGGACCAGCTACC – 3'
P2ins2 Reverse	5' – GGTAGCTGGTCCGGAATTGCCTT TCTGTATCATATTTGTTACTTGGCTCATTGC – 3'
P2ins3 Forward	5' – GCAATGAGCCAAGTAACAAATC CAGCTACCATAATGATACAGAAAGGCAATTCCGG – 3'
P2ins3 Reverse	5' – CCGGAATTGCCTTTCTGTATCAT TATGGTAGCTGGATTTGTTACTTGGCTCATTGC – 3'
Primers and probes for qPCR	
Total viral DNA (Late RT) forward	5' – TGTGTGCCCGTCTGTTGTGT – 3'
Total viral DNA (Late RT) Reverse	5' – GAGTCCTGCGTCGAGAGAGC – 3'
Total viral DNA (Late RT) Probe	(FAM) 5' – CGCCCGAACAGGGACTTGAA – 3' (TAMRA)
Total viral DNA (WPRE) forward	5' – TGGATTCTGCGCGGGA – 3'
Total viral DNA (WPRE) Reverse	5' – GAAGGAAGGTCCGCTGGATT – 3'
Total viral DNA (WPRE) Probe	(FAM) 5' – CTTCTGCTACGTCCCTTCGGCCCT – 3'

2LTR circles forward	5' – AACTAGAGATCCCTCAGACCCTTTT – 3'
2LTR circles reverse	5' – CTTGTCTTCGTTGGGAGTGAATT – 3'
2LTR circles probe	(FAM) 5' – CTAGAGATTTTCCACACTGAC – 3'
Titin forward	5' – AAAACGAGCAGTGACGTGAGC – 3'
Titin reverse	5' – TTCAGTCATGCTGCTAGCGC – 3'
Titin probe	(FAM) 5' – TGCACGGAAGCG TCTCGTCTCAGTC – 3' (TAMRA)
Primers for sequencing	
att seq LA 1R	5' – TTCAGCAAGCCGAGTCCTGCG – 3'
att seq LA 2F	5' – TCGAGGCAATTCGAGCTCGG – 3'
att seq LA 3R	5' – CCTTATGTATCATAACATACG – 3'
int seq LA 0F	5' – GCTGTGATAAATGTCAGC – 3'
int seq LA 1R	5' – CCTCTGTGCGAGTAACGCC – 3'
int seq LA 2R	5' – ATTGTGGATGAATACTGCC – 3'
int seq LA 3F	5' – GACAGGTAAGAGATCAGGC – 3'
P2 seq LA 1F	5' – GGATGACAGAAACCTTGTGG – 3'

2.2 – Molecular Biology Techniques

2.2.1 – Growth and maintenance of *E.coli*

Escherichia coli (*E.coli*) were grown in liquid LB media at 37°C with agitation at 250rpm or streaked out on solid LB plates containing 1.5% bacto agar. *E.coli* transformed with plasmids were grown on the same media supplemented with ampicillin (50µg/ml).

For long-term storage, bacterial cultures were stored in 20% volume for volume (v/v) of glycerol in LB media at -80°C.

2.2.2 – Production and transformation of electro-competent *E.coli*

1L of LB media without antibiotics was inoculated 1:1000 from a fresh overnight culture of *E.coli* DH5 α . The culture was grown at 37°C with agitation until the optical density at 600nm (OD600) reached between 0.6-0.8. The cells were harvested by centrifugation at 4500rpm in a superspeed centrifuge for ten minutes, the supernatant discarded and the cell pellet resuspended in 1L of cold sterile 10% glycerol. The spin was repeated, the supernatant discarded and the cell pellet was resuspended in 250mL of cold sterile 10% glycerol. The cells were centrifuged again, after which the supernatant was carefully discarded and the cells resuspended in the 10mL of 10% glycerol. The cells were transferred to a smaller centrifuge tube and spun at 7000rpm for ten minutes. The supernatant was removed and the cell pellet resuspended in 2mL cold sterile 10% glycerol, 100µL of cells were transferred to microfuge tubes and snap-frozen in a dry ice-ethanol bath. The electro-competent cells were subsequently stored at -80°C.

To introduce DNA in electro-competent *E.coli*, cells were thawed slowly on ice and mixed with 1-10ng of desalted DNA. The cell/DNA mixture was then pipetted into pre-chilled disposable micro-electroporation cuvettes (0.2cm) and electroporated at 2.5kV, 200Ω, 25µF using a Bio-Rad Gene Pulser. 1mL of room temperature SOC media was added to the cuvette and the media/bacterial mixture was transferred to a 5mL tube. The cells recovered at 37°C, with agitation at 250rpm for one hour. The cells were pelleted by centrifugation at 2000rpm for five minutes and the supernatant discarded. The cells were

then resuspended in 100 μ L and spread on LB agar plates containing the appropriate antibiotics. The plates were incubated overnight at 37°C.

2.2.3 – Production and transformation of chemical-competent *E.coli*

100mL LB media without antibiotics was inoculated with 1:1000 from a fresh overnight culture of *E.coli* DH5 α . This was grown in a shaking incubator at 37 °C until the OD600 was 0.6. The culture was cooled on ice for five minutes, then the bacteria were pelleted and re-suspended in 50mL ice-cold TFB-I buffer. After five minutes incubation on ice, bacteria were pelleted again, re-suspended in 5mL ice-cold TFB-II buffer and incubated on ice for fifteen minutes. 100 μ L aliquots were prepared and stored at -80°C.

To transform chemical-competent cells, bacteria were slowly thawed on ice and DNA added to each aliquot. This mixture was left on ice for thirty minutes. Bacteria were heat shocked for forty five seconds at 42°C and immediately put back on ice for two minutes. 1mL of SOC medium was added and cells recovered in a shaking incubator at 37°C for one hour. The cells were pelleted and the supernatant discarded. Bacteria were then resuspended in 100 μ L of LB media, spread on LB agar plates containing the appropriate antibiotics and incubated overnight at 37°C.

2.2.4 – Plasmid DNA preparation

Small-scale plasmid DNA was extracted from a 5mL bacteria culture grown overnight at 37°C using the min-prep kit from Quiagen as per manufacturer's instructions. Large-scale plasmid DNA was obtained from 100mL or 500mL bacteria cultures using a Quiagen maxi-prep or mega-prep kit, respectively, following the manufacturer's instructions.

The plasmid DNA concentration was calculated by measuring the absorbance of light with a wavelength of 260nm using a NanoDrop ND-1000 spectrophotometer with a 0.2mm path length, at this wavelength 50 μ g/mL of double-stranded DNA has an absorbance of 1.

2.2.5 – Cloning

2.2.5.1 – Restriction enzyme digestion

Plasmid DNA (typically 1-2 μ g) was digested in a final volume of 10 μ L of 1x buffer (supplied by the manufacturer) and bovine serum albumin (0.1mg/ml). The amount of enzyme used varied depending of the concentration of the enzyme stock and the amount of DNA, but never exceeded 10% (v/v) of the total reaction volume. The endonuclease reaction was carried out at the manufacturer's recommended temperature for 1-3 hours and DNA digestion was verified by agarose gel electrophoresis. Double or triple digestions were performed either in compatible buffers or sequentially, after clean-up of DNA by ethanol precipitation using the Qiaquick PCR purification kit and following the manufacturer's protocol.

2.2.5.2 – Polymerase Chain Reaction (PCR)

PCR reactions were performed in a total volume of 50 μ L containing template DNA (10ng), forward and reverse primers listed in table 2.2 (0.5 μ M each), dNTPs (each at 200 μ M), MgCl₂ (1.5mM), buffer (1x) and *Taq* DNA polymerase (1U). Typically cycling conditions were as follows: 30 cycles of 94°C for 1min, 60°C for 1min, 72°C for 1min.

2.2.5.3 – Isolation of DNA fragments

DNA fragments were resolved by electrophoresis through 0.8-2% agarose gels in 1x TAE buffer. To prepare the gels, agarose was dissolved in 1x TAE buffer by boiling in a microwave. After cooling, ethidium bromide was added (to obtain a final concentration of 0.5 μ g/ml) for visualisation of DNA. DNA samples were mixed with DNA loading buffer before loading onto agarose gels. A 1 kb plus DNA ladder was included in each gel to enable size determination of DNA fragments. Gels were electrophoresed using a voltage of 50-100V (up to 150mA) and the separated DNA fragments subsequently visualised by exposure to ultra-violet light using an UVIdoc gel documentation system.

Following electrophoresis, DNA fragments were excised from agarose gels using a clean scalpel blade under ultra-violet light. The DNA was then extracted from the agarose using a QIAquick gel extraction kit as per the manufacturer's instructions.

2.2.5.4 – Ligation of DNA fragments

When the digested vector DNA ends had compatible termini with the insert DNA, the digested plasmid vector DNA was treated with calf intestinal alkaline phosphatase (CIP) to dephosphorylate the DNA ends prior to ligation. Dephosphorylation reactions were performed directly in restriction endonuclease buffers, immediately following digestion by adding 10 units of CIP enzyme to the reaction mixture and incubating at 37°C in a water bath for one extra hour.

Ligation of DNA fragments were performed using vector to insert ratios of 1:1 or 1:4 (v/v) in a final volume of 10µL of 1x T4 DNA Ligase buffer containing 1 unit of T4 DNA Ligase. The vector DNA concentration used was typically 20-50ng of DNA. Ligation reactions were incubated overnight at room temperature or 16°C. The ligation reactions were either immediately transformed into *E.coli* or stored at -20°C for later use.

2.2.6 – Genomic DNA Isolation

DNA was extracted from samples using the salting out method (Miller *et al.*, 1988). Cells were pelleted by centrifugation at 1500rpm for five minutes. The supernatant was discarded and it was added 3mL of nuclei lysis buffer, 0.5mL of proteinase K buffer, 0.2mL of SDS (10%) and proteinase K (1mg). The mixture was digested overnight at 37°C. 1mL of saturated NaCl (6M) was added and the tubes were shaken vigorously for at least fifteen seconds. The tubes were centrifuged at 4500rpm for fifteen minutes and the supernatant transferred to a clean tube. Two volumes of absolute ethanol were added and the tubes were inverted several times until the DNA precipitated. The DNA was then removed, passed through ethanol (70%), transferred to a clean eppendorf and resuspended in 100µL of H₂O. The final concentration of DNA was accessed by quantification the absorbance at 260nm using a NanoDrop ND-1000 spectrophotometer.

2.2.7 – Total RNA Isolation

1-10x10⁶ cells were well resuspended in 1mL of TRI reagent and incubated at room temperature for ten minutes. 0.2mL of chloroform was added to each sample after which, the tubes were shaken vigorously and then left resting at room temperature for fifteen minutes. The tubes were then centrifuged at 12000rpm for fifteen minutes. The upper aqueous phase was transferred to a new tube and 0.5mL of isopropanol was added to each sample to precipitate the RNA. The samples were incubated for one hour at –20°C and then the RNA was pelleted by centrifugation at 12000rpm for ten minutes. The supernatant was discarded and the RNA was washed with 1mL of ethanol. The ethanol was removed by centrifugation (12000rpm, ten minutes) and the RNA resuspended in H₂O. The RNA was quantified by measuring the absorbance at 260nm using a NanoDrop ND-1000 spectrophotometer, at this wavelength 40µg/mL of RNA has an absorbance of 1.

2.2.8 – Site directed mutagenesis

Mutagenesis of plasmid DNA was performed using the Quickchange II XL site-directed mutagenesis kit from stratagene. The primer design was performed according to the manufacturer directions. Table 2.2 shows the primers used for each reaction. The PCR reaction was performed in a total volume of 50µL, containing a total of 100ng of template DNA, buffer (1x), 125ng of primer forward, 125ng of primer reverse, 1µL of dNTP mix, 3µL of quick solution reagent and 2.5U of *Pfu Turbo* DNA polymerase. All reactions followed the thermal cycling parameters: 1min at 95°C followed by 16 cycles of 30s at 95°C, 1min at 55°C and 11min at 68°C. After the PCR reaction, the parental DNA was digested with *Dpn* I enzyme for three hours. The resulting DNA was then transformed in chemical competent bacteria. DNA was extracted from single clones and sequenced to confirm correct mutagenesis.

2.3 – Tissue Culture Techniques

2.3.1 – Cell lines propagation

The following cell lines were used: 293T, C2C12 and HT1080. 293T cells are a highly transfectable cell line derived from the 293 human embryonic kidney cell line that express the SV40 large T antigen. The C2C12 cell line is a mouse myoblast cell line, which can be terminally differentiated into muscle fibres by serum starvation (Yaffe and Saxel, 1977). HT1080 is a human cell line derived from a fibrosarcoma (Rasheed *et al.*, 1974) that maintains a close karyotypic similarity to the initial karyotype over cell passages (Chen *et al.*, 1983).

All cell lines were maintained in Dulbecco's modified eagle medium (DMEM) containing GlutaMAX supplemented with foetal calf serum (10% v/v), penicillin (100UI/mL) and streptomycin (100µM/mL). Cells were grown in 37°C incubators in a 5% CO₂ atmosphere. Cells were passaged when confluency reached 90-95%, the monolayers were first washed with Dulbecco's phosphate buffered saline (PBS) and then incubated for 2-5 minutes at 37°C with trypsin/EDTA, the cells were subsequently diluted 1:3 up to 1:10 in fresh complete culture media and transferred to new tissue culture flasks.

C2C12 cells were differentiated using DMEM supplemented with 2% horse serum, penicillin (100UI/mL) and streptomycin (100µM/mL) to maintain the cells.

2.3.2 – Cell lines storage

Cells from a 90% confluent monolayer in a 175cm² tissue culture flask were pelleted by centrifugation at 1200rpm in a tabletop centrifuge for five minutes, resuspended in 10mL freezing medium (90% FCS, 10% dimethylsulfoxide) and transferred to a cryovial. Cells were frozen slowly overnight to -80°C in an isopropanol freezing box and then transferred to liquid nitrogen.

To revive frozen cells, aliquots were thawed rapidly in a 37°C water bath and transferred to 4mL growth medium. The cells were pelleted at 1200rpm in a centrifuge to remove the dimethylsulfoxide and then resuspended in 5mL culture media and transferred to a 80cm² tissue culture flask.

2.3.3 – Lentiviral vector production

Viral vectors were produced by transient transfection. 10^7 293T cells were seeded in 175cm² tissue culture flasks one day before transfection. The transfection media was prepared by mixing 50µg of vector plasmid, 37.5µg of packaging plasmid and 17.5µg of envelope plasmid in 5mL of Optimem and 5mL of Optimem containing 2nM of Polyethylenimine (PEI). This transfection media was incubated twenty minutes at room temperature prior to addition onto cells. Cells were washed with PBS and the transfection media was added to the cells. Cells were incubated for three hours at 37°C and 5% CO₂. The transfection media was removed and fresh culture media was added to the cells. On the following day, the media was removed and 30mL of fresh culture media was added to each flask. Two and three days after transfection, the media was collected and cell debris were removed by filtration through a 0.22µm filter. The media was then centrifuged at 150000g in an ultracentrifuge. The supernatant was removed and 150µL of Optimem was added to each tube and the tubes were incubated on ice for twenty minutes. The pellet was resuspended and 20µL aliquots were prepared and rapidly frozen at -80°C.

2.3.4 – Adeno-Associated vector production

Viral vectors were produced by transient transfection. 10^7 293T cells were seeded in ten 20cm tissue culture plates one day prior to transfection. The transfection media was prepared by mixing 50mL of Optimem containing 300µg of vector plasmid and 1400µg of packaging plasmid (pDG) with 50mL of Optimem containing 2nM of PEI. This mixture was then incubated twenty minutes at room temperature. Cells were washed with PBS and 10mL of transfection media was added to each plate. Cells were incubated for four hours at 37°C and 5% CO₂. The transfection media was removed and fresh culture media was added to the cells. Cells were harvested two days after. Cells were then pelleted and resuspended in 15mL of 50mM Tris (pH 8.5) containing 150mM NaCl. Cells were then lysed by three freeze-thaw cycles. Benzonase was then added to the mixture at a final concentration of 50U-mL and incubated at 37°C for one hour.

An iodixanol gradient was used to purify the vectors. The iodixanol gradient was prepared in 35mL ultracentrifuge tubes, overlaying (from bottom to top) 5mL of 60%

iodixanol, 5mL of 40% iodixanol, 6mL of 25% iodixanol and 9mL of 15% iodixanol. The cell lysate was then transferred to this tube and the tube sealed with rotor cap accessory. The tubes were then centrifuged at 350000g for two hours. The vectors were then removed from the clear 40% layer by making an air hole on the top of the tube and removing the layer with a needle. The vectors were further purified using a HiTrap heparin column. Aliquots were then prepared and frozen at -80°C .

2.3.5 – Vector transduction *in vitro*

Cells were seeded one day before transduction. Typically, 10^5 cells/well in 24 well plates or 5×10^5 cells/well in 6 well plates were used. The vectors were diluted to obtain the desired final vector concentration in half the typical culture volume used to grow cells and added to the cells. One day after infection, the media was removed and fresh culture media added to the cells when cells were grown more time.

2.3.6 – Vector infection *in vivo*

This procedure was performed by Dr Simon Waddington. 1-day old neonatal mice were placed on ice slush until they no longer moved when touched. For intravenous injection, the heads were illuminated and 40 μL of concentrated lentiviral vectors injected into the superficial temporal vein (a prominent vessel located on either side of the head, visible just below the eye) using a 33G Hamilton needle. For intramuscular injections, the hind legs were illuminated and 5 μL of concentrated lentiviral vectors injected in the tibialis anterior muscle using a 33G Hamilton needle. Following the procedure, the mice were allowed to return to normal temperature on a warming mat before being returned to their mothers.

2.3.7 – Integration efficiency

The level of integration of the different vectors was quantified by infecting 1×10^6 HT1080 cell with vectors expressing neomycin phosphotransferase (Neo). Three days after infection, the medium was replaced with medium containing 1mg/mL of G418 (Invivogen). Cells were grown for three weeks and colonies from single clones were counted.

2.4 – Analyses of *in vitro* and *in vivo* samples

2.4.1 – Flow cytometry

To calculate the titre of a vector batch expressing a fluorescent marker, 10^5 cells per well, in a 24 well plate, were infected with serial dilutions of the vector, in a final volume of 300 μ L. The cells were harvested two days after transduction, washed with PBS and pelleted in a FACS tube by centrifugation at 1500rpm for five minutes. The samples were then resuspended in 200 μ L of PBS and 200 μ L of PFA (2%) was added to each sample. Cells were then stored in the dark and analysed by flow cytometry within two days. The titre of the vector was then calculated by multiplying the percentage of fluorescent cells (less the percentage of positive cells in an untransduced well) by 10^5 and divided by the volume of vector used in that well. Only wells that had between 5 and 40% of fluorescent positive cells were used to quantify the titre.

To quantify the percentage or the intensity of fluorescent positive cells, cells were infected with a viral vector or transfected with a plasmid. Samples were harvested at a given time after transduction and an average of $2-5 \times 10^5$ cells were then used for flow cytometry. The cells were pelleted by centrifugation (1500rpm for five minutes), washed with PBS and resuspended in 200 μ L of PBS. 200 μ L of PFA (2%) was then added to each sample. The samples were then analysed by flow cytometry.

All samples were analysed using a CyAn ADP flow cytometry analyser and Summit version 4.1 software (DakoCytomation).

2.4.2 – Titration of p24 in lentiviral preparations

The quantity of p24 antigen in lentiviral supernatants was measured using a Coulter HIV-1 p24 antigen assay as per the manufacturer's instructions. Supernatants were diluted 1:1 x 10⁶, 1:5 x 10⁶ and 1:1 x 10⁷ and added to p24 antibody-coated wells together with diluted calibration standards. Samples were lysed, washed and then incubated with biotinylated anti-HIV-1. The samples were washed again and subsequently incubated with streptavidin-HRPO followed by the addition of tetramethylbenzidine and Coulter stop reagent (4N H₂SO₄). The absorbance of each plate was then read at 450nm and the quantity of p24 reagent calculated from a standard curve generated from the diluted calibration standards.

2.4.3 – Quantitative real-time PCR

Quantitative real-time PCR (qPCR) was used to determine the total cell number in a sample, the copy number of total viral DNA or 2LTR viral circles in transduced cells or the number of plasmids in transfected cells. Total DNA was isolated as described in 2.2.6. Between 10 and 100ng of DNA were then used as a template for each PCR. Reactions were performed in triplicate in a final volume of 25µL containing forward and reverse primers at 0.9µM (table 2.2), a fluorescently-labelled probe at 0.2µM and a pre-mixed mastermix (Platinum qPCR SuperMix-UDG with ROX) diluted to 1x with sterile distilled water. Standard curves of diluted plasmids harbouring the target DNA were used for quantification. Reactions were performed using an ABI Prism 7000 Sequence Detection System (Applied Biosystems) using the default cycling parameters – 1 cycle of 50°C for 2min, 1 cycle of 95°C for 10min followed by 40 cycles of 95°C for 15s and 60°C for 1min.

2.4.4 – Enzyme-linked immunosorbent assay (ELISA)

2.4.4.1 – Factor IX (FIX) ELISA

Approximately 90µL of peripheral blood was collected from the tail vein of mice and mixed immediately with 10µL of sodium citrate buffer (0.1M). Samples were then

centrifuged at 13000rpm in a micro-centrifuge for ten minutes and the serum transferred to fresh eppendorfs and stored at -20°C.

Maxisorp 96-well NUNC-Immuno ELISA plates (NUNC) were coated with 100µL/well with FIX capture antibody diluted 1:100 in coating buffer, wrapped in cling film and stored overnight at 4°C. The coated plates were washed three times with 100µL of ELISA wash buffer and blocked for one hour at room temperature in ELISA blocking buffer. The plates were washed a further three times and 100µL of diluted samples in ELISA sample diluent added in triplicate along with 100µL of diluted human plasma. Plates were incubated at room temperature for ninety minutes and washed three times. 100µL/well of FIX detection antibody diluted 1:100 in ELISA sample diluent was added to each well and plates were incubated at room temperature for ninety minutes in the dark. The plates were subsequently washed three times and 100µL/well of OPD substrate solution was added. The plates were incubated for eight minutes at room temperature in the dark, at which time the reaction was stopped by adding 50µL stop solution (2N H₂SO₄). The absorbance of each well was then read at 490nm using a FLUOstar Optima plate reader. Standard curves of diluted human plasma were used to quantify the percentage of hFIX contained in each sample.

2.4.4.2 – Enhanced green fluorescent protein (eGFP) ELISA

Samples were prepared from muscle or cultured cells. Protein was extracted in lysis buffer using a homogenizer. Total protein was assayed using the Biorad protein assay kit with bovine serum albumin as a standard and following the manufacturer's instructions. The eGFP was quantified in triplicate against a recombinant eGFP standard using the monoclonal antibody ab1218-100 (1:10000) as primary antibody, and the biotin labelled secondary antibody ab5688 (1:5000). Streptavidine-peroxidase was used to develop the ELISA, and the absorbance was read at 450nm using a FLUOstar Optima plate reader.

2.4.5 – Immunohistochemistry

Samples from the tibialis anterior muscle were fixed in paraformaldehyde and sections (5µm) were cut from paraffin-embedded tissue blocks. Slides were placed in histoclear solution and then hydrated in graded ethanol solutions. Antigen retrieval was performed using sodium citrate buffer and heating. Endogenous peroxidase was blocked by incubation with 0.3% H₂O₂ in phosphate-buffered saline for ten minutes. Slides were incubated with 2% normal mouse serum, then with avidin and biotin. Slides were incubated overnight with rabbit anti-GFP serum, and then with biotinylated goat anti-rabbit secondary antibody for forty minutes. After being thoroughly washed in PBS, the slides were incubated with Vectastain ABC kit in accordance with the manufacturer's instructions. AEC substrate was used for detecting peroxidase activity, in accordance with the manufacturer's instructions. Hematoxylin was used for counterstaining and slides were then mounted in aquamount. Images were obtained on a Leica DMLS upright microscope using a ×40 objective with a Canon Coolpix 4500 digital camera, and processed with Adobe Photoshop software.

2.4.6 – Western Blot

Cells were washed in PBS and pelleted in 1.5mL eppendorf tubes by centrifugation at 1500rpm for five minutes in a table-top centrifuge. Protein was harvested from cell pellets by resuspension in 100µL lysis buffer supplemented with 1µL HALT Protease inhibitor cocktail and incubation on ice for ten minutes with occasional shaking. Viral lysates were prepared by resuspending 25µL of concentrated virus in lysis buffer with 1µL HALT Protease inhibitor cocktail and incubating on ice for ten minutes. The lysates were spun at 13000rpm in a microcentrifuge at 4°C for fifteen minutes and the supernatants containing the nuclear and cytoplasmic protein transferred to fresh eppendorfs. The samples were diluted 1:1 with loading buffer (100mM Tris-HCl, 4% SDS, 0.2% bromophenol blue, 20% glycerol and 10% β-mercaptoethanol) and heated to 95°C for five minutes. Samples were then stored at -20°C until run in a polyacrylamide gel.

Protein samples were loaded onto 4-12% Bis-Tris Novex pre-cast gels along with a pre-stained protein standard. Proteins were separated by polyacrylamide gel electrophoresis at 200V for fifty minutes in an XCell *SureLock* Mini-Cell (Invitrogen) in 1x NuPAGE MOPS SDS running buffer. The separated proteins were subsequently immunoblotted onto PVDF membrane using the XCell II blot module (Invitrogen) at 30V for one hour. The PVDF membrane was washed with methanol for thirty seconds, rinsed in distilled water and then washed in 1x NuPAGE transfer buffer prior to immunoblotting.

Following immunoblotting, PVDF membranes were incubated overnight with blocking buffer (PBS, 0.1% (v/v) Tween 20, 5% (w/v) milk powder) to prevent non-specific antibody staining. Membranes were subsequently washed five times in PBS-T (PBS, 0.1% (v/v) Tween 20) and incubated with primary antibody in 3mL blocking buffer for two hours at room temperature with agitation. The membranes were washed a further five times in PBS-T and then incubated with an HRP-conjugated secondary antibody in 3mL blocking buffer for forty five minutes at room temperature with agitation. The membranes were washed for a final five times after which 1mL of ECL western blotting substrate was added directly to the membrane and incubated for one minute at room temperature. Protein bands were visualised using an UVIchemi (UVItec) chemiluminescence documentation system and UVIsoft (UVItec) software.

2.4.7 – Southern Blot

Genomic DNA was digested overnight at 37°C with a selected restriction enzyme. DNA fragments were then separated by overnight gel electrophoresis on a 0.6% agarose gel at 20V. The gel was sequentially submerged in 0.25M HCl for ten minutes, 0.5M NaOH with 1.5M NaCl for thirty minutes and 1M Tris-HCl + 1.5 M NaCl (pH 7.4) for thirty minutes. Each of these incubations was done on a revolving platform at room temperature and the gel was rinsed with distilled water in between incubations with the different solutions. DNA fragments were blotted overnight by neutral transfer, using 20X SSC, onto a positively charged nylon membrane (HybondTM-XL, Amersham Bioscience). The membrane was then washed in 2X SSC. The membrane was pre-hybridised in 15mL Rapid-HybTM Buffer supplemented with 10µg/mL Salmon Testis

DNA (Sigma) overnight at 65°C. DNA was probed with a 32P labelled GFP probe (PCR product generated using primers GFP Forward 1 and GFP Reverse 2). The probe was made using the Rediprime II Random Prime Labelling System. Each probe was labelled with 50µCi (1.85MBq) of Redivue [α -32P] dCTP. Unincorporated nucleotides were removed using the QIAquick Nucleotide Removal kit (Qiagen). Prior to addition to the hybridisation buffer, the probe was denatured by heating it to 100°C for five minutes. The membrane was hybridised in the presence of the probe overnight at 65°C. The following day, the membrane was washed sequentially in 2% SSC + 0.1% SDS, 1% SSC + 0.1% SDS, 0.5% SSC + 0.1% SDS and 0.25% SSC + 0.1% SDS. Each wash was performed at 65°C for twenty minutes. The membranes were then exposed to Kodak BioMax film.

2.4.8 – Confocal microscopy

One day prior before infection, 10^4 cells were plated on poly-L-Lysine coated slides. Cells were then transduced with serial dilutions of vectors and incubated at 37°C and 5% CO₂, for eighteen hours. Samples were then prepared for confocal microscopy. Cells were fixed in 4% PFA for thirty minutes at room temperature. Cells were then permeabilised with 0.5% Triton X for five minutes and stained with Alexa Fluor 488 Phalloidin for five minutes. Samples were always washed with PBS between each step. A cover slip was then placed on top of the cells and sealed with Vectashield hard set mounting medium containing Dapi. Samples were then visualized in a confocal microscope TCS SP2 (Leica) and data collected was analysed with Leica confocal software.

2.4.9 – Statistical Analysis

Data from the integration efficiency quantification assays and eGFP ELISA measurements were analysed using a Kuskal–Wallis test for statistical significance ($P < 0.05$), and particular groups were compared using a Mann–Whitney test.

Chapter Three

Development of Non-integrating Lentiviral Vectors

3 – Development of Non-Integrating Lentiviral Vectors

3.1 – Introduction

Vectors based on lentiviruses are able to deliver DNA into a broad range of cell types, both dividing and non-dividing. Furthermore, their characteristic of genome integration into the host genome makes them particularly useful when prolonged transgene expression is required in dividing cells. However, an integration event nearby an oncogene or a tumour suppressor gene may cause malignant transformation. The lymphoproliferative condition that arose in two patients in a successful clinical trial demonstrates the risk of insertional mutagenesis that may be caused by integrating vectors (Hacein-Bey-Abina *et al.*, 2003a; Hacein-Bey-Abina *et al.*, 2003b).

Integration may not be necessary in non-dividing cells. Also, transient transgene expression can have useful applications (Nightingale *et al.*, 2006). Upon reverse transcription, viral DNA enters the nucleus in order to be integrated into the host genome. Nevertheless, circular DNA molecules are also created as by-products of integration. These molecules can have one LTR (formed by homologous recombination between the two LTRs) or two LTRs produced by non-homologous end-joining of the double stranded blunt viral DNA produced upon reverse transcription (Butler *et al.*, 2002). The formation of circular genomes is thought to occur in the nucleus. As a consequence, these molecules have been used to quantify nuclear viral DNA. It was believed that these molecules were by-products of integration, but recent findings suggest that this non-integrated viral DNA could support transcription (Wu and Marsh, 2003; Brussel and Sonigo, 2004; Poon and Chen, 2003). Indeed, two reports have shown stable transgene expression in post-mitotic tissues *in vivo*, from non-integrating lentiviral vectors (NILVs) (Philippe *et al.*, 2006; Yanez-Munoz *et al.*, 2006).

There has been some controversy about these recent findings since first generation NILVs showed little transgene expression (Blomer *et al.*, 1997; Naldini *et al.*, 1996b). Bayer *et al.* have recently published a study showing that the deletion of a large sequence within the unique in 3' (U3) region of the virus (self inactivating (SIN) vectors) is responsible for the different expression profiles of these NILVs, compared with NILVs

containing the whole U3 region. The removal of that sequence in SIN vectors allows efficient expression from episomal lentiviral vectors (Bayer *et al.*, 2008).

In order to produce NILV that express the transgene of interest, integration has to be blocked specifically. Integrase (IN) cannot be deleted because it is involved in reverse transcription (Zhu *et al.*, 2004) and possibly in nuclear entry (Gallay *et al.*, 1996). Nevertheless, several groups have demonstrated that mutations in IN or in the *cis*-acting sequences interacting with IN (att sites) block integration.

The catalytic core of IN contains a motif (DD35E) that is conserved among retroviral integrases and transposable elements (Johnson *et al.*, 1986) and is absolutely necessary for integration. Mutation of any of the residues impairs integration (Leavitt *et al.*, 1996). Other mutations in residues can block specifically one of the catalytic steps of the integration reaction. For example, the Q148 residue is implicated in IN binding to viral DNA and 3'processing (Esposito and Craigie, 1998; Gerton *et al.*, 1998). Lysines at position 264, 266 and 273 are involved indirectly in strand transfer: Cereseto and colleagues demonstrated that these residues are acetylated in the target cells and that this modification is required for strand transfer (Cereseto *et al.*, 2005). A mutation in residue N120, which is near the catalytic site, but implicated in DNA binding, also blocks integration (Leavitt *et al.*, 1996; Lu *et al.*, 2005). Tryptophan at position 235 is also related with DNA binding and mutation of this amino acid blocks integration (Leavitt *et al.*, 1996). The conserved dinucleotides CA (att sites) in the viral genome are essential for integration, as they are the recognition sequence for integrase binding to the genomic viral DNA. Mutations in either U3 or U5 att sites reduce integration 60 and 30%, respectively. Most importantly, a synergistic effect is observed when the viral genome has both mutations, where integration represents 1% of the wt genome (Masuda *et al.*, 1998).

The aim of this work was the development and assessment of NILVs. Mutations that target different mechanisms of integration were studied and combined in a single vector. This may contribute towards improving the safety of using NILVs in gene therapy.

3.2 – Production of non-integrating lentiviral vectors

In order to produce non-integrating vectors, two approaches can be taken: mutation of the viral IN or mutation in the vector genome in the attachment sites (*att*), where IN binds. Figure 3.1 shows where point mutations were introduced. The mutations were generated by site-directed mutagenesis and confirmed by sequencing.

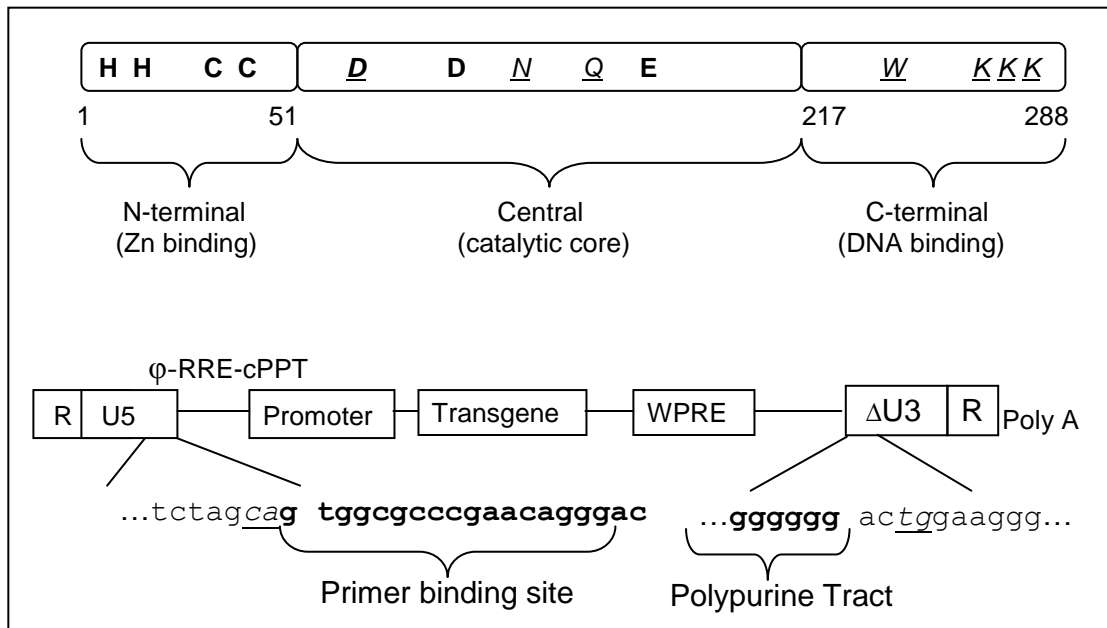


Figure 3.1: Schematic illustration of the mutations created in the integrase and in the viral genome.

A - Integrase domains representation. Bold represents conserved amino acids (Zn finger (HHCC) and catalytic core (DD(35)E), italic and underlined residues represents mutations done in the enzyme. **B** - Viral genomic RNA representation. Mutations produced in the attachment sites (conserved dinucleotides) are shown in italic and underlined

To confirm that the introduction of the mutations did not affect vector production, viral particles were assessed by quantification of the p24 protein, using an ELISA kit. As shown in Table 3.1, the vector stocks had similar order of magnitude of viral p24 protein for all of the mutant and the integrating vectors, implying that the mutations did not affect vector production.

Infectivity was measured by assessing the concentration of transducing units. 293T cells were infected with several dilutions of the viral vectors expressing eGFP from the p5 promoter. This is a promoter derived from adeno-associated virus. Two days after infection, cells were harvested and the number of eGFP expressing cells was accessed by flow cytometry. Table 3.1 shows that high titre was achieved for all of the vector mutants.

Table 3.1: Quantification of viral particles and infectivity

Vectors	Viral particles (pg (p24)/mL) ^(a)	Infectivity (TU/mL) ^(b)
Integrating vector	1.95E+07	3.51E+08
D64V	4.39E+07	5.12E+08
N120L	4.11E+07	7.61E+08
W235E	1.43E+07	1.84E+08
Q148A	4.82E+07	4.63E+08
K264R	6.03E+07	2.92E+08
K264,266,273R	3.28E+07	9.56E+07
Δ att	2.53E+07	1.25E+09

(a) – Viral particles were quantified by measuring the concentration of p24 in serial dilutions of each vector using a p24 ELISA kit. An average of at least 3 different vector productions is shown.

(b) – To assess infectivity of vector preparations, 293T cells were infected with serial dilutions of vectors expressing eGFP. On the second day post infection, cells were scored for eGFP by flow cytometry, using untransduced cells as negative control. Titres were then calculated (#infected cells x % transduced cells / volume). The values shown are the mean of at least three experiments.

3.3 – Assessment of integration from NILVs *in vitro*

To evaluate the vector expression over time, 293T cells were infected with the same MOI (0.3 – to ensure low vector copy numbers at start) and eGFP expressing cells were then accessed by flow cytometry at several time points.

