

# **Development of a novel *in vitro* 3D model to identify cancer genes via Insertional Mutagenesis**

**Sharmin Alhaque MSc, MPharm**

**Department of Life Sciences  
College of Health, Medicine and Life Sciences  
Brunel University of London**

May 2021



**Brunel**  
University  
London

Thesis submitted to Brunel University of London for the degree of Doctor of Philosophy

## **Acknowledgments**

I would like to thank my PhD supervisor Professor Michael Themis for providing laboratory space and giving me the opportunity to conduct this research project. Also, I am grateful for his constant encouragement, expertise and support endowed upon me. I would also like to express my special gratitude to Professor David Hay for his support and helpful discussion during my PhD training program. Special thanks to Dr Hassan Rashidi for sharing his wealth of practical expertise, endless technical advice and emotional support.

I would like to dedicate this thesis to my amazing family in Iran and Kurdistan. There is no doubt without their continued support I could not have come this far.

This Study was supported by generous studentship award from Brunel University London which I am forever grateful.

## **Declaration**

I declare that I have prepared this thesis and not submitted, in whole or in part for the award any degree to the Brunel University London or other learning research institutions. Except where stated otherwise by reference or acknowledgement, the work presented is entirely my own.

Sharmin Alhaque

## Abstract

The aim of the presented research project was to develop a personalised *in vitro*-based human model to test the genotoxicity of gene therapy viral vectors. For this purpose, human induced pluripotent stem cells (hiPSCs) and human induced pluripotent stem cell-derived 3D heps were used in combination with a number of lentiviral and adeno-associated viral vectors and molecular analysis was performed to detect insertion sites (IS) to assess the predictive power of 3D heps model to predict insertional mutagenesis.

In this study, two previously designed plasmid were used, one of which contained the U3 region of the LTR from the wild type (pHV) and considered as “Unsafe” vector, and the other was a self-inactivating (SIN) lentiviral vector with no U3 region of the LTR (pHR) and used as a “Safe” vector. Initially, high-titre production of both lentiviral vectors was achieved in the HEK293 cell line. The titre was calculated using flow cytometry analysis. The recombinant AAV serotype-2 vectors were kindly provided by our collaborators in Australia. These vectors constructed with “Clean ITR” with strong and weak promoters, driving the green fluorescent protein (GFP) expression.

In parallel, P106i, as an integration-free hiPSCs, was expanded and fully characterised, using flow cytometry, immunostaining and qPCR analysis. Expression of pluripotency cell surface markers such as SSEA4, TRA-1-81 and TRA-1-60 was detected in more than 90% of the cells using flow cytometry. The expression of POU domain class 5 transcription factor 1 (POU5F1 or OCT4) and SRY-Box Transcription Factor 2 (SOX2) as two major transcription factors regulating pluripotency were also confirmed at gene and protein levels by quantitative polymerase chain reaction (qPCR) and immunostaining, respectively. The high level of pluripotency markers confirmed the pluripotent state and the identity of P106i, which was used in this study.

Then, a recently published protocol to generate phenotypically stable 3D heps from human embryonic stem cells (hESCs) were amended and optimised to generate 3D heps from hiPSC efficiently. This protocol addresses issues surrounding previous 3D heps protocols, such as scalability and long-term *in vitro* phenotypic stability. Notably, hiPSC-derived 3D heps displayed liver functions for an extended period. Standard characterisation tests were performed on day 20 of differentiation using a range of molecular and cell biology techniques, including immunofluorescence analysis of liver-

specific markers such as HNF4-alpha and secretion of serum albumin (ALB) and alpha-fetoprotein (AFP) using ELISA assay. The result revealed a high level of ALB and a low level of AFP in hiPSC-derived 3D heps compared to conventional 2D hepatocyte-like cells.

Following characterisation, hiPSCs and 3D heps were transduced with recombinant AAV serotype-2 vectors (at MOI 1E5) and lentiviral vectors (at MOI 20). Modifications were made to enhance transduction efficiency in 3D heps. The outer layer of the 3D heps exhibited the highest level of GFP expression, which transduced with rAAV serotype-2 vectors. In 3D heps transduced with pHR and pHV lentiviral vectors, the same pattern was observed; however, a higher level of GFP expression was observed in cells transduced with the pHR lentiviral vector. The result of successful transduction was confirmed by PCR analysis. hiPSCs were also transduced with pHR and pHV lentiviral vectors at two different time points, and the result was confirmed via PCR analysis.

In order to see the effect of gene expression in the proximity of viral insertion at the individual cell level, single-cell cloning (SCC) was performed on hiPSCs which were transduced with pHR and pHV lentiviral vectors. The results showed positive GFP expression in hiPSCs, confirming transduction. The DNA and RNA of the single-cell clones, hiPSCs and 3D heps were collected and sent for downstream analysis of the insertion site (IS).

Upon viral integration, it is crucial to detect IS in the setting of clonal dominance. For bioinformatic analysis, lentiviral IS was retrieved by EPTS/LM-PCR and CIS were detected in P106i and 3D heps samples. This result revealed an overall decrease in IS overtime in transduced cells for both lentiviral vectors by comparison of the days 3 and 30 time points. The number of identified IS in 3D heps transduced with pHR and pHV lentiviral was low compared to P106i cells. There was also an apparent reduction in IS in P106i cells transduced with pHR and pHV lentiviral vectors over time, suggesting possible cell death during propagation and culture expansion.

Following EPTS/LM-PCR, identification of cancer-related genes IS, and CIS were performed in P106i cells and 3D heps. The result indicates that in P106i, the number of proto-oncogenes was higher in samples transduced with pHV lentiviral vectors, suggesting the possible effect of the full LTR vector on the number of cancer-related

genes. In addition, polyclonality of pHV and pHR was observed in analysed single-cell clones of P106i. However, in clone G transduced with pHV lentiviral vector, the gene *LINC01249* exhibited 93.4% of the viral sequence count. In addition, despite the polyclonality of the samples, the qPCR result revealed that the genes near the IS have increased in level of relative gene expression. In conclusion, the 3D heps provide a valuable in vitro tool to assess genotoxicity associated with viral vectors.

# Table of Contents

Acknowledgments.....	2
Declaration.....	3
Abstract.....	4
Table of Figures .....	14
List of Tables .....	19
Abbreviation.....	21
Chapter 1.....	25
1. Introduction.....	25
Prologue .....	25
1.1. Gene Therapy .....	25
1.1.1. History of Gene Therapy .....	25
1.1.2. Challenges of Gene Therapy .....	26
1.1.3. Gene Therapy vectors and their properties.....	27
1.1.4. Gene Therapy delivery routes ( <i>in vivo</i> & <i>ex vivo</i> ) .....	29
1.1.5. Stem Cells as “Carriers” for Gene Therapy .....	30
1.1.5.1. Adult Stem Cells (ASC) in Gene Therapy .....	30
1.1.5.2. Embryonic Stem Cell (ESC) in Gene Therapy.....	31
1.1.6. Challenges for Stem-Cell based gene therapy .....	33
1.1.7. Chimeric antigen recipient T (CAR-T) cell therapy.....	33
1.1.8. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas) technique ....	34
1.1.9. Different vector delivery systems for gene therapy .....	34
1.1.9.1. Non-viral delivery.....	35
1.1.9.2. Viral vector delivery .....	35
1.1.9.3. Adenoviral vectors (ADVs) .....	36
1.1.9.4. Adeno-associated virus (AAV).....	38
1.1.9.5. Retrovirus.....	41
1.1.9.6. Gamma-retrovirus and lentivirus life cycle.....	41
1.1.10. Generation of lentivirus-based vector .....	43
1.1.11. Generation of replication- incompetent viral vectors .....	43
1.1.12. Advantages and disadvantages of retroviral based vectors .....	43
1.1.13. Mechanisms of Insertional Mutagenesis .....	45
1.1.14. Other influencers of viral vector-mediated genotoxicity.....	47

1.1.15.	Genotoxicity Prevention Strategies .....	49
1.1.16.	Insights from key human gene therapy clinical trials .....	50
1.1.16.1.	HSC gene therapy trials for SCID-X1.....	51
1.1.16.2.	HSC gene therapy trials for ADA-SCID.....	51
1.1.16.3.	HSC gene therapy trials for X-CGD.....	52
1.1.16.4.	HSC gene therapy trials for WAS.....	52
1.1.16.5.	HSC gene therapy trials for X-ALD.....	53
1.1.17.	Importance of gene therapy preclinical models.....	53
1.1.17.1.	Examples of current <i>in vivo</i> genotoxicity preclinical models.....	54
1.1.18.	Challenges for genotoxicity screening .....	58
1.1.19.	Replacement, Reduction, Refinement (3Rs) benefits.....	58
1.2.	The Liver.....	60
1.2.1.	Human liver development .....	61
1.2.1.1.	Mammalian Embryonic Development.....	61
1.2.1.2.	Hepatic Endoderm Specification.....	63
1.2.2.	Formation of the Hepatic Bud from Hepatic Endoderm.....	65
1.2.2.1.	Proliferation and Migration .....	65
1.2.2.2.	Liver Bud Growth .....	66
1.2.2.3.	Hepatocyte Specification .....	67
1.2.2.4.	Hepatocyte Maturation .....	67
1.2.3.	Liver Architecture.....	69
1.3.	Liver cells as alternative sources for <i>in vitro</i> for toxicological and pharmacological screening .....	71
1.3.1.	Primary liver hepatocytes.....	72
1.3.2.	HepG2 cell line .....	73
1.3.3.	HepaRG cell line.....	73
1.4.	Pluripotent stem cells (PSCs).....	74
1.4.1.	Overview .....	74
1.4.2.	Key milestones in PSCs history and classifications .....	74
1.4.3.	PSCs definition & properties.....	75
1.4.4.	Embryonic Stem Cells (ESCs).....	75
1.4.5.	Limitations of ESCs.....	76
1.4.6.	Historical background of iPSCs research technology.....	77
1.4.6.1.	First generation.....	77
1.4.6.2.	Second generation .....	78



1.4.6.3. Human iPSCs .....	78
1.4.7. Reprogramming strategies and pluripotency induction .....	78
1.4.7.1. Integrating viral systems .....	79
1.4.7.2. Non-integrating viral systems .....	79
1.4.8. Properties of iPSCs .....	80
1.4.8.1. Morphological features.....	81
1.4.8.2. Genetic profiles .....	81
1.4.8.3. Differentiation potential .....	82
1.4.9. Available sources of adult somatic cells for reprogramming.....	83
1.4.10. Maintenance of iPSCs in feeder-free and <i>xeno</i> -free culture environments .....	83
1.4.10.1. Feeder-free cell culture.....	83
1.4.10.2. <i>xeno</i> -free culture medium .....	84
1.4.11. Applications and future directions of iPSCs.....	86
1.4.12. Generation of <i>in vitro</i> hepatocyte-like cells from PSCs .....	89
1.4.12.1. Direct differentiation of hESC into hepatocyte-like cells (HLC) .....	89
1.4.12.2. Endoderm Specification .....	90
1.4.12.3. Hepatic Specification and differentiation .....	91
1.4.12.4. Hepatic differentiation of iPSCs .....	92
1.5. 3D cell culture system .....	92
1.5.1. Overview .....	92
1.5.2. Two dimensional (2D) versus Three dimensional (3D) systems .....	93
1.5.3. Scaffold-free 3D methods .....	93
1.5.4. What are 3D organoids? .....	94
1.5.5. iPSC-derived 3D Liver Organoid Models .....	95
1.5.5.1. Comparison of different protocols for generation of 3D iPSC-derived hepatic organoids .....	95
1.5.5.2. Limitations of 3D <i>in vitro</i> models .....	97
1.5.5.3. 3D organoids versus 3D spheroids.....	99
1.6. Research study hypothesis .....	101
1.7. Main objectives: .....	101
2. Materials & Methods .....	102
2.1. Chemicals, Reagents, supplements and Buffers .....	102
2.1.1. Chemicals .....	102
2.1.2. Buffers.....	102
2.1.3. Consumables.....	103

2.1.4. Commercial Kits .....	104
2.2. Human Pluripotent Stem Cell culture and hepatic differentiation.....	104
2.2.1. Induced pluripotent stem cell lines .....	104
2.2.2. Matrix Coating using BD Matrigel™ (MG).....	105
2.2.3. Vitronectin (VT) coating .....	105
2.2.4. Coating of the plate with Recombinant Laminin 521 (LN-521) .....	106
2.2.5. Culturing hiPSCs .....	106
2.2.6. Passaging hiPSCs .....	106
2.2.7. Freezing and thawing hiPSCs .....	107
2.3. Hepatic differentiation of hiPSCs in 2D monolayer and 3D culture.....	108
2.3.1. 2D monolayer differentiation .....	108
2.3.2. Formation of self-aggregated hiPSCs spheroids (3D-Spheroids) .....	111
2.3.3. 3D Hepatic Differentiation .....	112
2.4. Characterisation.....	113
2.4.1. Immunofluorescent staining of 2D monolayer hiPSC and hiPSC-derived HLCs.....	113
2.4.2. Immunofluorescence (3D Spheroids) .....	114
2.4.3. Histology (3D spheroids).....	115
2.4.4. Imaging and acquisition .....	115
2.4.5. Flow cytometry .....	115
2.4.6. Enzyme-Linked Immunosorbent Assay (ELISA).....	116
2.4.7. Cytochrome P450 Assay.....	117
2.4.8. Cytochrome P450 drug inducibility.....	117
2.5. Molecular Biology Techniques .....	118
2.5.1. RNA extraction .....	118
2.5.2. DNA extraction.....	118
2.5.3. RNA purification .....	119
2.5.4. Reverse transcription (RT) .....	119
2.5.5. Quantitative polymerase chain reaction (qPCR).....	119
2.6. Protein Biochemistry Techniques .....	121
2.6.1. Cellular protein extraction .....	121
2.6.2. Measuring protein concentration.....	121
2.7. Vector production.....	121
2.7.1. Lentiviral vector constructs.....	121
2.7.2. Adeno associated virus (AAV) viral vector constructs .....	122
2.7.3. Bacterial culture and plasmid amplification .....	124
2.7.4. Restriction enzyme digest.....	124

2.7.5. Polymerase Chain Reaction (PCR).....	124
2.7.6. Growth and storage of bacterial cultures.....	125
2.7.7. Isolation of plasmid DNA by Miniprep.....	125
2.7.8. Isolation of plasmid DNA by Megaprep.....	126
2.7.9. Vector production in mammalian cell line.....	126
2.7.10. Lentiviral titration using flow cytometry.....	127
2.7.11. Preparation of lentivirus vector using GeneJuice™ transfection reagent.....	128
2.7.12. H1V-1 -derived lentiviral production using HEK293 cells and GeneJuice™.....	129
2.8. Transduction.....	129
2.8.1. Transduction of JHP106i hiPSCs.....	129
2.8.2. Transduction of 3D heps.....	130
2.8.3. Long-term outgrowth 3D culture for genotoxicity.....	130
2.9. Single-Cell Cloning (SCC) of the transduced hiPSCs.....	131
2.9.1. Cloning efficiency.....	131
2.9.2. Preparation of conditioned medium.....	131
2.9.3. Single cell cloning.....	131
2.10. Statistics and Bioinformatic.....	132
Chapter 3.....	134
3.1. Introduction.....	134
3.1.1. Integrating viral vectors and Insertional Mutagenesis.....	134
3.1.2. <i>in vivo</i> & <i>in vitro</i> genotoxicity models.....	135
3.1.3. Liver and mature hepatocytes.....	136
3.1.4. 3D hepatic differentiation.....	137
3.2. Objectives.....	137
3.3. Results.....	138
3.3.1. Characterisation of the hiPSCs population.....	138
3.3.2. Culture and characterisation of hiPSCs maintained in mTeSR1™.....	138
3.3.3. Morphological analysis of healthy and unhealthy colonies.....	138
3.3.4. Flow cytometry analysis.....	141
3.3.5. Stemness nature of hiPSCs.....	141
3.3.6. 2D hepatic differentiation of human induced pluripotent stem cells, using a serum-free approach.....	144
3.3.7. Development and optimisation of 3D heps for genotoxicity assay.....	149
3.3.8. Development of the 3D platform to generate hPSC-derived 3D heps.....	150
3.3.9. Morphological analysis of 2D and 3D hepatic differentiation of P106-iPSC line.....	152

