Development of a viral and a non-viral based gene transfer systems using the yeast Saccharomyces cerevisiae

A thesis submitted for the degree of Doctor of Philosophy by

Jonathan Kirk Bowden

September 2016

Abstract

VSV-G has been used for several years to pseudotype reteroviral and lentiviral vectors to increase the range of cell types that these vectors can be targeted to as well as increasing transfection efficiency and serum resistance. It has previously been shown that purified VSV-G protein can be added to several types of non-viral complexes to produce these same advantages. VSV-G therefore holds great potential in gene therapy for both viral and non-viral vectors.

Due to the cellular toxicity of VSV-G in mammalian cells VSV-G pseudotyped viral vectors are generally produced from transiently transfected cells which greatly limit the scale of viral production. VSV-G for non-viral vectors is also limited in the same manner but also suffer from expensive and time consuming methods to purify the VSV-G from the expression media.

To address these problems with production we attempted to generate strains of the yeast *Saccharomyces cerevisiae* that can produce VSV-G pseudotyped lentivirus and VSV-G protein from inducible integrated vectors. We theorised that the cell wall of *Saccharomyces cerevisiae* would prevent syncytia and cellular toxicity of VSV-G during production, allowing the continuous production of virus or protein.

In this report we show that this new production method allows us to produce and purify VSV-G from yeast using simple and scalable methods and that this produces a greater enhancement of transfection efficiency than mammalian derived VSV-G. However we were not able to demonstrate the production of VSV-G pseudotyped virus, seemingly due to the genotoxic effects of viral integrase.

Acknowledgements

The work described in these pages would not have been possible without the help and support of my colleagues, friends and family. I would like to thank my supervisor, Mike Themis, for his guidance over the last four years as well as his invaluable scientific advice.

I would also like to thank the technical staff of the biosciences department at Brunel University for their help in sourcing equipment and reagents and especially to Mathew Themis, without whom many of my experiments would have been impossible.

Lastly I would like to thank my family and friends. You helped keep a roof over my head, food in my stomach and tea in my thermos. Without you I wouldn't have even been able to start, let alone finish, this work.

Thank you all.

Contents

List of Figures and Tables

Abbreviations

AAV	Adeno associated virus
BLAST	Basic local alignment search tool
bp	Base pair
CA	Capsid
cDNA	Complementary deoxyribonucleic acid
CIP	Calf intestinal phosphatase
CMV	Cytomegalovirus
cPPT	Central polypurine tract
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphospates
DOPE	Dioleoyl-L-α-phosphatidylethanolamine
DOTMA	N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride
ELISA	Enzyme linked immunosorbent assay
FACS	Fluorescent activated cell sorting
GFP	Green fluorescent protein
HBV	Hepatitis B virus
HBx	Hepatitis B virus X protein
HCC	Hepatocellular carcinoma
hFIX	Human clotting factor IX
HIV	Human immunodeficiency virus
HSV	Herpes simplex virus
IN	Integrase
IU	Infectious unit
kb	Kilobase pairs

kDa	Kilodaltons
LB	Luria-Bertani
LTR	Long terminal repeat
MA	Matrix
$MgCl_2$	Magnesium chloride
mRNA	Messenger ribonucleic acid
NC	Nucleocapsid
NLS	Nuclear localisation signal
OD600	Optical density at 600 nanometres
ORF	Open reading frame
PBS	Primer binding site
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEI	Polyethyleneimine
polyA	Polyadenylation
PPT	Polypurine tract
PR	Protease
RNA	Ribonucleic acid
rpm	Rotations per minute
RRE	Rev response element
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase polymerase chain reaction
SCID	Severe combined immunodeficiency
SFFV	Spleen focus forming virus
SIN	Self-inactivating
siRNA	Small interfering ribonucleic acid

tRNA	Transfer ribonucleic acid
VSV-G	Vesicular stomatitis virus envelope glycoprotein
WPRE	Woodchuck hepatitis virus posttranscriptional regulatory element
YP	Yeast Peptone
YPD	Yeast Peptone Dextrose
YPG	Yeast Peptone Galactose
YPR	Yeast Peptone Raffinose

1. Introduction

1.1. Gene Therapy

Genetic disorders are a diverse group of diseases directly caused by abnormalities in the host genome leading to abnormal expression of genes. Gene therapy aims to treat or correct these abnormalities by the introduction of genetic material into the host cell to either add genes to replace non-functional or missing host proteins or to modulate the expression of host genes. Gene therapy may be applied to many simple monogenic disorders by the introduction of a functional copy of the faulty gene and in the future may be used to treat more complex polygenic disorders. Figure 1 shows the distribution of disorders amenable to gene therapy and the number of clinical trials that have been performed between 1989 and 2015.

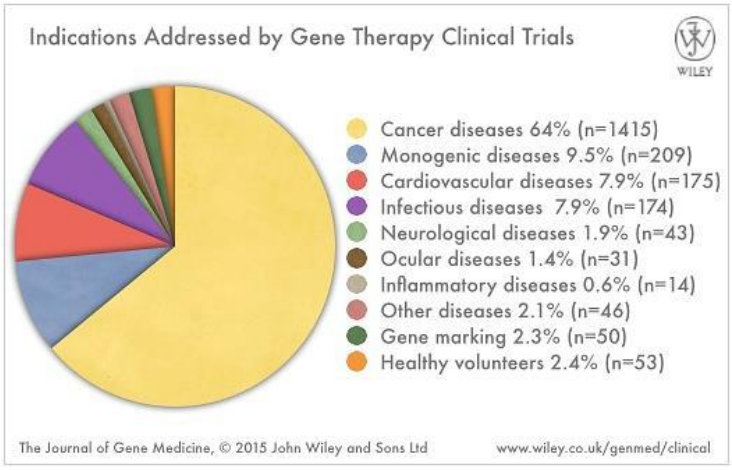

Figure 1. Pie chart showing the distribution of disorders that have been treated in gene therapy trials

Data taken from The Journal of Gene Medicine Clinical Trial site (http://www.abedia.com/wiley/index.html)

For gene therapy to be successful in the clinic several goals need to be achieved; Delivery of therapeutic exogenous genetic material into the correct target population of cells, to reach therapeutic levels of protein needed for treatment or even correction, not to be deleterious to cell survival, to avoid immune-rejection of the therapeutic vector

used to deliver the gene or protein product and not to cause long term side effects or death to the host. To reach target organs and cells, an array of vectors have been designed and developed that are able to carry and deliver therapeutic genes. These vectors can be categorized as either non-viral or viral based and have different properties suited to the correction of different disorders. Organisms and their cells have evolved a host of strategies that prevent exogenous genetic material from entering cells and establishing residence. These defences make the therapeutic application of genetic material problematic. It is not surprising therefore, that many different gene therapy vectors have been developed and tested in the past 30 years to overcome these defences. Before treatment, there is a choice of vector designed to suit each intended application in order to achieve successful gene delivery and a therapeutic outcome. Each of these vectors has different advantages and disadvantages (summarised in Table 1) to suit their clinical use.

Table 1. Properties of commonly used gene therapy vectors

Vector	Maximum insert size	Integration	Duration of expression	Genotoxicity	Inflammatory response
Naked DNA/RNA	No limit	No	Transient	Very low to none	Low to none
Chemical Carriers	No limit	No	Transient	Very low to none	Low to none
Hybrid Methods	No limit	No	Transient	Very low to none	Low
Adenovirus	Upto 30kb	Low frequency	Transient	Low	High
Adeno-associated virus	Upto 4.7kb	Yes	Long Term	Low	Low
Herpes simplex virus	>30kb	No	Transient	Low	Low
Retrovirus	Upto 7kb	Yes	Long Term	High	Low
Lentivirus	Upto 8kb	Yes	Long Term	Medium	Low

1.2. Non-Viral gene therapy Vectors

Non-viral gene therapy vectors are a diverse group of pharmacological agents capable of transferring a wide range of nucleic acids. They offer distinct advantages over viral vectors due to their low immunogenicity and reduced potential for mutagenicity. However, many non-viral vectors suffer from high cytotoxicity and low transfer efficiency. Gene expression of non-viral vectors is normally transient which can be viewed as an advantage or disadvantage depending on its intended use. These vectors have been used in a number of clinical trials as shown in Table 2.

Table 2. Types of non-viral gene therapy vectors used in clinical trials

Vector	No. of trials	%
Naked DNA/RNA	446	77.70%
Chemical Carriers	115	20.03%
Hybrid Methods	13	2.26%
Total	574	

Data taken from The Journal of Gene Medicine Clinical Trial site (http://www.abedia.com/wiley/index.html)

1.2.1. Naked DNA gene delivery

The most basic non-viral gene transfer is the direct application of naked (uncomplexed) DNA to cells. Many cell types can be transfected this way such as melanoma cells (Vile & Hart 1993), liver hepatocytes (Hickman *et al*. 1994) and cardiomyocytes of the heart (Ardehali *et al*. 1995) however gene transfer is very inefficient with only 1% of cells being successfully transfected (Wolff *et al*. 1991). The low efficiency of naked DNA transfection *in vivo* is believed to be due to the presence of endonucleases in the cells and the extracellular space which rapidly degrade the DNA within only 10 minutes (Kawabata *et al*. 1995). Due to this low gene transfer efficiency very few clinical studies

have been performed and with little real success. In one study where patients with chronic ischemic neuropathy were treated via intramuscular injections of plasmid DNA carrying the gene for vascular endothelial growth factor (VEGF) showed a small but significant improvement of their symptoms due to revascularization of the tissues (Simovic *et al*. 2001). While studies like these are encouraging, alternative gene transfer regimes were required to increase transfection rates to achieve a therapeutic and long-lasting outcome to patients in the clinic.

Electroporation is one of the earliest methods used in *in vitro* gene transfer (Neumann *et al*. 1982). This method uses a pulse of electric across the cell membrane resulting in the formation of transient pores through which the DNA can pass. Whilst this method has been used *in vivo* the need to insert electrodes and precisely control the magnitude and frequency of the electric pulses for each tissue type limits the possible uses for this technique clinically (Trezise 2002). A similar method useful for membrane permeabilization applies pulses of ultrasound to cells (Gambihler *et al*. 1994). This technology has been primarily used *in vitro*, however the potential for future clinical application is promising due to the relative ease of use of ultrasound compared with the method of electroporation.

1.2.2. Delivery of DNA complexed with chemical carriers

While naked DNA delivery has had limited success in the past, as mentioned this method suffers from low gene transfer efficiency. Complexing DNA with chemical carriers, on the other hand, provides higher transfection efficiency and have therefore been developed extensively over the past 25 years (Yin *et al*. 2014). There are two main types of chemical carriers, these are polyplex and lipid-based. Both of these agents are able to condense DNA into small particles and in doing so protect the DNA

from degradation by nucleases present in the bloodstream of an organism (Rolland & Mumper 1998; Mumper *et al*. 1998).

Polyplexes

One of the first polyplexes reported to enhance the transfection of mammalian cells was the compound poly-L-lysine (PLL) which was originally shown to be able to bind to and condense DNA (Olins *et al*. 1967). While not capable of transfecting DNA on its own, PLL can be coupled with other compounds to target specific receptors leading to absorption via receptor-mediated endocytosis. An example of this is the coupling of PLL to asialoglycoprotein to target DNA to asialoglycoprotein receptor-bearing cells (Wu & Wu 1987). This makes PLL one of the few non-viral compounds capable of specific cell targeting although the range of cell types successfully targeted is limited.

One of the most studied polymeric compounds used in gene transfer is polyethylenimine (PEI) (Figure 2). PEI is a polymer of one amine group and two carbon aliphatic spacers in either linear or branched forms. This structure leads PEI to have a mild positive charge that allows it to bind to and condense DNA into small particles like PLL. Unlike PLL however, the overall positive charge on PEI allows it to bind to negatively charged cell membranes followed by entry into the cell via endocytosis. PEI also has the advantage of offering high buffering capacity due to the large number of protanable amines (Boussif *et al*. 1995; Lungwitz *et al*. 2005) that results in neutralising the acidic environment within endosomes thus preventing degradation of the PEI/DNA complex. This also leads to influx of negatively charged chloride ions, increased osmotic pressure and rupture of the endosome with release of the PEI/DNA complex into the cytoplasm (Sonawane *et al*. 2003). So far only one clinical trial has been performed using PEI for gene transfer to pancreatic cancer cells (Buscail *et al*. 2015). This phase 1 study found that administration of PEI complexed with a vector

expressing the SSTR2, DCK and UMK genes into advance tumours was well tolerated with limited side effects and that 12 out of 13 patients treated remained free of metastasis following gene therapy (92%)

Figure 2. The chemical structures of poly-L-lysine monomer (left) and polyethylenimine monomer (right)

Lipid based

One of the first successful lipid-based vectors was phosphatidylserine (Fraley *et al*. 1980) a phospholipid that was used to transfect monkey kidney cells with viral SV40 DNA. The preparation phosphatidylserine liposomes encapsulating DNA is complex and requires several rounds of purification, concentration and sonication to produce a small amount of usable vector. This problem was addressed by using an alternative cationic lipid transfection reagent N-[1-(2,3-dioleyloxy)propylJ-N,N,N-trimethylammonium chloride (DOTMA) (Felgner *et al*. 1987). This compound was shown to spontaneously encapsulate DNA there by reducing the preparation of the vector to simply mixing purified DNA with DOTMA in solution. Spontaneously encapsulated DNA is then attracted electrostatically to negatively charged DNA molecules via the positively charged DOTMA cation. This electrostatic attraction between the DOTMA/DNA complex and the negatively charged cell membrane assists with membrane binding and internalisation and ultimately increases transfection efficiency.

Whilst there have been many advances in cationic lipid formulations that have increased the effectiveness of these reagents, they are still ineffective when used *in*

vivo due to poor stability, rapid clearance and inflammatory responses induced by the vectors (Lonez *et al*. 2008; Whitehead *et al*. 2009). Positively charged liposomes naturally bind to negatively charged compounds including negatively charged serum proteins. These interactions can cause reduced cell membrane interaction (Li *et al*. 1999), aggregation and clearance by macrophages (Morille *et al*. 2008) and vector disintegration (Wiethoff *et al*. 2001). Liposome base gene therapy vectors have been much more successful at reaching clinical trials with at least one phase III trial having recently been completed using Allovectin-7 (a plasmid/lipid complex containing the DNA sequences encoding HLA-B7 and ß2 microglobulin) (Bedikian & Del Vecchio 2008). Allovectin-7 increases the expression of major histocompatibility complex class I from transfected cells leading to an increased ability of the immune system to target and kill these cells. In this phase 3 trial Allovectin-7 plus dacarbazine was shown to increase the median survival durations of treated patients to 10.75 months compared to 9.24 months on dacarbazine alone.

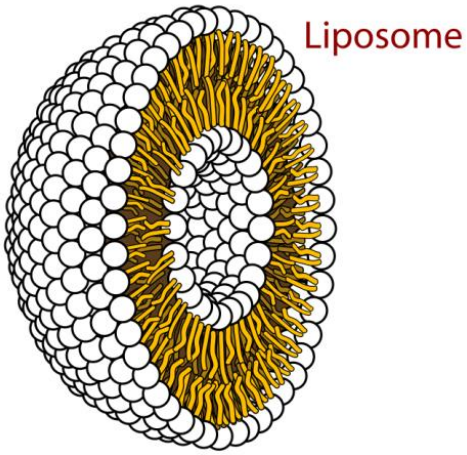

Figure 3. Cross-sectional view of the structure of a liposome vector.

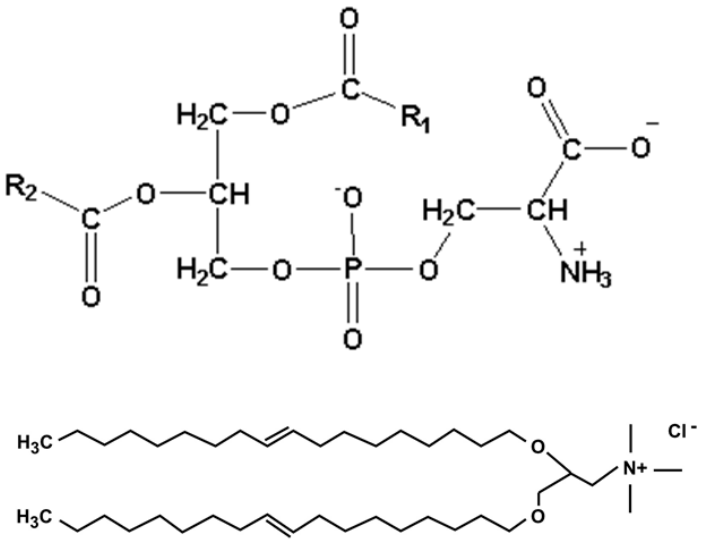

Figure 4. The chemical structures of phosphatidylserine monomer (top) and DOTMA (bottom)

1.2.3. Hybrid methods

There are three main barriers to achieving effective transfection by non-viral vectors into cells. The vector system must be capable of efficient cellular binding, internalization and endosomal escape (Pouton *et al*. 2001; Bally *et al*. 1999). Hybrid methods have been developed with viral components incorporated into non-viral formulations. These have been shown to overcome some of the limitations to non-viral based gene transfer alone. The majority of hybrid methods use envelope proteins from a viral source complexed to DNA to increase cellular binding or proteins containing nuclear import signals to increase expression.

An example of hybrid vectors is the use of adenoviral protein hexon in PEI/DNA complexes (Carlisle *et al*. 2001) to increase nuclear delivery and transgene expression in target cells. In this method hexon protein has to be covalently bonded to PEI before being mixed with DNA to form complexes. While this did increase the transfection efficiency, the need to covalently bond hexon to PEI limits the production level of these complexes. PEI and purified hexon protein needs to be modified with 3-(2-Pyridyldithio)propionic acid N-hydroxysuccinimide ester (SPDP) to make then capable

of forming a disulphide bond. After being bonded the PEI-hexon, conjugates have to be repurified before they can be used making this process less cost effective.

Another example of a hybrid method is the complexing of DNA to the VSV-G envelope glycoprotein. Unlike hexon protein, VSV-G does not need to be covalently conjugated to chemical carriers. As VSV-G is a lipid soluble protein it will naturally self-assemble into complexes with lipid based chemical carriers. It is also partially negatively charged allowing it to be complexed with positively charged chemical carriers such as PEI (Okimoto *et al*. 2001; Miyanohara 2012).

1.3. VSV-G enhances non-viral transfection

Many aspects of VSV-G make it a promising transfection reagent for non-viral gene therapy. VSV-G has been used for many years to pseudotype lentiviral and retroviral vectors to give these vectors broad target cell range tropism and hence fusion to many cell types including those normally refractory to other methods of gene transfer (Yee *et al*. 1994; Burns *et al*. 1993). VSV-G has therefore been shown to be effective at cellular binding (Wehland *et al*. 1982; Schlegel *et al*. 1982) and endosomal escape (Pastan & Willingham 1983) which are two of the main barriers to non-viral transfection. VSV-G has also been shown to be able to be purified and added to viral particles lacking viral envelope rendering them infectious by pseudotyping them (Abe, Chen, *et al*. 1998)

Many non-viral vectors, like polybrene and lipofectin, suffer from serum inactivation and clearance thereby significantly reducing their ability to aid DNA to efficacy transfect cells *in vivo* via introduction into the blood (Yang & Huang 1997; Escriou *et al*. 1998; Tandia *et al*. 2003). Conversely both polybrene and lipofectin complexes have been found to be resistant to serum inactivation when complexed with VSV-G *in vitro* (Abe, Miyanohara, *et al*. 1998) and *in vivo* (Hirano *et al*. 2002) potentially making these hybrid complexes far more effective in a clinical setting. VSV-G has also been complexed with polybrene (Okimoto *et al*. 2001), Lipofectin (Abe, Miyanohara, *et al*. 1998) and liposomes (Imazu *et al*. 2000; Shoji *et al*. 2004) to greatly increase these agents to assist vector transfection efficiency.

While these properties make VSV-G a good transfection reagent, its large-scale use is limited by the tedious and expensive production regime needed for this glycoprotein to be generated from mammalian cells via transfection. Production is also problematic due to the fusogenic nature of this glycoprotein, which can only be produced by

transient transfecting 293T cells. Cells expressing VSV-G form syncytia and die within 3-4 days post transfection and therefore every time the glycoprotein is required, repeat transfection of cells by VSV-G carrying plasmid in needed (Eslahi *et al*. 2001). This inefficient method of production requires an alternative approach to generate VSV-G glycoprotein. During VSV-G expression the glycoprotein is released into the cell culture media and VSV-G is purified by ultra-centrifugation (Miyanohara 2012). This produces protein with contaminating components of the culture medium, which is not suitable for transfection and especially not safe for clinical application. In addition, scale-up is difficult when using mammalian transfection protocols. None-the-less VSV-G may be able to provide a robust transfection agent that could deliver DNA at a far greater transfection frequency than chemical carriers alone.

1.4. Viral Gene Therapy Vectors

There are many barriers to efficient transfection of DNA that a good gene therapy vector must overcome to be effective. While there has been much research into non-viral methods to increase their efficiency they still suffer from low transfection efficiency. On the other hand there are many viruses that have evolved to overcome these barriers to transfect human cells. This section deals with viral gene therapy vectors that are derived from these viruses and have been engineered to deliver therapeutic DNA.

Table 3. Types of viral gene therapy vectors used in clinical trials

Vector	No. of trials	%
Adenovirus	503	40.43%
Adeno-associated virus	137	11.01%
Herpes simplex virus	73	5.87%
Retrovirus	417	33.52%
Lentivirus	114	9.16%
Total	1244	

Data from The Journal of Gene Medicine Clinical Trial site (http://www.abedia.com/wiley/index.html)

1.4.1. Adenovirus vectors

Adenoviruses are a large family of viruses consisting of many different subtypes. All adenoviruses are double stranded DNA viruses with regular icosahedral capsids and no plasma envelope. There have been 57 different serotypes that have been isolated from humans and another 27 serotypes from simians. They have 34-48kb linear genomes flanked by two origins of replication (Wides *et al*. 1987) called inverted terminal repeats (ITRs). The left hand ITR also contains a packaging signal to allow genomes to be packaged into virions (Hearing *et al*. 1987). The viral genome encodes

two classes of genes; early stage and late stage. Early stage genes (E1-E4) are expressed as soon as the virus enters the nucleus of the host cell. The E1a and E1b genes are the first to be transcribed and are required by the rest of the early stage genes for efficient transcription (Jones & Shenk 1979). E2a and E2b encode proteins that regulate the levels of adenoviral DNA replication throughout infection (Nevins & Winkler 1980). The E3 region produces four unique gene products that regulate host immune responses to adenovirus infection by supressing the expression of class 1 MHC antigens, tumour necrosis factor and epidermal growth factor receptor (Ginsberg *et al*. 1989). The E4 region has several effects on viral DNA production, translation and host cell lysis (Marcellus *et al*. 1998; Halbert *et al*. 1985; Huang & Hearing 1989). The late stage genes are transcribed together in the major late transcription unit (MLTU) from the major late promoter (MLP) approximately 8 hours after transfection (Le Moullec *et al*. 1983). The L1 gene product is required for the packaging of viral DNA into empty capsids (Gustin *et al*. 1996). The L2 region encodes the capsid protein penton (Zubieta *et al*. 2005) and L3 encodes the capsid protein hexon (Jörnvall *et al*. 1981) and viral protease (Webster & Kemp 1993). The L4 region encodes 2 major functions required for productive infection, a RNA splicing factor that regulates late mRNA splicing and a protein involved in viral assembly (Wu *et al*. 2013). Finally the L5 region encodes a fibre knob domain (Schoggins *et al*. 2003) that is part of the viral capsid and that mediates viral attachment to host cells via the coxsackievirus and adenovirus receptor (CAR) (Bergelson *et al*. 1997).

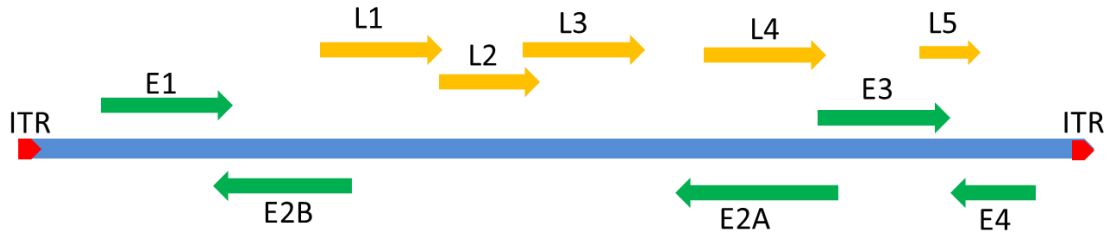

Figure 5. The structure and organisation of type 5 adenovirus genome

E represents early genes and L represents late genes. The ITR's are shown at the end of the vector in red.

To make adenoviruses viable gene therapy vectors several modifications have been made to their genome. The E1 region is required for viral replication so are deleted to make vectors replication deficient and the E3 region is involved in pathogenicity and is also deleted(Kelly & Lewis 1973; McConnell & Imperiale 2004). These modifications also reduce the size of the adenovirus genome which, because of the limited capacity of the capsid (34-48kb), allows larger transgenes to be inserted. Parts of the E2 and E4 regions are deleted to abrogate viral gene expression and therefore partially circumvents an immune response to cells transduced by the vector (Armentano *et al*. 1995; Gorziglia *et al*. 1996) and increases the duration of gene expression by these cells (Engelhardt *et al*. 1994; Hu *et al*. 1999). This also reduces the chance of replication competent virus forming from homologous recombination. While these modifications greatly reduce several problems associated with adenovirus vectors there is still a large amount of sequence homology and viral gene expression that could lead to adverse effects being caused by these vectors in the form of a strong immune response by the host against virus antigens being presented due to continued expression of viral genes (Nunes *et al*. 1999; Schnell *et al*. 2001). Strong innate immune responses can cause severe systemic reactions in treated patients leading to high fevers, systemic damage and even death (Raper *et al*. 2002).

In attempt to reduce these problems helper dependent adenoviral vectors (HDAd) have been generated that contain no viral sequences except the ITRs and packaging sequence on the virus backbone (Kochanek *et al*. 1996). HDAd production is achieved by transfecting producer cells with two linear DNA vectors. The first contains the full length adenovirus sequence with deletions of the packaging signal and E1a and E3 sequences. This produces adenoviral particles without any genome because of the lack of packaging signal needed for genome insertion into the capsid. The second DNA

vector contains the viral ITRs, viral packaging signal and the transgene of interest. This allows the packaging of only the second DNA vector with no viral protein sequences and greatly increases the capacity of the adenovirus vector. These HDAd vectors produce lower immunotoxicity allowing for high levels of transduction with long term expression of transgene; however innate immune reactions can still occur against the viral capsid proteins. Viral capsid proteins can activate Toll-like receptors 3, 7 and 9 which leads to increased expression of inflammatory cytokines and type I interferons (Yamaguchi et al. 2007). This innate response limits the amount of viral vector that can safely be used *in vivo* and could lead to severe to fatal reactions in patients if not administered correctly.

Due to the extensive development of adenovirus vectors since 1989, 28% of gene therapy trials have used these vectors, with five currently in phase III trials. However many of these trials have repeatedly highlighted the problem of the host immune response to these vectors and the resulting cytotoxicity caused (Raper *et al*. 2002).

1.4.2. Adeno-associated virus vectors

Another type of viral vector that has been adapted for gene therapy is the adeno-associated virus (AAV) vector. Similar to adenoviruses, AAVs are non-enveloped icosahedral capsid viruses with linear DNA genomes. Unlike adenoviruses, AAVs genomes are single stranded and only 4.7kb is size (Srivastava *et al*. 1983). The genome of AVVs are flanked by ITRs that are similar to those found at the 5' and 3' ends of adenovirus genomes, however these ITRs are unique in the fact that they fold over into T-shaped hairpins to act as self-priming sites for DNA synthesis (Bohenzky *et al*. 1988). The AVV genome is very compact with only 2 open reading frames encoding all the viral proteins. The Rep open reading frame encodes 4 regulatory proteins (Rep78, Rep68, Rep52 and Rep40) required for several functions of AVVs. Rep78

regulates the replication of the viral genome as well as taking part in viral integration (Smith & Kotin 2000) and inducing S phase arrest in infected cells (Berthet *et al*. 2005). Rep68 is a site specific endonuclease that, along with Rep78, is involved with the site specific integration of AAVs (Young *et al*. 2000). Finally Rep52 and Rep40 are 3'-5' DNA helicases that are required for accumulation and packaging of viral genome into capsids (King *et al*. 2001). The Cap open reading frame encodes 3 viral proteins (VP1, VP2 and VP3) that make up the icosahedral capsid of AAVs. VP1 and VP2 contain domains that affect host cell tropism, endosome escape and trafficking within the host cell (Sonntag *et al*. 2006).

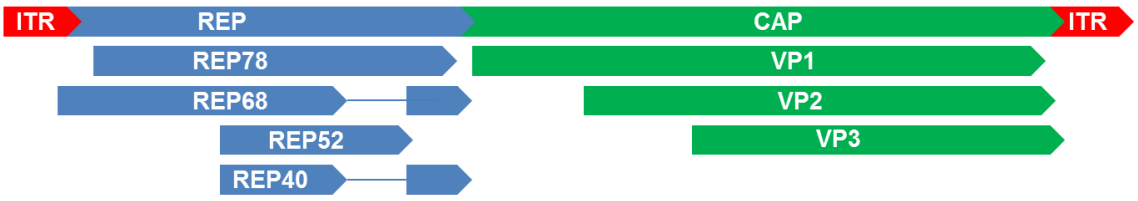

Figure 6. The structure and organisation of the adeno-associated virus genome

ITR's are represented in red, the Rep sequences in blue and the Cap sequences in green

For the production of AAV vectors for gene therapy the gene of interest can be packaged into the viral particle with just the flanking ITRs and the Cap and Rep genes are provided *in trans* (Samulski *et al*. 1989). This is an advantage for gene therapy vectors as it allows the integration of therapeutic DNA with a minimal amount of viral DNA and no viral genes are transferred. The integration of AAVs is targeted to specific site in human chromosome 19q13.42 named AAVS1 (Kotin *et al*. 1990). This site specific integration is achieved by the interaction of Rep68 and Rep78 proteins with Rep binding site (RBS) and terminal resolution site (TRS) found within the ITRs of the AVV genome and at several loci within the human genome (Hüser *et al*. 2010).

One of the major advantages of AAV vectors are their lower immunogenicity compared to adenovirus. AAV vectors do not elicit the robust type I IFN response as is seen for adenoviral vectors (McCaffrey et al. 2008) and nor do they induce strong cytokine responses in transduced tissues (Zhu et al. 2009).

As AAV vectors are a relatively new development in gene therapy very few clinical trials have been performed so far. The majority of these are phase I and II trials with only 8 phase III underway. These trials have been fairly successful however, while less severe than adenovirus vectors, AAV trials have shown a similar problem of immunotoxicity reducing efficiency and safety (Mingozzi & High 2011).

While the use of AAV vectors in the clinic has so far been limited, the advantages of AAVs have seen their development as delivery vector for CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats associated protein 9) genome-editing technologies. The CRISPR/Cas9 system is composed of 2 components that must be delivered together for its correct functioning, a DNA endonuclease capable of site specific double strand cleavage (Cas9) along with a guide RNA (gRNA) that targets the enzyme to a complementary DNA sequence (Jinek et al. 2012). CRISPR/Cas9 systems have seen much development in recent years for targeted genome-editing *in vitro* but there use in the clinic will require delivery methods capable of transducing cells *in vivo*. AAVs are a promising candidate for this type of delivery system due to their diverse tissue-targeting profiles, low pathogenicity and their ability to transduce both dividing and nondividing cells. It has been shown (Senís et al. 2014) that the coding sequence for Cas9 and a gRNA coding sequence can be incorporated into a AAV genome and successfully modify the miR122 target in a mouse liver. While this was a success the total size for the packaged genome was larger than the standard 4.7kb used with AAV vectors. In general oversized AAV vectors have shown very

inconsistent results (Lai et al. 2010; Dong et al. 2010) and so an alternative method has been developed. In a study by Swiech et al. 2014 the Cas9 and gRNA coding sequences were packaged into a separate AAV vectors and delivered together to the mouse brain. The study showed robust modification of a single (Mecp2) and multiple (Dnmt1, Dnmt3a and Dnmt3b) gene targets from this duel delivery method.

1.4.3. Herpes simplex virus vectors

Herpes simplex viruses (HSV) are a family of enveloped double stranded DNA viruses. They consist of four main structures:

- At their core is a ~152kb genome
- This core is then contained in an icosadeltahedral protein capsid
- The capsid is then connected to the viral envelope via a collection of matrix proteins called the tegument
- Finally the viral particle is encapsulated in a trilaminar lipid envelope embedded with viral glycoproteins

The viral genome is organised into two main segments called the long and short unique segments (U_L and U_S) (Perry & McGeoch 1988) encoding 84 viral genes. Approximately half of these genes are nonessential to viral production in tissue culture allowing a large amount of genome sequence to be removed and large (<20kB) therapeutic sequences to be delivered in these vectors.

HSV encodes several glycoproteins that mediate cell attachment and entry. The two surface glycoproteins gC and gB are the main mediators of viral attachment to cells via heparan sulfate proteoglycans on the host cell surface (Laquerre *et al*. 1998). This gives HSV a wide cell tropism that can be useful when a range of cell types need to be

transduced, however these genes can be deleted and replaced with other envelope proteins to pseudotype these vectors to specific cells (Anderson *et al*. 2000).

Once in the cell HSV can either undergo lytic infection or become latent. In lytic infections the viral proteins are expressed and the viral genome is replicated leading to the build-up of viral particles inside the host cell until they are released by cell lysis (Reviewed in, Everett 2000). For HSV to be used as a gene therapy vector the lytic activity of HSV needs to be eliminated in target cells. Three main methods have been developed to achieve this. The first method is called conditionally replicating vectors in which specific genes are deleted to allow the viral vector to replicate *in vitro* but not *in vivo*. The ICP34.5 gene can be deleted allowing HSV to replicate lytically in rapidly dividing cancer cells but not in non-dividing cells such as neurons (Andreansky *et al*. 1997).

The second method is the replication defective vectors in which ICP4 or ICP27 genes are deleted to produce a virus that is incapable of replicating. To produce these vectors producer cells are required that supply either ICP4 or ICP27 genes *in trans* to allow lytic replication of the vector (Marconi *et al*. 1996). This method allows the easy preparation of high-titer replication deficient vector but the vector has the potential to become replication competent if cells become infected with wild type HSV (Ozuer *et al*. 2002).

The final method is to supply almost the entire viral genome *in trans*. In this method the transgenes are placed in a plasmid containing the packaging signals and origin of replication of HSV while a bacterial artificial chromosome contains all the coding sequences of HSV without any packaging signals (Horsburgh *et al*. 1999; Suter *et al*. 1999). This method produces virus vectors that contain no viral encoded proteins and

which cannot become replication competent in target cells however they suffer from difficulties in producing high yield clinical grade viral preparations.

1.4.4. Retrovirus and lentivirus vectors

Retroviruses are a family of viruses with the defining feature of being able to stably integrate their genome into the genome of their host organism. They consist of 7–12 kb positive sense single stranded RNA genomes encoding a protein capsid, envelope glycoprotein and reverse transcription/integration enzymes. Retroviruses are obligate parasites that require the transcription and translation machinery of a host cell to replicate and complete their life cycle. The retrovirus genome encodes three main functional polyproteins group-specific antigen (*gag)*, polymerase (*pol)* and envelope (*env)* and the lentivirus genome also encodes *gag*, *pol* and *env* and many other accessory proteins.

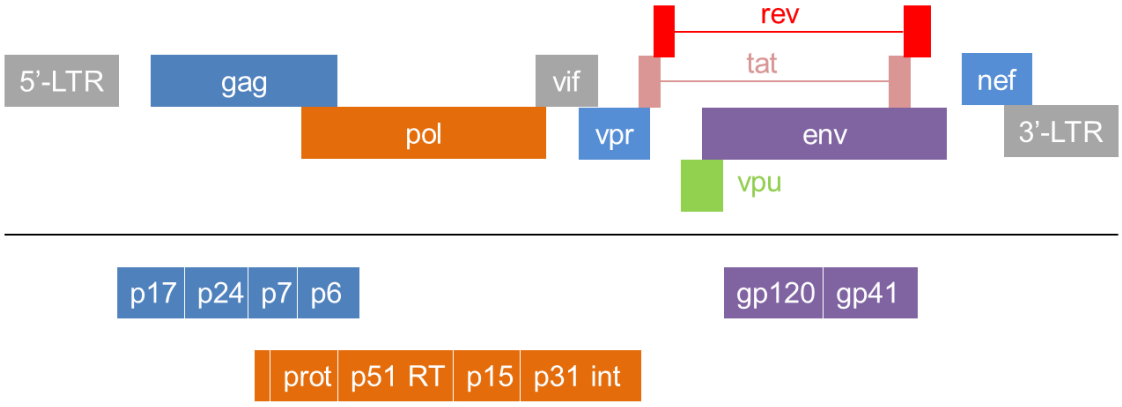

Figure 7. The structure and organisation of the lentiviral genome (top) and the division of the gag, pol and env genes into their separate proteins (bottom)

The function of the each gene is described in the following text

The *gag* gene encodes the major components of the viral capsid as a single 55kDa precursor that is cleaved by viral protease into its functional parts. The first component is the matrix protein p17 attaches to the lipid envelop of budded virus to stabilise the

viral particle and also helps mediate nuclear import in infected cells (Gallay *et al*. 1995). The p24 capsid protein is also structural and forms the central core of the viral particle (Zhang *et al*. 1996). The p24 protein recruits the host protein cyclophilin A during virion formation to stabilise the viral capsid (Liu et al. 2016). The nucleocapsid protein p7 contains two zinc-finger motifs that recognise the packaging signal within the viral genome allowing packaging of the viral genome into forming particles (Lapadat-Tapolsky *et al*. 1993). The final part of the gag protein is the p6 region which is essential for successful budding (Solbak *et al*. 2013) and incorporation of the Vpr protein into the budding virion (Paxton *et al*. 1993). During Viral budding p6 recruits the host protein TSG101 (tumor susceptibility gene 101) which targets the virion to ESCRT-1 (endosomal sorting complexes required for transport) complex. ESCRT-1 then initiates the formation of multivesicular bodies and viral budding.