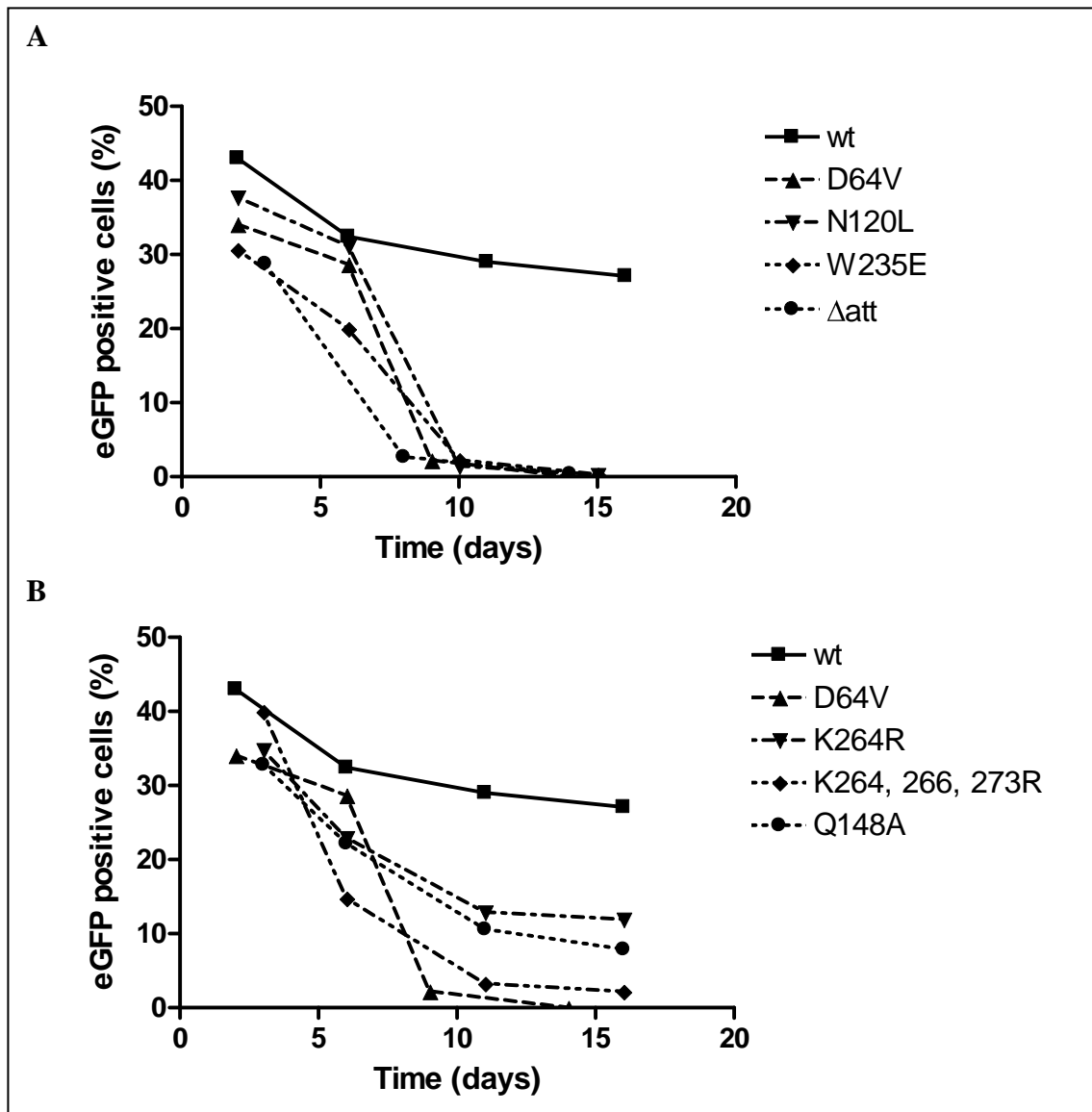


Figure 3.2: Transgene expression profile over time of cells infected with mutant vectors. 293T cells were infected with the same MOI (0.3) with vectors expressing eGFP from the p5 promoter. The percentage of transgene expressing cells was then monitored by flow cytometry analysis over time. For easy comparison, wt and D64V vectors were plotted in all graphs. **A** - cells infected with N120L, W235E and Δatt vectors; **B** - cells infected with Q148A, K264R and K264,266,273R mutant vectors.

Cells infected with the D64V integrase mutant vector express the transgene transiently (Figure 3.2). The kinetics of transgene expression for N120L and W235E mutants is similar (Figure 3.2 A). For the three mutants, the percentage of eGFP expressing cells decreases over time, being indistinguishable from the negative control on the 14th day post infection. The kinetics of expression from the mutants Q148A, K264R and K264,266,273R differ from the D64V vector (Figure 3.2 B). Although the cells lose expression over time, the percentage of GFP positive cells reaches a plateau. Expression was 1.5, 2 and 5 fold lower than expression from the integrating vector for the mutants K264R, Q148A and K264,266,273R, respectively.

Therefore, it was demonstrated that these vectors are able to efficiently transduce cells. However, expression is lost over time indicating that integration of the vector into the host cell genome is blocked.

3.4 – Mutant vector expression characterisation

In order to verify the levels of transgene expression, the mean intensity of fluorescence was measured in cells infected with the same MOI (0.3). This value is correlated with transgene expression. Table 3.2 shows that the mutations had minimal impact on the level of transgene expression, as MFI is no more than 2 fold lower, compared with the integrating vector.

Table 3.2: Expression levels (MFI) of eGFP positive cells infected with vectors

Vectors	MFI^(a)
Integrating vector	1.00
D64V	0.50
N120L	0.51
W235E	0.52
Q148A	0.87
K264R	0.97
K264,266,273R	0.86
Δ att	0.96

(a) – 293T cells were infected with the vectors at MOI 0.3. At the second day post infection, cells were harvested and expression levels were scored by flow cytometry. The mean intensity fluorescence (MFI) obtained was standardised using the MFI of the negative control and then compared with the value obtained for the integrating vector.

The accumulation of viral DNA in cells was also studied. After reverse transcription, the linear viral DNA enters the nucleus where it becomes integrated into the host genome. However, through non-homologous end joining, viral DNA circles containing 2LTR are formed. Viral DNA circles containing 1LTR are also created, probably due to homologous recombination between the viral LTRs. Therefore, to investigate the kinetics of these processes, 293T cells were infected with the integrating vector and the mutant vector D64V. Total DNA was harvested from infected cells at several time points and a Southern Blot was performed to detect the viral DNA forms. When the viral DNA is digested with *EcoR* I (an enzyme that cuts the DNA between the viral LTRs), and the DNA probed with the transgene, it is possible to distinguish the 3 viral DNA forms. Figure 3.3 shows a schematic representation of this process.

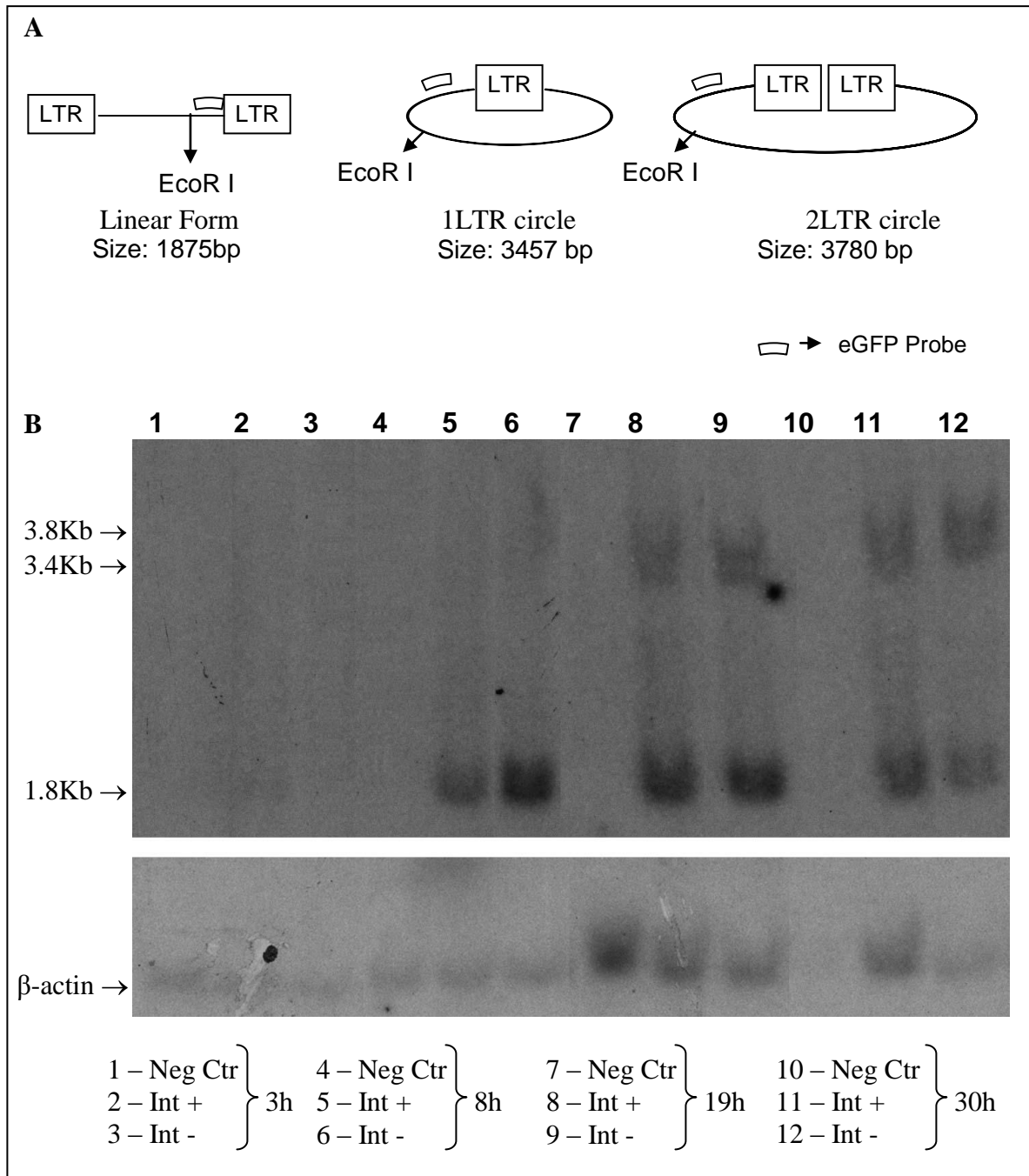


Figure 3.3: Southern Blot of DNA harvested from cells infected with integrating and non-integrating vectors. 293T cells were infected with the integrating vector (int +) and the non-integrating vector D64V (int -). The DNA was extracted at several time points, digested with *EcoR* I and probed for the transgene (eGFP). A – Schematic representation of the viral DNA forms detected with the southern blot. B – Southern blot of 7μg of total genomic DNA probed for eGFP and β-actin.

Figure 3.3 shows that viral DNA can be clearly detected as early as 8h, for both integrating and non-integrating vectors, although a faint band can also be identified at 3h for both vectors. In cells infected with the integrating vector, the concentration of the viral linear DNA form increases from the 8h to the 19h time point, but it decreases slightly afterwards. A similar pattern is also seen for the D64V vector, nonetheless, the decrease of the concentration of the linear form is more pronounced for this vector. Due to the similar size, it is not possible to clearly distinguish between 1LTR and 2LTR circles, so the corresponding bands will be regarded as one, corresponding to the total viral DNA circles. The viral DNA circles were detected 8h post infection in cells infected with both vectors, the concentration increased in the 19h time point and then it remains constant at the 30h time point. It is important to mention that 30h after infection, the relative concentration of DNA circles is higher in cells infected with the D64V vector than with the integrating vector.

The viral DNA forms were then investigated in more detail. 293T cells were infected with the vectors and DNA was harvested two days after infection. Total viral DNA and 2LTR circles were then measured by qPCR. The results (Figure 3.4) show that total viral DNA levels were similar for all vectors tested ($p > 0.05$ at a significance level of 95%, tested by one-way ANOVA). The mean of 2LTR circles levels was also not statistically different (tested by one-way ANOVA, at a significance level of 95%), although a pattern can be detected. NILVs have higher levels of 2LTR circles, on average, than the integrating vector. One exception is the 2LTR DNA circles from the mutants K264R and K264,266,273R, which appear to have levels comparable to the integrating vector.

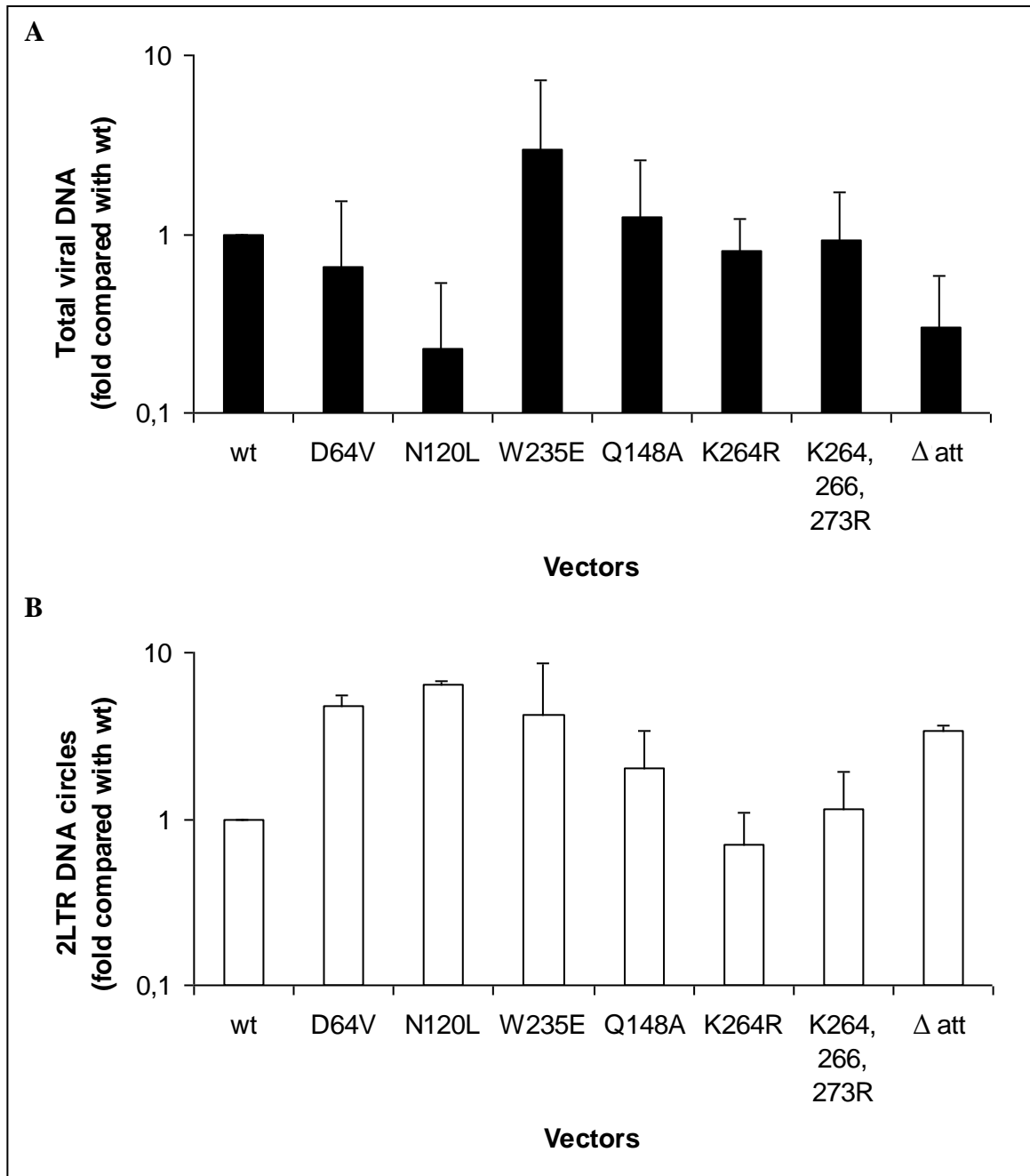


Figure 3.4: Viral DNA quantification in infected cells with lentiviral vectors. 293T cells were infected at the same MOI (0.3) with vectors. On the second day after infection, total DNA was harvested from cells. qPCR was performed to the samples, in parallel with a standard derived from serial dilutions of plasmid DNA containing the sequence amplified by the primers. **A** – Quantification of total viral DNA (amplifying a region between U5 and the packaging signal) **B** – Quantification of 2LTR circles (amplifying the sequence between U5 on the 5’LTR and U3 on the 3’LTR). The values obtained were then normalised so the integrating vector (wt) would have the value 1. Bars represent the obtained mean and error bars the standard error of the mean.

3.5 – Quantification of background integration

The study of transgene expression over time in cells infected with NILVs demonstrated that these vectors are not being integrated into the host genome. Next, to quantify the background level of integration of the mutants, HT1080 cells were infected with vectors expressing the protein neomycin phosphotransferase that confers resistance to an antibiotic. Cells expressing the transgene survive when cultured in medium containing G418, but cells that do not produce this protein will die. Therefore, only cells that have an integrated viral DNA and express the transgene will survive under these conditions over time. G418 resistant cell colonies were quantified 3 weeks after infection. This provides a defined level of background integration that can be compared with the integration efficiency of the integrating vector.

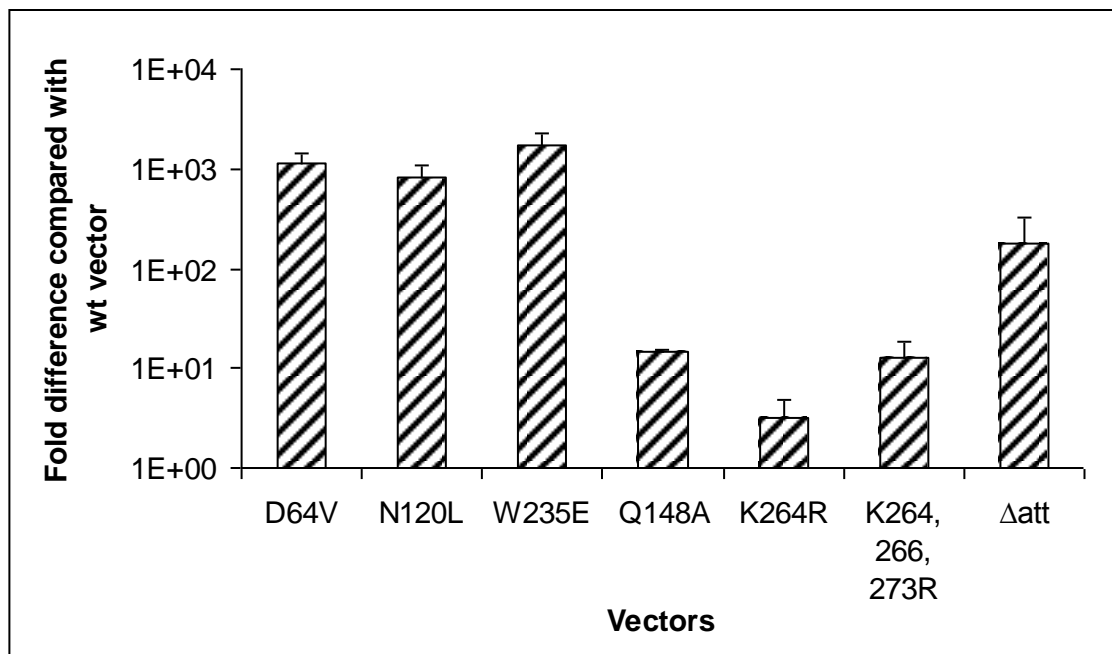


Figure 3.5: Quantification of background integration of mutant vectors. HT1080 cells were infected with serial dilutions of vectors expressing neomycin phosphotransferase. Cells were incubated in medium supplemented with G418 and colonies counted 3 weeks after infection. The number of colonies obtained from cells infected with NILVs was compared with the ones infected with the integrating counterpart. Bars represent the average of at least two different experiments. Error bars represent the standard error of the mean.

As shown in Figure 3.5, the D64V mutant integrates 10^3 fold less efficiently than the integrating vector. Similar levels were found for the mutants N120L and W235E. Integration levels for K264R, Q148A and K264,266,273R mutants were 3, 13 and 14 times lower than the integrating vector, respectively. The Δatt mutation gives a background level of integration of 100 fold compared with the wild type genome vector.

3.6 – Multiple mutant vectors

The vectors harbouring the single mutations D64V, N120L and W235E showed the lowest integration efficiency. Therefore, the next step was to incorporate the mutations in the same integrase and combine the mutant integrase with the mutant att sites in the vector. The mutations were introduced by direct site mutagenesis and confirmed by sequencing the DNA. The production, infectivity and level of expression were then accessed. As described previously, the production was evaluated by measuring the concentration level of p24, the infectivity by measuring the concentration of infectious particles and the expression level by accessing the MFI of infected cells.

Table 3.3: Quantification of viral particles, infectivity and expression levels (MFI) of multiple mutant vectors

Vectors	Viral particles (pg (p24)/mL) ^(a)	Infectivity (TU/mL) ^(b)	MFI ^(c)
D64V + N120L	3.22E+07	2.30E+08	0.49
D64V + W235E	1.20E+07	2.05E+08	0.49
D64V + N120L + W235E	1.91E+07	9.03E+08	0.69
D64V + Δ att	4.53E+06	1.82E+08	0.97
D64V + N120L + W235E + Δ att	2.73E+07	9.35E+07	0.69

All values represent the average of at least three different vector productions.

(a) – Particles were quantified by measuring the p24 protein levels using an ELISA kit.

(b) – 293T cells were infected with serial dilutions of vectors expressing eGFP controlled by the p5 promoter. Two days after infection, cells were harvested and eGFP positive cells were quantified by flow cytometry, using untransduced cells as a negative control. Infectious titre was then calculated.

(c) – eGFP production was assessed two days after infecting 293T cells with lentiviral vectors (using MOI 0.3). In order to measure expression, the mean intensity fluorescence (MFI) was scored by flow cytometry. The MFI obtained was normalised with the negative control and then compared with the value obtained for the integrating vector.

Table 3.3 shows that, as for single mutants, all multiple mutants had similar levels of p24 and infectivity, compared with the integrating vector. The expression from these NILVs was 2 fold lower than the integrating counterpart.

The lack of integration was confirmed by studying the expression of cells infected with these multiple mutant vectors over time. Figure 3.6 demonstrates that 293T cells express the transgene transiently and that, similarly to the D64V mutant, expression is not detected after day 14 after infection.

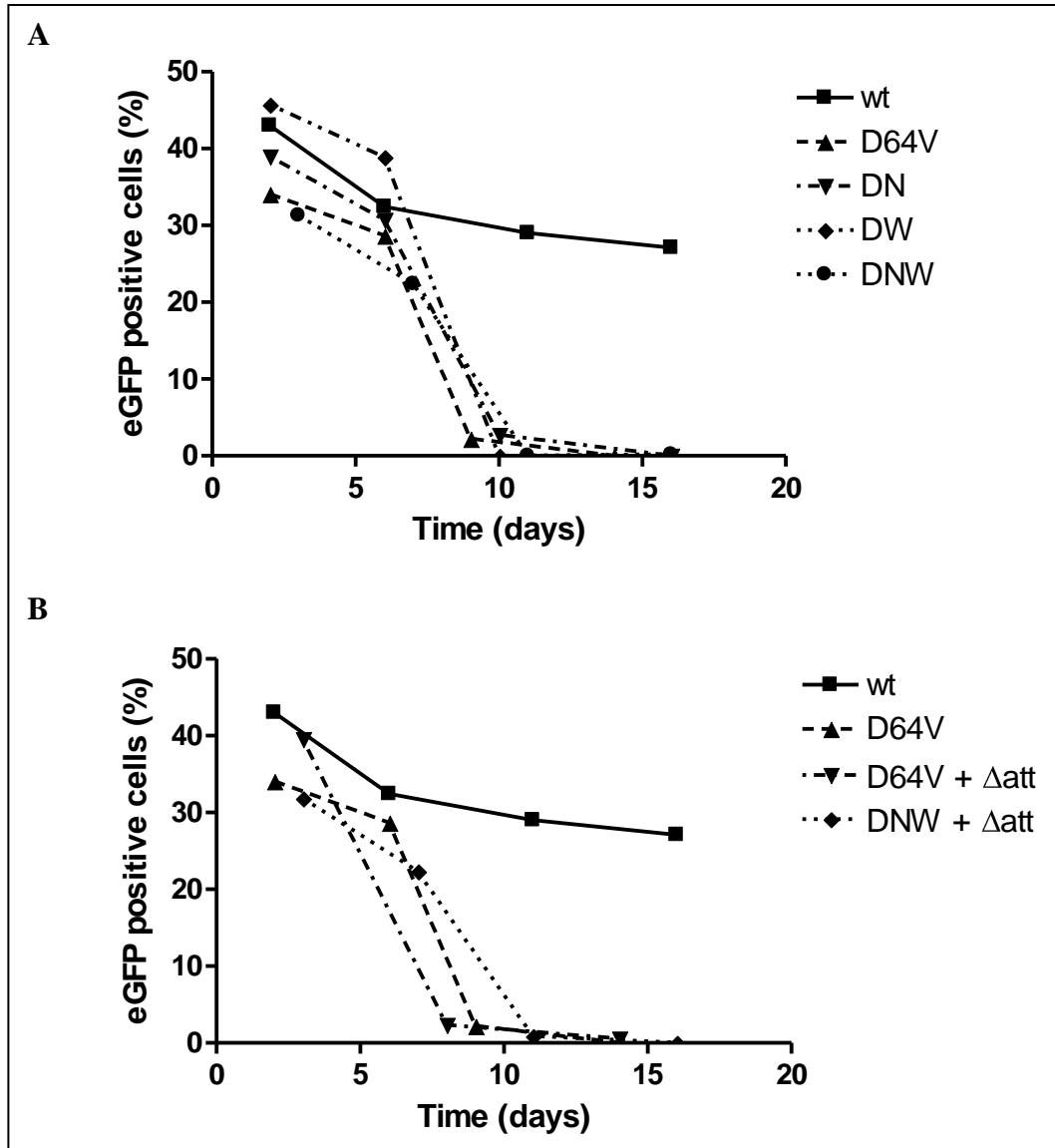


Figure 3.6: Transgene expression profile over time of cells infected with multiple mutant vectors.

293T cells were infected with the same MOI (0.3) with vectors expressing eGFP from the p5 promoter. Green positive cells were monitored by flow cytometry over time. For the ease of comparison, integrating (wt) and D64V vectors were plotted in both graphs. **A** - cells infected with D64V + N120L (DN), D64V + W235E (DW) and D64V + N120L + W235E (DNW); **B** - cells infected with combination of mutant integrases and mutant viral genomes: D64V + Δatt and D64V + N120L + W235E + Δatt (DNW + Δatt).

The background integration was then measured by the antibiotic resistance assay described previously in section 3.5. For the ease of comparison, the results of the levels of integration of the mutant D64V IN mutant and the Δatt vector mutant are also shown in figure 3.7. Interestingly, the combination of mutations D64V+N120L and D64V+W235E in IN did not diminish further the level of background, compared with the mutations alone. The triple mutant D64V+N120L+W235E (DNW) background was 1.4 fold lower than the D64V mutation alone, but not statistically significant. This implies that these mutations do not have a synergistic or additive effect on background integration. The Δatt mutation gives a background level of integration of 100 fold compared with the wt genome vector. In combination with the D64V IN mutant, the level is similar to the one of the D64V mutant alone (10^3 fold). In combination with the DNW mutant, the level was lowered further, being 3×10^3 fold less integrated than the integrating vector but not statistically different from the D64V vector (tested by one-way ANOVA, with a significance level of 95%).

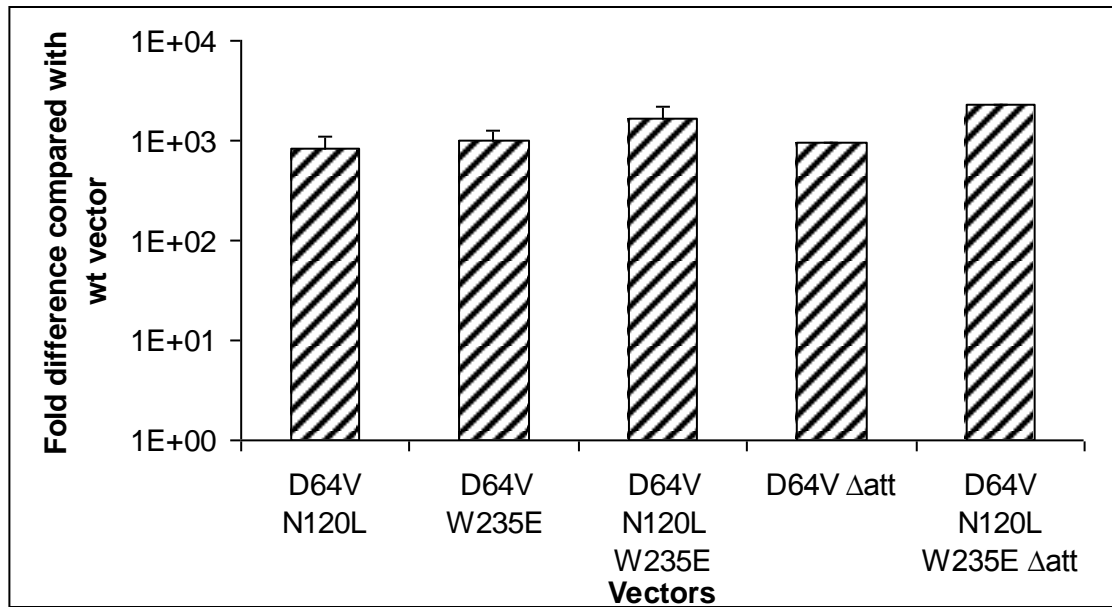


Figure 3.7: Quantification of background integration of multiple mutant vectors. HT1080 cells were infected with serial dilutions of vectors expressing neomycin phosphotransferase. Cells were incubated in medium supplemented with G418 and colonies counted 3 weeks after infection. The number of colonies obtained for cells infected with NILVs was compared with the integrating counterpart. Bars represent the average of at least 2 different experiments and error bars the standard error of the mean.

3.7 – Discussion

Non-integrating lentiviral vectors (NILVs) use the advantages of lentiviral based vectors without the potential problem of insertional mutagenesis. As these vectors do not integrate into the host chromosomes and lack a replication signal, the viral DNA will be diluted out during cell division. Therefore, NILVs will be particularly useful, for example, in post mitotic tissue gene therapy.

NILVs were designed by mutating integrase, the key viral enzyme in the integration process and the *cis*-acting sequences (att sites) in the vector DNA where IN binds. It was demonstrated that the generated mutants could be produced at similar levels as the integrating vector and that infectivity is also similar (Table 2), implying that the mutations introduced did not affect vector production (assembly and budding). As shown in Figure 3.2, expression diminishes over time for all mutant vectors generated. This implies that viral DNA is degraded or diluted out with cell division indicating that these vectors fail to integrate the viral DNA into the host chromosomes.

The viral DNA accumulation in cells infected with integrating and non-integrating vectors was then studied. For the integrating and non-integrating vectors under analysis, the viral DNA can be detected in infected cells as soon as 3h after infection. The concentration of the linear form of the viral DNA increases until 19h but then decreases in the last time point. Most importantly, this effect is more pronounced in the D64V mutant vector. Regarding the viral DNA circles, it is possible to detect these forms 8h after infection. The concentration increases afterwards, in the 19h time point and is maintained in the last time point. Also, the relative concentration of DNA circles is higher for the D64V mutant compared with the integrating counterpart vector. These results imply that reverse transcription occurs early after infection and possibly the viral DNA detected at the 3h time point is still in the cell cytoplasm. LTR circles are an indication of nuclear viral DNA, as recombination only occurs in the nucleus, so it is possible to infer that nuclear entry has happened between 3 and 8h after infection. The decrease of linear DNA concentration from 19h to 30h can be explained by assuming that this is the substrate for integration and viral DNA circles formation. Viral DNA forms were quantified 48h after infection by qPCR, and the results show that the total viral DNA concentration of NILVs is

similar to the integrating counterpart vector, implying that reverse transcription is not impaired. Also, the concentration of viral 2LTR circles was even higher for NILVs, indicating that nuclear entry is also not blocked. For the NILVs, integration is impaired, so the linear viral DNA may be converted to DNA circles at a rate higher than the integrating vector, explaining the higher concentration of circle forms for the NILVs.

The integration background was then quantified. Single mutations in residues D64V, N120 and W235 of Integrase affected integration, lowering it to 10^3 fold less than the integrating vector. Similarly, Leavitt and colleagues show that a vector harbouring the N120L mutation in IN integrates 10^{-3} and that vectors containing the mutations D64V and W235E integrate at 10^4 fold lower than the integrating vector (Leavitt *et al.*, 1996). The mutations K264R, Q148A and K264,266,273R were less effective. Integration was lowered 3, 13 and 14 fold, respectively when compared with the integrating counterpart. This is correlated with the plateau of eGFP expression seen in 293T transduced with these mutants (Figure 3.2). The mutation of the att sites diminishes integration by 100 fold. This result is similar to the one obtain by Masuda and colleagues (Masuda *et al.*, 1998).

NILVs were then further developed to incorporate double and triple mutations in integrase and combined with the att sites mutant vector. Results show that integration could not be reduced much further. Although the quadruple mutant DNW + Δ att integrated at a 3×10^3 rate less than the integrating counterpart vector, this value was not statistically significant. Nightingale and colleagues recently publish similar results, combining the D64V mutation and the Δ att mutations (Nightingale *et al.*, 2006). These results suggest that the combination of mutations is neither additive nor synergetic.

Linear double stranded DNA can be integrated non-specifically in the host genome. When there is a double strand break in the host genome, the DNA is repaired either by homologous recombination or non-homologous end joining. The last pathway involves the ligation of two double stranded DNA molecules, not requiring any homology between them (reviewed in van Gent *et al.*, 2001). Using this mechanism, the cells may then integrate the viral linear DNA produced upon reverse transcription or viral DNA circles that were broken at double stranded break points in the genome. Therefore, it is important to consider that the cells used in this assay may not be the best model because they are

very recombinogenic – HT1080 have a random integration rate of 4×10^{-4} (Yanez and Porter, 2002). Possibly, primary cells would be more suitable models for this experiment.

The occurrence of mutations during or before vector production in packaging constructs that encodes an integration deficient enzyme may result in vectors carrying active INs. Hence, it is important to emphasise that the triple mutant packaging plasmid (DNW) lowers the risk of reversion back to an integrating phenotype. Furthermore, the mutation of the att sites within the viral genome reduces the risk of rescue or integration of the vector if a patient treated with NILVs is infected with a wild type HIV. Therefore, the multiple mutations in these NILVs represent a safer vector for the use in gene therapy.

Chapter Four

Application of Non-integrating Lentiviral Vectors

4 – Application of Non-Integrating Lentiviral Vectors

4.1 – Introduction

In the previous chapter, it was shown that non-integrating lentiviral vectors (NILVs) are produced in a similar manner to their integrating counterpart. It was also demonstrated that NILVs are infective and the transgene is efficiently expressed in the target cells. However, the viral DNA of the developed NILVs does not contain a replication origin. Therefore, expression decays when these vectors are used to transduce dividing cells, since the viral DNA does not integrate into the host genome, as depicted in Figures 3.2 and 3.6. The application of these vectors will then be studied in this chapter.

Transient expression of a transgene can have several applications, such as the introduction of a signal to stimulate cell proliferation or to provide a differentiation signal. Another application of the developed NILVs would be suicide gene therapy. The application of these NILVs in gene therapy of monogenic diseases may require prolonged expression of the transgene. This can only be achieved in non-dividing cells that do not degrade the viral episomal DNA.

It was previously shown that NILVs are able to sustain expression in non-dividing tissues, such as brain and retina for long periods of time. Prolonged transgene expression was seen in cortical primary neurons infected with NILVs *in vitro* and in brain after a stereotactical injection. The well established model of retinal dystrophies Rpe65^{rd12/rd12} was also used to show that NILVs are able to efficiently transduce retinal cells and rescue Rpe65 deficiency (Philippe *et al.*, 2006; Yanez-Munoz *et al.*, 2006). The aim of the work presented in this chapter is to extend this research to study the application of NILVs in muscle. Muscle cells do not divide unless there is a lesion and thus, they provide a good target for stable expression from NILVs.

4.2 – Test of NILVs in non-dividing cells

C2C12 cells are a good model to study muscle cells *in vitro*. These cells are easily differentiated into non-dividing fibres by serum starvation. The kinetic of expression from NILVs was studied in dividing and non-dividing C2C12 cells. Dividing and differentiated cells were infected with similar concentrations of vectors expressing eGFP from the p5 promoter. The GFP positive cell percentage was then quantified over time by flow cytometry.

As shown in Figure 4.1, all the vectors tested are able to infect muscle cells. The expression in dividing C2C12 cells with NILVs has similar kinetics as in 293T cells: for all vectors studied, expression is transient and not detected after 12 days. In contrast, transgene expression is stably maintained in non-dividing cells for up to 12 days. In the differentiated C2C12 cells, the percentage of transgene expressing cells was higher at the second time point, five days after infection than at the first time point. This may indicate that expression in these cells is delayed compared to 293T cells, where the highest percentage of transgene expressing cells was detected two days after transduction. A decrease of the percentage of eGFP positive cells is then seen at the last time point for most of the vectors, including the integrating vector. This may be attributed to cell death due to cytotoxicity of the transgene.

