An Investigation to Optimise Viral Vector Production and to Generate an *in vitro* Model for Genotoxicity Assessment.

A thesis submitted for the degree of Doctor of

Philosophy by

Saqlain Mehboob Amanullah Suleman

Department of Life Sciences, Brunel University London

Abstract

Viral vectors, including adenoviruses and lentiviruses, have been extensively developed for gene therapy. Gene addition using these delivery mechanisms have proven a successful method in correction of genetic diseases. The rapid expansion in clinical and non-clinical trials in this field indicate a present need to optimise viral vector production to achieve pure, high titre vectors for downstream applications. The current protocols for adenoviral vector purification remain time consuming and present with low recovery yields. We explore aqueous two-phase systems as a cheaper, low maintenance alternative for adenoviral and lentiviral vector purification. These data show high recovery of adenoviral vectors using a PEG 600/ $(NH_4)_2SO_4$ system. However, lentiviral vectors are not successfully partitioned using these systems. HIV-1 based lentiviruses have been shown to contribute towards genotoxicity *in vivo*. While models exist, there is need to provide a sensitive, humanised platform of genotoxicity. Here, we present data using induced pluripotent stem cells and their hepatocyte and cardiomyocyte derivatives. Using HIV-1 vectors carrying transgenes flanked by either the native or self-inactivating LTR configurations, we see clear differences in integration site profiling between these vectors in these stem cell and their hepatocyte derivatives. Examination of integration sites over time in replicating stem cells also provides a rapid model for clonal outgrowth. These data suggest the usefulness of using both a naïve and terminally differentiated cell types to sensitively understand lentiviral mediated genotoxicity further. These data also provide evidence for transcription factor binding site tethering of lentiviruses to common sites in the host genome. Comparative analysis to differentiated cardiomyocytes suggest the importance of examining vector mediated adverse events further to understand tissue specific side effects. This work provides for optimised vector production and a

model of lentiviral vector genotoxicity which can be used to improve the safety profile of these vectors before clinical applications.

Content's page

Declaration

I, Saqlain Mehboob Amanullah Suleman, hereby declare that all work presented in this thesis is original and has been prepared by myself, unless otherwise stated. This work has not, and will not, be submitted in whole or in part to another University for the award of any other Degree. The research presented in this thesis was conducted pursuant to the University Code of Research Ethics and all compulsory training required has been completed appropriately.

Acknowledgments

The opportunity to complete a Doctorate has been a fantastic experience and I would like to acknowledge the people who have helped me in this journey.

My thanks go to my primary supervisor, Dr Michael Themis, for his invaluable support in creating this project, as well as his time and assistance over the course of my doctoral research.

I would also like to thank the other researchers in my lab group and University staff for their help in conducting these experiments. I thank Dr Jonathan Bowden for lending his experience to training a new researcher in the lab and aiding me in analysis of *in vivo* mouse and patient data. My thanks also go to Dr Svetlana Ignatova, Dr Peter Hewitson and Dr Jonathan Huddleston for lending their expertise in aqueous two phase systems and high-performance liquid chromatography. I am also grateful to Professor Simon Waddington and Dr Rajvinder Karda, of the EGA Institute for Women's Health at UCL, for their help with performing adenoviral *in vivo* gene transfer experiments. I acknowledge Dr Terry Roberts for his premise to test various reagents in the production of lentiviruses and Dr Yasuhiro Takeuchi, of the Division of Infection and Immunity at UCL, for kindly providing the pHV plasmid and training in various methodologies. I thank Dr Arshad Khan and Dr Annette Payne for their training in bioinformatic analyses. I am grateful to the Centre for Genomic Research, University of Liverpool and Genewerk GmbH, Heidelberg, for providing RNA and DNA sequencing platforms for my research.

I also express my gratitude for funding various projects from N4Pharma, Ataxia UK and GlaxoSmithKline.

Finally, I would like to thank my family and friends for supporting me in my pursuit to getting my doctorate, in particular Tahera Datoo and Murtaza Datoo. Without you, I would not have been able to complete this feat.

Dedication

This thesis is dedicated to three shining starts in my life. You are the inspirations behind my every move and forever in my heart.

Mehboob Amanullah Suleman

1966-2005

Musarrat Yusuf Suleman

1972-2005

Sibtain Mehboob Suleman

1995-2005

Chapter I
Literature Review

Gene therapy is a promising field with 3180 clinical trials registered as ongoing in 2021 (Wiley 2021). The basic principle of gene therapy is correction of the underlying genetic mutation contributing towards disease progression through gene addition. Addition of recombinant nucleic acids to replace mutated sequences in the host genome fulfils the central dogma of DNA to correct genetic disease (Crick 1970, Mammen, Ramakrishnan et al. 2007). The expansion of this field is highlighted by the 11 gene therapy products which are currently clinically approved worldwide, with the majority of these products using a viral vector system (Wiley 2021). This method has been used in the correction of multiple genetic diseases, including X- linked SCID, cardiomyopathies and retinal disorders (Thornhill, Schambach et al. 2008, Salabarria, Nair et al. 2020, Hassall, Barnard et al. 2017). Indeed, 66% of clinical trials ongoing in 2020 are reported to use some form of viral vector for gene delivery (Wiley 2021).

1.1. Principals of Gene Therapy

While the development of gene therapy products is relatively recent, the premise of gene therapy can be traced back to the early 20th century. Griffith's experiment demonstrating viral pneumococcal properties of non-virulent pneumococcal Type I after combining Type I and II strains together seminally demonstrated the transformation principal (Griffith 1928, Dawson, Sia 1931). This was later discovered to be due to DNA transfer between cells (Avery, Macleod et al. 1944, Alloway 1933, Alloway 1932). This transfer of genetic material was further discovered to also occur through bacterial conjugation or transduction (Tatum, Lederberg 1947, Zinder, Lederberg 1952). Importantly from a gene therapy perspective, the transfer of genetic material was shown to allow transfer of genetic traits to rescue cells and that this transfer is heritable

(Szybalska, Szybalski 1962). This suggested the principals of successful gene therapy, whereby addition of a gene back into cells of patients may allow correction of genetic diseases which is permanent in the host. The effectiveness of using viruses was first discussed with an aim towards removing the pathological effects of viruses manipulated for therapeutic gene delivery, an important safety control procedure (Tatum 1966). This was shown to be successful using the tobacco mosaic and Shope papilloma viruses, albeit with a limited degree of success (Rogers, S., Pfuderer 1968, Rogers, S., Lowenthal et al. 1973, Terheggen, Lowenthal et al. 1975). However, the principal of genetic transfer using viral vectors was shown. The first clinical trial of gene therapy in humans was recorded in the latter half of the 20th century, targeting tumour infiltrating leukocytes against gliomas and ADA-SCID, showing successful gene transfer using a retrovirus (RV) vector and long-term expression after 4 years (Rosenberg, Aebersold et al. 1990, Bordignon, Notarangelo et al. 1995, Blaese, Culver et al. 1995). Similarly, *in vivo* expression of reporter genes mediated through adenovirus (Ad) vectors showed successful gene transfer in patients, though with limited success (Puumalainen, Vapalahti et al. 1998). This feat demonstrates the efficacy of therapeutic research, with gene delivery proceeding from conceptual design to clinical testing in under 70 years.

The advantage of using gene therapy as an alternative to vaccines is for permanent correction of genetic diseases (Smith, J., Lipsitch et al. 2011). While episomal viral vectors exist (i.e. Ad), these are more useful in oncolytic gene therapy (Wold, Toth 2013). This platform is not useful for permanent correction of genetic diseases which can be achieved through integrative viral vectors (i.e. RV, LV). However, this introduces the risk of insertional mutagenesis. Thus, alterations to viral vectors must be made to improve transduction efficiency and prevent adverse side effects.

1.2. Non-Viral and Viral Vectors for Gene Therapy

Multiple viral and non-viral vectors have been used for gene therapy. These have various advantages but also pitfalls. While initial use of non-viral vectors to mediate genetic transfer proved unsuccessful due to low efficiency rates, improved protocols for transfection have been developed. These are promising due to their low manufacturing cost, high production capacity, broad expression, and safety profile (Ramamoorth, Narvekar 2015). Naked plasmid DNA delivery has been optimised using various sequence modifications and transfection procedures to improve gene transfer (Mitsui, Nishikawa et al. 2009, Nafissi, Sum et al. 2014). Nanoparticle complexation with pDNA has also been shown to successfully deliver genes to target cells, with improved expression outputs (Kobelt, Schleef et al. 2013, Kulkarni, Myhre et al. 2017). While delivery mechanisms have been improved, the issue remains that pDNA is commonly produced in prokaryotic cells. While this leads to increased yields of plasmid DNA, common bacterial sequences present can be recognised by human host cell machinery and trigger degradation of the nucleic acids. These sequences also contribute towards heterochromatic silencing of transgene expression independent of CpG island configuration (Chen, Z. Y., Riu et al. 2008). Little progress has been made in replacing pDNA manufactured in bacterial cells. DoggyBone™ DNA has been developed as linear DNA capable for production to high yields in eukaryotic cells (Scott, V. L., Patel et al. 2015). This product has also been shown to successfully produce LV vectors for downstream use in gene therapy (Karda, Counsell et al. 2019). While progress has been made in moving away from prokaryotic production, the majority of plasmid DNA is harvested from bacterial cells.

Viral vectors have been well studied and developed for use in gene therapy. This relies on the biology and life cycle of virions to effectively deliver a therapeutic payload to cells. These vectors include Ad, LV, and adeno-

associated viruses (AAV). Each virus has different advantages for use in a gene therapy platform.

1.3. Adenoviral Vectors- Life Cycle and Genome

Ad are known to cause the common cold and have therefore been well studied, so much so that this vector represents 17.5% of viral vectors used in clinical trials (Hilleman, Werner 1954, Rowe, Huebner et al. 1953, Ginsberg, Pereira et al. 1966, Wiley 2021). Wild type Ad are class I, hexagonal, non-enveloped viruses that contain a 37 kb dsDNA genome (Figure 1.1). The virion presents as 70-100 nm wide, with over 50 human adenovirus serotypes present (Reddy, Natchiar et al. 2010, Epstein, Holt 1960). The triangular surface of the adenovirus is composed of hexons, trimers of polypeptide I (Fields, Knipe et al. 2007). Fibres extend from each of the vertexes of the capsid which are composed of triple β fold of polypeptide II. The fibres associate with the capsid of the virus with a penton base (Pereira, Valentine et al. 1968). The extension of the fibres itself doubles the diameter of the virus generating larger molecules for host infection (Reddy, Natchiar et al. 2010, Mautner, Pereira 1971, van Raaij, Mitraki et al. 1999).The dsDNA and protein core are surrounded by a capsid composed of multiple copies of three major (II, III, IV) and five minor (IIIa, Iva2, VI, VIII, IX) proteins (Pereira, Valentine et al. 1968). The biology of the Ad makes it useful in targeting cells through interactions which can be manipulated for use in gene therapy.

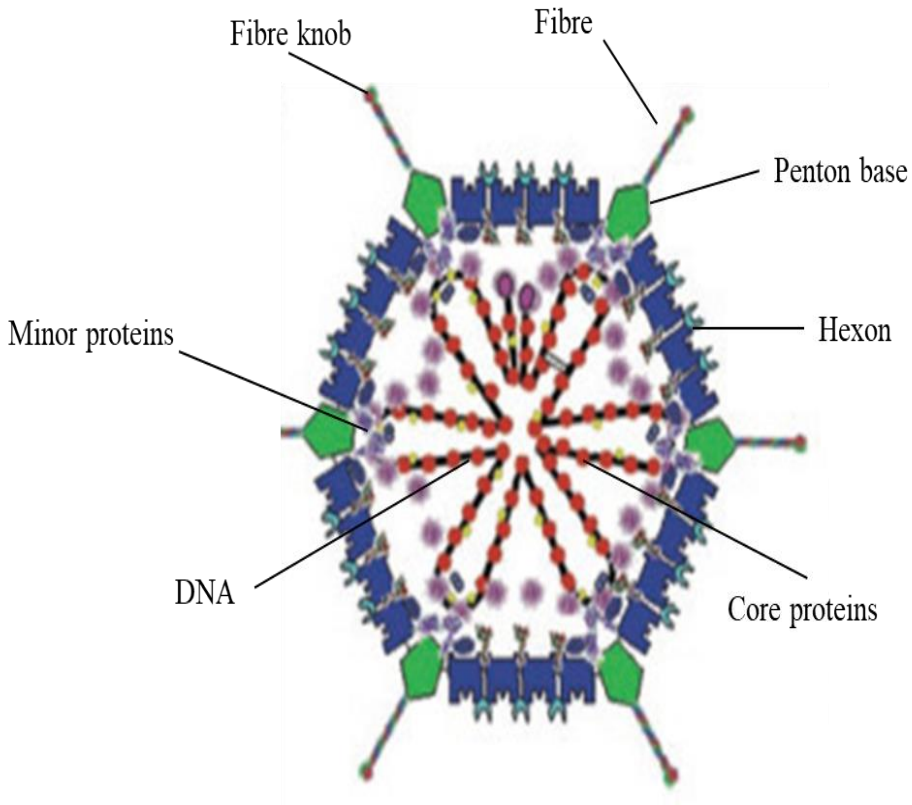


Figure 1.1- Schematic of Adenovirus biology. The hexagonal shape of the virion is shown to be comprised of hexon proteins, with a penton base and fibre (including fibre knob) extending from the surface. The genome packaged within the virus is surrounded by a capsid (comprised of minor proteins) and complexed with major proteins to package the DNA efficiently within the virion (Kennedy, Parks 2009)

The lifecycle of the Ads also makes these virions a useful vector for therapeutic transgene delivery. The Ad infects cells through interactions of the fibre with receptors on the cell surface membrane, including the coxsackie virus and adenovirus (CAR). Other serotypes also bind to CD46 and heparin sulphate glycans (HSG) receptors using conserved amino acids in the AB region of fibres (Gaggar, Shayakhmetov et al. 2003, Roelvink, Mi Lee et al. 1999, Bergelson, Cunningham et al. 1997, Smith, T. A., Idamakanti et al. 2003). The tight binding of the Ad fibre to CAR receptors outcompete traditional binding to promote Ad cell entry (Freimuth, Springer et al. 1999). Hexons also bind to C4 proteins triggering the complement cascade to remove virus infection (Xu, Z., Tian et al. 2008). The virus is carried into cells through receptor mediated

endocytosis through interactions of five RGD motifs on the fibre knob and αvβ3/ αvβ5 integrins on the cell surface. This process is also triggered through interactions of side loops present on fibres binding to the D1 domain of CAR receptors specifically (Bewley, Springer et al. 1999, Roelvink, Lizonova et al. 1998, Freimuth, Springer et al. 1999).

Adenovirus fibres are subsequently released from the capsid, commencing degradation of the penton bases and capsid by the endosome (Greber, Willetts et al. 1993). The release of protein VI from the capsid trigger lysis of the endosome and Ad escape (Leopold, Ferris et al. 1998, Maier, Galan et al. 2010). The virus is shuttled through the cell using dynamin, attaching to microtubules to migrate towards the nucleus (Greber, Willetts et al. 1993). The adenovirus binds to nuclear factors to disassociate from the microtubule and binds to N-terminal CAN/NUP214 cytoplasmic filament on nuclear pores to mediate nuclear entry via histone H1 proteins (Bailey, Crystal et al. 2003, Strunze, Engelke et al. 2011, Trotman, Mosberger et al. 2001, Leopold, Kreitzer et al. 2000). The Ad genome enters the nucleus through binding to Klc1/2 chain on kinensin-1, which itself is bound to NUP358 and NUP214. Kinensin-1 dissociates nucleoporins through disrupting NUP214, increasing the permeability of the nuclear envelope, and enhancing entry of the viral genome still further (Strunze, Engelke et al. 2011). These principals of Ad entry and transcription allow manipulation for therapeutic transgene delivery for gene therapy.

The genome of the Ad is highly organised into a number of early and late genes (Figure 1.2). The 30-36 kb genome is flanked by ~100 bp 5' and 3' ITRs, containing the origin of replication and separately a *cis* packaging signal (Freyer, Katoh et al. 1984).

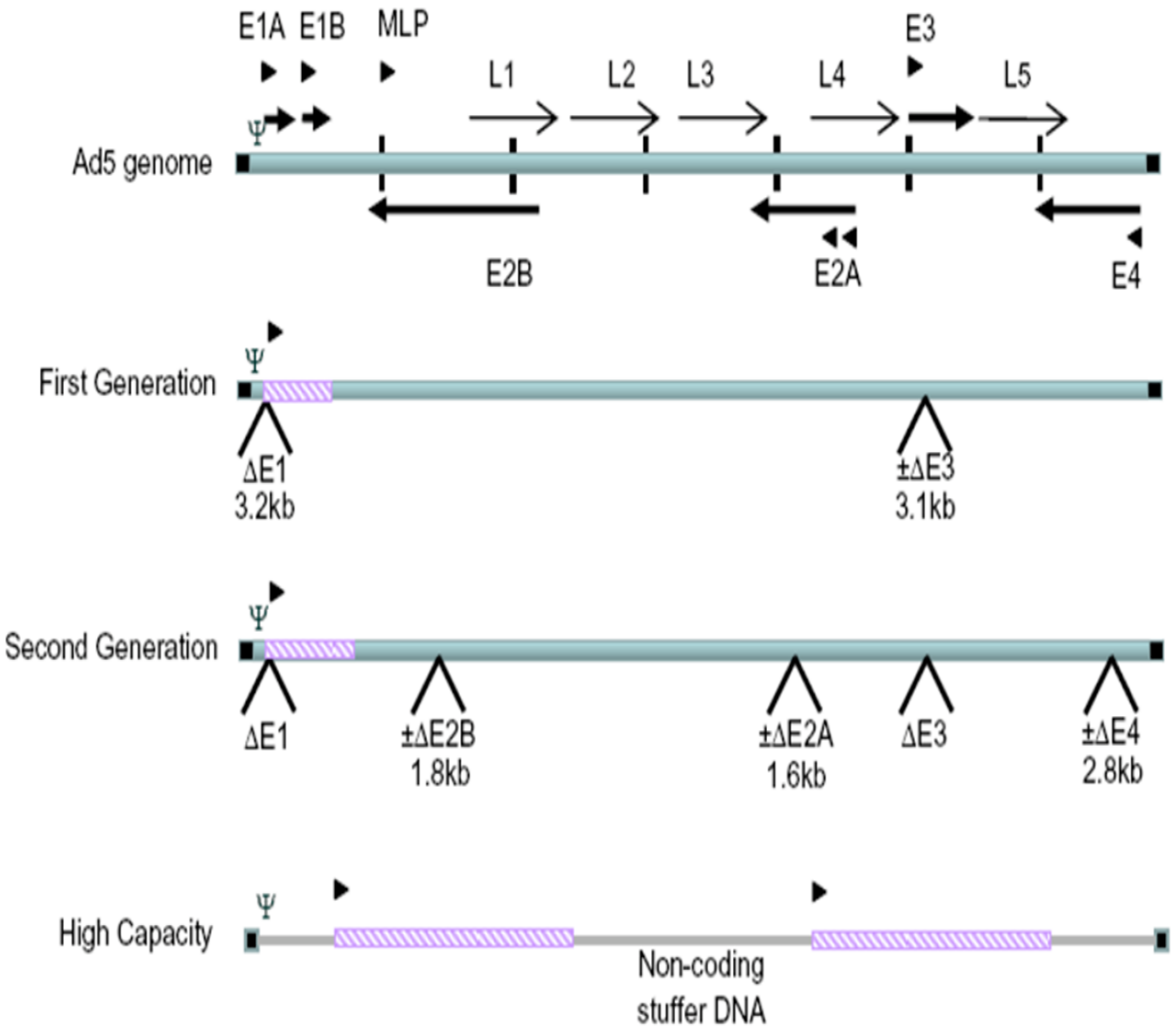


Figure 1.2- Adenovirus genome. The wild type AV genome (Ad5 genome) is shown to be comprised of Early and Late genes involved in viral replication and pathogenesis. First, second and high-capacity generations of these viruses contain deleted regions of specific early genes. First generation vectors contain E1 and E3 deletions, second generation contain additional deletions in E2A, E3 and E4. High-capacity vectors contain no viral genes but the transgene is flanked by ITRs 5' and 3'. These genes have been sequentially deleted in subsequent generations of AV vectors to increase safety and packaging capacity of vectors (Kennedy, Parks 2009)

The six early genes (*E1A*, *E1B*, *E2A*, *E2B*, *E3* and *E4*) are expressed sequentially to allow viral DNA transcription. *E1A* codes for 13S and 12S proteins which are used to activate early transcriptional units and to trigger mitosis of the host cell (Freyer, Katoh et al. 1984, Feldman, Nevins 1983, van Ormondt, Maat et al. 1980). 13S E1A is the only protein involved in active transcription of early genes and thus is independent of 12S E1A (Montell, Courtois et al. 1984). 12S E1A is used to activate six transcriptional units, containing conserved region 3. 13S associates with transcription factors (including ATF2) to localise E1A to promoters in host cells allowing transcription of the viral nucleic acids (Feldman, Nevins 1983, Montell, Courtois et al. 1984, Liu, F., Green 1990). Squelching assays indicate factors, such as SUR2, bind to conserved region 3 to allow transcription of E1A (Boyer, T. G., Martin et al. 1999). E1A is also involved in apoptosis mediated through activation of p53 as well as repression of Rb to promote cellular proliferation (Teodoro, Shore et al. 1995, Lowe, Ruley 1993, Nielsch, Fognani et al. 1991, Zamanian, La Thangue 1992). While this may prove contradictory to each other, apoptosis is inhibited by wild type viral proteins. E1B encodes for 19 kDa and 53 kDa proteins which are both involved in viral transformation. 19 kDa E1B is vital to viral transformation (Bernards, de Leeuw et al. 1986). Both these proteins are involved in blocking host mRNA transport allowing for viral transcript transport, translation and to protect the genome against nucleases (Freyer, Katoh et al. 1984). 53 kDa E1B is specifically involved in promoting cell cycle transition to S phase and preventing apoptosis by inhibiting p53 and Blc2 (Sassone, Margulets et al. 2015, Freyer, Katoh et al. 1984, Wang, J., Gao et al. 2015). E2 proteins are involved in viral DNA replication. *E2* produces a 72 kDa DNA binding, 80 kDa precursor terminal and 140kDa DNA polymerase proteins. 72 kDa E2 is involved in DNA replication through strand displacement, repression of *E4* and assembly of the virion (Nicolas, Sarnow et al. 1983, Babich, Nevins 1981, Nevins, Winkler 1980). 80 kDa E2 binds to the

first nucleotide in the Ad genome, acting as a primer for DNA replication which is cleaved during maturation (Smart, Stillman 1982). The functional properties of the *E3* gene allow the Ad to evade the immune system. The E3 19K glycoprotein has been shown to be vital to Ad survival (Kapoor, Wold et al. 1981). This glycoprotein forms a CR1 complex with other glycoproteins, such as RIDα, which associate with HLA proteins to aid evasion of infected cells by the host immune system (Frietze, Campos et al. 2012, Robinson, Rajaiya et al. 2011, Sester, Koebernick et al. 2010). 6.7K E3 inhibits apoptosis by inhibiting TNF-induced arachidonic acid release and stabilising Ca^{2+} homeostasis (Frietze, Campos et al. 2012, Moise, Grant et al. 2002). TNFR1 degradation is prevented by 14.5K E3 and a cell membrane complex is formed by 10.4K, 14.5K and 6.7K E3 to downregulate tumour necrosis factor-related apoptosis induced ligand (TRAIL). This downregulates TRAIL death receptors to prevent apoptosis of infected cells (Klingseisen, Ehrenschwender et al. 2012, Benedict, Norris et al. 2001). Thus, the Ad employs multiple pathways to promote cell survival which is useful for propagation of therapeutic transgenes. E4 proteins are also involved in DNA replication and promoting late gene expression by inhibiting host protein synthesis. E4ORF1 proteins activate epidermal growth factor receptor (*EGFR*), insulin-like growth factor receptor 1 (*IGF1R*) and PI3K to express Myc to promote cell division (Kong, Kumar et al. 2015).

Late genes are activated subsequent to transcription of early genes by the major late promoter (MLP), itself activated by E113S (Nevins 1981, Nevins, Wilson 1981). Low concentrations of L1-L5 are therefore produced during early phase; they are constitutively enhanced by L4 22K (Backstrom, Kaufmann et al. 2010). Late gene expression is also activated by *IAV2* which transcribes DEF-B. This protein binds to DEF-A, allowing the formation of new virions (Lutz, Kedinger 1996). Multiple mRNA transcripts are coded for by late genes owing to alternative splicing (Logan, Shenk 1984, Furuichi 1977). However, it has been

found that MLP does not work alone in inducing late gene transcription as the concentration of CStF is also vital (Mann, K. P., Weiss et al. 1993). Late genes code for structural proteins and are involved in inhibiting apoptosis, with L4 22K inhibiting early gene expression (Wu, K., Orozco et al. 2012). The wild type genome itself associates with proteins V, VII and Mu. Protein VII attaches to DNA to protect and condense the nucleic acids (Karen, Hearing 2011, Brown, Westphal et al. 1975). Therefore, the early and late genes are sequentially transcribed to allow for DNA replication and production of progeny virions. Replication occurs with the viral genome associating with various proteins including NF1 transcription factors (Hay, R. T. 1985). Understanding the function of the genome of Ad allows targeting of specific genes for deletion in vectors for use in gene therapy.

Due to the non-integrative nature of these viruses, Ads are commonly used in oncolytic gene therapy (Sharon, Kamen 2017). Advexin and ONYX- 015 both target and lyse p53 deficient cells, a hallmark of cancerous tissue. This has been shown to effectively reduce the size of tumours after administration (Shen, Kitzes et al. 2001, Nemunaitis, Ganly et al. 2000, Bischoff, Kirn et al. 1996).

1.4. Alternative Approaches to Host Responses to Adenovirus Delivery

While the immune response to adenovirus vectors is an issue faced in gene therapy, advanced generations of these vectors have been manipulated to elicit a lower response. This is poignantly highlighted by the death of Jessie Gelsinger. Pattern recognition receptors adhere to pathogen associated recognition patterns on the virion. Complement proteins also adhere to infected cells marking them for cell lysis (Hartman, Kiang et al. 2007, Jiang, Wang et al. 2004). The release of cytokines and chemokines from Ad infection also triggers an inflammatory response (Lieber, He et al. 1997). Adaptive immunity forms another barrier to

efficient Ad vector delivery. The humoral response occurs following antigen presenting cells phagocytosing Ad. Viral peptides presenting MHC class II molecules are detected by CD4+ lymphocytes, which are also complementary to adenoviral hexons (Serangeli, Bicanic et al. 2010). B lymphocyte activation and antibody production, such as IgM against human Ad5 primes the immune system against further Ad infection (Yang, Y., Greenough et al. 1996, Qiu, Xu et al. 2015). Antibodies primarily target the fibre knobs and hypervariable regions of hexon proteins (Pichla-Gollon, Drinker et al. 2007, Gahery-Segard, Farace et al. 1998). The cellular immune response removes infected cells by binding to MHC class I receptors (Yang, Y., Su et al. 1996). The prior immunity against Ads poses an issue for vector delivery to sensitised individuals.

Methods have been developed to reduce host immune responses to vector delivery. Antibodies are serotype specific due to the heterogeneity of the hypervariable region. Therefore, alternative species serotypes (i.e. canine Ad2) may be used in oncolytic gene therapy (Keriel, Rene et al. 2006, Sumida, Truitt et al. 2005). Chimpanzee Ad serotypes have also been employed as gene delivery mechanisms, including in vaccine production (Folegatti, Ewer et al. 2020, Dicks, Spencer et al. 2012). However, caution is required when using other species as an immune reaction this may also be elicited in humans. This is highlighted by antibodies against chimpanzee Ad identified in humans (Xiang, Li et al. 2006). Genomic deletions in successive generations of Ad vectors also reduce the host immune response to vectors. Moreover, coating Ad in polyethylene glycol reduces detection by the immune system and increases the half-life of vectors due to decreased cytokine and antibody production (Hiwasa, Nagaya et al. 2012, Croyle, Le et al. 2005, Eto, Yoshioka et al. 2008). Subsequent generations of Ad vectors have been generated to reduce immunological response and increase size of insert capacity (Figure 1.2). First

generation Ad vectors contain a partial deletion in *E1* and *E3* genes (Kratzer, Kreppel 2017). Second and third generation Ads contain further deletions to reduce host immune responses and increase the size of expression cassettes. Deletions within early genes also improve the general safety profile of vectors through production of replication deficient viruses (Yang, Y., Nunes et al. 1994). Second generation adenoviruses contain further deletions in E2 and E4 which reduce detection of infected cells and increase the size of expression cassettes (Amalfitano, Hauser et al. 1998). Finally, helper-dependent adenoviruses are extremely useful in gene therapy as these vectors contain the transgene containing the packaging signal and flanked by ITRs at the 5' and 3' end. As these vectors do not generate viral proteins, these must be provided in *trans* to form fully functional but replication deficient virions (Parks, Chen et al. 1996). These deletions increase insert capacity, with the virus effectively delivering 105% of the wild-type genome to cells (Bett, Prevec et al. 1993). These vectors lyse infected cells after replication and thus are helpful in destroying cancer cells. However, it is important to avoid deletions of E3 19 kDa and 14.7 kDa proteins. These proteins reduce MHC class I migration to cell membranes and inhibit arachidonic acid, inhibiting detection of infected cells and apoptosis, respectively. Therefore, these viral vectors should be developed further for clinical treatment.

Ads are commonly used in gene therapy clinical trials, presenting as one of the most common viral vectors used in the field (Wiley 2021). The Merck STEP trial evaluated the use of Ad5 vaccines carrying *gag*, *pol* and *nef* for immunization against human immunodeficiency virus (HIV) infection. While the trial was unsuccessful, with 24 vaccinated individuals subsequently infected with HIV, this was due to the vaccine preventing viral replication rather than preventing infection (Buchbinder, Mehrotra et al. 2008). Experiments have shown that Ads can effectively deliver 105% of the wild-type genome to cells

(30-38KB) (Bett, Prevec et al. 1993). Conditionally replicating Ads, which contain partial genomic deletions, are particularly useful in oncolytic gene therapy. These vectors target tumour cells by expressing suicide genes or specific receptors. Moreover, a tumour derived promoter can be used to drive E1A expression (Rodriguez, Schuur et al. 1997). Oncolytic vectors such as ONYX-015, which contain E1B deletions, selectively replicate within and lyse host cancer cells (Figure 1.3). E1B 53kDa inhibits p53 and, therefore, E1B deficient Ads selectively replicate and lyse p53 deficient cells (Rodriguez, Schuur et al. 1997, Bischoff, Kirn et al. 1996). Oncolytic Ads containing both *E1A* and *E1B* mutations effectively target tumour cells (i.e. AxdAdB-3) (Fukuda, Abei et al. 2003). Gendicine™ was approved for treatment of head- and squamous cell carcinoma, which replaced *E1* with p53 cDNA (Peng, Wu et al. 2006). Therefore, Ads can be successfully used in oncolytic gene therapy application.

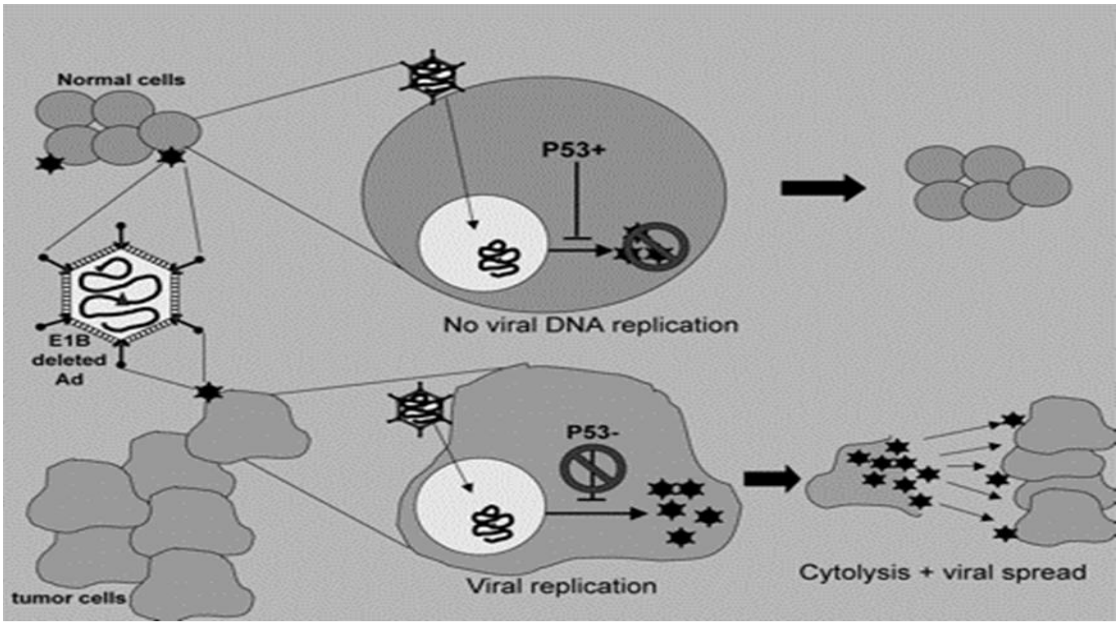


Figure 1.3- Oncolytic Adenovirus methodology. Schematic representation of oncolytic AV (i.e. ONYX-015) methodology. Cancer cells charecterised by depletion in P53 are selectively lysed by oncolytic AV which contain deletions in *E1B*. This promotes replication of the virion and lysis of host cells. Prescence of P53 inhibits replication of AV so as to not lyse non-cancerous cells (Chu, Post et al. 2004)

1.5. Production Adenoviral Vectors for Gene Therapy

Multiple viral vectors, including Ad and LV based vectors, have been used in clinical and non-clinical gene therapy studies. The primary safety feature of these vectors is their replication deficient nature. These viral vectors are produced to carry therapeutic transgenes and effectively transduce cells but do not contain the necessary genes to replicate and generate progeny virions i*n vivo*. This is required to reduce immunogenicity reactions by the host to ensure efficient gene delivery by the vector.

The production of Ad vectors requires three essential processes. Amplification of vectors occurs in a complementary cell line; purification of vectors removes unwanted substances; and quantification determines functional concentrations (Armendariz-Borunda, Bastidas-Ramirez et al. 2011). Replication-deficient Ad vectors cannot replicate outside these producer cells allowing for effective

oncolytic gene therapy. Thus, various complementary cell lines grow these vectors by providing replication proteins in *trans*.

HEK293 cells are useful for Ad production due to high transfection rates (50-80%) using calcium phosphate or chloroquine (Kwak, Yang et al. 2015, Gavrilescu, Van Etten 2007). The HEK293 cell line was created through transfection with Ad5 DNA. HEK293 cells contain 11% of 5' hAd5 DNA, including E1 genes (Graham, Smiley et al. 1977). This cell line highly expresses E1 proteins allowing effective growth of Ads (10,000/cell) compared to other cell lines (e.g. HC11). Expression of SV40T antigen allows high viral DNA replication and transfection (Lauren 2011, Patel, Tikoo 2006, Sullivan, Pipas 2002, Hurko, McKee et al. 1986). Accordingly, these suggest major advantages to using HEK293T cells for Ads vector production.

The process of growing Ad vectors has been developed using bioprocessors such as iCELLis 500, although the initial cell input (500m2) poses a challenge in producing large quantities of Ads for industrial use (Lesch, Heikkila et al. 2015). Ad cloning kits (i.e. AdEasy™XL) also provide a successful method of growing viral vectors. Furthermore, advances have been made in growing HEK293 cells in suspension for Ad vector production to high levels (Moreira, Silva et al. 2020, Ao, Gao et al. 2020). Growing cells as spheroids in suspension cultures is advantageous as it increases surface volume ratio to increase space for cells to grow. These cultures require lower maintenance than traditional monolayer culture of cells and occupy less space in incubators. However, this emerging technology requires further optimisation and validation to become readily accessible to general laboratories.

1.6. Purification of Adenovirus Vectors

Purification of Ad vectors involves downstream processing (DSP). DSP removes cellular debris, excess chemicals and maintains vector integrity by removing free nucleic acids (i.e. using Benzoase) (Vicente, Mota et al. 2011, Altaras, Aunins et al. 2005). A common method of vector purification is separation using a density gradient. Ultracentrifugation allows the purification of vectors based on density using a CsCl gradient and pressure separation on anisotropic membranes (0.5-1000 kDa) (Tkacik, Michaels 1991, Rose, Hoggan et al. 1966). CsCl is effective in removing contaminants and can be scaled up for bulk processing (up to 2kl) (Farid 2007, Boychyn, Yim et al. 2001). However, this process is not ideal due to many factors. Ultracentrifugation is time-consuming (~1.5 hours), limits the initial viral vector input and CsCl is toxic so must undergo dialysis before end product usage (Benyesh-Melnick, Probstmeyer et al. 1966). Therefore, other DSP methods need to be employed.

An effective method of DSP is aqueous two-phase systems (ATPS). This uses two immiscible water-soluble polymers and salts to purify products based on density, hydrophobic interactions, and interfacial tension (Figure 1.4) (Albertsson, Andersson et al. 1982).

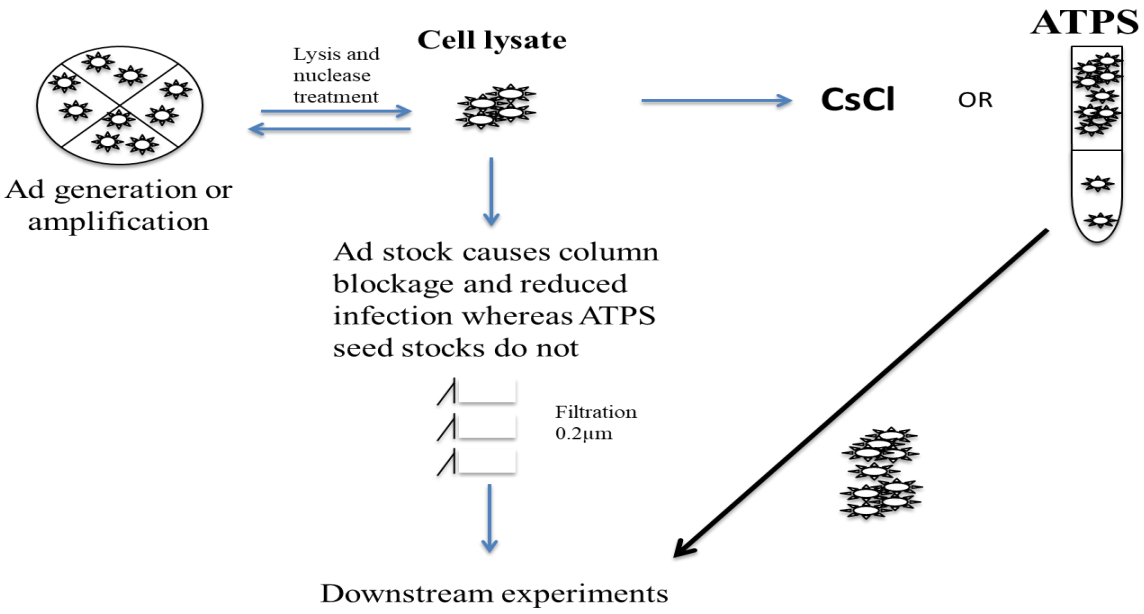


Figure 1.4- AV purification using ATPSs. AV are generated through cell transfection or amplified and lysed to generate cell lysates containing AV particles. Cell lysates can be clarified through filtration for product characterisation and downstream applications. Alternatively, cell lysates containing AV can be purified using CsCl density gradients or ATPS formulations for direct product characterisation or use in downstream experiments. Adapted from unpublished data currently under review (Suleman et al, 2021)

Commonly, polyethyl glycol (PEG) is combined with a kosmotropic salt such as phosphate (Guo, P., El-Gohary et al. 2012). ATPS is a gentle system (70-90% water), which produces high product yield and demonstrates a 10-fold increase in yields compared to chromatography (Oelmeier, Dismer et al. 2011, Negrete, Ling et al. 2007, Braas, Walker et al. 2000). As such, ATPS is an effective method of vector purification but research should be carried out in scaling ATPS for bulk vector purification.

Many factors need to be taken into consideration for effective partitioning of Ad vectors using ATPS. These include temperature, pH and hydrophobicity of the system to ensure maximum recovery yields of the vector (Baskir, Hatton et al. 1989, Tan, C. H., Show et al. 2015). Surface hydrophobicity is suggested to be directly proportional to the molecular weight of PEG (Oelmeier, Dismer et al. 2012). Considering these factors, DSP is generally dependent on cell density,

stability and viability as many purification methods isolate viruses based on density (Altaras, Aunins et al. 2005).

Previous studies have demonstrated the effectiveness of ATPS in purification of vectors and biomolecules. These systems are formed by commonly formed by dissolving a polymer and a salt to form two immiscible phases. The physical properties of these systems have been shown to be effective at purifying macromolecules. This system has also proven effective at separating adeno-associated virus (AAV) vectors for gene therapy using PEG8000- sodium chloride formulation (Guo, El-Gohary et al. 2012). Ad vectors have also previously been shown to recover highly infectious particles after separation using a PEG300- phosphate system (Negrete, Ling et al. 2007). These systems are commonly used in protein and macromolecule purification. PEG600 has been shown to purify 100% of a 6.1 kb plasmid DNA vector (Trindade, Diogo et al. 2005). PEG1000- potassium phosphate has been used to purify *Bacillus subtilis* NS99 alkaline protease and ovalbumin (Wongmongkol, Prichanont 2006, Zhi, Deng et al. 2005). PEG400- ammonium sulphate allows enzymatic reactions to occur while purifying the products of these reactions (Zhu, Wei et al. 2001). Thus, further development of ATPS purification of Ad vectors is required to streamline the purification procedure of these vectors further.

Various phase systems have been used to successfully purify biomolecules and vectors to high yields. ATPS has been shown effective for AV and AAV purification though no such application has been applied to LV, indicating a need to further develop these methodologies.

1.7. Retroviral and Lentiviral Vectors in Gene Therapy and Life Cycle

The issue of short-term gene expression as seen in Ad vectors is overcome by alternative viral vectors, including RV and HIV-1 based LV. The wild-type

function of these virus to integrate the transgene into host gDNA ensures permanent transfer of a large therapeutic payload (7.5 kb) to the host (Sharon, Kamen 2017). Simple RV such as γ- RV or Murine Leukemia Virus (MLV) contain common characteristics with LV, containing two copies of ssRNA encoding *gag*, *pol* and *env* genes (Escors, Breckpot et al. 2012, Warnock, Daigre et al. 2011).

Complex RV, or LV vectors, are extremely well studied and are a major vector used in gene therapy. LV, deriving from the Latin for slow, were first discovered after studies of sheep flocks in Iceland that had developed a slow, progressive pathology leading to pneumoencephalopathy (Sigurdsson, Grimsson et al. 1952). The primary difference between RV and LV is the presence of accessory genes in LV (Warnock, Daigre et al. 2011). A wide variety of species-specific LV have been manipulated for gene therapy, including human (HIV), simian (SIV), feline (FIV) and equine infectious anaemia virus (EIAV). These viruses are extremely useful for the delivery of therapeutic transgenes due to their natural biology. Gene therapy studies, conducted with the RV Maloney Murine Leukemia Virus (MoMLV), indicated the efficacy of gene transfer in patients (Mann, R., Mulligan et al. 1983). However, these viruses do not infect non- dividing cells. Therefore, RVs provided the basis for LV production. LV have a range of benefits above RV, including triggering reduced immune responses in humans as these viruses (SIV, EIAV) are not natural hosts (Durand, Cimarelli 2011).

HIV-1 has been well studied after the autoimmune deficiency disease (AIDS) epidemic in the 1980s. As a result, viral vectors based on HIV-1 are commonly used in gene therapy.

There is a common structure to HIV-1 based LV (Figure 1.5). These are class IV enveloped viruses, containing two copies of positive sense RNA. The genome is surrounded by a capsid containing approximately 2000 copies of p24

proteins and tightly condensed with p7 proteins. The capsid also contains essential enzymes used in the regulation and life cycle of lentiviruses including reverse transcriptase, proteases, ribonuclease, and integrase. The integrity of the capsid is maintained as it is surrounded by a matrix of p17 proteins. A phospholipid bilayer surrounding the capsid contains glycoproteins extending from the surface of the virus, for receptor mediated endocytosis (Durand, Cimarelli 2011). The envelope receptors allow the virus to specifically bind to and deliver its payload to cells. This can be engineered to present specific receptors for specific cell infection (Duverge, Negroni 2020).

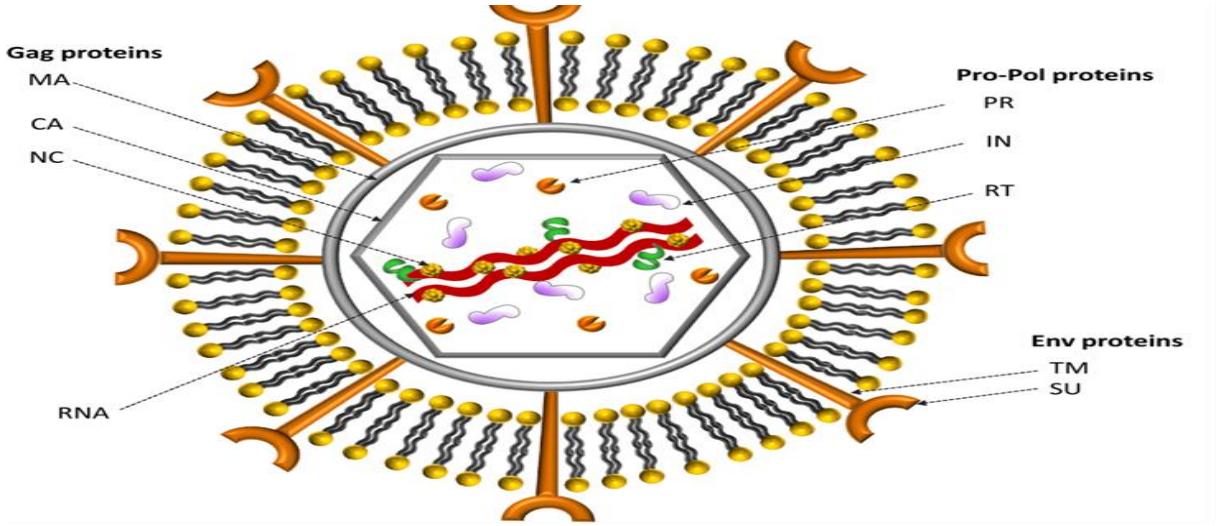


Figure 1.5- Schematic representation of HIV-1. Biology of HIV-1 is shown to contain two copies of ssRNA with wild type proteins for integration (gag-matrix (MA), capsid (CA) nucleocapsid (NC) and pro-pol- protease (PR), integrase (IN) and reverse transcriptase (RT)) surrounded by a capsid. The capsid is surrounded by a phospholipid bilayer and receptor (env) proteins (transmembrane, TM and surface, SU) to promote infection of cells (Rodrigues, Alves et al. 2011)

The biology of lentiviruses is further elucidated by its life cycle. Once the virus enters a host, the envelope receptor allows fusion to specific cells. The virus is then carried into the cell, typically through clathrin coated receptor mediated

endocytosis and broken down in the endosome due to the acidic environment. The capsid is then released into the cytoplasm of cells and degraded, allowing the release of RNA. Reverse transcription of the viral RNA is catalysed by the enzyme present in the virion into complementary deoxyribonucleic acid (cDNA) through forming a pre-initiation complex with RNA by complexing to integrase, matrix and nucleocapsid proteins. While the pre-initiation complex forms in the cytoplasm of cells, cDNA must be present in the nucleus of cells to allow integration to occur. This complex migrates into the nucleus through interactions with nucleoporins and hijacking nuclear localisation sequences of proteins (Fassati, Goff 1999, Lee, M. S., Craigie 1998, Farnet, Bushman 1997, Bukrinsky, Sharova et al. 1993, Farnet, Haseltine 1991). Here, LV have a distinct advantage over other viral vectors. Due to the ability of LV to cross nuclear pores, they are able to efficiently transduce quiescent and dividing cells. Alternatively, RV must wait for cell division and the break-down of the nuclear membrane before integration can take place (Durand, Cimarelli 2011).

cDNA conversion of the viral RNA genome is mediated by the reverse transcriptase enzyme through multiple priming of the viral nucleic acids (Baltimore 1970, Temin, Mizutani 1992). While the reverse transcriptase activity is active, the RNAase H activity of this enzyme degrades the negative sense viral cDNA to form short fragments of strong stop copy cDNA. These fragments are transferred to the 3’ end of the viral DNA to act as a primer for synthesis of the negative-sense viral DNA. This process is necessary to restore the function of the 3’ LTR region. The positive strand is synthesised through priming of the strand by polypurine tracts which are RNAase H resistant (Miller, Wang et al. 1995, Hughes 2015). The use of multiple central polypurine tracts allows the formation of DNA flaps due to multiple priming of the positive strand, which may protect the DNA from degradation and aid crossing the nuclear membrane (Arhel, Souquere-Besse et al. 2006, Arhel, Souquere-Besse

et al. 2007). Integration of the genome is mediated by the integrase enzyme. Integrase, a 288-amino acid protein in HIV-1, contains three functional domains: N- terminal domain containing a HHCC zinc binding motif, a catalytic core domain containing the magnesium binding D-D-35-E motif active site and the C- terminal domain (Park, Yun et al. 2019). While there are many models of viral integration, including the cell cycle and tethering protein models, the general consensus is that integration occurs within three distinct phases: terminal chain reaction, strand transfer reaction and DNA repair (Craigie, Bushman 2012). The process begins with a dinucleotide removed at the 3' end of DNA and the integrase molecule breaking the DNA 5 base pairs apart. This break is repaired using the using the 3' end of viral DNA. Host DNA repair mechanism are required to repair the five-nucleotide gap which removes two 5' nucleotides from the end of viral DNA. This results in the duplication of 5 bp flanking the integrated provirus. The integrase molecule allows the 3' OH group to attach to host nucleotides by integration within the opposite sense strand (Ciuffi 2008). The importance of this enzyme was highlighted through mutagenesis studies, where the genome is converted to cDNA but fails to integrate within the genome (Murphy, De Los Santos et al. 1993). The integrase association with cDNA forms the pre integration complex (PIC) (Bowerman, Brown et al. 1989, Ellison, Abrams et al. 1990, Farnet, Haseltine 1990). These PICs contain a variety of host cellular proteins, including NUP153 (Bukrinsky, Sharova et al. 1993, Farnet, Haseltine 1991, Lee, K., Ambrose et al. 2010, Li, L., Olvera et al. 2001).

1.8. HIV-1 based Lentivirus Integration

Integration of HIV-1 into the host genome occurs in a semi random manner. Various studies have suggested that HIV-1 preferentially integrates within

active genes (Ciuffi, Llano et al. 2005). Mapping of viral integration sites have shown a preference for LV to integrate within the outer region of the nucleus, containing open chromatin and actively transcribed genes. This suggests an evolutionary mechanism to target genes that are actively transcribed to ensure prorogation within the host (Jordan, Defechereux et al. 2001). Studies have also shown favourable integration within distorted major groove sites of DNA, correlating with chromatin looped out to the outer edges of the nucleus (Pruss, Bushman et al. 1994, Pruss, Reeves et al. 1994). Integrase and Gag proteins are still a major influence on the site of viral integration. HIV-1 virions containing murine leukaemia virus (MLV) integrase have shown an alternative insertion profile comparable to MLV, indicating the importance of these proteins for virus integration (Cohn, Silva et al. 2015, Ciuffi 2008). PIC recognition of components or features of the host cell chromatin is specific (Bushman, F., Lewinski et al. 2005, Bushman, F. D. 2003, Barr, Ciuffi et al. 2006, Lewinski, Yamashita et al. 2006) and clear differences have been demonstrated between promoter regions chosen by RV in contrast to LV, that favour insertion into the active transcription unit of genes (Schroder, Shinn et al. 2002, Barr, Ciuffi et al. 2006, Barr, Leipzig et al. 2005, Lewinski, Yamashita et al. 2006). This difference has been attributed to cellular proteins that interact with the viral integrase. Tethering of the HIV LV to the host genome is generally thought to be mediated by the mammalian PSIP1/LEDGF/p75 protein (Cherepanov, Maertens et al. 2003, Vandegraaff, Devroe et al. 2006). Cells depleted of PSIP1/LEDGF/p75 have significantly reduced HIV integration. However, integration sites (IS) are still favoured in transcription units, suggesting other factors may influence target site selection (Ciuffi, Llano et al. 2005). While various LVs integrate within the host genome, HIV has been shown as the effective at transducing cells and allowing long term gene expression ideal for gene therapy.

While less is known about MLV tethering to the host genome, it has been shown that this RV preferentially integrate near specific transcription factor binding sites (TFBS). For both LV and RV, the U3 region of the virus long terminal repeat (LTR) appears important for host transcription factor binding, which is integral to the synthesis of viral proteins of the native virus (Felice, Cattoglio et al. 2009). This is not exclusive to these viruses as it has also been found that yeast Ty3 retrotransposable elements, depend on specific host transcription factors TFIIIB and TFIIIC for Ty3 integration to the transcription initiation sites of RNA polymerase III transcribed genes (Kirchner, Connolly et al. 1995). It is plausible, therefore that HIV LV IS selection and possibly PIC tethering may also involve recognition of host TFBSs that mirror the predicted TFBS (pTFBS) in the HIV LTR region. Approximately 27.7% of the human genome has been found to associate with TFBS regions (Chen, H., Li, Liu, Zheng, Wang, Bo, and Shu 2015). Directing IS selection through TFBS present in the host may be important for viral replication by inserting into actively transcribed host genes.

Multiple transcription factors have been found to bind to the HIV genome. These include tumour necrosis factor alpha (TNF-α), which has been shown to activate HIV-1 transcription in chronically infected T-cells via binding of the transcription factor NF-κβ to specific sites in the U3 region (Duh, Maury et al. 1989). In addition, binding sites for c-myb (Dasgupta, Saikumar et al. 1990) and AP1 (Canonne-Hergaux, Aunis et al. 1995) transcription factors can be found in LTR regions and support viral transcription. These transcription factors are known to be involved in the HIV lifecycle including viral transcription, latency, and infection of non-activated T- cells (Tacheny, Michel et al. 2012, Chao, Walker et al. 2003, Ma, C., Dong et al. 2013, Hohne, Businger et al. 2016, Venkatachari, Zerbato et al. 2015).

Once the viral genome has undergone replication, it is packaged in capsid proteins and buds off cells, carrying the phospholipid bilayer from host cells, coated in receptor glycoproteins. Knowledge of the wild type LV life cycle has allowed manipulation of viruses to further benefit individual patients.

1.9. Genome of HIV-1 based Lentiviruses

During the lifecycle of the LV, the genome is replicated ensuring further viral progeny are produced. The genome of LVs demonstrates the capacity of these viruses to be used as efficient vectors in gene therapy (Figure 1.6).

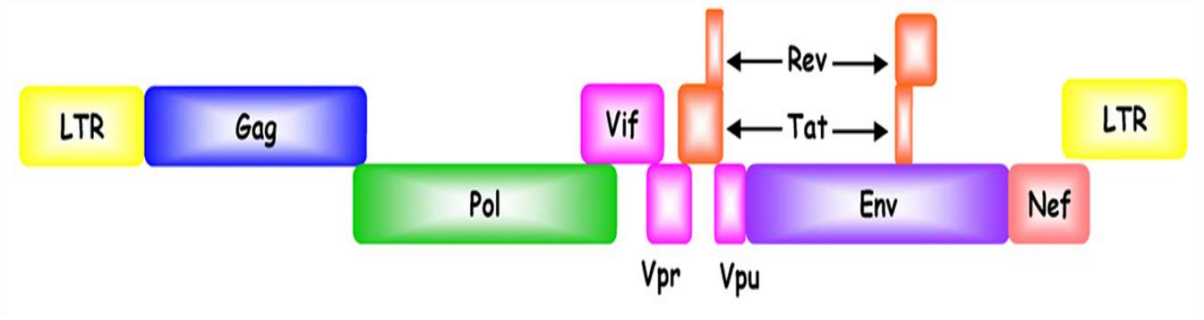


Figure 1.6- Wild type HIV genome. Schematic representation of the HIV-1 genome. These include *gag*, *pol*, accessory (*Vif*, *Vpr*, *Vpu*, *Tat*, *Rev*, *Nef*) and *Env* genes. Deletions in host genes through successive generations of vectors has indicated genes vital to HIV-1 production *in vitro* for gene therapy (Shum, Zhou et al. 2013)

The HIV-1 genome is flanked by 5' and 3' long terminal repeats (LTR) consisting of a U3, R and U5 region. The U3 region acts as an enhancer, with the R repeat region linking the region to U5, which acts as a promoter sequence. A primer binding sequence is present adjacent to the 5' U5 region. This 18-base pair sequence allows binding of tRNA to allow transcription of viral RNA. The HIV LV codes for a number of genes including *Gag*, *Pol* and *Env*. *Gag* genes

code for supporting viral proteins including matrix and nucleocapsid proteins. *Pol* genes produce reverse transcriptase, integrase, and protease enzymes. *Env* codes for surface and transmembrane glycoproteins. Further accessory proteins include tat, rev, nef, vi, spr and vpu are encoded for by the wild type HIV genome. These accessory proteins are involved in regulating HIV infection and replication within cells. A vital feature of the genome of LV is the central polypurine tract (cPPT), present at the 3' end. This resists RNAase H mediated degradation and also acts as a further internal primer for DNA synthesis. The cPPT central termination sequence also resists RNAase H mediated degradation and acts as an internal primer. This provides discontinuity for viral replication, allowing the formation of a triple helix structure at the 3' end of the viral genome, termed a flap region. This is used to increase infectivity of the virus but subsequently removed. The cPPT- central termination sequence also has an important role within nuclear import and promotes increased kinetics of the reverse transcriptase enzyme (Durand, Cimarelli 2011). Thus, knowledge of the virus genome allows manipulation for use in gene therapy.

The genome of the virus can be manipulated to ensure safe infection and delivery of transgenes within host cells. LV can carry a transgene of up to 8 kb and demonstrate a low immunogenicity response upon infection (Nikolaitchik, Dilley et al. 2013, Kumar, Keller et al. 2001). Various deletion studies have demonstrated specific genes that are vital for LV production. As such, to produce replication defective LV vectors, these biological particles are grown in human embryonic kidney 293 (HEK293) producer cells. These cells receive genes in *trans*. While these cells were initially produced following exposure to Ad type 5 DNA, allowing replication deficient Ad vector packaging, the versatility of this cell line allows packaging of LV vectors. HEK293T cells can be transfected effectively with plasmids containing various genes required to produce lentiviral vectors, in particular due to expression of SV40 T antigen

(Thomas, Smart 2005, Graham, Smiley et al. 1977). The transgene to be packaged within the LV must be encoded for on one plasmid, flanked by 3' and 5' LTRs and psi (ψ) packaging sequence. *Gag*, *Pol* and *Env* genes encoded for on two separate plasmids allow essential proteins to be packaged within the virus, in *trans*. Transfection of HEK293T cells allows the efficient transcription of these plasmids by hijacking host cell machinery. The transgene can be packaged within the capsid due to the ψ packaging sequence along with essential proteins, including reverse transcriptase, integrase, and protease. The virus buds from cells, enveloped in the phospholipid bilayer coated and envelope proteins (Sena-Esteves, Gao 2018). However, to ensure therapeutic safety this viral vector is replication deficient. While the virus can infect cells due to the presence of receptor glycoproteins, the transgene does not code for essential gag, pol or envelope proteins required for replication of the virus. Thus, upon infection and integration, the therapeutic vector is not capable of replication. This ensures the safety lentiviral vectors in gene therapy.

Multiple generations of lentivirus vectors have been created through deletions of wild type genes, allowing a larger insertion cassette. These deletion studies have found which genes are essential to viral production. The first HIV-1 replication incompetent vectors were produced expressing all viral genes, with specific deletions in 3' *Env*. The expression cassette also contained sequences for viral export including rev response element, *nef, vif, vpu* and *vpr* (Terwilliger, Godin et al. 1989). Along with further deletions, LV vectors were produced using fewer wildtype genes which identified genes vital for lentiviral production. The result was the creation of third generation vectors which solely contain the transgene flanked by a 5' and 3' LTRs. All essential proteins are provided on plasmids in *trans* to increase the size of transgene insertion (Dull, Zufferey et al. 1998). This allows the production of LV vectors which can be grown at a lower biosafety level.

Advancements have been made in the way of ensuring the safety of replication defective viral vectors. The issue with insertional mutagenesis is that due to an active 3' and 5' LTR segments, the U3 fragment acts as an enhancer promoting the transcription and dysregulation of nearby genes. Self-inactivating LTRs (SIN- LTR) have been created by inactivating the U3 fragment through mutations or deletions (Miyoshi, Blömer et al. 1998). While this does not allow the transcription of the transgene, an internal promoter is included to ensure transcription of the transgene. This has been shown to be effective for use in gene therapy demonstrated by X-SCID trails using a SIN-LTR vector (Zhou, S., Mody et al. 2010). Thus, progress has been made in ensuring the safety of these therapeutic vectors.

1.10. Pseudotyping Lentiviral Vectors with Vesicular Stomatitis Virus Protein G

A major advantage of LVs is pseudotyping viruses with various envelopes. As the Env gene is provided in *trans* on a separate plasmid, various glycoproteins can be coded for to pseudotype the viral vector to target specific cells (Duverge, Negroni 2020). This allows more effective infection of specific differentiated cell types. A common envelope of pseudotyped lentivirus vectors is the vesicular stomatitis virus glycoprotein (VSV-G), typically of the Indiana serotype. VSV comprises a class V, negative sense, single stranded RNA encoding five proteins. The glycoprotein consists of 511 amino acids, including a N- terminal sequence of 16 amino acids which is cleaved during insertion of the glycoprotein in the endoplasmic reticulum. The mature glycoprotein present on the virus envelope is responsible for the attachment and fusion of the virus with host cells (Lyles, Rupprecht 2007, Lawson, Stillman et al. 1995). The fusion of VSV-G to endosomes trigger conformational changes within the

glycoprotein. Consequently, increases in the acidity of endosomes triggers the release of VSV-G proteins and insertion into the cell membrane (Sun, Roth et al. 2010). The VSV-G protein is particularly attractive as it has a ubiquitous insertion profile and allows concentration of virions via ultracentrifugation despite there being few cell types which are not well transduced by this protein (Burns, Friedmann et al. 1993, Amirache, Levy et al. 2014). Upon infection, VSV-G binds to specific receptors present on the cell surface membrane allowing the virus to enter the cell by clathrin mediated endocytosis. Multiple studies have demonstrated that exposure of cells to various chemicals, including trypsin enhances glycoprotein binding (Schloemer, Wagner 1975). These suggests that the binding site of VSV-G is saturable, further indicated by inhibition of VSV-G binding to Vero cells after phospholipase C and phosphatidylserine treatment but not subsequent to protease or heat shock. Therefore, it is suggested that VSV-G interacts with phospholipase C upon binding in the inner cell membrane, which is more effective in an acidic environment (pH 5.0-6.2). However, as phospholipase C lies within the inner membrane of the cell, is not the initial receptor that VSV-G binds to. Multiple studies have strongly indicated that phosphatidylserine and low-density lipoprotein (LPL) are receptors for VSV-G. This goes some way to explaining the ubiquity of VSV-G as phosphatidylserine and LDL are present on multiple differentiated cell lines. Various assays have shown that the histidine rich amino acid region 145-168 binds to phosphatidylserine with the G protein. This further enhances the entry of the virus into host cells (Coil, Miller 2004, Estepa, Rocha et al. 2001, Seganti, Superti et al. 1986, Schlegel, Tralka et al. 1983, Schlegel, Willingham et al. 1982). The stability of VSV-G is further demonstrated by its ability to withstand concentration by ultracentrifugation (Cronin, Zhang et al. 2005). A further benefit of VSV-G is that is triggers activation signals upon insertion of the transgene to allow expression of the therapeutic transgene especially in quiescent cells (Durand, Cimarelli 2011). This demonstrates the

versatility to using VSV-G to psuedotype lentivirus vectors for use in gene therapy.

However, there are several pitfalls with using VSV-G as an envelope. The infection of specific cell lines using the VSV-G envelope, such as lymphocytes, can still prove difficult due to the low expression of phosphatidylserine and LDL (Amirache, Levy et al. 2014). VSV-G has also been shown to be toxic to cells, resulting in the formation of syncytia (Hoffmann, Wu et al. 2010). Moreover, due to the wide tropism of VSV-G unspecific cells may also be infected with these vectors after administration.

Further developments to VSV-G have been made to allow specific targeting of cells and to reduce cytotoxicity. The VSV-G gene has been successfully fused with Rabies virus glycoprotein (RV-G). The Rabies virus has been well studied, in particular after various outbreaks. It has been shown to transmit transsynaptically through a retrograde fashion in the central nervous system. This glycoprotein has been shown to effectively bind to phosphatidylserine, a common receptor with VSV-G. However, many proteins have been proposed as receptors of the Rabies virus, including nicotinic acetylcholine receptor (NAchR), $p75^{NTR}$ and neural cell-adhesion molecule (NCAM). The sole glycoprotein present on the Rabies virus envelope is of particular interest in targeting neuronal populations specifically through gene therapy. VSV-G has been fused with the Rabies virus glycoprotein (FuG-B) with the extracellular and transmembrane domains of the RV- G and the cytoplasmic domain of VSV-G. FuG-B has been shown to effectively pseudotype LV vectors to high titers and efficiently transfer genes via retrograde transport along axons, including motor and dopaminergic neurons (Kato, Kobayashi et al. 2011, Federici, Kutner et al. 2009, Mentis, Gravell et al. 2006). This raises the possibility of correcting degenerative neurological diseases (Poewe, Seppi et al. 2017).

Therefore, LV vectors can be successfully produced using through split genome systems, pseudotyped to increase cell specificity, for gene therapy.

The elegant systems of delivering transgenes to host cells via viral vectors allow cells to transcribe wild type proteins to correct metabolic function. This has already been shown as successful in the treatment of a wide variety of monogenetic disorders, including cystic fibrosis. The delivery of the cystic fibrosis transmembrane regulator (*CFTR*) gene has been shown as effective in various mouse models (Duchesneau, Besla et al. 2017). Gene therapy has proven to successfully treat and cure patients of genetic diseases using a variety of vectors suited to the treatment of specific diseases.

1.11. Lentivirus Purification Methods for Gene Therapy

ATPS has been previously reported to successful purify Ad and AAV based vectors. However, no such application exists towards LV vectors and should be investigated further.

The purification and concentration of LV vectors remains an issue to recover high titre particles for use in downstream applications. LVs are commonly generated through split genome transfer, with transfection of HEK293T cells using a number of plasmids to generate replication deficient viral particles. Alternatively, stable cell lines generated through integration of DNA encoding viral components (i.e transgene, *gag*, *pol*, *rev*, *env*) have been generated to produce vectors continuously. This is useful for continuous production of high titre LV but poses an issue with the cost and length of time required for generation of stable clones and drop in titre over an extended period of time (Xu, K., Ma et al. 2001). Commonly, VSV-G is used to psuedotype LV

particles due to their ubiquitous insertion profile for cells. However, only a few stable cell lines have been generated using VSV-G envelopes due to the cytotoxicity of this glycoprotein (Qiao, Moreno et al. 2006). Alternative approaches have been sought for long term expression of high titre LV vectors. Pseudotyping bald viruses with VSV-G have been shown effective at increasing the level of gene transfer (Tijani, Munis et al. 2018). Furthermore, inducible cell lines have been generated for stable vector production. These cell lines are generated by integrating DNA for viral components in cells under inducible promoters. Upon exposure to specific chemicals (e.g. doxycycline), these genes are expressed to produce viral vectors. However, the viral titre in these cell lines may drop over time due to selection pressures required selection of clones through antibody exposure (Ikeda, Takeuchi et al. 2003). Both stable and inducible cells are difficult to clone resulting in a time consuming and cost ineffective procedures. The particles remain unstable at 37°C, with a half-life of 9-12 hours (Bandeira, Peixoto et al. 2012).

As well as various production methods of these vectors, protocols have been developed for LV purification. Conditioned medium of transfected cells are commonly filtered to remove cell debris. However, contaminating proteins remain which are concentrated upon ultracentrifugation (Soldi, Sergi et al. 2020). These have been removed using various protocols. Ultracentrifugation using a 10% sucrose cushion has shown optimal vector recovery, greater than observed with PEG precipitation or protein column ultracentrifugation (Boroujeni, Gardaneh 2018). Chromatographic purification has been shown successful at purifying LV vectors. Anion exchange chromatography has shown promise with 70% of viral particles recovered and 90% of DNA impurities removed (Merten, O. W., Charrier et al. 2011). However, issues persist with LV capture and shearing due to particle stress after binding to specific beads. These issues have been optimised to varies degrees of success, with LV particles

harvested from WinPac cells have shown a high degree of concentration using ion exchange chromatography combined with cellulose nanofibres to increase surface area for virion capture (Ruscic, Perry et al. 2019, Lesch, Laitinen et al. 2011). A combination of these techniques has shown success. Microfiltration, anion exchange chromatography and ultracentrifugation of vectors has shown recovery of 80% of particles (Bandeira, Peixoto et al. 2012). Therefore, multiple successful protocols have been developed to remove impurities from LV vector preparations.

1.12. Sequencing of Lentivirus RNA

Wild type HIV-1 viral RNA has been sequencing to near full. length coverage using a variety of RNA sequencing platforms (Di Giallonardo, Zagordi et al. 2013, Dauwe, Staelens et al. 2016, Hebberecht, Vancoillie et al. 2019). Sequencing of these viruses have revealed more than 50 splice variants packaged within these virions. The packaged RNA transcripts are generally spliced from the 5' end of the genome and classed into three groups: unspliced (encoding *gag*, *pol*), spliced (the excision of *gag*, *pol* and encoding for *env*, *vpu*, *vif*, *vpr* and truncated *tat*) and double spliced variants (excision of major introns overlapping *gag*, *pol*, *vpu* and *env*). Rare variants spliced from the 3' end of the genome have also been identified (Smith, J., Azad et al. 1992, Carrera, Pinilla et al. 2010, Ocwieja, Sherrill-Mix et al. 2012). As such, the diversity of sequenced packaged within these virions reveals the use of various splice donor and acceptor sites identified in the HIV genome (Vega, Delgado et al. 2016, Emery, A., Zhou et al. 2017). These transcription of splice variants are temporally regulated through wild type HIV infection. This is highlighted by transport of the genome across the nuclear membrane dependent on the concentration of Rev. Double spliced variants are rev independent and expressed initially, with

single- and un-spliced variants expressed after the Rev threshold is reached (Kim, S. Y., Byrn et al. 1989). This is highlighted for gene therapy with cryptic splice sites in the genome shown to alter transcript production (Moiani, Paleari et al. 2012). These splice variants have been associated with insertional mutagenesis (Almarza, Bussadori et al. 2011, Nilsen, Maroney et al. 1985). Therefore, these need to be considered in the production of LV vectors.

1.13. Adverse Side Effects Identified in Gene Therapy

There are major safety concerns facing gene therapy. The use of integrative LV vectors raises the issue of insertional mutagenesis. The semi-random insertion of a viral transgene into the host genome can trigger oncogenesis due to the dysregulation of tumour suppressor, protooncogenes or oncogenes (Ciuffi, Llano et al. 2005). This has been reported in a number of clinical trials (David, Doherty 2017). Due to integration, the viral transgene may contribute to promoter or enhancer activation of protooncogenes/oncogenes, dysregulation of host tumour suppressor genes or indeed modification of epigenetic mechanisms of host genes. Gene dysregulation may then in turn contribute towards oncogenesis. This has been illustrated by a number of historic cases, including in X-SCID trials. CD34 T- cells derived from patients were treated with a retrovirus carrying the *γc* transgene before autologous cell transfer (Hacein-Bey-Abina, Von Kalle et al. 2003). While this corrected the disease in 90% of patients, five patients developed lymphoma due to integration of the transgene within the promoter of the *LMO2* protooncogene (Hacein-Bey-Abina, Garrigue et al. 2008, Hacein-Bey-Abina, Von Kalle et al. 2003). This phenomenon is not unique to LV vectors, with insertional mutagenesis reported in adeno-associated vector (AAV) clinical trials (Chandler, LaFave et al. 2015, Donsante, Miller et

al. 2007). This demonstrates the need to ensure the safety of viral vectors before proceeding to the clinic.

1.14. Mechanisms of Genotoxicity

Insertional mutagenesis and associated oncogenesis are adverse side effects undesirable in gene therapy. The integration of therapeutic transgenes in specific loci may cause gene dysregulation contributing towards clonal outgrowth. This has been reported in a number of clinical trials, notably X-SCID trials where a number of patients developed leukemias due to integration in the *LMO2* oncogene (Hacein-Bey-Abina, Garrigue et al. 2008, Hacein-Bey, Cavazzana-Calvo et al. 1996, Hacein-Bey-Abina, Von Kalle et al. 2003). As such, it is important to develop models of integration of various viral vectors to assess and mitigate contributing factors to avoid insertional mutagenesis in the clinic.

Multiple factors affect the genotoxicity of a viral vector such as LV, including dominant gain of function mutations due to integrations, promoter activation, enhancer activation and gene truncations, primarily due to the site of integration (Aiuti, Cossu et al. 2013). Studies have shown common integration profiles of various viral vectors. RV insert in proximity to transcription start sites, gene regulatory elements, DNAase I hypersensitive sites and CpG islands. Hotspots for these vectors have shown integration in genes common to biological pathways (i.e. cell growth) (Suerth, Labenski et al. 2014, Cattoglio, Pellin et al. 2010, Lewinski, Yamashita et al. 2006, Niederer, Bangham 2014). Moreover, MLV has been found to integrate within sites spanning only 2% of the human genome (LaFave, Varshney et al. 2014). HIV has shown a different insertion profile, with common integration within transcriptional units of highly expressed genes (Mitchell, Beitzel et al. 2004, Schroder, Shinn et al. 2002). This

virus is known to integrate within the major groove of outward facing DNA, primarily at the nuclear periphery where active genes are located (Wang, G. P., Levine et al. 2009). Due to the expression levels of genes effecting the sites of integration, this implies IS are dependent on cell status, as highlighted by differences detected in insertions detected between in resting and activated CD4+ T cells (Brady, Agosto et al. 2009).

Dominant gain of function mutations due to integrations of the vector within sites cause gene dysregulation (Figure 1.7). Aberrant gene expression may be due to a number of factors. Promoter or enhancer activation are caused through the LV LTR region acting to aberrantly regulate promoters or enhancers of genes (Touw, Erkeland 2007, Aiuti, Cossu et al. 2013). This gene dysregulation of protooncogenes and oncogenes may contribute towards cellular transformation and oncogenesis. Enhancer activation has been detected in clinical trials and mouse models of LV integration where insertional mutagenesis has been identified. Conversely if enhancer activity is reduced, this contributes towards inactivation tumour suppressor genes which is also undesirable (Cesana, Ranzani et al. 2014).

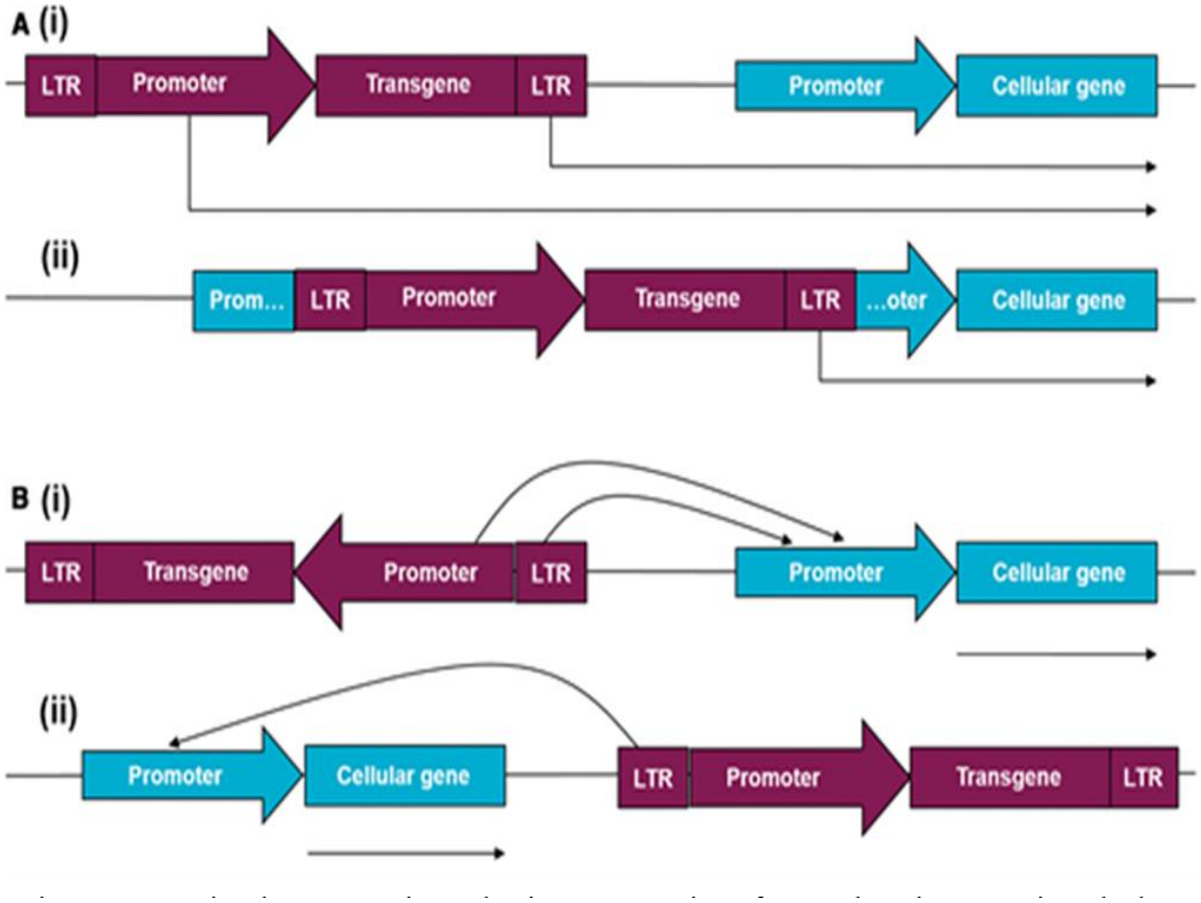


Figure 1.7- Insertional mutagenesis mechanisms. Integration of HIV-1 based genome into the host genome may cause insertional mutagenesis based on a variety of factors due to integration. A: Promoter activation of upstream transcriptional units (i) or in proximity of the 5' promoter region of a gene (ii) contributes to gene dysregulation and potential oncogenesis. B: Enhancer mediated activation of nearby target gene (i) or downstream integration of the virus may mediate activation of upstream genes through the LTR (ii). (David, Doherty 2017)

As well as these, gain of function may cause truncations of gene transcripts. These can be subdivided into two distinct classes: truncations of gene transcripts and readthrough from the 3' LTR of the LV vector genome. These have been found more commonly in LV vectors carrying a native LTR with fusions occurring between a splice donor or acceptor site between the LTR and exons in genes contributing towards truncated protein formation (Cesana, Sgualdino et al. 2012). Truncations are important to understanding vector genotoxicity as these may alter the characteristics of mRNA transcribed. This

has been highlighted by the insensitivity of cells to miRNA let-7 degradation of alternate HMGA2 transcripts (Cavazzana-Calvo, Payen et al. 2010). Epigenetic modifications can also be caused through LV integrations, for instance histone and DNA methylation, contributing towards promoter silencing (Touw, Erkeland 2007, Aiuti, Cossu et al. 2013). The importance of IS in understanding genotoxicity demonstrates vector design may be modified to reduce these risks including LTR, splice donor and acceptor sites and poly adenylation signal.

Non-viral factors also contribute towards genotoxicity which must be considered for clinical therapeutic gene delivery. The transgene used only has a partial effect on genotoxic events as demonstrated by MLV based RV inserting into *LMO2* in treating X-SCID. However, in ADA-SCID trials, *LMO2* was identified as an IS but leukaemia was only observed in the former trials (Aiuti, Cassani et al. 2007, Niederer, Bangham 2014).

The epigenetic pattern of cells can also contribute towards insertional mutagenesis events. Integration profiles are more clustered regions in CD34+ cells transduced with MLV, with strong preference for H3K4me3 (Biasco, Ambrosi et al. 2011). Methylation silencing of genes has been shown in RV treatment of chronic granulomatous disease, with RV promoter mediated methylation correlated with global changes. No methylation was detected at the LTR enhancer sites suggesting the promoter region is responsible for this effect (Stein, Ott et al. 2010).

Tissue type and maturity of cells also plays a role. HIV preferentially integrates within active genes and therefore insertion in specific cell types is varied. Fewer common insertion sites were identified in immature haematopoietic stem cells compared to their mature peripheral blood leukocyte derivatives (Biasco, Ambrosi et al. 2011). This is highlighted by a foetal mouse model being more sensitive to genotoxicity than adult mice (Themis, Waddington et al. 2005).

Gene dysregulation through these non-viral methods must be considered when assessing the genotoxic risk of gene therapy to patients.

1.15. Models for Analysing Gene Therapy Safety

The importance of determining the safety of gene therapy is paramount. This has been highlighted by the European Medicines Agency (EMA) guidelines on the development and manufacture of LV vectors. These guidelines, detailing the nature of lentiviral production, transgene efficacy and biosafety testing are paramount for pharmaceutical production of vectors for gene therapy. This includes determining the absence of replication competent LV and oncogenesis. The EMA has published strict guidelines on ensuring viral vectors produced for gene therapy are not replication competent. LV vectors produced for gene therapy are manufactured to be replication deficient. Therefore, these viral vectors are not able to replicate and spread within the patient. However, due to spontaneous recombination events, replication competent viruses may be produced. These viruses, while carrying the therapeutic transgene, still contains the wild type function of replication and therefore, the safety of vectors is compromised. The EMA has provided guidelines detailing that should any replication competent viruses be observed within a batch of vectors, then the whole batch must be discarded due to the contamination present within this batch (Committee for Medical Products for Human Use 2005). The need for these guidelines has been highlighted by detecting insertional mutagenesis and associated oncogenesis in a number of clinical trials (Cavazzana-Calvo, Payen et al. 2010, Biffi, Montini et al. 2013). The strict international guidelines surrounding the production and safety testing of lentiviral vectors demonstrates the importance of determining the safety of these vectors before administering to patients.

Genotoxicity models are required to study the safety and efficacy of viral vectors *in vivo* and *in vitro*. Multiple *in vivo* animal models are useful to indicate the genotoxic effects of LVs, with this vector implicated in hepatocellular carcinoma (HCC) (Nowrouzi, Cheung et al. 2013, Chandler, LaFave et al. 2015, Themis, Waddington et al. 2005). A sensitive murine *in utero* model of LV genotoxicity identified specific LV vector insertions and dysregulation of oncogenes contributing towards hepatocellular carcinoma (Nowrouzi, Cheung et al. 2013). HCC was demonstrated *in utero* after delivery using a range of non-primate LV, including EIAV and FIV. This model is highly sensitive in targeting a wide array of highly expressed genes at a foetal stage. The genotoxicity of RV and HIV based LV have shown an oncogenic risk in an *in vivo* model using cdkn2a$^{-/-}$ mice (Montini, Cesana et al. 2006). These mice have been shown to be prone to tumour development and are therefore biased towards oncogenetic events and a useful model to understanding genotoxicity. The *cdkn2a* gene encodes two proteins, p16^{ink4a} and p19^{arf} which act as a regulator of cdk4/6 mediated by Rb1 phosphorylation and mdm2 mediated degradation of p53, respectively. Thus, the knockout of *cdkn2a* expression leads to a deficiency of Rb1 and p53, both of which are involved in pathways regulating cell cycle (Sherr 2004). This makes the mice more prone to developing tumours. RV and LV mediated gene transfer in these mice have identified IS common to genes known to be involved in oncogenesis and enriched for pathways involved in cell cycle. *Ex vivo* gene transfer using RV or LV to haematopoietic stem cells from Cdkn2a$^{-/-}$ mice and translation in wild type mice revealed tumorigenesis quickly in mice. However, this model is limited to assessing genotoxicity with a high vector copy number and both these models target cells of a murine lineage (Montini, Cesana et al. 2006). While the genotoxic risk of RV was determined, this model did not reveal LV prone to tumorigenesis and IS sites were not enriched for specific pathways despite delivering a higher viral load, suggesting a higher dose of LV can be tolerated

by animals (Follenzi, Ailles et al. 2000). Furthermore, RV showed more targeted integrations than their LV counterparts (Montini, Cesana et al. 2006). These mice have also been used to assess the *in vivo* risk of Maloney murine leukaemia virus (MoMLV), which identified insertional mutagenesis and tumour formation in oncogenes associated with *cdkn2a* (Lund, Turner et al. 2002).

A different approach to understanding genotoxicity has been developed through *ex vivo* transplantation models in animals. Transduced haematopoietic stem cells have been transplanted back into cdkn2a$^{-/-}$ mice leading to tumour formation. Analysis of integration sites and oncogenesis have supported the hypothesis that LV carrying a self-inactivating LTRs carry a 10-fold lower risk of genotoxicity than RV. The functional regions for genotoxicity were determined as the LTRs, as assessed by swapping the LV based LTR into RV vectors (Montini, Cesana et al. 2009).

However, these models are expensive, time consuming and cause animal suffering. Genotoxicity in animals may not fully translate to a human platform. Mice have been shown to be, at minimum, 10 times more resistant to drugs than humans. Similarly, rats and dogs have been shown to tolerate a 4.5 to 100-fold increase in concentrations of specific drugs in comparison to humans (Wong, Budha et al. 2012). Coupled with an increasing emphasis on the replacement, reducing and refinement of the use of animals in research, cellular models of genotoxicity are paramount (Percie du Sert, Robinson 2018).

Cellular models are more beneficial in reducing the use of animals in research as well as proving a useful model of cellular genotoxicity, time and cost effective. Through using cellular models to determine genotoxicity, the safety profile of vectors may be determined before the transition into the clinic.

The *in vitro* immortalisation (IVIM) assay has been developed as a rapid cellular assay for determining the genotoxicity of viral vectors avoiding the use of animals. This has been developed using a murine cell line developed from C57BL6 mice, specifically chosen due to the absence of background transformation. Cellular transformation and molecular characteristics were observed with increasing doses of gene transfer. This provides a rapid model for analysing insertions of RV and LV vectors. A lower risk of insertional mutagenesis was found with LV vectors, in particular when carrying constructs flanked by SIN LTRs and cellular promoters. However, this assay identified a low genotoxic risk of SIN vectors with transformation above background (Molich, Navarro et al. 2009). Further, IL-3 dependent murine cell lines transduced with γRV have been studied to determine MLV integration in proximity to transcription start sites (Wu, X., Li, Crise, and Burgess 2003). This assay provides a rapid *in vitro* method with less intensive maintenance of cells to assess genotoxicity and provides a direct comparison between RV and LV vectors. However, these *in vitro* models are biased due to species specificity, are not highly sensitive and have a predisposed genetic background. Therefore, there is a need for an unbiased human based *in vitro* model of genotoxicity that sensitively captures integration site preferences.

1.16. Insertion Site Sequencing Technologies

There are multiple assays that have been developed to sequence the DNA next to IS. This is of importance to assess the insertion profile of cells and has been used to detect RV and LV insertion sites, including in relation to insertional mutagenesis (Woychik, Alagramam 1998). Two step PCR methods have been developed including inverse PCR, but these methods are limited in sensitivity. Inverse PCR is used for *in vitro* identification of IS from clonal outgrowth of one IS which limits downstream applications (Silva, Santos et al. 2017).

Ligation mediated PCR can be used to detect multiple IS but the sensitivity must be improved for various applications (Ciuffi, Barr 2011).

Extension primer tag selection ligation mediated PCR (EPTS/LM PCR) has been developed using selection of extension tags and magnetic capture prior to adapter ligation and direct sequencing. This has proved a highly sensitive method (up to 1 in 1000 cells) to determine the IS of RV and LV transduced cells, *in vitro* and *in vivo* (Schmidt, Hoffmann et al. 2001).

1.17. Lentivirus modifications for Safer Gene Delivery

A number of modifications have been made of LV vectors in an effort to make them safer vectors for gene delivery. Of these, the most notable one is the development of self-inactivating (SIN) LTRs. The long terminal repeat region flanking the viral genome contains a U5, R and U3 regions which act as an enhancer and promoter respectively. In a SIN LTR configuration, the U3 region is deleted, indicating that the LTR cannot act as a promoter of genes close to the site of integration, thus reducing the risk of insertional mutagenesis (Zufferey, Dull et al. 1998). These vectors have been shown to carry a lower oncogenic risk than the native LTR (Zychlinski, Schambach et al. 2008). Due to the inability of the 5' LTR to activate gene expression, an internal promoter is commonly used for transgene expression.

A further modification to LV vectors is the inclusion of chromatin insulators. These elements included in the vector genome function to lock interactions between the vector and host genome (Emery, D. W. 2011). These can be classed as barrier or enhancer blocking insulators which act to prevent shutdown of the internal promoter or prevents the enhancer region of the LTR from activating the promoter region. With the development of SIN LTRs, enhancer mediated oncogenesis is commonly observed with these vectors and therefore it is not

surprising that enhancer blocking insulators are commonly used (Browning, Trobridge 2016). While the inclusion of these regions may decrease the titre of LV production 6-20-fold, conflicting reports have been published (Nielsen, Jakobsson et al. 2009, Liu, M., Maurano et al. 2015).

Modification of vector dose has also been used to reduce the risk of insertional mutagenesis. While a lower vector dose results in fewer IS and risk of oncogenesis, this poses the risk of not correcting the targeted disease. In cystic fibrosis, cells have shown to be corrected with only 25% of cells targeted by gene delivery (Zhang, L., Button et al. 2009). A further study has shown targeting of 20-30% of sinus epithelial cells from $CFTR^{-/-}$ pigs showed correction to 50% of normal levels. This was shown even more sensitively, with as little as 7% gene transfer showing detectable levels of CFTR protein production, 6% above that with no gene delivery (Potash, Wallen et al. 2013). Therefore, modification of viral dosage may be helpful in reducing the risk of insertional mutagenesis but the levels of gene transfer for correction must be determined for specific diseases. There is a present need for an alternative to *in vivo* and *in vitro* animal models to understand the safety profile of viral vectors in a humanised platform.

1.18. Induced Pluripotent Stem Cells and Their Applications

Induced pluripotent stem cells (iPSCs) provide an unlimited source of cells for differentiation into multiple cell types. This work has built on previous knowledge of growing embryonic stem cells, with these established in culture demonstrating the capacity of stem cells to grow *in vitro* (Evans, Kaufman 1981, Thomson, Itskovitz-Eldor et al. 1998, Takahashi, K., Yamanaka 2006). iPSCs are derived from somatic cells, reprogrammed into a pluripotent state using various factors (Takahashi, Yamanaka 2006). The potential of these cells

in therapeutics and regenerative therapy was recognised by the Nobel committee (Nobel Assembly 2012). This is especially useful with the increased emphasis on personalised medicine, with the capacity for patient cells to be reprogrammed and autologously transplanted. These cells were originally reprogrammed using *Oct4*, *Sox2*, *c-Myc* and *KLF4*. The functions of these genes demonstrate their important role in maintaining pluripotency. *Oct4* is a transcription factor involved in maintenace of pluripotency. *Sox2* is involved in mediating expression of *Oct4* implying both genes are vital for pluripotency (Masui, Nakatake et al. 2007). The protooncogene *c-Myc* has been implicated in a number of cancers. The transcribed protein recruits chromatin modifying proteins to trigger transcriptional activation (Dang 2012). However, *c-Myc* is not vital for reprogramming and thus reduced the oncogenic potential of these cells (Nakagawa, Koyanagi et al. 2008). *KLF4* itself activates *Sox2* (Niwa, Ogawa et al. 2009). Therefore, the induction of these pathway triggers reprogramming of cells to a pluripotent cell state. The pluripotency of these cells can be identified using various markers, such as TRA-1-60 and TRA-1-81 (Schopperle, DeWolf 2007).

The ability to reprogram a patient's somatic cells reduces the need for invasive procedures to collect host cells. These cells can be readily differentiated into specific cell types for the patient before correction via gene therapy and autologous cell transfer. This has been highlighted by autologous skin graft transplants in patients suffering from recessive dystrophic epidermolysis bullosa. RV carrying *COL7A1* were used to correct autologous keratinocytes, with patients demonstrated an improvement in wound healing (Siprashvili, Nguyen et al. 2016). This indicates the potential for further autologous treatment using iPS cells.

Several protocols have been developed for the modification of iPS cells. Typically, viruses expressing *OCT4*, *SOX2*, *KL4* and *c- MYC* infect cells,

triggering cellular reprogramming. A range of viruses have been tested for infection. While γRV vectors encoding each factor individually were initially used, the efficiency of iPS cell formation was poor and there was a higher incidence of insertional mutagenesis due to multiple integration points (Takahashi, K., Yamanaka 2006). Further progress has been made in creating polycistronic lentiviral vectors, expressing all reprogramming factors which have successful produced iPS cells (Gonzalez, Barragan Monasterio et al. 2009). However, integration events still pose a risk of insertional mutagenesis. Thus, episomal transfer of these factors has been utilised with integration deficient lentiviruses. Due to a mutation in the integrase gene, the viral genome can still be reverse transcribed to cDNA but is not able to integrate within the host genome. Therefore, 5' and 3' LTRs form circular pieces of DNA, which transiently express these reprogramming factors. This has been shown to be sufficient for the generation of iPS cells and indicates a potential safety mechanism, as the expression of these factors is deregulated in daughter populations (Wu, Y., Marsh 2001).

Non integrative viruses have been used for iPSC reprogramming. Ad has been used to highly transduce the donor cell type using *Oct4*, *Sox2*, *c-Myc* and *Klf4* under the control of promoters (e.g. hCMV-IE) to reprogram mouse tail tip fibroblasts. This vector shows high levels of transduction efficiency (~40%) while the rate of iPSC clones formed remains low (0.001-0.0001%) (Stadtfeld, Nagaya et al. 2008, Zhou, W., Freed 2009). Sendai viruses are commonly used for reprogramming of somatic to iPS cells. This is a RNA virus which, as it is non integrative, is diluted out after 10 cellular passages and rapidly transcribes and translates the nucleic acids to a high concentration of proteins. These effectively reprogrammed a range of cells to a pluripotent state (Fusaki, Ban et al. 2009, Seki, Yuasa et al. 2014). Sendai virus transduction has shown rapid

reprogramming (~25 days) at a high efficiency (1% fibroblasts) (Malik, Rao 2013).

The differences in integrative and non-integrative methods of forming iPSCs has been highlighted in their differences in gene expression. Cells reprogrammed using a non-integrative method were found to be closer to the human embryonic stem cell state than their integrative counterparts (Liu, Y., Cheng et al. 2012). Therefore, the safety of iPSC generation has been developed to an extent.

Recent studies have highlighted the importance of the epigenome in treating patients with underlying genetic conditions. Epigenetic modifications to the host genome, including methylation, acetylation and sumoylation have been shown as important for the expression of specific genes and dysregulation has been implicated in a number of pathologies (Godini, Lafta et al. 2018). Therefore, the epigenetic profile of iPSCs should be assessed before further clinical progression.

Epigenetic modification in iPSCs has been studied to an extent and revealed the modifications subsequent to initial gene transfer (Ankam, Rovini et al. 2019). The epigenome of these pluripotent reprogrammed cells is similar to that of the host. For this reason, it is of paramount importance in choosing the correct host tissue type to reprogram to iPSCs. Due to host immune responses upon viral infection and integration of DNA, changes in the methylation profile of transgenes occur and global methylation of host DNA increases (Fang, Xiao 2001). This may result in the silencing of multiple host genes which increases the risk of oncogenesis. Therefore, epigenetic modifications must be considered when modifying cells for treatment to ensure safety of cells, especially in autologous cell transfer.

The fetal mouse model has been shown as a sensitive model of LV vector mediated genotoxicity (Nowrouzi, Cheung et al. 2013). Human derived iPSC provides an alternative source of tissue to recapture a similar naïve cell state in a human *in vitro* platform to understand LV vector adverse side effects.

1.19. Induced Pluripotent Stem Cells as an Alternative Model for Toxicity

iPSCs have been shown as a viable model for screening various drugs (Sachlos, Risueno et al. 2012). An estimated 90% of drugs that fail in clinical trials are primarily for safety or efficacy issues (Hay, M., Thomas et al. 2014). This presents the need to model the various toxicity effectively prior to moving to clinical trials to provide a more cost and time efficient production pipeline.

iPSCs and their neuronal derivatives have been used for toxicity analysis. In particular, the response of these cells to drug treatments for Parkinsons, Alzheimers and Huntington's disease, all of which affect neuronal cells in degenerative conditions (Buzanska, Jurga et al. 2006, Takahashi, J. 2013).

Cardiomyocytes derived from iPSCs have been used to mimic pathologies effecting the heart. These cells have shown a similar response to embryonic stem cells to assess the safety and efficacy of cardioactive drugs, demonstrating their potential in therapeutic drug discovery (Lian, Zhang et al. 2010, Laustriat, Gide et al. 2010). For instance, four drugs (Isoproterenol, E-4031, Verapamil, Metoprolol) have been studied for their response in cardiomyocytes. The response of these drugs on iPSC derived cardiomyocytes shows their promise in assessing the drug efficiency *in vitro* (Csobonyeiova, Polak et al. 2016).

The promise of these cells in assessing drug toxicology suggests these pluripotent cells and their cardiomyocyte derivatives, may be useful in assessing the genotoxicity of lentiviral mediated gene therapy *in vitro*. This is useful in

understanding the safety profile of the vector prior to downstream clinical development.

While iPS cells are a useful source of material for disease modelling, research has shown the genomic instability of these reprogrammed cells (Blasco 2011). Various mutations have been identified in iPS cells which have been attributed to a number of factors. These include pre-existing mutations in parental cells. The genomic background of the parental cells used for reprogramming to a pluripotent cell state is important as these mutations can be carried over. These may be particularly difficult initially due their low prevalence in somatic cells (Abyzov 2012, Young, M. A. 2012). Further, reprogramming induced point mutations are common (Sugiura 2014). Due to these, a number of global cellular changes lead to increased instability with prolonged culture of cells *in vitro* (Gore 2011). Point mutations may contribute to mitotic segregation deficiency due to chromosome fusion, DNA damage and laminin depletion (Henry, M. 2019). These mutations must be considered for downstream approaches, with tumorigenesis identified after allogenic stem cell transplantation (Yasuda, T. 2014).

1.20. Stem Cell Differentiation to Therapeutic Targets

iPSCs have been successfully differentiated into multiple lineages to elucidate developmental biology and for therapeutic use. Hepatocyte differentiation of these pluripotent cells provide a rapid alternative source of tissue compared to the standard non-human based models (Fukuda-Taira 1981, Zorn, Mason 2001). Various protocols have been developed, using a range of growth factors and small molecules such as Activin A, FGF, BMP4, OSM and HGF for hepatic induction and maturation. These differentiated liver cells have been well characterised to recapture *in vivo* primary human hepatocytes through CYP3A activity and gene expression analysis (Takayama, Morisaki et al. 2014, Si-

Tayeb, Noto et al. 2010, Lu, Einhorn et al. 2015). While these differentiated hepatocytes are immature, targeted maturation protocols have been developed, including spheroid organoid culture (Baxter, Withey et al. 2015, Rashidi, Luu et al. 2018).

Cardiac differentiation from pluripotent stem cells has been widely developed to improve differentiation efficiency. Pluripotent stem cells were initially reported to differentiate in cardiomyocytes using embryoid bodies and serum. While it was possible to differentiate to cardiomyocytes, these protocols showed a low efficiency (Kehat, Kenyagin-Karsenti et al. 2001). Other protocols use co culturing of END2 cells to improve differentiation towards a cardiomyocyte lineage (Mummery, Ward-van Oostwaard et al. 2003). Understanding of *in vivo* cardiomyocyte development has improved knowledge of cytokines to supply in differentiation protocols, such as BMP4 and Activin A as well as inhibition of Wnt signalling pathways (Laflamme, Chen et al. 2007, Yang, L., Soonpaa et al. 2008). Wnt signalling pathways have been shown to be both beneficial and deleterious to improve differentiation efficiency (Ren, Lee et al. 2011, Tran, Wang et al. 2009).

Limitations do persist with differentiation protocols. Heterogenous populations of cardiomyocytes (i.e. atrial, ventricular, and nodal) are differentiated using standard protocols, with these generally biased towards the ventricular subtype. These have various electrophysiological properties which may be disease specific (Ma, J., Guo et al. 2011, Blazeski, Zhu et al. 2012). While in general it is useful to target a heterogenous population to recapture *in vivo* biology, manipulation of various molecules delivered in the differentiation protocols can improve subtype specific differentiation for specific assays (Devalla, Schwach et al. 2015, Protze, Liu et al. 2017). The development of multiple enhanced protocols has led to the development of highly efficient commercially available kits for cardiomyocyte differentiation (Baci, Chirivi et al. 2020).

iPSC derived cardiomyocytes have also been seen to present an immature phenotype. These differentiated cells present low levels of markers and ion transport genes (i.e. *MYL2*, *MYH7*, *TCAP*, *MYOM2*, *KCNJ2*, *RYR2*) known to be present in mature cardiomyocytes (Funakoshi, Miki et al. 2016). This poses a problem downstream for assaying various drugs. Optimisation of protocols have been developed to improve maturation of cells using chemicals to varied success (Kamakura, Makiyama et al. 2013, Yang, X., Rodriguez et al. 2014, Sun, Nunes 2016). Indeed, spheroid culture of cardiomyocytes have been shown to mature quicker than traditional monolayer culture (Branco, Cotovio et al. 2019). Thus, work has been performed to enhance the maturity of this terminally differentiated cell type. The differentiated cells must be characterised as cardiomyocytes to validate downstream assays. This is commonly performed through morphological analysis and detection of cardiac specific proteins (i.e. cardiac troponin T, SERCA2, Actinin). Further molecular analyses include calcium ion imaging and analysis of contractile function, all characteristic of the muscle cells of the heart, to further validate the cell identity (Seki, Yuasa et al. 2014, Sayed, Liu et al. 2020).

Gene transfer to terminally differentiated cardiomyocytes has been shown possible using various viral vectors. Ad gene transfer is effective, but this presents limited opportunities in targeting aging cells (Communal, Huq et al. 2003, Kirshenbaum 1997). The episomal nature of this virus means it is not useful for permanent correction of genetic diseases. AAV and LV have also been shown to successfully transduce these cells to high levels of gene transfer (Kohama, Higo et al. 2020, Zhang, J., Ho et al. 2014). These vectors are especially useful for gene therapy, as highlighted by their common use in the field. These vectors successfully deliver genes to cells for therapeutic outputs. However, the safety aspect of LV mediated gene transfer still needs to be addressed.

The work presented in this thesis will seek to address the gaps in Ad purification and LV vector production. This work also aims to understand LV mediated genotoxicity in iPSC derived hepatocyte like and cardiomyocyte cells to explore both these derivatives as potential *in vitro* models to better understand the integration of LV in the host genome and assess deleterious effects in a human based platform.

Chapter II

Materials and Methods

Abbreviations

$(NH_4)_2SO_4$	Ammonium sulphate
AAV	Adeno associated vector
AD/AV	Adenovirus
ADA	Adenosine deaminase
Amp	Ampicillin
ATPS	Aqueous two-phase system
BLAST	Basic local alignment search tool
BLAT	BLAST-like alignment tool
BRL	Bio-rad Laboratories
BSA	Bovine Serum Albumin
BVC	Beckman Coulter Life Sciences
CAR	Coxsackie and adenovirus receptor
cDNA	Complementary DNA
CFTR	Cystic fibrosis transmembrane conductance regulator
CO_2	Carbon dioxide
CsCl	Caesium chloride
CPPT	Central polypurine tract
cTnT	Cardiac troponin T
DAVID	Database for annotation, visualization and integrated discovery
dH_2O	Distilled water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNMTs	Deoxyribonucleic acid methyltransferases
dNTPs	Deoxyribonucleoside triphosphate
DPBS	Dulbecco's phosphate buffered saline
DTT	Dithiothreitol

DTT	Dithiothreitol
E. Coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EIAV	Equine infectious anaemia virus
EMA	European medicine agency
ESC	Embryonic stem cells
FBS	Fetal bovine serum
FIV	Feline immunodesfficency virus
FS	Fisher Scientific
gDNA	Genomic deoxyribonucleic acid
GEO	Gene expression omnibus
GFP/eGFP	Green fluoresence protein
H_2O	Water
HCl	Hydrochloric acid
HIV	Human immunodeficiency virus
HLC	Hepatocyte like cells
HPLC	High performance liquid chromatography
HRP	Horseradish peroxidase
HSG	Heparin sulphate glycan
IGFR	Insulin growth factor receptor
IIG	Inactive ingredients database
iPSC	Induced pluripotent stem cell
IS	Insertion site
ITR	Inverted terminal repeats
IVIM	*In vitro* immortalisation assay
K_2HPO_4	Potassium phosphate

KCl	Potassium chloride
KCl	Potassium chloride
LB	Lysogeny broth
LCB	Leica Biosystems
LTR	Long terminal repeat
Luc	Luciferase
LV	Lentivirus
$MgCl_2$	Magnesium chloride
MHC	Major histocompatability complex
MLV	Murine leukaemia virus
MoMLV	Moloney murine leukaemia virus
NaOH	Sodium hydroxide
NEB	New England Biolabs
NGS	Next generation sequencing
NTC	Non template control
OD	Optical density
PACBIO	Pacific Biosciences
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDM	Plasmid DNA medium
pDNA	Plasmid DNA
PEG	Polyethylene glycol
PEI	Polyethyleneimine
PFA	Paraformaldehyde
pTFBS	Predicted transcription factor binding sites
QG	Qiagen
RCL	Replication competent lentivirus
RID	Retrovirus integration database

RNA	Ribonucleic acid
ROI	Regions of interest
RSV	Rous sarcoma virus
RT	Reverse transcriptase
RV	Retrovirus
SA	Sigma Aldrich
SCID	Severe combined immunodeficiency disease
SCT	Stemcell Technologies
SEM	Standard error of the mean
SFFV	Spleen focus-forming virus
SIN	Self-inactivating
SIV	Simian immunodeficiency virus
TBE	Tris-borate EDTA
TBST	Tris buffered saline and Tween 20
TE	Tris- EDTA
TFS	Thermo Fisher Scientific
TRRAP	Transformation/transcription domain-associated protein
Tris-Cl	Tris chloride
UCSC	University of California, Santa Cruz
VSV-G_{Ind}	Vesicular stomatitis virus glycoprotein Indiana serotype
WPRE	Woodchuck Hepatitis Vius Posttranscriptional Regulatory Element

2.1. Materials

Below are lists of reagents used in the experiments reported and their origin of purchase.

2.1.1. General Reagents and Consumables

1 Kb Plus DNA Ladder	TFS
38.5 mL, open-top thinwall ultra-clear tube, 25 x 89 mm	BC
7.5% BSA, in PBS	SA
Absolute ethanol, >99%	FS
Agar	FS
Agarose	FS
Alfa Aesar™ 4% Paraformaldehyde, in PBS	FS
Ammonium chloride	BDH Laboratory Supplies
Ampicillin sodium salt	FS
Bemis™ Parafilm™ M Laboratory Wrapping Film	FS
Bovine serum albumin	NEB
CD1 mice	Charles River
D- Luciferin	Gold Biotechnologies
DH5α/Stbl3 competent cells	FS
Dimethyl Sulphoxide	SA
Di-Potassium hydrogen orthophosphate	FS
Disposable scalpel	FS
Fisherbrand™ comfort nitrile gloves	FS
Fisherbrand™ Sterile PES Syringe Filter	FS
Formaldehyde, 4% in PBS	FS
Galactose	FS
Gel filtration standards	BRL
GeneRuler 100 bp DNA ladder	FS
Glucose	FS
Glycerol, ≥99.5 %	FS
Granulated Miller LB Broth	FS
HCl, 37%	SA

Isoflurane	Abbot Laboratories
Kanamycin sulphate	SA
Macherey-Nagel™ Tritest™ pH Paper Roll	FS
Magnesium chloride, ≥98%	SA
Magnesium sulphate	East Anglia Chemicals
Microcentrifuge Tubes	FS
Motic™ Microscope Lens Paper	FS
Nail varnish	Essence Cosmetics
Peptone	FS
Poly(ethylene glycol) average Mn 600, 950-1,050, 8000	SA
Potassium chloride	FS
Potassium dihydrogen orthophosphate	FS
Potassium ferricyanide (III)	SA
Potassium hexacyanoferrate (II) trihydrate	SA
Propan-2-ol	FS
Quick-Load® 1 kb DNA ladder	NEB
Rectangular cover glasses	FS
RNase-Free DNase Set	QG
RNaseZap™ RNase decontamination solution	FS
S.N.A.P.™ UV-free gel purification kit	TFS
Shandon™ TPX filter cards	FS
Snowcoat microscope slides	LCB
Sodium azide	FS
Sodium chloride	FS
Sodium hydroxide	SA
Sodium phosphate	FSA Life Supplies
Sodium sulphate	FS

Stericup 500ml Durapore® 0.20µm/0.45µm PVDF sterilisation column	Merck Millipore
SYBR® safe DNA gel stain	TFS
T4 DNA ligase	NEB
TBE buffer, 10X	FS
Tris base	FS
Triton-X	SA
TriTrack DNA loading dye, 6X	TFS
Tryptone	FS
TSK-gel 5µm G3000SWXL TSOH 08451 TSK gel column	SA
Tween™ 20	FS
X-Gal powder	FS
Yeast extract	FS
Yeast nitrogen base	FS

2.1.2. Plasmids for Viral Vector Production

All plasmids were provided by Dr Michael Themis, Brunel University London unless stated otherwise (Demaison, Parsley et al. 2002, Ikeda, Takeuchi et al. 2003, Naldini, Blomer et al. 1996). Plasmids were visualised using Serial Cloner (v2.6.1).