3.3.10. Functional characterisation of 2D P106 and 3D heps.....	153
3.3.11. Characterisation of P106-derived 3D heps.....	155
3.4. Discussion.....	161
Chapter 4.....	168
4.1. Introduction .....	168
4.2. Objectives .....	173
4.3. Results.....	173
4.3.1. High titer GFP expressing lentiviral vector preparation .....	173
4.3.2. Transduction of P106i cell line with pHR lentiviral vector at MOIs 10, 20 .....	175
4.3.3. Cell viability analysis for P106i cells transduced with pHR lentiviral vector .....	176
4.3.4. Brightfield microscopy analysis of the P106i cells transduced with pHR lentiviral vector ...	177
4.3.5. Transduction of P106i cell line with pHV lentiviral vector, using MOIs 10 and 20. ....	178
4.3.6. Recombinant AAV/GFP serotype 2 transgene (GFP) expression in P106i cells. ....	179
4.3.7. Brightfield microscopy analysis of the P106i cell line following transduction with rAAV/GFP serotype 2 vector. ....	180
4.3.8. Cell viability and fluorescent microscopy analysis for P106i cells transduced with rAAV/GFP serotype-2 vector.....	181
4.3.9. Transduction of <i>in vitro</i> 3D heps, using the rAAV/GFP serotype 2 vectors with different liver specific capsids.....	184
4.3.10. Transduction of <i>in vitro</i> 3D heps with pHR and pHV lentiviral vectors at MOI 20. ....	189
4.3.11. DNA/RNA extraction of P106i hiPSCs and 3D heps transduced with rAAV/GFP and lentiviral vectors.....	189
4.3.12. Vector copy number .....	193
4.3.13. Detection of pHR and pHV lentiviral and rAAV/GFP insertions in P106i and 3D heps .....	195
4.3.14. Single-Cell Cloning (SCC) of P106i hiPSCs transduced with pHR and pHV lentiviral vectors .....	199
4.3.15. Cloning efficiency of P106i transduced with pHR and pHV lentiviral vectors .....	200
4.3.16. Detection of pHR and pHV lentiviral insertions in P106i transduced clones.....	205
4.4 Discussion.....	211
Chapter 5.....	216
5.1. Introduction .....	216
5.1.1. Gene Transfer Technology in Gene Therapy .....	216
5.1.2. Key mechanisms of IM-induced Oncogenesis in the SCID-X1 trials.....	217
5.1.3. Preclinical Studies ( <i>in vivo</i> & <i>in vitro</i> ) for IM.....	220
5.1.4. Strategies to minimise the incidence of IM .....	221

5.1.5. Methods for tracking virus Integration Sites .....	223
5.1.6. LM-PCR & LAM-PCR methods .....	224
5.1.7. Non-restrictive linear-amplification-mediated PCR (nrLAM-PCR) .....	225
5.1.8. Solid-phase ligation-mediated PCR (EPTS/LM-PCR) technique .....	225
5.2. Objectives .....	226
5.3. Results .....	226
5.3.1. Integration site analysis (IS) on bulk P106i iPSC line and P106i-derived 3D heps transduced with pHR and pHV lentiviral vectors. ....	226
5.3.2. Insertion site analysis on P106i samples transduced with pHR and pHV lentiviral vectors at two different time points.....	228
5.3.3. Identification of total insertion site in genes found in P106i cells transduced with pHR and pHV lentiviral vectors at two time points .....	236
5.3.4. Identification of insertion sites in genes found in P106i derived 3D heps transduced with pHR and pHV lentiviral vectors .....	238
5.3.5. Identification of cancer related genes IS and common IS in P106i cells and 3D heps.....	241
5.3.6. Gene enrichment analysis.....	243
5.3.7. Single-Cell clones Insertion Site analysis.....	249
5.3.8. Relative gene expression of genes in proximity to the insertion sites in P106i clones .....	259
5.4. Discussion.....	260
6. 1 Discussion & Final Remarks .....	265
7. References .....	269
8. Supplementary Data .....	308
8.1. Publications.....	308

## Table of Figures

FIGURE 1.1: GENE THERAPY TREATMENT. ....	25
FIGURE 1.2: SCHEMATIC REPRESENTATION OF SOMATIC CELL GENE THERAPY. ....	27
FIGURE 1.3: GENE THERAPY VECTOR. VECTORS ARE USED FOR GENE TRANSFER APPLICATION. ....	28
FIGURE 1.4: GENE THERAPY DELIVERY ROUTES (IN VIVO & EX VIVO). ....	30
FIGURE 1.5: DIFFERENT APPROACHES FOR DELIVERING THERAPEUTIC TRANSGENES INTO PATIENTS. ....	32
FIGURE 1.6: ATTACHMENT AND REPLICATION OF ADV. ....	38
FIGURE 1.7: GENOMIC STRUCTURE OF WILD TYPE ADENO-ASSOCIATED VIRUS. ....	39
FIGURE 1.8: STRUCTURE OF RECOMBINANT ADENO-ASSOCIATED DERIVED VECTOR. ....	40
FIGURE 1.9: ADENO-ASSOCIATED VIRAL VECTOR SEROTYPES. ....	41
FIGURE 1.10: THE GENERAL LIFE CYCLE OF RETROVIRUS/LENTIVIRUS. ....	42
FIGURE 1.11. THE WILD TYPE LENTIVIRUS GENOME ORGANISATION AND GENERATION OF LENTIVIRUS DERIVED VECTOR. ....	44
FIGURE 1.12: DIFFERENT MECHANISMS OF VIRAL VECTOR-MEDIATED GENOTOXICITY. ....	47
FIGURE 1.13: EARLY EMBRYONIC GUT DEVELOPMENT. ....	62
FIGURE 1.14: HEPATIC ENDODERM SPECIFICATION. ....	65
FIGURE 1.15: THE LIVER. ....	72
FIGURE 1.16: SCHEMATIC OF DIFFERENT TYPES OF PSCs. ....	76
FIGURE 1.17: THE FOUR DIFFERENT KEY TECHNIQUES FOR DELIVERING REPROGRAMMING FACTORS. ....	80
FIGURE 1.18: TECHNIQUES FOR GENERATING 3D MICROTISSUES. ....	94
FIGURE 1.19: LIVER ORGANOIDS CAN BE PRODUCED FROM PSCs BY REGULATING SIGNALLING PATHWAYS DURING DIFFERENTIATION. ....	98
FIGURE 1.20: MORPHOLOGY OF ORGANOIDS VERSUS SPHEROIDS. ....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
FIGURE 2.1: FABRICATION OF AGAROSE MOLD USING COMMERCIALY AVAILABLE MICROTISSUE MOLD. ....	112
FIGURE 2.2: THE HIV TYPE 1 (HIV-1) DERIVED LENTIVIRAL VECTORS. ....	122
FIGURE 2.3: TOXICOLOGY DETAILS OF RAAV VECTOR DESIGN CONSTRUCTS EMPLOYED IN THIS STUDY. ....	123
FIGURE 3.1: MORPHOLOGY OF HIPSC COLONIES. ....	139
FIGURE 3.2: MORPHOLOGY ANALYSIS OF DIFFERENT TYPES OF HIPSCs COLONIES. ....	140
FIGURE 3.3: EXPRESSION OF STEM CELL SURFACE MARKERS IN HIPSCs PASSAGE 15. ....	142
FIGURE 3.4: EXPRESSION OF STEM CELL SURFACE MARKERS IN HIPSCs PASSAGE 17. ....	143
FIGURE 3.5: STEMNESS ANALYSIS OF HIPSCs POPULATION. ....	144

FIGURE 3.6: FLOW DIAGRAM OF THE HEPATOCYTE DIFFERENTIATION PROTOCOL IN SERUM-FREE MEDIA.....	145
FIGURE 3.7: MORPHOLOGY ANALYSIS OF DIFFERENT CELL DENSITY ON 33D6 HEPATIC DIFFERENTIATION. ....	146
FIGURE 3.8: MORPHOLOGY OF 2D HEPATIC DIFFERENTIATION OF 33D6-IPSC LINE.....	147
FIGURE 3.9: STANDARD CHARACTERISATION TESTS WERE PERFORMED ON DAY 20 MATURE HIPSC-DERIVED HLCs. ....	148
FIGURE 3.10: LIVER FUNCTIONALITY TESTS WERE PERFORMED ON 33D6-DERIVED MATURE HLCs. ....	148
FIGURE 3.11: STATIC SUSPENSION CULTURE. ....	150
FIGURE 3.12: FORMATION OF SPHERES IN AGAROSE MICROPLATES USING THE 3D PETRI DISH® MOULD. ....	151
FIGURE 3.13: THREE DIFFERENT SIZE RANGES FOR 3D SPHEROIDS .....	151
FIGURE 3.14: MORPHOLOGY OF 2D HEPATIC DIFFERENTIATION OF P106I CELL LINE. ....	152
FIGURE 3.15: MORPHOLOGY OF 3D HEPATIC DIFFERENTIATION OF P106-IPSC HEPATOSPHERES. ....	153
FIGURE 3.16: MORPHOLOGY OF 3D HEPS AND MATURATION PHASE. ....	154
FIGURE 3.17: ALB PROTEIN PRODUCTION OF 2D HLCs AND 3D HEPS.....	155
FIGURE 3.18: REPRESENTATIVE HISTOLOGY OF 3D P106-DERIVED HEPATOSPHERES. ....	156
FIGURE 3.19: GENE EXPRESSION ANALYSIS OF 3D P106-DERIVED SPHEROIDS AT DIFFERENT STAGES OF DIFFERENTIATION.....	157
FIGURE 3.20: IMMUNOFLUORESCENCE ANALYSIS OF 3D P106-DERIVED SPHEROIDS.....	160
FIGURE 4.1: HIGH TITRE PHR LENTIVIRAL PREPARATION AND TITRATION, USING FLOW CYTOMETRY ANALYSIS.....	174
FIGURE 4.2: TRANSDUCTION OF P106I CELL LINE USING PHR LENTIVIRAL VECTOR AT MOIS 10 AND 20. ....	176
FIGURE 4.3: REPRESENTATIVE BRIGHTFIELD IMAGE OF P106I CELLS TRANSDUCED WITH PHR LENTIVIRAL VECTOR. ....	178
FIGURE 4.4: TRANSDUCTION OF P106I CELL LINE, USING PHV LENTIVIRAL VECTOR.....	179
FIGURE 4.5: REPRESENTATIVE BRIGHTFIELD IMAGE OF P106I CELLS TRANSDUCED WITH RAAV/GFP SEROTYPE 2 AAV VECTOR. ....	180
FIGURE 4.6: TRANSDUCTION OF P106I CELL LINE, USING RAAV/GFP SEROTYPE-2 VECTOR....	182
FIGURE 4.7: TRANSDUCTION OF P106I CELL LINE, USING THE RAAV/CB7/GFP SEROTYPE-2 VECTOR.....	183
FIGURE 4.8: TRANSDUCTION OF P106I CELL LINE, USING THE RAAV/APOE/GFP SEROTYPE-2 VECTOR.....	184

FIGURE 4.9: TRANSDUCTION OF IN VITRO 3D HEPS, USING RAAV/APOE/GFP SEROTYPE-2 VECTOR WITH AAV2, LK03, AAV8 AND NP59 LIVER SPECIFIC CAPSIDS.....	185
FIGURE 4.10: MODIFICATIONS TO ENHANCE TRANSDUCTION EFFICIENCY IN 3D HEPS. ....	186
FIGURE 4.11: TRANSDUCTION OF IN VITRO 3D HEPS, USING THE RAAV/APOE/GFP SEROTYPE 2 VECTORS WITH AAV2, AAV8 AND DJ LIVER SPECIFIC CAPSIDS. ....	187
FIGURE 4.12: TRANSDUCTION OF IN VITRO 3D HEPS, USING RAAV/APOE/GFP SEROTYPE 2 VECTORS WITH LK03 AND NP59 LIVER SPECIFIC CAPSIDS. ....	188
FIGURE 4.13: TRANSDUCTION OF IN VITRO 3D HEPS, USING RAAV/APOE/GFP AND RAAV/CB7/GFP SEROTYPE 2 VECTORS.....	190
FIGURE 4.14: TRANSDUCTION OF IN VITRO 3D HEPS, USING PHR AND PHV LENTIVIRAL VECTORS. REPRESENTATIVE FLUORESCENT IMAGES SHOWING 3D HEPS INFECTED WITH GFP EXPRESSING PHR AND PHV LENTIVIRAL VECTORS AT MOI 20. ....	191
FIGURE 4.15: QUANTIFICATION OF VECTOR COPY NUMBER (VCN) IN P106I AND 3D HEPS, USING RAAV/GFP AND LENTIVIRAL VECTORS. ....	194
FIGURE 4.16: DETECTION OF LENTIVIRAL INSERTIONS IN P106I HIPSCs TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS. ....	196
FIGURE 4.17: DETECTION OF RAAV VIRAL INSERTIONS IN P106I HIPSCs TRANSDUCED WITH RAAV/APOE AND RAAV/CB7 VECTORS, USING PCR METHOD. ....	197
FIGURE 4.18: DETECTION OF LENTIVIRAL INSERTIONS IN 3D HEPS TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS. ....	198
FIGURE 4.19: DETECTION OF RAAV VIRAL INSERTIONS IN 3D HEPS TRANSDUCED WITH RAAV/APOE AND RAAV/CB7 VECTORS USING PCR. ....	199
FIGURE 4.20: CLONING EFFICIENCY OF P106I HIPSCs TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS.....	200
FIGURE 4.21: FLOW CYTOMETRY ANALYSIS OF P106I CELLS TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS PRIOR TO SCC.....	201
FIGURE 4.22: MORPHOLOGY OF SINGLE CLONE FOLLOWING SCC.....	202
FIGURE 4.23: MORPHOLOGY OF SINGLE CLONES DURING EXPANSION.....	202
FIGURE 4.24: MORPHOLOGY OF SINGLE CLONES DURING EXPANSION.....	203
FIGURE 4.25: FLOW CYTOMETRY ANALYSIS OF SINGLE CLONES TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS.....	204
FIGURE 4.26: DETECTION OF LENTIVIRAL INSERTIONS IN SINGLE CLONES TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS. ....	206
FIGURE 4.27: FLUORESCENT MICROSCOPY AND FLOW CYTOMETRY ANALYSIS OF P106I CELLS TRANSDUCED WITH PHR LENTIVIRAL VECTOR.....	207
FIGURE 4.28: REPRESENTATIVE FLUORESCENT IMAGES OF SINGLE CLONES A-Q TRANSDUCED WITH PHR LENTIVIRAL VECTOR SHOWING POSITIVE GFP EXPRESSION. ....	208