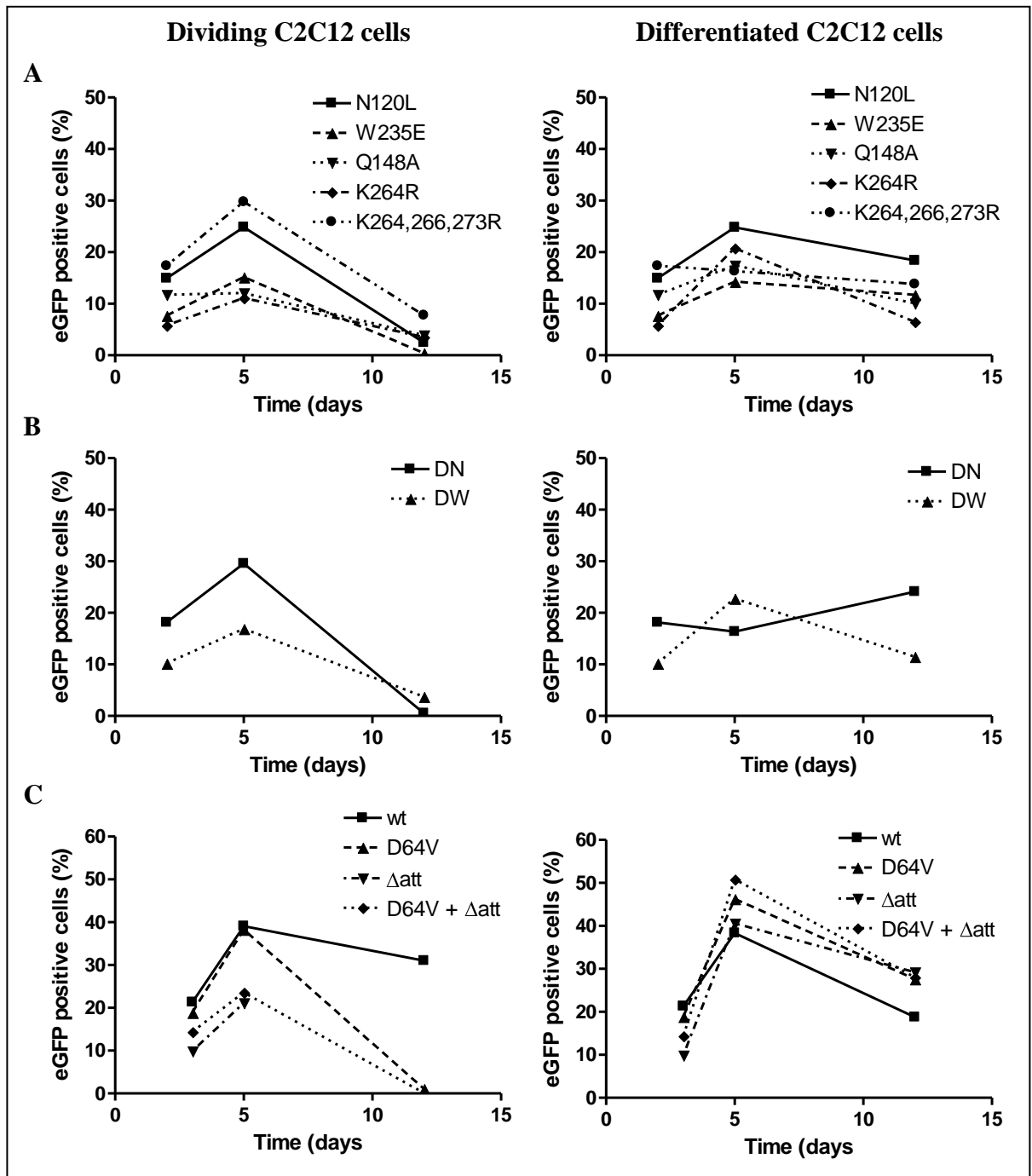


Figure 4.1: Expression profile in dividing and non-dividing muscle cells over time. Dividing and non-dividing C2C12 cells were infected with similar MOI (10) with vectors expressing eGFP from the p5 promoter. At different time points, transgene expression was assessed by flow cytometry. Graphs on the left hand side represent dividing cells and on the right, differentiated cells. **A** - Cells infected with vectors harbouring IN mutations N120L, W235E, Q148A, K264R and K264,266,273R; **B** - Cells infected with combination of IN mutations D64V + N120L and D64V + W235E; **C** - Cells infected with integrating vector, the catalytic mutant D64V, the viral genome mutant Δatt and the combination of D64V + Δatt .

4.3 – Viral DNA forms in non-dividing cells

It was demonstrated that NILVs are able to infect muscle cells and the transgene was detected for the whole life span of the cells. For prolonged transgene expression, the viral episomal DNA has to be maintained in the cell. Consequently, it is important to study the kinetic of viral circles in non-dividing cells. Viral DNA forms in non-dividing cells were then assessed at three different time points (3, 6 and 12 days post infection) by qPCR.

Figure 4.2 shows that total viral DNA concentration was similar for all of the vectors tested. Importantly, viral DNA was not degraded over time, as the levels were maintained to the last time point (12 days post infection). 2LTR circles were also quantified. It was found that 2LTR circle levels for NILVs were higher compared with the integrating vector. Also, 2LTR circles accumulated in cells infected by NILVs. At the 12th day post infection, the mutant vectors Δatt , D64V and D64V + Δatt accumulated 4, 5 and 7 fold more 2LTR circles than the integrating vector.

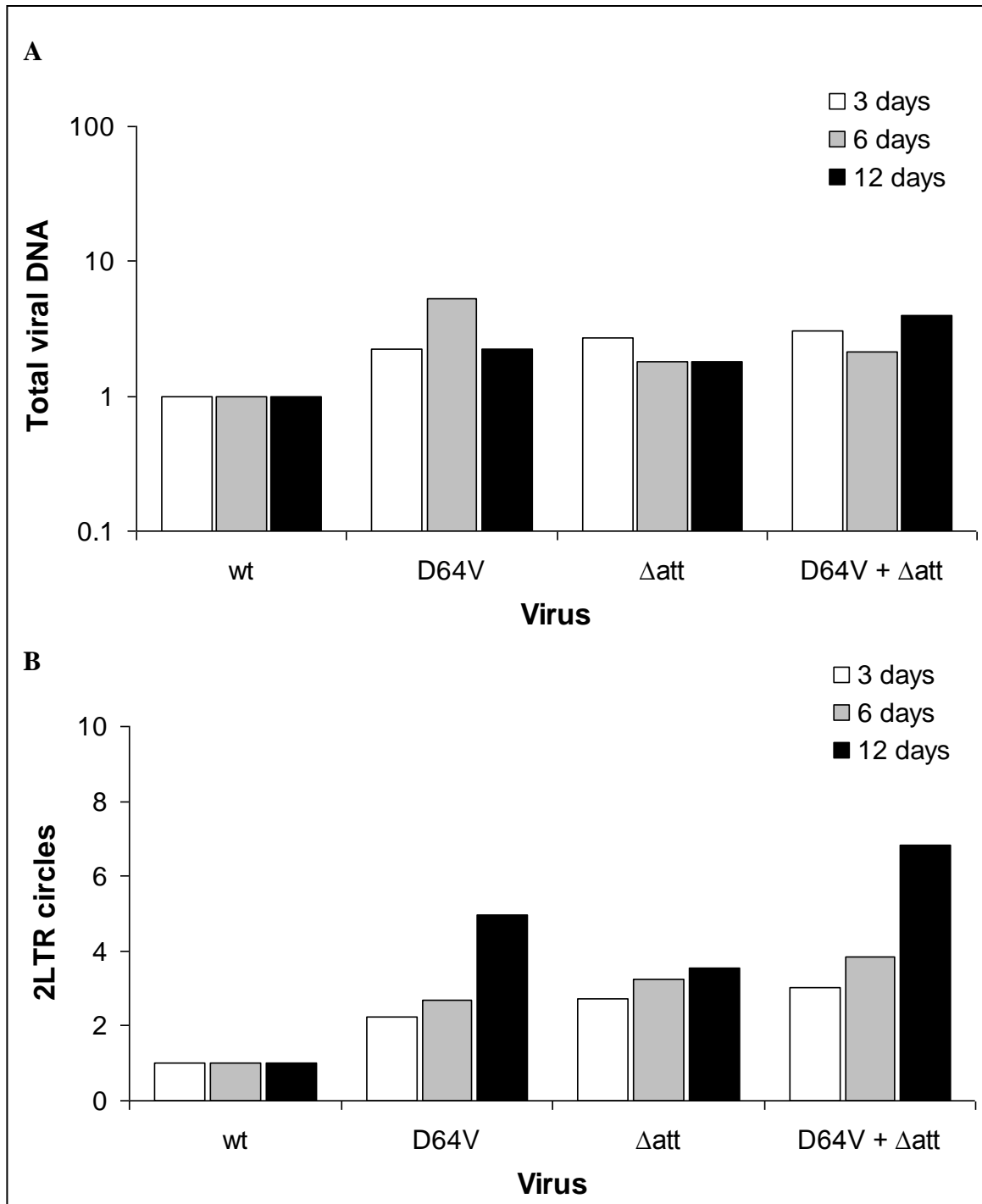


Figure 4.2: Viral DNA quantification in non dividing cells.

C2C12 cells were infected at the same MOI (10) with vectors and growth arrested in medium containing 2% horse serum. This induces differentiation of these cells into myotubes. On the 3rd, 6th and 12th day after infection, total DNA was harvested from cells. Total viral DNA (**A**) and 2LTR circles (**B**) were quantified by qPCR using serial dilutions of plasmid DNA containing the sequence amplified by the primers as a standard. The values obtained were then compared to the integrating vector (wt).

4.4 – Expression of NILVs *in vivo* in muscle

It was demonstrated that NILVs are able to infect muscle cells *in vitro* and that expression in differentiated cells is stable, without loss or degradation of the viral DNA. The vectors were then tested *in vivo*. The model chosen were neonatal mice, as these mice do not develop an immune response against injected antigens. The expression required *in vivo* will be higher than *in vitro* so, it was chosen to use the strong viral-derived SFFV promoter, instead of the p5 promoter, which is a weaker promoter. To allow an easy detection of protein expression from the vectors, eGFP was used as a transgene.

The transfer vector was constructed by cloning the SFFV promoter in the p5 promoter place. Lentiviral vectors were then produced. The infectious titre was measured *in vitro* in 293T cells and vector particles were assessed by quantifying the p24 protein concentration in vector preparations. MF1 neonatal mice were then injected with 5µL of each vector intramuscularly in the tibialis anterior muscle. The mice were sacrificed at several time points and eGFP positive tissue was visualised in a stereomicroscope. Samples from the green muscle area were then taken for further analysis.

All mice injected with integrating and non-integrating vectors were eGFP positive (Figure 4.3). Most importantly, the transgene was detected in all the samples at all time points, even 8 months after injecting the vectors. This results indicates stable and prolonged transgene expression from both integrating and NILVs.

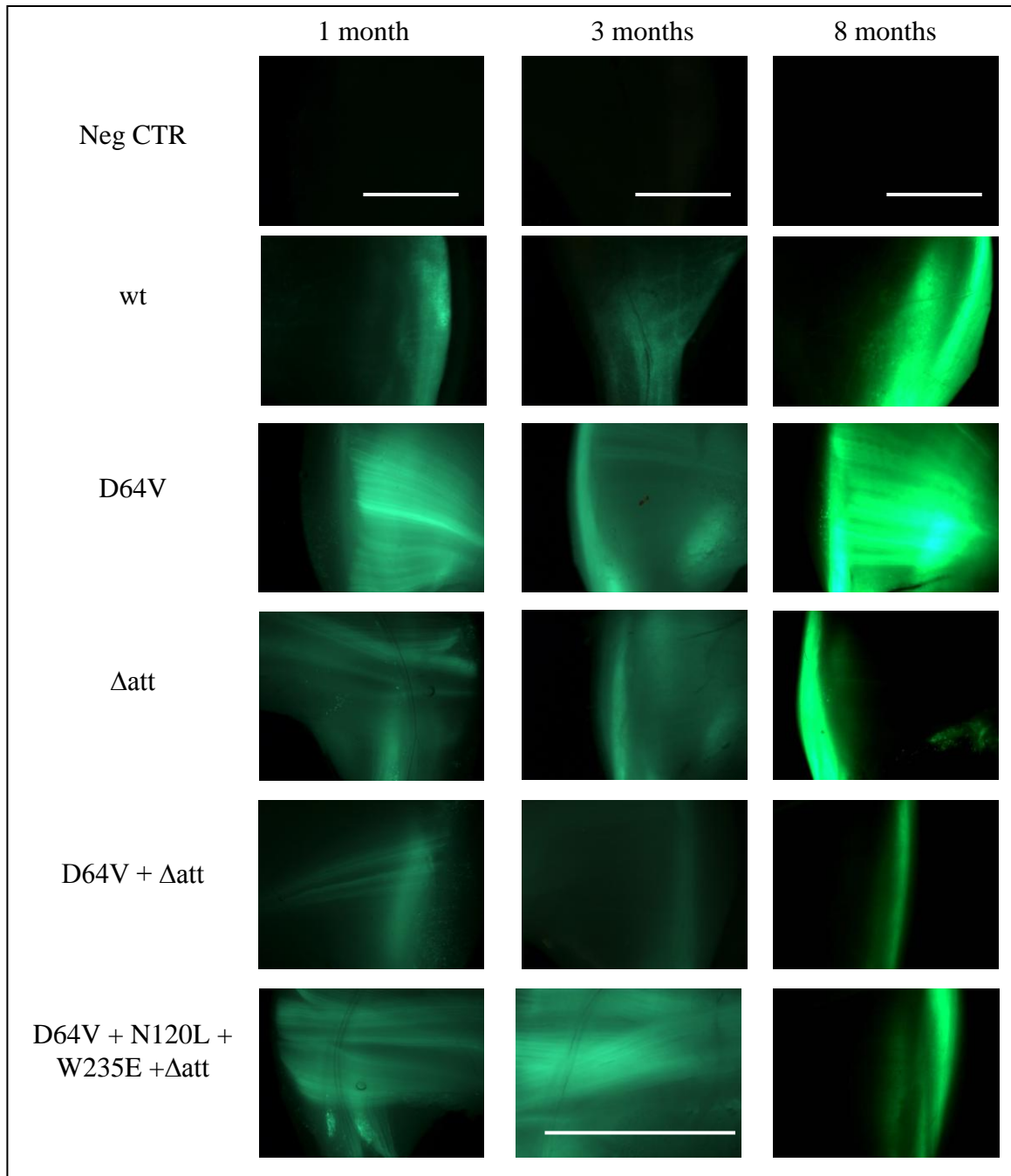


Figure 4.3: Transgene expression from muscle *in vivo*.

MF1 neonatal mice were injected intramuscularly with 5 μ L of vectors (integrating vector (wt), D64V, Δ att, D64V + Δ att and D64V + N120L + W235E + Δ att) expressing eGFP from the SFFV promoter. Vector titres (TU/mL) were: wt – 1.4x10⁹, D64V – 9.9x10⁹, Δ att – 6.4x10⁹, D64V + Δ att – 3.8x10⁹ and D64V + N120L + W235E + Δ att – 5.3x10⁹. Mice were sacrificed one, three and eight months after injection and transgene expression was visualised in whole legs under a fluorescence microscope. Pictures were taken using the same parameters (gain = 1, exposure time = 0.5s, amplification = 1x). One representative picture out of four replicates is shown. scale bars = 4mm.

In order to confirm that muscle cells were indeed expressing the transgene, tissue samples taken from mice sacrificed one month after injection were sectioned and eGFP was visualised in cells by immunohistochemical staining. Figure 4.4 demonstrates that muscle fibres were eGFP positive for all mice injected with integrating and non-integrating vectors.

It is also important to quantify the levels of expression. Protein was extracted from samples taken from the green part of the muscle and eGFP was then quantified by ELISA.

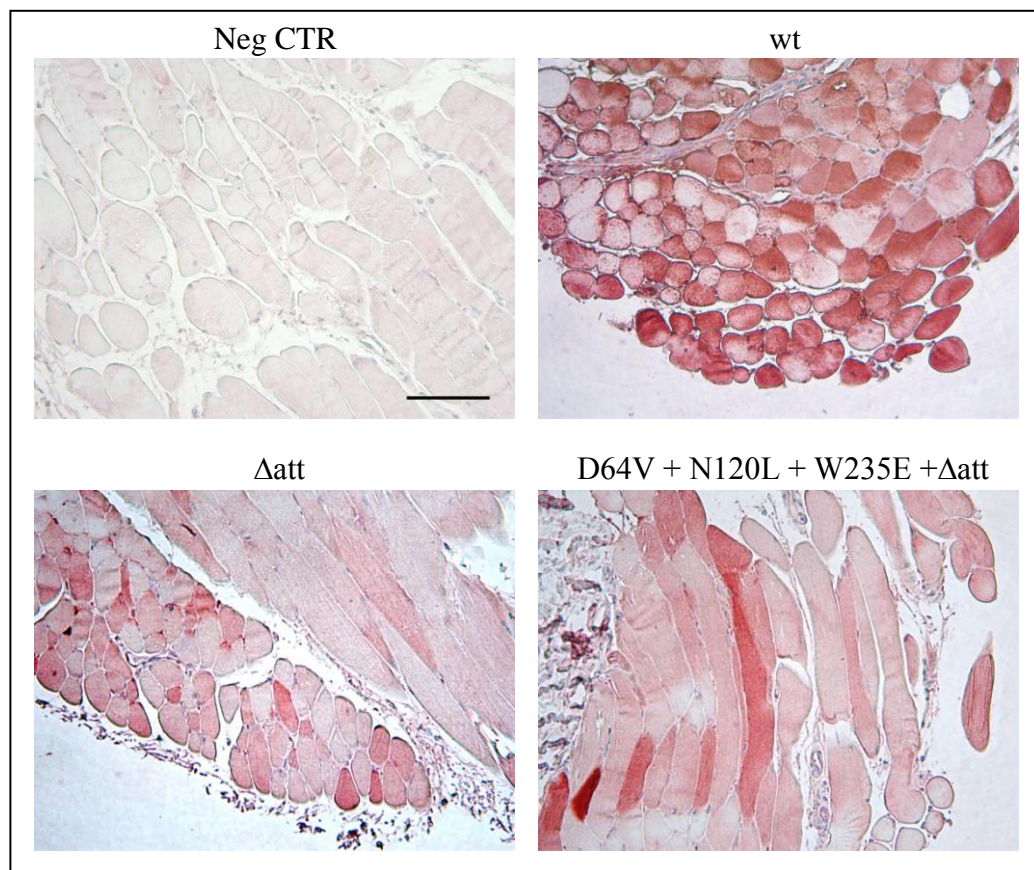


Figure 4.4: Identification of eGFP expressing muscle cells *in vivo* by immunohistochemistry.

Vectors were injected intramuscularly in one day old MF1 mice using 5 μ L of each vector per mouse. Vector titres (TU/mL) were: wt – 1.4x10⁹, D64V – 9.9x10⁹, Δ att – 6.4x10⁹, D64V + Δ att – 3.8x10⁹ and D64V + N120L + W235E + Δ att – 5.3x10⁹. Tibialis anterior muscle sections were taken one month after injection. Intracellular eGFP was stained with 3-amino-9-ethylcarbazole (red), scale bar = 100 μ m. One representative photograph of at least five replicates is shown.

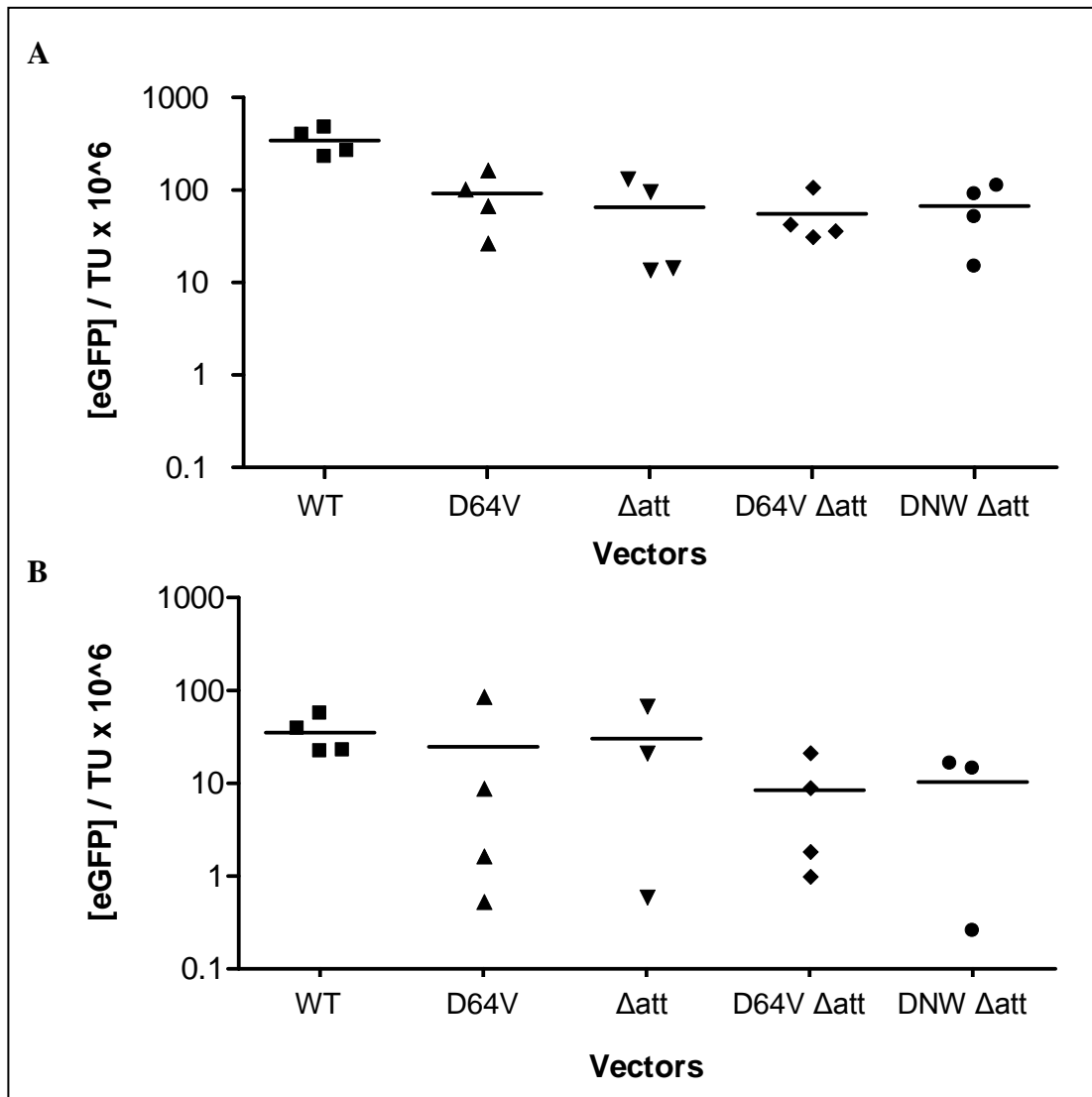


Figure 4.5: Quantification of eGFP protein from *in vivo* samples. Vectors (5 μ L) were injected intramuscularly in neonatal MF1 mice. Vector titres (TU/mL) were: wt – 1.4×10^9 , D64V – 9.9×10^9 , Δ att – 6.4×10^9 , D64V + Δ att – 3.8×10^9 and D64V + N120L + W235E + Δ att (DNW Δ att) – 5.3×10^9 . One and three months after injection, samples from the eGFP positive muscle were collected. Total protein was extracted by grinding 20mg of tissue in lysis buffer. Total protein was quantified using a Bradford based assay. eGFP was quantified by ELISA in parallel with a commercial eGFP standard. eGFP concentration ($\text{pg}_{\text{eGFP}}/\text{mg}_{\text{(total protein)}}$) was then divided by the injected transducing units (TU). Each point represents the average obtained for one mouse. **A** – one month samples; **B** – three months samples. Statistical analysis was performed using one way ANOVA to detect statistical differences between means and a Tukey test in order to detect which group of vectors was statistically different.

High levels of transgene expression were detected (0.01- 0.1% of total protein) at both 1 and 3 months post injection in all samples. At the first time point, the expression from the integrating vector was 5 fold higher than in the mutants, but levels between mutants were not statistically different ($p>0.05$, tested by one-way ANOVA). At 3 months post injection, all vectors expressed similar levels of eGFP ($p>0.05$, tested by one-way ANOVA). It is also important to mention that from 1 to 3 months post injection, a 10 fold reduction of the eGFP concentration for both integrating and non-integrating vectors was seen. Also, the concentration of eGFP in mice injected with NILVs is more scattered compared with samples from mice injected with the integrating counterpart.

The viral 2LTR circle concentration was also studied for both time points. Total DNA was harvested from samples and quantified by qPCR. From Figure 4.6 it can be seen that it was possible to detect 2LTR circles in all samples. It also shows that the levels are higher for NILVs compared with the integrating vector, although not statistically significant.

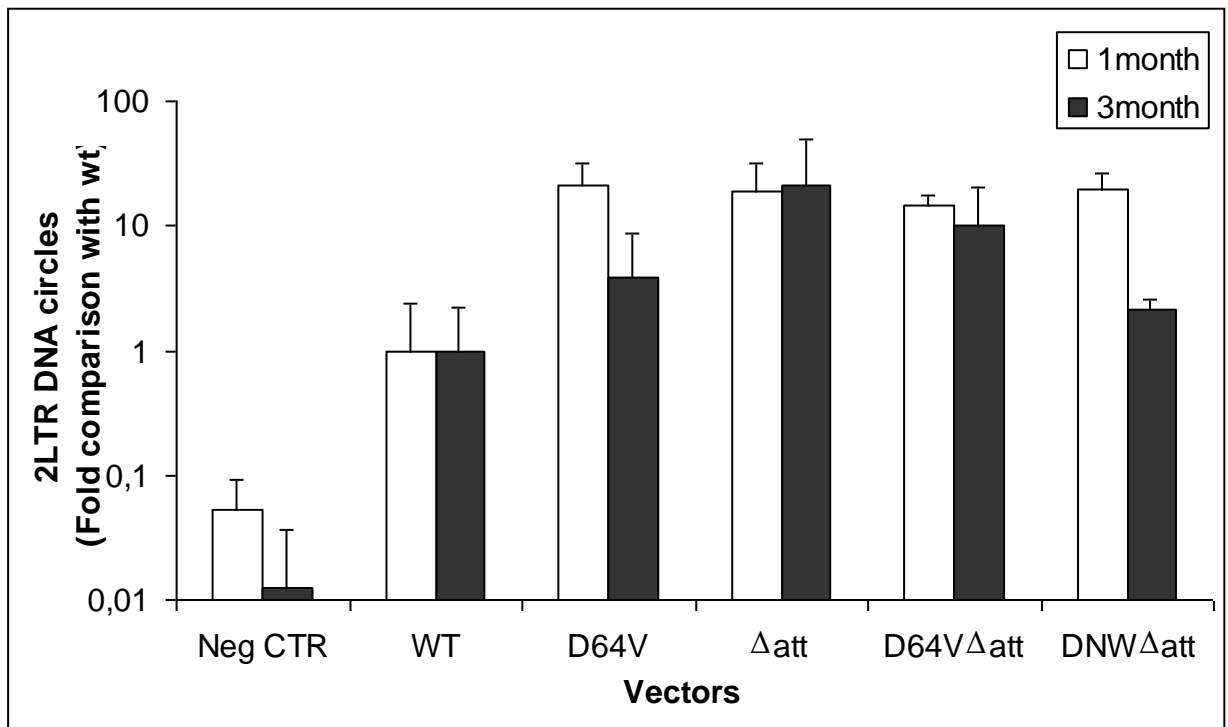


Figure 4.6: Quantification of viral 2LTR circles in *in vivo* samples.

Total DNA was extracted from mice muscle samples one and three months after injection. 2LTR circles were quantified by qPCR along plasmid standards. The values obtained were then compared with the integrating vector (wt). Bars represent the average obtain and error bars the standard error of the mean.

4.5 – Discussion

NILVs are able to efficiently transduce retina and brain, mediating prolonged transgene expression (Philippe *et al.*, 2006; Yanez-Munoz *et al.*, 2006). This work was extended by studying the infection of muscle cells by NILVs. C2C12 muscle cells were used as a model *in vitro*. It was demonstrated that NILV are able to transduce these cells (Figure 4.1). In dividing C2C12 cells, eGFP expression diminishes over time with comparable kinetics to transduced 293T cells. In contrast, when cells were allowed to differentiate, hence stopping cell division, transgene expression was maintained through the course of the experiment (Figure 4.1).

Viral DNA was quantified at several time points in non-dividing C2C12 cells (Figure 4.2). For all the mutants tested, values of total viral DNA were similar. Interestingly, 2LTR levels in NILV were higher and increasing over time, compared with the integrating vector, even at the later time point (12 days). Saenz and colleagues show similar results in growth arrested cells infected with catalytic IN mutants, where episomal DNA is maintained (Saenz *et al.*, 2004). Taken together, these results imply that episomal DNA is not degraded in non-dividing cells, and that they are templates for gene expression. Also, the results suggest that linear viral DNA that does not integrate undergoes circularisation. It should be noticed that the ratio between 1LTR and 2LTR circles is 9:1 (Butler *et al.*, 2001). Therefore, the quantification of 2LTR circles may be an underestimate of the total episomal viral circles.

Selected vectors were then injected intramuscularly in neonatal mice. One, three and eight months after injection, mice were sacrificed and transgene production was detected by microscopy. All mice injected with lentiviral vectors were eGFP positive at all time points. Upon quantification, it was shown that eGFP was expressed at high levels one and three months after injection. At the first time point, the integrating vector expressed more transgene than injected NILVs. This is correlated with the MFI measured in 293T cells *in vitro*. Most importantly, expression from D64V, Δatt , D64V + Δatt and the quadruple mutant D64V + N120L + W235E + Δatt were similar. At three months, expression decreased for all vectors. This could be due to promoter silencing, cytotoxicity of the transgene or a host immune response. It has been proposed that the SFFV promoter

may be silenced in mice bone marrow cells 3 months after infection (Zhang *et al.*, 2007a), which is in accordance with the results obtained in muscle cells. Therefore, this promoter may not be the best choice for prolonged transgene expression from muscle tissues. Interestingly, the reduction for the integrating vector was higher (8 fold) than for the NILVs (3 fold on average). It can be hypothesised that expression and possibly promoter silencing of integrated viral DNA is not similar for all the viral DNA forms. Also, it is possible that the higher level of transgene produced may induce an immune response against the cells with integrated vectors.

This work demonstrates that NILVs are able to infect muscle cells and that viral DNA is not degraded, allowing expression to be maintained for very long periods of time in non-dividing cells, both *in vitro* and *in vivo* (up to 8 months). NILVs may therefore be used for gene therapy of muscle diseases or when it is necessary to express a transgene from muscle cells.

Chapter Five

Correction of a Disease Model

with Non-integrating Lentiviral Vectors

5 –Correction of a Disease Model with Non-Integrating Lentiviral Vectors

5.1 – Introduction

In the previous chapters, it was demonstrated that the developed non-integrating lentiviral vectors (NILVs) are able to sustain transgene expression from non-dividing muscle cells, both *in vitro* and *in vivo*. The vectors are then suitable to be used in a gene therapy context, by assessing disease correction of an animal disease model.

There are muscle disease models that would be useful to study the application of NILVs in gene therapy, such as Duchenne muscular dystrophy. This is a well studied x-linked recessive genetic disorder. It is characterised by the lack of functional dystrophin in muscle cells, leading to progressive weakness, loss of skeletal muscle and cardiomyopathy. It was demonstrated that expression of microdystrophin in the mdx mice disease model prevents the development of the skeletal muscle pathology (Gregorevic *et al.*, 2006). Also, a lentiviral vector was used to deliver a truncated form of dystrophin. The minidystrophin gene was expressed for up to 6 months but the levels of expression were relatively low (Li *et al.*, 2005). NILVs seem to have a lower expression level compared with the integrating vector indicating that this model may not be adequate to the application of NILVs.

It is known that muscle can be genetically modified in order to produce proteins (Arruda *et al.*, 2001; Herzog *et al.*, 1997; Song *et al.*, 1998). There is a concern that a protein not usually produced by muscle fibres will not be active due to a lack of or incorrect posttranslational modifications. However, Arruda and colleagues have shown that these modifications in Factor IX (FIX) produced by myotubes or in liver (natural producer) are similar and specific activity is not affected. In order to use the muscle potential as a producer of plasma proteins, other disease models where non-integrating vectors can be used are haemophilia B and α 1-antitrypsin (AAT) deficiency. Both enzymes are produced naturally in liver cells and then secreted into plasma.

AAT's function is to protect lung alveolar tissues from destruction by neutrophil elastase (Primhak and Tanner, 2001). In patients with this monogenetic deficiency, AAT is

not found at normal levels in the blood. Although gene delivery into the liver is the obvious rationale, it has been demonstrated that secretion of the transgene by muscle transduced by adeno-associated virus can yield therapeutic levels of protein (Lu *et al.*, 2006). Nevertheless, direct delivery into the lung constitutes a more promising approach because the expression levels required for a therapeutic benefit may be lower (Stecenko and Brigham, 2003).

Haemophilia B is a bleeding disorder resulting from the mutation of the FIX gene. This protein is required for the generation of a fibrin clot and the clinical severity of the disease correlates well with the circulating levels of FIX. Low plasma concentrations of 1-5% of normal levels are sufficient for disease amelioration (Ljung, 1998). It has been previously shown that intramuscular injections in immunodeficient mice with adeno-associated vectors (AAV) expressing FIX generate 5-7% of normal FIX levels in plasma (Herzog *et al.*, 1997). This work was extended to study the expression levels achieved by the vectors in a canine model. In dogs, the levels of normal FIX levels reached 1.4% (Herzog *et al.*, 1999). These vectors were then used in a clinical trial. Although gene transfer was demonstrated in the trial, expression of FIX was low (Manno *et al.*, 2003). Also, in another clinical trial targeting liver using similar vectors, expression was transient due to an immune response against the AAV capsid (Manno *et al.*, 2006; Mingozi *et al.*, 2007). Therefore, it will be necessary to use vectors less immunogenic than the AAV vectors, such as lentiviral vectors.

The correction of haemophilia B in a disease model may be an ideal candidate to test the use of the developed NILVs in gene therapy: the vectors have low immunogenicity and the levels of expressed protein necessary to reach disease correction are relatively low.

5.2 – Production of vectors expressing FIX

It was demonstrated that NILVs are able to infect muscle cells with high levels of expression and that the transgene is detected for up to 8 months. The NILVs will then be suitable to deliver a vector for expression of FIX from muscle cells.

The transgene used in the previous experiments shown in chapter 3 and 4 was eGFP, which allows the rapid detection and quantification in expressing cells by flow

cytometry. However, for the application of the vectors to the haemophilia B model, the vector was modified by replacing eGFP with the human FIX gene, being expressed from the SFFV promoter. The human FIX was cloned into the vector with and without the att site mutation and sequenced. To insure that the vector expressed FIX and that the transgene was detectable, the integrating vector was produced and tested *in vitro* in 293T and dividing and differentiated C2C12 cells. These cells were infected with the vector and three days after infection the supernatant was removed. The FIX concentration in the supernatant was then measured by ELISA.

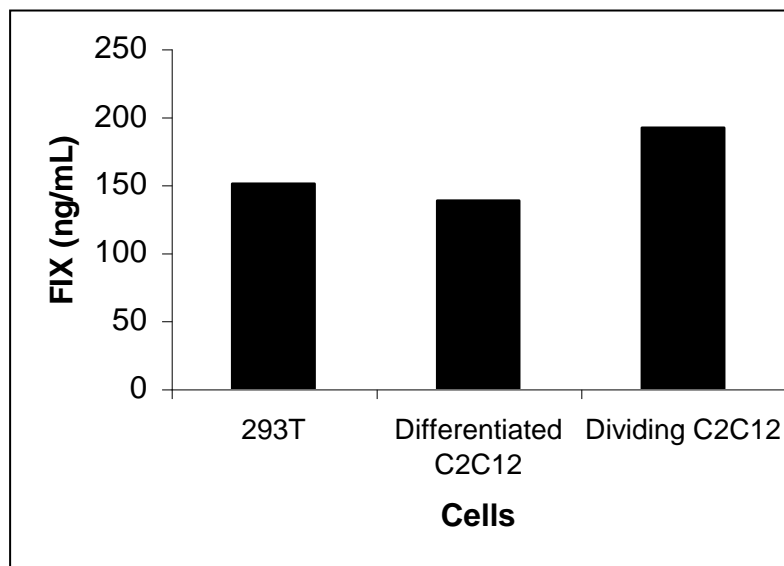


Figure 5.1: Expression of hFIX *in vitro*.

293T cells were infected with 4.5ng_(RT) of an integrating lentiviral vector expressing hFIX. Dividing and differentiated C2C12 cells were infected with 45ng_(RT) of the same vector. 3 days after infection, the supernatant was collected and hFIX was quantified by ELISA.

Figure 5.1 shows that the transgene can be detected in both cell lines. This indicates that the vector is efficiently expressed, secreted and able to be quantified by ELISA. Also, this data shows that muscle cells are able to express and secrete the transgene.

5.3 – Vector expression of hFIX *in vivo*

There has not been any report of expression of FIX by muscle cells infected with lentivectors *in vivo*. Hence, it was decided to perform a pilot experiment *in vivo* where immunocompetent MF1 mice would be injected intramuscularly with the integrating vector expressing hFIX from the SFFV promoter. Neonatal mice were injected in the tibialis anterior muscle of one leg with 5 μ L of concentrated vector (35.1ng_(RT)/ μ L). At several time points, a blood sample was taken from mice and plasma hFIX was quantified by ELISA.

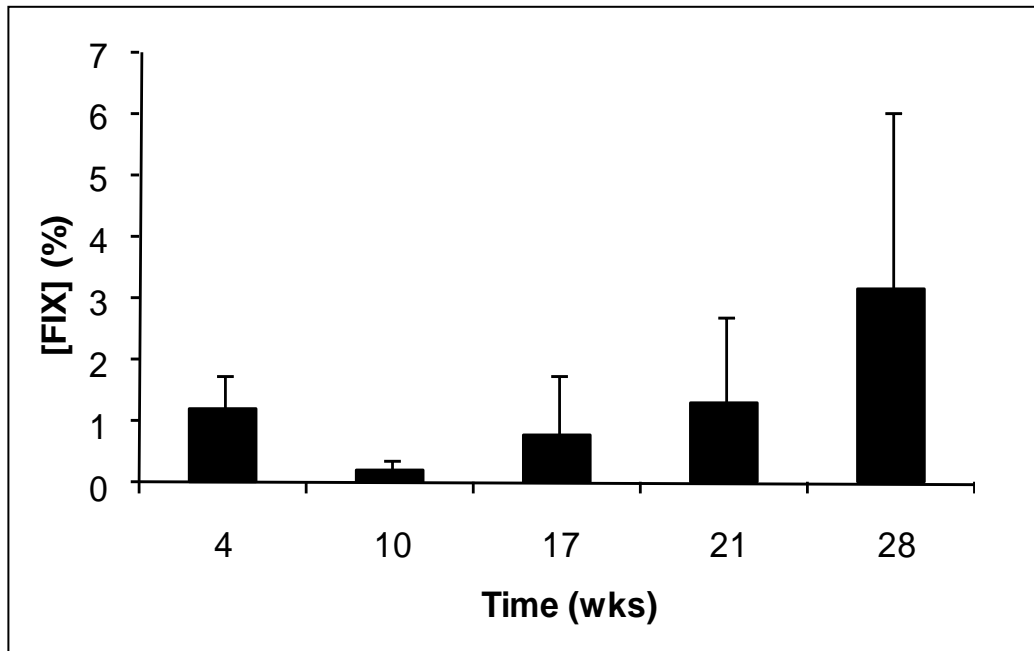


Figure 5.2: hFIX expression *in vivo*.