LNTSbetagalW	Encodes *LacZ* reporter gene
pCMVR8.74	Encodes HIV1 accessory proteins
HR'SIN-cPPT-SEW	Encodes *GFP* reporter gene with SIN LTR configuration (pHR)

pHV	Encodes *GFP* reporter gene with full LTR configuration. Kindly provided by Dr Yasuhiro Takeuchi, University College London
pMD2.G	Encode VSV-G_{Ind} glycoproteins

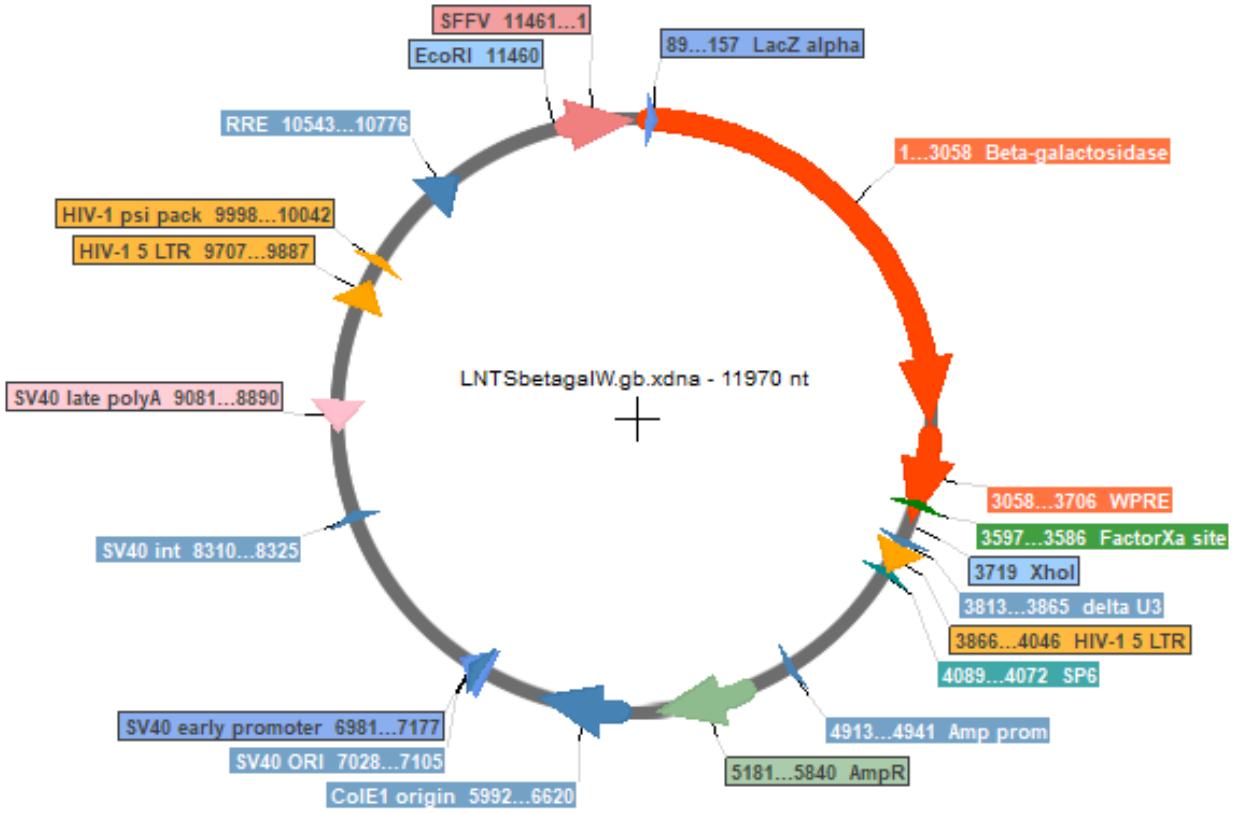


Figure 2.1- LNTSbetagalW plasmid construct. Schematic representation of LNTSbetagalW plasmid containing *β-galactosidase* gene under SFFV promoter. Other gene fragments also shown. 0.5µg LNTSbetagalW pDNA linearised and double cut using *XhoI* and/or *EcoRI* enzymes. *XhoI* and *EcoRI* cut the plasmid at 3,719nt and 11,460nt respectively, generating a 11,460bp fragment. Double cuts by both enzymes generate a 11,460bp and 3719bp fragments. Digest sites shown on schematic map. DNA bands were separated on a 0.6% agarose gel for 40 minutes at 70V, using uncut plasmid as a negative control and sizes compared to a 1 Kb Plus DNA ladder. All band sizes were found to be the correct size determining plasmid sequence integrity

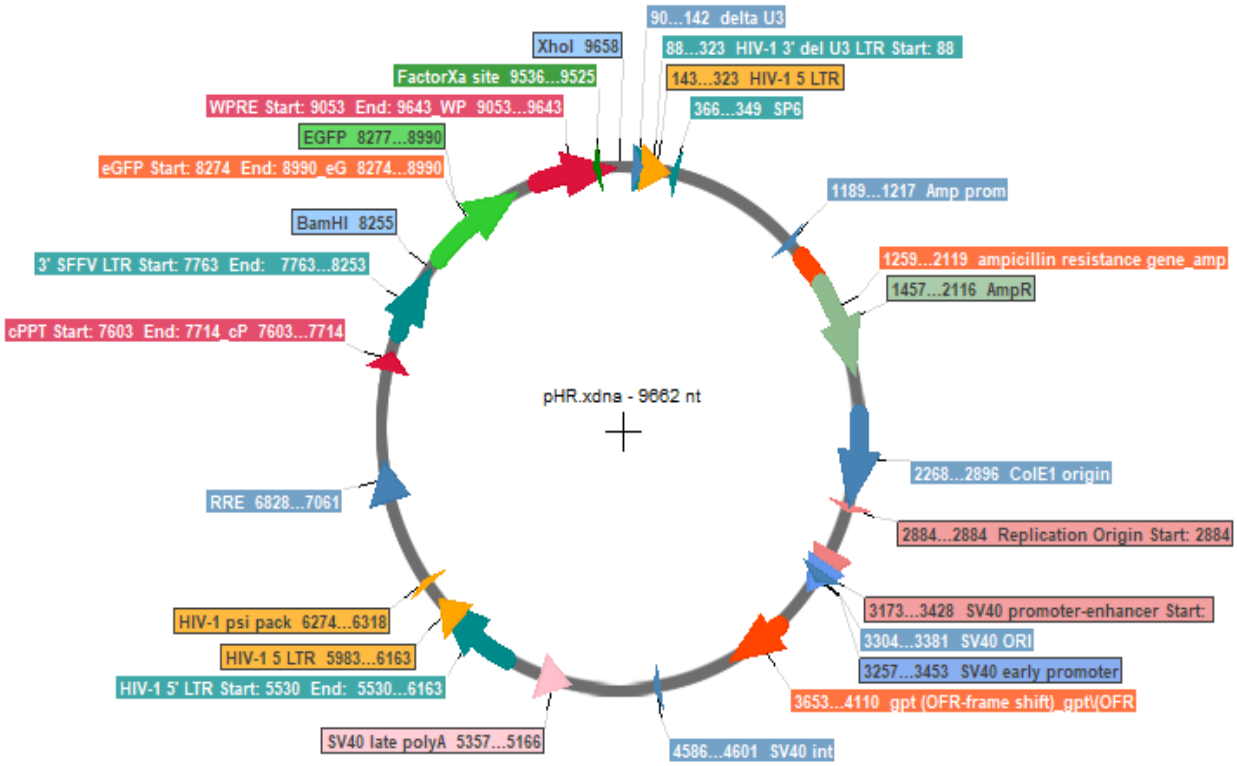


Figure 2.2- pHR plasmid construct. Schematic representation of pHR plasmid containing *eGFP* gene under SFFV promoter flanked by SIN LTRs. Other gene fragments included on plasmid are also shown. pHR pDNA was linearised and double cut using *XhoI* and/or *BamHI* enzymes. *XhoI* and *BamHI* cut at 9,658nt and 8,255nt respectively generating 9,662bp fragments. Double cuts using both enzymes generate 8,259bp and 1,403bp fragments. DNA bands were separated on a 0.6% agarose gel for 40 minutes at 70V, using uncut plasmid as a negative control and sizes compared to a 1 Kb DNA ladder. Cut sites shown on schematic representation. All band sizes were found to be the correct size determining plasmid sequence integrity

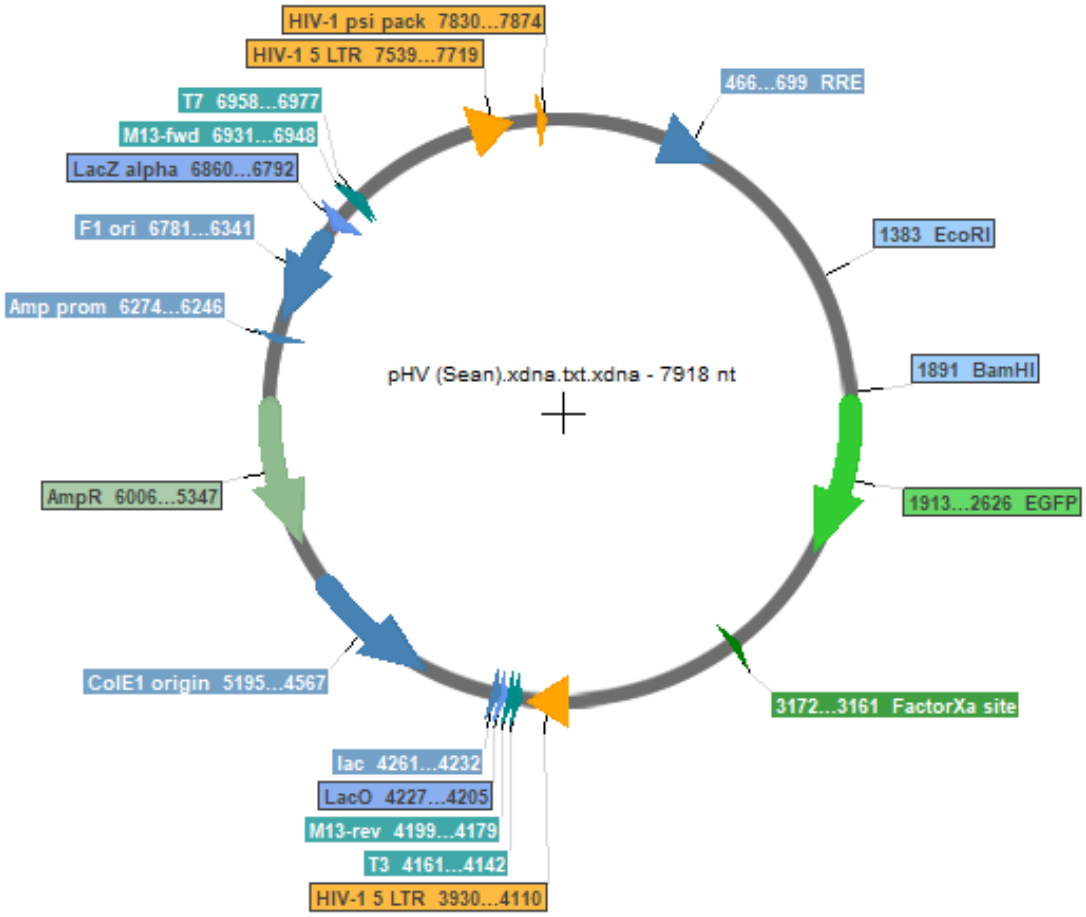


Figure 2.3- pHV plasmid construct. Schematic representation of pHV plasmid containing *eGFP* gene under SFFV promoter flanked by native LTRs. Other gene fragments included on plasmid are also shown. pHR pDNA was linearised and double cut using *EcoRI* and/or *BamHI* enzymes. *XhoI* and *BamHI* cut at 1,383nt and 1,891nt respectively generating 7,918bp fragments. Double cuts using both enzymes generate 7,410bp and 508bp fragments. DNA bands were separated on a 0.6% agarose gel for 40 minutes at 70V, using uncut plasmid as a negative control and sizes compared to a 1 Kb DNA ladder. Cut sites shown on schematic representation. All band sizes were found to be the correct size determining plasmid sequence integrity

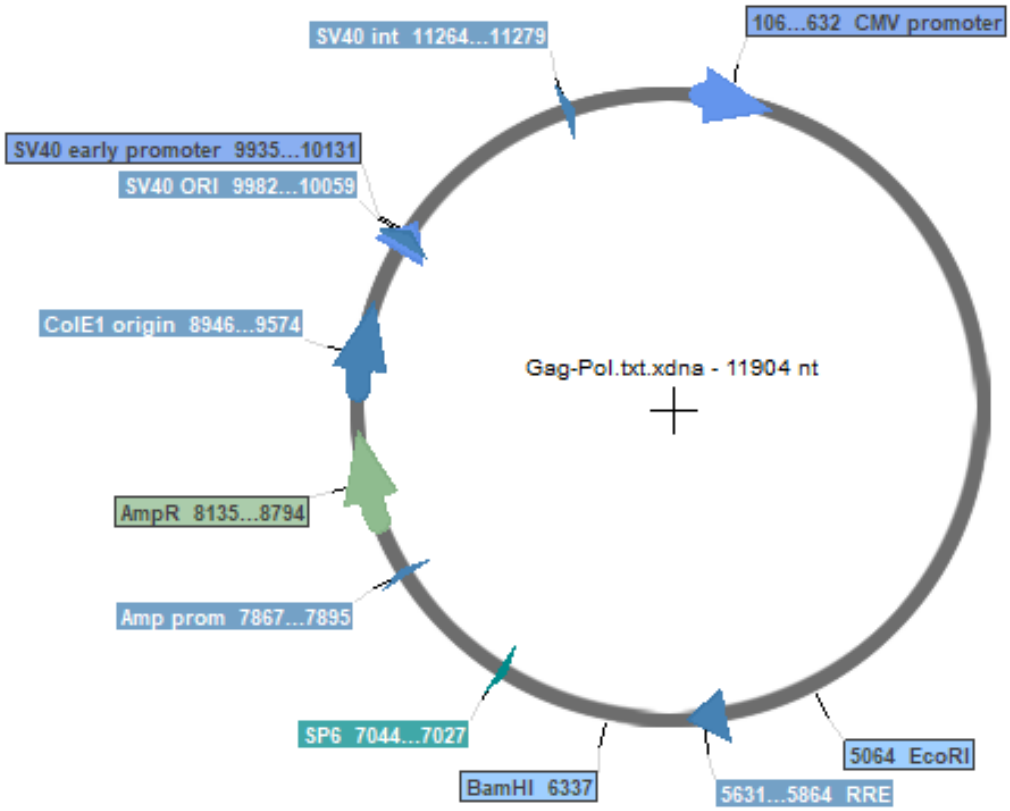


Figure 2.4- pCMVR8.74 plasmid construct. Schematic representation of pCMVR8.74 plasmid. Plasmid contains *Gag* and *Pol* genes under CMV promoter. Other genes sequences shown on schematic. pCMVR8.74 pDNA was linearised and double cut using *EcoRI* and/or *BamHI* enzymes. DNA bands were separated on a 0.6% agarose gel for 40 minutes at 70V, using uncut plasmid as a negative control and sizes compared to a 1 Kb Plus DNA ladder. Digest sites shown on schematic. All band sizes were found to be the correct size determining plasmid sequence integrity

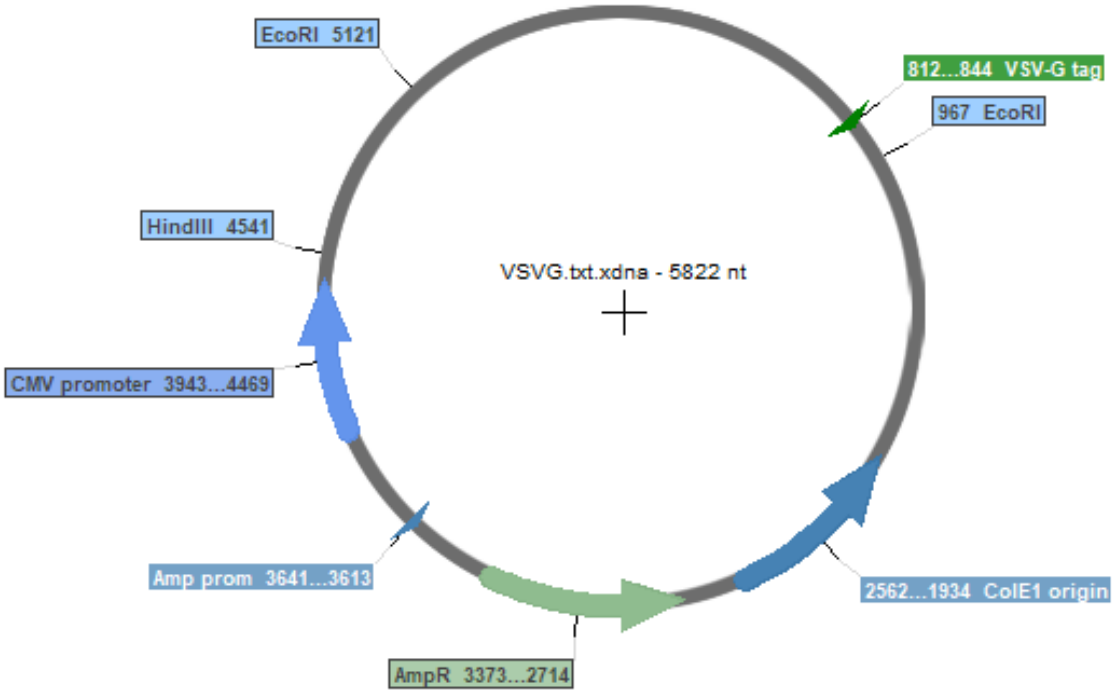


Figure 2.5- pMD2.G plasmid construct. Schematic representation of pMD2.G, with *VSV-G* under CMV promoter. Other genes also shown on schematic. pMD2.G pDNA was linearised and double cut using *HindIII* or *EcoRI* enzymes *HindIII* cuts at 4541nt generating a 5822bp fragment. *EcoRI* cuts at 5121nt and 961nt generating 4,154 and 1668bp fragments. Sample DNAs were separated on a 0.6% agarose gel for 40 minutes at 70V, either as uncut plasmid controls or digested using restriction enzymes run with a 1 Kb DNA ladder. Digest sites shown on schematic map. Linearised plasmids were found correct in size as shown above and the VSV-G insert was digested out of the plasmid, with the correct size of fragments determined

2.1.3. Restriction Enzymes

All enzymes and buffers were purchased from NEB.

BamHI

CutSmart® buffer

EcoRI

HindIII

XhoI

2.1.4. Commercially Available Kits

DNeasy® blood & tissue	QG
GeneJET plasmid miniprep	TFS
GoScript™ reverse transcriptase	Promega
Lenti-X™ p24 rapid titer	Takara Bio
PureLink™ expi endotoxin-free mega plasmid purification	TFS
QIAamp® viral RNA mini	QG
RNeasy® mini	QG
S.N.A.P™ UV free gel purification	FS
STEMdiff™ cardiomyocyte differentiation	SCT
STEMdiff™ cardiomyocyte dissociation	SCT
Taq PCR master mix	QG
TOPO TA cloning kit, with One Shot™ TOP10 E. coli	FS

2.1.5. Tissue Culture Reagents and Consumables

CellAdhere™ Laminin-521	SCT
Corning® Matrigel® hESC-Qualified Matrix, LDEV-free	Corning Incorporated
Countess™ cell counting chamber slides	FS
Dimethyl sulfoxide	SA
Dulbecco's phosphate buffered saline, without calcium chloride and magnesium chloride	SA
Falcon™ disposable aspirating pipets	FS
Fetal bovine serum	FS
Fisherbrand™ Cell Scrapers	FS
Fisherbrand™ Easy Reader™ conical polypropylene centrifuge tubes	FS
Fisherbrand™ sterile polystyrene disposable serological pipets with magnifier stripe	FS
Fisherbrand™ SureOne™ aerosol barrier pipette tips	FS
Gentle cell dissociation reagent	SCT
Gibco™ DMEM, high glucose, GlutaMAX™ Supplement, pyruvate	FS
Laminin- 521	SCT
mTeSR™1 kit	SCT
Nunc™ Biobanking and cell culture cryogenic tubes	FS
Nunc™ cell-culture Treated multidishes	FS
Nunc™ EasYFlask™ cell culture flasks	FS
Opti-MEM™ I reduced serum medium, no phenol red	FS
Penicillin-Streptomycin (10,000 U/mL)	FS
Phosphate buffered saline, 1X, pH 7.4	FS

ReLeSR™	SCT
Trypan blue stain (0.4%)	FS
TrypLE™ express enzyme (1X), no phenol red	FS
Y-27632	Cambridge Bioscience

2.1.6. Transfection and Transduction Reagents

FuGENE® 6	Promega
GeneJuice®	Merck Millipore
Nuvec® (NV00100028, NV00100026-28, NV00100032 & NV00100033)	N4Pharma
Polybrene	SA
Polyethylenimine, branched, average Mw ~25,000	SA

2.1.7. Antibodies

Alexa Flour 488 signal amplification kit, anti-mouse	FS
Cardiac troponin T monoclonal antibody (13-11)	FS
CD15	Biolegend
SSEA4	Biolegend
TRA-1-60 (Podocalyxin) monoclonal, PE	FS
TRA-1-81 (Podocalyxin) monoclonal, APC	FS
VECTASHIELD® Vibrance™ antifade mounting medium with DAPI	Vector labs

2.1.8. Mammalian Cell Lines

All cell lines used were provided by Dr Michael Themis, Brunel University London

HepG2 (Harjumaki, Nugroho et al. 2020)
HEK293T (Rio, Clark et al. 1985, DuBridge, Tang et al. 1987)
JHU106i (Chou, Gu et al. 2015)
TELCeB6 (Duisit, Salvetti et al. 1999, Cosset, Takeuchi et al. 1995)

2.1.9. Buffer and solutions

Blocking buffer

5% BSA in 100 ml PBS

Solution was made fresh each time and stored at 4°C until further use.

Fixative solution

In 100 ml PBS

10 ml 4% PFA
90 ml PBS

Fixative solution made fresh prior to use and kept at 4ºC until use.

Glycerol solution

50% (w/w) glycerol solution was prepared diluting appropriate concentration of glycerol powder in dH_2O. Solution was stored at room temperature for future use.

HPLC Buffer A

In 500 ml 0.1M phosphate buffer:

1% PEG 600

0.05% sodium azide

0.1M sodium sulphate

Solution was filtered using a 0.45µm unit and stored at 4°C until use.

LB medium

10g LB granules were added to 500ml dH_2O before sterilisation. Alternatively, LB medium was prepared using the following formula:

5 g peptone

2.5 g yeast extract

2.5 g sodium chloride

After sterilisation, the desired antibiotic was added to correct concentrations. Medium was kept in a dark, cool dry place until needed.

Permeabilisation buffer

0.2% Triton X in 100 ml PBS

Solution was made fresh before use and stored at 4°C until needed.

PDM

In 500 ml dH_2O:

2.21 g yeast extract
3.97 g tryptone
5 g glucose
6.40 g sodium phosphate
1.5 g potassium phosphate
0.25 g ammonium chloride
0.12 g magnesium sulphate

Medium was sterilised through a 0.45µm filter and appropriate antibiotic added. Solution was stored at 4°C for future use.

TBST, pH 7.4

In 1 L dH_2O:

8.8 g sodium chloride
0.2 g potassium chloride
3 g tris base
500 µl tween 20

pH was measured and adjusted using 1M HCL or NaOH until a reading of pH 7.4 was obtained. Solution was stored at room temperature until needed.

X Gal staining solution

The following salts were dissolved in 100 ml PBS:

4 mM potassium ferrocyanide (III)
4 mM potassium hexacyanoferrate (II) trihydrate
0.1 mM magnesium chloride

A working staining solution was prepared by diluting 4 ml of the above solution with 400 µl of 0.4 mg/ml X-Gal, dissolved in DMSO. Working staining solution was made prior to use and stored at 4°C, protected from sight.

2.1.10. Oligonucleotides

The oligonucleotides listed below were used in PCR amplification of specific DNA regions to determine genetic integrity. All oligonucleotides were purchased from TFS.

Primer name	**Sequence (5'-3')**	**Tm (°C)**
GAAB-F	AATGGATTTCCTGGCAGGACGC	66.2
GAAB-R	GCATTGGGCGATCTTGGCTTAA	64.1
LTR-F	GAGCTCTCTGGCTAACTAGG	60.2
LTR-R	GCTAGAGATTTTCCACACTG	55.9
SY100216195-080	AAGAGAGGCATCCTCACCCT	64.1
SY100216195-081	TACATGGCTGGGGTGTTGAA	66.3

2.1.11. Viral vectors

First generation adenoviral vector	AdRSVβ was used for *in vitro* studies (Castro, Goya et al. 1997). This is a first-generation adenoviral vector based on human Ad type 5. This was originally provided by A. Pavarani, Transgene, France. The vector contains the E. coli *LacZ* reporter gene bound to the SV40 nuclear localisation signal under the control of RSV promoter.

	Ad-Luc is similarly a first- generation human adenovirus type 5 vector, carrying the *Luciferase* (*Luc*) transgene (Hogg, Garcia et al. 2010). This was kindly provided by S. Waddington, University College London and used for *in vivo* studies.
pHR and pHV lentiviral vectors	LV vectors carrying *eGFP* transgene, under SFFV promoter were generated as described in Section 2.2.2.10. LVs are identical except the LTR configuration. pHV LV carries an *eGFP* transgene flanked by native LTRs (Strang, Ikeda et al. 2004). pHR LV carries an *eGFP* transgene flanked by SIN LTR configurations (Demaison, Parsley et al. 2002).

2.1.12. Equipment and Software

Biorad molecular imager gel doc XRS	BRL
Countess™ automated cell counter	FS
DAVID v6.7.	(Huang da, Sherman et al. 2009a, Huang da, Sherman et al. 2009b)
DM400 microscope	LCB
DMi inverted microscope	LCB
Elx808 absorbance reader	BioTek
Environmental shaker10X 400	Gallenkamp
Fisherbrand™ GT1 benchtop Centrifuge	FS
FLoid™ cell Imaging Station	FS

FunRich v3.1.3.	(Pathan, Keerthikumar et al. 2015, Pathan, Keerthikumar et al. 2017)
Gen5 v2.06.10	BioTek
Genevenn	http://genevenn.sourceforge.net/
Image Lab v5.2.1	BRL
IVIS machine	Perkin Elmer
Leica application suite advanced flourescence v4.0.0.11706	LCB
Living image software	Perkin Elmer
MJ research PTC-240 DNA engine TETRAD2 peltier thermal cycler	Conquer Scientific
Multi experiment viewer v4.9.0	(Saeed, Sharov et al. 2003)
Nanodrop 2000C UV-Vis v1.2.1	FS
NovoCyte flow cytometer	ACEA Biosciences Incorporated
NovoExpress software v1.2.5	ACEA Biosciences Incorporated
Opossum v3.0	(Kwon, Arenillas et al. 2012, Ho Sui, Fulton et al. 2007)
Optima XPN	BC
pH reader	Hanna Instruments
Pulse video analysis, v3.0	(Maddah, Heidmann et al. 2015)
SCL-10A WP system controller	Shimadzu Corporation
Serial cloner v2.6.1	http://serialbasics.free.fr/Serial_Cloner.html
Shandon cytospin 2	Marshall Scientific
Sigma laboratories centrifuge 6K15	SA
SW32Ti rotor	BC
UCSC genome browser	https://genome.ucsc.edu/
Uniprot	(UniProt Consortium 2019)
Waters 2695 system	LabX

Zetasizer Nano S	Malvern Panalytical

2.2. Methods

2.2.1. Molecular Biology Techniques

2.2.1.1. PDM Preparation

0.5 L PDM was created as mentioned in Section 2.1.9 (Danquah, Forde 2007). Stock ampicillin antibiotic was added to generate a 100 µg/ml solution. PDM was stored at 4°C for future use.

2.2.1.2. LB- Agar Plate Preparation

LB medium was prepared as described in Section 2.1.9. with 7.5 g agar powder added before sterilisation. Once the solution had cooled slightly, the desired antibiotic was added (100µg/ml) and mixed well before pouring into plates to set overnight. LB- antibiotic plates were stored at 4°C until needed.

2.2.1.3. Ampicillin Antibiotic Preparation

Ampicillin (100 mg/ml) stock solution were prepared by diluting salts in 5 ml dH2O before diluting with 5 ml 100% ethanol. The solution was vortexed until dissolved and sterile filtered through a 0.22 µm filter before storage at -20°C for further use.

2.2.1.4. Plasmid Transformation

All plasmids were provided as glycerol stocks of transformed cells containing appropriate plasmids transformed in DH5α or Stbl3 *E. Coli* cells. pHV was provided as pDNA by Dr. Y. Takeuchi. pHV pDNA was transformed into DH5α cells by incubating DH5α *E. Coli* cells with 100 ng pDNA on ice for 30 minutes. Cells were subsequently heat shocked at 42°C for 45 seconds before chilling on ice for 2 minutes. Cells were added to a suspension culture containing 1ml PDM, plated on Amp^+ plates and grown overnight at 37°C. A single colony was picked the following day and added to a suspension culture of PDM and grown at 37°C, shaking at 200 rpm overnight before growing up cultures as appropriate for plasmid purification.

2.2.1.5. Plasmid Glycerol Stock Preparation

Glycerol stocks of bacteria containing plasmids were prepared first by diluting 100 µl bacterial stocks with 100 µl 50% (w/w) glycerol solution diluted in dH2O. Stocks were frozen at -80°C for further use.

2.2.1.6. Plasmid Harvesting

Suspension cultures of bacterial cells containing required plasmids purified via anion exchange chromatography using PureLink™ expi endotoxin-free mega plasmid purification kit, according to manufacturer's instructions. Briefly, glycerol stocks of bacterial cultures were initially grown 10ml suspension cultures of PDM at 37°C overnight at 200 rpm. The following day, bacterial cultures were grown up in 500 ml PDM for a further 48 hours at 37°C, rotating at 140 rpm. Bacterial cells were pelleted through centrifugation at 1500 rpm for

5 minutes at 4°C. Pellets were resuspended in appropriate buffer containing RNAase A and homogenised. Cells underwent lysis and precipitation using appropriate buffers before binding to filtration column. Endotoxins were removed from the clarified lysate using appropriate buffers and further purified through a chromatographic column. pDNA was eluted in buffer provided and precipitated using isopropanol and ethanol via centrifugation at 1500 rpm, 4°C for 30 minutes. pDNA pellet was air dried and resuspending in TE buffer before storage at -80°C for future use. Plasmid concentration and purity was measured through OD readings using a NanoDrop2000 UV-Vis Spectrophotometer.

2.2.1.7. Restriction Digestion

Restriction enzyme digest was run to verify the integrity of plasmid constructs in 10µl reactions using the following reagents. Samples were incubated at 37°C for 1 hour to ensure DNA digestion:

1 µl Cutsmart® buffer
1µl restriction enzyme
500 ng DNA
dH_2O

2.2.1.8. Agarose Gel Electrophoresis

0.6%, 1% and 2% agarose gels were prepared agarose in 100ml 1X TBE for five minutes. 1 µl 10,000X SYBR® safe DNA gel stain was added to the cooled solution before casting the gel. Nucleic acid samples were resuspended with TriTrack dye before loading in wells alongside a DNA ladder. Gels were run at 70V for 35-45 minutes to separate bands fully. Gels were imaged on Biorad molecular imager gel doc XRS and analysed via ImageLab software.

2.2.1.9. PEI Preparation

10 mM PEI was prepared by diluting 10.3 ml PEI in 10ml ultrapure dH_2O. This solution was titrated to a neutral pH using HCl and NaOH and brought up to a final volume of 41.2 ml with dH_2O. Aliquots were frozen at -80°C for future use.

2.2.1.10. pH Analysis

pH of solutions was determined using a pH reader. Solutions were titrated to correct pH using HCl and NaOH solutions. The pH probe was washed well in dH_2O afer reading each sample.

Alternatively, samples were analysed using Macherey-Nagel™ Tritest™ pH paper roll and compared to stock colours provided.

2.2.1.11. P24 Gag Titration

Viruses were titrated using Lenti-X™ p24 Rapid Titer Kit, according to manufacturer's instructions. Briefly, buffers and standards were prepared by diluting reagents in deionised water and tissue culture medium respectively. Lysis buffer was added to ELISA coated wells provided with standards and biological samples aliquoted into appropriate wells, in duplicate. After incubation at 37°C for one hour, the wells were washed in and overlaid with Anti- P25 biotin conjugate. This was washed of after incubation and layered with streptavidin-HRP conjugate. After incubation, this was washed to remove excess background and substrate solution added for 30 minutes. Stop solution

was added to each well and absorbance read on Elx808 absorbance reader at 450nm and analysed using Gen5 software.

2.2.1.12. DNA Purification

Cellular gDNA was purified using DNeasy Mini kit according to manufacturer's instructions. Briefly, up to $5x10^6$ cells were used for gDNA purification. Cells were resuspended in 200 µl PBS, with protinase K (1:10). Samples were homogenised and incubated at 56ºC for 10 minutes. Nucleic acids were precipitated using ethanol and loaded on the column provided. Samples were washed as appropriate, and DNA eluted in buffer provided. Nucleic acid concentration and purity was measured through optical density readings using a NanoDrop2000 UV-Vis Spectrophotometer. Samples were stored at -80°C for future use.

2.2.1.13. Cellular RNA Purification

Cellular RNA was purified using RNeasy Mini kit according to manufacturer's instructions. Up to $1x10^7$ cells were detached from substratum and pelleted via centrifugation at 1500 rpm for 5 minutes. The cell pellet was resuspended in 350µl buffer RLT and homogenised via vortexing and manual pipetting before adding an equal volume of 70% ethanol. The sample was passed through the column provided, washed, and eluted in appropriate buffers. Nucleic acid concentration and purity was measured through OD readings using a NanoDrop2000 UV-Vis Spectrophotometer. Samples were aliquoted and stored at -80°C for future use.

2.2.1.14. Viral RNA Purification

LV RNA was purified for further analysis using QIAamp Viral RNA Mini kit according to manufacturer's instructions. LV samples were resuspended in PBS and added to lysis buffer containing carrier RNA. Samples were pulse vortexed and incubated for 10 minutes at room temperature. 100% ethanol was added to the sample before binding to the column provided through centrifugation. The column was washed using appropriate buffers and on column DNA digestion was performed, as described in Section 2.2.1.15. RNA was eluted in RNAase free H_2O provided. Nucleic acid concentration and purity was analysed using NanoDrop2000 UV-Vis Spectrophotometer.

2.2.1.15. DNase Treatment of RNA

Rnase- free DNase set was prepared by resuspending lyophilised DNase I in 500 µl RNAase free water provided (2.7 Kunitz units/µl). Aliquots were frozen at -20°C for future use. 10 µl resuspended DNase I was added to 70 µl buffer RDD before aliquoting directly onto RNeasy column. Column was incubated at room temperature for 15 minutes before washing with appropriate buffers to remove excess DNase. RNA was eluted in RNAase free water provided and frozen at -80°C for future use.

2.2.1.16. cDNA Synthesis

Reverse transcription of RNA was performed using GoScript™ Reverse Transcription System according to manufacturer's instructions. Briefly, first strand cDNA synthesis was carried out using purified RNA, Oligo$(dT)_{15}$ primer and nuclease free water in 5 µl reactions. RNA samples were heat shocked at

70°C for 5 minutes and chilled on ice for a minimum of 5 minutes. RNA mix was added to a master mix containing GoScript™ 5x reaction buffer, 25mM $MgCl_2$, 0.5mM PCR nucleotide mix, 1.0 µl GoScript™ reverse transcriptase and water in 20 µl reactions. Samples underwent cDNA synthesis in a thermocycler under the following programme.

Stage	**Temperature (°C)**	**Duration (minutes)**
Anneal	25	5
Extension	42	60
Enzyme inactivation	70	15

2.2.1.17. PCR

Lyophilised primers were hydrolysed in appropriate volumes of dH_2O to create 10mM stock solutions, which was further diluted to prepare 10µM working solutions for PCR reactions. Nucleic acid samples were added to a 25 µl reactions containing 2x Taq master mix containing Taq DNA polymerase (5U/µl in 20mM Tris-Cl, 100mM KCl, 1mM DTT, 0.1mM EDTA, 0.5% (v/v) Nonidet® P-40, 0.5% (v/v) Tween® 20, 50% (v/v) glycerol, pH 8.0) in 10X buffer (Tris·Cl, KCl, $(NH_4)_2SO_4$, 400mM dNTPs, pH 8.7), 0.5 µl 10µM forward and reverse primers and dH2O. Primers are described in Section 2.1.10. Samples were amplified for specific regions in a thermocycler using the following programme.

Cycles	Stage	Temperature (ºC)	Time
1x	Initial denature	94ºC	2 minutes
10x	Denature	94ºC	10 s
	Anneal	Primer specific temperature	30 s
	Extension	68ºC	45 s
20x	Denature	94ºC	10 s
	Anneal	Primer specific temperature	30 s
	Extension	68ºC	60 s, increase by 20 s increments per cycle
1x	Final extension	68ºC	10 minutes

10 µl per sample were run on agarose gel electrophoresis as described in Section 2.2.1.8.

2.2.1.18. DNA Purification from Agarose Gel

DNA bands were separated on agarose gel electrophoresis as described in Section 2.2.1.8. Samples were excised and purified using S.N.A.P.™ UV- free gel purification kit, according to manufacturer's instructions. Gel fragment was dissolved in 2.5x volume sodium iodide solution, at 50°C, vortexing at regular intervals. After the gel dissolved, the solution was mixed with 1.5x volume binding buffer and loaded on the column provided through centrifugation. The column was washed with appropriate buffers and DNA fragments eluted in TE buffer.

2.2.1.19. Nuvec® Preparation

Lyophilised Nuvec® was prepared followed the manufacturer's instructions. Briefly, 1 µg/µl solutions of Nuvec® were reconstituted in PBS. Nuvec® was washed in 2 ml 70% ethanol and pelleted by centrifugation at 1500 rpm for 5 minutes. Nuvec® was resuspended in appropriate volumes of PBS and sonicated for 30 minutes at room temperature, dispersing sediment through pipetting.

Nuvec production batches provided included:

NV00100028
NV00100026-28
NV00100032
NV00100033

2.2.1.20. Nucleic Acid Loading Capacity of Nuvec®

1 mg/ml stock concentrations of Nuvec® were prepared as described in Section 2.2.1.19. 1 µg pDNA carrying the *eGFP* gene was resuspended with 5 µg Nuvec®. A negative control was prepared by adding 5 µg Nuvec® to 5 µl PBS. As a positive control, 1 µg pDNA was resuspended in a total of 10 µl with PBS. Solutions were incubated at 4°C for four hours before centrifugation at 12,000 rpm for 10 minutes. Supernatant was measured for nucleic acid concentration using a Nanodrop 2000c UV-Vis Spectrophotometer. The concentration of pDNA bound to 5 µg Nuvec® was calculated by correcting samples for background and multiplying out by a factor of 10.

2.2.1.21. Generation of ATPS phases

Polymers and salts were weighed and dissolved in deionised water at 37°C, 1.0MΩcm^{-1} (Purite, Thames, Oxfordshire) to generate systems below. Each ATPS was vortexed and left to separate for 20 mins at room temperature.

PEG 1000 - K_2HPO_4 (14%-14% w/w)
PEG 8000 - K_2HPO_4 (14% - 14% w/w)
PEG 600 - $(NH_4)_2SO_4$ (20% - 20% w/w)

2.2.1.22. Size Exclusion Chromatography

HPLC separations of Ad samples were performed using a Waters 2695 system using a SCL-10A WP system controller. Mobile phases were run in HPLC Buffer A at room temperature, 0.5 ml/min flow rate, 6.89 pressure limit using TSK-Gel 5 µm G3000SWXL TSOH 08451 column at 260nm and 280nm wavelengths. The column was equilibrated with buffer and 50µl of samples run for ~40 minutes. Various samples were run on diluted in dH_2O at a 1:2 ratio, including Ad purified in upper phase of PEG600/$(NH_4)_2$SO4, Ad purified in lower phase of PEG600/$(NH_4)_2$SO4, PEG600/$(NH_4)_2$SO4 upper and lower phase, pure Ad previously purified through standard CsCl ultracentrifugation (1:5) and HEK293T lysate containing Ad supernatant (1:40). Protein gel filtration standards were assayed to ensure calibration of column, using according to manufacturer's instructions. Column was flushed using HPLC Buffer A after each run. Column efficiencies were measured and compared to spectrophotometry readings. Readings were found to be diluted two times, allowing correction of data.

2.2.2. Tissue culture techniques

2.2.2.1. Immortalised Cell Culture Media Preparation

Complete medium for immortalised cell culture was prepared containing 10% FBS and 100U/ml Penicillin Streptomycin in 500 ml Gibco™ DMEM high glucose GlutaMAX™ Supplement, pyruvate.

2.2.2.2. Immortalised Cell Culture

Immortalised, mammalian cancer cells (HEK293T, HepG2, TelCeb6) were thawed from liquid nitrogen and resuspended in a total of 5ml complete medium. Cells were pelleted by centrifugation at 1500 rpm for five minutes and plated in 10 ml fresh complete medium. Cells were incubated at 37°C, with 5% CO_2. Cells were passaged regularly when 70-80% by aspirating complete medium and washing cells in PBS to excess remove serum. Cells were detached by incubating with TrypLE™ Express enzyme for 5 minutes at 37ºC. Cells were pelleted by centrifugation at 150 0rpm for 5 minutes in fresh medium and resuspended in various volumes of complete medium to passage between cell culture vessels.

2.2.2.3. Cell Viability

Cells were counted to quantify cell number and viability assessed using a trypan blue exclusion assay and Countess™ Automated Cell Counter, according to manufacturer's instructions. Briefly, a 1:1 solution of cell sample and 0.4% (w/w) trypan blue solution was loaded in chamber slides and quantified using the automated cell counter.

2.2.2.4. Cryopreserving Mammalian Cells

Mammalian cells were cryopreserved in the vapour phase of liquid nitrogen in 1ml aliquots in freezing medium (10% FBS, 10% DMSO). Cells were cryopreserved for future use.

2.2.2.5. β Galactosidase Staining

X Gal staining solution and fixative solution was prepared as described in Section 2.1.9. Transduced cells were overlaid with fixative solution for 5 minutes at room temperature. Cells were washed twice with PBS to remove excess PFA and staining solution added. Cells were stored in a cool, dark place for 24 hours before quantification of positive (blue) cells.

2.2.2.6. AV Production

1.5×10^7 HEK293T were seeded per T175 flask and incubated at 37°C, 5% CO_2 overnight to ensure 70-80% confluency after 24 hours. Cells were transduced with stock AV using a MOI 100-200 in fresh complete medium. Cells were incubated for 72 hours at 37°C, 5% CO_2. Cells were mechanically sheared and detached from the substratum and pelted at 1500rpm for 10 minutes. Cell pellet was resuspended in fresh complete medium and lysed through three cycles of freeze thawing. Cell fragments were pelleted at 2000 rpm for 5 minutes and supernatant collected.

2.2.2.7. AV Genome Quantification

Genome titre of AV particles (vg/ml) were calculated using absorbance readings using Nanodrop 2000C UV-Vis, using the following calculation, as previously described (Mittereder, March et al. 1996):

Abs_{260} X 1.1 X 10^{12}

2.2.2.8. ATPS Purification of Viral Vectors

1ml of the upper and lower phase of each solution were combined and warmed at 37°C. Viral aliquots were added to ATPS and centrifuged at 3000 rpm for 15 minutes.

2.2.2.9. AV Transduction post ATPS Purification

$2x10^5$ HepG2 cells were seeded in a 12 well plate and incubated at 37ºC, 5% CO_2 overnight to allow attachment. 10µl of the upper and lower phase after AV purification were added HepG2 cells in complete medium. Cells were incubated at 37ºC, 5% CO_2 for 48 hours, with medium changed after 24 hours. After 48 hours cells were analysed for reporter gene expression (blue) using the β galactosidase assay, as described in Section 2.2.2.5.

2.2.2.10. LV Production

Generation of replication deficient HIV-1 based LV vectors was performed by seeding $1.5x10^7$ HEK293T per T175 flask and incubating at 37°C, 5% CO_2 overnight. Cells were transfected using the following concentrations of plasmids in Opti-MEM™. The concentration of plasmids and transfection time are dependent of transfection reagent used.

Transfection reagent	**PEI**	**FuGENE® 6**	**GeneJuice®**
Reagent volume	1µl	4µl	66µl
Transgene plasmid	40µg	6µg	16µg
pCMVR8.74	30µg	4µg	12µg
pMD2.G	10µg	4µg	4µg
Transfection time	4hr	24hr	24hr

Complete medium was replaced with complete medium after initial transfection time period and harvested daily over 72 hours. Conditioned medium was filtered through 0.45µM filters or pelleted at 1500rpm for 5 minutes to remove cell debris.

Total conditioned medium was concentrated via ultracentrifugation at 23,000 rpm at 4°C, using an SW32Ti rotor using Optima XPN. Alternatively, viral vectors were pelleted by high-speed centrifugation at 5000 rcf (4800g) for 17 hours at 4°C. Viral pellet was resuspended in 200µl Opti-MEM™ and stored at -80°C for future use.

2.2.2.11. Viral Vector Infectious Titration

Infectious titre (TU/ml) of viral vectors were performed as previously described (Gay, Moreau et al. 2012). 2 x 10^5 HEK293T cells were seeded in a 12 well plate and incubated at 37°C, 5% CO^2 overnight to adhere. Serial dilutions of crude or concentrated virus were prepared and incubated in complete cell culture medium with 5µg/ml polybrene, for 20 minutes at room temperature before transduction. The quantity of cells in one well were determined, as described in Section 2.2.2.3. for representation of the average number of cells infected. Medium was replaced after 24 hours and incubated for a further 48 hours before flow cytometry analysis using ACEA Novocyte flow cytometer. Dilutions expressing 1-30% GFP expression were analysed as accurate representations of viral titre. Virus titre was calculated at each dilution point as shown below and averaged for virus titre.

Titre (TU/ml) = ((Cell count * (Percentage GFP expression/100))/Volume) * DF

2.2.2.12. Gene Rescue Assay

LV vectors produced are replication deficient. However, LVs produced from cells transfected for greater than 3 days must be analysed for replication competent lentiviruses. HEK293T cells were transduced with pHR LV at MOI 10 for 24 hours before medium replacement. Cells were serially passaged over 14 days and conditioned medium from infected cells were collected. Conditioned medium was added to fresh HEK293T cells at a 50% (v/v) ratio, with 5µg/ml polybrene and incubated at room temperature for 20 minutes. Cells were incubated for 72 hours before flow cytometry to detect GFP expression.

2.2.2.13. Whole Cell PCR

1.4 x 10^4- 2.8 x 10^3 HEK293T cells were added to Taq PCR master mix as described in Section 2.2.1.17 using primers against β- actin (SY100216195-080, SY100216195-081) (Mercier et al., 1990). Thermocycler conditions were under the following programme:

Cycles	Stage	Temperature (ºC)	Time
3x	Initial denature	94ºC	3 minutes
		55°C	3 minutes
10x	Denature	94ºC	10s
	Anneal	59°C	30s
	Extension	68ºC	45s
20x	Denature	94ºC	10s
	Anneal	59°C	30s
	Extension	68ºC	60s, increase by 20s increments per cycle
1x	Final extension	68ºC	10 minutes

2.2.2.14. mTeSR™ 1 Preparation

mTeSR™1 medium was prepared for JHU106i hiPSC growth by addition of *5X* supplement to basal medium. Medium was aliquoted and stored at 4°C for further use.

2.2.2.15. Laminin 521 Plate Preparation

6 and 12 well tissue culture treated plates were coated in 2.5-5µg/well laminin-521 as a matrix for stem cell attachment, according to manufacturer's instructions in DPBS, without calcium chloride and magnesium chloride. Laminin- 521 coated plates were sealed and stored at 4°C for further use. In preparation for stem cell plating, laminin-521 coated plates were warmed at 37°C for 20 minutes.

2.2.2.16. JHU106i Cell Culture

JHU106i (P106) cells are a human iPS cells line derived from the blood of a 28-year old Caucasian male, reprogrammed using episomal vectors (Chou, Gu et al. 2015). These hiPSC were grown in mTeSR™ 1 -medium and passaged regularly when 70-80% confluent. Morphologically differentiated cells were manually cleared through aspiration. Cells were passaged by washing cells in DPBS. Cells were incubated in gentle cell dissociation reagent for 6 minutes at 37°C before aspiration of medium. Cells were manually resuspended in *mTeSR*™ 1 medium and diluted between laminin-521 coated wells.

Alternatively, cells were detached using ReLeSR™ to passage pure colonies of stem cells. 70-80% confluent P106 cells were washed with DPBS and subsequently washed using ReLeSR™. Medium was aspirated and cells incubated for 6 minutes at 37°C. 1ml of complete medium was added to cells before manual shearing from plate to selectively detach pluripotent stem cells and passaging.

2.2.2.17. Cardiac Differentiation

JHU106i cells were differentiated into cardiomyocytes using STEMdiff™ cardiomyocyte differentiation kit, according to manufacturer's instructions. STEMdiff™ cardiomyocyte differentiation supplement A, B and C were diluted 1:10 using STEMdiff™ cardiomyocyte differentiation basal medium. STEMdiff™ cardiomyocyte maintenance supplement was diluted 1:50 in using STEMdiff™ cardiomyocyte maintenance medium. Mediums were stored at 4°C for future use. Per confluent well, serial additions of differentiation basal medium A (supplemented with 1:100 Matrigel®) and B were added for 48 hours, with differentiation basal medium C added to cells for 96 hours. Complete maintenance medium was then added to cells for a minimum of 168 hours to derive terminal differentiation to cardiomyocytes in 15 days. Maintenance medium was replaced every 48 hours.

2.2.2.18. Cardiomyocyte Dissociation

Cardiomyocytes differentiated from JHU106i cells were dissociated using STEMdiff™ cardiomyocyte dissociation kit, according to manufacturer's instructions. Cardiomyocytes were washed with D-PBS and incubated with 1ml/well cardiomyocyte dissociation medium at 37°C for 10-12 minutes. Cardiomyocyte support medium was added to each well and gently pipetted using a 10ml serological pipette.

2.2.2.19. Immunofluorescence

Harvested cells were collected and resuspended in 0.1% BSA in PBS. Diluted antibodies were added to cells at 4°C, for 30 minutes. Cells were pelleted and resuspended in 100µl PBS before analysis via flow cytometry to quantify fluorescence.

2.2.2.20. Immunohistochemistry

Cells were harvested and attached to microscope slides using the Shandon cytospin 2, at 300rpm for 5 minutes. Cells were fixed in fixative solution (Section 2.1.9.) for 15 minutes. Cells were subsequently permeabilised in 0.2% Triton- X diluted in dH_2O for 10 minutes at 4ºC. Non- specific antibody binding sites were blocked using 100µl/slide blocking buffer (Section 2.1.9.) for 60 minutes, in a humidified, dark area. Primary antibody was added after blocking for 60 minutes, stored in a humidified, dark place. Slides were subsequently washed 3x5 minutes in TBST solution (described in Section 2.1.9.). 100µl secondary antibody were then added, with slides stored in a humidified, dark place for 60 minutes. Slides were then washed 3x5 minutes in TBST and 3x5 minutes in PBS solutions. Slides were dehydrated in 70%, 90% and 100% ethanol, each for 5 minutes. Slides were subsequently air dried before addition of 15µl DAPI and sealing of cover slip with clear nail varnish. Cells were analysed via fluorescence microscopy.

2.1.3. Animal Procedures

All animal work was completed by Prof. Simon Waddington and Dr Rajvinder Karda, University College London. All animal experiments conducted agreed with the United Kingdom Home Office guidelines, approved by the ethical review committee, and following the institutional guidelines at University College London.

2.1.3.1. Tail Vein Injections

Outbred male and female CD1 mice were first placed in a chamber at 37°C. They were anaesthetised with isoflurane with 100% oxygen. 100µl containing 10^9 Ad- Luc (vg/ml) was administered to mice through intravenous tail vein injections. Mice were allowed to recover in the heated chamber before, returning to their dams.

2.1.3.2. Whole-Body Bioluminescence Imaging

72 hours post AV delivery, mice were anaesthetised with isoflurane with 100% oxygen and received an intraperitoneal injection of 15µg/mL of D-luciferin. Mice were imaged after 5 minutes using a cooled charged-coupled device camera (IVIS machine) for between 1 second and 5 minutes. The ROI were measured Living Image Software and expressed as photons per second per centimetre squared per steradian (photons/second/cm2/sr).

2.1.4. Bioinformatic Analyses

2.1.4.1. Gene Expression Data

Published gene expression data was derived from GEO deposited microarray data of iPSCs and HLCs in GEO under GSE61287 (Takayama, Morisaki et al. 2014).

RNAseq gene expression data of iPSCs differentiating and differentiated cardiomyocyte derivatives were derived from GEO deposited data under GSE116574 (Branco, Cotovio et al. 2019).

2.1.4.2. Gene Enrichment and Analysis

Gene IDs provided in raw expression data were converted to gene symbols using DAVID (v6.7.) (Huang da, Sherman et al. 2009, Huang da, Sherman et al. 2009). Comparisons of gene sets were compared using Genevenn (available at http://genevenn.sourceforge.net/). Enrichment of gene sets was performed using FunRich software (v3.1.3.) (Pathan, Keerthikumar et al. 2015, Pathan, Keerthikumar et al. 2017). Gene ontology classification was determined using Uniprot (UniProt Consortium 2019). Single site analysis of sequences to determine pTFBS were determined using oPossum software (v3.0) under default settings, using an 85% matrix match threshold (Kwon, Arenillas et al. 2012, Ho Sui, Fulton et al. 2007).

Heat maps of differential gene expression were generated using Multi Experiment Viewer (v4.9.0) (Saeed, Sharov et al. 2003).

Data was extracted using Microsoft Excel for Office 365 using the following formulae:

1) =INDEX (DATA ARRAY, MATCH (ID, DATA ARRAY, 0))
2) =INDEX (DATA ARRAY, SMALL (IF (DATA ARRAY=ID, ROW (ID)-ROW (INITIAL CELL) +1), COLUMNS (FIRST RETURNING CELL: FIRST RETURNING CELL)
3) =INDEX (DATA ARRAY), SMALL (IF (DATA ARRAY=>0, ROW (INITIAL CELL) +1), COLUMNS (FIRST RETURNING CELL: FIRST RETURNING CELL)

2.1.4.3. Analysis of Predicted Transcription Factor Binding Sites at Integration Loci

20 bp sequences upstream and downstream of integration loci were retrieved through BLAT alignment to human genome build hg38 using UCSC genome browser tools (https://genome.ucsc.edu/). Sequences were enriched for pTFBS were identified using oPOSSUM3 Single Site Analysis v3.0 (http://opossum.cisreg.ca/) using default parameters. Target sequence hits were quantified through filtered for ≥1 hit per transcription factor name identified.

2.1.4.4. Contractility Analysis of Beating Cardiomyocytes

iPSC derived cardiomyocytes were analysed for contractility using Pulse Video Analysis software, v3.0 (Maddah, Heidmann et al. 2015). Pulse Video Analysis uses proprietary algorithms for assessment of cardiomyocyte contraction in video applications. Deep learning is employed for fully automated detection and removal of noise to ensure accurate measurement of signals. Video images are processed to segment image sequences into blocks to allow extraction of motion signals. Beating signals are quantified and clustered for analysis of contraction velocity and magnitude.

2.1.4.5. Statistical Tests

Two tailed student's t-tests were used for normally distributed data, to generate P values to determine significance. For non-normalised data sets, a Man Whitney U test was performed to analyse significance between samples. For pTFBS data sets, Z scores were calculated using binomial distributed data to determine significance in relation to background data sets.

Chapter III Purification of Viral Vectors using Aqueous Two Phase Systems

ATPS has been shown as a successful alternative purification method for viral vectors. Here, we optimise ATPS purification of AV vectors to show the successful *in vitro* and *in vivo* recovery of adenoviral particles, directly from cell lysates. These systems show low toxicity and reduced cellular contaminants in comparison to crude lysates. However, when used for LV partitioning, no increase in transduction rates were quantified. The use of ATPS is rapid, inexpensive and uses a simple protocol in comparison to other AV vector purification methods. These results indicate the usefulness of developing ATPS for mass scale purification of AV vectors for downstream applications.

3.1. Background

AVs are commonly used in the gene therapy field for episomal gene transfer allowing short term expression of transgenes. Replication defective AV can be generated through transient transfection of HEK293T or Per.C6 cells using plasmids encoding for various viral components. Alternatively, HEK293T cells can be transduced with replication defective AV, which can replicate within these producer cells due to the presence of the Ad5 genome (Graham, Smiley et al. 1977). Both production methods require downstream purification of AV from cell lysates to produce clean products for further applications. AV are commonly purified using CsCl ultracentrifugation, which has been used for removal of organic compounds from biological samples. This system uses density gradients to separate compounds based on molecular size and weight, with the density of the CsCl solution greater towards the lower end of the tube due to Cs+ ions being forced to the outer edge and top of the solution. This gradient allows the separation of organic molecules based on density under high centrifugal force to the neutral buoyancy of the solution (Carr, Griffith 1987, Russell, Sambrook 2001). While CsCl itself is inexpensive, the need for ultracentrifugation and multistep protocols requiring further dialysis to remove organic waste is time consuming and results in low recovery yields of AV vectors. Thus, development of other purification methods for AV particles is needed for sustainable and scalable generation of adenoviral vectors for gene therapy.

3.1.1. Aqueous Two Phase System Dynamics

ATPS is a method of liquid fractionation developed for biomolecule separation, more than 60 years after it was first accidentally discovered (Albertsson 1962, Albertsson, Andersson et al. 1982). This system uses two immiscible solutions

dissolved in water to drive separation of biomolecules. While use of two polymers (i.e. PEG, dextran) have been applied for ATPS generation, the majority of systems combine a polymer and an inorganic salt (phosphate, sulphate or citrate) to provide solutions of varying ionic strength to enhance biomolecule portioning. The formation of two phases is primarily due to the poor miscibility of polymers. The addition of salts absorb water in systems, with both characteristics contributing towards the formation of two immiscible phases (Hatti-Kaul 2000). While salts are generally found to be hydrophilic, the use of different chemical formulas alter the dynamics of systems, indicating the need for optimisation to determine the most effective system for biomolecule portioning and purification.

Three main forces act on ATPS for phase separation. Primarily, gravitational forces drive biphasic formation, which are dependent on the density of the solution. Flotation and frictional forces also affect the flow of molecules within these systems contributing towards the surface properties of each system, which cause the polymer to partition to two phases. Other factors which affect separation include electrophysical properties of the system as determined by salts used, hydrophobicity, affinity to polymers and size exclusion (Hatti-Kaul 2001). These factors influence separation of biomolecules, though not all factors contribute equally in each system.

Multiple conditions affect the separation of biomolecules into the upper or lower phases. Primarily, this depends on the concentration of the chemicals used in ATPS generation and molecular weight of PEG used. Increasing the molecular weight of the polymer used in the ATPS is inversely proportional to the concentration required for phase formation, as shown in generating tie lines for various formulations. Additionally, the increase in molecular weight of PEG increases the hydrophobicity of the solution, which contributes to biomolecule separation (Asenjo, Andrews 2012). The changes in hydrophobicity can be

determined through generating a phase diagram of the system, generally indicating less polymer is required for two phase formation. The addition of salts to polymers for phase formation also drives portioning of biomolecules into the PEG rich phase (Raja, Murty et al. 2011). Understanding of these characteristics' aids development of ATPS formations to optimise biomolecule purification.

Other characteristics of systems also affect biomolecule separation. The pH of the solution alters the net charge of the ATPS affecting protein separation. This is dependent on the isoelectric point of proteins. The positive dipole moment results in pH being proportional to an increase in separation into the polymer rich upper phase (Andrews, Schmidt et al. 2005). While an increase in pH may be useful for separation, solutions must be buffered carefully to ensure no molecule degradation.

Temperature can also affect the separation efficiency of systems. The temperature of the ATPS can affect the viscosity and density of the system, affecting the portioning of biomolecules based on size exclusion. While higher temperatures are more effective for the phase formation of polymer- salt systems, the opposite is observed when using two polymers for ATPSs (Walter, Johansson 1994). This is likely due to the characteristics of the molecules used in ATPS preparation. Therefore, multiple conditions must be optimised to determine efficacy of ATPS purification of biomolecules.

3.1.2. Applications of Aqueous Two Phase Systems

ATPS has been developed for multiple purposes. The polymer and salt solutions are dissolved primarily in water, with this forming the main constituent part of each system. This provides a gentle environment for the stabilisation of various compounds and prevents molecule degradation due to absence of organic

solvents (Asenjo, Andrews 2012). Systems have been further developed to include micellar ATPS for stable separation of proteins from cell membranes using non-ionic surfactants (Asenjo, Andrews 2012, Hatti-Kaul 2000). Due to the properties of proteins, these biomolecules typically separate into the lower phase of systems (Asenjo, Andrews 2012). Varying the molecular weights of polymers used and ionic strength of solutions allows protein separation to occur more effectively. Furthermore, the use of various salts allows increased protein separation from the interphase enhancing recovery yields. This technique has also been applied for antibody separation to success (Sulk, Birkenmeier et al. 1992). These systems have been studied in great depth for protein separation due to their ease of use, low cost and scalability which allow optimisation for continuus operation of ATPS.

3.1.3. Use of Aqueous Two Phase Systems for Viral Vector Recovery

Despite advancements in biomolecule separation for pharmaceutical development, ATPS has not been developed extensively for recovery of virus particles. Various studies have shown ATPS can be used for the successful purification of AAV vectors. Combinations of various molecular weights of PEG (1450, 4000 or 8000) with salts (NaCl, Na_2CO_3 or $(NH_4)_2SO_4$) have been used to successfully portion AAV8. However, an issue that persists is the copurification of hydrophobic proteins in cell lysates with viral particles, requiring batches to undergo further distillation. Altering the composition of systems has allowed removal of the majority of contaminating proteins to a greater extent when using 8% PEG8000- $(NH_4)_2SO_4$. This system shows high recovery of multiple AAV serotypes direct from cell lysates. However, despite viral purification, an additional treatment step using chloroform is still required to precipitate virus particles from the top of the upper phase of ATPS and

removal of hydrophilic proteins (Guo, El-Gohary et al. 2012, Guo, P., Xiao et al. 2013).

Previous studies have also shown ATPS can successfully recover AV to high yields using PEG300 – phosphate (Negrete, Ling et al. 2007, Negrete, Ling et al. 2007). It has also been shown previously that AV RNA can be recovered using ATPS using 6.3% $(C_6H_{10}O_6)_n$ - 3.5% PEG in 10 mM phosphate buffer (Mak, Oberg et al. 1976). While AV can be successfully recovered, further optimisation of ATPS formulations are required to enhance recovery yields.

Addition of PEG has been shown to concentrate Ad vectors up to 200-fold (Wesslen, Albertsson et al. 1959). PEGylation, the addition of PEG to vector supernatants, causes these polymers to coat external receptors on the virus membrane. Further, this process enhances *in vivo* gene transfer, particularly of AV, by reducing steric interference between vector glycoproteins and the cell surface membrane (Croyle, Yu et al. 2000, Croyle, Le et al. 2005). In addition to AV, PEG has previously been shown to concentrate LV vectors. Incubation of lentiviral supernatants with PEG6000 concentrates LV vectors after ultracentrifugation (Kutner, Zhang et al. 2009). However, while this can be used to concentrate LV vectors, ATPS does not concentrate particles greater than other methods including ultracentrifugation, where an average four log higher titre can be observed (Lee, J. Y., Lee 2018). PEGylation of pseudo typed lentiviral vectors, in particular carrying VSV-G is also beneficial for enhancing *in vivo* recovery (Croyle, Callahan et al. 2004). The coating of receptors using PEG prevents serum inactivation of viruses, increasing the half- life of LV particles in the blood stream. While this technique enhances *in vivo* gene transfer, contaminating proteins also purify with viral particles indicating the need for further optimisation.

Here, we have optimised AV recovery *in vitro* by ATPS using both AV previously purified through CsCl density gradients and direct from HEK293T

cell lysates. Furthermore, samples after ATPS purification have been analysed for toxicity and purity to assess the viability of this purification system for large scale AV purification. *In vivo* recovery using the optimal PEG600/$(NH_4)_2SO_4$ system shows comparable rates of gene transfer between ATPS and CsCl density purification indicating the viability of AV particles after separation. As well as analysing AV recovery, we also analyse LV separation using ATPS and *in vitro* recovery to assess the suitability of ATPS for both AV and LV vector purification. These show ATPS as a successful alternative method of purification of AV particles.

Aims

The aims of this study are as follows:

- To assess the rate of purification and toxicity of viral vectors using various ATPSs *in vitro*
- To determine the efficacy of AV vector purification direct from cell lysates
- To analyse the purity of AV vector purification via ATPS
- Comparison of *in vivo* gene transfer of AV vectors purified through ATPS and other methods
- To analyse the efficacy of ATPS for lentiviral vector purification *in vitro*

3.2. Results

3.2.1. AV Genome and *in vitro* Recovery after ATPS Purification

First generation AdRSVβ, carrying the *LacZ* reporter gene under RSV promoter control, was provided having been purified previously through CsCl density gradients (CsCl AV). CsCl AV were directly applied to three ATPSs: 14% (w/w) PEG1000/ K_2PO_4, 14% (w/w) PEG8000/ K_2PO_4 and 20% (w/w) PEG600/ $(NH_4)_2SO_4$, prepared as reported in Section 2.2.1.21. 3.5 x 10^{12} AV particles (vg/ml) were added to ATPS solutions and purified through centrifugation at 3000 rpm for 15 minutes (Section 2.2.2.8.). Quantification of AV genomes were detected via spectrophotometry analysis at Abs_{260nm}. Vector genomes (vg/ml) can be used to calculate titres (Abs_{260nm} x 1.1 x 10^{12}), with percentage yield determined. This work was completed in conjunction with colleagues in my research group.

Analysis of the three ATPS indicates successful portioning of vectors into the upper phase of each system (Table 3.1). Negligible AdRSVβ genomes were detected in the lower phase of each system, which is likely due to the hydrophobicity of the solution. Therefore, the lower phase was not analysed further. High yields of vector genomes were analysed in the upper phase of all ATPS, though there is variability in yields between systems. The presence of this variability is likely due to the various polymer- salt combinations, altering the phase system dynamics and changes in hydrophobicity. The upper phase of PEG600/ $(NH_4)_2SO_4$ was shown as the most successful system for recovery of AdRSVβ vectors genome (83%), showing an 8% and 26% increase in AdRSVβ yield in comparison to the upper phase of PEG1000/ K_2PO_4 and PEG8000/ K_2PO_4 ATPS formulations respectively.

ATPS	Output Titre (vg/ml)	Yield/%
PEG 1000/ K_2PO_4	2.6×10^{12}	75
PEG 8000/ K_2PO_4	1.6×10^{12}	47
PEG 600/ $(NH_4)_2SO_4$	2.9×10^{12}	83

Table 3.1- AV genome recovery after separation by ATPS. An initial input of $3.5x10^{12}$ AV particles were separated in PEG1000/ K_2PO_4, PEG8000/ K_2PO_4 and PEG600/ $(NH_4)_2SO_4$ systems. Optical density measurements were taken after separation to determine the number of AV genomes present in the upper phase of each ATPS, determining the titre and yield of AV particles. Negligible virus was quantified in the lower phases of each system. (n=6)

While genome recovery is possible, it must be assessed if these partitioned vectors are viable for gene transfer. Aggregation of AV particles can affect gene transfer capacity (Galdiero 1979). As reported in Section 2.2.2.8 and 2.2.2.9, aliquots of upper phase containing partitioned CsCl AV were used to transduce HepG2 cells and stained 48 hours post transduction for β- galactosidase activity (Section 2.2.2.5). Cells were analysed for positive transduction (blue cells) 24 hours post staining to indicate successful gene transfer, in conjunction with colleagues (Figure 3.1 & 3.3).

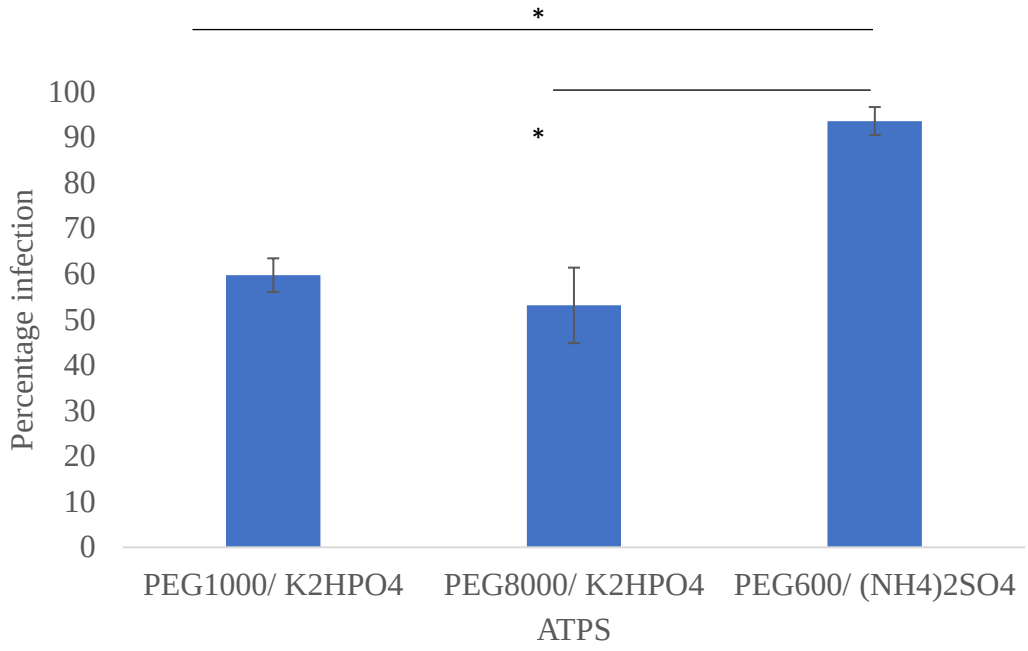


Figure 3.1- Percentage AV infection of HepG2 cells with CsCl AV after ATPS purification. The upper phase of each ATPS was used for infection of the HepG2 cells and quantified after β- galactosidase staining for percentage infection. HepG2 cells were infected at 59.75%±3.69 using PEG1000/ K_2HPO_4, 53.11%±8.29 using PEG8000/ K_2HPO_4 and 93.57%±3.08 using PEG600/ $(NH_4)_2SO_4$ (N=7). SEM shown. Significance denoted by asterix ($P<0.05$)

Transduction of HepG2 cells with AV purified through the upper phase of PEG1000/ K_2HPO_4, PEG8000/ K_2HPO_4 and PEG600/ $(NH_4)_2SO_4$ systems indicates greater than 53% positive infection of cells. The results of infection studies mirrors that of genome recovery, with the highest infection rates observed in PEG600/ $(NH_4)_2SO_4$. The rate of infection using this system is 33.82% and 40.46% higher than PEG1000/ K_2HPO_4, PEG8000/ K_2HPO_4 respectively. However, poor infection was observed after transduction of HepG2 cells using the lower phase of each system. These results are expected from genome recovery analysis and indicate the majority of AV partitioning in

the upper, PEG rich phase. Infection rates between PEG 1000/ K_2HPO_4 and PEG 8000/ K_2HPO_4 independently compared to PEG 600/ (NH4)2SO4 were found to be significantly different ($P<0.05$). These viral particles were analysed via spectrophotometry analysis to determine aggregation, as previously described (Lefesvre, van Bekkum 2003). Abs_{320nm}/Abs_{230nm} and Abs_{340nm}/Abs_{260nm} were all below 0.3, indicating no virus aggregation. These results suggest AV particles are purified to high levels using ATPS, in particular PEG 600/$(NH_4)_2SO_4$. These viruses demonstrate high levels of *in vitro* recovery and do not form aggregates, suggesting this method of purification is successful at partitioning AV vectors, in particular using PEG600/$(NH_4)_2SO_4$.

3.2.2. *In vitro* AV Recovery after ATPS Purification Direct from Cell Lysates

We have shown that AV recovery is possible using ATPS and next investigated whether these viral vectors can be purified directly from transduced HEK293T cell lysates. AV can be grown up in large quantities through transduction of HEK293T cells. HEK293 cells have been transformed with AV *E1A* and *E1B* genes from the human AV5, which produce AV early genes in *cis*. This allows packaging of replication deficient AV in these cells to generate large quantities of virus for purification and downstream applications (Graham, Smiley et al. 1977).

AV vectors generated through HEK293T cell transduction (Section 2.2.2.6) were quantified through optical density to determine titre ($3.43x10^{12} \pm 1.80x10^{11}$ vg/ml, n=5). 3×10^{12} vg/ml AdRSVβ particles from HEK293T cell lysates were added to ATPS solutions, purified and aliquots of the upper and lower phase were used to transduce HepG2 and IMRS2 cells. Reduced infection rates were observed in the lower phase of each system and were not explored

further. Aliquots of the upper phase were taken for infection of HepG2 and IMRS2 cells to determine successful gene transfer, in conjunction with colleagues. This indicates the successful portioning of AV direct from cell lysates (Figure 3.2 & 3.3).

These results mirrored that of genome and *in vitro* recovery of CsCl purified AV, with PEG600/ $(NH_4)_2SO_4$ demonstrating the highest quantification of positive infection on both HepG2 and IMRS2 cells. This system shows 22.50% and 27.50% higher infection on HepG2 cells than PEG1000/ K_2PO_4 and PEG8000/ K_2PO_4 respectively. Infection rates of HepG2 cells observed between PEG1000/ K_2PO_4 and PEG600/ $(NH_4)_2SO_4$ were found to be significantly different ($P<0.05$). Similarly, infection of IMRS2 cells after direct purification of AV from cell lysates shows an 8.34% and 63.67% higher rate of infection than PEG1000/ K_2PO_4 and PEG8000/ K_2PO_4 systems, respectively. As observed with CsCl purified AV partitioning, the highest rate of purification of AdRSVβ direct from cell lysates was observed using PEG600/ $(NH_4)_2SO_4$. However, higher rates of infection can be observed in IMRS2 cells compared to HepG2 after AV transduction using PEG1000/ K_2PO_4 (34.43%) and PEG600/ $(NH_4)_2SO_4$ (20.27%) systems. However, a 5.90% higher transduction rate is observed in HepG2 cells when using in PEG8000/ K_2PO_4. These results indicate that AV can be successfully purified direct from HEK293T lysates after AV transduction to high levels using ATPS, in particular from the upper phase of PEG600/ $(NH_4)_2SO_4$.

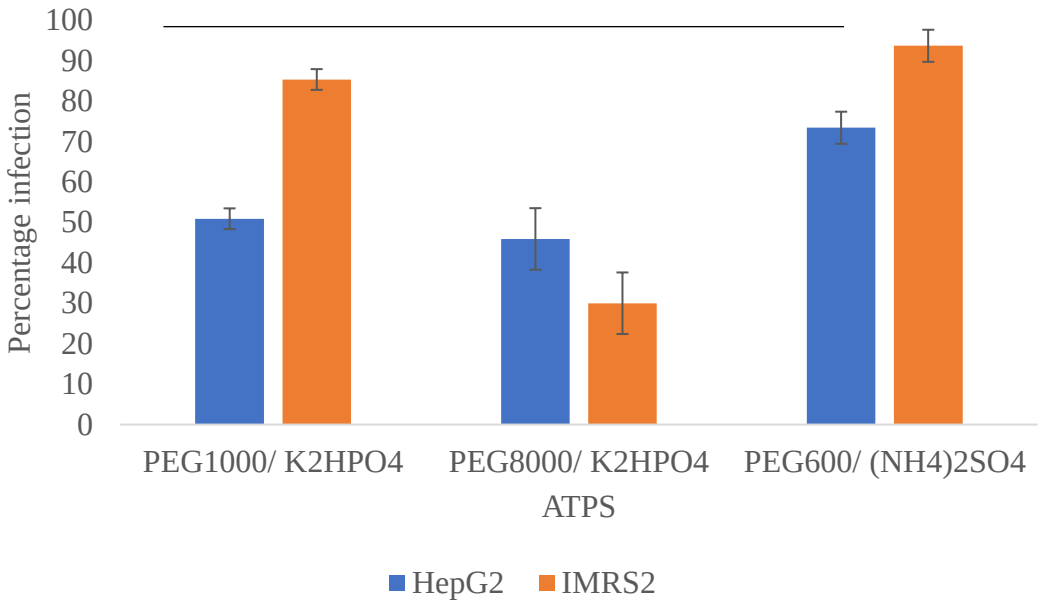


Figure 3.2- Percentage infection of HepG2 (blue, n=5) and IMRS2 cells (orange, n=3) cells using AV purified directly from HEK293T cell lysate. The upper phase presented the highest recovery of AV particles and was used for infection. 50.90±2.56% and 85.33±2.56% infection was observed after purification using PEG 1000/ K_2HPO_4. 45.90±7.61% and 30.00±7.61% infection was observed using PEG 8000/ K_2HPO_4. Purification using PEG 600/ $(NH_4)_2SO_4$ showed 73.40±3.97% and 93.67±3.97% infection, respectively. This demonstrates the high recovery yield of AV directly from 293T lysate, with the highest recovery using upper phase PEG 600/ $(NH_4)_2SO_4$. SEM shown. Significance shown by asterix ($P<0.05$)

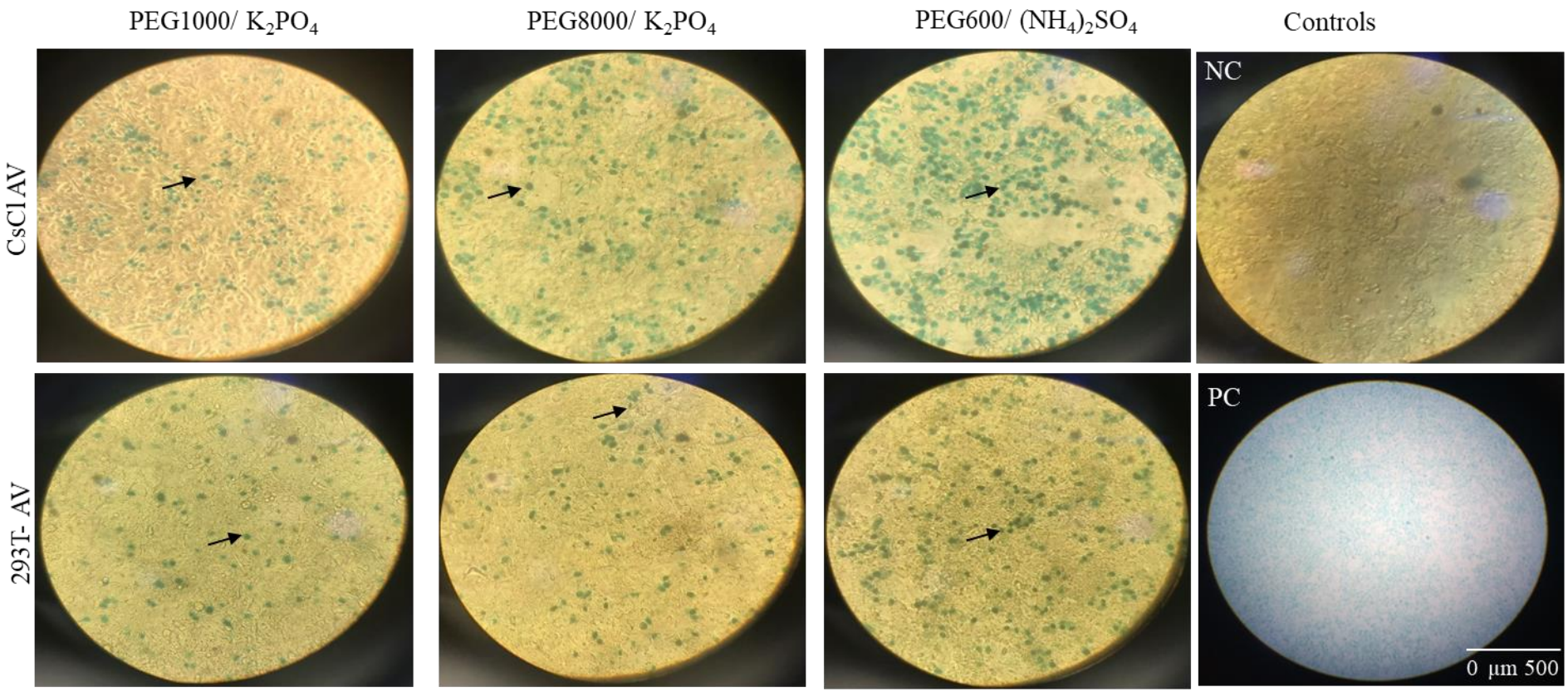


Figure 3.3- Infection of HepG2 cells using ATPS purified AV. AV purified were from stock AV (CsCl AV) or directly from 293T lysate (293T- AV) using ATPS. Infection of AV was characterised by staining for β- galactosidase activity (blue) and analysed via brightfield microscopy. Uninfected HepG2 and TelCeb6 cells stained as a negative (NC) and positive control (PC) respectively. Cells shown at x10 magnification and scale bar shown between 0 to 500µm

The differences in infection rates between IMRS2 and HepG2 cells may be due to the permissiveness of the neuroblastoma cells to AV infection compared to hepatocellular carcinoma cell lines. Despite differences in the rates of infection, PEG 600/ $(NH_4)_2SO_4$ demonstrates the highest rates of infection of all ATPS screened, with 73±3.97% and 93.67±3.97% infection in HepG2 and IMRS2 cells respectively.

This indicates successful purification of AV particles direct from HEK293T cell lysate. Infectious recovery assays on both neuronal and hepatic cell lines indicate the upper phase of PEG 600/ $(NH_4)_2SO_4$ as the most successful for purification. High recovery of infectious AV particles from this system indicates the potential of this simplified protocol for rapid purification of AV vectors in a benchtop centrifuge. AV partitioning direct from cell lysates further increases the speed of manufacturing without the need for multiple rounds of purification and dialysis. However, the rate of AV infection direct from cell lysates are lower than observed from pure AV purification using ATPS. Higher rates of infection of HepG2 cells are quantified after direct AV partitioning, with an 8.85%, 7.21% and 20.17% greater rate of infection using PEG1000/ K_2PO_4, PEG8000/ K_2PO_4 and PEG600/ $(NH_4)_2SO_4$ systems than quantified from cell lysates. This suggests an increased viral input is needed to gain higher rates of infection.

3.2.3. AV *in vitro* Toxicity Analysis

High gene transfer can be observed subsequent to AdRSVβ separation using ATPS, particularly in the PEG rich upper phase of systems. However, the toxicity of ATPS must be analysed to determine whether the chemicals used in preparation of these ATPS, or AV transduction is toxic *in vitro*. Low toxicity is important to ensuring cell survival post gene transfer. HepG2 cell viability was

quantified via a trypan blue exclusion assay (Section 2.2.2.3) after treatment using the upper phase of ATPS under various conditions: ATPS only, ATPS containing CsCl purified AdRSVβ and ATPS containing AsRSVβ purified direct from HEK293T lysate (Figure 3.4). The lower phase of ATPS was not assessed due to poor genome and *in vitro* recovery of AV.

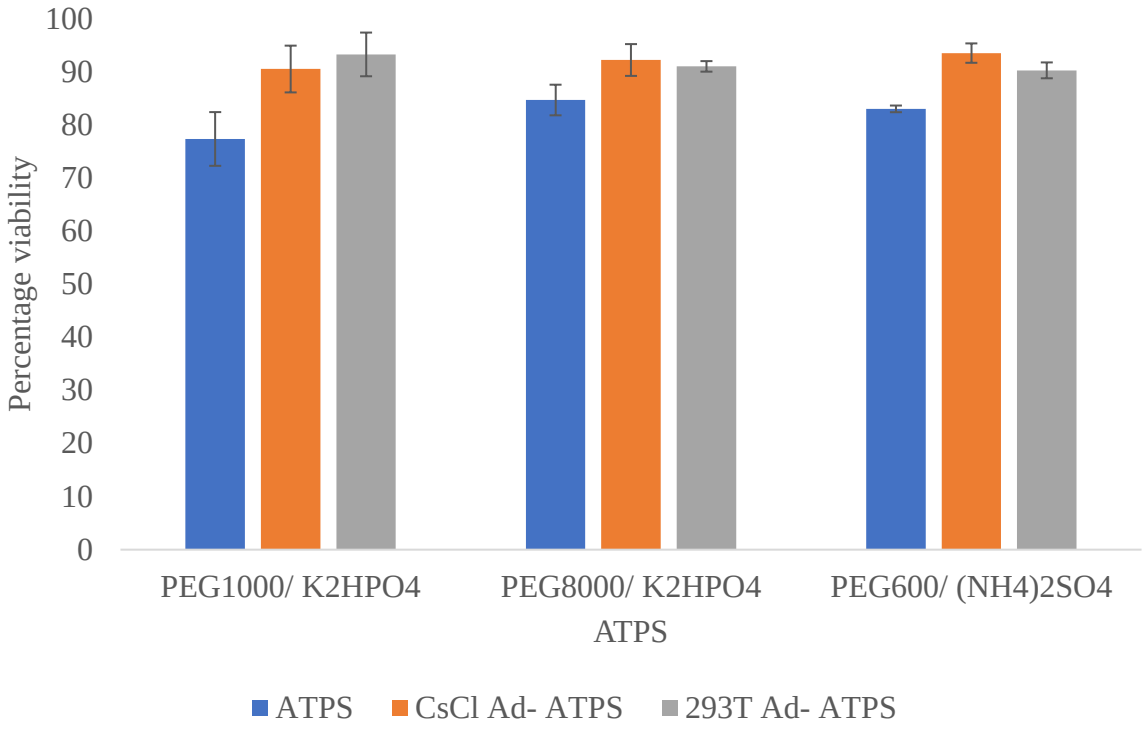


Figure 3.4- Viability of HepG2 cells after upper phase ATPS treatments. Treatments include: + ATPS only (blue, n=6), + CsCl AV (orange, n= ≥5) and + AV purified from 293T lysate (grey, n=4). PEG 1000/ K_2HPO_4 treatment demonstrated cell viability of 77.33±5.06% (+ ATPS), 90.50±2.88% (+ CsCl AV) and 93.25±0.63% (+293T AV). PEG 8000/ K_2HPO_4 showed a viability of 84.67±4.40% (+ATPS), 92.20±2.99% (+CsCl AV) and 91.00±1.83% (+293T AV). PEG 600/ $(NH_4)_2SO_4$ upper phase showed a viability of 83.00±4.12% (+ATPS), 93.50±0.99% (+ CsCl AV) and 90.25±1.49% (+293T AV). No significant difference (P>0.05) was found between treated cells in comparison to untreated HepG2 cells (80.56±2.73%). SEM shown

Cell viability post treatment with the upper phase of ATPS only ranged from 77.33±5.09% – 83.00±4.12%. Cell survival post transduction with CsCl AdRSVβ purified through ATPS ranged between 90.50±2.88% - 93.50±0.99%. Cells infected with AdRSVβ purified direct from HEK293T cell lysate, in the upper phase of each ATPS ranged from 90.25±1.49% to 93.25±0.63%. No significant decrease in cell survival is observed in any sample sets in

comparison to untreated HepG2 cells (80.56±2.73%, n=9, P>0.05). IMRS2 cells treated with ATPS purified AV direct from cell lysates exhibit 70.50±1.50% - 78.00% viability 48 hours after treatment. IMRS2 cells exhibit 13.00% - 19.58% greater toxicity compared to HepG2 cells. Additionally, IMRS2 cells do not show a significant drop in cell viability after exposure to AdRSVβ purified direct from HEK293T cell lysates in comparison to untreated cells (data not shown). This suggests no cytotoxicity after ATPS treatment or AV transduction.

Comparison of HepG2 cell viability after AdRSVβ transduction purified direct from cell lysates using the upper phase of PEG1000/ K_2PO_4, PEG8000/ K_2PO_4, PEG600/ $(NH_4)_2SO_4$ show no significant difference in comparison to untreated HepG2 cells or cell treated to ATPS alone. This suggests low toxicity of HepG2 and IMRS2 cells to AV transduction purified direct from cell lysates and ATPS treatments, over a 48-hour period. This suggests that ATPS treatment and AV transduction using the upper phase of various ATPS formulations do not cause cell toxicity.

3.2.4. AdRSVβ Transduction Sensitivity of IMRS2 cells

To understand the differences in the rate of infection of HepG2 and IMRS2 cells, we analysed the permissiveness of neuroblastoma cells (IMRS2) to AV infection. At a range of logarithmic dilutions of AV, ranging between 1×10^6 – 1×10^{11} vg/ml, between 1±0.88% - 100±0% of cells stained positive for β-galactosidase expression, performed in conjunction with colleagues (Figures 5 & 6). This demonstrates the sensitivity of these cells, with a dose dependent rate of infection of IMRS2 cells. Transduction of IMRS2 cells is observed at low

titres of AdRSVβ (1x10^6 vg/ml). This suggests that higher infection rates may be observed in IMRS2 compared to HepG2 cells due to increased sensitivity.

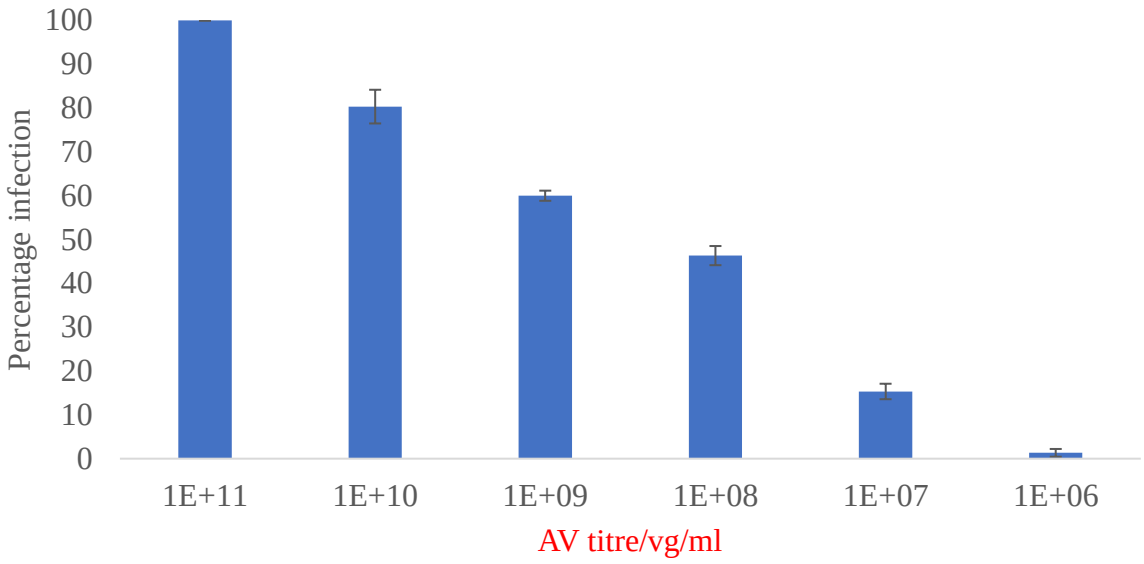


Figure 3.5- Sensitivity of IMRS2 cells to AV infection. Logarithmic dilutions of CsCl AV (1 x 10^{11} – 1 x 10^6 vg/ml) were used to infect IMRS2 cells and quantified after β-galactosidase staining. 100±0%, 80.33±3.84%, 60±1.15, 46.33±2.19%, 15.33±1.76% and 1.33±0.88% of IMRS2 cells were infected at dilutions ranging between 1x10^{11}- 1x10^6 AV (n=3). SEM shown.

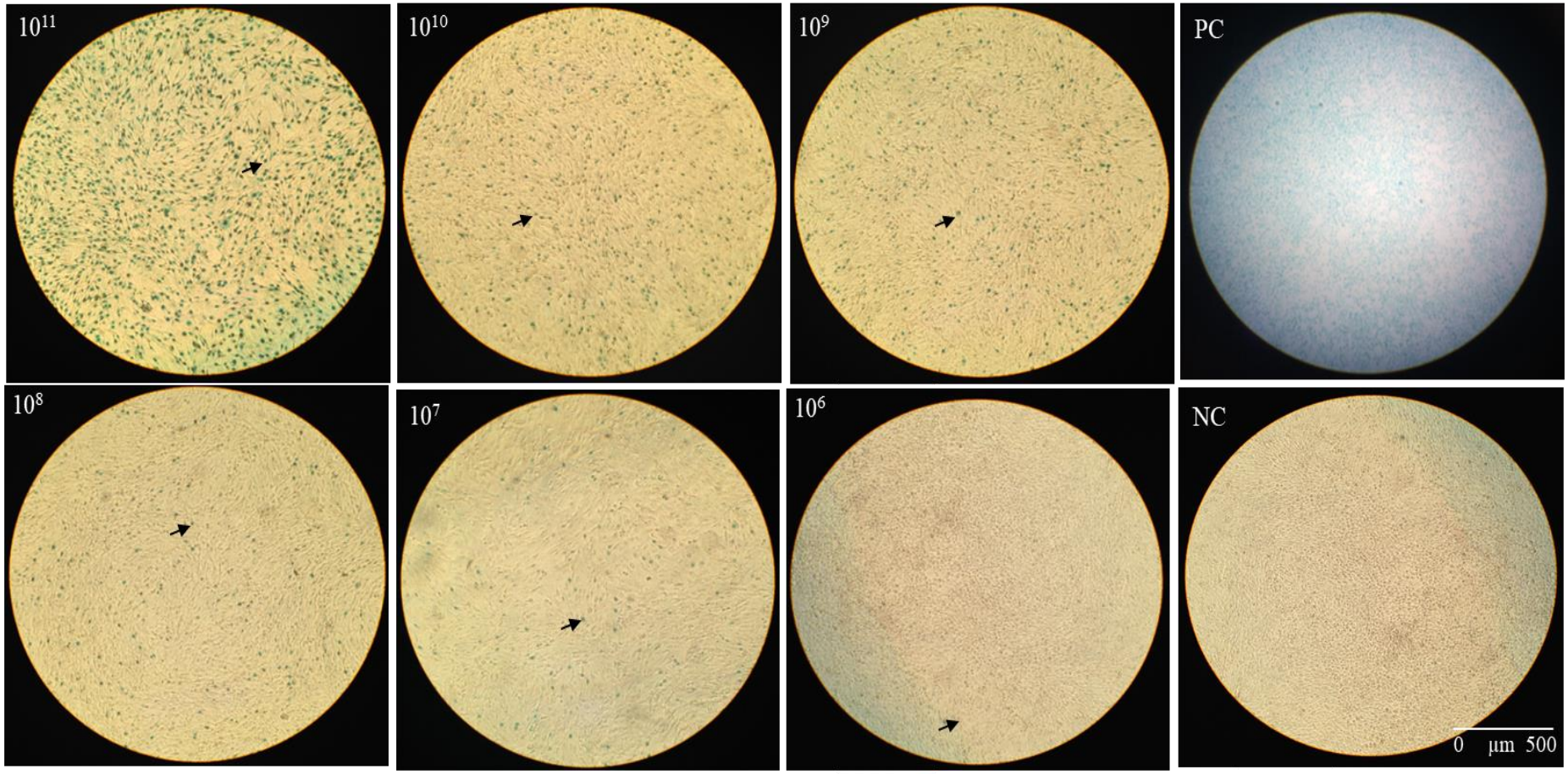


Figure 3.6- IMRS2 cells transduced at various dilutions of AV particles. Infection was detected by staining for β- galactosidase activity and brightfield microscopy analysis (x10 magnification). Positive infection shown as blue cells with examples highlighted by arrows. Uninfected IMRS2 and TelCeb6 cells were used as a negative (NC) and positive (PC) controls, respectively. Dose dependent AV infection demonstrates the sensitively of IMRS2 cells to AV infection at low titre AV particles.

3.2.5. *In vivo* Recovery of AV Purified Through ATPS

Successful *in vitro* recovery and low cytotoxicity of AV purified through ATPS direct from HEK293T cell lysates has been shown above. To assess if AV purified from ATPS can be used for *in vivo* gene transfer, recovery was measured from AV carrying the luciferase transgene (AV-Luc), kindly provided by Professor S. Waddington (UCL). AV-Luc (2.9×10^{12} vg/ml) were grown using HEK293T cell transduction to generate 2.88×10^{13} vg/ml. Direct partitioning of AV-Luc from cell lysates was performed using the PEG600/ $(NH_4)_2SO_4$ ATPS. Subsequent to partitioning, 10^9 vg/ml AV-Luc particles from the upper phase of this system were injected into the tail veins of white mice, as described in Section 2.1.3.1 (Figure 3.7).

Detection of luciferase transgene expression determined successful *in vivo* gene transfer after purification of AV direct from cell lysates. Qualatative comparison of ROI generated from injection of AV-Luc purified through ATPS to *in vivo* expression of CsCl purified AV- Luc show comparable rates of expression in mice. This indicates that ATPS can successfully purify AV vectors to high yields for *in vivo* gene transfer. The comparable qualitative expression of *in vivo* gene transfer to traditional purification methods suggest that ATPS purification of vectors do not inhibit vector viability *in vivo*. No adverse side effects were observed in treated mice demonstrating the low toxicity of PEG600/ $(NH_4)_2SO_4$ in animals. This further suggests the usefulness of ATPS, in particular PEG6000/ $(NH_4)_2SO_4$ ATPS, for successful partitioning of AV for *in vitro* and *in vivo* gene transfer.

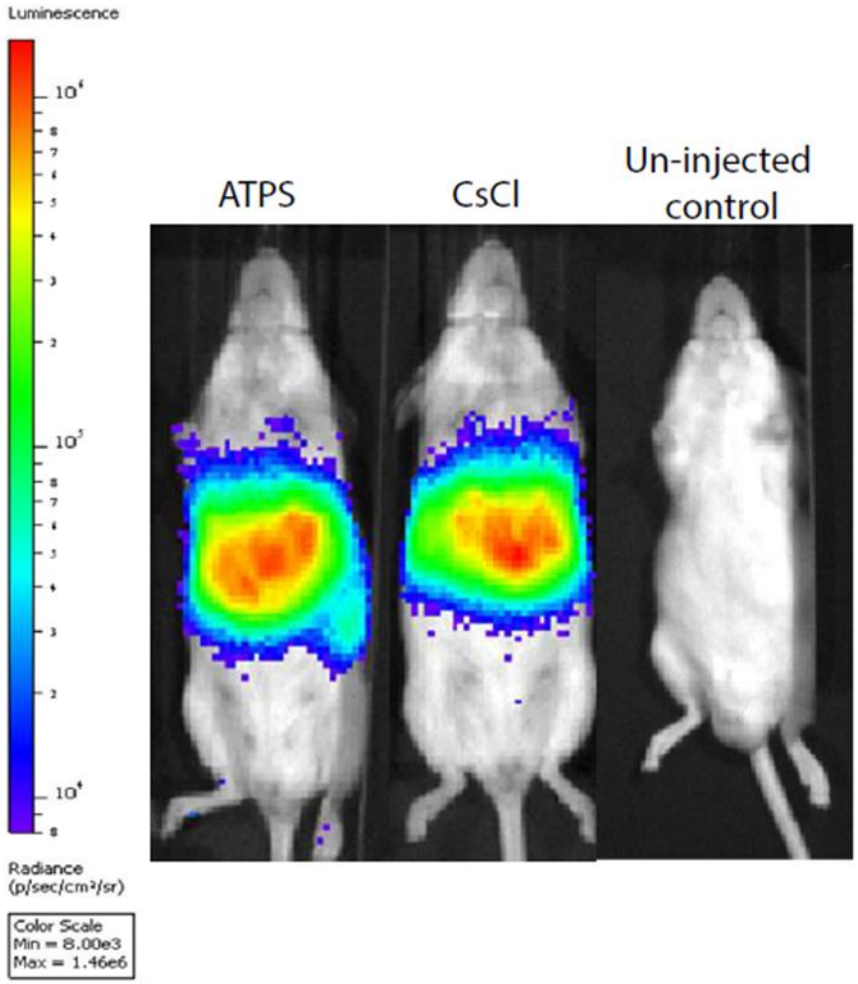


Figure 3.7- *In vivo* recovery of AV- Luc virus purified through upper phase of PEG600/ $(NH_4)2SO_4$ ATPS. 1 x 10^9 (vg/ml) AV- Luc particles purified through upper phase of PEG600/$(NH_4)_2SO_4$ ATPS or CsCl density gradient were injected into mice to determine *in vivo* gene transfer. AdLuc particles purified through the upper phase of PEG600/$(NH_4)_2SO_4$ ATPS show a comparable rate of gene transfer to traditional CsCl preparation of Ad- Luc. This demonstrates successful *in vivo* recovery of AV particles. N=3

3.2.6. Purity Analysis of AV Samples

Previously, we have shown no toxicity of AV transduction after ATPS purification *in vitro* or *in vivo* using PEG6000/ $(NH_4)_2SO_4$ ATPS. To further evaluate the sample purity of purified AV, samples were analysed via size exclusion chromatography (Section 2.2.1.22). Samples were run on a TSK gel®

column at Abs_{260nm} and Abs_{280nm} wavelengths for 40 minutes after column equilibration.

To ensure calibration of the column, protein standards were run in HPLC Buffer A without PEG. While appropriate peaks were observed, according to manufacturer's instructions, the subsequent run of ATPS purified AV contained contamination. Similar peaks were observed throughout the chromatogram to the previously run protein standards. This suggests the incomplete elution of the initial run of standards. Therefore, 1% PEG600 was added to HPLC Buffer A.

This optimised buffer was successful in eluting all of the sample, with no contaminating peaks observed in subsequent runs. Therefore, all samples were run in an optimised HPLC Buffer A, containing 1% PEG 600. Protein standards were run in the optimised HPLC Buffer A for calibration of the column. Distinct peaks were detected, for thryoglobulin, γ globulin, ovalbumin, myoglobulin and vitamin B12 as indicated by manufacturer's instructions (Figure 3.8).

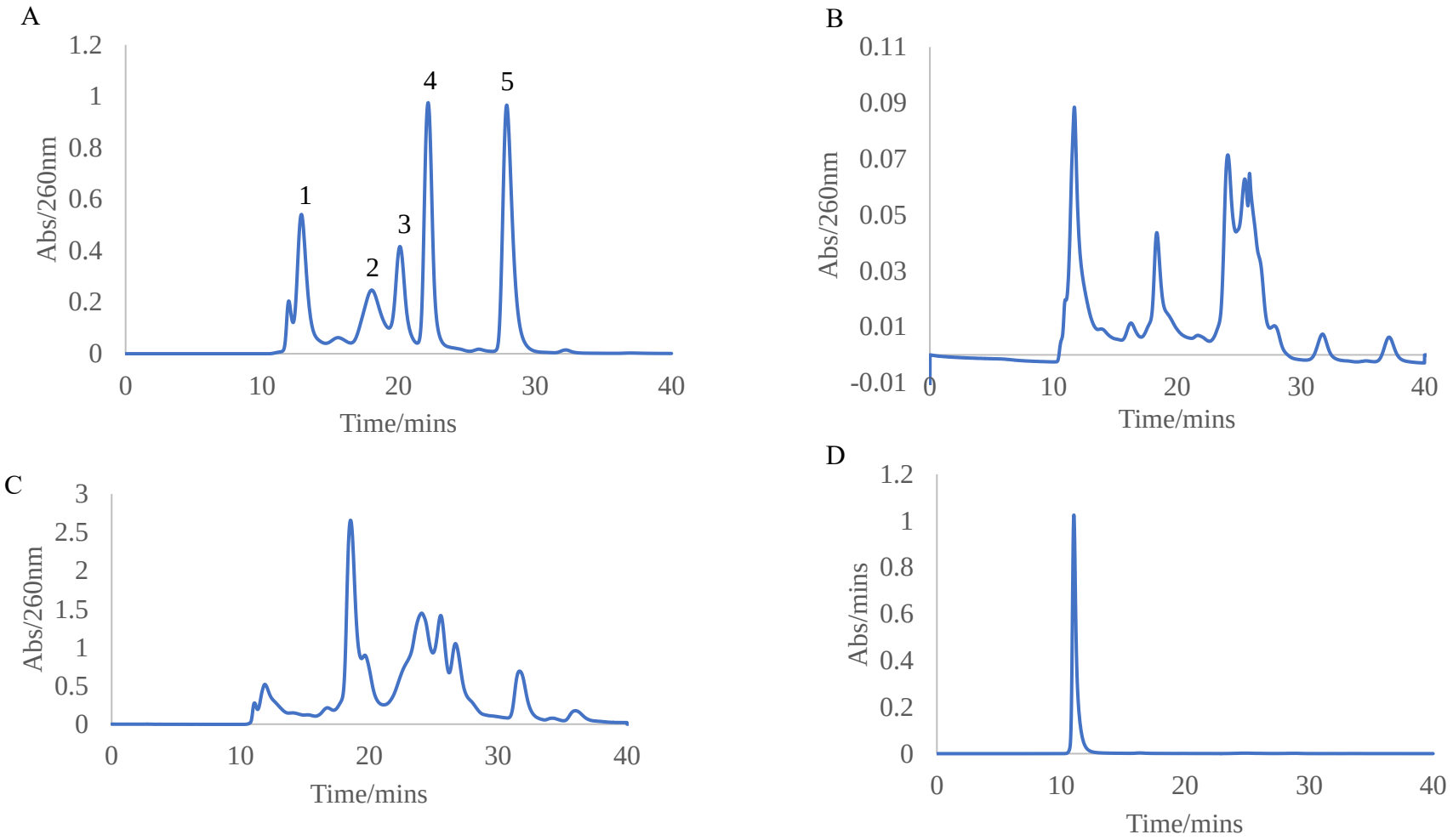


Figure 3.8- Chromatogram of AV purified samples using HPLC Buffer A with 1% PEG600 for 40 minutes. A- Protein standards: 1: thryoglobulin, 2: γ globulin, 3- ovalbumin, 4- myoglobulin, 5- vitamin B12. B- AV samples in upper phase of PEG600/$(NH_4)_2SO_4$ ATPS. AV eluate observed after 11.68 minutes. Permutations in chromatogram show contaminating proteins present. C- Unpurified HEK293T cells lysate containing AV. AV eluate at 11.96 minutes. D- AV stock purified through CsCl density gradients. AV elute seen after 11.02 minutes. Permutations in chromatogram show contaminating proteins found in upper phase preps of cell lysates compared to a clean AV prep purified through CsCl gradients

When running AV purified through CsCl density gradients, AV retention time was found to be 11.03 minutes, with no perturbations observed in the rest of the run suggesting no other biological contaminations in the viral prep. In comparison to this, AV purified through ATPS upper phase, direct from HEK293T cell lysates, also showed a retention time of 11.71 minutes. The retention time of AV are within 1.04 minutes between all samples (CsCl AV: 11.03 minutes, ATPS AV: 11.71, 293T lysate- AV: 12.07 minutes) suggesting consistent elution from the column. However, contaminating peaks are seen in the chromatogram when running AV purified direct from cell lysate (Figure 3.8). These peaks correspond well with the peaks present when running crude cell lysate containing AV and are also found at similar elution times as observed in the protein standards. As no perturbations are observed when running the upper phase alone through the HPLC column, the contaminating peaks on the chromatogram are seen to derive from the cell lysate.

While the width of all peaks remains consistent between all runs, the height of the AV peak is higher in ATPS samples in comparison to crude cell lysate. Conversely, contaminating protein peaks are shorter in ATPS samples compared to cell lysate chromatograms. This suggests concentration of AV particles and exclusion of contaminating cellular proteins through ATPS partitioning.

Molecular size analysis, using the Zetasizer Nano S, under 30% sucrose buffer setting due to the viscosity of PEG, determined 85.7% of the viral fraction collected to be 93.63 nm, corresponding with the average size of AV particles.

These results suggest that while PEG600/ $(NH_4)_2SO_4$ partitions AV particles to high genome recovery yields and infectivity, the sample is not pure with contaminating proteins present in the cell lysate partitioning with AV in the upper phase. However, AV is concentrated through ATPS and no toxicity is observed *in vitro* or *in vivo* after AV treatment. This suggests that while samples

are not free of contaminating cellular proteins, these do not appear to cause a detrimental effect for *in vitro* and *in vivo* assays.

3.2.7. LV Purification through ATPS

Due to the successful purification of AV particles, including directly from cell lysates, we then proposed to analyse if ATPS can also be used to purify LV. Aliquots of pHR LV, carrying the *eGFP* transgene under SFFV promoter and flanked by SIN LTRs, were added to ATPS solutions and purified as previously described (Section 2.2.2.8). Aliquots were used to treat HEK293T cells after complexation with 5 µg/ml polybrene for 20 minutes at room temperature before transduction. Cells were analysed for GFP expression via Novocyte flow cytometry and analysis on Novoexpress software, 72 hours post transduction.

Low GFP expression was detected in all three ATPS. LV transduction is not successfully partitioned in the upper phase (Figure 3.9) with the majority of expression seen in the lower phase treatment of each ATPS (Figure 3.9). 0.89±0.17% and 0.30±0.13% GFP expression is detected post partitioning with the upper and lower phase of PEG1000/ K_2HPO_4, respectively. 0.72±0.06% and 1.19±0.54% GFP expression was detected using the upper and lower phase of PEG8000/ K_2HPO_4. 0.21±0.04% and 1.22±0.09% GFP expression can be detected using the upper and lower phases of PEG600/ $(NH_4)_2SO_4$, respectively. The level of transduction after ATPS purification is lower than observed when transducing cells with concentrated LV alone (9.72±0.71%). Although this difference is not found to be significant ($P>0.05$), the low rate of viral quantification suggests ATPS is not a successful method for LV recovery.

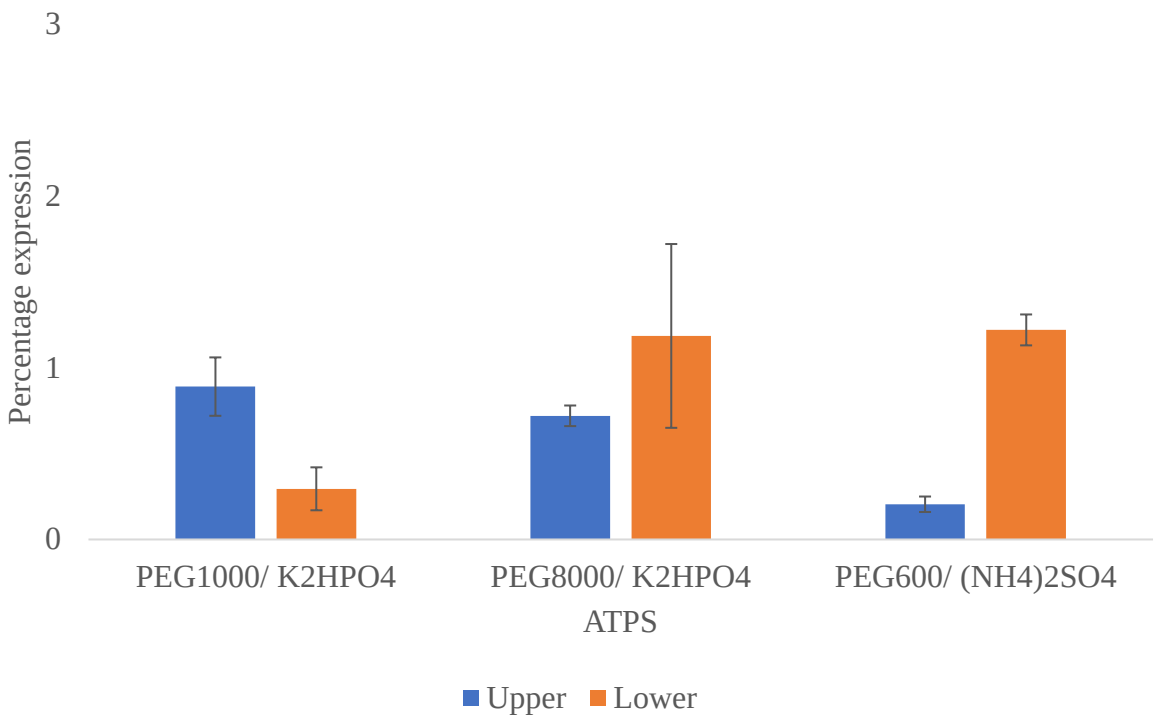


Figure 3.9- Percentage GFP expression in HEK293T cells after transduction using LV purified through ATPS upper (blue) and lower phases (light grey). PEG 1000/ K_2HPO_4: 0.89±0.17% & 0.30±0.13%. PEG 8000/ K_2HPO_4: 0.72±0.06% and 1.19±0.54%. PEG 600/ $(NH_4)_2SO_4$: 0.21±0.04% and 1.22±0.09% respectively. Positive control shows 9.72±0.71% GFP expression. This indicates poor separation of LV particles by ATPS (n=2). SEM shown

3.2.8. *In vitro* LV Toxicity Analysis

Cells were analysed post treatment to determine cytotoxicity using a trypan blue exclusion assay, as previously described (Figure 3.10). ATPS treatment showed between 89.00±1.00% and 86.50±3.50% viability using the upper and lower phase of PEG1000/ K_2HPO_4, respectively. 92.50±0.50% and 87.00±3.00% cell viability was observed using PEG8000/ K_2HPO_4. HEK293T cells treated using PEG600/ $(NH_4)_2SO_4$ demonstrated a viability of 94.00±0% and 82.00±4% respectively. LV transduction after ATPS purification presented viabilities of 89.50±3.28% and 92.72±1.11 using PEG1000/ K_2HPO_4. 79.5±2.53% and 79±2.86% viability was measured after PEG8000/ K_2HPO_4 treatment.