FIGURE 4.29: FLUORESCENT MICROSCOPY IMAGING OF EXPANDED SINGLE CLONES .....	209
FIGURE 4.30: DETECTION OF LENTIVIRAL INSERTIONS IN SINGLE CLONES TRANSDUCED WITH PHR LENTIVIRAL VECTORS.....	210
FIGURE 5.1: SCHEMATIC REPRESENTATION OF FIRST MECHANISM OF <i>RETROVIRIDAE</i> -BASED GENE TRANSFER VECTORS. ....	219
FIGURE 5.2: SCHEMATIC REPRESENTATION OF FIRST MECHANISM OF <i>RETROVIRIDAE</i> -BASED GENE TRANSFER VECTOR INDUCED INSERTIONAL MUTAGENESIS.....	220
FIGURE 5.3: SCHEMATIC REPRESENTATION OF THE SECOND MECHANISM OF <i>RETROVIRIDAE</i> - BASED GENE TRANSFER VECTORS INDUCED INSERTIONAL MUTAGENESIS. ....	220
FIGURE 5.4: INTEGRATION SITE ANALYSIS ON BULK P106I iPSCs AND 3D HEPS.....	227
FIGURE 5.5: TOP 10 INSERTIONS SITES (IS) IN P106I iPSCs TRANSDUCED WITH PHR LENTIVIRAL VECTOR AT DAY 3. ....	229
FIGURE 5.6: TOP 10 INSERTIONS SITES (IS) IN P106I iPSCs TRANSDUCED WITH PHR LENTIVIRAL VECTOR AT DAY 30. ....	22931
FIGURE 5.7: TOP 10 INSERTIONS SITES (IS) IN P106I iPSCs TRANSDUCED WITH PHV LENTIVIRAL VECTOR AT DAY 3. ....	233
FIGURE 5.8: TOP 10 INSERTIONS SITES (IS) IN P106I iPSCs TRANSDUCED WITH PHV LENTIVIRAL VECTOR AT DAY 30. ....	235
FIGURE 5.9: COMPARISON OF TOP 10 GENES IN P106I CELLS TRANSDUCED WITH PHV LENTIVIRAL VECTOR AT DAYS 3 AND 30 TIMEPOINTS.....	237
FIGURE 5.10: COMPARISON OF TOP 10 GENES IN P106I CELLS TRANSDUCED WITH PHR LENTIVIRAL VECTOR AT DAYS 3 AND 30 TIMEPOINTS. ....	238
FIGURE 5.11: COMPARISON OF TOP 10 GENES IN P106I-DERIVED 3D HEPS TRANSDUCED WITH PHV AND PHR LENTIVIRAL VECTORS. ....	239
FIGURE 5.12: COMPARISON OF TOP 10 GENES IN P106I CELLS AND 3D HEPS. ....	240
FIGURE 5.13: COMPARISON OF TOP 10 GENES IN P106I CELLS AND 3D HEPS. ....	241
FIGURE 5.14: THE QUANTIFICATION OF AVERAGE ASQ PER IS GENE IN P106I SAMPLES TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS AT TWO TIME POINTS.....	248
FIGURE 5.15: TOP 10 INSERTION SITES IN P106I iPSCs CELLS WHICH WERE TRANSDUCED WITH PHV LENTIVIRAL VECTOR AND SINGLE CELL CLONED MANUALLY. ....	252
FIGURE 5.16: TOP 10 INSERTION SITES IN P106I iPSCs CELLS WHICH WERE TRANSDUCED WITH PHR LENTIVIRAL VECTOR AND SINGLE CELL CLONED MANUALLY. ....	253
FIGURE 5.17: TOP 10 INSERTION SITES IN P106I iPSCs CELLS WHICH WERE TRANSDUCED WITH PHR LENTIVIRAL VECTOR.....	254
FIGURE 5.18: TOP 10 INSERTION SITES IN P106I iPSCs TRANSDUCED WITH PHR LENTIVIRAL VECTOR.....	255

FIGURE 5.19: TOP 10 INSERTION SITES IN P106I IPSCs TRANSDUCED WITH PHR LENTIVIRAL VECTOR.....	256
FIGURE 5.20: TOP 10 INSERTION SITES IN P106I IPSCs TRANSDUCED WITH PHR LENTIVIRAL VECTOR.....	257
FIGURE 5.21: THE QUANTITY OF TOTAL IS AND CIS IN P106I CLONES TRANSDUCED WITH PHV LENTIVIRAL VECTORS.....	258
FIGURE 5.22: THE QUANTITY OF TOTAL IS AND CIS IN P106I CLONES TRANSDUCED WITH PHR LENTIVIRAL VECTORS.....	258
FIGURE 5.23: THE RELATIVE GENE EXPRESSION OF P106I CLONES TRANSDUCED WITH PHR LENTIVIRAL VECTOR. ....	259
FIGURE 5.24: THE RELATIVE GENE EXPRESSION OF P106I CLONES TRANSDUCED WITH PHV LENTIVIRAL VECTOR. ....	260

## List of Tables

TABLE 1.1: SUMMARY OF DEFINED CULTURE MEDIA AND COMPOSITIONS FOR MAINTENANCE OF PLURIPOTENT STEM CELLS.....	87
TABLE 1.2: COMPARISON OF KEY HEPATIC DIFFERENTIATION PROTOCOLS FOR IPSC DERIVED 3D LIVER ORGANOID AND HEPATOSPHERES.....	100
TABLE 2.1: LIST OF CHEMICALS.....	102
TABLE 2.2: LIST OF BUFFERS.....	102
TABLE 2.3: LIST OF CONSUMABLES.....	103
TABLE 2.4: LIST OF COMMERCIAL KITS.....	104
TABLE 2.5: LIST OF CELL CULTURE SUPPLEMENTS.....	104
TABLE 2.6: VOLUME OF MATRIGEL REQUIRED FOR EACH SPECIFIC PLATE FORMATS.....	105
TABLE 2.7: SUMMARY OF ESSENTIAL MEDIA AND REAGENTS USED FOR STEM CELL CULTURE MAINTENANCE.....	107
TABLE 2.8: VOLUME OF mTESR™1 REQUIRED FOR PASSAGING hiPSCs IN DIFFERENT PLATE FORMATS.....	107
TABLE 2.9: THE SURFACE AREA (SA) BASED ON DIFFERENT PLATE FORMATS.....	109
TABLE 2.10: ENDODERM INDUCTION MEDIUM (EIM) COMPOSITION.....	109
TABLE 2.11: HEPATOBLAST DIFFERENTIATION MEDIUM (HDM) COMPOSITION.....	110
TABLE 2.12: HEPATOCYTE MATURATION MEDIUM (HMM) COMPOSITION.....	110
TABLE 2.13: THE REQUIRED VOLUME OF RPMI MEDIUM WITH ESSENTIAL GROWTH FACTORS DURING THE FIRST PHASE OF HEPATIC DIFFERENTIATION FOR DIFFERENT PLATE FORMATS.....	111
TABLE 2.14: THE REQUIRED VOLUME OF HDM AND HMM DURING THE SECOND PHASE OF HEPATIC DIFFERENTIATION FOR DIFFERENT PLATE FORMATS.....	111
TABLE 2.15: SUMMARY OF 2D HEPATIC DIFFERENTIATION STAGES AND THE DAYS THAT REQUIRE MEDIUM CHANGING OR MEDIUM SWITCHING.....	111
TABLE 2.16: 3D MEDIUM COMPOSITION.....	113
TABLE 2.17. SUMMARY OF PRIMARY ANTIBODIES USED FOR CHARACTERISATION OF hiPSCs AND MATURE HEPATOCYTES IN THIS STUDY.....	114
TABLE 2.18. SUMMARY OF SECONDARY ANTIBODIES USED IN THIS STUDY.....	114
TABLE 2.19. SUMMARY OF ANTIBODIES USED FLOW CYTOMETRY ANALYSIS.....	116
TABLE 2.20. SUMMARY OF THE OLIGONUCLEOTIDES AND THEIR SEQUENCES USED IN THIS STUDY.....	120
TABLE 2.21. SUMMARY OF AAV VECTORS USED IN THIS STUDY.....	123
TABLE 2.22: LIST OF RT-PCR REAGENTS.....	125
TABLE 2.23: LIST OF TRANSFECTION REAGENTS.....	128

TABLE 4.1. SUMMARY OF THE AVERAGE NUMBER OF CELLS, THE CELL VIABILITY AND THE LEVEL OF GFP EXPRESSION IN P106 HIPSC LINE TRANSDUCED WITH PHR VECTOR AT DIFFERENT TIMEPOINTS, USING MOIS 10 AND 20.....	177
TABLE 4.2. THE AVERAGE NUMBER OF CELLS, THE CELL VIABILITY AND THE LEVEL OF GFP EXPRESSION IN P106i CELL LINE TRANSDUCED WITH RAAV/GFP SEROTYPE-2 VECTOR AT DIFFERENT TIME POINTS USING MOI (1E5).....	181
TABLE 4.3: THE DNA SAMPLES FOR P106i SAMPLES TRANSDUCED WITH THE LENTIVIRAL AND RAAV/GFP VECTORS.....	192
TABLE 4.4: THE RNA SAMPLES FOR P106i SAMPLES TRANSDUCED WITH THE LENTIVIRAL AND RAAV/GFP VECTORS.....	192
TABLE 4.5: THE DNA SAMPLES FOR 3D HEPS TRANSDUCED WITH LENTIVIRAL AND RAAV/GFP VECTORS.....	192
TABLE 4.6: THE RNA SAMPLES FOR 3D HEPS TRANSDUCED WITH LENTIVIRAL AND RAAV/GFP VECTORS.....	193
TABLE 4.7: THE DNA SAMPLES FOR SINGLE CLONES TRANSDUCED WITH LENTIVIRAL VECTORS	204
TABLE 4.8: THE RNA SAMPLES FOR SINGLE CLONES TRANSDUCED WITH LENTIVIRAL VECTORS	205
TABLE 5.1. IDENTIFICATION OF NUMBER OF CANCER RELATED GENES IN P106i AND 3D HEPATOSPHERE SAMPLES TRANSDUCED WITH PHR AND PHV LENTIVIRAL VECTORS. ....	242
TABLE 5.2. THE TOP 10 BIOLOGICAL PATHWAYS IDENTIFIED IN P106i SAMPLES TRANSDUCED WITH PHR LENTIVIRAL VECTOR AT DAY 3 AND DAY 30 TIME POINTS. ....	244
TABLE 5.3. THE TOP 10 BIOLOGICAL PATHWAYS IDENTIFIED IN 3D HEPS TRANSDUCED WITH PHR LENTIVIRAL VECTOR AT DAY 3 POST TRANSDUCTION.....	245
TABLE 5.4. THE TOP 10 BIOLOGICAL PATHWAYS IDENTIFIED IN P106i SAMPLES TRANSDUCED WITH PHV LENTIVIRAL VECTOR AT DAY 3 AND DAY 30 TIME POINTS. ....	246
TABLE 5.5. THE TOP 10 BIOLOGICAL PATHWAYS IDENTIFIED IN 3D HEPS TRANSDUCED WITH PHV LENTIVIRAL VECTOR AT DAY 3 POST TRANSDUCTION.....	247
TABLE 5.6. THE NINE IDENTIFIED PROTO-ONCOGENES IN P106i SAMPLES TRANSDUCED WITH PHR LENTIVIRAL VECTOR. ....	249
TABLE 5.7. THE NINE IDENTIFIED PROTO-ONCOGENES IN P106i SAMPLES TRANSDUCED WITH PHV LENTIVIRAL VECTOR. ....	249

## Abbreviation

2D	Two-dimensional
3D	Three-dimensional
3D heps	3D hepatospheres
3Rs	Replacement, Reduction, Refinement
AIDS	Acquired Immunodeficiency Disease Syndrome
AFP	Alpha-fetoprotein
ALB	Albumin
Alpha-SMA	Alpha-Smooth Muscle Actin
AP	Alkaline Phosphatase
ApoE	Apolipoprotein E
ASC	Adult Stem Cells
ASLV	Avian Sarcoma Leukosis Virus
ASQ	Absolute Sequence Count
ADV	Adenovirus
AAV	Adeno-Associated virus
AAP	Assembly Activating Proteins
ADA-SCID	Adenosine Deaminase Severe Combined Immunodeficiency
ALDP	Adrenoleukodystrophy Protein deficiency
BSA	Bovine Serum Albumin
BCA	Bicinchoninic Acid
CAR-T	Chimeric Antigen Receptor-T
CDKN2A	Cyclin-dependant kinase 2 A
CIS	Common Insertion Sites
cGMP	Current Good Manufacturing Practice
CRISPR-CAS	Clustered Regulatory Interspaced Short Palindromic Repeats-Associated Proteins
cPPT	Central Polypurine tract element
CT	Cycle Threshold
CAR	Adenovirus Receptor
CM	Conditioned Medium
CNS	Central Nervous System
D3	Day three post transduction
D30	Day thirty post transduction
DS	Double Stranded
DNA	Deoxyribonucleic acid

DMSO	Dimethyl Sulfoxide
EIM	Endoderm Induction Medium
eGFP	Enhanced Green Fluorescent Protein
ELISA	Enzyme-Linked Immunosorbent Assay
ESC	Embryonic Stem Cells
ENV	Envelope gene
Epi	Episomal Reprogramming
EPTS/LM-PCR	Solid-Phase Ligation Mediated Polymerase Chain Reaction
EB	Embryoid Body
ECM	Extracellular Matrix
FACS	Fluorescence-Activated Cell Sorting
FGF	Fibroblast Growth Factor
FCS	Foetal Calf Serum
GFP	Green Fluorescent Protein
GT	Gene Therapy
GALV	Gibbon Ape Leukaemia Virus
GDF3	Growth and Differentiation Factor 3
HepG2	Human hepatocarcinoma cell line
HDM	Hepatoblast Differentiation Medium
HMM	Hepatocyte Maturation Medium
HGF	Hepatocyte Growth Factor
HSC	Hematopoietic Stem Cells
HLA	Human Leukocyte Antigen
HLC	Hepatocyte-like Cells
HIV-1	Human Immunodeficiency Virus type-1
hAAT	Human Alpha-1 Antitrypsin
hESCs	Human Embryonic Stem Cells
HEK293	Human Embryonic Kidney 293
HSS	DNAase-I Hypersensitive Sites
HCC	Hepatocellular Carcinoma
hiPSCs	Human Induced Pluripotent Stem Cells
hPSCs	Human Pluripotent Stem Cells
IM	Insertional Mutagenesis
ITR	Inverted Terminal Repeats
IL2RG	Interlukin-2 Receptor Subunit Transgene
IS	Insertion Site/Integration Site

IP	Ingenuity Pathway
iPSCs	Induced Pluripotent Stem Cells
KOSR	Knock-Out Serum Replacement
KO	Knock-Out
LB	Luria-Bertani media
LDL	Low Density Lipoprotein
LM-PCR	Ligation-Mediated Polymerase Chain Reaction
LAM-PCR	Linear Amplification Mediated Polymerase Chain Reaction
LN-521	Laminin-521
LSP	Liver specific promoter
LTR	Long Terminal Repeats
LV	Lentivirus
MHC-1	Major Histocompatibility Complex-1
MLV	Murine Leukaemia Virus
MSC	Mesenchymal Stem Cells
MEFs	Mouse Embryonic Fibroblasts
MOI	Multiplicity Of Infection
NEAA	Non-Essential Amino Acids
nr-LAM-PCR	Non-restrictive linear amplification mediated polymerase chain reaction
OSKM	Yamanaka Factors
OCT-4	Octamer-Binding Transcription factor-4
OSM	Oncostatin M
PCR	Polymerase Chain Reaction
PS	Packaging Signal
PGK	Phosphoglycerate Kinase
PSCs	Pluripotent Stem Cells
PO	Proto-Oncogenes
Poly A	Polyadenylation Signal
qPCR	Quantitative Polymerase Chain Reaction
rAAV	Recombinant AAV
RGD	Arginine-Glycine-Aspartic acid sequence
RV	Retrovirus
RNA	Ribonucleic acid
RSQ	Relative Sequence Count
RT	Reverse Transcription
SA	Surface Area