Neonatal MF1 mice were injected in the tibialis anterior muscle with 5 μ L of the integrating vector expressing hFIX from the SFFV promoter (1×10^8 pg_(p24)/mL). Blood samples were taken from mice at several time points, and hFIX was quantified in plasma by ELISA. Bars represent the mean and error bars the standard error of the mean obtained in three animals.

As shown in Figure 5.2, hFIX was detected in the plasma of all injected mice. The concentration of the transgene in the plasma was 1-2% of normal human values one month after injection. At the second time point, the levels of FIX decreased, yet were still detectable in all mice. After the fourth month after injections, the levels of FIX raised, with the achievement of 7.3% of normal levels of hFIX in the plasma in one mouse.

Knowing that it was possible to detect hFIX in the plasma of mice injected with lentiviral vectors, the integrating and the non-integrating vectors expressing hFIX were then compared *in vivo*. MF1 mice were injected with 5 μ L of concentrated vectors of the integrating (35.2ng_(RT)/ μ L), D64V (28.2ng_(RT)/ μ L) and the DNW Δ att (38ng_(RT)/ μ L) vectors. Blood samples of injected mice were taken at several time points and the concentration of hFIX in the plasma was accessed by ELISA.

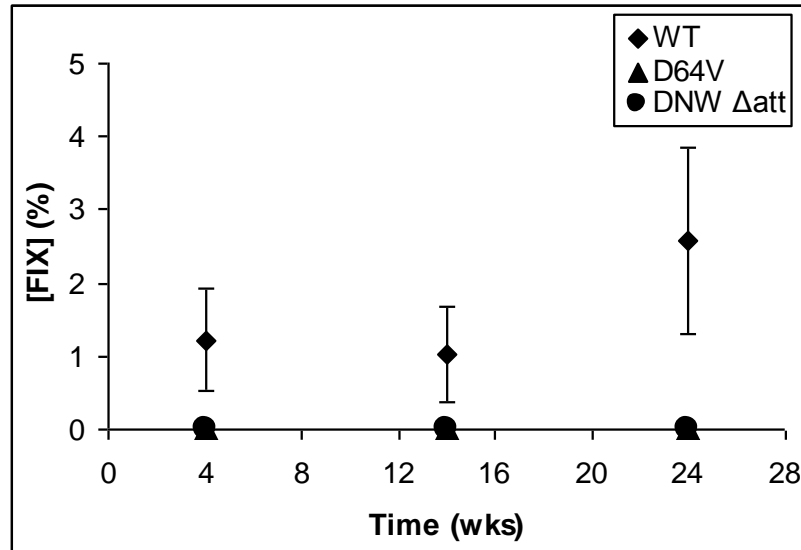


Figure 5.3: Comparison of hFIX expression from integrating and non-integrating lentiviral vectors *in vivo*.

Neonatal MF1 mice were injected with 5 μ L of concentrated vectors in the tibialis anterior muscle (titres were: integrating vector (wt) – 35.2ng_(RT)/ μ L, D64V – 28.2ng_(RT)/ μ L and D64V + N120L + W235E + Δ att (DNW Δ att) – 38ng_(RT)/ μ L. Plasma hFIX was assessed by ELISA in blood samples taken at several time points. Points in the graph represent the average obtained for three animals and error bars the standard error of the mean.

Human FIX was identified in all mice injected with the integrating vector, but no plasma hFIX was detected in samples from mice injected with the NILV (Figure 5.3). The average hFIX concentration levels in mice injected with the integrating vector were similar to the previous experiment (Figure 5.2). An average of 1.2% of normal levels were detected one month after injection, a lower level obtained at the three months time point, and the highest levels reached at the last time point.

The integrating and non-integrating vectors were also tested using the mouse disease model of haemophilia B. Since it was known that the levels of circulating hFIX

were not very high, it was decided to increase the vector dose injected into these animals. Haemophiliac mice were injected with a total of 15 μ L of concentrated vectors in the tibialis anterior muscle in both legs. The particle concentration of the injected integrating and the non-integrating vector were different, being 61.8ng_(RT)/ μ L and 25.6ng_(RT)/ μ L, respectively. It was decided not to dilute the integrating vector to match the NILV titre in order to maximise the concentration of transgene in the plasma. Blood samples from the injected mice were then taken at several time points, and the concentration of hFIX was measured by ELISA.

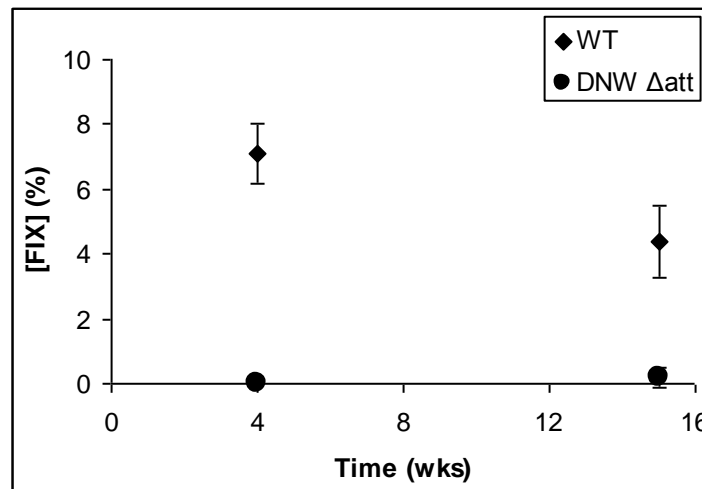


Figure 5.4: Expression of hFIX in haemophiliac mice
Haemophilia B mice were injected with 15 μ L of integrating and non-integrating lentiviral vectors in the tibialis anterior muscle of both legs. Titres were 61.8ng_(RT)/ μ L and 25.6ng_(RT)/ μ L for the integrating (wt) and non-integrating (DNW Δ att) vectors, respectively. Blood samples were taken at several time points and hFIX was then quantified in plasma by ELISA. Points in the graph represent the average obtained for three animals and error bars the standard error of the mean.

The levels of hFIX found in mice injected with the integrating vector were approximately 7% one month after injection (Figure 5.4), but the concentration of FIX in these mice diminished in the fourth month after injection, as it happened in normal mice (Figure 5.3). It is important to note that these levels are sufficient to ameliorate the disease. In mice injected with the NILV DNW Δ att, no FIX was detected four weeks after injection and only 0.5% in one mouse was detected in the fourth month. This reflects the difference

of vector particles delivered in each injection, but it also confirms that expression of hFIX from the vectors tested needs to be improved in order to achieve disease amelioration.

5.4 – Optimisation of hFIX expression from muscle cells

In order to improve expression from the developed vectors, it is possible to optimise the codon sequence of the transgene. There has been a report of a codon optimised version of the hFIX transgene that potentially increases expression (Nathwani *et al.*, 2006). The hFIX coding sequence in this report was modified by using a subset of codons most frequently found in highly expressed eukaryotic genes ("codon optimisation"), and then adjusted to reduce the potential for inappropriate splicing and CpG methylation to augment transgene expression.

FIX produced in muscle cells accumulates in the extracellular space. Possibly, this is due to an interaction with collagen IV which is regarded as a natural binding for FIX (Wolberg *et al.*, 1997), trapping the protein in the extracellular space. Schuettrumpf and colleagues demonstrated that the plasma concentration of a mutated form of FIX with low affinity for collagen IV (FIX K5A/V10K) expressed from an AAV vector injected intramuscularly in mice is 2-5 fold higher than the wt FIX. It was also published in the same report that the mutation R338A increases the specific activity of the FIX protein (Schuettrumpf *et al.*, 2005). Therefore, it was decided to use the codon optimised version of the hFIX (hFIXco) containing the two mutations that would increase the circulating levels of the expressed protein. The mutations K5A and V10K were introduced in the FIXco transgene by site directed mutagenesis obtaining the hFIXcoKV transgene. The mutations were confirmed by sequencing the plasmid. The hFIX transgene in the vector was then replaced by hFIXco and hFIXcoKV, and vectors were produced. The vectors were then tested *in vitro*. 293T cells and dividing and non-dividing C2C12 cells were infected with an integrating vector expressing FIXco or FIXcoKV. The levels of FIX in the supernatant were then measured two days after infection by ELISA.

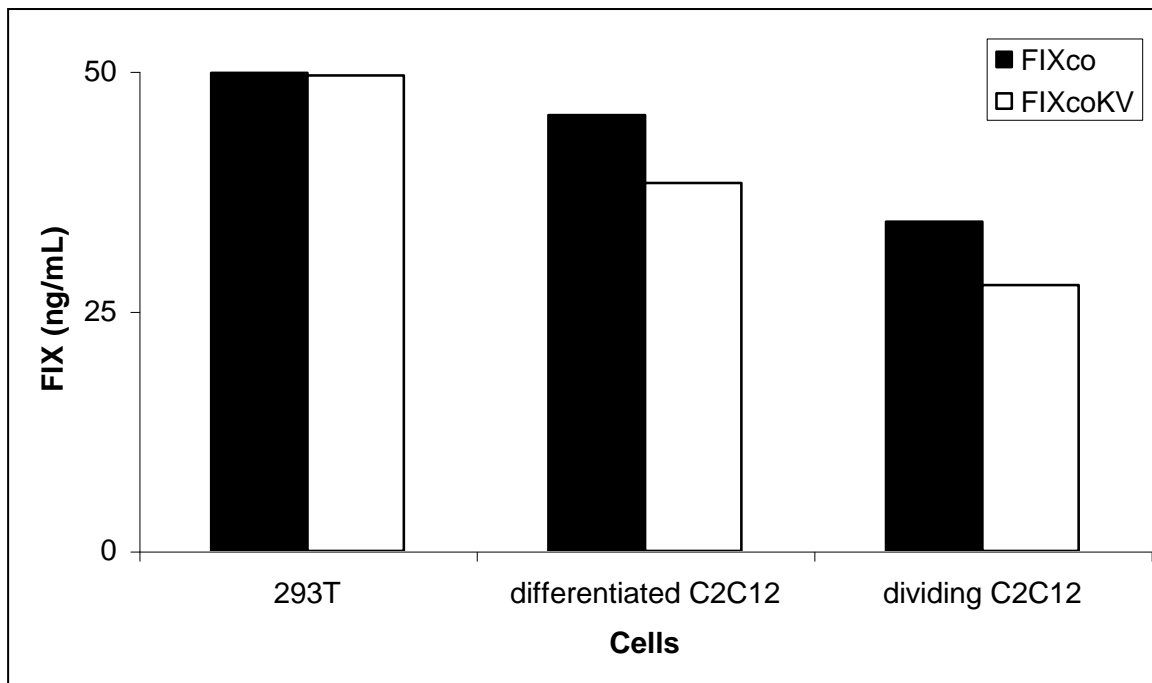


Figure 5.5: Comparison of expression of different codon optimised hFIX versions *in vitro*. Integrating vectors were produced carrying the hFIX codon optimised cDNA with and without the mutations K5A and V10K (FIXcoKV and FIXco, respectively). 293T cells were then infected with 4.5 and 3.8 ng_(RT) with the vectors FIXco and FIXcoKV, respectively. Dividing and differentiated C2C12 cells were also infected but with 10 fold more vectors than 293T cells. The concentration of hFIX was assessed in the supernatant 3 days after infection by ELISA

hFIXco was expressed in all cell lines tested (Figure 5.5), but it was not clear if the mutated form of FIXco is expressed in higher concentrations. These vectors were then tested *in vivo*. MF1 neonatal mice were injected with 33μL of vectors in several muscle groups (tibialis anterior and intercostal muscles), maximising the amount of vector able to be injected in such mice. The concentration of injected vectors was again different, since it was decided not to dilute the vector stocks. The concentration of injected integrating vector expressing FIXco and FIXcoKV were 77.5ng_(RT)/μL and 34.8ng_(RT)/μL, respectively and the concentration of the DNW Δatt vector injected expressing FIXco and FIXcoKV were 55.8ng_(RT)/μL and 66.1ng_(RT)/μL, respectively. Plasma hFIX was subsequently quantified by ELISA in blood samples taken from the mice at several time points.

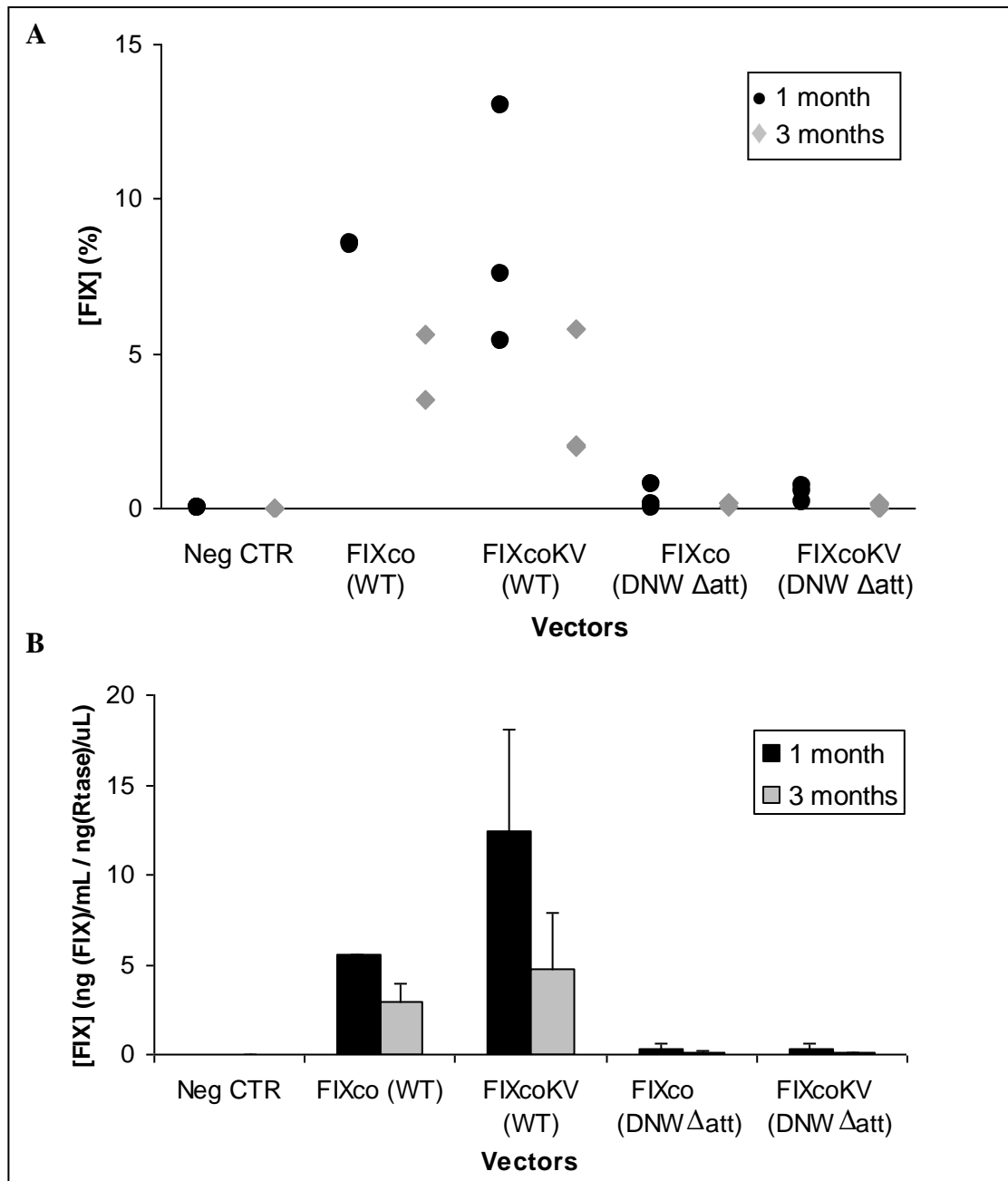


Figure 5.6: Comparison of expression of different hFIX transgenes *in vivo*. Integrating (wt) and non-integrating (DNW Δatt) lentiviral vectors expressing the codon optimised version of hFIX (FIXco) and the codon optimised hFIX harbouring the mutations K5A and V10K (FIXcoKV) were injected in neonatal MF1 mice. A total of 33μL of each vector was injected intramuscularly in each mouse. The titres (ng_(RT)/mL) of the vectors injected were: FIXco (wt) – 77.5, FIXco (DNW Δatt) – 55.8, FIXcoKV (wt) – 34.8 and FIXcoKV (DNW Δatt) – 66.1. Blood samples were then taken at several time points and the hFIX plasma concentration was quantified by ELISA. The data collected is shown in two configurations. The concentration of plasma hFIX found in each mouse is represented in **A**. Bars in **B** represent the average FIX concentration obtained for each mice group taking into account the concentration of injected particles and error bars the standard error of the mean.

The quantity of injected viral particles for each vector was different, so it was decided to show two representations of the data collected from this experiment. Figure 5.6 B takes into account the different titre of the injected vectors and permits a direct comparison of expression from the vectors. Figure 5.6 A allows the estimation of disease amelioration by showing the percentage of normal FIX circulation reached by the vectors in the blood of the animals. It is clear that the codon optimised version of hFIX harbouring the two mutations K5A and V10K increases the circulating FIX levels compared with the transgene without mutations (Figure 5.6 B). The levels of plasma hFIX obtained by injecting the mutated improved form of FIX were two fold higher, on average, than in mice injected with the version not mutated of FIX, both in the case of integrating and non-integrating vectors. FIXcoKV is therefore the best transgene to reach the maximum levels of hFIX in the plasma of mice.

The levels of expressed and secreted FIX in mice achieved by the injection of integrating vectors expressing hFIXco and hFIXcoKV were 8.5% and 8.6%, respectively, the highest value obtained in all experiments. These values are clearly higher than the threshold value for disease amelioration. The levels of hFIX obtained in mice injected with the non-integrating DNW Δ att vector were lower, being 0.3% for the vector expressing hFIXco and 0.5% for the vector expressing hFIXcoKV.

5.5 – Discussion

Data presented in the previous chapter demonstrated that NILVs are able to infect muscle cells, both *in vitro* and *in vivo*, with sustained transgene expression for up to 8 months *in vivo*. This research was then extended by testing the developed NILVs in gene therapy.

There are several models where NILVs can be tested for disease correction of genetic diseases. The chosen model – correction of haemophilia B by expression of FIX from muscle – possessed several characteristics where NILVs would be ideal candidates to be the vector of choice for the treatment of this disease. The vector only needs to achieve expression and secretion of FIX up to 1% of normal levels in plasma to ameliorate the

disease (Ljung, 1998). The published approaches to achieve disease correction rely on the use of AAV vectors to deliver the transgene to muscle or liver cells. It was reported that when this vector was delivered to the liver, there was an immune response against the vector in humans (Manno *et al.*, 2006; Mingozi *et al.*, 2007). When the vector was delivered to muscle, the expression of the transgene was low and transient (Manno *et al.*, 2003). Up until this moment, there has not been any report of the use of lentivirus based vectors to deliver the transgene to muscle cells, so the investigation of the use of lentivectors for this purpose would be very interesting.

The vectors that were used in the previous experiments had eGFP as a transgene, so it was necessary to change this transgene into the clinically relevant hFIX transgene. The vectors were then produced and tested *in vitro*. It was demonstrated that the integrating lentiviral vector expressing hFIX from the SFFV promoter was able to deliver the transgene to 293T and C2C12 muscle cells and that the cells were able to secrete the transgene to the supernatant (Figure 5.1). Most importantly, the produced hFIX was readily detectable and quantifiable by ELISA.

Next, it was necessary to perform pilot experiments *in vivo*, as there was no report of the use of these vectors to deliver the hFIX transgene to muscle cells *in vivo*. The integrating lentivector was injected in MF1 neonatal mice, and hFIX was monitored over time. Data from this experiment showed that hFIX was detected in blood, at 1-2% of normal FIX levels. Although these levels were low, the concentration of plasma FIX was higher than the therapeutic threshold.

In a second experiment, also performed in immunocompetent MF1 mice, expression from muscle cells by the integrating lentivector, the D64V and the DNW Δ att mutants were compared. Unfortunately, expression from the NILVs was not detected at any time point after injection. It is possible that the NILVs tested were able to deliver the DNA to muscle cells and that expression of the transgene was present, although at such low concentration levels that were not detected due to the sensitivity of the ELISA assay. Nevertheless, disease amelioration levels were again achieved by the injection of integrating lentiviral vectors, as the concentration of FIX in plasma was higher than the 1% of normal levels threshold.

The integrating and non-integrating vectors were also tested in the animal disease model of haemophilia B. In this experiment, the vector dose injected was increased in order to achieve a higher level of plasma FIX. The FIX expressed from the NILV was not detected one month after injection, but it was detected in one mouse four months after injection, although at the low concentration of 0.5% of normal levels. In addition, the concentration of plasma FIX achieved by the integrating vector was higher than in the previous experiments. In these mice, 7% of normal levels of hFIX in the plasma were quantified in the blood samples, reflecting the higher concentration of particles injected in these mice.

It was therefore demonstrated that integrating lentiviral based vectors were able to achieve the target FIX concentration for disease amelioration, both in immunocompetent MF1 and haemophilic mice. Nevertheless, it was apparent that expression and secretion of FIX had to be improved to achieve detectable levels in plasma by NILVs. In order to augment the concentration of hFIX in blood, the hFIX sequence was codon optimised to increase expression. Secretion to plasma was improved by mutating residues to disrupt interaction of this protein with collagen IV without compromising enzymatic activity, as reported previously (Schuettrumpf *et al.*, 2005).

Integrating vectors expressing the codon optimised version of FIX with and without the mutations K5A and V10K were produced and tested *in vitro*. It was demonstrated that both 293T and C2C12 muscle cells were able to produce and secrete this transgene, although it was not clear whether the optimisation of FIX would increase expression or secretion by the cells.

Next, the improved vectors were tested *in vivo*. Comparing the values obtained with the vector expressing the codon optimised and the normal version of hFIX, it was not clear if the codon optimisation increases the expression levels of the transgene in muscle cells. For the integrating vector, the average value reached by the vector expressing hFIX_{co} was 8.5%, at the one month time point, when 2556ng_(RT) were injected in MF1 mice. For the integrating vector expressing hFIX, the levels reached in MF1 mice at the same time point were 1.2% when 176ng_(RT) were injected and 7.1% when 920ng_(RT) were injected in haemophilic mice. The haemophilic mice used were of BL6 background, so a direct comparison with the MF1 mice is not correct, because there can be differences of

expression not due to the vector but due to the biological phenotype of the animals. Also, the amount of injected vector particles in the MF1 mice was very different. Hence, it is possible that the comparison is not appropriate since the dependence between injected particles and the percentage of hFIX obtained may not be linear. If there was a correlation, and the codon optimisation improved expression from muscle cells, it would be expected to obtain more than 17.5% with the higher amount of vector injected. However, this was not the case, as only half of that concentration was obtained.

For the DNW Δ att vector, the comparison is even more difficult as hFIX was not detected at the one month time point in mice injected with the vector expressing the normal hFIX. But when mice were injected with the vector expressing the codon optimised version, at the same time point, the levels of hFIX in the plasma were 0.3%. It may be possible that it is necessary to reach a certain threshold to be able to quantify secreted FIX in the plasma, and that it was not possible to reach it by injecting only 140ng_(RT) of the vector.

Comparing the vectors expressing the mutated and non-mutated forms of hFIX, it was shown that there was a 2 fold increase of expression when the mutated form was used as a transgene. These results confirm the previous data published, where AAV vectors were used to deliver the wild type and mutated forms of hFIX to muscle cells (Schuettrumpf *et al.*, 2005). It was therefore proven that the best transgene to achieve the highest concentration of hFIX in plasma by production of the protein by muscle cells is the FIX transgene mutated to inhibit interaction with collagen IV. Haemophilia B patients may therefore be treated with muscular injections of integrating lentiviral vectors expressing this transgene. However, in order to increase the levels of hFIX in plasma and reach the therapeutic threshold, it may be necessary to increase the volume of injected vectors or even repeat injections at several time points.

Despite the efforts taken to improve expression from the NILVs, it was concluded that there was a problem with expression from these vectors. It is therefore important to study the reasons for it, in order to improve expression. Nevertheless, it should be emphasised that the integrating lentiviral vector is able to sustain expression of hFIX in mice at a higher level than the therapeutic threshold, indicating disease amelioration of haemophilia B.

Chapter Six

Comparison of Expression from Episomal and Integrated vectors

6 – Comparison of Expression from Episomal and Integrated Vectors

6.1 – Introduction

Gene therapy is a powerful technique that involves the use of vectors to deliver a transgene of interest into cells. Vectors for gene therapy have been developed and improved throughout the last years. Some vectors have been engineered based on viruses in order to use their natural capabilities to deliver genetic material into cells. All developed vectors have advantages and disadvantages. The choice of vector has to be careful to achieve the appropriate goal. It relies on several different aspects such as vector packaging capacity, efficiency of transduction of the target cell, immunological reaction to the vector, infection of dividing or non-dividing cells, integration of the genetic material into the cell chromosomes for prolonged transgene expression when cells are dividing, among others characteristics.

Vectors based on adenoviruses have the advantages of a large packaging capacity and infection of non-dividing cells, but elicit a major immune response. Adeno-associated viral vectors (AAV) also infect non-dividing cells, but have a small packaging capacity (4.9kb) and do not integrate the vector DNA into the cell chromosomes. Retroviruses and lentiviruses have a relatively large packaging capacity (9kb) and integrate the viral genome into the host DNA. Also, lentiviral vectors have the ability to infect non-dividing cells, contrary to retrovirus vectors. Clinical trials have demonstrated the efficiency of retroviral vectors, but uncovered also a potential drawback of these vectors. In the X-linked severe combined immunodeficiency trials, several patients developed leukaemia due to insertional mutagenesis of the vector (Hacein-Bey-Abina *et al.*, 2003b).

Here, NILVs were developed in order to exploit the capacity of lentiviral vectors of efficiently infecting non-dividing cells while reducing the risk of insertional mutagenesis. This study has demonstrated that NILVs are able to transduce non-dividing cells *in vitro* and *in vivo* with prolonged transgene expression. Most importantly, it has been shown that the episomal viral forms are stable and not degraded in cells (chapter 4). However, the

level of expression from NILVs was lower compared to the integrating vector. A secreted transgene being expressed from muscle cells transduced with NILVs was found in the blood stream at a 10 fold lower concentration compared with cells infected with the integrating counterpart (Figure 5.6).

It is therefore important to study the reason for this discrepancy in order to optimise expression levels. In this chapter, it will be studied in which step of expression NILVs are impaired, by comparison to the integrating counterpart. It is also important to compare the performance of these NILVs with other episomal vectors. Here, the first comparative study of expression from integrating and non-integrating lentiviral vectors with an AAV vector and also with plasmid transfection will be presented.

6.2 – Comparison of expression from integrated and non-integrated lentiviral vectors

It was previously shown in chapter 4, that NILVs are able to sustain expression in muscle cells for up to 8 months *in vivo*. When eGFP was quantified in muscle from injected mice, it was found that expression from NILVs was statistically not different to integrating vectors, three months after injection. eGFP is a very stable protein that is accumulated in the cytoplasm of cells. When NILVs were tested in the haemophilia model, expressing a secreted protein, it was found that NILVs expressed 10 fold less protein compared with their integrating counterpart. FIX is secreted and has a short half-life in circulation. These characteristics show that the comparison of expression from integrating and non-integrating lentivectors in the haemophilia model does not have the bias of analysing a random area of injected muscle to assess intracellular protein production and the accumulation of a stable protein intracellularly.

From the previous experiments, there was a difference of expression when different promoters were used, so it was necessary to determine if lower expression from NILVs was promoter independent. Since expression from NILVs occurs from episomal DNA, the SFFV promoter was compared with the p5 promoter. p5 is a promoter derived from adeno-associated virus (AAV), which is a natural episomal virus. Expression from the

ubiquitously acting chromatin opening element (UCOE) was also compared, as this enhancerless promoter is involved in chromatin assembly and re-configuration. This promoter is derived from the human *HNRPA2B1-CBX3* locus and it has been shown that linked to CMV it prevents silencing and increases median levels of expression (Zhang *et al.*, 2007a; Williams *et al.*, 2005). The 293T cell line was chosen as target cells to perform this study because they are a permissive cell line, where the promoters chosen do not have a particular advantage. 293T cells were infected with integrating and non-integrating vectors expressing eGFP from the different promoters. Two days after infection, the cells were harvested and assessed for eGFP production by flow cytometry analysis. The transgene production was estimated by the mean intensity fluorescence (MFI) of infected cells, in populations where only 30% of the cells were expressing, to insure that cells were not infected with more than one vector copy, on average.

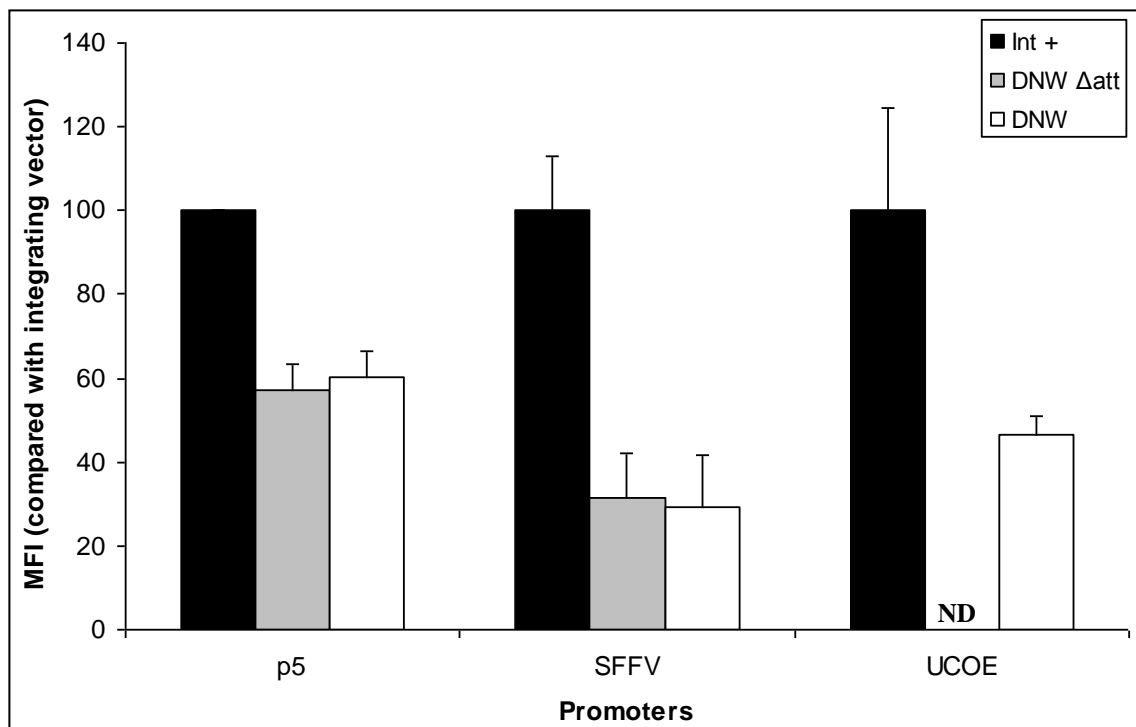


Figure 6.1: Comparison of expression from different promoters in integrating and non-integrating lentiviral vectors *in vitro*.

Integrating and non-integrating lentiviral vectors expressing eGFP driven by different promoters were produced and used to infect 293T cells, at MOI 0.3. Three days after infection, cells were harvested and the eGFP produced was quantified by measuring the mean intensity fluorescence (MFI) by flow cytometry. Bars represent an average of 3 experiments and error bars the standard error of the mean.

It was found that regardless of the promoter used, NILVs always expressed less transgene than the integrating vector (Figure 6.1). Nevertheless, expression from NILVs was only 1.7 fold lower when p5 promoter was used, compared with expression from the integrating counterpart. The drop of expression is higher when the SFFV promoter was employed in the vectors, being 3.4 fold, on average, in this case. It is interesting to note that expression from the UCOE promoter in the context of NILVs is similar to the p5 promoter, as the expression from these promoters is only 2 fold less compared with the integrating promoter.

In order to study the reason for lower expression 293T cells were infected with integrating and non-integrating lentivectors expressing eGFP from the SFFV promoter, the promoter with highest difference in expression between the integrated and non-integrated vectors. The experiment was designed in order to deliver similar quantities of DNA from the vectors into cells. Two days after infection, cells were harvested. Total DNA was extracted from cells by salting out. Total RNA was extracted using TRI reagent and cDNA was prepared by reverse transcription using random primers. The vector DNA and RNA concentration was then assessed by qPCR with primers designed against the WPRE sequence. The production of protein was accessed by measuring the MFI of eGFP positive cells by flow cytometry.

For the ease of comparison, the integrating vector DNA and RNA concentration was assumed to be 1 and the other vectors were compared to this value. In Figure 6.2 it is shown that when similar viral DNA concentration of integrating and non-integrating lentivectors are delivered into cells, there is a lower amount of transgene produced from NILVs. Furthermore, it can be seen that the vector RNA concentration in NILV infected cells was lower compared with the integrating counterpart. It is therefore demonstrated that the episomal lentivectors are less transcriptionally active than the integrating vectors.

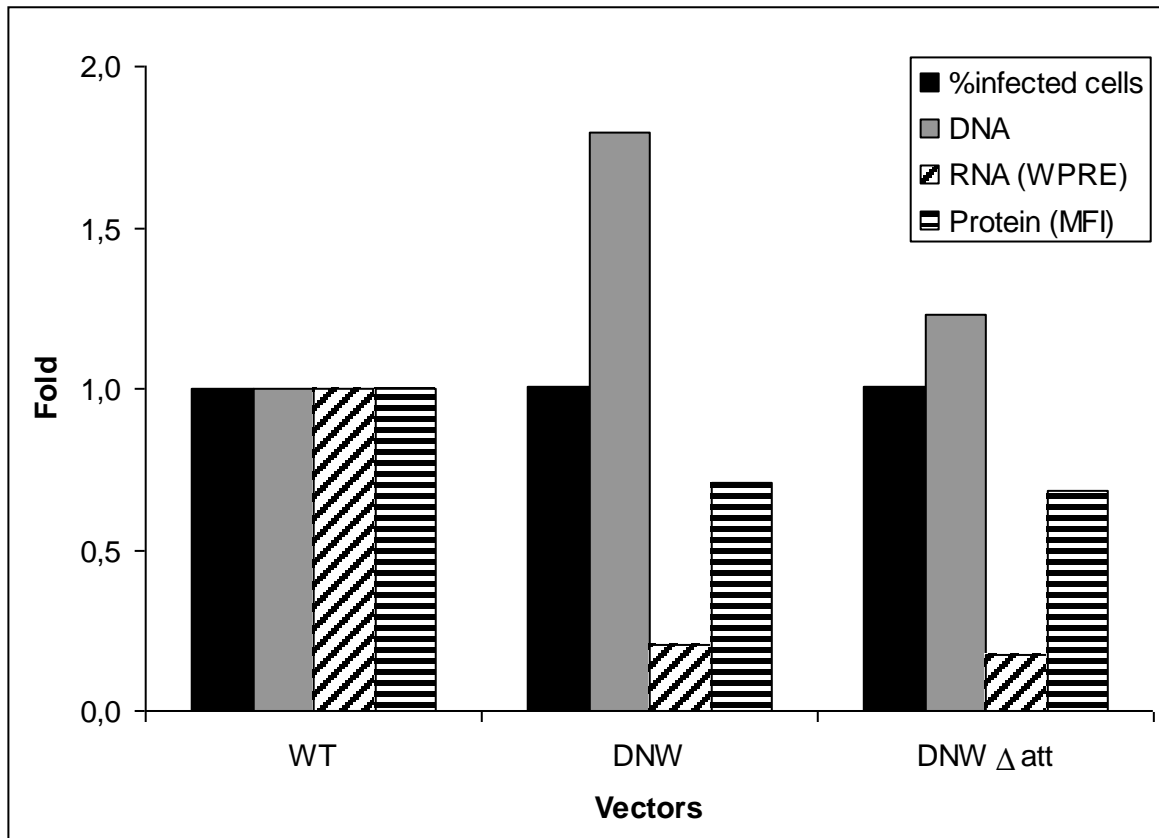


Figure 6.2: Analysis of DNA, RNA and Protein in cells infected with lentiviral vectors. 293T cells were infected with integrating (wt) and non-integrating (DNW and DNW Δ att) lentiviral vectors expressing eGFP from the SFFV promoter at MOI 0.3. Two days after infection, cells were harvested. DNA and RNA were extracted and quantified by qPCR. eGFP was quantified by flow cytometry. The values obtained were then compared with the integrating vector.

6.3 – Comparison of expression from episomal and integrated viral vectors

It was proven that NILVs express less transgene than their integrating counterparts and the reason for it is a lower transcription activity from NILVs. Therefore, it is important to extend this research and compare expression of NILVs with other episomal vectors.

AAV is a promising vector for gene therapy. It has been used in clinical trials to deliver FIX to the liver and muscle of patients with haemophilia B. Unfortunately, the expression was transient in liver delivery due to an immunological reaction and expression levels were low when delivered to muscle. Nevertheless, there is scope to improve this vector and use it in genetic diseases, where the transgene can be accommodated into the low capacity of this vector. This viral vector delivers a single stranded DNA (ssDNA) copy to the cell. After second-strand synthesis, circles and concatemers are formed and maintained in non-dividing cells as episomal DNA. This vector was then chosen to compare expression between integrated and episomal vectors.

The AAV and lentivector systems are very different, and a direct comparison is difficult. Both systems have particular need of sequences in the vector to be incorporated into the virus particles. Also, while the lentiviral based vector is incorporated in the virion as RNA, the AAV vector genome is a single stranded DNA (ssDNA). Furthermore, these vectors infect cells with different efficiencies, so a permissive cell line to both vectors has to be used to allow a comparison. 293T cells are a permissive cell line for both AAV and lentivectors. C2C12 cells are a good model to study infection of muscle cells *in vitro*. So these cell lines were chosen to perform this study.

In order to compare the expression performance from lentivirus and AAV based vectors, the promoter and transgene have to be the same. Retroviruses manage to efficiently export unspliced RNA into the cell cytoplasm. This is achieved by *cis*-acting elements, such as the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). The WPRE element is commonly used in lentiviral vectors, to achieve higher titres. This element may also influence RNA stability in infected cells. Therefore, the SFFV-eGFP-WPRE sequence used in the lentivectors was cloned into an AAV vector.