The Pol genes are expressed as a Gag-Pol polyprotein encoding the Gag proteins and the four Pol proteins. Viral protease p10 is an aspartyl protease that cleaves all of the Gag and Gag-Pol polyproteins into their functional components (Ashorn *et al*. 1990). The p51 region encodes one of the most important proteins to the viral life cycle, reverse transcriptase, which converts the viral RNA genome into DNA ready for integration (Zack *et al*. 1990) with the help of p15 RNase H (Becerra *et al*. 1990). Finally the viral integrase p31 mediates the attachment, site selection and integration of the viral DNA genome into the host (Bushman *et al*. 1990).

The two envelope proteins gp120 and gp41 are expressed as a single 160kDa sequence separately from Gag-Pol. The precursor protein is then cleaved by the cellular protease Furin into its components. The gp120 incorporates into the plasma envelop of viral particles and mediates the binding of virus to CD4 receptors (Paxton *et al*. 1993) and the CXCR4 and CCR5 co-receptors (Deng *et al*. 1996). The fusogenic

domain of gp41 then allows the release of the viral particle into the host cell (Camerini & Seed 1990).

The accessory proteins of lentiviruses consist of Rev, Tat (transcriptional transactivator), Nef (negative factor), Vpr (Viral Protein R) and Vif (Viral infectivity factor). Rev is a sequence-specific RNA binding protein that is required for the export of unspliced viral genomic RNA from the nucleus to the cytoplasm. Rev binds to 240-base region called the Rev response element (RRE) within the viral genome and targets the complex to the Exportin 1 (XPO1) nuclear export receptor via a leucine-rich nuclear export signal (NES) in the Rev protein (Fischer et al. 1995). The Tat viral protein is essential for lentiviral replication (Ruben et al. 1989) where it acts to promote the production of full-length viral transcripts. Tat recruits Cyclin T1 and CDK9 (Cyclin-dependent kinase 9) to phosphorylate the carboxylterminal domain of RNA polymerase II leading to full length transcripts of viral RNA (Tiley et al. 1992).

Nef is a 27-kD myristoylated protein that greatly increases HIV infectivity and disease progression in infected hosts. The expression of Nef in infected cells leads to an increased rate of CD4 and MHC (major histocompatibility complex) Class I endocytosis and lysosomal degradation (Aiken et al. 1994; Schwartz et al. 1996). The lower levels of CD4 cause an increased level of Env incorporation and virion budding while lower MHC Class I reduces the efficiency of the killing of HIV infected cells by cytotoxic T cells. The viral protein Vpr is essential for viral genome integration into non-dividing cells. Vpr is incorporated into viral particles via specific interactions with the carboxyl-terminal region of Gag (Zack et al. 1990) and is released into infected cells with the viral genome. Once in the infected cells Vpr combines with the preintegration complex (PIC) and helps import the complex through the nuclear pore via a nuclear localization signal (Heinzinger et al. 1994). The accumulation of Vpr expressed in infected cells

also blocks the activation of the p34cdc2/cyclin B complex, a key regulator of the cell cycle, leading to cell cycle arrest at the G2 phase (Jowett et al. 1995). Vif is a 23-kDa polypeptide that is essential for the replication of HIV in certain cell lines (Strebel et al. 1987). In infected blood lymphocytes and macrophages a host enzyme called APOBEC3G (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G) is incorporated into virions where they interfere with reverse transcription by inducing numerous deoxycytidine to deoxyuridine mutations and rendering the viral particle non-infectious (Donahue et al. 2008). The expression of Vif in these cells prevents the incorporation of APOBEC3G into viral particles and promotes destruction of the enzyme (Stopak et al. 2003).

The ability of retroviruses to stably integrate exogenous genetic material into the host genome was first described in 1981 (Wei *et al*. 1981; Shimotohno & Temin 1981) were herpes simplex virus thymidine kinase gene was integrated into the genome of NIH-3T3 TK- cells with the Harvey murine sarcoma virus to produce TK+ transformants. This was the first evidence that retroviruses could be used for gene delivery. Soon after retroviruses like murine leukaemia virus (MLV) were being modified for gene delivery (Mann *et al*. 1983), they suffered from one drawback in that they remained self-replicating in infected cells. This remained a problem of retroviral vectors until it was discovered that the processes of cell entry and genome integration did not require viral gene expression in the host cell. This allowed the production of viral vectors that did not code for these viral sequences in there genome by providing these viral components *in trans* during viral assembly in the producer cells (Miller & Rosman 1989).

With viable vectors for gene transfer being developed it wasn't long before clinical trials using vectors designed to correct genetic diseases were underway. Over the years

hundreds of clinical trials have been performed on retroviral and lentiviral vectors against a variety of disease targets (summarised in Table 4). One of the earliest of these trials integrated the adenosine deaminase (ADA) gene via a retroviral vector into the T-cells of patients suffering from severe combined immunodeficiency (SCID), a disease caused by a faulty ADA gene (Anderson *et al*. 1990). Whilst successful at first it was soon found that these simple retroviral vectors had drawbacks with an inability to infect non-dividing cells (Roe *et al*. 1993) and the ability to cause oncogenesis and insertional mutagenesis of host genes (Nienhuis *et al*. 2006). The LTRs of retroviruses contain strong and ubiquitously active promoter and enhancer elements that drive rapid viral production in infected cells. In the process of gene therapy the viral genome can integrate into the host genome in or around proto-oncogenes. Depending on the gene and site of integration this can lead to unregulated expression of full or truncated oncogenes from the viral LTR (van Lohuizen & Berns 1990). In addition, to this viral integration can inactivate genes however this is normally a recessive event unless the loci effected exists as a single copy haploid gene (King *et al*. 1985).

Table 4. Number of gene therapy trials registered for Retroviral and Lentiviral Vectors

Targets of Therapy	No. of Retroviral Trials	No. of Lentiviral Trials
Cancer diseases	281	74
Cardiovascular diseases	3	1
Infectious diseases	44	18
Inflammatory diseases	5	0
Monogenic diseases	54	41
Neurological diseases	2	4
Ocular diseases	1	4
Others	11	2

This table shows the number of retroviral and lentiviral gene therapy trials that have been approved/initiated for each category of therapy target between 1989-2016. Data from The Journal of Gene Medicine Clinical Trial site (http://www.abedia.com/wiley/index.html)

Unlike their simple retroviral counterparts lentiviruses, like the human immunodeficiency virus (HIV-1), are capable of integrating into non-dividing cells

(Lewis *et al*. 1991). This is achieved by a number of adaptations specific to lentiviruses. These include nuclear import signals within the matrix protein (Bukrinsky *et al*. 1993) a viral integrase (Bouyac-Bertoia *et al*. 2001) that has an increased association with the viral capsid that locates to nuclear pores and a central polypurine tract (cPPT) sequence that promotes nuclear entry (Zennou *et al*. 2000) of the pre-integration complex via an unknown mechanism. Together these adaptations greatly enhance the ability of lentivirus vectors to integrate into non-dividing cells. While lentiviruses also cause insertional oncogenesis they do so at a much lower rate than simple retroviruses (Modlich *et al*. 2009) mainly due to the difference in preferred integration site of the 2 types of virus (Hematti *et al*. 2004) and the design of the LTRs used in each vector (Kafri *et al*. 1997).

1.5. Retrovirus and lentivirus biology

Simple retroviruses and lentiviruses share very similar life cycles within a cell, differing only in their mechanism of nuclear entry. With simple retroviruses the disassembly of the nuclear envelope is required to access the host genome whereas lentiviruses can be actively transported through nuclear pores. The life cycle of each virus begins with the binding of the viral envelope to a specific cell surface receptor and fusion of the viral particle to the cell membrane. The specific receptor and route of entry is slightly different for each specific envelope glycoprotein but in general the binding of the envelope protein to the receptor causes the target cell to take up and release the viral core into the cytoplasm. In the case of VSV-G, the most common glycoprotein used to pseudotype viral vectors, the receptor for cellular binding is unknown but VSV-G appears to bind to many mammalian and even insect cell types suggesting that its target is ubiquitously present. The binding of VSV-G to its target induces clathrin-mediated endocytosis of the viral particle into an endosome(Aiken 1997; Sun *et al*. 2005). Once the viral endosome has entered the cytoplasm VSV-G interacts with phosphatidlyserine in the phospholipid membrane of the endosome causing fusion and release of the viral core into the cytoplasm(Coil & Miller 2004).

Once the viral core has been released into the cytoplasm, the reverse transcriptase enzyme begins the process of converting the viral RNA genome into double stranded DNA ready for integration into the host genome. DNA polymerisation beings at the primer binding site (PBS) of the viral genome, using a cellular tRNA carried over from the producer cell as the primer. As the template is reverse transcribed from the PBS to 5’ LTR of the viral genome, the RNA template is removed by the RNAse H activity of RT. Once the RT enzyme has reached the 5’ LTR the newly synthesised DNA strand can act as a primer for the 3’ end of the viral genome. This is possible due to the

homology found between the 5' and 3' LTR's, specifically in the terminal repeat sequences of the R region. The rest of the viral genome is then reverse transcribed into DNA and the RNA template is degraded by the RNAse H activity of RT apart from a short purine rich sequence called the polypurine tract (PPT).

Once RT has completed synthesis of the first strand, the PPT sequences left behind by the RT enzyme act as primers for the synthesis of the complementary second strand from PPT to the 3' end of the second DNA strand. Once RT has reached the end of the DNA template the newly synthesised second strand is transferred to the 3' end of the first strand due to homology between the 2 PBS sites. Synthesis then continues along both strands from the 5' to 3' ends to produce a complete double stranded DNA viral genome.

Once reverse transcription is complete most of the viral core disassembles leaving only the double stranded DNA genome associated with the pre-integration complex. The pre-integration complex (PIC) consists of viral protein R (Vpr), matrix antigen (MA), viral integrase and several associated host proteins. The PIC is actively transported to the nuclear envelope via interactions with host cell proteins. The nuclear localisation sequence within MA was considered to be the main mechanism of nuclear import but Popov *et al*. 1998 have shown that while MA has some effect on nuclear import Vpr is the main molecule required for import. Once the PIC has been successfully imported into the nucleus the viral DNA genome can be integrated into the host genome. The enzyme integrase binds to the LTR regions of the viral DNA and digests away 2 nucleotides from the 3' end of each strand. The PIC will then attach to the target site in the host genome based on interactions with cellular proteins associated with the target regions of the host genome. The integrase causes a double strand break at the site of integration and uses the energy released from breaking the phosphodiester bonds to

join the 3’ ends of the viral DNA to the host (Engelman *et al*. 1991). Host cellular components then fill in and ligate the single stranded gaps between the virus and host.

Once in the genome, full viral genomic RNA is stably expressed and spliced to produce the coding mRNAs required for viral protein production (Palù *et al*. 2000). Viral proteins bind to the cell membrane and begin to bud as unspliced viral RNA genomes containing the Psi packaging signal (ψ) are packaged into the viral particle (Katz *et al*. 1986).

1.6. Side effects of gene therapy - Genotoxicity

While vectors that can stably insert into the host genome are very promising for long term expression of therapeutic genes, the potential for genotoxicity has limited the deployment of these vectors in the clinic (Smith *et al*. 1996). All integrating vectors have the potential to insert into the host genome in a manner that can disrupt normal gene function however insertion can take place without causing genotoxicity. The sequence of the inserted DNA, the insertion site preference of the vector and the regulatory sequences inserted all effect the frequency and type of genotoxicity. Because of this several models of genotoxicity and insertional mutagenesis have been developed to elucidate the different mechanisms of genotoxicity and to study the safety profile of different methods.

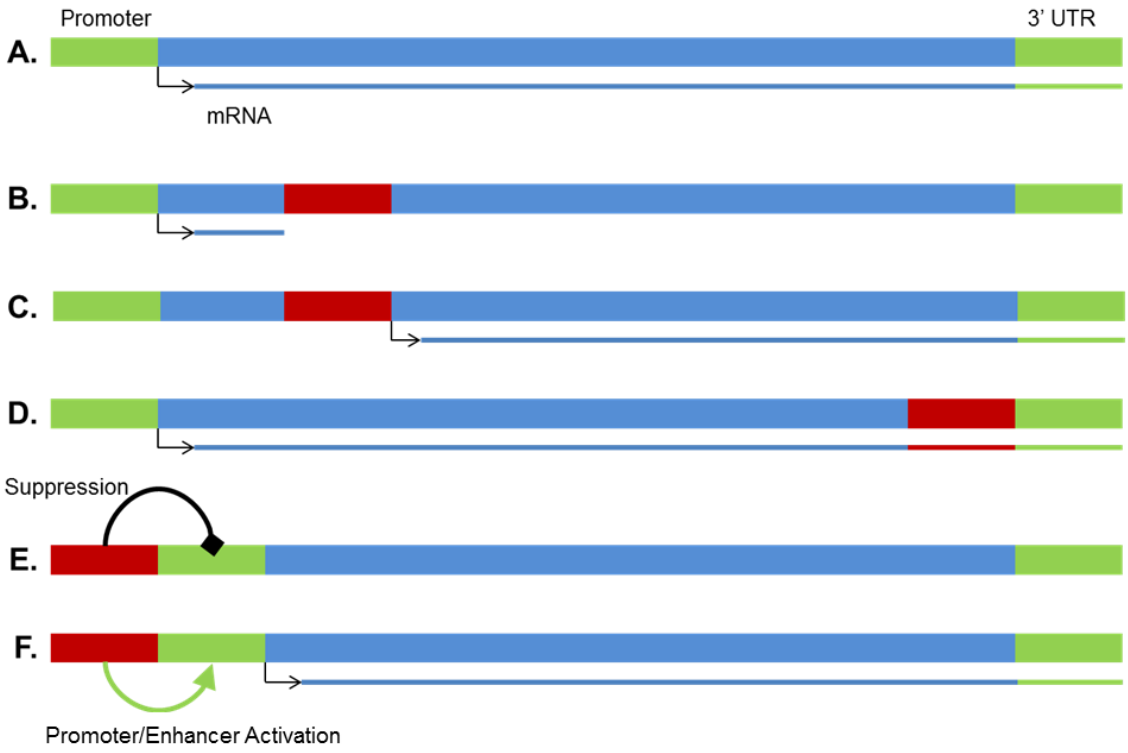

Figure 8. Mechanisms of genotoxicity from viral insertion

A. The wild-type arrangement of promoter, sequence and untranslated region (UTR) of a hypothetical gene and the expression of its mRNA. **B.** Shows how the insertion of a vector into the 5' region of a gene can cause early termination of transcription and inactivation of the gene. **C**. Shows the expression of a

truncated gene being expressed from a viral promoter sequence. This truncated protein could have abnormal levels of activity if the truncation causes the loss of a functional regularity domain. **D**. Shows the integration of a viral sequence in the 3' end of the gene near the UTR. As the UTR determines the stability of an mRNA in the cell, this can lead to increased or decreased levels of mRNA transcripts and protein product. **E**. Shows how integration upstream of the promoter can lead to changes in regulatory sequences and suppression of a gene. **F**. Shows how insertion of the viral genome upstream of the promoter can lead to Promoter/Enhancer activation and increasing transcription levels of this gene.

One of the early models of insertional mutagenesis was developed from a model for carcinogen testing (Themis *et al*. 2003). In this model the frequency of inactivating mutants of a specific locus can be measured *in vitro*. The *hprt* locus carries an X-linked gene that induces sensitivity to the selection agent 6-thioguanine. Male V79 cells are used because they have only one X gene to inactivate. HAT medium is used to purge *hprt* negative cells from the starting population followed by 6-thioguanine selection for cells that become mutated by the carcinogen. This model was adapted to test RV by making the target V79 cells permissive to infection. While this is a good model for finding the rate of inactivating mutations it does not show other types of mutation involving genes that become inactivated or activated by retroviral integration. This study found that there was no increase in 6-thioguanine resistance over background levels in cultures with only 1-2 proviral insertions per cell, whereas there was a 2.3 fold increase of resistance in cultures with multiple RV insertions (~10). This suggests that retroviral based gene therapy methods should limit the dose of viral vector so that only insertions only occur at 1-2 copies per genome to reduce the risk of inactivating mutations.

A separate model was designed by Bokhoven *et al*. to look for gain of function and activating mutations caused by insertional mutagenesis. In this cell based assay, viral vectors are applied to IL-3 dependent cells before being grown in media lacking IL-3. If the insertion of the viral vector leads to IL-3 independence the cells would continue to

grow in its absence allowing the identification of insertional mutagenesis (Bokhoven *et al*. 2009). This study found that infection by lentiviral or retroviral vectors produced IL-3 independent mutations at similar rates but found that the mechanisms behind the mutations were different for the two vectors. The retroviral vector inserted downstream of the IL-3 gene and enhanced the production of IL-3 mRNA and protein. Lentiviral vectors were found to insert upstream of the growth hormone receptor (GHR) gene and the viral LTR region induced ghr production leading to IL-3 independence. This study also showed that SIN vectors have a much lower rate of activating mutations.

While these two models are good at showing whether a vector is capable of causing mutations, they both select for mutations with little or no clinical relevance. A better model would show any mutation that could lead to clinical complications. Immortalisation of cells is a key mechanism in the development of cancers which has led to the development of the *in vitro* immortalisation assay to find the rate that specific vectors convert primary cells into immortalised cells (Modlich *et al*. 2009). This model has the advantage of being able to produce mutagenesis frequencies between different vectors allowing comparisons to be made between their safety profiles. This model has shown that self-inactivating (SIN) vectors are still capable of causing enhancer mediated gene upregulation leading to immortalisation and that this effect is much higher in SIN gamma-retroviral vectors than SIN lentiviral vectors.

To test for the potential of vectors to induce oncogenesis *in vivo* animal models have been developed. A tumour-prone mouse model has been developed with deletions to the Cdkn2 and Ifnar1 genes that make these mice predisposed to liver cancer. This makes this mouse highly sensitive to the effects of insertional mutagenesis and allows the identification of cancer associated genes after vector insertion (Ranzani *et al*. 2013). This model found that lentiviral vectors can insert into the *Braf* gene leading to

the loss of a regulatory domain resulting in the production of a constitutive active protein. As these mice are already prone to liver oncogenesis from *Cdkn2* and *Ifnar1* deletions this model suffers from a bias towards oncogenic pathways involving these two genes which may obscure other insertion sites that can lead to oncogenesis. To this end studies have been performed in mice models from a unperturbed genetic background that are not predisposed to oncogenesis. In one such model several different lentivirus vectors were administered to the mouse liver *in utero*. This study found that HIV-derived vectors did not produce any tumours and these vectors have a low rate of genotoxicity. Conversely, vectors developed from non-primate sources like EIAV had a different insertion profile with insertions taking place around several cancer associated genes (Nowrouzi *et al*. 2013). While potentially less sensitive, testing vectors in mice without a predisposition to oncogenesis will produce results that more closely resemble the effect that would be seen in the clinic. Together these models have advanced our understanding of the mechanisms of insertional mutagenesis and increase the safety of integrating vectors.

1.7. Development of lentiviruses as gene transfer agents

The first generation of vectors that were developed had three components (Naldini *et al*. 1996):

- A transfer vector containing the HIV *cis*-acting elements LTRs, (ψ) packaging signal and the insert gene under the control of a CMV promoter
- A packaging vector containing all the HIV *trans*-acting elements consisting of all the viral proteins except Env
- And an envelope vector containing the envelope protein VSV-G.

These vectors are provided separately so that *trans* viral elements enable production of the viral particle within the 293T cells but these viral particles would contain none of the *trans* elements within their genomes. This would make the viral particles incapable of further infection after integration into the target cells without several rounds of rearrangement and recombination to reproduce infective particles.

Lentiviruses like HIV have long been known to be able to incorporate envelope proteins of multiple viruses when they are co-infected into cells. Studies on co-infection of cells with HIV and xenotropic murine leukaemia virus (Canivet *et al*. 1990), amphotropic MLV (Chesebro *et al*. 1990) or herpes simplex virus (Zhu *et al*. 1990) give rise to virions with expanded host ranges. It was soon found that envelope proteins could be supplied from a separate expression plasmid to pseudotype a viral vector (Page *et al*. 1990). These early pseudotyped vectors were an improvement over previous vectors, due to the ability to choose the host range using envelopes from other viruses, however they suffered from an inability to concentrate the vectors without loss of the envelope. It was found that the use of VSV-G as the pseudotyping envelope was able to greatly increase the resistance of the viral particles to ultracentrifugation (Burns *et al*.

1993; Bartz *et al*. 1996). This allows VSV-G pseudotyped viruses to be concentrated to titers of $>10^9$ colony forming units/ml which facilitates there use in gene therapy research. The target for VSV-G envelope also appears to be ubiquitous in cells (Schlegel *et al*. 1983) allowing there use across a broad host-cell range.

In the second generation of vectors the viral genes vif, vpr, vpu and nef are deleted from the packaging vector so that even in the rare event of recombinant infective viral particles being produced they would lack these four virulence factors further increasing the biosafety of this system. This also reduces the homology with wild type HIV thereby reducing the chance of recombination within an infected host. This vector system was shown to package viral particles just as efficiently as the previous system and was capable of transfecting all most all cell lines with the same efficiency as first generation vectors (Zufferey *et al*. 1997).

The third generation of vectors added two new components to the previously described vector systems. Lentivectors normally require the expression of the viral Tat gene for efficient full length transcription from the 5' LTR of the transfer vector. Deletion of the Tat gene greatly diminished the ability of this vector to transfect cells due to the low levels of viral RNA genomes available to be package into the viral particles. To compensate for this a constitutive promoter was added into the 5' LTR to make a Tat independent transfer vector. The packaging vector was also modified into two vectors to further reduce the possibility of recombination. The viral protein Rev is required to export viral RNA containing the Rev Response Element (RRE) from the nucleus to the cytoplasm before it can be spliced. The first packaging vector contains Gag-Pol-RRE under a constitutive promoter with the Rev gene being delivered by a second vector and also driven by a constitutive promoter (Dull *et al*. 1998).

As well as the development of three generations of vectors several other modifications have been made to these vector systems to reduce the risks of recombination and to increase efficiency of vector production. Gag-Pol-RRE in the original packaging vectors contain AU rich codons that make the RNA transcripts unstable and so require viral Rev protein for efficient translation (Schwartz *et al*. 1992). To make the Gag-Pol construct Rev independent the codon sequence of Gag-Pol was optimised for human gene expression and the RRE sequence was deleted (Kotsopoulou *et al*. 2000). This greatly reduced the homology between the packaging vector and the transfer vector (which contains part of Gag and RRE for efficient packaging of the genome).

Another development on the system was self-inactivating vectors (SIN). During reverse transcription the 3' LTR is used as a template to generate the 5' LTR in the proviral DNA so that modifications of the 3' LTR are copied to the 5' LTR during infection. By deleting the U3 sequence in the 3' LTR, full length RNA can be produced and packaged in the producer cell line but infected cells would be unable to produce any full length RNA from the 5' LTR. This reduces the chances of replication-competent recombinants from emerging as well as reducing the possibility of aberrant gene expression around the insertion site due to promoter or enhancer sequences in the LTR's (Zufferey *et al*. 1998)

The routine method for lentivector production for research is by the transient transfection of the 3-4 plasmid vectors into human kidney 293T cells in low serum media. After 60-72 hours the media can be harvested and the cell debris removed by low speed centrifugation and filtration through a 0.45-micron filter. This media then contains low concentrations of the viral vector with some contaminating proteins. While this extract is often sufficient for *in vitro* studies in order to be used clinically the stock of viral vector should be clean of any contaminants and be of a much higher

concentration. One of the most commonly used methods to concentrate and purify viral stocks is by density gradient ultracentrifugation. In this method, the crude stock is passed through a sucrose gradient in a high speed centrifuge (50,000-150,000 g) and the virus is recovered from a band using ~35% sucrose concentrations and after pelleting is then resuspended in a physiological buffer. This method is fairly effective for the purification of VSV-G pseudotyped virus with high concentration factors (100-300 fold) and high recovery (94-100%) (Burns *et al*. 1993) due to the resistance to the sheer stresses that VSV-G gives to the viral particle. When it comes to other viral envelops however this method becomes less effective due to loss of the envelope proteins or damage to the viral membrane caused by the concentration.

While the production of vectors from transient transfection has been successful in gene therapy research there are problems with scaling this system up for the high standards that are required for clinical trials. Variations in plasmid quality or transfection efficiency hinder the reproducibility of viral stocks, they require large quantities of reagents and technical expertise to repeatedly transfect cell lines for production, which leads to high production costs. There is a higher risk of recombination events with the large quantities of plasmid DNA, which could lead to replication competent virus and the use of transformed cell lines requires the removal of any putative transforming agents from the clinical stocks. To alleviate this problem packaging cell lines have been produced by stably integrating the 3-4 vectors into the genome by transduction (Sanber *et al*. 2015). Each vector is inserted separately to reduce the risk of recombination events. Previously described SIN vectors cannot be transduced into a packaging cell line in this manner as they lose the ability to produce full length RNA genomes due to the copying of the 3' deletion to the 5' end during transduction. To make SIN vectors that could function in packaging cell lines the Tet-responsive element was inserted into the 3' U3. After transduction this conditional SIN (cSIN) would be able to produce full length RNA

genomes only in cells expressing the Tet-regulated transactivator (tTA) but not in target cells which lack tTA (Xu *et al*. 2001). VSV-G (Li *et al*. 1993; Hoffmann *et al*. 2010) and lentiviral proteases are cytotoxic (Konvalinka *et al*. 1995; Nie *et al*. 2002) if constitutively expressed and so most packaging cell lines change the promoters of the envelope and packaging vectors to a synthetic tetracycline inducible promoter so that cells can be grown up before production is induced.

1.8. Yeast production systems

Yeasts are a popular choice of production cells for both medical and industrial interests. Yeast, and in particular *Saccharomyces cerevisiae*, has been used in baking and brewing for thousands of years and with the advent of molecular cloning in the 1970s has become an important producer of recombinant proteins. The first effective vaccine against hepatitis B was produced from recombinant hepatitis B surface antigen expressed from *S. cerevisiae* (McAleer *et al*. 1984) and even today the majority of recombinant therapeutics are produced from this species of yeast (Huang *et al*. 2010). The popularity of *S. cerevisiae* over other organisms for the production of recombinant proteins comes from the specific advantages that this species have over other production systems (Summarised in Table 5)

Table 5. Advantages and disadvantages of different host systems for protein production data from (Demain & Vaishnav 2009)

	e. coli	S. cerevisiae	Mammalian Cells
Advantages	High yields and cost effective	High yields and cost effective	Produces high quality proteins
	High density growth	Generally regarded as safe strain	Produces humanized proteins
	Ease of culture and modification	Ease of culture and modification	Correct protein folding
	Whole genome sequenced	Whole genome sequenced	
		Stable expression and secretion	
		Can perform post-translational modifications	
		Correct protein folding	
Disadvantages	Lacks post-translational modifications	Secretory pathway varies from human	Slow growth and low yield
	Lacks correct protein folding		Expensive cultivation
	Intracellular production		Potential viral contamination
	Pyrogens and endotoxins		

One of biggest advantages of yeast production of recombinant proteins over bacterial production is post-translational modifications (Tokmakov *et al*. 2012). Mammalian derived recombinant proteins produced in yeast will undergo disulfide bond formation, subunit assembly and proteolytic processing of signal peptides in the same way as they would in their mammalian host. They will also undergo phosphorylation and glycosylation, which can be required for the correct functioning of the protein. However the glycosylation patterns of proteins are culture condition, species-, tissue- and cell-type-specific (Parekh & Patel 1992) meaning that in some cases the glycosylation systems in yeast don't produce therapeutically useful proteins. Human-type glycosylation has been achieved in *S. cerevisiae* (Amano *et al*. 2008) and *Pichia pastoris* (Jacobs *et al*. 2009) by replacing yeast glycosylation genes with human homologs. Yeast also has the advantage over bacterial production as yeast doesn't produce Pyrogens or endotoxins that which has led to *S. cerevisiae* being given GRAS (Generally regarded as safe) status by the FDA (US Food and Drug Administration).

S. cerevisiae was the first eukaryote to have its genome fully sequenced (Goffeau *et al*. 1996) which has led to the development of a large array genomic tools with which to study gene expression (Table 6). This gives *S. cerevisiae* one of the largest experimental datasets of any organism (Petranovic *et al*. 2010) with the Saccharomyces Genome Database (SGD) (Cherry *et al*. 1998) and Yeast Proteome Database (YPD) (Hodges *et al*. 1999) collecting data on the function and interaction of yeast genes and proteins. Together these tools and experimental data have been used to design a wide variety of expression vectors and protocols for recombinant protein production.

Table 6. The genomic tools that have been developed for high-through put data collection of yeast genomics

Genomic tools	References
Transcriptome analysis	(Lashkari *et al*. 1997)
Proteome analysis	(Usaite *et al*. 2008; Zhu *et al*. 2001)
Metabolome analysis	(Jewett *et al*. 2006; Villas-Bôas *et al*. 2005)
Flux analysis	(Sauer 2006)
Interactome analysis	(Harbison *et al*. 2004; Lee *et al*. 2002; Uetz *et al*. 2000)
Locasome analysis	(Huh *et al*. 2003)

Vector Design

The design of yeast expression vectors greatly influences their function and efficiency for the production of specific recombinant proteins. Yeast expression vectors consist of a yeast promoter sequence, gene of interest, transcriptional termination sequence and a selectable marker. Integrating yeast vectors also contain DNA homologous to a target locus in the yeast genome whereas episomal vectors contain an origin of replication. As high production of recombinant protein is generally the goal of yeast expression systems, the 2 micron origin of replication is usually used in yeast episomal plasmid as it produces a high copy number of with the yeast cells thereby increasing the expression level. However this strategy can lead to the saturation of the secretory pathway, which is required for expressed proteins to be released from the yeast, leading to a low yield. In situations where the secretory pathway can become saturated low copy number plasmids have been used to produce higher yields (Robinson *et al*. 1996; Parekh *et al*. 1995). To ensure plasmid stability episomal plasmids need to be maintained in selective media. Auxotrophic markers like Leu2, Trp1, Ura3, and His3 (Corrales-Garcia *et al*. 2011) allow yeast strains lacking these genes to grow in media lacking specific amino acids. Auxotrophic markers are the most common used in the production of recombinant proteins in yeast as they allow a simple method of selection without the use of antibiotics. While defined auxotrophic selection media is cheaper

than antibiotic media and doesn't run the risk of producing antibiotic resistant organisms, it is still an expensive media to maintain for large scale production. Integrative plasmids can also use these markers to both insert DNA into the host genome (Via homology between the host and vector auxotrophic sequence) and to provide selection. Integrative strategies also have the advantage of not needing to maintain selection pressure after production of a pure transformed strain as the rate of vector loss for integrated vectors is very low (Janowicz *et al*. 1991).

The choice of promoter for gene expression is determined by the properties of the expressed gene. Over the years, several inducible and constitutive promoters have been developed by isolating yeast promoters from the yeast genome (Table 7). Gene products that are toxic to cells can be expressed from inducible promoters so that the expression of toxic product is only transient. *GAL* promoters have become the most widely applied inducible promoters as they provide strong expression while being easily regulated by the addition or removal of a simple sugar (galactose). However, as galactose acts as a carbon source for yeast these promoters will suffer from reduced expression in culture as galactose levels drop (Hovland *et al*. 1989). These promoters are also limited by the levels of the Gal4p transcription factor within the cell which limits the expression from *GAL* promoters (Johnston & Hopper 1982). For non-toxic gene products constitutive promoters can be used to drive high level expression with the choice of specific promoter being based solely on the level of expression they can drive. This has led to *TPI1* becoming one of the most used constitutive yeast promoters due to its proven ability to produce very high levels of gene expression (Egel-Mitani *et al*. 2000).

In addition to promoters, the selection of terminators can drastically increase the transcription level of genes up to 35-fold over no terminator (Curran *et al*. 2013).

Table 7. List of promoters developed for yeast expression systems

Promoter	Regulation	Reference
TPI1	Constitutive	(Egel-Mitani *et al*. 2000)
ADH1	Constitutive	(Hitzeman *et al*. 1981)
GAPDH	Constitutive	(Rosenberg *et al*. 1990)
PGK1	Constitutive	(Dobson *et al*. 1982)
ENO	Constitutive	(Holland *et al*. 1981)
PYK1	Constitutive	(Burke *et al*. 1983)
GAL1-10	Inducible (galactose)	(Johnston & Davis 1984)
ADH2	Inducible (ethanol)	(Price *et al*. 1990)
CUP1	Inducible (copper)	(Karin *et al*. 1984)
PHO5	Inducible (phosphate)	(Meyhack *et al*. 1982)

The final modification that can be considered for vector design is codon optimisation. It has been known for a while that while each amino acid can be coded for by 1-4 different codons the distribution of codon usage is not random and that these distributions affect protein expression levels (Sharp *et al*. 1986). This has led to the process of optimising codon sequences of the insert gene to match the high expression codons of the production organism to increase recombinant protein production (Chang *et al*. 2006).

Culture conditions

Once the vector design has been optimised for protein production culture conditions can be optimized for further improvements. Temperature has several effects on cellular metabolism, cell wall composition, protein folding and secretion (Gasser *et al*. 2008). The best temperature for yeast cell growth is $30^{O}C$ however the best temperature for

the expression of a specific protein is dependent on the properties of the specific protein. Lower culture temperatures have been shown reduce accumulation of misfolded protein (Tøttrup & Carlsen 1990), reduce proteolysis (Jahic *et al*. 2003) and increase secretion of un-glycosylated proteins (Ferro-Novick *et al*. 1984). However higher temperatures have been shown to increase the production of other proteins (Huang *et al*. 2008; Smith *et al*. 2005) suggesting that temperatures above and below the standard growth temperature to find the best expression levels. The pH of media affects the secretion levels, protein stability and enzyme activity and also has to be adjusted for specific systems. Membrane bound proteins show optimal production at neutral to alkaline conditions (Sarramegna *et al*. 2002; Bonander *et al*. 2005) while most other types of proteins are best expressed at a lower pH (Canonaco *et al*. 2002)

Production of Lentiviral proteins

Yeast systems have a long history in the production of therapeutic proteins with many years of research and development for this purpose. While yeast has become the main method for producing singular proteins, much less research has gone into more complex protein products such as the production of virus particles. As was described in section 1.7 lentiviral vectors are produced from three main components in mammalian cells. The envelope vector VSV-G has not been expressed in yeast previously but, being a simple single protein, a yeast based system could be amenable to the production of this protein. The next component is the packaging vector expressing the polyprotein Gag-Pol. For lentiviral vector production in yeast to be successful the Gag-Pol polyprotein would need to be produced, assemble into virus like particles (VLP), bud and undergo maturation. The Gag region has been successfully expressed in *Saccharomyces cerevisiae* and has been shown to self-assemble into VLPs (Sakuragi *et al*. 2002). These VLPs were also capable of budding from the yeast when the cell wall was removed but could not undergo maturation without the Pol genes. The final

component of the lentiviral vector system is the transfer vector, which needs to express RNA capable of being packaged into VLPs, reverse transcribed and integrated into a host genome. One study (Tomo *et al*. 2013) has shown that the full HIV-1 genome can be packaged into VLPs produced from a separate plasmid expressing only the Gag genes and that this packaging was selective. This study also showed that the full Gag-Pol RNA could be expressed but did not test whether this produced full length Gag-Pol polyprotein or if particles from this construct could mature into infectious particles. Without the Pol genes these VLPs would be unable to mature into infectious particles and would lack reverse transcriptase and integrase necessary for genomic integration. Taken together, these studies suggest it may be possible to produce lentiviral vectors from yeast if full length Gag-Pol polyprotein can be expressed and the VLPs they produce can mature.

1.9. Hypothesis and Aims

Hypothesis 1:

Integrating lentiviral vectors have huge potential in the field of gene therapy for their ability to modify gene expression within a cell however there are draw backs to large scale production of VSV-G pseudotyped particles for the clinic because VSV-G pseudotyped vectors causes syncytia formation and cytotoxicity. The first hypothesis of this thesis is that it should be possible to use the yeast saccharomyces cerevisiae to generate functional replication defective HIV particles pseudotyped with VSV-G. Because particles would be held under the cell wall, no syncytia formation would cause cell fusion and death.

Aim 1:

- To determine whether yeast can be used to produce high titre lentiviral vectors pseudotyped with VSV-G and show these lentiviral vectors can infect target cells efficiently

Hypothesis 2:

The low transfection efficiency that is characteristic of non-viral gene therapy vectors can be increased using VSV-G. To help improve the efficacy of non-viral transfections, VSV-G will be generated from yeast to be used as a DNA transfection agent. This would avoid repeat transfection of mammalian cells to generate VSV-G.

Aims 2:

- To demonstrate that VSV-G can be continuously produced from yeast
- To show that yeast produced VSV-G can be an effective transfection reagent

2. Materials & Methods

2.1. Materials

2.1.1. General Reagents

1 kb DNA Ladder	New England Biolabs
Acetic acid 1M	Fisher Scientific
Agar	Fisher Scientific
Agarose	Fisher Scientific
Ampicillin	Fisher Scientific
BSA (bovine serum albumin)	New England Biolabs
Calcium Chloride	Fisher Scientific
Canavanine	Sigma
DMSO (dimethyl sulfoxide)	Fisher Scientific
dNTPs (deoxyribonucleotide triphosphates)	New England Biolabs
dsRNA Ladder	New England Biolabs
Glycerol	Fisher Scientific
IMS (industrial methylated spirit)	Fisher Scientific
Lipofectamine® LTX	Invitrogen
Manganese(II) Chloride	Fisher Scientific
Mercaptoethanol	Fisher Scientific
$MgCl_2$ Solution	New England Biolabs
MOPS	Fisher Scientific
Polyethyleneimine	Sigma
Potassium Acetate	Fisher Scientific
RNaseZap	Applied Biosystems
Rubidium Chloride	Fisher Scientific
Sodium hydroxide 1M	Fisher Scientific

SYBR® Safe		Invitrogen
Tris-Base		Fisher Scientific
UltraPure™ Alcohol	Phenol:Chloroform:Isoamyl	Invitrogen

2.1.2. Enzymes + Buffers

2.1.2.1. Restriction Enzymes

All Restriction Enzymes and their respective buffers were sourced from New England Biolabs.

2.1.2.2. Other Enzymes and buffers

All the following buffers and enzymes were sourced from New England Biolabs:

Taq Polymerase

Taq 5X Master Mix

Vent Polymerase

Q5® High-Fidelity 2X Master Mix

Taq 10X Buffer

T4 DNA Ligase

T4 DNA Ligase 10X buffer

Calf Intestinal phosphatase (CIP)

Zymolyase and buffer were sourced from Zymo Research

2.1.3. Plasmids

2.1.3.1. pRS integrating yeast vector plasmids

All the following plasmids where originally developed by Sikorski & Hieter 1989. All three vectors contain; an ampicillin marker and bacterial origin for propagation in

bacteria and a GAL1 promoter and CYC1 terminator for galactose inducible protein expression in yeast.

pRS303: contains a HIS3 marker for integration and selection in yeast. This plasmid was used as the backbone for inducible VSV-G expression in yeast.