88.25±1.70% and 84.5±3.23% viability was measured after PEG600/ $(NH_4)_2SO_4$ transduction. Cells transduced using LV with 5 µg/ml polybrene presented 92.00±2.83% viability. The viability of treated cells showed no significant difference (P>0.05) to untreated cells (90.25±1.78%) suggesting no increase in toxicity due to ATPS treatment or LV infection. However, poor LV vector gene transfer was shown *in vitro* after ATPS purification.

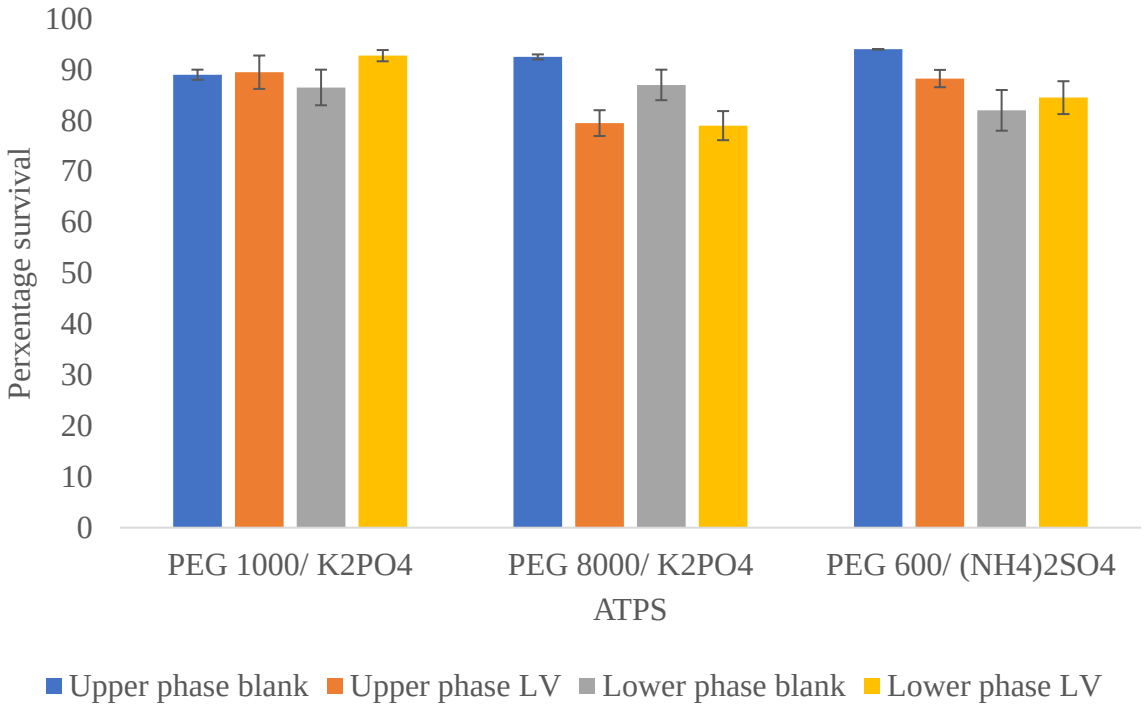


Figure 3.10- Percentage viability of HEK293T cells post transduction using LV purified through ATPS. SEM shown. No significant drop in viability after treatment with ATPS or LV purified in ATPS in comparison to untreated cells. This indicates low toxicity of ATPS chemicals and LV (N≥2). SEM shown

3.3. Discussion

AV are useful gene transfer vectors that are under development as licensed bioproducts for therapeutic applications, including gene therapy of genetic diseases (Albertsson 1962, McNally, Darling et al. 2014, Negrete, Ling et al. 2007, Negrete, Ling et al. 2007, Oelmeier, Dismer et al. 2011). However, manufacturing and purification of large quantities of recombinant AV particles remains a challenge for laboratories, with greater than 10^{15} virus particles typically required for clinical trials. Due to AV5 recognition by the human immune systems, AV from other species has been developed, including AV5, AV26 and AVs based on chimpanzee and gorilla adenoviruses are in clinical trials (Cheng, Wang et al. 2015, Baden, Liu et al. 2015). This suggests the need for alternative time and cost-effective techniques for AV manufacture and downstream purification.

AV vector production typically employs adherent cultures of producer cells transfected with plasmids carrying *cis* and *trans* elements required for virus particle production. Alternatively, HEK293 producer cells can be infected with attenuated replication defective particles for further replication. Replication deficient AV are able to replicate in these producer cells due to the presence of E genes in the HEK293 genome (Graham, Smiley et al. 1977). Cell lysates containing AV vectors are contaminated with intracellular proteins and gDNA which may have negative effects in downstream applications. CsCl density gradients and ultracentrifugation are commonly used for AV vector purification from cell lysates. However, this process is both time consuming and expensive. This process has also been shown to yield poor recovery of virus particles and may cause particle aggregation, reducing the infectivity of samples (Cheng, Wang et al. 2015, Baden, Liu et al. 2015, Berkner 1988). Other methods of purification may be used, including using chromatography. This technique

elutes viral samples over columns. However, columns can easily become blocked and limits the volume of lysate input (Berkner 1988).

PEG has been found especially useful in separating biomolecules and its hydrophobicity is directly correlated to its molecular weight. PEG300-phosphate ATPS has previously been shown to purify AV particles to high titres, with the majority of vectors identified in the upper aqueous phase (Negrete, Ling et al. 2007, Negrete, Ling et al. 2007). AAV vectors have also been successfully partitioned using 10% PEG8000/13.2% $(NH_4)_2SO_4$ ATPS, to high recovery yields and infectivity (Guo, Xiao et al. 2013, Rogers, R. D., Bond et al. 1996). This suggests the usefulness of these systems for viral vector purification.

14% - 14% w/w PEG1000 - K_2HPO_4, and 14% - 14% w/w PEG8000 - K_2HPO_4 and 20% - 20% w/w PEG 600 - $(NH_4)_2SO_4$ showed successful and high recovery of AV vectors *in vitro* (47-83%). These vectors were detected mainly in the upper, PEG rich phase of these systems and infected HepG2 and IMRS2 cells efficiently. No significant cytotoxicity was observed in these cells, including after infection with AV purified direct from cell lysates. Both spectrophotometry and *in vitro* assays indicate the most effective ATPS was PEG 600-$(NH_4)_2SO_4$, with the highest rates of infections in both liver and neuronal cell lines. Analysis of these products via size exclusion chromatography indicates the presence of contaminating protein particles in ATPS- AV preparations, though at a lower concentration than in crude cell lysates. This suggests hydrophobic proteins are also forced into the upper phase of the ATPS along with AV vectors as observed with AAV (Guo, Xiao et al. 2013, Guo, El-Gohary et al. 2012). As no toxicity is observed *in vitro* or *in vivo*, this suggests the usefulness of the PEG600 ATPS for downstream applications, but further optimisation is required for the purification of AV vectors direct from cell lysates for clinical translation.

The successful recovery of AV and low toxicity suggests the potential for ATPS for AV purification, including direct of cell lysates providing a relatively quick method in comparison to using CsCl density gradients for purification. It is suggested that as well as the different rates of genome recovery, the PEGylation of vectors enhance *in vitro* and *in vivo* gene transfer which may indicate the reason for these vectors remaining viable after a year of storage at -80°C in ATPS solutions.

Despite successful AV purification through ATPS, this is not observed when partitioning LV. LV gene transfer is greatly reduced after ATPS purification compared to traditional concentration methods. The low rate of purification of LV possibly may be due to the difference in characteristics of LV to AV particles. The VSV-G protein on the surface of LV membranes are hydrophobic which may result in aggregation upon exposure to hydrophobic forces. In comparison, hexons present on the capsid of AV exhibit both hydrophobic and hydrophilic characteristics (Crawford-Miksza, Schnurr 1996, Carneiro, Bianconi et al. 2002). While PEG itself has been shown to concentrate LV vectors, the formation of ATPS using PEG and salt combinations is likely to alter system characteristics with LV not able to partition successfully into the upper PEG rich phase. This suggests differences in AV and LV characteristics contribute to successful AV purification. While this technique does not cause toxicity *in vitro*, due to the poor concentration of LV vectors we determined to continue with standard concentration methods.

Further optimisation of the PEG600/ $(NH_4)_2SO_4$ ATPS through tie line analysis is required to determine the most effective ratio of chemicals for exclusion of cellular contaminants. This system has been shown as the most effective for AV purification *in vitro* and has comparable *in vivo* recovery to standard purification practises. However, quantification of in vivo recovery should be performed. Further, contaminants should be excluded after purification direct

from cell lysates to produce a pure product for clinical downstream applications. Further analysis of tie lines and alterations of the proportion of PEG600 and $(NH_4)_2SO_4$ in the system may allow enhanced exclusion of contaminates. Fractionation of viral peak on the chromatogram and hollow fibre filtration can also be employed to isolate the virus in elutes. These avenues should be explored further to improve the purity of the virus obtained through this system.

The upper phase of PEG600/ $(NH_4)_2SO_4$ was determined as the most successful phase for AV purification. PEG600 is commonly used in a clinical setting as an emulsifier, commonly used in ointments and toothpaste and can readily be cleared by the renal system. The concentrations of PEG600 used for infections remain within IGD limits and can be readily cleared by kidneys or oxidised to form acidic metabolites (D'souza, Shegokar 2016). However, these systems were not found suitable for LV vectors partitioning. This suggests the suitability of this system and can therefore be readily translated to research and clinical based applications downstream for purification of AV vectors.

Chapter IV Optimisation of Lentiviral Vector Protocols and Sequencing Analysis

We have previously shown the optimised purification of AV vectors directly from infected cell lysates using ATPS. However, as this does protocol does not partition LV vectors effectively, standard harvesting and concentration protocols were used. AV vectors are episomal and therefore have not been reported in causing genotoxicity. In contrast, LV vectors integrate transgenes into the host genome. While this allows for permanent correction of genetic diseases, this can contribute towards adverse side effects. Gene transfer and related genotoxic adverse side effects will therefore be studied in human iPSCs and their derivatives. Due to the common use of LV vectors in non-clinical and clinical applications, there is a need for high titre LV production. Here, we show through reagent optimisation, an enhanced protocol for LV vector production through transient transfection. We further sequence these vectors to understand the potential for aberrant packaging by LV vectors which may be delivered upon gene transfer.

4.1. Background

iPSCs have been used in toxicity testing and therefore, these cells and their derivatives may provide a valuable model for genotoxicity. LV mediated gene therapy is a promising tool for correction of genetic diseases. 313 clinical trials have been registered using LV for gene delivery in 2021 to date, accounting for 10.1% of vectors used in clinical trials (Wiley 2021). LV particles are enveloped Class IV viruses containing positive sense, single stranded RNA. These viruses have multiple characteristics which are ideal for permanent gene transfer and present low immune responses, allowing increasing doses to be administered with little host inactivation. LV have also been found to package transgenes up to 8 kb, useful for large therapeutic transgene packaging. These viruses also allow permanent gene transfer due to transgene integration within the host gDNA with success in clinical trials, including correction of adenosine deaminase deficiency and transgene persistence shown up to 12 years post treatment (Muul, Tuschong et al. 2003). A further advantage of using LV above RV for gene therapy, is that the virions transduce both proliferating and quiescent cells (Lewis, Hensel et al. 1992, Naldini, Blömer et al. 1996). The manipulation of these viruses for gene therapy is aided by pseudotyping LV for cell specific tropisms. Commonly, VSV-G is used for target of a broad range of cell types (Burns, Friedmann et al. 1993). Specific cell targeting has been shown as effective, for instance pseudotyping LV using the Rabies envelope glycoprotein (RV-G) which allows successful retrograde axonal transduction of murine neurons (Hislop, Islam et al. 2014). These manipulations and characteristics of LV make them ideal candidates for use in gene therapy.

4.1.1. LV Vector Safety

The generation of LV for safe use in gene therapy employs split genome packaging to produce replication deficient vectors. Provision of genes in *trans* allows transgene and accessory proteins to be packaged within producer cells. This system indicates the methodology employed for modifications of envelope proteins, with specific pseudotyping of the virus particles dependent on envelope proteins expressed. The split genome packaging method for vector production also reduces the risk of producing replication competent LVs (RCLs). A major safety feature of these vectors is their replication deficient nature. RCLs pose a risk to patients due to the risk of associated oncogenesis. This was first identified in non-human primate cells transduced with RV vectors in preparation for bone marrow transplantation. Due to the propagation of RCLs, 38% of animals treated developed T cell neoplasms (Donahue, Kessler et al. 1992). As such, the presence of RCLs in clinical gene therapy products is tightly regulated before clinical approval of use in patients (Farley, McCloskey et al. 2015). This is an important feature to examine, despite the presence of RCLs being spontaneous and rare. Probability analysis of 375 T cell samples have shown patients would need to be observed for 52.8 years to observe a single RCL event. (Marcucci, Jadlowsky et al. 2018). With sensitive gene rescue assays in use for RCL detection, split genome packaging of LV addresses immediate safety concerns of using these viruses for gene therapy.

The development of split genome packaging has provided a further understanding of HIV-1 biology and has provided the basis for the removal of unnecessary genes for HIV-1 production. Wild type HIV-1 contains structural genes, namely *gag*, *pol* and *env* and two regulatory genes involved in viral replication (*tat* and *rev*). *Tat* and *rev* both regulate expression of viral genes. Tat binds to the Transactivation Response Element present in RNA transcripts to recruit cyclin kinases allowing transcriptional elongation via RNA Polymerase

II complex (Wei, Garber et al. 1998). Subsequent to rev production reaching a threshold, this protein binds to the rev response elements in unspliced and spliced viral RNA transcripts to bind to nuclear export proteins, allowing shuttling of transcripts to the cytoplasm for translation. HIV also contains four accessory genes; *vif*, *vpr*, *vpu* and *nef*, required for *in vivo* replication but which are not essential to replication (Fields, Knipe et al. 2007). These elements have been removed as LV vector production has been developed. First generation LV are generated through expression of three plasmids encoding for the transgene, core, and accessory proteins and the third encoding for envelope proteins (Naldini, Blömer et al. 1996). Second generation LV provides the vector transgene on one plasmid, *gag*, *pol*, *tat*, and *rev* on a second and envelope proteins on a third (Zufferey, Nagy et al. 1997). Third generation LV system provides an additional packaging plasmid to provide rev in *trans*. *Tat* has been shown to be unnecessary if replaced by constituently active promoter sequences and in *trans* expression of *rev* contributes to high titre LV production (Naldini, Blömer et al. 1996, Dull, Zufferey et al. 1998). The removal of these genes has contributed to safer LV systems of gene delivery and increased the space for transgene insertion.

Further modifications, including the development of SIN LTRs and inducible systems increase the safety of LV. SIN LTRs contain mutated or deleted U3 regions of the LTR, preventing the region acting as a promoter. This reduces the risk of insertional mutagenesis, with an internal promoter allowing transgene expression (Zufferey, Dull et al. 1998). Inducible systems allow the expression of the viral transgene when exposed to specific chemicals, such as tetracycline (Blomer, Naldini et al. 1997). This on/off system allows the expression of the transgene in specific conditions. These modifications enhance of LV safety, by reducing the risk of viral spread, genotoxicity, and aberrant expression of transgenes.

4.1.2. LV Genotoxicity

While the biology of LV, such as HIV-1, has been understood very well there is a need to understand the method of integration and associated genotoxic effects further. While it has been shown that LV integrate in a semi random manner in host gDNA, these viruses have been shown to integrate preferentially in transcriptional units (Wu, X., Li, Crise, and Burgess 2003). The integration profile of LV is understood to be safer than RV, in particular using SIN configuration LTRs (Montini, Cesana et al. 2009). However, oncogenesis has been observed with these vectors in various models. In patients, it has been shown that insertional mutagenesis can lead to oncogenesis in treated X-SCID patients developing leukemias (Howe, Mansour et al. 2008, Themis, Waddington et al. 2005) This indicates the need to understand the safety of these vectors further for use in gene therapy, as while the underlying genetic condition may be corrected, insertional mutagenesis and oncogenesis is undesired.

4.1.3. LV Vector Production and Purification

A major bottleneck for use of LVs in production of high titre preps for downstream applications (Segura, Mangion et al. 2013). Transfection protocols for LV vector generation generally use adherent cell culture which limits space due to cell culture vessel surface area and incubator seizes. This method requires intensive maintenace of cells, which is not ideal for large scale vector production for the clinic (Segura, Mangion et al. 2013, Merten, O. W. 2015). Therefore, attempts should be made to generate high titre vectors in a cost and time effective manner.

Protocols for suspension cultures of HEK293T cells have been developed to enhance scale up of LV production process for downstream applications. HEK293T cells adapted for suspension growth (HEK293TF) have been developed and adapted for growth in serum free conditions (Ansorge, Lanthier et al. 2009). While suspension cultures allow increased surface area for cell growth and viral production, there are still issues that persist in scaling up the process. HEK293TF cells have shown a maximum scalability of up to 3L and as such, further developments are required for clinical advancement.

LV production through plasmid transfection still requires downstream processing for concentration and purification of viral particles. Plasmid transfection of HEK293T cells uses various reagents to carry pDNA into cells. Various reagents such as PEI and calcium phosphate can be used to transfect cells with pDNA for LV vector production (Ding, Kilpatrick 2013)**.** However, these methods remain toxic to cells. Various commercial reagents have been marketed as successful and low toxic alternatives and should be explored to determine their efficacy in LV vector production.

Concentration of LV particles usually occurs via ultracentrifugation and resuspension of the viral pellet in a reduced volume of serum free medium (Ichim, Wells 2011). Other methods of concentration include ion exhange, size exclusion or affinity absorbtion chromatography (Segura, Garnier et al. 2007, Kutner, Puthli et al. 2009, Kutner, Zhang et al. 2009, Transfiguracion, Jaalouk et al. 2003). However, issues persist with low viral yields and limited scalability using these alternative methods of concentration.

While concentration of LV provides smaller volumes of virus required to reach desired MOIs, impurities must be removed for clinical grade vector production. Presence of contaminating DNA and proteins from cell lysates can be detected through molecular biology techniques, including silver staining and western blots (Merten, O-W, Charrier et al. 2011). Multiple techniques have been

published for removal of contaminants in LV preps. The addition of a sucrose cushion when concentrating viral supernatants through ultracentrifugation has been shown to concentrate vectors to high titres and remove the majority of impurities (Boroujeni, Gardaneh 2018). Anion exchange chromatography, using convective interaction media monolithic columns, purify clarified lysates with 80% infectious particles remaining. Low concentrations of impurities are detected, with an additional benzoase step removing 99% of contaminating DNA. Despite relative success in removing impurities, issues with yield remain with only 36% of input virus being purified (Bandeira, Peixoto et al. 2012). Hollow fibre filtration has also been shown to successfully purify contaminants from LV vectors in supernatant. This system allows the automated production LV through the culture and filtration of transfected HEK293T cells. Cells remain viable and high yields of vectors have been purified, with up to 98% of initial viral input. However, the capacity of the Terumo Quantum Cell Expansion System capacity is 180 ml and is difficult to scale up. Multiple runs of the system are also required to purify all contaminants (Sheu, Beltzer et al. 2015). This identifies advantages to using systems for purification of contaminants from LV in supernatant but further development is required to increase yield and capacity for large scale development.

Stable cell lines have been generated for long term LV production. Plasmid transfection of HEK293T cells is transient and thus does not allow continuous LV production. Stable cell lines have been generated through integration of viral genes within host gDNA. A number of thes cell lines have been generated including one for the production of LV carrying SIN LTRs flanking ILR2G for treatment of X-SCID (Throm, Ouma et al. 2009). STAR cells generate HIV-1 pseudotyped with RD114 at a titre of 10^7 TU/ml, which have shown efficient transduction of HSCs (Ikeda, Takeuchi et al. 2003, Relander, Johansson et al. 2005). WINPAC cells have also been generated to stably generate 10^6 TU/ml

SIN configuration HIV pseudotyped with RDpro. A further advantage of stable cell lines is that antibiotic selection can ensure a pure population of cells for LV production, demonstrated by blasticidin selection of positive WINPAC cells (Sanber, Knight et al. 2015). However, VSV-G is commonly used as to pseudotype HIV-1 due to its ubiquitous infection profile. As VSV-G has been shown to be toxic to cells, it has been difficult to establish a stable cell line for production of these vectors (Burns, Friedmann et al. 1993). Hu and colleagues overcame these issues to produce a stable inducible cell line producing VSV-G pseudotyped HIV, carrying a transgene flanked by SIN LTRs using tight regulation of VSV-G through tetracycline (Hu, Li et al. 2015). However, despite the advantages of stable producer cell line, the time scale for cloning cells (8-10 weeks), scalability, instability of clones and drop in titre over time (3 months) remain issues for expansion of these cells (Chaudhary, Pak et al. 2012).

This identifies challenges in large scale production and purification in the development of LV vectors which need to be addressed to enhance vector production.

LV vector pseudotyping enhances cell infection. VSV-G is commonly used to pseudotype vectors due the ubiquitous insertion profile of this glycoprotein. However, enhancers are often used to improve viral transduction through negation of electrostatic repulsion between the virion and cell surface membrane (Swaney, Sorgi et al. 1997). Polybrene is commonly used to complex with LV. This compound is useful as it presents low toxicity and an average three-fold increase in transduction efficiencies (Barrilleaux, Knoepfler 2011). Other compounds have been shown as successful for enhancing LV transduction. DAEA dextran has been shown to increase efficiencies in primary cells, greater than that of polybrene (Denning, W., Das et al. 2013). Commercially available reagents have also been marketed for low toxicity. Lentiboost (Sirion Biotech, Germany) has recently been shown as effective for

transduction of HSCs (Delville, Soheili et al. 2018). However, commercial reagents often are less cost effective than polybrene and require independent optimisation. The range of compounds available allows effective transduction of LV using polycations and chemicals to enhance LV transduction and increase gene transfer to target cells.

Sequencing of LV RNA is vital for diagnostics. LV RNA has been purified and sequenced for detection of viral infection in HIV patients to understand biosafety protocols and develop reference standards for infectivity (Bagutti, Alt et al. 2011, Mattiuzzo, Ashall et al. 2015). As HIV RNA has been sequenced, analysis of viral RNA packaged in vectors for gene therapy should be understood further to determine transcripts delivered alongside transgene in traditional viral preparations, which may be transferred to patients alongside therapeutic transgenes.

These bottlenecks in LV production and transduction will be addressed in the studies below.

Aims

The aims of this study are as follows:

- Optimisation of lentiviral vector production to produce high titre vectors
- Optimisation of lentiviral vector transduction to enhance efficiency of gene transfer
- Sequencing of lentiviral vector RNA to analyse contents of packaged viral RNA

4.2. Results

4.2.1. Plasmid Harvest Optimisation

In preparation for LV generation, plasmids carrying the *LacZ* transgene (LNTSbetagalW), *eGFP* transgene flanked by SIN LTRs (pHR), *eGFP* transgene flanked by full LTRs (pHV), accessory genes (pCMVR8.74) and *VSV-G$_{Ind}$* (pMD2.G) were harvested from transformed E. Coli bacterial cells. All plasmids were provided by Dr Michael Themis (Brunel University London) except pHV, which was a kind gift from Dr Y. Takeuchi (University College London). pHV pDNA was transformed into *DH5α* E. coli cells, as described in previously (Sections 2.2.1.4). Briefly, 500 ng pDNA were transformed into bacteria via a heat shock protocol before plating on LB- agar, using an ampicillin antibiotic marker. Single colonies were picked and grown in LB medium overnight before harvesting (Section 2.2.1.6).

Plasmid harvesting was optimised to obtain high yields of pDNA. Briefly, glycerol stocks of transformed *DH5α* or *Stbl3 E. coli* cells were initially grown in 10ml PDM overnight and grown up in 500ml medium for a further 24 or 48 hours before harvesting.

Plasmid growth was optimised for maximum yields of pDNA (Figure 4.1). Yields of LNTSbetagalW, pCMVR8.74 and pMD2.G were compared after incubation for a total of 48 or 72 hours. Analysis of DNA concentrations via spectrophotometry determined higher yields were obtained after 72 hours. However, the increase in plasmid concentration varies between plasmid constructs. A 2.86-fold greater concentration of LNTSbetagalW was harvested after 72 hours growth (2366.63 ng/µl) in comparison to 48 hours (613.10 ng/µl). pCMVR8.74 was harvested at 8.82-fold greater concentration after 72 hours growth (739.82±284.94 ng/µl) in comparison to 48 hours growth (75.33±24.38

ng/µl). For pMD2.G, a 21.35-fold increase in pDNA was harvested after 72 hours growth (1348.15±1206.68 ng/µl) in comparison to 48 hours (60.32±46.53 ng/µl). While there is no significant difference between concentrations, the higher yield of plasmid DNA is useful for downstream assays. Using this optimised protocol, pHR was harvested at 1231.22±456.60 ng/µl and pHV at 821.83 ng/µl.

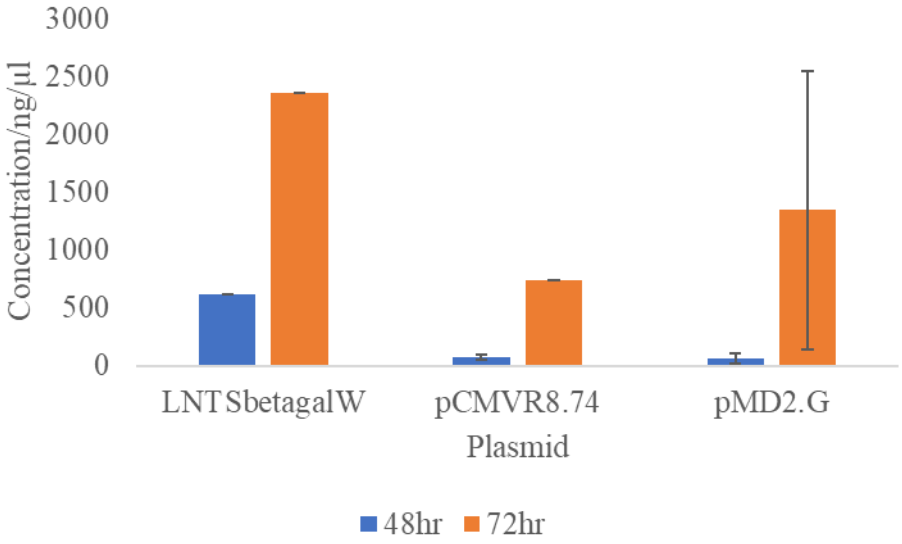


Figure 4.1- Plasmid harvest optimisation. Average optical density readings of plasmid harvests optimised for incubation times. Growth of bacterial cells for 48 hours increase plasmid yield. LNTSbetabalW concentration determined as 613ng/µl (48 hours) and 2366.63ng/µl (72 hours) (n=1). pCMVR8.74 was quantified as 75.33±24.38 (48 hours) and 739.82ng/µl (72 hours) (n=6). pMD2.G was quantified as 60.32±46.53 (48 hours) and 1348.15±1206.68 (72 hours) (n=3)

The purity of pDNA is not altered by increasing incubation times. The Abs_{260}/Abs_{280} ratio for LNTSbetagalW is calculated as 1.92 (48 hours) and 1.85 (72 hours). Similarly, pCMVR8.74 measured a 1.91±0.03 Abs_{260}/Abs_{280} ratio (48 hours) and 1.88±0.01 (72 hours). The Abs_{260}/Abs_{280} ratio for pMD2.G was also found to be similar with 1.66±0.24 (48 hours) and 1.86±0 (72 hours) calculated. This suggests that the pDNA purity is not affected from increasing

incubation times with all Abs_{260}/Abs_{280} ratios falling within the acceptable range (1.8-2.0) (Hassan, Husin et al. 2015, Glasel 1995).

This optimisation of incubation times for bacterial growth demonstrates a higher yield of pDNA is obtained after 72 hours growth. This is useful for harvesting a greater concentration of pDNA for plasmid transfection.

4.2.2. Plasmid Verification

Plasmid integrity was verified by restriction enzyme digestion and gel electrophoresis. Plasmids were digested and separated on gel electrophoresis using a 0.6% agarose gels, as described in Sections 2.2.1.7-2.2.1.8. Uncut pDNA was run as in parallel as a negative control.

The integrity of all plasmids harvested were verified through restriction enzyme digestion and separation using agarose gel electrophoresis. All restriction digests produced correct fragment sizes suggesting that extended growth of bacteria for a total of 72 hours does not decrease the sequence integrity of plasmids.

4.2.3. Genejuice® Transfection Reagent Optimisation

Three transfection reagents, formulated using various chemicals, were compared to optimise LV production. PEI, a polymer, Fugene® 6, a proprietory non-liposomal reagent and Genejuice®, a proprietory polyamine and cellular protein reagent were compared for plasmid transfection and LV production.

While transfection protocols for LV production using PEI and Fugene® 6 have been previously reported and optimised, few protocols have been developed for Genejuice® transfection of cells and therefore, this reagent requires

optimisation. Optimisation of Genejuice® as a transfection reagent for LV production was performed as described in Section 2.2.2.10. HEK293T cells were transduced using 8 µg in total of pHR, pCMVR8.74 and pMD2.G plasmids in a 4:3:1 ratio, with varying ratios of pDNA to Genejuice® (1:1, 1:2, 1:3, 1:6 and 1:8). The medium was replaced 24 hours post transfection and collected every 24 hours for 72 hours post medium change. Transfected cells exhibit varying levels for fluorescence emission over time suggesting different levels of plasmid transfection (Figure 4.2).

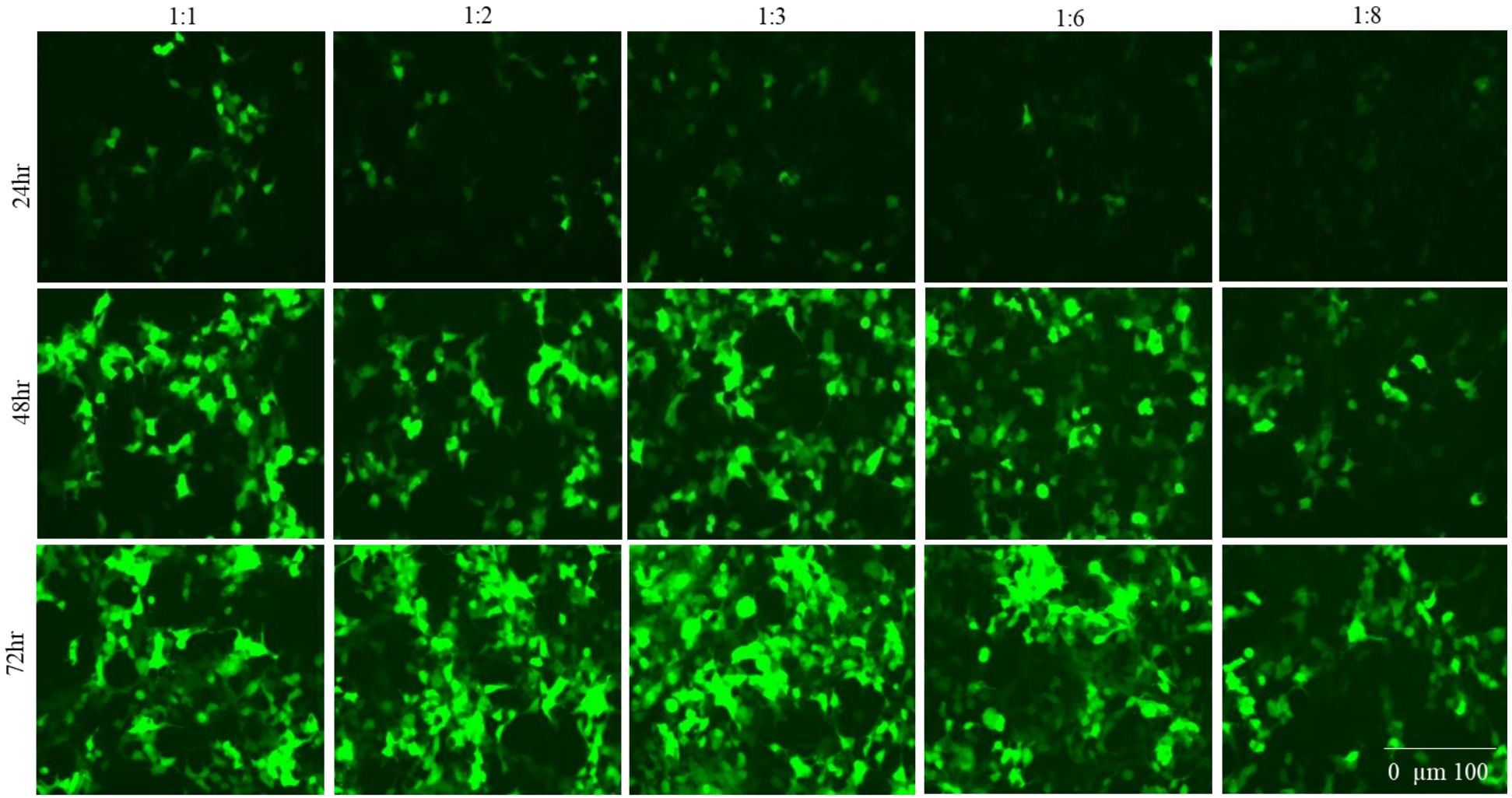


Figure 4.2- Optimisation of Genejuice® transfection reagent. Various ratios of pDNA: reagent were used for transfection of HEK293T cells for LV production. Fluorescent microscopy of transfected cells shown over 72hrs for virus collection. Scale bar shown between 0 and 100µm, x20 magnification

Cell viability was analysed 72 hours post transfection to determine optimal concentration to maintain cellular health (Section 2.2.2.3). Analysis of viabilities indicates increasing cocentrations of Genejuice® increase cellular toxicity over time (Figure 4.3). While the viability of transfected cells is unchanged between 1:1 (73.50±6.50%) and 1:2 (73.5±2.47%) ratios of pDNA to reagent, this steadily decreases with increasing concentrations of Genejuice® added. 63.50±13.50% (1:3), 51.00±5.00% (1:6) and 34.00±2.00% (1:8) viability was measured after 72 hours. In comparison to untreated cells (86.00±1.00%), viabilities taken using cells transfected with a 1:3 ratio or greater decrease consistently. This suggests that a 1:1 and 1:2 ratio of pDNA to Genejuice reagent is the least toxic to cells and useful for long term culture of transfected cells.

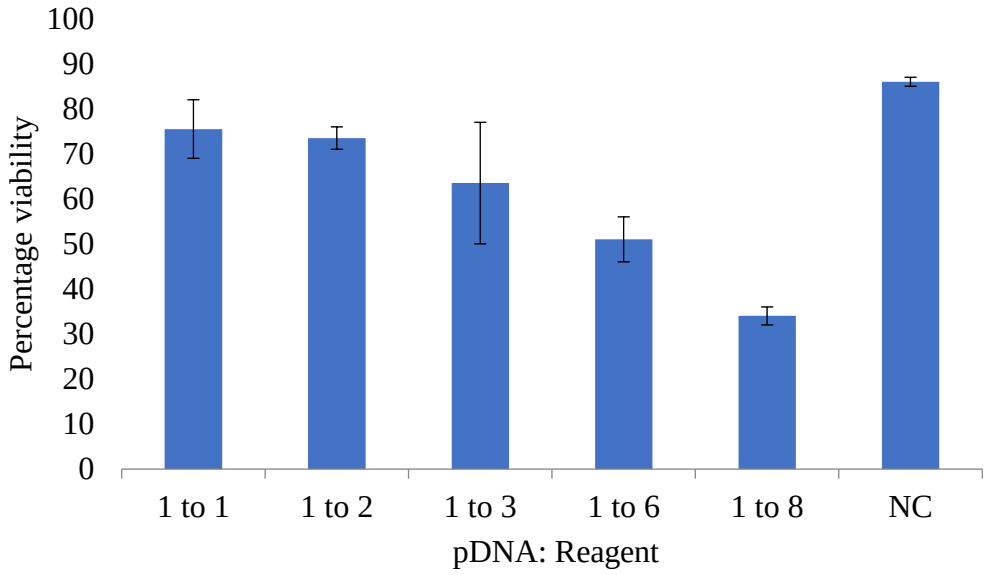
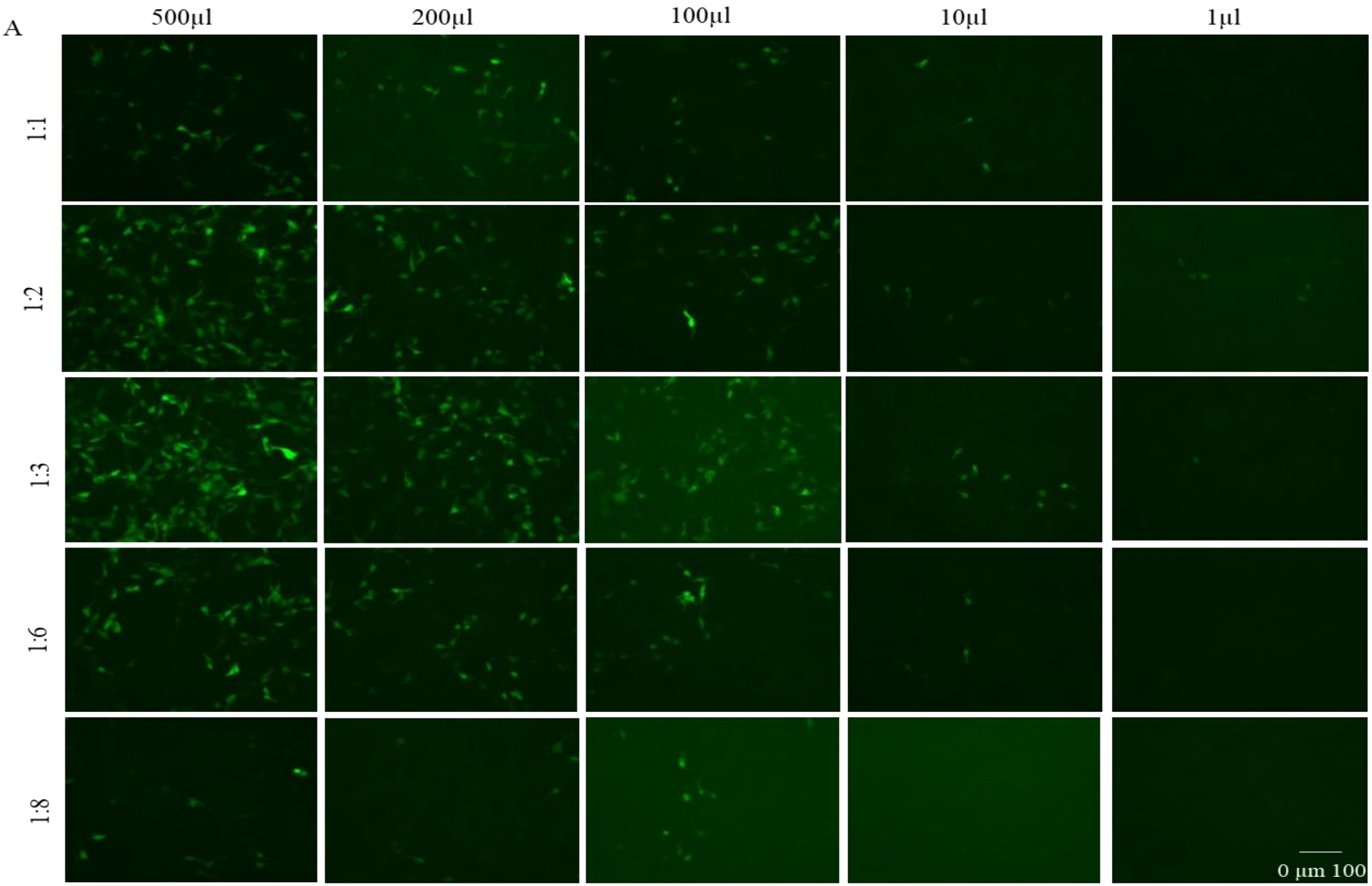


Figure 4.3- Viability of transfected HEK293T cells for optimisation of Genejuice ® transfection reagent. Various ratios of pDNA: reagent were used for transfection of HEK293T cells for LV production. Viability of transfected cells measured 72hrs post transfection via trypan blue exclusion assay. SEM shown (n=2)

Conditioned mediums containing LV were titrated to determine LV titre using various pDNA: Genejuice® (Figure 4.3). Crude titration of LV was performed as described in Section 2.2.2.11. LV titres were determined as $5.69x10^6$ (1:1), $3.19x10^6$ (1:2), $1.23x10^7$ (1:3), $5x10^5$ (1:6), $2.77x10^5$ (1:8) TU/ml. While the highest concentration of LV was produced using a 1:3 ratio, this ratio also showed an increase in cellular toxicity. Low titres estimated from 1:6 and 1:8 production is likely due to an increase in cell death. Therefore, a 1:2 ratio was determined as the optimal concentration for high titre LV production and maintaining health of transfected HEK293T cells over time.


A
500µl
200µl
100µl
10µl
1µl
1:1
1:2
1:3
1:6
1:8
0 µm 100

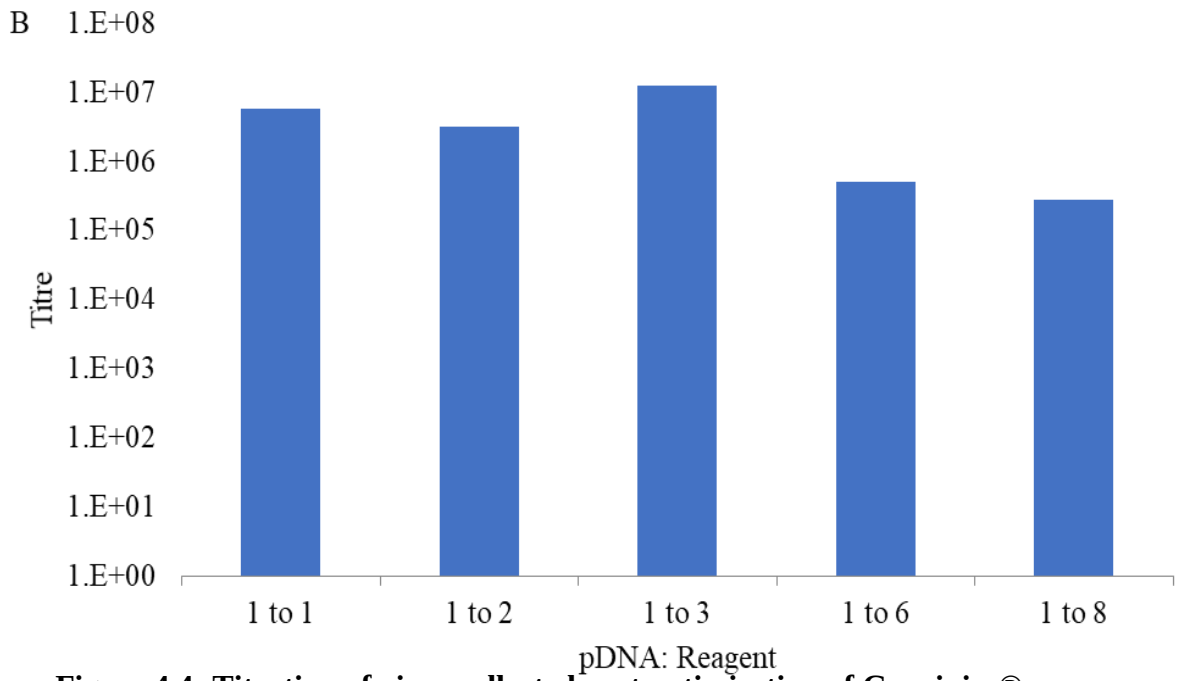


Figure 4.4- Titration of virus collected post optimisation of Genejuice® transfection reagent. Various ratios of pDNA: reagent were used for transfection of HEK293T cells for LV production. A- Various aliquots of crude supernatant were used for transduction of cells to determine titre, with cells images under fluorescent light after 72 hours, x20 magnification and scale bar shown between 0 and 100µm. B- Cells were analysed for GFP expression using flow cytometry 72hrs post transduction to determine titre (n=2)

4.2.4. Transfection Reagent Viability

LV vectors can be generated through transfection using PEI, Fugene® 6 or Genejuice®, in a scalable method for downstream applications. The viability of transfected HEK293T cells must be analysed to determine cytotoxicity after transfection using these various reagents. Viability of HEK293T cells transfected for the production of LacZ LV using PEI, Fugene® 6 and Genejuice® were assayed over 72 hours post transfection in parallel (Figure 4.5). HEK293T cells transfected to produce LV using PEI measured a viability of 88.63±3.29% (24 hours), 88.50±0.71% (48 hours) and 88.88±1.43% (72 hours) over harvest time. When using Fugene® 6 as a transfection reagent, cells showed a viability of 94.88±0.90% (24 hours), 86.63±5.69% (48 hours) and

84.13±4.67% (72 hours). HEK293T cells transfected for LV production using Genejuice® presented a viability of 95.75±0.97% (24 hours), 96.50±0.79% (48 hours) and 93.00±0.87% (72 hours). Comparison of transfected cell viability to untreated cells indicates a significant drop in viability in cells transfected using PEI or Fugene® 6, though this decrease in viability is not as drastic in comparison to PEI ($P<0.05$). However, cells transfected using Genejuice® did not show a statistically significant reduction in viability. Viability of untreated HEK293T cells were measured as 94.88±0.89% (24 hours), 95.96±0.42% (48 hours) and 92.96±0.49% (72 hours). While it is widely accepted that PEI is toxic to cells and is therefore not suitable for long term viral production, Fugen®e 6 and Genejuice® are commercially available transfection reagents which are specifically targeted as non-toxic. Both Genejuice® and Fugene® 6 transfection present less toxicity than PEI. Genejuice® is shown as a successful transfection reagent for long term viral production.

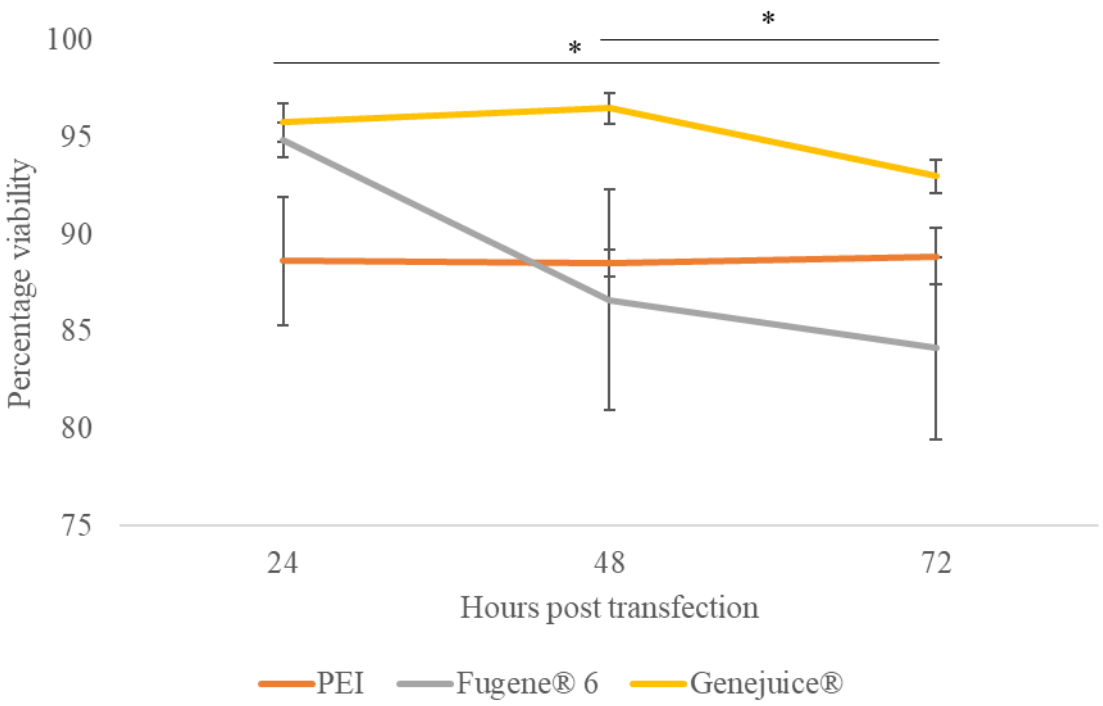


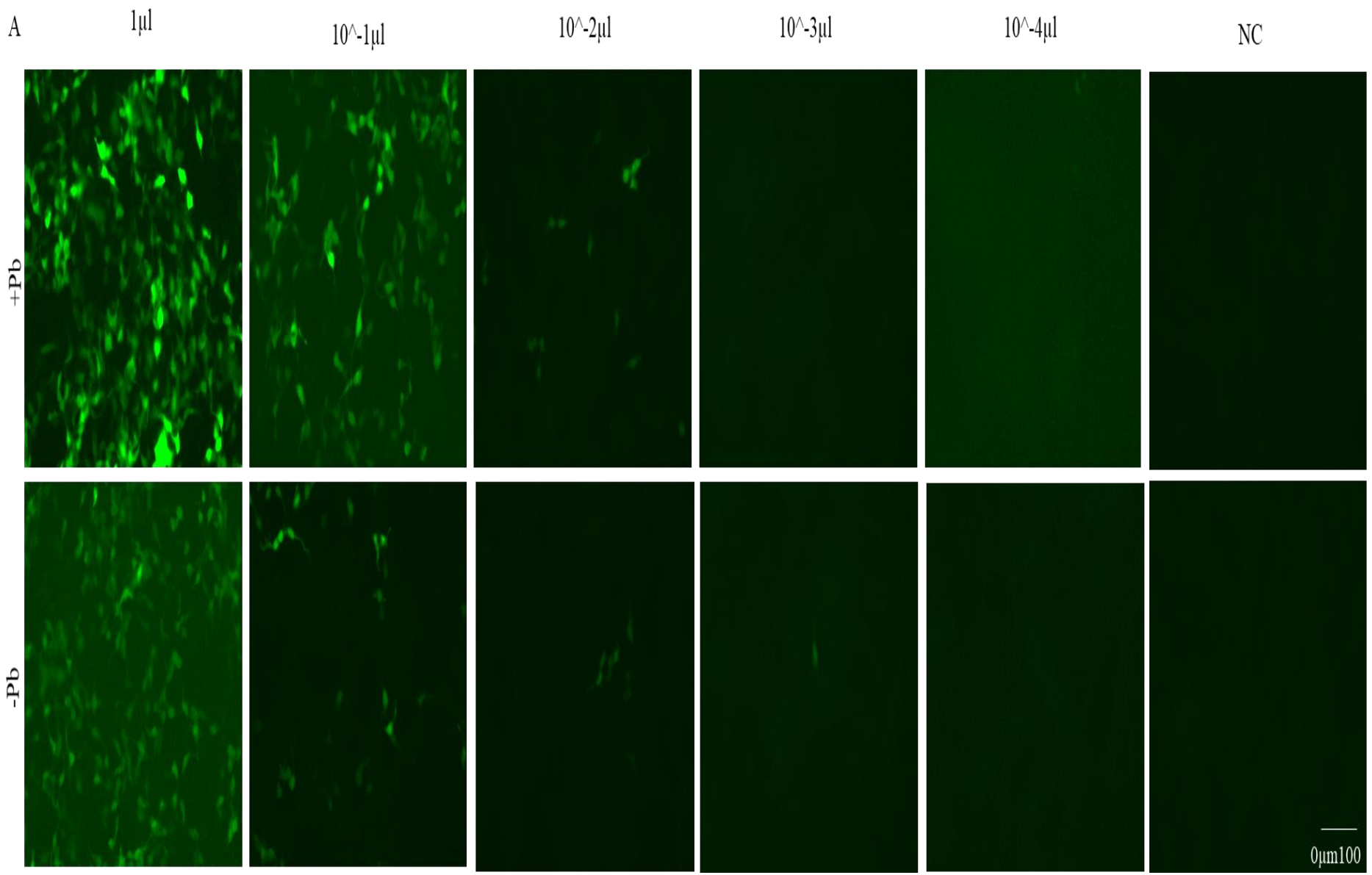
Figure 4.5- Viability of HEK293T cells after LV transfection using various reagents. Cells transfected to produce LV using three transfection reagents, with viability assayed over 72 hours post transfection using a trypan blue exclusion assay. SEM shown and significance denoted by an asterix ($P<0.05$) (n=4)

Both these results demonstrate the usefulness of Genejuice® for LV vector production, in producing high titre virus and maintain cell health.

4.2.5 Polybrene Optimisation

LV transduction of cells has previous been shown to be enhanced through use of a number of various chemicals. Polybrene has been used extensively to enhance LV vector transduction of cells. To confirm if polybrene does increase viral transduction in target HEK293T cells, serial dilutions of LV were added to HEK293T cells complexed with and without 5µg/ml polybrene (Section 2.1.2.11). Flow cytometry analysis of transduced cells indicates a higher percentage transduction when complexing virus with polybrene for transduction (Figure 4.6). Average percentage transduction after polybrene complexation

was determined as 66.47±0.16% (1µl), 20.21±1.10% (10^{-1}µl), 3.39±0.08% (10^{-2} µl), 2.86±0.12% (10^{-3}µl) and 1.36±0.29% (10^{-4}µl). Similar dilutions of virus without polybrene complexation presented decreased transduction efficiencies with 43.29±0.76% (1µl), 5.72±0.40% (10^{-1}µl), 1.40±0.40% (10^{-2} µl), 0.73±0.16% (10^{-3}µl) and 0.73±0.14% (10^{-4}µl). While the differences are not significant ($p>0.05$), the 2.92-0.54-fold increase in transduction efficiency after complexation with polybrene demonstrates its efficacy in improving viral transduction.


A
1µl
10^-1µl
10^-2µl
10^-3µl
10^-4µl
NC
+Pb
-Pb
0µm100

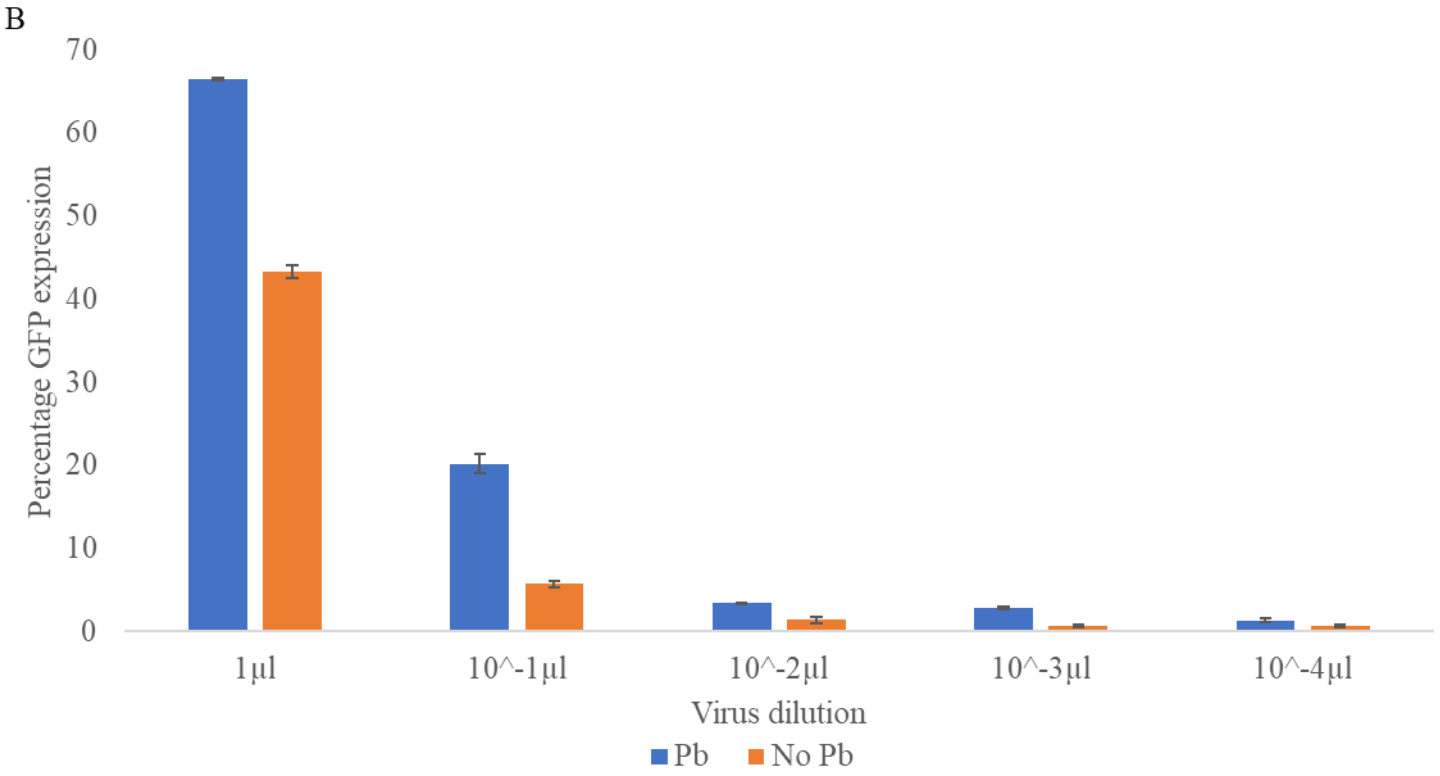
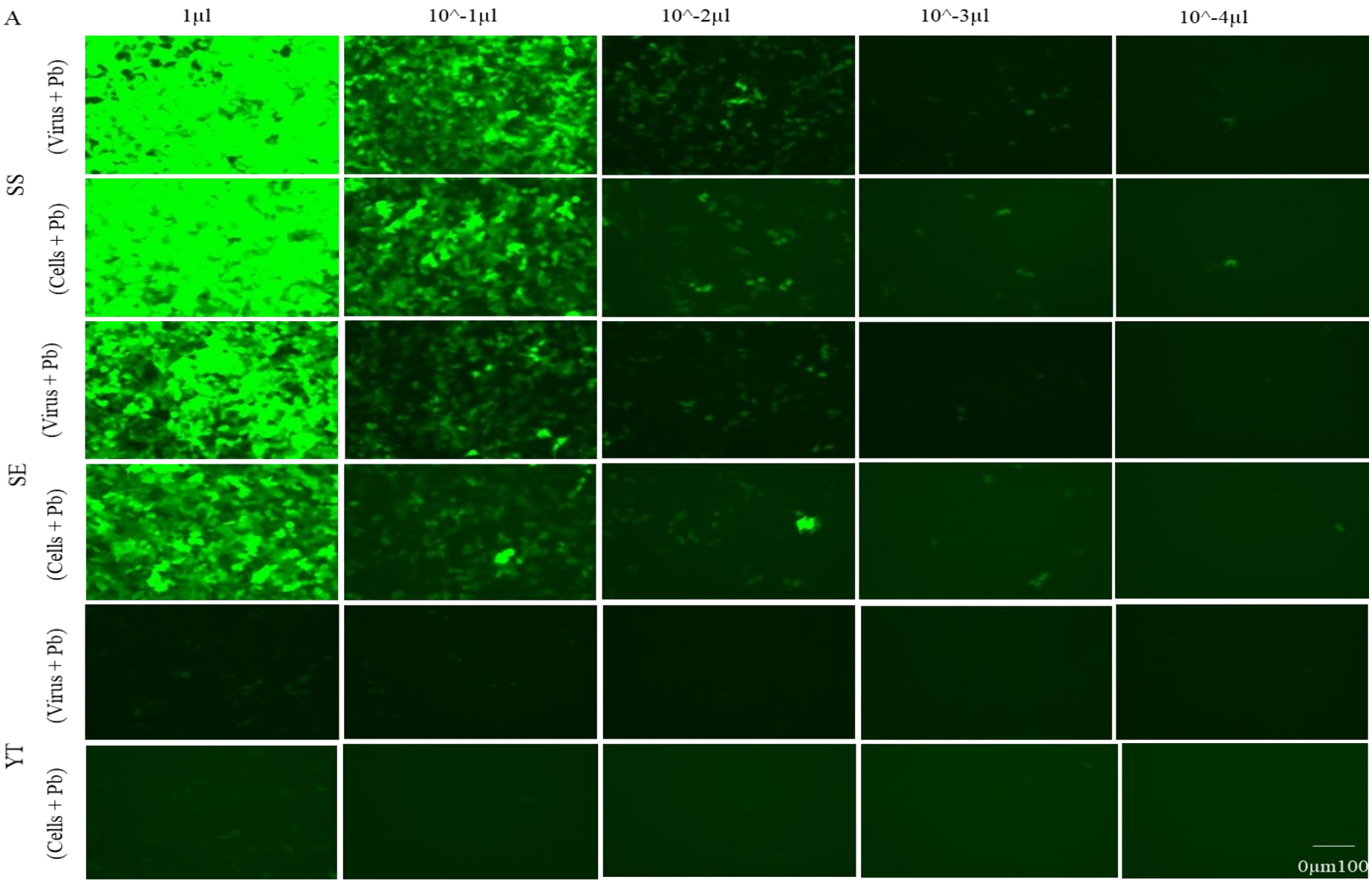


Figure 4.6- LV transduction enhancement using polybrene (Pb). A- Fluorescent microscopy images of cells transduced with serial dilutions of pHR LV complexed with and without 5 µg/ml polybrene at logarithmic dilutions of virus, imaged 72 hours post transduction. Scale bar between 0 and 100µm shown, x20 magnification. B- Transduced cells were analysed using flow cytometry to quantify GFP expression. n=2

Alternatively, polybrene incubation with cells can increase LV transduction (Barrilleaux and Knoepfler, 2011). This alternative protocol for lentivirus transduction was explored using three different batches of virus, including two batches produced by independent laboratories (Figure 4.7). Cells were incubated with 5 µg/ml polybrene upon seeding and transduced with serial dilutions of virus. Similarly, serial dilutions of virus were prepared and incubated with 5 µg/ml polybrene for 20 minutes at room temperature before transducing cells. Cells were incubated for 72 hours before end point analysis of GFP expression via flow cytometry to determine titre. Comparison of titre after seeding cells with polybrene and incubating virus with polybrene indicates higher rates of transduction after complexing virus with polybrene. Comparison of viral samples made in house indicated a 2.39-fold higher titre when complexing the virus with polybrene (3.09×10^{10} TU/ml) compared to incubating cells with polybrene (9.12×10^{9} TU/ml). The titre of virus provided by Prof. Brian Bigger (University of Manchester) showed a 0.91-fold higher titre of virus complexed with polybrene (9.07×10^{9} TU/ml) compared to cells incubated with polybrene (4.75×10^{9} TU/ml). Similar results were also obtained with virus provided by Dr Yasu Takeuchi (UCL), with a 0.81-fold increase in titre when complexing virus with polybrene (1.23×10^{8} TU/ml) compared to cells (6.80×10^{7} TU/ml). Comparative results of viral titrations indicate a higher rate of transduction can be achieved from complexing virus samples with polybrene.


A
1µl
10^-1µl
10^-2µl
10^-3µl
10^-4µl
SS
(Virus + Pb)
(Cells + Pb)
SE
(Virus + Pb)
(Cells + Pb)
YT
(Virus + Pb)
(Cells + Pb)
0µm100

B

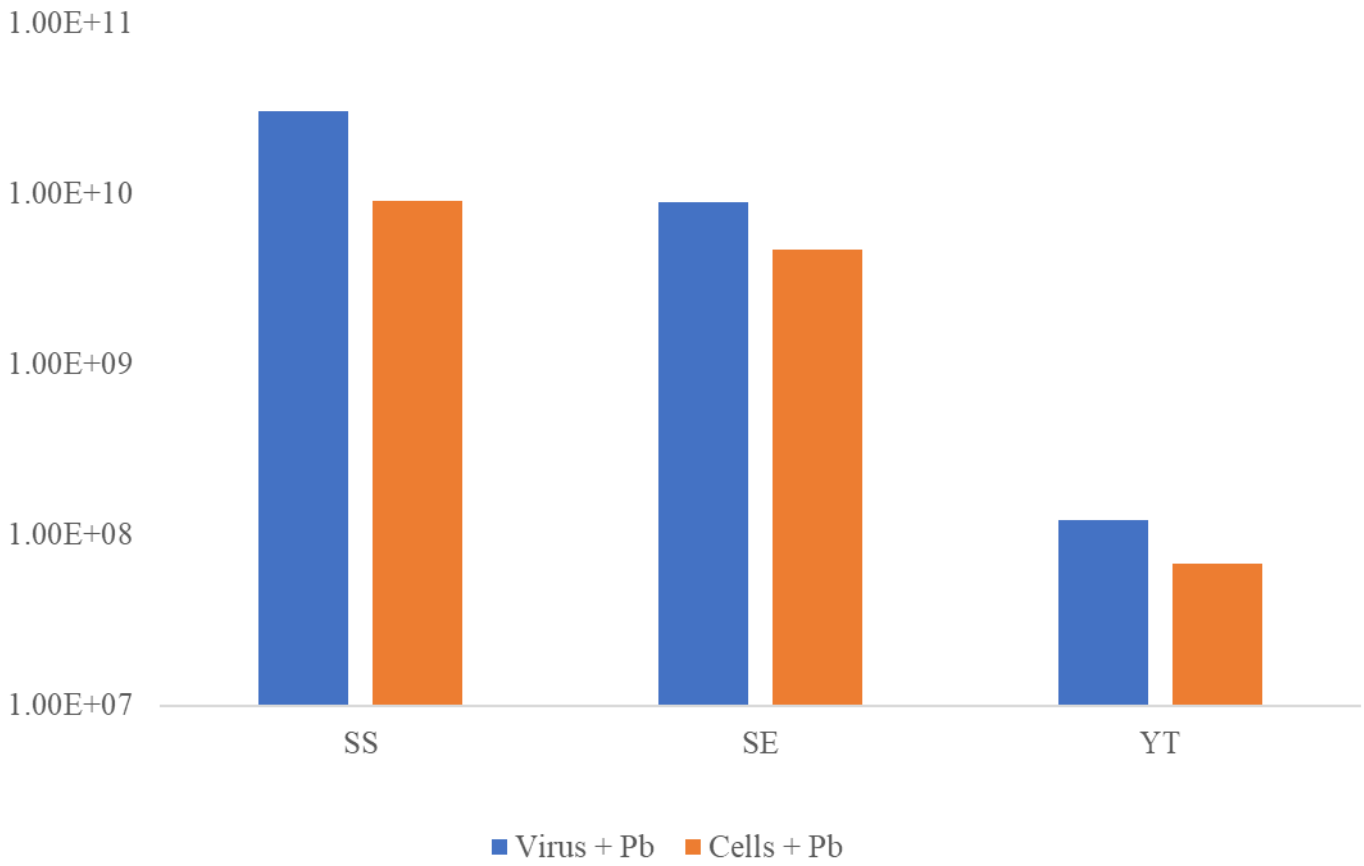
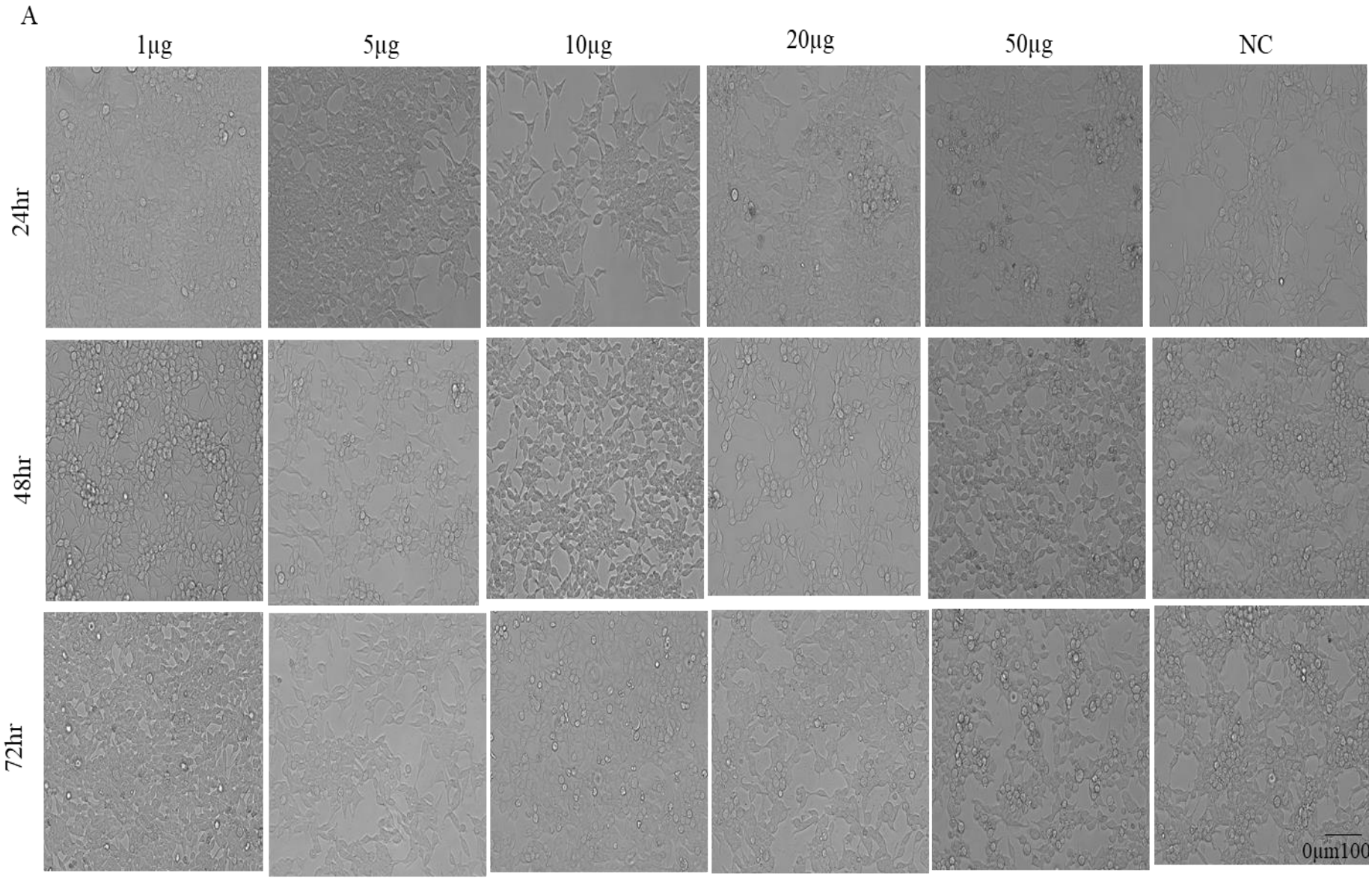


Figure 4.7- Pb complexation optimisations for LV titration. Cells were incubated with Pb upon seeding or virus complexed with polybrene prior to transduction. A- Fluorescent images of cells taken 72 hours post transduction, x20 magnification and scale bar between 0 to 100µm shown. B- Transduced cells were analysed via flow cytometry to determine viral titre. Variable titres were estimated between conditions, with higher titres observed after complexation of LV with Pb prior to transduction. Various batches of LV were analysed, including virus prepared in house (SS LV), and virus previously produced by other labs (SE & YT) (n=2)

While polybrene is effective as an enhancer to viral transduction, it must be assessed if this chemical is also toxic to cells to determine safe concentrations of polybrene exposure to use for *in vitro* assays. 1 µg, 5 µg, 10 µg, 20 µg and 50 µg/0.5ml polybrene were added to HEK293T cells and incubated as previously described (Section 2.2.2.11). Cells were analysed for cell viability every 24 hours, for a total of 72 hours using a trypan blue exclusion assay in comparison to untreated cells (Section 2.2.2.3). No statistically significant drop in viability was observed between in comparison to untreated cells ($P>0.05$, Figure 4.8). Furthermore, cell viability remained stable over 72 hours. The viability of untreated cells was measured as 98±0.41% (24 hours), 94±1.45% (48 hours) and 98±0.48% (72hours). Viability of cells treated with 1 µg polybrene were quantified as 98±0.33% (24 hours), 97±0.41% (48 hours) and 98±0.41% (72hours). Treatment of cells using 5 µg polybrene were shown to not be toxic to cells with 99±0.29% (24 hours), 98±0.33% (48 hours) and 97±0.58% (72hours) viability over time. Cells treated using 10 µg polybrene were 98±0% (24 hours), 91±4.59% (48 hours) and 95±0.48% (72 hours) viable. 20µg polybrene treatment was not toxic to cells, with viability measured as 98±0.33% (24 hours), 95±1.33% (48 hours) and 96±0.50% (72 hours) over time. Cells treated with 50 µg polybrene were also healthy, with 96±0.25% (24 hours), 95±1.31% (48 hours) and 94±0.63% (72 hours) viability determined. These results indicate polybrene is non-toxic and a suitable enhancer for viral transduction, with the highest level of viral gene transfer observed after complexing LV samples with 5 µg/ml polybrene before transduction.


A
1μg
5μg
10μg
20μg
50μg
NC
24hr
48hr
72hr
0μm100

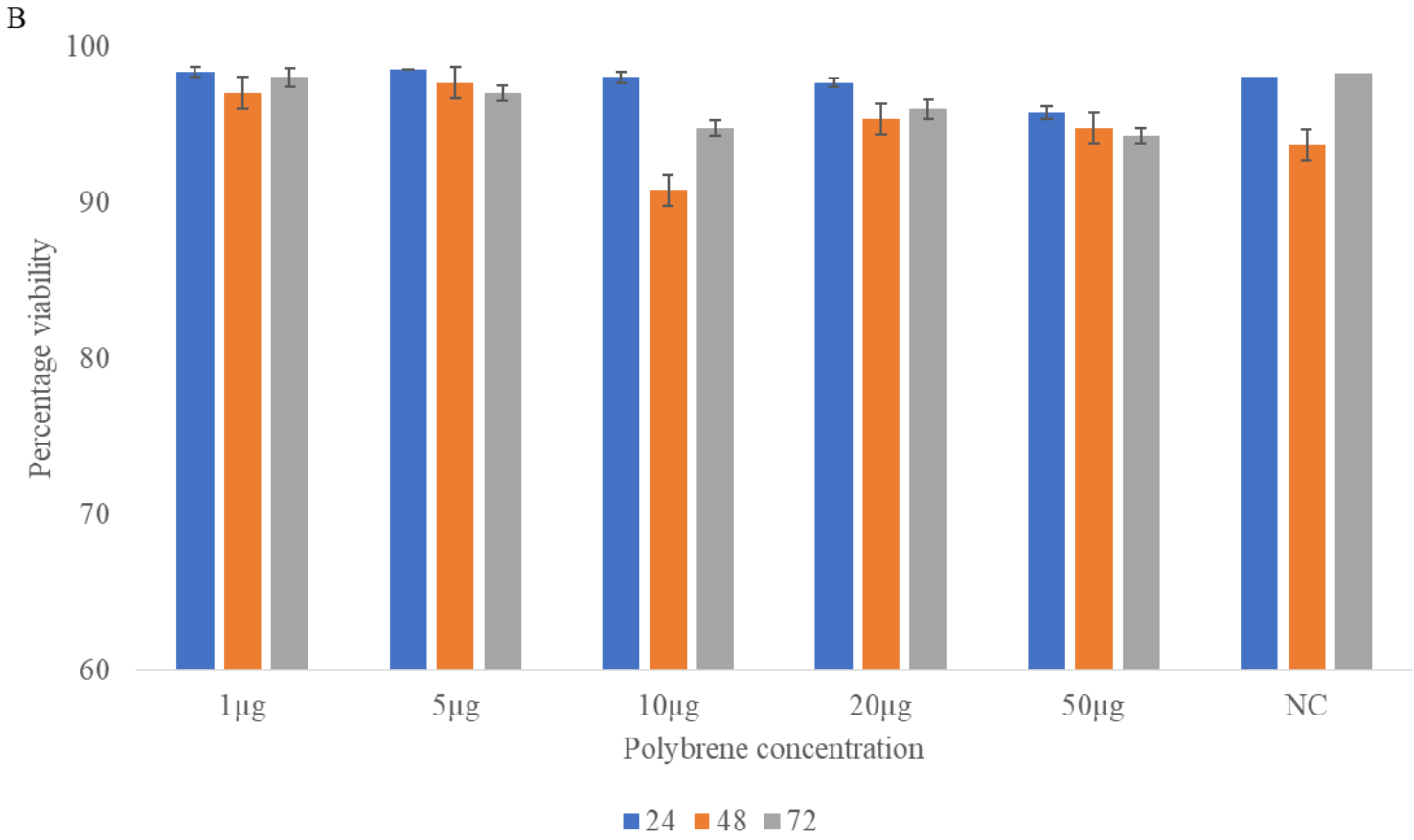


Figure 4.8- Toxicity of polybrene exposure. A- HEK293T indicator cells exposed to various concentrations of polybrene (µg/ml) to analyse toxicity over 72 hrs, imaged under brightfield light, x20 magnification with scale bar shown between 0 and 100µm. B- Viability measured every 24 hours, for 72 hours post treatment using a trypan blue exclusion assay. No decrease in cell viability was measured in comparison to untreated cells (n=4)

4.2.6. Nuvec® Cytotoxicity

Silicon nanoparticles have been shown to be useful in delivery of nucleic acids into cells for use in therapeutic avenues (Pearce, Mai et al. 2008, Park, Jeong et al. 2016). Nuvec® is a proprietory spiked silicon nanoparticle 80-500 nm wide, produced by N4Pharma, which has shown promise in clinical trials in delivery of nucleic acids for therapeutic avenues *in vitro* and *in vivo*. Due to the success of Nuvec® in nucleic acid transfection, we propose analysing the potential of Nuvec® to LV transduction to assess if Nuvec® can be used as an alternative enhancer of LV gene transfer.

Due to the unknown cytotoxic effects of Nuvec® in vitro on HEK293T cells, initial viability assays were used to confirm toxicity and recovery of cells over a 72-hour period, post Nuvec treatment. Four different production batches were provided by N4Pharma (NV00100028, NV00100026-28, NV00100032, NV00100033) which were produced using various methods though this study was performed blind. Nuvec® was reconstituted as previously described (Section 2.2.1.19). HEK293T cells were exposed to various concentrations (1-80µg/0.5ml) of Nuvec® for 24 hours in complete medium before replacement of medium and growth for a further 72 hours (Figure 4.9).

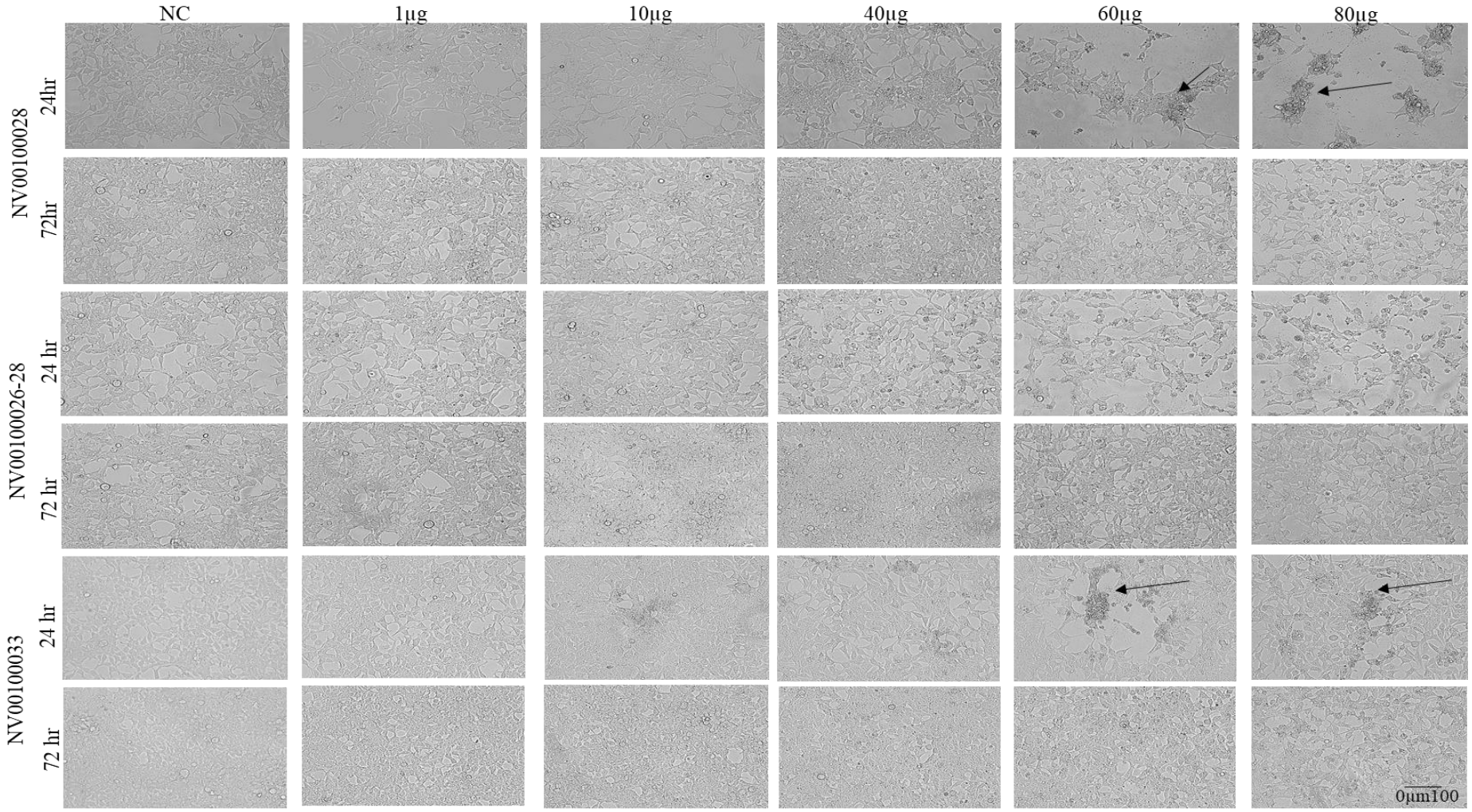
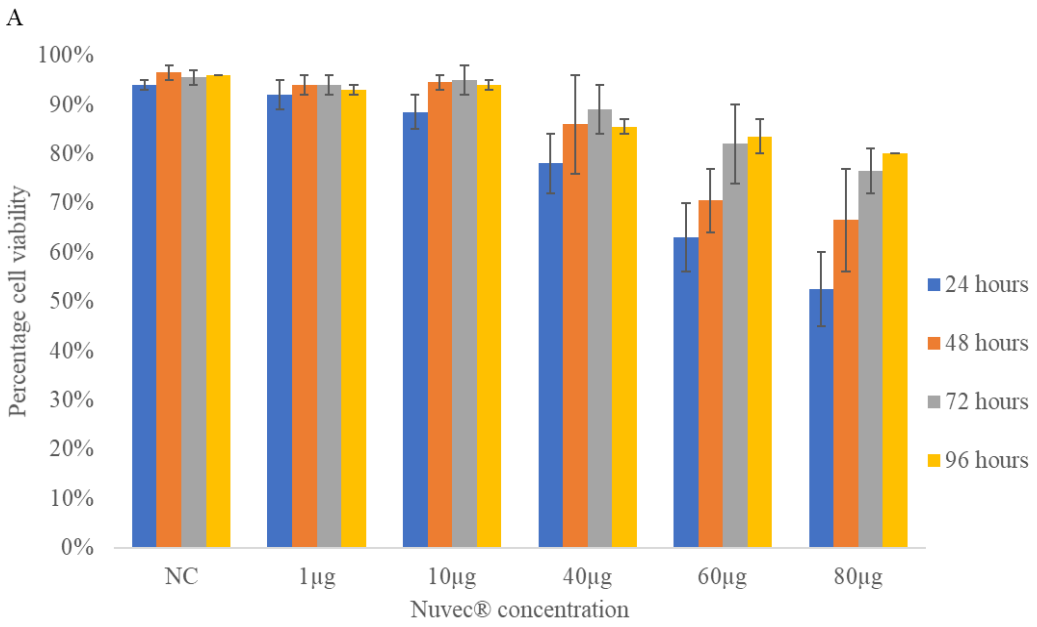


Figure 4.9- Toxicity analysis of Nuvec on HEK293T cells. Brightfield microscopy of cells treated NV00100028 for 24 hours, x20 magnification with scalebar showing 0 to 100µm, at 24 hours and 72 hours post treatment. Cell viabilities were taken every 24 hours for 72 hours post treatment. Cytotoxicity was observed at higher concentrations of Nuvec used (arrows), as determined by cell morphology. Cells recovered well over 72 hrs. (n=2)

Cell viability was quantified every 24 hours post treatment to assess cytotoxicity of Nuvec® and cell recovery (Figure 4.10). These results show an increase in toxicity is proportional to the concentration of Nuvec used in all batches of Nuvec®. HEK293T cell exposure to concentrations of Nuvec® show low cytotoxicity observed between 1-40 µg Nuvec® treatment. However, while 60-80 µg Nuvec® treatment shows increased cell toxicity, cells do recover over time. While the rate of recovery is consistent between batches of Nuvec® assayed, the initial decrease in cell viability after 24 hours post treatment differs. 60-80 µg NV00100028 and NV00100032 treatment present a <70% cell viability post 24 hours. NV00100026-28 treatment does not present a <70% cell viability 24 hours post treatment, except using 80 µg. Cell viability does increase over time, with all concentrations of Nuvec® used between varius production batches. It is important to note Nuvec® toxicity over the initial 24-hour period, with the aim to reduce initial cell death upon Nuvec® treatment. These results suggest that 1-40 µg of Nuvec® should be used for downstream assays to prevent cytotoxicity.



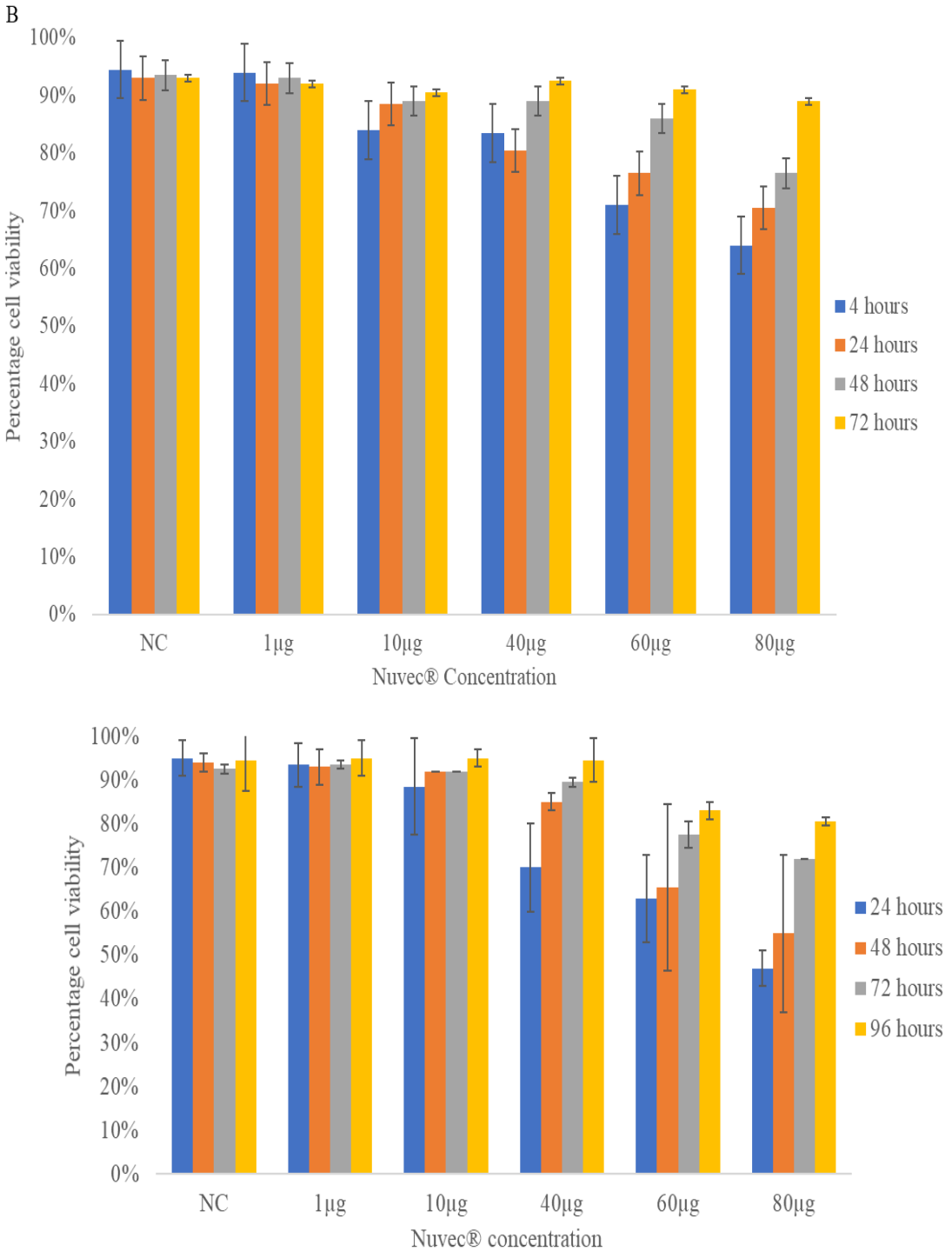


Figure 4.10- Time course analysis of cell viability after Nuvec® treatment. Viability measured every 24 hours for 96 hours post treatment with Nuvec®, using a trypan blue exclusion assay. Various batches of Nuvec® were used, including NV00100028 (A), NV00100026-28 (B) and NV00100032 (C). Cells treated with 60- 80µg showed the highest level of cellular toxicity though cells were able to recover to a healthy population over time (n=4)

4.2.7. Nuvec® Transfection and Transduction Efficiency

The plasmid binding capacity of various Nuvec® batches were analysed, according to manufacturer's instructions (Section 2.2.1.20). Briefly, 1 µg pDNA was incubated with 5 µg Nuvec® for four hours, before analysis. This showed the concentrations of bound pDNA to Nuvec® for various batches (Figure 4.11). pDNA concentrations bound to Nuvec® were quantified as 168±4.47 ng/µl (NV00100028), 174±12 ng/µl (NV001000026-28), 122±1 ng/µl (NV001000032) and 106 ng/µl (NV001000033). All results are statistically significant in comparison to the negative control ($P<0.05$). This suggests all Nuvec® batches successfully bind to pDNA, with NV001000026-28 as the most effective.

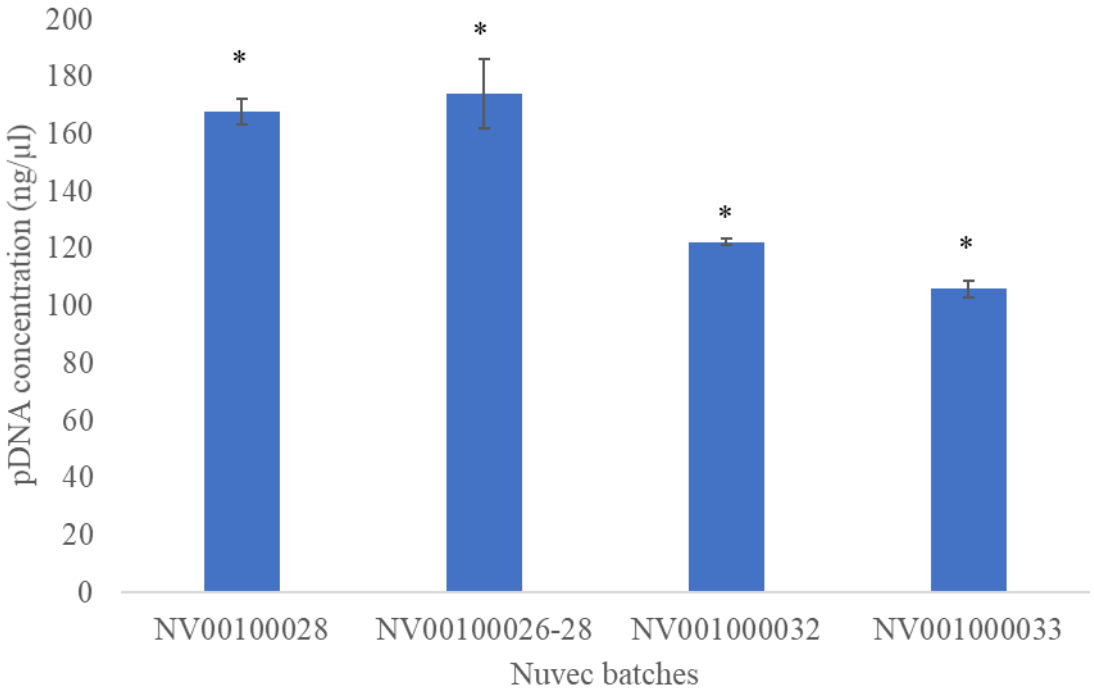


Figure 4.11- Nuvec®-pDNA (ng/µl) binding efficiency. Analysis of pDNA binding efficiency to various batches of Nuvec®. This demonstrates successful pDNA binding to Nuvec® nanoparticles, with variation between production batches. NV00100028: 168±4 ng /µl, NV00100026-28: 174±12 ng /µl, NV00100032: 122±1 ng /µl, NV00100033: 106±3 ng/µl. Significance shown by asterix (n=3)

To assess the potential of silicon nanoparticles to improve transduction efficiencies, LV was complexed with Nuvec® for HEK293T cell transduction. All batches of Nuvec® provided were prepared as previously described (Section 2.2.1.19). 1 µg, 10 µg, 40 µg, 60 µg and 80 µg/0.5ml Nuvec® was complexed with LV to transduce cells at a MOI of 20. As it has not previously been shown if Nuvec® can complex with LV, various incubation times (10, 20 and 30 minutes) were also analysed for each concentration of Nuvec® (Figure 4.12-16). GFP expression was analysed via flow cytometry to determine transduction efficiencies. Transduction using Nuvec® as an enhancer was compared to polybrene complexation as a standard enhancer (2.61±0.46% GFP expression).

When complexing LV with various concentrations of Nuvec®, an increase in transfection efficiencies can be observed. The general trend shown is an increase in the level of GFP expression after complexing with increasing concentrations of Nuvec®. The duration of incubation time does not increase transduction efficiencies, with the similar percentage GFP expression detected between using each production batches at similar concentrations. Comparison of transduction efficiencies after complexing with various production batches of Nuvec® identified the most successful batch for transduction efficiency enhancement. All production batches tested indicate 10 µg Nuvec® matches or slightly enhances the transduction efficiency in comparison to polybrene transduction. Complexation of LV with ≥40 µg Nuvec® enhances GFP expression observed above that detected through polybrene complexation alone. Comparison of transduction efficiencies and incubation periods measured using various batches of Nuvec® suggests that NV00100032 is the most successful batch of Nuvec® for viral complexation for 20 minutes incubation. Complexation with virus enhances transduction efficiencies by 2.86% (1 µg), 3.36 (10 µg), 13.38% (40 µg), 24.36% (60 µg), 28.59% (80 µg) above that of polybrene complexation. This data suggests that complexation of LV with 10-

80 µg Nuvec® is suitable for enhancing transduction efficiencies of LV above standard transduction protocols.

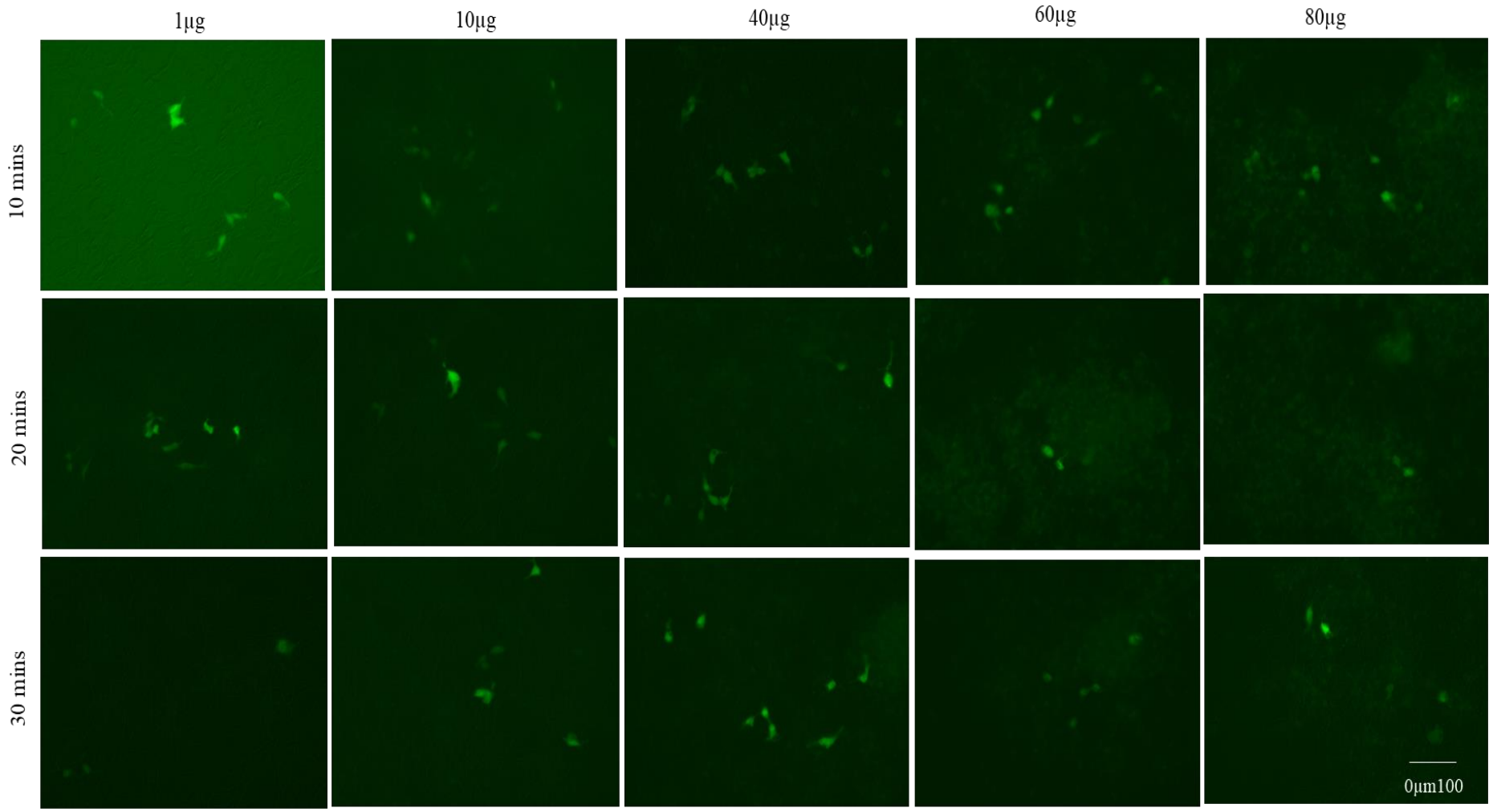


Figure 4.12- Fluorescent microscopy analysis of LV transduction efficiency on HEK293T using NV00100028. Complexation of LV (MOI 20), with various concentrations of NV00100028 (µg/0.5ml) and various incubation times. Cells were analysed via flow cytometry to quantify gene expression. Cells imaged under fluorescent microscopy 72hrs post transduction, x20 magnification and scale bar shown between 0 and 100µm (n=4)

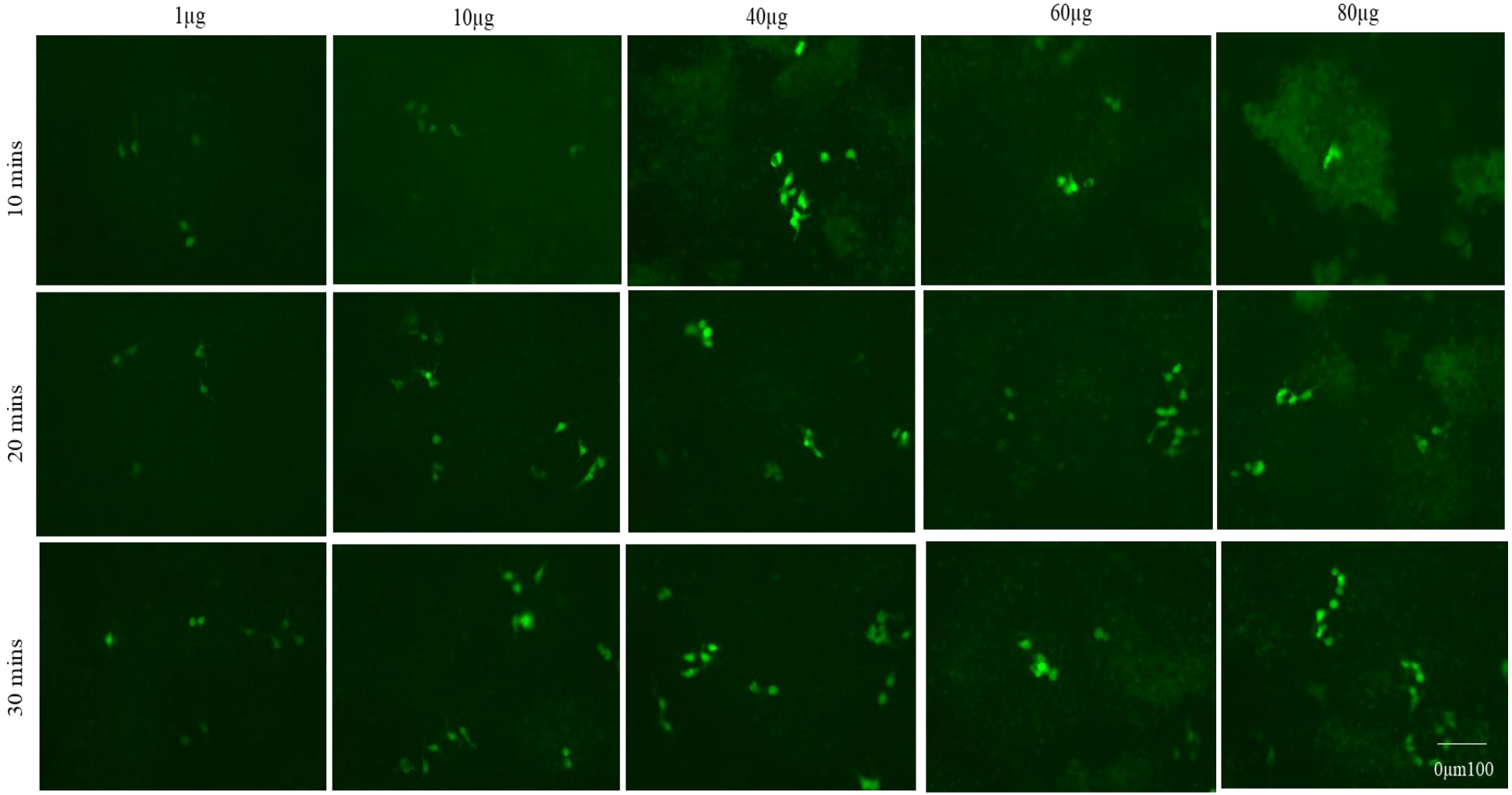


Figure 4.13- Fluorescent microscopy analysis of LV transduction efficiency on HEK293T using NV00100026-28. Complexation of LV (MOI 20), with various concentrations of NV00100026-28 (µg/0.5ml) and various incubation times. Cells were analysed via flow cytometry to quantify gene expression. Cells imaged under fluorescent microscopy 72hrs post transduction, x20 magnification with 0 to 100 µm scale bar shown (n=4)

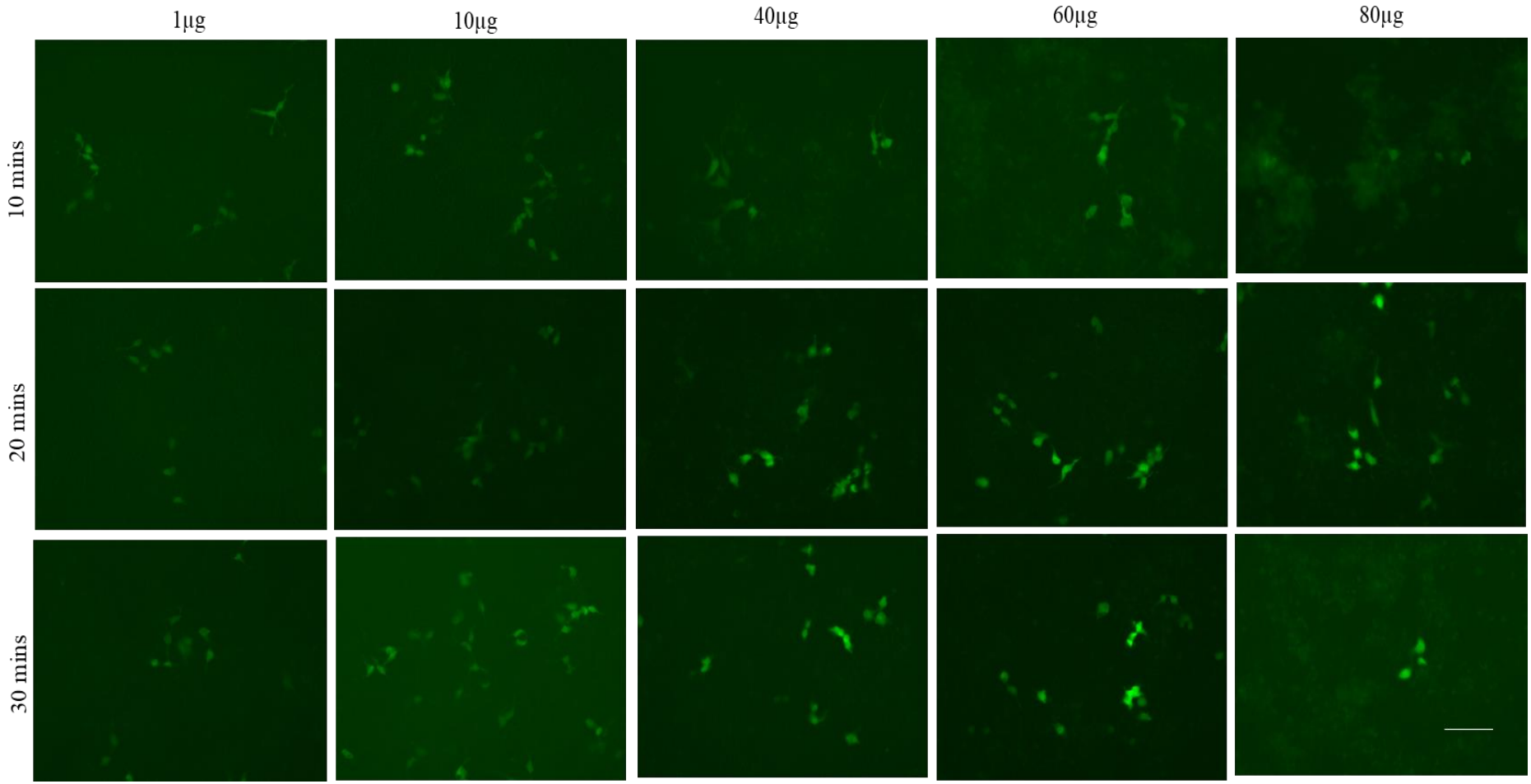


Figure 4.14- Fluorescent microscopy analysis of LV transduction efficiency on HEK293T using NV00100032. Complexation of LV (MOI 20), with various concentrations of NV00100032 (µg/0.5ml) and various incubation times. Cells were analysed via flow cytometry to quantify gene expression. Cells imaged under fluorescent microscopy 72hrs post transduction, x20 magnification and scale bar shown between 0 to 100 µm (n=4)

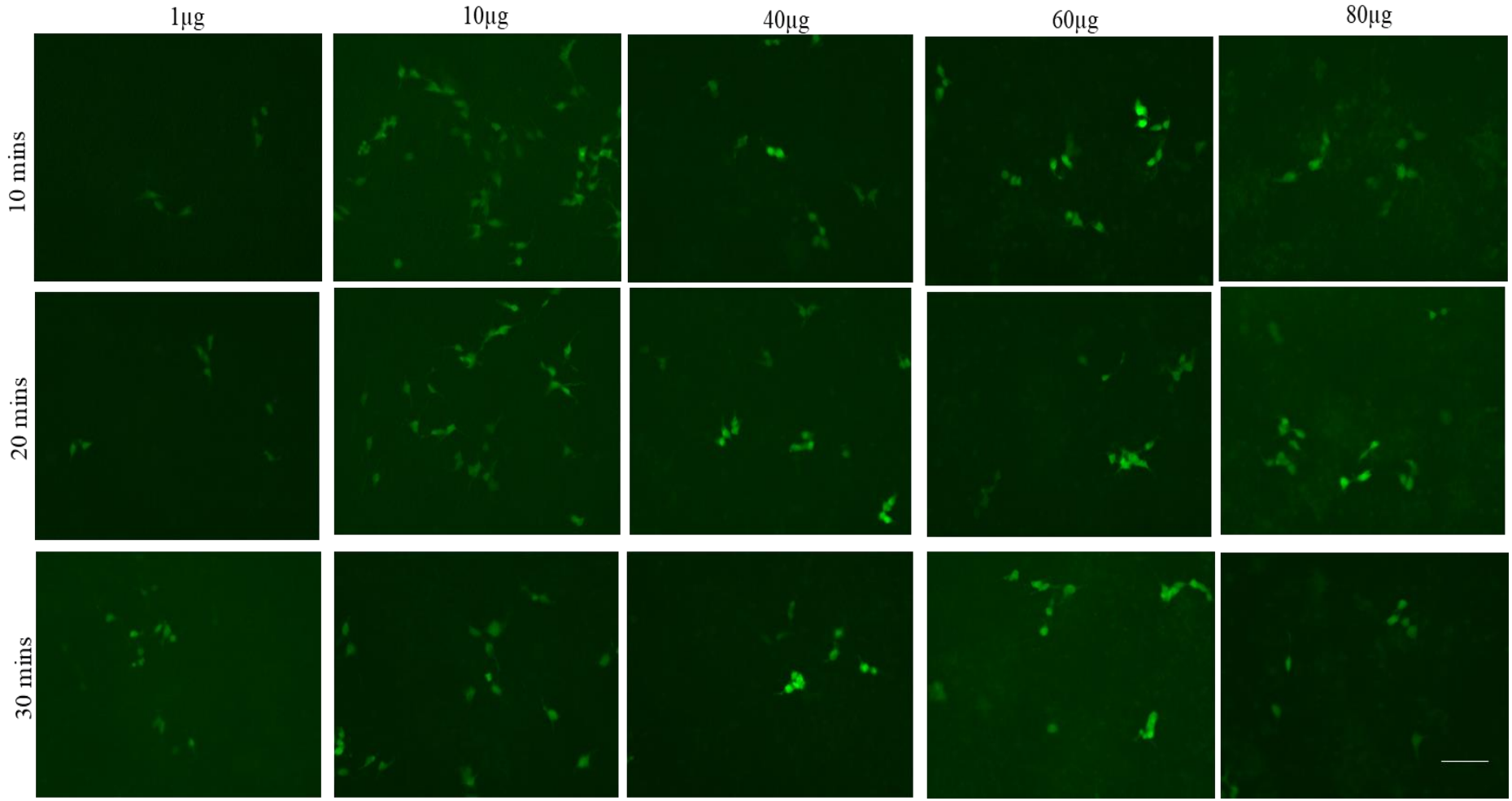
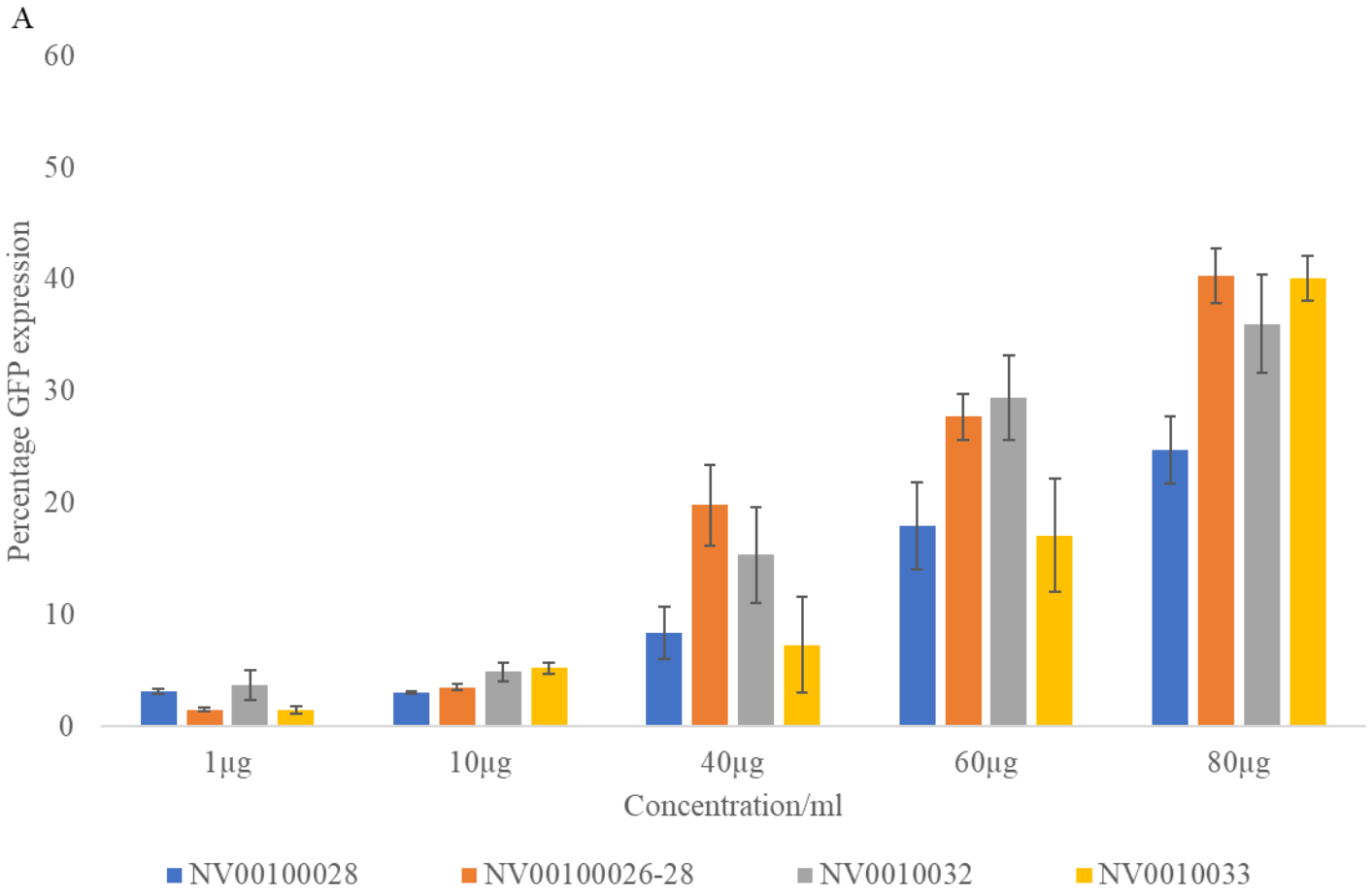
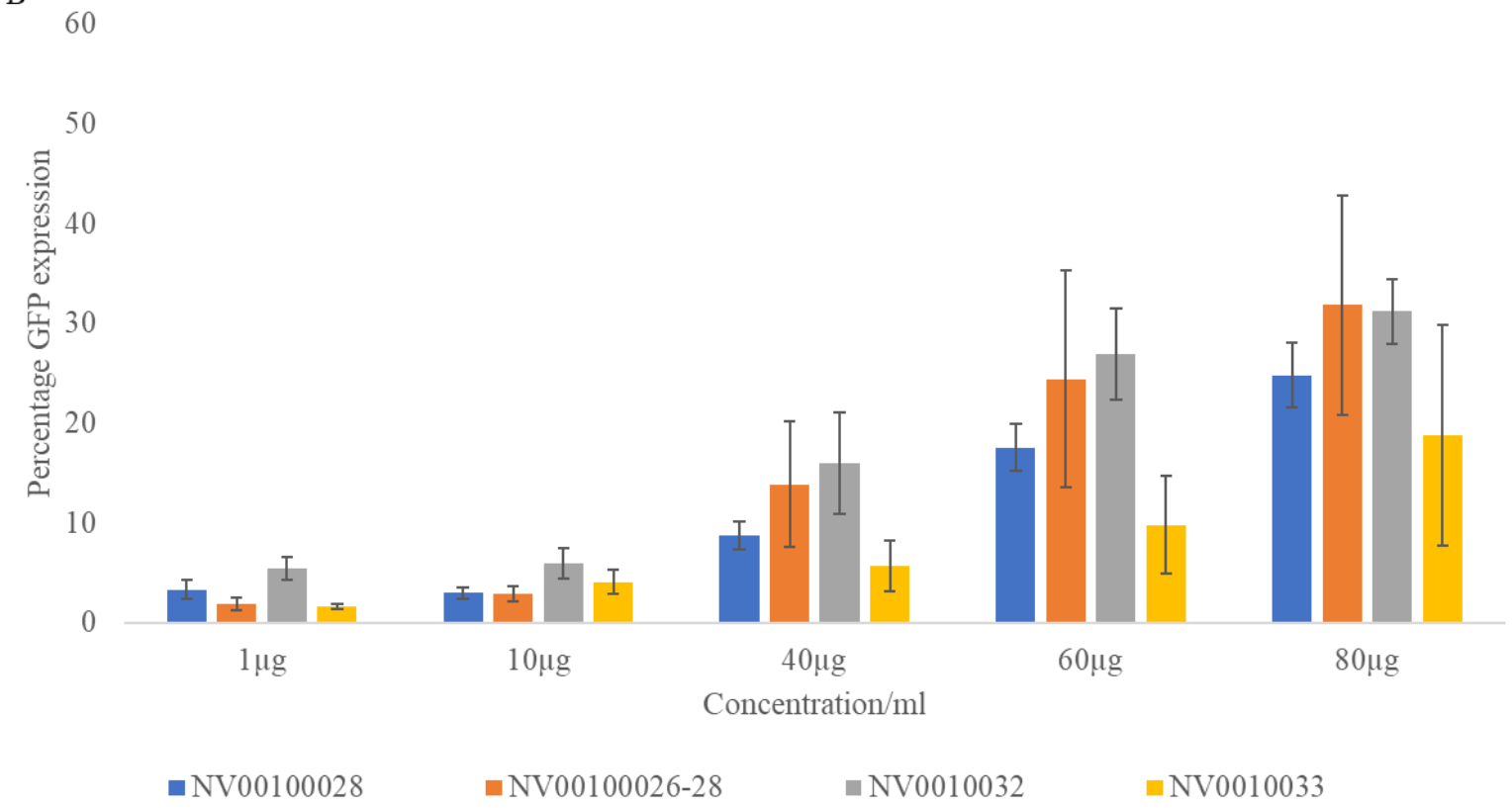


Figure 4.15- Fluorescent microscopy analysis of LV transduction efficiency on HEK293T using NV00100033. Complexation of LV (MOI 20), with various concentrations of NV0010003 (µg/0.5ml) and various incubation times. Cells were analysed via flow cytometry to quantify gene expression. Cells imaged under fluorescent microscopy 72hrs post transduction, x20 magnification and scale bar between 0 and 100µm shown (n=4)


A
60
50
40
30
20
10
0
Percentage GFP expression
1µg
10µg
40µg
60µg
80µg
Concentration/ml
NV00100028
NV00100026-28
NV0010032
NV0010033

B


60
50
40
30
20
10
0
Percentage GFP expression
1µg
10µg
40µg
60µg
80µg
Concentration/ml
NV00100028
NV00100026-28
NV0010032
NV0010033

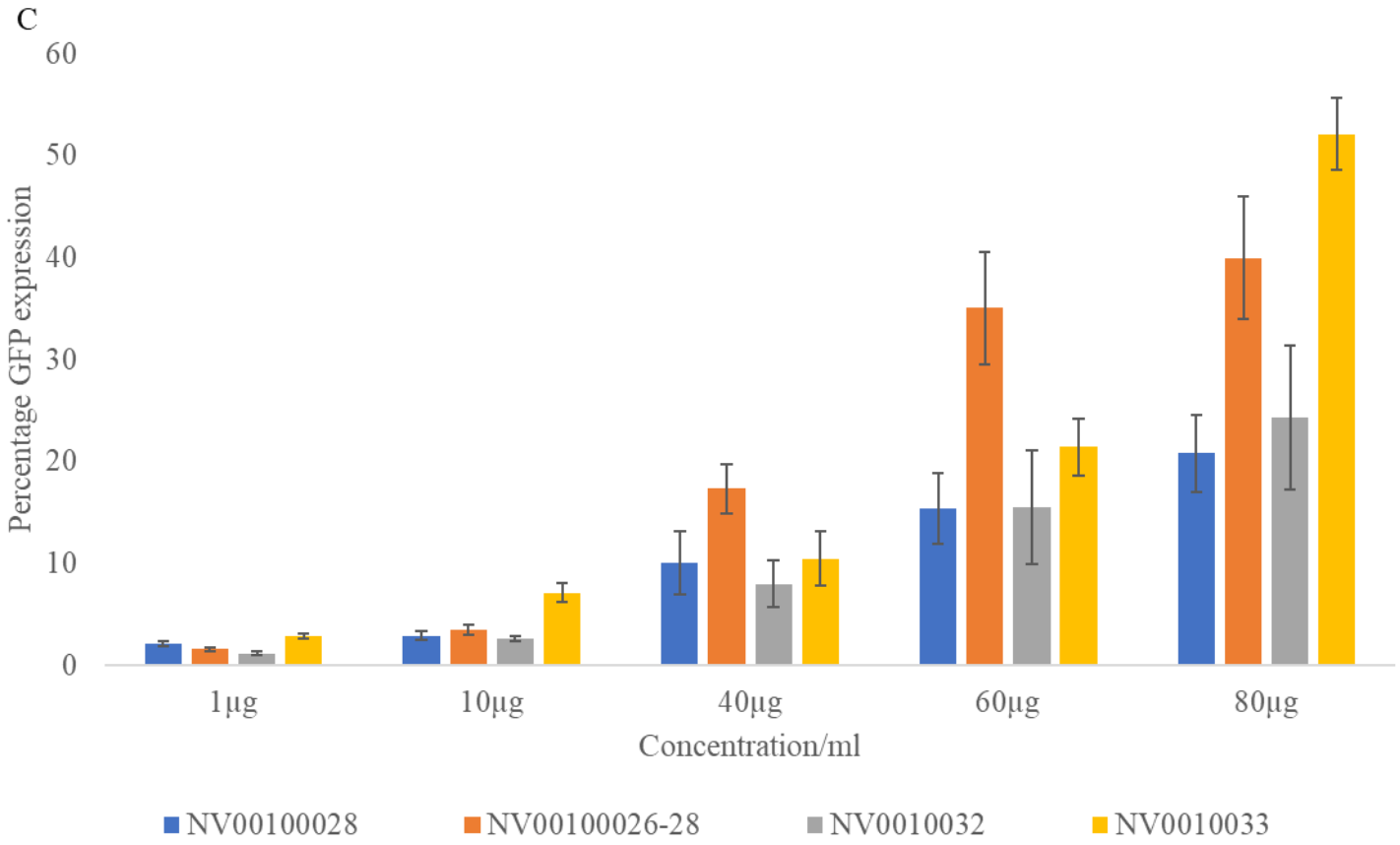


Figure 4.16- Transduction efficiencies using Nuvec. Flow cytometry analysis was used to determine percentage GFP expression of HEK293T cells transduced with LV (MOI 20) complexed with various batches of Nuvec after complexation for 10 (A), 20 (B) and 30 (C) minutes (n=4)

4.2.8. Nuvec®- LV Complexation Cytotoxicity

We next assessed the toxicity of Nuvec® complexation with LV *in vitro*. HEK293T cells transduced with LV complexed with polybrene (MOI 20) show little cytotoxicity, with cell viability measured as 91±1.41% 72 hours post transduction. Cell viability was not altered between incubation times validating transduction efficiency results (Figure 4.17). Greater than 90% viability is observed after transduction using 1-40 µg Nuvec®. However, a significant increase in cell death is observed after transduction using 60-80 µg Nuvec®, with less than 85% viability measured after 72 hours post transduction ($P<0.05$). These results suggest that while Nuvec® does not decrease cell viability below 70% over a 72-hour period, the greater concentration of Nuvec® used results in increase cell toxicity. The optimal concentration of Nuvec® in reducing cytotoxicity remains between 1-40 µg Nuvec®. These results closely mirror that of cell viability with Nuvec® treatment alone, suggesting that LV transduction does not increase cell toxicity.