An experiment was designed where differences in infection were not taken into account. Instead, it only relied on the delivery of a similar quantity of DNA containing the same expression cassette into the cells. With these characteristics, the protein production would only vary from intrinsic differences of the DNA or viral proteins carried in the vector and, with a similar expression cassette, a difference would derive from the particular *cis*-acting sequences used specifically in each vector.

All vectors were first titred in 293T cells to quantify the concentration of transducing units (TU) present in each vector preparation. 293T and dividing C2C12 cells were then infected with different concentrations of the vectors. Two days after infection, DNA was harvested and the viral vector copy number was quantified by qPCR. The protein production was assessed at the same time point analysing the MFI by flow cytometry. In order to standardise between different experiments, the MFI of the eGFP positive cells obtained for each sample was normalised against the MFI value obtained for untransduced cells in the same experiment.

The expression from the AAV vectors can be delayed due to the second-strand synthesis, which is a slow process. Therefore, differentiated C2C12 cells were also infected with the vectors and vector copy number and expression analysed 12 days after infection.

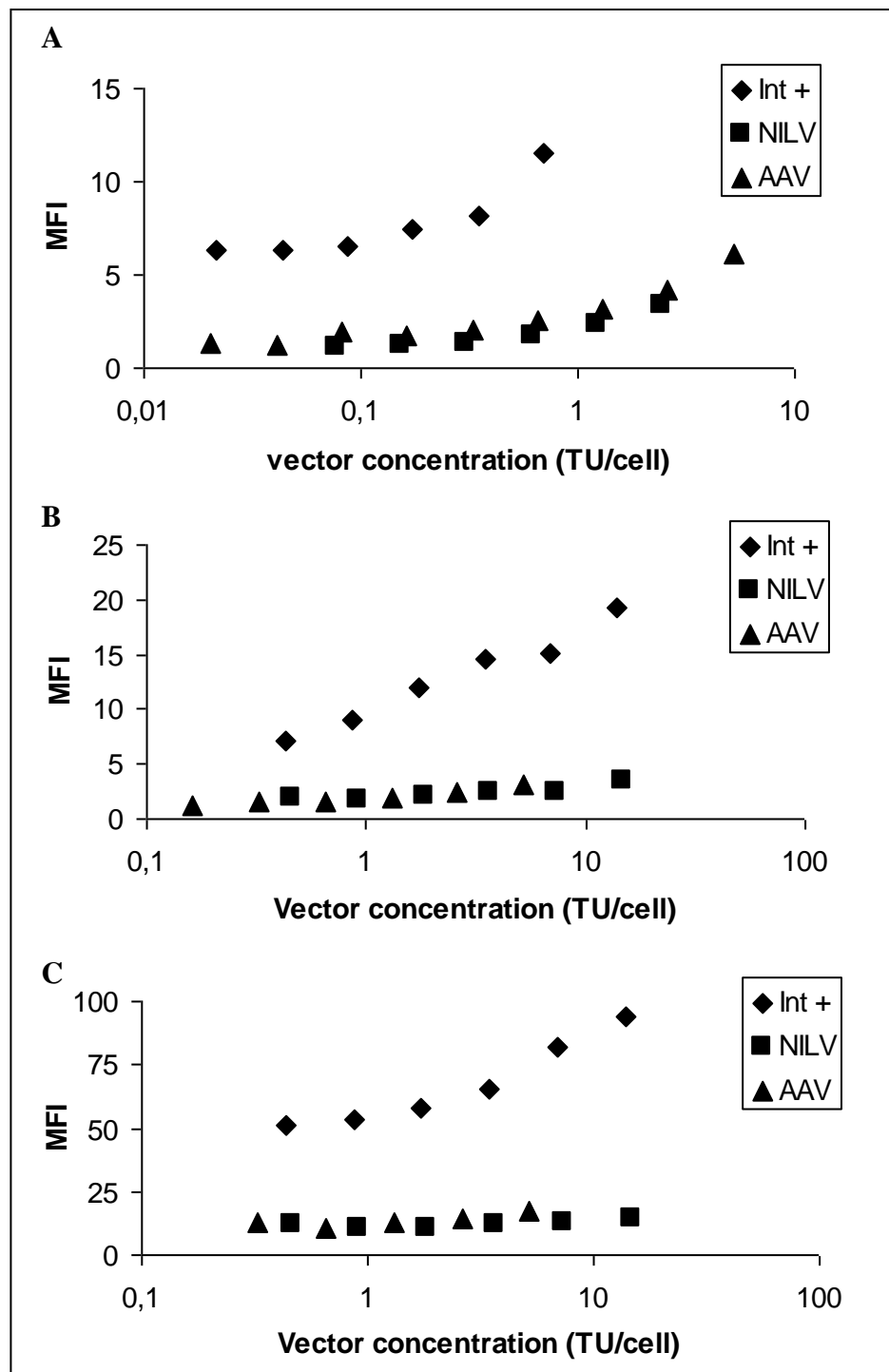


Figure 6.3: Expression from different integrating and episomal viral vectors *in vitro*. Cell lines were infected with different concentrations of integrating (Int +) and non-integrating (NILV) lentiviral vectors and with AAV vectors (AAV). All vectors expressed eGFP from the SFFV promoter. Transgene production was assessed by measuring the MFI by flow cytometry. The value obtained for each sample was normalised against the MFI of untransduced cells analysed at the same time. For 293T (A) and dividing C2C12 cells (B), cells were harvested 2 days after infection. For differentiated C2C12 cells (C), the analysis was performed 12 days after infection.

Figure 6.3 shows that, in all cell lines tested, the increase of concentration of vectors used during infection increases the MFI obtained in eGFP expressing cells. Also, in 293T cells, in order to obtain a value of MFI in cells transduced with the AAV vector similar to the integrating lentivector, it is necessary to infect cells with at least 100 fold more transducing units per cell. In dividing and differentiated C2C12 cells, the level of MFI obtained in cells infected with the AAV vector never reached the one obtained for the integrating vector. The data also indicates that there is eGFP accumulation in C2C12 cells over time, as the MFI levels in differentiated cells (analysed 12 days after infection) are higher compared with the values obtained 2 days after infection in dividing C2C12 cells.

Transgene expression was then compared in samples where 20-40% of the cells were eGFP positive. To achieve such percentage in dividing and differentiated C2C12 cells, the concentration of integrating and non-integrating lentiviral vectors used was 10TU/cell and 2TU/cell for the AAV vector.

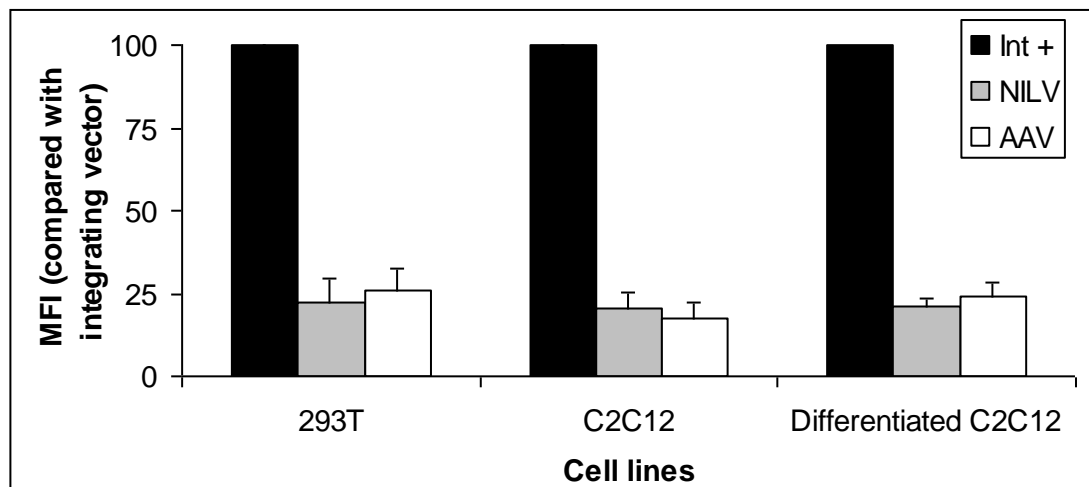


Figure 6.4: Expression comparison between integrated and episomal viral vectors
 293T cells were infected with integrating (Int +) and non-integrating (NILV) lentiviral vectors and AAV vectors (AAV) at a concentration of 0.3TU/cell. Dividing and differentiated C2C12 cells were infected with 10TU/cell of Integrating and NILVs and with 2TU/cell of an AAV vector. All vectors expressed eGFP from the SFFV promoter. Expression was analysed 2 days after infection for dividing cells and 12 days after infection for the differentiated cells by measuring the MFI by flow cytometry. The values obtained were standardised with the MFI of a negative control (untransduced cells) and then divided by the average value of cells infected with the integrating lentiviral vector. Bars represent the mean of MFI obtained for at least three replicates of three independent experiments. Error bars represent the standard error of the mean.

Figure 6.4 demonstrates that the amount of eGFP produced by the integrating vector is higher than the one produced by the episomal vectors. Both NILV and AAV vectors transgene expression levels were similar and only 20-25% of the value obtained with the integrating lentiviral vector. This result was similar in all the cell lines tested.

The relation between the DNA delivered and the protein produced was also studied. The number of vector copies obtained by qPCR was normalised against the concentration of extracted DNA (that was standardised to obtain the number of cells in each sample) and plotted against the corresponding MFI values.

Figure 6.5 shows that, in all the cells tested, the increment of DNA delivered to the cell increases the amount of transgene being produced. For the 293T cells, when similar concentrations of vector DNA were delivered to the cells, the concentration of protein produced from the integrating lentiviral vector was always higher compared with the episomal vectors. Both episomal AAV and non-integrating lentiviral vectors had a similar expression pattern, which was 3-5 fold lower than the integrating vector. Similar results were obtained for both dividing and differentiated C2C12 cells infected with the vectors, where the cells infected with the integrating vector express more transgene than with the episomal vectors NILV and AAV.

Since the formation of DNA is different for the vectors studied, and the rate of second-strand synthesis for AAV vectors is slow, it is also important to compare the number of vector copies obtained in dividing C2C12 cells (which were assessed two days after infection) and in differentiated C2C12 cells (which were measured 12 days after infection). The copy number in cells infected with the same concentration of integrating lentiviral vectors decreased two fold. Similarly, the number of AAV vectors was also two fold lower in differentiated C2C12 cells. The copy number obtained in cells infected with the NILV also decreased, but at the higher rate of 14 fold.

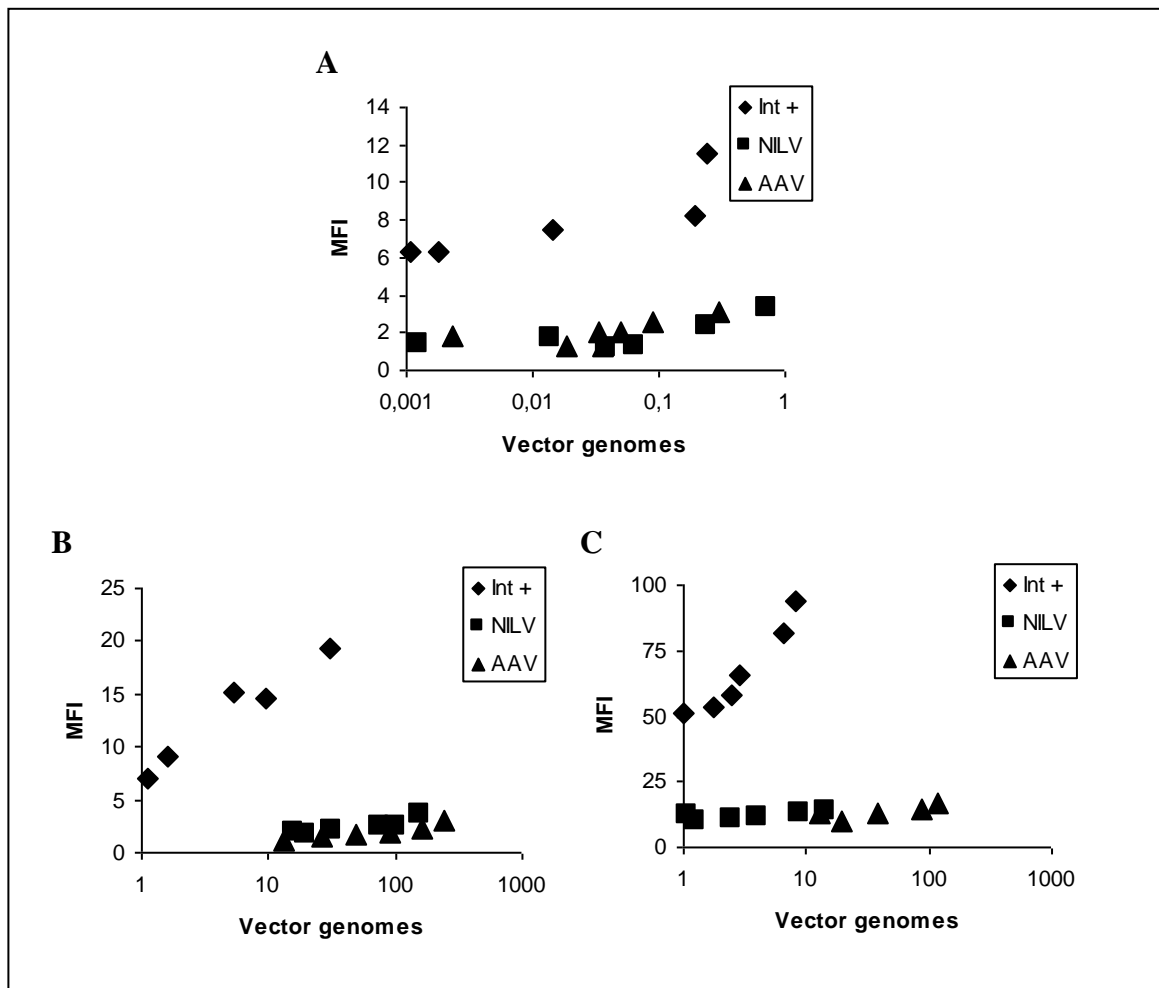


Figure 6.5: Relation between expression and vector DNA concentration *in vitro*. 293T (A) and dividing (B) and differentiated (C) C2C12 cells were infected with different concentrations of integrating and episomal viral vectors expressing eGFP from the SFFV promoter. Two days after infection, 293T and dividing C2C12 cells were harvested. Differentiated C2C12 cells were harvested 12 days after infection. A sample was used to assess transgene production by measuring the MFI of eGFP expressing cells by flow cytometry. The value obtained was then normalised against the MFI of untransduced cells analysed at the same time. Total DNA was also extracted and used as a template to measure vector genomes. The concentration of DNA obtained was compared with standards to evaluate the number of cells in each sample and then used to normalise the number of genomes obtained by qPCR.

6.3 – Comparison of expression from lentiviral vectors and transfected DNA

The comparison of NILV and AAV vectors revealed that both have similar expression patterns, and that the level of protein produced is 4-5 fold lower compared with an integrating lentiviral vector. It was then decided to extend this study by analysing the expression from lentivectors and a transfected plasmid.

In order to compare expression from these systems, it is necessary to have a similar expression cassette. Therefore, the transfected plasmid DNA chosen (SEW) was the lentivector plasmid used in the production of integrating and NILVs, that contains eGFP being expressed from the SFFV promoter and all the *cis* elements used in the lentivectors.

293T cells were used as a model for this study. The transfection reagent chosen to deliver the plasmid DNA to the cells was Lipofectamine 2000. This is a commercial reagent that complexes with the DNA and allows an efficient delivery of the plasmid to the target cell nucleus, when cells are dividing. To investigate the expression from the transfected plasmid, the DNA-Lipofectamine complexes were diluted before transfection of the cells. 293T cells were also infected with different concentration of integrating and non-integrating lentiviral vectors. Two days after infection or transfection, cells were harvested and expression levels assessed by measuring the MFI of eGFP expressing cells by flow cytometry. Total DNA was also extracted from the cells and the vector copy number was assessed by qPCR. As described in the section 6.4, the MFI was standardised with the negative control and the vector copy number with the DNA concentration.

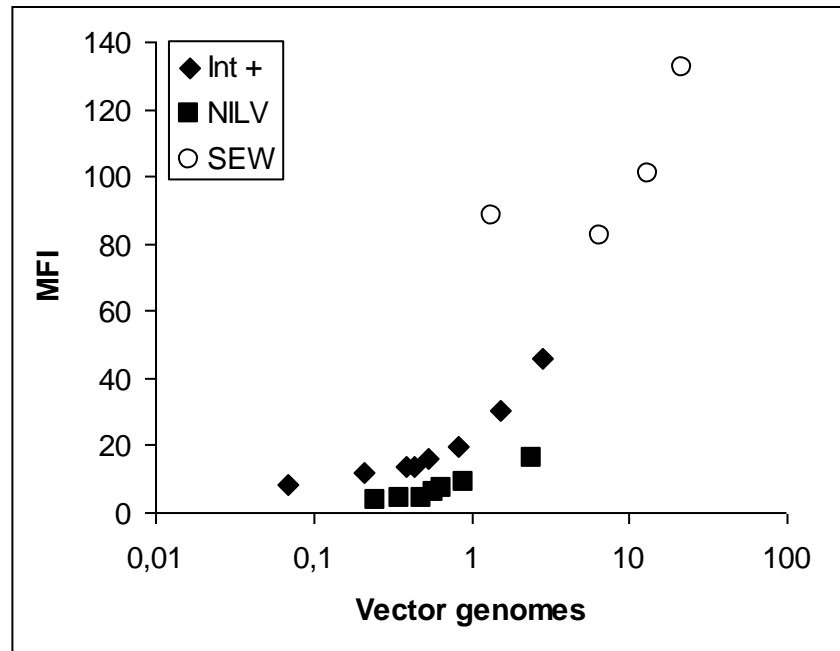


Figure 6.6: Expression comparison between integrating lentivectors, NILVs and transfected DNA in 293T cells

293T cells were infected with different concentrations of integrating (Int +) and NILVs or transfected with serial dilutions of DNA-Lipofectamine complexes (SEW). Two days after infection, cells were harvested the MFI of eGFP expressing cells was assessed by flow cytometry. The value obtained was then normalised against the MFI of untransduced cells analysed at the same time. Total DNA was also extracted and used as a template to measure vector DNA. The concentration of DNA obtained was compared with standards to evaluate the number of cells in each sample and then used to normalise the number of genomes obtained by qPCR.

The transfection of a plasmid delivers large quantities of plasmids to cells, whereas lentiviral vectors mediate the transfer of genomes at low copy numbers. Therefore, the comparison of the results obtained with these systems may not be accurate. Figure 6.6 indicates that the expression from a transfected plasmid is higher compared with both integrating and non-integrating vectors. When similar concentration of DNA was delivered to the cells, the transfected plasmid expressed 2 fold more protein compared with the integrating lentivector and 2-5 fold more compared with the NILV.

6.4 – Discussion

In this work, NILVs were developed and applied in the gene therapy context, showing that these vectors are able to sustain prolonged expression in non-dividing cells, both *in vitro* and *in vivo*. It was also shown that the vector particle production is not affected, that it is possible to achieve high infectious titres, and that viral DNA kinetics are similar to the integrating vectors. These results imply that infectivity, reverse transcription and nuclear entry of viral DNA are not altered by the introduced mutations. Nevertheless, it was also apparent that NILVs have lower expression levels compared with the integrating counterpart.

When the vector expression cassette used had the p5 promoter driving eGFP as a transgene, the expression levels were within the same range, although there was a 2 fold difference in 293T cells infected with integrating and non-integrating vectors. This difference became more apparent when the expression cassette was changed into the hFIX transgene being expressed from the SFFV promoter. *In vivo*, muscle cells infected with the DNW Δ att vector expressed and secreted the transgene into the blood stream, but the concentration of plasma hFIX was more than 10 fold lower compared with mice injected with the integrating vector.

Therefore, in order to optimise expression from the NILVs, it was necessary to determine why there was a difference in expression. The previous experiments suggested that there was a difference of expression when different promoters were used. Thus, an experiment designed to compare expression of the vectors from different promoters was performed. Figure 6.1 shows that, regardless of the promoter used, NILVs expressed lower concentrations of the transgene. Comparing the differences of expression of eGFP from integrating and non-integrating vectors, it becomes apparent that the highest difference arises when the SFFV promoter is used (3.4 fold) and the lowest when the p5 promoter is used (1.7 fold). Interestingly, the difference of expression of the vectors containing the UCOE promoter is similar to the p5 promoter (2 fold). The SFFV promoter is derived from the spleen focus-forming virus. It is regularly used in vectors since it is a strong promoter producing high concentrations of the transgene. The spleen focus-forming virus is a replication-defective murine type C virus, and it is incorporated into the target genome

(Ruscetti *et al.*, 1980). On the contrary, the p5 promoter is derived from the AAV, an episomal virus. Although the AAV can be found integrated in the target genome, it is commonly found in the episomal form. It is possible that both promoters have different characteristics that influence expression when integrated or not into a genome. Hence, it is plausible to assume that the p5 promoter would produce similar quantities of RNA whether or not it is integrated into the host genome, contrarily to the SFFV promoter. This would explain the reason for the observed difference between these promoters in the context of integrating and non-integrating vectors. The SFFV promoter was then used in further experiments to determine the reason for such discrepancies of expression because of the larger difference between integrated and non-integrating vectors.

In order to investigate why there is a difference of expression between integrating and non-integrating lentiviral vectors, it was necessary to determine at which level there was an impairment of the expression from NILVs. 293T cells were then infected with both vectors and viral DNA, RNA and produced transgene were quantified. Figure 6.2 demonstrates that the reduced expression from NILVs is due to a lower transcription level. When similar quantities of viral DNA are delivered into the cells, less RNA is produced when the viral DNA is episomal, and consequently less protein is detected.

There are several hypotheses to explain why NILVs have a lower transcription activity. There could be a faster promoter methylation in the episomal forms preventing the promoter to produce high concentration of RNA over time. Although valid, this explanation is unlikely. Experiments performed with different promoters were done 48h after transduction, a very early time point, where episomal promoters are unlikely to have become methylated. Also the UCOE promoter is not easily silenced (Zhang *et al.*, 2007a). It is possible that the episomal DNA is not located in an optimal place within the nucleus for access by the transcription machinery. Another explanation is the contribution of surrounding sequences when the vector is integrated into the genome. Although the integration of lentiviral vectors is not the same for each individual DNA, the vectors tend to integrate within highly expressed regions of the genome (Mitchell *et al.*, 2003; Schroder *et al.*, 2002). It is therefore possible that on average, the integrated forms of the vectors produce more RNA due to the enhancement of the promoter from nearby regions.

It was important to extend this study and compare expression with other episomal vectors. The AAV vector is a promising vector for gene therapy, and it has been also used to deliver FIX cDNA to muscle cells. Therefore, this natural episomal vector was chosen to evaluate expression between integrating and episomal viral vectors. To compare expression between these systems, it is necessary to have the same expression cassette, so the SFFV promoter, eGFP and WPRE sequences used in the lentivectors were cloned in the AAV vector. The vectors were produced and first titred in 293T cells by measuring the concentration of transducing units. This was performed by analysing eGFP expressing cells two days after infection.

293T cells, dividing and differentiated C2C12 cells were used as models for the expression comparison. 293T cells were chosen, as they are permissive for both viral vectors. The expression of AAV vector may be delayed due to the slow second strand synthesis process. Therefore, differentiated C2C12 cells were transduced with the vectors and analysed 12 days after and for comparison, infected dividing C2C12 cells were assessed 2 days after infection. It is important to mention that to achieve a similar percentage of expression positive cells, it was necessary to infect C2C12 cells with 5 fold more lentiviral vectors than with AAV vectors. This result indicates that these cells are more permissive towards the AAV vectors.

Cells were then infected with serial dilutions of the vectors. Expression was evaluated by measuring the MFI of eGFP positive cells. Vector genomes delivered to the cells were also quantified in the same samples. From the Figures 6.3 and 6.5, it is possible to identify a positive relation between the concentration of transduced vectors or the vector copies delivered to the cells and the protein being produced and accumulated in the cells. This result is similar in all cells and for all vectors tested. The data suggests that the level of vector concentration used in the experiments did not saturate the production of eGFP.

The cells were then infected with the vectors to obtain a population of 20-40% of eGFP expressing cells. By comparing the MFI of the eGFP positive cells, it is possible to see that the level of expression from both episomal vectors AAV and NILV is similar and lower compared with the integrating vector (Figure 6.4). Most important, the expression per copy of DNA delivered by the integrating lentiviral vector is 3-5 fold higher compared with both episomal vectors NILV and AAV, regardless of the target cells (Figure 6.5).

The comparison of the concentration of vector copies obtained in C2C12 cells two and 12 days after infection shows that there is a reduction of 14 fold for the NILVs but only two fold for the integrating lentiviral vectors. The differentiated C2C12 cells are a mixed population of terminally differentiated cells and a minority of dividing cells that may still undergo cellular division over time. The integrating vector will not be diluted out with cell division as the NILV, so this may explain the reduction of the amount of vector DNA. The AAV vector is also episomal so, if the rate of dilution due to cell division is similar to the NILV, the concentration of AAV vector should also be reduced in the same proportion. Nonetheless, the reduction of DNA concentration is similar to the integrating vector and not to the NILV. This data implies that although some of the ssDNA delivered to the cells may have been lost due to cell division, second-strand synthesis has happened and possibly, concatemers and higher order concatemer circles (through a rolling circle replication process) have also been formed. Taken together, these results suggest that the different sequences necessary and used in the production of both NILV and AAV vectors do not influence expression.

Expression from the integrating and NILVs was also compared with a transfected plasmid. Results indicate that the expression from the transfected plasmid is higher than from the integrating and NILVs. However, the vector copy number delivered by each system is very different and a comparison may not be accurate as only the highest dilution of plasmid transfection and the highest concentration of vector transduction resulted in comparable copy numbers being delivered to the cells.

Chapter Seven

Protein Delivery with Lentiviral Vectors

7 – Protein Delivery with Lentiviral Vectors

7.1 – Introduction

The correction of genetic diseases relies mainly in the addition of a correct gene to target cells. However, this method has several disadvantages. Viral vectors have a limited packaging capacity, restricting the size of the transgene. An example is the dystrophin gene, one of the longest genes known, covering 2.4 megabases. Mutations in this gene give rise to Duchenne muscular dystrophy. However, the incorporation of such gene is impossible in the available viral vectors. Another disadvantage is the control of transgene expression. The transcription of a gene can be heavily regulated by surrounding sequences. An example that demonstrates the importance of such sequences is locus control region of the β -globin, which is involved in the β -thalassemia genetic disease. This anaemic condition arises from the lack of β -globin protein and can be caused by deletions or mutations of the coding region or its promoter (Grosveld *et al.*, 1987; Van Der Ploeg *et al.*, 1980). A different approach to the sub-optimal gene addition strategy is gene repair. In this technique, the mutation or deletion of a specific sequence is repaired by a defined replacement with a correct sequence, by a process called homologous recombination (HR). This enables the production of a correct gene with the control of expression maintained from the endogenous regulatory sequences.

Homologous recombination events are rare. Early reports show that there are 1-10 events in 10^6 cells transfected with a substrate plasmid (Finn *et al.*, 1989; Thyagarajan *et al.*, 1995). However, this process can be enhanced by introducing a double strand break (DSB) in the chromosomal target. There are several ways to achieve a targeted DSB, such as the use of modified triplex-forming oligonucleotides (shown to increase HR events 3 to 10 fold, (Faruqi *et al.*, 1996; Sandor and Bredberg, 1995), modified peptide nucleic acid (shown to increase HR events 10 to 60 fold, (Faruqi *et al.*, 1998; Rogers *et al.*, 2002), modified homing endonucleases or zinc finger nucleases (ZFN).

NILVs can be used to transiently express homing nucleases or ZFN enhancing HR events. In one report, two NILVs were used to deliver the template and the homing nuclease *I-Sce I*, resulting in gene correction in 12% of the target cells (Cornu and

Cathomen, 2007). Using a similar approach, Lombardo and colleagues have shown that combining one vector providing the template and two other NILVs to transiently express the two required ZFN (the endonuclease used is only active as a dimer) resulted in up to 39% of gene editing, with HR accounting for 74% of those events (Lombardo *et al.*, 2007). The strategy presented in both reports relies on the infection of one cell with all the required vectors. One important aspect to improve the technology would be to decrease the number of vectors required to carry the template and the enzymes necessary to increase HR events, thus improving the probability of gene repair. This chapter is going to focus on the development of a new approach to deliver the template and an enzyme to enhance HR within the same vector. The goal of this vector is the delivery to the target cell of a protein contained in the viral particle and the template within the vector genome.

There have been several reports about the incorporation of foreign proteins into the lentiviral particle. Vpr is incorporated into the virion by the interaction with the gag P6 region during assembly (Paxton *et al.*, 1993). It was shown that a Vpr-GFP fusion protein could be efficiently incorporated into viral particles (McDonald *et al.*, 2002). Other fusion proteins that allow the incorporation of foreign proteins into the virion particles were also constructed, such as Nef-GFP (Welker *et al.*, 1998), several gag-GFP fusions (Sandefur *et al.*, 1998; Sherer *et al.*, 2003) and *E.coli* *lex A* repressor fused to integrase (Holmes-Son and Chow, 2002). The work presented here relies on the insertion of a foreign protein in gag by fusion with the P2 protein. This is a 14 amino acid protein that separates capsid (CA) and nucleocapsid (NC). During maturation, the viral protease (Pro) cleaves gag and gag-pol generating the structural proteins matrix (MA), CA, NC and P6 that assemble and form the virion particle, also producing the spacer proteins P1 and P2. It has been hypothesised that P2 has a role in the regulation of the proteolytic process. Nevertheless, the deletion of this protein had little effect on particle release (Krausslich *et al.*, 1995; Pettit *et al.*, 1994). Also, when 5 amino acids of the P2 protein were deleted (aa 6-10), virion particle production was similar to the wild type virus although viral replication was delayed (Accola *et al.*, 1998).

7.2 – Production of vectors for delivery of proteins

In order to maximise the number of foreign protein molecules within the virion, the strategy adopted here was to incorporate the foreign protein within gag, fusing it in the middle of the viral protein P2 (Figure 7.1 – P2mut). The hypothesis is that during maturation, the viral protease will cleave gag releasing the foreign protein inside the virion. However, it has been proposed that the P2 protein is necessary for the correct maturation of the virion particles. Therefore, a different strategy was also planned in order to introduce the foreign protein in gag but leaving an intact P2 protein. In this approach, the amino acid sequence necessary for the recognition and cleavage of gag in the boundary of P2 and NC proteins was duplicated in the P2-NC region (Figure 7.1 - P2ins) and the foreign protein will then be incorporated between P2 and NC. During maturation, the viral protease may recognise and cleave gag between CA and P2 and the junctions P2-X and X-NC will potentially be recognised and cleaved as a normal P2-NC boundary, releasing the foreign protein (X) and the native P2 protein. Therefore, in the P2mut vector, the foreign protein will carry the extra amino acids AEAMSS in the c-terminus and GTNPATIM in the n-terminus resulting from the fusion with the P2 protein. In the P2ins vector, the cleaved protein will carry IQNGNS in the c-terminus and GPATIM in the n-terminus.

The packaging plasmid of the lentiviral vector was then mutated to facilitate the cloning of the foreign protein. Since there was no *BspE* I restriction site within the plasmid, amino acids 6 and 7 of the P2 protein were mutated from valine and glutamine to serine and glycine to produce a unique restriction sequence within gag in the region of P2, creating the packaging plasmid P2mut. For the second strategy, several sequential rounds of site directed mutagenesis were done in the packaging plasmid to introduce the required sequences in gag to create the packaging plasmid P2ins. The packaging plasmids were then sequenced to confirm the correct mutations or nucleotide insertions.

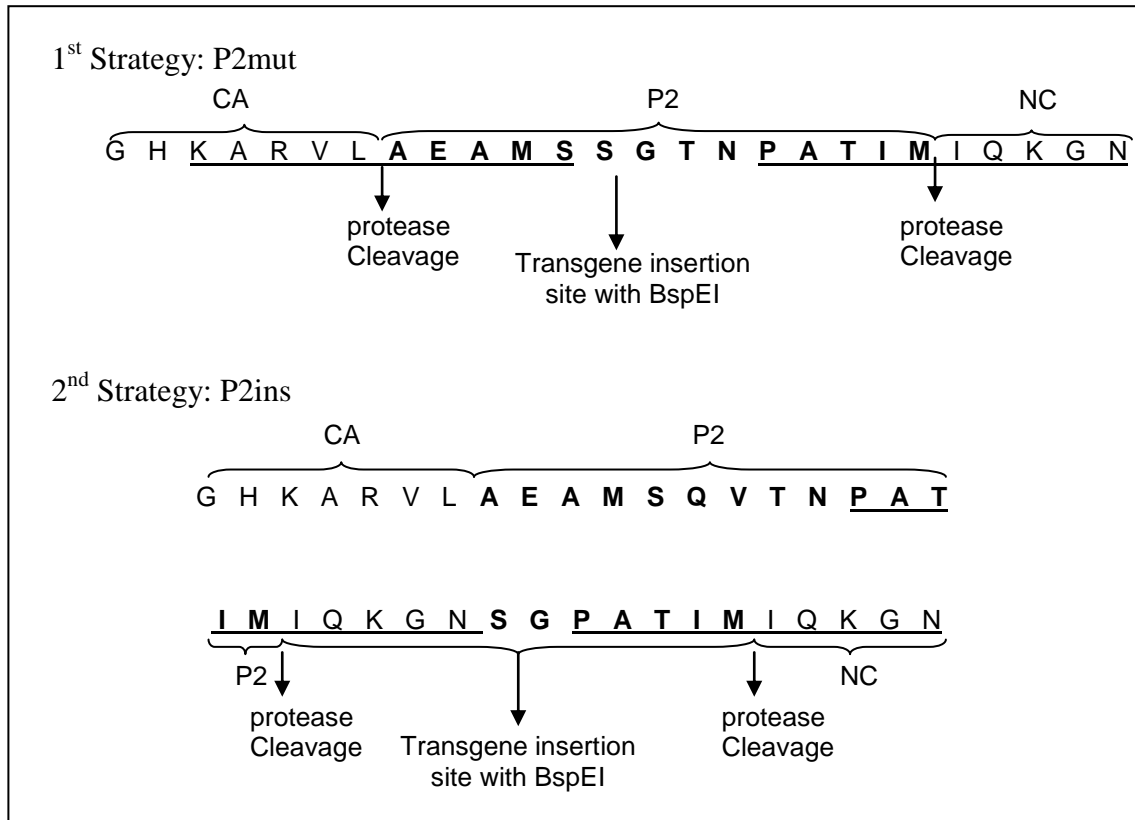


Figure 7.1: Schematic representation of the strategy adopted to insert a foreign protein in the P2 region of gag.

Lentiviral vectors were produced using the native P2 (P2wt), P2mut and P2ins packaging plasmids and an expression cassette containing eGFP being driven by the p5 promoter. Vectors carrying normal and mutated forms of gag were also produced by transfecting the produced cells with a 1:1 ratio of normal packaging plasmid P2wt, and the P2mut or P2ins packaging plasmids. Production was assessed by measuring the concentration of RT in the vector preparations. Infectivity was measured by infecting 293T cells with 50ng_(RT) of vectors and assessing the percentage of eGFP positive cells by flow cytometry 3 days after transduction.

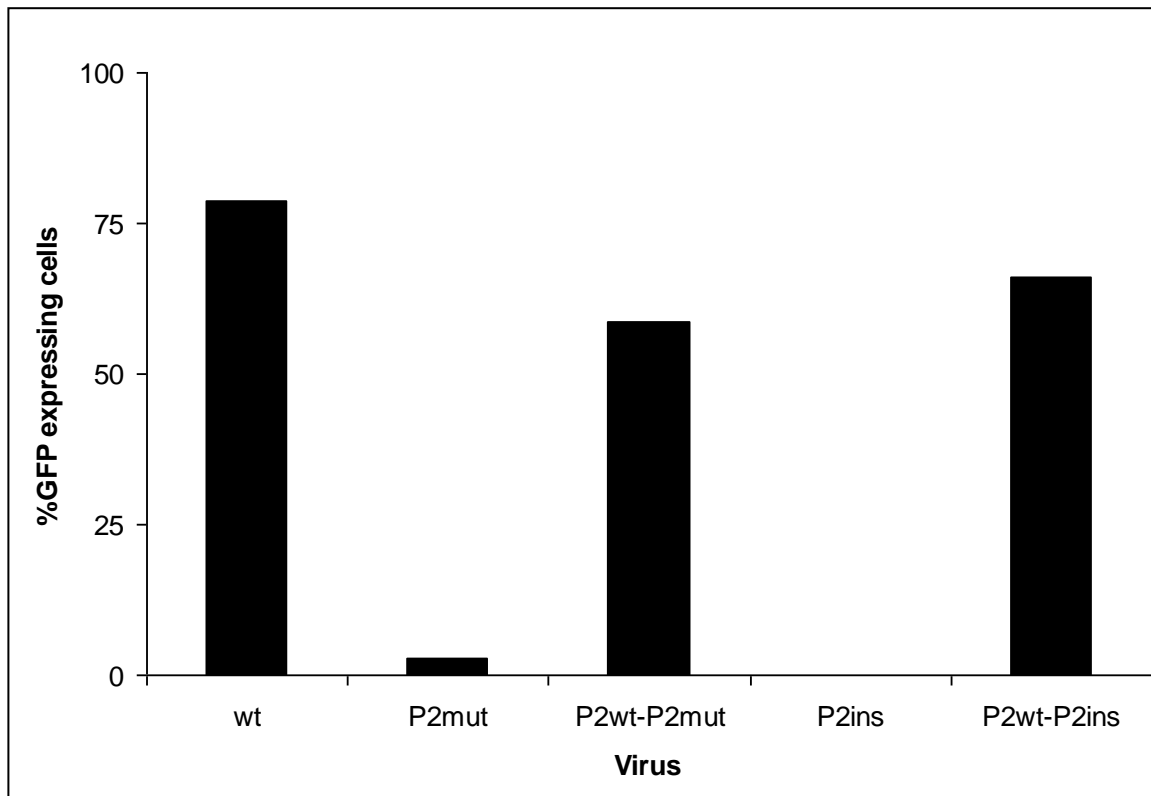


Figure 7.2: Infectivity of vectors harbouring mutations in the P2 region of gag. The P2 region of gag was mutated in order to introduce a unique restriction site (P2mut) or to introduce a second P2/NC protease cleavage site (P2ins). Vectors were produced with these packaging plasmids or with a combination of the packaging plasmids expressing a native gag protein, in a ratio 1:1. All vector expressed eGFP from the p5 promoter. 293T cells were infected with the same RT concentration of each vector ($50\text{ng}_{(\text{RT})}$). The percentage of cells expressing eGFP was assessed by flow cytometry 3 days after infection.