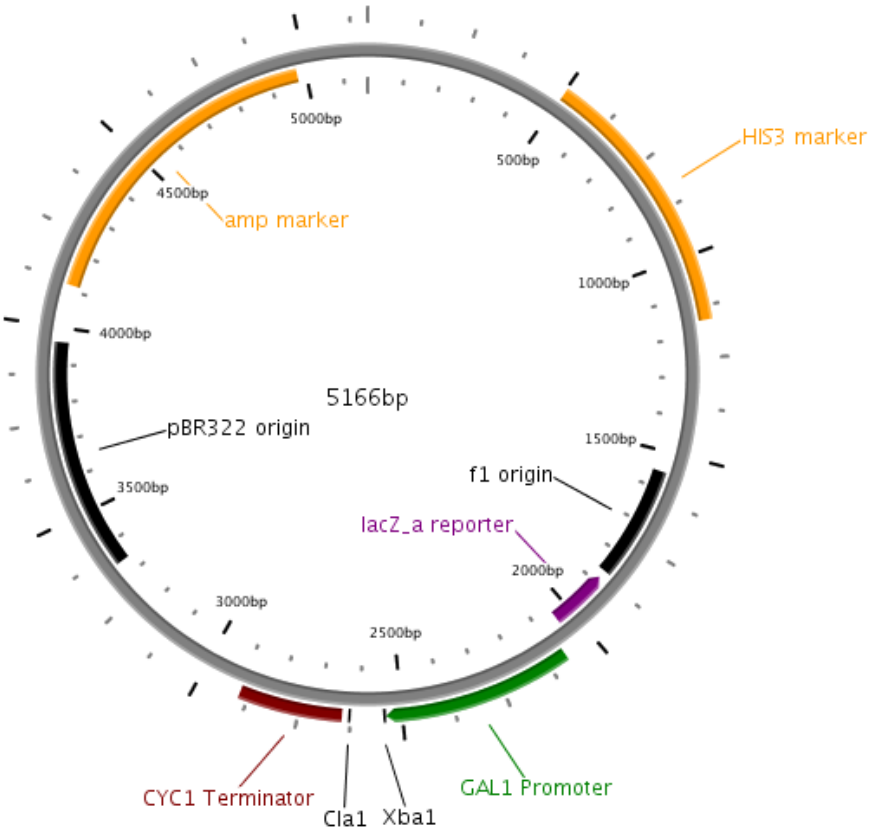

Figure 9. Detailed plasmid map of the pRS303 vector

pRS305 contains a LEU2 marker for integration and selection in yeast. This plasmid was used as the backbone for inducible Gag-Pol expression in yeast.

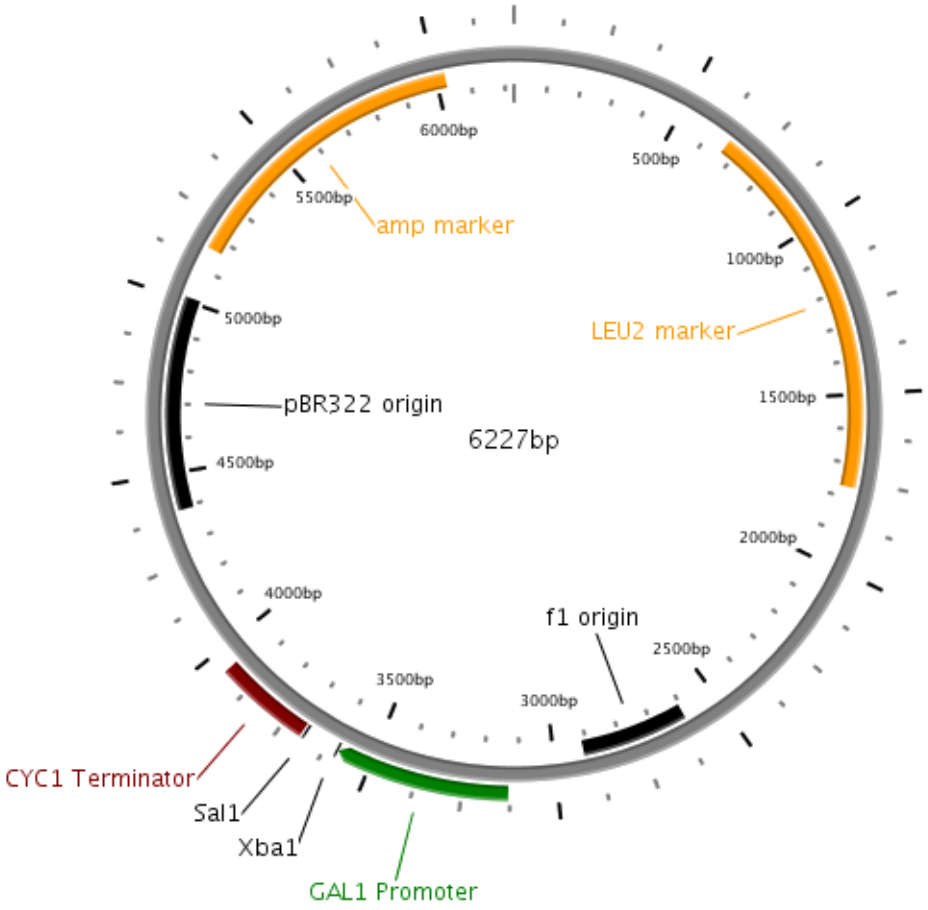

Figure 10. Detailed plasmid map of the pRS305 vector

pRS306 contains a URA3 marker for integration and selection in yeast. This plasmid was used for the inducible expression of the lentiviral backbone with GFP transgene

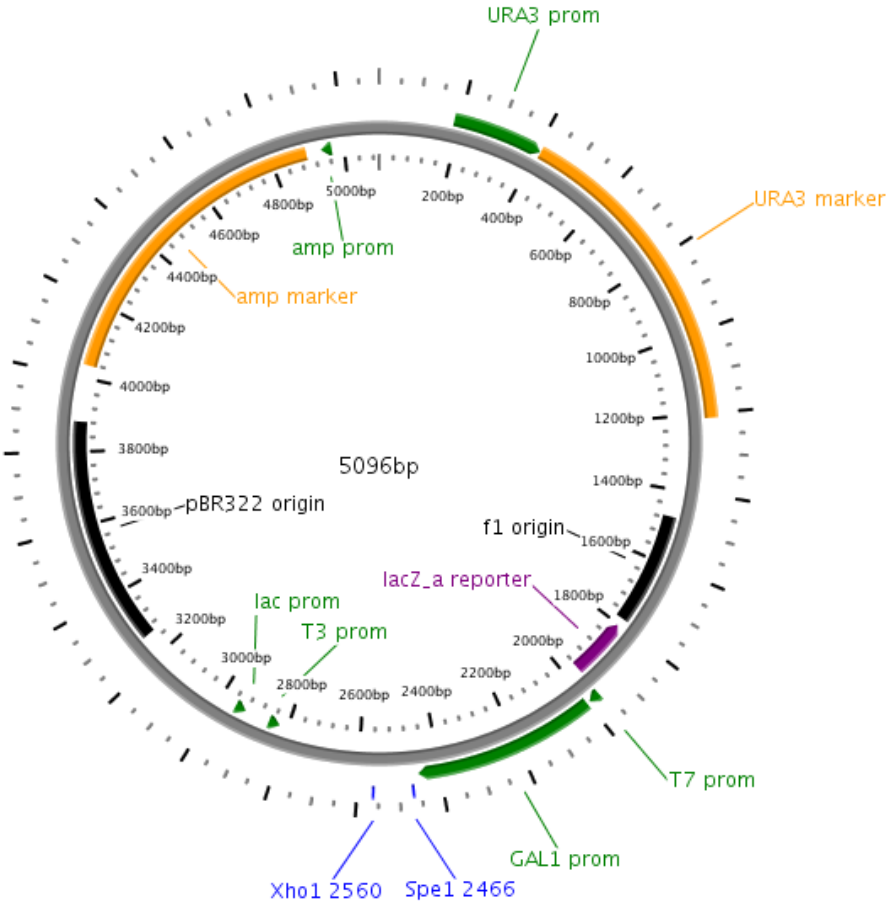

Figure 11. Detailed plasmid map of the pRS306 vector

2.1.3.2. Plasmids used as the source of lentiviral expression components

pCAG-kGP4.1R is a mammalian expression vector used in the production of lentiviral vectors and originally developed by Kato et al. 2007. It was used as the source for the Gag-Pol sequence.

Figure 12. Detailed plasmid map of the pCAG-kGP4.1R vector

pHR'SIN.cPPT-SEW is a mammalian expression vector used in the production of lentiviral vectors and originally developed by Demaison et al. 2002. It was used as the source of the lentiviral backbone sequence containing the GFP transgene.

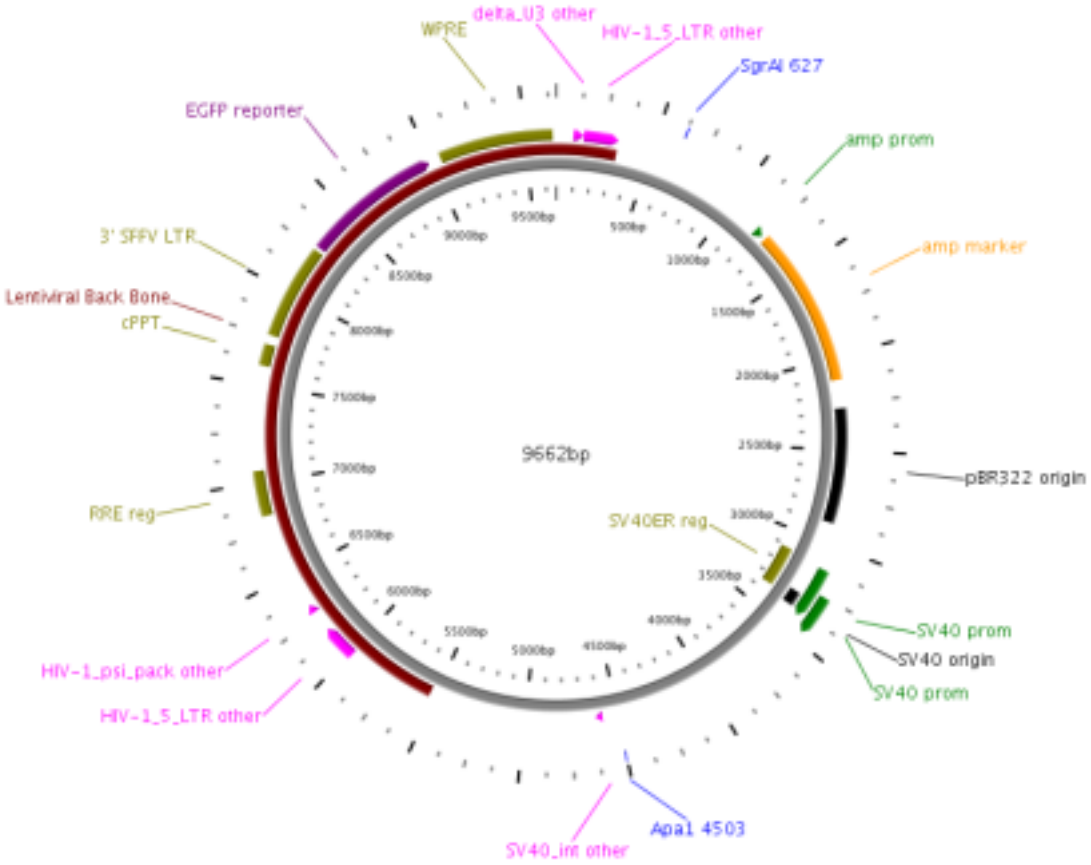

Figure 13. Detailed plasmid map of the pHR'SIN.cPPT-SEW vector

pMD2.G-VSVG is a mammalian expression vector used in the production of VSV-G. pMD2.G-VSVG was a gift from Didier Trono (Addgene plasmid # 12259)

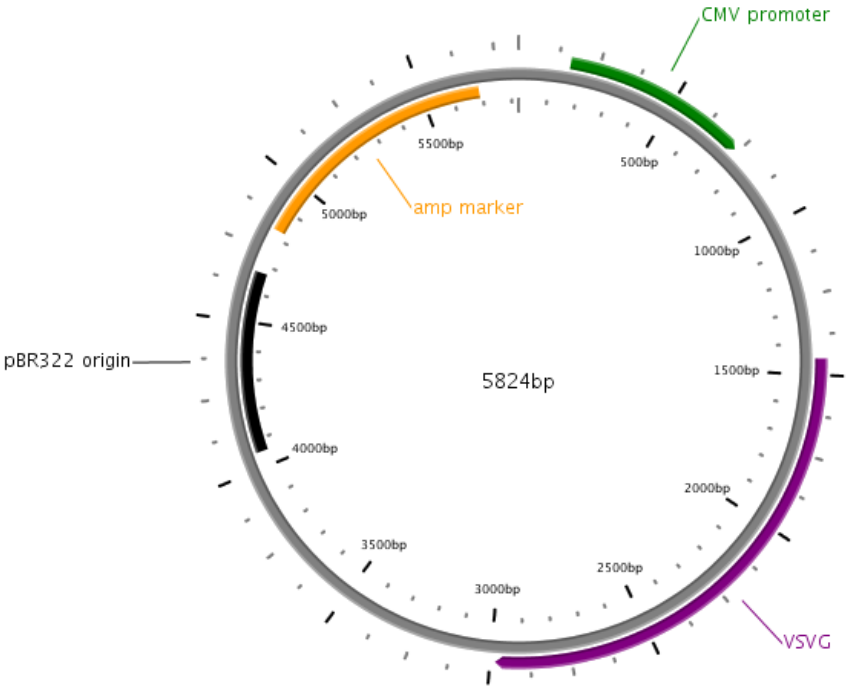

Figure 14. Detailed plasmid map of the pMD2.G-VSVG vector

2.1.4. Primers and oligonucleotides

Primers were designed with the aid of Primer3 software package so that primer pairs had similar T_m values (Untergasser *et al*. 2012; Koressaar & Remm 2007)

2.1.4.1. PCR Primers

Table 8. Primer sequences used for PCR amplification in this study

Name	Sequence
Gag-Pol Xba1	5'-CTTCTCTAGAAATTCGAGGGGTTACCATGG-3'
Gag-Pol Sal1	5'-CTTCGTCGACCCTCCAGGTCTGAAGATCTCTATCT-3'
CAN1-F	5'-CTTCTACTCCGTCTGCTTTC-3'
CAN1-R	5'-CAGAGTTCTTCAGACTTC-3'

The primers Gag-Pol Xba1 and Gag-Pol Sal1 where used to amplify the Gag-Pol fragment and add the Xba1 and Sal1 Restriction sites. The Primers CAN1-F and CAN1-R were used to amplify the CAN1 gene from the yeast genome.

2.1.4.2. Sequencing Primers

Table 9. Primer sequences used in sequencing reactions in this study

Name	Sequence
Gag-Pol 1-1001	5'-ATAACCACTTTAACTAATACTTTCA-3'
Gag-Pol 901-1901	5'-TGACACATAATCCACCTATCCC-3'
Gag-Pol 1801-2801	5'-ACTCATAGAAATCTGCGGACA-3'
Gag-Pol 2701-3701	5'-ATTGAATTGGGCAAGTCAGA-3'
Gag-Pol 3601-4509	5'-TGGAATAGATAAGGCCCAAGA-3'
CAN1-SeqF1	5'-TCGAAAACTTTGTCACCACC-3'
CAN1-SeqF2	5'-CCAATAACGGAATCCAACTG-3'
CAN1-SeqF3	5'-GCCATCCTCCATAGAGAACG-3'
CAN1-SeqR1	5'-GGCATAGCAATGACAAATTC-3'
CAN1-SeqR2	5'-TTGGACGTACAAAGTTCCAC-3'
CAN1-SeqR3	5'-CAAGTTGGCTCCTAAATTCC-3'

2.1.4.3. Linkers

Table 10. Sequences of DNA linkers used to add specific restriction sites to a PCR product for cloning.

Name	**Sequence**
LNT Linker 1-1	5'-CTAGTCGGTATCCAGGCCC-3'
LNT Linker 1-2	5'-TGGATACCGA-3'
LNT Linker 2-1	5'-CGTGAGGCGTGGCC-3'
LNT Linker 2-2	5'-ACGCCTCACGAGCT-3'

These linkers were used to make Spe1 to Apa1 and Sgra1 to Xho1 restriction site linkers for ligation of the LNT fragment to the pRS306 vector.

2.1.5. Bacterial, yeast and mammalian

2.1.5.1. Bacterial Strains

DH5α (F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk$^-$, mk$^+$) *pho*A *sup*E44 λ$^-$*thi*$^-$1 *gyr*A96 *rel*A1)

This bacterial cell line is designed for high efficiency transformation and high yield propagation of bacterial plasmids

Dam-/Dcm- (*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10)* Tet^S *endA1 rspL136* (Str^R) *dam13::Tn9* (Cam^R) *xylA-5 mtl-1 thi-1 mcrB1 hsdR2*)

This bacterial cell line has had the Dam and Dcm methyltransferase genes deleted to allow the production of plasmids without methylation so that they can be digested by methylation sensitive endonucleases

Stbl3 (F$^-$*mcr*B *mrr hsd*S20(r_B^-, m_B^-) *rec*A13 *sup*E44 *ara*-14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20(Str^R) *xyl*-5 λ$^-$*leu mtl*-1)

This bacterial cell line has been designed for efficient propagation of plasmid vectors containing long terminal repeats by deleting genes involved in homologous recombination

2.1.5.2. Yeast Strains

BY4742 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0)

A yeast strain containing deletions in the histidine, leucine, methionine and uracil synthesis pathways to allow auxotrophic selection

2.1.5.3. Mammalian cell lines

All mammalian cell lines where originally obtained from the American Type Culture Collection (ATCC; Manassas, VA, USA)

HEK293T	A highly transfectable derivative of the 293 cell line (a permanent line of primary human embryonal kidney cells transformed by sheared human adenovirus type 5 DNA) into which the gene for SV40 T-antigen has been inserted.
BT474	A human breast cancer cell line derived from a ductal carcinoma

2.1.6. Media

Bacterial media

Lennox Luria-Bertani (LB) broth	Fisher Scientific

Yeast Media

YPD Broth	Fisher Scientific

Yeast Extract	Fisher Scientific
Peptone	Fisher Scientific
Yeast Nitrogen Base without Amino Acids	Sigma
D-(+)-Glucose powder	Sigma
D-(+)-Galactose powder	Sigma
D-Sorbitol powder	Sigma
Raffinose	Sigma
All Drop-Out Supplements	CloneTech

YP	900ml	Selective Media	1 Litre	YP-Sorbitol	900ml
Yeast extract	10g	Yeast Nitrogen Base	6.7g	Yeast extract	10g
Peptone	20g	Glucose	5g	Peptone	20g
		Amino Acid Drop out mix	0.8g	Sorbitol	182.17g

YPD	1 Litre	GSS Media	1 Litre
Yeast extract	10g	Yeast Nitrogen Base	6.7g
Peptone	20g	Galactose	5g
Glucose	20g	Amino Acid Drop out mix	0.8g
		Sorbitol	182.17g

YP, YPD and YP-Sorbitol media was made as above and autoclaved. To make YPG, YPR and YPG-Sorbitol 100ml of 20% W/V Galactose or Raffinose stock was added to the respective media to a final concentration of 2%. GSS (Galactose Sorbitol Selective) Media was made as above and filter sterilised. For solid media 15g agar was added before autoclaving.

2.1.7. Commercial Kits

Nucleospin ® Plasmid Extraction Kit	Macherey-Nagel
Nucleospin® Gel and PCR Clean-up kit	Macherey-Nagel
Nucleospin® RNA Isolation Kit	Macherey-Nagel
Nucleospin® RNA Virus Kit	Macherey-Nagel
Frozen EZ Yeast Transformation II Kit	Zymo Research
Purelink™ HiPure Plasmid Maxiprep Kit	Invitrogen
ProtoScript® II First Strand cDNA Synthesis Kit	New England Biolabs
TA Cloning® Kit (with pCR™2.1 Vector)	Invitrogen
Lenti-X qRT-PCR Titration Kit	CloneTech
PureLink™ HiPure Plasmid (Filter) Maxiprep Kit	Invitrogen

2.1.8. Consumables

0.22µm Vacuum Filter 250mL 45mm neck	Fisher Scientific
0.45µm Vacuum Filter, 50ml, Steriflip	Fisher Scientific
Microcentrifuge tube (0.5 & 1.5 ml)	Fisher Scientific
Centrifuge tube (50 ml)	Fisher Scientific
Glass pipettes	Fisher Scientific
Lens cleaning tissues	Fisher Scientific
Pipette tips	Fisher Scientific
Petri Dishes	Sigma
Glass beads	Sigma
Latex and Nitrile Gloves	Fisher Scientific
Glass measuring Cylinder (10ml & 50mL)	Fisher Scientific
Erlenmeyer Flask narrow neck 250mL	Fisher Scientific
PCR tubes	Fisher Scientific
Autoclave tape	Fisher Scientific

2.2 Methods

2.2.1. Polymerase Chain Reaction (PCR)

2.2.1.1. Taq/Vent Polymerases

The PCR conditions in each experiment were dependent on the properties of the templates and primers used but the general protocol was as follows:

Denature at 95^{o}C for 1 minute, annealing at primer specific temperature for 30 seconds, elongation at 72^{o}C 1 minute per kb of expected product, repeat for 30-40 cycles then keep at 4^{o}C. The final volume of the reactions were always 100µl for amplification of cloning fragments and 10µl for colony PCRs. Reactions were made to a final concentration of 0.2mM dNTP, 1-5mM $MgCl_2$ and 0.5µM of each primer with 2 units per 100µl of Taq and 0.4 units per 100µl of Vent Polymerase. When cloning fragments for ligation into the TOPO® vector Vent was excluded to preserve the A overhangs of the PCR product.

For colony PCR of yeast, colonies were suspended in 10µl 0.02M NaOH heated to 99^{o}C for 10 minutes to release the genomic DNA before being put on ice. 0.5µl of this solution was then added to PCR mix. For bacterial colony PCR the colonies were resuspended directly in the PCR mix.

2.2.1.2. Q5® High-Fidelity 2X Master Mix

The reaction conditions for Q5® are similar to that for Taq/Vent with these changes; Denature at 98^{O}C for 30 seconds, anneal for 15 seconds, elongation for 30 seconds per kb of expected product. The reaction mix was made of 0.5 volume Q5® 2X Master Mix and 0.5µM of each primer up to final volumes of 50µl for cloning or 10µl for colony PCR

For colony PCR of yeast, colonies were suspended in 10µl 0.02M NaOH heated to 99^{o}C for 10 minutes to release the genomic DNA before being put on ice. 0.5µl of this solution was then added to PCR mix. For bacterial colony PCR the colonies were resuspended directly in the PCR mix.

2.2.1.3. First Strand cDNA Synthesis and RT-PCR

For cDNA Synthesis the ProtoScript® II First Strand cDNA Synthesis Kit was used according to the manufactures instructions. In brief,~1µg of total RNA in a volume 1-6µl was added to 2µl of random hexamer primer mix and made up to a volume of 8µl with distilled water. This mixture was then denatured at 65^{O}C for 5 minutes before putting on ice. To the RNA/primer mix 10ul of ProtoScript II Reaction mix (2X) was added and 2ul of ProtoScript II Enzyme Mix (10X) before incubation at 42^{o}C for one hour. After incubation the enzyme was inactivated at 80^{o}C for 5 minutes and the reaction diluted to 50µl with water. For negative controls ProtoScript II Enzyme Mix was replaced with distilled water.

PCR amplification of the cDNA was carried out as described in section 2.2.1.2. Q5® High-Fidelity 2X Master Mix

2.2.2. Enzymatic modification of DNA

2.2.2.1. Restriction Enzyme Digest

PCR or vector DNA was mixed with water and 0.1 volumes of the appropriate 10X reaction buffer for the enzyme. 1 unit per ul of enzyme was then added and the reaction mixture was incubated at the recommended temperature for an hour for vectors or overnight for PCR products. A reaction volume of 30µl was used for vectors and 100µl for PCR products.

2.2.2.2 DNA Ligation

For DNA ligation digested and purified Vector DNA was mixed with cut insert at a 2:1-5:1 molar ratios depending on the relative size of fragment to vector. For cloning with the TA-TOPO® vector fresh PCR was used instead of cut insert. 80 units of T4 DNA ligase was added per µl of 1x T4 Ligation Buffer and the reaction was left at room temperature for an hour or at 4^{o}C overnight.

2.2.2.3 Removal of 5' terminal phosphate

To reduce self-ligation of cut plasmid during ligation reactions the DNA was dephosphorylated with calf intestinal phosphatase (CIP). After restriction enzyme digestion 1 unit of CIP was added per µl of restriction digest reaction mix and incubated for 1 hour at 37^{O}C. The DNA was then purified before ligation.

2.2.3. DNA Purification from PCR or Enzyme Digestion

DNA Purification of PCR reactions or Enzyme Digestions was performed with NucleoSpin® Gel and PCR Clean-up kit from Macherey-Nagel according to the manufactures instructions. One volume of sample was mixed with 2 volumes of binding buffer before being added to the spin column and allowed to bind to the silica membrane. The columns were then washed with washing buffer, dried before the DNA was eluted in 10-50µl elution buffer heated to 70^{o}C.

2.2.4. Agarose Gel

One percent Agarose gel was made by dissolving 0.5g Agarose in 50ml TAE buffer before adding 3µl ethidium bromide or 3µl SYBR® Safe. Gels were run at 100 volts for an hour in a *PerfectBlue Gel System Mini S* gel electrophoresis apparatus (PeqLab Ltd). DNA samples were mixed with 6X bromophenol blue loading buffer and run with

1kb DNA ladder. DNA fragments were visualised on a *ChemiDoc™ XRS+* and images taken with the ImageLab™ software package.

2.2.5. DNA Purification from Agarose Gel

Agarose gels were illuminated under UV light and the fragments of interest were extracted with a scalpel. The fragment was then purified with the NucleoSpin® Gel and PCR Clean-up kit from Macherey-Nagel according to the manufactures instructions. The exercised gel fragment was dissolved in 2 volumes of binding buffer at 50^{o}C for 5-10 minutes. After the gel had completely dissolved the DNA was allowed to bind to the spin columns before being washed with washing buffer, dried and the DNA eluted in 10-50µl elution buffer heated to 70^{o}C.

2.2.6. NanoDrop quantification of DNA and RNA

DNA and RNA samples were quantified with a NanoDrop 2000c UV-Vis Spectrophotometer from Thermo Scientific using elution buffer as a blank. 1µl of sample was loaded onto the NanoDrop before recording the ng/µl concentration reading and A_{260}/A_{280} and A_{260}/A_{230} ratio readings. A_{260}/A_{280} readings of ~1.8 where considered pure for DNA and readings of ~2.0 where considered pure for RNA. A_{260}/A_{230} readings of 2.0-2.2 where considered pure for both DNA and RNA.

2.2.7. Preparation of chemical competent E.coli

RF1	250ml	RF2	250ml
RbCl	3g	RbCl	0.3g
$MnCl_2$	2.47g	MOPS	0.52g
KAc	0.74g	CaCl2	2.75g
$CaCl_2$	0.38g	Glycerol (50%)	74ml
Glycerol (50%)	74ml		
1M acetic acid	488µl		

The selected E.coli strain was inoculated in 10ml LB media and grown overnight at 37^{o}C in a shaking incubator. Then 100ml of LB media was inoculated to an OD600 of 0.05 and allowed to grow to an OD of 0.2-0.3 in the shaking incubator. Once the culture had reached the desired OD it was incubated on ice for 15 minutes. The cells were then pelleted by centrifugation for 5 minutes at 5100RPM at 4^{o}C. The supernatant was discarded and the cells were resuspended in 20ml RF1 before being incubated on ice for 15 minutes. The cells were then repelleted for 5 minutes at 5100RPM at 4^{o}C and resuspended in 4ml RF2. The cells were incubated on ice for another 15 minutes before making 100µl aliquots and freezing on dry ice. The competent E.coli was then stored at -80^{o}C until needed.

2.2.8. Bacterial Transformation by Heat Shock

Chemically Competent E.coli was defrosted on ice before the addition of 5ul of ligation mixture or 1ul of purified plasmid. The competent cells were then incubated on ice for 45 minutes before being heat shocked at 42^{o}C for 1 minute. The cells were allowed to recover for one minute on ice before being plated on LB agar plates supplemented with appropriate antibiotic for transformant selection and incubated overnight at 37^{o}C.

2.2.9. Isolation of high-copy number plasmid DNA from E. coli

2.2.9.1. Isolation via Nuclueospin® columns

DNA Purification of high-copy number plasmid DNA was performed with the Nuclueospin® Plasmid Extraction Kit from Macherey-Nagel according to the manufactures instructions. *E.coli* was grown overnight in 1-5ml of LB media before cells were pelleted in a microcentrifuge at 11000x g. The supernatant was discarded and the pellet resuspended in 250µl resuspension buffer. The cells were then lysed in 250µl lysis buffer at room temperature for 5 minutes before adding 300µl naturalization buffer. The lysate was then clarified by spinning the solution at 11000x g for 5-10 minutes and the resulting supernatant was added to the spin column and allowed to bind. The column was then washed with washing buffer, dried and the DNA was eluted in 50µl elution buffer heated to 70^{o}C.

2.2.9.2. Isolation via Phenol:Chloroform:Isoamyl Alcohol (25:24:1)

E.coli was grown overnight in 1-5ml of LB media before cells were pelleted in a microcentrifuge at 11000x g. The supernatant was discarded and the pellet was resuspended in 400µl of water. 500µl of the UltraPure™ Phenol:Chloroform:Isoamyl Alcohol mixture was added to the suspension and briefly vortexed to form an emulsion. The mixture was centrifuged for 5 minutes at 13,000RPM. After centrifugation the top aqueous phase was pipetted off without disturbing the interphase layer and then added to 500µl of isopropanol prepared in a fresh microcentrifuge tube. The solution was mixed by gentle inversion before centrifugation at 13,000RPM for 5 minutes. The supernatant was discarded and the pellet washed with 70% ethanol several times. As much ethanol was removed as possible by pipetting before the pellet was dried in a heating block at 70^{O}C. The pellet was then resuspended in an appropriate amount of distilled water.

2.2.9.3. Isolation via PureLink™ HiPure Plasmid (Filter) Maxiprep Kit

Larger scale DNA Purification of high-copy number plasmid DNA was performed using the PureLink® HiPure Plasmid Maxiprep Kit according to the manufactures instructions. *E.coli* was grown overnight in 100ml of LB media before cells were pelleted in a microcentrifuge at 4000x g for 10 minutes. At the same time 30ml Equilibration Buffer (EQ1) was added to the HiPure Maxi Column and allowed to drain by gravity flow. The cell pellet was resuspended in 10ml of Resuspension Buffer (R3) containing RNase A. 10ml Lysis buffer was added and mixed by gentle inversion before incubating at room temperature for 5 minutes. 10ml of Precipitation Buffer (N3) was added and mixed by gentle inversion. The cell lysate was added directly to the filter/column and allowed to drain by gravity flow. The filter was removed and 60mL Wash Buffer (W8) was added to the column and allowed to drain. A sterile 50ml tube was placed under the column and 15ml of Elution Buffer (E4) was added to the column and allowed to drain into the tube. 10.5ml of isopropanol was added to the eluate to precipitate the DNA. The tube was centrifuged at 12,000x g for 30 minutes at 4^{o}C before discarding the supernatant. The DNA pellet was air-dried for 10 minutes then resuspended in 200-500ul of TE Buffer (TE).

2.2.10 Chemically competent S. cerevisiae

2.2.10.1. Preparation of chemically competent S. *cerevisiae*

Chemically Competent S. *cerevisiae* was prepared using the Frozen EZ Yeast Transformation II Kit according to the manufactures instructions. In brief 10ml of YPD media was inoculated with a single colony of yeast and grown overnight in a shaking incubator at 30^{o}C. The next morning 1ml of saturated culture was added to 10ml of YPD media and grown to an OD of 0.6 to 1.0 in the shaking incubator. The culture was then centrifuged at 500xG for 5 minutes to pellet the cells. The supernatant was discarded and the cells resuspended in 10ml of solution 1. The culture was centrifuged

again at 500xG for 5 minutes, the supernatant was discarded and the cells resuspended in 1ml of solution 2. Aliquots of 50µl of resuspended cells were made into sterile microcentrifuge tubes before being slowly frozen in a -80^{o}C freezer.

2.2.10.2. Transformation of chemically competent S. *cerevisiae*

Chemically competent cells were removed from the freezer and allowed to defrost at room temperature. Approximately 1µg of DNA, in no more than 5µl of solution, was added to the cells along with 500µl of solution 3. This mixture was then vortexed vigorously before being incubated for 1-3 hours at 30^{o}C, vortexing every 15 minutes. After incubation 100-200µl of transformation reaction was spread on selective plates and incubated for 2-4 days at 30^{o}C.

2.2.11. RNA and DNA Extraction from yeast

2.2.11.1. Total RNA Extraction from yeast

RNA Extraction from yeast was performed with Nucleospin® RNA Isolation Kit from Macherey-Nagel according to the manufactures instructions. In brief 5-50ml of yeast culture was harvested and centrifuged at 5,000RPM for 10 minutes, the supernatant was discarded and the pellet resuspended in 1ml Spheroplast buffer (1M sorbitol/100mM EDTA) to which 50U Zymolyase was added. The reaction was incubated at 30^{o}C for 30 minutes and the resulting spheroplasts were pelleted at 5,000RPM for 10 minutes and the supernatant carefully discarded. 350µl of buffer RA1 and 3.5µl of ß-mercaptoethanol was added to the pellet and then vortexed vigorously to lyse the cells. The lysate was filtered through NucleoSpin® Filters at 11,000g for 1 minute. 350µl of 70% ethanol was added to the filtered lysate and mixed by vortexing before being loaded into a NucleoSpin® RNA Column and centrifuged at 11,000g for 30 seconds. 350µl of MDB was added to the column and spun at 11,000g for 1 minute to dry the membrane. 95µl of freshly made DNase reaction mixture was added to the

column and incubated at room temperature for 15 minutes. 200µl of RAW2 buffer was added and the column centrifuged for 30 seconds at 11,000g before 600µl of RA2 was added and centrifuged again. 250µl of fresh RA2 was then added and spun for 2 minutes at 11,000g to dry the membrane. Finally the RNA was eluted in 60µl of RNAse free water at 11,000g for 1 minute before storage at -20^{o}C.

2.2.11.2. Genomic DNA extraction from yeast for PCR

Single yeast colonies were picked from a plate and suspended in 100µl of 200mM lithium acetate 1% SDS solution. The suspension was vortexed then incubated for 5 minutes at 70^{O}C. 300µl of ice cold 100% ethanol was added and mixed by gentle inversion. Precipitated DNA was collected by centrifugation 11,000g for 5 minutes. Supernatant was removed and the DNA pellet was washed in 500µl of 70% ethanol. The pellet was then suspended in 100µl of TE buffer and cell debris was removed by centrifugation at 11,000g for 1 minute. 1µl of supernatant was then used for PCR.

2.2.12. Viral Titration

2.2.12.1. Viral RNA extraction

Viral RNA Extraction from supernatant was performed with the Nucleospin® RNA Virus Kit from Macherey-Nagel according to the manufactures instructions. In brief 200µl of cell suspension was centrifuged at 5,000g for 1 minute and 150µl of supernatant was transferred to a new tube. 600µl of buffer RAV1 was added to the sample before vortexing and incubation for 5 minutes at 70^{o}C. After incubation 600µl of 100% ethanol was added and the solution vortexed. The solution was loaded onto a NucleoSpin® RNA virus column and centrifuged for 1 minute at 8,000xg. 500µl of buffer RAW was added to the column and centrifuged for 1 minute at 8,000xg. 600µl of buffer RAV3 is then added and centrifuged before another 200µl was added and centrifuged for 5 minute at 11,000xg to remove any trace ethanol. The column was then placed in a

sterile 1.5ml microcentrifuge tube and 50µl RNase free water preheated to 70^{o}C was added and incubated for 2 minutes before being centrifuged for 1 minute at 11,000xg. Quality of extracted RNA was assessed by A_{260}/A_{230} and A_{260}/A_{230} ratios after NanoDrop quantification.

2.2.12.2. qRT-PCR Titration

Viral titration was performed using the Lenti-X™ qRT-PCR Titration kit from CloneTech according to the manufactures instructions. For each RNA sample 12.5µl eluted RNA, 2.5µl DNase 1 buffer (10X), 4µl DNase 1 (5 units/µl) and 6µl RNase-Free water was added to a PCR tube. The reaction was incubated at 37^{o}C for 30 minutes then 70^{o}C for 5 minutes on a thermocycler. qRT-PCR master mix was made from 8µl RNase-Free water, 12.5µl Quant-X Buffer (2X), 0.5µl Lenti-X Forward and Reverse primers (10 µM each), 0.5µl ROX Reference Dye LSR, 0.5µl Quant-X enzyme (5 units/µl) and 0.5µl RT Enzyme mix per well needed. Control RNA was 10-fold serially diluted in EASY Dilution Buffer to produce a $5x10^{7}$-$5x10^{3}$ copies/µl standard curve. Sample RNA was diluted in a similar manner to produce 1x to 0.001x dilutions. 2µl of each RNA sample dilution and control dilution was added to their respective wells in duplicated. Six wells are used for no template controls (NTC) containing 23µl master mix and 2µl Easy dilution buffer only. The PCR plate was briefly centrifuge to remove any bubbles before being loaded into the qPCR machine. The qPCR machine was programed to run 42^{o}C for 5 minutes, 95^{o}C for 10 seconds then to do 40 cycles of 95^{o}C for 5 seconds and 60^{o}C for 30 seconds before finally doing a dissociation curve.