SB	Sleeping Beauty
SS	Single-Stranded
SFFV	Spleen Focus Forming Virus
SSEA	Stage-Specific Embryonic Antigen
SCC	Single-Cell cloning
SCID-X1	X-linked Severe Combined Immunodeficiency
SCNT	Somatic Cell Nuclear Transfer
SCI	Spinal Cord Injury
SeV	Sendai Virus
SIV	Simian Immunodeficiency Virus
SIN vectors	Self-Inactivating retroviral vectors
Sox-2	Sex determining region Y-box 2
TBG	Thyroxine-Binding Globulin
Tert	Telomerase Reverse Transcriptase
TSS	Transcriptional Start Sites
TSG	Tumour Suppressor Gene
T-ALL	T-cell Acute Lymphoblastic Leukaemia
TRA	Tumour Related Antigen
UTR	Untranslated Region
VCN	Vector Copy Number
VSV-G	Vesicular Stomatitis Virus-G
VSELS	Very Small Embryonic Like Stem Cells
VG	Vector Gram
VT	Vitronectin
WPRE	Woodchuck Post-transcriptional Regulatory Element
WAS	Wiskott-Aldrich Syndrome
X-CGD	X-linked Chronic Granulomatous Disease
X-ALD	X-linked adrenoleukodystrophy



# Chapter 1

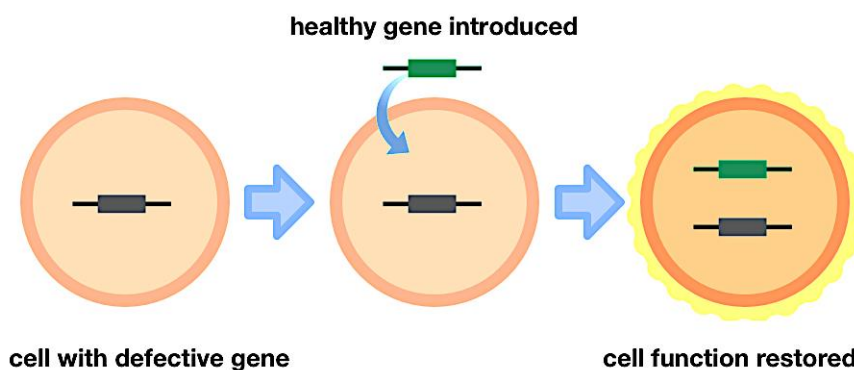
## 1. Introduction

### Prologue

The main introduction chapter to this study is divided into two parts: 1) Gene therapy, Insertional Mutagenesis (IM) and safety of viral vectors and 2) The liver, tissue of interest and Pluripotent Stem Cell Technology. The interplay of stem cells and gene therapy will be further explained which laid the core foundation for the proposed research study. At the end, the main objectives are elucidated.

### 1.1. Gene Therapy

Gene therapy is an advanced therapeutic approach in modern medicine. This field is still highly experimental but there is a great potential to become a treatment regimen when the pharmacological or surgical interventions are not effective. The main aim of gene therapy is to transfer the corrected genetic information into patient cells, tissues and organs. Following gene transfer, diseased genes can be eliminated, or normal function of the cells restored. In addition, new instructions can be added such as production of immune system mediator proteins that has a key role in defending cancer or other diseases (Figure 1.1) (Zwaka, 2006).



**Figure 1.1: Gene therapy treatment. The main aim of gene therapy is to correct defective genes and to restore cell functionality (the-gist.org).**

#### 1.1.1. History of Gene Therapy

Since 1950, performing site-specific alterations to the human genome has always been an objective in medicine. Francis Crick and James Watson successfully

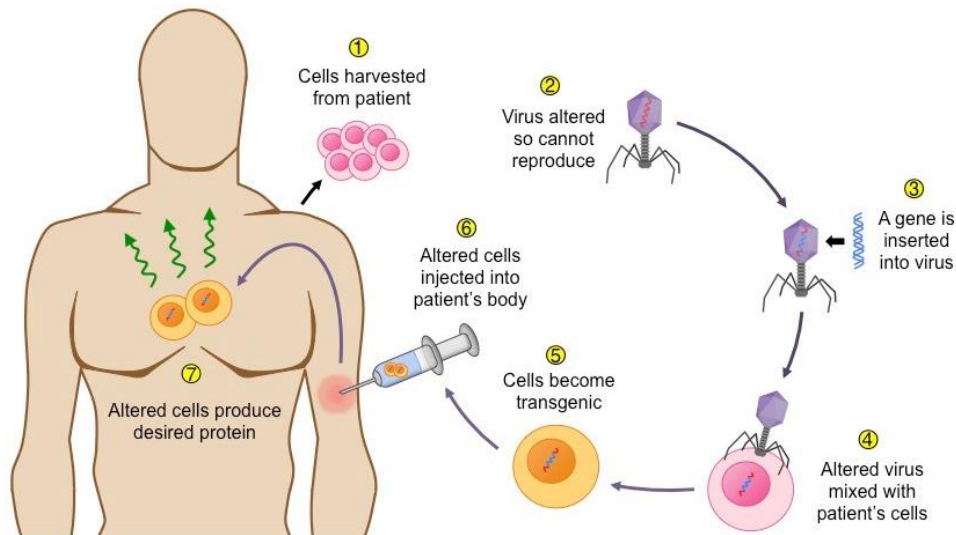
developed the revolutionary model of the double strand deoxyribonucleic acid (DNA) in 1970. Following this development, researchers identified a series of enzyme. These enzymes facilitated separation of the genes in preselected sites along the DNA molecule. These techniques paved the way for the introduction of genetic engineering technology and also production of new drugs and antibodies and gene therapy (Misra, 2013).

The genetic engineering technology dates back to 1980 and was used for the treatment of human diseases (Misra, 2013). Most clinical trials were performed in the United States, Europe, and Australia. The major focus of the gene therapy was the treatment of diseases caused by recessive gene disorders such as haemophilia, cystic fibrosis, and muscular dystrophy. The field has also progressed into more complex diseases such as cancer and certain viral infections, such as acquired immunodeficiency syndrome (AIDS) (Ginter, 2000, Misra, 2013).

One of the most often used techniques applied in this field was the recombinant DNA technology. In this technology, the gene of interest is inserted into a vector which can be either plasmidial or in viral form. Efficiency of viral vectors were much higher, therefore, it was used extensively (Misra, 2013).

### **1.1.2. Challenges of Gene Therapy**

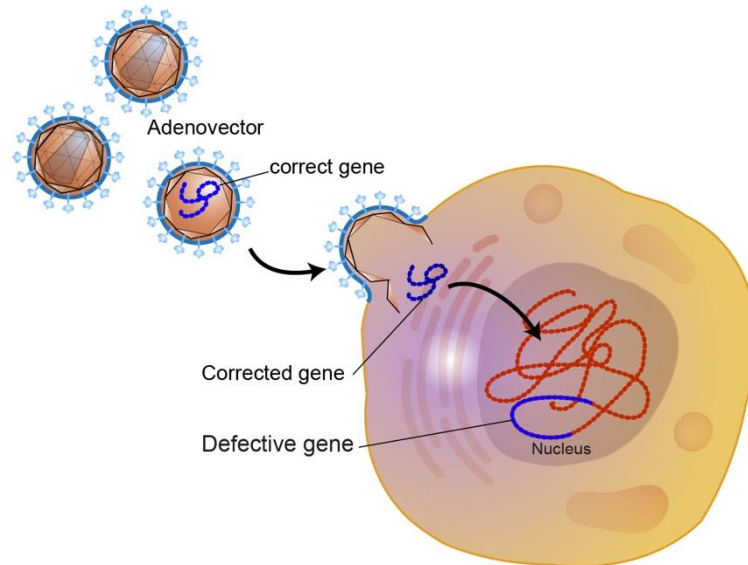
To date, several gene therapy protocols have been established. However, gene therapy is still facing technical difficulties and requires further research and developments. Some of these challenges are including identification of target cells, efficient gene transfer to the cells and understanding underlying disease mechanisms (Goncalves and Paiva, 2017). Gene therapy target cells are broadly divided into two main groups: gene therapy of the germline (Matthews and Curiel, 2007) and gene therapy of somatic cells (Bank, 1996). In germline gene therapy defective genes are altered. These modifications are integrated into the genome and will be transferred to next generations. This approach is more useful in the treatment of genetic and hereditary diseases. For somatic cell gene therapy, the patient's somatic cells are treated with therapeutic genes. Hence, treatment is more patient specific and will not be inherited by future generations (Figure 1.2) (Bank, 1996, Matthews and Curiel, 2007).



**Figure 1.2: Schematic representation of somatic cell gene therapy. In this method, patient’s somatic cells are harvested. Collected cells are transduced with corrected therapeutic genes. Under laboratory conditions, corrected cells are propagated and transferred back into the patient (ib.com.au).**

### **1.1.3. Gene Therapy vectors and their properties**

The aim of gene therapy is to treat certain diseases by substituting defective gene with the healthy gene in the genome. However, the effective gene transfer into a cell remains challenging. Therefore, a molecular carrier is required which is called a “vector”. Vector is applied for gene transfer application and needs to be very specific. Other properties of gene therapy vectors are including, an ability to transfer one or more genes of the sizes suitable for clinical applications, an efficient transfer to the cells while bypassing the immune system and large scale up process. Following treatment, a good gene therapy vector should not induce any allergic or other inflammatory reactions; it is expected to increase normal functional activity of the cells and correct defective genes; and more importantly it should be safe for the patient and also for the professionals who modify the changes prior to application (Misra, 2013, Goncalves and Paiva, 2017). A schematic representative of a gene therapy adenoviral vector has shown in (Figure 1.3).



**Figure 1.3: Gene therapy vector. Vectors are used for gene transfer application. A good gene therapy vector should not trigger allergic or other inflammatory reactions. The role of vector is to increase the normal functional activity of the cells and correct defective genes. In this figure an adenoviral vector is shown which is used to replace defective gene inside the cell (genome.gov).**

Several studies have reported efficacy of viral vectors and their ability for gene transfer. However, only few studies demonstrated existing technical challenges presented with these carriers. The remaining viral genetic material in the vector may pose a risk and can induce immune responses. In addition, there is a possible risk of oncogenic activation and cellular transformation which will be explained. (Goncalves and Paiva, 2017). There are two main techniques to deliver gene into the cells which are mediated by virus or physical mechanisms. The latter is made by advanced nanotechnology techniques (Linden, 2013). These techniques employ polymers as a base to form networks that encapsulate a gene. Upon cellular penetration, encapsulated gene is released. Some of these techniques are DNA microinjections (McDonnell and Askari, 1996), cationic liposomes (Caplen et al., 1995), cationic polymers (Plank et al., 1999) and particle bombardment (Yang et al., 1994). Each gene transfer technique is different with varying degree of efficiency and can be chosen based on the required application. It is crucial to develop new viral vector systems with optimum efficiency rate and high specificity for target cells. Improvement of vector safety and managing the inflammatory response will optimise the clinical translational work more effectively (Goncalves and Paiva, 2017).

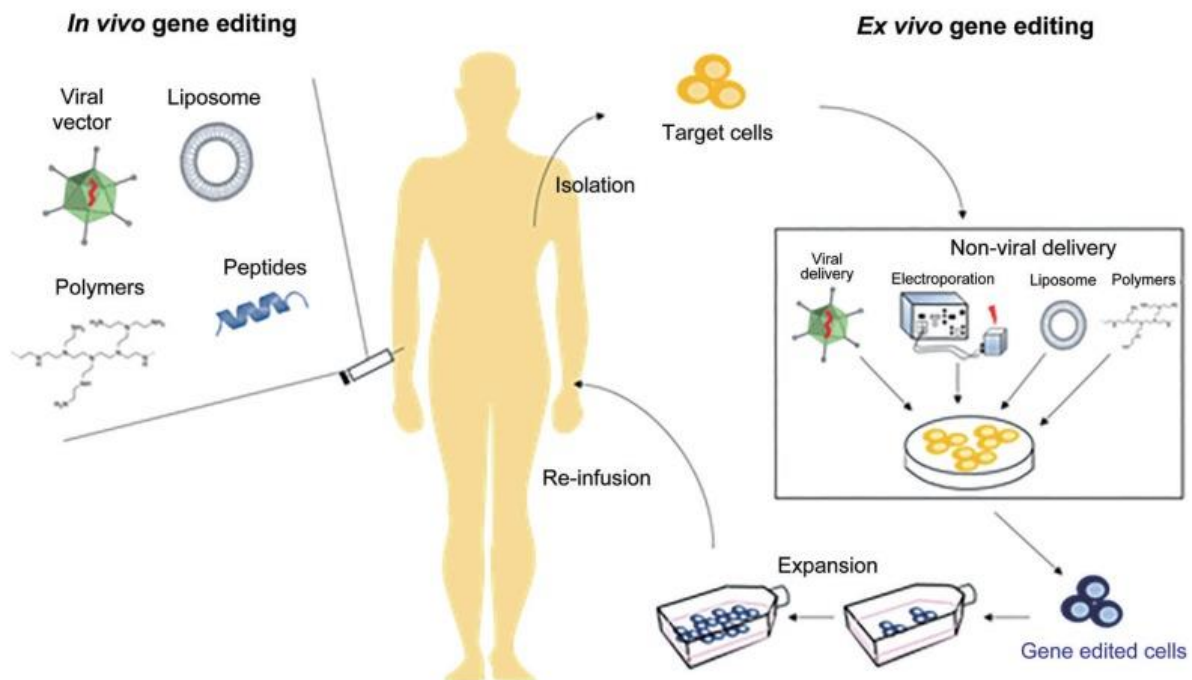
#### **1.1.4. Gene Therapy delivery routes (*in vivo* & *ex vivo*)**

The success of gene therapy technology relies upon two important and challenging factors. Firstly, the successful delivery of the therapeutic transgene into the desired human target cells. Secondly, the ability of the gene to function appropriately in the target cell. There are two ways to transfer genes into the patient which are including direct transfer or using living cells as vehicles to transfer the desired genes. Both routes have certain advantages and disadvantages (Zwaka, 2006).

Direct gene transfer (*in-vivo* gene editing) is a simple process in which genes are encapsulated into liposomes or other form of microparticles. These encapsulated genes are directly transferred into a patient's tissues or bloodstream. Alternatively, genes can be transferred through genetically-engineered viruses such as retroviruses or adenoviruses (ADVs). These viruses are genetically modified to ensure they are not toxic or infectious and also, that they are replication incompetent. These basic changes have been extensively studied over the past 10 years. Although direct gene transfer is simple, there are some disadvantages associated with. In direct gene transfer, there is not a sophisticated control over the transfer of therapeutic gene. This is occurred due to either random integration into the patient's genome or occasionally remains unintegrated for a short period of time in the target tissue. Furthermore, the target or tissue in many cases does not readily accessible for direct gene transfer application (Zwaka, 2006).