Both transduction efficiencies and cytotoxicity assays suggest complexation of LV with 40 µg NV00100032 for 20 minutes is optimal for increasing transduction efficiencies above that observed using standard transduction protocols alone whilst maintaining cell health.

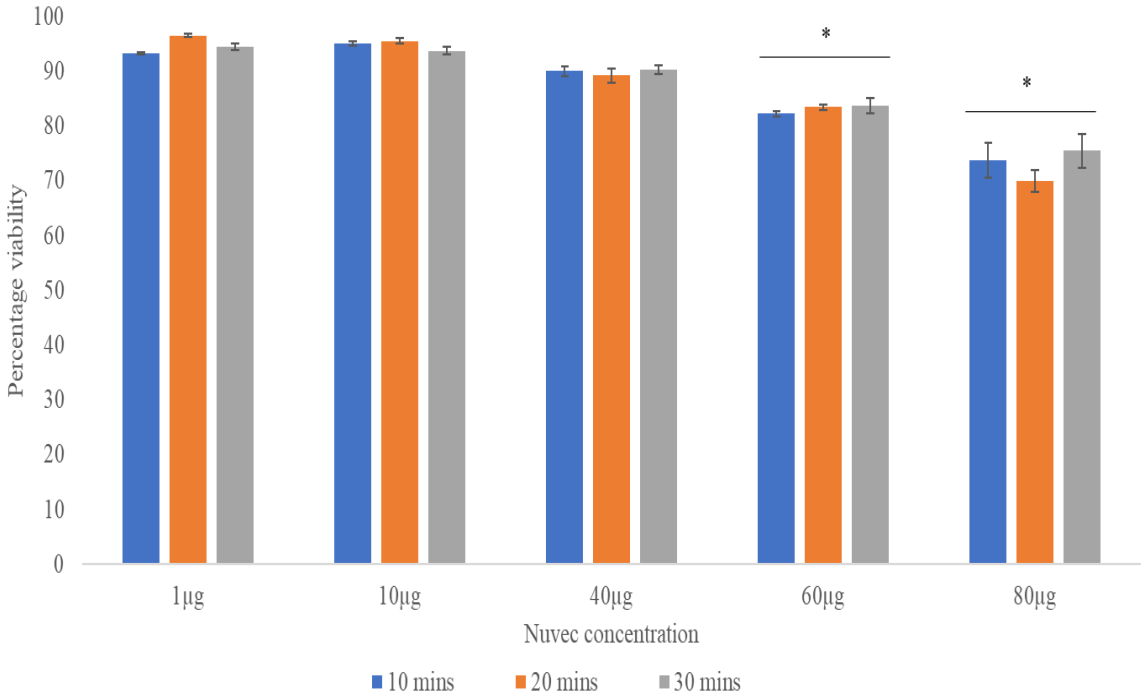
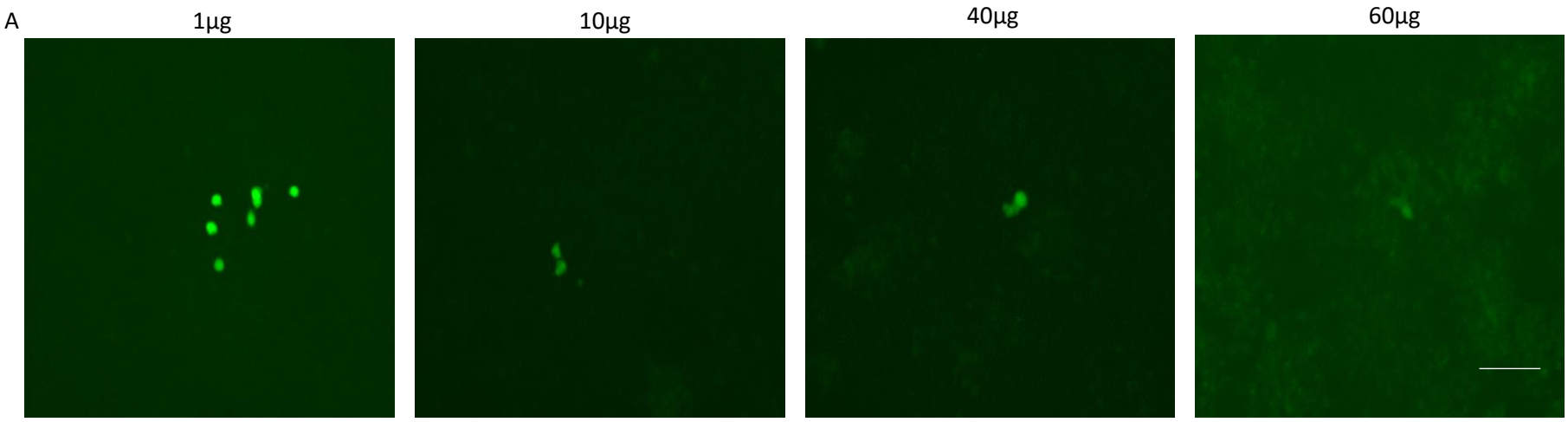


Figure 4.17- Percentage viability of HEK293T cells were transduced with LV (MOI 20) complexed with various concentrations of Nuvec and incubation times. Viability was analysed using trypan blue exclusion assay, with results of treatments compared to untreated cells to show a drop in viability using 60-80µg Nuvec. Significance shown by asterix ($P<0.05$), n=2

4.2.9. Nuvec® and Polybrene Complexation Optimisation

Whilst complexation of LV with Nuvec® alone is shown to increase transduction efficiencies above polybrene complexation alone, it is hypothesised that both polybrene and Nuvec® complexation may increase transduction efficiencies further. To investigate this, complexation of LV (MOI 20) with 5 µg/ml polybrene and various concentrations of Nuvec® (NV00100033) for 20 minutes were used to transduce HEK293T cells. GFP expression after complexation with polybrene and 1 µg Nuvec® is 1.27±0.28%, 10 µg: 1.29±0.22%, 40 µg 12.39±1.07%, 60 µg: 23.22±1.14% (Figure 4.18). Comparison of percentage GFP expression indicates no increase in transduction efficiencies after complexation using 1-10µg Nuvec®. However, a 6.66% and 13.42% increase in GFP expression is detected when complexing virus with 40

and 60 µg Nuvec® in comparison to Nuvec® complexation alone. As 60 µg Nuvec® shows cytotoxicity, complexation of LV with 40 µg Nuvec® and 5 µg polybrene is optimal to increase transduction efficiencies further.


A
1µg
10µg
40µg
60µg

B

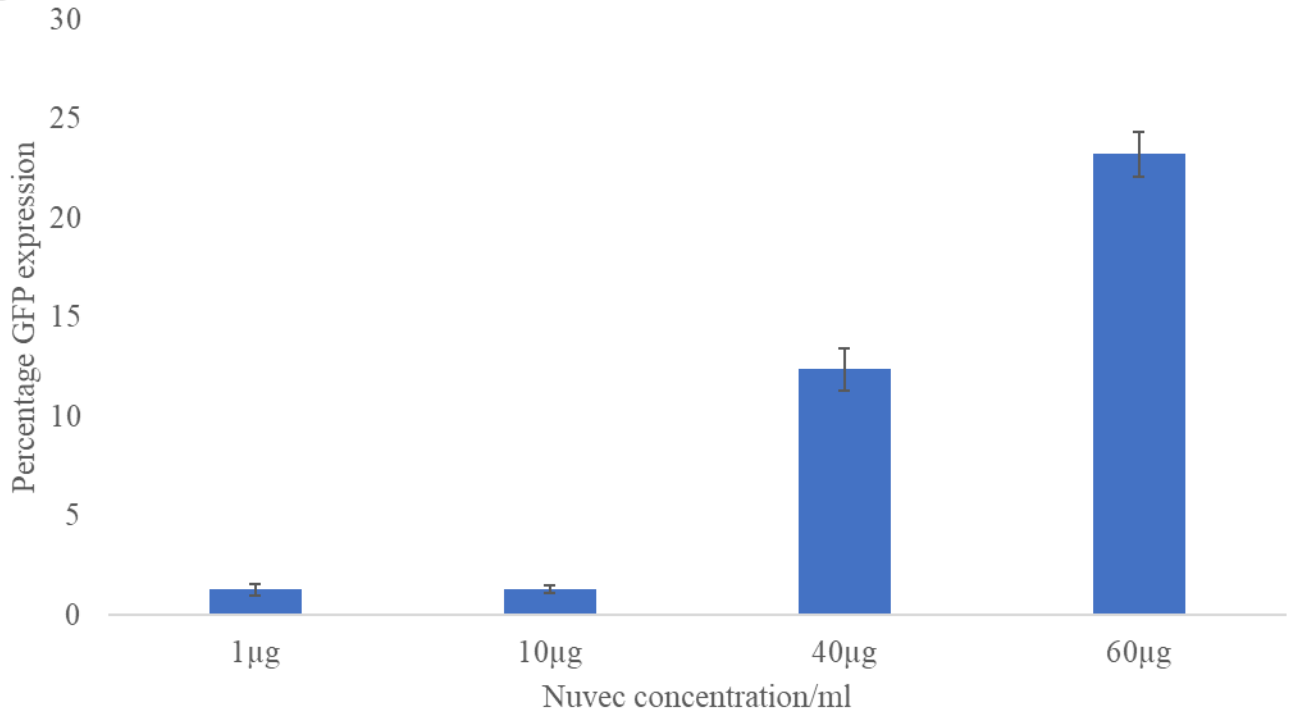
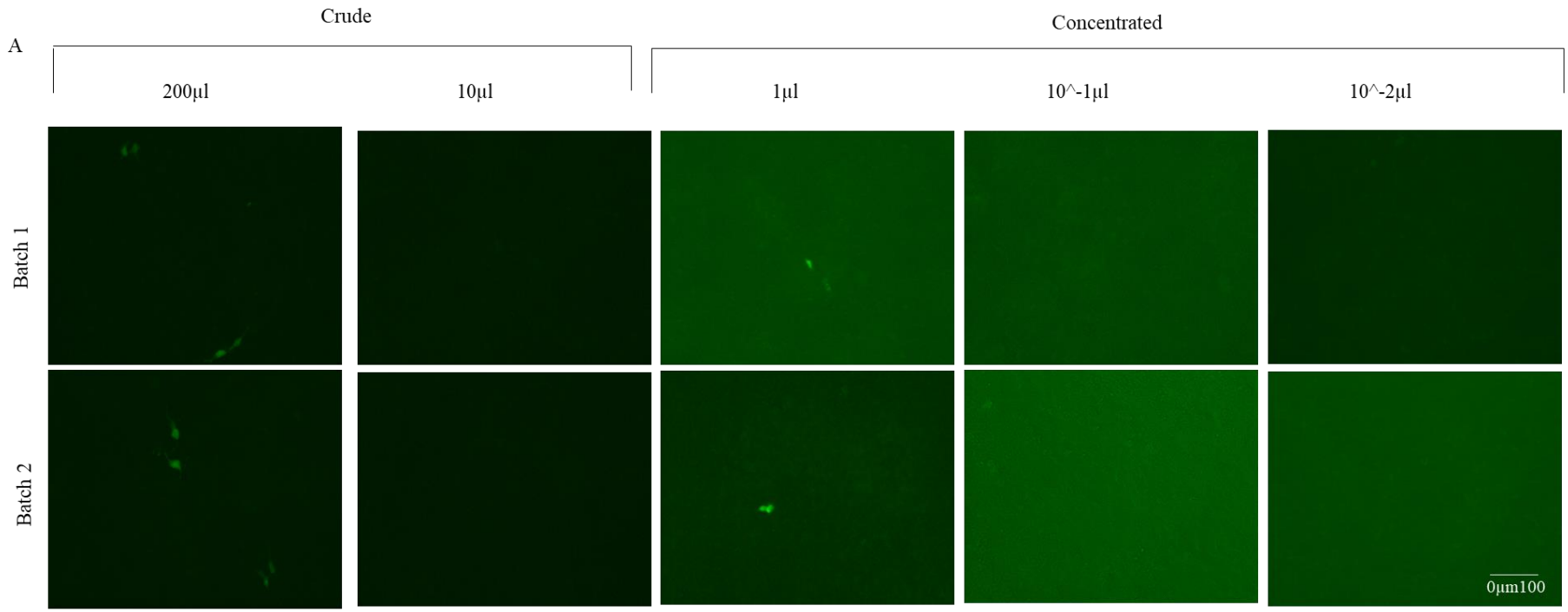


Figure 4.18- LV transduction efficiency using Nuvec and polybrene. A- HEK293T cells transduced using LV (MOI 20) complexed with various concentrations of NV00100033 (µg/0.5ml) and 5µg/ml polybrene, complexed for 20 minutes and visualised under fluorescent microscopy 72 hours post transduction, x20 magnification and scale bar shown between 0 and 100µm. B- GFP expression detected 72 hours post transduction via flow cytometry to quantify transduction efficiency, n=4

4.2.10. Analysis of LV Produced through Nuvec® Transfection

Previous experiments have shown the success of Nuvec® in transfection of DNA. pHR LV production using Nuvec® as a transfection reagent was optimised by Ms Serena Fawaz, Brunel University London. However, upon collection and concentration of supernatant containing LV particles, it was suggested whether free floating Nuvec® remained in the supernatant which may be collected and concentrated with the virus prep (Spinnrock, Martens et al. 2019). pHR LV produced using Nuvec® (N-pHR LV) was titrated initially using 5 µg/ml polybrene (Figure 4.19). Crude batch 1 of N- pHR LV was titrated as $2.83x10^6$ TU/ml crude, which increased to $4.33x10^7$ TU/ml upon concentration via ultracentrifugation. Batch 2 produced a similar titre of N-pHR LV with $4.02x10^6$ TU/ml estimated as crude virus, which increased to $1.87x10^7$ TU/ml upon concentration. This indicates successful production of N-pHR LV using Nuvec as a transfection reagent.



B

	Crude	Concentrated
Batch 1	$2.83x10^6$	$4.33x10^7$
Batch 2	$4.02x10^6$	$1.87x10^7$

Figure 4.19- Titration of crude and concentrated LV produced using NV00100028. A- Fluorescent microscopy of HEK293T cells transduced using serial dilutions of crude and concentrated virus, imaged 72 hours post transduction (x20 magnification and scale bar shown between 0 to 100µm). B- Flow cytometry quantification of GFP expression was used to determine titre of virus (TU/ml), n=2

HEK293T cells were transduced with N-pHR LV (MOI 20) without an enhancer and with 1-40 µg Nuvec (NV00100033) for 30 minutes. Both batches of N- pHR LV showed similar transduction efficiencies when complexed to Nuvec®, with an average of 1.35±0.18% (1 µg), 2.80±0.27% (10 µg) and 20.91±1.15% (40 µg) GFP expression (Figure 4.20). This is shown as an increase in transduction efficiency compared to LV transduction alone without enhancer complexation. Percentage transduction is found similar to transduction of LV produced through transient transfection using other transfection reagents. This suggests that Nuvec has not been concentrated with virus upon ultracentrifugation of the virus supernatant.

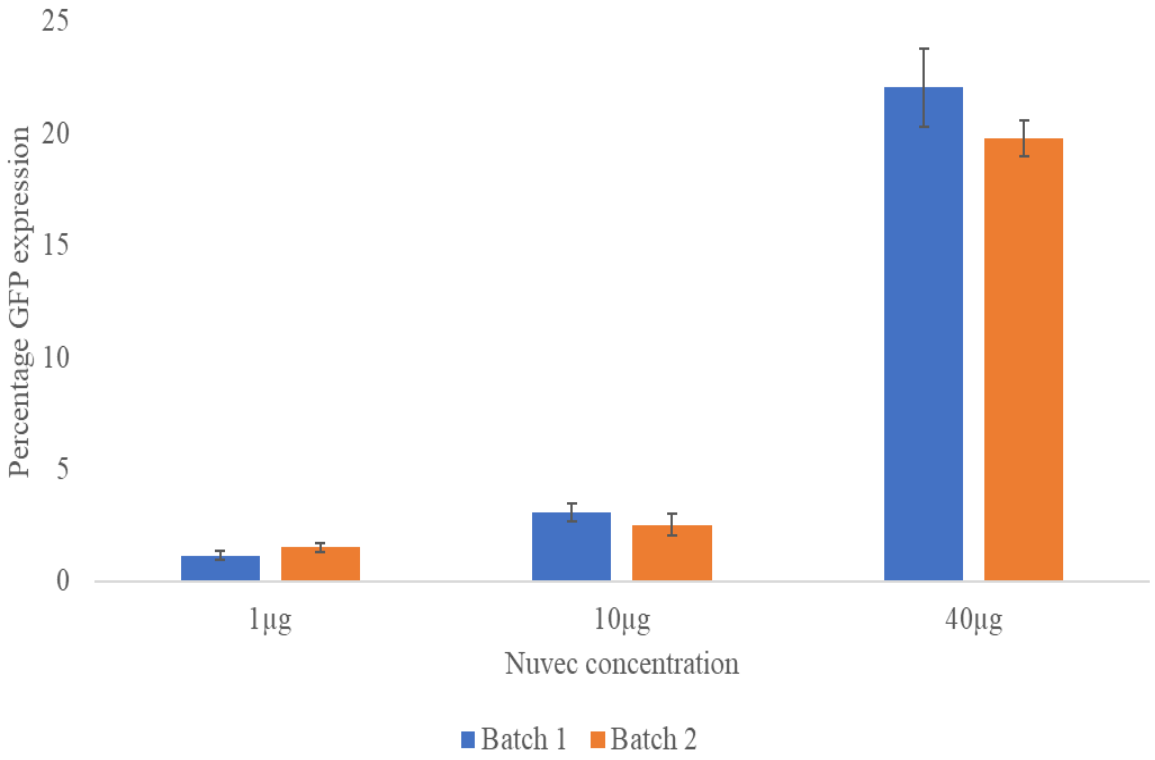


Figure 4.20- Analysis of N-LV transduction efficiency. Percentage GFP expression analysed via flow cytometry, to determine transduction efficiencies using flow cytometry after transduction using various concentrations of NV100033, complexed with pHR LV (MOI 20) for 20 minutes. Cells analysed 72 hours post transduction to show, n=4

4.2.11. LV Production Optimisation

LV carrying the *LacZ* transgene, pseudotyped with VSV-G_{Ind} were generated through split genome packaging of pLNTSbetagalW, pCMVR8.74 and pMD2.G (Section 2.2.2.10). LV titre produced using three various transfection reagents (PEI, Fugene® 6 and Genejuice®) were compared to determine optimised protocol for high titre LV production. These three reagents were used to analyse efficacy of viral production. LV were titrated (Figure 4.26) on HEK293T cells and stained for β- galactosidase activity as previously described (2.2.2.11).

The titre of LacZ LV generated through PEI transfection of plasmid DNA are estimated as an average of $5.29x10^8$ ($6x10^8$, $4.7x10^8$, $5.18x10^8$). LV produced through Fugene® 6 transfection is estimated to have an average titre of $4.13x10^9$, with individual batches of virus titrated as $7.63x10^9$, $1.97x10^9$ and $2.81x10^9$. Genejuice® transfection of pDNA for LV production generated an average titre of $2.13x10^{10}$ ($1.59x10^{10}$, $3.83x10^{10}$ and $9.82x10^9$). While these titres are not significantly different ($P>0.05$), these results indicate Genejuice® as the most effective transfection reagent for LV production (Figure 4.23). The Genejuice® protocol produces LV of 4.16 -fold higher titre than with Fugene 6 and 39.34-fold greater titre than when using PEI for LV production.

Further sensitive quantification of LV is required to determine accurate titres for LV titre. As β- galactosidase is not being a live cell stain, variability may occur in the staining assay. pHR LV, pseudotyped with VSV-G_{Ind} were produced through triple plasmid transfection of HEK293T cells using split genome packaging to generate replication deficient LV. These viruses can be quantified using flow cytometry, a sensitive method for detection of fluorescence.

Optimised protocols for LV production were used for transient transfection using PEI, Fugene® 6 or Genejuice® reagents. LV were produced as previously

described in 2.2.2.10. Briefly, HEK293T cells were transfected using specific concentrations of plasmids and conditioned medium harvested over 72 hours before concentration (Figure 4.21). Cells exhibit increasing green fluorescence, indicating increased gene expression over time. LV were titrated for infectious titre (TU/ml) in HEK293T cells (Figure 4.22).

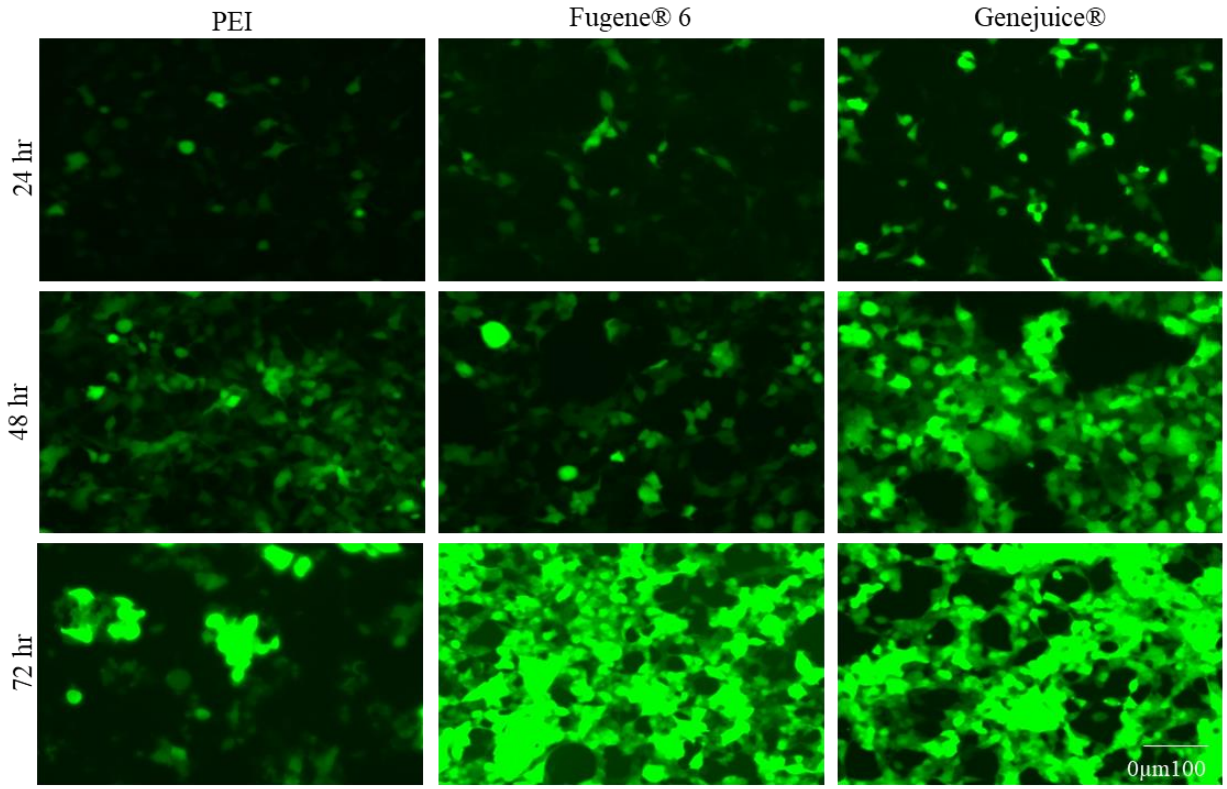


Figure 4.21- Comparison of LV production using various transfection reagents. PEI, Fugene® 6 and GeneJuice® were used as transfection reagents for pHR LV production over time. Fluorescent microscopy of transfected cells over 72 hours for virus collection show an increase in fluorescence over time (x20 magnification, scale bar shown between 0 and 100µm)

Titration of LV indicate Genejuice® generates the highest titre (Figure 4.23). LV titre produced from PEI transfection were estimated on average as $1.28x10^8$. Individual batches of LV were titrated as $6.19x10^7$, $4.02x10^7$, $1.26x10^7$, $6.24x10^7$ and $4.64x10^8$. Titres generated through transient transfection of HEK293T cells using Fugene® 6 were titrated as an average of $1.51x10^8$. Production batches of LV were titrated as $2.54x10^8$, $1.11x10^8$, $1.50x10^8$,

$2.37x10^8$ and $3.48x10^7$. LV generated through Genejuice® transfection were averaged as $2.06x10^9$. Separate batches of LV were titrated as $2.18x10^9$, $3.98x10^9$, $1.04x10^9$, $1.55x10^9$ and $1.54x10^9$. The titres generated through Genejuice® transfection and PEI or Fugene 6 transfection are statistically significant ($P<0.05$).

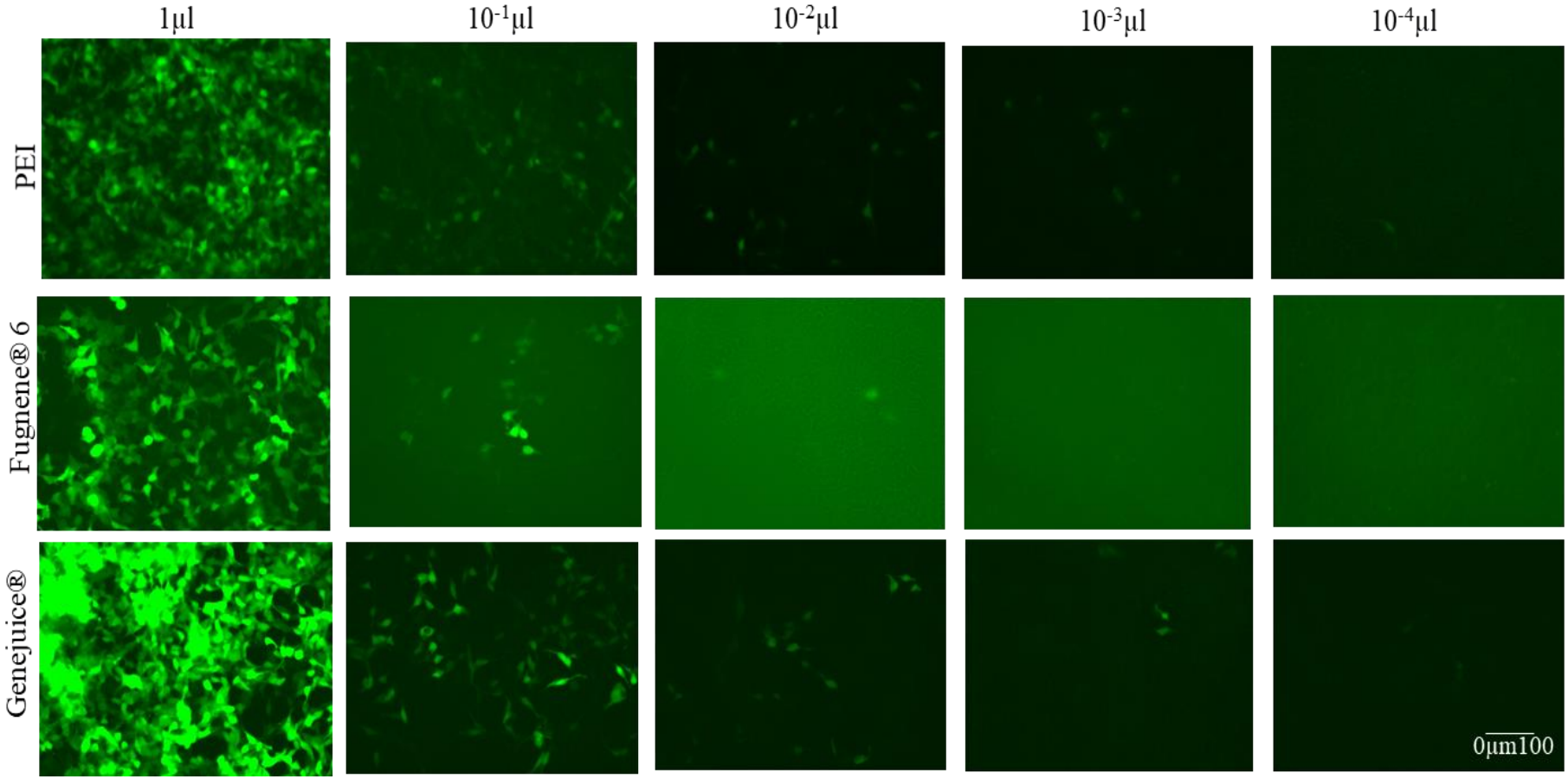


Figure 4.22- Infectious titration of pHR LV generated using three transfection reagents. Fluorescent microscopy images of HEK293T cells transduced with serial dilutions of pHR LV produced using PEI, Fugene® 6 and GeneJuice® shown 72hrs post transduction, x20 magnification with scale bar (0 to 100µm). Flow cytometry quantification of fluorescent cells were used to determine titre of various batches of LV

P24 titration of virus batches indicate high titre production of LV particles. Particle titre was calculated as $2.8x10^{12}$, $1.37x10^{12}$ and $2.62x10^{13}$ LP/ml after LV generation using PEI, Fugene® 6 and Genejuice®, respectively. In contrast, infectious titres were determined as $1x10^{9}$, $3.48x10^{7}$ and $3.98x10^{9}$ TU/ml respectively. This demonstrates the high concentration of LV particles present in production batches, though a reduced number of particles remain infectious.

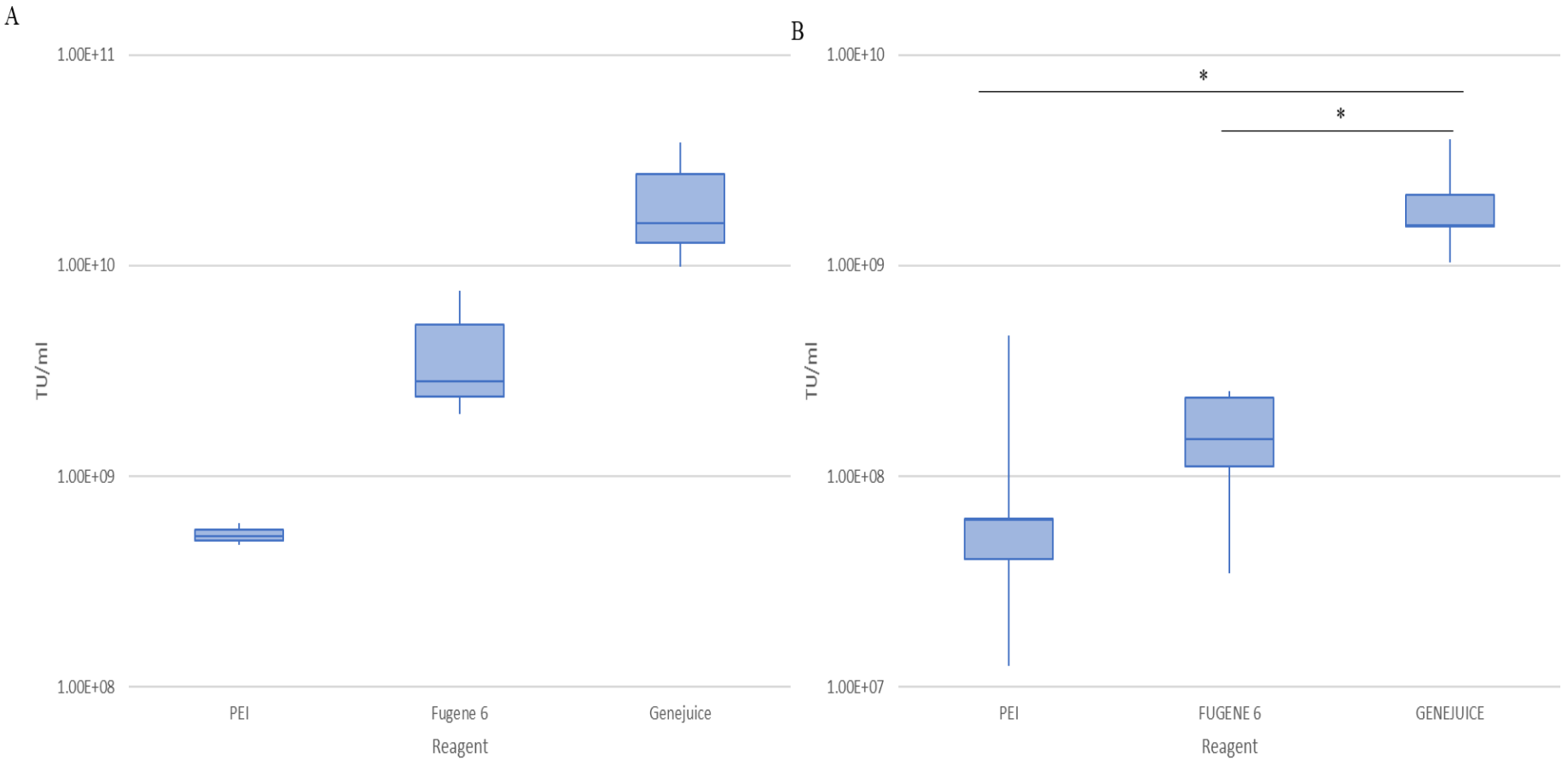


Figure 4.23- LV titres produced through transient transfection of HEK293T cells using various transfection reagents. A- LacZ LV titres determined subsequent to β- galactosidase staining of transduced cells (n=3). B- pHR LV titres, determined through flow cytometry quantification of GFP expression. Significance denoted by an asterix, n=5

4.2.12. pHV LV Production

Genejuice® has been shown as the optimal reagent for high titre LV production with low toxicity in transfected cells. pHV LV, pseudotyped with VSV-G_{Ind} were produced through transient transfection of HEK293T cells using Genejuice® transfection of pHV, pCMVR8.74 and pMD2.G pDNA. pHV LV are similar to pHV LV, except that they carry an *eGFP* transgene flanked by full LTRs compared to SIN LTRs. These viruses are required to determine genotoxic effects of LV carrying the full or SIN LTR.

Titration of pHV LV indicates successful production of high titre vectors for use in downstream assays (Figure 4.24). Average infectious titre of pHV LV were $4.72x10^9$ TU/ml. Individual production batches of pHV LV were shown to be titrated to $5.8x10^9$ TU/ml, $4.57x10^9$ TU/ml and $3.8x10^9$ TU/ml.

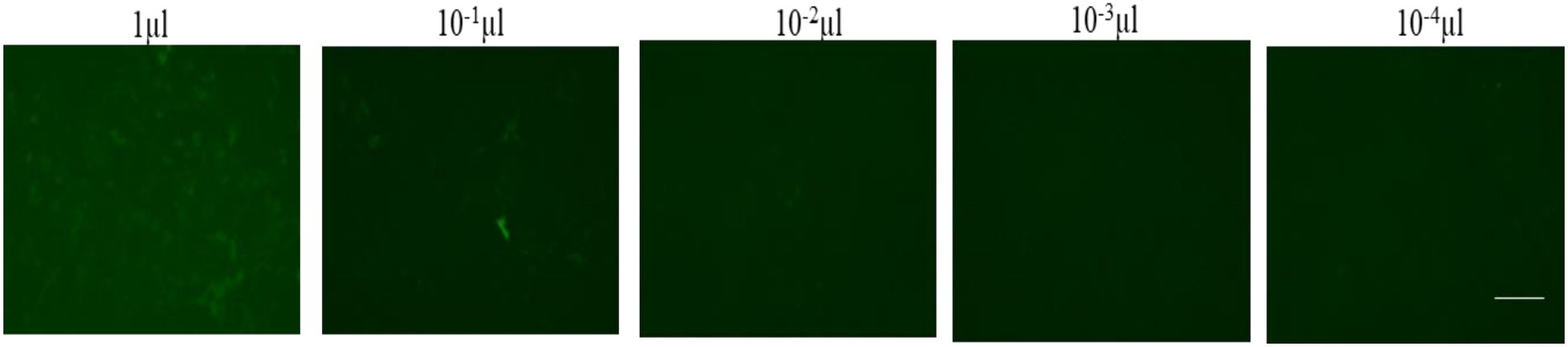


Figure 4.24- pHV LV titration. Fluorescent images of HEK293T cells transduced with serial dilutions of pHV LV, imaged 72 hours post transduction (x20 magnification, scale bar (0 to 100µm) shown). Cells were analysed by flow cytometry to determine GFP expression and titre of virus batches, n=3

4.2.13. Scalable LV production

There is a need for high titre LV production for use in downstream assays in gene therapy in a cost-effective manner. We have shown the generation of high titre pHR LV through an optimised transfection protocol using Genejuice®, which also exhibits low cytotoxicity. Therefore, optimisation of LV production was carried out on a large scale to determine the efficacy of this reagent in producing high titre pHR LV over a longer period of time.

HEK293T cells were transfected with pHR, pCMVR8.74 and pMD2.G, as previously described (Section 2.1.2.10) at a multiple of 10. Daily virus harvests were collected over 15 days and cells were passaged upon confluency (Figure 4.25). High levels of GFP expression indicate high gene transfer and expression. While conditioned medium was collected for 15 days post transfection, virus was only quantified for 11 days, with virus titre falling consistently over this period. This is expected due to the transient nature of plasmid expression, with average expression for 7 days.

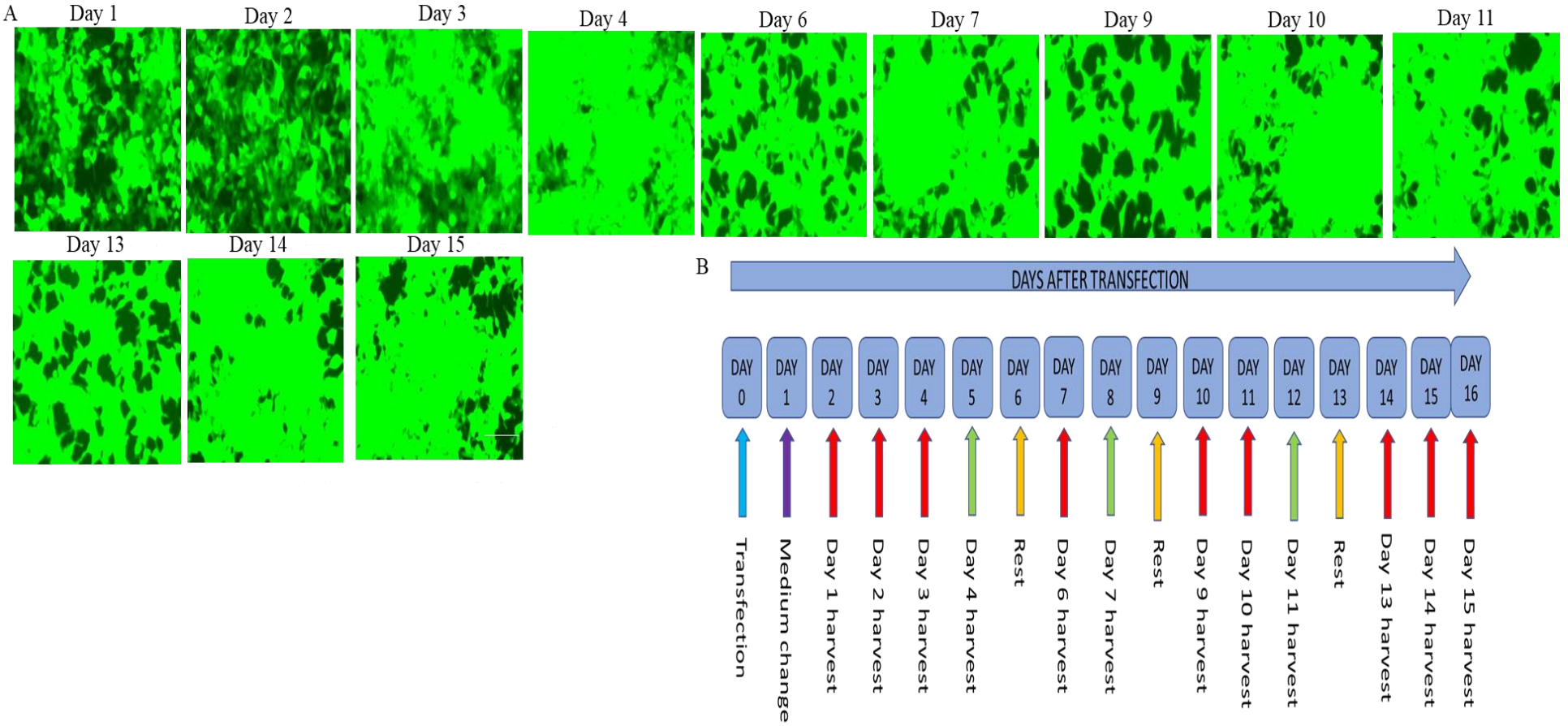
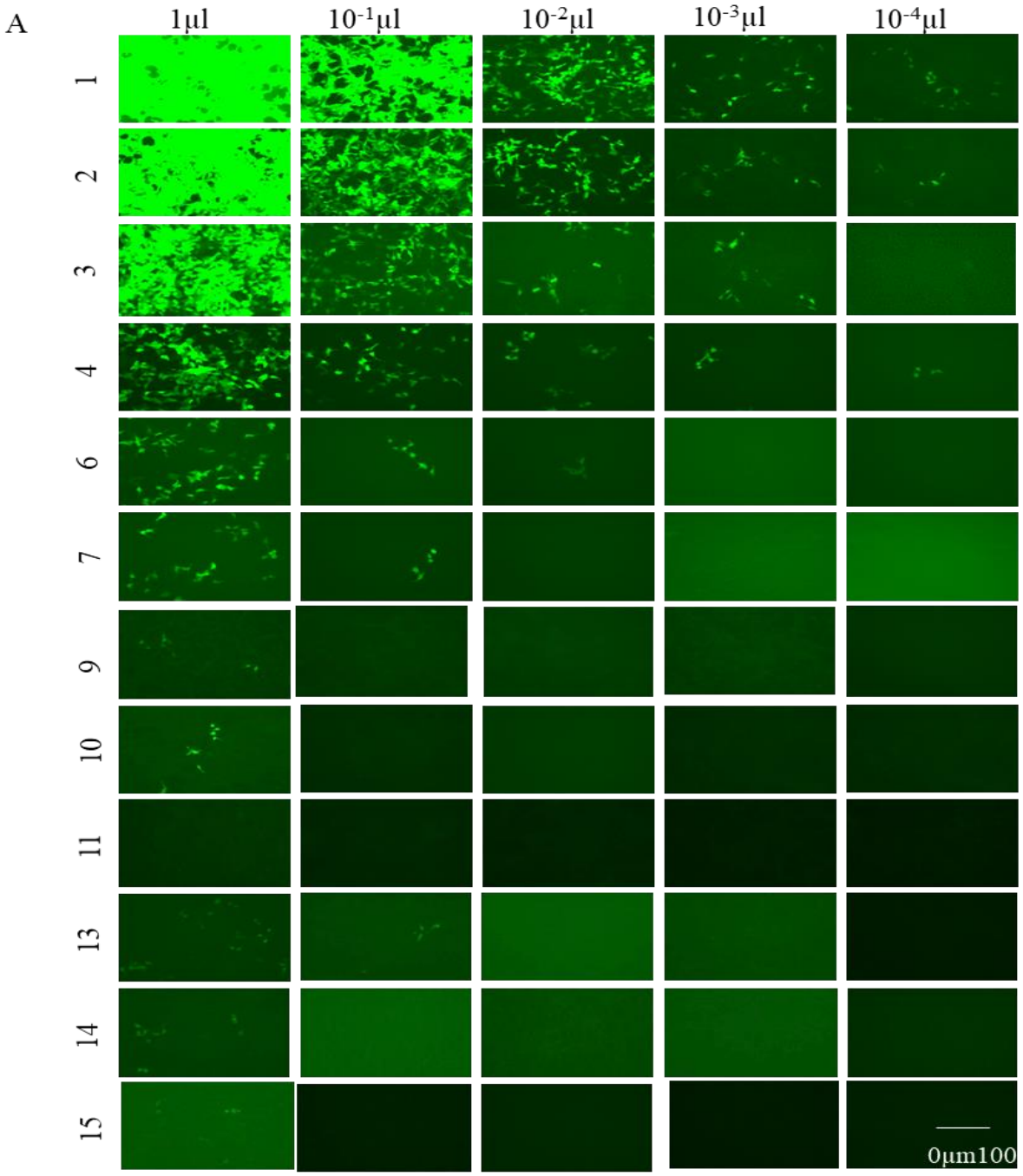


Figure 4.25- Time course of scalable pHR LV production. HEK293T cells transfected to produce LV using Genejuice® on scalable production protocol were grown and serially passaged over 16 days. A- Fluorescent microscopy of transfected cells imaged over time indicate increasing GFP plasmid expression over 15 days (x20 magnification, scale bar (0 to 100µm) shown). B- Schematic diagram of methodology used. Various stages and steps shown above and colour coded; Red: Harvest, Green: Split cells, Orange: No collection

High titre LV titre was harvested over 10 days. (Figure 4.26) with the highest concentration of LV quantified in the initial 96-hour period post transfection. pHR LV titre was estimated as $3.09x10^{10}$ TU/ml (day one) with titration verified on EL4 cells as $1.11x10^{9}$ TU/ml (completed by Dr S. Ellison, University of Manchester). Further daily harvests were titrated as $2.42x10^{11}$ TU/ml (day two), $7.10x10^{10}$ TU/ml (day three) and $6.49x10^{9}$ TU/ml on day 4 post transfection. While the viral titres drop over this 96-hour period, this can be explained by the transient nature of plasmid transfection. Post 96 hours, viral titres remained significant, though titre continue to drop until day 11. Vector titre was estimated as $1.32x10^{9}$ TU/ml (day 6), $1.6x10^{8}$ TU/ml, (day 7), $9.69x10^{7}$ TU/ml (day 9), $4.25x10^{7}$ TU/ml (day 10). Conditioned medium from day 11-15 was harvested but no significant LV titre was quantified. This indicates that in a larger scale production of LV for 10 days, that can be adopted readily by labs, high titre LV can be produced in a cost and time effective manner.


A
1μl
10-1μl
10-2μl
10-3μl
10-4μl
1
2
3
4
6
7
9
10
11
13
14
15
0μm100

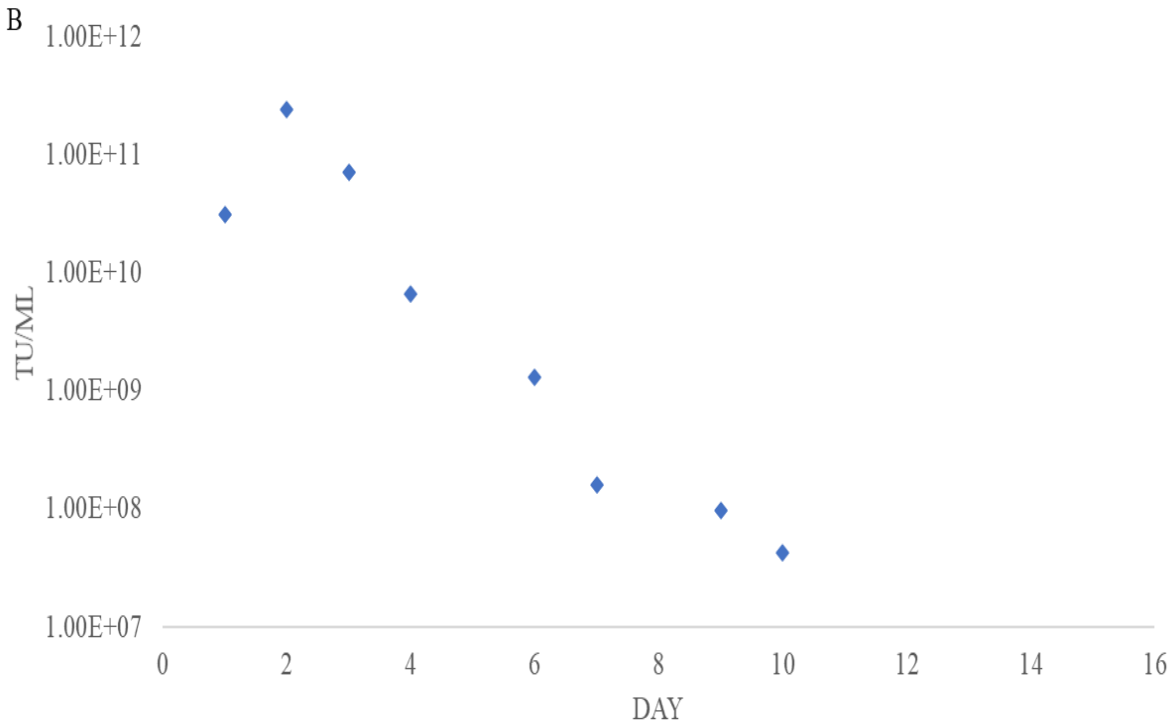


Figure 4.26- Time course LV titre of scalable production. A- Fluorescent microscopy images of HEK293T cells transduced with limiting dilutions of pHR LV virus harvests, imaged 72hrs post transduction (x20 magnification, scale bar between 0 to 100 µm shown). B- Infectious titre (TU/ml) of daily harvests of scalable transfection. Transduced cells were analysed using flow cytometry to quantify GFP expression to calculate infectious titre. Daily harvests were titrated to estimate LV concentration, indicating a drop in titre over 10 days but with high titre LV produced throughout the incubation period (n=2)

The viability of transfected cell was analysed upon splitting cells (1:3) when confluent (Figure 4.27). Cell viability was analysed using a trypan blue exclusion assay (Section 2.2.2.3). A small drop (-10.33%) in viability was measured over 15 days ($P>0.05$). Viability at day 4 was measured as 96.50±0.50%, day 7: 95.17±1.17%, day 11: 96.17±1.17% and day 15: 86.17±3.61%. As a significant drop in viability was only mmeasured after 15 days, this suggests successful long term culture of cells transfected for LV production using Genejuice®.

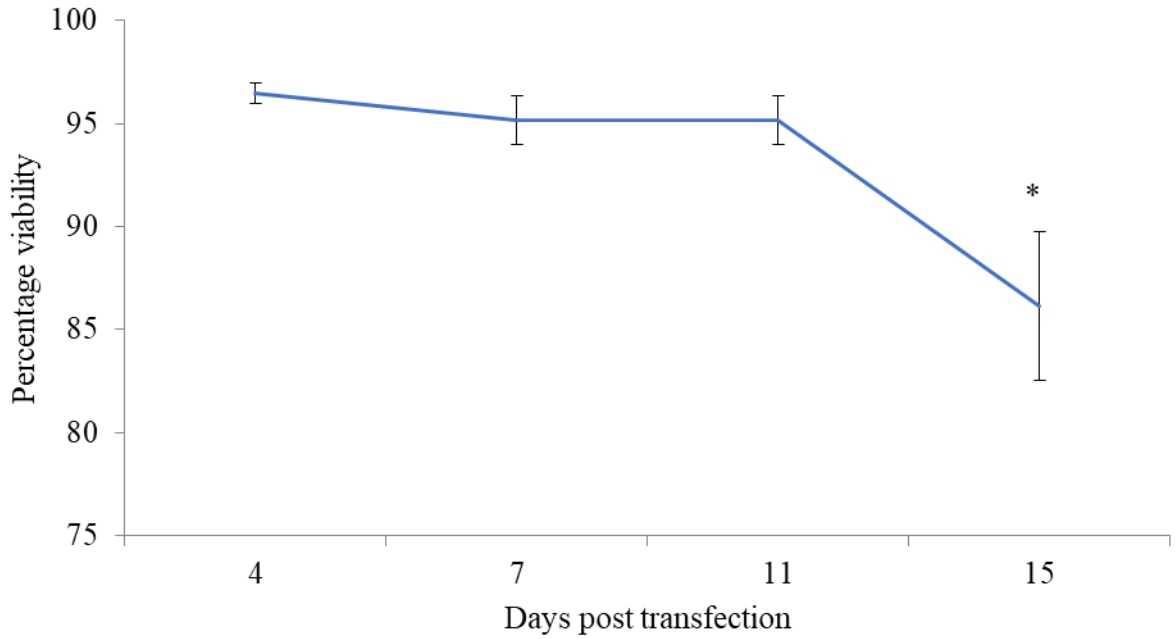


Figure 4.27- Viability of scalable pHR LV transfected cells over time. Viability of HEK293T cells transfected to produce LV, on a scalable methodology. Cells were grown and serially passaged over 16 days. Viability measured on days where cells were split using a trypan blue exclusion assay. SEM shown and significance denoted by asterix ($P<0.05$) (n=3)

These results indicate long term, high titre production of pHR LV in a scalable model for effective lab maintenace with low toxicity of transfected cells over time.

4.2.14. Gene Rescue Assay

Replication defective LV are generated using a split genome method for use in downstream assays. However, the presence of RCLs needs to be assessed to determine safety of vectors produced before downstream applications *in vitro* to prevent adverse side effects. Gene rescue assays for LV samples from long term production were performed (Section 2.2.2.12). Briefly, cells were transduced with LV harvested from day 7 and day 15 of long-term cultures and serially passaged over two weeks. Conditioned medium was added to cells before analysis for GFP expression via flow cytometry 72 hours post transduction. No GFP expression was detected in HEK293T cells transduced with conditioned medium. Cells transduced with conditioned medium of cells transduced using virus harvested from day 7 detected 0.32±0.30% GFP expression and 0.36±0.32% from day 15 virus harvest. The absence of RCLs, both at an early and late harvests indicate the safety of scalable long-term LV production.

4.2.15. Reverse Transcription Optimisation

Reverse transcription is a biological process by which RNA is converted into cDNA. This has been shown to occur naturally in HIV LV, with the reverse transcriptase enzyme packaged in HIV virions which catalyse viral RNA to cDNA prior to integration into the host genome (Zhang, H., Dornadula et al. 1996). This enzyme has been extracted and is used commercially in *in vitro* assays to quantify purified RNA (Pekrun, Petry et al. 1995, Koller, Graumann et al. 1995). This method can be used to analyse the purity of samples for gDNA.

Reverse transcription was performed using the GoScript™ Reverse Transcription System (Promega) according to manufacturer's instructions and was used to analyse the purity of purified RNA samples (Section 2.2.1.16).

However, several conditions were required to be optimised to ensure sensitive conversion to cDNA.

Initially, the concentration of $MgCl_2$ used in each reaction was required to be optimised to determine the optimal concentration to enhance reverse transcriptase activity. Various aliquots of $MgCl_2$ were added to a reverse transcription reaction of 5 µg HEK293T RNA (DNAase treated) and cDNA synthesis performed under standard manufacturer's instructions. 5 µl cDNA product was added to a PCR reaction (Section 2.2.1.17) using GAAB forward and reverse primers (Section 2.1.10) to amplify the GAA repeat of the *FXN* housekeeping gene. PCR products were run on a 2% agarose gel electrophoresis (Section 2.2.1.8) to separate products (Figure 4.28). Analysis of samples in comparison to negative controls determines the highest concentration of cDNA using 3.5-4 µl $MgCl_2$. This is determined as the bands with greatest intensity are present in these samples. Negative controls, including PCR of RNA and non-template controls indicate absence of PCR products and contamination. Amplification of GAA repeats in HEK293T gDNA indicates successful amplification of PCR products, as a positive control. Therefore, 3.5 µl $MgCl_2$ was used in further reverse transcription assays.

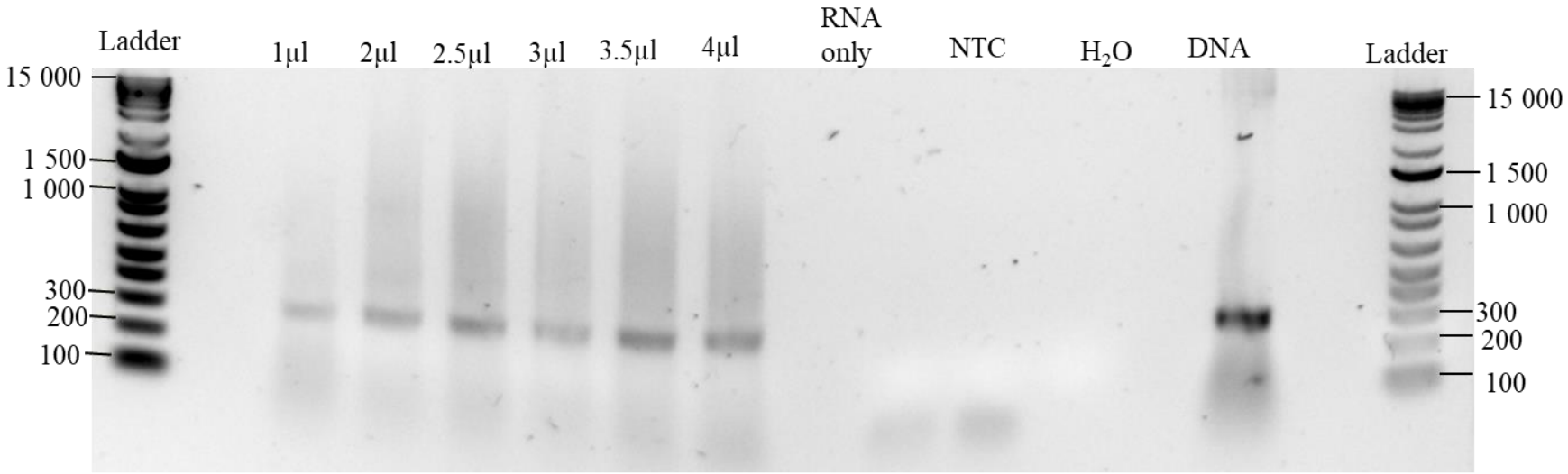


Figure 4.28- $MgCl_2$ optimisation for reverse transcriptase reaction. Various concentrations of $MgCl_2$ were added to a reverse transcriptase reaction. PCR fragments were amplified for GAA repeats, in cDNA to determine optimal concentration of $MgCl_2$ for reactions. Samples were run on a 2% agarose- TBE gel at 70V for 40 minutes using a 1 Kb+ DNA ladder. This indicates successful production of cDNA at all concentrations with optimal production with 3.5µl $MgCl_2$. Correct band sizes were identified in all cDNA and gDNA samples in comparison to negative controls

Similarly, the extension temperature used in the annealing stage of cDNA synthesis was optimised. 5 µg HEK293T RNA was used in an optimised reverse transcription reaction (Section 2.2.1.16) using 3.5 µl $MgCl_2$. A range of temperatures between the known minimum and maximum temperature for activity of the reverse transcriptase enzyme (37°C, 42°C and 55°C) were assayed. cDNA samples were analysed for presence of the *FXN* housekeeping gene by PCR amplification of the GAA repeat region. Separation of PCR fragments on a 2% agarose gel (Section 2.2.1.8) indicates presence of housekeeping gene after extension using all three parameters (Figure 4.29). The comparison of the bands produced under each extension temperature indicates the strongest band intensity is observed after extension using 42°C, in comparison to positive and negative controls.

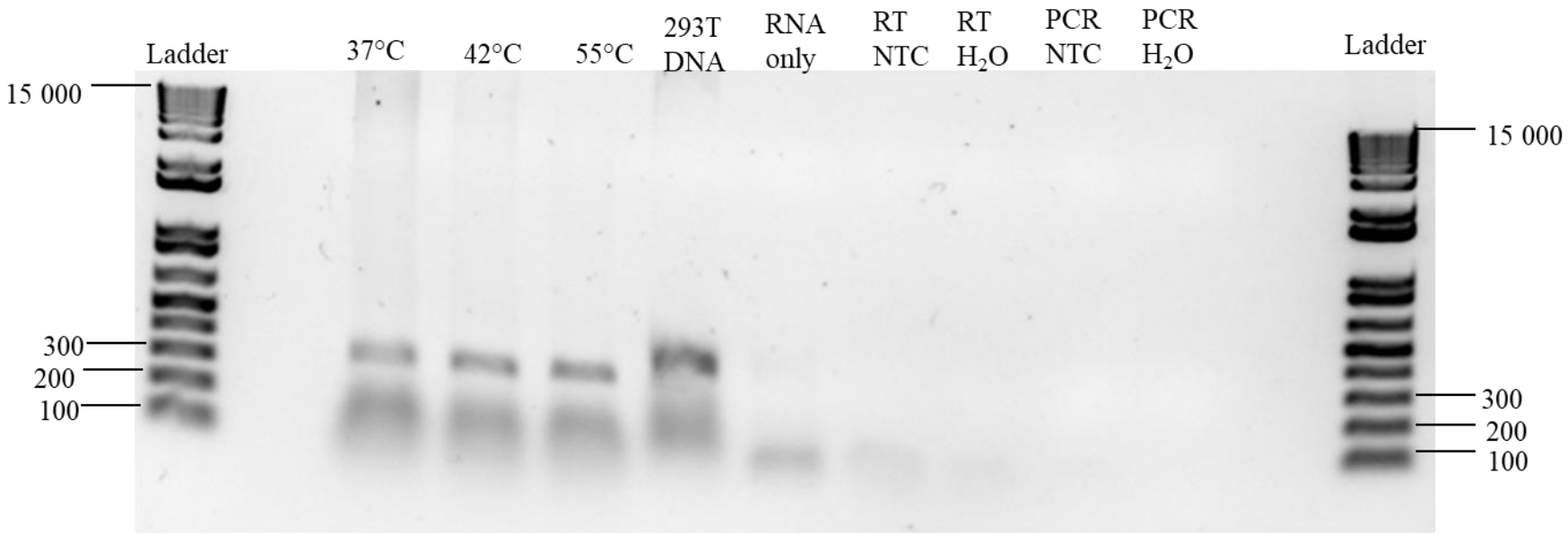


Figure 4.29- Optimisation of extension stage temperature of reverse transcriptase reaction. Reverse transcriptase reaction was extended at various temperatures (37-55°C) for cDNA production. cDNA samples were amplified using PCR for GAA repeats. Samples run on 2% agarose- TBE gel at 70V for 40 minutes. Gel indicates all extension temperatures are successful for cDNA production, with products identified in cDNA and DNA samples, in comparison to negative control

These experiments demonstrate the optimal conditions for cDNA synthesis, using 3.5 µl $MgCl_2$ per reaction and extension temperature at 42°C.

The sensitivity of this reaction was assayed to assess what concentration of RNA is needed in downstream experiments. Reverse transcription under optimised conditions (Section 2.2.1.16) was used for cDNA production using 5µg-0.1µg HEK293T RNA. GAA fragments within the *FXN* gene were amplified in samples using PCR and analysis via gel electrophoresis (Sections 2.2.1.8 and 2.2.1.17.). This indicates the reverse transcriptase reaction converts 0.1 µg RNA to cDNA (Figure 4.30). PCR products are visible in all input RNA parameters, including 0.1 µg. However, band intensity observed decreases with decreasing concentration of input RNA, as expected. Comparison to negative and positive controls indicates successful PCR reactions and absence of contamination. This demonstrates the sensitivity of the reverse transcription reaction, to 0.1 µg of RNA.

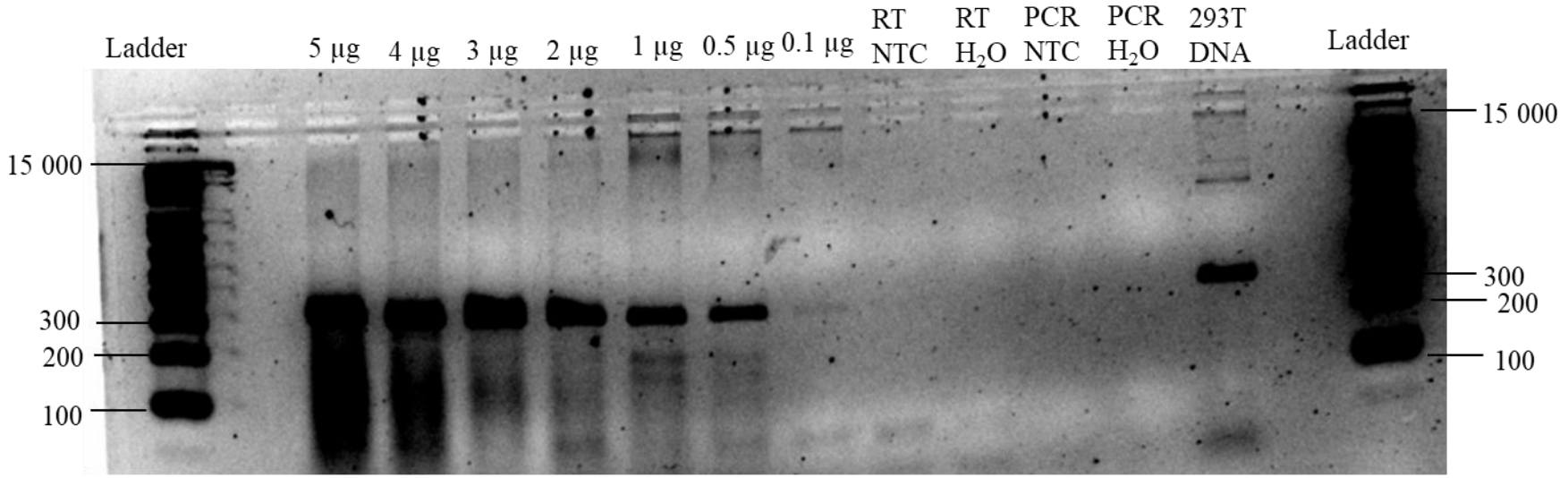


Figure 4.30- Sensitivity of reverse transcriptase reaction. Reverse transcriptase sensitivity determined using various concentrations of RNA for cDNA production. 5 µg-0.1 µg input RNA was used in reactions and amplified using PCR for GAA repeats. Samples run on 2% agarose- TBE gel at 70V for 40 minutes. Overexposure of gel indicated 0.1 µg RNA converted to cDNA

4.2.16. Nucleic Acid Purification

Nucleic acid purification of cellular gDNA or RNA were performed using DNeasy and RNeasy mini kits, according to manufacturer's instructions (Section 2.2.1.12 & 2.2.1.13). Briefly, HEK293T cell lysates were bound to a column before removal of contaminants and elution. High concentrations of pure gDNA were purified through the kit, as shown through quantification via spectrophotometry analysis. 297 ±71.67 ng/µl gDNA was purified with low concentrations of contaminating RNA present, as shown by Abs_{260}/Abs_{280} ratio reading of 2.03±0. Purified gDNA was stored at -20°C for future use.

Similarly, cellular RNA from HEK293T cells were purified using RNeasy mini kit (Section 2.2.1.13). A high concentration of cellular RNA was purified, quantified through spectrophotometry as 1650.95 ±360.93 ng/µl. RNA was treated on column with DNAase to remove gDNA contamination, according to manufacturer instructions (Section 2.2.1.15). Purity of RNA, as determined by Abs_{260}/Abs_{280}, was 2.05±0.01. Eluted RNA was stored at -80°C for future use.

To determine successful activity of DNAase, 422 ng/µl of gDNA was treated with DNAase treatment and eluted, according to manufacturer's instructions. Analysis of eluted sample after DNAase treatment via spectrophotometry indicated only 19.30 ng/µl gDNA remained. The 7.75-fold reduction in gDNA after DNAase treatment indicates successful activity of DNAase enzyme for successful removal of gDNA.

The purification of HEK293T gDNA and RNA to high concentrations and purity allows further use of cellular nucleic acids in downstream assays.

4.2.17. Lentiviral RNA Purification and Contamination Analysis

HIV-1 based lentiviral RNA has been sequenced to show wild type packaging of nucleic acids. The RNA transcripts packaging within HIV-1 based LV vectors used for gene therapy should be sequenced to assess transcripts packaged within these viruses which may have a deleterious effect to the host.

Lentiviral RNA was purified to analyse sequences packaged within virions. pHR LV and LV carrying no transgene (empty LV) was purified and DNAase treated prior to PACBIO sequencing using QIAamp viral RNA mini kit (Section 2.2.1.14). pHR LV titre was estimated as $7.10x10^{10}$ TU/ml. Empty particles are known to be generated from transient transfection of HEK293T cells when generating LV for gene transfer (Committee for Medical Products for Human Use 2005). Empty viruses were titrated using a P24 gag assay (Section 2.2.1.11) as $4.62x10^{13}$ LP/ml. Both LV carrying a transgene (pHR) and with no packaged transgene (empty LV) were sequencing to compare differences between viral packaging.

Viral RNA was treated with on column DNAase digestion to remove contaminating gDNA and quantified via spectrophotometry. Analysis of samples indicates 139.23±63.10 ng/µl pHR LV RNA was purified while only 74.50±2.47 ng/µl RNA was purified from empty LV. While 0.46-fold (64.73 ng) less RNA is purified from empty LV, as expected, RNA can be detected in both full and empty particles.

Viral RNA samples were analysed for gDNA contamination to assess purity of samples prior to sequencing (Figure 4.31). RNA samples underwent reverse transcription and PCR amplification of β- actin housekeeping gene (SY100216195-080 & SY100216195-081 primers) and of the LTR present in LV transgenes (LTR primers). Positive and negative controls were used to determine successful cDNA production, DNAase treatment and absence of

contamination. Similar concentrations of HEK293T RNA (untreated with DNAase), HEK293T cDNA (from DNAase treated RNA), HEK293T DNA and DNAase treated HEK293T DNA were used in conjunction with non-template controls to validate results. Absence of PCR products in non-template controls indicates absence of contamination in samples. The presence of B- actin band in HEK293T RNA (untreated with DNAase) and absence of PCR products in HEK293T DNA treated with DNAase indicates successful activity of DNAase to remove contaminating gDNA. The presence of a β- actin PCR product in HEK293T gDNA samples indicates successful PCR amplification. The presence of a β- actin PCR product in HEK293T cDNA samples indicate successful reverse transcriptase reaction and cDNA production. The negative control for the LTR fragment was successfully indicated by the absence of the LTR fragment in cellular cDNA, RNA and gDNA samples.

Viral samples were all shown as not contaminated by gDNA. In pHR LV samples, both RNA and cDNA samples were devoid of β- actin PCR product. The LTR fragment is present in both viral RNA and cDNA samples. While this was not expected in viral RNA samples, it has been reported that endogenous reverse transcription occurs in HIV particles prior to infection and integration into the host gDNA (Zhang, Dornadula and Pomerantz, 1996). Empty LV RNA samples are similarly absent for the β- actin housekeeping gene PCR product. However, LTR fragments can be observed in both RNA and cDNA sample suggesting packaging of endogenous samples. These samples all suggest the integrity of viral RNA samples for sequencing.

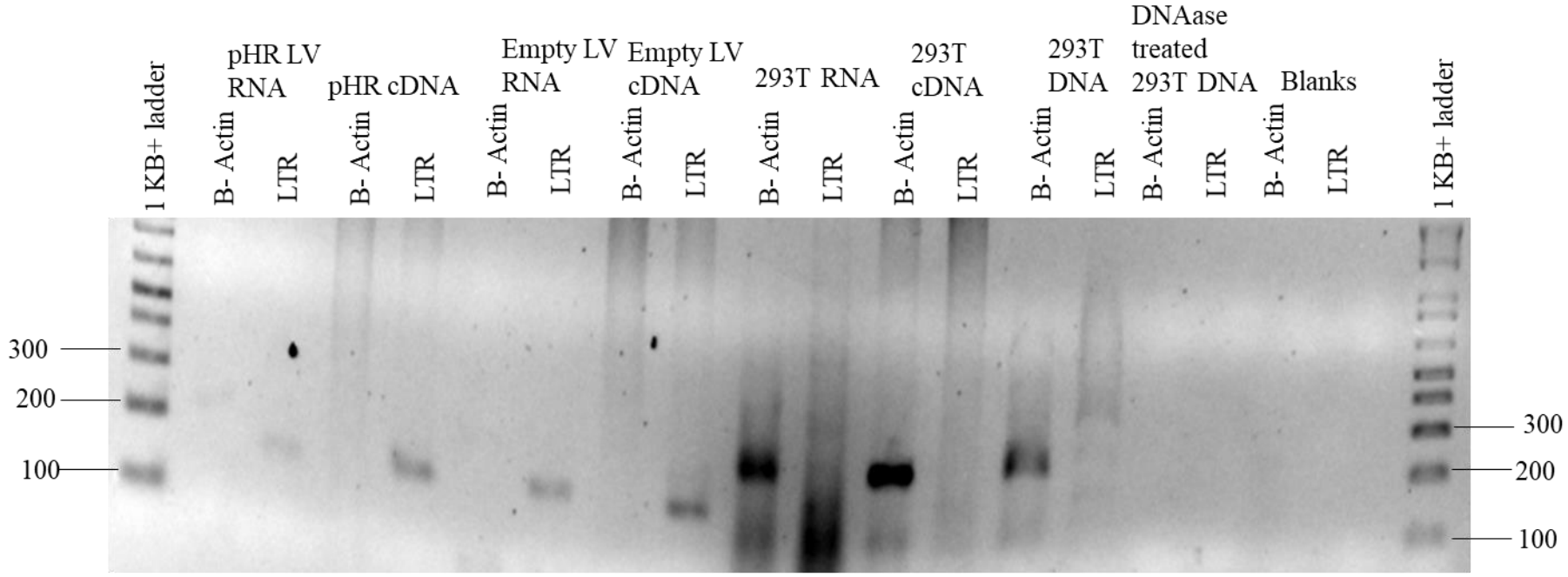


Figure 4.31- RNA contamination analysis of viral RNA samples. Viral RNA samples before and after reverse transcription were analysed for the presence of gDNA using B- actin primers and viral sequences (LTR). Absence of B actin in viral samples and presence of B actin in cellular nucleic acid samples indicates viral samples are pure of gDNA contamination. Samples were run on a 2% agarose- TBE gel at 70V for 35 minutes to separate bands in the DNA marker

4.2.18. LV RNA Sequencing

LV contains ssRNA which is converted to cDNA prior to integration in the host. Therefore, the packaging contents of LV particles should be understood to determine RNA segments possibly being transferred to hosts in gene therapy. The RNA from two LV vectors were sequenced and analysed by the Centre for Genomic Research (CGR) at the University of Liverpool using PACBIO sequencing. Both pHR and empty LV were sequenced as empty viral particles can be produced and concentrated in viral preparations produced through transient expression of plasmids via split genome packaging (Committee for Medical Products for Human Use 2005). Therefore, both LV carrying the transgene and without are likely to be delivered in treatment and the differences between the RNA packaged within each virus should be understood.

Viral RNA provided to CGR were DNAase treated and examined for gDNA contamination, as requested. Sample RNA was initially quality controlled to examine integrity of RNA provided prior to sequencing against known positive and negative controls (Figure 4.32). Quality control analysis of pHR (18170_1) and empty (18170_2) LV samples indicate the presence of rRNA. 18S and 28S rRNA was quantified as 0.3% and 0.9% of the total RNA sample in pHR LV RNA. Empty LV RNA sample was shown to contain 2.2% and 2.8% 18S and 28S rRNA respectively. Therefore, samples were treated by CGR with a proprietory ribozyme depletion enzyme for human rRNA. Subsequently, pHR RNA samples were shown to contain no rRNA and empty LV RNA quantified to contain a reduced quantity of rRNA, with 1.1% and 1.7% 18S and 28S rRNA detected in the treated sample. As expected, rRNA depletion decreased the overall quantity of RNA in the sample, with pHR LV RNA decreased by 8084 pg/µl, from 8,409 pg/µl to 325 pg/µl and empty LV RNA decreased by 5306 pg/µl, from 5,859 pg/µl to 553 pg/µl. Viral RNA was considered suitable for poly(A) tailing and library preparation for PACBIO sequencing.

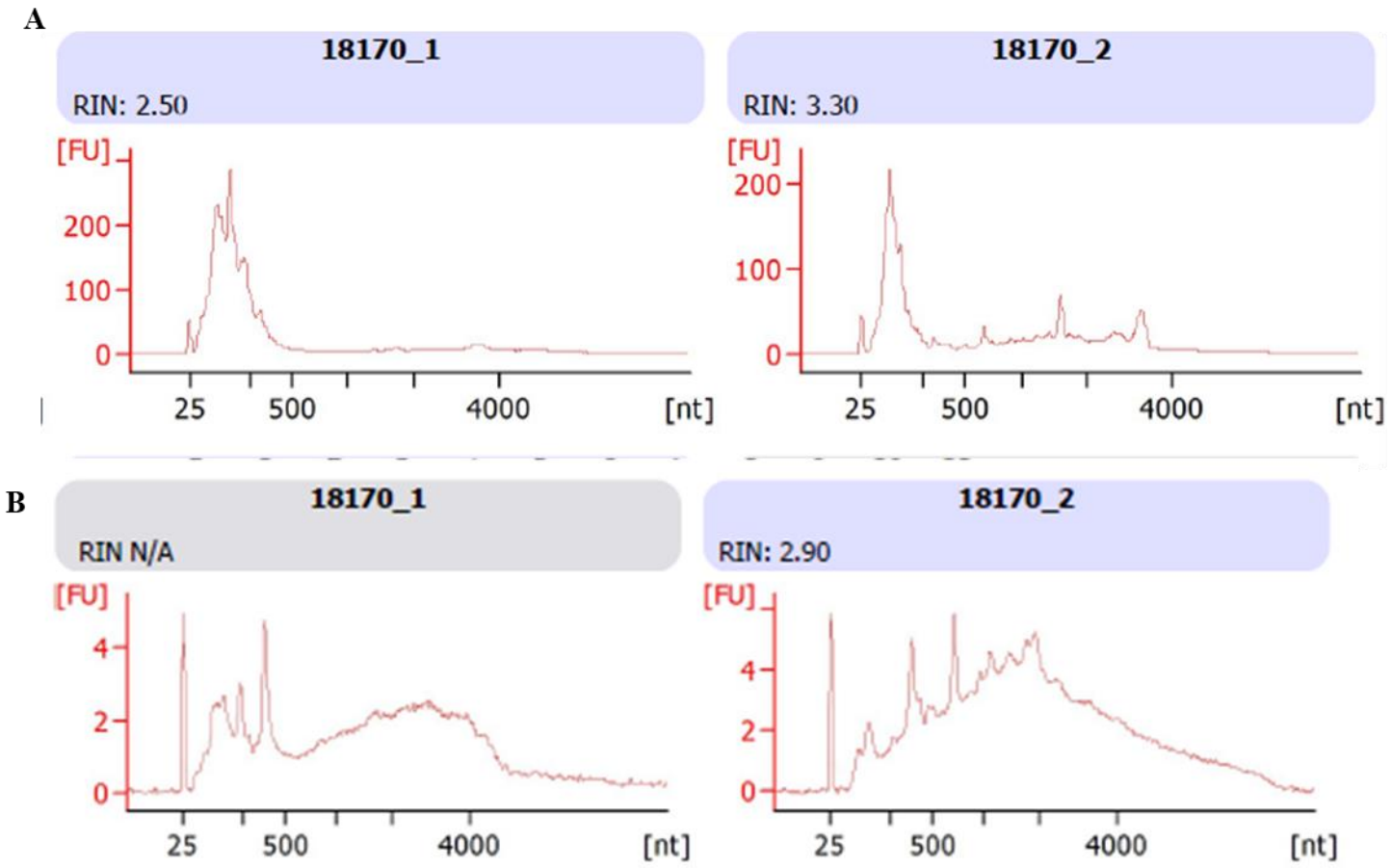


Figure 4.32- Quality control of pHR (18170_1) and Empty (18170_2) LV RNA samples. Initial analysis detected the presence of 18S and 28S rRNA (A). Ribozyme treatment of samples and analysis detected a decrease in rRNA detection (B). Integrity of rRNA depleted samples were sufficient for PACBIO sequencing. Figures kindly provided by CGR, University of Liverpool

Sequence data was processed, and intermediate circular consensus (CCS) reads identified through an IsoSeq pipeline within SMRT Link software suite (v5.1.0.26412) by CGR to generate highly sensitive reads. 3,956,793 greater nucleic acid reads can be identified in pHR (6,617,371) compared to empty (2,660,578) LV RNA. However, these identify with 33,336 greater CCS reads in empty LV (691,011) compared to pHR (657,675) LV. The mean CCS length is similar between pHR (1,272 bp) and empty (1823 bp) LV. Sequences were filtered via multiple criteria. Sequence reads were initially filtered for the single best hit that were ≥80bp, deriving from sequence reads that were ≥100bp. Sequences were aligned to the RefSeq database and filter criteria included an alignment match for ≥10% of the CCS read length. Post filtering, 23.6% (155, 352) and 4.3% (3,000) CCS reads for pHR and empty LV sequences were taxonomically aligned. The reduced alignment of empty LV RNA is likely due to the reduced number of initial nucleic acid reads due to the reduced viral RNA content in empty LV, as expected.

Filtered CCS sequences which aligned to the viral RefSeq database indicate a range of RNA identified in HIV particles (Table 4.1). Whist Woodchuck hepatitis virus is identified as the most abundant sequences in pHR LV, this is likely due to a contamination of the RefSeq database as alignment of these reads with the full nucleotide database identify the most significant hit aligned to cloning vector pHR-hSyn-EGFP. This is expected due to the nature of the transgene construct used for packaging within this vector. No sequences present in empty LV align with this taxon class, as expected due to no transgene provided in *trans* in viral production. The most abundant sequences in empty LV and second most abundant aligned sequences present in pHR LV aligned with HIV-1. This is expected due to the backbone of both pHR and empty LV particles. However, there are other sequences in the viral samples which align with other viruses, including polyomavirus, murine leukemia virus and

salmonella virus SP6. This suggests that other viral sequences are packaged endogenously in HIV vectors generated for use in gene therapy.

Species	Taxon ID	pHR	Empty	pHR-RelativeAbundance	Empty-RelativeAbundance
Woodchuck hepatitis virus	35269	69357	0	4.46E-01	0.00E+00
Human immunodeficiency virus 1	11676	51539	2446	3.32E-01	8.15E-01
Macaca mulatta polyomavirus 1	1891767	21023	160	1.35E-01	5.33E-02
Friend murine leukemia virus	11795	9164	0	5.90E-02	0.00E+00
Salmonella virus SP6	194966	1875	21	1.21E-02	7.00E-03
BeAn 58058 virus	67082	946	274	6.09E-03	9.13E-02
Vesicular stomatitis Indiana virus	11277	574	17	3.69E-03	5.67E-03
Human adenovirus 5	28285	190	43	1.22E-03	1.43E-02
Spleen focus-forming virus	11819	132	0	8.50E-04	0.00E+00
Chrysochromulina ericina virus	455364	120	3	7.72E-04	1.00E-03
Human betaherpesvirus 5	10359	75	3	4.83E-04	1.00E-03
Moloney murine leukemia virus	11801	61	0	3.93E-04	0.00E+00
Abelson murine leukemia virus	11788	59	0	3.80E-04	0.00E+00
Pandoravirus dulcis	1349409	57	0	3.67E-04	0.00E+00
Finkel-Biskis-Jinkins murine sarcoma virus	353765	40	1	2.57E-04	3.33E-04
Dishui lake phycodnavirus 1	2079134	25	4	1.61E-04	1.33E-03
Melbournevirus	1560514	16	0	1.03E-04	0.00E+00
Cotesia congregata bracovirus	39640	14	2	9.01E-05	6.67E-04
Penguinpox virus	648998	10	0	6.44E-05	0.00E+00
Micromonas pusilla virus 12T	755272	9	0	5.79E-05	0.00E+00
Harvey murine sarcoma virus	11807	7	4	4.51E-05	1.33E-03
Human endogenous retrovirus K113	166122	7	4	4.51E-05	1.33E-03
Human mastadenovirus C	129951	7	3	4.51E-05	1.00E-03
Marseillevirus marseillevirus	694581	7	0	4.51E-05	0.00E+00
Pandoravirus salinus	1349410	6	1	3.86E-05	3.33E-04
Human adenovirus 1	10533	5	2	3.22E-05	6.67E-04
Pandoravirus quercus	2107709	5	1	3.22E-05	3.33E-04
Escherichia virus P1	10678	3	0	1.93E-05	0.00E+00
Molluscum contagiosum virus subtype 1	10280	3	0	1.93E-05	0.00E+00
Ectropis obliqua nucleopolyhedrovirus	59376	3	1	1.93E-05	3.33E-04
Y73 sarcoma virus	11884	1	2	6.44E-06	6.67E-04
Avian leukosis virus - RSA	363745	1	0	6.44E-06	0.00E+00
Pestivirus giraffe-1 H138	119222	1	2	6.44E-06	6.67E-04
Moloney murine sarcoma virus	11809	1	0	6.44E-06	0.00E+00
Shigella phage SfIV	1407493	1	0	6.44E-06	0.00E+00
Macropodid alphaherpesvirus 1	137443	1	0	6.44E-06	0.00E+00
Yokapox virus	1076255	1	0	6.44E-06	0.00E+00
Camelpox virus	28873	1	0	6.44E-06	0.00E+00
Monkeypox virus Zaire-96-I-16	619591	1	0	6.44E-06	0.00E+00
Canarypox virus	44088	1	0	6.44E-06	0.00E+00
Brazilian marseillevirus	1813599	1	0	6.44E-06	0.00E+00
Only Syngen Nebraska virus 5	1917232	1	0	6.44E-06	0.00E+00
Pandoravirus neocaledonia	2107708	1	0	6.44E-06	0.00E+00
Bovine viral diarrhea virus 3 Th/04_KhonKaen	402584	0	2	0.00E+00	6.67E-04
Cafeteria roenbergensis virus BV-PW1	693272	0	4	0.00E+00	1.33E-03

Table 4.1- Analysis of PACBIO sequenced RNA from viral samples. Taxonomic assignment based on filtered BLAST hits. Quantification of sequences aligned with taxonomic classes shown in pHR and Empty columns. Relative abundance of sequences in sample set shown for each viral sample, representing the proportion of reads mapping to each species. Provided by CGR, University of Liverpool

Sequences reads are unlikely to have come from contaminating human samples. cDNA synthesis and PCR analysis confirm the absence of gDNA contamination. Analysis of CCS reads also indicates no significant hits to HIV-1 psi (ψ) packaging sequence in empty LV. This is expected due to no transgene provided for packaging. pHR RNA identify 3,370 CCS reads containing significant hits to the HIV packaging sequence. Of these 3,347/3,370 (99.32%) identify with pHR, HIV and human reference sequences. 3,289 (98.27%) sequences align with pHR, 26 (0.78%) with HIV and 32 (0.96%) with human reference sequences. This suggests the viral RNA samples do not derive from a human source, with the majority of ψ packaging sequences deriving from pHR.

The ψ packaging sequence is known to be involved in packaging of sequences into lentiviral capsids (Kuzembayeva, Dilley et al. 2014). Sequence identity to the packaging sequence identified only 3,370/657,675 (0.51%) CCS reads contain the ψ packaging sequence in pHR LV. This suggests that the majority of sequences identified in these virions are aberrantly packaged. No CCS reads identify significantly with the ψ packaging sequence in empty LV CCS reads, as expected due to the absence of a packaged transgene. The relative abundance of sequence aligning to woodchuck hepatitis virus and HIV-1 are found to be higher in comparison to other sequences identified in pHR compared to empty LV. pHR LV contains 0.22 relative abundant sequences greater compared to other sequences identified. Empty LV contains 0.18 relative abundance sequences greater in non- HIV sequences compared to CCS reads which identify with HIV. This suggests that while the quantity of sequences packaged depend on transgene packaging, both full and empty LV particles contain aberrant sequences.

Mining of sequence reads identifies unique sequences to either pHR or empty LV. Two aligned sequences, namely *Bovine viral diarrhea* virus 3 Th/04 KhonKaen and *Cafeteria roenbergensis* virus BV-PW1 are not identified in

pHR samples while these sequences are detected to a relative abundance in empty LV (0.000667 and 0.00133 respectively). However, 23 taxonomical sequences are not identified in empty LV in comparison to full viral particles (*Woodchuck hepatitis* virus, *Friend murine leukemia* virus, *Spleen focus-forming* virus, *Moloney murine leukemia* virus, *Abelson murine leukemia* virus, *Pandoravirus dulcis*, *Melbournevirus*, *Penguinpox* virus, *Micromonas pusilla* virus 12T, *Marseillevirus marseillevirus*, *Escherichia virus* P1, *Molluscum contagiosum* virus subtype 1, *Avian leukosis* virus – RSA, *Moloney murine sarcoma* virus, *Shigella phage SfIV*, *Macropodid alphaherpesvirus 1*, *Yokapox* virus, *Camelpox* virus, *Monkeypox* virus Zaire-96-I-16, *Canarypox* virus, *Brazilian marseillevirus*, *Only Syngen Nebraska* virus 5, *Pandoravirus neocaledonia*). While these are shown not to contain the HIV ψ packaging sequence, the packaging of these RNAs suggests that HEK293T cells may package other sequences independent of the HIV packaging sequence. As a greater number of unique taxonomical alignments are identified in empty LV, this suggests the absence of a transgene induces packaging of aberrant sequences into viral particles.

This identifies clear differences between full and empty LV particles and suggests other aberrant sequences are packaged, which may be deleterious to genomic integrity if transferred to the host.

4.3. Discussion

LV are commonly used in gene therapy for permanent correction of genetic diseases. LV vector production faces several bottlenecks for expanding development for clinical production. This includes producing high yields of plasmids for use in transfection assays for LV production through split genome packaging. Various steps have been used to optimise pDNA yield, including optimisation of E. Coli origin of replication and carbon source fermentation protocols (Carnes, Hodgson et al. 2004). Mediums for bacterial growth have been developed to optimise carbon source concentrations (Xu, Z. N., Shen et al. 2005). Various other methods have been used to improve plasmid yield, including Doggybone™ DNA. This closed linear DNA is generated through enzymatic digestion to generate long concatamer repeats (Scott, V., Patel et al. 2015). Doggybone™ DNA can generate LV particles to a titre comparable with traditional plasmid transfection. This system eliminates contamination of pDNA samples with bacterial components and can be produced rapidly (Karda, Counsell et al. 2019). Here we optimise plasmid harvest protocols using PDM, which has been shown to grow bacteria to high yields in a cost-effective manner in comparison to LB medium (Danquah, Forde 2007). Increasing the incubation time for growth of virus from 48 to 72 hours also increases the yield of plasmid harvested by a minimum of 2-fold. While differences persist between plasmid concentrations harvested after 48 and 72 hours of bacterial growth, this is likely to be due to the size of the plasmid (Norman, Riber et al. 2014). Both pCMVR8.74 and LTNSbetagalW plasmid sizes are greater than 11,000 bp while the length of pMD2.G is 5822 bp. While extended bacterial growth can result in an accumulation of point mutations in plasmids, verification of plasmid sequences through enzymatic digestion demonstrates sequence integrity.

Multiple protocols have been developed for the production of high titre LV, to varying degrees of success. These include optimising culture conditions,

concentrations of plasmid used and transfection reagents (i.e., PEI and $CaPO_4$) (Pham, Kamen et al. 2006). Scalable LV production to ~$1x10^9$ TU/ml were generated in fixed bed bioreactors (Valkama, Leinonen et al. 2017). While transfection reagents have been optimised for plasmid transfection, comparisons are required between reagents to determine the optimal transfection reagent for LV vector production. Protocols have been published for transfection of cells using PEI or Fugene® 6 (Besnier, Takeuchi et al. 2002). However, only one protocol can be identified for production of LV using Genejuice® and thus requires optimisation to determine the most effective ratio to pDNA for transfection (Mekkaoui, Parekh et al. 2018). Cells transfected using various ratios of Genejuice® to a consistent concentration of pDNA demonstrate that a 1:2 and 1:3 ratio produce the highest titre of HIV-1 based LV compared to 1:1, 1:6 and 1:8 ratios and maintain cell health. While a 1:3 ratio does produce the highest titre of LV ($1.23x10^7$), this still decreases viability 10% lower than observed with a 1:2 reagent transfection, suggesting a 1:2 ratio of pDNA to Genejuice is optimal for long term culture of transfected cells for viral production.

Comparison of optimised LV production protocols using PEI, Fugene® 6 or Genejuice® indicates that the use of Genejuice® is most effective for high titre LV generation. Analysis of LV carrying *LacZ* transgene demonstrates a 4.16-fold higher titre than Fugene® 6 and a 39.34-fold greater titre than when using PEI. More sensitive analysis of LV titration using flow cytometry of pHR LV also indicates transfection using Genejuice® produces a significantly higher titre of LV particles. LV produced from Genejuice® transfection is quantified as 0.94-fold greater than PEI and 0.93-fold greater than Fugene® 6 transfection. High titre pHV LV production is also observed with through Genejuice® transfection. The differences in titre observed between pHR and LacZ LV are likely due to the increase in sensitivity detection of pHR titration using flow

cytometry. Comparison of cell viability over time for LV production using PEI, Fugene® 6 or Genejuice® demonstrate that Genejuice® does not cause cytotoxicity. PEI is known to be toxic to cells, indicating the need for transfection reagent replacement after 4 hours. PEI is a branched molecule which complexes DNA to condense molecules. Complexes are protected from degradation from serum nucleases (Intra, Salem 2008). These complexes have net positive charges to efficiently bind to glycoproteins, proteoglycans and sulfated proteoglycans on the cell membrane, which occurs within three hours post transfection (Mislick, Baldeschwieler 1996). DNA-PEI complexes are released from the endosome through proton sponge effect, with PEI buffering the endosomal microenvironment to delay acidification, osmotic swelling and degradation of the endosome and release of DNA complexes (Pollard, Remy et al. 1997, Akinc, Thomas et al. 2005). However, PEI has been shown to aggregate when exposed to serum, reduce transfection efficiencies, and induce cellular apoptosis (Wightman, Kircheis et al. 2001, Ogris, Walker et al. 2003, Khonsari, Schneider et al. 2016). Therefore, alternative transfection reagents have been developed for gene delivery mechanisms. Polyamines and various polycations work in similar ways to PEI for transfection of cells using pDNA (Zhang, H., Vinogradov 2010). Both Fugene® 6 and Genejuice® have been marketed as low toxic alternatives for transfection. Analysis of transfected cells over time indicates cell viability for PEI drops significantly over a 72-hour period. Cells transfected for LV production using Fugene® 6 drop significantly after 48 hours but Genejuice® does not. Time course analysis of viability also indicates the low toxicity of Genejuice® transfection for LV production.

Scalable LV production demonstrates high titre LV vector production using Genejuice® over 11 days with low cytotoxicity. This is longer than expected, with transient transfection being reported to last an average of 7 days. The absence of RCLs is shown in scalable LV production suggesting this protocol is

acceptable for vector generation. The presence of RCLs is tightly regulated for the clinical use of vectors by multiple agencies to prevent adverse side effects from uncontrolled LV spread (Committee for Medical Products for Human Use 2005). These results indicate the usefulness of Genejuice® as a transfection reagent for LV production to high titre.

Pseudotyping vectors enhances the range of cells the virus can transduce. VSV-G is commonly used to pseudotype vectors as it has been shown to transduce a wide range of cells due to binding of a ubiquitous cell surface receptor (Schlegel, Tralka et al. 1983). While pseudotyping vectors is an effective method for increasing the range of cell targets, there is a need to enhance transduction efficiencies by reducing steric repulsion between the virus and cell membrane. Various enhancers have been developed for increasing LV transduction efficiencies (Schott, Leon-Rico et al. 2019). Polybrene is commonly used to transduce cells by complexation with viruses to negate the negative charge (Le Doux, Landazuri et al. 2004). Validation of enhancement indicates a maximum 2.92-fold increase in transduction with polybrene, with no observable cytotoxicity over 72 hours. Further optimisation of polybrene indicates a greater transduction efficiency after complexation of virus with polybrene, in comparison to seeding cells with polybrene. These conditions support previous work that show polybrene complexation with virus is successful in enhancing LV transduction without presenting cytotoxicity at low concentrations (Seitz, Baktanian et al. 1998).

Polybrene has shown poor transduction of certain cell types, including murine T- cells (Kerkar, Sanchez-Perez et al. 2011). Nuvec® is a reagent of proprietary spiked silicone nanoparticles complexed with PEI, produced by N4Pharma. Previously, silica nanoparticles have shown an increase in pDNA transfection (Pearce, Mai et al. 2008, Park, Jeong et al. 2016). Time course analysis of various concentrations of Nuvec® over 72 hours indicate 1-40 µg of Nuvec® is

the least toxic to HEK293T cells, with increasing concentrations decreasing cell viability. This is consistent with previous studies indicating toxicity of silica nanoparticles is dependent on particle size and dosage (Kim, I. Y., Joachim et al. 2015). Transduction of Nuvec® after complexation with LV at a range of times show an increase in transduction efficiency, greater than that of polybrene. This is dependent on the concentration of Nuvec® used. However, increasing the time of incubation does not appear to further increase percentage expression of GFP. Nanoparticles have been shown to sediment upon ultracentrifugation (Spinnrock, Martens et al. 2019). Nuvec® can be used to generate LV, however Nuvec® does not appear to be free floating in supernatant or concentrated with LV as transduction efficiencies remain consistent. The viability of cells is not altered in comparison to Nuvec® treatment alone, suggesting that LV transduction using Nuvec® does not decrease cellular health. These results suggest that LV complexation with 40 µg Nuvec® (NV00100032) for 20 minutes is the most effective method for enhancing viral transduction while maintaining cell health. Furthermore, complexation of LV with 40 µg Nuvec® and 5 µg polybrene increases transduction efficiencies still further, suggesting Nuvec® and polybrene can be used in conjunction for optimising LV transduction. Low concentrations of PEI have been shown to increase transduction efficiencies of LV (Castellani, Di Gioia et al. 2010). Thus, the effect of PEI coupling with Nuvec® should be understood further.

Purification of RNA from LV is effective in detecting sequences packaged within LV (Gall, Ferns et al. 2012, Henn, Boutwell et al. 2012, Brener, Gall et al. 2015). Contigs have previously been shown to recover up to 96% of the complete viral genome (Cornelissen, Gall et al. 2017). Purified LV RNA has shown to be free of gDNA contamination prior to sequencing through reverse transcription analysis. PCR amplification of the LTR fragment in pHR RNA

and cDNA is expected due to endogenous reverse transcription in HIV particles (Zhang, Dornadula et al. 1996). PACBIO sequencing of RNA is effective for long sequence reads of samples to detect intact transcripts (Rhoads, Au 2015). Empty particles have been shown to be generated through transient transfection of HEK293T cells (Committee for Medical Products for Human Use 2005). These particles will therefore also be concentrated along with vectors carrying the transgene and will be delivered *in vitro* and *in vivo*. Analysis of both pHR and empty LV determines CCS reads within both pHR and empty LV particles. This is despite no transgene provided in split genome packaging of empty particles. The presence of RNA in empty particles suggests aberrant transcript packaging with HIV-1 based LV. Analysis of sequences detect contaminating RNA present in both vectors, with 23 taxon classes of sequences unique to pHR suggesting packaging of these RNA transcripts is influenced by the pGFP expression. While contaminating sequences are present at a lower abundance in empty compared to full LV particles, both vectors contain aberrant RNA transcripts. 0.51% (pHR) and 0% (empty) of reads are found to contain the ψ packaging sequence suggesting the majority of packaged sequences are aberrantly packaged. The ability of these RNA transcripts to integrate within the host upon gene transfer should be understood further to determine genotoxic effects of aberrant transcripts.

Viral vector production can be optimised in a number of ways to overcome bottlenecks faced for production of high titre LV. Here, we show incubation of transformed bacteria for 72 hours before harvest increases plasmid yields for downstream processes. Genejuice® transfection of cells for LV production produces high titre LV, in a scalable fashion with low cytotoxicity and no RCL production. Furthermore, transduction of cells using LV complexed polybrene has been shown to increase transduction efficiencies with low cytotoxicity. In comparison to this, Nuvec® increases transduction efficiencies still further, in

particular after dual complexation with polybrene. Sequencing analysis of both full (pHR) and empty LV indicate the presence of contaminating RNA packaged within viruses. These results show optimised, scalable high titre production of viruses for use in downstream assays.

Chapter V Bioinformatic Analysis of Insertion Sites in iPSC and HLC Derivatives

LV vector mediated integration of the transgene into the host genome has been shown to cause adverse side effects. Understanding this genotoxicity in a human *in vitro* model is useful to improve the safety profile of these gene therapies further. As the liver is commonly used as a gold standard for toxicity analysis, here we provide a model for IS analysis based on HIV-1 based LV vector mediated insertion in iPS cells and their HLC derivatives. This proves a sensitive model for understanding vector related genotoxicity. Comparison of HIV-1 LV carrying either the native or SIN configuration LTR show further subtle differences between vectors.