To generate a standard curve, average Ct values from the control dilution were plotted against copy number ($5x10^{7}$-$5x10^{3}$). Average Ct values for each duplicate sample dilution were calculated and the standard curve was then used to

calculate the corresponding copy number value. The copy number of the original samples was then calculated using the formula:

$$\text{Copies/ml} = \frac{\text{(calculated copy number)(1000µl/ml)(2x DNase)(50µl elution)}}{\text{(150µl sample)(2µl added per well)}}$$

2.2.13. Electron Microscopy

2.2.13.1. Transmission Electron Microscopy

Transmission electron microscopy was performed on a JEOL2100 field emission gun transmission electron microscope at 100kV. To prepare the samples 1.5ml of the cell culture (at an OD_{600} of 1.0) was pelleted by centrifugation and the supernatant was discarded. The cells were resuspended in 1ml of phosphate buffer before being pelleted and the supernatant discarded. 100µl of 25% glucoaldhyde was added to the pellet and left for 1 hour at room temperature to fix then cells. The glucoaldhyde was aspirated off the cells and the pellet was washed 3 times in phosphate buffer to remove any residual glucoaldhyde. Approximately 200µl of 4% molten agar was added to the pellet and left to set. The agar plug with cells was removed from the tube and placed in osmium tetra oxide solution for 1 hour at room temperature. The agar plug was removed from the osmium tetra oxide solution and placed in distilled water for 15 minutes. The distilled water was replaced and the process repeated 2 more times before adding 70% ethanol for 10 minutes. The plug was then transferred to 100% ethanol for 10 minutes before the ethanol was replaced with fresh ethanol for 15 minutes. The plug was then washed twice in 100% propylene oxide for 15 minutes before being transferred to 50:50 embedding medium/propylene oxide for 30 minutes. Finally the plug was transferred to 100% embedding medium for 2 hours at room temperature before being set at 60^{o}C for 24 hours. After the resin had set, 100nm thick sections of the cells were cut using a microtome and mounted on copper grids for use in the electron microscope.

2.2.13.2. Scanning Electron Microscopy

To prepare samples 1.5ml of the cell culture (at an OD_{600} of 1.0was pelleted by centrifugation and the supernatant was discarded. The cells were washed in 1ml of phosphate before being pelleted and the supernatant discarded. 100µl of 25% glucoaldhyde was added to the pellet and left for 1 hour at room temperature to fix then cells. The glucoaldhyde was aspirated off the cells and the pellet was washed 3 times in phosphate buffer to remove and residual glucoaldhyde. 100µl of osmium tetra oxide was added to the cells and left for 1 hour before being washed 3 times in phosphate buffer. The resuspended cells were transferred to a 100nm pore filter membrane and dehydrated by passing 2ml of 50%, 4ml of 70% and then 4ml of 100% ethanol slowly through the membrane. The dehydrated membrane was placed directly into a critical point dryer to dry the cells without damaging them. Finally the membrane was attached to a mount with carbon glue and sputter coated with gold palladium alloy ready for SEM.

2.2.14. Formation of Yeast Spheroplasts

Yeast cells to be spheroplasted were collected from liquid media by centrifugation at 5,000RPM for 5 minutes and the supernatant was discarded. The cells were then resuspended in 1ml of Spheroplast buffer with 10ul of Zymolyase (5U/µl) and incubated for 30 minutes at 30^{o}C for spheroplasts to form.

2.2.15. SDS-PAGE and Western Blotting

Equal volumes of sample and 2x Laemmli sample buffer were mixed before being heated to 95^{o}C for 5 minutes to denature the proteins. A volume of sample containing ~20µg of protein was loaded into a precast 12% Tris-Glycine gel from Precise™ along with 5ul of color prestained protein standard. Gels were run in SDS-PAGE running

buffer at 100V for one hour. Proteins were transferred to Hybond-C membranes by wet transfer at 100V in ice-cold transfer buffer. The membrane was washed once in TBS before being blocked overnight at 4^{o}C in 10% BSA blocking buffer on an orbital shaker. After blocking the membrane was washed once in TBS before adding anti-VSV-G tag antibody (1:3000) in blocking buffer and incubated overnight at 4^{o}C on an orbital shaker. After overnight incubation the membrane was washed 4 times in TBS for 5 minutes for the first wash and 15 for each subsequent wash. The membrane was then incubated in 2ml of Pierce™ ECL Western Blotting Substrate for 1 minute at room temperature before being imaged on the BioRad imager.

2.2.16. Enzyme-Linked Immunosorbent Assay (ELISA)

For the standard curve synthetic VSV-G peptide was diluted to a concentration of 20µg/ml in carbonate coating buffer and 50µl was loaded into the top three wells of the ELISA plate. A 1/10 serial dilution down the wells was used to form a concentration curve from 20µg/ml to 2pg/ml. As the molecular weights of the standard, Yeast VSV-G and mammalian VSV-G differed slightly all calculations were done based on the molar concentration of VSV-G target epitope (14.9µM to 1.49pM).

Samples of VSV-G were diluted 1/10 in carbonate coating buffer and 50µl was loaded onto the plate in triplicate. As a negative control 50ul carbonate coating buffer was also loaded in triplicate. The plate was then covered and incubated overnight at 4^{o}C to allow the peptides to bind to the wells. The peptide solution was removed and the wells washed two times with 200µl of PBS. 200µl of blocking solution (10% BSA in coating buffer) was then added to each well and the plate was incubated at room temperature for 2 hours to block any remaining binding sites. The blocking buffer was removed and the wells were then washed twice with 200µl PBS. 100µl of anti-VSV-G tag antibody, diluted 1/5000 in blocking buffer, was added to each well before the plate was cover

and incubated overnight at 4^{o}C. The plate was then washed four times with PBS before 100µl of TMB substrate was added to each well. After sufficient colour development (approximately 30 minutes) 100µl of 3N sulphuric acid was added to each well to stop the reaction. The absorbance of each well at 450nm was then measured and recorded.

To calculate the concentration of the samples a standard curve was plotted using the averaged blank-corrected 450nm measurements of each standard vs. its molar concentration. A 4-Parameter Logistic Regression algorithm was used to fit the curve and the OD readings of each sample was compared to this curve to find its concentration.

2.2.17. BCA total protein Assay

BCA assays were performed using the Pierce™ BCA Protein Assay Kit according to the manufactures instructions. In brief, dilutions of BSA between 2000µg/ml and 25ug/ml were made in carbonate coating buffer and 25ul loaded in triplicate on to an ELISA plate. Each sample was diluted 1/10 in coating buffer and 25µl loaded in triplicated onto the plate and coating buffer was also loaded as a negative control. 200ul of working reagent was then added to each well before the plate was covered and incubated at 37^{o}C for 30 minutes. The plate was then cooled to room temperature and the absorbance read at 570nm.

To calculate the concentration of the samples a standard curve was plotted using the averaged blank-corrected 570nm measurements of each standard vs. its concentration. A 4-Parameter Logistic Regression algorithm was used to fit the curve and the OD readings of each sample was compared to this curve to find its concentration.

2.2.18. Mammalian Cell Culture

All mammalian cells were maintained in DMEM supplemented with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were grown in monolayers on 10cm tissue culture dishes in a 37^{o}C 5% CO_2 incubator.

To split confluent cells the cell media was aspirated off and the cells were washed once with PBS to remove any remaining media. 2ml of 0.25% trypsin was added to each plate and incubated for 2 minutes to dislodge the cells. 2ml of media was then added to neutralize the trypsin and the cells were pipetted up and down to produce a homogenous single cell suspension. Cells were split into fresh media at a ratio's of between 1:3 and 1:10 every few days depending on their growth characteristic.

To freeze cells for long term storage, cells were grown on 10cm dishes until confluent and trypsinised before being pelleted at 2,000 RPM for 5 minutes. Cells were then resuspended in 1ml freezing media (DMEM with 20% FBS and 10% DMSO) and aliquoted into labelled cryotubes. The cryotubes were placed in a -80^{o}C freezer for 24 hours to slowly freeze before being transferred to liquid nitrogen.

To recover cells from liquid nitrogen storage, a cryotube containing the cells required was removed from liquid nitrogen and quickly defrosted in a 37^{o}C water bath. The freezing media was then added to 10ml of growth media and the cells pelleted at 2,000 RPM for 5 minutes before aspirating off the supernatant to remove the DMSO. The pellet was then resuspended in 1ml of media and added to two 10cm plates containing 10ml growth media.

2.2.19. Mammalian Cell Transfection

The day before transfection cells were seeded at a density of $4.0X10^6$ cells per 10cm dish so that the cells would be >80% confluent at time of transfection. 8µg of Maxiprep plasmid DNA was diluted into 1ml of Opti-MEM and mix quickly by inverting the tube. 16µg of PEI was then added to the tube and quickly vortexed to mix. In experiments with VSV-G the protein was added with the PEI before vortexing. The solution was then incubated at room temperature for 30 minutes to allow DNA/PEI complexes to form. The DNA/PEI complex solution was then added drop wise to the growing cells and the dish was shaken gently to insure an even distribution.

2.2.20. Viability Assay

For cell viability assays 100µl of 0.4% (w/v) trypan blue solution was added to 100µl of mammalian cell suspension and mixed gently. The staining solution was allowed stand for 5 minutes before a small amount was applied to a hemacytometer counting chamber. The number of stained and unstained cells was counted and percentage viability was calculated for each sample.

2.2.21. Imaging Flow Cytometry

Cells used for imaging flow cytometry were grown for 48 hours after transfection before being collected by trypsinization and resuspended in AccuMAX™ for use in the Imagestream^X imaging flow cytometer. Using the Inspire™ data acquisition software, images of 100,000 cells were captured on channel 1 for bright field and channel 2 for eGFP expression. Excitation of eGFP was performed with a 488 nm laser at a power setting of 50mW and images were captured using a 40× objective.

The Ideas™ program was used to produce an analysis template that would be used on all our collected samples. The analysis template was generated from image data of

untransfected cells using the building blocks wizard. First single cells are identified and gated and then a region is drawn to select only those cells that are in the correct focal plane during imaging flow. Finally a region of light intensity on channel 2 was selected to represent the intensity above which cells would be considered GFP positive. This region was selected such that no cells from untransfected samples would be counted as GFP positive but any cells with a higher intensity would be counted. The transfection efficiency for each sample was then calculated as the percentage of in focus single cells that were GFP positive. The results were compared using the Student's unpaired *t*-Test and the data were regarded as significantly different if the P value was less than 0.05.

2.2.22. Fluctuation assay

In preparation for the fluctuation assay SC-Arg+Canavanine plates were prepared and dried overnight in a 30^{O}C incubator. Strains to be tested were pick from selective plates and used to inoculate appropriate selective media. The cells were cultured overnight in a shaking 30^{O}C incubator before being diluted 1:10,000 into YPD and YPG media. 500ul of diluted culture were added to each well of a 24 well plate before being sealed and incubated for 24 hours in a 30^{O}C incubator. Cells in each well were resuspended by gentle pipetting and 100ul of each well was spot plated onto dried SC-Arg+Canavanine. The spots were given enough time to dry before being placed in the incubator for 24 hours. The remaining culture from each plate was collected and diluted 1:100. 100ul was then spread onto YPD plates and placed in the incubator for 24 hours.

After 24 hours the YPD plates were removed and the number of colonies counted to find the average number of viable cells for each sample. The canavanine plates were

then removed and the number of spots without any resistant colonies were counted. The equation below was then used to find the mutation rate within each sample.

$$u = \frac{-\ln(\mathrm{P0})}{Nf'}$$

Where

u = mutation rate = probability of mutation per cell per division

$\mathrm{P0}$ = The fraction of cultures without any mutants

Nf' = number of viable cells per spot plated

Confidence limits (±95%) were then calculated using the statistical software program R and samples are statistically different ($p<0.01$) if their confidence limits did not overlap.

3. Results: Lentiviral vector production from yeast

3.1. Construction of Gal induction vectors for Lentivirus Production in Yeast

Pseudotyped lentivirus vectors are normally produced by transiently transfecting mammalian cells with 3 plasmids; The packaging plasmid containing the HIV derived Gag and Pol genes, The transfer plasmid containing the HIV genome packaging signal and the gene of interest under the control of a separate promoter, and the envelope plasmid containing the envelope that the vector is to be pseudotyped with. The quantity of viral particles produced by this method is limited by two factors; Firstly, transiently transfected cells can only produce these vectors for ~3 days before the cells die and secondly that the density of mammalian cells that can be grown in culture is always limited. In this chapter an attempt was made to address these problems by developing an inducible yeast expression system to generate lentivirus particles. By using a stably integrated inducible system, the potential to generate viral particles via induction, thereby enabling a permanent producer cell line to be created for several batches of viruses, was investigated. Hence, exploiting the fact that yeast cells can be grown to high density reaching more than $1x10^8$ cells per ml, which is a dramatic increase in the number of virus producer cells that can be grown per ml when compared to mammalian culture and provide an improved producer system for high virus titre.

For this work the pCAG-kGP4.1R vector was chosen as the source of our Gag-Pol insert due to its previous use in lentiviral production and pRS305 was chosen as the yeast expression vector due to the *leu2* selection marker/integration site and the strong inducible promoter GAL1. Due to the lack of compatible restriction sites between the

two plasmids we chose to use a PCR reaction to isolate the Gal-Pol sequence and to add the Xba1 and Sal1 we would use to ligate this fragment into pRS305 (Figure 15).

To isolate the Gag-Pol fragment from the pCAG-kGP4.1R vector, 35 cycles of PCR with a Taq/Vent polymerase mix was performed to amplify the fragment. The size of the fragment was confirmed by agarose gel electrophoresis as ~4.3kb before being ligated to a TA TOPO® vector (Figure 16.A). The ligation mixture was transformed into chemically competent Dam^-/Dcm^- *e.coli* and the resulting colonies grown in liquid culture before the plasmid was purified with the Nucleospin plasmid purification kit. Each sample of plasmid was checked by restriction digest with Xba1 and Sal1 followed by agarose gel electrophoresis (Figure 16.B). Samples containing the correct fragment size for the Gag-Pol insert (~4.3kb) were then sent for sequencing. After the sequence of the insert had been confirmed, the vector was digested again with Xba1 and Sal1 before gel purifying the Gag-Pol fragment. The integrating yeast vector pRS305 was also digested with the same enzymes and the enzymes heat inactivated. The purified Gag-Pol fragment and the digested pRS305 backbone were ligated overnight in equal molar amounts before being transformed into chemically competent *Stbl3 e.coli*. The resulting colonies were grown in liquid media and the plasmid was purified before being checked with a Cla1 digest and agarose gel to find inserts of the size of the Gag-Pol fragment (Figure 16.C).

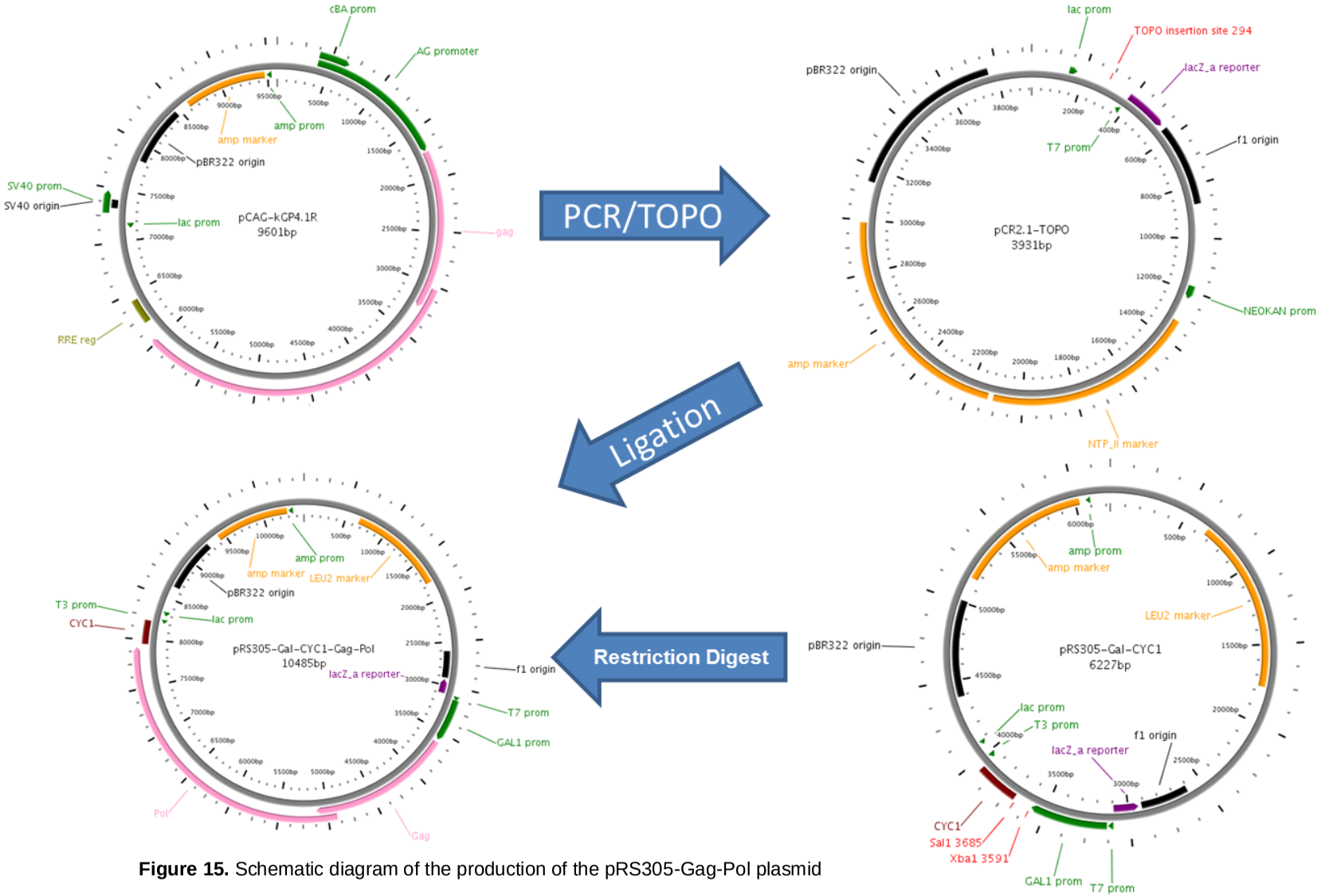

Figure 15. Schematic diagram of the production of the pRS305-Gag-Pol plasmid

The Gag-Pol coding region of pCAG-kGP4.1R was amplified by PCR using the primers Gag-Pol Xba1 and Gag-Pol Sal1 to generate a sequence that would have compatible restriction sites for cloning into pRS305. This fragment was first cloned into a TOPO sequencing vector as that the sequence of the insert could be confirmed. The sequence verified Gag-Pol insert was then extracted restriction digestion with Xba1 and Sal1 before gel purification. The insert was then ligated into p305 that had been digested with the same restriction sites.

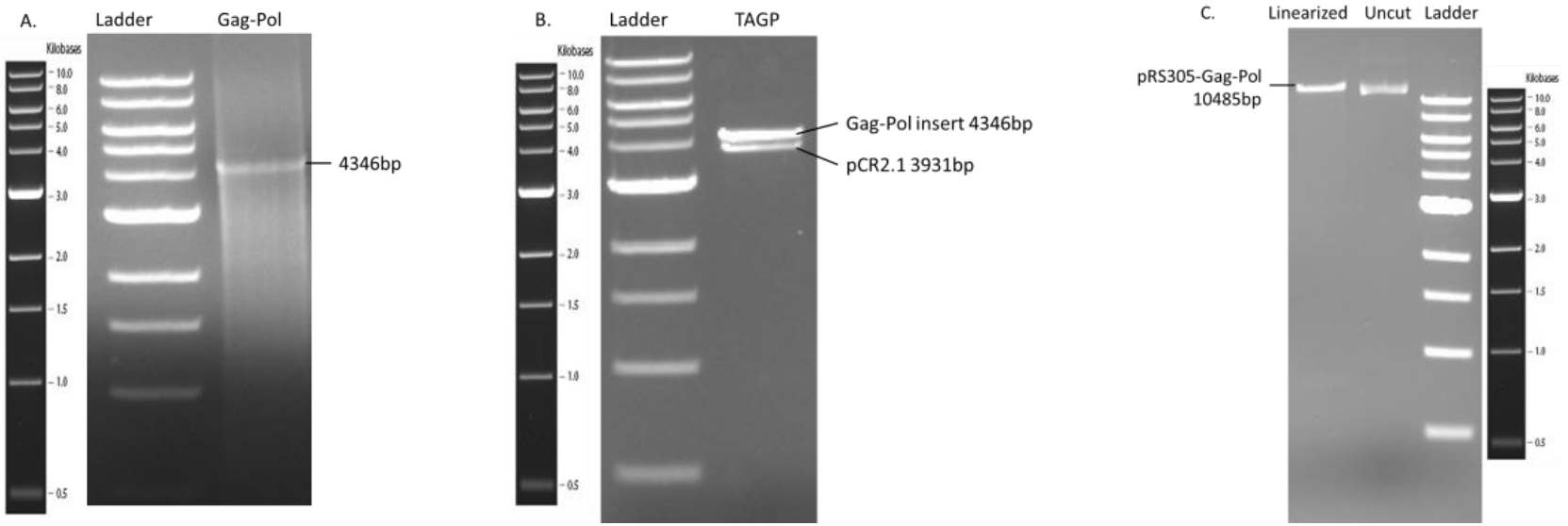

Figure 16. Agarose gels of the construction of the pRS305-Gag-Pol plasmid via PCR

A) Shows the PCR product of the a size of approximately ~4.3kb (Gag-Pol). B) The TOPO vector digested with Xba1 and Sal1 to release the Gag-Pol insert (~4.3kb) from the pCR2.1 vector backbone (~4kb) C) pRS305-Gag-Pol linearized with Cla1 (Linearized) against uncut vector (Uncut). A 1kb ladder (Ladder) was used as the marker in all three gels

The plasmid pHR'SIN.cPPT-SEW was chosen to provide the lentiviral backbone as it contained an eGFP insert gene that would allow testing of the infectivity of the viral product by flow cytometry. pRS306 was chosen as the yeast expression vector due to the *ura3* selection marker/integration site and the strong inducible promoter GAL1. As the LNT insert could be easily isolated due to its Apa1 and Sgra1 it was decided to use restriction digestion and gel purification to isolate the insert and to use DNA linkers to make the ends compatible with the Spe1 and Xho1 digestion sites in pRS306 (Figure 17).

To isolate the eGFP containing lentivector backbone from the pHR'SIN.cPPT-SEW plasmid, the plasmid was digested with Apa1 and Sgra1 to cut outside the LTR's. The digestion mixture was run on an agarose gel and the fragment of the correct size (~4.5Kb) for the insert was gel purified (Figure 18). At the same time the yeast integrating vector pRS306 was digested with Spe1 and Xho1 before being heat inactivated. To ligate the insert into the new vector two linkers were created (Spe1 to Apa1 and Sgra1 to Xho1) to make the digested ends compatible. Each linker was made by allowing two synthetic oligonucliotides to anneal at room temperature to form short double stranded DNA molecules with compatible overhangs to the cut sites. The Spe1 to Apa1 linker was designed so that the Apa1 site was lost after ligation as Apa1 is required to be a unique site within the *ura3* gene for later linearization. For the ligation cut insert, vector and linkers were mixed at a molar ratio of 1:1:5:5 in the ligation reaction. The ligation mixture was transformed into chemically competent *Stbl3 e.coli* and the resulting colonies were checked by colony PCR for the presence of the LNT insert before the positive colonies were grown in liquid media, the plasmids purified and the vector and insert checked by restriction digest with Apa1.

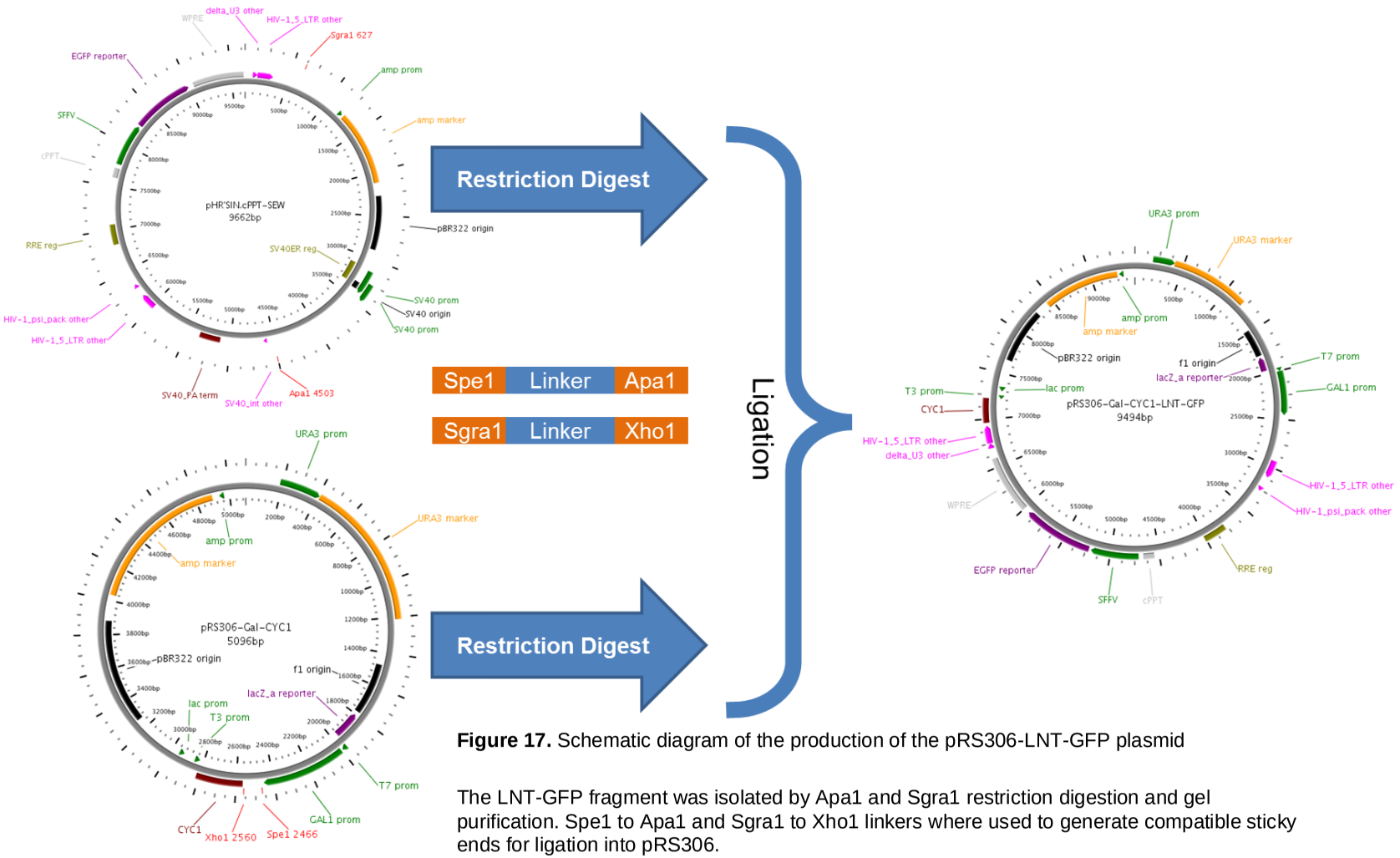

Figure 17. Schematic diagram of the production of the pRS306-LNT-GFP plasmid

The LNT-GFP fragment was isolated by Apa1 and Sgra1 restriction digestion and gel purification. Spe1 to Apa1 and Sgra1 to Xho1 linkers where used to generate compatible sticky ends for ligation into pRS306.

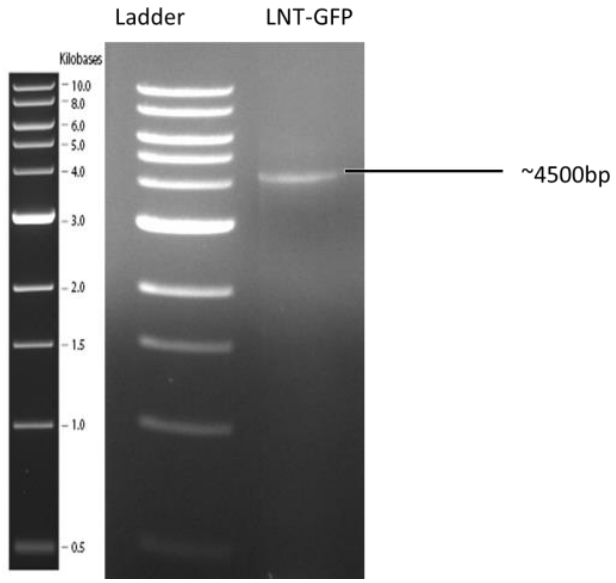

Figure 18. Gel purification of the LNT-GFP fragment.

The LNT-GFP fragment was gel purified after digestion of pHR'SIN.cPPT-SEW with Apa1 and Sgra1. The purified fragment was run on a gel (LNT-GFP) and a fragment of ~4.5kb was found. A 1kb ladder was used as the marker (Ladder).

The plasmid pMD2.G-VSVG was chosen as the source of our VSV-G insert due to its previous use in lentiviral production and pRS303 was chosen as the yeast expression vector due to the *his3* selection marker/integration site and the strong inducible promoter GAL1. As the VSV-G insert could be isolated with BsaAI and MscI restriction digestion to produce a blunt-ended product and there was a suitable site in pRS303 for a blunt-ended ligation (EcoRV) we decided to use this method to generate pRS303-VSVG plasmid (Figure 19).

The envelope gene VSV-G was cut from the pMD2.G-VSVG vector by digestion with BsaAI and MscI to produce a blunt-ended product. As the size of the VSV-G fragment (~2.3kb) was very close to that of the vector backbone (~2.6kb), gel purification would not be sufficient to remove the backbone. To aid in the purification of the VSV-G fragment the backbone was also digested with ApaI so that the two fragments of the backbone would be much smaller than the VSV-G fragment (Figure 20). The digestion mixture was run on an agarose gel and the ~2.3kb fragment was gel purified. The vector pRS303 was digested with EcoRV and the enzyme was then heat inactivated. The vector was then CIP treated to remove the terminal phosphates to prevent any self-religation of the vector before being column purified (Figure 21). Equal molar amounts of insert and vector were mixed in the ligation reaction and then transformed into chemically competent *Stbl3 e.coli*. The resulting colonies were grown in liquid culture and the plasmids purified via phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation. As blunt ended ligation can cause the insert to be ligated in either the forward or reverse orientation we needed to check the isolated plasmids for correct orientation. The plasmid was digestion with BamHI and EcoRI to check for correct orientation of insert as the EcoRI site appeared on one edge of the insert which would lead to different size fragments appearing on a gel depending on the orientation of the insert.

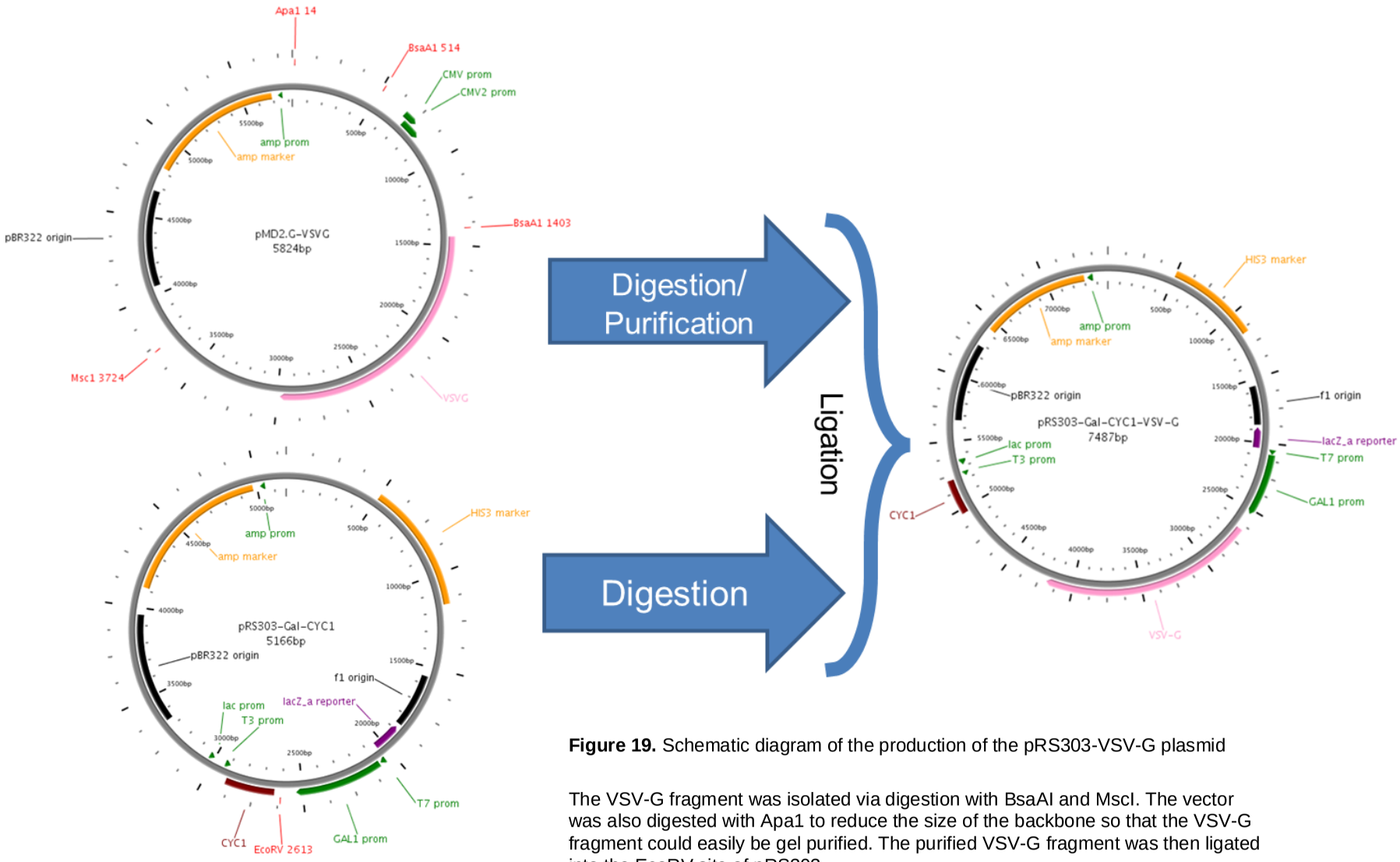

Figure 19. Schematic diagram of the production of the pRS303-VSV-G plasmid

The VSV-G fragment was isolated via digestion with BsaAI and MscI. The vector was also digested with Apa1 to reduce the size of the backbone so that the VSV-G fragment could easily be gel purified. The purified VSV-G fragment was then ligated into the EcoRV site of pRS303.

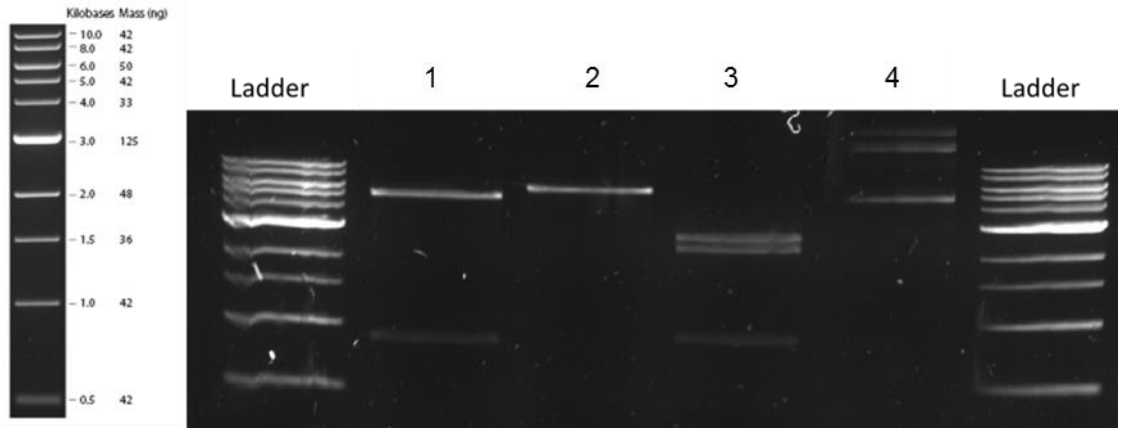

Figure 20. Removal of the VSV-G fragment from pMD2.G-VSVG vector.

Lane 1: pMD2.G-VSVG digested with BsaAI to produce 2 fragments of ~4.9kb and 0.9kb. Lane 2: pMD2.g-VSVG digested with MscI to produce a fragment of ~5.8kb. Lane 3: pMD2.g-VSVG digested with BsaAI and MscI to produce three fragments of ~2.6kb, ~2.3kb and ~0.9kb. Lane 4: uncut DNA.

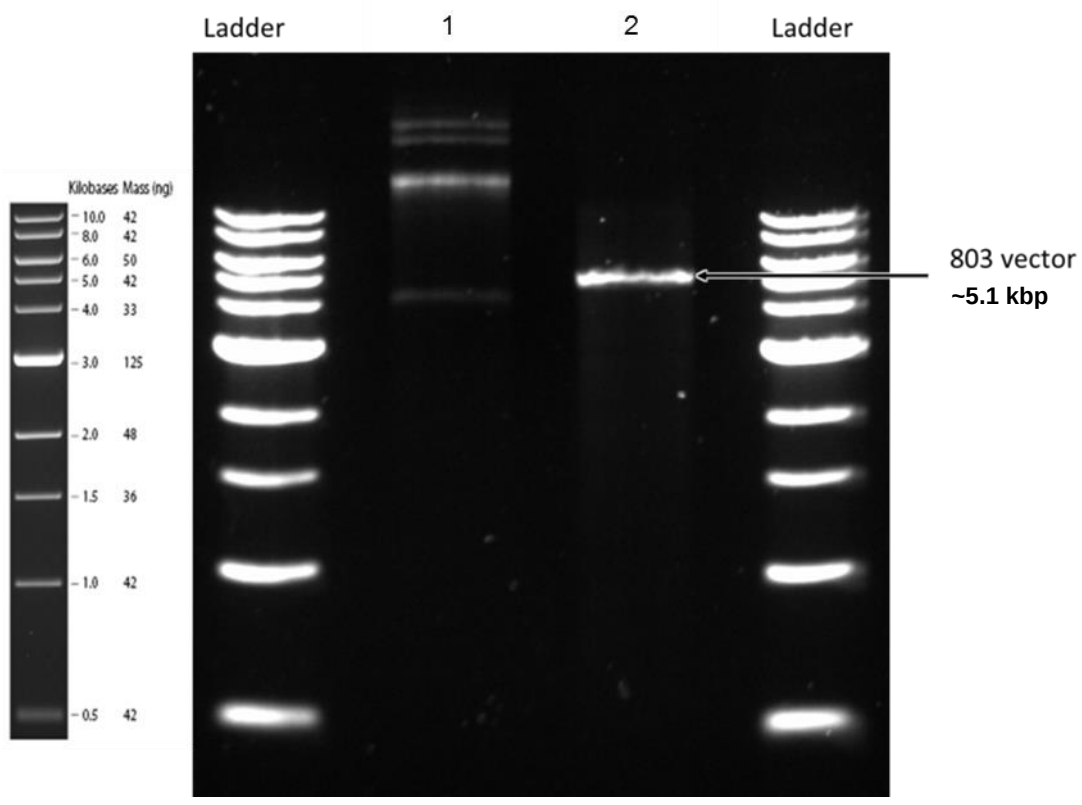

Figure 21. Agarose gel showing the pRS303 vector linearized with EcoRV.