The other route of gene transfer is by applying living cells as vehicles to direct therapeutic genes (*ex-vivo* gene editing). The process itself is relatively complex compared to other route and it is divided into three major steps. In the first step, cells from patient are harvested and propagated under laboratory condition. Second, the desired therapeutic gene is introduced into these cells. Finally, the transformed cells are transferred back to the patient. There are certain advantages associated with this method. Upon cell isolation and propagation in the laboratory, there is much more control over changes on cells. Some of the cell types can be expanded easily under laboratory conditions prior to transplantation. Further, some cell types are able to localise to specific site of the human body, such as hematopoietic stem cells which will be transferred to the bone marrow. This is particularly important if application of therapeutic gene to specific region is desired. Despite these key advantages, living cell can often be problematic. Successful isolation of specific cell type requires

extensive knowledge of biological markers and also understanding of how the cell can survive in culture and proliferate. Unfortunately, not all the markers are known for different cell types and many primary cells will not stay stable for long period of time *in vitro* without acquiring certain changes or mutations (Figure 1.4) (Zwaka, 2006).



**Figure 1.4: Gene therapy delivery routes (in vivo & ex vivo).** Direct gene transfer or *in vivo* gene editing is a simple process in which genes are encapsulated into liposomes or other forms of microparticles. Alternatively, genes can be transferred through genetically engineered viruses such as retroviruses. *Ex-vivo* gene editing is by applying living cells as vehicles to direct therapeutic gene. Cells from patient are harvested and propagated under laboratory conditions and transduced with therapeutic gene. Following modification, cells are transferred back into the patient (Shim et al., 2017).

## 1.1.5. Stem Cells as “Carriers” for Gene Therapy

### 1.1.5.1. Adult Stem Cells (ASC) in Gene Therapy

Stem cells are broadly classified into two major groups: embryonic and adult stem cells. The properties of these cells will be explained in greater details in Section 1.3. ASCs maintain adult tissue with high turnover rate such as blood, skin and intestinal epithelium. Hematopoietic stem cells (HSC) is a great example of ASCs and their use demonstrated in animal and human gene therapy applications. Although these cells do not present in large amount, they can be easily harvested from bone marrow.

Further, specific cell markers presented on the surface of these cells enable scientists to identify and enrich HSCs among a mixed population of peripheral blood cells. Following propagation and *in vitro* modification, cells are transplanted back into the same patient (autologous) or a different patient (allogenic). These cells have the ability to contribute to all types of mature blood cells once transplanted (Aiuti et al., 2002, Hacein-Bey-Abina et al., 2002). Mesenchymal stem cells (MSCs) are also used for gene transfer application in gene therapy. They have the ability to form cartilage, bone, adipose tissue (Gregory et al., 2005). For therapeutic gene transfer into hematopoietic stem cells from bone marrow a viral vector is used. It is derived from a class of viruses called retrovirus. Retroviral vectors are able to transfer therapeutic genes into actively dividing cells. Since the rate of proliferation in ASCs are low, efficiency of viral gene transfer was quite low. Efforts were made to use other types of retroviruses such as lentiviruses and ADVs to overcome low transduction efficiency issue. Therefore, non-dividing as well as dividing cells can be targeted more effectively (Figure 1.5) (Zwaka, 2006).

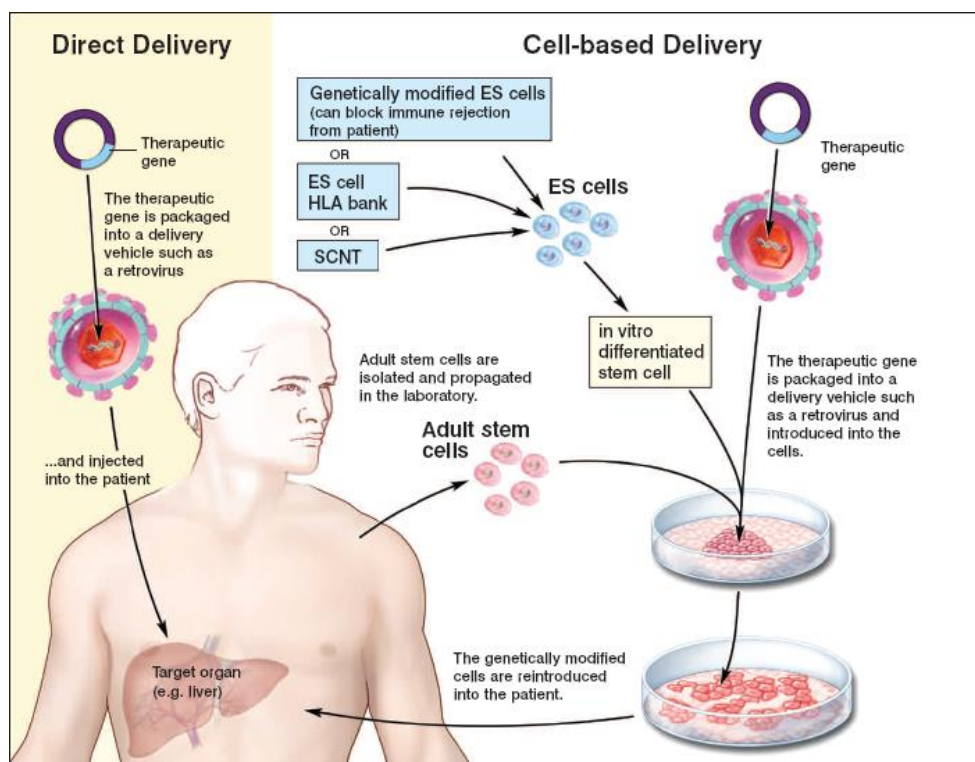
A major disadvantage of using stem viral vectors is that therapeutic gene can integrate randomly or semi-randomly into the chromosomes of the target cell. In gene therapy applications, this is not desirable as the carrier can potentially alter the activity of the neighbouring genes in close proximity to the insertion site. Further, they can inactivate host genes by full integration. These phenomena are known as IM. One of the major drawbacks of using ASCs in gene therapy is that it is difficult for them to remain as pluripotent (stem cell state) during *ex vivo* modifications. Under laboratory conditions, these cells tend to lose their stem cell properties and undergo a process called differentiation giving rise to mature cell types. However, significant advances in supportive culture conditions enable effective use of human HSCs for gene therapy applications (Zwaka, 2006).

#### **1.1.5.2. Embryonic Stem Cell (ESC) in Gene Therapy**

ESCs have unlimited self-renewal properties while maintaining the ability to differentiate into derivatives of three embryonic germ layers. The hESCs can be expanded indefinitely due to high expression of telomerase enzyme that prevents telomere shortening after cell division and senescence (de Wert and Mummery, 2003). Recent advances in culturing ESCs had a great impact on three important sections which are including, differentiation into various cell types such as hepatic, neurons,

vascular, and cardiac, etc; development of different cell lines and establishment of protocols which enable researchers to genetically alter these cells for gene therapy applications.

The potential of using human ESCs for gene therapy applications are extensive. These cell lines can be easily cultured and propagated. Following *in vitro* expansion, it can be applied for controlled and specific genetic modifications. Also, other properties such as high proliferation rate, stability and differentiation made them an ideal tool for *in vitro* gene therapy applications. These cells provide a constant *in vitro* cellular material and can be employed for optimising gene therapy protocols and genetic modification techniques. For research studies number of clinical samples is limited and therefore stem cells can be a great alternative (Evans and Kaufman, 1981, Martin, 1981).



**Figure 1.5: Different approaches for delivering therapeutic transgenes into patients. Direct delivery route which involves transfer of therapeutic gene via genetically engineered viruses such as a retrovirus. Cell-based delivery route is conducted in two different ways. The first route is to differentiate embryonic stem cells or genetically modified stem cells. Differentiated cells are transduced with therapeutic gene and transferred back into the patient. Alternatively, ASCs are isolated and propagated under laboratory conditions. Cells are transduced with therapeutic gene, using a viral vector**



**and genetically modified cells will be transferred back into the patient (Collins and Thrasher, 2015).**

### **1.1.6. Challenges for Stem-Cell based gene therapy**

Despite significant advances in using stem cells for gene therapy applications, few technical challenges still remain. Extensive modifications of stem cells require longer *in vitro* culture time. Therefore, it is argued that human ESCs may acquire several genetic and epigenetic changes which ultimately might harm the patient on long term. Epigenetic changes do not affect the genetic blueprint of the cells, while regulate gene activity. Also, few studies reported regular occurrence of sporadic chromosomal abnormalities when human ESCs are propagated as bulk populations. The result of these studies emphasises the importance of regular cell line monitoring and optimising culture conditions. Furthermore, undifferentiated embryonic stem cells are capable of forming a type of cancer called a teratocarcinoma. Recent protocols established and safety measures have been made to eliminate any remaining undifferentiated stem cells. This can be achieved by extensive purification techniques or by incorporating suicide genes that can be externally controlled (Draper et al., 2004, Gerecht-Nir and Itskovitz-Eldor, 2004).

Another important complication is the patient's immune system response. It is shown that transgenic genes as well as gene therapy carriers can potentially trigger immune system responses. If the stem cells are not autologous, immune rejection can be an undesired side effect of the transplanted cell type. Efforts to eliminate the risk include expression of immune system modulating genes by stem cells, immunotolerable bone marrow and inhibition of human leukocyte antigen (HLA) genes (Gerecht-Nir and Itskovitz-Eldor, 2004). Based on these strategies, stem cell-based gene therapies have potential novel therapeutic applications. There are other cell-based therapies which utilise gene-based modifications. Chimeric antigen recipient T (CAR-T) cell therapy and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and associated proteins (CAS) as gene editing tool. These techniques are briefly explained in the next section.

### **1.1.7. Chimeric antigen recipient T (CAR-T) cell therapy**

CAR-T cell therapy is a type of immunotherapy that involves alteration and/or reprogramming of immune cells particularly T lymphocytes of the patients. The

modified T cells will recognize and combat tumour T cells. Initially, Eshhar and colleagues developed the first generation of CAR. They fused a single chain variable (scFv) to a transmembrane domain and an intracellular signalling unit which is called chain CD3 zeta (Gross et al., 1989). Using this design, the recognition of the tumour-specific epitope was increased, and the T cells were activated. Although the third generation of CAR have been developed, there are still technical challenges. One of the main disadvantages of CAR-T therapy is the identification of non-tumour cells that express similar epitope by CAR. Another disadvantage of this technique is the cytokine release syndrome. Following CAR-T infusion, immune system is activated, and levels of inflammatory cytokines are elevated. An important consideration in CAR-T cell therapy is the design of vectors and with necessary experimentation the safety for clinical applications can be enhanced (Goncalves and Paiva, 2017).

#### **1.1.8. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas) technique**

Since 2012, this technique is used as one the important biotechnological tool for gene editing and is originated from a region in *Escherichia coli* genome. This technique enables scientists to edit the target specific DNA sequences of the genome, using three molecules: nuclease (Cas9), which has a role in double-strand DNA breakage; an RNA guide which guides the whole mechanism to the target; and the target DNA. This technique is less complex and more efficient compared to older methods such as Zinc-Finger Nucleases and Transcription Activator-like Effector Nuclease (TALEN). This system enables scientists to inactivate genes (Knock-out) or integrate genes (Knock-in) or edit genes (Marraffini and Sontheimer, 2010). Significant advances have been made to optimise delivery systems particularly the safety of the system for translational trials.

#### **1.1.9. Different vector delivery systems for gene therapy**

For successful generation of protein products of the introduced gene exogenous genetic material must be delivered to the cell's nucleus. This process can be done by two classes of vectors: non-viral and viral delivery which will be explained in greater details. Gene delivery via the viral vectors is called transduction whilst transfer via the non-viral vectors is called transfection (Mali, 2013).

#### **1.1.9.1. Non-viral delivery**

As mentioned before an ideal vector should have the following key properties: Firstly, to deliver transgene to a specific cell type. Secondly, an ability to accommodate suitable foreign gene sizes. Thirdly, stable transgene expression and correction of defected gene and finally to be non-immunogenic and safe. The process of delivering transgene by viral vector is called transduction while delivery by non-viral vectors called transfection.

Transfection can be performed by different methods such as chemical transfection, lipid vectors, physical transfection and electroporation therapy. For chemical transfection, calcium phosphate, lipid or other protein complexes are used to introduce DNA to the cells. Lipid vectors are liposomes which are produced using a combination of plasmid DNA and a lipid solution. This complex interacts with the cell membranes of a variety of cell types resulting in delivering the plasmid DNA into the cytoplasm and the nucleus for transient expression. Physical transfection can be performed by using electroporation, microinjection and use of ballistic particles. Electroporation therapy temporarily increases cell membrane permeability to large molecules allowing their entry to the cells.

Non-viral vectors are divided into two groups which are naked-DNA and liposomes. The basis of non-viral vectors is plasmid which is a circular DNA strand. In this method, therapeutic genes can be introduced into the plasmid. Following formation of recombinant plasmid, it can be introduced into cells via different routes. For instance, it can be injected directly into targeted tissues or organ. The advantages of using non-viral vectors are the low cost and ability for large-scale production. Further, these vectors have low risk of immunogenicity and this helps for potential re-dosing (Mali, 2013).

#### **1.1.9.2. Viral vector delivery**

The spectrum of viral vectors for gene therapy applications is quite broad and they are applied for transient and permanent gene expression. The viral vectors are classified into two types: Integrating and non-integrating viral vectors. Integrating viral vectors such as retroviral, lentiviral and adeno-associated viral vectors, can integrate into the human genome. Non-integrating vectors such as adenoviral vector can remain in the nucleus without integrating into the chromosomal DNA and expression of the foreign

gene is transient. Also, viral vectors can be classified based on RNA and DNA viruses with either can have single stranded (ss) or double-stranded (ds) genomes (Biceroglu and Memis, 2005). The main group of viral vectors are ADVs, Herpes Simplex Virus Adeno-Associated viruses, Retroviruses and Lentiviruses. For the purpose of this thesis, the focus will be on ADV, adeno-associated virus and retroviruses.

### **1.1.9.3. Adenoviral vectors (ADV)**

The ADV was the first viral vector employed for gene therapy and was approved for clinical trial studies in 1990. This virus is originated from human adenoid tissue-derived cell cultures which occurred in 1953 and therefore, it is called ADV. This type of virus belongs to a diverse family of non-enveloped dsDNA viruses known as *Adenoviridae* (Rowe et al., 1953). This virus rarely causes major complications in healthy individuals. However, immunocompromised individuals can develop a wide range of illnesses such as cold, sore throat, pneumonia, bronchitis and other neurologic diseases. To date, 57 human AD serotypes have been identified and they are divided into seven categories based on their unique properties (Rauschhuber et al., 2012, Crystal, 2014). AV contains a linear dsDNA ranging from 26 to 45kb in a medium sized (~100nm) non-enveloped icosahedral viral particle. The structure of this viral particle composed of penton and hexon subunits. The major part of the viral capsid coat is formed by hexon units which carry antigenic motifs. The other penton subunits have a key role in synthesising fiber and knob domains required for infection (Nemerow et al., 2012). AD infection is initiated by fiber knob domain and binds to a variety of proteins such as major histocompatibility complex (MHC-1) alpha2 subunit, CD46 and AD receptor (CAR) expressed on cell surface (Shayakhmetov and Lieber, 2000). Process of endocytosis and entry of the virus is performed by the interaction between arginine-glycine-aspartic acid (RGD) sequence of the fiber penton subunit and integrins on the cell surface (Bai et al., 1993). This bond is crucial for AV broad tissue tropism and also a way for better transduction efficiency. This enables scientists to modify binding sites for CAR and other ligands to inhibit AV reinfection for gene therapy application. This is an important factor and by deletion of AV pathogenic genes this virus has been repurposed (Figure 1.6) (Goswami et al., 2019).