All of the vectors produced with P2mut or P2ins (or with a combination with P2wt) had similar levels of RT (between $4.5 - 7.2\text{ng}_{(\text{RT})}/\mu\text{L}$), which were 10 fold lower compared with an integrating vector produced with the P2wt, implicating that the vectors could be produced, although at a lower concentration. Figure 7.2 demonstrates that all the vectors are able to infect 293T cells, but with different efficiencies. When the vectors contain only the mutant gag protein, the percentage of eGFP expressing cells was much lower than when the vector contained the native gag protein. The percentage of eGFP positive cells was 79% in cells infected with the P2wt vector and only 2.7% and 0.1% in cells infected

with the P2mut and P2ins vectors, respectively. Nevertheless, infectivity of vectors produced from cells transfected with a mixture of native gag and mutated gag was almost restored to P2wt levels, as cells infected with these vectors were 59-66% eGFP positive. In this experiment, no major differences were discovered between the two different approaches to incorporate a foreign protein into the lentiviral vectors. Therefore, it was decided to carry the study with only one of the developed vectors. The P2ins vector was chosen as the P2 protein would still be contained in the virion. eGFP was then cloned in the restriction site of P2ins, by PCR cloning. The atg start codon was introduced just after serine, and the stop codon of eGFP was mutated to glycine.

In order to determine if the eGFP protein was processed correctly within the viral particles, the vectors P2insGFP and P2wt-P2insGFP (proportion of 1:1 between native gag and eGFP fused to gag) were produced and a western blot was done to the viral particles to detect eGFP. Total protein in the viral stocks was determined and similar quantities of protein were loaded in the gel. As a positive control, cells stably infected with a lentivector expressing eGFP were also contained in the western blot.

Figure 7.3 demonstrates that the processed eGFP protein can be detected in both vectors. As expected, the eGFP protein contained in the viral particles was slightly heavier than eGFP produced in the cells. This shift in the eGFP weight was due to the extra 10 amino acids that the protein now carries as result of the fusion with gag. The eGFP protein in the produced vectors is well processed, as there are little heavier eGFP containing bands (unprocessed gag containing eGFP) being detected. Also, it is possible to see that the vector containing only gag-eGFP has more eGFP protein compared with the vector carrying a mixture of both native gag and gag-eGFP fusion.

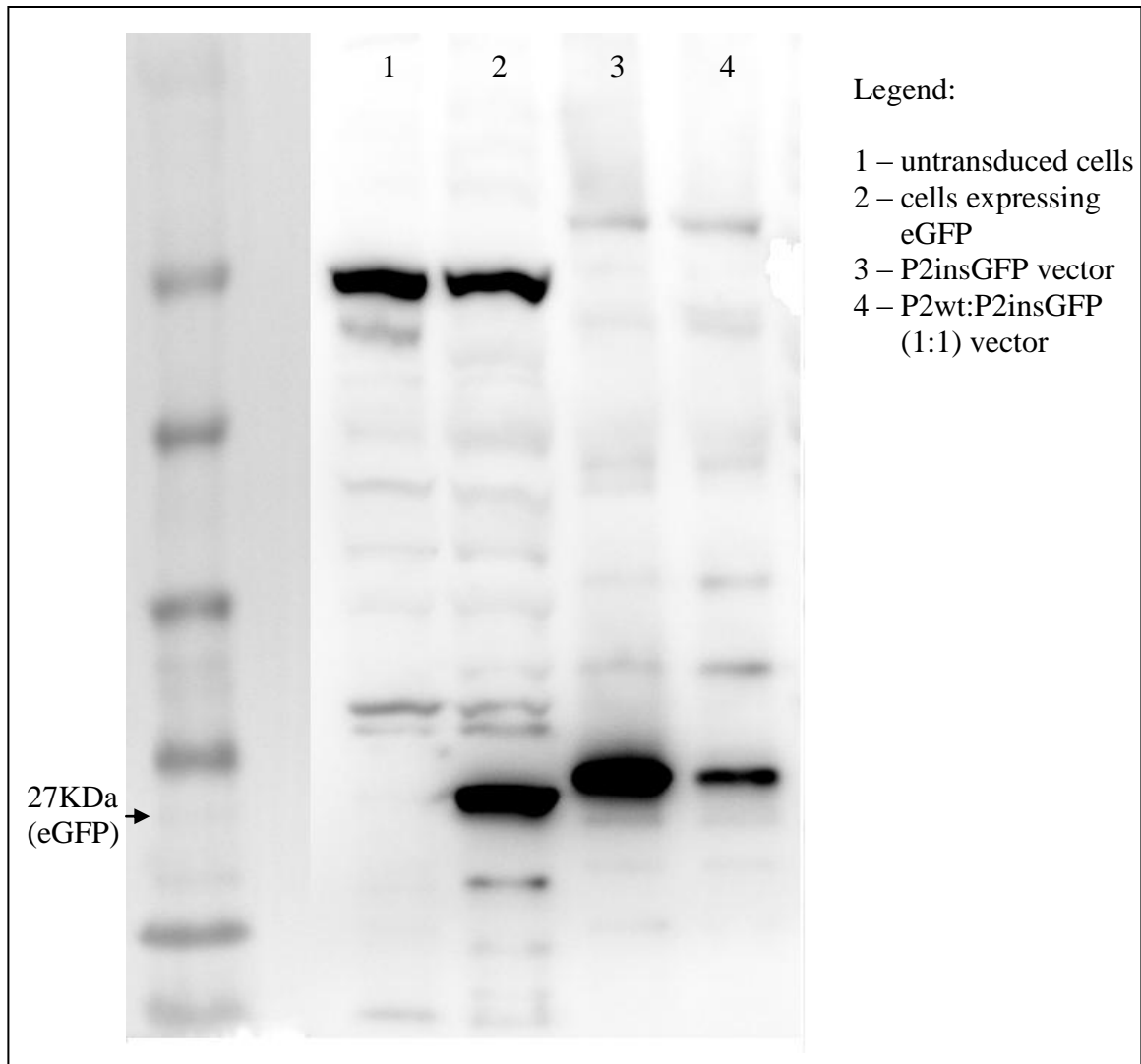


Figure 7.3: Western blot of lentiviral vectors carrying eGFP protein

The eGFP sequence was introduced in the packaging plasmid P2ins in the *BspE* I restriction site. Vectors were produced with this packaging plasmid or with a mixture containing the P2ins and P2wt packaging plasmids in a proportion 1:1. A western blot was then performed to 10 μ L of concentrated viral lysates (50ng_(RT)) and 12 μ g of total protein extracted from cells expressing eGFP. eGFP was detected using an antibody that binds to that protein.

7.2 – Protein delivery to cells by lentiviral vectors

It was established that it is possible to incorporate a foreign protein within gag, by fusion with the P2 protein, with good protease processing. Also, the two strategies designed to fuse the foreign protein with P2 were similar, so it was decided to use only one of the developed packaging plasmids (P2ins) to test protein delivery.

Vectors were then produced with DsRed as a transgene and tested *in vitro* in 293T cells to study protein delivery to cells. Since the fluorescent protein eGFP would be transported in the viral particle, a pilot experiment was performed to determine at which time after infection it would be best to assess protein delivery. Hence, 293T cells were infected with 10 μ L of concentrated vectors and harvested at several time points to assess eGFP positive cells by flow cytometry.

Figure 7.4 shows that it was possible to detect green cells as early as six hours, but the percentage of eGFP positive cells diminishes afterwards, for the vector produced with a mixture of the native and mutated gag protein - P2wt:P2insGFP (in a 1:1 proportion). Conversely, DsRed positive cells were only detected 24h after transduction, and the percentage of DsRed positive cells increased in the 48h time point. Therefore, the best time to assess protein delivery is less than 24h after infection. However, to test infectivity, it is better to wait at least 48h following transduction with the vectors.

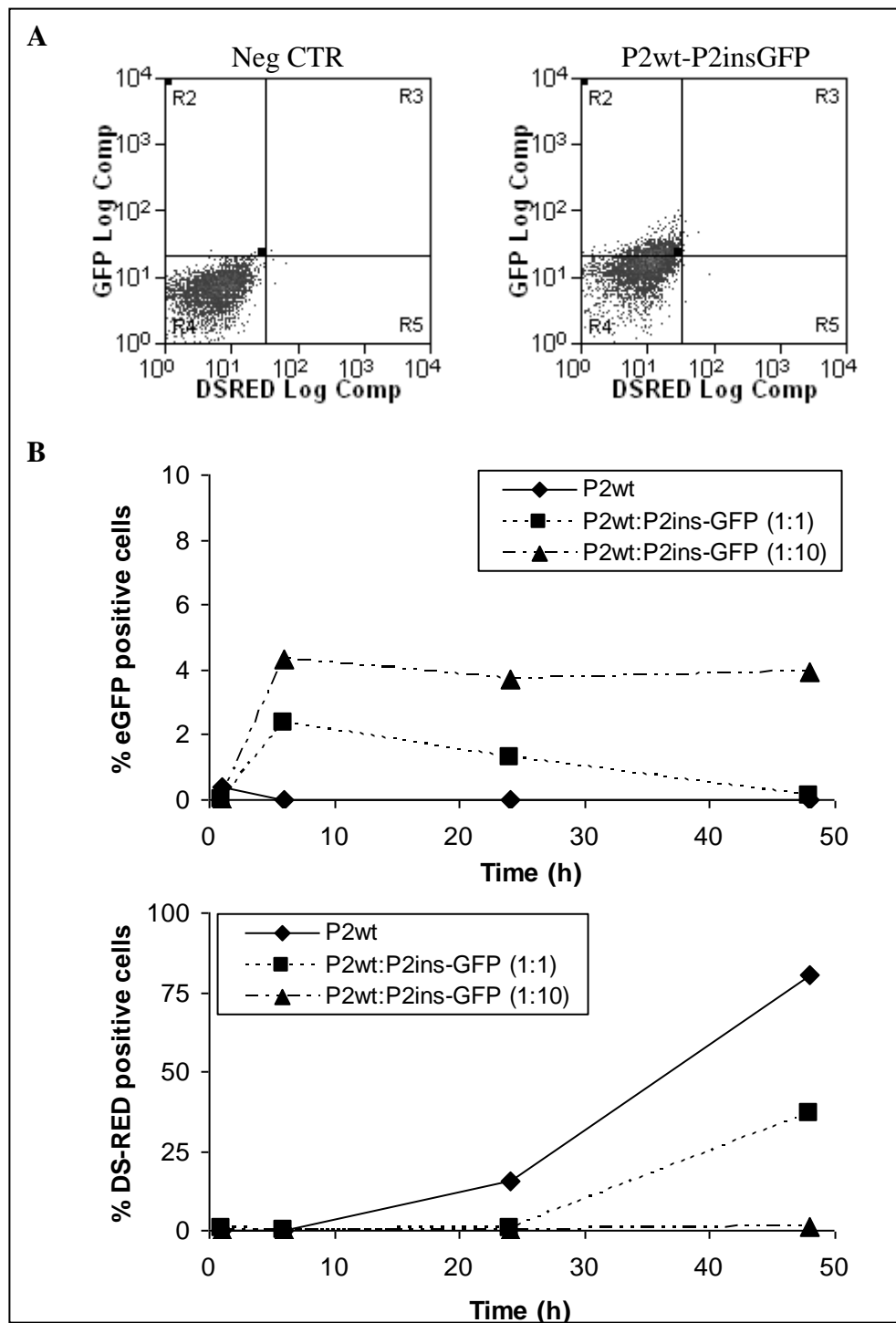


Figure 7.4: Protein delivery to cells from vectors carrying eGFP
 293T cells were infected with 50ng_(RT) of vectors carrying eGFP and expressing DsRed as a transgene. The percentage of eGFP and DsRed positive cells was assessed by flow cytometry at several time points after infection. Graphs with the data obtained over time were generated (**B**). A dot blot of a sample analysed is shown as an example (**A**)

In order to study if the vector is able to achieve protein delivery into the cell cytoplasm, an experiment was designed to visualise cells infected with vector carrying eGFP fused to gag. 293T cells were infected with 50ng_(RT) of P2insGFP and P2wt:P2insGFP vectors expressing Ds-Red as a transgene. Twenty hours after infection, cells were fixed with PFA, washed and stained with phalloidin. This peptide binds actin and can be conjugated with fluorescent analogs permitting the visualisation of actin in the cells by confocal microscopy. Actin is a protein found in the cytoplasm of cells. Therefore, actin was used as a marker to delimit the cytoplasm of cells in confocal microscopy. The infected cells were then visualised in a confocal microscope.

Figure 7.5 shows that it was possible to detect green spots, with a fluorescence intensity well above the background. These spots may correspond to viral particles, viral aggregates or VSVg vesicles carrying the eGFP protein. The vectors are probably able to attach the cells but internalisation may be poor as the majority of green spots were detected outside the cells. Green dots were seen inside cells but, this event was rare. It is possible that the concentration of eGFP delivered to cells is not sufficient to allow efficient detection of the protein in the cytoplasm of target cells.

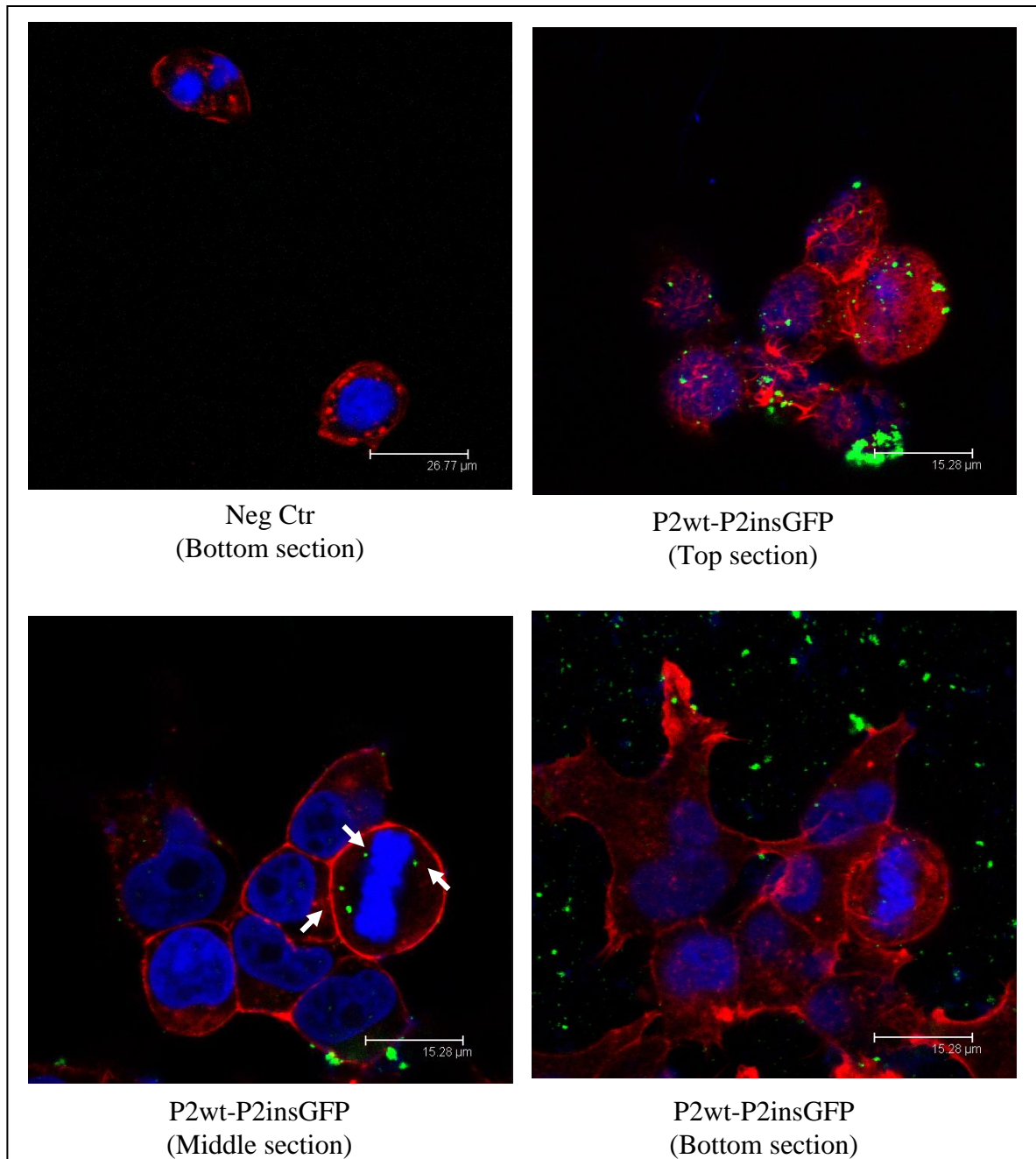


Figure 7.5: Visualisation of cells infected with a lentiviral vector carrying eGFP protein. 293T cells were infected with 50ng_(RT) of lentiviral vectors produced with a packaging plasmid containing eGFP protein inserted in the P2 region of the gag protein and a plasmid with wt gag protein, in a proportion 1:1 (P2wt-P2insGFP). Twenty hours after infection, the cells were fixed. β -actin was stained with phalloidin conjugated with Alexa Fluor 488 and the genetic material with Dapi. The cells were then visualised using a confocal microscope. An example of pictures taken at different z sections (height) of cells infected with the P2wt-P2insGFP vector and a picture taken to untransduced cells are shown. In the image, red represents β -actin, blue the DNA, green the eGFP and arrows indicate intracellular eGFP.

7.3 – Discussion

The aim of the work presented in this chapter was the improvement of vectors for gene repair. Gene repair can be accomplished in cells by homologous recombination, a rare event in cells that can be increased by introducing double strand breaks in the target gene. Recent reports have shown that it is possible to increase gene repair by delivery to the cells with a donor sequence and introducing a specific DSB using a zinc finger nuclease or a homing endonuclease. Nevertheless, both donor sequence and enzyme to increase HR have been delivered to cells in different vectors, thus relying on chance for the cells to be infected with all the vectors necessary to achieve gene repair. Here, a vector was engineered to achieve protein delivery. The developed vector would deliver a foreign protein within the viral particle, and the vector genome would be able to supply the donor sequence for homologous recombination. To incorporate the protein into the viral particle, there have been several reports that fuse foreign proteins with Vpr and other structural proteins of the virions. This is a novel vector where a foreign protein would be fused to gag in the P2 region.

Two approaches were conceived to achieve the goal. In one, the foreign protein would be fused in the middle of the viral P2 protein. In this case, the P2 protein would be disrupted and the foreign protein would be released in the virion upon cleavage by the viral protease in the maturation stage of the vector production. It has been shown that P2 is involved in correct maturation of the virus particle. However, deletion of 6 amino acids was well tolerated, and virus produced with this deletion had only delayed infectivity. In a second approach, the foreign protein was incorporated also in the P2 region, without disruption of this small protein. The protease cleavage site between P2 and NC was duplicated, and the foreign protein was introduced in gag between these protease cleavage sites. Therefore, during maturation, the foreign protein would be released by cleavage of gag, but an intact P2 protein would be present in the virion for correct maturation of the viral particle.

It was shown that when the P2 protein was mutated in order to introduce a unique restriction site, to easily clone the foreign protein in gag, the infectivity of the vector was

diminished (Figure 7.2). Also, when the P2-NC protease cleavage site was duplicated and the unique restriction site was introduced between the duplicated cleavage sites, the vector lost infectivity. Nevertheless, infectivity was restored almost to wild type levels when the production of vector was made with a mixture of wt gag and mutated gag. Since both approaches had similar results, it was opted to continue the study with only one of the developed vectors. The P2ins vector was chosen, as there are no deletions of viral proteins, minimising possible problems in maturation related to the lack of P2 protein in the virion.

eGFP was then fused as a model to study protein delivery with this vector. It was demonstrated that the eGFP is well processed as almost all the eGFP contained in the virions was cleaved (Figure 7.3). It was seen that when a mixture of wt gag and gag-eGFP fusion was used to produce the vectors, the quantity of eGFP in the virions was lower compared when virions were produced with only gag-eGFP fusion.

Protein delivery was then assessed by infecting 293T cells with vectors carrying eGFP protein in the viral particle. It was determined that the best time to assess protein delivery would be less than 24h after infection, as after this time point, the percentage of eGFP positive cells diminished (Figure 7.4). Protein delivery was studied in more detail, to determine if the vectors were able to efficiently deliver eGFP into the cytoplasm of cells. Cells were infected with viral particles carrying eGFP and visualised under a confocal microscope. The eGFP protein was easily detected, as the fluorescence level was well above the background. It was hypothesised that the fluorescence spots may correspond to vectors carrying the eGFP protein. Most of the eGFP spots were detected outside the cells, and only in some rare events, spots were detected in the cytoplasm. It is possible that the concentration of eGFP delivered to cells is not sufficient to distinguish fluorescent cells from background, so the result is inconclusive. However, it supports the hypothesis that the concentration of protein delivered to cells is not very high. Therefore, these vectors may not be efficient enough to be used as tools for gene repair. The measurement of gene repair efficiency with this system may elucidate this question.

Chapter Eight

Final Discussion

8 – Final Discussion

Gene therapy is a technique where a DNA sequence is introduced into a cell by the use of a vector. There have been enormous advances in the development vectors to optimise gene delivery. Vectors based on adenoviruses, adeno-associated viruses (AAV), retroviruses and lentiviruses among others, have been used to exploit the natural properties of infection. It has been demonstrated that gene delivery by such vectors is high and efficient, although with different levels of success, depending on the models used. Recently, results from clinical trials were reported, demonstrating efficient disease correction by gene therapy (Cavazzana-Calvo *et al.*, 2000; Gaspar *et al.*, 2004). Nevertheless, it was later reported that one of the patients had developed a monoclonal lymphoproliferative disease and since then, other patients have developed similar forms of leukaemia (Hacein-Bey-Abina *et al.*, 2003a). The T-cell clones from the patient were analysed revealing the integration of the vector in the proximity of the LMO2 proto-oncogene promoter, leading to aberrant transcription and translation of the gene product (Hacein-Bey-Abina *et al.*, 2003b). The leukaemias in this trial represented the first reports of insertional mutagenesis in humans by gene therapy vectors and prompted a detailed research to study how integration by viral vectors occurs. It has been shown that lentiviral vectors integrate preferentially within highly expressed genes, and gammaretroviruses integrate preferentially near transcription start sites (Mitchell *et al.*, 2004; Schroder *et al.*, 2002; Wu *et al.*, 2003), demonstrating that the integration preferences of these vectors are different and thus, the safety profile may also differ.

Integration is an important feature of retroviral and lentiviral vectors, enabling transgene expression in dividing cells infected with the vectors. Nevertheless, when only transient expression in dividing cells or stable expression in non-dividing cells is required, integration from the vectors may not be necessary. Therefore, the development of non-integrating lentiviral vectors may prove to be an important and safer tool for gene therapy.

In a previous report, mutations in the viral HIV integrase produced replication defective viruses. Although the mutant virus production was similar to the wild type virus, the viral gene expression was impaired (Wiskerchen and Muesing, 1995). Furthermore, the

development of non-integrating lentiviral vectors (NILVs) by the mutation of the viral integrase in first generation lentiviral vectors demonstrated that normal levels of viral DNA could be detected after infection of target cells (Leavitt *et al.*, 1996). However, expression from these vectors was negligible (Blomer *et al.*, 1997).

A major improvement towards the safety of vectors for gene therapy was the development of self inactivating (SIN) vectors (Zufferey *et al.*, 1998). In these vectors, a large part of the U3 region in the 3'LTR was deleted, abolishing the viral promoter activity and allowing transgene expression to be controlled by the incorporation of an internal promoter. Recently, it was demonstrated that the viral DNA generated by non-integrating lentiviral vectors using the SIN version is capable of efficient expression (Yanez-Munoz *et al.*, 2006). Another report demonstrated that use of non-SIN or SIN vectors is the reason for the discrepancy in expression from these non-integrating vectors, possibly due to a sequence element within the U3 region of the viral LTR that reduces expression (Bayer *et al.*, 2008).

The aim of the work presented in this thesis was to further develop the research in this new field of viral vectors for gene therapy. In order to produce NILVs, there are two possible approaches. The enzyme responsible for integration can be mutated, or the viral sequence for integrase recognition can be altered. In this context, the combination of several mutations may improve the safety of the developed vectors. Several mutations in integrase have been published in the literature where, although integration of the vector is impaired, infectivity is not affected. Any mutation in the catalytic motif of integrase (DD35E) impairs integration (Leavitt *et al.*, 1996). The vector used by Yanez-Munoz and colleagues to demonstrate that NILVs are able to sustain expression in cells has a mutation in the D64 residue. Other mutations have been suggested to affect specific steps of integration and were therefore chosen to be studied here. N120L and W235E mutations do not impair integration *in vitro*. Vectors harbouring these mutations are not able to complete integration in cells upon infection, but it was postulated that these residues are related to host DNA recognition (Leavitt *et al.*, 1996). The mutation Q148A blocks integration and has been implicated in the recognition of viral DNA (Esposito and Craigie, 1998; Gerton *et al.*, 1998). Also, the mutations K264,266,273R have been shown to reduce integrase activity due to the lack of acetylation on those residues. Additionally, the mutation of the

conserved dinucleotides CA (attachment sites) to TG in the vector genome has been shown to impair integration (Masuda *et al.*, 1998).

All of the mutations detailed were introduced in the lentiviral vectors by site directed mutagenesis. Vector production was then assessed by the measurement of viral particles (by quantifying the concentration of p24 protein) and infectivity by the measurement of transducing units. All vectors produced had similar concentrations of p24, indicating that vector production was not affected, as previously published. High titres were achieved for all mutant vectors, demonstrating that infectivity was also not altered. Dividing cells were then infected with the vectors and expression was assessed over time. If the viral DNA does not become integrated, the DNA is lost over time due to dilution or degradation of the DNA and thus, expression is transient. Expression from all the vectors was detected in 293T cells at an early time point but it decreased afterwards for all NILVs. In contrast, the cells infected with the integrating vector maintained expression throughout the course of the experiment. These results suggested that the mutations introduced impaired viral integration.

It was necessary to assess integration in more detail, so an assay was established to measure background integration enabling the quantification of integration and the comparison with the integrating vector. This assay relied on the expression of a transgene that confers resistance to G148. The dividing cells that have the vector integrated in the genome stably express the transgene and survive over time, but in contrast, cells that lost the viral DNA would die. The results from this experiment show that the background integration was different among the vectors. The D64V, N120L and W235E vectors exhibit the lowest level of integration, being 10^3 fold lower compared with the integrating vector. The Q148A, K264R and K264,266,273R vectors integrated the viral DNA 13, 3 and 14 fold lower, respectively, than the integrating vector. The vector carrying the att mutation had a 100 fold difference in integration compared with the integrating vector. All these results are in accordance with previous studies (Cereseto *et al.*, 2005; Esposito and Craigie, 1998; Gerton *et al.*, 1998; Leavitt *et al.*, 1996; Masuda *et al.*, 1998). Therefore, it is possible to conclude that the produced vectors are integration defective. These results correlate well with the expression data collected over time from infected cells. Transgene expression in cells infected with the low integrating NILVs D64V, N120L and W235E

could not be distinguished from background 15 days after infection. However, cells infected with the vectors Q148A, K264R and K264,266,273R were still expressing the transgene at that time point, although the percentage of expressing cells was lower compared with cells infected with the integrating vector.

The mutations that had the highest impact on integration (D64V, N120L and W235E) were then combined and vectors harbouring the 3 mutations in integrase with or without the mutant att sites were produced. As with the previously tested NILVs, production and infectivity of these vectors was similar to the integrating vector, and 293T cells infected with these vectors expressed the transgene transiently. The quantification of the background integration revealed that the DNW mutant vector integrates on average 1.4×10^3 fold lower than the integrating vector and in combination with the att sites integration is lowered 3×10^3 fold on average. Although this represents a decrease of 1.4 to 3 fold of the background integration compared with the single mutant vectors, it is not statistically different, so it is not possible to infer that the combination of the mutations has a synergetic or additive effect on integration impairment. This is in accordance with a previous study where the mutant att sites were combined with the D64V integrase mutation (Nightingale *et al.*, 2006). These multiple mutant vectors present an advantage in safety, compared with single mutant NILVs. Mutations may occur in the packaging plasmid during viral production or plasmid preparation. One of such mutations may restore the enzymatic activity of integrase. Therefore, the multiple mutation combination may reduce the potential for reversion back to an integrative phenotype. Also, the mutation in the att site has the advantage to prevent rescue of the vector if the target cell becomes infected with a wild type virus.

Viral DNA accumulation in cells infected with the lentivectors was also studied. It was shown that viral DNA could be detected as early as three hours post infection and that the formation of viral DNA circles occurs after that for both integrating and NILVs. Total viral DNA was quantified in cells, 48h after infection. Results demonstrate that the concentration of viral DNA is similar for all integrating and non-integrating vectors, so it is possible to conclude that the studied integrase defective vectors do not have an impact on reverse transcription.

It was shown that NILVs are able to deliver viral DNA to target cells but expression is transient. There are several applications for such vectors. An example is the temporary expression of HoxB4 in haematopoietic stem cells. This protein has been shown to increase self-renewal and promote expansion of these cells (Antonchuk *et al.*, 2002). However, the permanent production of such protein can have adverse effects on the cells (Zhang *et al.*, 2007b; Krosi *et al.*, 1998). Due to the capacity of lentiviruses to infect such cells, NILVs would make very good vector candidates to accomplish this task. Another interesting application of such vectors is vaccination. Lentiviral vectors efficiently stimulate cell-mediated and humoral immunity, following subcutaneous and intramuscular immunization. Indeed, NILVs have been shown to be efficient immunogens with long-lasting immune responses (Karwacz *et al.*, 2009; Negri *et al.*, 2007). If prolonged transgene expression is required, the NILVs could also be modified. One interesting development would be the incorporation of a replication signal or a scaffold/matrix associating region (S/MAR). In such a vector, the risk of insertional mutagenesis would not be present but the transgene would be stably expressed both in dividing and non-dividing cells. It has been previously shown that it is possible to incorporate the simian virus 40 (SV40) oriT into NILVs allowing the maintenance of the episomal DNA in dividing cells, without reversion back to an integrative phenotype. Replication and expression from that vector was sustained for up to 56 days in cells that express the T-antigen, but in cells that do not express the T-antigen, stable expression was not seen (Lu *et al.*, 2004; Vargas, Jr. *et al.*, 2008). Nonetheless, these vectors would carry a risk of mutagenesis due to the presence of the SV40 oriT. Up until now, no NILV incorporating an S/MAR that is retained in dividing cells has been reported. NILVs are able to deliver DNA to the cells without integration into the genome, so a different strategy for stable expression in dividing cells would be the integration of the vector genome in a safer way, such as the use of a different enzyme to mediate this process. One example is the Rep protein from AAV. This protein can site-specifically integrate DNA substrates into the human genome, including circular DNA (Kotin *et al.*, 1990). Therefore, an interesting development would be the combination of this enzyme with the delivery of the target DNA mediated by NILVs. Other examples include the phage Φ C31 integrase, the Cre and Fip recombinases and the Sleeping Beauty transposase. The DNA delivered by NILVs may

also be used as a template for homologous recombination. This strategy has been used by combining several NILVs to deliver the DNA template and enzymes to enhance homologous recombination demonstrating the potential of these vectors in gene repair (Cornu and Cathomen, 2007; Lombardo *et al.*, 2007). However, if the target cells are not dividing then the episomal DNA may be maintained in the cells. Indeed, NILVs have been previously tested in retina and brain tissues. It was demonstrated that cells in these post-mitotic tissues infected with NILVs display stable and prolonged transgene expression (Philippe *et al.*, 2006; Yanez-Munoz *et al.*, 2006). Here, it was decided to extend this work and study the application of NILVs for gene delivery to muscle cells.

The developed NILVs were first tested in C2C12 muscle cells *in vitro*. In contrast with dividing cells, differentiated C2C12 cells maintained expression over time, demonstrating that these vectors can stably express a transgene in non-dividing cells. The vectors were then tested *in vivo*, by injecting mice in the tibialis anterior muscle. It was demonstrated that the vectors are able to infect muscle cells *in vivo*. Most importantly, it was shown that the vectors are able to sustain transgene expression over time for up to eight months post injection. The quantification of transgene expression revealed that all vectors expressed the transgene at high levels (0.01-0.1% of total protein). The quantification of viral DNA in non-dividing cells revealed that the episomal vectors are not degraded over time and also that 2LTR circles levels were higher for NILVs, compared with the integrating vector. These results were well documented both *in vitro* and *in vivo*. This may be explained by assuming that the newly transcribed viral DNA, which does not serve as a template for integration, is therefore more available for circularisation. This result also indicates that the mutations in the vectors did not impair the viral DNA nuclear entry.

NILVs were then tested in a disease model. The model chosen was haemophilia B, due to the low immunogenicity of the vectors and the requirement of a relatively low protein production for disease amelioration. Also, it has been proven that it is possible to achieve disease amelioration by the production of FIX from muscle cells (Herzog *et al.*, 1997). Therefore, the aim of this work was to evaluate disease correction of haemophiliac mice by integrating and non-integrating lentiviral vectors injected intramuscularly.

It was demonstrated that it is possible to quantify FIX from the supernatant of infected 293T and C2C12 cells and also in the blood of injected mice. However, the concentration of circulating FIX obtained with the injection of the integrating vector in mice was low. It was then necessary to improve vector expression. To achieve that, a codon optimised version of hFIX was used. Also, this cDNA was mutated in order to decrease the binding of the secreted transgene to collagen IV. The results demonstrate that the mutations introduced to FIX increased the concentration of plasma FIX, but it was not clear if the codon optimisation had a positive effect on transgene expression. In mice injected with NILVs, the values of plasma FIX remained very low even using the best expressed and secreted transgene with the highest vector concentration. However, the levels of hFIX produced in mice infected with the integrating lentiviral vector were always above the therapeutic threshold, indicating that gene therapy with this vector may achieve disease amelioration. Muscle cells may not be very permissive towards lentiviral vectors pseudotyped with VSVg. An indication for this was the necessity to infect C2C12 cells with MOI 10 to obtain 30% of transgene expressing cells. Therefore, the use of a more suitable envelope may increase vector infectivity thus enhancing protein production from these cells. Another possibility to improve protein production would be the repetition of vector injection. Therefore, integrating lentiviral vectors may be useful tools for the treatment of haemophilia B.

Taking all the results obtained *in vivo* together, it was possible to conclude that the integrating vector expresses more transgene than the NILVs. This prompted a new project to study the reason for lower expression from NILVs compared with the integrating lentiviral vectors. The expression obtained from integrating and non-integrating vectors using different promoters to drive transgene expression was evaluated in 293T cells. The promoters chosen for this study were the strong viral promoter SFFV, the weak p5 promoter from the natural episomal virus AAV, and the UCOE element, an enhancerless promoter involved in chromatin organisation. The results obtained show that, regardless of the promoter used, NILVs express less transgene than the integrating vector. Nevertheless, the differences in expression were different. There was a 3.4 fold difference between integrating and NILVs using the SFFV promoter and only a 1.7 fold difference with the p5 promoter. As the first is derived from an integrating virus and the second from an episomal

virus, a hypothesis to explain the difference is that the p5 promoter may be less susceptible to influences from nearby sequences if the vector is integrated in the host genome, in contrast to the SFFV promoter. On the other side, it is possible to assume that the p5 promoter may be less affected than the SFFV promoter if the DNA is episomal and not in an optimal location for expression.

In order to investigate the reason for the lower expression from NILVs, DNA, RNA and transgene produced in cells were measured and compared with the integrating vector. It was demonstrated that when similar concentrations of DNA are delivered to cells, the levels of RNA produced from the episomal NILVs were lower compared with the concentrations obtained for the integrating counterpart. This lower RNA concentration related to the lower protein concentration being produced. Then, it is possible to conclude that NILVs display lower transcription activity compared with integrating lentiviral vectors.

This study was further developed by comparing expression from integrating and NILVs with AAV vectors and transfected DNA. Since these systems are very divergent, a comparison is difficult. The experiment had to be designed in order to take into account several differences. The lentivector genome is RNA and needs to be reverse transcribed upon cell entry whereas the AAV vector delivers a ssDNA that, upon a slow process of second-strand synthesis, forms a transcriptionally active DNA molecule. Both vectors infect dividing and non-dividing cells, but a plasmid DNA molecule is not efficiently transfected in terminally differentiated cells. Also, target cells can be more permissive towards one of the vector systems. Another difference between the viral vectors is the need of particular *cis* sequences. The AAV vector only requires the inverted terminal repeats (ITRs). Apart from the long terminal repeats (LTRs), packaging signal, PPT and the central polypurine tract (cPPT), lentivectors also use a sequence (WPRE) to export unspliced RNA to the cell cytoplasm. This element is used to increase the vector titre and is located between the transgene and the 5'LTR. Therefore, the SFFV promoter, eGFP and the WPRE sequence were cloned into an AAV vector. The plasmid chosen to be transfected was the lentivector so that all vectors had the same expression cassette.

Expression was compared in 293T cells, as these dividing cells are permissive for all the vector systems. Cells were transfected or infected with serial dilutions of the

vectors. Two days after, expression and vector copy number were assessed. Data shows that, when similar quantities of DNA are delivered into cells, both episomal viral vectors AAV and NILV had similar levels of expression, but 4-5 fold lower compared with the integrating lentivector. Cells transfected with a plasmid had a 2 fold higher expression compared with the integrating lentiviral vector. However, plasmid transfection delivers a very large quantity of DNA to cells, so the comparison may not be appropriate.