Lane 1: uncut pRS303 vector. Lane 2: pRS303 vector after digestion with the EcoRV enzyme and column purification showing a fragment size of ~5.1kb.

3.2. Transformation of yeast with Gal-Lenti production vectors

The yeast strain BY4742 was used for yeast transformations with the lentivirus vectors. The yeast strain was made competent using the Frozen EZ Yeast Transformation II Kit. To transform yeast with the integrating vectors had to be linearized within the yeast selectable marker to allow recombination into the yeast genome. The vector pRS305-Gag-Pol was digested with the Cla1 enzyme to linearize it within the *leu2* gene before being checked on a gel for complete linearization. The digested DNA was then concentrated by isopropanol precipitation and approximately 1µg of DNA was used to transform ~1x10^6 BY4742 cells. The transformed cells were plated on Leu2 selective plates, with cells transformed without DNA as a negative control, and incubated at 30^{o}C until colonies formed. Each colony was tested by colony PCR and positive colonies were grown in liquid culture.

The transformed yeast BY4742-GP yeast was then made competent. The vector pRS306-LNT-GFP was digested with Apa1 to linearize it within the *ura3* gene before being checked on a gel for complete linearization. The digested DNA was then concentrated by isopropanol precipitation and approximately 1µg of DNA was used to transform ~1x10^6 BY4742-GP cells. The transformed cells were plated on –URA –LEU selective plates, with cells transformed without DNA as a negative control, and incubated at 30^{o}C until colonies formed. Each colony was tested by colony PCR for both LNT and Gag-Pol and colonies positive for both constructs were grown in 10ml selective culture overnight.

To ensure that this strain was capable of RNA production of Gag-Pol and the LNT backbone, the strain was induced and total RNA was examined using a RT-PCR for these genes. First an overnight culture of BY4742-LNT-GP was used to inoculate 10ml

of YPG media to induce the production of the LNT and Gag-Pol mRNA. After 6 hours of induction total RNA was harvested with the Nucleospin® RNA Isolation Kit. The extracted RNA was then used for first strand cDNA synthesis using random hexamer primer mix followed by RT-PCR to show the production of both Gag-Pol and LNT mRNA in the same strain. Dilute plasmid DNA was used as a positive control and total RNA without reverse transcriptase was used as a negative control. As can be seen in Figure 22 the BY4742-LNT-GP strain does produce Gag-Pol and LNT mRNA.

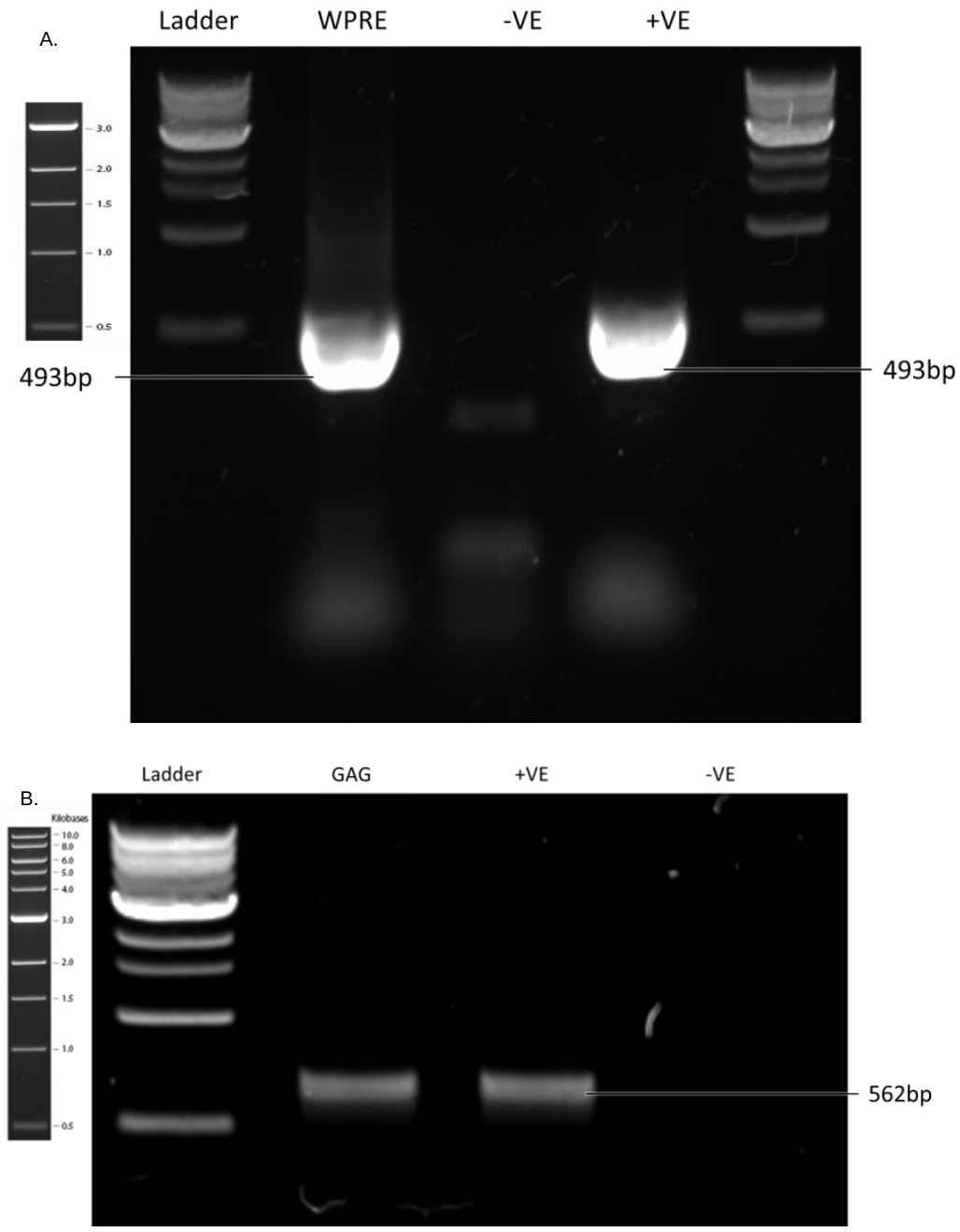

Figure 22. RT-PCR for a fragment of Gag and WPRE mRNA.

A) Shows a RT-PCR product of ~493bp from the sample RNA (WPRE) and a positive control (+VE) lanes with no product from the negative control (-VE) lane. B). Shows a RT-PCR product of ~562bp from the sample RNA (GAG) and a positive control (+VE) lanes with no product from the negative control (-VE) lane.

3.3. Transformation of Gal Lentiviral producing yeast with a VSV-G expression vector

The envelope plasmid pRS303-VSV-G was linearized in the *his3* marker with BsiWI before being checked on an agrose gel for complete linearization. The digested DNA was then concentrated by isopropanol precipitation. The BY4742-LNT-GP yeast was made competent as before and approximately 1µg of linearized pRS303-VSV-G was used for the transformations. The transformed cells were plated on –URA –LEU –HIS selective plates, with cells transformed without DNA as a negative control, and incubated at 30^{o}C until colonies formed. Colonies formed on all the plates including the negative control suggesting that the transformation was contaminated or that the strain had become His3+ before transformation. Stocks of BY4742-LNT-GP were streaked onto -URA –LEU plates and grown overnight at 30^{o}C. Single colonies were picked to inoculate an overnight culture of YPD and the cells were then transferred to YPG and fresh YPD media for 3 hours. After 3 hours of induction total RNA was extracted from the cells and RT-PCR was performed for the LNT and Gag-Pol constructs as before. The RT-PCR showed that the strain was producing LNT and Gag-Pol mRNA in the galactose induced media but not the YPD media as would be expected of the strain suggesting the stocks aren't contaminated. Stocks of BY4742, BY4742-GP, BY4742-LNT-GP were taken and streaked onto multiple selective plates to find out if the strains were His3+, the results of this can be seen in Table 11.

Table 11. Showing the growth characteristics of the strains under different selective conditions

	BY4742	BY4742-GP	BY4742-LNT-GP
-URA	-	-	+
-LEU	-	+	+
-HIS	-	+	+
-URA -HIS	-	-	+
-LEU -HIS	-	+	+
-URA -LEU	-	-	+
-URA -LEU -HIS	-	-	+

(-) represents no growth and (+) represents growth

The wild-type shows no growth on any of the selective media suggesting that it is His3-, Leu2- and Ura3- as expected for the BY4742 strain. Both BY4742-GP and BY4742-LNT-GP grew in the correct selective media for their strain but also in the absence of histidine. This suggests that there has been a His3+ reversion of the wild type strain at some point during the production of the BY4742-GP strain. Previous work in our lab (unpublished data) has suggested that viral integrase causes double strand breaks within the host genome which could have led to the reversion of the His3 gene in the BY4742-GP clone. This theory was further tested in chapter 4.

Without a His3 marker it would not be possible to integrate the envelope plasmid into the host genome. While an envelope protein would be necessary for cell entry of viral particles the production of mature particles only requires the gag-pol packaging polyprotein and the lentiviral backbone. To this end, it was decided to use the BY4742-LNT-GP strain to identify whether mature viral particles could be produced from yeast before attempting to pseudotype these particles with the VSV-G envelope glycoprotein.

3.4. Analysis of Yeast Lentiviral production

Once the production of mRNA had been confirmed, growth curves were generated to compare the growth rates of the BY4742-LNT-GP strain to the wild type BY4742 strain under different culture conditions. These growth curves would allow us to find the best conditions for our strain during lentiviral production. If the growth curve of one strain was lower than the wild type under induction it would suggest that the production of lentiviral particles where leading to reduced viability of the cells under those conditions.

Colonies of BY4742-LNT-GP and BY4742 were picked from a plate and used to inoculate 10ml of an appropriated selective media or YPR before being grown overnight. The next day 100ul of each suspension culture was diluted into 900ul of fresh media and had its absorbance at 600nm (OD_{600}) measured in a spectrophotometer. 1ml of fresh media was used as a blank for all readings. Strains grown in selective media were used to inoculate YPR or YPG media to an OD_{600} of 0.020 while strains grown in YPR were used to inoculate selective media with or without galactose to an OD_{600} of 0.025. These OD's were chosen so that these cultures would show significant changes of OD over 24 hours without reaching saturation.

The strains were grown for 24 hours in a shaking incubator at $30^{O}C$ and the OD_{600} was recorded at 3, 6, 12 and 24 hours. This was repeated 3 times for each test. When strains were grown in selective overnight before transferring to YP there was no significant difference between the growth rates for wild type and the BY4742-LNT-GP strain with or without induction over the 24 hours (Figure 23.A). This suggests that the production of lentiviral particles under these conditions do not significantly reduce viability of the cells.

When strains were grown in YPR overnight and transferred to selection and induced, the BY4742-LNT-GP strain showed significantly lower growth then the other strains (Figure 23.B), however, there was minimal difference without induction. From this data it was decided that for subsequent experiments BY4742-LNT-GP would be grown up in selective media before being induced in YPG media.

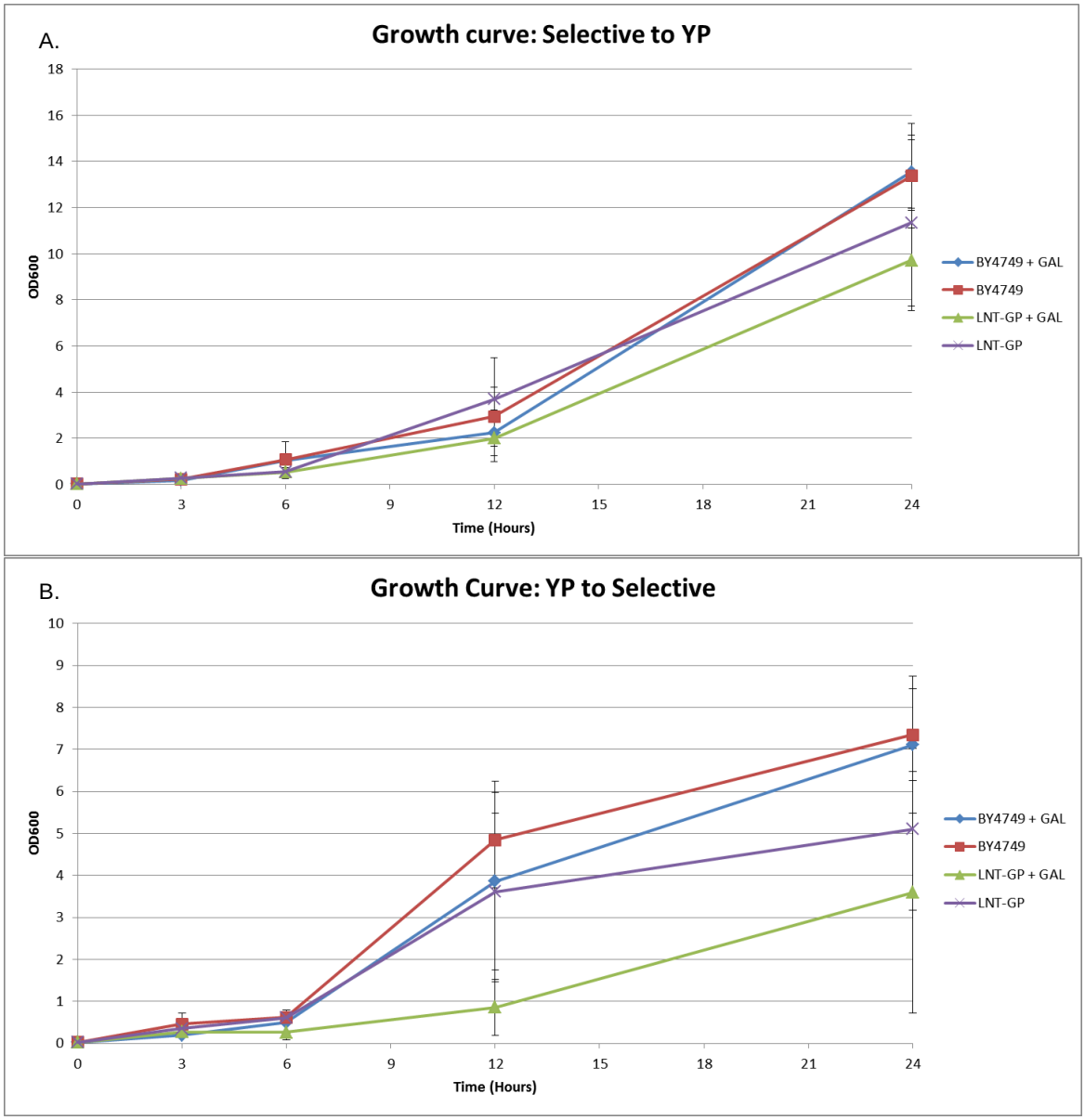

Figure 23. Growth curves for BY4742-LNT-GP and BY4742.

A) Shows the growth curves for BY4742-LNT-GP and BY4742 when going from selective culture to rich media with and without +GAL induction. B) shows the growth curves for BY4742-LNT-GP and BY4742 when going from selective culture to rich media with and without +GAL induction. Error bars show standard error of the mean (SE) at each sample point.

Once the growth curves were generated, viral production and titre was checked using Lenti-X™ qRT-PCR Titration Kit to measure the concentration of lentiviral genomes in the culture supernatant. This kit was chosen as it has previously been used to titre virus from cell free supernatant. As yeast cells have a cell wall lentiviral particles would be unable to bud. To allow viral particles to bud Zymolyase was used to digest away the cell wall and so that spheroplasts would form. Spheroplasts are osmotically sensitive, therefore, 1M sorbitol was used to prevent cell lysis. Two methods of viral particle release were tested; firstly, by the induction of viral production before Zymolyase treatment and secondly by culturing yeast spheroplasts in YPG-Sorbitol.

BY4742-LNT-GP was grown overnight in selective media to an OD_{600} of 10 (~$1x10^8$ cells/ml) before being transferred to YPG media and grown for 24 hours in a shaking incubator at $30^{O}C$. Samples were taken at 12 and 24 hours and treated with Zymolyase for 30 minutes in Spheroplast buffer to release viral particles. Concurrently BY4742-LNT-GP was grown overnight in selective media to an OD_{600} of 10 (~$1x10^8$ cells/ml) before being treated with Zymolyase for 30 minutes in Spheroplast buffer to Spheroplast the cells. The Spheroplasts were then suspended in YPG-Sorbitol media supplemented with Zymolyase and grown for 12 hours. Samples of supernatant were taken at 3, 6 and 12 hours for qRT-PCR. Viral RNA was extracted from 150µl of supernatant samples using the NucleoSpin RNA Virus Kit and DNase treated. These RNA samples were then tested in duplicate in accordance with the Lenti-X™ qRT-PCR Titration Kit instructions. A standard curve (Figure 24) was generated using dilutions of control RNA of a known titre and dilutions of the RNA samples were used to generate an amplification plot. Wells without any RNA where used as no template controls (NTC) to ensure that any amplification seen was caused by viral RNA in the samples and not from contamination of the reagents. The disassociation curves show that specific amplification only occurred in the positive controls and the samples retrieved from the

spheroplasts cultured in YPG-Sorbitol (Figure 25) however the titre from these cells was below the detectable range.

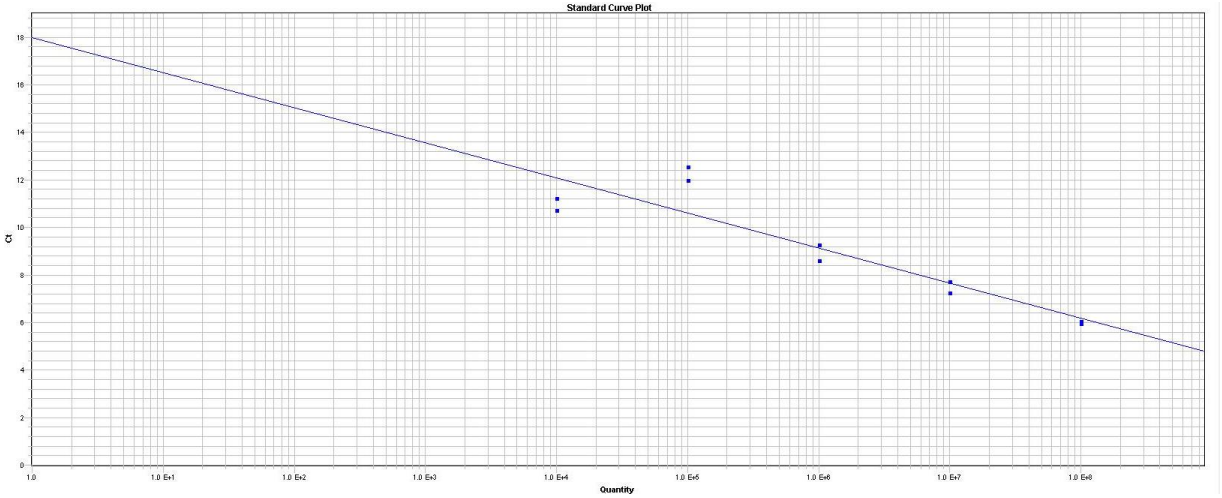

Figure 24. Plot of the Lenti-X™ qRT-PCR standard curve

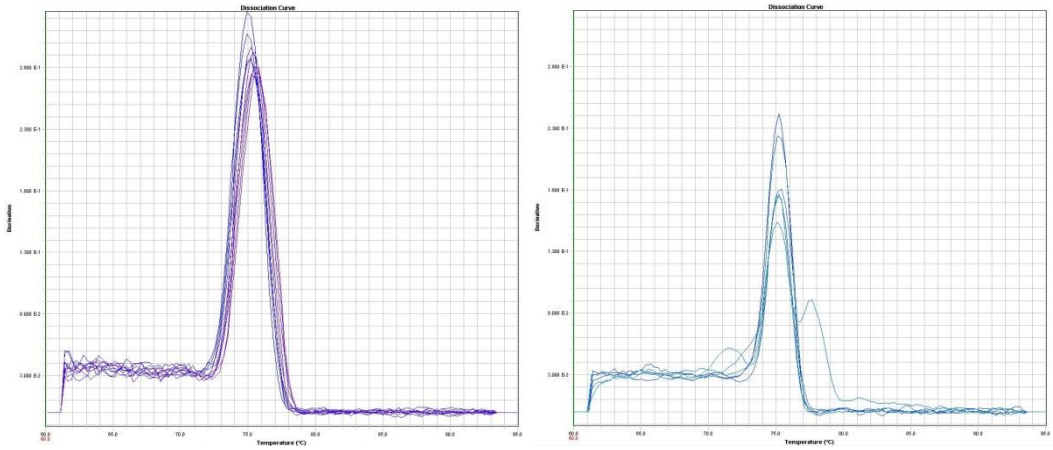

Figure 25. Comparison of the dissociation curves for the control RNA and RNA samples.

These graphs show the dissociation curves generated by the qRT-PCR for the control RNA (left) and YPG-Sorb 30^{o}C RNA sample (right). Both curves have their peaks at 75^{o}C showing that the products being amplified and measured by qRT-PCR in both the sample and control are the same.

Lentiviral vectors are normally produced in mammalian culture at 37^{o}C and it was suspected that the low titre could be caused by reduced activity of the viral protease when cultured at 30^{o}C. To test this theory BY4742-LNT-GP was grown overnight in selective and YPD media at 30^{o}C. The cells were pelleted by centrifugation before being resuspended in Spheroplast buffer and Zymolyase treated for 30 minutes at

30^{o}C. The spheroplasts were then transferred to YPG-Sorb and GSS Media, respectively, and grown for 12 hours at 34^{o}C and 37^{o}C in a shacking incubator. Samples of viral supernatant were taken at 12 hours and viral RNA was extracted from the samples using the NucleoSpin RNA Virus Kit followed by DNase treatment. These samples were then tested in duplicate in accordance with the Lenti-X™ qRT-PCR Titration Kit using the manufacturer's instructions. A standard curve was generated (Figure 26) using dilutions of control RNA of a known concentration representing a specific titre and dilutions of the RNA samples were used to generate an amplification plot. Wells without any RNA where used as no template controls (NTC). The results of this qRT-PCR show that viral RNA does appear in the supernatant at a higher concentration when cultures are grown at 34^{o}C (but not 37^{o}C) than when grown at 30^{o}C. The titre was highest with cells grown in YPG-Sorb with a titre of 1-2 X10^{4} copies per ml. This is however a much lower concentration than expected when compared to mammalian viral cultures which have titres of 10^{5} to 10^{7} copies per ml. The dissociation curves (Figure 27) for the samples and control RNA show that the amplified product in the samples is the same specific product as the control RNA and not spurious amplification. These results suggest that there may be packaged viral genomes being produced and released into the supernatant. To test whether mature viral particles are in fact budding from these spheroplasts, electron microscopy was performed to visually identify viral particles.

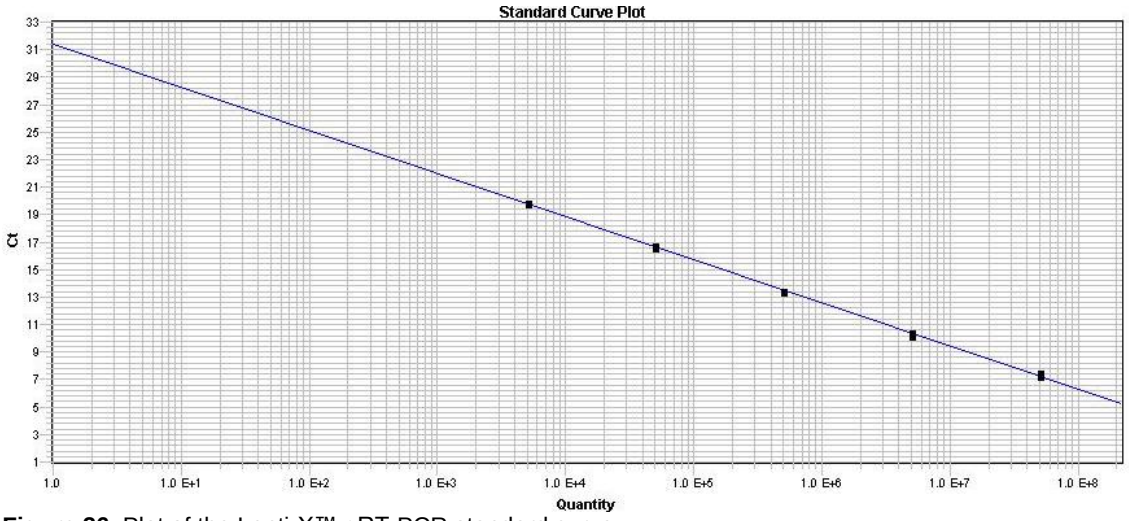

Figure 26. Plot of the Lenti-X™ qRT-PCR standard curve

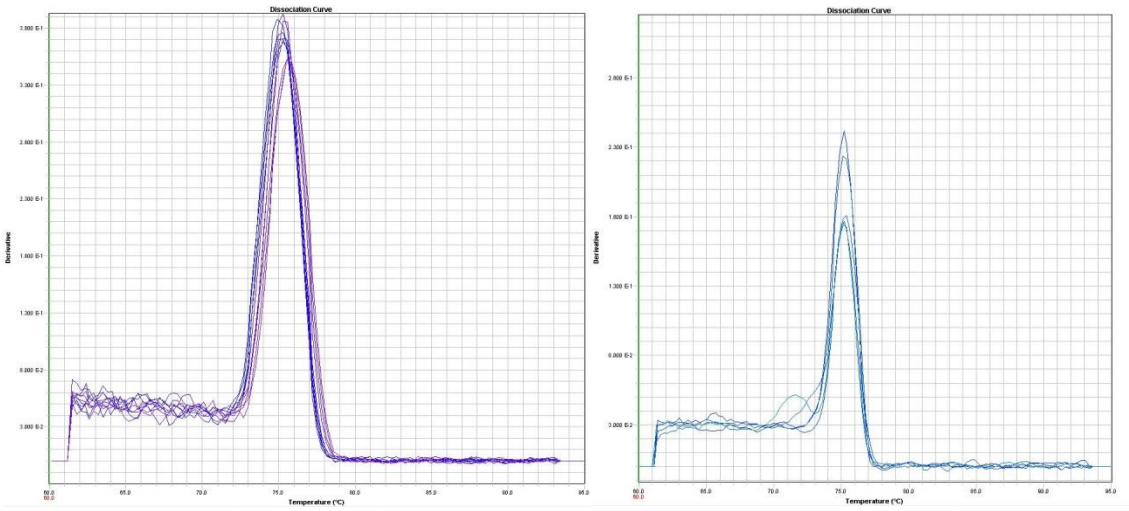

Figure 27. Comparison of the dissociation curves for the control RNA and RNA samples.

These graphs show the dissociation curves generated by the qRT-PCR for the control RNA (left) and YPG-Sorb 34^{o}C RNA sample (right). Both curves have their peaks at 75^{o}C showing that the products being amplified and measured by qRT-PCR in both the sample and control are the same.

3.5. Electron microscopy of Lentiviral particle formation

After producing a virus, cells were prepared for scanning and transmission electron microscopy to see if the yeast were producing mature lentiviral particles and if they were successfully budding from the cell surface. BY4742 and BY4742-LNT-GP were grown overnight in YPD at 30^{o}C before being pelleted and resuspended in Spheroplast buffer and Zymolyase treated for 30 minutes at 30^{o}C. The spheroplasts were then transferred to YPG-Sorb and GSS respectively and grown for 12 hours at 34^{o}C. The cells were then prepared for scanning electron microscopy (SEM) and transmission electron microscopy (TEM) as described in the methods and materials.

SEM was used for its ability to image surface structures of the yeast cell that would allow the identification of the characteristic blebbing of the cell membrane associated with lentiviral budding (Figure 28 and Figure 29). TEM was used for its ability to image cross sectional slices of prepared cells to look at internal structures and would allow the identification of lentiviral particles as well as their location and level of maturation (Figure 30).

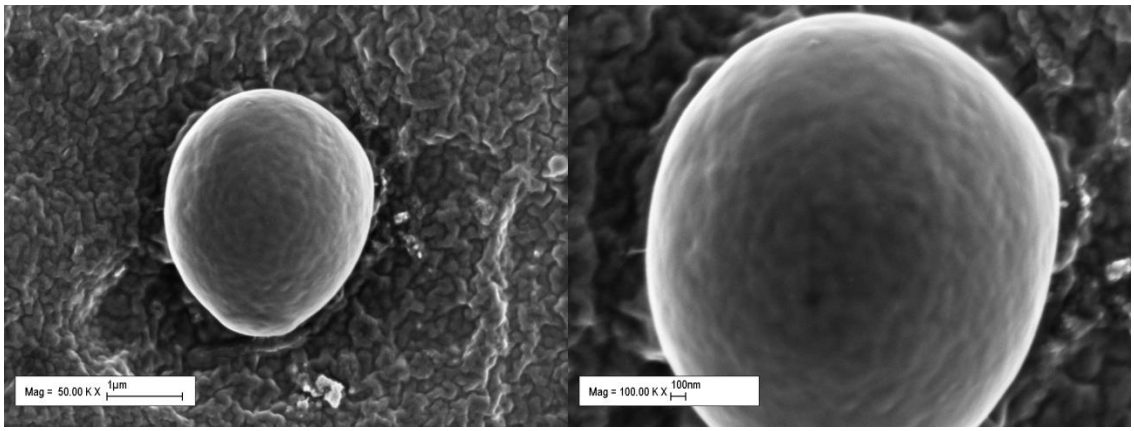

Figure 28. A representative example of a single BY4742 cell showing smooth membrane using SEM at 50,000x and 100,000x magnifications.

SEM of BY4742 clearly showed the smooth surface of these yeast cells after cell wall removal as would be expected.

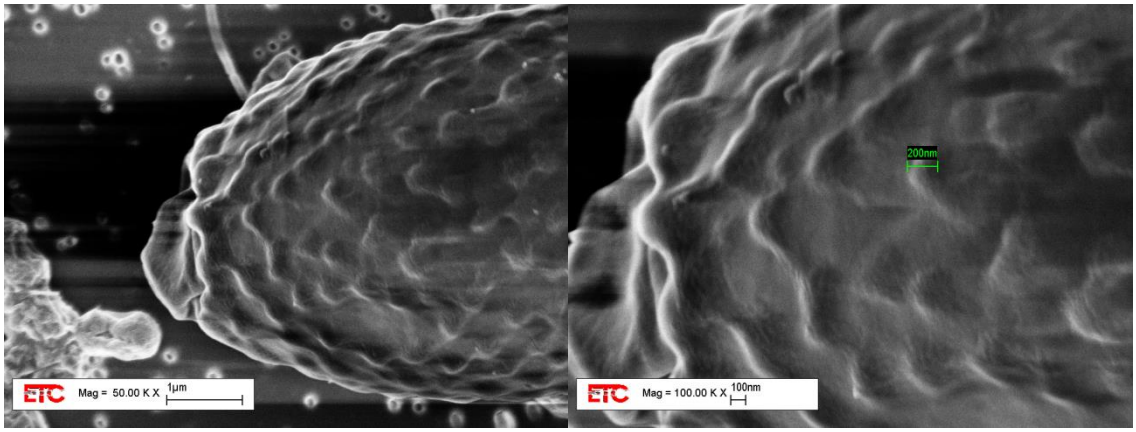

Figure 29. A representative example of a single BY4742-LNT-GP cell after transformation with virus packaging components and induction to generate virus particles.

SEM shows blebbing of the cell surface ranging in size from 150 to 200nm at 50,000X and 100,000x magnifications. No virus particles can be seen leaving the cell surface.

SEM images of BY4742 (Figure 28) after cell wall removal shows the smooth cell surface membrane that would be expected of yeast spheroplasts. SEM images of BY4742-LNT-GP (Figure 29), however, showed distinct blebbing of the cell surface ranging in size from 150 to 200nm which is highly suggestive of the early stages of lentiviral budding. Using TEM, electron dense particles of approximately 100nm in size could be seen possibly representing lentiviral particles, however, these particles appear to be retained within the yeast cells. Particles that appeared outside the cells were suspected to have derived from cell lysis as no budding was observed using SEM or TEM. Also no virus particles were identified within the intercellular space (Figure 30). When lentiviral particles mature, capsid proteins within the virus condense into an electron dense barrel shaped structure around the viral genome at the centre of the particle (Höglund *et al*. 1992). The viral particles in these images all lack this central structure showing that these particles were still in the immature stage of lentiviral formation.

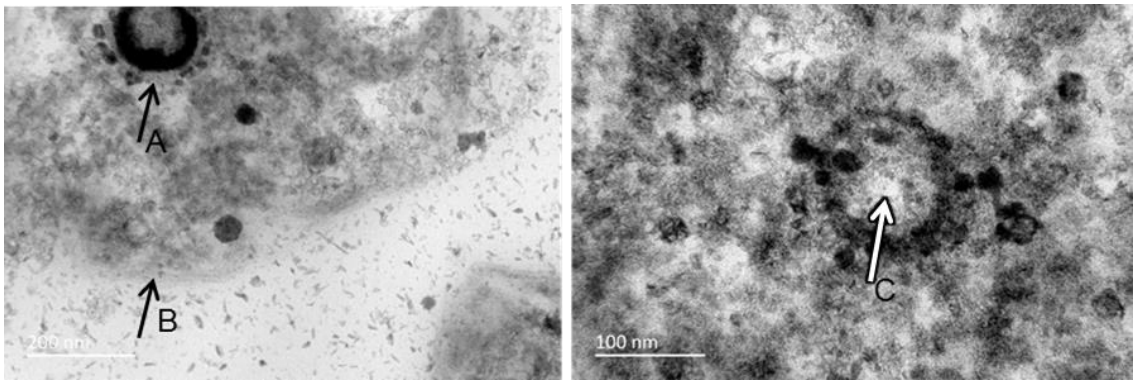

Figure 30. TEM images of suspected viral particles in yeast cells.

A) Shows an immature viral particle found in a cell. B) shows the lipid bilayer of the plasma membrane. The cytoplasm of the cell is darker than intracellular space due to its protein content. C) shows a viral particle lacking an electron dense barrel shaped structure indicative of an immature viral particle

Together SEM and TEM suggested that while viral particles are forming and attempting to bud from the surface of the yeast cell, these particles do not appear to be able to leave the cell surface as mature particles. These observations suggests that mature lentiviral vectors were not produced by these yeast cells and that the low viral titre produced is likely due to cell lysis as no budding could be observed.

4. Results: Suspected involvement of the virus integrase produced from the Gag-Pol expression vector in Yeast mutagenesis

4.1. Fluctuation Assay to determine the rate of mutation of the CAN1 gene caused by Gag-Pol expression

As shown in Chapter 3, the yeast strain BY4742-GP had spontaneously reverted at the *his3* gene locus. This was suspected to possibly be due to leaky expression of viral integrase from the Gag-Pol insert leading to mutagenesis of *his3* after DNA damage and repair of the genome following integrase activity. To test this possibility, a yeast fluctuation assay was performed to find the rate of mutation in wild type compared to Gag-Pol expressing yeast.

Fluctuation assays work by assaying multiple parallel cultures to find the fraction that do not develop resistance to selection over a fixed period of time. The fraction that does not develop resistance (P0) is proportional to the rate of mutation (u) of the gene of interest and the number of viable cells plated onto selective media (Nf'). Under specific conditions, by assaying control and treated cultures in parallel, the mutation rate of the strain can be calculated and attributed to the suspected mutagenic agent, which in this case is suspected to be the lentivirus integrase.

For these experiments canavanine was chosen as the selective agent. Canavanine is an arginine analog that is toxic to yeast cells by its incorporation into proteins. Canavanine selects for cells that have loss of function mutations in the CAN1 gene that is required for the import of canavanine into the cell. For this experiment, culture

conditions needed to be optimised so that colonies formed on the spot plates could be countable representing the effect of induced versus spontaneous CAN 1 mutagenesis.

For the optimisation experiment, BY4742 was first streaked onto SC-Arg and SC-Arg+Canavanine plates to ensure the wild type strain was sensitive canavanine. No colonies grew on the selective plates while they did grow on SC-Arg plates showing that the staring culture did not have a CAN1 mutation. Single colonies were then picked from the SC-Arg plate and used to inoculate complete synthetic media and grow overnight at 30^{o}C in a shaking incubator. The culture was then diluted 1:1000, 1:10,000 and 1:100,000 into YPD media. For each dilution, 500ul of diluted culture was added to each well of two 24 well plates. The plates were then incubated at 30^{o}C for 24 or 48 hours. 100ul of each culture was then spot plated onto dried SC-Arg+canavanine plates and incubated until colonies formed. After 24 hours colonies had formed on the spot plates and each set was analysed to find the best conditions for the main experiment. All the colonies from the 24 well plates that had been cultured for 48 hours were overgrown. Out of the 24 hour plates only the 1:10,000 dilutions produces spot plates with countable colonies with a ratio of null cultures to mutant cultures. Therefore for the main experiment a 1:10,000 dilution was used with 24 hour incubation.

To ensure the accuracy of the measured mutation rates, freshly transformed BY4742-GP yeast were used to ensure that any mutations that formed had occurred only after transformation. Firstly, single colonies from the transformation plate were streaked onto SC-Arg and SC-Arg+Canavanine to ensure the transformation procedure did not induce any CAN1 mutants. No colonies grew on the selective plates while they did grow on SC-Arg plates showing that the starting colony used did not have a CAN1 mutation. For the fluctuation assay, the appropriate selective media was inoculated with BY4742 and BY4742-GP taken from the SC-Arg plates. Each culture was then

grown overnight before being diluted 1:10,000 into YPD and YPG media. 500ul of diluted culture was then added to each well of a 24 well plate and incubated at 30^{o}C. After 24 hours of incubation 100ul of each culture was spot plated onto dried SC-Arg+canavanine plates. The remaining culture was collected, diluted 1:100 and spread on YPD plates. The plates were incubated for 24 hours to allow single colonies to form. The number of colonies on the YPD plates were then counted to find the concentration of viable cells. This was then used to calculate the number of viable cells per spot plated. The fraction of cultures that did not produce colonies on SC-Arg+canavanine plates was counted and used to find the mutation rate of each culture condition.