In terms of gene expression during infection and multiplication, AV genome is classified into early (E1, E2a, E2b, E3 and E4), intermediate (IVA2 and IX), and late genes (L1, L2, L3, L4 and L5). Further, AV genomes contains non-coding inverted

terminal repeat (ITR) sequences, packaging sequences and several viral RNAs (Nemerow et al., 2012).

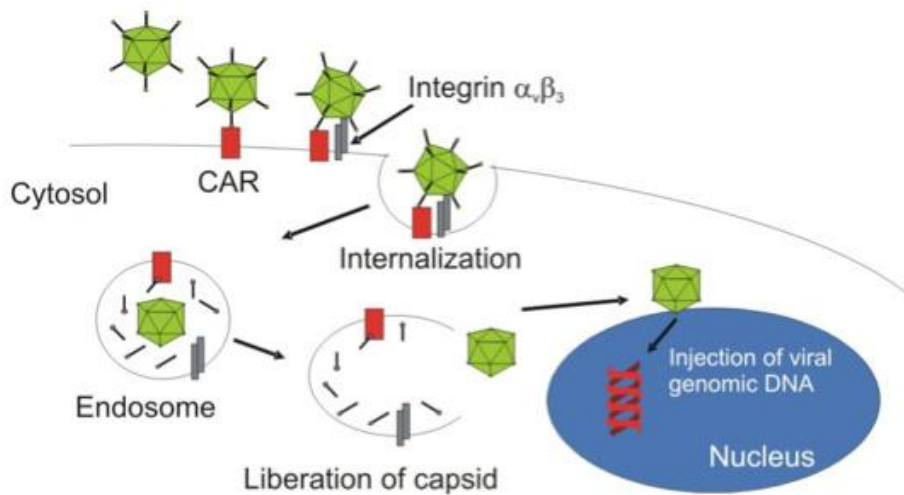
To develop safe and efficient gene therapy vectors AV genome has been modified several times. They are organised into three groups: first-generation vectors, second-generation and third generation vectors.

The first-generation vectors were made by eliminating E1 or E3 genes and therefore became replication incompetent. There are few disadvantages with first-generation vectors including small transgene, inefficient expression of viral proteins and contamination with replication-competent virus (Liu and Muruve, 2003).

Second generation vectors were generated by elimination of E2A, E2B and E4 from the genome of the first-generation vectors. However, due to the leaky expression of viral proteins and loss of transgene expression their application was not favoured (Amalfitano et al., 1998).

Third generation AV vectors were called gutless or helper-dependent AV vectors. This generation was lacking all the viral genes except for packaging and ITR sequences. They received considerable attention for gene therapy applications as they could offer capacity for larger therapeutic transgene (up to 37 Kb in size). They also offered long term transgene expression and less contamination with replication competent virus particles. They were also less immunogenic compared to previous predecessors. This recombinant AV vector were successfully used in for two years in animals with no undesired side effects (Sakhuja et al., 2003, Alba et al., 2005).

AV vectors enable episomal or permanent insertion of therapeutic genes and can integrate into cellular DNA effectively. This achievement was important as it allowed clinicians to target different diseases.

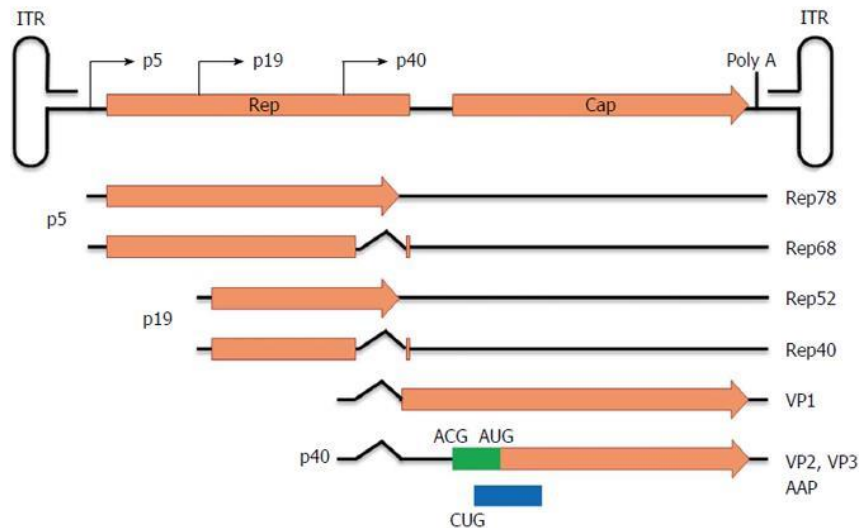


**Figure 1.6: Attachment and replication of ADV.** The ADV replicates in the nucleus of host cells, using the replication machinery of host cells. For virus entry, two types of interactions between the virus and host cell are required. Process of endocytosis and entry of the virus is performed by the interaction between arginine-glycine-aspartic acid (RGD) sequence of the fiber penton subunit and integrins on the cell surface. After the entry, the endosome acidifies and capsid is released. Following release, the viral DNA enters the nucleus of host cells via nuclear pore. The life cycle of ADV is classified into two phases: the early and the late phase. The early phase involves with the expression of non-structural, regulatory proteins. The late phase has a role in the expression of structural proteins and packaging of all the genetic materials generated by DNA replication (Diagnostic, Goswami et al., 2019).

#### 1.1.9.4. Adeno-associated virus (AAV)

The AAVs is used extensively in gene therapy. This microbe is originated as a contaminant in the simian AV preparation and later found in wide range of animal and human samples (Atchison et al., 1965). AAV carries a 4.7 kb long ssDNA genome and it is packed within a non-enveloped viral particle. Viral particle contains p5,p19 and p40 promoters in addition to *rep* and *cap* genes flanked by two 145 nucleotide-long ITR sequence and no polymerase gene (Laughlin et al., 1979). The palindromic sequence base pairs of ITRs synthesise cDNA. Also, *rep* and *cap* genes undergo splicing events to express replication proteins (Rep78, Rep68, Rep52, and Rep40), capsid or other virion proteins (VP1, VP2 and VP3) and an assembly activating proteins (AAP) (Grieger and Samulski, 2012). This protein helps virion proteins to generate capsid. Further, p40 promoter expresses VP1, VP2 and VP3 and at a ratio

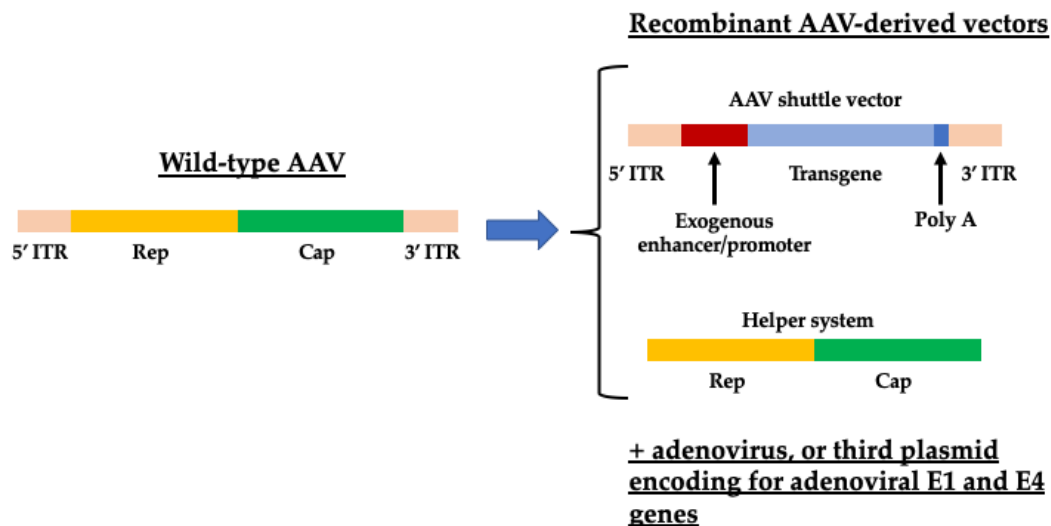
of 1:1:10 produce the external capsid of the virion. Capsid proteins hold phospholipase domain which has an important role in protecting virions from intracellular protease system attack (Figure 1.7). AAV virus for infection and completion of replication requires other helper proteins, viruses such as AV, herpes simplex virus type I and II, cytomegalovirus and UV radiation (Jay et al., 1981, Carter, 2004).



**Figure 1.7: Genomic structure of wild type adeno-associated virus. AAV viral particle contains p5, p19 and p40 promoters in addition to rep and cap genes flanked by two 145 nucleotide-long ITR sequence and no polymerase gene. Rep and cap genes undergo splicing events to express replication proteins (Rep78, Rep68, Rep52, and Rep40), capsid or other virion proteins (VP1, VP2 and VP3) and an assembly activating proteins (AAP). This protein helps virion proteins to generate capsid. Capsid proteins hold phospholipase domain which has an important role in protecting virions from intracellular protease system attack (Carter, 2004).**

Recombinant AAV-derived vectors can be made by providing essential genes such as E1a, E2a or E4 which are required for viral infection. To generate AAV vectors, these genes along with AAV expression plasmid (*rep-cap* plasmid) are applied to co-transfect HEK-293 cells (Grieger and Samulski, 2012). In the AAV expression plasmid therapeutic transgenes are cloned which carry ITR sequences. Their size capacity can be increased by using separate plasmid carrying *rep-cap* genes or by production of virus in *rep-cap* stable cell lines. Recombinant AAV provides long term transgene expression. Due to lack of *rep-cap* genes in recombinant AAV constructs, they can deliver therapeutic transgene in the episomal form without permanent integration into

the genome. Following delivery of transgene, it can remain quiescent in cells for long period of time without causing side adverse effects (Figure 1.8). Therefore, AAV based viral vectors are considered as potential candidates to treat a wide range of diseases (Goswami et al., 2019).

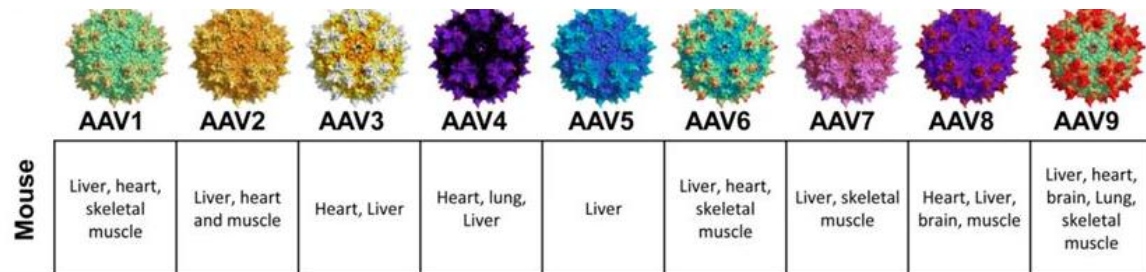


**Figure 1.8: Structure of recombinant adeno-associated derived vector. Recombinant adeno-associated derived vector is made by co-transfection of essential genes and AAV helper plasmid (rep-cap plasmid) in HEK-293 cells. In the AAV expression plasmid therapeutic transgenes are cloned which carry ITR sequences. Their size capacity can be increased by using separate plasmid carrying rep-cap genes or by production of virus in rep-cap stable cell lines. Recombinant AAV provides long term transgene expression (Romano, 2019).**

Recombinant AAV based vectors have broad tissue tropism. However, for efficient transduction and precise delivery of therapeutic transgenes they do require expression of heparin sulphate proteoglycan, alpha beta integrins, fibroblast growth factor receptor 1, platelet-derived growth factor receptor, hepatocyte growth factor receptor, epidermal growth factor receptor, laminin receptor on the surface of target cells. To date, 13 AAV serotypes have been isolated and each serotype can have better affinity toward a specific receptor (Figure 1.9) (Gao et al., 2004, Pillay et al., 2016). This unique property has made them an ideal candidate to target specific cell or tissue types. For instance, AAV serotype 1 (AAV1) holds higher transduction efficiency for muscles, neurons, heart and retinal pigment epithelium. AAV serotype 2 has demonstrated higher transduction efficiency for many types of cancer cells, kidney and



retinal pigment epithelium (Zincarelli et al., 2008). In fact, it is the only serotype that can transduce and deliver transgene therapeutic gene to kidney. AAV serotype 8 can transduce pancreas and it was widely used to deliver therapeutic *FIX* gene to the liver to treat haemophilia in clinical trials (Nathwani et al., 2014).



**Figure 1.9: Adeno-associated viral vector serotypes. To date, 13 AAV serotypes have been isolated and each serotype can have better affinity toward a specific receptor. This unique property has made them an ideal candidate to target specific cell or tissue types (Gao et al., 2004).**

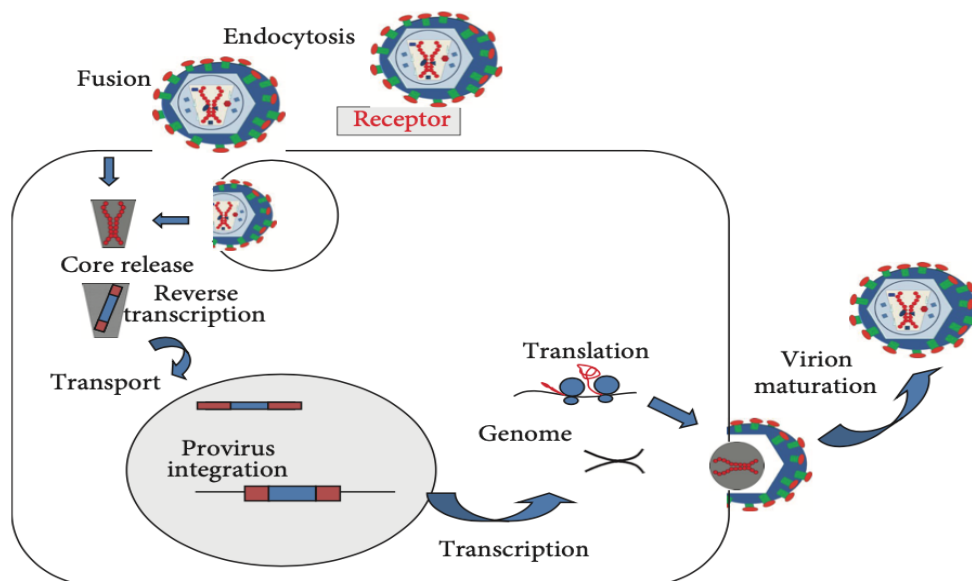
#### 1.1.9.5. Retrovirus

Retroviruses (RV) are spherical in shape and approximately 100 nm in diameter. They belong to *Retroviridae* family. This type of virus is classified into different viruses which are including: foamy virus, human immunodeficiency virus (HIV-1), simian immunodeficiency virus (SIV) and murine leukaemia virus (MLV). One of the main properties of these viruses are their ability to reverse transcribe their genetic blueprint of positive and ssDNA into dsDNA and to integrate into the host cell genome. The two important subtypes of RV are gammaretrovirus and lentivirus which are derived from MLV and HIV-1, respectively. RV life cycle is a multistep process which is including: binding, fusion, reverse transcription, integration, replication, assembly and budding (Maetzig et al., 2011b). The next section will focus on life cycle of gammaretrovirus and lentivirus in greater details.