5.1. Introduction

5.1.1. Induced Pluripotent Stem Cells

iPSCs are a promising alternative for various therapeutic methodologies, including autologous transplantation. These cells are reprogrammed to a pluripotent cell state through expression of various factors, such as Oct4, Sox2, Klf4 and c-Myc (Takahashi, Yamanaka 2006). iPSCs have the potential to be differentiated into various lineages, including cardiomyocytes and hepatocyte-like cells (HLC). This affords scientists an indefinite source of biological material for patient specific tissue engineering.

Pluripotent stem cells, including induced pluripotent stem (iPS) cells are a promising source of cells for various therapeutic avenues, including gene therapy. The premise of somatic reprogramming was first demonstrated with somatic nuclear transplantation. Nuclear transfer from Xenopus somatic cells into oocytes show successful characteristics of the somatic nuclear cell, with cells expressing endoderm in marker after nuclear transfer from the late neurula stage (King, Briggs 1955, Briggs, King 1952). Even somatic nuclear transfer from late stage endodermal cells into oocytes still form functional nerve and muscle cells, though at low efficiency (Gurdon 1960). These experiments suggest that reprogramming factors reside with the nucleus of differentiated cells to promote cell differentiation.

iPS cells are of huge promise as somatic cells have been reprogrammed into a pluripotent cell state, overcoming ethical issues faced using embryonic stem cells. These cells were initially derived from mouse fibroblasts and after exposure to *Oct3/4*, *Sox2*, *c-Myc* and *Klf4* genes were reprogrammed to iPS cells. While it was unknown which candidate genes were vital to pluripotency reprogramming, the screening of 24 candidate genes using a drug resistance

platform and β- galactosidase reporter gene indicated no one gene induced pluripotency by itself. Screening of combinations of factors initially indicated *Oct3/4*, *Sox2*, *c-Myc* and *Klf4*, where required for reprogramming of cells to an undifferentiated cell state (Takahashi, Yamanaka 2006). These cells were fully characterised as pluripotent through morphological analysis, gene expression and expressing surface markers similar to that of human embryonic stem cells, with the differentiation potential of these cells demonstrated.

The four Yamanaka factors have been known to be involved in pluripotency. Both *Oct3/4* and *Sox2* are transcription factors known to be involved in pluripotency, with high expression of both factors at the embryonic stem cell stage (Boyer, L. A., Lee et al. 2005). *c-Myc* is a known oncogene, with downstream targeting for cellular proliferation (Adhikary, Eilers 2005). *c-Myc* has been mapped for an excess of 25,000 targets in the human genome, which has been suggested to trigger histone acetylation and unwinding and transcription of DNA, with c-Myc binding to TRRAP unit of histone acetyltransferase complexes (McMahon, Van Buskirk et al. 1998). *Klf4* has also been implicated in oncogenesis and is known to directly repress the *p53* tumour suppressor gene (Rowland, Bernards et al. 2005). Low levels of p53 were initially observed by Yamanaka in iPS cells. Similarly, activation of p21CIP1 by Klf4 has been shown to suppress proliferation. While Myc has been found to suppress p21CIP1 and promote proliferation, the interplay of these two factors in cells triggers reprogramming (Zhang, W., Geiman et al. 2000, Seoane, Le et al. 2002). These genes are therefore useful for reprogramming of somatic cells.

Since Yamanaka's work, developments have been made in improving the efficiency of reprogramming cells. Initially, RVs were used to deliver the four genes to somatic cells. While this system reprogrammed cells into a pluripotent state, the low efficiency and integrative nature of retroviruses causes bias towards differentiation towards specific cell types and tumorigenesis (Yu,

Vodyanik et al. 2007, Okita, Ichisaka et al. 2007). Methods have been developed to remove integrated sequences to prevent insertional mutagenesis. Cre/LoxP systems have been developed to excise DNA with integrated sequences flanked by LoxP sites which are excised after transient transfection of cells with Cre recombinase (Kadari, Lu et al. 2014). Other systems to remove integrated genes include employing PiggyBAC vectors or through use of oriP/EBNA1 based plasmids (Woltjen, Michael et al. 2009, Kaji, Norrby et al. 2009, Yu, Hu et al. 2009). These plasmids have been derived from the Epstein-Barr virus, with integrative sequences lost with further proliferation without a drug selective marker. These methods have been shown as successful in removing integrated sequences used for reprogramming.

Due to the danger of integrative methods of reprogramming, episomal vectors have been developed for reprogramming to iPS cell state. Episomal DNA has been shown to reprogram human somatic cells using Yamanka's factors and alternatively *OCT4*, *SOX2*, *NANOG* and *LIN28*. Episomal delivery of Yamanaka factors using Sendai viruses has been shown to successfully reprogram both neonatal and adult fibroblasts at a high efficiency (1%) (Fusaki, Ban et al. 2009). Optimisation of virus infection using alternative incubation temperatures has been shown to successful reprogram iPS cells from haemocytes (Seki, Yuasa et al. 2010). While the efficiency of temperature manipulations was 10-fold lower than observed previously, this rate is still comparable to efficiencies observed when using integrative methods of reprogramming.

An alternative method for episomal reprogramming employs adenoviruses. Gene delivery using these non-integrative viruses has been shown to successfully generate iPS cells. However, it has been found that a quarter of adenoviral iPS cell lines are tetraploid, possibly due to cell fusion from to viral

infection (Stadtfeld, Nagaya et al. 2008). The rate of efficiency was also lower than when using RVs and LVs.

Insertional mutagenesis predisposes integrative iPS cells to oncogenesis. Embryonic stem cells have been reported to form teratomas in murine hearts suggesting further issues with iPS cells (Gertow, Cedervall et al. 2007, Nussbaum, Minami et al. 2007). While developments have been made in episomal reprogramming of cells, the oncogenic potential of these cells persist. The cocktail of genes which are used for cellular reprogramming are known to be involved in oncogenesis. Upregulation of *Oct3/4* has been implicated in multiple cancers (Klimczak 2015). While there are multiple common genes expressed at the embryonic and induced pluripotent cell states, iPS cells have a high quantity of unique genes activated predisposing these cells towards oncogenesis irrespective of the somatic cell source. This has further been shown in comparison to mesenchymal stem cells, with iPS cells shown a 4.5 and 1.5-fold higher expression of *c-Myc* and *KLF4* respectively (Foroutan, Najmi et al. 2015). While *OCT4*, *SOX2*, *NANOG* and *LIN28* were found expressed higher in mesenchymal stem cells, iPS cells still show a higher oncogenic potential due to its pluripotent state. These observations need to be accounted for in downstream assays.

5.1.2. Hepatocyte-Like Cells as a Toxicity Model

Hepatocytes are commonly used as the gold standard for understanding biochemical toxicity. Tissue specific toxicity (i.e. cardiomyocyte) have been assessed for target specific compounds (Sinnecker, Laugwitz et al. 2014). Liver toxicity remains an important issue to address, with 40% of drug compounds failing in clinical trials due to this reason (MacDonald, Robertson 2009). Hepatocellular carcinoma (HCC) remains a prevalent disease, accounted as the

fourth most common cause of cancer related deaths worldwide. While the incidence of HCC varies greatly worldwide, this carcinoma represents the most common cause of hepatic cancers diagnosed (Singal, Lampertico et al. 2020). The shift from the use of animal models indicates the necessity for robust *in vitro* models of disease working to reduce the use of animals in scientific research). Oncogenic hepatic cell lines, such as HepG2 cells, may be considered as a cheaper cell line that is easier to maintain and manipulate. However, these cells show a limited enzymatic activity of these cells in comparison to primary hepatocytes (Gerets, Tilmant et al. 2012). Moreover, there are significant differences in the gene expression of HepG2 cells in comparison to *in vivo* hepatocytes, both demonstrating the need for an alternative cellular model (Costantini, Di Bernardo et al. 2013). iPSC derived hepatocyte- like cells (HLC) are useful to understanding the molecular phenotype and toxicity and have the potential to show patient specific responses (Takayama, Morisaki et al. 2014, Choudhury, Toh et al. 2017). These are useful alternatives to understanding hepatotoxicity *in vitro*.

5.1.3. LV-Mediated Integration and Insertional Mutagenesis

LV- mediated gene therapy is an expanding field, with more than 180 clinical trials registered in 2020 (ClinicalTrials gov). These viruses have distinct beneficial characteristics for use in gene therapy. Reverse transcription of the ssRNA catalyses the production of double-stranded cDNA which forms pre-integration complexes (PICs). PICs associate with host chromatin to integrate the viral genome into host gDNA (Coffin, Hughes et al. 1997). HIV-1 based LV integrate semi-randomly but 70-80% preferentially within the gene body of highly transcribed genes (Ciuffi, Llano et al. 2005, Schroder, Shinn et al. 2002, Mitchell, Beitzel et al. 2004, Rossetti, Cavarelli et al. 2013). These viruses show

differences in integration site selection to other RVs. MLV integrates near transcriptional start sites (TSS) whereas the Avian Sarcoma-Leukosis virus neither prefers transcription start sites or actively transcribed genes suggesting each viral integration complex selective binds to chromatin complexes for integration (Daniel, Smith 2008, Beard, Dickerson et al. 2007, Rossetti, Cavarelli et al. 2013, Hematti, Hong et al. 2004). LV PIC recognition of the host cell chromatin is specific (Bushman, Lewinski et al. 2005, Bushman 2003, Lewinski, Yamashita et al. 2006, Barr, Ciuffi et al. 2006). This is primarily due to cellular proteins that interact with the viral integrase, including PSIP1/LEDGF/p75 protein (Cherepanov, Maertens et al. 2003, Vandegraaff, Devroe et al. 2006).

There is a need to understand the genotoxicity of LV vectors further. Various models have been developed to understand insertional mutagenesis. While these are not essential for clinical products unless there is a serious adverse effect reported, these assays provide valuable data to develop the safety profile of the LV based product (Aiuti, Cossu et al. 2013, European Medicines Agency 2019). Mouse implantation studies have been used with the generation of useful murine models expressing human cytokines to recapitulate the *in vivo* human response (Cosgun, Rahmig et al. 2014). However, mice still do not fully recapture human responses and the long period of study (~12 months) and issues surrounding transduction efficiency mean an alternative model is sought (McCarthy 2003).

Insertional mutagenesis occurs through multiple factors. The basic principle of this serious adverse effect is the integration of a therapeutic transgene delivered by HIV-1-based LV causing host gene dysregulation and oncogenesis. This process can be caused by multiple mechanisms. Insertion of the transgene is of most concern when near to or within protooncogenes, oncogenes, or tumour suppressor genes. Dysregulation of these genes through gain of function

mutations whereby promoter or enhancer activation, gene truncation or epigenetic silencing all contribute towards insertional mutagenesis and the oncogenic potential of the cell (Aiuti, Cossu et al. 2013, Suerth, Labenski et al. 2014, Touw, Erkeland 2007, Rae, Trobridge 2013)**.** Read through of the viral LTR into gene bodies may occur when transgenes are inserted upstream of transcriptional units (Suerth, Labenski et al. 2014). Thus, increasing the safety profile of LV-based vectors is of paramount importance to prevent adverse side effects in patients.

A major safety component of HIV-1 LV is the generation of replication-deficient viruses through split genome packaging. SIN LTRs are generated using a mutation or deletion of the U3 region in the virus LTR, silencing promotion of the transgene which is instead under the control of an internal promoter. This is an important step in reducing the risk of insertional mutagenesis (Miyoshi, Blömer et al. 1998, Zufferey, Dull et al. 1998).

5.1.4. Methylation Responses to LV Integration

HIV-1 transgene silencing can also be caused by epigenetic changes, primarily methylation of the genes of insertion or the 5' LTRs (Rae, Trobridge 2013). This has been shown in wild type HIV-1 where CpG hypermethylation near to the HIV genome contributes towards latency (Trejbalova, Kovarova et al. 2016, Bednarik, Cook et al. 1990). This is further validated by the reversal and reactivation of HIV transcription through azacytidine treatment mediated histone hypomethylation (Blazkova, Trejbalova et al. 2009). Two DNA methyltransferases (DMNTs) are known to be involved in CpG methylation of two regions in the 5' LTR (Jordan, Bisgrove et al. 2003). DMNT expression is upregulated upon HIV-1 infection with wild type HIV infection causing hypermethylation of the LTR (Mikovits, Young et al. 1998, Fang, Mikovits et

al. 2001). However, hypermethylation response appears to be host specific, with varying levels of methylation detected in genome wide studies (Zhang, X., Justice et al. 2016). While CpG island methylation near to the HIV genome is most common, intragenic hypermethylation is known to occur (Kint, Trypsteen et al. 2020). Thus, the silencing of the transgene is indicative of a poor outcome downstream.

One issue presented in using iPSCs as an *in vitro* model is the epigenetic memory from donor tissue, which must be accounted for when considering the methylome of treated cells in developing genotoxicity out reads (Zhou, Wang et al. 2017, Kim, K., Doi et al. 2010)

5.1.5. LV-Mediated Clonal Outgrowth

The consequence of insertional mutagenesis is clonal outgrowth of oncogenic cells, as reported in multiple clinical trials (David, Doherty 2017). The development of next generation sequencing (NGS) technologies has improved the identification and quantification of integrated viral transcripts (Giordano, Appelt et al. 2015, Schmidt, Hoffmann et al. 2001). Patients developed T cell lymphomas after an attempt to correct X-SCID, due to RV integration within the *LMO2* oncogene (Hacein-Bey-Abina, Hauer et al. 2010, Hacein-Bey-Abina, Garrigue et al. 2008, Gaspar, Cooray et al. 2011). Initial correction of chronic granulomatous disease was shown but insertional activation of *EVI1*, *PRDM16* and *SETBP1* contribute towards pathologies (Stein, Ott et al. 2010, Ott, Schmidt et al. 2006). LV integration within *HMGA2* has also contributed towards oncogenesis in a clinical trial targeted towards β- thalassemia. These serious adverse effects have also been reported in licenced clinical products (i.e. Strimvelis) (FiercePharma 2020).

This is not exclusive to humans as murine suffering from β- thalassemia also presented with sarcomas and lymphomas subsequent to gene correction (Lidonnici, Paleari et al. 2018). Another study has shown that while clonal dominance was not observed after serial transplantations of livers in mice subsequent to gene therapy, clusters of genes were identified in actively transcribed genes and the prevention of insertional mutagenesis was only likely due to the low MOI used (Rittelmeyer, Rothe et al. 2013).

iPSCs are a promising alternative for genotoxicity studies as they recapture patient specific traits *in vitro* and have the capacity to be differentiated into multiple lineages, including cardiomyocytes. iPSC derived cardiomyocytes have also been reported to be successfully transduced with VSVG pseudotyped LV (Sato, Higuchi et al. 2016, Sato, Kobayashi et al. 2015).

5.1.6. Transcription Factor Binding Sites and LV Integration

The PSIP1/LEDGF/p75 protein has a vital role in tethering the PIC to the host genome. However, IS in cells depleted in this protein still favour transcriptional units, suggesting other factors may influence viral genome IS selection (Ciuffi, Llano et al. 2005). For RV, the U3 region appears important for host transcription factor binding (Felice, Cattoglio et al. 2009). It is plausible, therefore that HIV LV IS selection and possibly PIC tethering may also involve recognition of host TFBSs that mirror the pTFBS in the HIV LTR region. Approximately 27.7% of the human genome has been found to associate with TFBS regions (Chen, H., Li, Liu, Zheng, Wang, Bo, and Shu 2015b). The hypothesis that directing IS selection through common TFBS may be important for viral survival in actively transcribed genes.

Multiple transcription factors have previously been identified to bind to the HIV genome including tumour necrosis factor alpha (TNF-α), which has been shown to activate HIV-1 transcription in chronically infected T-cells via binding of the transcription factor NF-κB to specific sites in the U3 region (Duh, Maury et al. 1989). c-myb and AP1 TFBS have been identified in the LTR, and these proteins are involved in HIV latency, transcription and infection (Dasgupta, Saikumar et al. 1990, Canonne-Hergaux, Aunis et al. 1995, Tacheny, Michel et al. 2012, Chao, Walker et al. 2003, Ma, Dong et al. 2013). Therefore, it is important to study the pTFBS identified in HIV IS to understand the mechanism of insertion further.

Here, we analyse insertion site data to show the subtle differences are found between iPSC and their HLC derivatives, including between cells transduced with either the native or SIN LTR configurations. Outgrowth of infected iPSCs for 30 days also reveal a rapid method for analysing potential clonal outgrowth. We also identify common pTFBSs are targeted for integration with the HIV-1 LTR and survival of the virus. These data shows the usefulness of using iPSC and their HLC derivatives as a model for LV mediated genotoxicity.

Aims

- Analysis of iPSC & HLC gene expression
- Analysis and enrichment of IS in LV infected iPSC & HLCs using a native or SIN configuration vector
- Comparison of pTFBS indented in infected iPSC and HLCs and LTR configurations

5.2. Results

5.2.1. iPSC and Derived HLC Gene Expression Analysis

Initially, the gene expression from iPSCs and their HLC derivatives were analysed. Hepatocytes are commonly used for pharmacological toxicity analysis (MacDonald, Robertson 2009).

Published microarray gene expression data of primary hepatocyte derived iPSCs and their HLC derivatives (GSE 61287) were analysed to determine differentially regulated gene expression, gene enrichment and gene ontologies, as previously described in Section 2.1.4.2. (Takayama, Morisaki et al. 2014)

Log_2 expression arrays were generated from normalised microarray data sets ($P<0.05$) in comparison to a primary hepatocyte control. 36,475 genes were identified on the microarray chip. Expression array values above one represents upregulation of genes and conversely, values below one represents downregulated genes (Figure 5.1). This identified that of all genes identified, 73% were upregulated in iPSC, with 27% of genes found to be downregulated. Similarly, in HLC 59% of genes identified were found to be upregulated while 41% were identified as downregulated. This indicates differences in the cell stages, with 4,483 more upregulated genes expressed at the pluripotent stem cell stage.

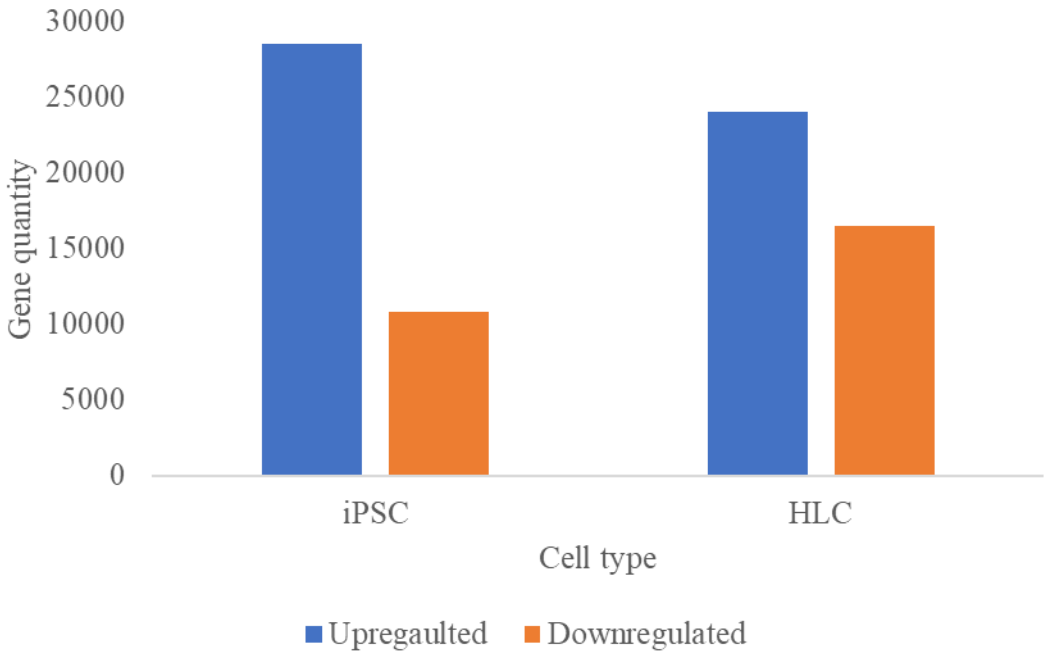
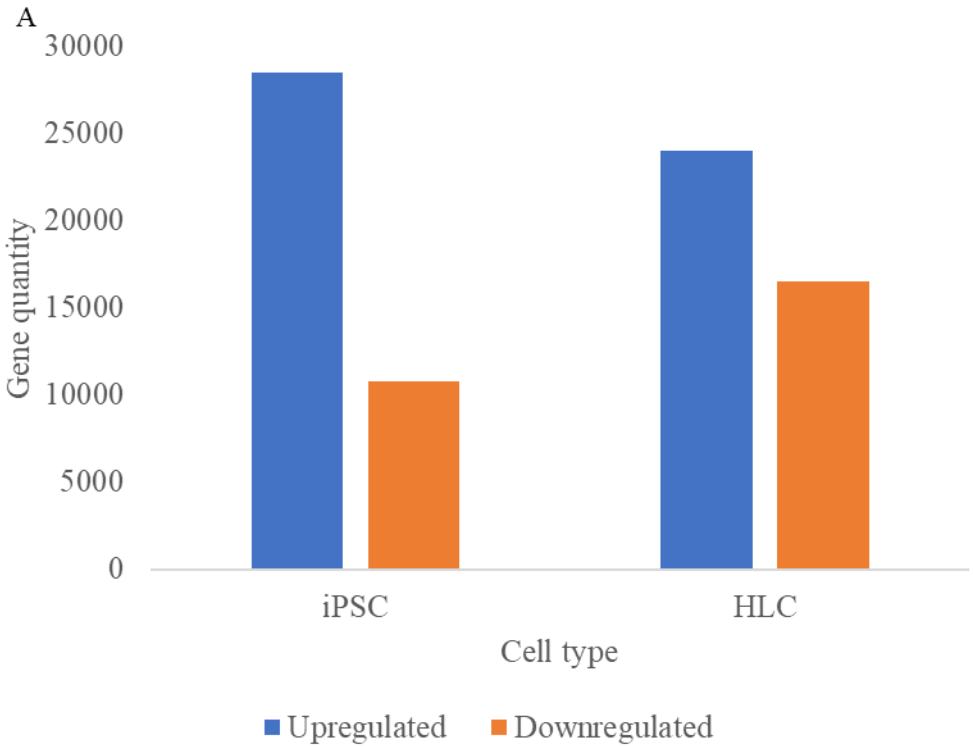


Figure 5.1- Gene regulation in iPSC and HLC derivatives. Quantity of genes found as upregulated (blue) or downregulated (orange) in iPSC or HLC derivatives through Log_2 gene expression array analysis of microarray gene expression data in comparison to primary hepatocytes

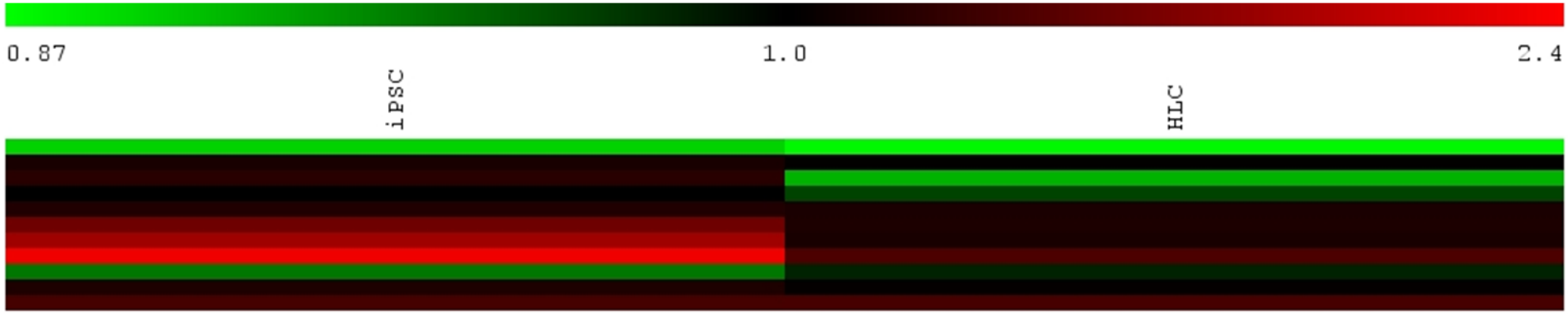
Gene ontology analysis identified cancer related genes expressed at both cell types- namely protooncogenes, oncogenes and tumour suppressor genes (UniProt Consortium 2019). 165 tumour suppressor genes, 223 protooncogenes and 11 oncogenes were identified (Figure 5.2). Analysis of expression arrays indicated differences in the expression of the genes at the iPSC and HLC stages. 82% (9) of oncogenes were upregulated in iPSC while 18% (2) were downregulated. In HLC, 64% (7) and 36% (4) of oncogenes were upregulated and downregulated, respectively. For tumour suppressor genes identified in iPSC gene expression data set, 64% (105) and 36% (60) were upregulated and downregulated, respectively. In HLC, 56% (93) and 44% (72) tumour suppressor genes were identified as upregulated and downregulated. 71% (158) of the protooncogenes identified in iPSC were upregulated while 29% (65) were found to be downregulated. In comparison, in HLC 58% (130) and 42% (93) protooncogenes were found to be upregulated and downregulated, respectively.

This further demonstrates that the majority of cancer related genes are upregulated at a greater proportion in naïve iPSCs in comparison to the terminally differentiated HLC stage.

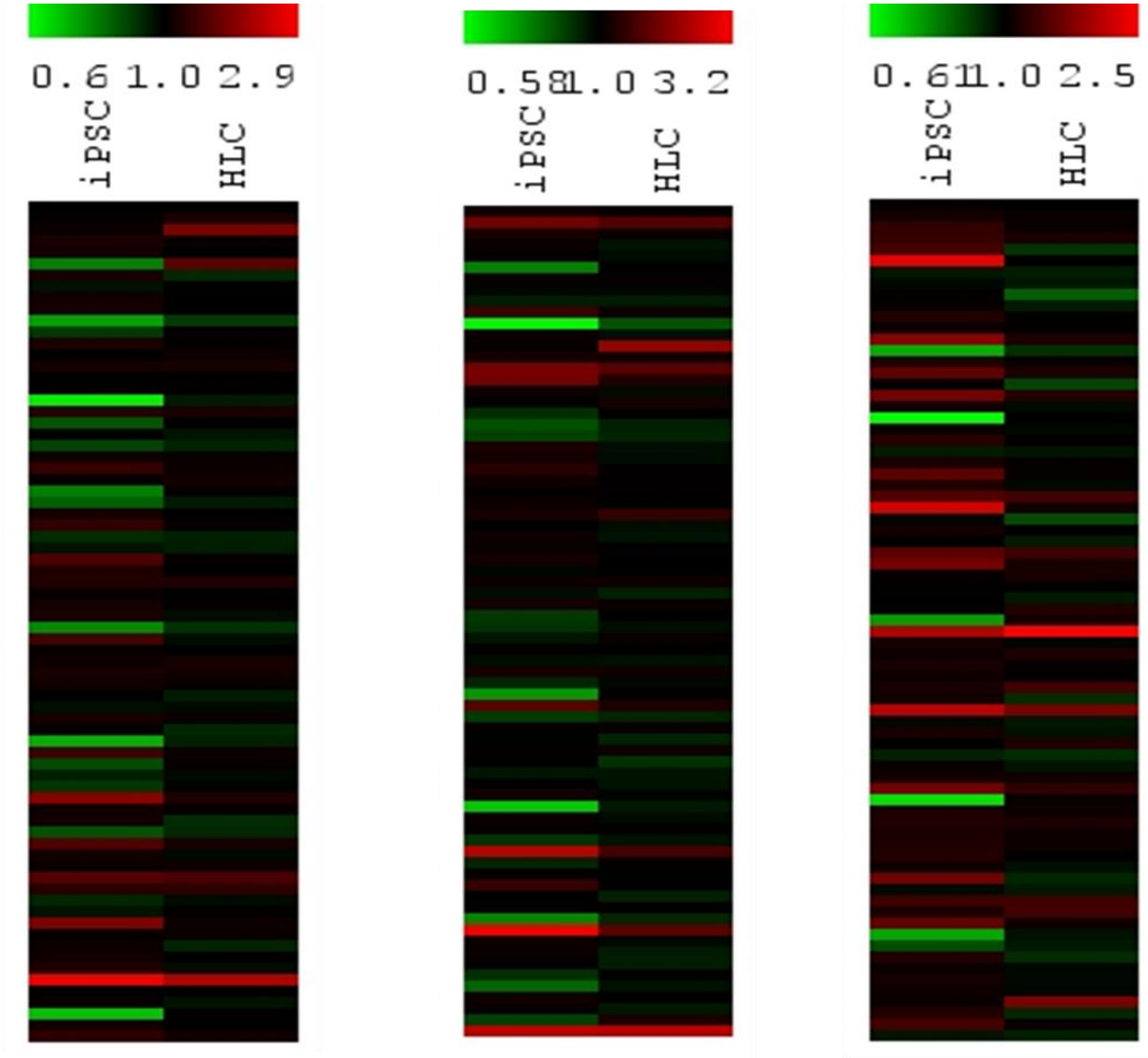
A


30000
25000
20000
15000
10000
5000
0
Gene quantity
iPSC
HLC
Cell type
Upregulated
Downregulated

B


0.87
1.0
2.4
iPSC
HLC

C



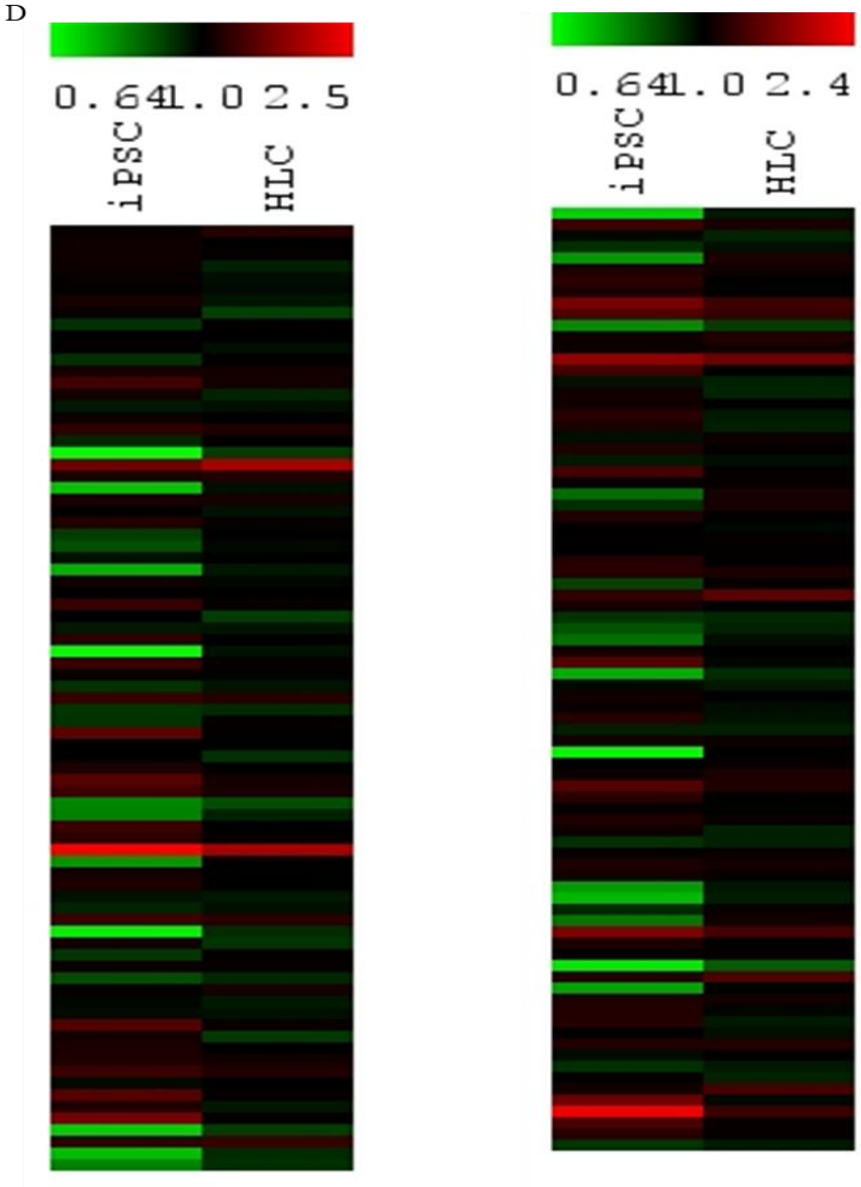


Figure 5.2- iPSC and HLC gene expression (P<0.05). A- Quantification of the number of genes found differentially regulated in iPSC and HLC, as determined by Log_2 expression arrays in comparison to primary hepatocytes. Heat maps of expression arrays identified oncogenes (B), protooncogenes (C) and tumour suppressor genes (D) gene expression. Key bars above each heat map indicate levels of expression

To further reveal similarities and differences between the two cell stages, similarly expressed genes were compared between iPSC and HLC. 78% (7) upregulated oncogenes are shared between iPSC and HLC while 22% (2) of genes are uniquely upregulated at the iPSC stage. Conversely, when examining downregulated oncogenes, 50% (2) genes are shared between iPSC and HLC, with a further 50% (2) genes uniquely downregulated in HLC. Comparison of upregulated tumour suppressor genes found 60% (74) genes are shared, with 25% (31) and 15% (19) genes uniquely expressed in iPSC and HLC stages, respectively. Comparison of downregulated tumour suppressor genes identifies 45% (41) of genes shared with 21% (19) and 34% (31) genes only expressed in the iPSC and HLC, respectively. 61% (109) upregulated protooncogenes are shared between iPSC and HLC, with 27% (49) and 12% (21) of genes only expressed at the iPSC and HLC stages. Upon examination of downregulated protooncogenes, 41% (46) are shared with 17% (19) and 42% (47) genes uniquely downregulated in iPSC and HLC, respectively. This suggests that while there are cancer related genes that share expression values between the two cell types, differences persist in the quantity and levels of expression between naïve and terminally differentiated cell stages indicating the need to examining HIV-1 LV insertion site preferences at both stages. LV containing either a native or SIN LTR configuration may show subtle differences in insertion profiles and associated oncogenic effects between iPSC and HLC.

Gene enrichment analysis using FunRich software (v3.1.3) (Pathan, Keerthikumar et al. 2015, Pathan, Keerthikumar et al. 2017) identified overrepresented significant biological pathways in gene data sets. This identifies pathways enriched at both the iPSC and HLC stage (Table 5.1). All differentially regulated genes were inputted to identify significant biological pathways enriched. Comparison of the top 10 significant pathways analysed ($P<0.05$) between cell types identifies only one common pathway, eukaryotic

translation elongation, indicating further differences between the two cell stages.

	Upregulated	Downregulated
iPSC	Gene Expression	Immune System
	Processing of Capped Intron-Containing Pre-mRNA	Metabolism of lipids and lipoproteins
	Influenza Infection	Biological oxidations
	Nonsense-Mediated Decay	Phase 1 - Functionalization of compounds
	Nonsense Mediated Decay Enhanced by the Exon Junction Complex	Interferon gamma signaling
	Influenza Life Cycle	Cytokine Signaling in Immune system
	mRNA Processing	Interferon Signaling
	Formation and Maturation of mRNA Transcript	Endogenous TLR signaling
	Nonsense Mediated Decay Independent of the Exon Junction Complex	Lipoprotein metabolism
	Eukaryotic Translation Elongation	Formation of Fibrin Clot (Clotting Cascade)
HLC	Epithelial-to-mesenchymal transition	Formation and Maturation of mRNA Transcript
	Peptide chain elongation	HIV Infection
	Eukaryotic Translation Elongation	mRNA Processing
	Signaling by GPCR	Gene Expression
	Viral mRNA Translation	Processing of Capped Intron-Containing Pre-mRNA
	Eukaryotic Translation Termination	Metabolism
	Formation of a pool of free 40S subunits	Host Interactions of HIV factors
	Developmental Biology	mRNA Splicing - Major Pathway
	Regulation of beta-cell development	mRNA Splicing
	Regulation of gene expression in beta cells	Regulation of mRNA Stability by Proteins that Bind AU-rich Elements

Table 5.1- Gene enrichment of differentially regulated genes found in iPSC and HLC derivatives. Enrichment for biological pathways using FunRich identifies the top 10 most significant pathways identified in differentially regulated genes sets between cell types, with comparative analysis revealing further differences ($P<0.05$)

Further comparative analysis of the top 10 significant biological pathways enriched in differentially related cancer genes revealed further subtle differences (Table 5.2-3). In iPSC, 100% of biological pathways enriched for upregulated or downregulated oncogenes are in common with similarly expressed genes in HLC. 50% of biological pathways for protooncogenes found to be upregulated or downregulated are in common with equivalent analysis in HLC. Examination of the top 10 biological pathways ($P<0.05$) enriched in tumour suppressor genes found 70% of upregulated and 40% of downregulated biological pathways common between iPSC and HLC. This reveals that while overall pathways expressed are unique to each stage, common pathways are expressed for cancer related genes expressed there are common pathways targeted. However, integration of the LV genome may target common pathways and the global differences demonstrate the need to assess LV genotoxicity at both the pluripotent and terminally differentiated cell stages.

iPSC	Oncogene	Proto oncogene	Tumour suppressor gene
Upregulated	PLK1 signaling events	CDC42 signaling events	ATM pathway
	Polo-like kinase signaling events in the cell cycle	Regulation of CDC42 activity	ATR signaling pathway
	Validated transcriptional targets of deltaNp63 isoforms	Integrin-linked kinase signaling	p53 pathway
		AP-1 transcription factor network	LKB1 signaling events
		Glypican pathway	ErbB receptor signaling network
		Plasma membrane estrogen receptor signaling	Glypican pathway
		Endothelins	Proteoglycan syndecan-mediated signaling events
		TRAIL signaling pathway	Arf6 downstream pathway
		S1P1 pathway	mTOR signaling pathway
		ErbB1 downstream signaling	Insulin Pathway
Downregulated	Mesenchymal-to-epithelial transition	PAR1-mediated thrombin signaling events	Cyclin D associated events in G1
		Thrombin/protease-activated receptor (PAR) pathway	G1 Phase
		Signaling events mediated by PTP1B	Interferon Signaling
		Insulin Pathway	p53 pathway
		Arf6 signaling events	CXCR4-mediated signaling events
		Arf6 downstream pathway	Netrin mediated repulsion signals
		Class I PI3K signaling events mediated by Akt	TCR signaling in naïve CD4+ T cells
		PDGFR-beta signaling pathway	Direct p53 effectors
		S1P1 pathway	LKB1 signaling events
		Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling	ErbB receptor signaling network

Table 5.2- Gene enrichment of differentially regulated cancer genes found in iPSC. Enrichment for biological pathways using FunRich identifies the top 10 most significant pathways identified in differentially regulated genes sets between cell types ($P<0.05$) with comparative analysis revealing further similarities and differences

HLC	Oncogene	Proto oncogene	Tumour suppressor gene
Upregulated	Validated transcriptional targets of deltaNp63 isoforms	Glypican pathway	ATM pathway
	PLK1 signaling events	Plasma membrane estrogen receptor signaling	p53 pathway
	p63 transcription factor network	Endothelins	ATR signaling pathway
	Polo-like kinase signaling events in the cell cycle	Arf6 downstream pathway	E2F transcription factor network
		mTOR signaling pathway	LKB1 signaling events
		Arf6 trafficking events	ErbB receptor signaling network
		ErbB1 downstream signaling	Glypican pathway
		Signaling events mediated by focal adhesion kinase	Proteoglycan syndecan-mediated signaling events
		S1P1 pathway	Arf6 signaling events
		Arf6 signaling events	Internalization of ErbB1
Downregulated	Mesenchymal-to-epithelial transition	CDC42 signaling events	Direct p53 effectors
	FOXA1 transcription factor network	Regulation of CDC42 activity	ATM pathway
	Activation of Chaperones by IRE1alpha	PAR1-mediated thrombin signaling events	ATR signaling pathway
	Unfolded Protein Response	Thrombin/protease-activated receptor (PAR) pathway	LKB1 signaling events
	Validated nuclear estrogen receptor alpha network	TRAIL signaling pathway	ErbB receptor signaling network
	FOXA transcription factor networks	EGF receptor (ErbB1) signaling pathway	p53 pathway
		Internalization of ErbB1	Class I PI3K signaling events
		Class I PI3K signaling events mediated by Akt	EGF receptor (ErbB1) signaling pathway
		Arf6 signaling events	PDGFR-beta signaling pathway
		PDGFR-beta signaling pathway	Insulin Pathway

Table 5.3- Gene enrichment of differentially regulated cancer genes found in iPSC. Enrichment for biological pathways using FunRich identifies the top 10 most significant pathways identified in differentially regulated genes sets between cell types ($P<0.05$) with comparative analysis revealing further similarities and differences

Hepatocellular carcinoma (HCC) is the fourth most common cancer worldwide (Singal, Lampertico et al. 2020). This disease commonly originates from liver cell pathology and therefore, the expression of genes associated with HCC were analysed in HLC (Figure 5.3). 228 genes are known to be involved in HCC, of which 85% (193) can be identified in the gene expression data set used. Examination of the expression array of these genes finds 50.3% (97) are upregulated and 49.7% (96) are downregulated. This demonstrates the gene expression data set is appropriate for analysis of LV mediated HCC pathologies.

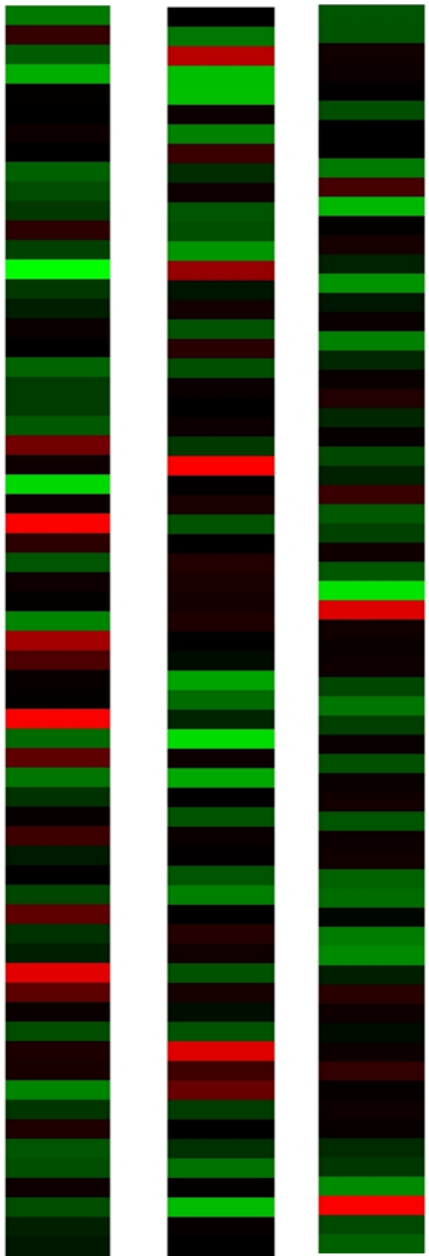


Figure 5.3- HLC expression levels of genes associated with HCC. Gene expression array derived from microarray of iPSC and HLC gene expression. 50.26% of these genes are found upregulated, with 49.74% downregulated. Gradient colour scheme from green to red between 0.1 to 9.6

5.2.2. iPSC & HLC LV Insertion Site Analysis

HIV-1 based LV carrying either a native (pHV) or SIN (pHR) configuration LTR were used to transduce P106 iPSCs and terminally differentiated HLCs at a MOI of 20. HLC differentiation of cells were performed by colleagues as previously described (Lucendo-Villarin, Rashidi et al. 2019, Wang, Y., Alhaque et al. 2017, Rashidi, Luu et al. 2018). Cells were harvested for gDNA at 72 hours post transduction as described in Section 2.2.1.12. P106 iPSC transduced with pHR or pHV were also grown out for 30 days post transduction before gDNA harvest. Robust gene expression was detected before harvest of gDNA 72 hours post transduction in three biological repeats. Multiple ISs were identified in all samples by EPTS/LM-PCR (Genewerk, Heidelberg). 118,701 insertion site loci in P106 iPSC transduced using pHR harvested 72 hours post transduction. 60,157 loci were identified in these cells harvested 30 days post transduction. 121,160 loci and 55,753 loci were identified in P106 iPSC transduced using pHV LV harvested 72 hours and 30 days post transduction, respectively. In HLC samples transduced with pHR and pHV LV, 22,874 and 34,142 loci were identified, respectively. Between biological triplicates, common ISs were found (Figure 5.4). In P106 iPSC transduced with pHR, 12,542 common insertion sites were identified after harvesting cells 72 hours post transduction. 7,002 common ISs remained 30 days post transduction. Similar analysis in pHV transduced P106 iPSC identified 13,636 and 9,724 common insertion sites after day 3 and day 30 harvesting timepoints, respectively. In HLC samples, 8,182 and 9,205 common ISs were identified after transduction with pHR or pHV LV, respectively. The reduced quantity of ISs in HLC is expected.

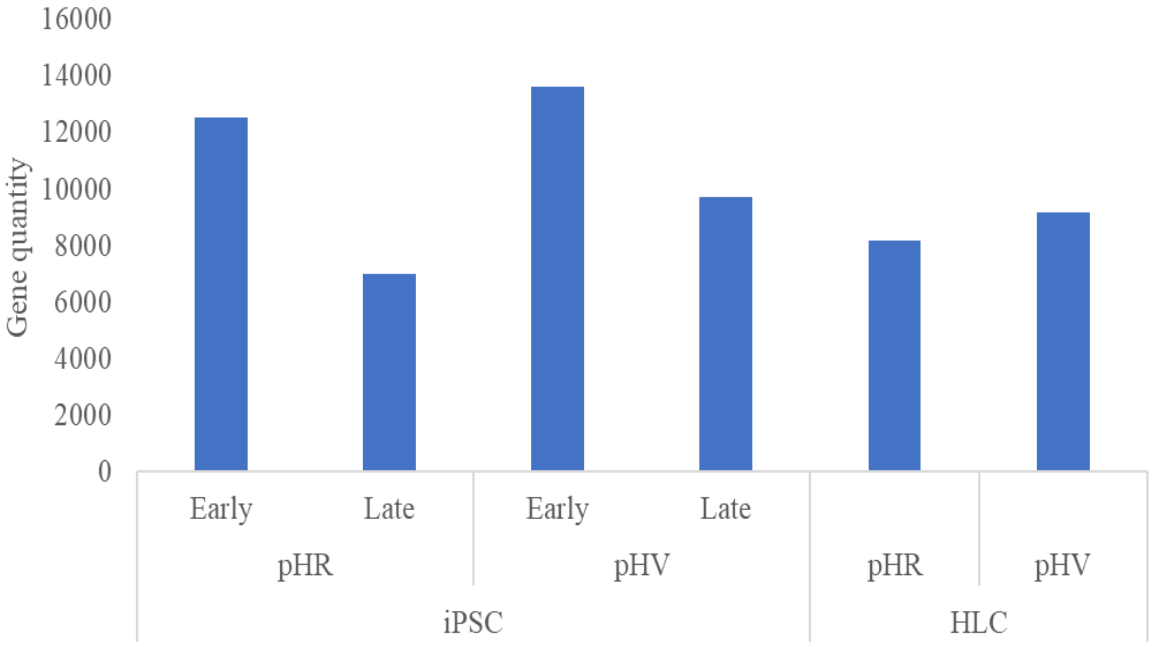


Figure 5.4- Common ISs identified. iPSC and HLC samples transduced with pHR or pHV were sequenced to identify ISs. Common IsS were identified between three biological repeats of each sample. Fewer ISs were identified in HLC in comparison to iPSC (day 3 timepoints) with a reduction in clonality observed in iPSC samples over time

Comparison between biological repeats found consistent numbers of common sites of insertion (Figure 5.5). The number of common ISs in iPSC pHR samples at day 3 harvest was 9,152±226, with this reducing to 5,964±132 at day 30 harvests. 9,670±279 and 5,819±95 unique ISs were found on average in iPSC samples transduced with pHV LV at day 3 and day 30 harvests respectively. Similar analysis in HLC samples found 4,478±234 and 5,452±346 unique ISs in pHR and pHV samples, harvested at day 3 post transduction, respectively. This indicates that while differences persist in the total quantity of ISs between cell types, the average number of unique ISs across biological repeats remains consistent. Interestingly, the percentage of common IS between biological repeats remains consistent between cell types, independent of vector configuration.

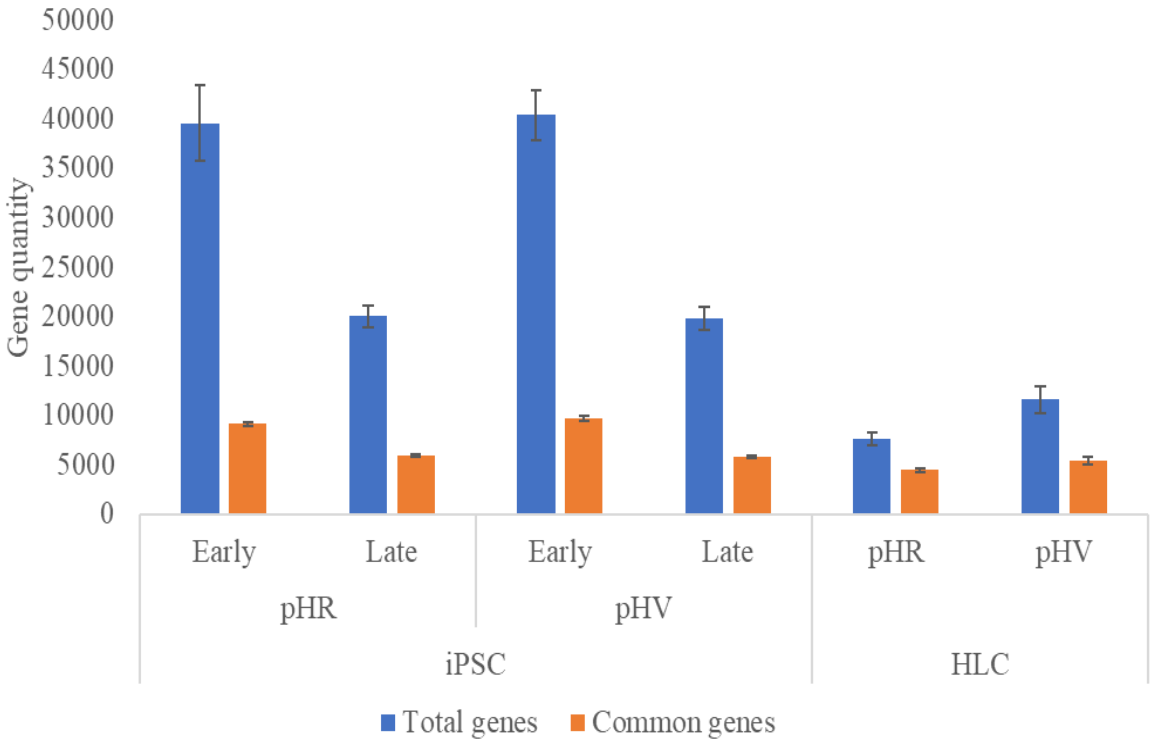


Figure 5.5- Average quantity of ISs identified between iPSC and HLC samples. The average quantity of total gene ISs are shown in comparison to the quantity of common IS identified in each sample. These remain consistent between samples (n=3)

P106 iPSC samples transduced with pHR or pHV LV were grown out to 30 days to examine clonal outgrowth in genes over time, with cells split every 72 hours, as performed by colleagues. 52.1% (6,695) of sites are common between pHR samples, with 45.5% (5,847) and 2.4% (307) sites uniquely identified in the day 3 and day 30 timepoints. Similar analysis of pHV samples identified 59.5% (8,712) sites are common between day 3 and day 30 sampling timepoints, with 33.6% (4,924) and 6.9% (1,012) only identified in day 3 or day 30 timepoints, respectively. This suggests that clonal outgrowth can be identified to an extent over a 30-day time period.

Comparison of insertion site selection between vector configurations indicate differences in insertion site preference. 69.4% (10,722) of IS are common between iPSC pHR and pHV samples harvested at day 3, with 11.8% (1,820)

and 18.9% (2,914) unique to each vector, respectively. Upon examination of iPSC grown out for 30 days post transduction, 72.0% (7,002) of sites are common between vector configurations, with 28% (2,722) only identified in pHV samples. Similar analysis in HLC samples showed 52.0% (5,945) are shared between pHR and pHV LV insertion sites, with 19.5% (2,237) and 28.5% (3,261) unique to each vector, respectively. 72% (7,002) of sites are common. This indicates subtle differences in IS preferences between the native and SIN LTR configurations.

Further comparative analysis shows differences in insertion sites between cell types. Comparison of pHR samples in iPSC and HLC, harvested at similar timepoints indicate that 50.2% (6,924) of insertion sites are common, with 40.7% (5618) and 9.1% (1,258) unique to iPSC and HLC samples, respectively. Analysis after pHV transduction reveal 54.3% (8,041) insertion sites are common between cell types, with37.8% (5,595) and 7.9% (1,165) unique to iPSC and HLC samples. Due to the differences in IS detected, it is important to understand the effects of LV mediated integration in both cell stages.

5.2.3. Gene Expression Alignment of IS in iPSC and HLC Samples

To identify the expression levels of genes targeted by pHR and pHV LV, genes identified as IS were aligned with a published microarray dataset (GSE 61287). This found the majority of IS genes (>76%) to be identified in a gene expression array. Assessment of the expression levels of these genes identified that 57-76% of IS genes are upregulated in the equivalent cell stages (Figure 5.6). This validates previous work that indicates LV prefers to integrate within highly expressed genes (Ciuffi, Llano et al. 2005, Schroder, Shinn et al. 2002, Mitchell, Beitzel et al. 2004, Rossetti, Cavarelli et al. 2013). Negligible differences are observed in IS over time or between vector configurations in

similar cell types suggesting that the viral preference for integration in highly expressed genes is independent of U3 region. However, the proportion of upregulated genes targeted for insertion in HLC is lower than identified in iPSC, suggesting the differences in gene expression levels between cell stages effects IS choice.

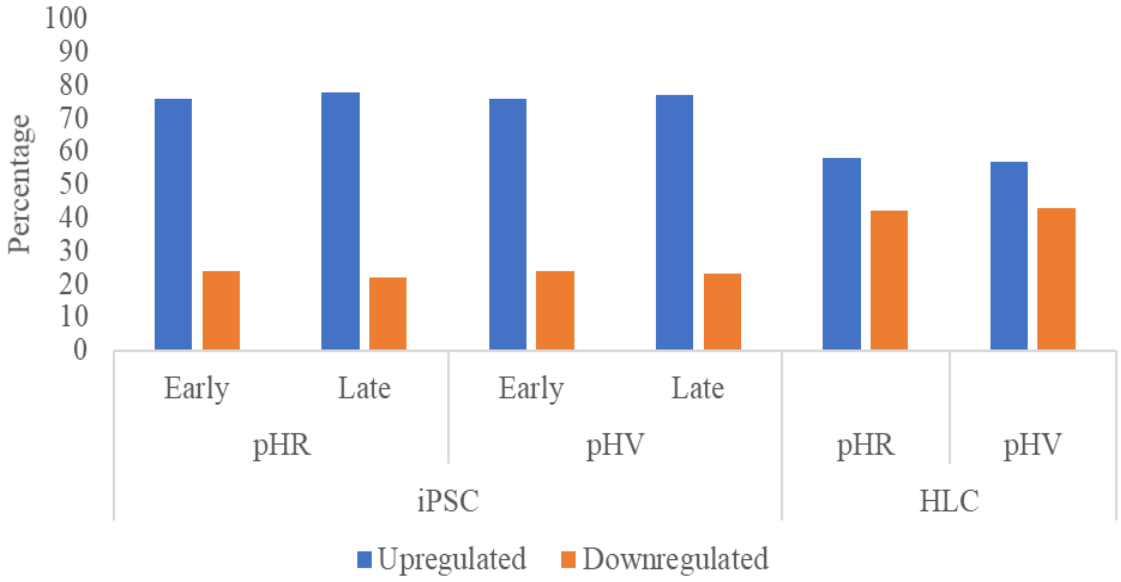


Figure 5.6- IS alignment with gene expression. ISs identified in iPSC and HLC samples after transduction with pHR or pHV LV were identified as upregulated or downregulated using an expression array generated using a microarray gene expression data se

5.2.4. Gene Ontology Analysis of IS Identified in iPSC and HLC Samples

Gene ontology analysis of IS in iPSC and HLC samples identify targeted genes known to be involved in cancer (Table 5.4). A greater number of protooncogenes and tumour suppressor genes are identified in infected iPSC, with 35 and 29 extra genes identified targeted in pHR LV samples. 40 and 25 greater protooncogenes and tumour suppressor genes are identified targeted in iPSC pHV samples, respectively. While the same number of oncogenes are identified as insertion site targets in pHR samples, one further oncogene is identified in iPSC pHV samples. This suggests both iPSC and HLC are valuable

targets to understanding insertion site mechanisms of LV integration, of both LTR configurations.

		PO		TSG		On	
		Early	Late	Early	Late	Early	Late
iPSC	pHR	154	108	122	80	6	6
	pHV	169	141	126	105	9	6
HLC	pHR	119		93		6	
	pHV	129		101		8	

Table 5.4- Gene ontology analysis of IS. Analysis of ISs in iPSC and HLC samples transduced with pHR or pHV LV identified protoonocogenes (PO), tumour suppressor genes (TSG) and oncogenes (On) targeted by each vector. Further comparisons reveal subtle differences between IS targeting based on vector configuration

Comparison of cancer related genes identified between cell stages indicate subtle differences in vector targeting important to understanding genotoxic effect of LV. Comparison of all cancer gene IS in pHR samples reveals 66.7% (198) are shared, with 27.3% (81) and 6.1% (18) identified only in iPSC and HLC (day 3) respectively. Comparison of oncogene IS identified 50% (4) are shared, with 25% (2) oncogenes found only in iPSC and HLC samples. Similar analysis using protooncogenes identified 67.9% (110) in common, with 26.5% (43) and 5.6% (9) identified uniquely in iPSC and HLC samples, respectively. Comparison of tumour suppressor gene IS finds 66.4% (85) are shared between cell types, with 28.1% (36) and 5.5% (7) identified only in iPSC and HLC samples respectively. Similar comparison of pHV samples for all cancer gene IS finds 68.6% (216) IS are shared, with 26.0% (82) and 5.4% (17) found only in iPSC or HLC samples, respectively. Comparison of cancer gene subtypes finds 70% (7) of oncogenes are common and 20% (2) and 10% (1) oncogene

identified in iPSC and HLC samples, respectively. Examination of protooncogenes finds 69.7% (122) gene IS shared, with 46 (26.3%) and 40% (7) uniquely identified in iPSC or HLC samples respectively. Comparative analysis of tumour suppressor genes finds 67.9% (91) are common between cell types, with 25.4% (34) and 6.7% (9) only identified in iPSC or HLC samples. This suggests that subtle differences persist in transduction of iPSC and HLC suggesting the need for both platforms.

Comparison of cancer gene IS between vector configurations reveal differences between the native and SIN LTR IS preferences. Comparison of all cancer gene IS in iPSC (day 3) identified 77.7% (251) are common between pHR and pHV samples, with 7.7% (25) and 14.6% (47) identified only in respective samples. After growing these cells out for 30 days, 77.0% (191) remained in common, with 23.0% (57) identified only in pHV samples. We next looked individually at each class of cancer genes. Comparison of protooncogene IS identified in iPSC (day 3) samples found 78.3% (141) common between pHR and pHV, with 6.7% (12) and 15.0% (27) identified uniquely in data sets. At day 30, 76.4% (107) protooncogene IS were shared, with 23.6% (33) found only in pHV samples. Similarly, 66.7% (6) oncogenes were identified in iPSC samples (day 3) both after pHR or pHV transduction, with 33.3% (3) only identified in pHV samples. 100% (6) oncogenes were common at day 30 samples. For tumour suppressor gene IS, 78.3% (108) were common between pHR and pHV iPSC (day 3) samples, with 9.4% (13) and 12.3% (17) found only in respective samples. After growing the cells out for 30 days, 76.9% (80) protooncogene IS were common between vector configurations, with 23.1% (24) found only in pHV samples.

A similar level of shared cancer gene IS are identified in HLC samples. Between all cancer gene IS, 67.4% (182) are common between pHR and pHV samples, with 12.6% (34) and 20.0% (54) identified only in respective samples. Upon further examination of each individual class of cancer genes, 75.0% (6)

oncogene IS are common, with 25.0% (2) found only in pHV samples. 72.2% (104) protooncogene IS were common between vectors, with 10.4% (15) and 17.4% (25) found only in pHR and pHV samples respectively. Upon comparison of tumour suppressor gene IS, 61.3% (73) are common between vectors, with 16.0% (19) and 22.7% (27) found only in pHR and pHV samples. This reveals subtle differences between vector integrations in iPSC and HLC between the vector configurations.

Transduced P106 iPSC samples were grown out for 30 days to examine clonal outgrowth. Comparison of IS in cancer related genes identifies clonality of samples. For all cancer gene IS identified between pHR samples at day 3 and Day 30, 64% (183) were in common, with 33.2% (95) and 2.8% (8) identified only in day 3 or day 30 samples, respectively. Examination of each class of cancer genes revealed further similarities. 71.4% (5) oncogenes are common in pHR samples, with 14.3% (1) only identified in both day 3 and day 30 samples. 66.7% (104) protooncogene IS are common between the early and late stage, with 31.4% (49) and 1.9% (3) found only in day 3 and day 30 samples, respectively. Similar analysis of tumour suppressor gene IS found 60.8% (76) shared between day 3 and day 30 samples, with 36.0% (45) and 3.2% (4) only identified in the respective samples.

A similar result was identified in P106 iPSC samples transduced with pHV. 69.4% (222) of all cancer gene IS identified are common between day 3 and day 30 samples, with 23.4% (75) and 7.2% (23) identified uniquely at each stage, respectively. Similarly, comparison of oncogene IS identified 66.7% (6) shared between the early and late stage, with 33.3% (3) identified only at day 3. Comparison of protooncogenes found 74.0% (131) shared between day 3 and day 30 samples, with 20.9% (37) and 5.1% (9) identified only at each stage, respectively. 64.7% (90) tumour suppressor genes are common between the day 3 and day 30 timepoints, with 25.2% (35) and 10.1% (14) identified only in

each stage, respectively. The incidence of unique IS identified at day 30 can be explained by the sensitivity of sequencing protocol. These results suggest clonal outgrowth over this time period.

5.2.5. Gene Enrichment and Expression Analysis of IS Identified in iPSC and HLC

Gene enrichment of IS identified biological pathways significantly associated with sites of LV insertion ($P<0.05$). Comparison of the top 10 significant biological pathways reveal further similarities between vector IS choice (Table 5.4). Comparison between vector configurations indicates 80% (iPSC, day 3), 70% (iPSC, day 30) and 80% (HLC) are shared between pHR and pHV samples. Comparison of enriched pathways over time in iPSC identified 60% (pHR) and 70% (pHV) are similar. Both results suggest that despite vector configurations and clonal outgrowth, common biological pathways are targeted by the LV vector and persist over time. Comparison between cell types (day 3) found 60% (pHR) and 80% (pHV) of biological pathways in common suggesting the vector targets similar pathways regardless of differences in IS.

	iPSC		HLC
	Early	Late	
pHR	ErbB receptor signaling network	PAR1-mediated thrombin signaling events	PAR1-mediated thrombin signaling events
	PAR1-mediated thrombin signaling events	Thrombin/protease-activated receptor (PAR) pathway	Thrombin/protease-activated receptor (PAR) pathway
	Thrombin/protease-activated receptor (PAR) pathway	Plasma membrane estrogen receptor signaling	LKB1 signaling events
	Plasma membrane estrogen receptor signaling	Endothelins	ErbB receptor signaling network
	CDC42 signaling events	LKB1 signaling events	mTOR signaling pathway
	Endothelins	CDC42 signaling events	Internalization of ErbB1
	LKB1 signaling events	ErbB receptor signaling network	Arf6 signaling events
	Arf6 downstream pathway	Regulation of CDC42 activity	Class I PI3K signaling events mediated by Akt
	S1P1 pathway	TRAIL signaling pathway	Class I PI3K signaling events
	Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling	PDGFR-beta signaling pathway	Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling
pHV	LKB1 signaling events	PAR1-mediated thrombin signaling events	ErbB receptor signaling network
	ErbB receptor signaling network	Thrombin/protease-activated receptor (PAR) pathway	PAR1-mediated thrombin signaling events
	CDC42 signaling events	Plasma membrane estrogen receptor signaling	Thrombin/protease-activated receptor (PAR) pathway
	PAR1-mediated thrombin signaling events	Endothelins	Plasma membrane estrogen receptor signaling
	Thrombin/protease-activated receptor (PAR) pathway	LKB1 signaling events	Endothelins
	Plasma membrane estrogen receptor signaling	ErbB receptor signaling network	LKB1 signaling events
	Endothelins	Class I PI3K signaling events mediated by Akt	CDC42 signaling events
	Regulation of CDC42 activity	EGF receptor (ErbB1) signaling pathway	EGF receptor (ErbB1) signaling pathway
	S1P1 pathway	Signaling events mediated by focal adhesion kinase	PDGFR-beta signaling pathway
	Signaling events mediated by focal adhesion kinase	Arf6 signaling events	Signaling events mediated by focal adhesion kinase

Table 5.4- Biological pathways associated with IS gene enrichment. Total IS gene for iPSC and HLC samples transduced with pHR or pHV LV were enriched for biological pathways represented. Top 10 significant pathways shown ($P<0.05$)

A similar result is seen after enrichment of cancer gene IS for biological pathways (Table 5.5). When comparing biological pathways for cancer gene IS after pHR transduction between the early and late timepoints in iPSC found 70-75% of biological pathways are common (Day 3). At day 30, between 70-100% of biological pathways are common with the early timepoint. Enrichment of cancer related gene IS in pHV iPSC samples (day 3) found 70-100% of pathways are common with the later timepoint. Similar analysis of enriched cancer gene IS found a high level of similarity between pathways between vector configurations, over time and between cell types. These results suggest that despite differences in IS in cancer related genes, common pathways are targeted and persist despite vector configuration and clonal outgrowth.

A	iPSC			
	Early		Late	
	pHR	pHV	pHR	pHV
TSG	ATM pathway ATR signaling pathway p53 pathway ErbB receptor signaling network LKB1 signaling events Glypican pathway Proteoglycan syndecan-mediated signaling events Signaling events mediated by focal adhesion kinase Insulin Pathway EGF receptor (ErbB1) signaling pathway	ATM pathway ATR signaling pathway p53 pathway Direct p53 effectors LKB1 signaling events ErbB receptor signaling network Glypican pathway Proteoglycan syndecan-mediated signaling events Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling Arf6 signaling events	ATM pathway ATR signaling pathway p53 pathway Direct p53 effectors LKB1 signaling events ErbB receptor signaling network Proteoglycan syndecan-mediated signaling events p73 transcription factor network Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling Signaling events mediated by focal adhesion kinase	ATM pathway ATR signaling pathway p53 pathway LKB1 signaling events ErbB receptor signaling network Direct p53 effectors Proteoglycan syndecan-mediated signaling events mTOR signaling pathway S1P1 pathway Signaling events mediated by focal adhesion kinase
PO	CDC42 signaling events Regulation of CDC42 activity PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway Plasma membrane estrogen receptor signaling Endothelins TRAIL signaling pathway AP-1 transcription factor network Arf6 signaling events mTOR signaling pathway	CDC42 signaling events Regulation of CDC42 activity PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway Plasma membrane estrogen receptor signaling Endothelins TRAIL signaling pathway Internalization of ErbB1 ErbB1 downstream signaling Arf6 trafficking events	CDC42 signaling events Regulation of CDC42 activity PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway Plasma membrane estrogen receptor signaling Endothelins TRAIL signaling pathway Class I PI3K signaling events Insulin Pathway S1P1 pathway	CDC42 signaling events Regulation of CDC42 activity PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway Plasma membrane estrogen receptor signaling Endothelins TRAIL signaling pathway Arf6 downstream pathway Signaling events mediated by focal adhesion kinase Arf6 trafficking events
On	Validated transcriptional targets of deltaNp63 isoforms PLK1 signaling events p63 transcription factor network Polo-like kinase signaling events in the cell cycle Epithelial-to-mesenchymal transition RAC1 signaling pathway Regulation of RAC1 activity Regulation of RhoA activity RhoA signaling pathway CDC42 signaling events	PLK1 signaling events Polo-like kinase signaling events in the cell cycle Validated transcriptional targets of deltaNp63 isoforms p63 transcription factor network CDC42 signaling events Regulation of CDC42 activity Regulation of nuclear beta catenin signaling and target gene transcription Canonical Wnt signaling pathway Noncanonical Wnt signaling pathway Epithelial-to-mesenchymal transition	Validated transcriptional targets of deltaNp63 isoforms PLK1 signaling events p63 transcription factor network Polo-like kinase signaling events in the cell cycle Epithelial-to-mesenchymal transition RAC1 signaling pathway Regulation of RAC1 activity Regulation of RhoA activity RhoA signaling pathway CDC42 signaling events	Validated transcriptional targets of deltaNp63 isoforms PLK1 signaling events p63 transcription factor network Polo-like kinase signaling events in the cell cycle Epithelial-to-mesenchymal transition RAC1 signaling pathway Regulation of RAC1 activity Regulation of RhoA activity RhoA signaling pathway CDC42 signaling events

B	HLC	
	pHR	pHV
TSG	ATM pathway ATR signaling pathway Polo-like kinase signaling events in the cell cycle p53 pathway LKB1 signaling events ErbB receptor signaling network Direct p53 effectors PIP3 activates AKT signaling Glypican pathway Proteoglycan syndecan-mediated signaling events	ATM pathway ATR signaling pathway p53 pathway ErbB receptor signaling network Direct p53 effectors Polo-like kinase signaling events in the cell cycle LKB1 signaling events Glypican pathway Proteoglycan syndecan-mediated signaling events mTOR signaling pathway
PO	PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway TRAIL signaling pathway CDC42 signaling events Regulation of CDC42 activity Arf6 trafficking events EGF receptor (ErbB1) signaling pathway PDGFR-beta signaling pathway Internalization of ErbB1 Insulin Pathway	CDC42 signaling events Regulation of CDC42 activity PAR1-mediated thrombin signaling events Thrombin/protease-activated receptor (PAR) pathway Plasma membrane estrogen receptor signaling Endothelins TRAIL signaling pathway Arf6 signaling events Class I PI3K signaling events S1P1 pathway
On	PLK1 signaling events Polo-like kinase signaling events in the cell cycle RhoA signaling pathway Regulation of RhoA activity Regulation of RAC1 activity RAC1 signaling pathway Mesenchymal-to-epithelial transition CDC42 signaling events Regulation of CDC42 activity	PLK1 signaling events Polo-like kinase signaling events in the cell cycle Validated transcriptional targets of deltaNp63 isoforms p63 transcription factor network Epithelial-to-mesenchymal transition RAC1 signaling pathway Regulation of RhoA activity Regulation of RAC1 activity RhoA signaling pathway Mesenchymal-to-epithelial transition

Table 5.5- Biological pathways associated with cancer gene IS enrichment. Tumour suppressor genes (TSG), protooncogenes (PO) and oncogenes (On) identified as IS in iPSC (A) and HLC (B) samples transduced with a naïve (pHV) or SIN (pHR) configuration LTR. Top 10 significant pathways shown ($P<0.05$)

It has previously been reported that HIV-1 based LV integrate into upregulated genes (Ciuffi, Llano et al. 2005, Schroder, Shinn et al. 2002, Mitchell, Beitzel et al. 2004, Rossetti, Cavarelli et al. 2013). To identify the levels of expression in targeted genes, IS were aligned with a published gene expression data set (GSE 61287). Examination of the expression levels of IS genes found the majority of genes to be upregulated. Further examination of the expression levels of cancer related genes found similar results (Figure 5.8). In iPSC, 100% of oncogene targets, 73-83% of protooncogene IS and 62-69% of tumour suppressor genes IS are found to be upregulated across vector configuration and sampling timepoints. In HLC, 60% of oncogene and protooncogene IS and 57% of tumour suppressor gene targets are found to be upregulated in this cell type. This validates the majority of IS, in particular cancer genes, are found to be upregulated genes. Differences persist between iPSC and HLC targets, possibly due to the increased levels of gene expression in iPSC.

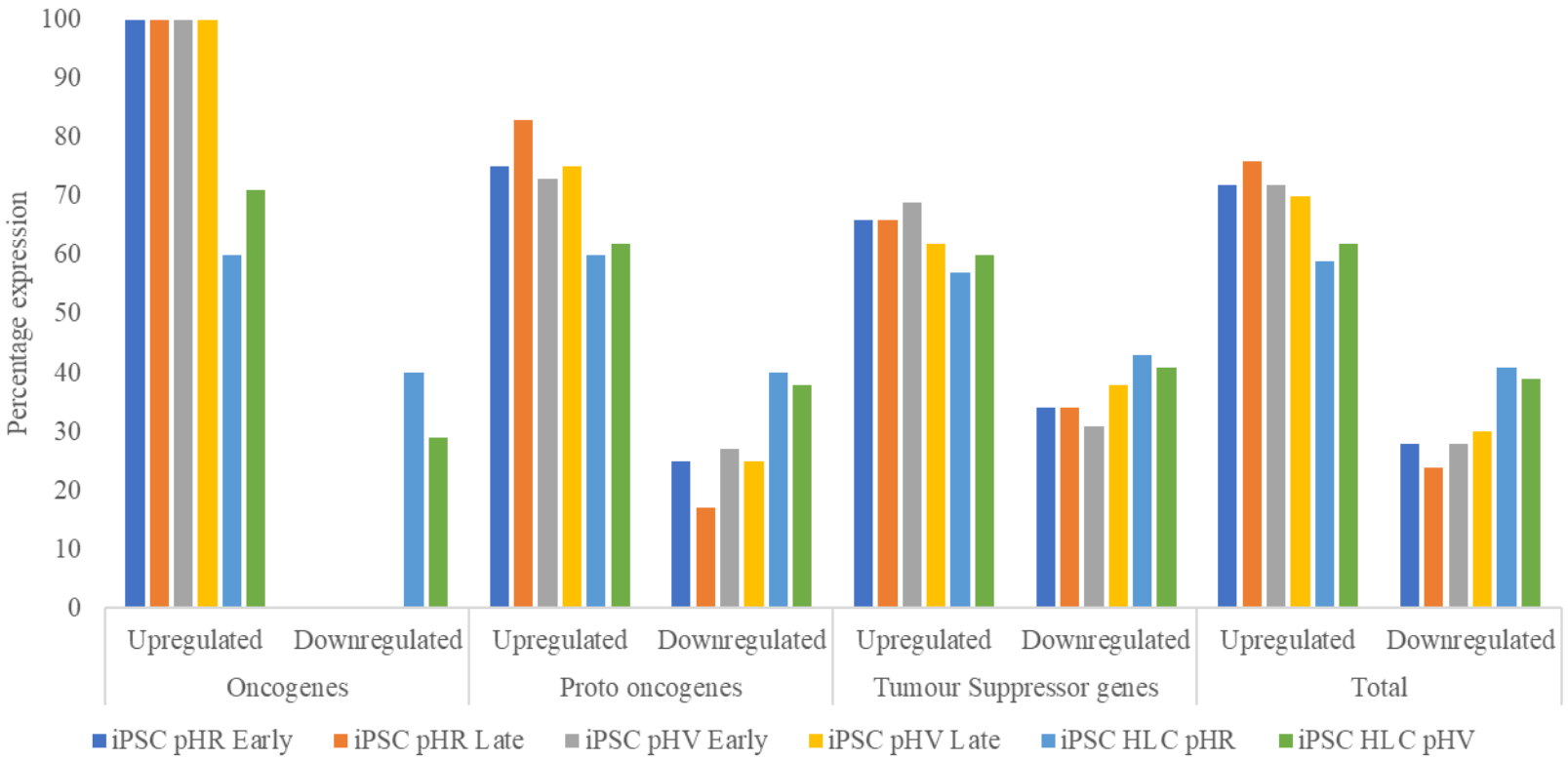


Figure 5.8- Cancer IS gene expression alignment. Percentage of IS identified in oncogenes, protooncogenes, tumour suppressor genes and all cancer gene IS aligned with a microarray gene expression data set to identify genes as upregulated or downregulated in comparison to primary hepatocytes

5.2.6. IS Alignment with Hepatocellular Carcinoma

IS were also aligned with genes involved in HCC (Figure 5.9). Of the 228 genes known to be associated with this disease, 54% and 31% were identified as IS in iPSC transduced with pHR LV at day 3 and 30 timepoints, respectively. 57% and 30% of genes were identified in pHV transduced iPSC samples over time. In HLC samples, 34% (pHR) and 41% (pHV) of genes known to be involved in HCC were identified. These targets represented 1% or less of total IS gene targets in samples. The majority of these genes are upregulated in comparison to the gene expression data set (GSE 61287). 72% and 74% of these genes are upregulated at day 3 and 30 timepoints in pHR LV transduced iPSC. In iPSC transduced with pHV LV, 71% and 76% of HCC associated gene IS are found as upregulated at both timepoints. In HLC, 55% and 57% of HCC gene targets are found upregulated after pHR and pHV transduction, respectively. This demonstrates the usefulness of this cellular model in understanding genotoxicity, with multiple targets towards cancer related genes, including specifically genes associated with HCC.

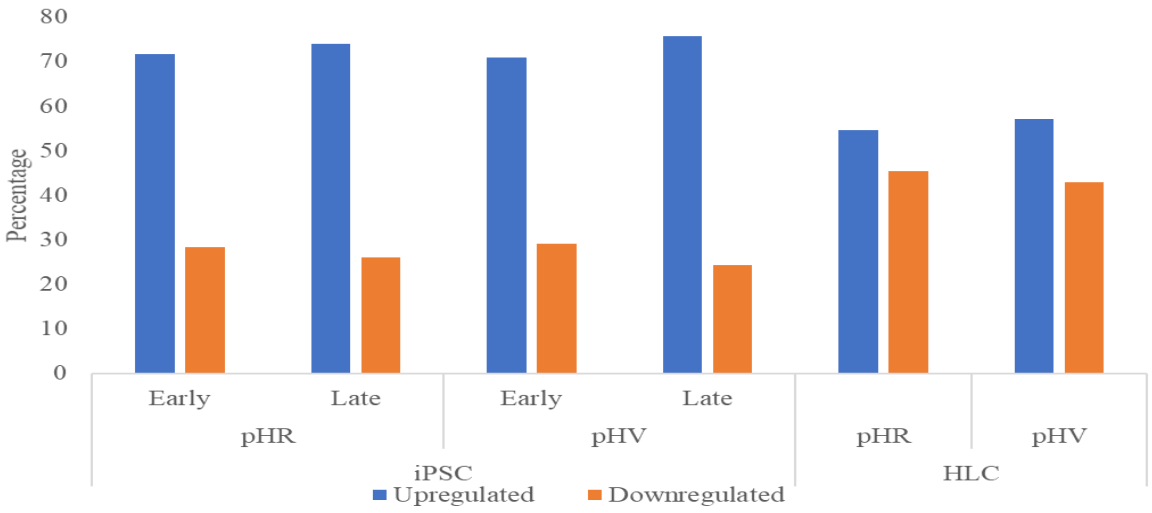


Figure 5.9- HCC associated IS gene expression alignment. Percentage of expression status of IS identified in genes associated with HCC shown in iPSC and HLC transduced with pHR or pHV LV. Sites of insertion were identified to be upregulated (blue) or downregulated (orange) from published microarray data set

Comparison of HCC gene targets in transduced samples reveals subtle differences between cell types and vector configurations. Comparison of IS over time in iPSC revealed 50.8% (65) genes are common between pHR timepoints, with 44.5% (57) and 4.7% (6) identified only at day 3 and 30 timepoints, respectively. In pHV iPSC samples over time, 56.6% (81) genes are shared, with 35.0% (50) and 8.4% (12) genes identified only in day 3 and 30 samples, respectively. Comparison of gene targets between vector configurations in iPSC identified 67.5% (102) (day 3) and 50.5% (47) (day 30) in common. Similar analysis in HLC found 52.7% (59) genes common between vectors, with 16.1% (8) and 31.3% only identified in pHR or pHV samples, respectively. Comparison between cell types identified further differences, with 48.5% (65) genes common between pHR samples, with 42.5% (57) and 9% (12) identified only in iPSC and HLC, respectively. In pHV samples, 57.3% (82) are common between iPSC and HLC, with 34.3% (49) and 84% (12) identified uniquely in each sample, respectively. This further validates the requirement for both cell types to be studies for genotoxicity to capture a wide array of gene targets.

5.2.7. Proliferation Assay in iPSCs

To further assess clonal outgrowth over time, transduced iPSC samples were analysed for an increase in sequence counts of the virus over time. These are identified as the number of viral transcripts detected at each integration site.

The average sequence count per gene integration was measured in each transduced sample (Figure 5.10). An average of 22.63±4.19 and 29.26±3.27 sequences were detected per gene in iPSC pHR samples at day 3 and day 30 sampling timepoints. Examination of iPSC pHV samples indicated a larger increase over time, with 27.47±1.47 and 44.27±2.01 sequences per gene, respectively. Both increases in pHR and pHV samples suggest clonal outgrowth

over time. The greater increase in pHV samples suggests the adverse effect of the native LTR in comparison to SIN. Similar analysis of HLC samples found a similar sequence count to iPSC at similar timepoints. 30.80±1.76 and 24.84±4.30 sequences per gene were detected in HLC samples after pHR and pHV transduction, respectively.

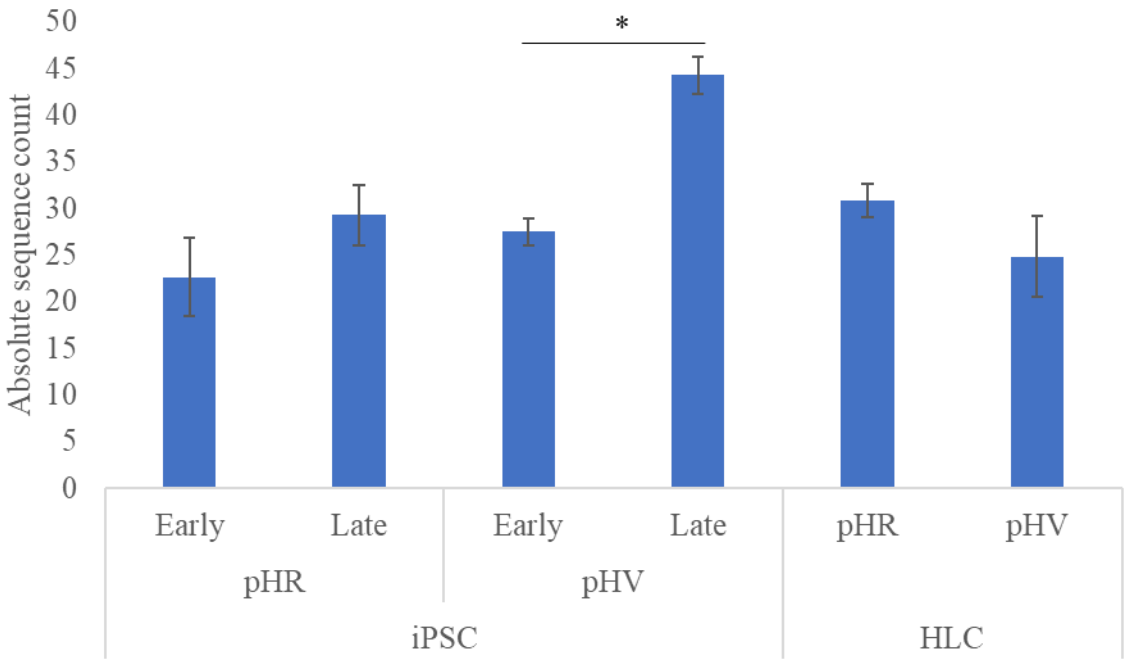


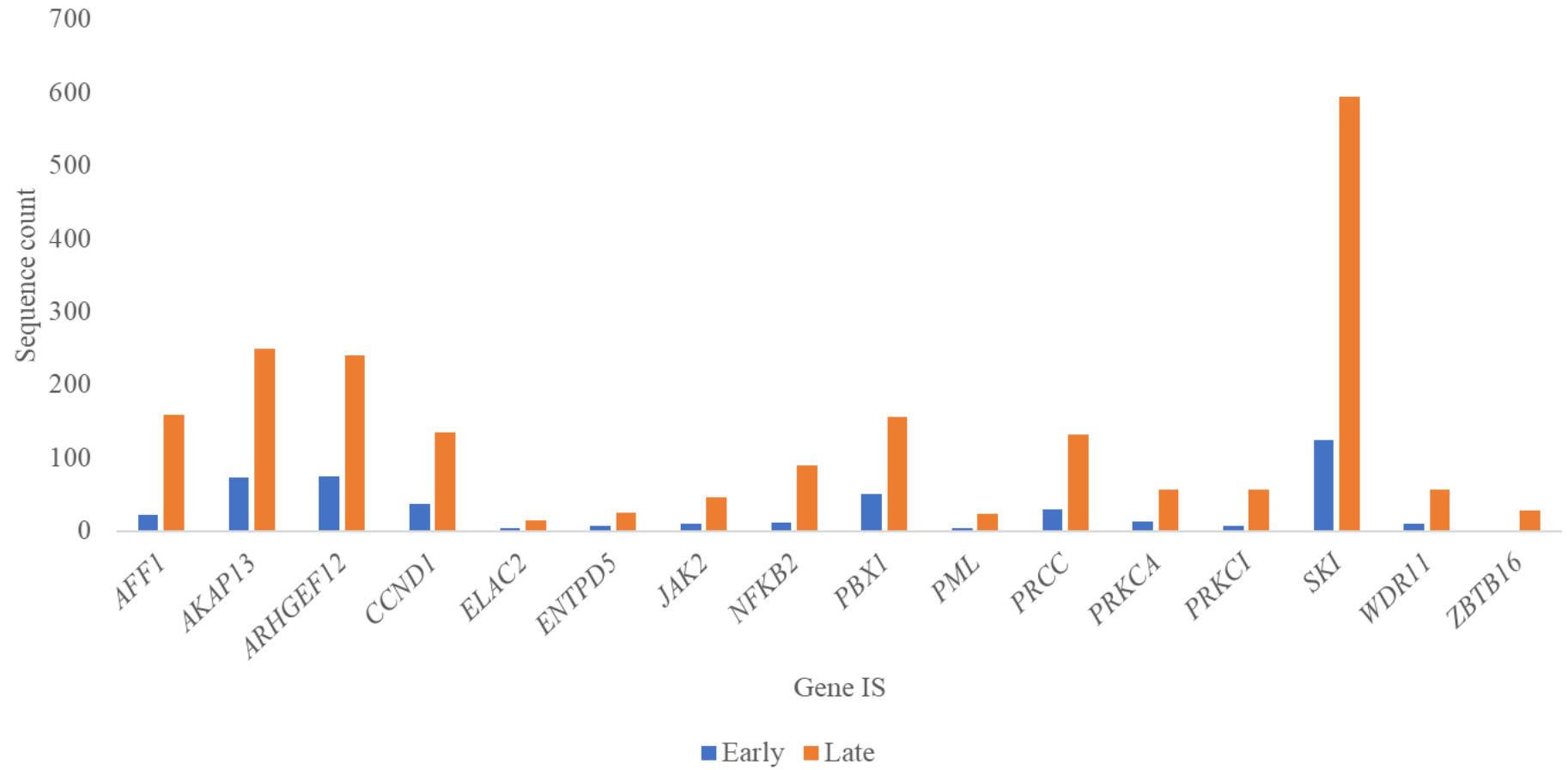
Figure 5.10- Sequence count per gene IS. Average absolute sequence counts detected per gene insertion detected in iPSC and HLC samples after pHR or pHV LV transduction. An increase in sequence counts is seen over time in iPSC IS found in samples, with a significant increase observed in pHV samples. Significance denoted by astirx ($P<0.05$).

To analyse clonal outgrowth over time, the sequence count in transduced iPSC samples were analysed. Overall, there was a reduction in clonality in both pHR and pHV samples. Total sequence count in pHR samples reduced by 16% over 30 days (208,930±42,321 (day 3) to 174,315±18,861 (day 30)). A reduction in clonality was also observed in pHV samples, though to a lesser extent. The total

sequence count reduced by 3% over time (264962±8,887 (day 3) to 257605±12,309 (day 30)). This further suggests clonal dominance of the native LTR in comparison to the SIN configuration. Average sequence counts in HLC samples were determined as 34% lower than in iPSC samples at the day 3 timepoint (pHR, 137,144±2,412) and 24% (pHV, 132,705±15,515). Comparison of sequence counts in IS shared between iPSC and HLC found that while both samples were treated identically, these differences may persist due to the proliferative nature of iPSCs.

Analysis of protooncogenes and oncogenes IS in iPSC over time revealed further differences (Figure 5.11). Protooncogenes and oncogenes are known activators of cellular replication and thus an increase in sequence count in these genes are expected if insertions are contributing towards clonal outgrowth (Hacein-Bey-Abina, Hauer et al. 2010, Hacein-Bey-Abina, Garrigue et al. 2008, Hacein-Bey-Abina, Von Kalle et al. 2003). 12% (13 out of 109) genes increased in sequence count over 30 days in pHR samples by equal to or greater than 2-fold. A similar number of protooncogenes and oncogenes (12%, 16 out of 137 genes) increased over time in pHV sample by 2-fold or greater. No genes were found in common between both pHR and pHV datasets. Furthermore, enrichment of these genes found no similar biological pathways associated.

A


700
600
500
400
300
200
100
0
Sequence count
AFF1
AKAP13
ARHGEF12
CCND1
ELAC2
ENTPD5
JAK2
NFKB2
PBX1
PML
PRCC
PRKCA
PRKCI
SKI
WDR11
ZBTB16
Gene IS
Early
Late

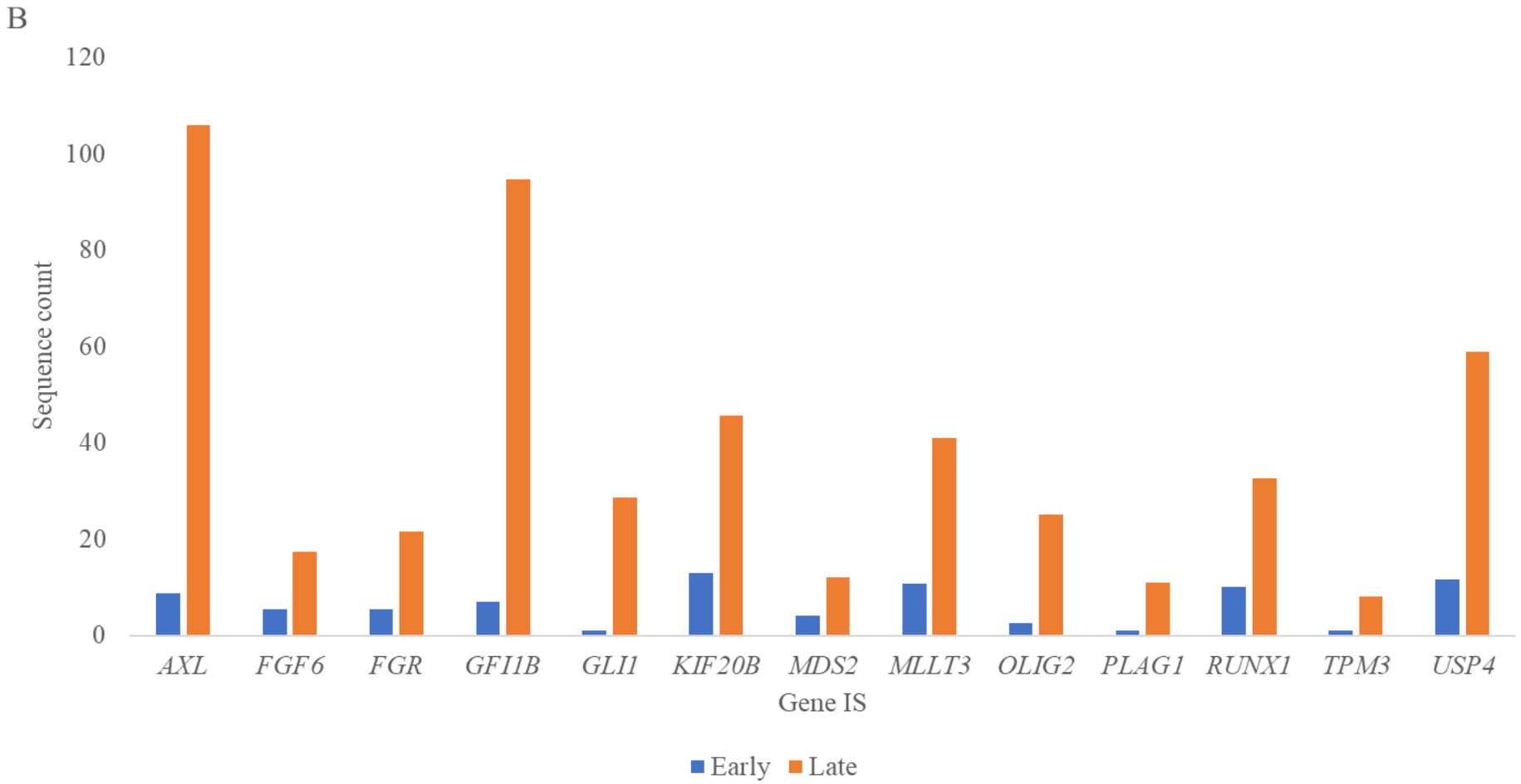
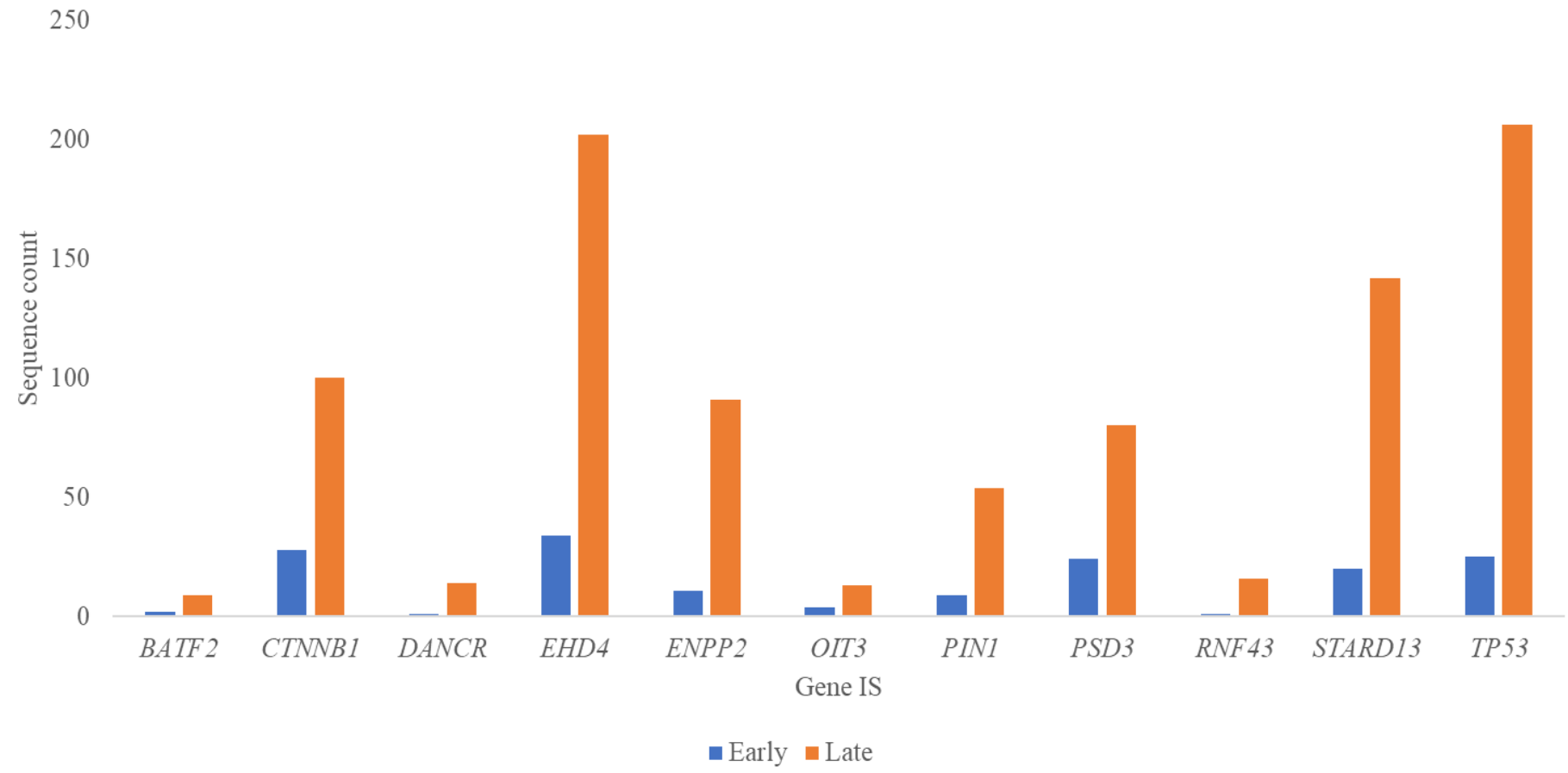


Figure 5.11- Proliferation assay gene identity. Gene IS identified as oncogenes or protooncogenes that increased by 2-fold or greater in sequence count over time in iPSC transduced with pHV (A) or pHR (B) LV between the early (day 3, blue) and late (orange, day 30) sample timepoints

These genes are seen to increase in sequence count similar to that of non-cancer genes. Analysis of all IS genes present in both the early (day 3) and late (day 30) iPSC samples transduced with pHV LV found 14% (1,257 out of 8,712) of IS genes to increase by 2-fold or greater in sequence count. Similar analysis of pHR samples found 16% (1074 out of 6695) of genes to increase in sequence counts by 2-fold or greater. This suggests the clonal dominance of oncogenes and protooncogenes over time are driving cellular proliferation.

Further analysis of genes known to be involved in HCC found that 58% (pHR) and 45% (pHV) increased in sequence counts over 30 days (Figure 5.12). Of these genes, 37% (14, pHR) and 44% (11, pHV) increased by 2-fold or greater in sequence count over time. These results suggest vector mediated clonal outgrowth.

A


250
200
150
100
50
0
Sequence count
BATF2
CTNNB1
DANCR
EHD4
ENPP2
OIT3
PIN1
PSD3
RNF43
STARD13
TP53
Gene IS
Early
Late

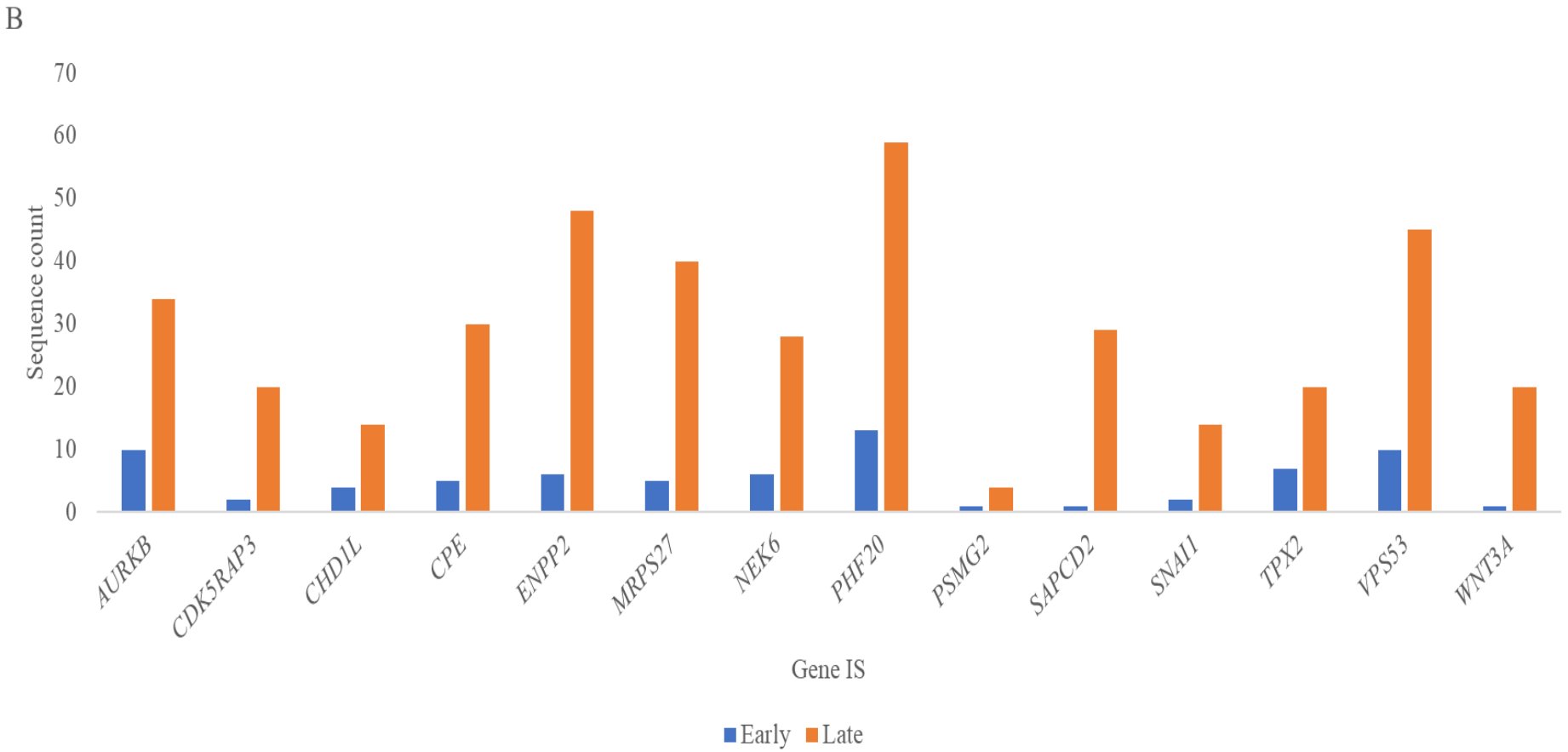


Figure 5.12- HCC proliferation assay gene identity. Gene IS identified associated with HCC that increased by 2-fold or greater in sequence count over time in iPSC transduced with pHV (A) or pHR (B) LV between the early (blue, day 3) and late (orange, day 30) sample harvest timepoints

5.2.8. Transcription Factor Binding Site Analysis of LV Tethering to Host Genome

TFBS have been reported to be involved in the integration of RV (Felice, Cattoglio et al. 2009). We proposed mining iPSC and HLC IS data after transduction using pHR or pHV LV, to assess differences between the SIN and native LTR configurations. To analyse similarities in HIV-1 LV IS preference in iPSC and HLC with native (pHV LV) or SIN (pHR LV) LTR configuration, we analysed pTFBS using a 20 bp window 3' and 5' of IS using UCSC BLAT genome browser (http://genome.ucsc.edu) alignment to hg38 build. This analysis found 118,701 IS loci in iPSC transduced using pHR harvested 72 hours post transduction. 60,157 loci were identified in these cells harvested 30 days post transduction. 121,160 loci and 55,753 loci were identified in iPSC transduced using pHV LV harvested 72 hours and 30 days post transduction, respectively. In HLC samples transduced with pHR and pHV LV, 22,874 and 34,142 loci were identified, respectively. oPOSSUM v3.0 Single Site Analysis (http://opossum.cisreg.ca/) identified pTFBS for 115 unique pTFBSs present in iPSC pHR transduction (Figure 5.13). After growing these cells out for 30 days, 112 sites remained. After pHV transduction of iPSC, 115 unique pTFBS were identified at both time points. pTFBS identification in HLC samples 3 days post transduction found 113 and 114 sites were identified after pHR and pHV transduction, respectively. Comparison of pTFBS identified between day 3 and day 30 sampling of iPSC pHR transduction found 112 sites (97%) were common. 115 (100%) sites were common between equivalent iPSC pHV samples (Figure 5.13).

To analyse similarities between the two LTR configurations, we compared pTFBS identified near insertion sites after iPSC and HLC transduction with pHR and pHV LV vectors (Figure 5.13). Comparison between sites found in iPSC harvested after 72 hrs post transduction indicated 114 sites (98%) are

shared between vector types. In HLC samples, 112 (97%) pTFBS were found to be in common. Furthermore, in iPSC cells grown to 30 days of culture, 112 sites (97%) are found to be common between vector types.

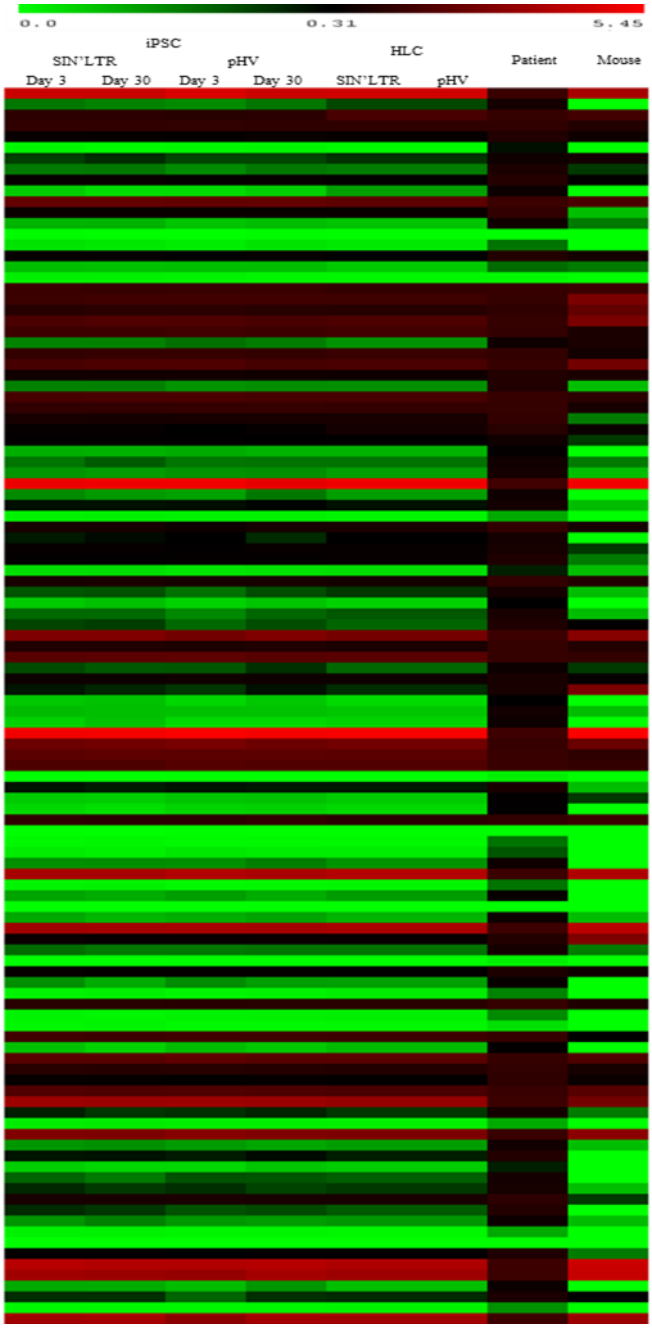


Figure 5.13- Heat map of pTFBS identified in *in vitro* iPSC and HLC samples infected with pHR (SIN) or pHV (native) LV and *in vivo* mouse (pHR) and HIV patients (native LTR) IS samples. Percentage of pTFBS present in 20 bp upstream and downstream of insertion sites are shown, as a proportion of all sites identified

Analysis of the top 10 most common pTFBS found in samples by sequence counts (hits) revealed further similarities. 82% (9) pTFBS are common between iPSC pHR LV infected samples identified between the early and 30-day time points. The same is shown for iPSC samples transduced with pHV LV. 9 (82%) sites are common between HLC pHR and pHV LV data sets. Comparison of the top 10 pTFBS between cell types demonstrates 9 (82%) of sites are common between pHR iPSC and HLC transduction and 100% of sites are common between pHV transduction. This reveals common pTFBS targeting, independent of cell type or U3 modification in the LTR.

To assess similarities between pTFBS identified between cell types, we next compared sites between iPSC and HLC samples transduced using the same LV LTR configuration. This analysis revealed that the majority of binding sites are found to be similar between the cell types. Comparison of iPSC and HLC transduced found 112 sites (97%) are common between pHR samples and 114 sites (99%) common between pHV transduced samples (Figure 5.13). These analyses suggest that LV IS for each vector highly associated with common pTFBS present in the DNA of the cell.

5.2.9. 5’ LTR pTFBSs Influence *in vitro* IS selection

pTFBS present in the 5’ LTR sequences of pHR and pHV LV vectors were analysed. 5’ LTR sequences were analysed via oPOSSUM v3.0 Single Site Analysis (http://opossum.cisreg.ca/) to identify these predicted binding sites. 42 pTFBS were identified in the pHV LTR and 17 in the pHR SIN LTR (Figure 5.14). Interestingly, the deltaU3 introduces one additional pTFBS (Hand1::Tcfe2a) not found in the native configuration.

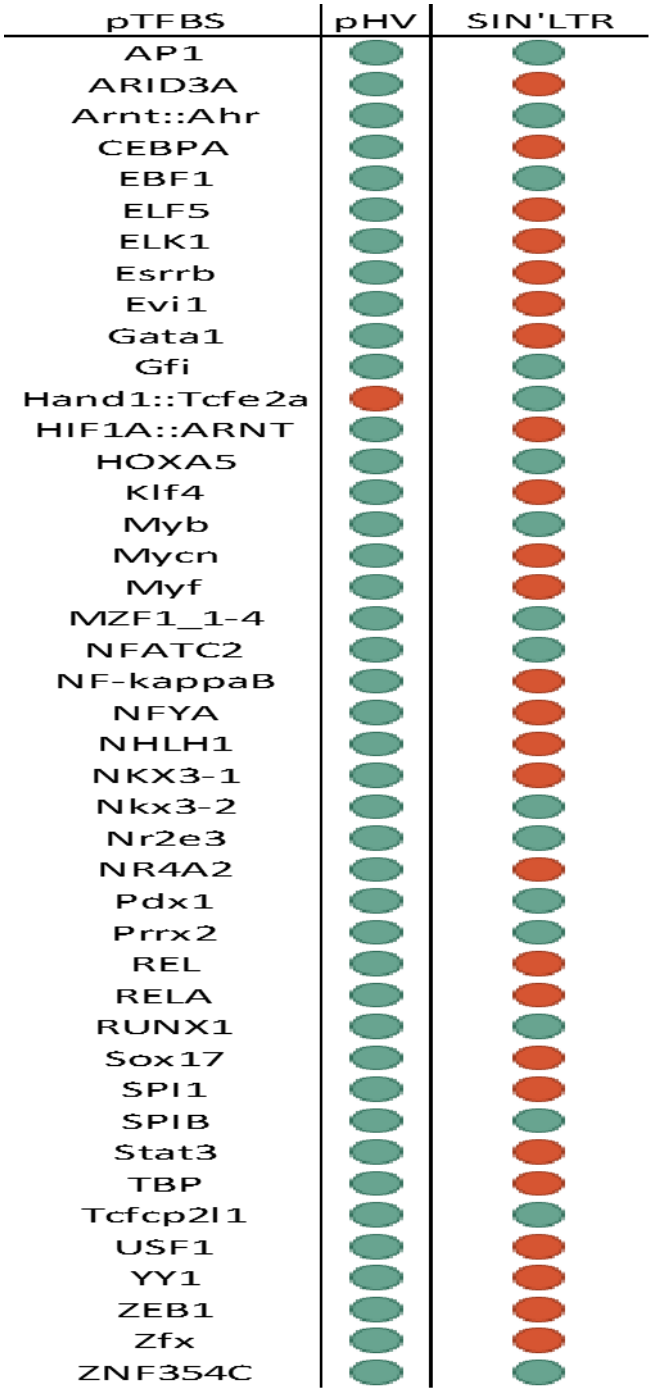


Figure 5.14. pTFBS present in native (pHV) and SIN (pHR) 5' LTR sequences. Presence of pTFBS in sequence shown in green while absence of site in sequence is shown in red. This indicates 42 sites are identified in pHV, with 17 sites found in pHR. All sites in pHR are common with pHV except Hand::Tcfe2a

pTFBS identified in pHR SIN and pHV complete LTR account for only 17 out of 170 (10%) and 42 out of 170 (25%) of families in the Jasper database, respectively (Portales-Casamar, Thongjuea et al. 2010).