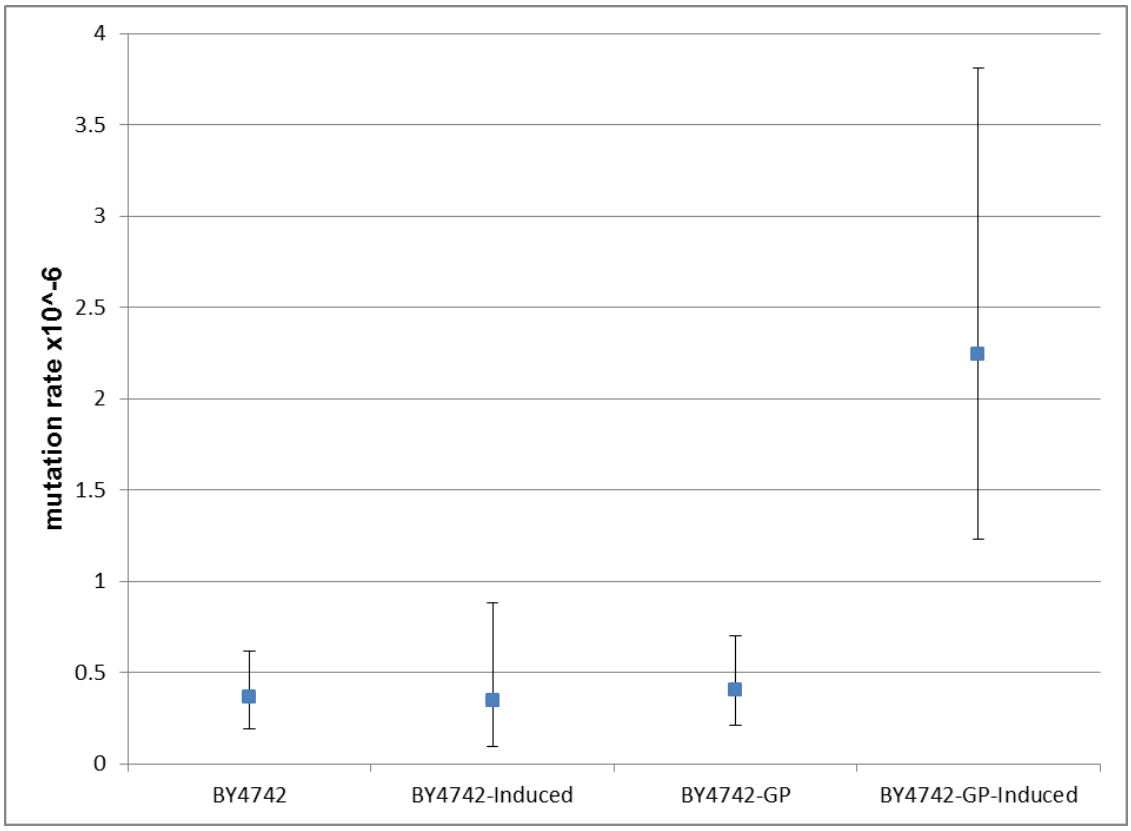

Figure 31. Graph showing the calculated mutation rates of BY4742 and BY4742-GP with and without induction. (Bars show 95% confidence intervals)

As can be seen in Figure 31 the rate of mutation in the yeast strain expressing the Gag-Pol construct was 6 fold higher than that of uninduced and wild type yeast

suggesting that Gag-Pol expression maybe causing mutagenesis of the yeast genome. Confidence intervals where used to find statistical significance with non-overlapping confidence intervals showing a significance of $p<0.01$ between the values.

4.2. Sequencing of mutated CAN1 gene locus

To further analyse these mutations we attempted to sequence the CAN1 gene from the mutated colonies to attempt to identify possible sequence motifs. HIV integrase has been shown to generate double strand breaks with 5-bp overhangs that can generate distinctive 5-bp duplications in the host genome (Sinha & Grandgenett 2005). Alternately DNA repair by non-homologous end joining can lead to addition, deletion or substitutions of base pairs around the ligation site (Heidenreich *et al*. 2003). By sequencing these mutations we hoped to identify the mechanism by which the CAN1 gene was mutated. 20 colonies were picked from the SC-Arg+canavanine drop plates and were sent to Genewiz, Inc (South Plainfield, NJ) for their Yeast Colony Sequencing service. At Genewiz the CAN1 gene locus was isolated from each colony via PCR amplification using the CAN1-F and CAN1-R primers. The PCR product was then purified and sequenced on an AB3730 sequencer. Of the 20 colonies sent for sequencing, 19 produced high quality reads that can be used to analyse for mutations as can be seen on the chromatogram (Figure 32).

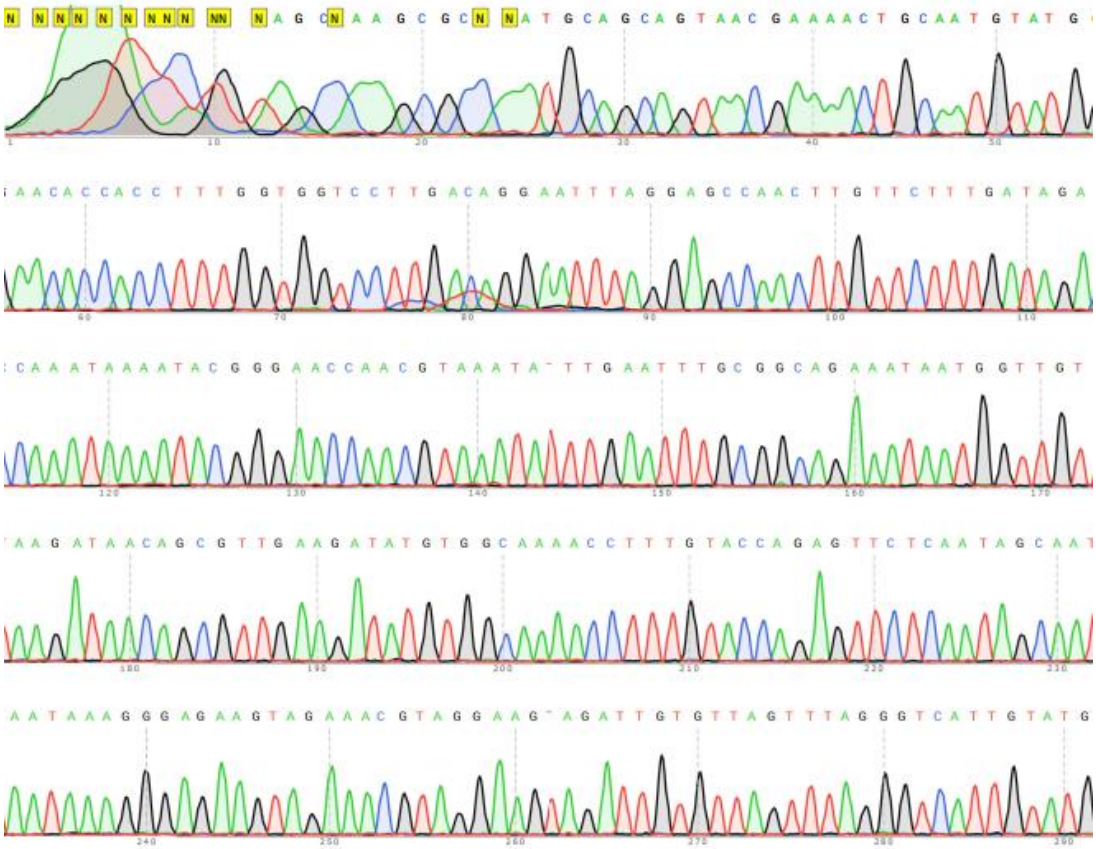

Figure 32. The first 5 lines of a successful chromatogram showing distinct clear peaks

The full sequence contig of CAN1 for each sample was generated from 3 forward and 3 reverse primer (Primer sequence in section 2.1.4.2.) sequencing reactions using the DNA Sequence Assembler v4 (2013, Heracle BioSoft, www.DnaBaser.com) software package. The entire length of the CAN1 locus was covered by at least 2 sequencing reactions. The sequence for each sample was then aligned to the reference sequence for CAN1 to look for mutations (Figure 33). From the 19 sequences, mutations where found in every sequence clustered into 8 unique mutations. All the mutations found where single base pair substitutions. The most common sequence motif appeared to be WWSWW with 47% (9/19) showing this motif with a single TTCAA sequence accounting for 26% of the mutations (5/19). The regularity of these mutations across different samples suggests that these mutations are targeted to these sequences.

```
   1  ATGACAAATT CAAAAGAAGA CGCCGACATA GAGGAGAAGC ATATGTACAA TGAGCCGGTC   60
  61  ACAACCCTCT TTCACGACGT TGAAGCTTCA CAAACACACC ACAGACGTGG GTCAATACCA   120
 121  TTGAAAGATG AGAAAAGTAA AGAATTGTAT CCATTGCGCT CTTTCCCGAC GAGAGTAAAT   180
 181  GGCGAGGATA CGTTCTCTAT GGAGGATGGC ATAGGTGATG AAGATGAAGG AGAAGTACAG   240
 241  AACGCTGAAG TGAAGAGAGA GCTTAAGCAA AGACATATTG GTATGATTGC CCTTGGTGGT   300
                             A(1,6)
 301  ACTATTGGTA CAGGTCTTTT CATTGGTTTA TCCACACCTC TGACCAACGC CGGCCCAGTG   360
              (10)C   G(2,7,18)
 361  GGCGCTCTTA TATCATATTT ATTTATGGGT TCTTTGGCAT ATTCTGTCAC GCAGTCCTTG   420
 421  GGTGAAATGG CTACATTCAT CCCTGTTACA TCCTCTTTCA CAGTTTTCTC ACAAAGATTC   480
 481  CTTTCTCCAG CATTTGGTGC GGCCAATGGT TACATGTATT GGTTTTCTTG GGCAATCACT   540
                                                      A(13)
 541  TTTGCCCTGG AACTTAGTGT AGTTGGCCAA GTCATTCAAT TTTGGACGTA CAAAGTTCCA   600
 601  CTGGCGGCAT GGATTAGTAT TTTTTGGGTA ATTATCACAA TAATGAACTT GTTCCCTGTC   660
 661  AAATATTACG GTGAATTCGA GTTCTGGGTC GCTTCCATCA AAGTTTTAGC CATTATCGGG   720
 721  TTTCTAATAT ACTGTTTTTG TATGGTTTGT GGTGCTGGGG TTACCGGCCC AGTTGGATTC   780
 781  CGTTATTGGA GAAACCCAGG TGCCTGGGGT CCAGGTATAA TATCTAAGGA TAAAAACGAA   840
                                A(4,5)
 841  GGGAGGTTCT TAGGTTGGGT TTCCTCTTTG ATTAACGCTG CCTTCACATT TCAAGGTACT   900
 901  GAACTAGTTG GTATCACTGC TGGTGAAGCT GCAAACCCCA GAAAATCCGT TCCAAGAGCC   960
 961  ATCAAAAAAG TTGTTTTCCG TATCTTAACC TTCTACATTG GCTCTCTATT ATTCATTGGA   1020
        (11,15,16)T      A(12,17)
1021  CTTTTAGTTC CATACAATGA CCCTAAACTA ACACAATCTA CTTCCTACGT TTCTACTTCT   1080
1081  CCCTTTATTA TTGCTATTGA GAACTCTGGT ACAAAGGTTT TGCCACATAT CTTCAACGCT   1140
1141  GTTATCTTAA CAACCATTAT TTCTGCCGCA AATTCAAATA TTTACGTTGG TTCCCGTATT   1200
1201  TTATTTGGTC TATCAAAGAA CAAGTTGGCT CCTAAATTCC TGTCAAGGAC CACCAAAGGT   1260
1261  GGTGTTCCAT ACATTGCAGT TTTCGTTACT GCTGCATTTG GCGCTTTGGC TTACATGGAG   1320
1321  ACATCTACTG GTGGTGACAA AGTTTTCGAA TGGCTATTAA ATATCACTGG TGTTGCAGGC   1380
1381  TTTTTTGCAT GGTTATTTAT CTCAATCTCG CACATCAGAT TTATGCAAGC TTTGAAATAC   1440
1441  CGTGGCATCT CTCGTGACGA GTTACCATTT AAAGCTAAAT TAATGCCCGG CTTGGCTTAT   1500
1501  TATGCGGCCA CATTTATGAC GATCATTATC ATTATTCAAG GTTTCACGGC TTTTGCACCA   1560
1561  AAATTCAATG GTGTTAGCTT TGCTGCCGCC TATATCTCTA TTTTCCTGTT CTTAGCTGTT   1620
1621  TGGATCTTAT TTCAATGCAT ATTCAGATGC AGATTTATTT GGAAGATTGG AGATGTCGAC   1680
                T(3,8,9,14,19)
1681  ATCGATTCCG ATAGAAGAGA CATTGAGGCA ATTGTATGGG AAGATCATGA ACCAAAGACT   1740
1741  TTTTGGGACA AATTTTGGAA TGTTGTAGCA TAG                               1773
```

Figure 33. Full sequence of the CAN1 gene with the 19 mutations highlighted

This figure shows the full sequence of the CAN1 gene. The 19 mutation sites found during sequencing of the 19 samples are highlighted in yellow with the mutated base and the number of the sample it was found in annotated directly below.

Taken together, the data suggests that the expression of Gag-Pol in our yeast does increase the rate of mutation and that these mutations appear to be targeted to a specific sequence motif. As the mutations all appear to be single nucleotide substitutions it's possible that they are being generated by non-homologous end joining of double strand breaks (Heidenreich *et al*. 2003). While this data does support our theory that HIV integrase is generating double strand breaks it does not explain the sequence specificity. HIV integrase has been expressed in yeast (Parissi *et al*. 2003) and shown to cause mutations, but the mutations where not sequenced. It is possible that yeast host genes are being incorporated into the pre-integration complex that are targeting integrase to this sequence but further studies would be required to elucidate the exact mechanism.

5. Results: Production of VSV-G from yeast for use as an enhancer of DNA transfection

5.1. Transformation of yeast with Gal-VSV-G production vectors

VSV-G protein has been shown previously to be an enhancer of several types of non-viral transfection (Okimoto *et al*. 2001; Abe, Miyanohara, *et al*. 1998; Imazu *et al*. 2000). Its ability to bind to cells, induce uptake across most cell lines, mediate endosomal escape and provide serum resistance to chemical carriers makes VSV-G a promising transfection agent. However its use is severely hampered by its production which requires repeated transfection of mammalian cells for each batch of VSV-G and time consuming methods to purify the protein. In this chapter, an attempt was made to develop a continuous yeast production system of VSV-G for use in DNA transfections.

The wild type yeast strain BY4742 was grown overnight from tested stocks and the yeast was made competent using the Frozen EZ Yeast Transformation II Kit. The vector pRS303-VSV-G was digested with the BsiWI enzyme to linearize it within the *his3* gene before being checked on a gel for complete linearization. Approximately 1µg of DNA was used to transform BY4742 cells before being plated on -HIS selective plates and incubated at 30^{o}C until colonies formed. Cells transformed without DNA were used as a negative control. Each colony was then streaked onto -LEU, -HIS and -URA and incubated at 30^{o}C to check that the transformants had the correct phenotype. All the colonies grew on the -HIS plate while none grew on -URA or -LEU plates.

To confirm the production of VSV-G RNA RT-PCR was performed. Single colonies of BY4742-VSVG were picked to inoculate an overnight culture of YPD and the cells were then transferred to 50ml YPG media for 3 hours. After 3 hours of induction, total RNA

was extracted from the cells with the Nucleospin® RNA Isolation Kit and RT-PCR was performed for a 500bp fragment of VSV-G (Figure 34).

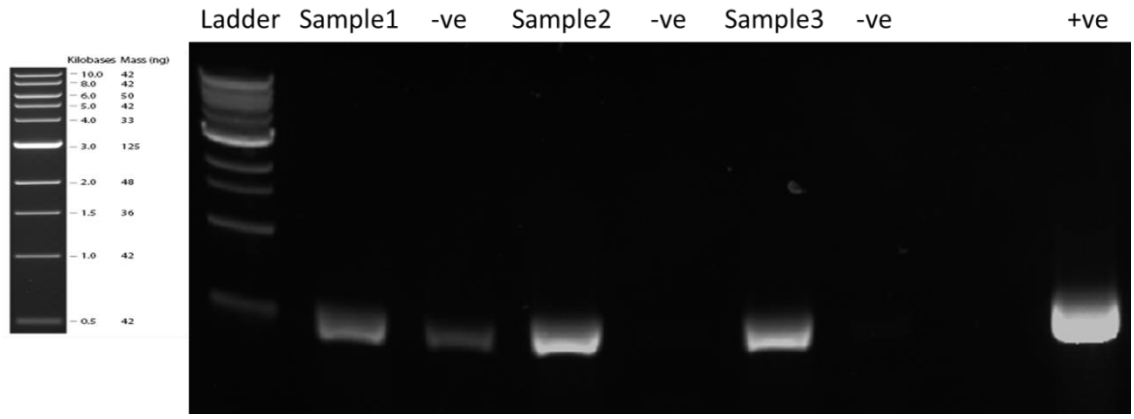

Figure 34. RT-PCR to determine VSV-G expression by the yeast BY4742-VSVG.

RT-PCR was performed on 3 BY4742-VSV-G expanded colonies. Total RNA was extracted from these samples after induction using galactose. In all three samples, positive amplifications of 500bp were obtained from each. In sample 1, a positive amplification was identified in the –ve control and therefore this sample was excluded from further study. Sample 2 and 3 also gave positive amplifications and following this, sample 2 was chosen to generate VSV-G further.

5.2. Toxicity of VSV-G production in yeast

Once the VSV-G RNA had been confirmed to be correct, growth curves were generated to test the toxicity of VSV-G production in the yeast strain. VSV-G is known to cause cytotoxicity in mammalian production systems due to syncytia formation and his cytotoxicity limits prolonged VSV-G productivity by mammalian cells. It was hypothesised that the cell wall of yeast would block cell fusion and greatly reduce the observed cytotoxicity allowing prolonged expression of VSV-G. This would allow the build up of VSV-G in the yeast periplasmic space which could then be released by Zymolyase treatment.

Colonies of BY4742 and BY4742-VSVG were then picked from selective plates and used to inoculate 10ml of YPD media before being grown overnight. The next day the optical density of the cultures were measured at 600nm. These cultures were used to inoculate YPG and YPD to an OD of 0.020 before being grown for 72 hours in a shaking incubator at 30^{O}C. OD600 readings were recorded at 3, 6, 12, 24, 48, 60 and 72 hours. This was repeated 3 times to obtain an average from the data and its statistical relevance. The data shows that over a period of 72 hours induction, VSV-G did not produce any toxicity, compared to wild type yeast under the same conditions.

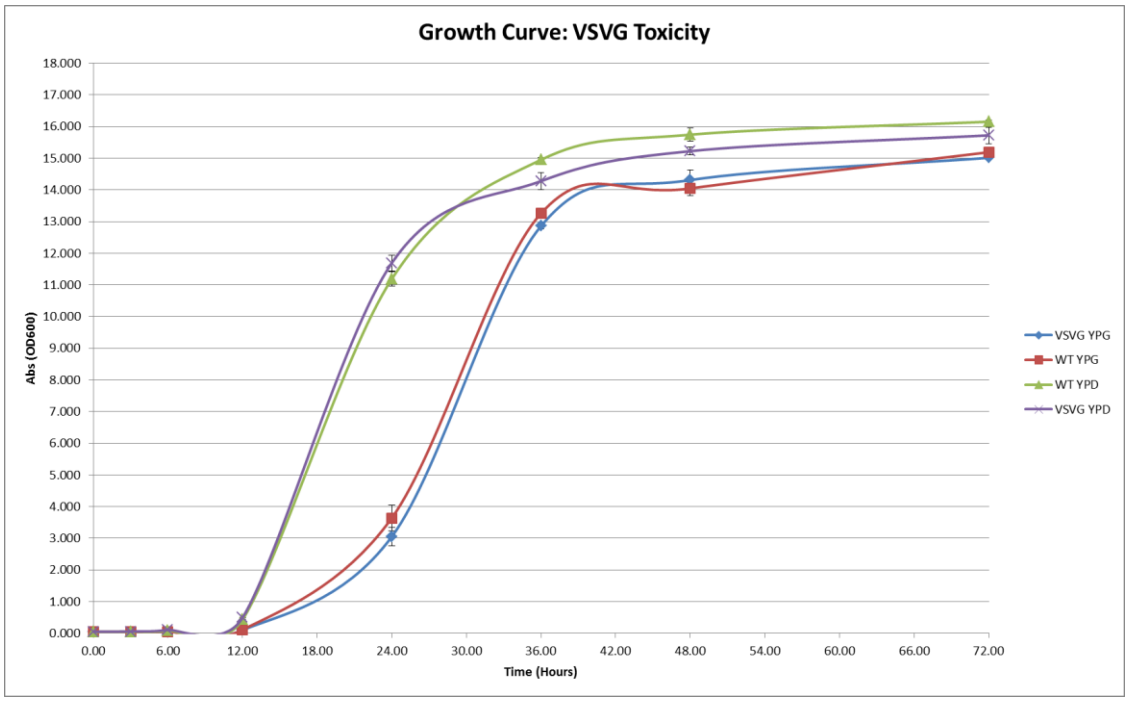

Figure 35. Growth curves of WT and VSV-G strains of BY4247

This graph compares the growth curves of WT and VSV-G strains of BY4247 with and without induction. The growth rate of both strains where not statistically different ($P < 0.05$) after 3 repeats showing that there is very little toxicity for VSVG production.

Once it was shown that VSV-G production in yeast was not toxic, VSV-G was generated and purified from both yeast and mammalian cells before being tested for quantity, concentration and purity.

5.3. VSV-G purification and quantification

It was hypothesised that the cell wall of yeast would stop the release of the expressed VSV-G into the culture media and therefore allows concentration of VSV-G by pelleting yeast cells at low speed centrifugation. This would allow quicker and cheaper concentration and purification of VSV-G compared to its production from mammalian cells and circumvents the requirements of VSV-G ultracentrifugation in the procedure. For yeast VSV-G, a 50ml YPD culture was inoculated with BY4247-VSV-G and grown overnight at 30^{o}C in a shaking incubator. The culture was then pelleted at 5000RPM for 5 minutes before being resuspended in 50ml of YPG and incubated overnight to allow expression of VSV-G. The cells were again pelleted and resuspended in 1ml Spheroplast buffer with 50U Zymolyase and incubated at 30^{o}C for 30 minutes. The spheroplasts were then pelleted for 10 minutes at 5000RPM and the VSV-G containing supernatant collected.

For VSV-G production by mammalian cells, HEK293T cells in a total of 50ml of culture media were transformed with PEI (as described in the material and methods). The conditioned media was collected at 24, 48 and 72 hours post transfection and pooled. The media was then spun down overnight at 90,000xg and the pellet resuspended in 100ul of PBS.

Once both yeast and mammalian VSV-G had been purified SDS-PAGE and western blot analysis was performed to identify the VSV-G protein. The samples from both mammalian and yeast VSV-G was run on a 12% polyacrylamide gel and transferred onto a membrane and stained with the HRP conjugated anti-VSV-G tag antibody. This antibody recognises the amino acid sequence 501-511 (YTDIEMNRLGK) in the VSV-G protein. As can be seen in Figure 36 both samples showed specific bands for VSV-G

matching the theoretical weight (46kDa) of VSV-G but mammalian VSV-G shows higher glycosylation then the yeast derived VSV-G (58kDa).

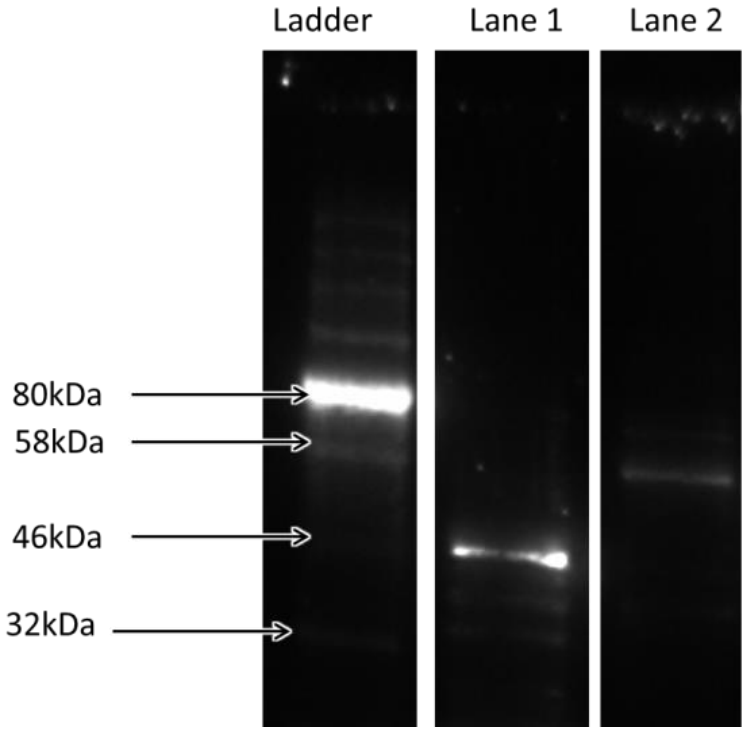

Figure 36. Western blot analysis of VSV-G produced from yeast and mammalian cells.

Lane 1 detects a protein of approximately 46kDa from the yeast VSV-G sample. Lane 2 detects a protein of approximately 58kDa from the mammalian VSV-G sample.

Once the presence of VSV-G in the samples had been confirmed, the quantity and purity of the VSV-G samples was measured. To get the quantity of VSV-G an ELISA assay was performed against a standard curve generated from dilutions of synthetic VSV-G peptide to estimate the concentration of VSV-G in each sample. To get the purity of the VSV-G samples the total protein concentration of the samples where compared to the results of the ELISA assay to find the percentage of VSV-G in the total protein. To do this a BCA assay was performed to find the total protein concentration of the samples.

Table 12. The Total protein concentration, VSV-G concentration and approximate purity of the yeast and mammalian VSV-G samples

	BCA Assay (ug/ml)	VSVG ELISA (ug/ml)	~% Purity
Yeast VSVG	325.600	220.100	67.598
Mammalian VSVG	135.161	78.571	58.131

Table 12 shows the purity of VSV-G from mammalian and yeast samples where within 10 percentage points of each other suggesting that these samples are equivalent in terms of purity, however the concentration of VSV-G in the yeast sample extracted from 50ml of yeast culture was nearly 3 fold that of the VSV-G produced from the same volume of transfected mammalian cells. This suggests that the yeast production method was much more efficient then the mammalian method for the production of VSV-G.

Once the quantity and purity of both mammalian and yeast VSV-G had been determined, experiments were performed to analyse the efficiency of mammalian and yeast VSV-G as a transfection reagent.

5.4. Comparing VSV-G produced by yeast and mammalian cells for transfection efficiency

As it was suspected that the levels of glycosylation of VSV-G when produced by yeast and mammalian cells was different (Figure 36), it was important to determine whether VSV-G generated by yeast was as efficient as VSV-G generated by mammalian cells. To determine this, a series of transfections were performed in the BT474 cell line. The BT474 cell line was chosen because these cells are generally resistant to transfection and would, therefore be ideal to show enhancement of transfection by non-viral VSV-G mediated gene transfer. Varying quantities of PEI were used to aid VSV-G transfection with a fixed amount of VSV-G protein (1ug) and GFP DNA (8ug).

Each combination of PEI and VSV-G was repeated six times and analysed separately before averaging the results and statistical evaluation of significance. 48 hours after transfection the cells were harvested by trypsinization, washed in PBS and resuspended in 200ul of accuMAX™. The total number of cells and GFP positive cells were then counted by imaging flow cytometry using an Amnis® ImageStream®X Mark II Imaging Flow Cytometer.

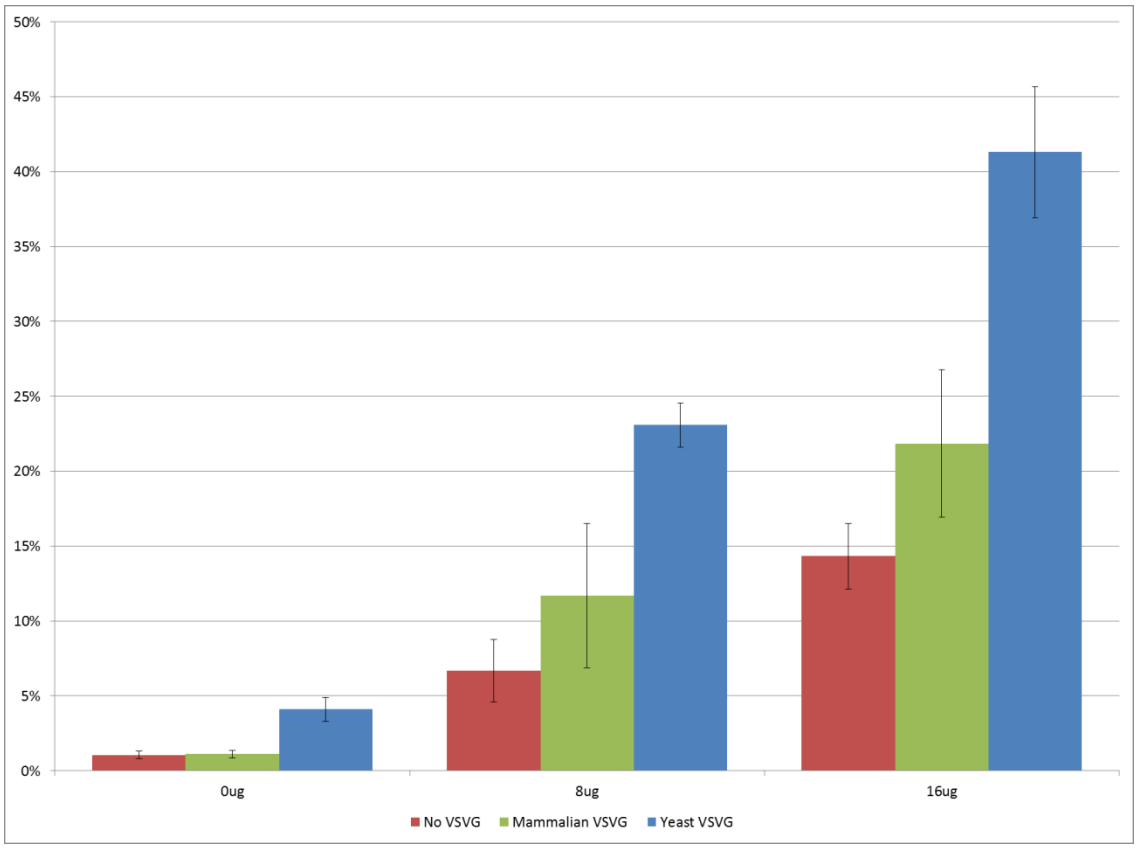

Figure 37. Transfection efficiency of yeast and mammalian VSV-G

A comparison of the transfection efficiency of a range of PEI/DNA complexes with either yeast, mammalian or no VSV-G added to the transfection mixture. Efficiency is given as the percentage of GFP positive cells 48 hours after transfection with error bars showing +/- SE.

This shows that VSV-G had very little effect on transfection efficiency on its own without PEI, however yeast derived VSV-G did significantly increase the transfection efficiency of PEI at concentrations of either 8ug or 16ug of PEI above that of PEI alone ($P < 0.05$) and above that of mammalian VSV-G ($P < 0.05$).

To ensure that this apparent difference of transfection efficiency was not caused by cellular toxicity and loss of cell viability, a viability assay was performed on cells transfected with and without yeast VSV-G. Trypan blue only stains dead cells allowing the visual identification of viable and non-viable cells which can be used as a measure of viability. Cells were transfected as stated previously and harvested 48 hours after transfection. 100ul of harvested cell suspension was mixed with 100ul of 0.4% (w/v) trypan blue solution and incubated for 5 minutes. A small amount of this solution was then applied to a hemacytometer counting chamber and the number of stained and unstained cells was counted and percentage viability was calculated for each sample. Each transfection and viability assay was repeated 6 times.

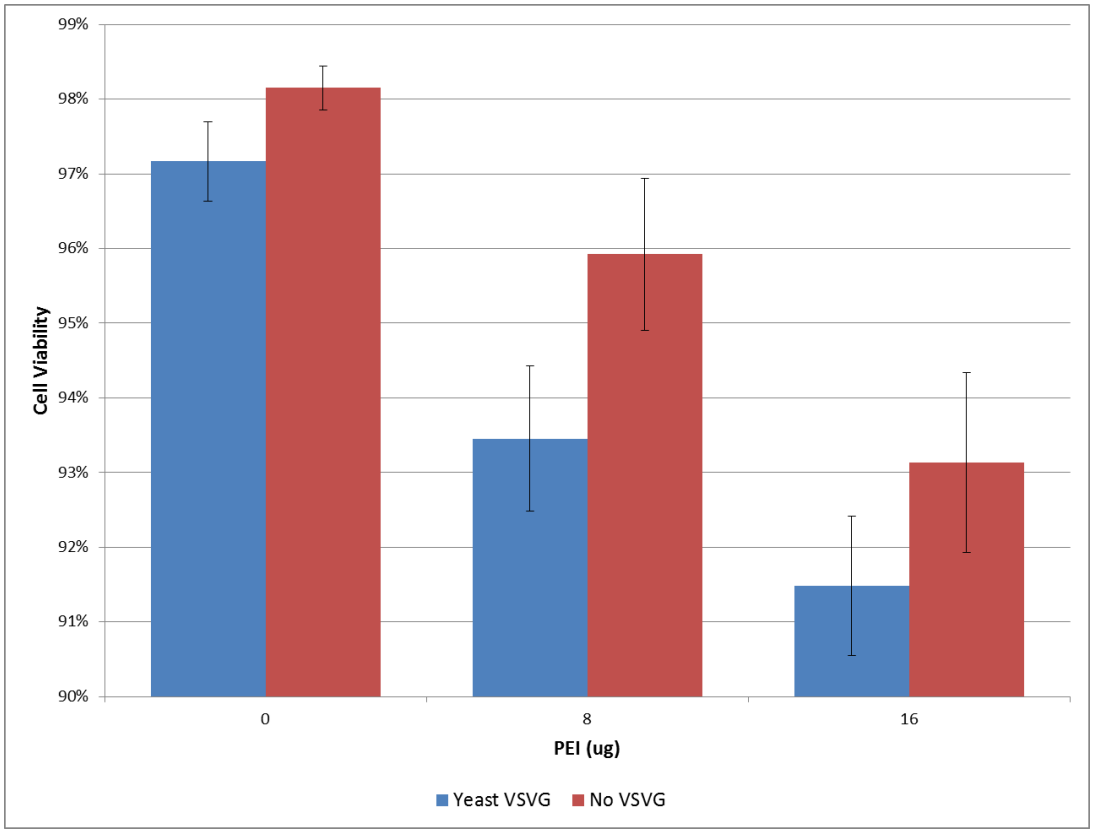

Figure 38. Viable cells 48 hours after transfection.

This graph shows the viability of cells transfected with PEI and PEI/yeast VSVG 48 hours after transfection. Viability is given as a percentage with error bars showing SE.

Figure 38 shows cell viability at around 98%+/-SE in the absence of PEI of VSV-G. Cell viability fell slightly with increasing concentrations of PEI but remained greater than 90% even at the highest concentration of PEI used. The addition of yeast VSV-G did also appear to decrease the viability of transfected cells slightly compared to PEI alone, however this effect was not statistically significant after 6 experiments ($P < 0.05$ at all three concentrations) and the difference was fairly small at 1-2%.

6. Discussion

6.1. Generation of lentivirus particles from *Saccharomyces cerevisiae*

Integrating retroviral and lentiviral vectors have huge potential in the field of gene therapy for their ability to permanently integrate therapeutic genes into patient cells. In one of the earliest trials of its kind, 10 patients were treated with an amphotropic murine leukemia virus based vector carrying the γ_c gene coding for the interleukin-2 receptor subunit gamma in order to correct severe combined immunodeficiency-X1 (Cavazzana-Calvo *et al*. 2000). At first this trial had seemed very successful with 9 of the 10 patients treated having stably restored immunity, however it soon emerged that integration of the retroviral genome near the LIM domain–only 2 (LMO2) proto-oncogene in 4 of the patients had caused them to the development of leukaemia (Hacein-Bey-Abina *et al*. 2008). While all integrating vectors carry the risk of insertional oncogenesis, it has been shown that lentiviruses do so at a much lower rate than retroviruses (Modlich *et al*. 2009) mainly due to the difference in preferred integration site of these viruses with lentiviruses preferring to insert into the gene transcription unit and retroviruses targeting the promoter region of the gene (Hematti *et al*. 2004). The lower genotoxicity of lentivirial vectors has led to a greater use of this vector system in clinical trials with 114 trials having been performed to date.

Production of lentiviral particles has mainly been achieved using transient transfection 293T human embryonic kidney cells with plasmids carrying the packaging and envelope trans component genes gag/pol and env, respectively and the cis backbone vector sequences that include vector LTRs, and an internal promoter to drive transgene expression. Transient transfection has mainly been used due to the toxicity of VSV-G

expression when it is used to pseudotype viral particles. Alternate envelopes, like RD114 (Sandrin *et al*. 2002), have been developed for pseudotyping vectors that do not generate syncytia and cell toxicity when expressed by producer cells. However, the ability concentrate VSV-G pseudotyped virus via ultracentrfugation (Burns *et al*. 1993; Bartz *et al*. 1996) as well as the broad host range (Schlegel *et al*. 1983) that VSV-G offers has often made it the envelope of choice.

This method of virus production is difficult to scale up for clinical production. Packaging cell lines have been developed that stably integrate the genes needed for lentiviral vector production avoiding the need for transient transfection of 293T producer cells (Sanber *et al*. 2015). However, whilst this improves scalability, subsequent transfection of the VSV-G gene to pseudotype the vector particles still limits prolonged virus production due to producer cell syncytia and toxicity. In addition, the production of lentivirus vectors by mammalian cells is still limited by the slow growth of these cells with low vector yield and the cost of mammalian cell culture reagents.