The expression was also compared between integrating and episomal viral vectors in C2C12 cells. These cells were analysed 2 days (dividing cells) and 12 days (differentiated cells) after infection in order to compensate for a possible delay of the expression from the AAV vector due to the slow second-strand synthesis. Regardless of the time point, expression from the AAV vector was similar to the NILV and both were lower compared with expression from the integrating lentiviral vector. The viral DNA concentration in cells, 12 days after infection with the NILV, diminished 14 fold compared to cells tested 2 days after infection. Possibly, this happened due to dilution by some cell division. However, for the AAV vectors, the DNA concentration was only two fold lower in the last time point. If the rate of dilution due to cell division is similar between NILV and AAV vectors, then it is possible to infer that the AAV vector DNA has completed the second-strand process and may even have formed concatemers (or higher order concatemer circles by rolling circle replication). Therefore, the lower expression levels compared with the integrating vector cannot be explained by a lack of second-strand synthesis.

AAV vectors are natural episomal vectors nonetheless, the results demonstrate that these vectors do not possess any advantage in expression over NILVs and even have a lower expression compared with the integrating lentivectors. Therefore, it is possible to conclude that the sequences present in the AAV vector, namely, the ITRs, do not influence expression. The discrepancy in expression between integrating and NILVs may be explained by the influence of sequences nearby the insertion sites of the integrating vectors or the chromatin structure, as it is known that lentiviral vectors tend to integrate preferentially within highly transcribed regions of the genome. It is important to mention that C2C12 cells are more permissive towards AAV vectors than to lentiviral vectors, as it was necessary to infect cells with 5 fold more lentivectors to obtain the same percentage of transgene expressing cells. This result implies that muscle cells may be more permissive

towards AAV vectors, which in turn may explain the lower levels of FIX obtained injecting the lentivectors intramuscularly in mice, compared with the FIX concentration achieved by injection of AAV vectors (Herzog *et al.*, 1997).

A different application of NILVs is the enhancement of gene repair. This strategy relies on homologous recombination (HR) events between similar sequences using the host cell mechanisms. These events are rare, but can be increased by introducing a double strand break (DSB) in the target sequence which can be achieved by zinc finger nucleases or homing endonucleases. As aforementioned, the approaches published until now rely on the transient expression of such proteins using NILVs (Cornu and Cathomen, 2007; Lombardo *et al.*, 2007). Two or more NILVs are necessary to achieve the goal of gene repair enhancement. Therefore, these approaches rely on the infection of the target cell with all the vectors. Here a new approach where only one NILV would be necessary to achieve enhanced HR was presented. This vector was designed to carry a protein to enhance HR inside the virion. Several reports have previously demonstrated that it is possible to incorporate additional proteins within the viral particles, by fusing the foreign protein to a viral protein such as Vpr, Nef, gag or integrase (Holmes-Son and Chow, 2002; McDonald *et al.*, 2002; Sandefur *et al.*, 1998; Sherer *et al.*, 2003; Welker *et al.*, 1998). A new approach was developed where the foreign protein would be fused to the viral P2 protein. Two strategies were designed. In one, the P2 protein was mutated so that a unique restriction site would be introduced in the middle of the P2 sequence facilitating the fusion of the foreign protein to gag. In a second approach, the native P2 protein in gag would be maintained. In this vector, the cleavage site between P2 and NC was duplicated so that the foreign protein could be inserted within these sites. Upon maturation, viral protease would cleave gag, releasing the foreign protein within the virion. The viral vector would supply the DNA template upon reverse transcription of the vector genome, and the protein necessary for HR augmentation (a meganuclease, such as *I-Sce I*) would be carried within the viral particle and delivered in the target cell after infection.

The two developed packaging plasmids were used to produce vectors carrying eGFP as a transgene. Additionally, a mixture of native gag and mutated gag in a proportion of 1:1 were used to produce vectors. 293T cells were infected with the vectors and infectivity was assessed by measuring the percentage of eGFP positive cells. It was

demonstrated that the vectors produced from mutated gag plasmids lost infectivity. Although the concentration of the vector particles produced was similar to the wt levels (assessed by measuring the p24 protein), it is possible that the vectors produced were not able to incorporate the genomic viral RNA, or that another step in the viral life cycle would be impaired, such as reverse transcription or nuclear entry. A report has been published implicating that the P2 region may be important for viral RNA recognition, supporting the first hypothesis (Kaye and Lever, 1998). Nevertheless, the vectors produced with a mixture of wt and mutated gag were almost as infectious as the wt vectors. If the mutated gag proteins are not being specifically excluded from the viral particles, it is possible to assume that these vectors would be able to achieve protein delivery without disrupting infectivity.

Since both strategies had similar results, the P2ins vector was chosen to study protein delivery to cells. In this context, eGFP was inserted in gag. Vectors were produced and protein incorporation in the virions was assessed by western blot. It was shown that the eGFP is incorporated in the virion particles and that the protein is well processed by the viral protease since most of the eGFP present in the vectors is cleaved. Cells infected with the vector transporting the fluorescent protein were eGFP positive (assessed by flow cytometry). In order to study protein delivery in more detail, cells were infected with vectors carrying eGFP and visualised by confocal microscopy. It was shown that green spots, which may correspond to vectors carrying eGFP, remain mostly in the periphery of the cells probably attached to the cellular membrane. It is important to refer that eGFP was detected within the cytoplasm of some cells, but this event was rare. The fact that cells infected with the eGFP carrying vectors were positive by flow cytometry may be explained by the inability of the machine to differentiate between cells that have the protein within the cell or only attached to them. Nevertheless, it is also possible that the concentration of delivered eGFP was not sufficient to be above background in order to allow the visualisation of intracellular eGFP by confocal microscopy. This implies that the concentration of delivered protein is low and thus, these vectors may not deliver enough protein to enhance homologous recombination. It is possible that the developed vectors are just not able to infect cells, so a different strategy could be developed. One approach could be the fusion of the foreign protein in the end of the viral integrase with the introduction of an extra protease cleavage site between IN and the enzyme. In conclusion, the

development of such vectors would potentially create the ideal tool for the treatment of genetic diseases. These vectors would alleviate the risk of mutagenesis correcting mutations regardless of whether the target cells were dividing or not. In addition, the target gene would still be expressed from the endogenous promoter with all the cellular regulatory mechanisms.

References

1. Accola, M.A., S.Hoglund, and H.G.Gottlinger. 1998. A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. *J. Virol.* 72:2072-2078.
2. Adam, M.A., N.Ramesh, A.D.Miller, and W.R.Osborne. 1991. Internal initiation of translation in retroviral vectors carrying picornavirus 5' nontranslated regions. *J. Virol.* 65:4985-4990.
3. Aiuti, A., S.Slavin, M.Aker, F.Ficara, S.Deola, A.Mortellaro, S.Morecki, G.Andolfi, A.Tabucchi, F.Carlucci, E.Marinello, F.Cattaneo, S.Vai, P.Servida, R.Miniero, M.G.Roncarolo, and C.Bordignon. 2002. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* 296:2410-2413.
4. Amalfitano, A., C.R.Begy, and J.S.Chamberlain. 1996. Improved adenovirus packaging cell lines to support the growth of replication-defective gene-delivery vectors. *Proc. Natl. Acad. Sci. U. S. A* 93:3352-3356.
5. Amalfitano, A., and J.S.Chamberlain. 1997. Isolation and characterization of packaging cell lines that coexpress the adenovirus E1, DNA polymerase, and preterminal proteins: implications for gene therapy. *Gene Ther.* 4:258-263.
6. Antonchuk, J., G.Sauvageau, and R.K.Humphries. 2002. HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. *Cell* 109:39-45.
7. Arhel, N., A.Genovesio, K.A.Kim, S.Miko, E.Perret, J.C.Olivo-Marin, S.Shorte, and P.Charneau. 2006. Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nat. Methods* 3:817-824.
8. Armentano, D., J.Zabner, C.Sacks, C.C.Sookdeo, M.P.Smith, J.A.St George, S.C.Wadsworth, A.E.Smith, and R.J.Gregory. 1997. Effect of the E4 region on the persistence of transgene expression from adenovirus vectors. *J. Virol.* 71:2408-2416.
9. Arruda, V.R., J.N.Hagstrom, J.Deitch, T.Heiman-Patterson, R.M.Camire, K.Chu, P.A.Fields, R.W.Herzog, L.B.Couto, P.J.Larson, and K.A.High. 2001. Posttranslational modifications of recombinant myotube-synthesized human factor IX. *Blood* 97:130-138.

10. Arthur, L.O., J.W.Bess, Jr., R.C.Sowder, R.E.Benveniste, D.L.Mann, J.C.Chermann, and L.E.Henderson. 1992. Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. *Science* 258:1935-1938.
11. Atchison, R.W., B.C.Casto, and W.M.Hammon. 1965. ADENOVIRUS-ASSOCIATED DEFECTIVE VIRUS PARTICLES. *Science* 149:754-756.
12. Baltimore, D. 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226:1209-1211.
13. Barboric, M., R.M.Nissen, S.Kanazawa, N.Jabrane-Ferrat, and B.M.Peterlin. 2001. NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* 8:327-337.
14. Barre-Sinoussi, F., J.C.Chermann, F.Rey, M.T.Nugeyre, S.Chamaret, J.Gruet, C.Dauguet, C.Axler-Blin, F.Vezinet-Brun, C.Rouzioux, W.Rozenbaum, and L.Montagnier. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220:868-871.
15. Bartholomew, C., and J.N.Ihle. 1991. Retroviral insertions 90 kilobases proximal to the Evi-1 myeloid transforming gene activate transcription from the normal promoter. *Mol. Cell Biol.* 11:1820-1828.
16. Bayer, M., B.Kantor, A.Cockrell, H.Ma, B.Zeithaml, X.Li, T.McCown, and T.Kafri. 2008. A large U3 deletion causes increased in vivo expression from a nonintegrating lentiviral vector. *Mol. Ther.* 16:1968-1976.
17. Beitzel, B., and F.Bushman. 2003. Construction and analysis of cells lacking the HMGA gene family. *Nucleic Acids Res.* 31:5025-5032.
18. Berlioz, C., and J.L.Darlix. 1995. An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. *J. Virol.* 69:2214-2222.
19. Berns, K.I. 1990. Parvovirus replication. *Microbiol. Rev.* 54:316-329.
20. Bischoff, J.R., D.H.Kirn, A.Williams, C.Heise, S.Horn, M.Muna, L.Ng, J.A.Nye, A.Sampson-Johannes, A.Fattaey, and F.McCormick. 1996. An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. *Science* 274:373-376.

21. Blomer, U., L.Naldini, T.Kafri, D.Trono, I.M.Verma, and F.H.Gage. 1997. Highly efficient and sustained gene transfer in adult neurons with a lentivirus vector. *J. Virol.* 71:6641-6649.
22. Bokhoven, M., S.L.Stephen, S.Knight, E.F.Gevers, I.C.Robinson, Y.Takeuchi, and M.K.Collins. 2009. Insertional gene activation by lentiviral and gammaretroviral vectors. *J. Virol.* 83:283-294.
23. Boshoff, C., and R.Weiss. 2002. AIDS-related malignancies. *Nat. Rev. Cancer* 2:373-382.
24. Braaten, D., E.K.Franke, and J.Luban. 1996. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *J. Virol.* 70:3551-3560.
25. Bradac, J., and E.Hunter. 1986. Polypeptides of Mason-Pfizer monkey virus. II. Synthesis and processing of the env gene products. *Virology* 150:491-502.
26. Bray, M., S.Prasad, J.W.Dubay, E.Hunter, K.T.Jeang, D.Rekosh, and M.L.Hammarskjold. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. U. S. A* 91:1256-1260.
27. Brown, B.D., M.A.Venneri, A.Zingale, S.L.Sergi, and L.Naldini. 2006. Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer. *Nat. Med.* 12:585-591.
28. Brown, P.O., B.Bowerman, H.E.Varmus, and J.M.Bishop. 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc. Natl. Acad. Sci. U. S. A* 86:2525-2529.
29. Brussel, A., and P.Sonigo. 2004. Evidence for gene expression by unintegrated human immunodeficiency virus type 1 DNA species. *J. Virol.* 78:11263-11271.
30. Bukovsky, A.A., T.Dorfman, A.Weimann, and H.G.Gottlinger. 1997. Nef association with human immunodeficiency virus type 1 virions and cleavage by the viral protease. *J. Virol.* 71:1013-1018.
31. Bukrinsky, M. 2004. A hard way to the nucleus. *Mol. Med.* 10:1-5.
32. Bukrinsky, M.I., N.Sharova, T.L.McDonald, T.Pushkarskaya, W.G.Tarpley, and M.Stevenson. 1993. Association of integrase, matrix, and reverse transcriptase

antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. U. S. A* 90:6125-6129.

33. Burke, C.J., G.Sanyal, M.W.Bruner, J.A.Ryan, R.L.LaFemina, H.L.Robbins, A.S.Zeft, C.R.Middaugh, and M.G.Cordingley. 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. *J. Biol. Chem.* 267:9639-9644.
34. Burns, J.C., T.Friedmann, W.Driever, M.Burrascano, and J.K.Yee. 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci. U. S. A* 90:8033-8037.
35. Bushman, F.D. 1994. Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc. Natl. Acad. Sci. U. S. A* 91:9233-9237.
36. Bushman, F.D., and R.Craigie. 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. U. S. A* 88:1339-1343.
37. Bushman, F.D., and M.D.Miller. 1997. Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J. Virol.* 71:458-464.
38. Butler, S.L., M.S.Hansen, and F.D.Bushman. 2001. A quantitative assay for HIV DNA integration in vivo. *Nat. Med.* 7:631-634.
39. Butler, S.L., E.P.Johnson, and F.D.Bushman. 2002. Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J. Virol.* 76:3739-3747.
40. Camaur, D., and D.Trono. 1996. Characterization of human immunodeficiency virus type 1 Vif particle incorporation. *J. Virol.* 70:6106-6111.
41. Cavazzana-Calvo, M., S.Hacein-Bey, B.G.de Saint, F.Gross, E.Yvon, P.Nusbaum, F.Selz, C.Hue, S.Certain, J.L.Casanova, P.Bousso, F.L.Deist, and A.Fischer. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* 288:669-672.
42. Ceci, J.D., C.P.Patriotis, C.Tsatsanis, A.M.Makris, R.Kovatch, D.A.Swing, N.A.Jenkins, P.N.Tsichlis, and N.G.Copeland. 1997. Tpl-2 is an oncogenic kinase that is activated by carboxy-terminal truncation. *Genes Dev.* 11:688-700.

43. Cepko, C.L., B.E.Roberts, and R.C.Mulligan. 1984. Construction and applications of a highly transmissible murine retrovirus shuttle vector. *Cell* 37:1053-1062.
44. Cereseto, A., L.Manganaro, M.I.Gutierrez, M.Terreni, A.Fittipaldi, M.Lusic, A.Marcello, and M.Giacca. 2005. Acetylation of HIV-1 integrase by p300 regulates viral integration. *EMBO J.* 24:3070-3081.
45. Chang, F., S.Syrjanen, and K.Syrjanen. 1995. Implications of the p53 tumor-suppressor gene in clinical oncology. *J. Clin. Oncol.* 13:1009-1022.
46. Charneau, P., and F.Clavel. 1991. A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *J. Virol.* 65:2415-2421.
47. Charneau, P., G.Mirambeau, P.Roux, S.Paulous, H.Buc, and F.Clavel. 1994. HIV-1 reverse transcription. A termination step at the center of the genome. *J. Mol. Biol.* 241:651-662.
48. Chejanovsky, N., and B.J.Carter. 1989. Replication of a human parvovirus nonsense mutant in mammalian cells containing an inducible amber suppressor. *Virology* 171:239-247.
49. Chen, T.R., R.J.Hay, and M.L.Macy. 1983. Intercellular karyotypic similarity in near-diploid cell lines of human tumor origins. *Cancer Genet. Cytogenet.* 10:351-362.
50. Cherepanov, P., G.Maertens, P.Proost, B.Devreese, J.Van Beeumen, Y.Engelborghs, E.De Clercq, and Z.Debyser. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* 278:372-381.
51. Ciliberto, G., L.Dente, and R.Cortese. 1985. Cell-specific expression of a transfected human alpha 1-antitrypsin gene. *Cell* 41:531-540.
52. Ciuffi, A., M.Llano, E.Poeschla, C.Hoffmann, J.Leipzig, P.Shinn, J.R.Ecker, and F.Bushman. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* 11:1287-1289.
53. Coil, D.A., and A.D.Miller. 2004. Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. *J. Virol.* 78:10920-10926.
54. Cornu, T.I., and T.Cathomen. 2007. Targeted genome modifications using integrase-deficient lentiviral vectors. *Mol. Ther.* 15:2107-2113.

55. Cosset, F.L., Y.Takeuchi, J.L.Battini, R.A.Weiss, and M.K.Collins. 1995. High-titer packaging cells producing recombinant retroviruses resistant to human serum. *J. Virol.* 69:7430-7436.
56. Coyne, C.B., and J.M.Bergelson. 2005. CAR: a virus receptor within the tight junction. *Adv. Drug Deliv. Rev.* 57:869-882.
57. Craigie, R. 2001. HIV integrase, a brief overview from chemistry to therapeutics. *J. Biol. Chem.* 276:23213-23216.
58. D'Halluin, J.C., G.R.Martin, G.Torpiere, and P.A.Boulanger. 1978a. Adenovirus type 2 assembly analyzed by reversible cross-linking of labile intermediates. *J. Virol.* 26:357-363.
59. D'Halluin, J.C., M.Milleville, P.A.Boulanger, and G.R.Martin. 1978b. Temperature-sensitive mutant of adenovirus type 2 blocked in virion assembly: accumulation of light intermediate particles. *J. Virol.* 26:344-356.
60. Dalgleish, A.G., P.C.Beverley, P.R.Clapham, D.H.Crawford, M.F.Greaves, and R.A.Weiss. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 312:763-767.
61. Danos, O., and R.C.Mulligan. 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U. S. A* 85:6460-6464.
62. Darlix, J.L., C.Gabus, M.T.Nugeyre, F.Clavel, and F.Barre-Sinoussi. 1990. Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* 216:689-699.
63. De Noronha, C.M., M.P.Sherman, H.W.Lin, M.V.Cavrois, R.D.Moir, R.D.Goldman, and W.C.Greene. 2001. Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr. *Science* 294:1105-1108.
64. Deng, H., R.Liu, W.Ellmeier, S.Choe, D.Unutmaz, M.Burkhart, P.Di Marzio, S.Marmon, R.E.Sutton, C.M.Hill, C.B.Davis, S.C.Peiper, T.J.Schall, D.R.Littman, and N.R.Landau. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381:661-666.
65. DePolo, N.J., J.D.Reed, P.L.Sheridan, K.Townsend, S.L.Sauter, D.J.Jolly, and T.W.Dubensky, Jr. 2000. VSV-G pseudotyped lentiviral vector particles produced in human cells are inactivated by human serum. *Mol. Ther.* 2:218-222.

66. Desmaris, N., A.Bosch, C.Salaun, C.Petit, M.C.Prevost, N.Tordo, P.Perrin, O.Schwartz, H.de Rocquigny, and J.M.Heard. 2001. Production and neurotropism of lentivirus vectors pseudotyped with lyssavirus envelope glycoproteins. *Mol. Ther.* 4:149-156.
67. Dobner, T., N.Horikoshi, S.Rubenwolf, and T.Shenk. 1996. Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. *Science* 272:1470-1473.
68. Dragic, T., V.Litwin, G.P.Allaway, S.R.Martin, Y.Huang, K.A.Nagashima, C.Cayanan, P.J.Maddon, R.A.Koup, J.P.Moore, and W.A.Paxton. 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 381:667-673.
69. Dranoff, G., E.Jaffee, A.Lazenby, P.Golumbek, H.Levitsky, K.Brose, V.Jackson, H.Hamada, D.Pardoll, and R.C.Mulligan. 1993. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. *Proc. Natl. Acad. Sci. U. S. A* 90:3539-3543.
70. Du, Y., S.E.Spence, N.A.Jenkins, and N.G.Copeland. 2005. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* 106:2498-2505.
71. Dull, T., R.Zufferey, M.Kelly, R.J.Mandel, M.Nguyen, D.Trono, and L.Naldini. 1998. A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* 72:8463-8471.
72. Dvorin, J.D., P.Bell, G.G.Maul, M.Yamashita, M.Emerman, and M.H.Malim. 2002. Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import. *J. Virol.* 76:12087-12096.
73. Edvardsson, B., E.Everitt, H.Jornvall, L.Prage, and L.Philipson. 1976. Intermediates in adenovirus assembly. *J. Virol.* 19:533-547.
74. Egan, M.A., L.M.Carruth, J.F.Rowell, X.Yu, and R.F.Siliciano. 1996. Human immunodeficiency virus type 1 envelope protein endocytosis mediated by a highly conserved intrinsic internalization signal in the cytoplasmic domain of gp41 is suppressed in the presence of the Pr55gag precursor protein. *J. Virol.* 70:6547-6556.
75. Eglitis, M.A., P.Kantoff, E.Gilboa, and W.F.Anderson. 1985. Gene expression in mice after high efficiency retroviral-mediated gene transfer. *Science* 230:1395-1398.

76. Einfeld, D., and E.Hunter. 1988. Oligomeric structure of a prototype retrovirus glycoprotein. *Proc. Natl. Acad. Sci. U. S. A* 85:8688-8692.
77. Emerman, M., and H.Temin. 1984. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. *Cell* 39:449-467.
78. Emiliani, S., A.Mousnier, K.Busschots, M.Maroun, B.Van Maele, D.Tempe, L.Vandekerckhove, F.Moisant, L.Ben Slama, M.Witvrouw, F.Christ, J.C.Rain, C.Dargemont, Z.Debyser, and R.Benarous. 2005. Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J. Biol. Chem.* 280:25517-25523.
79. Engelman, A., and R.Craigie. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* 66:6361-6369.
80. Engelman, A., G.Englund, J.M.Orenstein, M.A.Martin, and R.Craigie. 1995. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J. Virol.* 69:2729-2736.
81. Esposito, D., and R.Craigie. 1998. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J.* 17:5832-5843.
82. Fallaux, F.J., A.Bout, d.V.van, I, D.J.van den Wollenberg, K.M.Hehir, J.Keegan, C.Auger, S.J.Cramer, H.van Ormondt, A.J.van der Eb, D.Valerio, and R.C.Hoeben. 1998. New helper cells and matched early region 1-deleted adenovirus vectors prevent generation of replication-competent adenoviruses. *Hum. Gene Ther.* 9:1909-1917.
83. Fan, H., and D.Baltimore. 1973. RNA metabolism of murine leukemia virus: detection of virus-specific RNA sequences in infected and uninfected cells and identification of virus-specific messenger RNA. *J. Mol. Biol.* 80:93-117.
84. Farnet, C.M., and F.D.Bushman. 1997. HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* 88:483-492.
85. Faruqi, A.F., M.Egholm, and P.M.Glazer. 1998. Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells. *Proc. Natl. Acad. Sci. U. S. A* 95:1398-1403.

86. Faruqi, A.F., M.M.Seidman, D.J.Segal, D.Carroll, and P.M.Glazer. 1996. Recombination induced by triple-helix-targeted DNA damage in mammalian cells. *Mol. Cell Biol.* 16:6820-6828.
87. Fassati, A., and S.P.Goff. 1999. Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. *J. Virol.* 73:8919-8925.
88. Fassati, A., and S.P.Goff. 2001. Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J. Virol.* 75:3626-3635.
89. Fassati, A., D.Gorlich, I.Harrison, L.Zaytseva, and J.M.Mingot. 2003. Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *EMBO J.* 22:3675-3685.
90. Felsenstein, K.M., and S.P.Goff. 1988. Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing. *J. Virol.* 62:2179-2182.
91. Feng, Y., C.C.Broder, P.E.Kennedy, and E.A.Berger. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272:872-877.
92. Fessler, S.P., F.Delgado-Lopez, and M.S.Horwitz. 2004. Mechanisms of E3 modulation of immune and inflammatory responses. *Curr. Top. Microbiol. Immunol.* 273:113-135.
93. Finn, G.K., B.W.Kurz, R.Z.Cheng, and R.J.Shmookler Reis. 1989. Homologous plasmid recombination is elevated in immortalized transformed cells. *Mol. Cell Biol.* 9:4009-4017.
94. Finston, W.I., and J.J.Champoux. 1984. RNA-primed initiation of Moloney murine leukemia virus plus strands by reverse transcriptase in vitro. *J. Virol.* 51:26-33.
95. Fisher, R.D., H.Y.Chung, Q.Zhai, H.Robinson, W.I.Sundquist, and C.P.Hill. 2007. Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. *Cell* 128:841-852.
96. Flomenberg, P.R., M.Chen, and M.S.Horwitz. 1987. Characterization of a major histocompatibility complex class I antigen-binding glycoprotein from adenovirus type 35, a type associated with immunocompromised hosts. *J. Virol.* 61:3665-3671.

97. Follenzi, A., L.E.Ailles, S.Bakovic, M.Geuna, and L.Naldini. 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* 25:217-222.
98. Fornerod, M., M.Ohno, M.Yoshida, and I.W.Mattaj. 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90:1051-1060.
99. Franke, E.K., H.E.Yuan, and J.Luban. 1994. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* 372:359-362.
100. Friedman, T., and R.Roblin. 1972. Gene Therapy for Human Genetic Disease. *Science* 175:949-&.
101. Furger, A., J.Monks, and N.J.Proudfoot. 2001. The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation. *J. Virol.* 75:11735-11746.
102. Gallay, P., V.Stitt, C.Mundy, M.Oettinger, and D.Trono. 1996. Role of the karyopherin pathway in human immunodeficiency virus type 1 nuclear import. *J. Virol.* 70:1027-1032.
103. Garber, K. 2006. China approves world's first oncolytic virus therapy for cancer treatment. *J. Natl. Cancer Inst.* 98:298-300.
104. Garrus, J.E., U.K.von Schwedler, O.W.Pornillos, S.G.Morham, K.H.Zavitz, H.E.Wang, D.A.Wettstein, K.M.Stray, M.Cote, R.L.Rich, D.G.Myszka, and W.I.Sundquist. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107:55-65.
105. Gaspar, H.B., E.Bjorkegren, K.Parsley, K.C.Gilmour, D.King, J.Sinclair, F.Zhang, A.Giannakopoulos, S.Adams, L.D.Fairbanks, J.Gaspar, L.Henderson, J.H.Xu-Bayford, E.G.Davies, P.A.Veys, C.Kinnon, and A.J.Thrasher. 2006. Successful reconstitution of immunity in ADA-SCID by stem cell gene therapy following cessation of PEG-ADA and use of mild preconditioning. *Mol. Ther.* 14:505-513.
106. Gaspar, H.B., K.L.Parsley, S.Howe, D.King, K.C.Gilmour, J.Sinclair, G.Brouns, M.Schmidt, C.von Kalle, T.Barington, M.A.Jakobsen, H.O.Christensen, A.Al Ghonaium, H.N.White, J.L.Smith, R.J.Levinsky, R.R.Ali, C.Kinnon, and A.J.Thrasher. 2004. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* 364:2181-2187.

107. Gaylord, W.H.Jr. 1955. Virus-like particles associated with the Rous sarcoma as seen in sections of tumor. *Cancer Res.* 15:80-83.
108. Germann, U.A., K.V.Chin, I.Pastan, and M.M.Gottesman. 1990. Retroviral transfer of a chimeric multidrug resistance-adenosine deaminase gene. *FASEB J.* 4:1501-1507.
109. Gerritsen, M.E., A.J.Williams, A.S.Neish, S.Moore, Y.Shi, and T.Collins. 1997. CREB-binding protein/p300 are transcriptional coactivators of p65. *Proc. Natl. Acad. Sci. U. S. A* 94:2927-2932.
110. Gerton, J.L., S.Ohgi, M.Olsen, J.DeRisi, and P.O.Brown. 1998. Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions. *J. Virol.* 72:5046-5055.
111. Gheysen, D., E.Jacobs, F.de Foresta, C.Thiriart, M.Francotte, D.Thines, and M.De Wilde. 1989. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* 59:103-112.
112. Gianni, A.M., D.Smotkin, and R.A.Weinberg. 1975. Murine leukemia virus: detection of unintegrated double-stranded DNA forms of the provirus. *Proc. Natl. Acad. Sci. U. S. A* 72:447-451.
113. Gilboa, E., M.Kolbe, K.Noonan, and R.Kucherlapati. 1982. Construction of a mammalian transducing vector from the genome of Moloney murine leukemia virus. *J. Virol.* 44:845-851.
114. Gossen, M., and H.Bujard. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A* 89:5547-5551.
115. Gottlinger, H.G., T.Dorfman, J.G.Sodroski, and W.A.Haseltine. 1991. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc. Natl. Acad. Sci. U. S. A* 88:3195-3199.
116. Grable, M., and P.Hearing. 1990. Adenovirus type 5 packaging domain is composed of a repeated element that is functionally redundant. *J. Virol.* 64:2047-2056.
117. Graham, F.L., J.Smiley, W.C.Russell, and R.Nairn. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36:59-74.

118. Grave, L., D.Dreyer, A.Dieterle, P.Leroy, A.I.Michou, C.Doderer, A.Pavirani, M.Lusky, and M.Mehtali. 2000. Differential influence of the E4 adenoviral genes on viral and cellular promoters. *J. Gene Med.* 2:433-443.
119. Green, M., M.Pina, R.Kimes, P.C.Wensink, L.A.MacHattie, and C.A.Thomas, Jr. 1967. Adenovirus DNA. I. Molecular weight and conformation. *Proc. Natl. Acad. Sci. U. S. A* 57:1302-1309.
120. Gregorevic, P., J.M.Allen, E.Minami, M.J.Blankinship, M.Haraguchi, L.Meuse, E.Finn, M.E.Adams, S.C.Froehner, C.E.Murry, and J.S.Chamberlain. 2006. rAAV6-microdystrophin preserves muscle function and extends lifespan in severely dystrophic mice. *Nat. Med.* 12:787-789.
121. Grosveld, F., G.B.van Assendelft, D.R.Greaves, and G.Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51:975-985.
122. Hacein-Bey-Abina, S., C.von Kalle, M.Schmidt, F.Le Deist, N.Wulffraat, E.McIntyre, I.Radford, J.L.Villeval, C.C.Fraser, M.Cavazzana-Calvo, and A.Fischer. 2003a. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 348:255-256.
123. Hacein-Bey-Abina, S., C.von Kalle, M.Schmidt, M.P.McCormack, N.Wulffraat, P.Leboulch, A.Lim, C.S.Osborne, R.Pawliuk, E.Morillon, R.Sorensen, A.Forster, P.Fraser, J.I.Cohen, B.G.de Saint, I.Alexander, U.Wintergerst, T.Freboung, A.Aurias, D.Stoppa-Lyonnet, S.Romana, I.Radford-Weiss, F.Gross, F.Valensi, E.Delabesse, E.Macintyre, F.Sigaux, J.Soulier, L.E.Leiva, M.Wissler, C.Prinz, T.H.Rabbitts, F.Le Deist, A.Fischer, and M.Cavazzana-Calvo. 2003b. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302:415-419.
124. Hallenberger, S., V.Bosch, H.Angliker, E.Shaw, H.D.Klenk, and W.Garten. 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* 360:358-361.
125. Hatzioannou, T., E.Delahaye, F.Martin, S.J.Russell, and F.L.Cosset. 1999. Retroviral display of functional binding domains fused to the amino terminus of influenza hemagglutinin. *Hum. Gene Ther.* 10:1533-1544.
126. Hauswirth, W.W., and K.I.Berns. 1977. Origin and termination of adeno-associated virus DNA replication. *Virology* 78:488-499.
127. Hayward, W.S., B.G.Neel, and S.M.Astrin. 1981. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. *Nature* 290:475-480.

128. Heise, C., A.Sampson-Johannes, A.Williams, F.McCormick, D.D.Von Hoff, and D.H.Kirn. 1997. ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemotherapeutic agents. *Nat. Med.* 3:639-645.
129. Henderson, A., M.Bunce, N.Siddon, R.Reeves, and D.J.Tremethick. 2000. High-mobility-group protein I can modulate binding of transcription factors to the U5 region of the human immunodeficiency virus type 1 proviral promoter. *J. Virol.* 74:10523-10534.
130. Henderson, L.E., H.C.Krutzsch, and S.Oroszlan. 1983. Myristyl amino-terminal acylation of murine retrovirus proteins: an unusual post-translational proteins modification. *Proc. Natl. Acad. Sci. U. S. A* 80:339-343.
131. Hermonat, P.L., M.A.Labow, R.Wright, K.I.Berns, and N.Muzyczka. 1984. Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. *J. Virol.* 51:329-339.
132. Herrmann, C.H., and A.P.Rice. 1995. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. *J. Virol.* 69:1612-1620.
133. Herzog, R.W., J.N.Hagstrom, S.H.Kung, S.J.Tai, J.M.Wilson, K.J.Fisher, and K.A.High. 1997. Stable gene transfer and expression of human blood coagulation factor IX after intramuscular injection of recombinant adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A* 94:5804-5809.
134. Herzog, R.W., E.Y.Yang, L.B.Couto, J.N.Hagstrom, D.Elwell, P.A.Fields, M.Burton, D.A.Bellinger, M.S.Read, K.M.Brinkhous, G.M.Podsakoff, T.C.Nichols, G.J.Kurtzman, and K.A.High. 1999. Long-term correction of canine hemophilia B by gene transfer of blood coagulation factor IX mediated by adeno-associated viral vector. *Nat. Med.* 5:56-63.
135. Holmes-Son, M.L., and S.A.Chow. 2002. Correct integration mediated by integrase-LexA fusion proteins incorporated into HIV-1. *Mol. Ther.* 5:360-370.
136. Horne, R.W., S.Brenner, A.P.Watson, and P.Wildy. 1959. The icosahedral form of an adenovirus. *J. Mol. Biol.* 1:84-86.
137. Hu, H., D.Serra, and A.Amalfitano. 1999. Persistence of an [E1-, polymerase-] adenovirus vector despite transduction of a neoantigen into immune-competent mice. *Hum. Gene Ther.* 10:355-364.

138. Huang, M.M., and P.Hearing. 1989. Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *J. Virol.* 63:2605-2615.
139. Hughes, S.H., P.R.Shank, D.H.Spector, H.J.Kung, J.M.Bishop, H.E.Varmus, P.K.Vogt, and M.L.Breitman. 1978. Provirus of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. *Cell* 15:1397-1410.
140. Ikeda, Y., Y.Takeuchi, F.Martin, F.L.Cosset, K.Mitrophanous, and M.Collins. 2003. Continuous high-titer HIV-1 vector production. *Nat. Biotechnol.* 21:569-572.
141. Index of Viruses - Retroviridae. 2006. ICTVdB - The Universal Virus Database. Columbia University, New York.
142. Jacks, T., and H.E.Varmus. 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* 230:1237-1242.
143. Jacque, J.M., and M.Stevenson. 2006. The inner-nuclear-envelope protein emerlin regulates HIV-1 infectivity. *Nature* 441:641-645.
144. Jamjoom, G.A., R.B.Naso, and R.B.Arlinghaus. 1977. Further characterization of intracellular precursor polyproteins of Rauscher leukemia virus. *Virology* 78:11-34.
145. Janik, J.E., M.M.Huston, K.Cho, and J.A.Rose. 1989. Efficient synthesis of adeno-associated virus structural proteins requires both adenovirus DNA binding protein and VA I RNA. *Virology* 168:320-329.
146. Jiang, M., J.Mak, A.Ladha, E.Cohen, M.Klein, B.Rovinski, and L.Kleiman. 1993. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. *J. Virol.* 67:3246-3253.
147. Johnson, M.S., M.A.McClure, D.F.Feng, J.Gray, and R.F.Doolittle. 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc. Natl. Acad. Sci. U. S. A* 83:7648-7652.
148. Jowett, J.B., D.J.Hockley, M.V.Nermut, and I.M.Jones. 1992. Distinct signals in human immunodeficiency virus type 1 Pr55 necessary for RNA binding and particle formation. *J. Gen. Virol.* 73 (Pt 12):3079-3086.
149. Ju, G., and A.M.Skalka. 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. *Cell* 22:379-386.

150. Kafri, T., H.van Praag, L.Ouyang, F.H.Gage, and I.M.Verma. 1999. A packaging cell line for lentivirus vectors. *J. Virol.* 73:576-584.
151. Kalpana, G.V., S.Marmon, W.Wang, G.R.Crabtree, and S.P.Goff. 1994. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* 266:2002-2006.
152. Kao, S.Y., A.F.Calman, P.A.Luciw, and B.M.Peterlin. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330:489-493.
153. Kaplan, A.H., and R.Swanstrom. 1991. The HIV-1 gag precursor is processed via two pathways: implications for cytotoxicity. *Biomed. Biochim. Acta* 50:647-653.
154. Karwacz, K., S.Mukherjee, L.Apolonia, M.P.Blundell, G.Bouma, D.Escors, M.K.Collins, and A.J.Thrasher. 2009. Nonintegrating lentivector vaccines stimulate prolonged T-cell and antibody responses and are effective in tumor therapy. *J. Virol.* 83:3094-3103.
155. Kashiwakura, Y., K.Tamayose, K.Iwabuchi, Y.Hirai, T.Shimada, K.Matsumoto, T.Nakamura, M.Watanabe, K.Oshimi, and H.Daida. 2005. Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J. Virol.* 79:609-614.
156. Katzman, M., R.A.Katz, A.M.Skalka, and J.Leis. 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. *J. Virol.* 63:5319-5327.
157. Kaye, J.F., and A.M.Lever. 1998. Nonreciprocal packaging of human immunodeficiency virus type 1 and type 2 RNA: a possible role for the p2 domain of Gag in RNA encapsidation. *J. Virol.* 72:5877-5885.
158. Keith, J., and H.Fraenkel-Conrat. 1975. Identification of the 5' end of Rous sarcoma virus RNA. *Proc. Natl. Acad. Sci. U. S. A* 72:3347-3350.
159. Kim, Y.K., C.F.Bourgeois, R.Pearson, M.Tyagi, M.J.West, J.Wong, S.Y.Wu, C.M.Chiang, and J.Karn. 2006. Recruitment of TFIID to the HIV LTR is a rate-limiting step in the emergence of HIV from latency. *EMBO J.* 25:3596-3604.
160. Klatzmann, D., E.Champagne, S.Chamaret, J.Gruet, D.Guetard, T.Hercend, J.C.Gluckman, and L.Montagnier. 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 312:767-768.