5.2.10. pTFBS in LTR LV Configurations are Present in the IS of Pluripotent and Differentiated Cells

We next analysed more specifically the frequency of pTFBS common to pHR and pHV LTRs and IS sequences in P106 iPSCs and derived HLCs. To do this, we chose a 20 bp window around LV IS (Section 2.1.4.3). This analysis showed that in pHR transduced iPSC harvested at day 3 and day 30 time points, pTFBS occurring in the SIN LTR account for 227,297 out of 633,929 (36%) and 116,539 out of 322,340 (36%) of all IS related pTFBS, respectively. In pHV transduced iPSC at these time points, 213,474 out of 329,674 (65%) and 194,721 out of 297,681 (65%) of IS sequences, respectively were identified as LTR pTFBS. Analysis of pHR and pHV transduced HLC (early time point only) found 44,600 out of 122,760 (36%) and 119,700 out of 185,552 (65%) IS were recognised within the pTFBS present in the each LTR. Importantly, the percentage pTFBSs common to LV LTRs and host IS was maintained regardless of vector, cell type and culture time.

To show that pTFBS present within 20 bp of insertion sites in both iPSC and HLC are at a level greater than would be expected in random 20 bp segments of human DNA and that LV insertions occur predominantly near these pTFBS, we generated 100,000 random 20 bp sequences using Perl script and BEDTools for use as background sequences for Opossum analyses of pTFBS. Calculation of Z scores derived from pTFBS analysis using the random sequences as the background data set demonstrates that >95% of sites are statistically

significantly enriched above background levels, as defined by a positive Z score (Figure 5.15).

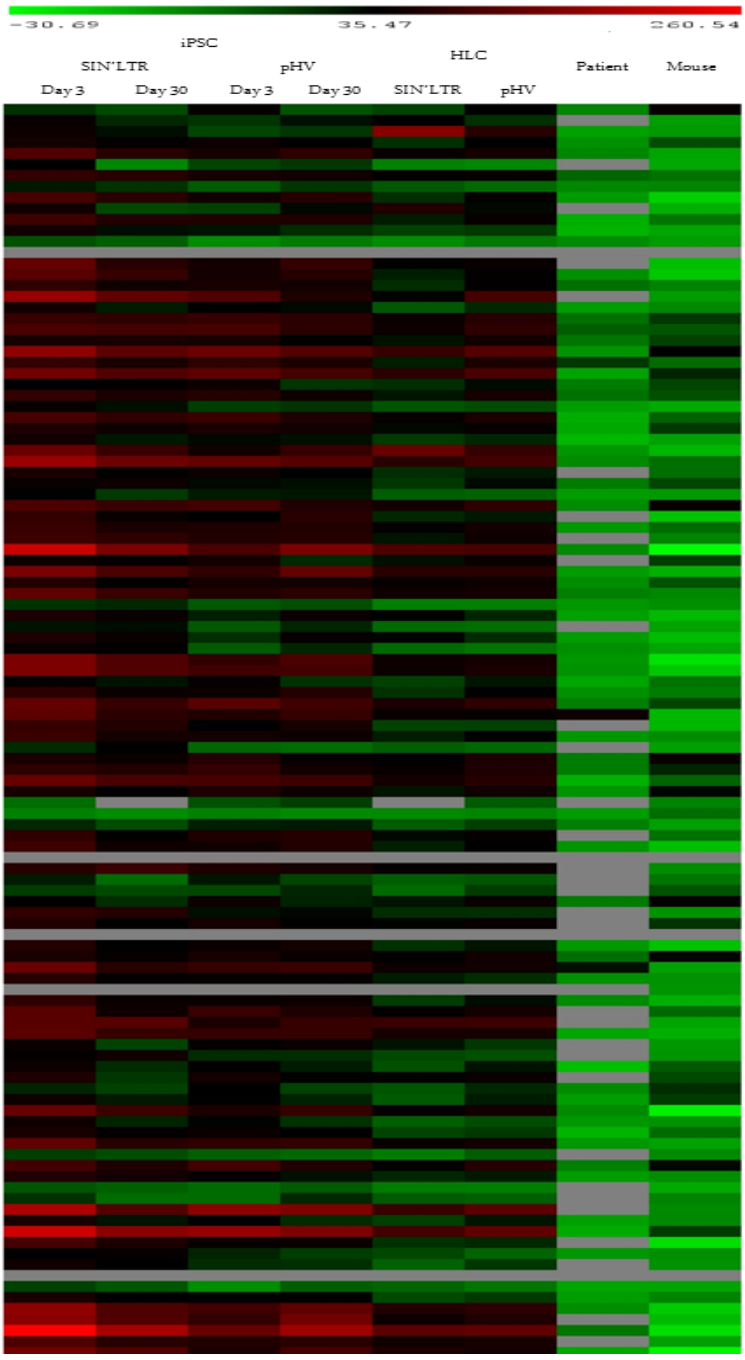


Figure 5.15. Heat map of enrichment of pTFBS present within 20bp upstream and downstream of IS in *in vitro* iPSC and HLC pHR and pHV infected samples and *in vivo* murine and HIV patient data sets. Associated significance against randomly generated background data sets is shown using Z scores, with pTFBS shown in alphabetical order. Positive Z score is taken as enrichment in comparison to background. Patterns are observed with more pTFBS present in the LTR more significantly enriched and pTFBS present in the LTR and known to be involved in HIV lifecycle to be moderately enriched

In iPSC pHR samples analysed at day 3, 110 out of 110 (100%) of sites are found to be enriched. In the 30-day iPSC pHR sample, 108 out of 109 (99%) sites enriched above background levels. Analysis of pHV transduction of iPSC indicates that 108 out of 110 (98%) and 109 out of 110 (99%) of sites are found above background at 3 and 30 days, respectively. HLC transduction reveals that 107 out of 109 (98%) and 110 out of 110 (100%) of sites are found above background levels in pHR and pHV samples, respectively. This suggests enrichment of pTFBS in the proximity of IS after LV transduction in both cell types.

Enrichment of pTFBS found in full and SIN LTR configurations using the Z scores were analysed after comparison of insertion site sequences to background sequences as above. Analysis of positive Z scores identified the majority of pTFBS found in LTR regions to be enriched. 17 out of 17 (100%) and 16 out of 17 (94%) of pTFBS were enriched across both iPSC transduced using pHR LV, at both day 3 and day 30 time points, respectively. Analysis in iPSC transduced with pHV LV demonstrates 41 out of 42 (100%) of pTFBS found in the LTR are enriched in both day 3 and day 30 samples respectively. Analysis of HLC samples indicates pTFBS in the LTR are enriched above background with of 16 out of 17 (94%) and all 42 sites (100%) in pHR and pHV LV samples, respectively.

These data suggest that pTFBS in the LTR influences IS selection with sites near pTFBS found enriched above the random control data sets.

5.2.11. pTFBS Appear Associated with LV Life Cycle

Analysis of specific transcription factors associated with the HIV lifecycle and are known to contribute to viral latency and expression. NFATC2 has been

shown to interact with U3 and U5 regions to activate transcription of the viral genome (Romanchikova, Ivanova et al. 2003). PBX1 and ZEB1 have been identified as transcription factors responsible for regulating transcription of the HIV genome (Venkatachari, Zerbato et al. 2015). AP1 has been shown to contribute to HIV-1 latency (Duverger, Wolschendorf et al. 2013). NF-κβ and SP1 sites present are both involved in expression of the viral genome (Stroud, Oltman et al. 2009, Harrich, Garcia et al. 1989).

Binding sites for PBX1, ZEB1, NFATC2, AP1, NF-κβ and SP1 transcription factors are identified near ISs in cultures of iPSC and HLC for complete and SIN LTR configurations. Statistical analysis of these pTFBS using Z scores compared to background control confirmed enrichment of these sites (Table 5.5).

pTFBS	iPSC				HLC		Patient	Mouse
	pHR		pHV					
	Day 3	Day 30	Day 3	Day 30	pHR	pHV		
AP1	60.78	42.3	40.49	48.09	26.54	25.4	-6.63	5.12
NFATC2	40.42	30.47	42.05	22.5	18.65	29.91	-7.73	5.56
NF-κB	118.4	74.52	65.77	84.65	43.64	43.13	50.41	-12.37
PBX1	19.95	16.08	16.7	25.69	7.61	19.87	NI	13.87
SP1	127.3	90.61	61.87	83	42.76	53.05	-0.5	-27.08
ZEB1	163.1	106.5	94.4	119.4	66.92	76.48	-1.21	-17.37

Table 5.5- Enrichment of pTFBS present within 20bp of IS in *in vitro* iPSC and HLC pHR and pHV transduced samples and *in vivo* murine and patient data sets, known to be involved in HIV lifecycle. Associated significance against background is shown using Z scores. pTFBS are shown in alphabetical order to observe trends across all data sets. A positive Z score is taken as enrichment in comparison to background and is highlighted. Trends can be observed with all pTFBS enriched across data sets.

In iPSC transduced using pHR at day 3, Z scores for PBX1 is 19.948, ZEB1 is 163.11, NFATC2 is 40.421, AP1 is 60.78, NF-κβ is 118.42 and SP1 is 127.3. In iPSC pHR samples at day 30, these scores were determined as 16.082 (PBX1), 106.49 (ZEB1), 30.473 (NFATC2), 42.298 (AP1), 74.521 (NF-κβ) and 90.611 (SP1). Analysis of iPSC pHV transduced samples at day 3 identified Z scores of these pTFBS as 16.702 (PBX1), 34.401 (ZEB1), 42.051 (NFATC2), 40.489 (AP1), 65.771 (NF-κβ) and 61.868 (SP1). At day 30 iPSC pHV samples, Z score of 25.685 (PBX1), 119.43 (ZEB1) and 24.499 (NFATC2), 48.089 (AP1), 84.647 (NF-κβ) and 82.997 (SP1) was determined. In HLC samples transduced using pHR LV, Z scores of these pTFBS were determined as 7.611 (PBX1), 66.921 (ZEB1) and 18.652 (NFATC2), 26.538 (AP1), 43.64 (NF-κβ) and 42.755 (SP1). In HLC samples transduced using pHV, Z scores were determined as 19.871 (PBX1), 76.478 (ZEB1) and 29.91 (NFATC2), 25.395 (AP1), 43.129 (NF-κβ) and 53.053 (SP1). This indicates these sites are highly enriched above background suggesting specific targeting of these pTFBS by each LV vector.

The frequency of overall pTFBS hits decrease from 633,929 to 322,340 (49% decrease) over time observed at the day 3 and day 30 time points in iPSC transduced with pHR. In iPSC transduced with pHV, 329,674 to 297,681 (10% decrease) pTFBS hits were observed between the early and 30-day sampling time points. A decrease is also observed in the frequency of pTFBS known to be involved in HIV lifecycle. Analysis of PBX1, ZEB1, NFATC2, AP1, NF-κβ and SP1 show in iPSC transduced with pHR determine 60,039 to 30,086 (50% decrease) hits were identified at the day 3 and day-30 time points. In iPSC samples transduced with pHV, 30,104 to 28,489 (5% decrease) hits to these pTFBS between day 3 and day-30 time points were identified. Sequence identity to pTFBS decreases consistently between the early and 30-day sampling time points of iPSC pHR samples for PBX1, ZEB1 and NFATC2,

AP1, NF-κβ and SP1. These specific pTFBS decrease from 848 to 441 (48% decrease) for PBX1, 22,744 to 11,192 (51% decrease) for ZEB1, 13,450 to 6,912 for NFATC2 (49% decrease), 18,192 to 9,167 for AP1 (50% decrease), 1,792 to 848 (53% decrease) for NF-κβ and 3,013 to 1,526 for SP1 (49% decrease) respectively in pHR samples. In pHV iPSC samples, an increase in sequence hits is identified for PBX1 with 448 to 460 (3% increase), 801 to 858 (7% increase) for NF-κβ and 1302 to 1395 (7% increase) for SP1. However, ZEB1, NFATC2 and AP1 hits decrease from 10,944 to 10,833 (1% decrease), 7,307 to 6,207 (15% decrease) and 9,252 to 8,736 (6% decrease), respectively over time. This indicates with both vectors that the majority of pTFBS decrease over time. However, differences are identified between the decrease in pTFBS hits between pHR and pHV vectors. In pHR iPSC samples, pTFBS involved in the HIV lifecycle decrease by a consistent proportion to total pTFBS found in pHR. However, in pHV iPSC samples, the transcription factors known to be involved in HIV lifecycle decrease by a lower percentage, suggesting the complete LTR may drive prevalence within these sites and that the pTFBS in HIV LTR have an influence over integration site selection by the vector, in particular to sites involved in viral replication, which appeared to persist over time.

5.2.12. Analysis of pTFBS in the IS of HIV-1 Infected Patients

To investigate HIV targeting in patients undergoing ART after several years *in vivo* clonal establishment, pTFBS around HIV-1 IS of infected patients was analysed using data obtained from the Retrovirus Integration Database (https://rid.ncifcrf.gov/intro.php) (RID).

To analyse pTFBS in common with the 5’ LTR region sequence of HIV-1, RID provided a total of 525,342 IS loci that were input into BEDTools

(https://bedtools.readthedocs.io/en/latest/) to generate 20 bp sequences 5' and 3' from each site. oPOSSUM v3.0 Single Site Analysis (http://opossum.cisreg.ca/) found 78 significant pTFBS families within 20 bp of the patient ISs (Figure 5.15). Comparative analysis of pTFBS also present in 5' HIV-1 LTR found 39 (representing 50% of the 78) sites predicted in the LTR were identified within IS data (Figure 5.13). Out of the total of 1,248 individual pTFBS hits for all TFBS families, 812 (65%) are found in common with the HIV LTR region. This implies that the LTR configuration influences IS selection *in vivo* and may direct the retention of certain IS in cells that progress towards clonal outgrowth as this data derives from clones established in patients living with HIV for several years (Wagner, McLaughlin et al. 2014).

This data was supported by comparison with a randomly generated data set provided using Perl script and BEDTools to generate 100,000 random 20 bp sequences for oPOSSUM analysis. Comparison between sequence hits between patient data and background samples reveal 22 out of 78 (28%) pTFBS with a positive Z score supporting enrichment (Figure 5.15). However, only 8 out of 42 (19%) of pTFBS present in the LTR are found to be enriched, suggesting long term clonal outgrowth is not favoured by IS with common pTFBS present in the LTR.

5.2.13. Characterisation of SIN Configuration HIV-1 LV IS Used to Infect Mice Also Reveal pTFBS Common with the LTR

To characterise IS selection of HIV-1 based LV carrying SIN configuration in a non-human, *in vivo* setting with respect to pTFBS motifs, fully immunocompetent mice (n=3) were injected by colleagues at early stage in development (day 16) to try to capture a wide array of vector insertions in highly expressed genes. Mice were injected with $8x10^{10}$ IU/ml of pHR LV and

sacrificed 4 weeks later to obtain DNA from their livers, where we have shown previously the majority of gene transfer occurs (Nowrouzi, Cheung et al. 2013). Mouse livers were subjected to immunohistochemistry for GFP expression. None of the mice showed adverse effects or tumour development and had normal liver morphology. Colleagues have previously used ligation adaptor mediated (LAM) PCR to determine LV vector insertion site selection in the mouse. These mouse injections have been previously reported (Themis, Waddington et al. 2005). DNA samples from the infected mice were subjected to this procedure and each successfully amplified fragment was sequenced and subjected to BLAST (http://www.ncbi.nlm.nih.gov/genome/seq/MmBlast.html) and BLAT (http://genome.ucsc.edu) alignment to the murine genome (and (http://genome.ucsc.edu). From this, IS loci (n=193) were mapped to 112 pTFBS motifs (Figure 5.15). All 17 (15%) pTFBS identified in 5' SIN LTR can be identified within IS. 10,654 pTFBSs hits were identified in total within 20 bp of the IS of which 2,524 pTFBSs hits are also present in the LTR (24%). A similar percentage was identified in human and the *in vitro* iPSC and HLC data above. This suggests in a murine model that the IS selection is partially influenced by LTR configuration with common pTFBS found between the LTR and IS data.

Murine background sequences for Opossum analysis were generated using Perl script and BEDTools to generate 100,000 random 20 bp sequences. Comparison of Z score significance of sequence hits to pTFBS in murine IS data sets compared background indicates 50 out of 112 (45%) sites are enriched, with 9 out of 17 (53%) of pTFBS found in SIN LTR significantly enriched in murine samples (Figure 5.15).

5.2.14. Comparison between *in vitro* and *in vivo* Data Sets

To assess whether *in vitro* data generated mirrored the *in vivo* data generated in this study, we compared pTFBS identified between *in vitro* samples following pHV infection with patient data and found all 78 sites in the human patient data in common with the 115 iPSC site data, for both the early and 30-day timepoints (68%). Similarly, all 78 sites in the patient data are in common with the 114 pTFBS (68%) identified in HLC *in vitro* infection data.

Comparison of pTFBS identified between *in vitro* pHR samples and murine data finds 97% (111, day 3) or 98% (110, day 30) sites in the murine data can also be found in common with the 115 and 112 pTFBS identified in *in vitro* iPSC samples. Similarly, all 111 sites in the murine data are in common with the 98% pTFBS (113) in the HLC pHR *in vitro* data.

In both patient and mouse data sets, the total number of pTFBS matches was less than that identified in the iPSC and HLC *in vitro* data sets. In mouse insertion site data, 10,654 sequences were identified with pTFBS in 244 ISs and 1,248 sequences in 534 ISs in patient data. This is significantly lower than sequence hits found *in vitro,* in both iPSC and HLC samples, with 633,929 and 322,340 sequences identified in 118,701 and 60,157 ISs at day 3 and day 30 in iPSC transduced with pHR. In pHV iPSC, 329,674 and 27,681 sequence hits are identified in 121,160 and 55,753 ISs. In HLC, 122, 760 and 185, 522 sequence hits are found in 22,874 and 34,142 ISs in pHR and pHV transduced cells, respectively. This is possibly due to the maturity of the cells analysed or the time of sample harvest post transduction.

Enriched pTFBS, indicated by a positive Z score, for PBX1 (13.867), NFATC2 (5.562) and AP1 (5.115) can be found to be near to murine ISs and are significantly enriched in comparison to background sequences. While ZEB1, NF-κβ and SP1 can be detected, these pTFBS do not have positive Z scores and

are therefore not enriched. In HIV patient data only NF-κβ (50.411) can be found as significantly enriched. This shows that differences exist between the *in vivo* data and the iPSC/HLC in vitro data where all pTFBSs can be identified as significantly enriched.

When comparing enriched pTFBS in the LTR or that are known to be involved in HIV lifecycle, we observe the majority of these sites are highly enriched. Sites identified in the LTR and known to be involved in HIV lifecycle (AP1, NFATC2) are observed as moderately enriched as indicated by Z scores (Figure 5.15). The majority of sites are enriched across most data sets (Table 5.5). 10 out of 20 (50%) of these pTFBS are significantly enriched in all data sets or excluding one in vivo data set. NK3-1, PDX1 and PRRX2 are significantly enriched across in vitro iPSC and HLC data sets and *in vivo* mouse and patient data. AP1, GFI, HOXA5, NFATC2, PBX1, SPIB and NF-κβ are enriched in iPSC and HLC in vitro data and one *in vivo* data set. Most pTFBS were identified in the LV LTRs and are known to be involved in HIV lifecycle with significantly enrichment both *in vitro* and *in vivo*.

5.2.15. LTR pTFBS align with genes previously reported to be involved in insertional mutagenesis

Analysis of pTFBS present in 8 genes known to be involved in insertional mutagenesis in lentiviral clinical trials, namely *LMO2*, *PRDM16*, *CCND2*, *MECOM*, *HMGA2*, *BMI1*, *BCL2* and *PRDM1*. oPOSSUM v3.0 Single Site Analysis (http://opossum.cisreg.ca/) found 116 pTFBS, of which 42 out of 43 (98%) of sites present in the native and SIN LTRs are present within these genes (Figure 5.16). Analysis of enrichment of these genes against all 24,752 genes stored in the oPOSSUM database using positive Z scores indicate that

32/42 (82%) and 14/17 (82%) of sites in the pHV and pHR LTR are enriched above background levels, respectively.

LTR pTFBS	Native LTR	SIN'LTR	Pooled Z scores
AP1			8.25
ARID3A			18.60
Arnt::Ahr			5.87
CEBPA			10.93
EBF1			0.34
ELF5			18.40
ELK1			5.14
Esrrb			5.23
Evi1			0.78
Gata1			9.73
Gfi			12.11
Hand1::Tcfe2a			-0.27
HIF1A::ARNT			4.99
HOXA5			22.73
Klf4			-3.59
Myb			5.51
Mycn			7.05
Myf			-4.74
MZF1_1-4			5.31
NFATC2			9.17
NFYA			-6.21
NF-kappaB			-1.49
NHLH1			-0.17
NKX3-1			13.32
Nkx3-2			8.73
Nr2e3			4.63
NR4A2			4.73
Pdx1			9.31
Prrx2			12.59
REL			9.14
RELA			-0.15
RUNX1			2.59
Sox17			16.40
SPI1			9.78
SPIB			13.88
Stat3			1.65
TBP			14.25
Tcfcp2l1			-8.94
USF1			1.44
YY1			3.44
ZEB1			-0.70
Zfx			-10.21
ZNF354C			-2.86

Figure 5.16. Enrichment of native and SIN LTR specific pTFBS within *LMO2*, *PRDM16*, *CCND2*, *MECOM*, *HMGA2*, *BMI1*, *BCL2* and *PRDM1* genes (pooled). Presence of pTFBS in LTR sequence shown in green while absence of site in sequence is shown in red. Associated significance against background is shown using Z scores, with pTFBS shown alphabetically. A positive Z score is taken as enrichment in comparison to background and is highlighted.

These comparative analysis results indicate LV IS selection is partially influenced by cellular proteins such as transcription factors that may influence site selection via tethering the vector to preferred sites in the genome. Sites present in the LTR may therefore influence binding to genes known to be involved in insertional mutagenesis. Modifications to LTRs may be proposed to ensure the removal of adverse pTFBS.

5.3. Discussion

LVs integrate in the host genome, predominantly into active genes (Mitchell, Beitzel et al. 2004, Barr, Ciuffi et al. 2006, Barr, Leipzig et al. 2005, Lewinski, Yamashita et al. 2006). Multiple models have been developed to better understand HIV-1 based integration and genotoxicity. IS selection must be understood further due to the risk of insertional mutagenesis and associated oncogenesis, as observed in in a number of non-clinical and clinical settings (Hacein-Bey-Abina, Garrigue et al. 2008, Howe, Mansour et al. 2008, Knyazhanskaya, Anisenko et al. 2019). LV vectors are believed to carry a lower genotoxic risk and are therefore seen as safer than retroviral vectors (Baum, Kustikova et al. 2006). However, IS selection still contribute towards the safety profile of these vectors with insertional mutagenesis being deleterious to host survival. Therefore, there is a need to better understand IS selection to improve the design of LV vectors for gene delivery. The current understanding of viral integration suggests tethering of the viral genome and PICs to host cellular proteins (i.e. hRad1, PSIP1/LEDGF/p75), which are involved in repairing nicks in DNA post integration (Knyazhanskaya, Anisenko et al. 2019, Mulder, Chakrabarti et al. 2002, Cherepanov, Maertens et al. 2003, Vandegraaff, Devroe et al. 2006). However, integration in cells where PSIP1/LEDGF/p75 has been knocked out, still show preferential integration of HIV-1 into active transcription units suggesting other factors must influence target site selection (Ciuffi, Llano et al. 2005).

Few models exist to aid the understanding of LV integration and cellular transformation. There are existing animal models that have shown that LV mediated insertional mutagenesis within known cancer related genes contribute to oncogenesis (Cesana, Ranzani et al. 2012, Montini, Cesana et al. 2009, Nowrouzi, Cheung et al. 2013). However, animal models are being increasingly discouraged due to the aim to reduce animal cruelty and are seen as not fully

representative of the human system. These models also often employ mice with specific genetic deficiencies which may bias the system towards tumour formation (Jackson, Thomas 2017). Few *in vitro* models have been published. The IVIM is a useful alternative to using animals identifying the role of enhancer regions in immortalisation, the integration profile of SIN configuration LV vectors and indicates no significant role of WPRE in clonal transformation (Modlich, Navarro et al. 2009). Similarly, a model has been generated using BAF3 and Bc115 cell lines (Bokhoven, Stephen et al. 2009). However, these methods still employ murine cells so there is a need for a human based model of genotoxicity.

Human iPSC and their derivatives have been reported as useful models for various diseases. Analysis of the gene expression of iPSC and their HLC derivatives indicate distinct gene expression profiles at both stages, implying the need to understand LV IS selection at both stages. As iPSC present a naïve cell state, with multiple genes expressed at high levels, this was proposed as a useful model for understand LV genotoxicity.

HIV-1 based LV integrate semi-randomly but 70-80% preferentially within the gene body of highly transcribed genes (Ciuffi, Llano et al. 2005, Schroder, Shinn et al. 2002, Mitchell, Beitzel et al. 2004, Rossetti, Cavarelli et al. 2013). Analysis of IS by LV carrying a SIN (pHR) or native (pHV) LTR show subtle differences in integration profiles suggesting the benefits of using the SIN LTR, but that this does not fully avoid cancer related gene which may contribute to insertional mutagenesis and clonal outgrowth. Furthermore, enrichment of IS genes to assess overrepresented biological pathways show clear differences in genes targeted by each vector. The differences in IS shown at both the iPSC and HLC stage validate the need to understand genotoxicity in both cell types. While a greater number of IS are found in iPSC, this is likely due to higher levels of gene expression in this cell type (Gao, Zhang et al. 2017, Schroder,

Shinn et al. 2002, Mitchell, Beitzel et al. 2004). A reduction in clonality is observed over time, with fewer ISs observed in cells sampled after long term growth, similar to previous reports (Ronen, Negre et al. 2011).

IS in cancer genes and associated insertional mutagenesis is a serious adverse event, mediated by LV gene therapy (Hacein-Bey-Abina, Hauer et al. 2010, Hacein-Bey-Abina, Garrigue et al. 2008, Hacein-Bey-Abina, Von Kalle et al. 2003, Ronen, Negre et al. 2011, Themis, Waddington et al. 2005). To understand the effects of LV integration on clonality further, P106 iPSCs were infected with pHR or pHV LV and assayed over time. Analysis over 30 days presented evidence for examining clonal outgrowth, with a significant increase in viral sequences detected in oncogenes and protooncogenes. HCC presents epidemiologically as the fourth most common cancer worldwide (Singal, Lampertico et al. 2020). Analysis of IS in P106 iPSC and HLC show multiple insertions in genes known to be involved in this cancer subtype, presenting this model as a useful tool in understanding HCC further.

pTFBSs that appear in the U3 region of RV LTRs are known to associate with pTFBSs in the host genome (Felice, Cattoglio et al. 2009). Work by Felice and colleagues identified TFBS of HIV-1 based LV did not associate with the full LTR, independent of the U3 region, within 1kb of IS except after substitution of the LV with MLV U3 region. This modified vector then showed a positive bias towards TFBS, as seen in MLV integration suggesting the U3 is important in tethering the vector genome to similar sites in the host.

Analysis of pTFBS significantly represented around IS (20bp window) in an immature proliferative (iPSC) and terminally differentiated (HLC) cell types, identified pTFBS significantly associated with these insertions suggesting these sites may be involved in tethering the vector genome to the host. This has been shown to be independent of cell type and LTR configuration. Little difference is observed in associated pTFBS over time in P106 iPSC suggesting the LV

tethering to common pTFBS to ensure survival. However, the quantity of pTFBS around IS with the LTR decrease over time in iPSC, possibly due to cell death and clonal dominance. Statistical analysis of pTFBS identified found enrichment above background levels of these sites and in particular, the pTFBS identified in the LTRs. The PIC is known to interact with nuclear import proteins, to influence IS selections suggesting nucleoporins may have a role in tethering the viral genome towards chromatin rich regions (Mitchell, Beitzel et al. 2004). Proteins such as Sus1 are involved in tethering of active genes to nuclear pore complexes to aid in transcriptional activation and nuclear export of nucleic acids (Pascual-Garcia, Govind et al. 2008, Rodríguez-Navarro, Fischer et al. 2004). These results suggest that there is an involvement of the pTFBS in the LTR tethering to IS site selection with the range of pTFBS present allowing adaptation of the virus to target cells.

The pTFBS data analysis in iPSC and HLC cultures was highly comparable with pTFBS identified in the IS of HIV-1 in patient genomes with full LTR that also suggests HIV-1 could be directing site selection. This was also found in mice with the SIN LTR configuration LV.

We show that pTFBS integration sites are found for TFs known to be associated with HIV transcription, namely PBX1, ZEB1, NFATC2, AP1, and SP1. PBX1 has been shown to be involved with viral transcription (Chao, Walker et al. 2003, Tacheny, Michel et al. 2012, Ma, Dong et al. 2013), ZEB1 and AP1 have been shown to be involved in HIV latency (Venkatachari, Zerbato et al. 2015, Duverger, Wolschendorf et al. 2013). NFATC2 is essential for productive infection of non-activated T-cells (Hohne, Businger et al. 2016). NF-κβ and SP1 sites in the LTR have both been shown to be involved in transcription of the HIV genome (Stroud, Oltman et al. 2009, Harrich, Garcia et al. 1989).

The pTFBS present in the full and SIN LTR can be identified in eight genes (*LMO2*, *PRDM16*, *CCND2*, *MECOM*, *HMGA2*, *BMI1*, *BCL2*, *PRDM1*) known

to be involved in insertional mutagenesis from clinical trials (David, Doherty 2017). Interestingly, the unique pTFBS identified in the SIN LTR can be identified in all eight genes. As previous analysis of *in vitro* iPSC and HLC and *in vivo* patient and murine data after pHR and pHV LV transduction indicate that the sites present in the LTR influence insertion, the presence of these pTFBS in genes known to be involved in insertional mutagenesis suggest that tethering of the viral genome may occur to genes deleterious to survival and vector safety. Thus, further modifications to the LTR sequence to eliminate these pTFBS recognition sequences in LV vectors is recommended in the development of safer lentiviral vectors for gene therapy.

In summary, we present data to show the integration profile of LV vectors, with either native or SIN LTR configuration in a non-biased, proliferative cell type (iPSC) and its derivatives (HLC). The differences in IS selection and gene enrichment signify the need for both cell types to understand IS choice further. These results also show the LTR playing a significant role in pTFBS tethering to the genome. These results are useful in understanding vector design for the development of LV vectors for downstream applications in gene therapy. While we have developed a model of HIV-1 based LV-mediated genotoxicity in hepatocytes, it may also be useful to assess vector related side effects in other tissue types.

Chapter VI Cardiomyocyte Differentiation and Lentiviral transduction

The analysis of IS data derived from iPSCs and HLCs infected with pHR or pHV LV vectors has shown a sensitive model for lentiviral genotoxicity. The liver has been used as a standard model for toxicity. Here, we explore alternative tissue types, namely cardiomyocytes, as LV- mediated genotoxicity platforms. Various tissue specific differentiation may be used as models for tissue specific genotoxicity.

6.1. Introduction

6.1.1. Pluripotent Stem Cells and Cardiomyocyte Derivatives

Pluripotent stem cells, including induced pluripotent stem (iPS) cells are a promising source of cells for various therapeutic avenues, including gene therapy. These cells can be differentiated into multiple cell types, including HLC and cardiomyocytes. iPSC derived cardiomyocytes have shown success in animal models in remodelling of hearts post myocardial infarction, raising the possibility of human transplantation in patients (Rojas, Kensah et al. 2017, Guan, Xu et al. 2020). These cells also provide an opportunity to model diseases *in vitro* (Csobonyeiova, Polak et al. 2016, Pang 2020). Multiple cardiac syndromes have been modelled successfully using iPSC derived cardiomyocytes, including familial cardiac hypertrophy and LEOPARD syndrome hypertrophy (Lan, Lee et al. 2013, Carvajal-Vergara, Sevilla et al. 2010). Work has been completed to show correction of genetic diseases in iPSC derived cardiomyocytes through LV mediated gene transfer (Sato, Kobayashi et al. 2015, Guo, F., Sun et al. 2019). The use of iPSC and their derivatives are therefore a promising therapeutic avenue in gene therapy.

Various protocols have been developed for cardiomyocyte differentiation with improved yields (Balafkan, Mostafavi et al. 2020). These include co-culture with stromal cells, embryoid body formation or 2D monolayer differentiation (Batalov, Feinberg 2015). The development of these protocols has resulted in the availability of commercially optimised kits for highly efficient cardiomyocyte differentiation (Baci, Chirivi et al. 2020).

While iPSC derived cardiomyocytes are well characterised for cardiac specific markers, the maturity of these cells have been assessed to validate their recapitulation of phenotypes for downstream applications. Comparative gene

expression analysis shows that while these cells closely resemble the gene expression identified in *in vitro* cardiac tissue than other tissues, the gene expression profile remains more closely aligned with the immature cardiac cells (Pavlovic, Blake et al. 2018). Various cardiac specific genes (i.e. *MHY6*, *N2A*) have been shown to be expressed at high levels in these cells (Denning, C., Borgdorff et al. 2016)**.** However, these are not similar to genes upregulated *in vivo* adult cardiomyocytes. Genes found highly regulated *in vivo* are quantified at lower levels in iPSC derived cardiomyocytes (van den Berg, C W, Okawa et al. 2015, Zhou, Y., Wang et al. 2017). Gene ontology analysis identifies a reliance of these cells on glycolysis for energy production, rather than fatty acid metabolism as identified in mature cardiomyocytes. However, distinct differences are identified in iPSC and their cardiomyocyte derivatives dependent on the donor tissue which may need to be considered for downstream applications (Zhou, Wang et al. 2017).

As well as transcriptomic and metabolic differences, iPSC derived cardiomyocytes also exhibit structural and contractile alterations in comparison to adult heart tissue. Protocols have been developed to improve the maturity of these terminally differentiated cells. Increasing the length of time of culture has shown a marked improvement in maturity of these cells, though this is not always feasible for downstream applications and cost efficiency (Kamakura, Makiyama et al. 2013). Spheroid culture differentiation of cardiomyocytes have also been shown to mature quicker (Branco, Cotovio et al. 2019). Manipulation of the microenvironment, including through electrical stimulation and/or mechanical stress have all been shown to successful improve the maturity status of cells (Ruan, Tulloch et al. 2016, Jaleel, Nakayama et al. 2008) While these cells are a useful model for *in vitro* analysis, further development of these cells are required *in vitro* to fully recapitulate the *in vivo* microenvironment.

6.1.2. Cardiomyocyte Differentiation Protocols

Multiple protocols have been developed to differentiate these pluripotent cells into terminal cell types, including hepatocytes and cardiomyocytes (Burridge, Keller et al. 2012, Kehat, Kenyagin-Karsenti et al. 2001, Alhaque, Themis et al. 2018). Hepatocyte differentiation of these pluripotent stem cells has provided a wealth of material for *in vitro* toxicity testing, recapitulating the *in vivo* microenvironment. As a wide range of drugs fail in clinical trials due to liver toxicity, these cells are useful in modelling toxicity in vitro before progressing down pipelines further (MacDonald, Robertson 2009).

iPSC can be differentiated into cardiomyocytes. The importance of cardiovascular research can be seen with these diseases presenting as one of the leading pathologies worldwide, with 27.75% of deaths attributed to cardiovascular disease in 2017 (BHF 2019). The loss of myocardium function has been shown to contribute to cardiac failure, indicating the need for further studies in cardiomyocytes. While different methodologies have quantified cardiomyocytes to different levels, it is generally accepted that these cells occupy 70-85% of the volume of a mammalian heart. shown various (Zhou, P., Pu 2016, Altara, Manca et al. 2016). With greater emphasis on precision medicine, iPSCs can be differentiated to cardiomyocytes for use in downstream assays (Mordwinkin, Lee et al. 2013). Initial work showed human embryonic stem cells were able to differentiate *in vitro* into cardiomyocytes using serum, though at a low efficiency (5-10%) (Kehat, Kenyagin-Karsenti et al. 2001). Seminal studies showed the differentiation potential of iPS cells to cardiomyocytes. Stepwise protocols of differentiation have revealed Activin A and BMP4 signalling along with inhibition of Wnt signalling allows mesodermal and cardiomyocyte differentiation (Yang, Soonpaa et al. 2008, Laflamme, Chen et al. 2007). Therefore, there is a present need to study iPS cell

derived cardiomyocytes for further development in therapeutic and regenerative technologies.

iPS cell derived cardiomyocytes have been well characterised to show contractile function, homogenous cell culture and response to cardiac specific drugs equivalent to that of cardiomyocytes differentiated from embryonic stem cells (Laustriat, Gide et al. 2010). These cells have advantages over using embryonic stem cells as they can be derived from cells gained in minimally invasive methods, provide an indefinite supply of replicating cells and are personalised to patients, allowing more sensitive assay of various disease treatments. Therefore, iPS cell derived cardiomyocytes may provide an effective in vitro method for assessing various therapeutics before clinical progression.

There is a present need for further human models of therapeutics, to determine toxicity and side effects. This was highlighted by the incidence of postnatal deformities after mothers were prescribed thalidomide. However, while this drug was tested in rodents showed no prenatal effects (Warkany 1988).

6.1.3. Downstream Application of Induced Pluripotent Stem Cell Derived Cardiomyocytes

Cardiomyocytes differentiated from iPS cells have been used for drug screening and toxicity assays. With approximately 90% of drugs reported to fail candidate screening, more sensitive methods are required for determine drug efficacy and safety before further clinical progression in a cost and time effective manner (Marchetto, Carromeu et al. 2010). Cardiac toxicity is identified as a factor of multiple failed drugs. Toxicity screening of treated iPSC derived cardiomyocytes indicates development of reactive oxygen species, impaired contractibility, and changes in heart rate. Various drugs have been tested *in vitro*

on these cells, such as procainamide, mexiletine and flecainide, showing a similar effect on cardiomyocytes as found in a clinical setting (Yokoo, Baba et al. 2009). Therefore, these terminally differentiated cells are useful for downstream assays, recapturing clinical effects.

These terminally differentiated cardiomyocytes resemble human cardiomyocytes *in vivo*. These cells express cardiac specific markers, including *SIRPα* and *VCAM1*. Cardiac troponin T (cTnt) is a commonly used marker specific towards cardiomyocytes (Pushp, Sahoo et al. 2020, Adamcova, Skarkova et al. 2019). iPS cell derived cardiomyocytes contain a heterologous population of muscle cells, closely aligned with ventricular, atrial and nodal cardiomyocytes. The distinct physiological and electrophysiological properties of each derivative population should be accounted for in downstream applications towards specific therapeutics (Ma, Guo et al. 2011). This is further validated by identification of expressed genes specific to atrial (*ATP2A2*, *HEY1*, *MYH6*, *NR2F1*, *NR2F2* and *TBX5*) and ventricular (*GJA1*, *MYL2*, *KCNJ2* and *MYH7*) cardiomyocytes (Cyganek, Tiburcy et al. 2018). However, these cells are more closely aligned with immature cardiomyocytes *in vivo*, with the notable absence of T tubules and underdeveloped contractile machinery. This is also observed with the immature metabolic profile of cardiomyocytes derived from iPS cells, with a predisposition to using glycolysis for energy production rather than lipid oxidation, as observed in mature cardiac muscle cells (Zhou, Wang et al. 2017). While the immaturity of these cells may indicate they do not fully represent *in vivo* heart function, the maturity of these cells improves over longer cultures, with an increase in the maturity of cardiomyocytes identified by improved morphology, density, myofibril alignment, Z disks, A bands, I bands and H zones (Gherghiceanu, Barad et al. 2011, Kamakura, Makiyama et al. 2013, Lundy, Zhu et al. 2013). The use of adipogenic cocktails, addition of lactate and glucose free mediums have all been shown independently to

improve maturity of cardiomyocytes (Kim, C., Wong et al. 2013, Drawnel, Boccardo et al. 2014, Tohyama, Hattori et al. 2013). Further protocols have been developed for specific atrial or ventricular cardiomyocytes which may be useful for specific downstream assays when drug targeting or disease modelling in specific subtypes of cardiac muscle (Cyganek, Tiburcy et al. 2018).

Gene expression analysis of iPS cell derived cardiomyocytes have also enhanced the understanding of cardiac development *in vitro*. This has revealed enrichment towards stepwise molecular pathways throughout development, including genes involved in mesoderm formation (*BRY*, *MIXL1*), cardiac mesoderm (*ISL1*, *KDR*, *MESP1*), cardiac progenitors (*GATA4*, *HAND1*, *MEF2C*, *NKX2.5*, *TBX5*) and structural genes involved in sarcomere production (*MYH6*, *MYL2*, *MYL7*, *TNNT2*) (Kattman, Witty et al. 2011, Burridge, Matsa et al. 2014). Further analysis of expressed genes reveals expression of sodium and potassium ion channels used in cardiac conductivity, including *SCN5A*, *KCNH2* and *KCNQ1* (Liang, Lan et al. 2013). Calcium ion signalling is shown to be present through expression of *IP3R*, *RYR2* and *SERCA2A* (Jung, Moretti et al. 2012, Rao, Prodromakis et al. 2013, Itzhaki, Rapoport et al. 2011). The expression of these cells is closely aligned with that of human cardiomyocytes *in vivo* (Pavlovic, Blake et al. 2018). This evidence validating their usefulness in downstream assay and model development.

6.1.4. Non-Viral and Viral Treatments of Differentiated Cardiomyocytes

Transfection and transduction of iPS cell derived cardiomyocytes has been shown as an effective methodology to model disease treatment outcomes. Models have been developed for various cardiac related pathologies, including long QT syndrome (Moretti, Bellin et al. 2010). Generation of models can be

used to assess the effectiveness of various treatments *in vitro* before clinical progression.

Non-viral methods of transfection of iPS cell derived cardiomyocytes have yielded poor efficacy. While progression has been made to optimise these methods, including using PEI or lipoplexes, viral transduction is still providing benefits with a highly efficient system for cellular transduction (Tan, S., Tao et al. 2019). Various viral vectors have been shown to successfully transduce these cell types. The COVID-19 pandemic was shown to be caused by infection of the SARS-CoV-2 virus (Yuen, Ye et al. 2020). The importance of disease modelling was indicated through severe cardiac failure presenting in COVID-19 patients. The binding of the SARS-CoV-2 virus to the ACE2 receptor creates susceptibility to infection of cardiac cells (Walls, Park et al. 2020, Sharma, Garcia et al. 2020). Cardiomyocytes derived from iPS cells were shown to be successfully infected using Sars-CoV-2 and modelling disease progression to indicate apoptosis of cells and gene dysregulation. Thus, disease modelling could aid in treatment development for COVID-19.

Ad can transduce rat heart cells in a dose dependent manner, with the virus localising to the border of non-infracted regions (Takahashi, K., Ito et al. 2003). Adenoviral infection has been further optimising through alterations of culture medium to enhance CAR receptor binding (Li, Z., Sharma et al. 2003). However, due to the transient nature of transgene expression and the poor infection of Ads to the heart, with the majority of these vectors targeting the liver, other viral vectors have been analysed for infectivity assays (Kang, Yun 2010).

AAVs have been shown as effective transducers of cardiomyocytes, dependent on capsid proteins. AAV2 has been shown to deliver *VEGF* to murine cardiomyocytes, under the control of the myosin light chain 2v promoter (Su, Joho et al. 2004). Various serotypes of AAV, including AAV6, 8 and 9 have

been observed to robustly transduce heart cells (Inagaki, Fuess et al. 2006, Palomeque, Chemaly et al. 2007). Specifically, AAV9 has been seen to preferentially target cardiomyocytes post reperfusion injury (Konkalmatt, Wang et al. 2012). These studies show the success of various AAV serotypes in transduction of heart cells for employment in gene therapy.

LV can also transduce terminally differentiated cardiomyocytes. LV pseudotyped with VSV-G has been shown to transduce rat cardiomyocytes in vitro and in vivo, indicating the success of these vectors therapeutically (Bonci, Cittadini et al. 2003). These vectors have been seen to efficiently deliver transgenes at a lower MOI for comparable expression levels to AAV, at a MOI of 10^4-10^6 and Ads, at a MOI of 10^3 (Rapti, Stillitano et al. 2015). These vectors have also been used in the rescue of iPS cells and derived cardiomyocytes from Pompe disease (Sato, Higuchi et al. 2016, Sato, Kobayashi et al. 2015). These indicators suggest the usefulness of LV as a gene therapy vector, over that of adenoviral or AAV vectors to efficiently transduce cells *in vitro* and *in vivo*.

Here we show the successful differentiation of P106 iPS cells to cardiomyocytes. Time course gene expression analysis of differentiating iPS reveal further differences between cell types which is the basis for infections at various timepoints. Transduction using LVs carrying an eGFP transgene were optimised and shown for successful, non-cytotoxic integration of the virus in the host gDNA. These transduction experiments form the basis for analysis of integration sites to determine genotoxicity of viral integration in a personalised, human, *in vitro* cellular model.

Aims

- Determination of transduction timepoints during differentiation through gene expression analysis
- Comparative analysis of cardiomyocyte and liver gene expression
- Differentiation of iPSC to cardiomyocytes and verification
- Transduction of iPSC and cardiomyocytes and verification

6.2. Results

6.2.1. Gene Expression Analysis

To determine the most effective timepoints of LV transduction of cardiomyocytes differentiating from iPSC, we first analysed published data to determine expression profiles over time. This was particularly in relation to genes known to be involved in cancer, namely oncogenes, protooncogenes, and tumour suppressor genes. These genes are known to be involved in insertional mutagenesis and related oncogenesis due to LV integration.

Published gene expression data from RNA Seq analysis were derived from GSE116574 (Branco, Cotovio et al. 2019) as previously described (Section 2.1.4.1). Normalised gene expression data was enriched and analysed over the stages of differentiation (Section 2.1.4.2). Time course analysis of differentiation iPSC were analysed at Days 0 (iPSC), 1, 3, 5, 7, 9, 12, 15, 18 and 20 (mature cardiomyocytes). Initially, expressed genes identified between triplicate were compared between each timepoint data set. At day 0, 59,897 gene transcripts were identified in common between triplicate repeats, with 311 identified in individual samples. 60,206 gene transcripts were identified in common in day 1 samples, with 2 transcripts identified in other samples. At day 7, 60,204 transcripts were identified in common, with 6 transcripts uniquely identified in individual samples. Similarly, 60,207 gene transcripts were identified in common between samples taken at day 12 of differentiation, with 1 gene transcript uniquely identified in one sample. These comparisons were used to generate master list of all genes identified as expressed at each stage for further analyses. 60,199 transcripts were identified at day 5 of differentiation, 60,207 transcripts were identified at day 9 and 60,207 transcripts were expressed at days 15, 18 and 20. These identify multiple transcripts of genes expressed throughout differentiation.

Using master lists of gene transcripts, common genes were identified as expressed through cardiomyocyte differentiation. 21,572 genes are expressed at day 0, 1, 3 5, 12 and 20. 20,201 genes were identified at day 7. 6,687 genes were expressed at day 9 of cardiomyocyte differentiation. 10,097 genes were identified as expressed at day 15 and 12,295 at day 18.

Comparison of genes expressed at the iPSC stage (day 0) and other timepoints through cardiomyocyte differentiation indicate the majority of expressed genes are shared throughout differentiation. All expressed genes are shared between day 0 and day 1, 3, 5, 12 and 20. While 20,201 (93.6%) expressed genes are shared between day 0 and 7, 1,371 (6.4%) of genes are uniquely expressed at day 0. Similarly, 6,687 (31%) of genes are commonly expressed at day 0 and 9 but the 14,885 (69%) of genes are uniquely expressed at day 0. While 10,097 (46.8%) of genes are commonly expressed between day 0 and 15, 11,475 (53.2%) are only expressed at the iPSC stage. Comparison between day 0 and 18 reveals 12,295 (57%) of genes are shared but 9,277 (43%) of genes are uniquely expressed at the iPSC stage. This indicates slight differences in gene expression throughout cardiomyocyte differentiation. This suggests the need to understand time course expression further to assess the best timepoint for LV transduction to ensure maximum exposure to genes, including cancer related genes to develop an effective genotoxicity model.

Gene enrichment of expressed genes further reveals differences in biological pathways expressed through cardiomyocyte differentiation (Table 6.1). Genes expressed in each sample time point were enriched for biological pathways using FunRich software (Section 2.1.4.2). The top 10 most significant pathways ($P<0.05$) were compared to pathways expressed at the iPSC stage. This indicates that while all 10 (100%) of these pathways were shared between day 0 and days 1, 3, 5, 12 and 20, differences were found at other stages of differentiation. While 1 (10%) pathway was shared between the iPSC stage and

day 7, no pathways were shared between day 0 and days 9, 15 or 18. These differences are expected with further significance on mesodermal differentiation throughout cardiomyocyte development. Similar pathways are likely to be expressed throughout stages of differentiation due to common genes expressed between various stages. It is important to note that gene enrichment does not account for genes which are highly expressed. This suggests similarities in the biological pathways of genes expressed at various stages during differentiation indicate the need to further assess gene expression to determine the most effective timepoint for LV integration.

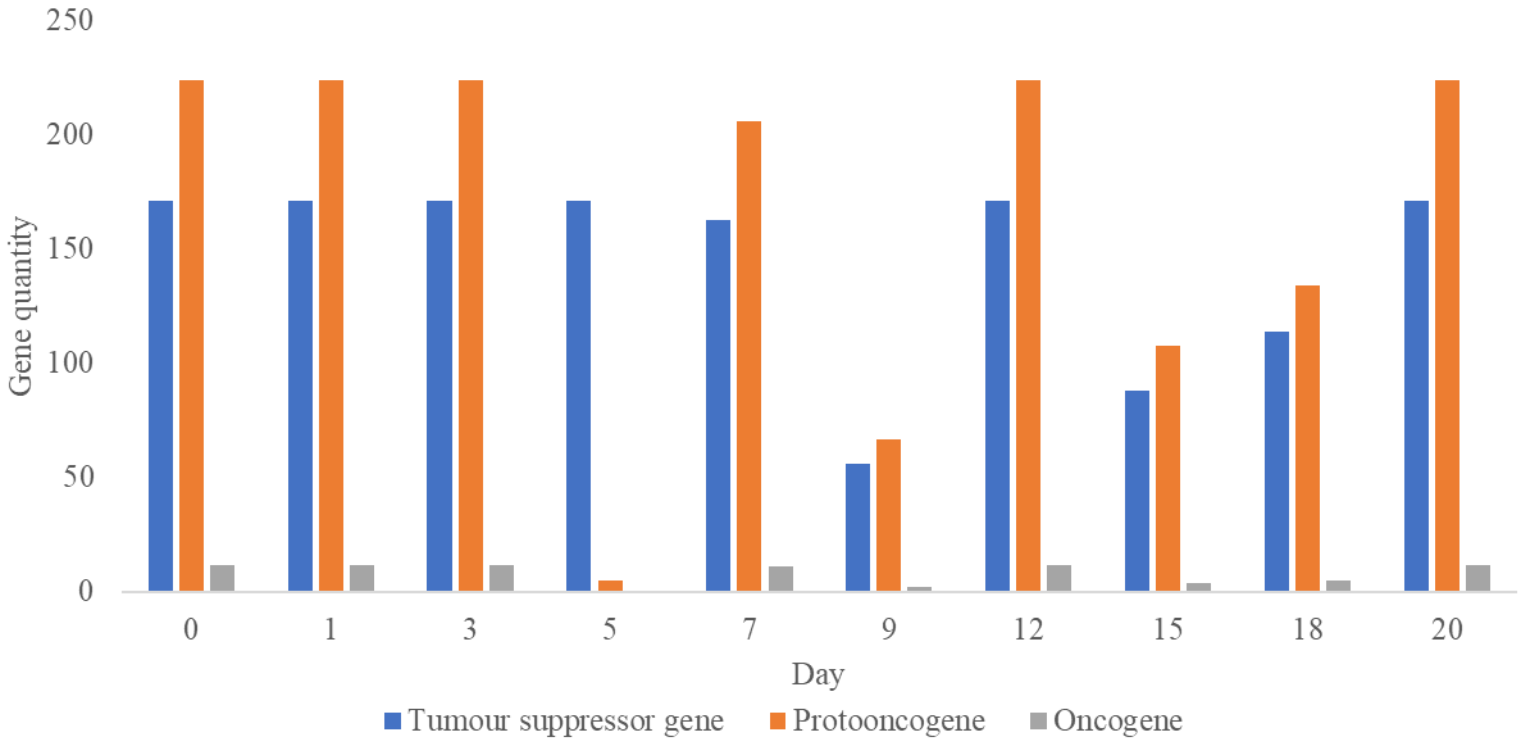
Day									
0	1	3	5	7	9	12	15	18	20
S Phase	S Phase	S Phase	S Phase	Signaling events mediated by TCPTP	superpathway of methionine degradation	S Phase	Direct p53 effectors	ALK1 signaling events	S Phase
Cell Cycle Checkpoints	Cell Cycle Checkpoints	Cell Cycle Checkpoints	Cell Cycle Checkpoints	IL4-mediated signaling events	Direct p53 effectors	Cell Cycle Checkpoints	superpathway of methionine degradation	superpathway of methionine degradation	Cell Cycle Checkpoints
Mitotic G1-G1/S phases	Mitotic G1-G1/S phases	Mitotic G1-G1/S phases	Mitotic G1-G1/S phases	Glucagon signaling in metabolic regulation	Plasmalogen biosynthesis	Mitotic G1-G1/S phases	choline biosynthesis III	Regulation of cytoplasmic and nuclear SMAD2/3 signaling	Mitotic G1-G1/S phases
Olfactory Signaling Pathway	Olfactory Signaling Pathway	Olfactory Signaling Pathway	Olfactory Signaling Pathway	Ion channel transport	urea cycle	Olfactory Signaling Pathway	Proteoglycan syndecan-mediated signaling events	TGF-beta receptor signaling	Olfactory Signaling Pathway
Synthesis of DNA	Synthesis of DNA	Synthesis of DNA	Synthesis of DNA	Pyruvate metabolism and Citric Acid (TCA) cycle	L1CAM interactions	Synthesis of DNA	Endothelins	Regulation of nuclear SMAD2/3 signaling	Synthesis of DNA
M/G1 Transition	DNA Replication Pre-Initiation	DNA Replication Pre-Initiation	DNA Replication Pre-Initiation	Mitotic G1-G1/S phases	Transport of Mature mRNA derived from an Intron-Containing Transcript	DNA Replication Pre-Initiation	Glypican pathway	ALK1 pathway	DNA Replication Pre-Initiation
DNA Replication Pre-Initiation	M/G1 Transition	M/G1 Transition	M/G1 Transition	FOXA2 and FOXA3 transcription factor networks	G alpha (12/13) signalling events	M/G1 Transition	Glypican 1 network	Endothelins	M/G1 Transition
G1/S Transition	G1/S Transition	G1/S Transition	G1/S Transition	AP-1 transcription factor network	Transport of Mature Transcript to Cytoplasm	G1/S Transition	Nectin adhesion pathway	PDGF receptor signaling network	G1/S Transition
Cyclin A:Cdk2-associated events at S phase entry	Cyclin A:Cdk2-associated events at S phase entry	Cyclin A:Cdk2-associated events at S phase entry	Cyclin A:Cdk2-associated events at S phase entry	Epithelial-to-mesenchymal transition	ATR signaling pathway	Cyclin A:Cdk2-associated events at S phase entry	PDGF receptor signaling network	Alpha9 beta1 integrin signaling events	Cyclin A:Cdk2-associated events at S phase entry
GPCR downstream signaling	GPCR downstream signaling	GPCR downstream signaling	GPCR downstream signaling	Syndecan-4-mediated signaling events	tryptophan degradation III (eukaryotic)	GPCR downstream signaling	Signaling events mediated by Hepatocyte Growth Factor Receptor (c-Met)	IFN-gamma pathway	GPCR downstream signaling

Table 6.1- Time course gene enrichment analysis of cardiomyocyte differentiation. Top 10 biological pathways ($P<0.05$) associated with expressed genes identified through throughout iPS cell differentiation to cardiomyocytes. Gene expression over 20 days of differentiation shown. Comparison of pathways identified reveals further similarities.

Cancer genes, namely tumour suppressor genes, oncogenes and proto-oncogenes are most likely to be involved in insertional mutagenesis subsequent to LV transduction, we next investigated the time course expression of these genes in cardiomyocyte development. Gene ontology analysis identified a range of cancer genes expressed throughout cardiomyocyte differentiation (Figure 6.1). Between 56 to 171 tumour suppressor genes are identified between day 0 and 20 of cardiomyocyte development. Similarly, between 5- 224 protooncogenes are identified and 0-12 oncogenes throughout cardiomyocyte development from iPSC. This identifies differences in cancer gene expression throughout differentiation.

While cancer genes represent a small proportion of the total genes expressed, this differs throughout cardiomyocyte development (Figure 6.1). Tumour suppressor genes represent between 0.68-0.81% of all genes between the various sampling time points. Protooncogenes represent between 0.02-1.04% and oncogenes represent 0-0.06% of all genes identified at various timepoints. While cancer related genes do not represent a significant proportion of all genes expressed throughout development, it is necessary to understand the time course expression of these genes due to their potential to contribute towards insertional mutagenesis mediated oncogenesis.

A



B

% as proportion of total genes				
Day	**Total genes**	**Tumour suppressor gene**	**Protooncogene**	**Oncogene**
0	21536	0.79	1.04	0.06
1	25121	0.68	0.89	0.05
3	25121	0.68	0.89	0.05
5	25121	0.68	0.02	0.00
7	20317	0.80	1.01	0.05
9	7640	0.73	0.88	0.03
12	25121	0.68	0.89	0.05
15	11541	0.76	0.94	0.03
18	14123	0.81	0.95	0.04
20	25121	0.68	0.89	0.05

Figure 6.1- Ontologies of expressed genes throughout cardiomyocyte development. A- Analysis of cancer specific genes reveal varying quantities expressed throughout cardiomyocytes differentiation from iPS cells. B- Percentage of cancer genes as a proportion of total genes expressed at various stages reveal subtle differences throughout cardiomyocyte differentiation.

Individual classes of cancer genes represent a proportion of total cancer genes identified throughout development (Figure 6.2). Tumour suppressor genes, protooncogenes and oncogenes represent between 42.01-97.16%, 2.84-55.04% and 0-2.95% of all cancer genes identified, respectively. This identifies more significant differences in cancer genes identified at each timepoint throughout cardiomyocyte development.

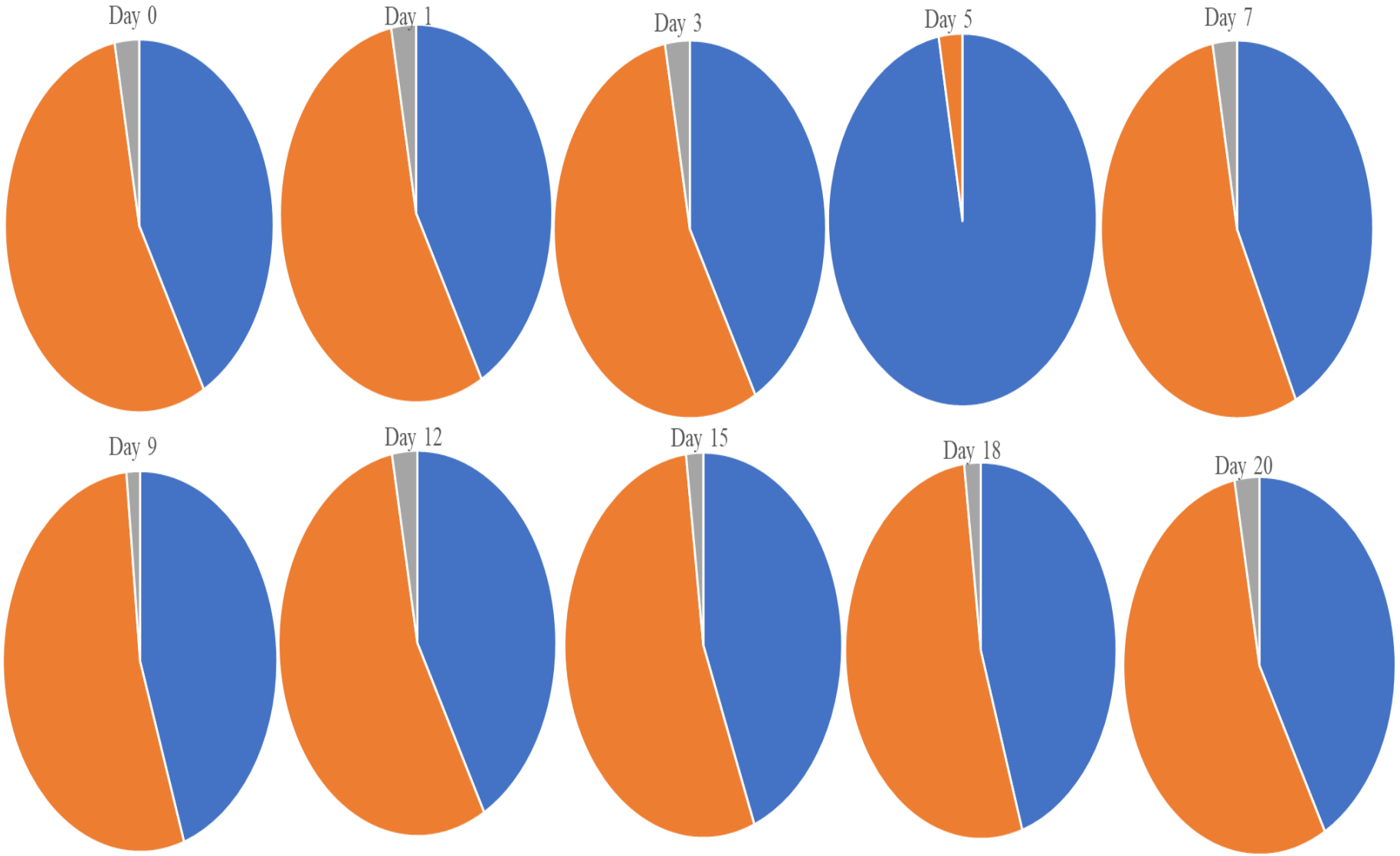


Figure 6.2- Proportion of cancer genes subtypes identified through cardiomyocyte differentiation. Subtypes of cancer genes, tumour suppressor genes (blue), protooncogenes (orange) and oncogenes (grey) identified as expressed at timepoint throughout cardiomyocyte differentiation, shown as a percentage of total cancer genes expressed in each stage.

The majority of cancer related genes are seen to be shared throughout cardiomyocyte differentiation. Comparison of tumour suppressor genes between day 0 and various other timepoints indicate the majority of genes are present at the iPSC stage and throughout development. All 171 genes (100%) are shared between day 0 and genes expressed at days 1, 3, 5, 12 and 20. Comparison between tumour suppressor genes expressed between day 0 and day 7 show 163 genes (95.3%) are shared, with 8 genes (4.7%) uniquely expressed at the iPSC stage. Similarly, 56 tumour suppressor genes (32.7%) are expressed at day 0 and day 9, with 115 genes (67.3%) expressed uniquely at the iPSC stage. 88 tumour suppressor genes (51.5%) are expressed both at the iPSC stage and day 15, with 83 genes (48.5%) expressed only at the iPSC stage. 114 genes (66.7%) are shared between days 0 and 18, with 57 genes (33.3%) only expressed at the iPSC stage.

Similar analysis of protooncogenes reveal the majority of these genes are expressed at the iPSC stage. 224 protooncogenes (100%) are expressed both at day 0 and day 1, 3, 12 and 20. Comparison of protooncogenes expressed at day 0 and day 5 reveal 5 genes (2.2%) are shared but 219 genes (97.8%) are uniquely expressed in iPSC. Similarly, 206 protooncogenes (92.0%) are expressed both at day 0 and day 7 with 18 genes (8%) expressed uniquely in iPSC. 67 of these genes (29.9%) are expressed both at day 0 and day 9, with 157 genes (70.1%) expressed uniquely at day 0. Comparison of protooncogenes expressed at day 0 and day 15 indicate 108 genes (48.2%) are commonly expressed at both stages, with 116 genes (51.8%) only expressed at day 0. 134 protooncogenes (59.8%) are expressed in both iPSC and day 18 of cardiomyocyte differentiation, with 90 genes (40.2%) uniquely expressed at the iPSC stage.

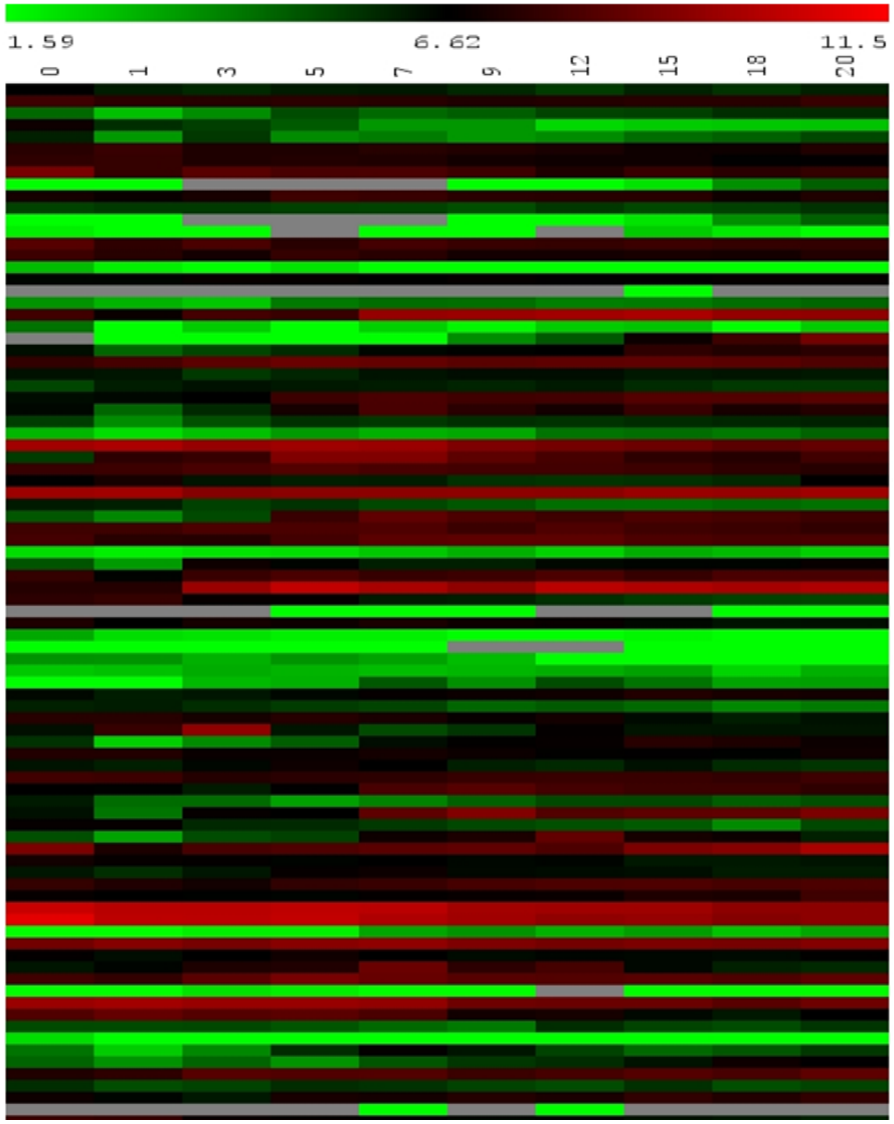
Comparison of oncogenes expressed through cardiomyocyte differentiation indicate genes commonly expressed between various stages. 12 oncogenes

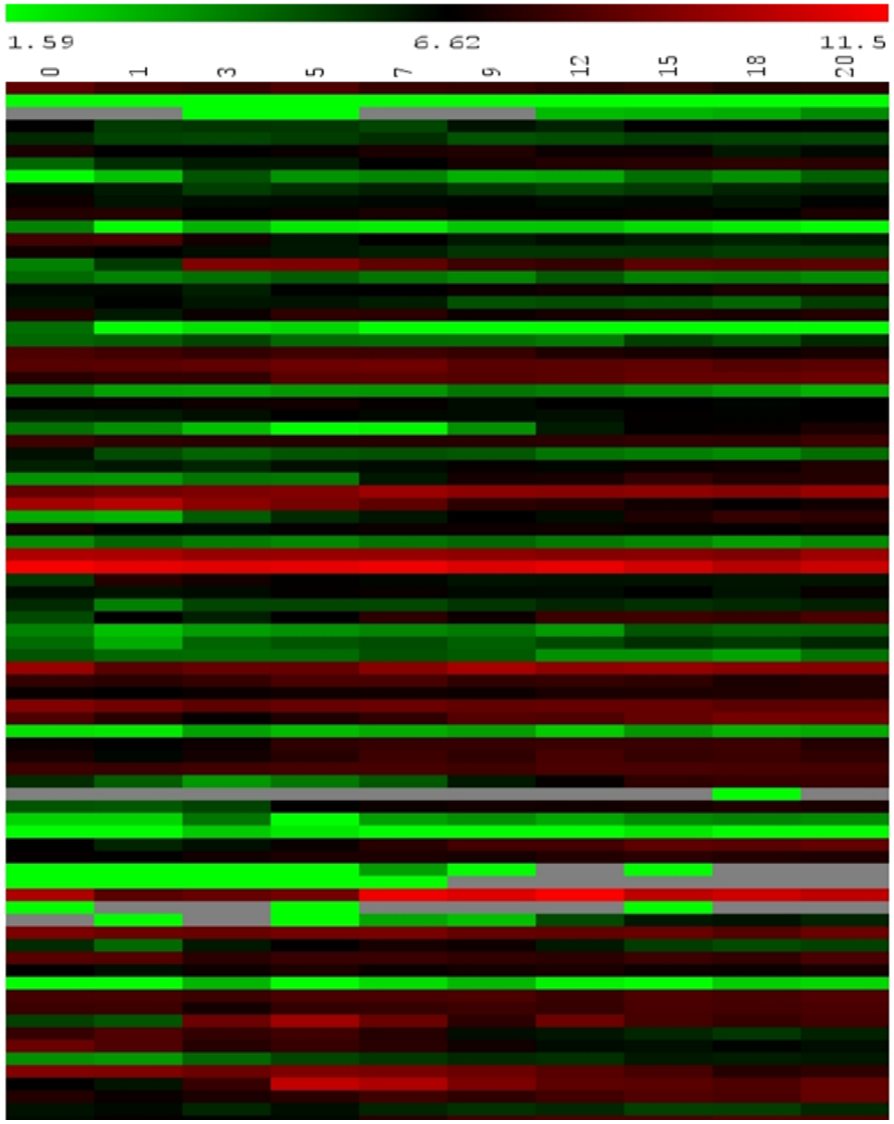
(100%) are expressed between day 0 and days 1, 3, 12, 15 and 20. 11 oncogenes (91.7%) are expressed both at day 0 and day 7, with 1 oncogene (8.3%) only identified in iPSC. Comparison between oncogenes expressed at day 0 and day 9 indicate 2 (16.7%) genes are commonly expressed with 10 (83.3%) oncogenes only expressed in iPSC. 4 oncogenes (33.3%) are commonly expressed at both day 0 and 15, with 8 genes (66.7%) only expressed in iPSC. Similarly, analysis between oncogenes expressed at day 0 and 18 reveal 5 (41.7%) common genes expressed at both timepoints and 7 (58.3%) oncogenes uniquely expressed in iPSC. This comparative analysis of cancer related genes indicates slight differences in genes expressed throughout differentiation, with the greatest number of common genes expressed at both the iPSC and terminally differentiated cardiomyocyte stages.

While the quantity of genes expressed, in particular oncogenes, protooncogenes and tumour suppressor genes, are similar between various stages of differentiation, the level of gene expression differs (Figure 6.3). Time course analysis of gene expression of cancer related genes indicates a wide range of differences between stages. As the greatest quantity of cancer related genes are expressed at the iPSC and mature cardiomyocyte stages, fold changes in expression between these two stages was analysed to determine genes that were upregulated or downregulated at the iPSC or cardiomyocyte stage. This identifies 89 (53.3%) and 78 (46.7%) tumour suppressor genes which are downregulated and upregulated in iPSCs compared to cardiomyocytes, respectively. Similar comparative expression analysis for protooncogenes determine 100 (47.2%) and 112 (52.8%) of genes are downregulated and upregulated in iPSC, respectively. 6 (50%) oncogenes are downregulated, and 6 (50%) oncogenes are upregulated in iPSC in comparison to genes expressed at the cardiomyocyte stage. These results are found to be vice versa when analysing gene expression in cardiomyocytes in comparison to iPSCs. This

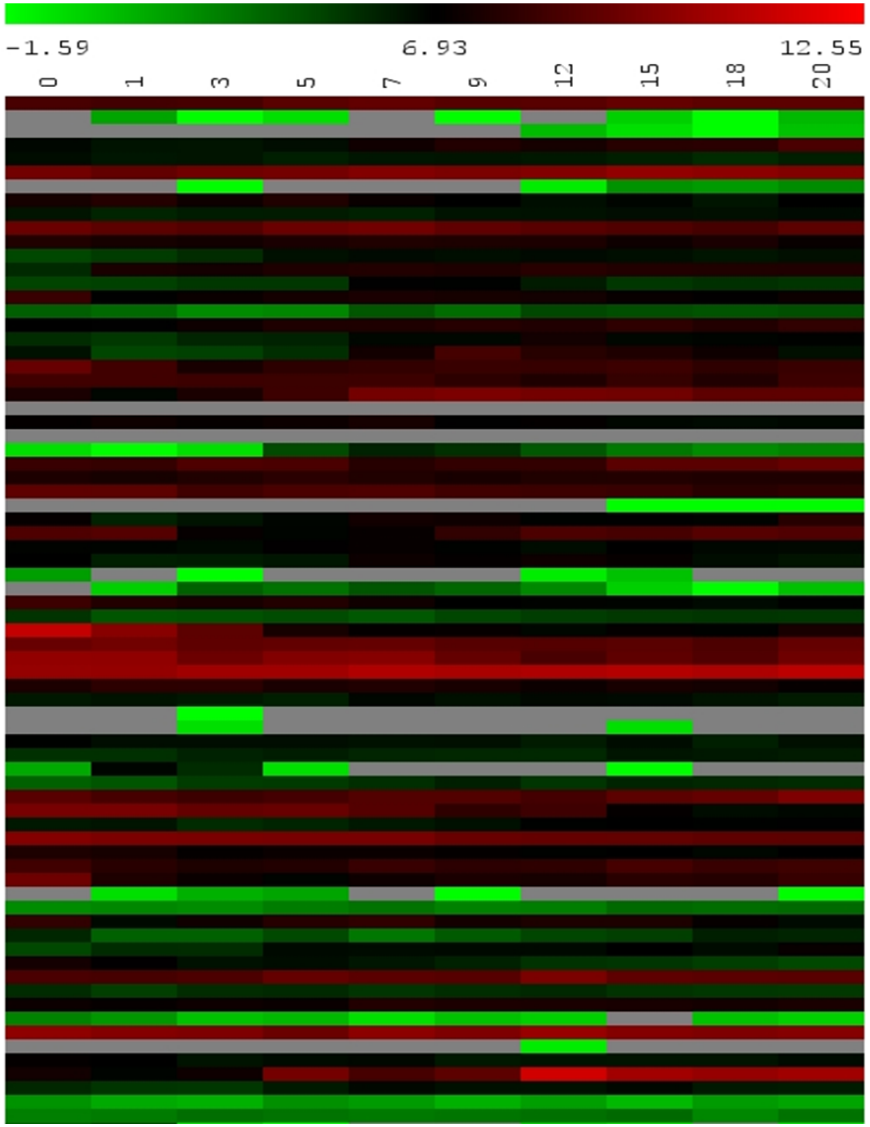
identifies differences in gene expression of cancer related genes which may be exploited by LV upon transduction of cells suggesting the need for analysis of LV transduction at both iPSC and terminally differentiated cardiomyocyte stages to capture a wide array of expressed genes.

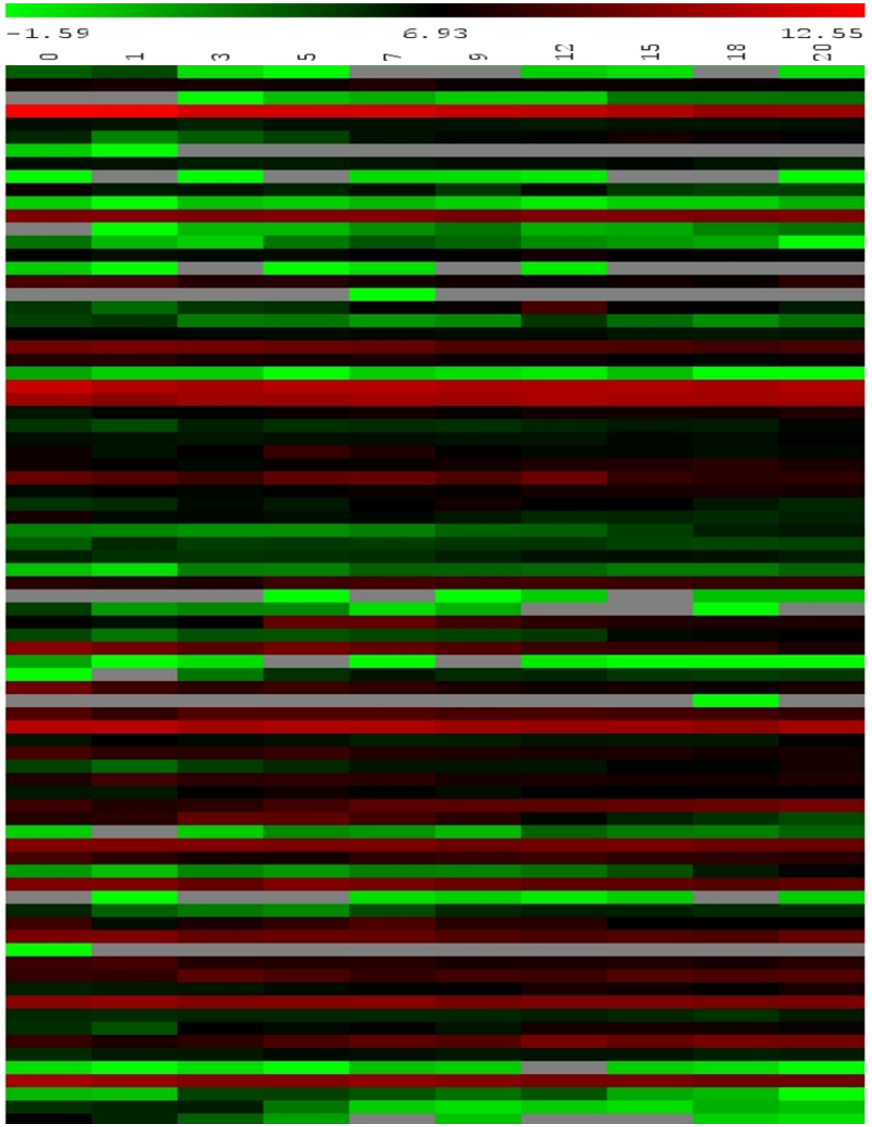
A

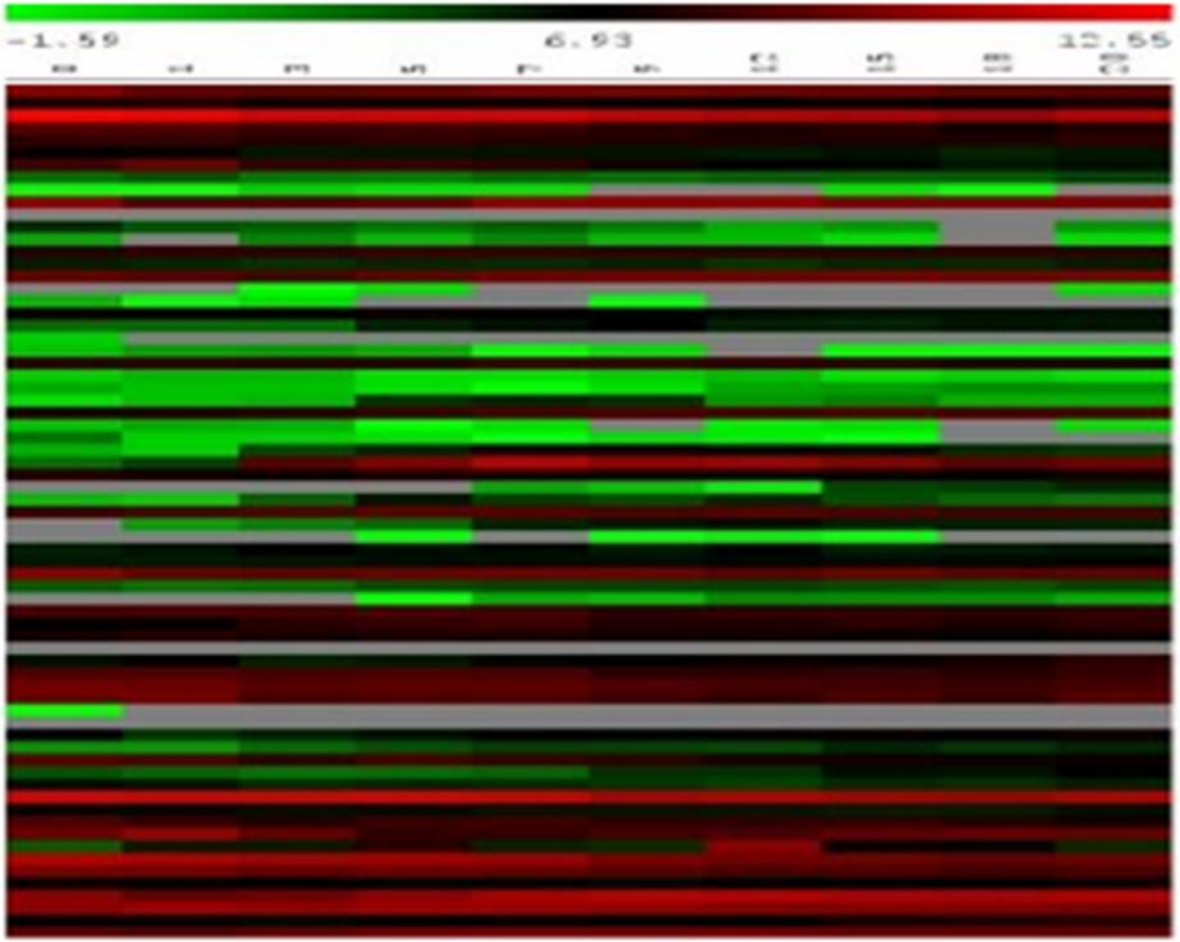




B






-1.59
6.93
15.55

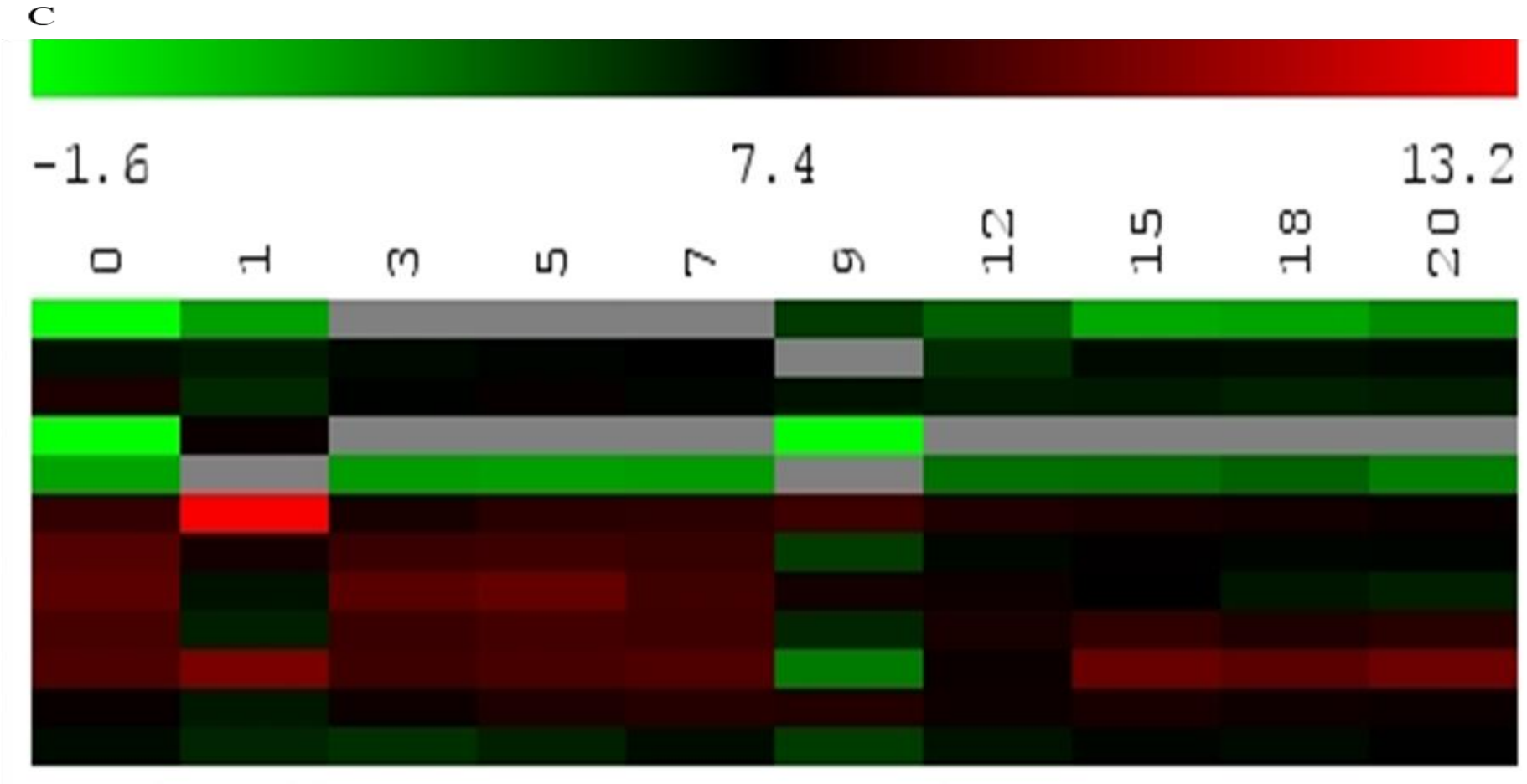


Figure 6.3- Time course gene expression throughout cardiomyocyte differentiation. Heat maps indicate expression array of highly (red) and reduced (green) expression of tumour suppressor genes (A), protooncogenes (B) and oncogenes (C) identified throughout time course cardiomyocyte development. Scale bar of gene expression and day of analysis shown. Unidentified genes shown in grey. Expression throughout Day 0-20 based on Log_2 values

To further assess differences in gene expression, cancer related genes found to be upregulated or downregulated were enriched for overrepresented biological pathways. A comparison of the top 10 significant biological pathways ($P<0.05$) associated with upregulated genes in both cell types indicate that no common pathways were found in tumour suppressor genes or oncogenes. 3 out of 10 (30%) of pathways identified in upregulated protooncogenes are shared between iPSC and cardiomyocytes. The same is identified for pathways enriched for downregulated protooncogenes. This further illustrates differences between cancer gene expression at the iPSC and cardiomyocyte stages which may be targeted by LV.

These analyses suggest the need to assay LV transduction at both the iPSC and terminally differentiated cardiomyocyte timepoints. A varying number of genes are expressed throughout cardiomyocyte differentiation. While these sets of genes are commonly expressed between iPSC and cardiomyocytes, which are enriched for common biological pathways, the expression of genes alter throughout differentiation. Specifically, tumour suppressor genes, oncogenes and protooncogenes are identified as expressed in the greatest quantities at the iPSC and cardiomyocyte stages. The levels of expression of these genes differ at both the iPSC and cardiomyocyte cell types. Further, these genes are enriched for different biological pathways. This illustrates differences at both the iPSC and cardiomyocyte cell types. As different genes and pathways are activated at the iPSC and cardiomyocyte cell stages, the LV may integrate within various genes and manipulate pathways depending on the time of infection. Therefore, it is necessary to assess LV integration at both the immature and mature cardiomyocyte cell types to provide a thorough readout of tissue specific genotoxicity of LV mediated gene therapy.

6.2.2. Comparative Gene Expression Analysis of iPSC Derived Cardiomyocyte and HLC

HLC IS has been previously analysed. To determine potential tissue specific IS differences between cardiomyocyte and HLC, gene expression profiles between cell types were compared.

Comparison of upregulated genes in the cardiomyocyte and HLC stages found only 25% (,9750 out of 39,018) genes shared. The profile of genes potentially involved in insertional mutagenesis remains more consistent. 52% (204 out of 396) of all upregulated cancer genes identified are common between HLCs and cardiomyocytes. 47% (75 out of 160), 58% (7 out of 12) and 55% (126 out of 228) of tumour suppressor genes, oncogenes and protooncogenes respectively are upregulated in both terminally differentiated cell types. Comparison of enriched genes find that 70% (all cancer genes), 70% (tumour suppressor genes), 50% (protooncogenes) and 40-50% (oncogenes) of biological pathways are common between upregulated genes in HLCs and cardiomyocytes. The difference on a global level and subtle differences between expression of cancer related genes similarities suggest cardiomyocyte specific IS related genotoxicity should be understood further.

6.2.3. P106 iPS Cell Culture Using mTeSR™1

P106 iPS cells were cultured as previously described (Section 2.2.2.16). These cells were obtained from the WiCell stem cell bank, deposited from the laboratory of Dr Lewis Becker, John Hopkins University. The cells were derived from the cord blood of a healthy 28-year-old, Caucasian male. These cells were reprogrammed using a non-integrative episomal plasmid method, encoding OCT4, SOX2, KLF4, c- MYC and Bcl-xl (Chou *et al.*, 2015). Briefly,

cells were grown in feeder and xenofree mTeSR1 medium adhering to plates coated with 2.5-5 µg laminin-521 matrix proteins per well. These conditions are ideal for iPS cell growth, allowing cells to form tight colonies while maintaining adherence and pluripotency (Lu *et al.*, 2014). P106 cells were grown and spontaneous differentiation cleared manually, as previously described. Upon microscopy analysis, pluripotent cells morphologically formed dense colonies with defined, tight borders and a large nucleus (Figure 6.4). However, contaminating differentiated cells are less dense without a tight border. These cells appear to replicate at a faster rate than P106 cells, thus overtaking the culture and reducing pluripotency potential. This indicates the need to remove differentiated cells. These cells were cleared as previously described, through manual marking of colonies subsequent to microscopy analysis and removing cells through manual aspiration. Alternatively, ReLeSR™ can be used to detach pluripotent cells from matrix using a selective, enzyme free technique. Cells were serially passaged over time when 70-90% confluency as previously described, for a minimum of 5 passages prior to use after recovering frozen cells.

.

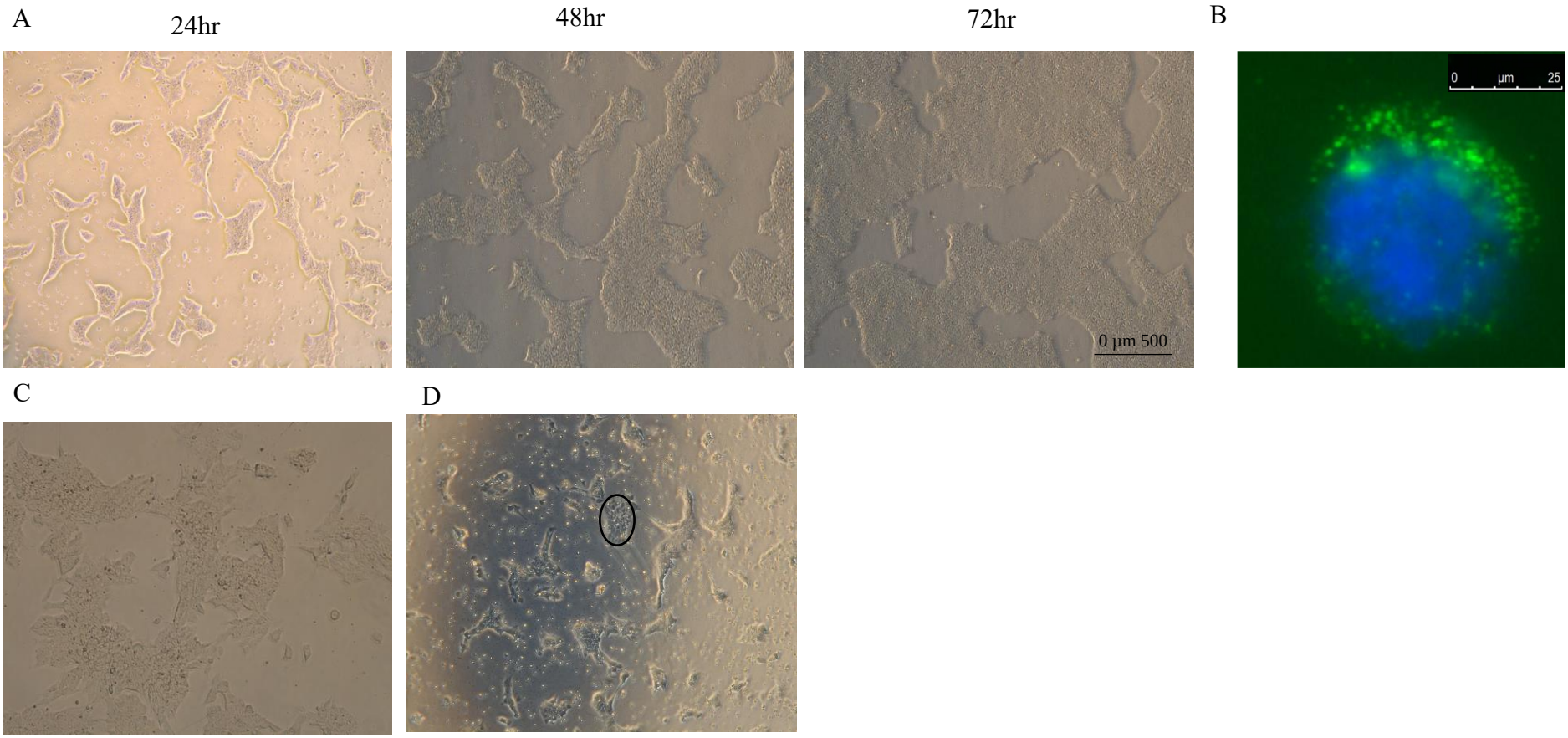


Figure 6.4- Morphological and immunohistochemistry analysis of P106 cell line. A- Brightfield microscopy of P106 cell culture over 72 hours, post seeding, showing tight colonies and morphology of pluripotent cells required for differentiation and downstream assays (x5 magnification, scale bar shown between 0 and 500µm). B- Cells were characterised morphologically and using pluripotent cell markers. Cells were stained using TRA-1-60, indicating protein expression and localisation at the periphery of cell nuclei, as expected (x100 magnification). C- Higher magnification brightfield microscopy analysis of P106 cells indicate correct morphological characteristics, with dense colonies and defined borders (x20 magnification, scale bar shown between 0 and 500µm). D- Impure colonies of iPS cells, shown after passaging cells. Highlighted region show differentiated cells, morphologically different to iPS cells (x5 magnification, scale bar shown between 0 and 500µm)

6.2.4. Characterisation of P106 pluripotent cell line

P106 iPS cells were fully characterised as pluripotent by colleagues prior to expansion and banking (Figure 6.4). Cells were analysed periodically for pluripotency using antibodies against pluripotent cell markers prior to flow cytometry quantification or immunohistochemistry (Sections 2.2.2.20). P106 iPSCs stained for TRA-1-60 (1:1000, n=4) and TRA-1-81 (1:1000, n=4) and SSEA4 (1:100, n=2) presented 88.79±3.96%, 87.60±1.74% and 96.25±3.75% expression, respectively. Due to the nature of iPSCs to spontaneously differentiate, cells were analysed for CD15 (1:25, n=2) expression as a marker for differentiation. Cells were quantified to express 13.08±3.23% after CD15 staining. Immunohistochemistry staining against TRA-1-60 (1:200) show concentration of these proteins around the periphery of the nucleus, as expected. This indicates the pluripotent state of P106 cells, through high expression of proteins known to be pluripotent markers and low expression of CD15, as a marker for differentiation. These results characterise the pluripotency and undifferentiated state of P106 iPS cell line indicating their potential for downstream differentiation.

6.2.5. Seeding Density Optimisation for 2D cardiomyocyte differentiation

Differentiation of iPSC to cardiomyocytes requires ≥95% confluent cells. Confluent cells are required for differentiation to increase stemness gene expression and maximise cardiac specific protein expression (Hamad *et al.*, 2019; Abo-Aziza and Zaki, 2017). Cell seeding density was assayed to determine the optimal concentration for seeding (Figure 6.5). In a 6 well plate, $7x10^5$, $9x10^5$ and $1x10^6$ P106 cells were seeded with 10 µM/ml ROCK inhibitor and grown at 37°C, supplemented with 5% CO_2. After 48 hours, it was observed

at wells were at 70%, 80% and 95% confluency, respectively. A similar protocol was performed using a 12 well plate format using $1x10^5$, $2x10^5$, $3x10^5$, $4x10^5$, $8x10^5$, $1x10^6$ cells. After 48 hours, wells showed 40%, 60%, 70%, 95%, 100% and 100% confluency, respectively. Seeding $1x10^6$ and $4x10^5$ P106 iPS cells were determined as the optimal density in 6 well and 12 well plate formats respectively, to achieve ≥95% confluency post 48 hours growth, in preparation for differentiation.

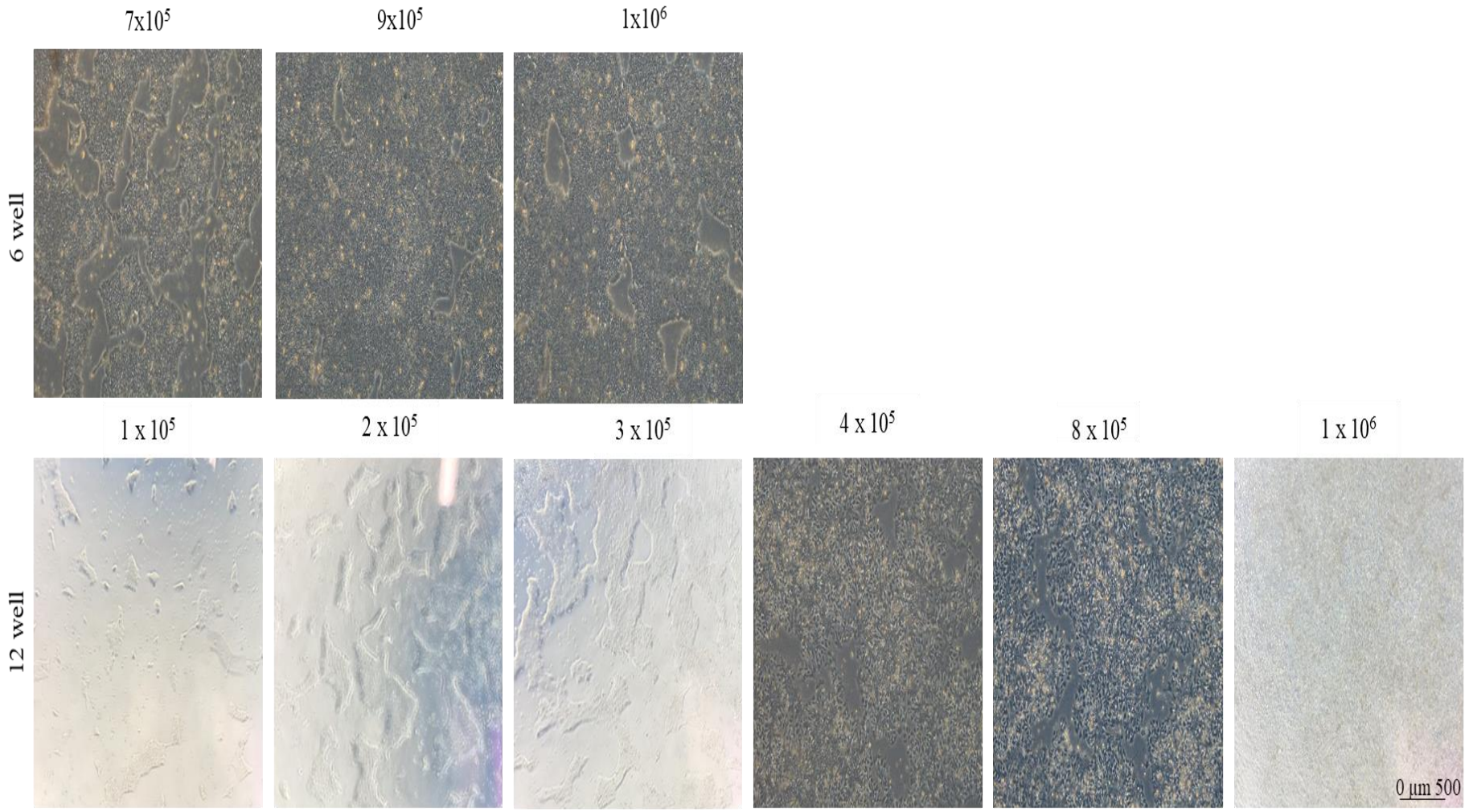


Figure 6.5- P106 cell seeding density assay. Brightfield microscopy analysis of various densities of P106 iPS cells were seeded in a 6 well and 12 well format and grown for 48 hours before analysis of confluency. $1x10^6$ and $4 x 10^5$ were determined as ~>90% confluent and the optimal cell seeding density required to each prior to commencing differentiation (x10 magnification, scale bar shown between 0 and 500µm)

5.3.5. Cardiomyocyte differentiation and characterisation

iPSCs are a promising source of terminally differentiated cells, with multiple protocols developed for stepwise protocols for cardiomyocyte differentiation. Stepwise differentiation of P106 iPSCs was performed over 15 days to mesodermal cells, cardiac mesoderm, cardiac progenitors, and mature cardiomyocytes (Section 2.2.2.17). This protocol produced a high yield of cTnT expressing cardiomyocytes with a stable excitation profile in an optimised, time and cost-efficient manner (Figure 5.6), as previously reported (Holliday *et al.*, 2018; Poulin *et al.*, 2019). P106 iPS cells were seeded at $1x10^6$ and $4x10^5$ using a 6- and 12- well plate, respectively. >95% confluent cells were differentiated step wise using STEMdiff™ cardiomyocyte differentiation kit into cardiomyocytes, with the most abundant subtype as atrial cardiomyocytes.

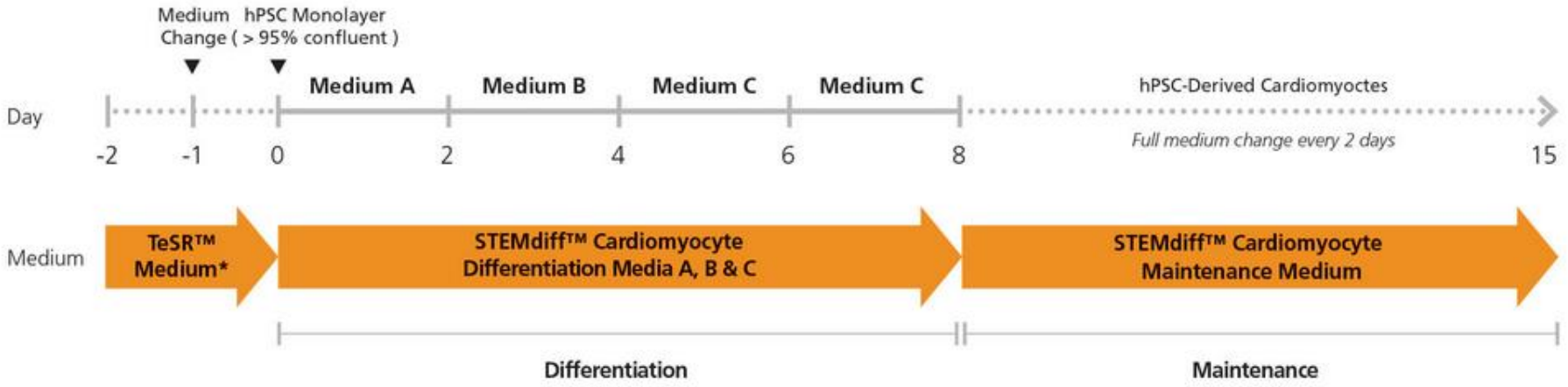


Figure 6.6- Cardiomyocyte differentiation using STEMdiff™ cardiomyocyte differentiation kit. 48 hours post seeding in mTeSR™1 , pluripotent stem cells are sequentially exposed to various growth factors. On day 0, medium is replaced with STEMdiff™ cardiomyocyte differentiation basal medium containing supplement A, commencing cardiomyocyte differentiation. On day 2, medium is replaced STEMdiff™ cardiomyocyte differentiation basal medium containing supplement B. Medium changes using STEMdiff™ cardiomyocyte differentiation basal medium containing supplement C were replaced on days 4 and 6. On day 8, medium is replaced with STEMdiff™ cardiomyocyte maintenance medium, replaced every 2 days to promote further cardiomyocyte differentiation with terminal cardiomyocytes derived at day 15 post differentiation and maintenance. Adapted from Stemcell Technologies

These terminally derived cells are characterised as cardiomyocytes for use in downstream assays. Cells were morphologically assessed using brightfield microscopy as morphologically to determine formation of syncytia. Upon examination, all stages throughout stepwise differentiation adhered morphologically to manufacturer's instructions (Figure 6.7). The development of syncytia throughout differentiation suggests the formation of cardiomyocytes.

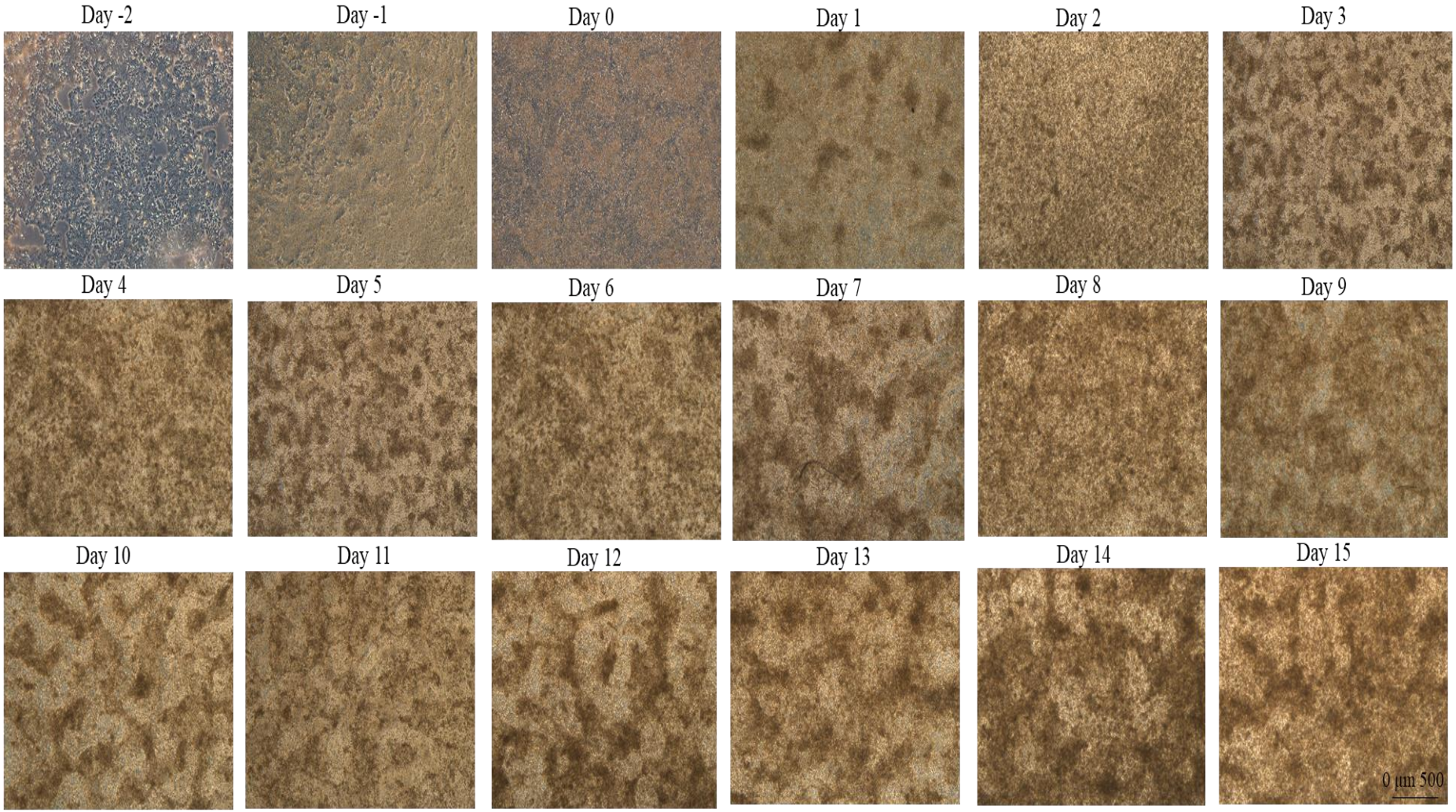


Figure 6.7- Morphology of 2D cardiomyocyte differentiation from P106 iPS cell line. Brightfield microscopy analysis of step wise and time of P106 iPS cells using STEMdiff™ cardiomyocyte differentiation kit. Analysis indicates morphological changes over time and increasing formation of syncytia. Clusters of beating cells were observed from day 8 post commencement of differentiation, with an increase of beating clusters observed until day 15 (x5 magnification, scale bar shown between 0 and 500µm)

Contractility of cardiomyocytes is an essential characteristic of these cells. Cardiomyocytes derived from P106 cells exhibited contractile function from day 8 post differentiation protocol, with increasing contractility observed until day 15. Contractility was analysed using Pulse Video Analysis (v3.0), with .mp4 video outputs uploaded and analysed using proprietary algorithms (Section 2.1.4.4). Analysis of terminally differentiated cardiomyocytes determine successful and standardised beating to cardiomyocytes across samples. The average beat rate was determined as 32.25±5.25 bpm, with the maximum rate as 62.25±15.75 bpm and minimum rate as 16±2 bpm. Velocity of contraction was quantified as 232.16±88.99 pixels/second with contraction time lasting an average of 0.71±0.24 and relaxation time as 0.9±0.21 seconds. Minimal variation between samples was determined, with 22.78±9.64 beat rate variation quantified. These indicate the successful contractility and uniformity of P106 iPSC-derived cardiomyocytes.

For further characterise these differentiated cells, cells were quantified for cardiac troponin T (cTnT) expression (Figure 6.8). cTnT is widely used as a marker for cardiomyocyte cells, with protein expression observed in cardiomyocytes. Immunohistochemistry staining of cardiomyocytes for cTnT was performed using an antibody against cTnT. Staining was performed using a 1:100, 1:500 and 1:1000 dilution of the primary antibody. cTnT were detected at all dilutions of the primary antibody, demonstrating sensitive detection of cardiac specific proteins. The optimal dilution for cell characterisation is 1:500. cTnT detected in cellular samples indicates these cells are cardiomyocytes. The identification of cTnT proteins across multiple nuclei, counterstained with DAPI, demonstrate the presence of syncitia, with multinucleated cells identified, characteristic of cardiomyocytes.