Yeast systems like *Saccharomyces cerevisiae* have been used for many years in the production of recombinant proteins and the majority of these are produced from this species of yeast (Huang *et al*. 2010). *Saccharomyces cerevisiae* has previously been shown to be capable of producing HIV-1 based viral like particles containing virus core proteins generated using just the gag gene and these particles have been shown to successfully bud from the plasma membrane once the cell wall was removed (Sakuragi *et al*. 2002). These particles, however, could not undergo maturation without the Pol encoding genes. Yeast have also been shown to be capable of packaging the full HIV-1 genome into these Gag only particles (Tomo *et al*. 2013). Together these studies suggest that yeast could be used as a producer cell system for lentiviral vectors and, therefore, in this study we investigated that possibility. If yeast could produce HIV

based lentiviruses then we hypothesised that the yeast cell wall should allow pseudotyping via VSV-G without syncytia formation between vector producing cells.

To do this, we generated three integrating yeast plasmids containing the viral packaging vector, transfer vector and VSV-G envelope to generate a HIV-1 based yeast producer system. The plasmids pRS305-Gag-Pol and pRS306-LNT-GFP, that used LEU2 and URA3 selection respectively, were successfully integrated into the BY4742 strain to produce BY4742-LNT-GP rendering this strain Leu2 and Ura3 positive whilst remaining His3 negative. To ensure the functionality of the vectors, the BY4742-LNT-GP cell line was grown in induction media and tested for the production of Gag-Pol and the LNT backbone mRNA. While attempting to integrate the pRS303-VSV-G vector that carried the His3 marker into BY4742-LNT-GP it was discovered that BY4742-LNT-GP had become His3 positive. This was discovered when attempting to transform BY4742-LNT-GP with the pRS303-VSV-G plasmid carrying the His3 marker when colonies formed on the negative control plate and then following subsequent streak tests of our stocks of BY4742, BY4742-GP and BY4742-LNT-GP on selective plates to check each of these strains for the His3 marker. The results of these streak plates show that although our original unmodified strain BY4742 was His3 negative our stocks of BY4742-GP and BY4742-LNT-GP had become His3 positive. We next retested our BY4742-GP and BY4742-LNT-GP strains for LNT and Gag-Pol mRNA production to ensure our stocks were not cross contaminated, which they were not. These findings suggested that somehow our BY4742-GP and BY4742-LNT-GP strains had undergone a His3 positive reversion at a time point following the introduction of the Gag-Pol vector. This led to the work in this thesis to examine the potential for the BY4742-LNT-GP strain to produce lentiviral particles in parallel to an investigation of the cause of His3 positive reversion that we discovered (Discussed in section 6.2).

To find the best conditions for lentiviral production in yeast, we used qRT-PCR to titrate the number of viral RNA genomes released into the culture media by BY4742-LNT-GP. As yeast cells have a thick cell wall, viral particles would not be released into the supernatant. We hypothesised that partial or complete budding may still take place underneath the cell wall, leading to accumulation of viral particles in the periplamsic space. To test this we induced the production of viral particles from BY4742-LNT-GP for 12 and 24 hours to allow time for lentiviral particles to bud and accumulate in the periplamsic space. The cell wall was then digested with Zymolyase to release virus particles into the supernatant for collection and titration of viral genomes. Viral RNA was extracted from the sample supernatants and the number of viral genomes was assessed by qRT-PCR using primers recognising the WPRE region of the lentiviral genome. Our results show that the titre of viral particles were below the detectable range in these supernatants suggesting that viral particles were not generated or that they do not accumulate in the periplamsic space.

We then went on to test if viral particles would bud from cultured spheroplasts. To do this, BY4742-LNT-GP was Zymolyase treated to digest the cell wall and form spheroplasts. These spheroplasts where cultured for 3, 6 and 12 hours in induction media supplemented with Zymolyase to stop the cell wall regenerating. Viral RNA was again extracted from the supernatant samples and the number of viral genomes was assessed by qRT-PCR. This experiment found a very low titre of viral genomes in our 12 hour supernatants. This suggested that while viral particles maybe budding from yeast cells when the cell wall is removed this process is possibly being somehow inhibited.

Lentiviral vectors are normally cultured in mammalian cells at 37^{o}C were as our yeast strain had been cultured at 30^{o}C. As lentiviral proteins are likely to be optimised for the

higher temperature of mammalian cell culture we hypothesised that our low titre could be increased by adjusting the temperature of the culture process. BY4742-LNT-GP spheroplasts where generated before being cultured for 12 hours at 34^{o}C and 37^{o}C in induction media. Viral RNA was extracted from these supernatant samples and tested using the same qRT-PCR method as before. We found that cells cultured at 37^{o}C produced very low concentrations of viral genomes whereas at 34^{o}C we had a much higher concentration of viral genomes with titres of 1-2 X10^{4} copies per ml. While this was an improvement in titre this represents the production from ~1X10^{7} yeast cells compared to lentiviral production from mammalian cultures which generally produce titres in the range of 10^{5} to 10^{7} copies per ml from ~1X10^{5} cells before concentration.

To test whether mature viral particles are in fact budding from these spheroplasts we then used electron microscopy to visually identify viral particles. BY4742-LNT-GP and BY4742 cells were Zymolyase treated to produce spheroplasts and cultured for 12 hours in induction media at 34^{o}C before being prepared for both SEM and TEM. SEM was used for its ability to image surface structures that would allow the identification of the characteristic blebbing of the cell membrane associated with lentiviral budding. TEM was used for its ability to image cross sectional slices of cells to look at internal structures and would allow the identification of lentiviral particles as well as their location and level of maturation. SEM images of BY4742-LNT-GP showed distinct blebbing of between 150 and 200nm which is highly suggestive of lentiviral early stage budding and this blebbing was not found in BY4742. TEM images show that electron dense particles of approximately 100nm in size are found within the BY4742-LNT-GP cells close to the plasma membrane. These particles are of the correct size and shape for lentiviral particles but appear to be retained within the yeast cells with no virus particles identified within the intercellular space. When lentiviral particles mature, capsid proteins within the virus condense into an electron dense barrel shaped

structure around the viral genome at the centre of the viral particle (Höglund *et al*. 1992). For all the particles observed under TEM, none were found with the condensed viral core indicative of mature particles. Together the SEM and TEM images suggest that while viral particles are forming and attempting to bud from the surface of the yeast cell, these particles do not appear to be able to mature or bud from the cell surface. These observations suggests that mature lentiviral vectors were not produced by these yeast cells and that the low viral titre of vector genomes produced was possibly due to yeast cell lysis as no budding could be observed

Conclusions and future work

From the above results, it appears that while viral particles are forming and beginning to bud from the surface of our yeast strain they are unable complete this process. While the lentiviral Gag genes have previously been shown to produce virus like particles that successfully bud into the supernatant (Sakuragi *et al*. 2002) no previous study has shown expression of Gag-Pol in yeast. From the work of this thesis it seems likely that the expression of the full Gag-Pol construct is not compatible with viral budding in yeast and the absence of cleavage of the gag/pol polyprotein in yeast may be the cause. As the Pol genes are required for lentiviral particle maturation and viral genome integration into host cells, simple removal of the Pol region would not be an alternative to overcome the lack of mature virus production in yeast.

In future work, it would be useful to test whether separating Gag and Pol expression into two separate expression vectors is sufficient to rescue viral budding and maturation in yeast. Work by Westerman *et al*. 2007 has shown that the addition of the Vpr lentiviral protein sequence to the beginning of the Pol region allows for the Pol polyprotein to be expressed separately from Gag while still allowing the packaging of Pol into viral particles. By generating separate Gag and vpr-Pol yeast expression

vectors under different conditional promoters we could induce expression of these two components separately to test weather PoI expression in yeast blocks the budding of particles in yeast.

6.2. Measurement of suspected mutagenesis in yeast expressing Gag-Pol

Our findings show that our BY4742-GP strain had undergone a His3+ reversion which we surmised had been caused by expression of viral integrase in some way affecting the rate of genome mutation. We recently demonstrated in our laboratory (unpublished data) that the HIV integrase causes DNA double strand breaks by using γH2AX immunocytochemistry on mammalian cells following infection by LV. We hypothesised that yeast expressing high levels of HIV integrase from the Gag-Pol construct induced double strand breaks in the yeast genome. Double strand breaks close to the His3 locus have been shown to increase the reversion rate of the mutated gene to wildtype (Schiestl et al. 1988) via homologous recombination repair events. To our knowledge full length Gag-Pol has not previously been expressed in yeast with most studies only looking at Gag expression (Sakuragi *et al*. 2002; Tomo *et al*. 2013) however the overexpression of HIV integrase has been performed in yeast (Parissi *et al*. 2003). In this study HIV integrase was expressed in yeast cells and shown to increase the rate of mutation at the *LYS2* and *LYS3* loci twofold *in vivo* and that non-sequence-specific nuclease activity of HIV integrase lead to DNA double strand breaks *in vitro*. Together these data suggests that viral integrase from the pol region of our construct maybe increasing the rate of mutation in our yeast strain leading to the His3 positive reversion we observed.

To test if the expression of the Gag-Pol polyprotein led to an increased rate of mutation we performed a fluctuation assay. The fluctuation assay was originally developed to estimate the rate of mutation of bacterial cells in culture (Luria & Delbrück 1943). In this method cells are grown in multiple parallel non-selective liquid cultures before being plated onto selective solid media. The fraction of cultures that produce no resistant

colonies (P0) on selective media is proportional to the number of viable cells that where plated (*Nf'*) and the mutation rate (*u*). This allows the mathematical estimation of the mutation rate from the fraction of cultures that produce no resistant colonies using the equation below.

$$u = \frac{-\ln(\mathrm{P0})}{Nf'}$$

For our experiments canavanine was chosen as the selective agent as it has previously been used in a fluctuation assay in yeast (Lang & Murray 2008). Canavanine is an arginine analog that is toxic to yeast cells and selects for loss of function mutations in the CAN1 gene. The CAN1 gene expresses a plasma membrane arginine permease that is required for transport of basic amino acids, including canavanine. This is similar to *hprt* mammalian mutagenesis model in which the development of 6-thioguanine resistance is used to calculate the rate of insertional mutagenesis by retroviruses (King *et al*. 1985; Themis *et al*. 2003).

To ensure the accuracy of the measured mutation rates, single colonies from a freshly transformed BY4742-GP plate were streaked onto SC-Arg and SC-Arg+Canavanine to ensure the transformation procedure did not induce any CAN1 mutants. No colonies grew on the canavanine plates showing that the transformation procedure did not introduce any CAN1 mutations. For the fluctuation assay single colonies of BY4742 and BY4742-GP were taken from SC-Arg plates and used to inoculate appropriate selective media. Each culture was then grown overnight before being diluted 1:10,000 into YPG media for induction of Gag-Pol expression and YPD for suppression of Gag-Pol expression. The diluted cultures where split into 24 parallel cultures and incubated for 24 hours at 30^{o}C. After 24 hours of incubation 100ul of each culture was spot plated onto dried SC-Arg+canavanine plates. The fraction of the 24 cultures that produced no

resistant colonies gives us the value of P0. The remaining culture was collected, diluted 1:100, spread on YPD plates and incubated for 24 hours. The number of colonies was then counted to find the concentration of viable cells and this was used to calculate the number of viable cells that where plated (*Nf*'). These values where then used to calculate the rate mutation in the CAN1 gene of BY4742 and BY4742-GP with and without induction. We found that the induction of Gag-Pol expression in BY4742-GP led to a 6 times higher rate of mutation than the background levels. This shows that the expression of one or more of the Gag-Pol genes does increase the rate of mutation and potentially genomic damage. Previous studies have shown that viral integrase expression can cause mutation (Parissi *et al*. 2003) making it highly likely that the mutation rate in our yeast strain is caused by the expression of integrase from the Gag-Pol polyprotein.

To further analyse these mutations we attempted to sequence the CAN1 gene from the mutated colonies to attempt to identify possible sequence motifs. HIV integrase generates a double strand break with 5-bp overhang that can generate distinctive 5-bp duplications in the host genome (Sinha & Grandgenett 2005). Alternately DNA repair by non-homologous end joining can lead to addition or deletion of base pair around the ligation site (Heidenreich *et al*. 2003). By sequencing these mutations we hoped to identify the mechanism by which the CAN1 gene was mutated.

The CAN1 gene from 20 colonies of mutated BY4742-GP were picked from the SC-Arg+canavanine drop plates and were sent to Genewiz, Inc (South Plainfield, NJ) for their Yeast Colony Sequencing service. At Genewiz the CAN1 gene locus was isolated from each colony via PCR amplification and the PCR product was then purified and sequenced. The full sequence of CAN1 for each sample was generated from 3 forward and 3 reverse primer sequencing reactions and assembled into contigs using the DNA

Sequence Assembler v4 (2013, Heracle BioSoft, www.DnaBaser.com) software package. The entire length of the CAN1 locus was covered by at least 2 sequencing reactions. The sequence for each sample was then aligned to the reference sequence for CAN1 to look for mutations. From the 19 sequences, mutations where found in every sequence and were clustered into 8 unique mutations. All the mutations found where single base pair substitutions. The most common sequence motif for these mutations appeared to be WWSWW with 47% (9/19) of mutations showing this motif with a single TTCAA sequence accounting for 26% (5/19) of the mutations. The regularity of these mutations across different samples suggests that these mutations are targeted to these sequences. Taken together, the data suggests that the expression of Gag-Pol in our yeast does increase the rate of mutation and that these mutations appear to be targeted to a specific sequence motif.

Conclusions and future work

Several models developed to study the genotoxic effects of gene therapy and the modifications that have been made to reduce these effects. While these models have tested the ability of different vectors to cause inactivating mutations (Themis *et al*. 2003), activating mutations (Bokhoven *et al*. 2009), immortalisation (Modlich *et al*. 2009) and *in vivo* oncogenesis (Ranzani *et al*. 2013) they are all based on the assumption that the genotoxic effects of retroviral and lentiviral vectors are caused by the insertion of DNA into the genome, leading to abnormal gene expression. The data from our fluctuation assay suggests that Gag-Pol is also a source of mutation and potential genotoxicity, even in the absence of recombinant viral genomes. Our sequencing data suggests that the expression of Gag-Pol is generating single nucleotide substitutions in WWSWW sequences with a high specificity to the TTCAA sequence. DNA repair of double strand breaks by non-homologous end joining has been shown to generate base pair substitutions (Heidenreich *et al*. 2003) suggesting

that HIV integrase is generating double strand breaks in the yeast genome leading to the increase mutation rate.

While this data does support our theory that HIV integrase is generating double strand breaks it does not explain the sequence specificity we observed. HIV integrase has previously been expressed in yeast (Parissi *et al*. 2003) and shown to cause mutations, but the mutations where not sequenced. It is possible that yeast host genes are being incorporated into the pre-integration complex that are targeting integrase to this sequence leading to this sequence specificity. To elucidate if this is the mechanism behind the sequence specific mutations we observed further experiments would be done to find if yeast host genes and viral integrase are interacting. A yeast two-hybrid screening of yeast DNA binding proteins would be able to identify if any are incorporated into a pre-integration complex with integrase.

As well as identifying the mechanism behind the site specific mutations we would also test if lentiviral integrase could cause similar mutations in mammalian models. Lentiviral particles lacking viral genomes could be tested in the tumour-prone mouse (Ranzani *et al*. 2013) or the fetal mouse models (Nowrouzi *et al*. 2013) to test whether the increased rate of mutagenesis could lead to increased oncogenesis without insertional events. If this is found to be the case, our fluctuation assay could be used as a simple assay to test the genotoxicity of different integrating vectors and help in the development of safer future vectors.

6.3. Production of VSV-G from yeast as a potential transfection reagent

Non-viral gene therapy vectors offer distinct advantages over viral vectors due to their lower immunogenicity and mutagenicity but currently suffer from low transfer efficiency. Non-viral agents like polybrene and lipofectin have also been found to be sensitive to serum inactivation and clearance *in vivo*, greatly reducing there transfection efficiency (Yang & Huang 1997; Escriou *et al*. 1998; Tandia *et al*. 2003). VSV-G has been used for many years to pseudotype viral vectors to increase cell tropism and increase transfer efficiency to many cell types including those normally refractory to other methods of gene transfer (Yee *et al*. 1994; Burns *et al*. 1993). More recently purified VSV-G protein has been found to capable of complexing with non-viral transfection reagents, such as polybrene (Okimoto *et al*. 2001), Lipofectin (Abe, Miyanohara, *et al*. 1998) and liposomes (Imazu *et al*. 2000; Shoji *et al*. 2004), to greatly increase their transfection efficiency and cell tropism. It has also been shown that this VSV-G containing complexes become resistant to serum inactivation both *in vitro* (Abe, Miyanohara, *et al*. 1998) and *in vivo* (Hirano *et al*. 2002). This makes VSV-G is a promising transfection reagent for non-viral gene therapy.

Currently, VSV-G is produced by the transient transfection of 293T cells with a VSV-G expression vector followed by purification by ultra-centrifugation (Miyanohara 2012). The fusogenic properties of VSV-G limit its prolonged production due to formation of syncytia and cell toxicity (Eslahi *et al*. 2001). Large scale production is further restricted by the tedious and expensive production regime required by mammalian culture and ultra-centrifugation. It was hypothesised that the cell wall of yeast would prevent the fusogenic properties of VSV-G from leading to cell toxicity and would therefor allow for long term expression of VSV-G. The cell wall would also prevent VSV-G from being

released from the cell into the culture media allowing the concentration of VSV-G containing yeast cells by low speed centrifugation and release of VSV-G by Zymolyase treatment to digest the cell wall. Together this would allow the concentration of VSV-G from yeast without the need for ultra-centrifugation. In this study we attempted to address the problems associated with VSV-G production by creating a yeast production system.

The VSV-G gene was cloned into the pRS303 vector with a galactose inducible promoter to produce the pRS303-VSV-G vector. This vector was then integrated into BY4742 to produce the BY4742-VSV-G yeast strain. To ensure the functionality of the VSV-G vector the BY4742-VSV-G yeast strain was grown in induction media and tested for the production of VSV-G mRNA using RT-PCR. Once VSV-G mRNA production was confirmed the toxicity of VSV-G production was tested. It was hypothsised that the cell wall of yeast would prevent the fusogenic properties of VSV-G from leading to syncytia formation and toxicity. To test this, BY4742 and BY4742-VSVG were grown with and without induction for 72 hours and OD600 readings were recorded at 3, 6, 12, 24, 48, 60 and 72 hours as a measure of cell density at these time points. This would allow the generation and comparison of growth curves to identify if VSV-G expression caused cell toxicity and lower growth rate. Comparison of these growth curves show that there is no significant difference in growth rate of yeast between BY4742 and BY4742-VSVG under induction. This data shows that the production of VSV-G in yeast does not lead to increased cell toxicity suggesting that prolonged expression of this protein is viable.

We next began testing yeast produced VSV-G against mammalian produced. A saturated culture of BY4742-VSVG was grown overnight in induction media to induce VSV-G production. The cells were then concentrated using low-speed centrifugation

before removing the cell wall with Zymolyase. The VSV-G containing supernatant was separated and stored. For the mammalian VSV-G, HEK293T cells were transformed with VSV-G expression vector and conditioned media was collected at 24, 48 and 72 hours post transfection and pooled. The VSV-G was then concentrated away from the media using ultracentrifugation. Once both yeast and mammalian VSV-G had been purified we used SDS-PAGE and western blot analysis to characterise the proteins. One band was found for both yeast and mammalian produced VSV-G but the bands were of different weights. The theoretical weight of un-glycosylated VSV-G is 46kDa which is the weight of the band found in our yeast produced VSV-G sample. The mammalian VSV-G band had a weight of approximately 58kDa which matches the weight of glycosylated VSV-G produced in mammalian systems (Shoji *et al*. 2004).

Once we had confirmed the presence of VSV-G in our samples we began testing the concentration of VSV-G and the purity achieved in both methods. To find the concentration of VSV-G in our samples we used an ELISA assay using dilutions of synthetic VSV-G peptide as a standard curve to estimate the concentration of VSV-G in each sample. From this we found that the concentration of VSV-G from the yeast production method is 3 times higher than the mammalian sample despite only using low speed centrifugation to concentrate yeast VSV-G. To find the purity of the VSV-G in our samples we performed a BCA total protein assay to find total concentration of protein in our samples and, by comparing this to the concentration of VSV-G, we could calculate the purity. From this data we found that the purity of our samples are very similar at 58-68%. Together this data shows that this yeast based production system is capable of producing VSV-G on an equal or greater scale than the previously used mammalian system without compromising purity.

We then tested the effectiveness of our yeast VSV-G against mammalian produced VSV-G in transfecting a resistant cell line. The BT474 cell line was chosen because this cell line is generally resistant to transfection and would, therefore, be ideal to show enhancement of transfection by non-viral VSV-G mediated gene transfer. Varying quantities of PEI were used to aid VSV-G transfection with a fixed amount of VSV-G protein (1ug) and GFP DNA (8ug). 48 hours after transfection the cells were harvested and analysed using Imaging Flow Cytometer to the find the percentage of GFP expressing cells as a measure of transfection efficiency. Our data shows that the VSV-G produced from yeast increases the transfection efficiency of PEI above that of PEI on its own ($P < 0.05$) and mammalian VSV-G ($P < 0.05$). We further tested the yeast VSV-G against no VSV-G in a viability assay to ensure that the difference in transfection efficiency was not due to different toxicities. We found that, while yeast VSV-G does slightly increase the toxicity of PEI, this effect was not statistically significant after 6 experiments ($P > 0.05$ at all three concentrations) and the difference was fairly small at 1-2%.

Conclusions and future work

Much of the development of gene therapy vectors has been with viral methods of DNA transference due to the high transfection efficiency of these vectors. These vectors, however, suffer from higher immunogenicity and (and in the case of integrating vectors) higher genotoxicity. Non-viral gene therapy vector have the potential to be much safer than viral vectors but suffer from there low transfection efficiency. VSV-G has previously been shown to be effective at increasing transfection efficiency of non-viral vectors, but the production of VSV-G from mammalian cell culture restricts the scalability of its production. The data presented in this study shows that VSV-G can be produced and partially purified from yeast using low speed centrifugation and that this VSV-G is as effective a transfection reagent as VSV-G produced from mammalian cells

purified via ultracentrifugation. This means that VSV-G production from yeast could easily be scaled up to the levels required for clinical use.

In the future studies yeast produced VSV-G should be tested in other cell lines and in an *in vivo* model to further validate it's abilities to enhance the efficiency of transfection for use in clinical non-viral gene therapy. Once shown to work *in vivo* our VSV-G expression system could be further improved by optimising the system for large scale production. Codon optimisation (Chang *et al*. 2006) of the coding sequence of VSV-G could be tested for its ability to increase protein expression levels as well as testing several different promoters to find which is best for long term expression. Additionally several methods of further purification, for example aqueous two-phase systems (ATPS) (Merchuk *et al*. 1998), could be tested to find the most cost efficient method for producing clinical grade protein product for VSV-G

Purified VSV-G has previously been shown to be able to effectively pseudotype virus particles lacking an envelope protein in a cell free manner (Okimoto *et al*. 2001). If it is not possible to generate full lentiviral particles from yeast, due to the problems associated with Gag-Pol expression, lentiviral particles without envelope could be generated from a producer cell line continuously without VSV-G toxicity. These particles could then be pseudotyped with VSV-G produced from our yeast production system. This would overcome the problems of Gag-Pol expression in yeast that we discovered in the earlier part of this study, allowing the continuous production of viral particles from mammalian producer cell lines and the continuous production of VSV-G in yeast. In future work we will generate an envelope-less lentiviral producer cell line to test whether these two products could then be combined to produce infectious VSV-G pseudotyped particles without the toxicity with similar titre and transfection efficiency as current systems.

7. References

Abe, a, Chen, S.T., Miyanohara, a & Friedmann, T., 1998. In vitro cell-free conversion of noninfectious Moloney retrovirus particles to an infectious form by the addition of the vesicular stomatitis virus surrogate envelope G protein. *Journal of virology*, 72(8), pp.6356–6361.

Abe, a, Miyanohara, a & Friedmann, T., 1998. Enhanced gene transfer with fusogenic liposomes containing vesicular stomatitis virus G glycoprotein. *Journal of virology*, 72(7), pp.6159–6163.

Aiken, C., 1997. Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. *Journal of virology*, 71(8), pp.5871–7.

Aiken, C., Konner, J., Landau, N.R., Lenburg, M.E. & Trono, D., 1994. Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. *Cell*, 76(5), pp.853–64.

Amano, K., Chiba, Y., Kasahara, Y., Kato, Y., Kaneko, M.K., Kuno, A., Ito, H., Kobayashi, K., Hirabayashi, J., Jigami, Y. & Narimatsu, H., 2008. Engineering of mucin-type human glycoproteins in yeast cells. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9), pp.3232–7.

Anderson, D.B., Laquerre, S., Ghosh, K., Ghosh, H.P., Goins, W.F., Cohen, J.B. & Glorioso, J.C., 2000. Pseudotyping of glycoprotein D-deficient herpes simplex virus type 1 with vesicular stomatitis virus glycoprotein G enables mutant virus attachment and entry. *Journal of virology*, 74(5), pp.2481–7.

Anderson, W.F., Blaese, R.M. & Culver, K., 1990. The ADA human gene therapy

clinical protocol. *Human Gene Therapy*, 1(3), pp.327–362.

Andreansky, S., Soroceanu, L., Flotte, E.R., Chou, J., Markert, J.M., Gillespie, G.Y., Roizman, B. & Whitley, R.J., 1997. Evaluation of genetically engineered herpes simplex viruses as oncolytic agents for human malignant brain tumors. *Cancer research*, 57(8), pp.1502–9.

Ardehali, A., Fyfe, A., Laks, H., Drinkwater, D.C., Qiao, J.H. & Lusis, A.J., 1995. Direct gene transfer into donor hearts at the time of harvest. *The Journal of thoracic and cardiovascular surgery*, 109(4), pp.716-9-20.

Armentano, D., Sookdeo, C.C., Hehir, K.M., Gregory, R.J., St George, J.A., Prince, G.A., Wadsworth, S.C. & Smith, A.E., 1995. Characterization of an adenovirus gene transfer vector containing an E4 deletion. *Human gene therapy*, 6(10), pp.1343–53.

Ashorn, P., McQuade, T.J., Thaisrivongs, S., Tomasselli, A.G., Tarpley, W.G. & Moss, B., 1990. An inhibitor of the protease blocks maturation of human and simian immunodeficiency viruses and spread of infection. *Proceedings of the National Academy of Sciences of the United States of America*, 87(19), pp.7472–6.

Bally, M.B., Harvie, P., Wong, F.M.P., Kong, S., Wasan, E.K. & Reimer, D.L., 1999. Biological barriers to cellular delivery of lipid-based DNA carriers. *Advanced Drug Delivery Reviews*, 38(3), pp.291–315.

Bartz, S.R., Rogel, M.E. & Emerman, M., 1996. Human immunodeficiency virus type 1 cell cycle control: Vpr is cytostatic and mediates G2 accumulation by a mechanism which differs from DNA damage checkpoint control. *Journal of virology*, 70(4), pp.2324–31.

Becerra, S.P., Clore, G.M., Gronenborn, A.M., Karlström, A.R., Stahl, S.J., Wilson, S.H.

& Wingfield, P.T., 1990. Purification and characterization of the RNase H domain of HIV-1 reverse transcriptase expressed in recombinant Escherichia coli. *FEBS Letters*, 270(1–2), pp.76–80.

Bedikian, A.Y. & Del Vecchio, M., 2008. Allovectin-7 therapy in metastatic melanoma. *Expert opinion on biological therapy*, 8(6), pp.839–44.

Bergelson, J.M., Cunningham, J.A., Droguett, G., Kurt-Jones, E.A., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L. & Finberg, R.W., 1997. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science (New York, N.Y.)*, 275(5304), pp.1320–3.

Berthet, C., Raj, K., Saudan, P. & Beard, P., 2005. How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(38), pp.13634–9.

Bohenzky, R.A., Lefebvre, R.B. & Berns, K.I., 1988. Sequence and symmetry requirements within the internal palindromic sequences of the adeno-associated virus terminal repeat. *Virology*, 166(2), pp.316–327.

Bokhoven, M., Stephen, S.L., Knight, S., Gevers, E.F., Robinson, I.C., Takeuchi, Y. & Collins, M.K., 2009. Insertional gene activation by lentiviral and gammaretroviral vectors. *Journal of virology*, 83(1), pp.283–94.

Bonander, N., Hedfalk, K., Larsson, C., Mostad, P., Chang, C., Gustafsson, L. & Bill, R.M., 2005. Design of improved membrane protein production experiments: quantitation of the host response. *Protein science : a publication of the Protein Society*, 14(7), pp.1729–40.

Boussif, O., Lezoualc'h, F., Zanta, M. a, Mergny, M.D., Scherman, D., Demeneix, B. & Behr, J.P., 1995. A versatile vector for gene and oligonucleotide transfer into cells

in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences of the United States of America*, 92(16), pp.7297–7301.

Bouyac-Bertoia, M., Dvorin, J.D., Fouchier, R. a, Jenkins, Y., Meyer, B.E., Wu, L.I., Emerman, M. & Malim, M.H., 2001. HIV-1 infection requires a functional integrase NLS. *Molecular cell*, 7(5), pp.1025–35.

Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M. & Stevenson, M., 1993. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature*, 365, pp.666–669.

Burke, R.L., Tekamp-Olson, P. & Najarian, R., 1983. The isolation, characterization, and sequence of the pyruvate kinase gene of Saccharomyces cerevisiae. *The Journal of biological chemistry*, 258(4), pp.2193–201.

Burns, J.C., Friedmann, T., Driever, W., Burrascano, M. & Yee, J.K., 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 90(17), pp.8033–7.

Buscail, L., Bournet, B., Vernejoul, F., Cambois, G., Lulka, H., Hanoun, N., Dufresne, M., Meulle, A., Vignolle-Vidoni, A., Ligat, L., Saint-Laurent, N., Pont, F., Dejean, S., Gayral, M., Martins, F., Torrisani, J., Barbey, O., Gross, F., Guimbaud, R., Otal, P., Lopez, F., Tiraby, G. & Cordelier, P., 2015. First-in-man phase 1 clinical trial of gene therapy for advanced pancreatic cancer: safety, biodistribution, and preliminary clinical findings. *Molecular therapy : the journal of the American Society of Gene Therapy*, 23(4), pp.779–89.

Bushman, F.D., Fujiwara, T. & Craigie, R., 1990. Retroviral DNA integration directed by

HIV integration protein in vitro. *Science (New York, N.Y.)*, 249(4976), pp.1555–8.

Camerini, D. & Seed, B., 1990. A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site. *Cell*, 60(5), pp.747–54.

Canivet, M., Hoffman, A.D., Hardy, D., Sernatinger, J. & Levy, J.A., 1990. Replication of HIV-1 in a wide variety of animal cells following phenotypic mixing with murine retroviruses. *Virology*, 178(2), pp.543–51.

Canonaco, F., Schlattner, U., Pruett, P.S., Wallimann, T. & Sauer, U., 2002. Functional expression of phosphagen kinase systems confers resistance to transient stresses in Saccharomyces cerevisiae by buffering the ATP pool. *The Journal of biological chemistry*, 277(35), pp.31303–9.

Carlisle, R.C., Bettinger, T., Ogris, M., Hale, S., Mautner, V. & Seymour, L.W., 2001. Adenovirus Hexon Protein Enhances Nuclear Delivery and Increases Transgene Expression of Polyethylenimine / Plasmid DNA Vectors. , 4(5), pp.473–483.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., Bousso, P., Deist, F.L. & Fischer, A., 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science (New York, N.Y.)*, 288(5466), pp.669–72.

Chang, S.-W., Lee, G.-C. & Shaw, J.-F., 2006. Codon optimization of Candida rugosa lip1 gene for improving expression in Pichia pastoris and biochemical characterization of the purified recombinant LIP1 lipase. *Journal of agricultural and food chemistry*, 54(3), pp.815–22.

Cherry, J.M., Adler, C., Ball, C., Chervitz, S.A., Dwight, S.S., Hester, E.T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., Weng, S. & Botstein, D., 1998. SGD:

Saccharomyces Genome Database. *Nucleic acids research*, 26(1), pp.73–9.

Chesebro, B., Wehrly, K. & Maury, W., 1990. Differential expression in human and mouse cells of human immunodeficiency virus pseudotyped by murine retroviruses. *Journal of virology*, 64(9), pp.4553–7.

Coil, D.A. & Miller, A.D., 2004. Phosphatidylserine Is Not the Cell Surface Receptor for Vesicular Stomatitis Virus Phosphatidylserine Is Not the Cell Surface Receptor for Vesicular Stomatitis Virus. *Journal of Virology*, 78(20), pp.10920–10926.

Corrales-Garcia, L.L., Possani, L.D. & Corzo, G., 2011. Expression systems of human β-defensins: vectors, purification and biological activities. *Amino acids*, 40(1), pp.5–13.

Curran, K.A., Karim, A.S., Gupta, A. & Alper, H.S., 2013. Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic engineering*, 19, pp.88–97.

Demain, A.L. & Vaishnav, P., 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnology advances*, 27(3), pp.297–306.

Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M. & Thrasher, A.J., 2002. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Human gene therapy*, 13(7), pp.803–13.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R.E., Hill, C.M., Davis, C.B., Peiper, S.C., Schall, T.J., Littman, D.R. & Landau, N.R., 1996. Identification of a major co-receptor for

primary isolates of HIV-1. *Nature*, 381(6584), pp.661–6.

Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C. & Fothergill, L.A., 1982. Conservation of high efficiency promoter sequences in Saccharomyces cerevisiae. *Nucleic acids research*, 10(8), pp.2625–37.

Donahue, J.P., Vetter, M.L., Mukhtar, N.A. & D'Aquila, R.T., 2008. The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology*, 377(1), pp.49–53.

Dong, B., Nakai, H. & Xiao, W., 2010. Characterization of genome integrity for oversized recombinant AAV vector. *Molecular therapy : the journal of the American Society of Gene Therapy*, 18(1), pp.87–92.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D. & Naldini, L., 1998. A third-generation lentivirus vector with a conditional packaging system. *Journal of virology*, 72(11), pp.8463–71.

Egel-Mitani, M., Andersen, A., Diers, I., Hach, M., Thim, L., Hastrup, S. & Vad, K., 2000. Yield improvement of heterologous peptides expressed in yps1-disrupted Saccharomyces cerevisiae strains. *Enzyme and microbial technology*, 26(9–10), pp.671–677.

Engelhardt, J.F., Ye, X., Doranz, B. & Wilson, J.M., 1994. Ablation of E2A in recombinant adenoviruses improves transgene persistence and decreases inflammatory response in mouse liver. *Proceedings of the National Academy of Sciences of the United States of America*, 91(13), pp.6196–200.

Engelman, A., Mizuuchi, K. & Craigie, R., 1991. HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, 67(6), pp.1211–1221.

Escriou, V., Ciolina, C., Lacroix, F., Byk, G., Scherman, D. & Wils, P., 1998. Cationic lipid-mediated gene transfer: Effect of serum on cellular uptake and intracellular fate of lipopolyamine/DNA complexes. *Biochimica et Biophysica Acta - Biomembranes*, 1368(2), pp.276–288.

Eslahi, N.K., Muller, S., Nguyen, L., Wilson, E., Thull, N., Rolland, A. & Pericle, F., 2001. Fusogenic activity of vesicular stomatitis virus glycoprotein plasmid in tumors as an enhancer of IL-12 gene therapy. *Cancer gene therapy*, 8(1), pp.55–62.

Everett, R.D., 2000. ICP0, a regulator of herpes simplex virus during lytic and latent infection. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 22(8), pp.761–70.

Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M. & Danielsen, M., 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences of the United States of America*, 84(21), pp.7413–7417.

Ferro-Novick, S., Hansen, W., Schauer, I. & Schekman, R., 1984. Genes required for completion of import of proteins into the endoplasmic reticulum in yeast. *The Journal of cell biology*, 98(1), pp.44–53.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W. & Lührmann, R., 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*, 82(3), pp.475–83.

Fraley, R., Subramani, S., Berg, P. & Papahadjopoulos, D., 1980. Introduction of liposome-encapsulated SV40 DNA into cells. *Journal of Biological Chemistry*, 255(21), pp.10431–10435.

Gallay, P., Swingler, S., Song, J., Bushman, F. & Trono, D., 1995. HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. *Cell*, 83(4), pp.569–76.

Gambihler, S., Delius, M. & Ellwart, J.W., 1994. Permeabilization of the plasma membrane of L1210 mouse leukemia cells using lithotripter shock waves. *The Journal of membrane biology*, 141(3), pp.267–75.

Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodríguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino, M.L., Mattanovich, D. & Villaverde, A., 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial cell factories*, 7, p.11.

Ginsberg, H.S., Lundholm-Beauchamp, U., Horswood, R.L., Pernis, B., Wold, W.S., Chanock, R.M. & Prince, G. a, 1989. Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proceedings of the National Academy of Sciences of the United States of America*, 86(10), pp.3823–3827.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S.G., 1996. Life with 6000 genes. *Science (New York, N.Y.)*, 274(5287), pp.546, 563–7.

Gorziglia, M.I., Kadan, M.J., Yei, S., Lim, J., Lee, G.M., Luthra, R. & Trapnell, B.C., 1996. Elimination of both E1 and E2 from adenovirus vectors further improves prospects for in vivo human gene therapy. *Journal of virology*, 70(6), pp.4173–8.

Gustin, K.E., Lutz, P. & Imperiale, M.J., 1996. Interaction of the adenovirus L1 52/55-kilodalton protein with the IVa2 gene product during infection. *J Virol*, 70(9), pp.6463–6467.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B.H., Wintergerst, U., Velez, M.C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F.D., Fischer, A. & Cavazzana-Calvo, M., 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of clinical investigation*, 118(9), pp.3132–42.

Halbert, D.N., Cutt, J.R. & Shenk, T., 1985. Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *Journal of virology*, 56(1), pp.250–7.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J. & others, 2004. Transcriptional regulatory code of a eukaryotic genome. *Nature*, 431(7004), pp.99–104.

Hearing, P., Samulski, R.J., Wishart, W.L. & Shenk, T., 1987. Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome. *Journal of virology*, 61(8), pp.2555–8.

Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V. & Wintersberger, U., 2003. Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *The EMBO journal*, 22(9), pp.2274–83.