#### 1.1.9.6. Gamma-retrovirus and lentivirus life cycle

Genome of gammaretrovirus is composed of three important genes, *gag*, *pol* and *env*, and is flanked on two sides by long terminal repeats (LTRs). The role of *gag* is to insert viral genome mRNA into virions, *pol* gene encodes the reverse transcriptase, integrase and protease and *env* gene encodes necessary surface proteins. Further, the cis-acting psi packaging element of the RV genome has a role in packaging the

RV mRNA into viral capsid during replication stage. The presence of envelop surface proteins enable them to bind to several receptor molecules which is essential for infection initiation and binding. Following the initial step and conformational change proteins virus will enter through fusion or endocytosis. In the next stage endosome releases its RNA and process of reverse transcription will occur. The resulted viral DNA is integrated into the host cell genome. DNA is transcribed into RNA using 5' Cap and 3' poly (A) tail and translated into necessary viral proteins for assembly and budding from cell membrane. This completes retroviral life cycle in addition to extracellular maturation (Figure 1.10). The newly formed mature RV can transduce a range of somatic cells such as embryonic stem cells, hematopoietic and neural stem cells (Maetzig et al., 2011b). The only downside for using gammaretroviral based vectors are their broad species specificity. This may lead to transduction of undesired cell types, incomplete reverse transcription and possible risk of Insertional Mutagenesis. Compared to gammaretrovirus, lentivirus can transduce post-mitotic cells (Schambach et al., 2009).



**Figure 1.10: The general life cycle of retrovirus/lentivirus. The schematic representative of the lentivirus life cycle is shown in this figure. It is a multistep process which involves binding, fusion, reverse transcription, integration, replication and assembly. In the first step envelope surface proteins bind to several receptor molecules. Following entry of the virus, retrovirus core is released followed by reverse transcription. The core is composed of the cDNA virus genome and integrates into the host cell chromosome in the nucleus. The cDNA is transcribed into RNA using 5' Cap**

and 3' poly (A) tail and translated into necessary viral proteins for assembly and budding from cell membrane (Dufait et al., 2012).

#### **1.1.10. Generation of lentivirus-based vector**

For generation of lentivirus-based vector, four plasmids are required. The *gag* and *pol* plasmid, the *rev* plasmid which helps to transfer mRNA into the cytoplasm, the heterologous envelope protein such as vesicular stomatitis virus G glycoprotein (VSV-G) for membrane fusion and virus entry and the transgene of interest. Co-expression of all these plasmids in HEK293T packaging cell lines are used for transient and stable transfections. As a result, lentiviral vector particles are generated which carry no replication-competent virions and are employed for various research and therapeutic applications. Methods have been established to utilise different packaging cell lines to generate clinical grade RV vector particles at a concentration of  $10^6$  to  $10^7$ /ml (Schambach et al., 2009).

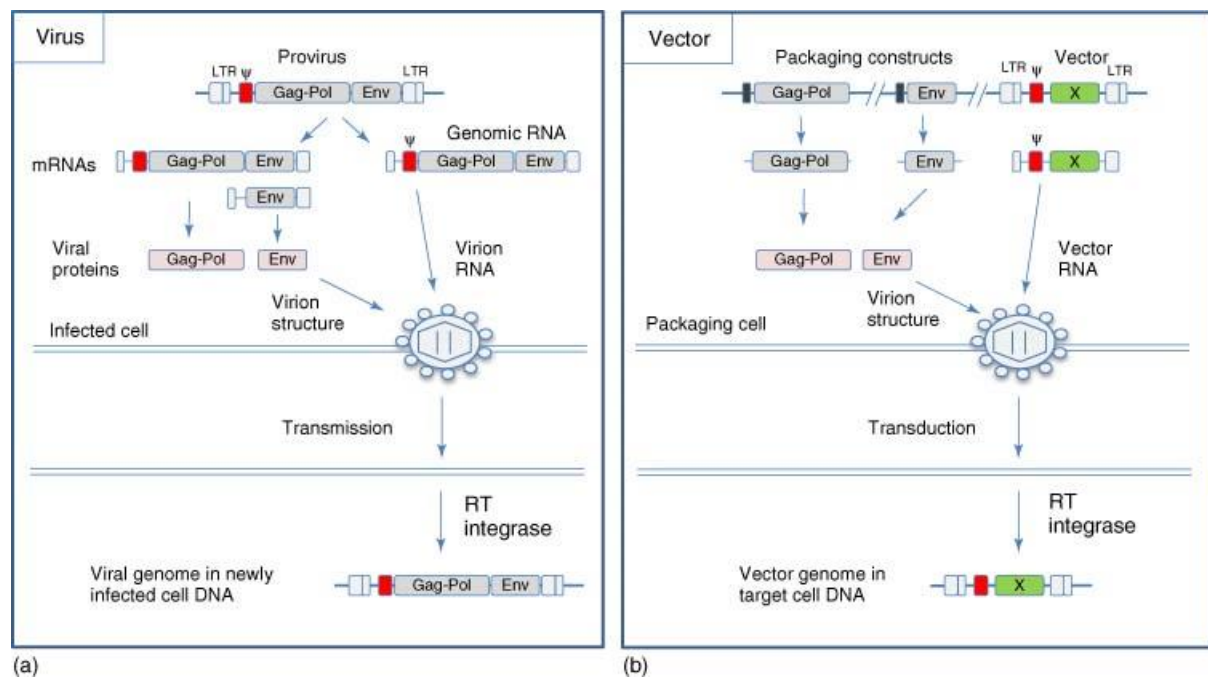
#### **1.1.11. Generation of replication- incompetent viral vectors**

To generate replication incompetent vectors, there is no psi packaging and RNA dimerization element in *gag/pol* and *env* expression constructs. The presence of viral structural proteins enables them to detect the psi-containing vector construct only and allows selective packaging of vector genomes into newly formed particles. Following entry of the vector particle into the host cell, only RNA of the vector construct is reverse transcribed. This will result in stable genomic integration into the host genome. This method helps to eliminate the risk of producing replication competent vector particles during clinical viral particle production. Lentiviral-based vector constructs have been employed to treat cancer, inherited and monogenic disorders (Figure 1.11).

#### **1.1.12. Advantages and disadvantages of retroviral based vectors**

Retroviruses can carry larger size transgene (9-12 Kb) in size. They can also be produced at high titers using established packaging cell lines. However, cellular non-specificity and possible risk of IM can be their potential adverse effects (Ibraheem et al., 2014). The integrase enzyme facilitates insertion of copies of retroviral genome into the host cell. Occasionally the genome copy might be inserted into undesired locations such as tumour suppressor gene or an oncogene. This will lead to uncontrolled cell division and possible cellular transformation (Misra, 2013). Thus, it is

important to evaluate the risk of IM for each retroviral vector construct prior to clinical applications.



**Figure 1.11. The wild type lentivirus genome organisation and generation of lentivirus derived vector. A) Genome structure of wild type HIV-1 is shown in this figure. It is composed of long terminal repeats (LTR), packaging signal (PS), GAG, POL genes, envelope gene (ENV) and the accessory genes such as Vif, Vpr, Rev, Tat. B) The design of the lentiviral gene transfer system. To make lentivirus derived vectors, genome is divided into three different plasmid constructs. The vector itself contains at least, the LTR, PS and the internal promoter modulating the transcription of the gene of interest. The packaging plasmid promotes the expression of the GAG-POL genes. Finally, the envelope plasmid expresses envelope glycoproteins which helps with the tropism aspect of the vector. The co-transfection of these three plasmids promote the production of lentiviral vectors with the ability to transduce cells (Dufait et al., 2012).**

The best preferential location for gammaretroviral vectors is integration near gene regulatory regions which is high risk (Hacein-Bey-Abina et al., 2008). In comparison, lentiviral vectors prefer the body of genes for integration and therefore as a result they have much lower risk of genotoxicity (Kaufmann et al., 2013). Efforts have been made to eliminate risk of IM with the development of self-inactivating retroviral vectors (SIN vectors). These vectors are transcriptionally inactive. Further research is required to

study vector mediated genotoxicity, development of new and specific vectors and evaluating the frequency of insertional mutagenesis. Nonetheless, many clinical trials have employed retrovirus as the vector of choice for gene therapy applications (Knopp et al., 2018). In the next section, mechanisms of IM and important human gene therapy clinical trials will be covered in more details. During some of these trials adverse events were observed which led to insertional activation of cancer related genes such as proto-oncogenes. Following these events, numerous studies were established to evaluate the molecular basis of vector related genotoxicity which can help to develop safer vectors with lower risk of genotoxicity.

### **1.1.13. Mechanisms of Insertional Mutagenesis**

A well-established safety concern of viral vector-based gene therapy is IM. The mechanism of IM is distinctly complex. Several genes can be initiated by various mechanisms. During this process, although the insertion into the genome can be biologically mute, it can significantly interfere with the host cell transcription or post-transcriptional activity of neighbouring genes. This might be biologically or clinically irrelevant. However, with additional mutagenic events this can lead to the process of cellular transformation (Aiuti et al., 2013). Activation of the proximal proto-oncogenes by the dominant gain-of-function mutations are the biggest concern. These mutations are generally occurred through promoter insertion, gene transcript truncation, promoter activation and epigenic gene silencing (Touw and Erkeland, 2007, Aiuti et al., 2013, Suerth et al., 2014) (Figure 1.12).

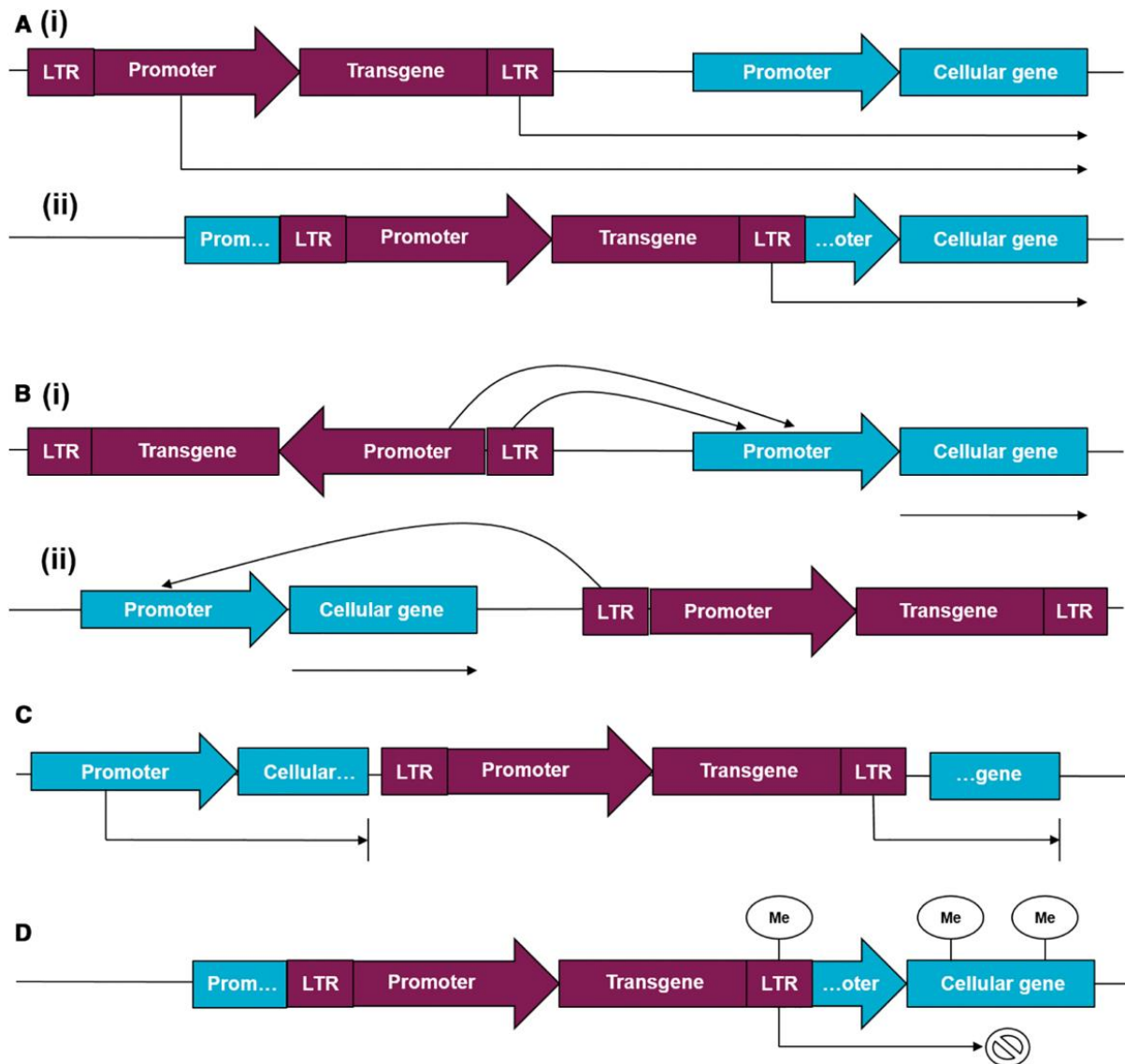
The first mechanism is viral promoter insertion which can occur in two different ways. A viral promoter can be inserted upstream of cellular transcription unit (Figure 1.12 Ai) (Suerth et al., 2014) or into the promoter 5' of a target gene (Figure 1.12 Aii) (Touw and Erkeland, 2007). This may result in read-through transcription in proximity to neighbouring genes and activation of gene via LTR promoter sequences, respectively. It can also induce overexpression of aberrant splice variants.

The second mechanism is promoter activation by viral vector enhancer or promoter sequences (Figure 1.12 Bi). If the viral vector integrates close to a target gene, level of gene expression may be increased via endogenous promoter activation. Further, if the vector integrates in the 3' untranslated region (UTR) this may enhance mRNA stability (Figure 1.12 Bii).

The third mechanism is gene transcript truncation. In this mechanism if a viral vector integrates within a gene, this can lead to disruption of mRNA and/or overexpression of truncated mRNA. This ultimately can generate transcripts lacking 3' or 5' sequences which may also inactivate a gene (Figure 1.12C).

The final mechanism is gene silencing. As a result of epigenetic changes, gene silencing event may occur. It is reported that virus-induced gene methylation can cause inhibition of the expression of flanking genes. This is occurred due to histone and DNA methylation (Figure 1.12D) (Touw and Erkeland, 2007).

The most prominent mechanism for proto-oncogene activation is viral vector enhancer-mediated promoter activation. Proto-oncogene activation is resulted from interaction of viral enhancer sequences with cellular promoters (Kool and Berns, 2009, Hachein-Bey-Abina et al., 2010, Cesana et al., 2012). Gene transcript truncation is commonly detected when viral vectors with active LTRs are used. In this mechanism vector-mediated transcripts can generate truncated proteins. This may lead to loss of function mutations in tumour suppressor genes (Cesana et al., 2014). There is strong evidence that elimination of regulatory regions of proto-oncogenes and leukaemia in mice was triggered by lentiviral mediated gene transcript truncations of 5' or 3' sequences (Montini et al., 2009).