161. Korman, A.J., J.D.Frantz, J.L.Strominger, and R.C.Mulligan. 1987. Expression of human class II major histocompatibility complex antigens using retrovirus vectors. *Proc. Natl. Acad. Sci. U. S. A* 84:2150-2154.
162. Kotin, R.M., M.Siniscalco, R.J.Samulski, X.D.Zhu, L.Hunter, C.A.Laughlin, S.McLaughlin, N.Muzyczka, M.Rocchi, and K.I.Berns. 1990. Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A* 87:2211-2215.
163. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125-8148.
164. Krall, W.J., D.C.Skelton, X.J.Yu, I.Riviere, P.Lehn, R.C.Mulligan, and D.B.Kohn. 1996. Increased levels of spliced RNA account for augmented expression from the MFG retroviral vector in hematopoietic cells. *Gene Ther.* 3:37-48.
165. Krausslich, H.G., M.Facke, A.M.Heuser, J.Konvalinka, and H.Zentgraf. 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *J. Virol.* 69:3407-3419.
166. Krosli, J., S.Baban, G.Krosli, S.Rozenfeld, C.Largman, and G.Sauvageau. 1998. Cellular proliferation and transformation induced by HOXB4 and HOXB3 proteins involves cooperation with PBX1. *Oncogene* 16:3403-3412.
167. Lai, M.M., and P.H.Duesberg. 1972. Adenylic acid-rich sequence in RNAs of Rous sarcoma virus and Rauscher mouse leukaemia virus. *Nature* 235:383-386.
168. Lazo, P.A., J.S.Lee, and P.N.Tsichlis. 1990. Long-distance activation of the Myc protooncogene by provirus insertion in Mlvi-1 or Mlvi-4 in rat T-cell lymphomas. *Proc. Natl. Acad. Sci. U. S. A* 87:170-173.
169. Leavitt, A.D., G.Robles, N.Alesandro, and H.E.Varmus. 1996. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *J. Virol.* 70:721-728.
170. Ledbetter, J.A. 1979. Two-dimensional analysis of murine leukemia virus gag-gene polyproteins. *Virology* 95:85-98.
171. Lee, J.S., H.E.Varmus, and J.M.Bishop. 1979. Virus-specific messenger RNAs in permissive cells infected by avian sarcoma virus. *J. Biol. Chem.* 254:8015-8022.
172. Lee, M.S., and R.Craigie. 1994. Protection of retroviral DNA from autointegration: involvement of a cellular factor. *Proc. Natl. Acad. Sci. U. S. A* 91:9823-9827.

173. Lee, M.S., and R.Craigie. 1998. A previously unidentified host protein protects retroviral DNA from autointegration. *Proc. Natl. Acad. Sci. U. S. A* 95:1528-1533.
174. Lewinski, M.K., M.Yamashita, M.Emerman, A.Ciuffi, H.Marshall, G.Crawford, F.Collins, P.Shinn, J.Leipzig, S.Hannenhalli, C.C.Berry, J.R.Ecker, and F.D.Bushman. 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS. Pathog.* 2:e60.
175. Lewis, P., M.Hensel, and M.Emerman. 1992. Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J.* 11:3053-3058.
176. Li, S., E.Kimura, B.M.Fall, M.Reyes, J.C.Angello, R.Welikson, S.D.Hauschka, and J.S.Chamberlain. 2005. Stable transduction of myogenic cells with lentiviral vectors expressing a minidystrophin. *Gene Ther.* 12:1099-1108.
177. Ljung, R.C. 1998. Can haemophilic arthropathy be prevented? *Br. J. Haematol.* 101:215-219.
178. Logan, A.C., D.L.Haas, T.Kafri, and D.B.Kohn. 2004. Integrated self-inactivating lentiviral vectors produce full-length genomic transcripts competent for encapsidation and integration. *J. Virol.* 78:8421-8436.
179. Lombardo, A., P.Genovese, C.M.Beausejour, S.Colleoni, Y.L.Lee, K.A.Kim, D.Ando, F.D.Urnov, C.Galli, P.D.Gregory, M.C.Holmes, and L.Naldini. 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* 25:1298-1306.
180. Lu, R., A.Limon, H.Z.Ghory, and A.Engelman. 2005. Genetic analyses of DNA-binding mutants in the catalytic core domain of human immunodeficiency virus type 1 integrase. *J. Virol.* 79:2493-2505.
181. Lu, R., N.Nakajima, W.Hofmann, M.Benkirane, K.T.Jeang, J.Sodroski, and A.Engelman. 2004. Simian virus 40-based replication of catalytically inactive human immunodeficiency virus type 1 integrase mutants in nonpermissive T cells and monocyte-derived macrophages. *J. Virol.* 78:658-668.
182. Lu, Y., Y.K.Choi, M.Campbell-Thompson, C.Li, Q.Tang, J.M.Crawford, T.R.Flotte, and S.Song. 2006. Therapeutic level of functional human alpha 1 antitrypsin (hAAT) secreted from murine muscle transduced by adeno-associated virus (rAAV1) vector. *J. Gene Med.* 8:730-735.
183. Lu, Y.L., P.Spearman, and L.Ratner. 1993. Human immunodeficiency virus type 1 viral protein R localization in infected cells and virions. *J. Virol.* 67:6542-6550.

184. Luban, J., K.L.Bossolt, E.K.Franke, G.V.Kalpana, and S.P.Goff. 1993. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell* 73:1067-1078.
185. Lusso, P., M.Di, V, B.Ensoli, G.Franchini, C.Jemma, S.E.Derocco, V.S.Kalyanaraman, and R.C.Gallo. 1990. Expanded HIV-1 cellular tropism by phenotypic mixing with murine endogenous retroviruses. *Science* 247:848-852.
186. Maddon, P.J., A.G.Dalgleish, J.S.McDougal, P.R.Clapham, R.A.Weiss, and R.Axel. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
187. Mangel, W.F., M.L.Baniecki, and W.J.McGrath. 2003. Specific interactions of the adenovirus proteinase with the viral DNA, an 11-amino-acid viral peptide, and the cellular protein actin. *Cell Mol. Life Sci.* 60:2347-2355.
188. Mann, D.A., I.Mikaelian, R.W.Zemmel, S.M.Green, A.D.Lowe, T.Kimura, M.Singh, P.J.Butler, M.J.Gait, and J.Karn. 1994. A molecular rheostat. Cooperative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J. Mol. Biol.* 241:193-207.
189. Mann, R., R.C.Mulligan, and D.Baltimore. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* 33:153-159.
190. Manno, C.S., A.J.Chew, S.Hutchison, P.J.Larson, R.W.Herzog, V.R.Arruda, S.J.Tai, M.V.Ragni, A.Thompson, M.Ozelo, L.B.Couto, D.G.Leonard, F.A.Johnson, A.McClelland, C.Scallan, E.Skarsgard, A.W.Flake, M.A.Kay, K.A.High, and B.Glader. 2003. AAV-mediated factor IX gene transfer to skeletal muscle in patients with severe hemophilia B. *Blood* 101:2963-2972.
191. Manno, C.S., G.F.Pierce, V.R.Arruda, B.Glader, M.Ragni, J.J.Rasko, M.C.Ozelo, K.Hoots, P.Blatt, B.Konkle, M.Dake, R.Kaye, M.Razavi, A.Zajko, J.Zehnder, P.K.Rustagi, H.Nakai, A.Chew, D.Leonard, J.F.Wright, R.R.Lessard, J.M.Sommer, M.Tigges, D.Sabatino, A.Luk, H.Jiang, F.Mingozzi, L.Couto, H.C.Ertl, K.A.High, and M.A.Kay. 2006. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* 12:342-347.
192. Masuda, T., M.J.Kuroda, and S.Harada. 1998. Specific and independent recognition of U3 and U5 att sites by human immunodeficiency virus type 1 integrase in vivo. *J. Virol.* 72:8396-8402.
193. Mathias, P., T.Wickham, M.Moore, and G.Nemerow. 1994. Multiple adenovirus serotypes use alpha v integrins for infection. *J. Virol.* 68:6811-6814.

194. Mazarakis, N.D., M.Azzouz, J.B.Rohll, F.M.Ellard, F.J.Wilkes, A.L.Olsen, E.E.Carter, R.D.Barber, D.F.Baban, S.M.Kingsman, A.J.Kingsman, K.O'Malley, and K.A.Mitrophanous. 2001. Rabies virus glycoprotein pseudotyping of lentiviral vectors enables retrograde axonal transport and access to the nervous system after peripheral delivery. *Hum. Mol. Genet.* 10:2109-2121.
195. McCarty, D.M. 2008. Self-complementary AAV vectors; advances and applications. *Mol. Ther.* 16:1648-1656.
196. McDonald, D., M.A.Vodicka, G.Lucero, T.M.Svitkina, G.G.Borisy, M.Emerman, and T.J.Hope. 2002. Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* 159:441-452.
197. McDonald, D., L.Wu, S.M.Bohks, V.N.KewalRamani, D.Unutmaz, and T.J.Hope. 2003. Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science* 300:1295-1297.
198. Miller, A.D., and C.Buttimore. 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell Biol.* 6:2895-2902.
199. Miller, A.D., J.V.Garcia, N.von Suhr, C.M.Lynch, C.Wilson, and M.V.Eiden. 1991. Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J. Virol.* 65:2220-2224.
200. Miller, M.D., B.Wang, and F.D.Bushman. 1995. Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. *J. Virol.* 69:3938-3944.
201. Miller, S.A., D.D.Dykes, and H.F.Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215.
202. Mingozzi, F., M.V.Maus, D.J.Hui, D.E.Sabatino, S.L.Murphy, J.E.Rasko, M.V.Ragni, C.S.Manno, J.Sommer, H.Jiang, G.F.Pierce, H.C.Ertl, and K.A.High. 2007. CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat. Med.* 13:419-422.
203. Mitchell, R., C.Y.Chiang, C.Berry, and F.Bushman. 2003. Global analysis of cellular transcription following infection with an HIV-based vector. *Mol. Ther.* 8:674-687.

204. Mitchell, R.S., B.F.Beitzel, A.R.Schroder, P.Shinn, H.Chen, C.C.Berry, J.R.Ecker, and F.D.Bushman. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS. Biol.* 2:E234.
205. Mitra, S.W., S.Goff, E.Gilboa, and D.Baltimore. 1979. Synthesis of a 600-nucleotide-long plus-strand DNA by virions of Moloney murine leukemia virus. *Proc. Natl. Acad. Sci. U. S. A* 76:4355-4359.
206. Mochizuki, H., J.P.Schwartz, K.Tanaka, R.O.Brady, and J.Reiser. 1998. High-titer human immunodeficiency virus type 1-based vector systems for gene delivery into nondividing cells. *J. Virol.* 72:8873-8883.
207. Modlich, U., J.Bohne, M.Schmidt, C.von Kalle, S.Knoss, A.Schambach, and C.Baum. 2006. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* 108:2545-2553.
208. Molling, K., D.P.Bolognesi, H.Bauer, W.Busen, H.W.Plassmann, and P.Hausen. 1971. Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat. New Biol.* 234:240-243.
209. Montano, M.A., K.Kripke, C.D.Norina, P.Achacoso, L.A.Herzenberg, A.L.Roy, and G.P.Nolan. 1996. NF-kappa B homodimer binding within the HIV-1 initiator region and interactions with TFII-I. *Proc. Natl. Acad. Sci. U. S. A* 93:12376-12381.
210. Montini, E., D.Cesana, M.Schmidt, F.Sanvito, M.Ponzoni, C.Bartholomae, S.L.Sergi, F.Benedicenti, A.Ambrosi, C.Di Serio, C.Dogliani, C.von Kalle, and L.Naldini. 2006. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat. Biotechnol.* 24:687-696.
211. Morizono, K., Y.Xie, G.E.Ringpis, M.Johnson, H.Nassanian, B.Lee, L.Wu, and I.S.Chen. 2005. Lentiviral vector retargeting to P-glycoprotein on metastatic melanoma through intravenous injection. *Nat. Med.* 11:346-352.
212. Morrison, T., C.McQuain, and L.McGinnes. 1991. Complementation between avirulent Newcastle disease virus and a fusion protein gene expressed from a retrovirus vector: requirements for membrane fusion. *J. Virol.* 65:813-822.
213. Mumm, S.R., and D.P.Grandgenett. 1991. Defining nucleic acid-binding properties of avian retrovirus integrase by deletion analysis. *J. Virol.* 65:1160-1167.
214. Nabel, G., and D.Baltimore. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 326:711-713.

215. Naldini, L., U.Blomer, F.H.Gage, D.Trono, and I.M.Verma. 1996a. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. U. S. A* 93:11382-11388.
216. Naldini, L., U.Blomer, P.Gallay, D.Ory, R.Mulligan, F.H.Gage, I.M.Verma, and D.Trono. 1996b. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
217. Nathwani, A.C., J.T.Gray, C.Y.Ng, J.Zhou, Y.Spence, S.N.Waddington, E.G.Tuddenham, G.Kemball-Cook, J.McIntosh, M.Boon-Spijker, K.Mertens, and A.M.Davidoff. 2006. Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood* 107:2653-2661.
218. Negri, D.R., Z.Michelini, S.Baroncelli, M.Spada, S.Vendetti, V.Buffa, R.Bona, P.Leone, M.E.Klotman, and A.Cara. 2007. Successful immunization with a single injection of non-integrating lentiviral vector. *Mol. Ther.* 15:1716-1723.
219. Neil, S.J., T.Zang, and P.D.Bieniasz. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451:425-430.
220. Nermut, M.V., and A.Fassati. 2003. Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes. *J. Virol.* 77:8196-8206.
221. Neville, M., F.Stutz, L.Lee, L.I.Davis, and M.Rosbash. 1997. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* 7:767-775.
222. Nightingale, S.J., R.P.Hollis, K.A.Pepper, D.Petersen, X.J.Yu, C.Yang, I.Bahner, and D.B.Kohn. 2006. Transient gene expression by nonintegrating lentiviral vectors. *Mol. Ther.* 13:1121-1132.
223. Niidome, T., and L.Huang. 2002. Gene therapy progress and prospects: Nonviral vectors. *Gene Therapy* 9:1647-1652.
224. Ott, M.G., M.Schmidt, K.Schwarzwaelder, S.Stein, U.Siler, U.Koehl, H.Glimm, K.Kuhlcke, A.Schilz, H.Kunkel, S.Naundorf, A.Brinkmann, A.Deichmann, M.Fischer, C.Ball, I.Pilz, C.Dunbar, Y.Du, N.A.Jenkins, N.G.Copeland, U.Luthi, M.Hassan, A.J.Thrasher, D.Hoelzer, C.von Kalle, R.Seger, and M.Grez. 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented

- by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat. Med.* 12:401-409.
225. Patience, C., Y.Takeuchi, F.L.Cosset, and R.A.Weiss. 1998. Packaging of endogenous retroviral sequences in retroviral vectors produced by murine and human packaging cells. *J. Virol.* 72:2671-2676.
 226. Patterson, S., and W.C.Russell. 1983. Ultrastructural and immunofluorescence studies of early events in adenovirus-HeLa cell interactions. *J. Gen. Virol.* 64:1091-1099.
 227. Paxton, W., R.I.Connor, and N.R.Landau. 1993. Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *J. Virol.* 67:7229-7237.
 228. Peters, G., F.Harada, J.E.Dahlberg, A.Panet, W.A.Haseltine, and D.Baltimore. 1977. Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. *J. Virol.* 21:1031-1041.
 229. Pettit, S.C., M.D.Moody, R.S.Wehtie, A.H.Kaplan, P.V.Nantermet, C.A.Klein, and R.Swanstrom. 1994. The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol.* 68:8017-8027.
 230. Philippe, S., C.Sarkis, M.Barkats, H.Mammeri, C.Ladroue, C.Petit, J.Mallet, and C.Serguera. 2006. Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 103:17684-17689.
 231. Philipson, L., and R.F.Pettersson. 2004. The coxsackie-adenovirus receptor--a new receptor in the immunoglobulin family involved in cell adhesion. *Curr. Top. Microbiol. Immunol.* 273:87-111.
 232. Phillips, A.N. 1996. Reduction of HIV concentration during acute infection: independence from a specific immune response. *Science* 271:497-499.
 233. Philpott, N.J., J.Gomos, K.I.Berns, and E.Falck-Pedersen. 2002. A p5 integration efficiency element mediates Rep-dependent integration into AAVS1 at chromosome 19. *Proc. Natl. Acad. Sci. U. S. A* 99:12381-12385.
 234. Pichlmair, A., S.S.Diebold, S.Gschmeissner, Y.Takeuchi, Y.Ikeda, M.K.Collins, and Reis e Sousa. 2007. Tubulovesicular structures within vesicular stomatitis virus

- G protein-pseudotyped lentiviral vector preparations carry DNA and stimulate antiviral responses via Toll-like receptor 9. *J. Virol.* 81:539-547.
235. Picker, L.J. 2006. Immunopathogenesis of acute AIDS virus infection. *Curr. Opin. Immunol.* 18:399-405.
236. Planelles, V., F.Bachelierie, J.B.Jowett, A.Haislip, Y.Xie, P.Banooni, T.Masuda, and I.S.Chen. 1995. Fate of the human immunodeficiency virus type 1 provirus in infected cells: a role for vpr. *J. Virol.* 69:5883-5889.
237. Poon, B., and I.S.Chen. 2003. Human immunodeficiency virus type 1 (HIV-1) Vpr enhances expression from unintegrated HIV-1 DNA. *J. Virol.* 77:3962-3972.
238. Primhak, R.A., and M.S.Tanner. 2001. Alpha-1 antitrypsin deficiency. *Arch. Dis. Child* 85:2-5.
239. Pryciak, P.M., and H.E.Varmus. 1992. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* 69:769-780.
240. Qing, K., C.Mah, J.Hansen, S.Zhou, V.Dwarki, and A.Srivastava. 1999. Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat. Med.* 5:71-77.
241. Rasheed, S., W.A.Nelson-Rees, E.M.Toth, P.Arnstein, and M.B.Gardner. 1974. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* 33:1027-1033.
242. Rein, A., M.R.McClure, N.R.Rice, R.B.Luftig, and A.M.Schultz. 1986. Myristylation site in Pr65gag is essential for virus particle formation by Moloney murine leukemia virus. *Proc. Natl. Acad. Sci. U. S. A* 83:7246-7250.
243. Rifkin, D., and R.W.Compans. 1971. Identification of the spike proteins of Rous sarcoma virus. *Virology* 46:485-489.
244. Roe, T., T.C.Reynolds, G.Yu, and P.O.Brown. 1993. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* 12:2099-2108.
245. Rogel, M.E., L.I.Wu, and M.Emerman. 1995. The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection. *J. Virol.* 69:882-888.

246. Rogers, F.A., K.M.Vasquez, M.Egholm, and P.M.Glazer. 2002. Site-directed recombination via bifunctional PNA-DNA conjugates. *Proc. Natl. Acad. Sci. U. S. A* 99:16695-16700.
247. Rosenberg, S.A., P.Aebersold, K.Cornetta, A.Kasid, R.A.Morgan, R.Moen, E.M.Karson, M.T.Lotze, J.C.Yang, S.L.Topalian, and . 1990. Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N. Engl. J. Med.* 323:570-578.
248. Ruhl, M., M.Himmelspach, G.M.Bahr, F.Hammerschmid, H.Jaksche, B.Wolff, H.Aschauer, G.K.Farrington, H.Probst, D.Bevec, and . 1993. Eukaryotic initiation factor 5A is a cellular target of the human immunodeficiency virus type 1 Rev activation domain mediating trans-activation. *J. Cell Biol.* 123:1309-1320.
249. Ruscetti, S., D.Troxler, D.Linemeyer, and E.Scolnick. 1980. Three laboratory strains of spleen focus-forming virus: comparison of their genomes and translational products. *J. Virol.* 33:140-151.
250. Saenz, D.T., N.Loewen, M.Peretz, T.Whitwam, R.Barraza, K.G.Howell, J.M.Holmes, M.Good, and E.M.Poeschla. 2004. Unintegrated lentivirus DNA persistence and accessibility to expression in nondividing cells: analysis with class I integrase mutants. *J. Virol.* 78:2906-2920.
251. Samulski, R.J., K.I.Berns, M.Tan, and N.Muzyczka. 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. U. S. A* 79:2077-2081.
252. Samulski, R.J., L.S.Chang, and T.Shenk. 1989. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J. Virol.* 63:3822-3828.
253. Samulski, R.J., and T.Shenk. 1988. Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J. Virol.* 62:206-210.
254. Sandefur, S., V.Varthakavi, and P.Spearman. 1998. The I domain is required for efficient plasma membrane binding of human immunodeficiency virus type 1 Pr55Gag. *J. Virol.* 72:2723-2732.
255. Sandor, Z., and A.Bredberg. 1995. Triple helix directed psoralen adducts induce a low frequency of recombination in an SV40 shuttle vector. *Biochim. Biophys. Acta* 1263:235-240.

256. Schaack, J., X.Guo, W.Y.Ho, M.Karlok, C.Chen, and D.Ornelles. 1995. Adenovirus type 5 precursor terminal protein-expressing 293 and HeLa cell lines. *J. Virol.* 69:4079-4085.
257. Schaack, J., W.Y.Ho, P.Freimuth, and T.Shenk. 1990. Adenovirus terminal protein mediates both nuclear matrix association and efficient transcription of adenovirus DNA. *Genes Dev.* 4:1197-1208.
258. Schambach, A., M.Galla, T.Maetzig, R.Loew, and C.Baum. 2007. Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors. *Mol. Ther.* 15:1167-1173.
259. Schlegel, R., M.C.Willingham, and I.H.Pastan. 1982. Saturable binding sites for vesicular stomatitis virus on the surface of Vero cells. *J. Virol.* 43:871-875.
260. Schroder, A.R., P.Shinn, H.Chen, C.Berry, J.R.Ecker, and F.Bushman. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110:521-529.
261. Schuettrumpf, J., R.W.Herzog, A.Schlachterman, A.Kaufhold, D.W.Stafford, and V.R.Arruda. 2005. Factor IX variants improve gene therapy efficacy for hemophilia B. *Blood* 105:2316-2323.
262. Schwartz, S., B.K.Felber, D.M.Benko, E.M.Fenyo, and G.N.Pavlikis. 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J. Virol.* 64:2519-2529.
263. Schwartzberg, P., J.Colicelli, M.L.Gordon, and S.P.Goff. 1984. Mutations in the gag gene of Moloney murine leukemia virus: effects on production of virions and reverse transcriptase. *J. Virol.* 49:918-924.
264. Selten, G., H.T.Cuyppers, and A.Berns. 1985. Proviral activation of the putative oncogene Pim-1 in MuLV induced T-cell lymphomas. *EMBO J.* 4:1793-1798.
265. Shen-Ong, G.L., H.C.Morse, III, M.Potter, and J.F.Mushinski. 1986. Two modes of c-myc activation in virus-induced mouse myeloid tumors. *Mol. Cell Biol.* 6:380-392.
266. Sherer, N.M., M.J.Lehmann, L.F.Jimenez-Soto, A.Ingmundson, S.M.Horner, G.Cicchetti, P.G.Allen, M.Pypaert, J.M.Cunningham, and W.Mothes. 2003. Visualization of retroviral replication in living cells reveals budding into multivesicular bodies. *Traffic.* 4:785-801.

267. Sherman, P.A., and J.A.Fyfe. 1990. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc. Natl. Acad. Sci. U. S. A* 87:5119-5123.
268. Shi, Y., E.Seto, L.S.Chang, and T.Shenk. 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* 67:377-388.
269. Shinnick, T.M., R.A.Lerner, and J.G.Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukaemia virus. *Nature* 293:543-548.
270. Shoemaker, C., S.Goff, E.Gilboa, M.Paskind, S.W.Mitra, and D.Baltimore. 1980. Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. *Proc. Natl. Acad. Sci. U. S. A* 77:3932-3936.
271. Smith, J.K., A.Cywinski, and J.M.Taylor. 1984. Specificity of initiation of plus-strand DNA by Rous sarcoma virus. *J. Virol.* 52:314-319.
272. Snyder, R.O., R.J.Samulski, and N.Muzyczka. 1990. In vitro resolution of covalently joined AAV chromosome ends. *Cell* 60:105-113.
273. Sodroski, J., C.Rosen, F.Wong-Staal, S.Z.Salahuddin, M.Popovic, S.Arya, R.C.Gallo, and W.A.Haseltine. 1985. Trans-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. *Science* 227:171-173.
274. Soneoka, Y., P.M.Cannon, E.E.Ramsdale, J.C.Griffiths, G.Romano, S.M.Kingsman, and A.J.Kingsman. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* 23:628-633.
275. Song, S., M.Morgan, T.Ellis, A.Poirier, K.Chesnut, J.Wang, M.Brantly, N.Muzyczka, B.J.Byrne, M.Atkinson, and T.R.Flotte. 1998. Sustained secretion of human alpha-1-antitrypsin from murine muscle transduced with adeno-associated virus vectors. *Proc. Natl. Acad. Sci. U. S. A* 95:14384-14388.
276. Sorge, J., and S.H.Hughes. 1982. Polypurine tract adjacent to the U3 region of the Rous sarcoma virus genome provides a cis-acting function. *J. Virol.* 43:482-488.
277. Spector, D.H., E.Wade, D.A.Wright, V.Koval, C.Clark, D.Jaquish, and S.A.Spector. 1990. Human immunodeficiency virus pseudotypes with expanded cellular and species tropism. *J. Virol.* 64:2298-2308.

278. Stecenko, A.A., and K.L.Brigham. 2003. Gene therapy progress and prospects: alpha-1 antitrypsin. *Gene Ther.* 10:95-99.
279. Stehelin, D., H.E.Varmus, J.M.Bishop, and P.K.Vogt. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260:170-173.
280. Stein, B.S., and E.G.Engleman. 1990. Intracellular processing of the gp160 HIV-1 envelope precursor. Endoproteolytic cleavage occurs in a cis or medial compartment of the Golgi complex. *J. Biol. Chem.* 265:2640-2649.
281. Straus, S.E., E.D.Sebring, and J.A.Rose. 1976. Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. *Proc. Natl. Acad. Sci. U. S. A* 73:742-746.
282. Stromberg, K., N.E.Hurley, N.L.Davis, R.R.Rueckert, and E.Fleissner. 1974. Structural studies of avian myeloblastosis virus: comparison of polypeptides in virion and core component by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Virol.* 13:513-528.
283. Summerford, C., J.S.Bartlett, and R.J.Samulski. 1999. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat. Med.* 5:78-82.
284. Summerford, C., and R.J.Samulski. 1998. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J. Virol.* 72:1438-1445.
285. Surosky, R.T., M.Urabe, S.G.Godwin, S.A.McQuiston, G.J.Kurtzman, K.Ozawa, and G.Natsoulis. 1997. Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *J. Virol.* 71:7951-7959.
286. Suzuki, Y., H.Yang, and R.Craigie. 2004. LAP2alpha and BAF collaborate to organize the Moloney murine leukemia virus preintegration complex. *EMBO J.* 23:4670-4678.
287. Szymczak, A.L., C.J.Workman, Y.Wang, K.M.Vignali, S.Dilioglou, E.F.Vanin, and D.A.Vignali. 2004. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* 22:589-594.
288. Takeuchi, Y., F.L.Cosset, P.J.Lachmann, H.Okada, R.A.Weiss, and M.K.Collins. 1994. Type C retrovirus inactivation by human complement is determined by both the viral genome and the producer cell. *J. Virol.* 68:8001-8007.

289. Takeuchi, Y., C.D.Porter, K.M.Strahan, A.F.Preece, K.Gustafsson, F.L.Cosset, R.A.Weiss, and M.K.Collins. 1996. Sensitization of cells and retroviruses to human serum by (alpha 1-3) galactosyltransferase. *Nature* 379:85-88.
290. Tanese, N., A.Telesnitsky, and S.P.Goff. 1991. Abortive reverse transcription by mutants of Moloney murine leukemia virus deficient in the reverse transcriptase-associated RNase H function. *J. Virol.* 65:4387-4397.
291. Temin, H. 1964. Homology between RNA from Rous Sarcoma Virus and DNA from Rous Sarcoma Virus-infected cells. *Proc. Natl. Acad. Sci. U. S. A* 52:323-329.
292. Temin, H., and S.Mizutani. 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226:1211-1213.
293. Terwilliger, E.F., E.A.Cohen, Y.C.Lu, J.G.Sodroski, and W.A.Haseltine. 1989. Functional role of human immunodeficiency virus type 1 vpu. *Proc. Natl. Acad. Sci. U. S. A* 86:5163-5167.
294. Thyagarajan, B., B.L.Johnson, and C.Campbell. 1995. The effect of target site transcription on gene targeting in human cells in vitro. *Nucleic Acids Res.* 23:2784-2790.
295. Tollefson, A.E., J.S.Ryerse, A.Scaria, T.W.Hermiston, and W.S.Wold. 1996. The E3-11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: characterization of cells infected with adp mutants. *Virology* 220:152-162.
296. Tratschin, J.D., I.L.Miller, and B.J.Carter. 1984. Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. *J. Virol.* 51:611-619.
297. Trono, D. 1992. Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J. Virol.* 66:4893-4900.
298. Valentine, RC., and H.G.Pereira. 2003. Antigens and structure of the adenovirus. Reprinted from *J. Mol. Biol.* 1965; 13: 13-20. *Rev. Med. Virol.* 13:71-82.
299. Valsamakis, A., S.Zeichner, S.Carswell, and J.C.Alwine. 1991. The human immunodeficiency virus type 1 polyadenylation signal: a 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc. Natl. Acad. Sci. U. S. A* 88:2108-2112.
300. Van Beveren, C., J.G.Goddard, A.Berns, and I.M.Verma. 1980. Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5' long terminal

- repeat and adjacent cellular sequences. *Proc. Natl. Acad. Sci. U. S. A* 77:3307-3311.
301. Van Der Ploeg, L.H., A.Konings, M.Oort, D.Roos, L.Bernini, and R.A.Flavell. 1980. gamma-beta-Thalassaemia studies showing that deletion of the gamma- and delta-genes influences beta-globin gene expression in man. *Nature* 283:637-642.
302. van Gent, D.C., J.H.Hoeijmakers, and R.Kanaar. 2001. Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.* 2:196-206.
303. Vargas, J., Jr., M.E.Klotman, and A.Cara. 2008. Conditionally replicating lentiviral-hybrid episomal vectors for suicide gene therapy. *Antiviral Res.* 80:288-294.
304. Varmus, H.E., R.V.Guntaka, W.J.Fan, S.Heasley, and J.M.Bishop. 1974. Synthesis of viral DNA in the cytoplasm of duck embryo fibroblasts and in enucleated cells after infection by avian sarcoma virus. *Proc. Natl. Acad. Sci. U. S. A* 71:3874-3878.
305. Varmus, H.E., S.Heasley, H.J.Kung, H.Oppermann, V.C.Smith, J.M.Bishop, and P.R.Shank. 1978. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. *J. Mol. Biol.* 120:55-82.
306. Vogt, V.M. 1996. Proteolytic processing and particle maturation. *Curr. Top. Microbiol. Immunol.* 214:95-131.
307. Weinberg, J.B., T.J.Matthews, B.R.Cullen, and M.H.Malim. 1991. Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J. Exp. Med.* 174:1477-1482.
308. Weitzman, M.D., S.R.Kyostio, R.M.Kotin, and R.A.Owens. 1994. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proc. Natl. Acad. Sci. U. S. A* 91:5808-5812.
309. Welker, R., M.Harris, B.Cardel, and H.G.Krausslich. 1998. Virion incorporation of human immunodeficiency virus type 1 Nef is mediated by a bipartite membrane-targeting signal: analysis of its role in enhancement of viral infectivity. *J. Virol.* 72:8833-8840.
310. Wells, D.J. 2004. Gene therapy progress and prospects: electroporation and other physical methods. *Gene Ther.* 11:1363-1369.

311. White, H., A.Thrasher, P.Veys, C.Kinnon, and H.B.Gaspar. 2000. Intrinsic defects of B cell function in X-linked severe combined immunodeficiency. *Eur. J. Immunol.* 30:732-737.
312. Whiteside, S.T., and A.Israel. 1997. I kappa B proteins: structure, function and regulation. *Semin. Cancer Biol.* 8:75-82.
313. Williams, S., T.Mustoe, T.Mulcahy, M.Griffiths, D.Simpson, M.Antoniou, A.Irvine, A.Mountain, and R.Crombie. 2005. CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. *BMC. Biotechnol.* 5:17.
314. Williams, S.A., L.F.Chen, H.Kwon, C.M.Ruiz-Jarabo, E.Verdin, and W.C.Greene. 2006. NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO J.* 25:139-149.
315. Wills, J.W., R.V.Srinivas, and E.Hunter. 1984. Mutations of the Rous sarcoma virus env gene that affect the transport and subcellular location of the glycoprotein products. *J. Cell Biol.* 99:2011-2023.
316. Wiskerchen, M., and M.A.Muesing. 1995. Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J. Virol.* 69:376-386.
317. Withers-Ward, E.S., Y.Kitamura, J.P.Barnes, and J.M.Coffin. 1994. Distribution of targets for avian retrovirus DNA integration in vivo. *Genes Dev.* 8:1473-1487.
318. Wolberg, A.S., D.W.Stafford, and D.A.Erie. 1997. Human factor IX binds to specific sites on the collagenous domain of collagen IV. *J. Biol. Chem.* 272:16717-16720.
319. Wu, X., Y.Li, B.Crise, and S.M.Burgess. 2003. Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300:1749-1751.
320. Wu, Y., and J.W.Marsh. 2003. Early transcription from nonintegrated DNA in human immunodeficiency virus infection. *J. Virol.* 77:10376-10382.
321. Wu, Z., A.Asokan, and R.J.Samulski. 2006. Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol. Ther.* 14:316-327.

322. Xiao, X., J.Li, and R.J.Samulski. 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* 72:2224-2232.
323. Yaffe, D., and O.Saxel. 1977. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* 270:725-727.
324. Yamashita, M., and M.Emerman. 2005. The cell cycle independence of HIV infections is not determined by known karyophilic viral elements. *PLoS. Pathog.* 1:e18.
325. Yamashita, M., and M.Emerman. 2006. Retroviral infection of non-dividing cells: old and new perspectives. *Virology* 344:88-93.
326. 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 Res.* 30:740-748.
327. Yanez-Munoz, R.J., K.S.Balaggan, A.MacNeil, S.J.Howe, M.Schmidt, A.J.Smith, P.Buch, R.E.MacLaren, P.N.Anderson, S.E.Barker, Y.Duran, C.Bartholomae, C.von Kalle, J.R.Heckenlively, C.Kinnon, R.R.Ali, and A.J.Thrasher. 2006. Effective gene therapy with nonintegrating lentiviral vectors. *Nat. Med.* 12:348-353.
328. Yee, J.K., J.C.Moores, D.J.Jolly, J.A.Wolff, J.G.Respess, and T.Friedmann. 1987. Gene expression from transcriptionally disabled retroviral vectors. *Proc. Natl. Acad. Sci. U. S. A* 84:5197-5201.
329. Yee, S.P., and P.W.Rigby. 1993. The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* 7:1277-1289.
330. Yoshinaka, Y., I.Katoh, T.D.Copeland, and S.Oroszlan. 1985. Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proc. Natl. Acad. Sci. U. S. A* 82:1618-1622.
331. Yu, H., A.B.Rabson, M.Kaul, Y.Ron, and J.P.Dougherty. 1996. Inducible human immunodeficiency virus type 1 packaging cell lines. *J. Virol.* 70:4530-4537.
332. Yu, S.F., T.von Ruden, P.W.Kantoff, C.Garber, M.Seiberg, U.Ruther, W.F.Anderson, E.F.Wagner, and E.Gilboa. 1986. Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. U. S. A* 83:3194-3198.

333. Yu, X., X.Yuan, Z.Matsuda, T.H.Lee, and M.Essex. 1992. The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J. Virol.* 66:4966-4971.
334. Yuan, X., X.Yu, T.H.Lee, and M.Essex. 1993. Mutations in the N-terminal region of human immunodeficiency virus type 1 matrix protein block intracellular transport of the Gag precursor. *J. Virol.* 67:6387-6394.
335. Zaiss, A.K., S.Son, and L.J.Chang. 2002. RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. *J. Virol.* 76:7209-7219.
336. Zaitseva, L., R.Myers, and A.Fassati. 2006. tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes. *PLoS. Biol.* 4:e332.
337. Zavada, J. 1982. The pseudotypic paradox. *J. Gen. Virol.* 63 (Pt 1):15-24.
338. Zennou, V., C.Petit, D.Guetard, U.Nerhbass, L.Montagnier, and P.Charneau. 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* 101:173-185.
339. Zhang, F., S.I.Thornhill, S.J.Howe, M.Ulaganathan, A.Schambach, J.Sinclair, C.Kinnon, H.B.Gaspar, M.Antoniou, and A.J.Thrasher. 2007a. Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. *Blood* 110:1448-1457.
340. Zhang, W., and M.J.Imperiale. 2000. Interaction of the adenovirus IVa2 protein with viral packaging sequences. *J. Virol.* 74:2687-2693.
341. Zhang, X.B., J.L.Schwartz, R.K.Humphries, and H.P.Kiem. 2007b. Effects of HOXB4 overexpression on ex vivo expansion and immortalization of hematopoietic cells from different species. *Stem Cells* 25:2074-2081.
342. Zheng, R., T.M.Jenkins, and R.Craigie. 1996. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci. U. S. A* 93:13659-13664.
343. Zhu, K., C.Dobard, and S.A.Chow. 2004. Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. *J. Virol.* 78:5045-5055.

344. Zhu, Z.H., S.S.Chen, and A.S.Huang. 1990. Phenotypic mixing between human immunodeficiency virus and vesicular stomatitis virus or herpes simplex virus. *J. Acquir. Immune. Defic. Syndr.* 3:215-219.
345. Zolotukhin, S., B.J.Byrne, E.Mason, I.Zolotukhin, M.Potter, K.Chesnut, C.Summerford, R.J.Samulski, and N.Muzyczka. 1999. Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.* 6:973-985.
346. Zufferey, R., J.E.Donello, D.Trono, and T.J.Hope. 1999. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.* 73:2886-2892.
347. Zufferey, R., T.Dull, R.J.Mandel, A.Bukovsky, D.Quiroz, L.Naldini, and D.Trono. 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* 72:9873-9880.
348. Zufferey, R., D.Nagy, R.J.Mandel, L.Naldini, and D.Trono. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* 15:871-875.
349. Zychlinski, D., A.Schambach, U.Modlich, T.Maetzig, J.Meyer, E.Grassman, A.Mishra, and C.Baum. 2008. Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol. Ther.* 16:718-725.