Heinzinger, N.K., Bukrinsky, M.I., Haggerty, S.A., Ragland, A.M., Kewalramani, V., Lee, M.A., Gendelman, H.E., Ratner, L., Stevenson, M. & Emerman, M., 1994. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proceedings of the*

National Academy of Sciences of the United States of America, 91(15), pp.7311–5.

Hematti, P., Hong, B.-K., Ferguson, C., Adler, R., Hanawa, H., Sellers, S., Holt, I.E., Eckfeldt, C.E., Sharma, Y., Schmidt, M., von Kalle, C., Persons, D. a, Billings, E.M., Verfaillie, C.M., Nienhuis, A.W., Wolfsberg, T.G., Dunbar, C.E. & Calmels, B., 2004. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS biology*, 2(12), p.e423.

Hickman, M.A., Malone, R.W., Lehmann-Bruinsma, K., Sih, T.R., Knoell, D., Szoka, F.C., Walzem, R., Carlson, D.M. & Powell, J.S., 1994. Gene expression following direct injection of DNA into liver. *Human gene therapy*, 5(12), pp.1477–83.

Hirano, M., Nakamura, S., Mitsunaga, F., Okada, M., Shimuzu, K. & Imamura, T., 2002. Transfer of maternally administered fusogenic liposome-DNA complexes into monkey fetuses in a pregnancy model. *Journal of Gene Medicine*, 4(5), pp.560–566.

Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, D. V., Ammerer, G. & Hall, B.D., 1981. Expression of a human gene for interferon in yeast. *Nature*, 293(5835), pp.717–722.

Hodges, P.E., McKee, A.H., Davis, B.P., Payne, W.E. & Garrels, J.I., 1999. The Yeast Proteome Database (YPD): a model for the organization and presentation of genome-wide functional data. *Nucleic acids research*, 27(1), pp.69–73.

Hoffmann, M., Wu, Y.-J., Gerber, M., Berger-Rentsch, M., Heimrich, B., Schwemmle, M. & Zimmer, G., 2010. Fusion-active glycoprotein G mediates the cytotoxicity of vesicular stomatitis virus M mutants lacking host shut-off activity. *The Journal of general virology*, 91(Pt 11), pp.2782–93.

Höglund, S., Ofverstedt, L.G., Nilsson, A., Lundquist, P., Gelderblom, H., Ozel, M. & Skoglund, U., 1992. Spatial visualization of the maturing HIV-1 core and its linkage to the envelope. *AIDS research and human retroviruses*, 8(1), pp.1–7.

Holland, M.J., Holland, J.P., Thill, G.P. & Jackson, K.A., 1981. The primary structures of two yeast enolase genes. Homology between the 5' noncoding flanking regions of yeast enolase and glyceraldehyde-3-phosphate dehydrogenase genes. *The Journal of biological chemistry*, 256(3), pp.1385–95.

Horsburgh, B.C., Hubinette, M.M., Qiang, D., MacDonald, M.L. & Tufaro, F., 1999. Allele replacement: an application that permits rapid manipulation of herpes simplex virus type 1 genomes. *Gene therapy*, 6(5), pp.922–30.

Hovland, P., Flick, J., Johnston, M. & Sclafani, R.A., 1989. Galactose as a gratuitous inducer of GAL gene expression in yeasts growing on glucose. *Gene*, 83(1), pp.57–64.

Hu, H., Serra, D. & Amalfitano, A., 1999. Persistence of an [E1-, polymerase-] adenovirus vector despite transduction of a neoantigen into immune-competent mice. *Human gene therapy*, 10(3), pp.355–64.

Huang, C.-J., Lowe, A.J. & Batt, C.A., 2010. Recombinant immunotherapeutics: current state and perspectives regarding the feasibility and market. *Applied microbiology and biotechnology*, 87(2), pp.401–10.

Huang, D., Gore, P.R. & Shusta, E. V, 2008. Increasing yeast secretion of heterologous proteins by regulating expression rates and post-secretory loss. *Biotechnology and bioengineering*, 101(6), pp.1264–75.

Huang, M.M. & Hearing, P., 1989. Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *Journal of virology*, 63(6),

pp.2605–15.

Huh, W.-K., Falvo, J. V, Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S. & O'Shea, E.K., 2003. Global analysis of protein localization in budding yeast. *Nature*, 425(6959), pp.686–91.

Hüser, D., Gogol-Döring, A., Lutter, T., Weger, S., Winter, K., Hammer, E.-M., Cathomen, T., Reinert, K. & Heilbronn, R., 2010. Integration preferences of wildtype AAV-2 for consensus rep-binding sites at numerous loci in the human genome. *PLoS pathogens*, 6(7), p.e1000985.

Imazu, S., Nakagawa, S., Nakanishi, T., Mizuguchi, H., Uemura, H., Yamada, O. & Mayumi, T., 2000. A novel nonviral vector based on vesicular stomatitis virus. *Journal of Controlled Release*, 68(2), pp.187–194.

Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N., 2009. Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nature protocols*, 4(1), pp.58–70.

Jahic, M., Gustavsson, M., Jansen, A.-K., Martinelle, M. & Enfors, S.-O., 2003. Analysis and control of proteolysis of a fusion protein in Pichia pastoris fed-batch processes. *Journal of biotechnology*, 102(1), pp.45–53.

Janowicz, Z.A., Melber, K., Merckelbach, A., Jacobs, E., Harford, N., Comberbach, M. & Hollenberg, C.P., 1991. Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, Hansenula polymorpha. *Yeast (Chichester, England)*, 7(5), pp.431–43.

Jewett, M.C., Hofmann, G. & Nielsen, J., 2006. Fungal metabolite analysis in genomics and phenomics. *Current opinion in biotechnology*, 17(2), pp.191–7.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. & Charpentier, E., 2012. A

Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096), pp.816–821.

Johnston, M. & Davis, R.W., 1984. Sequences that regulate the divergent GAL1-GAL10 promoter in Saccharomyces cerevisiae. *Molecular and cellular biology*, 4(8), pp.1440–8.

Johnston, S.A. & Hopper, J.E., 1982. Isolation of the yeast regulatory gene GAL4 and analysis of its dosage effects on the galactose/melibiose regulon. *Proceedings of the National Academy of Sciences of the United States of America*, 79(22), pp.6971–5.

Jones, N. & Shenk, T., 1979. An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proceedings of the National Academy of Sciences of the United States of America*, 76(8), pp.3665–3669.

Jörnvall, H., Akusjärvi, G., Aleström, P., von Bahr-Lindström, H., Pettersson, U., Appella, E., Fowler, A. V & Philipson, L., 1981. The adenovirus hexon protein. The primary structure of the polypeptide and its correlation with the hexon gene. *The Journal of biological chemistry*, 256(12), pp.6181–6.

Jowett, J.B., Planelles, V., Poon, B., Shah, N.P., Chen, M.L. & Chen, I.S., 1995. The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *Journal of virology*, 69(10), pp.6304–13.

Kafri, T., Blömer, U., Peterson, D.A., Gage, F.H. & Verma, I.M., 1997. Sustained expression of genes delivered directly into liver and muscle by lentiviral vectors. *Nature genetics*, 17(3), pp.314–7.

Karin, M., Najarian, R., Haslinger, A., Valenzuela, P., Welch, J. & Fogel, S., 1984. Primary structure and transcription of an amplified genetic locus: the CUP1 locus

of yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 81(2), pp.337–41.

Kato, S., Inoue, K., Kobayashi, K., Yasoshima, Y., Miyachi, S., Inoue, S., Hanawa, H., Shimada, T., Takada, M. & Kobayashi, K., 2007. Efficient gene transfer via retrograde transport in rodent and primate brains using a human immunodeficiency virus type 1-based vector pseudotyped with rabies virus glycoprotein. *Human gene therapy*, 18(11), pp.1141–51.

Katz, R. a, Terry, R.W. & Skalka, a M., 1986. A conserved cis-acting sequence in the 5' leader of avian sarcoma virus RNA is required for packaging. *Journal of virology*, 59(1), pp.163–7.

Kawabata, K., Takakura, Y. & Hashida, M., 1995. The fate of plasmid DNA after intravenous injection in mice: involvement of scavenger receptors in its hepatic uptake. *Pharmaceutical research*, 12(6), pp.825–30.

Kelly, T.J. & Lewis, A.M., 1973. Use of nondefective adenovirus-simian virus 40 hybrids for mapping the simian virus 40 genome. *Journal of virology*, 12(3), pp.643–52.

King, J.A., Dubielzig, R., Grimm, D. & Kleinschmidt, J.A., 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO journal*, 20(12), pp.3282–91.

King, W., Patel, M.D., Lobel, L.I., Goff, S.P. & Nguyen-Huu, M.C., 1985. Insertion mutagenesis of embryonal carcinoma cells by retroviruses. *Science (New York, N.Y.)*, 228(4699), pp.554–8.

Kochanek, S., Clemens, P.R., Mitani, K., Chen, H.H., Chan, S. & Caskey, C.T., 1996. A new adenoviral vector: Replacement of all viral coding sequences with 28 kb of DNA independently expressing both full-length dystrophin and beta-galactosidase.

Proceedings of the National Academy of Sciences of the United States of America, 93(12), pp.5731–5736.

Konvalinka, J., Litterst, M.A., Welker, R., Kottler, H., Rippmann, F., Heuser, A.M. & Kräusslich, H.G., 1995. An active-site mutation in the human immunodeficiency virus type 1 proteinase (PR) causes reduced PR activity and loss of PR-mediated cytotoxicity without apparent effect on virus maturation and infectivity. *Journal of virology*, 69(11), pp.7180–6.

Koressaar, T. & Remm, M., 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics (Oxford, England)*, 23(10), pp.1289–91.

Kotin, R.M., Siniscalco, M., Samulski, R.J., Zhu, X.D., Hunter, L., Laughlin, C.A., McLaughlin, S., Muzyczka, N., Rocchi, M. & Berns, K.I., 1990. Site-specific integration by adeno-associated virus. *Proceedings of the National Academy of Sciences of the United States of America*, 87(6), pp.2211–5.

Kotsopoulou, E., Kim, V.N., Kingsman, a J., Kingsman, S.M. & Mitrophanous, K. a, 2000. A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. *Journal of virology*, 74(10), pp.4839–52.

Lai, Y., Yue, Y. & Duan, D., 2010. Evidence for the Failure of Adeno-associated Virus Serotype 5 to Package a Viral Genome ≥8.2 kb. *Molecular Therapy*, 18(1), pp.75–79.

Lang, G.I. & Murray, A.W., 2008. Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. *Genetics*, 178(1), pp.67–82.

Lapadat-Tapolsky, M., De Rocquigny, H., Van Gent, D., Roques, B., Plasterk, R. & Darlix, J.L., 1993. Interactions between HIV-1 nucleocapsid protein and viral DNA

may have important functions in the viral life cycle. *Nucleic acids research*, 21(4), pp.831–9.

Laquerre, S., Argnani, R., Anderson, D.B., Zucchini, S., Manservigi, R. & Glorioso, J.C., 1998. Heparan sulfate proteoglycan binding by herpes simplex virus type 1 glycoproteins B and C, which differ in their contributions to virus attachment, penetration, and cell-to-cell spread. *Journal of virology*, 72(7), pp.6119–30.

Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O. & Davis, R.W., 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proceedings of the National Academy of Sciences*, 94(24), pp.13057–13062.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J.-B., Volkert, T.L., Fraenkel, E., Gifford, D.K. & Young, R.A., 2002. Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science (New York, N.Y.)*, 298(5594), pp.799–804.

Lewis, P., Hensel, M. & Emerman, M., 1991. Human immunodeficiency virus infection of cells arrested in the cell cycle. , 11(8), pp.3053–3058.

Li, S., Tseng, W.C., Stolz, D.B., Wu, S.P., Watkins, S.C. & Huang, L., 1999. Dynamic changes in the characteristics of cationic lipidic vectors after exposure to mouse serum: implications for intravenous lipofection. *Gene therapy*, 6(4), pp.585–94.

Li, Y., Drone, C., Sat, E. & Ghosh, H.P., 1993. Mutational analysis of the vesicular stomatitis virus glycoprotein G for membrane fusion domains. *Journal of virology*, 67(7), pp.4070–7.

Liu, C., Perilla, J.R., Ning, J., Lu, M., Hou, G., Ramalho, R., Himes, B.A., Zhao, G., Bedwell, G.J., Byeon, I.-J., Ahn, J., Gronenborn, A.M., Prevelige, P.E., Rousso, I., Aiken, C., Polenova, T., Schulten, K. & Zhang, P., 2016. Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nature Communications*, 7, p.10714.

van Lohuizen, M. & Berns, A., 1990. Tumurigenesis by slow-transforming retroviruses—an update. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1032(2–3), pp.213–235.

Lonez, C., Vandenbranden, M. & Ruysschaert, J.M., 2008. Cationic liposomal lipids: From gene carriers to cell signaling. *Progress in Lipid Research*, 47(5), pp.340–347.

Lungwitz, U., Breunig, M., Blunk, T. & Göpferich, a., 2005. Polyethylenimine-based non-viral gene delivery systems. *European Journal of Pharmaceutics and Biopharmaceutics*, 60(2), pp.247–266.

Luria, S.E. & Delbrück, M., 1943. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*, 28(6), pp.491–511.

Mann, R., Mulligan, R.C. & Baltimore, D., 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, 33(1), pp.153–159.

Marcellus, R.C., Lavoie, J.N., Boivin, D., Shore, G.C., Ketner, G. & Branton, P.E., 1998. The early region 4 orf4 protein of human adenovirus type 5 induces p53-independent cell death by apoptosis. *Journal of virology*, 72(9), pp.7144–53.

Marconi, P., Krisky, D., Oligino, T., Poliani, P.L., Ramakrishnan, R., Goins, W.F., Fink, D.J. & Glorioso, J.C., 1996. Replication-defective herpes simplex virus vectors for

gene transfer in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 93(21), pp.11319–20.

McAleer, W.J., Buynak, E.B., Maigetter, R.Z., Wampler, D.E., Miller, W.J. & Hilleman, M.R., 1984. Human hepatitis B vaccine from recombinant yeast. *Nature*, 307(5947), pp.178–80.

McCaffrey, A.P., Fawcett, P., Nakai, H., McCaffrey, R.L., Ehrhardt, A., Pham, T.-T.T., Pandey, K., Xu, H., Feuss, S., Storm, T.A. & Kay, M.A., 2008. The Host Response to Adenovirus, Helper-dependent Adenovirus, and Adeno-associated Virus in Mouse Liver. *Molecular Therapy*, 16(5), pp.931–941.

McConnell, M.J. & Imperiale, M.J., 2004. Biology of adenovirus and its use as a vector for gene therapy. *Human gene therapy*, 15(11), pp.1022–33.

Merchuk, J.C., Andrews, B.A. & Asenjo, J.A., 1998. Aqueous two-phase systems for protein separation: Studies on phase inversion. *Journal of Chromatography B: Biomedical Sciences and Applications*, 711(1), pp.285–293.

Meyhack, B., Bajwa, W., Rudolph, H. & Hinnen, A., 1982. Two yeast acid phosphatase structural genes are the result of a tandem duplication and show different degrees of homology in their promoter and coding sequences. *The EMBO journal*, 1(6), pp.675–80.

Miller, A.D. & Rosman, G.J., 1989. Improved Retroviral Vectors for Gene Transfer and Expression. *Biotechniques*, 7(9), pp.980–990.

Mingozzi, F. & High, K.A., 2011. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nature Reviews Genetics*, 12(5), pp.341–355.

Miyanohara, A., 2012. Preparation of vesicular stomatitis Virus-G (VSV-G) conjugate

and its use in gene transfer. *Cold Spring Harbor Protocols*, 7(4), pp.453–456.

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J., Bueren, J. & Baum, C., 2009. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, 17(11), pp.1919–28.

Morille, M., Passirani, C., Vonarbourg, A., Clavreul, A. & Benoit, J.-P., 2008. Progress in developing cationic vectors for non-viral systemic gene therapy against cancer. *Biomaterials*, 29(24–25), pp.3477–96.

Le Moullec, J.M., Akusjarvi, G., Stalhandske, P., Pettersson, U., Chambraud, B., Gilardi, P., Nasri, M. & Perricaudet, M., 1983. Polyadenylic acid addition sites in the adenovirus type 2 major late transcription unit. *J Virol*, 48(1), pp.127–134.

Mumper, R.J., Wang, J., L. Klakamp, S., Nitta, H., Anwer, K., Tagliaferri, F. & Rolland, A.P., 1998. Protective interactive noncondensing (PINC) polymers for enhanced plasmid distribution and expression in rat skeletal muscle. *Journal of Controlled Release*, 52(1–2), pp.191–203.

Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M. & Trono, D., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science (New York, N.Y.)*, 272(5259), pp.263–7.

Neumann, E., Schaefer-Ridder, M., Wang, Y. & Hofschneider, P.H., 1982. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *The EMBO journal*, 1(7), pp.841–5.

Nevins, J.R. & Winkler, J.J., 1980. Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc*

Natl Acad Sci U S A, 77(4), pp.1893–1897.

Nie, Z., Phenix, B.N., Lum, J.J., Alam, A., Lynch, D.H., Beckett, B., Krammer, P.H., Sekaly, R.P. & Badley, A.D., 2002. HIV-1 protease processes procaspase 8 to cause mitochondrial release of cytochrome c, caspase cleavage and nuclear fragmentation. *Cell death and differentiation*, 9(11), pp.1172–84.

Nienhuis, A.W., Dunbar, C.E. & Sorrentino, B.P., 2006. Genotoxicity of retroviral integration in hematopoietic cells. *Molecular therapy : the journal of the American Society of Gene Therapy*, 13(6), pp.1031–49.

Nowrouzi, A., Cheung, W.T., Li, T., Zhang, X., Arens, A., Paruzynski, A., Waddington, S.N., Osejindu, E., Reja, S., von Kalle, C., Wang, Y., Al-Allaf, F., Gregory, L., Themis, M., Holder, M., Dighe, N., Ruthe, A., Buckley, S.M., Bigger, B., Montini, E., Thrasher, A.J., Andrews, R., Roberts, T.P., Newbold, R.F., Coutelle, C., Schmidt, M. & Themis, M., 2013. The fetal mouse is a sensitive genotoxicity model that exposes lentiviral-associated mutagenesis resulting in liver oncogenesis. *Molecular therapy : the journal of the American Society of Gene Therapy*, 21(2), pp.324–37.

Nunes, F.A., Furth, E.E., Wilson, J.M. & Raper, S.E., 1999. Gene transfer into the liver of nonhuman primates with E1-deleted recombinant adenoviral vectors: safety of readministration. *Human gene therapy*, 10(15), pp.2515–26.

Okimoto, T., Friedmann, T. & Miyanohara, a, 2001. VSV-G envelope glycoprotein forms complexes with plasmid DNA and MLV retrovirus-like particles in cell-free conditions and enhances DNA transfection. *Molecular therapy : the journal of the American Society of Gene Therapy*, 4(3), pp.232–238.

Olins, D.E., Olins, a L. & Von Hippel, P.H., 1967. Model nucleoprotein complexes: studies on the interaction of cationic homopolypeptides with DNA. *Journal of*

molecular biology, 24(2), pp.157–176.

Ozuer, A., Wechuck, J.B., Goins, W.F., Wolfe, D., Glorioso, J.C. & Ataai, M.M., 2002. Effect of genetic background and culture conditions on the production of herpesvirus-based gene therapy vectors. *Biotechnology and bioengineering*, 77(6), pp.685–92.

Page, K.A., Landau, N.R. & Littman, D.R., 1990. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. *Journal of virology*, 64(11), pp.5270–6.

Palù, G., Parolin, C., Takeuchi, Y. & Pizzato, M., 2000. Progress with retroviral gene vectors. *Reviews in medical virology*, 10(3), pp.185–202.

Parekh, R., Forrester, K. & Wittrup, D., 1995. Multicopy overexpression of bovine pancreatic trypsin inhibitor saturates the protein folding and secretory capacity of Saccharomyces cerevisiae. *Protein expression and purification*, 6(4), pp.537–45.

Parekh, R.B. & Patel, T.P., 1992. Comparing the glycosylation patterns of recombinant glycoproteins. *Trends in biotechnology*, 10(8), pp.276–80.

Parissi, V., Caumont, A., de Soultrait, V.R., Desjobert, C., Calmels, C., Fournier, M., Gourgue, G., Bonneu, M., Tarrago-Litvak, L. & Litvak, S., 2003. The lethal phenotype observed after HIV-1 integrase expression in yeast cells is related to DNA repair and recombination events. *Gene*, 322, pp.157–168.

Pastan, I. & Willingham, M.C., 1983. Receptor-mediated endocytosis: coated pits, receptosomes and the Golgi. *Trends in Biochemical Sciences*, 8(7), pp.250–254.

Paxton, W., Connor, R.I. & Landau, N.R., 1993. Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *Journal of virology*, 67(12), pp.7229–37.

Perry, L.J. & McGeoch, D.J., 1988. The DNA sequences of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. *The Journal of general virology*, 69 (Pt 11, pp.2831–46.

Petranovic, D., Tyo, K., Vemuri, G.N. & Nielsen, J., 2010. Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. *FEMS yeast research*, 10(8), pp.1046–59.

Popov, S., Rexach, M., Zybarth, G., Reiling, N., Lee, M., Ratner, L., Lane, C.M., Moore, M.S. & Bukrinsky, M., 1998. Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. , 17(4), pp.909–917.

Pouton, C.W., Pouton, C.W., Seymour, L.W. & Seymour, L.W., 2001. Key issues in non-viral gene delivery. *Advanced drug delivery reviews*, 46(1–3), pp.187–203.

Price, V.L., Taylor, W.E., Clevenger, W., Worthington, M. & Young, E.T., 1990. Expression of heterologous proteins in Saccharomyces cerevisiae using the ADH2 promoter. *Methods in enzymology*, 185, pp.308–18.

Ranzani, M., Cesana, D., Bartholomae, C.C., Sanvito, F., Pala, M., Benedicenti, F., Gallina, P., Sergi, L.S., Merella, S., Bulfone, A., Doglioni, C., von Kalle, C., Kim, Y.J., Schmidt, M., Tonon, G., Naldini, L. & Montini, E., 2013. Lentiviral vector-based insertional mutagenesis identifies genes associated with liver cancer. *Nat Methods*, 10(2), pp.155–161.

Raper, S.E., Yudkoff, M., Chirmule, N., Gao, G.-P., Nunes, F., Haskal, Z.J., Furth, E.E., Propert, K.J., Robinson, M.B., Magosin, S., Simoes, H., Speicher, L., Hughes, J., Tazelaar, J., Wivel, N.A., Wilson, J.M. & Batshaw, M.L., 2002. A pilot study of in vivo liver-directed gene transfer with an adenoviral vector in partial ornithine transcarbamylase deficiency. *Human gene therapy*, 13(1), pp.163–75.

Robinson, A.S., Bockhaus, J.A., Voegler, A.C. & Wittrup, K.D., 1996. Reduction of BiP levels decreases heterologous protein secretion in Saccharomyces cerevisiae. *The Journal of biological chemistry*, 271(17), pp.10017–22.

Roe, T., Reynolds, T.C., Yu, G. & Brown, P.O., 1993. Integration of murine leukemia. , 12(5), pp.2099–2108.

Rolland, A. & Mumper, R., 1998. Plasmid delivery to muscle: Recent advances in polymer delivery systems. *Advanced drug delivery reviews*, 30(1–3), pp.151–172.

Rosenberg, S., Coit, D. & Tekamp-Olson, P., 1990. Glyceraldehyde-3-phosphate dehydrogenase-derived expression cassettes for constitutive synthesis of heterologous proteins. *Methods in enzymology*, 185, pp.341–51.

Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W.A. & Rosen, C.A., 1989. Structural and functional characterization of human immunodeficiency virus tat protein. *Journal of virology*, 63(1), pp.1–8.

Sakuragi, S., Goto, T., Sano, K. & Morikawa, Y., 2002. HIV type 1 Gag virus-like particle budding from spheroplasts of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), pp.7956–61.

Samulski, R.J., Chang, L.S. & Shenk, T., 1989. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *Journal of virology*, 63(9), pp.3822–8.

Sanber, K.S., Knight, S.B., Stephen, S.L., Bailey, R., Escors, D., Minshull, J., Santilli, G., Thrasher, A.J., Collins, M.K. & Takeuchi, Y., 2015. Construction of stable packaging cell lines for clinical lentiviral vector production. *Scientific Reports*, 5, p.9021.

Sandrin, V., Boson, B., Salmon, P., Gay, W., Nègre, D., Le Grand, R., Trono, D. & Cosset, F.-L., 2002. Lentiviral vectors pseudotyped with a modified RD114 envelope glycoprotein show increased stability in sera and augmented transduction of primary lymphocytes and CD34+ cells derived from human and nonhuman primates. *Blood*, 100(3), pp.823–32.

Sarramegna, V., Demange, P., Milon, A. & Talmont, F., 2002. Optimizing functional versus total expression of the human mu-opioid receptor in Pichia pastoris. *Protein expression and purification*, 24(2), pp.212–20.

Sauer, U., 2006. Metabolic networks in motion: 13C-based flux analysis. *Molecular systems biology*, 2, p.62.

Schiestl, R.H., Igarashi, S. & Hastings, P.J., 1988. Analysis of the mechanism for reversion of a disrupted gene. *Genetics*, 119(2), pp.237–47.

Schlegel, R., Dickson, R.B., Willingham, M.C. & Pastan, I.H., 1982. Amantadine and dansylcadaverine inhibit vesicular stomatitis virus uptake and receptor-mediated endocytosis of alpha 2-macroglobulin. *Proceedings of the National Academy of Sciences of the United States of America*, 79(7), pp.2291–2295.

Schlegel, R., Tralka, T.S., Willingham, M.C. & Pastan, I., 1983. Inhibition of VSV binding and infectivity by phosphatidylserine: is phosphatidylserine a VSV-binding site? *Cell*, 32(2), pp.639–46.

Schnell, M.A., Zhang, Y., Tazelaar, J., Gao, G.P., Yu, Q.C., Qian, R., Chen, S.J., Varnavski, A.N., LeClair, C., Raper, S.E. & Wilson, J.M., 2001. Activation of innate immunity in nonhuman primates following intraportal administration of adenoviral vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, 3(5 Pt 1), pp.708–22.

Schoggins, J.W., Gall, J.G.D. & Falck-Pedersen, E., 2003. Subgroup B and F fiber chimeras eliminate normal adenovirus type 5 vector transduction in vitro and in vivo. *Journal of virology*, 77(2), pp.1039–48.

Schwartz, O., Maréchal, V., Le Gall, S., Lemonnier, F. & Heard, J.M., 1996. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nature medicine*, 2(3), pp.338–42.

Schwartz, S., Felber, B.K. & Pavlakis, G.N., 1992. Distinct RNA sequences in the gag region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *Journal of virology*, 66(1), pp.150–9.

Senís, E., Fatouros, C., Große, S., Wiedtke, E., Niopek, D., Mueller, A.-K., Börner, K. & Grimm, D., 2014. CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. *Biotechnology Journal*, 9(11), pp.1402–1412.

Sharp, P.M., Tuohy, T.M.F. & Mosurski, K.R., 1986. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Research*, 14(13), pp.5125–5143.

Shimotohno, K. & Temin, H.M., 1981. Formation of infectious progeny virus after insertion of herpes simplex thymidine kinase gene into DNA of an avian retrovirus. *Cell*, 26(1 Pt 1), pp.67–77.

Shoji, J., Tanihara, Y., Uchiyama, T. & Kawai, A., 2004. Preparation of virosomes coated with the vesicular stomatitis virus glycoprotein as efficient gene transfer vehicles for animal cells. *Microbiology and immunology*, 48(3), pp.163–174.

Sikorski, R.S. & Hieter, P., 1989. A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces cerevisiae.

Genetics, 122(1), pp.19–27.

Simovic, D., Isner, J.M., Ropper, A.H., Pieczek, A. & Weinberg, D.H., 2001. Improvement in chronic ischemic neuropathy after intramuscular phVEGF165 gene transfer in patients with critical limb ischemia. *Archives of neurology*, 58(5), pp.761–8.

Sinha, S. & Grandgenett, D.P., 2005. Recombinant human immunodeficiency virus type 1 integrase exhibits a capacity for full-site integration in vitro that is comparable to that of purified preintegration complexes from virus-infected cells. *Journal of virology*, 79(13), pp.8208–16.

Smith, J.D., Richardson, N.E. & Robinson, A.S., 2005. Elevated expression temperature in a mesophilic host results in increased secretion of a hyperthermophilic enzyme and decreased cell stress. *Biochimica et biophysica acta*, 1752(1), pp.18–25.

Smith, K.T., Shepherd, A.J., Boyd, J.E. & Lees, G.M., 1996. Gene delivery systems for use in gene therapy: an overview of quality assurance and safety issues. *Gene therapy*, 3(3), pp.190–200.

Smith, R.H. & Kotin, R.M., 2000. An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. *Journal of virology*, 74(7), pp.3122–9.

Solbak, S.M.Ø., Reksten, T.R., Hahn, F., Wray, V., Henklein, P., Henklein, P., Halskau, Ø., Schubert, U. & Fossen, T., 2013. HIV-1 p6 - a structured to flexible multifunctional membrane-interacting protein. *Biochimica et biophysica acta*, 1828(2), pp.816–23.

Sonawane, N.D., Szoka, F.C. & Verkman, A.S., 2003. Chloride accumulation and

swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. *The Journal of biological chemistry*, 278(45), pp.44826–31.

Sonntag, F., Bleker, S., Leuchs, B., Fischer, R. & Kleinschmidt, J.A., 2006. Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *Journal of virology*, 80(22), pp.11040–54.

Srivastava, A., Lusby, E.W. & Berns, K.I., 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *Journal of virology*, 45(2), pp.555–64.

Stopak, K., de Noronha, C., Yonemoto, W. & Greene, W.C., 2003. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Molecular cell*, 12(3), pp.591–601.

Strebel, K., Daugherty, D., Clouse, K., Cohen, D., Folks, T. & Martin, M.A., 1987. The HIV “A” (sor) gene product is essential for virus infectivity. *Nature*, 328(6132), pp.728–30.

Sun, X., Yau, V.K., Briggs, B.J. & Whittaker, G.R., 2005. Role of clathrin-mediated endocytosis during vesicular stomatitis virus entry into host cells. *Virology*, 338(1), pp.53–60.

Suter, M., Lew, A.M., Grob, P., Adema, G.J., Ackermann, M., Shortman, K. & Fraefel, C., 1999. BAC-VAC, a novel generation of (DNA) vaccines: A bacterial artificial chromosome (BAC) containing a replication-competent, packaging-defective virus genome induces protective immunity against herpes simplex virus 1. *Proceedings of the National Academy of Sciences of the United States of America*, 96(22), pp.12697–702.

Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M. &

Zhang, F., 2014. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nature Biotechnology*, 33(1), pp.102–106.

Tandia, B.M., Vandenbranden, M., Wattiez, R., Lakhdar, Z., Ruysschaert, J.M. & Elouahabi, A., 2003. Identification of human plasma proteins that bind to cationic lipid/DNA complex and analysis of their effects on transfection efficiency: Implications for intravenous gene transfer. *Molecular Therapy*, 8(2), pp.264–273.

Themis, M., May, D., Coutelle, C. & Newbold, R.F., 2003. Mutational effects of retrovirus insertion on the genome of V79 cells by an attenuated retrovirus vector: implications for gene therapy. *Gene therapy*, 10(19), pp.1703–11.

Tiley, L.S., Madore, S.J., Malim, M.H. & Cullen, B.R., 1992. The VP16 transcription activation domain is functional when targeted to a promoter-proximal RNA sequence. *Genes & development*, 6(11), pp.2077–87.

Tokmakov, A.A., Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Fukami, Y. & Yokoyama, S., 2012. Multiple post-translational modifications affect heterologous protein synthesis. *The Journal of biological chemistry*, 287(32), pp.27106–16.

Tomo, N., Goto, T. & Morikawa, Y., 2013. Trans-packaging of human immunodeficiency virus type 1 genome into Gag virus-like particles in Saccharomyces cerevisiae. *Microbial cell factories*, 12(1), p.28.

Tøttrup, H. V & Carlsen, S., 1990. A process for the production of human proinsulin in Saccharomyces cerevisiae. *Biotechnology and bioengineering*, 35(4), pp.339–48.

Trezise, A.E.O., 2002. In vivo DNA electrotransfer. *DNA and cell biology*, 21(12), pp.869–77.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B.,

Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S. & Rothberg, J.M., 2000. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature*, 403(6770), pp.623–7.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. & Rozen, S.G., 2012. Primer3--new capabilities and interfaces. *Nucleic acids research*, 40(15), p.e115.

Usaite, R., Wohlschlegel, J., Venable, J.D., Park, S.K., Nielsen, J., Olsson, L. & Yates Iii, J.R., 2008. Characterization of global yeast quantitative proteome data generated from the wild-type and glucose repression saccharomyces cerevisiae strains: the comparison of two quantitative methods. *Journal of proteome research*, 7(1), pp.266–75.

Vile, R.G. & Hart, I.R., 1993. In Vitro and in Vivo Targeting of Gene Expression to Melanoma Cells Advances in Brief In Vitro and in Vivo Targeting of Gene Expression to Melanoma Cells. *Cancer Research*, 53, pp.962–967.

Villas-Bôas, S.G., Moxley, J.F., Akesson, M., Stephanopoulos, G. & Nielsen, J., 2005. High-throughput metabolic state analysis: the missing link in integrated functional genomics of yeasts. *The Biochemical journal*, 388(Pt 2), pp.669–77.

Webster, A. & Kemp, G., 1993. The active adenovirus protease is the intact L3 23K protein. *Journal of General Virology*, 74(7), pp.1415–1420.

Wehland, J., Willingham, M.C., Gallo, M.G. & Pastan, I., 1982. The morphologic pathway of exocytosis of the vesicular stomatitis virus G protein in cultured fibroblasts. *Cell*, 28(4), pp.831–841.

Wei, C.M., Gibson, M., Spear, P.G. & Scolnick, E.M., 1981. Construction and isolation of a transmissible retrovirus containing the src gene of Harvey murine sarcoma

virus and the thymidine kinase gene of herpes simplex virus type 1. *Journal of Virology*, 39(3), pp.935–944.

Westerman, K.A., Ao, Z., Cohen, E.A. & Leboulch, P., 2007. Design of a trans protease lentiviral packaging system that produces high titer virus. *Retrovirology*, 4, p.96.

Whitehead, K. a, Langer, R. & Anderson, D.G., 2009. Knocking down barriers: advances in siRNA delivery. *Nature reviews. Drug discovery*, 8(2), pp.129–38.

Wides, R.J., Challberg, M.D., Rawlins, D.R. & Kelly, T.J., 1987. Adenovirus origin of DNA replication: sequence requirements for replication in vitro. *Molecular and cellular biology*, 7(2), pp.864–74.

Wiethoff, C.M., Smith, J.G., Koe, G.S. & Middaugh, C.R., 2001. The potential role of proteoglycans in cationic lipid-mediated gene delivery. Studies of the interaction of cationic lipid-DNA complexes with model glycosaminoglycans. *The Journal of biological chemistry*, 276(35), pp.32806–13.

Wolff, J.A., Williams, P., Acsadi, G., Jiao, S., Jani, A. & Chong, W., 1991. Conditions affecting direct gene transfer into rodent muscle in vivo. *BioTechniques*, 11(4), pp.474–85.

Wu, G.Y. & Wu, C.H., 1987. Receptor-mediated in vitro gene transformation by a soluble DNA carrier system. *Journal of Biological Chemistry*, 262(10), pp.4429–4432.

Wu, K., Guimet, D. & Hearing, P., 2013. The adenovirus L4-33K protein regulates both late gene expression patterns and viral DNA packaging. *Journal of virology*, 87(12), pp.6739–47.

Xu, K., Ma, H., McCown, T.J., Verma, I.M. & Kafri, T., 2001. Generation of a stable cell line producing high-titer self-inactivating lentiviral vectors. *Molecular therapy : the*

journal of the American Society of Gene Therapy, 3(1), pp.97–104.

Yamaguchi, T., Kawabata, K., Koizumi, N., Sakurai, F., Nakashima, K., Sakurai, H., Sasaki, T., Okada, N., Yamanishi, K. & Mizuguchi, H., 2007. Role of MyD88 and TLR9 in the Innate Immune Response Elicited by Serotype 5 Adenoviral Vectors. *Human Gene Therapy*, 18(8), pp.753–762.

Yang, J.P. & Huang, L., 1997. Overcoming the inhibitory effect of serum on lipofection by increasing the charge ratio of cationic liposome to DNA. *Gene therapy*, 4(9), pp.950–960.

Yee, J.K., Miyanohara, A., LaPorte, P., Bouic, K., Burns, J.C. & Friedmann, T., 1994. A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 91(20), pp.9564–8.

Yin, H., Kanasty, R.L., Eltoukhy, A.A., Vegas, A.J., Dorkin, J.R. & Anderson, D.G., 2014. Non-viral vectors for gene-based therapy. *Nature reviews. Genetics*, 15(8), pp.541–555.

Young, S.M., McCarty, D.M., Degtyareva, N. & Samulski, R.J., 2000. Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination. *Journal of virology*, 74(9), pp.3953–66.

Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A. & Chen, I.S., 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell*, 61(2), pp.213–22.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. & Charneau, P., 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*, 101(2), pp.173–85.

Zhang, W.H., Hockley, D.J., Nermut, M. V, Morikawa, Y. & Jones, I.M., 1996. Gag-Gag interactions in the C-terminal domain of human immunodeficiency virus type 1 p24 capsid antigen are essential for Gag particle assembly. *The Journal of general virology*, 77 (Pt 4)(4), pp.743–51.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R.A., Gerstein, M. & Snyder, M., 2001. Global analysis of protein activities using proteome chips. *Science (New York, N.Y.)*, 293(5537), pp.2101–5.

Zhu, J., Huang, X. & Yang, Y., 2009. The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice. *Journal of Clinical Investigation*, 119(8), pp.2388–2398.

Zhu, Z.H., Chen, S.S. & Huang, A.S., 1990. Phenotypic mixing between human immunodeficiency virus and vesicular stomatitis virus or herpes simplex virus. *Journal of acquired immune deficiency syndromes*, 3(3), pp.215–9.

Zubieta, C., Schoehn, G., Chroboczek, J. & Cusack, S., 2005. The structure of the human adenovirus 2 penton. *Molecular Cell*, 17(1), pp.121–135.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, a, Quiroz, D., Naldini, L. & Trono, D., 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of virology*, 72(12), pp.9873–80.

Zufferey, R., Nagy, D. & Mandel, R., 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nature …*.