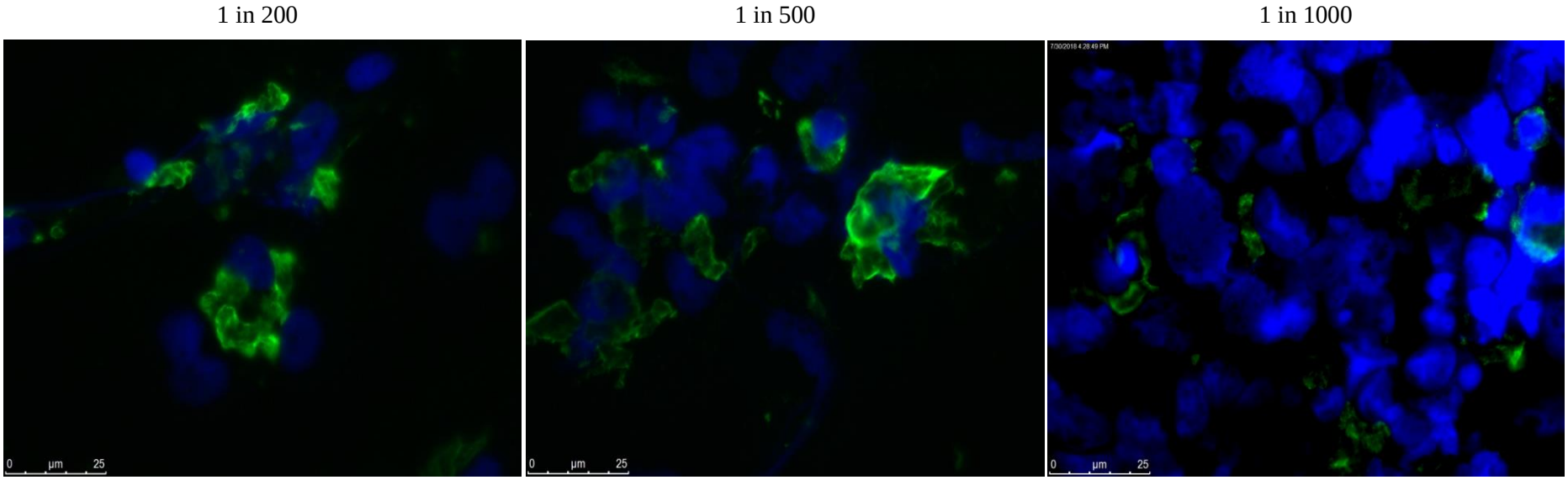


Figure 6.8- Immunofluorescence analysis of P106-derived cardiomyocytes. P106-derived cardiomyocytes were fixed with 4% PFA and stained against cTnT (green), at various dilutions of primary antibody. Optimal dilution for primary antibody use is shown as 1:500. Nuclei were counterstained using DAPI (blue), indicating the formation of multinucleated cells and syncytia, characteristic of functional cardiomyocytes (x100 magnification, scale bar shown between 0 and 25µm).

6.2.6. P106 iPSC and Derived Cardiomyocyte Transduction

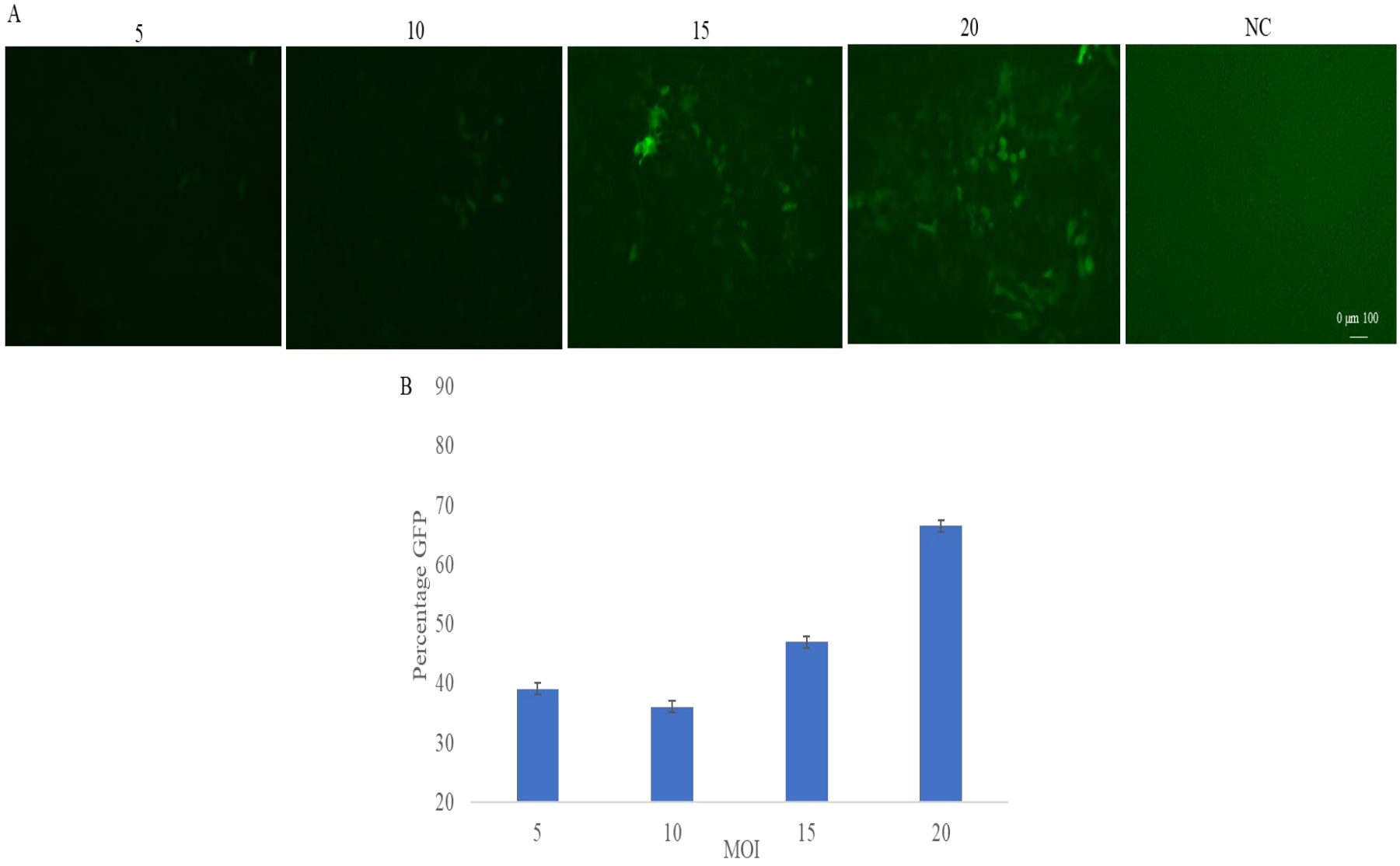
As we have characterised both undifferentiated P106 iPS cells and terminally differentiated cardiomyocytes, these cells have been verified as both pluripotent and are of cardiac lineage. Therefore, these cells can be used in downstream assays.

P106 iPS cell transduction was optimised to determine the most effective MOI for LV transduction. iPSCs were transduced with pHR LV, complexed with 5 µg/ml polybrene at a MOI of 5, 10, 15 and 20 (Figure 6.9). GFP expression analysed 72 hours post transduction determined transduction efficiencies as 39.17±9.92% (MOI 5), 36.11%±2.12 (MOI 10), 47.01%±3.23 (MOI 15) and 66.57%±3.59 (MOI 20). This dose dependent transduction efficiencies are expected with increasing LV infection.

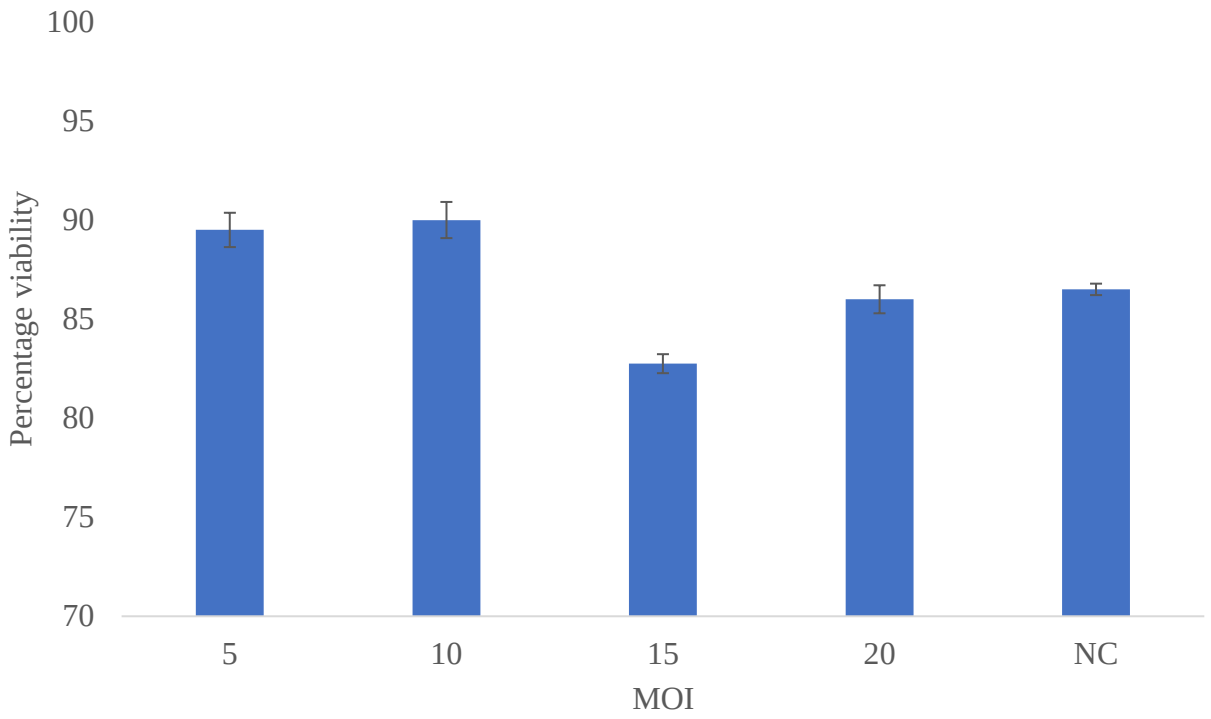
Cell toxicity was measured after LV transduction to observe potential cytotoxic effects. No morphological change and dead cells were observed suggesting low cytotoxicity. Cell viability was measured via a trypan blue exclusion assay (Section 2.2.2.3). Transduced cells were shown to have a viability of 89.5±.87% (MOI 5), 90±0.91% (MOI 10), 82.75±0.48% (MOI 15) and 86±0.71% (MOI 20) (Figure 6.10). The viability of transduced cells did not show a significant drop in viability ($p>0.05$) in comparison to uninfected cells (86.5±0.29%). This further suggests that LV transduction does not cause cytotoxicity in P106 cells

P106 iPS cell DNA was purified for downstream analysis of transduced cells. DNA was harvested from cells as previously described (Section 2.2.1.12) using DNEasy mini kit. A high quantity of gDNA was purified from cells. 95.53±14.47 ng/µl was purified from un-transduced cells. Transduction of iPSCs with a MOI of 5, 10, 15 or 20 yielded 106.07±8.93 ng/µl, 102.82±2.18 ng/µl, 115.86±3.79 ng/µl and 84.77±29.73 ng/µl.

Verification of LV transgene integration within the iPSC genomes was performed through PCR amplification of the LTR region as previously described (Section 2.2.1.17). Agarose gel electrophoresis of PCR samples determine successful integration of LV in iPSC genomes (Figure 6.9). The LTR fragment is present in P106 infected genomes with a MOI of 5, 10, 15 or 20. The presence of the β- actin housekeeping gene verifies efficiency of gDNA PCR. The presence of the β- actin fragment but absence of LTR fragment in gDNA derived from un-transduced cells demonstrates that the LTR fragment derives from LV integration, with the highest transduction efficiency observed using a MOI of 20.


A
5
10
15
20
NC
0 μm 100
B
90
80
70
60
50
40
30
20
Percentage GFP
5
10
15
20
MOI

C



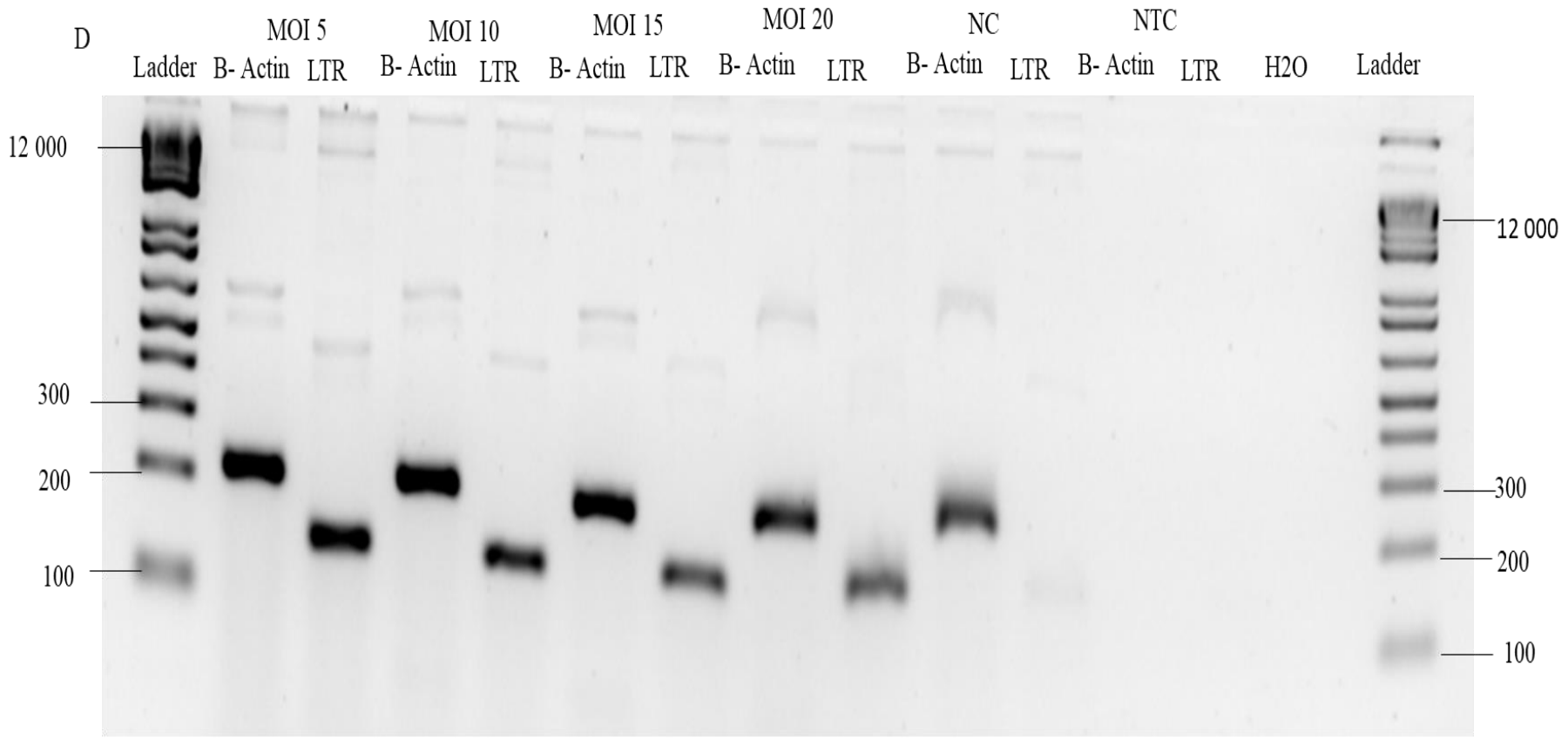
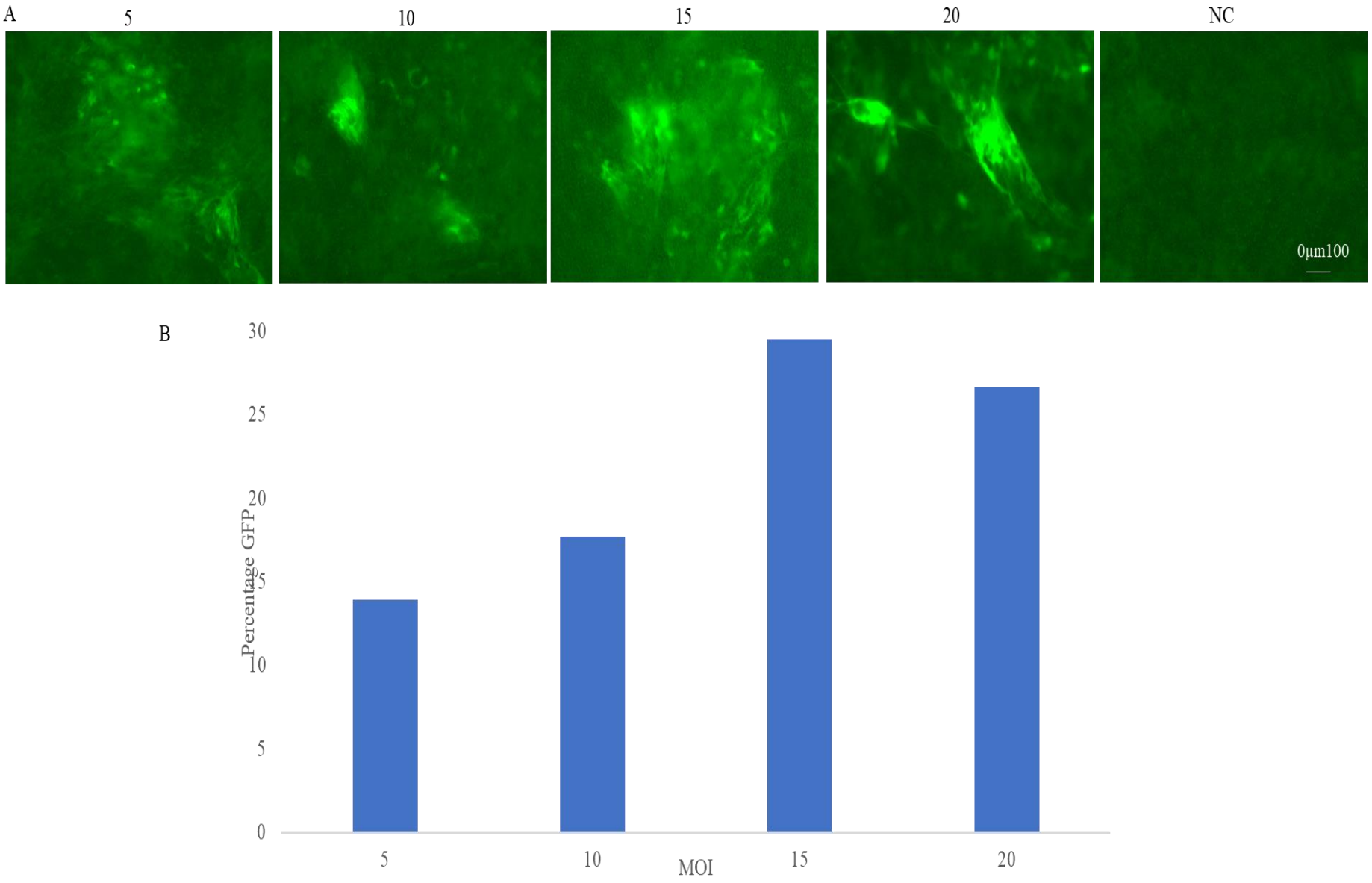
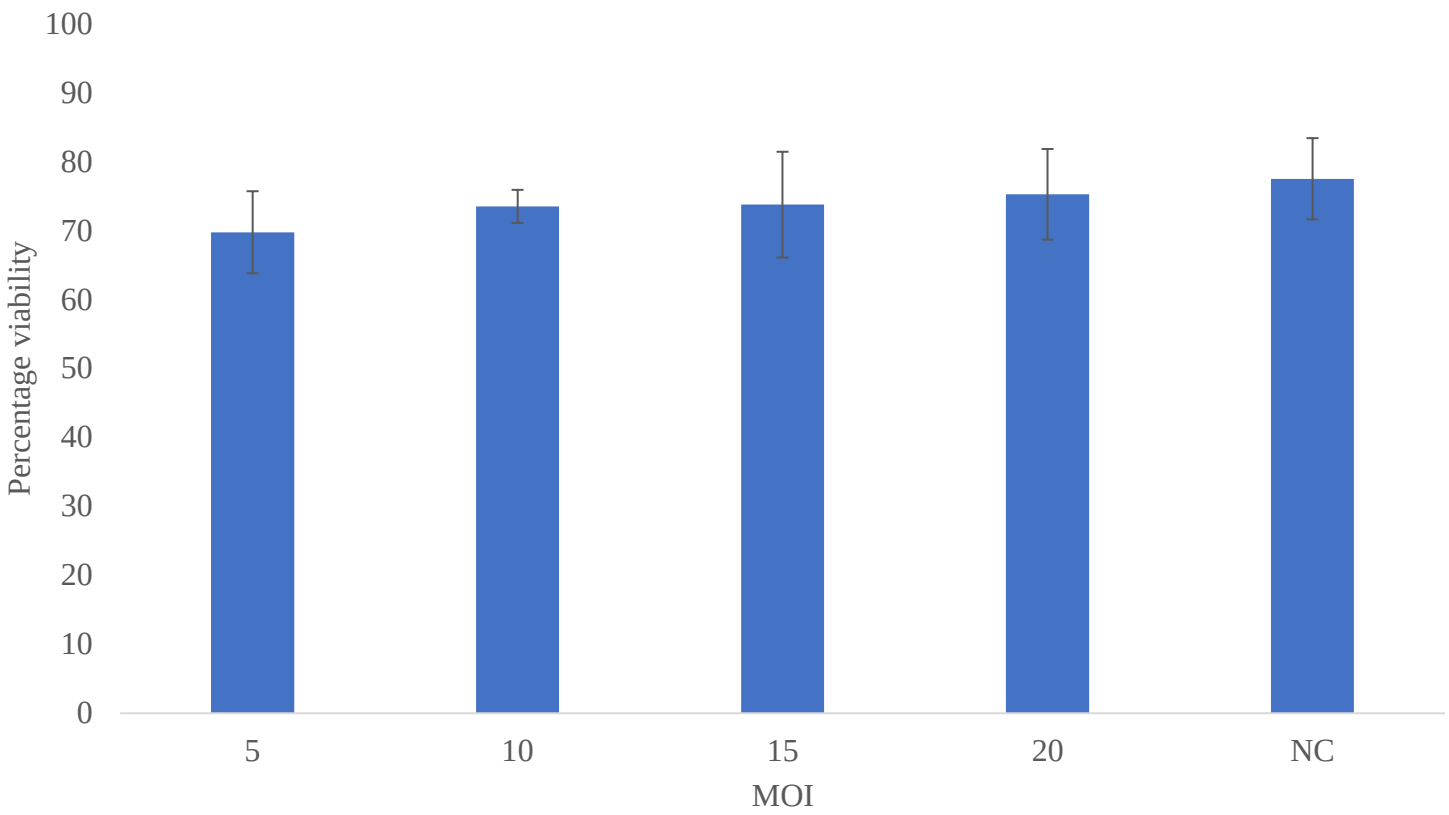


Figure 6.9- Lentiviral transduction optimisation of P106 iPS cell line. P106 cells were transduced using pHR LV with various MOIs of virus. A- Cells analysed 72hrs post transduction under fluorescent microscopy (x20 magnification, scale bar shown between 0 and 100µm), with an increase in GFP cells detected at MOI 15 & 20 transduction. B- Cells were quantified for GFP expression via flow cytometry, with dose dependent increases in GFP expression observed. Cells show as 39.17%±9.92 (5), 36.11%±2.12 (10), 47.01%±3.23 (15), 66.57%±3.59 (20), n=2. C- Cell viability after LV transduction. Transduction of iPSC at various MOI, in comparison to untreated cells (NC) show no significant drop in viability ($P>0.05$). MOI 5: 89.50±0.87%, MOI 10: 90.00±0.91%, MOI 15: 82.75±0.48%, MOI 20: 86.00±0.71%, NC: 86.50±0.29% (n=2). D- Validation of viral integration. PCR products separated on a 2% agarose gel at 70V for 40 minutes, using a 1KB+ ladder. Presence of fragment in transduced sample gDNA but absence in negative controls indicate LV integration

Transduction of P106 iPSC derived cardiomyocytes also demonstrates efficient LV infection of cells. Similar transduction of cardiomyocytes derived from iPSC at a MOI of 5, 10, 15 and 20 showed transduction efficiencies of 13.92%, 17.7%, 29.53% and 26.67%, respectively (Figure 6.10). Differences are observed between P106 iPSC and derived cardiomyocyte differentiation. While P106 cell transduction with LV is dose dependent, in derived cardiomyocytes transduction efficiency plateaus with infection at MOI 15 or 20. There is also lower transduction efficiencies in cardiomyocytes, with 25.25% (MOI 5), 18.41% (MOI 10), 17.48 (MOI 15) and 39.9% (MOI 20) greater GFP expression in P106 iPS cells.


A
5
10
15
20
NC
0μm100
B
30
25
20
5
10
5
0
Percentage GFP
5
10
MOI
15
20

C


100
90
80
70
60
50
40
30
20
10
0
Percentage viability
5
10
15
20
NC
MOI

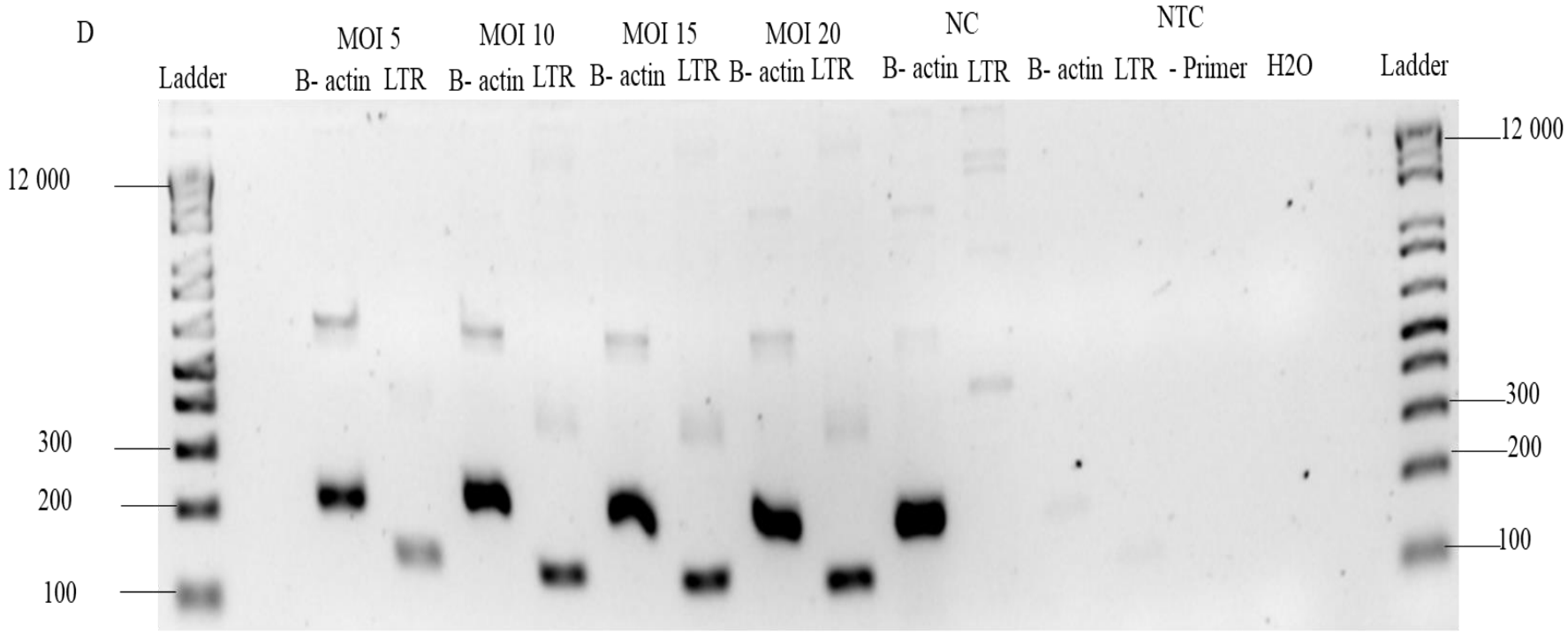


Figure 6.10- Lentiviral transduction of P106 derived cardiomyocytes optimisation. P106 derived cardiomyocytes were differentiated using STEMdiff™ cardiomyocyte differentiation kit were transduced using pHR LV with various MOIs of virus. A- Cells analysed 72hrs post transduction under fluorescent microscopy (x20, scale bar shown between 0 to 100µm) exhibited GFP expression, with an increase in expression detected after transduction using MOIs 15 & 20. B- Quantification of GFP expression in transduced P106 derived cardiomyocytes via flow cytometry as 13.92% (5), 17.70% (10), 29.53% (15), 26.67% (20). C- Viability of transduced cells using a trypan blue exclusion assay indicated no cytotoxicity in comparison to untreated cells ($P>0.05$). Viability was measured as 77.5%±5.89 (NC), 69.75%±5.95 (5), 73.5%±2.40 (10), 73.75%±7.67 (15), 73.25%±6.57 (20), n=2. D- Validation of LV integration in transduced samples through PCR amplification of LTR fragment. Presence of fragment in transduced sample gDNA but absence in negative controls indicate LV integration. Samples were run on a 2% agarose gel at 70V for 40 minutes.

Morphology of LV transduced cardiomyocytes derived from P106 cells remained unchanged. LV transduction do not appear to cause cytotoxicity in iPSC derived cardiomyocytes. The viability of transduced cells was analysed via a trypan blue exclusion assay, as previously described (Section 2.2.2.3). Transduced cells were shown to have a viability of 80±0% (MOI 5), 77±3% (MOI 10), 87±0% (MOI 15) and 86.5±1.5% (MOI 20) (Figure 6.10). The viability of transduced cells did not show a significant difference ($p>0.05$) in comparison to un-transduced cardiomyocytes (86±0%). This further suggests that LV transduction does not cause cytotoxicity in iPSC derived cardiomyocytes.

DNA purification from transduced cardiomyocytes were also harvested for use in downstream assays. gDNA purification was performed using DNEasy mini kit, as previously described. DNA yields were quantified as 50.03 ng/µl (MOI 5), 85.5 ng/µl (MOI 10), 58.23 ng/µl (MOI 15), 98.7 ng/µl (MOI 20) and 51.93 ng/µl (NC). This provided high quality gDNA for further use. However, this is between 17-58% lower than the concentration of gDNA harvested from iPSC. The lower yield of DNA may be due to the formation of syncytia in cardiomyocytes.

LV integration was also verified through PCR amplification of the LTR region and analysis via gel electrophoresis. All transduced samples were positive for the LTR fragment and β- actin housekeeping gene fragment (Figure 6.10). This indicates successful PCR amplification of gDNA and positive LV transduction. This is further verified in comparison to a negative control, where gDNA from untreated cardiomyocytes were amplified for the β- actin fragment but not the LTR fragment suggesting the LTR fragment is due to LV transduction.

Transductions in P106 iPSCs and derived cardiomyocytes indicate an MOI of 20 is the most efficient for with low cytotoxicity observed after LV infection.

These infected samples can be analysed to determine IS-mediated genotoxicity in a human based, *in vitro* iPSC-derived cardiomyocyte platform.

5.3. Discussion

iPS cells are a promising source of undifferentiated cells for use in therapeutic assays. These cells can be terminally differentiated into multiple cell types, including cardiomyocytes. These cells have shown promise in a clinical setting, with grafts of iPS cell derived cardiomyocytes into patients who suffered from cardiac arrest showing promise. Viral gene delivery to these cells has been shown as an effective technique in the treatment of various genetic diseases. While Ad and AAV have been shown to be effective transducers of iPSC derived cardiomyocytes, LV are successful vectors for gene delivery as they show comparable gene expression levels at lower MOIs, reducing cytotoxicity. However, there is a present need to understand the genotoxic side effects of LV mediated gene therapy. Animal models for gene therapy safety have been created but to date there is no published analysis of gene therapy safety to the heart in a human model. As iPS cells and cardiomyocyte derivatives recapture clinical symptoms, LV integration sites within these cells can be analysed as a model of lentiviral gene therapy safety to determine insertional mutagenesis and related oncogenesis. Time course analysis of cardiomyocyte differentiation indicates the greatest number of oncogenes, protooncogenes and tumour suppressor genes are expressed at both the iPS cell and terminally differentiated cell stage. While primary cancer of the heart is extremely rare with an incidence range of 1.38-30 per 100,000, sarcomas present the majority of these pathologies (Bonow, Mann et al. 2011, Al-Mamgani, Baartman et al. 2008). Large scale analysis of cardiac tumours indicates that *MDM2*, *CDK4*, *HMGA2*, *DDIT3*, *PDGFRA*, *EGFR* and *CDKN2A* are all involved in these tumours (Neuville, Collin et al. 2014). *HMGA2*, *DDIT3*, *CDKN2A* have been implicated in lentiviral insertional mutagenesis suggesting the safety of gene therapy to these cells must be understood further (Cavazzana-Calvo, Payen et al. 2010, Kustikova, Geiger et al. 2007).

Gene expression analysis suggests that the pluripotent and terminally differentiated cell stage would be useful timepoints to analyse LV integration as insertional mutagenesis of cancer related genes contributes towards oncogenesis. P106 iPS cells were determined as pluripotent through expression of various pluripotency markers, including TRA-1-60, TRA-1-81 and SSEA4. These markers are commonly used against pluripotency (Natunen, Satomaa et al. 2011). P106 cell staining against CD15 showed low levels of protein detection. CD15 is a common marker for differentiation, with specificity for detection of cells of neuronal lineage (Pruszak, Ludwig et al. 2009). iPS cells are important to remain in an undifferentiated state to enhance terminal differentiation capacity towards multiple cell types. These cells can also be grown in xeno and feeder free conditions. While feeder cells have been reported to maintain iPS cells, these cause complications in downstream processing (Takahashi, K., Narita et al. 2009). Xeno free conditions have shown to successfully maintain iPS cells in a pluripotent state and is important in satisfying GCP guidelines and producing a standardised human model of genotoxicity (Nakagawa, Taniguchi et al. 2014, Rodriguez-Piza, Richaud-Patin et al. 2010). Multiple protocols have been developed for highly efficient cardiomyocyte differentiation from pluripotent stem cells. cTnT characterised cardiomyocytes differentiated from iPS cells were used downstream for transduction assays. Morphological and physiological characteristics of these cardiomyocytes, including beat rate (32.25±5.25 bpm), velocity of contraction 232.16±88.99 pixels/second and duration of contraction (0.71±0.24s) were in line with published results (Del Alamo, Lemons et al. 2016).

Comparison of the gene expression of cardiomyocytes and HLC derivatives show differences in overexpressed genes. As the majority of IS are in overexpressed genes, few differences may be observed between the two pathways. P106 iPSCs were shown to successfully differentiate to

cardiomyocytes, identified morphologically and present expression of cTnT, a cardiac specific marker.

LV are commonly used in gene therapy for permanent gene transfer (Gouvarchin Ghaleh, Bolandian et al. 2020). Previously we have shown optimised high titre production of LV for use in downstream assays. These viruses also transduce replication and senescent cells, with VSV-G commonly used to pseudotyping the viral capsids. VSV-G allows infection of multiple cell types due to its affinity to the LDL receptor, a common receptor found on cellular membranes. These viruses have been found to successfully transduce iPS cells and cardiomyocyte derivatives. The optimal MOI for transduction was determined as 20, as shown through flow cytometry analysis of GFP expression. Viral integration was confirmed using PCR amplification. While transgene expression is observed in a dose dependant manner in iPS cells, this is not the case in derived cardiomyocytes, with similar expression after transduction at MOI 15 and 20.

These results indicate the successful of LV transduction of P106 cells and terminally differentiated cardiomyocytes based on gene expression analysis. Further characterisation of P106 cells and differentiated cardiomyocytes using antibody markers, karyotyping and molecular assays should be performed. These cells can be efficiently transduced with LV without observable cytotoxicity or change in morphology. As differences persist between iPSC derived HLC and cardiomyocyte gene expression, IS analysis in cardiomyocyte's should be performed to understand differences in tissue specific LV gene integration. Transduced samples can be used in downstream sequencing analysis to determine LV sites of integration to assess differences in integration sites and ontologies between LV carrying with full or SIN LTR configurations vector and determine.

Chapter VII
Discussion

Virus based vectors, including AV and LV, are commonly used as vehicles for gene transfer, with multiple products licenced for gene therapy (Wiley 2021). The rapid growth in the use of these vectors indicates a need to optimise vector production. Early large scale clinical studies require greater than 10^{15} AV particles (Cheng, Wang et al. 2015, Baden, Liu et al. 2015). However, current methods for the generation of high titre AV require fed-batch, adherent culture of cells that have been transfected or transduced to produce replication defective AV vectors (Lesch, Heikkila et al. 2015). The vectors can be obtained from cell lysates, but the current methods of purification show poor recovery yields of virus particles, result in aggregation of the virus, reducing infectivity downstream and is costly (Boychyn, Yim et al. 2001, Farid 2007, McNally, Darling et al. 2014). In contrast, ATPS employs a simpler, cost-effective system which has been shown to successfully purify biomolecules, including AV and AAV. Optimisation of these systems indicates the upper polymer rich phase of 20% - 20% w/w PEG 600 - $(NH_4)_2SO_4$, as the most successful for recovery of AV particles, including directly from cell lysates. This system shows high levels of recovery *in vitro*, in both liver and neuronal cell lines. While this process has been shown to isolate contaminating biomolecules from cell lysates in conjunction with AV vectors and should be optimised to avoid this, these contaminated systems have not had any adverse effects *in vitro*. The virus also shows comparable levels of gene transfer to CsCl purified particles *in vivo* without any adverse side effects observed. This method employs the use of a bench top centrifuge rather than expensive ultracentrifuge materials and can be used direct from the phase system, in stark contrast to CsCl density separation, which requires further dialysis to remove toxic salts. PEG 600 can be readily cleared by the kidneys and as such, is commonly used as an emulsifier in ointments and toothpastes. The concentrations used readily remain within IID limits (D'souza, Shegokar 2016). The use of PEG is also attributed to the longevity of infectious particles, even when frozen for a year at -80°C, with

PEGylation seen as an effective enhancer of AV and LV transduction (Croyle, Yu et al. 2000, Croyle, Le et al. 2005, Croyle, Callahan et al. 2004). Despite this, ATPS was not successful at recovering LV particles *in vitro* likely due to the alteration in system dynamics by biphasic formation.

While AV are largely episomal vectors and while no genotoxicity has been reported thus far, murine integration events have been reported *in vivo*, though extremely rare (Stephen, Montini et al. 2010). Similarly, effective high titre production of LV vectors is required for downstream applications. Protocols have been developed for high titre LV production, for instance using various transfection reagents, culture conditions and fixed bed bioreactors (Valkama, Leinonen et al. 2017, Besnier, Takeuchi et al. 2002). By developing an optimised protocol for LV production using Genejuice® transfection reagent, we generated at an average of $2.06x10^9$ TU/ml of LV in a small-scale adherent platform, significantly greater than that generated using PEI or Fugene® 6 transfection reagents This optimises high titre LV production as similar titres are gained from large scale fixed bed bioreactor transfection. This process can be scaled up efficiently for the majority of laboratory spaces to increase vector production over 11 days without cytotoxicity.

Polybrene is traditionally used to enhance LV transduction *in vitro* while silica nanoparticles increase pDNA transfection (Barrilleaux, Knoepfler 2011, Pearce, Mai et al. 2008, Park, Jeong et al. 2016) While polybrene does not enhance transduction in all cell types, we show that increasing concentrations of Nuvec® can increase transduction efficiencies greater than that quantified with polybrene alone (Denning, Das et al. 2013). While a dose dependent increase in toxicity is observed using Nuvec®, this is not unexpected, with the optimal concentration for both transduction efficiencies and cell health quantified as 40µg/0.5ml. This is further enhanced by dual complexation in conjunction with

polybrene. This indicates the promise of using silica nanoparticles of LV transduction enhancement.

Wild type HIV-1 particles have been sequenced commonly to characterise splice variants and transcripts present (Bagutti, Alt et al. 2011, Mattiuzzo, Ashall et al. 2015). Through sequencing HIV-1 based LV vectors using long range PACBIO methodologies, we were able to detect the presence of aberrant packaging of sequences found to have high sequence homology to other viruses (Rhoads, Au 2015). Similar taxon classes are identified in empty particles (created without a packaged transgene), albeit at a lower concentration. This is despite the majority of sequences not containing a ψ sequence. The potential for transfer of these aberrant sequences should be understood further and determine possible genotoxic outcomes.

While *in vitro* and *in vivo* models of genotoxicity have been previously published, these are largely based in animals indicating the need for a humanised platform of gene therapy safety (Montini, Cesana et al. 2009, Cesana, Ranzani et al. 2012, Jackson, Thomas 2017). The ability of iPSCs to differentiate into ectodermal, endodermal, and mesodermal tissue types has provided a platform for therapeutic use (Takahashi, Yamanaka 2006). This has particular promise in gene therapy. Gene transfer to these cells using AV and LV is an effective method to correct genetic disorders. As such, the potential for *ex vivo* gene therapy using personalised iPSC and their derivatives has never been closer. Somatic cells are typically reprogrammed using a cocktail of factors: *OCT4*, *SOX2*, *KLF4* and *c-MYC* (Takahashi, Yamanaka 2006). Using these cells derived from patients, specific disease models have been generated to further elucidate mechanisms and assess the success of various treatment regimens (Stein, Ott et al. 2010, Ott, Schmidt et al. 2006). This suggests these pluripotent cells and their derivatives can be used to sensitively analyse the

effects of LV vector mediated gene therapy. Analysis of IS identifies the majority of integration events in highly expressed genes, though this is greater in iPSC than HLC derivatives. Differences persist in integration profiles, including over time in iPSC, between cell types and vector configurations. Despite this, common biological pathways are enriched. This identifies common oncogenes, protooncogenes and tumour suppressor genes targeted by vectors carrying a native or SIN LTR configuration flanking the transgene. A proliferation assay was developed analysing IS 3- and 30- days post transduction. This identified rapidly various genes that contribute towards clonal outgrowth. This provides the first humanised safety model of LV vectors. While the SIN LTR is considered safer, the results shown here also show further modification of vectors are required to increase the safety profile of these vectors. The stability of P106 iPS cells should be determined to assess the background genetic variability *in vitro* which may impact vector integration and associated genotoxicity.

LV have been shown to preferentially integrate within active transcription units in the host DNA (Ciuffi, Llano et al. 2005, Schroder, Shinn et al. 2002, Mitchell, Beitzel et al. 2004, Rossetti, Cavarelli et al. 2013). Integration requires the formation of PICs which are chaperoned using various proteins, such as hRad1 and PSIP1/LEDGF/p75 (Cherepanov, Maertens et al. 2003, Vandegraaff, Devroe et al. 2006). However, knockout of PSIP1/LEDGF/p75 still results in preferential integration of LV transcripts in actively transcribed genes suggesting other proteins tethers the viral genome to the host (Ciuffi, Llano et al. 2005). The viral integrase has been shown as important as alternating species-specific integrase results in host IS to closely align with that of the integrase species (Modlich, Navarro et al. 2009). This is not exclusive to humans or viruses as yeast mediated bait selection of proteins indicates multiple proteins interacting with MLV integrase, including TF (Studamire, Goff 2008).

We presented here data to show that pTFBS present in the LTR closely align with pTFBS immediately surrounding IS. This is identified regardless of U3 status in LV genomes. These sites were enriched in comparison to expected pTFBS frequency in background data sets. Further, sites for TF proteins known to be involved in HIV-lifecycle (NKX3-1, PDX1, PRRX2, AP1, Gfi, HOXA5, NFATC2, SPIB and NF-κβ) were all seen to be preferentially targeted. PBX1 is known to be involved with viral transcription (Chao, Walker et al. 2003, Tacheny, Michel et al. 2012, Ma, Dong et al. 2013). ZEB1 and AP1 have been shown to be involved in HIV latency and NFATC2 is essential for productive infection of non-activated T-cells (Venkatachari, Zerbato et al. 2015, Duverger, Wolschendorf et al. 2013, Hohne, Businger et al. 2016). Also, NF-κβ and SP1 sites in the LTR have both been shown to be involved in transcription of the HIV genome (Stroud, Oltman et al. 2009, Harrich, Garcia et al. 1989). These results were closely mimicked *in vivo* in mice and human patients. The majority of these pTFBS were also identified in eight genes previously reported in insertional mutagenesis events (*LMO2*, *PRDM16*, *CCND2*, *MECOM*, *HMGA2*, *BMI1*, *BCL2*, *PRDM1*) (David, Doherty 2017). Interestingly, the additional site identified unique in the SIN LTR can be identified in all eight genes. These data suggests that LV insertion is tethered by proteins near common pTFBS to the LTR and may be useful to promote virus survival. We propose further modification to the LV LTR may be required to reduce the likelihood of integration into undesired sites to increase the safety of LV mediated gene therapy.

The liver is commonly used in toxicity analysis (MacDonald, Robertson 2009). Here, we show the efficient differentiation of P106 iPSCs to cardiomyocytes which can be transduced to a high degree using LV. Comparison of the gene expression between iPSC derived HLC and cardiomyocytes identifies

differences between the cell types suggesting the need for tissue specific toxicity platforms.

In summary, this thesis provides data to support the efficient separation of AV vectors using an optimised ATPS. This shows high recovery rates *in vitro* and *in vivo* with no observable toxicity. While these systems do not efficiently separate LV vectors, these viruses can be produced to high titres using an optimised and scalable production protocol using Genejuice ® transfection reagent. However, the sequencing of these vectors suggests the presence of aberrant viral transcripts which may be delivered upon gene transfer. The assessment of genotoxicity using LV in an iPSCs and terminally differentiated HLCs show subtle differences in integration based on the U3 configuration of the LTR. This model provides a useful out read of genotoxicity integration events in both a naïve and mature cell types which can be personalised to patients. This platform also provides evidence of pTFBS present in the LTR tethering to common sites in the host genome. This was also identified in human patients and mice *in vivo*. iPSCs can also be efficiently differentiated to cardiomyocytes which can be transduced to similar levels *in vitro*. Comparative gene expression analyses indicate subtle differences between HLC and cardiomyocytes suggesting the need for tissue specific genotoxicity platforms based on the primary tissue targeted for gene delivery.

These data provide an optimised protocol for efficient AV and LV vector production and the first human based platform for understanding LV mediated genotoxicity further. With these tools, the genotoxicity profile of these vectors can be identified. Modifications to these viruses can be made to ensure safe treatment of patients in the clinic.

Chapter VIII- References

ABYZOV, A., MARIANI, J., PALEJEV, D., ZHANG, Y., HANEY, M.S., TOMASINI, L., FERRANDINO, A.F., ROSENBERG BELMAKER, L.A., SZEKELY, A., WILSON, M., KOCABAS, A., CALIXTO, N.E., GRIGORENKO, E.L., HUTTNER, A., CHAWARSKA, K., WEISSMAN, S., URBAN, A.E., GERSTEIN, M. and VACCARINO, F.M., 2012. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature, 492(7429), pp. 438-442.

ADAMCOVA, M., SKARKOVA, V., SEIFERTOVA, J. and RUDOLF, E., 2019. Cardiac Troponins are Among Targets of Doxorubicin-Induced Cardiotoxicity in hiPCS-CMs. International journal of molecular sciences, 20(11), pp. 10.3390/ijms20112638.

ADHIKARY, S. and EILERS, M., 2005. Transcriptional regulation and transformation by Myc proteins. Nature reviews.Molecular cell biology, 6(8), pp. 635-645.

AIUTI, A., CASSANI, B., ANDOLFI, G., MIROLO, M., BIASCO, L., RECCHIA, A., URBINATI, F., VALACCA, C., SCARAMUZZA, S., AKER, M., SLAVIN, S., CAZZOLA, M., SARTORI, D., AMBROSI, A., DI SERIO, C., RONCAROLO, M.G., MAVILIO, F. and BORDIGNON, C., 2007. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. The Journal of clinical investigation, 117(8), pp. 2233-2240.

AIUTI, A., COSSU, G., DE FELIPE, P., GALLI, M.C., NARAYANAN, G., RENNER, M., STAHLBOM, A., SCHNEIDER, C.K. and VOLTZ-GIROLT, C., 2013. The committee for advanced therapies' of the European Medicines Agency reflection paper on management of clinical risks deriving from insertional mutagenesis. Human gene therapy. Clinical development, 24(2), pp. 47-54.

AKINC, A., THOMAS, M., KLIBANOV, A.M. and LANGER, R., 2005. Exploring polyethylenimine-mediated DNA transfection and the proton sponge hypothesis. The journal of gene medicine, 7(5), pp. 657-663.

ALBERTSSON, P.A., 1962. Partition of Cell Particles and Macromolecules. Fette, Seifen, Anstrichmittel, 64(8), pp. 701-702.

ALBERTSSON, P.A., ANDERSSON, B., LARSSON, C. and AKERLUND, H.E., 1982. Phase partition--a method for purification and analysis of cell organelles and membrane vesicles. Methods of biochemical analysis, 28, pp. 115-150.

ALHAQUE, S., THEMIS, M. and RASHIDI, H., 2018. Three-dimensional cell culture: from evolution to revolution. Philosophical transactions of the Royal Society of London.Series B, Biological sciences, 373(1750), pp. 10.1098/rstb.2017.0216.

ALLOWAY, J.L., 1933. Further Observations on the use of Pneumococcus Extracts in Effecting Transformation of Type in Vitro. The Journal of experimental medicine, 57(2), pp. 265-278.

ALLOWAY, J.L., 1932. The Transformation in Vitro of R Pneumococci into S Forms of Different Specific Types by the use of Filtered Pneumococcus Extracts. The Journal of experimental medicine, 55(1), pp. 91-99.

AL-MAMGANI, A., BAARTMAN, L., BAAIJENS, M., DE PREE, I., INCROCCI, L. and LEVENDAG, P.C., 2008. Cardiac metastases. International journal of clinical oncology, 14(3), pp. 369-372.

ALMARZA, D., BUSSADORI, G., NAVARRO, M., MAVILIO, F., LARCHER, F. and MURILLAS, R., 2011. Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional

read-through in keratinocytes transduced with self-inactivating lentiviral vectors. Gene therapy, 18(7), pp. 674-681.

ALTARA, R., MANCA, M., SABRA, R., EID, A.A., BOOZ, G.W. and ZOUEIN, F.A., 2016. Temporal cardiac remodeling post-myocardial infarction: dynamics and prognostic implications in personalized medicine. Heart failure reviews, 21(1), pp. 25-47.

ALTARAS, N.E., AUNINS, J.G., EVANS, R.K., KAMEN, A., KONZ, J.O. and WOLF, J.J., 2005. Production and formulation of adenovirus vectors. Advances in Biochemical Engineering/Biotechnology, 99, pp. 193-260.

AMALFITANO, A., HAUSER, M.A., HU, H., SERRA, D., BEGY, C.R. and CHAMBERLAIN, J.S., 1998. Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. Journal of virology, 72(2), pp. 926-933.

AMIRACHE, F., LEVY, C., COSTA, C., MANGEOT, P.E., TORBETT, B.E., WANG, C.X., NEGRE, D., COSSET, F.L. and VERHOEYEN, E., 2014. Mystery solved: VSV-G-LVs do not allow efficient gene transfer into unstimulated T cells, B cells, and HSCs because they lack the LDL receptor. Blood, 123(9), pp. 1422-1424.

ANDREWS, B.A., SCHMIDT, A.S. and ASENJO, J.A., 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: effect of surface hydrophobicity and charge. Biotechnology and bioengineering, 90(3), pp. 380-390.

ANKAM, S., ROVINI, A., BAHETI, S., HRSTKA, R., WU, Y., SCHMIDT, K., WANG, H., MADIGAN, N., KOENIG, L.S., STELZIG, K., RESCH, Z., KLEIN, C.J., SUN, Z. and STAFF, N.P., 2019. DNA methylation patterns in human iPSC-derived sensory neuronal differentiation. Epigenetics, 14(9), pp. 927-937.

ANSORGE, S., LANTHIER, S., TRANSFIGURACION, J., DUROCHER, Y., HENRY, O. and KAMEN, A., 2009. Development of a scalable process for high-yield lentiviral vector production by transient transfection of HEK293 suspension cultures. The journal of gene medicine, 11(10), pp. 868-876.

AO, D.S., GAO, L.Y., GU, J.H., QIAO, J.H., WANG, H., LIU, Y.F. and SONG, H., 2020. Study on Adenovirus Infection in vitro with Nanoself-Assembling Peptide as Scaffolds for 3D Culture. International journal of nanomedicine, 15, pp. 6327-6338.

ARHEL, N.J., SOUQUERE-BESSE, S. and CHARNEAU, P., 2006. Wild-type and central DNA flap defective HIV-1 lentiviral vector genomes: intracellular visualization at ultrastructural resolution levels. Retrovirology, 3, pp. 38-4690-3-38.

ARHEL, N.J., SOUQUERE-BESSE, S., MUNIER, S., SOUQUE, P., GUADAGNINI, S., RUTHERFORD, S., PREVOST, M.C., ALLEN, T.D. and CHARNEAU, P., 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. The EMBO journal, 26(12), pp. 3025-3037.

ARMENDARIZ-BORUNDA, J., BASTIDAS-RAMIREZ, B.E., SANDOVAL-RODRIGUEZ, A., GONZALEZ-CUEVAS, J., GOMEZ-MEDA, B. and GARCIA-BANUELOS, J., 2011. Production of first generation adenoviral vectors for preclinical protocols: amplification, purification and functional titration. Journal of bioscience and bioengineering, 112(5), pp. 415-421.

ASENJO, J.A. and ANDREWS, B.A., 2012. Aqueous two-phase systems for protein separation: phase separation and applications. Journal of chromatography.A, 1238, pp. 1-10.

AVERY, O.T., MACLEOD, C.M. and MCCARTY, M., 1944. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal

Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. The Journal of experimental medicine, 79(2), pp. 137-158.

BABICH, A. and NEVINS, J.R., 1981. The stability of early adenovirus mRNA is controlled by the viral 72 kd DNA-binding protein. Cell, 26(3 Pt 1), pp. 371-379.

BACI, D., CHIRIVI, M., PACE, V., MAIULLARI, F., MILAN, M., RAMPIN, A., SOMMA, P., PRESUTTI, D., GARAVELLI, S., BRUNO, A., CANNATA, S., LANZUOLO, C., GARGIOLI, C., RIZZI, R. and BEARZI, C., 2020. Extracellular Vesicles from Skeletal Muscle Cells Efficiently Promote Myogenesis in Induced Pluripotent Stem Cells. Cells, 9(6), pp. 10.3390/cells9061527.

BACKSTROM, E., KAUFMANN, K.B., LAN, X. and AKUSJARVI, G., 2010. Adenovirus L4-22K stimulates major late transcription by a mechanism requiring the intragenic late-specific transcription factor-binding site. Virus research, 151(2), pp. 220-228.

BADEN, L.R., LIU, J., LI, H., JOHNSON, J.A., WALSH, S.R., KLEINJAN, J.A., ENGELSON, B.A., PETER, L., ABBINK, P., MILNER, D.A.,JR, GOLDEN, K.L., VIANI, K.L., STACHLER, M.D., CHEN, B.J., PAU, M.G., WEIJTENS, M., CAREY, B.R., MILLER, C.A., SWANN, E.M., WOLFF, M., LOBLEIN, H., SEAMAN, M.S., DOLIN, R. and BAROUCH, D.H., 2015. Induction of HIV-1-specific mucosal immune responses following intramuscular recombinant adenovirus serotype 26 HIV-1 vaccination of humans. The Journal of infectious diseases, 211(4), pp. 518-528.

BAGUTTI, C., ALT, M., SCHMIDLIN, M., VOGEL, G., VÖGELI, U. and BRODMANN, P., 2011. Detection of adeno- and lentiviral (HIV1)

contaminations on laboratory surfaces as a tool for the surveillance of biosafety standards. Journal of Applied Microbiology, 111(1), pp. 70-82.

BAILEY, C.J., CRYSTAL, R.G. and LEOPOLD, P.L., 2003. Association of adenovirus with the microtubule organizing center. Journal of virology, 77(24), pp. 13275-13287.

BALAFKAN, N., MOSTAFAVI, S., SCHUBERT, M., SILLER, R., LIANG, K.X., SULLIVAN, G. and BINDOFF, L.A., 2020. A method for differentiating human induced pluripotent stem cells toward functional cardiomyocytes in 96-well microplates. Scientific reports, 10(1), pp. 18498-020-73656-2.

BALTIMORE, D., 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature, 226(5252), pp. 1209-1211.

BANDEIRA, V., PEIXOTO, C., RODRIGUES, A.F., CRUZ, P.E., ALVES, P.M., COROADINHA, A.S. and CARRONDO, M.J.T., 2012. Downstream Processing of Lentiviral Vectors: Releasing Bottlenecks. Human Gene Therapy Methods, 23(4),.

BARR, S.D., CIUFFI, A., LEIPZIG, J., SHINN, P., ECKER, J.R. and BUSHMAN, F.D., 2006. HIV integration site selection: targeting in macrophages and the effects of different routes of viral entry. Molecular therapy : the journal of the American Society of Gene Therapy, 14(2), pp. 218-225.

BARR, S.D., LEIPZIG, J., SHINN, P., ECKER, J.R. and BUSHMAN, F.D., 2005. Integration targeting by avian sarcoma-leukosis virus and human immunodeficiency virus in the chicken genome. Journal of virology, 79(18), pp. 12035-12044.

BARRILLEAUX, B. and KNOEPFLER, P., 2011. Transduction of human cells with polymer-complexed ecotropic lentivirus for enhanced biosafety. Journal of visualized experiments : JoVE, (53). pii: 2822. doi(53), pp. 10.3791/2822.

BASKIR, J.N., HATTON, T.A. and SUTER, U.W., 1989. Protein partitioning in two-phase aqueous polymer systems. Biotechnology and bioengineering, 34(4), pp. 541-558.

BATALOV, I. and FEINBERG, A.W., 2015. Differentiation of Cardiomyocytes from Human Pluripotent Stem Cells Using Monolayer Culture. Biomarker insights, 10(Suppl 1), pp. 71-76.

BAUM, C., KUSTIKOVA, O., MODLICH, U., LI, Z. and FEHSE, B., 2006. Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. Human Gene Therapy, 17(3), pp. 253-263.

BAXTER, M., WITHEY, S., HARRISON, S., SEGERITZ, C.P., ZHANG, F., ATKINSON-DELL, R., ROWE, C., GERRARD, D.T., SISON-YOUNG, R., JENKINS, R., HENRY, J., BERRY, A.A., MOHAMET, L., BEST, M., FENWICK, S.W., MALIK, H., KITTERINGHAM, N.R., GOLDRING, C.E., PIPER HANLEY, K., VALLIER, L. and HANLEY, N.A., 2015. Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes. Journal of hepatology, 62(3), pp. 581-589.

BEARD, B.C., DICKERSON, D., BEEBE, K., GOOCH, C., FLETCHER, J., OKBINOGLU, T., MILLER, D.G., JACOBS, M.A., KAUL, R., KIEM, H.P. and TROBRIDGE, G.D., 2007. Comparison of HIV-derived lentiviral and MLV-based gammaretroviral vector integration sites in primate repopulating cells. Molecular therapy : the journal of the American Society of Gene Therapy, 15(7), pp. 1356-1365.

BENEDICT, C.A., NORRIS, P.S., PRIGOZY, T.I., BODMER, J.L., MAHR, J.A., GARNETT, C.T., MARTINON, F., TSCHOPP, J., GOODING, L.R. and WARE, C.F., 2001. Three adenovirus E3 proteins cooperate to evade apoptosis

by tumor necrosis factor-related apoptosis-inducing ligand receptor-1 and -2. The Journal of biological chemistry, 276(5), pp. 3270-3278.

BENYESH-MELNICK, M., PROBSTMEYER, F., MCCOMBS, R., BRUNSCHWIG, J.P. and VONKA, V., 1966. Correlation between infectivity and physical virus particles in human cytomegalovirus. Journal of Bacteriology, 92(5), pp. 1555-1561.

BERGELSON, J.M., CUNNINGHAM, J.A., DROGUETT, G., KURT-JONES, E.A., KRITHIVAS, A., HONG, J.S., HORWITZ, M.S., CROWELL, R.L. and FINBERG, R.W., 1997. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. Science, 275(5304), pp. 1320-1323.

BERKNER, K.L., 1988. Development of adenovirus vectors for the expression of heterologous genes. BioTechniques, 6(7), pp. 616-629.

BERNARDS, R., DE LEEUW, M.G., HOUWELING, A. and VAN DER EB, A.J., 1986. Role of the adenovirus early region 1B tumor antigens in transformation and lytic infection. Virology, 150(1), pp. 126-139.

BESNIER, C., TAKEUCHI, Y. and TOWERS, G., 2002. Restriction of lentivirus in monkeys. Proc.Natl.Acad.Sci.U.S.A., 99(18), pp. 11920-11925.

BETT, A.J., PREVEC, L. and GRAHAM, F.L., 1993. Packaging capacity and stability of human adenovirus type 5 vectors. Journal of virology, 67(10), pp. 5911-5921.

BEWLEY, M.C., SPRINGER, K., ZHANG, Y.B., FREIMUTH, P. and FLANAGAN, J.M., 1999. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. Science (New York, N.Y.), 286(5444), pp. 1579-1583.

BHF, 2019. Heart and Circulatory Disease Statistics 2019. British Heart Foundation.

BIASCO, L., AMBROSI, A., PELLIN, D., BARTHOLOMAE, C., BRIGIDA, I., RONCAROLO, M.G., DI SERIO, C., VON KALLE, C., SCHMIDT, M. and AIUTI, A., 2011. Integration profile of retroviral vector in gene therapy treated patients is cell-specific according to gene expression and chromatin conformation of target cell. EMBO molecular medicine, 3(2), pp. 89-101.

BIFFI, A., MONTINI, E., LORIOLI, L., CESANI, M., FUMAGALLI, F., PLATI, T., BALDOLI, C., MARTINO, S., CALABRIA, A., CANALE, S., BENEDICENTI, F., VALLANTI, G., BIASCO, L., LEO, S., KABBARA, N., ZANETTI, G., RIZZO, W.B., MEHTA, N.A., CICALESE, M.P., CASIRAGHI, M., BOELENS, J.J., DEL CARRO, U., DOW, D.J., SCHMIDT, M., ASSANELLI, A., NEDUVA, V., DI SERIO, C., STUPKA, E., GARDNER, J., VON KALLE, C., BORDIGNON, C., CICERI, F., ROVELLI, A., RONCAROLO, M.G., AIUTI, A., SESSA, M. and NALDINI, L., 2013. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science (New York, N.Y.), 341(6148), pp. 1233158.

BISCHOFF, J.R., KIRN, D.H., WILLIAMS, A., HEISE, C., HORN, S., MUNA, M., NG, L., NYE, J.A., SAMPSON-JOHANNES, A., FATTAEY, A. and MCCORMICK, F., 1996. An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. Science (New York, N.Y.), 274(5286), pp. 373-376.

BLAESE, R.M., CULVER, K.W., MILLER, A.D., CARTER, C.S., FLEISHER, T., CLERICI, M., SHEARER. G., CHANG, L., CHIANG, Y., TOLSTOSHEV, P., GREENBLATT, J.J., ROSENBERG, S.A., KLEIN, H., BERGER, M., MULLEN, C.A., RAMSEY, W.J., MUUL, L., MORGAN, R.A. and ANDERSON, W.F., 1995. T lymphocyte-directed gene therapy for ADA-SCID: initial trial results after 4 years. Science, 270(5235), pp. 475-480.

BLASCO, M.A., SERRANO, M. and FERNANDEZ-CAPETILLO, O., 2011. Genomic instability in iPS: time for a break. The EMBO journal, 30(6), pp. 991-993.

BLAZESKI, A., ZHU, R., HUNTER, D.W., WEINBERG, S.H., BOHELER, K.R., ZAMBIDIS, E.T. and TUNG, L., 2012. Electrophysiological and contractile function of cardiomyocytes derived from human embryonic stem cells. Progress in biophysics and molecular biology, 110(2-3), pp. 178-195.

BLAZKOVA, J., TREJBALOVA, K., GONDOIS-REY, F., HALFON, P., PHILIBERT, P., GUIGUEN, A., VERDIN, E., OLIVE, D., VAN LINT, C., HEJNAR, J. and HIRSCH, I., 2009. CpG methylation controls reactivation of HIV from latency. PLoS pathogens, 5(8), pp. e1000554.

BOKHOVEN, M., STEPHEN, S.L., KNIGHT, S., GEVERS, E.F., ROBINSON, I.C., TAKEUCHI, Y. and COLLINS, M.K., 2009. Insertional gene activation by lentiviral and gammaretroviral vectors. Journal of virology, 83(1), pp. 283-294.

BONCI, D., CITTADINI, A., LATRONICO, M.V., BORELLO, U., AYCOCK, J.K., DRUSCO, A., INNOCENZI, A., FOLLENZI, A., LAVITRANO, M., MONTI, M.G., ROSS, J.,JR, NALDINI, L., PESCHLE, C., COSSU, G. and CONDORELLI, G., 2003. 'Advanced' generation lentiviruses as efficient vectors for cardiomyocyte gene transduction in vitro and in vivo. Gene therapy, 10(8), pp. 630-636.

BONOW, R.O., MANN, D.L., ZIPES, D.P. and LIBBY, P., 2011. Primary Tumours of the Heart. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 9 edn. Philadelphia: Elsevier Saunders, pp. 1638-1650.

BORDIGNON, C., NOTARANGELO, L.D., NOBILI, N., FERRARI, G., CASORATI, G., PANINA, P., MAZZOLARI, E., MAGGIONI, D., ROSSI, C.,

SERVIDA, P., UGAZIO, A.G. and MAVILIO, F., 1995. Gene therapy in peripheral blood lymphocytes and bone marrow for ADA- immunodeficient patients. Science (New York, N.Y.), 270(5235), pp. 470-475.

BOROUJENI, M.E. and GARDANEH, M., 2018. The Superiority of Sucrose Cushion Centrifugation to Ultrafiltration and PEGylation in Generating High-Titer Lentivirus Particles and Transducing Stem Cells with Enhanced Efficiency. Molecular biotechnology, 60(3), pp. 185-193.

BOWERMAN, B., BROWN, P.O., BISHOP, J.M. and VARMUS, H.E., 1989. A nucleoprotein complex mediates the integration of retroviral DNA. Genes & development, 3(4), pp. 469-478.

BOYCHYN, M., YIM, S.S.S., SHAMLOU, P.A., BULMER, M., MORE, J. and HOARE, M., 2001. Characterization of flow intensity in continuous centrifuges for the development of laboratory mimics. Chemical Engineering Science, 56(16), pp. 4759-4770.

BOYER, L.A., LEE, T.I., COLE, M.F., JOHNSTONE, S.E., LEVINE, S.S., ZUCKER, J.P., GUENTHER, M.G., KUMAR, R.M., MURRAY, H.L., JENNER, R.G., GIFFORD, D.K., MELTON, D.A., JAENISCH, R. and YOUNG, R.A., 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell, 122(6), pp. 947-956.

BOYER, T.G., MARTIN, M.E., LEES, E., RICCIARDI, R.P. and BERK, A.J., 1999. Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. Nature, 399(6733), pp. 276-279.

BRAAS, G.M., WALKER, S.G. and LYDDIATT, A., 2000. Recovery in aqueous two-phase systems of nanoparticulates applied as surrogate mimics for viral gene therapy vectors. Journal of chromatography.B, Biomedical sciences and applications, 743(1-2), pp. 409-419.

BRADY, T., AGOSTO, L.M., MALANI, N., BERRY, C.C., O'DOHERTY, U. and BUSHMAN, F., 2009. HIV integration site distributions in resting and activated CD4+ T cells infected in culture. AIDS (London, England), 23(12), pp. 1461-1471.

BRANCO, M.A., COTOVIO, J.P., RODRIGUES, C.A.V., VAZ, S.H., FERNANDES, T.G., MOREIRA, L.M., CABRAL, J.M.S. and DIOGO, M.M., 2019. Transcriptomic analysis of 3D Cardiac Differentiation of Human Induced Pluripotent Stem Cells Reveals Faster Cardiomyocyte Maturation Compared to 2D Culture. Scientific reports, 9(1), pp. 9229-019-45047-9.

BRENER, J., GALL, A., BATORSKY, R., RIDDELL, L., BUUS, S., LEITMAN, E., KELLAM, P., ALLEN, T., GOULDER, P. and MATTHEWS, P.C., 2015. Disease progression despite protective HLA expression in an HIV-infected transmission pair. Retrovirology, 12(55),.

BRIGGS, R. and KING, T.J., 1952. Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proceedings of the National Academy of Sciences of the United States of America, 38(5), pp. 455-463.

BROWN, D.T., WESTPHAL, M., BURLINGHAM, B.T., WINTERHOFF, U. and DOERFLER, W., 1975. Structure and composition of the adenovirus type 2 core. Journal of virology, 16(2), pp. 366-387.

BROWNING, D.L. and TROBRIDGE, G.D., 2016. Insulators to Improve the Safety of Retroviral Vectors for HIV Gene Therapy. Biomedicines, 4(1), pp. 10.3390/biomedicines4010004. Epub 2016 Feb 2.

BUCHBINDER, S.P., MEHROTRA, D.V., DUERR, A., FITZGERALD, D.W., MOGG, R., LI, D., GILBERT, P.B., LAMA, J.R., MARMOR, M., DEL RIO, C., MCELRATH, M.J., CASIMIRO, D.R., GOTTESDIENER, K.M., CHODAKEWITZ, J.A., COREY, L., ROBERTSON, M.N. and STEP STUDY PROTOCOL TEAM, 2008. Efficacy assessment of a cell-mediated immunity

HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet (London, England), 372(9653), pp. 1881-1893.

BUKRINSKY, M.I., SHAROVA, N., MCDONALD, T.L., PUSHKARSKAYA, T., TARPLEY, W.G. and STEVENSON, M., 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proceedings of the National Academy of Sciences of the United States of America, 90(13), pp. 6125-6129.

BURNS, J.C., FRIEDMANN, T., DRIEVER, W., BURRASCANO, M. and 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. Proc.Natl.Acad.Sci.U.S.A., 90(17), pp. 8033-8037.

BURRIDGE, P.W., KELLER, G., GOLD, J.D. and WU, J.C., 2012. Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. Cell stem cell, 10(1), pp. 16-28.

BURRIDGE, P.W., MATSA, E., SHUKLA, P., LIN, Z.C., CHURKO, J.M., EBERT, A.D., LAN, F., DIECKE, S., HUBER, B., MORDWINKIN, N.M., PLEWS, J.R., ABILEZ, O.J., CUI, B., GOLD, J.D. and WU, J.C., 2014. Chemically defined generation of human cardiomyocytes. Nature methods, 11(8), pp. 855-860.

BUSHMAN, F., LEWINSKI, M., CIUFFI, A., BARR, S., LEIPZIG, J., HANNENHALLI, S. and HOFFMANN, C., 2005. Genome-wide analysis of retroviral DNA integration. Nature reviews.Microbiology, 3(11), pp. 848-858.

BUSHMAN, F.D., 2003. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. Cell, 115(2), pp. 135-138.

BUZANSKA, L., JURGA, M., STACHOWIAK, E.K., STACHOWIAK, M.K. and DOMANSKA-JANIK, K., 2006. Neural stem-like cell line derived from a nonhematopoietic population of human umbilical cord blood. Stem cells and development, 15(3), pp. 391-406.

CANONNE-HERGAUX, F., AUNIS, D. and SCHAEFFER, E., 1995. Interactions of the transcription factor AP-1 with the long terminal repeat of different human immunodeficiency virus type 1 strains in Jurkat, glial, and neuronal cells. Journal of virology, 69(11), pp. 6634-6642.

CARNEIRO, F.A., BIANCONI, M.L., WEISSMULLER, G., STAUFFER, F. and DA POIAN, A.T., 2002. Membrane recognition by vesicular stomatitis virus involves enthalpy-driven protein-lipid interactions. Journal of virology, 76(8), pp. 3756-3764.

CARNES, A., HODGSON, C. and WILLIAMS, J., 2004. Optimization of E. coli Fermentation for Plasmid DNA Production. Molecular Therapy, 9(1), pp. S310.

CARR, S.M. and GRIFFITH, O.M., 1987. Rapid isolation of animal mitochondrial DNA in a small fixed-angle rotor at ultrahigh speed. Biochemical Genetics, 25(5-6), pp. 385-390.

CARRERA, C., PINILLA, M., PEREZ-ALVAREZ, L. and THOMSON, M.M., 2010. Identification of unusual and novel HIV type 1 spliced transcripts generated in vivo. AIDS Research and Human Retroviruses, 26(7), pp. 815-820.

CARVAJAL-VERGARA, X., SEVILLA, A., D'SOUZA, S.L., ANG, Y.S., SCHANIEL, C., LEE, D.F., YANG, L., KAPLAN, A.D., ADLER, E.D., ROZOV, R., GE, Y., COHEN, N., EDELMANN, L.J., CHANG, B., WAGHRAY, A., SU, J., PARDO, S., LICHTENBELT, K.D., TARTAGLIA, M., GELB, B.D. and LEMISCHKA, I.R., 2010. Patient-specific induced

pluripotent stem-cell-derived models of LEOPARD syndrome. Nature, 465(7299), pp. 808-812.

CASTELLANI, S., DI GIOIA, S., TROTTA, T., MAFFIONE, A.B. and CONESE, M., 2010. Impact of Lentiviral Vector-Mediated Transduction on the Tightness of a Polarized Model of Airway Epithelium and Effect of Cationic Polymer Polyethylenimine. BioMed Research International, 2010(The Epithelium–Molecular Landscaping for an Interactive Barrier),.

CASTRO, M.G., GOYA, R.G., SOSA, Y.E., ROWE, J., LARREGINA, A., MORELLI, A. and LOWENSTEIN, P.R., 1997. Expression of transgenes in normal and neoplastic anterior pituitary cells using recombinant adenoviruses: long term expression, cell cycle dependency, and effects on hormone secretion. Endocrinology, 138(5), pp. 2184-2194.

CATTOGLIO, C., PELLIN, D., RIZZI, E., MARUGGI, G., CORTI, G., MISELLI, F., SARTORI, D., GUFFANTI, A., DI SERIO, C., AMBROSI, A., DE BELLIS, G. and MAVILIO, F., 2010. High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. Blood, 116(25), pp. 5507-5517.

CAVAZZANA-CALVO, M., PAYEN, E., NEGRE, O., WANG, G., HEHIR, K., FUSIL, F., DOWN, J., DENARO, M., BRADY, T., WESTERMAN, K., CAVALLESCO, R., GILLET-LEGRAND, B., CACCAVELLI, L., SGARRA, R., MAOUCHE-CHRETIEN, L., BERNAUDIN, F., GIROT, R., DORAZIO, R., MULDER, G.J., POLACK, A., BANK, A., SOULIER, J., LARGHERO, J., KABBARA, N., DALLE, B., GOURMEL, B., SOCIE, G., CHRETIEN, S., CARTIER, N., AUBOURG, P., FISCHER, A., CORNETTA, K., GALACTEROS, F., BEUZARD, Y., GLUCKMAN, E., BUSHMAN, F., HACEIN-BEY-ABINA, S. and LEBOULCH, P., 2010. Transfusion

independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature, 467(7313), pp. 318-322.

CESANA, D., RANZANI, M., BARTHOLOMÄ, C., VOLPIN, M., MERELLA, S., BENEDICENTI, F., SERGI, L.S., SAN VITO, F., BROMBIN, C., NONIS, A., DISERIO, C., DOGLIONI, C., VON KALLE, C., SCHMIDT, M., NALDINI, L. and MONTINI, E., 2012. In Vivo Mouse Models for Vector Genotoxicity Testing and Lentiviral Vector-Based Cancer Gene Discovery. Molecular Therapy, 20(1), pp. S2-S3.

CESANA, D., RANZANI, M., VOLPIN, M., BARTHOLOMAE, C., DUROS, C., ARTUS, A., MERELLA, S., BENEDICENTI, F., SERGI SERGI, L., SANVITO, F., BROMBIN, C., NONIS, A., SERIO, C.D., DOGLIONI, C., VON KALLE, C., SCHMIDT, M., COHEN-HAGUENAUER, O., NALDINI, L. and MONTINI, E., 2014. Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. Molecular therapy : the journal of the American Society of Gene Therapy, 22(4), pp. 774-785.

CESANA, D., SGUALDINO, J., RUDILOSSO, L., MERELLA, S., NALDINI, L. and MONTINI, E., 2012. Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. The Journal of clinical investigation, 122(5), pp. 1667-1676.

CHANDLER, R.J., LAFAVE, M.C., VARSHNEY, G.K., TRIVEDI, N.S., CARRILLO-CARRASCO, N., SENAC, J.S., WU, W., HOFFMANN, V., ELKAHLOUN, A.G., BURGESS, S.M. and VENDITTI, C.P., 2015. Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. The Journal of clinical investigation, 125(2), pp. 870-880.

CHAO, S.H., WALKER, J.R., CHANDA, S.K., GRAY, N.S. and CALDWELL, J.S., 2003. Identification of homeodomain proteins, PBX1 and

PREP1, involved in the transcription of murine leukemia virus. Molecular and cellular biology, 23(3), pp. 831-841.

CHAUDHARY, S., PAK, J.E., GRUSWITZ, F., SHARMA, V. and STROUD, R.M., 2012. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. Nature protocols, 7(3), pp. 453-466.

CHEN, H., LI, H., LIU, F., ZHENG, X., WANG, S., BO, X. and SHU, W., 2015. An integrative analysis of TFBS-clustered regions reveals new transcriptional regulation models on the accessible chromatin landscape. Scientific reports, 5, pp. 8465.

CHEN, Z.Y., RIU, E., HE, C.Y., XU, H. and KAY, M.A., 2008. Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. Molecular therapy : the journal of the American Society of Gene Therapy, 16(3), pp. 548-556.

CHENG, C., WANG, L., KO, S.Y., KONG, W.P., SCHMIDT, S.D., GALL, J.G.D., COLLOCA, S., SEDER, R.A., MASCOLA, J.R. and NABEL, G.J., 2015. Combination recombinant simian or chimpanzee adenoviral vectors for vaccine development. Vaccine, 33(51), pp. 7344-7351.

CHEREPANOV, P., MAERTENS, G., PROOST, P., DEVREESE, B., VAN BEEUMEN, J., ENGELBORGHS, Y., DE CLERCQ, E. and DEBYSER, Z., 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. The Journal of biological chemistry, 278(1), pp. 372-381.

CHOU, B.K., GU, H., GAO, Y., DOWEY, S.N., WANG, Y., SHI, J., LI, Y., YE, Z., CHENG, T. and CHENG, L., 2015. A Facile Method to Establish Human Induced Pluripotent Stem Cells From Adult Blood Cells Under Feeder-

Free and Xeno-Free Culture Conditions: A Clinically Compliant Approach. Stem Cells Translational Medicine, 4(4), pp. 320-332.

CHOUDHURY, Y., TOH, Y.C., XING, J., QU, Y., POH, J., LI, H., TAN, H.S., KANESVARAN, R., YU, H. and TAN, M.H., 2017. Patient-specific hepatocyte-like cells derived from induced pluripotent stem cells model pazopanib-mediated hepatotoxicity. Scientific reports, 7, pp. 41238.

CHU, R.L., POST, D.E., KHURI, F.R. and VAN MEIR, E.G., 2004. Use of replicating oncolytic adenoviruses in combination therapy for cancer. Clinical cancer research : an official journal of the American Association for Cancer Research, 10(16), pp. 5299-5312.

CIUFFI, A., 2008. Mechanisms governing lentivirus integration site selection. Current gene therapy, 8(6), pp. 419-429.

CIUFFI, A. and BARR, S.D., 2011. Identification of HIV integration sites in infected host genomic DNA. Methods (San Diego, Calif.), 53(1), pp. 39-46.

CIUFFI, A., LLANO, M., POESCHLA, E., HOFFMANN, C., LEIPZIG, J., SHINN, P., ECKER, J.R. and BUSHMAN, F., 2005. A role for LEDGF/p75 in targeting HIV DNA integration. Nature medicine, 11(12), pp. 1287-1289.

CLINICALTRIALS.GOV, , Lentiviral vector trials. Available: https://clinicaltrials.gov/ct2/results?cond=&term=lentiviral+vector&cntry=&state=&city=&dist= [December/31, 2020].

COFFIN, J.M., HUGHES, S.H. and VARMUS, H.E., 1997. The Interactions of Retroviruses and their Hosts. In: J.M. COFFIN, S.H. HUGHES and H.E. VARMUS, eds, Retroviruses. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, .

COHN, L.B., SILVA, I.T., OLIVEIRA, T.Y., ROSALES, R.A., PARRISH, E.H., LEARN, G.H., HAHN, B.H., CZARTOSKI, J.L., MCELRATH, M.J.,

LEHMANN, C., KLEIN, F., CASKEY, M., WALKER, B.D., SILICIANO, J.D., SILICIANO, R.F., JANKOVIC, M. and NUSSENZWEIG, M.C., 2015. HIV-1 integration landscape during latent and active infection. Cell, 160(3), pp. 420-432.

COIL, D.A. and MILLER, A.D., 2004. Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. Journal of virology, 78(20), pp. 10920-10926.

COMMITTEE FOR MEDICAL PRODUCTS FOR HUMAN USE, 2005. Guideline on Development and Manufacture of Lentiviral Vectors. London: European Medicines Agency.

COMMUNAL, C., HUQ, F., LEBECHE, D., MESTEL, C., GWATHMEY, J.K. and HAJJAR, R.J., 2003. Decreased efficiency of adenovirus-mediated gene transfer in aging cardiomyocytes. Circulation, 107(8), pp. 1170-1175.

CORNELISSEN, M., GALL, A., VINK, M., ZORGDRAGER, F., BINTER, S., EDWARDS, S., JURRIAANS, S., BAKKER, M., ONG, S.H., GRAS, L., VAN SIGHEM, A., BEZEMER, D., DE WOLF, F., REISS, P., KELLAM, P., BERKHOUT, B., FRASER, C., VAN DER KUYL, A.C. and BEEHIVE CONSORTIUM, 2017. From clinical sample to complete genome: Comparing methods for the extraction of HIV-1 RNA for high-throughput deep sequencing. Virus research, 239, pp. 10-16.

COSGUN, K.N., RAHMIG, S., MENDE, N., REINKE, S., HAUBER, I., SCHAFER, C., PETZOLD, A., WEISBACH, H., HEIDKAMP, G., PURBOJO, A., CESNJEVAR, R., PLATZ, A., BORNHAUSER, M., SCHMITZ, M., DUDZIAK, D., HAUBER, J., KIRBERG, J. and WASKOW, C., 2014. Kit regulates HSC engraftment across the human-mouse species barrier. Cell stem cell, 15(2), pp. 227-238.

COSSET, F.L., TAKEUCHI, Y., BATTINI, J.L., WEISS, R.A. and COLLINS, M.K., 1995. High-titer packaging cells producing recombinant retroviruses resistant to human serum. Journal of virology, 69(12), pp. 7430-7436.

COSTANTINI, S., DI BERNARDO, G., CAMMAROTA, M., CASTELLO, G. and COLONNA, G., 2013. Gene expression signature of human HepG2 cell line. Gene, 518(2), pp. 335-345.

CRAIGIE, R. and BUSHMAN, F.D., 2012. HIV DNA integration. Cold Spring Harbor perspectives in medicine, 2(7), pp. a006890.

CRAWFORD-MIKSZA, L. and SCHNURR, D.P., 1996. Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. Journal of virology, 70(3), pp. 1836-1844.

CRICK, F., 1970. Central dogma of molecular biology. Nature, 227(5258), pp. 561-563.

CRONIN, J., ZHANG, X.Y. and REISER, J., 2005. Altering the tropism of lentiviral vectors through pseudotyping. Current gene therapy, 5(4), pp. 387-398.

CROYLE, M.A., YU, Q.C. and WILSON, J.M., 2000. Development of a rapid method for the PEGylation of adenoviruses with enhanced transduction and improved stability under harsh storage conditions. Human Gene Therapy, 11(12), pp. 1713-1722.

CROYLE, M.A., CALLAHAN, S.M., AURICCHIO, A., SCHUMER, G., LINSE, K.D., WILSON, J.M., BRUNNER, L.J. and KOBINGER, G.P., 2004. PEGylation of a vesicular stomatitis virus G pseudotyped lentivirus vector prevents inactivation in serum. Journal of virology, 78(2), pp. 912-921.

CROYLE, M.A., LE, H.T., LINSE, K.D., CERULLO, V., TOIETTA, G., BEAUDET, A. and PASTORE, L., 2005. PEGylated helper-dependent adenoviral vectors: highly efficient vectors with an enhanced safety profile. Gene therapy, 12(7), pp. 579-587.

CSOBONYEIOVA, M., POLAK, S. and DANISOVIC, L., 2016. Toxicity testing and drug screening using iPSC-derived hepatocytes, cardiomyocytes, and neural cells. Canadian journal of physiology and pharmacology, 94(7), pp. 687-694.

CYGANEK, L., TIBURCY, M., SEKERES, K., GERSTENBERG, K., BOHNENBERGER, H., LENZ, C., HENZE, S., STAUSKE, M., SALINAS, G., ZIMMERMANN, W.H., HASENFUSS, G. and GUAN, K., 2018. Deep phenotyping of human induced pluripotent stem cell-derived atrial and ventricular cardiomyocytes. JCI insight, 3(12), pp. 10.1172/jci.insight.99941. eCollection 2018 Jun 21.

DANG, C.V., 2012. MYC on the path to cancer. Cell, 149(1), pp. 22-35.

DANIEL, R. and SMITH, J.A., 2008. Integration site selection by retroviral vectors: molecular mechanism and clinical consequences. Human Gene Therapy, 19(6), pp. 557-568.

DANQUAH, M.K. and FORDE, G.M., 2007. Growth medium selection and its economic impact on plasmid DNA production. Journal of bioscience and bioengineering, 104(6), pp. 490-497.

DASGUPTA, P., SAIKUMAR, P., REDDY, C.D. and REDDY, E.P., 1990. Myb protein binds to human immunodeficiency virus 1 long terminal repeat (LTR) sequences and transactivates LTR-mediated transcription. Proceedings of the National Academy of Sciences of the United States of America, 87(20), pp. 8090-8094.

DAUWE, K., STAELENS, D., VANCOILLIE, L., MORTIER, V. and VERHOFSTEDE, C., 2016. Deep Sequencing of HIV-1 RNA and DNA in Newly Diagnosed Patients with Baseline Drug Resistance Showed No Indications for Hidden Resistance and Is Biased by Strong Interference of Hypermutation. Journal of clinical microbiology, 54(6), pp. 1605-1615.

DAVID, R.M. and DOHERTY, A.T., 2017. Viral Vectors: The Road to Reducing Genotoxicity. Toxicological sciences : an official journal of the Society of Toxicology, 155(2), pp. 315-325.

DAWSON, M.H. and SIA, R.H., 1931. In Vitro Transformation of Pneumococcal Types : I. a Technique for Inducing Transformation of Pneumococcal Types in Vitro. The Journal of experimental medicine, 54(5), pp. 681-699.

DELVILLE, M., SOHEILI, T., BELLIER, F., DURAND, A., DENIS, A., LAGRESLE-PEYROU, C., CAVAZZANA, M., ANDRE-SCHMUTZ, I. and SIX, E., 2018. A Nontoxic Transduction Enhancer Enables Highly Efficient Lentiviral Transduction of Primary Murine T Cells and Hematopoietic Stem Cells. Molecular therapy.Methods & clinical development, 10, pp. 341-347.

DEMAISON, C., PARSLEY, K., BROUNS, G., SCHERR, M., BATTMER, K., KINNON, C., GREZ, M. and THRASHER, A.J., 2002. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Human Gene Therapy, 13(7), pp. 803-813.

DENNING, C., BORGDORFF, V., CRUTCHLEY, J., FIRTH, K.S., GEORGE, V., KALRA, S., KONDRASHOV, A., HOANG, M.D., MOSQUEIRA, D., PATEL, A., PRODANOV, L., RAJAMOHAN, D., SKARNES, W.C., SMITH, J.G. and YOUNG, L.E., 2016. Cardiomyocytes from human pluripotent stem

cells: From laboratory curiosity to industrial biomedical platform. Biochimica et biophysica acta, 1863(7 Pt B), pp. 1728-1748.

DENNING, W., DAS, S., GUO, S., XU, J., KAPPES, J.C. and HEL, Z., 2013. Optimization of the transductional efficiency of lentiviral vectors: effect of sera and polycations. Molecular biotechnology, 53(3), pp. 308-314.

DEVALLA, H.D., SCHWACH, V., FORD, J.W., MILNES, J.T., EL-HAOU, S., JACKSON, C., GKATZIS, K., ELLIOTT, D.A., CHUVA DE SOUSA LOPES, S.M., MUMMERY, C.L., VERKERK, A.O. and PASSIER, R., 2015. Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. EMBO molecular medicine, 7(4), pp. 394-410.

DI GIALLONARDO, F., ZAGORDI, O., DUPORT, Y., LEEMANN, C., JOOS, B., KUNZLI-GONTARCZYK, M., BRUGGMANN, R., BEERENWINKEL, N., GUNTHARD, H.F. and METZNER, K.J., 2013. Next-generation sequencing of HIV-1 RNA genomes: determination of error rates and minimizing artificial recombination. PloS one, 8(9), pp. e74249.

DICKS, M.D., SPENCER, A.J., EDWARDS, N.J., WADELL, G., BOJANG, K., GILBERT, S.C., HILL, A.V. and COTTINGHAM, M.G., 2012. A novel chimpanzee adenovirus vector with low human seroprevalence: improved systems for vector derivation and comparative immunogenicity. PloS one, 7(7), pp. e40385.

DING, B. and KILPATRICK, D.L., 2013. Lentiviral vector production, titration, and transduction of primary neurons. Methods in molecular biology (Clifton, N.J.), 1018, pp. 119-131.

DONAHUE, R.E., KESSLER, S.W., BODINE, D., MCDONAGH, K., DUNBAR, C., GOODMAN, S., AGRICOLA, B., BYRNE, E., RAFFELD, M. and MOEN, R., 1992. Helper virus induced T cell lymphoma in nonhuman

primates after retroviral mediated gene transfer. The Journal of experimental medicine, 176(4), pp. 1125-1135.

DONSANTE, A., MILLER, D.G., LI, Y., VOGLER, C., BRUNT, E.M., RUSSELL, D.W. and SANDS, M.S., 2007. AAV vector integration sites in mouse hepatocellular carcinoma. Science (New York, N.Y.), 317(5837), pp. 477.

DRAWNEL, F.M., BOCCARDO, S., PRUMMER, M., DELOBEL, F., GRAFF, A., WEBER, M., GERARD, R., BADI, L., KAM-THONG, T., BU, L., JIANG, X., HOFLACK, J.C., KIIALAINEN, A., JEWORUTZKI, E., AOYAMA, N., CARLSON, C., BURCIN, M., GROMO, G., BOEHRINGER, M., STAHLBERG, H., HALL, B.J., MAGNONE, M.C., KOLAJA, K., CHIEN, K.R., BAILLY, J. and IACONE, R., 2014. Disease modeling and phenotypic drug screening for diabetic cardiomyopathy using human induced pluripotent stem cells. Cell reports, 9(3), pp. 810-821.

D'SOUZA, A.A. and SHEGOKAR, R., 2016. Polyethylene glycol (PEG): a versatile polymer for pharmaceutical applications. Expert opinion on drug delivery, 13(9), pp. 1257-1275.

DUBRIDGE, R.B., TANG, P., HSIA, H.C., LEONG, P.M., MILLER, J.H. and CALOS, M.P., 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Molecular and cellular biology, 7(1), pp. 379-387.

DUCHESNEAU, P., BESLA, R., DEROUET, M.F., GUO, L., KAROUBI, G., SILBERBERG, A., WONG, A.P. and WADDELL, T.K., 2017. Partial Restoration of CFTR Function in cftr-Null Mice following Targeted Cell Replacement Therapy. Molecular therapy : the journal of the American Society of Gene Therapy, 25(3), pp. 654-665.

DUH, E.J., MAURY, W.J., FOLKS, T.M., FAUCI, A.S. and RABSON, A.B., 1989. Tumor necrosis factor alpha activates human immunodeficiency virus

type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. Proceedings of the National Academy of Sciences of the United States of America, 86(15), pp. 5974-5978.

DUISIT, G., SALVETTI, A., MOULLIER, P. and COSSET, F.L., 1999. Functional characterization of adenoviral/retroviral chimeric vectors and their use for efficient screening of retroviral producer cell lines. Human Gene Therapy, 10(2), pp. 189-200.

DULL, T., ZUFFEREY, R., KELLY, M., MANDEL, R.J., NGUYEN, M., TRONO, D. and NALDINI, L., 1998. A third-generation lentivirus vector with a conditional packaging system. Journal of virology, 72(11), pp. 8463-8471.

DURAND, S. and CIMARELLI, A., 2011. The inside out of lentiviral vectors. Viruses, 3(2), pp. 132-159.

DUVERGE, A. and NEGRONI, M., 2020. Pseudotyping Lentiviral Vectors: When the Clothes Make the Virus. Viruses, 12(11), pp. 10.3390/v12111311.

ELLISON, V., ABRAMS, H., ROE, T., LIFSON, J. and BROWN, P., 1990. Human immunodeficiency virus integration in a cell-free system. Journal of virology, 64(6), pp. 2711-2715.

EMERY, A., ZHOU, S., POLLOM, E. and SWANSTROM, R., 2017. Characterizing HIV-1 Splicing by Using Next-Generation Sequencing. Journal of virology, 91(6), pp. 10.1128/JVI.02515-16. Print 2017 Mar 15.

EMERY, D.W., 2011. The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. Human Gene Therapy, 22(6), pp. 761-774.

EPSTEIN, M.A. and HOLT, S.J., 1960. Site and nature of adenovirus nucleic acid. Nature, 187, pp. 1050-1051.

ESCORS, D., BRECKPOT, K., ARCE, F., KOCHAN, G. and STEPHENSON, H., 2012. Chapter 2: Development of retroviral and lentiviral vectors. Lentiviral vectors and gene therapy. 1 edn. Switzerland: Springer Basel, pp. 11-28.

ESTEPA, A.M., ROCHA, A.I., MAS, V., PEREZ, L., ENCINAR, J.A., NUNEZ, E., FERNANDEZ, A., GONZALEZ ROS, J.M., GAVILANES, F. and COLL, J.M., 2001. A protein G fragment from the salmonid viral hemorrhagic septicemia rhabdovirus induces cell-to-cell fusion and membrane phosphatidylserine translocation at low pH. The Journal of biological chemistry, 276(49), pp. 46268-46275.

ETO, Y., YOSHIOKA, Y., MUKAI, Y., OKADA, N. and NAKAGAWA, S., 2008. Development of PEGylated adenovirus vector with targeting ligand. International journal of pharmaceutics, 354(1-2), pp. 3-8.

EUROPEAN MEDICINES AGENCY, 2019. Guideline on quality, non-clinical and clinical requirements for investigational advanced therapy medicinal products in clinical trials. European Medicines Agency.

EVANS, M.J. and KAUFMAN, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature, 292(5819), pp. 154-156.

FANG, J.Y., MIKOVITS, J.A., BAGNI, R., PETROW-SADOWSKI, C.L. and RUSCETTI, F.W., 2001. Infection of lymphoid cells by integration-defective human immunodeficiency virus type 1 increases de novo methylation. Journal of virology, 75(20), pp. 9753-9761.

FANG, J.Y. and XIAO, S.D., 2001. Alteration of DNA methylation in gastrointestinal carcinogenesis. Journal of gastroenterology and hepatology, 16(9), pp. 960-968.

FARID, S.S., 2007. Process economics of industrial monoclonal antibody manufacture. Journal of chromatography.B, Analytical technologies in the biomedical and life sciences, 848(1), pp. 8-18.

FARLEY, D.C., MCCLOSKEY, L., THORNE, B.A., TAREEN, S.U., NICOLAI, C.J., CAMPBELL, D.J., BANNISTER, R., STEWART, H.J., PEARSON, L.J., MOYER, B.J., ROBBINS, S.H., ZIELINSKI, L., KIM, T., RADCLIFFE, P.A., MITROPHANOUS, K.A., GOMBOTZ, W.R., MISKIN, J.E. and KELLEY-CLARKE, B., 2015. Development of a replication-competent lentivirus assay for dendritic cell-targeting lentiviral vectors. Molecular Therapy Methods, 2(15017),.

FARNET, C.M. and BUSHMAN, F.D., 1997. HIV-1 cDNA Integration: Requirement of HMG I(Y) Protein for Function of Preintegration Complexes In Vitro. Cell, 88, pp. 483-192.

FARNET, C.M. and HASELTINE, W.A., 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. Journal of virology, 65(4), pp. 1910-1915.

FARNET, C.M. and HASELTINE, W.A., 1990. Integration of human immunodeficiency virus type 1 DNA in vitro. Proceedings of the National Academy of Sciences of the United States of America, 87(11), pp. 4164-4168.

FASSATI, A. and GOFF, S.P., 1999. Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus. Journal of virology, 73(11), pp. 8919-8925.

FEDERICI, T., KUTNER, R., ZHANG, X.Y., KURODA, H., TORDO, N., BOULIS, N.M. and REISER, J., 2009. Comparative analysis of HIV-1-based lentiviral vectors bearing lyssavirus glycoproteins for neuronal gene transfer. Genetic vaccines and therapy, 7, pp. 1-0556-7-1.

FELDMAN, L.T. and NEVINS, J.R., 1983. Localization of the adenovirus E1Aa protein, a positive-acting transcriptional factor, in infected cells infected cells. Molecular and cellular biology, 3(5), pp. 829-838.

FELICE, B., CATTOGLIO, C., CITTARO, D., TESTA, A., MICCIO, A., FERRARI, G., LUZI, L., RECCHIA, A. and MAVILIO, F., 2009. Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. PloS one, 4(2), pp. e4571.

FIELDS, B.N., KNIPE, D.M. and HOWLEY, P.M., 2007. Fields virology. 5 edn. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

FOLEGATTI, P.M., EWER, K.J., ALEY, P.K., ANGUS, B., BECKER, S., BELIJ-RAMMERSTORFER, S., BELLAMY, D., BIBI, S., BITTAYE, M., CLUTTERBUCK, E.A., DOLD, C., FAUST, S.N., FINN, A., FLAXMAN, A.L., HALLIS, B., HEATH, P., JENKIN, D., LAZARUS, R., MAKINSON, R., MINASSIAN, A.M., POLLOCK, K.M., RAMASAMY, M., ROBINSON, H., SNAPE, M., TARRANT, R., VOYSEY, M., GREEN, C., DOUGLAS, A.D., HILL, A.V.S., LAMBE, T., GILBERT, S.C., POLLARD, A.J. and OXFORD COVID VACCINE TRIAL GROUP, 2020. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet, 396(10249), pp. 467-478.

FOLLENZI, A., AILLES, L.E., BAKOVIC, S., GEUNA, M. and NALDINI, L., 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nature genetics, 25(2), pp. 217-222.

FOROUTAN, T., NAJMI, M., KAZEMI, N., HASANLOU, M. and PEDRAM, A., 2015. Lower Oncogenic Potential of Human Mesenchymal Stem Cells Derived from Cord Blood Compared to Induced Pluripotent Stem Cells. International journal of organ transplantation medicine, 6(3), pp. 99-104.

FREIMUTH, P., SPRINGER, K., BERARD, C., HAINFELD, J., BEWLEY, M. and FLANAGAN, J., 1999. Coxsackievirus and adenovirus receptor amino-terminal immunoglobulin V-related domain binds adenovirus type 2 and fiber knob from adenovirus type 12. Journal of virology, 73(2), pp. 1392-1398.

FREYER, G.A., KATOH, Y. and ROBERTS, R.J., 1984. Characterization of the major mRNAs from adenovirus 2 early region 4 by cDNA cloning and sequencing. Nucleic acids research, 12(8), pp. 3503-3519.

FRIETZE, K.M., CAMPOS, S.K. and KAJON, A.E., 2012. No evidence of a death-like function for species B1 human adenovirus type 3 E3-9K during A549 cell line infection. BMC research notes, 5, pp. 429-0500-5-429.

FUKUDA, K., ABEI, M., UGAI, H., SEO, E., WAKAYAMA, M., MURATA, T., TODOROKI, T., TANAKA, N., HAMADA, H. and YOKOYAMA, K.K., 2003. E1A, E1B double-restricted adenovirus for oncolytic gene therapy of gallbladder cancer. Cancer research, 63(15), pp. 4434-4440.

FUKUDA-TAIRA, S., 1981. Hepatic induction in the avian embryo: specificity of reactive endoderm and inductive mesoderm. Journal of embryology and experimental morphology, 63, pp. 111-125.

FURUICHI, Y., 1977. An amazing sequence arrangement at the 5' ends of adenovirus-2 messenger RNA (author's transl). Tanpakushitsu kakusan koso.Protein, nucleic acid, enzyme, 22(10), pp. 1201-1204.

FUSAKI, N., BAN, H., NISHIYAMA, A., SAEKI, K. and HASEGAWA, M., 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proceedings of the Japan Academy, Series B, 85(8), pp. 348-362.

GAGGAR, A., SHAYAKHMETOV, D. and LIEBER, A., 2007. Identifying functional adenovirus-host interactions using tandem mass spectrometry. Methods in Molecular Medicine, 131, pp. 141-155.

GAHERY-SEGARD, H., FARACE, F., GODFRIN, D., GASTON, J., LENGAGNE, R., TURSZ, T., BOULANGER, P. and GUILLET, J.G., 1998. Immune response to recombinant capsid proteins of adenovirus in humans: antifiber and anti-penton base antibodies have a synergistic effect on neutralizing activity. Journal of virology, 72(3), pp. 2388-2397.

GALL, A., FERNS, B., MORRIS, C., WATSON, S., COTTEN, M., ROBINSON, M., BERRY, N., PILLAY, D. and KELLAM, P., 2012. Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes. Journal of clinical microbiology, 50(12), pp. 3838-3844.

GAO, Y., ZHANG, X., ZHANG, L., CEN, J., NI, X., LIAO, X., YANG, C., LI, Y., CHEN, X., ZHANG, Z., SHU, Y., CHENG, X., HAY, D.C., LAI, D., PAN, G., WEI, G. and HUI, L., 2017. Distinct Gene Expression and Epigenetic Signatures in Hepatocyte-like Cells Produced by Different Strategies from the Same Donor. Stem cell reports, 9(6), pp. 1813-1824.

GASPAR, H.B., COORAY, S., GILMOUR, K.C., PARSLEY, K.L., ADAMS, S., HOWE, S.J., AL GHONAIUM, A., BAYFORD, J., BROWN, L., DAVIES, E.G., KINNON, C. and THRASHER, A.J., 2011. Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. Science translational medicine, 3(97), pp. 97ra79.

GAVRILESCU, L.C. and VAN ETTEN, R.A., 2007. Production of replication-defective retrovirus by transient transfection of 293T cells. Journal of visualized experiments : JoVE, (10):550. doi(10), pp. 550.

GERETS, H.H., TILMANT, K., GERIN, B., CHANTEUX, H., DEPELCHIN, B.O., DHALLUIN, S. and ATIENZAR, F.A., 2012. Characterization of primary

human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. Cell biology and toxicology, 28(2), pp. 69-87.

GERTOW, K., CEDERVALL, J., UNGER, C., SZOKE, K., BLENNOW, E., IMREH, M.P. and AHRLUND-RICHTER, L., 2007. Trisomy 12 in HESC leads to no selective in vivo growth advantage in teratomas, but induces an increased abundance of renal development. Journal of cellular biochemistry, 100(6), pp. 1518-1525.

GHERGHICEANU, M., BARAD, L., NOVAK, A., REITER, I., ITSKOVITZ-ELDOR, J., BINAH, O. and POPESCU, L.M., 2011. Cardiomyocytes derived from human embryonic and induced pluripotent stem cells: comparative ultrastructure. Journal of Cellular and Molecular Medicine, 15(11), pp. 2539-2551.

GINSBERG, H.S., PEREIRA, H.G., VALENTINE, R.C. and WILCOX, W.C., 1966. A proposed terminology for the adenovirus antigens and virion morphological subunits. Virology, 28(4), pp. 782-783.

GIORDANO, F.A., APPELT, J.U., LINK, B., GERDES, S., LEHRER, C., SCHOLZ, S., PARUZYNSKI, A., ROEDER, I., WENZ, F., GLIMM, H., VON KALLE, C., GREZ, M., SCHMIDT, M. and LAUFS, S., 2015. High-throughput monitoring of integration site clonality in preclinical and clinical gene therapy studies. Molecular therapy.Methods & clinical development, 2, pp. 14061.

GLASEL, J.A., 1995. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. BioTechniques, 18(1), pp. 62-63.

GODINI, R., LAFTA, H.Y. and FALLAHI, H., 2018. Epigenetic modifications in the embryonic and induced pluripotent stem cells. Gene expression patterns : GEP, 29, pp. 1-9.

GONZALEZ, F., BARRAGAN MONASTERIO, M., TISCORNIA, G., MONTSERRAT PULIDO, N., VASSENA, R., BATLLE MORERA, L., RODRIGUEZ PIZA, I. and IZPISUA BELMONTE, J.C., 2009. Generation of mouse-induced pluripotent stem cells by transient expression of a single nonviral polycistronic vector. Proceedings of the National Academy of Sciences of the United States of America, 106(22), pp. 8918-8922.

GORE, A., LI, Z., FUNG, H.L., YOUNG, J.E., AGARWAL, S., ANTOSIEWICZ-BOURGET, J., CANTO, I., GIORGETTI, A., ISRAEL, M.A., KISKINIS, E., LEE, J.H., LOH, Y.H., MANOS, P.D., MONTSERRAT, N., PANOPOULOS, A.D., RUIZ, S., WILBERT, M.L., YU, J., KIRKNESS, E.F., IZPISUA BELMONTE, J.C., ROSSI, D.J., THOMSON, J.A., EGGAN, K., DALEY, G.Q., GOLDSTEIN, L.S. and ZHANG, K., 2011. Somatic coding mutations in human induced pluripotent stem cells. Nature, 471(7336), pp. 63-67.

GOUVARCHIN GHALEH, H.E., BOLANDIAN, M., DOROSTKAR, R., JAFARI, A. and POUR, M.F., 2020. Concise review on optimized methods in production and transduction of lentiviral vectors in order to facilitate immunotherapy and gene therapy. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 128, pp. 110276.

GRAHAM, F.L., SMILEY, J., RUSSELL, W.C. and NAIRN, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. Journal of General Virology, 36(1), pp. 59-74.

GREBER, U.F., WILLETTS, M., WEBSTER, P. and HELENIUS, A., 1993. Stepwise dismantling of adenovirus 2 during entry into cells. Cell, 75(3), pp. 477-486.

GRIFFITH, F., 1928. The Significance of Pneumococcal Types. The Journal of hygiene, 27(2), pp. 113-159.

GUO, F., SUN, Y., WANG, X., WANG, H., WANG, J., GONG, T., CHEN, X., ZHANG, P., SU, L., FU, G., SU, J., YANG, S., LAI, R., JIANG, C. and LIANG, P., 2019. Patient-Specific and Gene-Corrected Induced Pluripotent Stem Cell-Derived Cardiomyocytes Elucidate Single-Cell Phenotype of Short QT Syndrome. Circulation research, 124(1), pp. 66-78.

GUO, P., EL-GOHARY, Y., PRASADAN, K., SHIOTA, C., XIAO, X., WIERSCH, J., PAREDES, J., TULACHAN, S. and GITTES, G.K., 2012. Rapid and simplified purification of recombinant adeno-associated virus. Journal of virological methods, 183(2), pp. 139-146.

GUO, P., XIAO, X., EL-GOHARY, Y., PAREDES, J., PRASADAN, K., SHIOTA, C., WIERSCH, J., WELSH, C. and GITTES, G.K., 2013. A simplified purification method for AAV variant by polyethylene glycol aqueous two-phase partitioning. Bioengineered, 4(2), pp. 103-106.

GURDON , J.B., 1960. The developmental capacity of nuclei taken from differentiating endoderm cells of Xenopus laevis. Journal of embryology and experimental morphology, 8, pp. 505-526.

HACEIN-BEY, H., CAVAZZANA-CALVO, M., LE DEIST, F., DAUTRY-VARSAT, A., HIVROZ, C., RIVIERE, I., DANOS, O., HEARD, J.M., SUGAMURA, K., FISCHER, A. and DE SAINT BASILE, G., 1996. gamma-c gene transfer into SCID X1 patients' B-cell lines restores normal high-affinity interleukin-2 receptor expression and function. Blood, 87(8), pp. 3108-3116.

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. and 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. 3132-3142.

HACEIN-BEY-ABINA, S., VON KALLE, C., SCHMIDT, M., MCCORMACK, M.P., WULFFRAAT, N., LEBOULCH, P., LIM, A., OSBORNE, C.S., PAWLIUK, R., MORILLON, E., SORENSEN, R., FORSTER, A., FRASER, P., COHEN, J.I., DE SAINT BASILE, G., ALEXANDER, I., WINTERGERST, U., FREBOURG, T., AURIAS, A., STOPPA-LYONNET, D., ROMANA, S., RADFORD-WEISS, I., GROSS, F., VALENSI, F., DELABESSE, E., MACINTYRE, E., SIGAUX, F., SOULIER, J., LEIVA, L.E., WISSLER, M., PRINZ, C., RABBITTS, T.H., LE DEIST, F., FISCHER, A. and CAVAZZANA-CALVO, M., 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science (New York, N.Y.), 302(5644), pp. 415-419.

HARJUMAKI, R., NUGROHO, R.W.N., ZHANG, X., LOU, Y.R., YLIPERTTULA, M., VALLE-DELGADO, J.J. and OSTERBERG, M., 2020. Author Correction: Quantified forces between HepG2 hepatocarcinoma and WA07 pluripotent stem cells with natural biomaterials correlate with in vitro cell behavior. Scientific reports, 10(1), pp. 8803-020-65140-8.

HARRICH, D., GARCIA, J., WU, F., MITSUYASU, R., GONAZALEZ, J. and GAYNOR, R., 1989. Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. Journal of virology, 63(6), pp. 2585-2591.

HARTMAN, Z.C., KIANG, A., EVERETT, R.S., SERRA, D., YANG, X.Y., CLAY, T.M. and AMALFITANO, A., 2007. Adenovirus infection triggers a rapid, MyD88-regulated transcriptome response critical to acute-phase and adaptive immune responses in vivo. Journal of virology, 81(4), pp. 1796-1812.

HASSALL, M.M., BARNARD, A.R. and MACLAREN, R.E., 2017. Gene Therapy for Color Blindness. The Yale journal of biology and medicine, 90(4), pp. 543-551.

HASSAN, R., HUSIN, A., SULONG, S., YUSOFF, S., JOHAN, M.F., YAHAYA, B.H., ANG, C.Y., GHAZALI, S., CHEONG, S.K. and DEPARTMENT OF STANDARD MALAYSIA, MOSTI, 2015. Guidelines for nucleic acid detection and analysis in hematological disorders. The Malaysian journal of pathology, 37(2), pp. 165-173.

HATTI-KAUL, R., 2001. Aqueous two-phase systems. A general overview. Molecular Biotechnology, 19(3), pp. 269-277.

HATTI-KAUL, R., 2000. Aqueous Two-Phase Systems: Methods and Protocols. Berlin: Springer Science & Business Media.

HAY, M., THOMAS, D.W., CRAIGHEAD, J.L., ECONOMIDES, C. and ROSENTHAL, J., 2014. Clinical development success rates for investigational drugs. Nature biotechnology, 32(1), pp. 40-51.

HAY, R.T., 1985. The origin of adenovirus DNA replication: minimal DNA sequence requirement in vivo. The EMBO journal, 4(2), pp. 421-426.

HEBBERECHT, L., VANCOILLIE, L., SCHAUVLIEGE, M., STAELENS, D., DEMECHELEER, E., HARDY, J., MORTIER, V. and VERHOFSTEDE, C., 2019. Single genome sequencing of near full-length HIV-1 RNA using a limiting dilution approach. Journal of virological methods, 274, pp. 113737.

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. and CALMELS, B.,

2004. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS biology, 2(12), pp. e423.

HENN, M.R., BOUTWELL, C.L., CHARLEBOIS, P., LENNON, N.J., POWER, K.A., MACALALAD, A.R., BERLIN, A.M., MALBOEUF, C.M., RYAN, E.M., GNERRE, S., ZODY, M.C., ERLICH, R.L., GREEN, L.M., BERICAL, A., WANG, Y., CASALI, M., STREECK, H., BLOOM, A.K., DUDEK, T., TULLY, D., NEWMAN, R., AXTEN, K.L., GLADDEN, A.D., BATTIS, L., KEMPER, M., ZENG, Q., SHEA, T.P., GUJJA, S., ZEDLACK, C., GASSER, O., BRANDER, C., HESS, C., GUNTHARD, H.F., BRUMME, Z.L., BRUMME, C.J., BAZNER, S., RYCHERT, J., TINSLEY, J.P., MAYER, K.H., ROSENBERG, E., PEREYRA, F., LEVIN, J.Z., YOUNG, S.K., JESSEN, H., ALTFELD, M., BIRREN, B.W., WALKER, B.D. and ALLEN, T.M., 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS pathogens, 8(3), pp. e1002529.

HENRY, M., HAWKINS, J.R., BOYLE, J. and BRIDGER, J.M., 2019. The Genomic Health of Human Pluripotent Stem Cells: Genomic Instability and the Consequences on Nuclear Organization. Frontiers in genetics, 9(623),.

HILLEMAN, M.R. and WERNER, J.H., 1954. Recovery of new agent from patients with acute respiratory illness. Proceedings of the Society for Experimental Biology and Medicine.Society for Experimental Biology and Medicine (New York, N.Y.), 85(1), pp. 183-188.

HISLOP, J.N., ISLAM, T.A., ELEFTHERIADOU, I., CARPENTIER, D.C., TRABALZA, A., PARKINSON, M., SCHIAVO, G. and MAZARAKIS, N.D., 2014. Rabies virus envelope glycoprotein targets lentiviral vectors to the axonal retrograde pathway in motor neurons. The Journal of biological chemistry, 289(23), pp. 16148-16163.

HIWASA, K., NAGAYA, H., TERAO, S., ACHARYA, B., HAMADA, K., MIZUGUCHI, H. and GOTOH, A., 2012. Improved gene transfer into bladder cancer cells using adenovirus vector containing RGD motif. Anticancer Research, 32(8), pp. 3137-3140.

HO SUI, S.J., FULTON, D.L., ARENILLAS, D.J., KWON, A.T. and WASSERMAN, W.W., 2007. oPOSSUM: integrated tools for analysis of regulatory motif over-representation. Nucleic acids research, 35(Web Server issue), pp. W245-52.

HOFFMANN, M., WU, Y.J., GERBER, M., BERGER-RENTSCH, M., HEIMRICH, B., SCHWEMMLE, M. and 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. 2782-2793.

HOGG, R.T., GARCIA, J.A. and GERARD, R.D., 2010. Adenoviral targeting of gene expression to tumors. Cancer gene therapy, 17(6), pp. 375-386.

HOHNE, K., BUSINGER, R., VAN NUFFEL, A., BOLDUAN, S., KOPPENSTEINER, H., BAEYENS, A., VERMEIRE, J., MALATINKOVA, E., VERHASSELT, B. and SCHINDLER, M., 2016. Virion encapsidated HIV-1 Vpr induces NFAT to prime non-activated T cells for productive infection. Open biology, 6(7), pp. 10.1098/rsob.160046.

HOWE, S.J., MANSOUR, M.R., SCHWARZWAELDER, K., BARTHOLOMAE, C., HUBANK, M., KEMPSKI, H., BRUGMAN, M.H., PIKE-OVERZET, K., CHATTERS, S.J., DE RIDDER, D., GILMOUR, K.C., ADAMS, S., THORNHILL, S.I., PARSLEY, K.L., STAAL, F.J., GALE, R.E., LINCH, D.C., BAYFORD, J., BROWN, L., QUAYE, M., KINNON, C., ANCLIFF, P., WEBB, D.K., SCHMIDT, M., VON KALLE, C., GASPAR, H.B. and THRASHER, A.J., 2008. Insertional mutagenesis combined with

acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. The Journal of clinical investigation, 118(9), pp. 3143-3150.

HUANG DA, W., SHERMAN, B.T. and LEMPICKI, R.A., 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research, 37(1), pp. 1-13.

HUANG DA, W., SHERMAN, B.T. and LEMPICKI, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols, 4(1), pp. 44-57.

HUGHES, S.H., 2015. Reverse Transcription of Retroviruses and LTR Retrotransposons. Microbiology spectrum, 3(2), pp. MDNA3-0027-2014.

HURKO, O., MCKEE, L. and ZUURVELD, J.G., 1986. Transfection of human skeletal muscle cells with SV40 large T antigen gene coupled to a metallothionein promoter. Annals of Neurology, 20(5), pp. 573-582.

ICHIM, C.V. and WELLS, R.A., 2011. Generation of high-titer viral preparations by concentration using successive rounds of ultracentrifugation. Journal of translational medicine, 9, pp. 137-5876-9-137.

IKEDA, Y., TAKEUCHI, Y., MARTIN, F., COSSET, F.L., MITROPHANOUS, K. and COLLINS, M., 2003. Continuous high-titer HIV-1 vector production. Nature biotechnology, 21(5), pp. 569-572.

INAGAKI, K., FUESS, S., STORM, T.A., GIBSON, G.A., MCTIERNAN, C.F., KAY, M.A. and NAKAI, H., 2006. Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. Molecular therapy : the journal of the American Society of Gene Therapy, 14(1), pp. 45-53.

INTRA, J. and SALEM, A.K., 2008. Characterization of the transgene expression generated by branched and linear polyethylenimine-plasmid DNA

nanoparticles in vitro and after intraperitoneal injection in vivo. Journal of controlled release : official journal of the Controlled Release Society, 130(2), pp. 129-138.

ITZHAKI, I., RAPOPORT, S., HUBER, I., MIZRAHI, I., ZWI-DANTSIS, L., ARBEL, G., SCHILLER, J. and GEPSTEIN, L., 2011. Calcium handling in human induced pluripotent stem cell derived cardiomyocytes. PloS one, 6(4), pp. e18037.

JACKSON, S.J. and THOMAS, G.J., 2017. Human tissue models in cancer research: looking beyond the mouse. Disease models & mechanisms, 10(8), pp. 939-942.

JALEEL, N., NAKAYAMA, H., CHEN, X., KUBO, H., MACDONNELL, S., ZHANG, H., BERRETTA, R., ROBBINS, J., CRIBBS, L., MOLKENTIN, J.D. and HOUSER, S.R., 2008. Ca2+ influx through T- and L-type Ca2+ channels have different effects on myocyte contractility and induce unique cardiac phenotypes. Circulation research, 103(10), pp. 1109-1119.

JIANG, H., WANG, Z., SERRA, D., FRANK, M.M. and AMALFITANO, A., 2004. Recombinant adenovirus vectors activate the alternative complement pathway, leading to the binding of human complement protein C3 independent of anti-ad antibodies. Molecular therapy : the journal of the American Society of Gene Therapy, 10(6), pp. 1140-1142.

JORDAN, A., BISGROVE, D. and VERDIN, E., 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. The EMBO journal, 22(8), pp. 1868-1877.

JORDAN, A., DEFECHEREUX, P. and VERDIN, E., 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. The EMBO journal, 20(7), pp. 1726-1738.

JUNG, C.B., MORETTI, A., MEDEROS Y SCHNITZLER, M., IOP, L., STORCH, U., BELLIN, M., DORN, T., RUPPENTHAL, S., PFEIFFER, S., GOEDEL, A., DIRSCHINGER, R.J., SEYFARTH, M., LAM, J.T., SINNECKER, D., GUDERMANN, T., LIPP, P. and LAUGWITZ, K.L., 2012. Dantrolene rescues arrhythmogenic RYR2 defect in a patient-specific stem cell model of catecholaminergic polymorphic ventricular tachycardia. EMBO molecular medicine, 4(3), pp. 180-191.

KADARI, A., LU, M., LI, M., SEKARAN, T., THUMMER, R.P., GUYETTE, N., CHU, V. and EDENHOFER, F., 2014. Excision of viral reprogramming cassettes by Cre protein transduction enables rapid, robust and efficient derivation of transgene-free human induced pluripotent stem cells. Stem cell research & therapy, 5(2), pp. 47.

KAJI, K., NORRBY, K., PACA, A., MILEIKOVSKY, M., MOHSENI, P. and WOLTJEN, K., 2009. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature, 458(7239), pp. 771-775.

KAMAKURA, T., MAKIYAMA, T., SASAKI, K., YOSHIDA, Y., WURIYANGHAI, Y., CHEN, J., HATTORI, T., OHNO, S., KITA, T., HORIE, M., YAMANAKA, S. and KIMURA, T., 2013. Ultrastructural maturation of human-induced pluripotent stem cell-derived cardiomyocytes in a long-term culture. Circulation journal : official journal of the Japanese Circulation Society, 77(5), pp. 1307-1314.

KANG, E. and YUN, C.O., 2010. Current advances in adenovirus nanocomplexes: more specificity and less immunogenicity. BMB reports, 43(12), pp. 781-788.

KAPOOR, Q.S., WOLD, W.S. and CHINNADURAI, G., 1981. A nonessential glycoprotein is coded by early region E3 of adenovirus type 7. Virology, 112(2), pp. 780-784.

KARDA, R., COUNSELL, J.R., KARBOWNICZEK, K., CAPRONI, L.J., TITE, J.P. and WADDINGTON, S.N., 2019. Production of lentiviral vectors using novel, enzymatically produced, linear DNA. Gene therapy, 26(3-4), pp. 86-92.

KAREN, K.A. and HEARING, P., 2011. Adenovirus core protein VII protects the viral genome from a DNA damage response at early times after infection. Journal of virology, 85(9), pp. 4135-4142.

KATO, S., KOBAYASHI, K., INOUE, K., KURAMOCHI, M., OKADA, T., YAGINUMA, H., MORIMOTO, K., SHIMADA, T., TAKADA, M. and KOBAYASHI, K., 2011. A lentiviral strategy for highly efficient retrograde gene transfer by pseudotyping with fusion envelope glycoprotein. Human Gene Therapy, 22(2), pp. 197-206.

KATTMAN, S.J., WITTY, A.D., GAGLIARDI, M., DUBOIS, N.C., NIAPOUR, M., HOTTA, A., ELLIS, J. and KELLER, G., 2011. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. Cell stem cell, 8(2), pp. 228-240.

KEHAT, I., KENYAGIN-KARSENTI, D., SNIR, M., SEGEV, H., AMIT, M., GEPSTEIN, A., LIVNE, E., BINAH, O., ITSKOVITZ-ELDOR, J. and GEPSTEIN, L., 2001. Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. The Journal of clinical investigation, 108(3), pp. 407-414.

KENNEDY, M.A. and PARKS, R.J., 2009. Adenovirus virion stability and the viral genome: size matters. Molecular therapy : the journal of the American Society of Gene Therapy, 17(10), pp. 1664-1666.

KERIEL, A., RENE, C., GALER, C., ZABNER, J. and KREMER, E.J., 2006. Canine adenovirus vectors for lung-directed gene transfer: efficacy, immune

response, and duration of transgene expression using helper-dependent vectors. Journal of virology, 80(3), pp. 1487-1496.

KERKAR, S.P., SANCHEZ-PEREZ, L.P., YANG, S., BORMAN, Z.A., MURANSKI, P., JI, Y., CHINNASAMY, D., KAISER, A.D.M., HINRICHS, C.S., KLEBANOFF, C.A., SCOTT, C.D., GATTINONI, L., MORGAN, R.A., ROSENBERG, S.A. and RESTIFO, N.P., 2011. Genetic engineering of murine CD8+ and CD4+ T cells for preclinical adoptive immunotherapy studies. Journal of Immunotherapy, 34(4), pp. 343-352.

KHONSARI, H., SCHNEIDER, M., AL-MAHDAWI, S., CHIANEA, Y.G., THEMIS, M., PARRIS, C., POOK, M.A. and THEMIS, M., 2016. Lentivirus-meditated frataxin gene delivery reverses genome instability in Friedreich ataxia patient and mouse model fibroblasts. Gene therapy, 23(12), pp. 846-856.

KIM, K., DOI, A., WEN, B., NG, K., ZHAO, R., CAHAN, P., KIM, J., ARYEE, M.J., JI, H., ERLICH, L., YABUUCHI, A., TAKEUCHI, A., CUNNIFF, K.C., HONGGUANG, H., MCKINNEY-FREEMAN, S., NAVEIRAS, O., YOON, T.J., IRIZARRY, R.A., JUNG, N., SEITA, J., HANNA, J., MURAKAMI, P., JAENISCH, R., WEISSLEDER, R., ORKIN, S.H., WEISSMAN, I.L., FEINBERG, A.P. and DALEY, G.Q., 2010. Epigenetic memory in induced pluripotent stem cells. Nature, 467(7313), pp. 285-290.

KIM, C., WONG, J., WEN, J., WANG, S., WANG, C., SPIERING, S., KAN, N.G., FORCALES, S., PURI, P.L., LEONE, T.C., MARINE, J.E., CALKINS, H., KELLY, D.P., JUDGE, D.P. and CHEN, H.S., 2013. Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. Nature, 494(7435), pp. 105-110.

KIM, I.Y., JOACHIM, E., CHOI, H. and KIM, K., 2015. Toxicity of silica nanoparticles depends on size, dose, and cell type. Nanomedicine : nanotechnology, biology, and medicine, 11(6), pp. 1407-1416.

KIM, S.Y., BYRN, R., GROOPMAN, J. and BALTIMORE, D., 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. Journal of virology, 63(9), pp. 3708-3713.

KING, T.J. and BRIGGS, R., 1955. Changes in the nuclei of differentiating endoderm cells as revealed by nuclear transplantation. Proc.Natl.Acad.Sci.U.S.A., 41(5), pp. 321-325.

KINT, S., TRYPSTEEN, W., DE SPIEGELAERE, W., MALATINKOVA, E., KINLOCH-DE LOES, S., DE MEYER, T., VAN CRIEKINGE, W. and VANDEKERCKHOVE, L., 2020. Underestimated effect of intragenic HIV-1 DNA methylation on viral transcription in infected individuals. Clinical epigenetics, 12(1), pp. 36-020-00829-1.

KIRCHNER, J., CONNOLLY, C.M. and SANDMEYER, S.B., 1995. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. Science (New York, N.Y.), 267(5203), pp. 1488-1491.

KIRSHENBAUM, L.A., 1997. Adenovirus mediated-gene transfer into cardiomyocytes. Molecular and cellular biochemistry, 172(1-2), pp. 13-21.

KLIMCZAK, M., 2015. Oncogenesis and induced pluripotency - commonalities of signalling pathways. Contemporary oncology (Poznan, Poland), 19(1A), pp. A16-21.

KLINGSEISEN, L., EHRENSCHWENDER, M., HEIGL, U., WAJANT, H., HEHLGANS, T., SCHUTZE, S. and SCHNEIDER-BRACHERT, W., 2012.

E3-14.7K is recruited to TNF-receptor 1 and blocks TNF cytolysis independent from interaction with optineurin. PloS one, 7(6), pp. e38348.

KNYAZHANSKAYA, E., ANISENKO, A., SHADRINA, O., KALININA, A., ZATSEPIN, T., ZALEVSKY, A., MAZUROV, D. and GOTTIKH, M., 2019. NHEJ pathway is involved in post-integrational DNA repair due to Ku70 binding to HIV-1 integrase. Retrovirology, 16(1), pp. 30-019-0492-z.

KOBELT, D., SCHLEEF, M., SCHMEER, M., AUMANN, J., SCHLAG, P.M. and WALTHER, W., 2013. Performance of high quality minicircle DNA for in vitro and in vivo gene transfer. Molecular biotechnology, 53(1), pp. 80-89.

KOHAMA, Y., HIGO, S., MASUMURA, Y., SHIBA, M., KONDO, T., ISHIZU, T., HIGO, T., NAKAMURA, S., KAMEDA, S., TABATA, T., INOUE, H., MOTOOKA, D., OKUZAKI, D., TAKASHIMA, S., MIYAGAWA, S., SAWA, Y., HIKOSO, S. and SAKATA, Y., 2020. Adeno-associated virus-mediated gene delivery promotes S-phase entry-independent precise targeted integration in cardiomyocytes. Scientific reports, 10(1), pp. 15348-020-72216-y.

KONG, K., KUMAR, M., TARUISHI, M. and JAVIER, R.T., 2015. Adenovirus E4-ORF1 Dysregulates Epidermal Growth Factor and Insulin/Insulin-Like Growth Factor Receptors To Mediate Constitutive Myc Expression. Journal of virology, 89(21), pp. 10774-10785.

KONKALMATT, P.R., WANG, F., PIRAS, B.A., XU, Y., O'CONNOR, D.M., BEYERS, R.J., EPSTEIN, F.H., ANNEX, B.H., HOSSACK, J.A. and FRENCH, B.A., 2012. Adeno-associated virus serotype 9 administered systemically after reperfusion preferentially targets cardiomyocytes in the infarct border zone with pharmacodynamics suitable for the attenuation of left ventricular remodeling. The journal of gene medicine, 14(9-10), pp. 609-620.

KRATZER, R.F. and KREPPEL, F., 2017. Production, Purification, and Titration of First-Generation Adenovirus Vectors. Methods in molecular biology (Clifton, N.J.), 1654, pp. 377-388.

KULKARNI, J.A., MYHRE, J.L., CHEN, S., TAM, Y.Y.C., DANESCU, A., RICHMAN, J.M. and CULLIS, P.R., 2017. Design of lipid nanoparticles for in vitro and in vivo delivery of plasmid DNA. Nanomedicine : nanotechnology, biology, and medicine, 13(4), pp. 1377-1387.

KUMAR, M., KELLER, B., MAKALOU, N. and SUTTON, R.E., 2001. Systematic determination of the packaging limit of lentiviral vectors. Human Gene Therapy, 12(15), pp. 1893-1905.

KUTNER, R.H., PUTHLI, S., MARINO, M.P. and REISER, J., 2009. Simplified production and concentration of HIV-1-based lentiviral vectors using HYPERFlask vessels and anion exchange membrane chromatography. BMC biotechnology, 9, pp. 10-6750-9-10.

KUTNER, R.H., ZHANG, X.Y. and REISER, J., 2009. Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. Nature protocols, 4(4), pp. 495-505.

KWAK, Y.T., YANG, Y. and PARK, S.Y., 2015. Chloroquine-associated psychosis mimicking very late-onset schizophrenia: Case Report. Geriatrics & gerontology international, 15(8), pp. 1096-1097.

KWON, A.T., ARENILLAS, D.J., WORSLEY HUNT, R. and WASSERMAN, W.W., 2012. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. G3 (Bethesda, Md.), 2(9), pp. 987-1002.

LAFAVE, M.C., VARSHNEY, G.K., GILDEA, D.E., WOLFSBERG, T.G., BAXEVANIS, A.D. and BURGESS, S.M., 2014. MLV integration site

selection is driven by strong enhancers and active promoters. Nucleic acids research, 42(7), pp. 4257-4269.

LAFLAMME, M.A., CHEN, K.Y., NAUMOVA, A.V., MUSKHELI, V., FUGATE, J.A., DUPRAS, S.K., REINECKE, H., XU, C., HASSANIPOUR, M., POLICE, S., O'SULLIVAN, C., COLLINS, L., CHEN, Y., MINAMI, E., GILL, E.A., UENO, S., YUAN, C., GOLD, J. and MURRY, C.E., 2007. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. Nature biotechnology, 25(9), pp. 1015-1024.

LAN, F., LEE, A.S., LIANG, P., SANCHEZ-FREIRE, V., NGUYEN, P.K., WANG, L., HAN, L., YEN, M., WANG, Y., SUN, N., ABILEZ, O.J., HU, S., EBERT, A.D., NAVARRETE, E.G., SIMMONS, C.S., WHEELER, M., PRUITT, B., LEWIS, R., YAMAGUCHI, Y., ASHLEY, E.A., BERS, D.M., ROBBINS, R.C., LONGAKER, M.T. and WU, J.C., 2013. Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells. Cell stem cell, 12(1), pp. 101-113.

LAUREN, K., BINGNAM, 2011. Amplification of Cre- expressing Adenovirus in 293T cells. Malvern, USA: Vector Biolabs.

LAUSTRIAT, D., GIDE, J. and PESCHANSKI, M., 2010. Human pluripotent stem cells in drug discovery and predictive toxicology. Biochemical Society transactions, 38(4), pp. 1051-1057.

LAWSON, N.D., STILLMAN, E.A., WHITT, M.A. and ROSE, J.K., 1995. Recombinant vesicular stomatitis viruses from DNA. Proceedings of the National Academy of Sciences of the United States of America, 92(10), pp. 4477-4481.

LEE, J.Y. and LEE, H.H., 2018. A new chemical complex can rapidly concentrate lentivirus and significantly enhance gene transduction. Cytotechnology, 70(1), pp. 193-201.

LEE, K., AMBROSE, Z., MARTIN, T.D., OZTOP, I., MULKY, A., JULIAS, J.G., VANDEGRAAFF, N., BAUMANN, J.G., WANG, R., YUEN, W., TAKEMURA, T., SHELTON, K., TANIUCHI, I., LI, Y., SODROSKI, J., LITTMAN, D.R., COFFIN, J.M., HUGHES, S.H., UNUTMAZ, D., ENGELMAN, A. and KEWALRAMANI, V.N., 2010. Flexible use of nuclear import pathways by HIV-1. Cell host & microbe, 7(3), pp. 221-233.

LEE, M.S. and CRAIGIE, R., 1998. A previously unidentified host protein protects retroviral DNA from autointegration. Proceedings of the National Academy of Sciences of the United States of America, 95(4), pp. 1528-1533.

LEOPOLD, P.L., FERRIS, B., GRINBERG, I., WORGALL, S., HACKETT, N.R. and CRYSTAL, R.G., 1998. Fluorescent virions: dynamic tracking of the pathway of adenoviral gene transfer vectors in living cells. Human Gene Therapy, 9(3), pp. 367-378.

LEOPOLD, P.L., KREITZER, G., MIYAZAWA, N., REMPEL, S., PFISTER, K.K., RODRIGUEZ-BOULAN, E. and CRYSTAL, R.G., 2000. Dynein- and microtubule-mediated translocation of adenovirus serotype 5 occurs after endosomal lysis. Human Gene Therapy, 11(1), pp. 151-165.

LESCH, H.P., HEIKKILA, K.M., LIPPONEN, E.M., VALONEN, P., MULLER, A., RASANEN, E., TUUNANEN, T., HASSINEN, M.M., PARKER, N., KARHINEN, M., SHAW, R. and YLA-HERTTUALA, S., 2015. Process Development of Adenoviral Vector Production in Fixed Bed Bioreactor: From Bench to Commercial Scale. Human Gene Therapy, 26(8), pp. 560-571.

LESCH, H.P., LAITINEN, A., PEIXOTO, C., VICENTE, T., MAKKONEN, K.E., LAITINEN, L., PIKKARAINEN, J.T., SAMARANAYAKE, H., ALVES, P.M., CARRONDO, M.J., YLA-HERTTUALA, S. and AIRENNE, K.J., 2011. Production and purification of lentiviral vectors generated in 293T suspension cells with baculoviral vectors. Gene therapy, 18(6), pp. 531-538.

LEWINSKI, M.K., YAMASHITA, M., EMERMAN, M., CIUFFI, A., MARSHALL, H., CRAWFORD, G., COLLINS, F., SHINN, P., LEIPZIG, J., HANNENHALLI, S., BERRY, C.C., ECKER, J.R. and BUSHMAN, F.D., 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS pathogens, 2(6), pp. e60.

LEWIS, P., HENSEL, M. and EMERMAN, M., 1992. Human immunodeficiency virus infection of cells arrested in the cell cycle. The EMBO journal, 11(8), pp. 3053-3058.

LI, L., OLVERA, J.M., YODER, K.E., MITCHELL, R.S., BUTLER, S.L., LIEBER, M., MARTIN, S.L. and BUSHMAN, F.D., 2001. Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. The EMBO journal, 20(12), pp. 3272-3281.

LI, Z., SHARMA, R.V., DUAN, D. and DAVISSON, R.L., 2003. Adenovirus-mediated gene transfer to adult mouse cardiomyocytes is selectively influenced by culture medium. The journal of gene medicine, 5(9), pp. 765-772.

LIAN, Q., ZHANG, Y., ZHANG, J., ZHANG, H.K., WU, X., ZHANG, Y., LAM, F.F., KANG, S., XIA, J.C., LAI, W.H., AU, K.W., CHOW, Y.Y., SIU, C.W., LEE, C.N. and TSE, H.F., 2010. Functional mesenchymal stem cells derived from human induced pluripotent stem cells attenuate limb ischemia in mice. Circulation, 121(9), pp. 1113-1123.

LIANG, P., LAN, F., LEE, A.S., GONG, T., SANCHEZ-FREIRE, V., WANG, Y., DIECKE, S., SALLAM, K., KNOWLES, J.W., WANG, P.J., NGUYEN,

P.K., BERS, D.M., ROBBINS, R.C. and WU, J.C., 2013. Drug screening using a library of human induced pluripotent stem cell-derived cardiomyocytes reveals disease-specific patterns of cardiotoxicity. Circulation, 127(16), pp. 1677-1691.

LIDONNICI, M.R., PALEARI, Y., TIBONI, F., MANDELLI, G., ROSSI, C., VEZZOLI, M., APRILE, A., LEDERER, C.W., AMBROSI, A., CHANUT, F., SANVITO, F., CALABRIA, A., POLETTI, V., MAVILIO, F., MONTINI, E., NALDINI, L., CRISTOFORI, P. and FERRARI, G., 2018. Multiple Integrated Non-clinical Studies Predict the Safety of Lentivirus-Mediated Gene Therapy for beta-Thalassemia. Molecular therapy.Methods & clinical development, 11, pp. 9-28.

LIEBER, A., HE, C.Y., MEUSE, L., SCHOWALTER, D., KIRILLOVA, I., WINTHER, B. and KAY, M.A., 1997. The role of Kupffer cell activation and viral gene expression in early liver toxicity after infusion of recombinant adenovirus vectors. Journal of virology, 71(11), pp. 8798-8807.

LIU, Y., CHENG, D., LI, Z., GAO, X. and WANG, H., 2012. The gene expression profiles of induced pluripotent stem cells (iPSCs)generated by a non-integrating method are more similar to embryonic stem cells than those of iPSCs generated by an integrating method. Genetics and Molecular Biology, 35(3), pp. 693-700.

LIU, F. and GREEN, M.R., 1990. A specific member of the ATF transcription factor family can mediate transcription activation by the adenovirus E1a protein. Cell, 61(7), pp. 1217-1224.

LIU, M., MAURANO, M.T., WANG, H., QI, H., SONG, C.Z., NAVAS, P.A., EMERY, D.W., STAMATOYANNOPOULOS, J.A. and STAMATOYANNOPOULOS, G., 2015. Genomic discovery of potent

chromatin insulators for human gene therapy. Nature biotechnology, 33(2), pp. 198-203.

LOGAN, J. and SHENK, T., 1984. Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Proceedings of the National Academy of Sciences of the United States of America, 81(12), pp. 3655-3659.

LOWE, S.W. and RULEY, H.E., 1993. Stabilization of the p53 tumor suppressor is induced by adenovirus 5 E1A and accompanies apoptosis. Genes & development, 7(4), pp. 535-545.

LU, J., EINHORN, S., VENKATARANGAN, L., MILLER, M., MANN, D.A., WATKINS, P.B. and LECLUYSE, E., 2015. Morphological and Functional Characterization and Assessment of iPSC-Derived Hepatocytes for In Vitro Toxicity Testing. Toxicological sciences : an official journal of the Society of Toxicology, 147(1), pp. 39-54.

LUCENDO-VILLARIN, B., RASHIDI, H., ALHAQUE, S., FISCHER, L., MESEGUER-RIPOLLES, J., WANG, Y., O'FARRELLY, C., THEMIS, M. and HAY, D.C., 2019. Serum Free Production of Three-dimensional Human Hepatospheres from Pluripotent Stem Cells. Journal of visualized experiments : JoVE, (149). doi(149), pp. 10.3791/59965.

LUND, A.H., TURNER, G., TRUBETSKOY, A., VERHOEVEN, E., WIENTJENS, E., HULSMAN, D., RUSSELL, R., DEPINHO, R.A., LENZ, J. and VAN LOHUIZEN, M., 2002. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nature genetics, 32(1), pp. 160-165.

LUNDY, S.D., ZHU, W.Z., REGNIER, M. and LAFLAMME, M.A., 2013. Structural and functional maturation of cardiomyocytes derived from human pluripotent stem cells. Stem cells and development, 22(14), pp. 1991-2002.

LUTZ, P. and KEDINGER, C., 1996. Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter. Journal of virology, 70(3), pp. 1396-1405.

LYLES, D.S. and RUPPRECHT, C.E., 2007. Rhabdoviridae. In: D.M. KNIPE and P.M. HOWLE, eds, Fields Virology. Wilkins, PA, USA: Lippincott Williams, .

MA, C., DONG, X., LI, R. and LIU, L., 2013. A computational study identifies HIV progression-related genes using mRMR and shortest path tracing. PloS one, 8(11), pp. e78057.

MA, J., GUO, L., FIENE, S.J., ANSON, B.D., THOMSON, J.A., KAMP, T.J., KOLAJA, K.L., SWANSON, B.J. and JANUARY, C.T., 2011. High purity human-induced pluripotent stem cell-derived cardiomyocytes: electrophysiological properties of action potentials and ionic currents. American journal of physiology.Heart and circulatory physiology, 301(5), pp. H2006-17.

MACDONALD, J.S. and ROBERTSON, R.T., 2009. Toxicity testing in the 21st century: a view from the pharmaceutical industry. Toxicological sciences : an official journal of the Society of Toxicology, 110(1), pp. 40-46.

MADDAH, M., HEIDMANN, J.D., MANDEGAR, M.A., WALKER, C.D., BOLOUKI, S., CONKLIN, B.R. and LOEWKE, K.E., 2015. A non-invasive platform for functional characterization of stem-cell-derived cardiomyocytes with applications in cardiotoxicity testing. Stem cell reports, 4(4), pp. 621-631.

MAIER, O., GALAN, D.L., WODRICH, H. and WIETHOFF, C.M., 2010. An N-terminal domain of adenovirus protein VI fragments membranes by inducing positive membrane curvature. Virology, 402(1), pp. 11-19.

MALIK, N. and RAO, M.S., 2013. A review of the methods for human iPSC derivation. Methods in molecular biology (Clifton, N.J.), 997, pp. 23-33.

MAMMEN, B., RAMAKRISHNAN, T., SUDHAKAR, U. and VIJAYALAKSHMI, 2007. Principles of gene therapy. Indian Journal of Dental Research, 18(4), pp. 196-200.

MANN, K.P., WEISS, E.A. and NEVINS, J.R., 1993. Alternative poly(A) site utilization during adenovirus infection coincides with a decrease in the activity of a poly(A) site processing factor. Molecular and cellular biology, 13(4), pp. 2411-2419.

MANN, R., MULLIGAN, R.C. and BALTIMORE, D., 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell, 33(1), pp. 153-159.

MARCHETTO, M.C., CARROMEU, C., ACAB, A., YU, D., YEO, G.W., MU, Y., CHEN, G., GAGE, F.H. and MUOTRI, A.R., 2010. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell, 143(4), pp. 527-539.

MARCUCCI, K.T., JADLOWSKY, J.K., HWANG, W.T., SUHOSKI-DAVIS, M., GONZALEZ, V.E., KULIKOVSKAYA, I., GUPTA, M., LACEY, S.F., PLESA, G., CHEW, A., MELENHORST, J.J., LEVINE, B.L. and JUNE, C.H., 2018. Retroviral and Lentiviral Safety Analysis of Gene-Modified T Cell Products and Infused HIV and Oncology Patients. Molecular therapy : the journal of the American Society of Gene Therapy, 26(1), pp. 269-279.

MASUI, S., NAKATAKE, Y., TOYOOKA, Y., SHIMOSATO, D., YAGI, R., TAKAHASHI, K., OKOCHI, H., OKUDA, A., MATOBA, R., SHAROV, A.A., KO, M.S. and NIWA, H., 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nature cell biology, 9(6), pp. 625-635.

MATTIUZZO, G., ASHALL, J., DORIS, K.S., MACLELLAN-GIBSON, K., NICOLSON, C., WILKINSON, D.E., HARVEY, R., ALMOND, N.,

ANDERSON, R., EFSTATHIOU, S., MINOR, P.D. and PAGE, M., 2015. Development of Lentivirus-Based Reference Materials for Ebola Virus Nucleic Acid Amplification Technology-Based Assays. PloS one, 10(11), pp. e0142751.

MAUTNER, V. and PEREIRA, H.G., 1971. Crystallization of a Second Adenovirus Protein (the Fibre). Nature, 240, pp. 456-457.

MCCARTHY, K.F., 2003. Marrow frequency of rat long-term repopulating cells: evidence that marrow hematopoietic stem cell concentration may be inversely proportional to species body weight. Blood, 101(9), pp. 3431-3435.

MCMAHON, S.B., VAN BUSKIRK, H.A., DUGAN, K.A., COPELAND, T.D. and COLE, M.D., 1998. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell, 94(3), pp. 363-374.

MCNALLY, D.J., DARLING, D., FARZANEH, F., LEVISON, P.R. and SLATER, N.K., 2014. Optimised concentration and purification of retroviruses using membrane chromatography. Journal of chromatography.A, 1340, pp. 24-32.

MEKKAOUI, L., PAREKH, F., KOTSOPOULOU, E., DARLING, D., DICKSON, G., CHEUNG, G.W., CHAN, L., MACLELLAN-GIBSON, K., MATTIUZZO, G., FARZANEH, F., TAKEUCHI, Y. and PULE, M., 2018. Lentiviral Vector Purification Using Genetically Encoded Biotin Mimic in Packaging Cell. Molecular therapy.Methods & clinical development, 11, pp. 155-165.

MENTIS, G.Z., GRAVELL, M., HAMILTON, R., SHNEIDER, N.A., O'DONOVAN, M.J. and SCHUBERT, M., 2006. Transduction of motor neurons and muscle fibers by intramuscular injection of HIV-1-based vectors pseudotyped with select rabies virus glycoproteins. Journal of neuroscience methods, 157(2), pp. 208-217.

MERTEN, O.W., CHARRIER, S., LAROUDIE, N., FAUCHILLE, S., DUGUÉ, C., JENNY, C., AUDIT, M., ZANTA-BOUSSIF, M.A., CHAUTARD, H., RADRIZZANI, M., VALLANTI, G., NALDINI, L., NOGUIEZ-HELLIN, P. and GALY, A., 2011. Large-Scale Manufacture and Characterization of a Lentiviral Vector Produced for Clinical Ex Vivo Gene Therapy Application. Human Gene Therapy, 22(3), pp. 343-356.

MERTEN, O.W., 2015. Advances in cell culture: anchorage dependence. Philosophical transactions of the Royal Society of London.Series B, Biological sciences, 370(1661), pp. 20140040.

MIKOVITS, J.A., YOUNG, H.A., VERTINO, P., ISSA, J.P., PITHA, P.M., TURCOSKI-CORRALES, S., TAUB, D.D., PETROW, C.L., BAYLIN, S.B. and RUSCETTI, F.W., 1998. Infection with human immunodeficiency virus type 1 upregulates DNA methyltransferase, resulting in de novo methylation of the gamma interferon (IFN-gamma) promoter and subsequent downregulation of IFN-gamma production. Molecular and cellular biology, 18(9), pp. 5166-5177.

MILLER, M.D., WANG, B. and BUSHMAN, F.D., 1995. Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. Journal of virology, 69(6), pp. 3938-3944.

MISLICK, K.A. and BALDESCHWIELER, J.D., 1996. Evidence for the role of proteoglycans in cation-mediated gene transfer. Proceedings of the National Academy of Sciences of the United States of America, 93(22), pp. 12349-12354.

MITCHELL, R.S., BEITZEL, B.F., SCHRODER, A.R., SHINN, P., CHEN, H., BERRY, C.C., ECKER, J.R. and BUSHMAN, F.D., 2004. Retroviral DNA

integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS biology, 2(8), pp. E234.

MITSUI, M., NISHIKAWA, M., ZANG, L., ANDO, M., HATTORI, K., TAKAHASHI, Y., WATANABE, Y. and TAKAKURA, Y., 2009. Effect of the content of unmethylated CpG dinucleotides in plasmid DNA on the sustainability of transgene expression. The journal of gene medicine, 11(5), pp. 435-443.

MIYOSHI, H., BLÖMER, U., TAKAHASHI, M., GAGE, F.H. and VERMA, I.M., 1998. Development of a Self-Inactivating Lentivirus Vector. Journal of virology, 72(10), pp. 1850-1857.

MODLICH, U., NAVARRO, S., ZYCHLINSKI, D., MAETZIG, T., KNOESS, S., BRUGMAN, M.H., SCHAMBACH, A., CHARRIER, S., GALY, A., THRASHER, A.J., BUEREN, J. and 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. 1919-1928.

MOIANI, A., PALEARI, Y., SARTORI, D., MEZZADRA, R., MICCIO, A., CATTOGLIO, C., COCCHIARELLA, F., LIDONNICI, M.R., FERRARI, G. and MAVILIO, F., 2012. Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. The Journal of clinical investigation, 122(5), pp. 1653-1666.

MOISE, A.R., GRANT, J.R., VITALIS, T.Z. and JEFFERIES, W.A., 2002. Adenovirus E3-6.7K maintains calcium homeostasis and prevents apoptosis and arachidonic acid release. Journal of virology, 76(4), pp. 1578-1587.

MONTELL, C., COURTOIS, G., ENG, C. and BERK, A., 1984. Complete transformation by adenovirus 2 requires both E1A proteins. Cell, 36(4), pp. 951-961.

MONTINI, E., CESANA, D., SCHMIDT, M., SANVITO, F., BARTHOLOMAE, C.C., RANZANI, M., BENEDICENTI, F., SERGI, L.S., AMBROSI, A., PONZONI, M., DOGLIONI, C., DI SERIO, C., VON KALLE, C. and NALDINI, L., 2009. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. Journal of Clinical Investigation, 119(4), pp. 964-975.

MONTINI, E., CESANA, D., SCHMIDT, M., SANVITO, F., PONZONI, M., BARTHOLOMAE, C., SERGI, L.S., BENEDICENTI, F., AMBROSI, A., DI SERIO, C., DOGLIONI, C., VON KALLE, C. and NALDINI, L., 2006. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nature Biotechnology, 24(6), pp. 687-696.

MORDWINKIN, N.M., LEE, A.S. and WU, J.C., 2013. Patient-specific stem cells and cardiovascular drug discovery. Jama, 310(19), pp. 2039-2040.

MOREIRA, A.S., SILVA, A.C., SOUSA, M.F.Q., HAGNER-MCWHIRTERC, A., AHLENC, G., LUNDGREN, M., COROADINHA, A.S., ALVES, P.M., PEIXOTO, C. and CARRONDO, M.J.T., 2020. Establishing Suspension Cell Cultures for Improved Manufacturing of Oncolytic Adenovirus. Biotechnology journal, 15(4), pp. e1900411.

MORETTI, A., BELLIN, M., WELLING, A., JUNG, C.B., LAM, J.T., BOTT-FLUGEL, L., DORN, T., GOEDEL, A., HOHNKE, C., HOFMANN, F., SEYFARTH, M., SINNECKER, D., SCHOMIG, A. and LAUGWITZ, K.L., 2010. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. The New England journal of medicine, 363(15), pp. 1397-1409.

MULDER, L.C., CHAKRABARTI, L.A. and MUESING, M.A., 2002. Interaction of HIV-1 integrase with DNA repair protein hRad18. The Journal of biological chemistry, 277(30), pp. 27489-27493.

MUMMERY, C., WARD-VAN OOSTWAARD, D., DOEVENDANS, P., SPIJKER, R., VAN DEN BRINK, S., HASSINK, R., VAN DER HEYDEN, M., OPTHOF, T., PERA, M., DE LA RIVIERE, A.B., PASSIER, R. and TERTOOLEN, L., 2003. Differentiation of human embryonic stem cells to cardiomyocytes: role of coculture with visceral endoderm-like cells. Circulation, 107(21), pp. 2733-2740.

MURPHY, J.E., DE LOS SANTOS, T. and GOFF, S.P., 1993. Mutational analysis of the sequences at the termini of the Moloney murine leukemia virus DNA required for integration. Virology, 195(2), pp. 432-440.

MUUL, L.M., TUSCHONG, L.M., SOENEN, S.L., JAGADEESH, G.J., RAMSEY, W.J., LONG, Z., CARTER, C.S., GARABEDIAN, E.K., ALLEYNE, M., BROWN, M., BERNSTEIN, W., SCHURMAN, S.H., FLEISHER, T.A., LEITMAN, S.F., DUNBAR, C.E., BLAESE, R.M. and CANDOTTI, F., 2003. Persistence and expression of the adenosine deaminase gene for 12 years and immune reaction to gene transfer components: long-term results of the first clinical gene therapy trial. Blood, 101(7), pp. 2563-2569.

NAFISSI, N., SUM, C.H., WETTIG, S. and SLAVCEV, R.A., 2014. Optimization of a one-step heat-inducible in vivo mini DNA vector production system. PloS one, 9(2), pp. e89345.

NAKAGAWA, M., KOYANAGI, M., TANABE, K., TAKAHASHI, K., ICHISAKA, T., AOI, T., OKITA, K., MOCHIDUKI, Y., TAKIZAWA, N. and YAMANAKA, S., 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nature biotechnology, 26(1), pp. 101-106.

NALDINI, L., BLÖMER, U., GALLAY, P., ORY, D., MULLIGAN, R., GAGE, F.H., VERMA, I.M. and TRONO, D., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science, 12(272), pp. 263-267.

NATUNEN, S., SATOMAA, T., PITKANEN, V., SALO, H., MIKKOLA, M., NATUNEN, J., OTONKOSKI, T. and VALMU, L., 2011. The binding specificity of the marker antibodies Tra-1-60 and Tra-1-81 reveals a novel pluripotency-associated type 1 lactosamine epitope. Glycobiology, 21(9), pp. 1125-1130.

NEGRETE, A., LING, T.C. and LYDDIATT, A., 2007. Aqueous two-phase recovery of bio-nanoparticles: a miniaturization study for the recovery of bacteriophage T4. Journal of Chromatography B, 854(1, 2), pp. 13-19.

NEGRETE, A., LING, T.C. and LYDDIATT, A., 2007. Production of adenoviral vectors and its recovery. Process Biochemistry, 42(7), pp. 1107-1113.

NEMUNAITIS, J., GANLY, I., KHURI, F., ARSENEAU, J., KUHN, J., MCCARTY, T., LANDERS, S., MAPLES, P., ROMEL, L., RANDLEV, B., REID, T., KAYE, S. and KIRN, D., 2000. Selective replication and oncolysis in p53 mutant tumors with ONYX-015, an E1B-55kD gene-deleted adenovirus, in patients with advanced head and neck cancer: a phase II trial. Cancer research, 60(22), pp. 6359-6366.

NEUVILLE, A., COLLIN, F., BRUNEVAL, P., PARRENS, M., THIVOLET, F., GOMEZ-BROUCHET, A., TERRIER, P., DE MONTPREVILLE, V.T., LE GALL, F., HOSTEIN, I., LAGARDE, P., CHIBON, F. and COINDRE, J.M., 2014. Intimal sarcoma is the most frequent primary cardiac sarcoma: clinicopathologic and molecular retrospective analysis of 100 primary cardiac sarcomas. The American Journal of Surgical Pathology, 38(4), pp. 461-469.

NEVINS, J.R., 1981. Mechanism of activation of early viral transcription by the adenovirus E1A gene product. Cell, 26(2 Pt 2), pp. 213-220.

NEVINS, J.R. and WILSON, M.C., 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. Nature, 290(5802), pp. 113-118.

NICOLAS, J.C., SARNOW, P., GIRARD, M. and LEVINE, A.J., 1983. Host range temperature-conditional mutants in the adenovirus DNA binding protein are defective in the assembly of infectious virus. Virology, 126(1), pp. 228-239.

NIEDERER, H.A. and BANGHAM, C.R., 2014. Integration site and clonal expansion in human chronic retroviral infection and gene therapy. Viruses, 6(11), pp. 4140-4164.

NIELSCH, U., FOGNANI, C. and BABISS, L.E., 1991. Adenovirus E1A-p105(Rb) protein interactions play a direct role in the initiation but not the maintenance of the rodent cell transformed phenotype. Oncogene, 6(6), pp. 1031-1036.

NIELSEN, T.T., JAKOBSSON, J., ROSENQVIST, N. and LUNDBERG, C., 2009. Incorporating double copies of a chromatin insulator into lentiviral vectors results in less viral integrants. BMC biotechnology, 9, pp. 13-6750-9-13.

NIKOLAITCHIK, O.A., DILLEY, K.A., FU, W., GORELICK, R.J., TAI, S.H., SOHEILIAN, F., PTAK, R.G., NAGASHIMA, K., PATHAK, V.K. and HU, W.S., 2013. Dimeric RNA recognition regulates HIV-1 genome packaging. PLoS pathogens, 9(3), pp. e1003249.

NILSEN, T.W., MARONEY, P.A., GOODWIN, R.G., ROTTMAN, F.M., CRITTENDEN, L.B., RAINES, M.A. and KUNG, H.J., 1985. c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and promoter insertion

result in expression of an amino-truncated EGF receptor. Cell, 41(3), pp. 719-726.

NIWA, H., OGAWA, K., SHIMOSATO, D. and ADACHI, K., 2009. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature, 460(7251), pp. 118-122.

NOBEL ASSEMBLY, 2012-last update, The Nobel Prize in Physiology or Medicine 2012. Available: https://www.nobelprize.org/prizes/medicine/2012/press-release/ [February, 1, 2021].

NORMAN, A., RIBER, L., LUO, W., LI, L.L., HANSEN, L.H. and SORENSEN, S.J., 2014. An improved method for including upper size range plasmids in metamobilomes. PloS one, 9(8), pp. e104405.

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. and 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. 324-337.

NUSSBAUM, J., MINAMI, E., LAFLAMME, M.A., VIRAG, J.A., WARE, C.B., MASINO, A., MUSKHELI, V., PABON, L., REINECKE, H. and MURRY, C.E., 2007. Transplantation of undifferentiated murine embryonic stem cells in the heart: teratoma formation and immune response. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 21(7), pp. 1345-1357.

OCWIEJA, K.E., SHERRILL-MIX, S., MUKHERJEE, R., CUSTERS-ALLEN, R., DAVID, P., BROWN, M., WANG, S., LINK, D.R., OLSON, J., TRAVERS, K., SCHADT, E. and BUSHMAN, F.D., 2012. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. Nucleic acids research, 40(20), pp. 10345-10355.

OELMEIER, S.A., DISMER, F. and HUBBUCH, J., 2012. Molecular dynamics simulations on aqueous two-phase systems - Single PEG-molecules in solution. BMC biophysics, 5, pp. 14-1682-5-14.

OELMEIER, S.A., DISMER, F. and HUBBUCH, J., 2011. Application of an aqueous two-phase systems high-throughput screening method to evaluate mAb HCP separation. Biotechnology and bioengineering, 108(1), pp. 69-81.

OGRIS, M., WALKER, G., BLESSING, T., KIRCHEIS, R., WOLSCHEK, M. and WAGNER, E., 2003. Tumor-targeted gene therapy: strategies for the preparation of ligand-polyethylene glycol-polyethylenimine/DNA complexes. Journal of controlled release : official journal of the Controlled Release Society, 91(1-2), pp. 173-181.

OKITA, K., ICHISAKA, T. and YAMANAKA, S., 2007. Generation of germline-competent induced pluripotent stem cells. Nature, 448(7151), pp. 313-317.

OTT, M.G., SCHMIDT, M., SCHWARZWAELDER, K., STEIN, S., SILER, U., KOEHL, U., GLIMM, H., KUHLCKE, K., SCHILZ, A., KUNKEL, H., NAUNDORF, S., BRINKMANN, A., DEICHMANN, A., FISCHER, M., BALL, C., PILZ, I., DUNBAR, C., DU, Y., JENKINS, N.A., COPELAND, N.G., LUTHI, U., HASSAN, M., THRASHER, A.J., HOELZER, D., VON KALLE, C., SEGER, R. and GREZ, M., 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nature medicine, 12(4), pp. 401-409.

PALOMEQUE, J., CHEMALY, E.R., COLOSI, P., WELLMAN, J.A., ZHOU, S., DEL MONTE, F. and HAJJAR, R.J., 2007. Efficiency of eight different AAV serotypes in transducing rat myocardium in vivo. Gene therapy, 14(13), pp. 989-997.

PANG, L., 2020. Toxicity testing in the era of induced pluripotent stem cells: A perspective regarding the use of patient-specific induced pluripotent stem cell–derived cardiomyocytes for cardiac safety evaluation. Current Opinion in Toxicology, 23, pp. 50-55.

PARK, J.H., JEONG, H., HONG, J., CHANG, M., KIM, M., CHUCK, R.S., LEE, J.K. and PARK, C.Y., 2016. The Effect of Silica Nanoparticles on Human Corneal Epithelial Cells. Scientific reports, 6, pp. 37762.

PARK, J.H., YUN, J.H., SHI, Y., HAN, J., LI, X., JIN, Z., KIM, T., PARK, J., PARK, S., LIU, H. and LEE, W., 2019. Non-Cryogenic Structure and Dynamics of HIV-1 Integrase Catalytic Core Domain by X-ray Free-Electron Lasers. International journal of molecular sciences, 20(8), pp. 10.3390/ijms20081943.

PARKS, R.J., CHEN, L., ANTON, M., SANKAR, U., RUDNICKI, M.A. and GRAHAM, F.L., 1996. A helper-dependent adenovirus vector system: removal of helper virus by Cre-mediated excision of the viral packaging signal. Proceedings of the National Academy of Sciences of the United States of America, 93(24), pp. 13565-13570.

PASCUAL-GARCIA, P., GOVIND, C.K., QUERALT, E., CUENCA-BONO, B., LLOPIS, A., CHAVEZ, S., HINNEBUSCH, A.G. and RODRIGUEZ-NAVARRO, S., 2008. Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2. Genes & development, 22(20), pp. 2811-2822.

PATEL, A.K. and TIKOO, S.K., 2006. 293T cells expressing simian virus 40 T antigen are semi-permissive to bovine adenovirus type 3 infection. The Journal of general virology, 87(Pt 4), pp. 817-821.

PATHAN, M., KEERTHIKUMAR, S., ANG, C.S., GANGODA, L., QUEK, C.Y., WILLIAMSON, N.A., MOURADOV, D., SIEBER, O.M., SIMPSON, R.J., SALIM, A., BACIC, A., HILL, A.F., STROUD, D.A., RYAN, M.T., AGBINYA, J.I., MARIADASON, J.M., BURGESS, A.W. and MATHIVANAN, S., 2015. FunRich: An open access standalone functional enrichment and interaction network analysis tool. Proteomics, 15(15), pp. 2597-2601.

PATHAN, M., KEERTHIKUMAR, S., CHISANGA, D., ALESSANDRO, R., ANG, C.S., ASKENASE, P., BATAGOV, A.O., BENITO-MARTIN, A., CAMUSSI, G., CLAYTON, A., COLLINO, F., DI VIZIO, D., FALCON-PEREZ, J.M., FONSECA, P., FONSEKA, P., FONTANA, S., GHO, Y.S., HENDRIX, A., HOEN, E.N., IRACI, N., KASTANIEGAARD, K., KISLINGER, T., KOWAL, J., KUROCHKIN, I.V., LEONARDI, T., LIANG, Y., LLORENTE, A., LUNAVAT, T.R., MAJI, S., MONTELEONE, F., OVERBYE, A., PANARETAKIS, T., PATEL, T., PEINADO, H., PLUCHINO, S., PRINCIPE, S., RONQUIST, G., ROYO, F., SAHOO, S., SPINELLI, C., STENSBALLE, A., THERY, C., VAN HERWIJNEN, M.J.C., WAUBEN, M., WELTON, J.L., ZHAO, K. and MATHIVANAN, S., 2017. A novel community driven software for functional enrichment analysis of extracellular vesicles data. Journal of extracellular vesicles, 6(1), pp. 1321455.

PAVLOVIC, B.J., BLAKE, L.E., ROUX, J., CHAVARRIA, C. and GILAD, Y., 2018. A Comparative Assessment of Human and Chimpanzee iPSC-derived Cardiomyocytes with Primary Heart Tissues. Scientific reports, 8(1), pp. 15312-018-33478-9.

PEARCE, M.E., MAI, H.Q., LEE, N., LARSEN, S.C. and SALEM, A.K., 2008. Silicalite nanoparticles that promote transgene expression. Nanotechnology, 19(17), pp. 175103-4484/19/17/175103. Epub 2008 Mar 25.

PENG, H.H., WU, S., DAVIS, J.J., WANG, L., ROTH, J.A., MARINI, F.C.,3RD and FANG, B., 2006. A rapid and efficient method for purification of recombinant adenovirus with arginine-glycine-aspartic acid-modified fibers. Analytical Biochemistry, 354(1), pp. 140-147.

PERCIE DU SERT, N. and ROBINSON, V., 2018. The NC3Rs gateway: Accelerating scientific discoveries with new 3Rs models and technologies. F1000Research, 7, pp. 591.

PEREIRA, H.G., VALENTINE, R.C. and RUSSELL, W.C., 1968. Crystallization of an Adenovirus Protein (The Hexon). Nature, 219, pp. 946-947.

PICHLA-GOLLON, S.L., DRINKER, M., ZHOU, X., XUE, F., RUX, J.J., GAO, G.P., WILSON, J.M., ERTL, H.C., BURNETT, R.M. and BERGELSON, J.M., 2007. Structure-based identification of a major neutralizing site in an adenovirus hexon. Journal of virology, 81(4), pp. 1680-1689.

POEWE, W., SEPPI, K., TANNER, C.M., HALLIDAY, G.M., BRUNDIN, P., VOLKMANN, J., SCHRAG, A.E. and LANG, A.E., 2017. Parkinson disease. Nature reviews.Disease primers, 3, pp. 17013.

POLLARD, H., REMY, J., LOUSSOUARN, G., DEMOLOMBE, S., BEHR, J. and ESCANDE, D., 1997. Polyethylenimine but Not Cationic Lipids Promotes Transgene Delivery to the Nucleus in Mammalian Cells. Journal of Biological Chemistry, 273(13), pp. 7507-7511.

PORTALES-CASAMAR, E., THONGJUEA, S., KWON, A.T., ARENILLAS, D., ZHAO, X., VALEN, E., YUSUF, D., LENHARD, B., WASSERMAN, W.W. and SANDELIN, A., 2010. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic acids research, 38(Database issue), pp. D105-10.

POTASH, A.E., WALLEN, T.J., KARP, P.H., ERNST, S., MONINGER, T.O., GANSEMER, N.D., STOLTZ, D.A., ZABNER, J. and CHANG, E.H., 2013. Adenoviral gene transfer corrects the ion transport defect in the sinus epithelia of a porcine CF model. Molecular therapy : the journal of the American Society of Gene Therapy, 21(5), pp. 947-953.

PROTZE, S.I., LIU, J., NUSSINOVITCH, U., OHANA, L., BACKX, P.H., GEPSTEIN, L. and KELLER, G.M., 2017. Sinoatrial node cardiomyocytes derived from human pluripotent cells function as a biological pacemaker. Nature biotechnology, 35(1), pp. 56-68.

PRUSS, D., BUSHMAN, F.D. and WOLFFE, A.P., 1994. Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. Proceedings of the National Academy of Sciences of the United States of America, 91(13), pp. 5913-5917.

PRUSS, D., REEVES, R., BUSHMAN, F.D. and WOLFFE, A.P., 1994. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. The Journal of biological chemistry, 269(40), pp. 25031-25041.

PRUSZAK, J., LUDWIG, W., BLAK, A., ALAVIAN, K. and ISACSON, O., 2009. CD15, CD24, and CD29 define a surface biomarker code for neural lineage differentiation of stem cells. Stem cells (Dayton, Ohio), 27(12), pp. 2928-2940.

PUSHP, P., SAHOO, B., FERREIRA, F.C., SAMPAIO CABRAL, J.M., FERNANDES-PLATZGUMMER, A. and GUPTA, M.K., 2020. Functional

comparison of beating cardiomyocytes differentiated from umbilical cord-derived mesenchymal/stromal stem cells and human foreskin-derived induced pluripotent stem cells. Journal of biomedical materials research.Part A, 108(3), pp. 496-514.

PUUMALAINEN, A.M., VAPALAHTI, M., AGRAWAL, R.S., KOSSILA, M., LAUKKANEN, J., LEHTOLAINEN, P., VIITA, H., PALJÄRVI, L., VANNINEN, L. and YLÄ-HERTTUALA, A., 1998. Beta-galactosidase gene transfer to human malignant glioma in vivo using replication-deficient retroviruses and adenoviruses. Human Gene Therapy, 9(12), pp. 1769-1774.

QIAO, J., MORENO, J., SANCHEZ-PEREZ, L., KOTTKE, T., THOMPSON, J., CARUSO, M., DIAZ, R.M. and VILE, R., 2006. VSV-G pseudotyped, MuLV-based, semi-replication-competent retrovirus for cancer treatment. Gene therapy, 13(20), pp. 1457-1470.

QIU, Q., XU, Z., TIAN, J., MOITRA, R., GUNTI, S., NOTKINS, A.L. and BYRNES, A.P., 2015. Impact of natural IgM concentration on gene therapy with adenovirus type 5 vectors. Journal of virology, 89(6), pp. 3412-3416.

RAE, D.T. and TROBRIDGE, G.D., 2013. Retroviral genotoxicity, gene therapy - tools and potential applications. Intech.

RAJA, S., MURTY, V.R., THIVAHARAN, V., RAJASEKAR, V. and RAMESH, V., 2011. Aqueous Two Phase Systems for the Recovery of Biomolecules – A Review. Science and Technology, 1(1), pp. 2163-2167.

RAMAMOORTH, M. and NARVEKAR, A., 2015. Non viral vectors in gene therapy- an overview. Journal of clinical and diagnostic research : JCDR, 9(1), pp. GE01-6.

RAO, C., PRODROMAKIS, T., KOLKER, L., CHAUDHRY, U.A., TRANTIDOU, T., SRIDHAR, A., WEEKES, C., CAMELLITI, P., HARDING,

S.E., DARZI, A., YACOUB, M.H., ATHANASIOU, T. and TERRACCIANO, C.M., 2013. The effect of microgrooved culture substrates on calcium cycling of cardiac myocytes derived from human induced pluripotent stem cells. Biomaterials, 34(10), pp. 2399-2411.

RAPTI, K., STILLITANO, F., KARAKIKES, I., NONNENMACHER, M., WEBER, T., HULOT, J.S. and HAJJAR, R.J., 2015. Effectiveness of gene delivery systems for pluripotent and differentiated cells. Molecular therapy.Methods & clinical development, 2, pp. 14067.

RASHIDI, H., LUU, N.T., ALWAHSH, S.M., GINAI, M., ALHAQUE, S., DONG, H., TOMAZ, R.A., VERNAY, B., VIGNESWARA, V., HALLETT, J.M., CHANDRASHEKRAN, A., DHAWAN, A., VALLIER, L., BRADLEY, M., CALLANAN, A., FORBES, S.J., NEWSOME, P.N. and HAY, D.C., 2018. 3D human liver tissue from pluripotent stem cells displays stable phenotype in vitro and supports compromised liver function in vivo. Archives of Toxicology, 92(10), pp. 3117-3129.

REDDY, V.S., NATCHIAR, S.K., STEWART, P.L. and NEMEROW, G.R., 2010. Crystal structure of human adenovirus at 3.5 A resolution. Science (New York, N.Y.), 329(5995), pp. 1071-1075.

RELANDER, T., JOHANSSON, M., OLSSON, K., IKEDA, Y., TAKEUCHI, Y., COLLINS, M. and RICHTER, J., 2005. Gene transfer to repopulating human CD34+ cells using amphotropic-, GALV-, or RD114-pseudotyped HIV-1-based vectors from stable producer cells. Molecular therapy : the journal of the American Society of Gene Therapy, 11(3), pp. 452-459.

REN, Y., LEE, M.Y., SCHLIFFKE, S., PAAVOLA, J., AMOS, P.J., GE, X., YE, M., ZHU, S., SENYEI, G., LUM, L., EHRLICH, B.E. and QYANG, Y., 2011. Small molecule Wnt inhibitors enhance the efficiency of BMP-4-directed

cardiac differentiation of human pluripotent stem cells. Journal of Molecular and Cellular Cardiology, 51(3), pp. 280-287.

RHOADS, A. and AU, K.F., 2015. PacBio Sequencing and Its Applications. Genomics, Proteomics, Bioinformatics, 13(5), pp. 278-289.

RIO, D.C., CLARK, S.G. and TJIAN, R., 1985. A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction. Science (New York, N.Y.), 227(4682), pp. 23-28.

RITTELMEYER, I., ROTHE, M., BRUGMAN, M.H., IKEN, M., SCHAMBACH, A., MANNS, M.P., BAUM, C., MODLICH, U. and OTT, M., 2013. Hepatic lentiviral gene transfer is associated with clonal selection, but not with tumor formation in serially transplanted rodents. Hepatology (Baltimore, Md.), 58(1), pp. 397-408.

ROBINSON, C.M., RAJAIYA, J., ZHOU, X., SINGH, G., DYER, D.W. and CHODOSH, J., 2011. The E3 CR1-gamma gene in human adenoviruses associated with epidemic keratoconjunctivitis. Virus research, 160(1-2), pp. 120-127.

RODRIGUES, A., ALVES, P.M. and COROADINHA, A., 2011. Production of Retroviral and Lentiviral Gene Therapy Vectors: Challenges in the Manufacturing of Lipid Enveloped Virus. In: K. XU, ed, Viral Gene Therapy. InTech, .

RODRIGUEZ, R., SCHUUR, E.R., LIM, H.Y., HENDERSON, G.A., SIMONS, J.W. and HENDERSON, D.R., 1997. Prostate attenuated replication competent adenovirus (ARCA) CN706: a selective cytotoxic for prostate-specific antigen-positive prostate cancer cells. Cancer research, 57(13), pp. 2559-2563.

RODRIGUEZ-PIZA, I., RICHAUD-PATIN, Y., VASSENA, R., GONZALEZ, F., BARRERO, M.J., VEIGA, A., RAYA, A. and IZPISUA BELMONTE, J.C., 2010. Reprogramming of human fibroblasts to induced pluripotent stem cells under xeno-free conditions. Stem cells (Dayton, Ohio), 28(1), pp. 36-44.

ROELVINK, P.W., LIZONOVA, A., LEE, J.G.M., LI, Y., BERGELSON, J.M., FINBERG, R.W., BROUGH, D.E., KOVESDI, I. and WICKHAM, T.J., 1998. The Coxsackievirus-Adenovirus Receptor Protein Can Function as a Cellular Attachment Protein for Adenovirus Serotypes from Subgroups A, C, D, E, and F. Journal of virology, 72(10), pp. 7909-7915.

ROELVINK, P.W., MI LEE, G., EINFELD, D.A., KOVESDI, I. and WICKHAM, T.J., 1999. Identification of a conserved receptor-binding site on the fiber proteins of CAR-recognizing adenoviridae. Science (New York, N.Y.), 286(5444), pp. 1568-1571.

ROGERS, R.D., BOND, A.H., BAUER, C.B., ZHANG, J. and GRIFFIN, S.T., 1996. Metal ion separations in polyethylene glycol-based aqueous biphasic systems: correlation of partitioning behavior with available thermodynamic hydration data. Journal of chromatography.B, Biomedical applications, 680(1-2), pp. 221-229.

ROGERS, S., LOWENTHAL, A., TERHEGGEN, H.G. and COLUMBO, J.P., 1973. Induction of arginase activity with the Shope papilloma virus in tissue culture cells from an argininemic patient. The Journal of experimental medicine, 137(4), pp. 1091-1096.

ROGERS, S. and PFUDERER, P., 1968. Use of viruses as carriers of added genetic information. Nature, 219(5155), pp. 749-751.

ROJAS, S.V., KENSAH, G., ROTAERMEL, A., BARAKI, H., KUTSCHKA, I., ZWEIGERDT, R., MARTIN, U., HAVERICH, A., GRUH, I. and

MARTENS, A., 2017. Transplantation of purified iPSC-derived cardiomyocytes in myocardial infarction. PloS one, 12(5), pp. e0173222.

RONEN, K., NEGRE, O., ROTH, S., COLOMB, C., MALANI, N., DENARO, M., BRADY, T., FUSIL, F., GILLET-LEGRAND, B., HEHIR, K., BEUZARD, Y., LEBOULCH, P., DOWN, J.D., PAYEN, E. and BUSHMAN, F.D., 2011. Distribution of lentiviral vector integration sites in mice following therapeutic gene transfer to treat beta-thalassemia. Molecular therapy : the journal of the American Society of Gene Therapy, 19(7), pp. 1273-1286.

ROSE, J.A., HOGGAN, M.D. and SHATKIN, A.J., 1966. Nucleic acid from an adeno-associated virus: chemical and physical studies. Proceedings of the National Academy of Sciences of the United States of America, 56(1), pp. 86-92.

ROSENBERG, S.A., AEBERSOLD, P., CORNETTA, K., KASID, A., MORGAN, R.A., MOEN, R., KARSON, E.M., LOTZE, M.T., YANG, J.C. and TOPALIAN, S.L., 1990. Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. The New England journal of medicine, 323(9), pp. 570-578.

ROSSETTI, M., CAVARELLI, M., GREGORI, S. and SCARLATTI, G., 2013. HIV-derived vectors for gene therapy targeting dendritic cells. Advances in Experimental Medicine and Biology, 762, pp. 239-261.

ROWE, W.P., HUEBNER, R.J., GILMORE, L.K., PARROT, R.H. and WARD, T.G., 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proceedings of the Society for Experimental Biology and Medicine.Society for Experimental Biology and Medicine (New York, N.Y.), 84(3), pp. 570-573.

ROWLAND, B.D., BERNARDS, R. and PEEPER, D.S., 2005. The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. Nature cell biology, 7(11), pp. 1074-1082.

RUAN, J.L., TULLOCH, N.L., RAZUMOVA, M.V., SAIGET, M., MUSKHELI, V., PABON, L., REINECKE, H., REGNIER, M. and MURRY, C.E., 2016. Mechanical Stress Conditioning and Electrical Stimulation Promote Contractility and Force Maturation of Induced Pluripotent Stem Cell-Derived Human Cardiac Tissue. Circulation, 134(20), pp. 1557-1567.

RUSCIC, J., PERRY, C., MUKHOPADHYAY, T., TAKEUCHI, Y. and BRACEWELL, D.G., 2019. Lentiviral Vector Purification Using Nanofiber Ion-Exchange Chromatography. Molecular therapy.Methods & clinical development, 15, pp. 52-62.

RUSSELL, D.W. and SAMBROOK, J., 2001. Molecular cloning: a laboratory manual. 1 edn. New York: Cold Sprind Harbor Laboratory Press.

SACHLOS, E., RISUENO, R.M., LARONDE, S., SHAPOVALOVA, Z., LEE, J.H., RUSSELL, J., MALIG, M., MCNICOL, J.D., FIEBIG-COMYN, A., GRAHAM, M., LEVADOUX-MARTIN, M., LEE, J.B., GIACOMELLI, A.O., HASSELL, J.A., FISCHER-RUSSELL, D., TRUS, M.R., FOLEY, R., LEBER, B., XENOCOSTAS, A., BROWN, E.D., COLLINS, T.J. and BHATIA, M., 2012. Identification of drugs including a dopamine receptor antagonist that selectively target cancer stem cells. Cell, 149(6), pp. 1284-1297.

SAEED, A.I., SHAROV, V., WHITE, J., LI, J., LIANG, W., BHAGABATI, N., BRAISTED, J., KLAPA, M., CURRIER, T., THIAGARAJAN, M., STURN, A., SNUFFIN, M., REZANTSEV, A., POPOV, D., RYLTSOV, A., KOSTUKOVICH, E., BORISOVSKY, I., LIU, Z., VINSAVICH, A., TRUSH, V. and QUACKENBUSH, J., 2003. TM4: a free, open-source system for microarray data management and analysis. BioTechniques, 34(2), pp. 374-378.

SALABARRIA, S.M., NAIR, J., CLEMENT, N., SMITH, B.K., RABEN, N., FULLER, D.D., BYRNE, B.J. and CORTI, M., 2020. Advancements in AAV-mediated Gene Therapy for Pompe Disease. Journal of neuromuscular diseases, 7(1), pp. 15-31.

SANBER, K.S., KNIGHT, S.B., STEPHEN, S.L., BAILEY, R., ESCORS, D., MINSHULL, J., SANTILLI, G., THRASHER, A.J., COLLINS, M.K. and TAKEUCHI, Y., 2015. Construction of stable packaging cell lines for clinical lentiviral vector production. Scientific reports, 5, pp. 9021.

SASSONE, F., MARGULETS, V., MARASCHI, A., RODIGHIERO, S., PASSAFARO, M., SILANI, V., CIAMMOLA, A., KIRSHENBAUM, L.A. and SASSONE, J., 2015. Bcl-2/adenovirus E1B 19-kDa interacting protein (BNip3) has a key role in the mitochondrial dysfunction induced by mutant huntingtin. Human molecular genetics, .

SATO, Y., HIGUCHI, T., KOBAYASHI, H., MINAMISAWA, S., IDA, H. and OHASHI, T., 2016. Lentiviral Gene Transfer to iPS Cells: Toward the Cardiomyocyte Differentiation of Pompe Disease-Specific iPS Cells. In: T. NAKANISHI, R.R. MARKWALD, H.S. BALDWIN, B.B. KELLER, D. SRIVASTAVA and H. YAMAGISHI, eds, Etiology and Morphogenesis of Congenital Heart Disease: From Gene Function and Cellular Interaction to Morphology. Tokyo: The Author(s), pp. 341-343.

SATO, Y., KOBAYASHI, H., HIGUCHI, T., SHIMADA, Y., ERA, T., KIMURA, S., ETO, Y., IDA, H. and OHASHI, T., 2015. Disease modeling and lentiviral gene transfer in patient-specific induced pluripotent stem cells from late-onset Pompe disease patient. Molecular therapy.Methods & clinical development, 2, pp. 15023.

SAYED, N., LIU, C., AMEEN, M., HIMMATI, F., ZHANG, J.Z., KHANAMIRI, S., MOONEN, J.R., WNOROWSKI, A., CHENG, L., RHEE,

J.W., GADDAM, S., WANG, K.C., SALLAM, K., BOYD, J.H., WOO, Y.J., RABINOVITCH, M. and WU, J.C., 2020. Clinical trial in a dish using iPSCs shows lovastatin improves endothelial dysfunction and cellular cross-talk in LMNA cardiomyopathy. Science translational medicine, 12(554),.

SCHLEGEL, R., TRALKA, T.S., WILLINGHAM, M.C. and PASTAN, I., 1983. Inhibition of VSV binding and infectivity by phosphatidylserine: is phosphatidylserine a VSV-binding site? Cell, 32(2), pp. 639-346.

SCHLEGEL, R., WILLINGHAM, M.C. and PASTAN, I.H., 1982. Saturable binding sites for vesicular stomatitis virus on the surface of Vero cells. Journal of virology, 43(3), pp. 871-875.

SCHLOEMER, R.H. and WAGNER, R.R., 1975. Cellular adsorption function of the sialoglycoprotein of vesicular stomatitis virus and its neuraminic acid. Journal of virology, 15(4), pp. 882-893.

SCHMIDT, M., HOFFMANN, G., WISSLER, M., LEMKE, N., MUSSIG, A., GLIMM, H., WILLIAMS, D.A., RAGG, S., HESEMANN, C.U. and VON KALLE, C., 2001. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. Human Gene Therapy, 12(7), pp. 743-749.

SCHOPPERLE, W.M. and DEWOLF, W.C., 2007. The TRA-1-60 and TRA-1-81 human pluripotent stem cell markers are expressed on podocalyxin in embryonal carcinoma. Stem cells (Dayton, Ohio), 25(3), pp. 723-730.

SCHOTT, J.W., LEON-RICO, D., FERREIRA, C.B., BUCKLAND, K.F., SANTILLI, G., ARMANT, M.A., SCHAMBACH, A., CAVAZZA, A. and THRASHER, A.J., 2019. Enhancing Lentiviral and Alpharetroviral Transduction of Human Hematopoietic Stem Cells for Clinical Application. Molecular therapy.Methods & clinical development, 14, pp. 134-147.

SCHRODER, A.R., SHINN, P., CHEN, H., BERRY, C., ECKER, J.R. and BUSHMAN, F., 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell, 110(4), pp. 521-529.

SCOTT, V., PATEL, A., VILLARREAL, D.O., HENSLEY, S.E., RAGWAN, E., YAN, J., SARDESAI, N.Y., ROTHWELL, P.J., EXTANCE, J.P., CAPRONI, L.J. and WEINER, D.B., 2015. Novel synthetic plasmid and Doggybone™ DNA vaccines induce neutralizing antibodies and provide protection from lethal influenza challenge in mice. Human Vaccines & Immunotherapeutics, 11(8), pp. 1972-1982.

SEGANTI, L., SUPERTI, F., GIRMENIA, C., MELUCCI, L. and ORSI, N., 1986. Study of receptors for vesicular stomatitis virus in vertebrate and invertebrate cells. Microbiologica, 9(3), pp. 259-267.

SEGURA, M.M., GARNIER, A., DUROCHER, Y., COELHO, H. and KAMEN, A., 2007. Production of lentiviral vectors by large-scale transient transfection of suspension cultures and affinity chromatography purification. Biotechnology and bioengineering, 98(4), pp. 789-799.

SEGURA, M.M., MANGION, M., GAILLET, B. and GARNIER, A., 2013. New developments in lentiviral vector design, production and purification. Expert opinion on biological therapy, 13(7), pp. 987-1011.

SEITZ, B., BAKTANIAN, E., GORDON, E.M., ANDERSON, W.F., LABREE, L. and MCDONNELL, P.J., 1998. Retroviral vector-mediated gene transfer into keratocytes: in vitro effects of polybrene and protamine sulfate. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie, 236(8), pp. 602-612.

SEKI, T., YUASA, S., KUSUMOTO, D., KUNITOMI, A., SAITO, Y., TOHYAMA, S., YAE, K., KISHINO, Y., OKADA, M., HASHIMOTO, H.,

TAKEI, M., EGASHIRA, T., KODAIRA, M., KURODA, Y., TANAKA, A., OKATA, S., SUZUKI, T., MURATA, M., FUJITA, J. and FUKUDA, K., 2014. Generation and characterization of functional cardiomyocytes derived from human T cell-derived induced pluripotent stem cells. PloS One, 9(1),.

SENA-ESTEVES, M. and GAO, G., 2018. Production of High-Titer Retrovirus and Lentivirus Vectors. Cold Spring Harbor protocols, 2018(4), pp. 10.1101/pdb.prot095687.

SEOANE, J., LE, H.V. and MASSAGUE, J., 2002. Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. Nature, 419(6908), pp. 729-734.

SERANGELI, C., BICANIC, O., SCHEIBLE, M.H., WERNET, D., LANG, P., RAMMENSEE, H.G., STEVANOVIC, S., HANDGRETINGER, R. and FEUCHTINGER, T., 2010. Ex vivo detection of adenovirus specific CD4+ T-cell responses to HLA-DR-epitopes of the Hexon protein show a contracted specificity of T(HELPER) cells following stem cell transplantation. Virology, 397(2), pp. 277-284.

SESTER, M., KOEBERNICK, K., OWEN, D., AO, M., BROMBERG, Y., MAY, E., STOCK, E., ANDREWS, L., GROH, V., SPIES, T., STEINLE, A., MENZ, B. and BURGERT, H.G., 2010. Conserved amino acids within the adenovirus 2 E3/19K protein differentially affect downregulation of MHC class I and MICA/B proteins. Journal of immunology (Baltimore, Md.: 1950), 184(1), pp. 255-267.

SHARMA, A., GARCIA, G.,JR, WANG, Y., PLUMMER, J.T., MORIZONO, K., ARUMUGASWAMI, V. and SVENDSEN, C.N., 2020. Human iPSC-Derived Cardiomyocytes Are Susceptible to SARS-CoV-2 Infection. Cell reports.Medicine, 1(4), pp. 100052.

SHARON, D. and KAMEN, A., 2017. Advancements in the design and scalable production of viral gene transfer vectors. Biotechnology and bioengineering, .

SHEN, Y., KITZES, G., NYE, J.A., FATTAEY, A. and HERMISTON, T., 2001. Analyses of single-amino-acid substitution mutants of adenovirus type 5 E1B-55K protein. Journal of virology, 75(9), pp. 4297-4307.

SHERR, C.J., 2004. Principles of tumor suppression. Cell, 116(2), pp. 235-246.

SHEU, J., BELTZER, J., FURY, B., WILCZEK, K., TOBIN, S., FALCONER, D., NOLTA, J. and BAUER, G., 2015. Large-scale production of lentiviral vector in a closed system hollow fiber bioreactor. Molecular Therapy Methods, 2(1520),.

SHUM, K.T., ZHOU, J. and ROSSI, J.J., 2013. Aptamer-based therapeutics: new approaches to combat human viral diseases. Pharmaceuticals (Basel, Switzerland), 6(12), pp. 1507-1542.

SIGURDSSON, B., GRIMSSON, H. and PALSSON, P.A., 1952. Maedi, a chronic, progressive infection of sheep's lungs. The Journal of infectious diseases, 90(3), pp. 233-241.

SINGAL, A.G., LAMPERTICO, P. and NAHON, P., 2020. Epidemiology and surveillance for hepatocellular carcinoma: New trends. Journal of hepatology, 72(2), pp. 250-261.

SINNECKER, D., LAUGWITZ, K.L. and MORETTI, A., 2014. Induced pluripotent stem cell-derived cardiomyocytes for drug development and toxicity testing. Pharmacology & therapeutics, 143(2), pp. 246-252.

SIPRASHVILI, Z., NGUYEN, N.T., GORELL, E.S., LOUTIT, K., KHUU, P., FURUKAWA, L.K., LORENZ, H.P., LEUNG, T.H., KEENE, D.R., RIEGER, K.E., KHAVARI, P., LANE, A.T., TANG, J.Y. and MARINKOVICH, M.P., 2016. Safety and Wound Outcomes Following Genetically Corrected

Autologous Epidermal Grafts in Patients With Recessive Dystrophic Epidermolysis Bullosa. Jama, 316(17), pp. 1808-1817.

SI-TAYEB, K., NOTO, F.K., NAGAOKA, M., LI, J., BATTLE, M.A., DURIS, C., NORTH, P.E., DALTON, S. and DUNCAN, S.A., 2010. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology (Baltimore, Md.), 51(1), pp. 297-305.

SMART, J.E. and STILLMAN, B.W., 1982. Adenovirus Terminal Protein Precursor. Partial amnio acid sequence and the site of covalent linkage to virus DNA. Journal of Biological Chemistry, 257(22), pp. 13499-13506.

SMITH, J., AZAD, A. and DEACON, N., 1992. Identification of two novel human immunodeficiency virus type 1 splice acceptor sites in infected T cell lines. The Journal of general virology, 73 (Pt 7)(Pt 7), pp. 1825-1828.

SMITH, J., LIPSITCH, M. and ALMOND, J.W., 2011. Vaccine production, distribution, access, and uptake. Lancet (London, England), 378(9789), pp. 428-438.

SMITH, T.A., IDAMAKANTI, N., MARSHALL-NEFF, J., ROLLENCE, M.L., WRIGHT, P., KALOSS, M., KING, L., MECH, C., DINGES, L., IVERSON, W.O., SHERER, A.D., MARKOVITS, J.E., LYONS, R.M., KALEKO, M. and STEVENSON, S.C., 2003. Receptor interactions involved in adenoviral-mediated gene delivery after systemic administration in non-human primates. Human Gene Therapy, 14(17), pp. 1595-1604.

SOLDI, M., SERGI SERGI, L., UNALI, G., KERZEL, T., CUCCOVILLO, I., CAPASSO, P., ANNONI, A., BIFFI, M., RANCOITA, P.M.V., CANTORE, A., LOMBARDO, A., NALDINI, L., SQUADRITO, M.L. and KAJASTE-RUDNITSKI, A., 2020. Laboratory-Scale Lentiviral Vector Production and Purification for Enhanced Ex Vivo and In Vivo Genetic Engineering. Molecular therapy.Methods & clinical development, 19, pp. 411-425.

SPINNROCK, A., MARTENS, M., ENDERS, F., BOLDT, K. and COLFEN, H., 2019. Controlled Preparation of Nanoparticle Gradient Materials by Diffusion. Nanomaterials (Basel, Switzerland), 9(7), pp. 10.3390/nano9070988.

STADTFELD, M., NAGAYA, M., UTIKAL, J., WEIR, G. and HOCHEDLINGER, K., 2008. Induced pluripotent stem cells generated without viral integration. Science (New York, N.Y.), 322(5903), pp. 945-949.

STEIN, S., OTT, M.G., SCHULTZE-STRASSER, S., JAUCH, A., BURWINKEL, B., KINNER, A., SCHMIDT, M., KRAMER, A., SCHWABLE, J., GLIMM, H., KOEHL, U., PREISS, C., BALL, C., MARTIN, H., GOHRING, G., SCHWARZWAELDER, K., HOFMANN, W.K., KARAKAYA, K., TCHATCHOU, S., YANG, R., REINECKE, P., KUHLCKE, K., SCHLEGELBERGER, B., THRASHER, A.J., HOELZER, D., SEGER, R., VON KALLE, C. and GREZ, M., 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nature medicine, 16(2), pp. 198-204.

STRANG, B.L., IKEDA, Y., COSSET, F.L., COLLINS, M.K. and TAKEUCHI, Y., 2004. Characterization of HIV-1 vectors with gammaretrovirus envelope glycoproteins produced from stable packaging cells. Gene therapy, 11(7), pp. 591-598.

STROUD, J.C., OLTMAN, A., HAN, A., BATES, D.L. and CHEN, L., 2009. Structural basis of HIV-1 activation by NF-kappaB--a higher-order complex of p50:RelA bound to the HIV-1 LTR. Journal of Molecular Biology, 393(1), pp. 98-112.

STRUNZE, S., ENGELKE, M.F., WANG, I.H., PUNTENER, D., BOUCKE, K., SCHLEICH, S., WAY, M., SCHOENENBERGER, P., BURCKHARDT, C.J. and GREBER, U.F., 2011. Kinesin-1-mediated capsid disassembly and

disruption of the nuclear pore complex promote virus infection. Cell host & microbe, 10(3), pp. 210-223.

STUDAMIRE, B. and GOFF, S.P., 2008. Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. Retrovirology, 5, pp. 48-4690-5-48.

SU, H., JOHO, S., HUANG, Y., BARCENA, A., ARAKAWA-HOYT, J., GROSSMAN, W. and KAN, Y.W., 2004. Adeno-associated viral vector delivers cardiac-specific and hypoxia-inducible VEGF expression in ischemic mouse hearts. Proceedings of the National Academy of Sciences of the United States of America, 101(46), pp. 16280-16285.

SUERTH, J.D., LABENSKI, V. and SCHAMBACH, A., 2014. Alpharetroviral vectors: from a cancer-causing agent to a useful tool for human gene therapy. Viruses, 6(12), pp. 4811-4838.

SUGIURA, M., KASAMA, Y., ARAKI, R., HOKI, Y., SUNAYAMA, M., UDA, M., NAKAMURA, M., ANDO, S. and ABE, M., 2014. Induced pluripotent stem cell generation-associated point mutations arise during the initial stages of the conversion of these cells. Stem cell reports, 2(1), pp. 52-63.

SULEMAN, S., SCHRUBAJI, K., FILIPPOU, C., IGNATOVA, S., HEWITSON, P., HUDDLESTON, J., KARDA, R., WADDINGTON, S. and THEMIS, M., Unpublished. Rapid and inexpensive purification of adenovirus vectors using optimised aqueous two-phase technology. Journal of virological methods, .

SULK, B., BIRKENMEIER, G. and KOPPERSCHLAGER, G., 1992. Application of phase partitioning and thiophilic adsorption chromatography to the purification of monoclonal antibodies from cell culture fluid. Journal of immunological methods, 149(2), pp. 165-171.

SULLIVAN, C.S. and PIPAS, J.M., 2002. T antigens of simian virus 40: molecular chaperones for viral replication and tumorigenesis. Microbiology and molecular biology reviews : MMBR, 66(2), pp. 179-202.

SUMIDA, S.M., TRUITT, D.M., LEMCKERT, A.A., VOGELS, R., CUSTERS, J.H., ADDO, M.M., LOCKMAN, S., PETER, T., PEYERL, F.W., KISHKO, M.G., JACKSON, S.S., GORGONE, D.A., LIFTON, M.A., ESSEX, M., WALKER, B.D., GOUDSMIT, J., HAVENGA, M.J. and BAROUCH, D.H., 2005. Neutralizing antibodies to adenovirus serotype 5 vaccine vectors are directed primarily against the adenovirus hexon protein. Journal of immunology (Baltimore, Md.: 1950), 174(11), pp. 7179-7185.

SUN, X. and NUNES, S.S., 2016. Biowire: a platform for maturation of human pluripotent stem cell-derived cardiomyocytes. Methods, 15(101), pp. 21-26.

SUN, X., ROTH, S.L., BIALECKI, M.A. and WHITTAKER, G.R., 2010. Internalization and fusion mechanism of vesicular stomatitis virus and related rhabdoviruses. Future virology, 5(1), pp. 85-96.

SWANEY, W.P., SORGI, F.L., BAHNSON, A.B. and BARRANGER, J.A., 1997. The effect of cationic liposome pretreatment and centrifugation on retrovirus-mediated gene transfer. Gene therapy, 4(12), pp. 1379-1386.

SZYBALSKA, E.H. and SZYBALSKI, W., 1962. Genetics of human cell line. IV. DNA-mediated heritable transformation of a biochemical trait. Proceedings of the National Academy of Sciences of the United States of AmerJID - 7505876, 48, pp. 2026-2034.

TACHENY, A., MICHEL, S., DIEU, M., PAYEN, L., ARNOULD, T. and RENARD, P., 2012. Unbiased proteomic analysis of proteins interacting with the HIV-1 5'LTR sequence: role of the transcription factor Meis. Nucleic acids research, 40(21), pp. e168.

TAKAHASHI, J., 2013. Cell replacement therapy for Parkinson's disease using iPS cells. Rinsho Shinkeigaku, 53(11), pp. 1009-1012.

TAKAHASHI, K., ITO, Y., MORIKAWA, M., KOBUNE, M., HUANG, J., TSUKAMOTO, M., SASAKI, K., NAKAMURA, K., DEHARI, H., IKEDA, K., UCHIDA, H., HIRAI, S., ABE, T. and HAMADA, H., 2003. Adenoviral-delivered angiopoietin-1 reduces the infarction and attenuates the progression of cardiac dysfunction in the rat model of acute myocardial infarction. Molecular therapy : the journal of the American Society of Gene Therapy, 8(4), pp. 584-592.

TAKAHASHI, K., NARITA, M., YOKURA, M., ICHISAKA, T. and YAMANAKA, S., 2009. Human induced pluripotent stem cells on autologous feeders. PloS one, 4(12), pp. e8067.

TAKAHASHI, K. and YAMANAKA, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 126(4), pp. 663-676.

TAKAYAMA, K., MORISAKI, Y., KUNO, S., NAGAMOTO, Y., HARADA, K., FURUKAWA, N., OHTAKA, M., NISHIMURA, K., IMAGAWA, K., SAKURAI, F., TACHIBANA, M., SUMAZAKI, R., NOGUCHI, E., NAKANISHI, M., HIRATA, K., KAWABATA, K. and MIZUGUCHI, H., 2014. Prediction of interindividual differences in hepatic functions and drug sensitivity by using human iPS-derived hepatocytes. Proceedings of the National Academy of Sciences of the United States of America, 111(47), pp. 16772-16777.

TAN, C.H., SHOW, P.L., OOI, C.W., NG, E.P., LAN, J.C. and LING, T.C., 2015. Novel lipase purification methods - a review of the latest developments. Biotechnology journal, 10(1), pp. 31-44.

TAN, S., TAO, Z., LOO, S., SU, L., CHEN, X. and YE, L., 2019. Non-viral vector based gene transfection with human induced pluripotent stem cells derived cardiomyocytes. Scientific reports, 9(1), pp. 14404-019-50980-w.

TAN, Z.J., WANG, C.Y., YANG, Z.Z., YI, Y.J., WANG, H.Y., ZHOU, W.L. and LI, F.F., 2015. Ionic Liquid-Based Ultrasonic-Assisted Extraction of Secoisolariciresinol Diglucoside from Flaxseed (Linum usitatissimum L.) with Further Purification by an Aqueous Two-Phase System. Molecules (Basel, Switzerland), 20(10), pp. 17929-17943.

TATUM, E.L., 1966. Molecular biology, nucleic acids, and the future of medicine. Perspectives in Biological Medicine, 10(1), pp. 19-32.

TATUM, E.L. and LEDERBERG, J., 1947. Gene Recombination in the Bacterium Escherichia coli. Journal of Bacteriology, 53(6), pp. 673-684.

TEMIN, H.M. and MIZUTANI, S., 1992. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. 1970. Biotechnology (Reading, Mass.), 24, pp. 51-56.

TEODORO, J.G., SHORE, G.C. and BRANTON, P.E., 1995. Adenovirus E1A proteins induce apoptosis by both p53-dependent and p53-independent mechanisms. Oncogene, 11(3), pp. 467-474.

TERHEGGEN, H.G., LOWENTHAL, A., LAVINHA, F., COLOMBO, J.P. and ROGERS, S., 1975. Unsuccessful trial of gene replacement in arginase deficiency. Zeitschrift fur Kinderheilkunde, 119(1), pp. 1-3.

TERWILLIGER, E.F., GODIN, B., SODROSKI, J.G. and HASELTINE, W.A., 1989. Construction and use of a replication-competent human immunodeficiency virus (HIV-1) that expresses the chloramphenicol acetyltransferase enzyme. Proc.Natl.Acad.Sci.U.S.A., 86(10), pp. 3857-3861.

THEMIS, M., WADDINGTON, S.N., SCHMIDT, M., VON KALLE, C., WANG, Y., AL-ALLAF, F., GREGORY, L.G., NIVSARKAR, M., THEMIS, M., HOLDER, M.V., BUCKLEY, S.M., DIGHE, N., RUTHE, A.T., MISTRY, A., BIGGER, B., RAHIM, A., NGUYEN, T.H., TRONO, D., THRASHER, A.J. and COUTELLE, C., 2005. Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice. Molecular therapy : the journal of the American Society of Gene Therapy, 12(4), pp. 763-771.

THOMAS, P. and SMART, T.G., 2005. HEK293 cell line: a vehicle for the expression of recombinant proteins. Journal of pharmacological and toxicological methods, 51(3), pp. 187-200.

THOMSON, J.A., ITSKOVITZ-ELDOR, J., SHAPIRO, S.S., WAKNITZ, M.A., SWIERGIEL, J.J., MARSHALL, V.S. and JONES, J.M., 1998. Embryonic stem cell lines derived from human blastocysts. Science (New York, N.Y.), 282(5391), pp. 1145-1147.

THORNHILL, S.I., SCHAMBACH, A., HOWE, S.J., ULAGANATHAN, M., GRASSMAN, E., WILLIAMS, D., SCHIEDLMEIER, B., SEBIRE, N.J., GASPAR, H.B., KINNON, C., BAUM, C. and THRASHER, A.J., 2008. Self-inactivating gammaretroviral vectors for gene therapy of X-linked severe combined immunodeficiency. Molecular therapy : the journal of the American Society of Gene Therapy, 16(3), pp. 590-598.

THROM, R.E., OUMA, A.A., ZHOU, S., CHANDRASEKARAN, A., LOCKEY, T., GREENE, M., DE RAVIN, S.S., MOAYERI, M., MALECH, H.L., SORRENTINO, B.P. and GRAY, J.T., 2009. Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection. Blood, 113(21), pp. 5104-5110.

TIJANI, M., MUNIS, A.M., PERRY, C., SANBER, K., FERRARESSO, M., MUKHOPADHYAY, T., THEMIS, M., NISOLI, I., MATTIUZZO, G., COLLINS, M.K. and TAKEUCHI, Y., 2018. Lentivector Producer Cell Lines with Stably Expressed Vesiculovirus Envelopes. Molecular therapy.Methods & clinical development, 10, pp. 303-312.

TKACIK, G. and MICHAELS, S., 1991. A rejection profile test for ultrafiltration membranes & devices. Bio/technology (Nature Publishing Company), 9(10), pp. 941-3, 946.

TOHYAMA, S., HATTORI, F., SANO, M., HISHIKI, T., NAGAHATA, Y., MATSUURA, T., HASHIMOTO, H., SUZUKI, T., YAMASHITA, H., SATOH, Y., EGASHIRA, T., SEKI, T., MURAOKA, N., YAMAKAWA, H., OHGINO, Y., TANAKA, T., YOICHI, M., YUASA, S., MURATA, M., SUEMATSU, M. and FUKUDA, K., 2013. Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. Cell stem cell, 12(1), pp. 127-137.

TOUW, I.P. and ERKELAND, S.J., 2007. Retroviral insertion mutagenesis in mice as a comparative oncogenomics tool to identify disease genes in human leukemia. Molecular therapy : the journal of the American Society of Gene Therapy, 15(1), pp. 13-19.

TRAN, T.H., WANG, X., BROWNE, C., ZHANG, Y., SCHINKE, M., IZUMO, S. and BURCIN, M., 2009. Wnt3a-induced mesoderm formation and cardiomyogenesis in human embryonic stem cells. Stem cells (Dayton, Ohio), 27(8), pp. 1869-1878.

TRANSFIGURACION, J., JAALOUK, D.E., GHANI, K., GALIPEAU, J. and KAMEN, A., 2003. Size-exclusion chromatography purification of high-titer vesicular stomatitis virus G glycoprotein-pseudotyped retrovectors for cell and gene therapy applications. Human Gene Therapy, 14(12), pp. 1139-1153.

TREJBALOVA, K., KOVAROVA, D., BLAZKOVA, J., MACHALA, L., JILICH, D., WEBER, J., KUCEROVA, D., VENCALEK, O., HIRSCH, I. and HEJNAR, J., 2016. Development of 5' LTR DNA methylation of latent HIV-1 provirus in cell line models and in long-term-infected individuals. Clinical epigenetics, 8, pp. 19-016-0185-6. eCollection 2016.

TRINDADE, I.P., DIOGO, M.M., PRAZERES, D.M. and MARCOS, J.C., 2005. Purification of plasmid DNA vectors by aqueous two-phase extraction and hydrophobic interaction chromatography. Journal of chromatography.A, 1082(2), pp. 176-184.

TROTMAN, L.C., MOSBERGER, N., FORNEROD, M., STIDWILL, R.P. and GREBER, U.F., 2001. Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. Nature cell biology, 3(12), pp. 1092-1100.

UNIPROT CONSORTIUM, 2019. UniProt: a worldwide hub of protein knowledge. Nucleic acids research, 47(D1), pp. D506-D515.

VALKAMA, A.J., LEINONEN, H.M., LIPPONEN, E.M., TURKKI, V., MALINEN, J., HEIKURA, T., YLÄ-HERTTUALA, S. and LESCH, H.P., 2017. Optimization of lentiviral vector production for scale-up in fixed-bed bioreactor. Gene Therapy, 25, pp. 39-46.

VAN DEN BERG, C.W., OKAWA, S., CHUVA DE SOUSA LOPES, S.M., VAN IPEREN, L., PASSIER, R., BRAAM, S.R., TERTOOLEN, L.G., DEL SOL, A., DAVIS, R.P. and MUMMERY, C.L., 2015. Transcriptome of human foetal heart compared with cardiomyocytes from pluripotent stem cells. Development (Cambridge, England), 142(18), pp. 3231-3238.

VAN ORMONDT, H., MAAT, J. and VAN BEVEREN, C.P., 1980. The nucleotide sequence of the transforming early region E1 of adenovirus type 5 DNA. Gene, 11(3-4), pp. 299-309.

VAN RAAIJ, M.J., MITRAKI, A., LAVIGNE, G. and CUSACK, S., 1999. A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature, 401(6756), pp. 935-938.

VANDEGRAAFF, N., DEVROE, E., TURLURE, F., SILVER, P.A. and ENGELMAN, A., 2006. Biochemical and genetic analyses of integrase-interacting proteins lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in preintegration complex function and HIV-1 replication. Virology, 346(2), pp. 415-426.

VEGA, Y., DELGADO, E., DE LA BARRERA, J., CARRERA, C., ZABALLOS, A., CUESTA, I., MARINO, A., OCAMPO, A., MIRALLES, C., PEREZ-CASTRO, S., ALVAREZ, H., LOPEZ-MIRAGAYA, I., GARCIA-BODAS, E., DIEZ-FUERTES, F. and THOMSON, M.M., 2016. Sequence Analysis of In Vivo-Expressed HIV-1 Spliced RNAs Reveals the Usage of New and Unusual Splice Sites by Viruses of Different Subtypes. PloS one, 11(6), pp. e0158525.

VENKATACHARI, N.J., ZERBATO, J.M., JAIN, S., MANCINI, A.E., CHATTOPADHYAY, A., SLUIS-CREMER, N., BAR-JOSEPH, Z. and AYYAVOO, V., 2015. Temporal transcriptional response to latency reversing agents identifies specific factors regulating HIV-1 viral transcriptional switch. Retrovirology, 12, pp. 85-015-0211-3.

VICENTE, T., MOTA, J.P., PEIXOTO, C., ALVES, P.M. and CARRONDO, M.J., 2011. Rational design and optimization of downstream processes of virus particles for biopharmaceutical applications: current advances. Biotechnology Advances, 29(6), pp. 869-878.

WALLS, A.C., PARK, Y.J., TORTORICI, M.A., WALL, A., MCGUIRE, A.T. and VEESLER, D., 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell, 181(2), pp. 281-292.e6.

WALTER, H. and JOHANSSON, G., 1994. Methods in Enzymology: Aqueous Two-Phase Systems. Volume 228. 1 edn. Academic Press.

WANG, G.P., LEVINE, B.L., BINDER, G.K., BERRY, C.C., MALANI, N., MCGARRITY, G., TEBAS, P., JUNE, C.H. and BUSHMAN, F.D., 2009. Analysis of lentiviral vector integration in HIV+ study subjects receiving autologous infusions of gene modified CD4+ T cells. Molecular therapy : the journal of the American Society of Gene Therapy, 17(5), pp. 844-850.

WANG, J., GAO, Q. and LI, Q., 2015. Adenovirus type 12 E1B 55-kilodalton oncoprotein promotes p53-mediated apoptotic response of ovarian cancer to cisplatin. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine, 36(8), pp. 6569-6577.

WARKANY, J., 1988. Why I doubted that thalidomide was the cause of the epidemic of limb defects of 1959 to 1961. Teratology, 38(3), pp. 217-219.

WARNOCK, J.N., DAIGRE, C. and AL-RUBEAI, M., 2011. Introduction to viral vectors. Methods in molecular biology (Clifton, N.J.), 737, pp. 1-25.

WEI, P., GARBER, M.E., FANG, S.M., FISCHER, W.H. and JONES, K.A., 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell, 92(4), pp. 451-462.

WESSLEN, T., ALBERTSSON, P.A. and PHILIPSON, L., 1959. Concentration of animal viruses using two-phase systems of aqueous polymer solutions. Archiv fur die Gesamte Virusforschung, 9, pp. 510-520.

WIGHTMAN, L., KIRCHEIS, R., ROSSLER, V., CAROTTA, S., RUZICKA, R., KURSA, M. and WAGNER, E., 2001. Different behavior of branched and linear polyethylenimine for gene delivery in vitro and in vivo. The journal of gene medicine, 3(4), pp. 362-372.

WILEY, 2021-last update, Gene Therapy Clinical Trials Worldiwde. Available: https://a873679.fmphost.com/fmi/webd/GTCT [February,1, 2021].

WOLD, W.S. and TOTH, K., 2013. Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. Current gene therapy, 13(6), pp. 421-433.

WOLTJEN, K., MICHAEL, I.P., MOHSENI, P., DESAI, R., MILEIKOVSKY, M., HAMALAINEN, R., COWLING, R., WANG, W., LIU, P., GERTSENSTEIN, M., KAJI, K., SUNG, H.K. and NAGY, A., 2009. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature, 458(7239), pp. 766-770.

WONG, H., BUDHA, N.R., WEST, K., BLACKWOOD, E., WARE, J.A., YU, R., DARBONNE, W.C., GOULD, S.E., STEIGERWALT, R., ERICKSON, R., HOP, C.E., LORUSSO, P., ECKHARDT, S.G., WAGNER, A., CHAN, I.T., MAMOUNAS, M., FLYGARE, J.A. and FAIRBROTHER, W.J., 2012. Dogs are more sensitive to antagonists of inhibitor of apoptosis proteins than rats and humans: a translational toxicokinetic/toxicodynamic analysis. Toxicological sciences : an official journal of the Society of Toxicology, 130(1), pp. 205-213.

WONGMONGKOL, N. and PRICHANONT, S., 2006. Partition of alkaline protease in aqueous two-phase systems of polyethylene glycol 1000 and potassium phosphate. Korean Journal of Chemical Engineering, 23(1), pp. 71-76.

WOYCHIK, R.P. and ALAGRAMAM, K., 1998. Insertional mutagenesis in transgenic mice generated by the pronuclear microinjection procedure. The International journal of developmental biology, 42(7), pp. 1009-1017.

WU, C. and DUNBAR, C.E., 2011. Stem cell gene therapy: the risks of insertional mutagenesis and approaches to minimize genotoxicity. Frontiers of medicine, 5(4), pp. 356-371.

WU, K., OROZCO, D. and HEARING, P., 2012. The adenovirus L4-22K protein is multifunctional and is an integral component of crucial aspects of infection. Journal of virology, 86(19), pp. 10474-10483.

WU, X., LI, Y., CRISE, B. and BURGESS, S.M., 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science (New York, N.Y.), 300(5626), pp. 1749-1751.

WU, Y. and MARSH, J.W., 2001. Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. Science (New York, N.Y.), 293(5534), pp. 1503-1506.

XIA, Y., SHEN, S. and VERMA, I.M., 2014. NF-kappaB, an active player in human cancers. Cancer immunology research, 2(9), pp. 823-830.

XIANG, Z., LI, Y., CUN, A., YANG, W., ELLENBERG, S., SWITZER, W.M., KALISH, M.L. and ERTL, H.C., 2006. Chimpanzee adenovirus antibodies in humans, sub-Saharan Africa. Emerging infectious diseases, 12(10), pp. 1596-1599.

XU, K., MA, H., MCCOWN, T.J., VERMA, I.M. and KAFRI, T., 2001. Generation of a stable cell line producing high-titer self-inactivating lentiviral vecto. Molecular Therapy, 3(1), pp. 97-104.

XU, Z., TIAN, J., SMITH, J.S. and BYRNES, A.P., 2008. Clearance of adenovirus by Kupffer cells is mediated by scavenger receptors, natural antibodies, and complement. Journal of virology, 82(23), pp. 11705-11713.

XU, Z.N., SHEN, W.H., CHEN, H. and CEN, P.L., 2005. Effects of medium composition on the production of plasmid DNA vector potentially for human gene therapy. Journal of Zhejiang University.Science.B, 6(5), pp. 396-400.

YANG, L., SOONPAA, M.H., ADLER, E.D., ROEPKE, T.K., KATTMAN, S.J., KENNEDY, M., HENCKAERTS, E., BONHAM, K., ABBOTT, G.W.,

LINDEN, R.M., FIELD, L.J. and KELLER, G.M., 2008. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. Nature, 453(7194), pp. 524-528.

YANG, X., RODRIGUEZ, M., PABON, L., FISCHER, K.A., REINECKE, H., REGNIER, M., SNIADECKI, N.J., RUOHOLA-BAKER, H. and MURRY, C.E., 2014. Tri-iodo-l-thyronine promotes the maturation of human cardiomyocytes-derived from induced pluripotent stem cells. Journal of Molecular and Cellular Cardiology, 72, pp. 296-304.

YANG, Y., GREENOUGH, K. and WILSON, J.M., 1996. Transient immune blockade prevents formation of neutralizing antibody to recombinant adenovirus and allows repeated gene transfer to mouse liver. Gene therapy, 3(5), pp. 412-420.

YANG, Y., NUNES, F.A., BERENCSI, K., FURTH, E.E., GONCZOL, E. and WILSON, J.M., 1994. Cellular immunity to viral antigens limits E1-deleted adenoviruses for gene therapy. Proceedings of the National Academy of Sciences of the United States of America, 91(10), pp. 4407-4411.

YANG, Y., SU, Q. and WILSON, J.M., 1996. Role of viral antigens in destructive cellular immune responses to adenovirus vector-transduced cells in mouse lungs. Journal of virology, 70(10), pp. 7209-7212.

YASUDA, T., UENO, T., FUKUMURA, K., YAMATO, A., ANDO, M., YAMAGUCHI, H., SODA, M., KAWAZU, M., SAI, E., YAMASHITA, Y., MURATA, M., KIYOI, H., NAOE, T. and MANO, H., 2014. Leukemic evolution of donor-derived cells harboring IDH2 and DNMT3A mutations after allogeneic stem cell transplantation. Leukemia, 28(2), pp. 426-428.

YOKOO, N., BABA, S., KAICHI, S., NIWA, A., MIMA, T., DOI, H., YAMANAKA, S., NAKAHATA, T. and HEIKE, T., 2009. The effects of cardioactive drugs on cardiomyocytes derived from human induced pluripotent

stem cells. Biochemical and biophysical research communications, 387(3), pp. 482-488.

YOUNG, M.A., LARSON, D.E., SUN, C.W., GEORGE, D.R., DING, L., MILLER, C.A., LIN, L., PAWLIK, K.M., CHEN, K., FAN, X., SCHMIDT, H., KALICKI-VEIZER, J., COOK, L.L., SWIFT, G.W., DEMETER, R.T., WENDL, M.C., SANDS, M.S., MARDIS, E.R., WILSON, R.K., TOWNES, T.M. and LEY, T.J., 2012. Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. Cell stem cell, 10(5), pp. 570-582.

YU, J., HU, K., SMUGA-OTTO, K., TIAN, S., STEWART, R., SLUKVIN, I.I. and THOMSON, J.A., 2009. Human induced pluripotent stem cells free of vector and transgene sequences. Science (New York, N.Y.), 324(5928), pp. 797-801.

YU, J., VODYANIK, M.A., SMUGA-OTTO, K., ANTOSIEWICZ-BOURGET, J., FRANE, J.L., TIAN, S., NIE, J., JONSDOTTIR, G.A., RUOTTI, V., STEWART, R., SLUKVIN, I.I. and THOMSON, J.A., 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science (New York, N.Y.), 318(5858), pp. 1917-1920.

YUEN, K.S., YE, Z.W., FUNG, S.Y., CHAN, C.P. and JIN, D.Y., 2020. SARS-CoV-2 and COVID-19: The most important research questions. Cell & bioscience, 10, pp. 40-020-00404-4. eCollection 2020.

ZAMANIAN, M. and LA THANGUE, N.B., 1992. Adenovirus E1a prevents the retinoblastoma gene product from repressing the activity of a cellular transcription factor. The EMBO journal, 11(7), pp. 2603-2610.

ZHANG, J., HO, J.C.Y., CHAN, Y.C., LIAN, Q., SIU, C.W. and TSE, H.F., 2014. Overexpression of myocardin induces partial transdifferentiation of

human-induced pluripotent stem cell-derived mesenchymal stem cells into cardiomyocytes. Physiological Reports, 2(2),.

ZHANG, H., DORNADULA, G. and POMERANTZ, R.J., 1996. Endogenous reverse transcription of human immunodeficiency virus type 1 in physiological microenviroments: an important stage for viral infection of nondividing cells. Journal of virology, 70(5), pp. 2809-2824.

ZHANG, H. and VINOGRADOV, S.V., 2010. Short biodegradable polyamines for gene delivery and transfection of brain capillary endothelial cells. Journal of controlled release : official journal of the Controlled Release Society, 143(3), pp. 359-366.

ZHANG, L., BUTTON, B., GABRIEL, S.E., BURKETT, S., YAN, Y., SKIADOPOULOS, M.H., DANG, Y.L., VOGEL, L.N., MCKAY, T., MENGOS, A., BOUCHER, R.C., COLLINS, P.L. and PICKLES, R.J., 2009. CFTR delivery to 25% of surface epithelial cells restores normal rates of mucus transport to human cystic fibrosis airway epithelium. PLoS biology, 7(7), pp. e1000155.

ZHANG, W., GEIMAN, D.E., SHIELDS, J.M., DANG, D.T., MAHATAN, C.S., KAESTNER, K.H., BIGGS, J.R., KRAFT, A.S. and YANG, V.W., 2000. The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the p21WAF1/Cip1 promoter. The Journal of biological chemistry, 275(24), pp. 18391-18398.

ZHANG, X., JUSTICE, A.C., HU, Y., WANG, Z., ZHAO, H., WANG, G., JOHNSON, E.O., EMU, B., SUTTON, R.E., KRYSTAL, J.H. and XU, K., 2016. Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals. Epigenetics, 11(10), pp. 750-760.

ZHI, W.B., DENG, Q.Y., SONG, J.N. and OUYANG, F., 2005. Purification of ovalbumin from hen egg white by high-speed counter-current aqueous two-

phase chromatography. Sheng wu gong cheng xue bao = Chinese journal of biotechnology, 21(1), pp. 129-134.

ZHOU, P. and PU, W.T., 2016. Recounting Cardiac Cellular Composition. Circulation research, 118(3), pp. 368-370.

ZHOU, S., MODY, D., DERAVIN, S.S., HAUER, J., LU, T., MA, Z., HACEIN-BEY ABINA, S., GRAY, J.T., GREENE, M.R., CAVAZZANA-CALVO, M., MALECH, H.L. and SORRENTINO, B.P., 2010. A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. Blood, 116(6), pp. 900-908.

ZHOU, W. and FREED, C.R., 2009. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. Stem cells (Dayton, Ohio), 27(11), pp. 2667-2674.

ZHOU, Y., WANG, L., LIU, Z., ALIMOHAMADI, S., YIN, C., LIU, J. and QIAN, L., 2017. Comparative Gene Expression Analyses Reveal Distinct Molecular Signatures between Differentially Reprogrammed Cardiomyocytes. Cell reports, 20(13), pp. 3014-3024.

ZHU, J., WEI, D., CAO, X., LIU, Y. and YUAN, Z., 2001. Partitioning behaviour of cephalexin and 7-aminodeacetoxicephalosporanic acid in PEG/ammonium sulfate aqueous two-phase systems. Journal of Chemical Technology and Biotechnology, 76(11), pp. 1194-1200.

ZINDER, N.D. and LEDERBERG, J., 1952. Genetic exchange in Salmonella. Journal of Bacteriology, 64(5), pp. 679-699.

ZORN, A.M. and MASON, J., 2001. Gene expression in the embryonic Xenopus liver. Mechanisms of development, 103(1-2), pp. 153-157.

ZUFFEREY, R., NAGY, D., MANDEL, R.J., NALDINI, L. and TRONO, D., 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nature biotechnology, 15(9), pp. 871-875.

ZUFFEREY, R., DULL, T., MANDEL, R.J., BUKOVSKY, A., QUIROZ, D., NALDINI, L. and TRONO, D., 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. Journal of virology, 72(12), pp. 9873-9880.

ZYCHLINSKI, D., SCHAMBACH, A., MODLICH, U., MAETZIG, T., MEYER, J., GRASSMAN, E., MISHRA, A. and BAUM, C., 2008. Physiological promoters reduce the genotoxic risk of integrating gene vectors. Molecular therapy: the journal of the American Society of Gene Therapy, 16(4), pp. 718-725.