**Figure 1.12: Different mechanisms of viral vector-mediated genotoxicity. A) Promoter insertion (i) either by upstream of transcription units or (ii) in the promoter region of 5' of the target gene. B) Promoter activation either by (i) virus integration in a close proximity of a target gene as a result of enhancer sequences in the viral LTR or (ii) virus integration in the 3' UTR. (C) Integration of virus may generate truncated proteins which subsequently lead to disruption of mRNA and gene inactivation. (D) Certain epigenetic changes such as virus-induced gene methylation can prohibit the expression of flanking genes (David and Doherty, 2017).**

### 1.1.14. Other influencers of viral vector-mediated genotoxicity

There are also other influencers of vector-mediated genotoxicity which can be classified into two main groups of viral and non-viral factors. Viral factors are divided

into two subgroups which are including vector design and the integration site profile of the vector. It has been reported that vector construct modulates the genotoxic potential of retroviruses (Montini et al., 2009). A choice of transcriptionally active LTR containing strong enhancer and promoter sequences vector construct has been known as the major determinant of genotoxicity. It is reported that integration of such vector into the genome can trans-activate neighbouring genes (Montini et al., 2009). Furthermore, a selection of strong enhancers or promoters can play a role in genotoxicity/carcinogenicity risk. It was shown that following treatment with AAV vectors containing thyroxine-binding globulin (TBG) and chicken beta-actin enhancer strong promoters high level of hepatocellular carcinoma was observed. However, it was not detected in weaker promoters such as human alpha-1 antitrypsin (hAAT)(Chandler et al., 2015). Hence, choice of promoters can influence AAV genotoxicity.

Integration profile of the viral vector can be regarded as another influencer of genotoxicity. It is argued that RV integration is not a random process and each viral integrase segment has a weak tendency toward a short genomic sequence (Niederer and Bangham, 2014). For instance, gamma-retroviruses have a selective preference for transcriptional start sites (TSS), gene regulatory elements such as promoters/enhancers, DNAase-I hypersensitive sites (HSS) and CpG island (David and Doherty, 2017). It has also been reported that gamma-retroviruses have an integration bias for specific “hot spots”. Hot spots are specific genes which are related with cell growth and cancer. This may further increase risk of IM (Montini et al., 2009). Further, HIV-1 based lentiviral vector has a strong bias for transcriptional units in actively transcribed genes (Niederer and Bangham, 2014, Suerth et al., 2014). Therefore, selective integration of viral vectors within the genome can influence genotoxicity and ultimately IM.

Non-viral factors can be divided into transgene, target cell, epigenetics, age of recipient and disease state prior to gene therapy treatment. In terms of transgene, it is extremely important to consider the cellular role of the transgene product and its effect on genotoxicity. For instance, MLV-based gamma-retroviral vectors were used in both X-linked severe combined immunodeficiency (SCID-X1) and adenosine deaminase-SCID (ADA-SCID) human gene therapy trials. For these trials, Interlukin-2 receptor subunit transgene (IL2RG) and ADA transgenes were used. Interestingly,



LMO2 integrations were detected at a similar frequency in both trials. However, leukaemia was only reported in the SCID-X1 trial. These findings suggest that integration is not the only reason for cellular transformation and the vectors containing the transgene may play a role. ADA is an enzyme which involves with survival whilst IL2RG is a potentially oncogenic growth factor receptor (Aiuti et al., 2007, Niederer and Bangham, 2014). Details of each of these gene therapy trials will be discussed in Section 1.1.16.

It is known that integration of viral vectors can also be cell specific (Aiuti et al., 2013). For example, MLV-based gamma-retroviral vectors can transduce hematopoietic stem cells and mature peripheral blood lymphocytes in two different ways. As a result, different integration patterns can be detected such as differential gene expression profile. Also, it can affect the overall accessibility of the genome during transduction (Biasco et al., 2011). Cell or tissue maturity can influence risk of oncogenesis. For instance, in vivo foetal and neonatal mice exhibited more sensitivity to the effects of integrating vectors. This resulted in a higher risk of oncogenesis compared to older mice in the study (Nowrouzi et al., 2013, Themis et al., 2005). Other factors such as epigenetic, age of recipient and disease background may also influence genotoxicity (David and Doherty, 2017). In the next sections, genotoxicity prevention strategies and key human gene therapy clinical trials will be covered.

### **1.1.15. Genotoxicity Prevention Strategies**

Over the years new strategies have been devised to prevent or reduce risk of vector-mediated genotoxicity. These strategies were focused on vector design such as development of Self-Inactivating (SIN) vectors and chromatin insulators. Other strategies such as incorporation of polyadenylation signals, tissue-specific promoters, titration of vector dose can improve gene therapy safety while reducing risk of genotoxicity (David and Doherty, 2017). For the purpose of this thesis, SIN vectors will be discussed in greater details.

SIN vectors including lentivirus and gamma-retroviruses have been designed to restrict the effects of the viral LTR promoter on flanking genes. This strategy reduces genotoxicity mediated risks commonly seen with integrating retroviral vectors. To construct SIN vectors, U3 region in the plasmid is deleted. Hence, SIN vectors are lacking viral promoter and enhancer activity in their 3'LTR region. It has been reported

that SIN lentiviral vectors have a lower risk of IM compared to gamma-retroviruses. In SIN lentiviral vectors polyadenylation signals and splice sites are still present. However, it has been shown that generation of splice-variant transcripts of genes surrounding LV integration sites were reduced. By employing this strategy genotoxicity risk has been reduced. However, other problems such as promoter activation and gene transcript truncation are remaining. For therapeutic gene expression, internal enhancers and promoters have been incorporated in SIN lentiviral vectors. This may ultimately affect non-target genes. Those vectors which utilise strong enhancers or promoters can modify cells by insertional activation, though their transforming potential was shown to be much lower compared to full LTR vectors. It is argued that there is a strong link with formation of tumour and promoter strength and is directly proportional. Therefore, it is advised to employ weaker internal promoters with moderate activity to reduce the risk of IM. However, it has been shown that even reduction of enhancer activity by moderate promoters is not solely beneficial and tumour suppressor gene inactivation has been detected. Few studies reported that using weak internal promoters can form aberrant or truncated transcripts. It is argued that the complex interaction between the location, distance, strength of promoters and mRNA polyadenylation signals can produce unique aberrant mRNAs. Although efforts have been made to reduce risk of genotoxicity there are still concerns regarding SIN viral vectors (David and Doherty, 2017).

#### **1.1.16. Insights from key human gene therapy clinical trials**

Over the past two decades, haematopoietic stem cells (HSC) have been used for gene therapy based clinical trials. They were employed for the treatment of severe inherited diseases such as X-linked severe combined immunodeficiency (SCID-X1)(Cavazzana-Calvo et al., 2000, Hacein-Bey-Abina et al., 2010), adenosine deaminase deficiency (ADA-SCID)(Aiuti et al., 2002, Aiuti et al., 2009), X-linked chronic granulomatous disease (X-CGD) (Ott et al., 2006, Kang et al., 2010), X-linked adrenoleukodystrophy (X-ALD) (Cartier et al., 2009) and Wiskott-Aldrich syndrome (WAS) (Boztug et al., 2010). Initially there was a problem with low transduction efficiency rate which later improved. However, gene therapy faced a major setback in 2003 when clonal vector mediated leukaemias were reported in several patients during SCID-X1 trial (Hacein-Bey-Abina et al., 2003, Stein et al., 2010). Therefore, it is important to develop sensitive and stringent methods to assess viral vector mediated

genotoxicity prior to gene therapy treatment. It is also crucial to design safer integrating vectors for genetic modification of diseases. In the next section, some of pivotal HSC gene therapy human clinical trials and their associated complications will be covered in greater details.

#### **1.1.16.1. HSC gene therapy trials for SCID-X1**

SCID-X1 is an X-linked inherited disorder caused by cytokine receptor common gene mutations which is located on the X-chromosome. Patients develop major complications such as severe immunodeficiency and they die in early childhood (Leonard, 1996). The first report of successful gene therapy for SCID-X1 was reported in 2000 and provided a great confidence in the field (Cavazzana-Calvo et al., 2000). However, three years after gene therapy, uncontrolled clonal proliferation of transduced T cells was detected in the two youngest patients. The results shown the leukemic cells containing proviral insertions have made aberrant *LMO2* gene expression (Hacein-Bey-Abina et al., 2003). A total of 20 patients with SCID-X1 were treated during HSC gene therapy from 1999 to 2009 using corrective gamma-retroviral vectors in France and England (Gaspar et al., 2004, Fischer et al., 2010). Unfortunately, five out of 20 patients were developed T cell leukaemias between 23 and 68 months, following treatment. As a result, one patient died and other four patients treated with chemotherapy and/or allogenic transplantation (Hacein-Bey-Abina et al., 2008, Howe et al., 2008). The researchers concluded that proviral enhancer activated proto-oncogene *LMO2* which was documented in all four cases and *CCND2* proto-oncogene for the fifth patient. This report raised concerns about the safety aspect of gene therapy treatment and also sparked a detailed analysis of the genotoxicity of viral vectors.

#### **1.1.16.2. HSC gene therapy trials for ADA-SCID**

A more successful gene therapy trial for ADA-SCID was conducted in Milan which they used gamma-retroviral MLV-based vectors to deliver therapeutic gene (Aiuti et al., 2002). As opposed to the SCID-X1 clinical trial, none of the patients (19 in total) with ADA deficiency developed any side effects (Cassani et al., 2009, Fischer et al., 2010). Further, no clonal expansion or hematopoietic malignancies were detected. Insertion site analysis was carried out in transduced cells before and up to 47 months following transplantation. Results revealed that retroviral insertion sites were

concentrated more in gene-dense regions, promoters and transcriptionally active genes both in pre transplanted CD34<sup>+</sup> and *in vivo* cells. Insertion site locations were close to or within proto-oncogenes and also in genes controlling cell growth and proliferation such as *LMO2*. However, in these patients the corrected T cells carrying *LMO2* insertion did not clonally proliferate and did not transform into leukaemia (Aiuti et al., 2007). Results from SCID-X1 and ADA-SCID trials indicated that the characteristics of the transduced cells, *in vivo* proliferative behaviour of the cells, underlying diseases, therapeutic transgene, vector and transduction could affect the likelihood of transformation of potentially high risk clones to clonal dominance and leukaemia (Wu and Dunbar, 2011).

#### **1.1.16.3. HSC gene therapy trials for X-CGD**

X-CGD is caused by mutations of the *CYBB* gene. The normal function of this gene is required by neutrophils to generate microbial oxidants. Patients lacking this gene develop complications such as life threatening bacterial and fungal infections (Heyworth et al., 2003). In a clinical trial, two patients received corrected CD34<sup>+</sup> cells. In this trial, they employed gamma retroviral vector expressing corrected gene (Ott et al., 2006). Initially, high levels of corrected neutrophils and monocytes were detected, and patients displayed notable clinical improvement with less serious chronic infections. However, several months following infusion treatment rapid expansion of transduced myeloid cells occurred in both patients. Therefore, multiple expansion of clones with retroviral insertional activation of *MDS1/EVI1*, *SETBP1* or *PRDM16* were detected. Both patients following 15- and 28-months post infusion treatment developed myelodysplasia/acute myeloid leukaemia and one patient died 27 months after receiving treatment (Stein et al., 2010). Another gene therapy trial was also conducted, using another type of retroviral vector. However, no evidence for clonal dominance or leukaemia during three years follow up was detected. The difference in insertional genotoxicity between these two trials was related to the vector type. For the first trial a potent pro-viral enhancer was incorporated in the vector backbone and therefore it was more likely to activate nearby proto-oncogenes (Kang et al., 2010).

#### **1.1.16.4. HSC gene therapy trials for WAS**

Two patients underwent gene therapy trial for WAS which is an X-linked recessive primary immunodeficiency-thrombocytopenia disorder (Boztug et al., 2010). They

employed a retroviral vector expressing the WAS protein and was pseudotyped with the gibbon ape leukaemia virus (GALV) envelope protein. Following infusion of corrected CD34<sup>+</sup> cells, clinical condition of both patients was improved. Following two and half years post-infusion treatment clonal distribution and expansion of corrected cells *in vivo* were studied. However, retroviral insertions in the *LMO2*, *CCND2* and *BMI1* genes were detected in corrected T cells for both patients (Boztug et al., 2010). Previously it has been shown that activation of these genes by vector insertions in patients with SCID-X1 can cause malignant transformation. Later on it was reported that one patient developed T-cell acute lymphoblastic leukaemia (T-ALL) as a result of vector mediated activation of *LMO2* gene, similar to the T-ALL complication detected in the SCID-X1 trial (Wu and Dunbar, 2011).

#### **1.1.16.5. HSC gene therapy trials for X-ALD**

X-ALD is a severe neurodegenerative disorder caused by *ABCD1* gene mutations. This will lead to adrenoleukodystrophy protein deficiency (ALDP) and inability to break down very long chain fatty acids (Moser et al., 2007). The best treatment option was to replace brain HSC-derived microglial cells by HSC gene therapy. A pilot study was conducted in 2009 in which a replication incompetent lentiviral vector was used to transduce autologous CD34<sup>+</sup> cells. Clinical improvements were observed for the 24 to 30 months of follow up. Further, large-scale analysis of IS revealed a high number of unique ISs and polyclonal distribution of transduced corrected HSC cells over time. Although insertions were located mainly in gene coding regions, no evidence for clonal expansion was reported in patients (Cartier et al., 2009).

Based on these trials, efforts have been made to develop sensitive preclinical safety tests prior to gene therapy applications. Different *in vivo* genotoxicity models and the insights from these models are covered in the following section.

#### **1.1.17. Importance of gene therapy preclinical models**

Human gene therapy clinical trials can provide relevant insights regarding gene therapy efficacy and genotoxicity of viral vectors. Nonetheless, our information is hampered by the small number and size of the clinical studies. Also, underlying disease might affect the clonal reconstitution pattern following gene therapy treatment. Hence, integration analysis of different viral vectors in cell lines and in preclinical animal models can provide useful information to assess vector genotoxicity. A decade

ago, the full sequence of human genome became available. Since then it is possible to investigate IS distribution, which was not previously feasible. Following this success, the first large scale analysis of HIV-1 based lentivirus IS in a human T cell line was published. In this study, ligation-mediated PCR (LM-PCR) was used to identify IS in transduced cell line (Schroder et al., 2002). A useful technique was developed to assess vector-mediated genotoxicity *in vitro*. In this method high level transduction of primary murine bone marrow cells was used. Following expansion, limiting dilution plating step was used to detect immortalised clones (Modlich et al., 2006). This technique was a further improvement to a previous published protocol (Du et al., 2005). In the old protocol, MLV vectors derived immortalised myeloid cells were generated and the clones had vector insertions close to *Evi1* or related proto-oncogenes. This study revealed SIN-MLV vectors with a strong internal enhancer/promoter may also modify cells by IM event. However, transformation capacity of the vector was significantly reduced compared to normal full LTR containing vector construct (Modlich et al., 2006). This method was also employed to study genotoxicity effect of other classes of viral vectors (Zychlinski et al., 2008, Modlich et al., 2009).

#### **1.1.17.1. Examples of current *in vivo* genotoxicity preclinical models**

Integrating viral vectors are widely used in gene therapy and may induce oncogenesis. Gamma-retroviral vector gene transfer into HSCs resulted in high frequency of clonal proliferation and leukaemia in HSC clinical trials. Therefore, safer vector constructs and sensitive preclinical safety assays are urgently required to encounter this major hurdle. A useful *in vivo* HSC mouse genotoxicity model for murine leukaemia virus-derived (MLV-derived) (gamma-retroviral and lentiviral vectors) is tumour-prone *Cdkn2a*<sup>-/-</sup> mice (Lund et al., 2002, Montini et al., 2006). The *Cdkn2a* locus has an important role in regulation of senescence. It also inhibits cell modification caused by aberrant oncogene expression. Therefore, *Cdkn2a* inactivation promote several types of cancer-triggering lesions. The tumour-prone *Cdkn2a*<sup>-/-</sup> mice have been indispensable in IM studies for identifying cancer genes, many of which are highly applicable in human oncogenesis. It was reported that expression of *CDKN2A* locus was lost as a secondary mutation in two patients with X-SCID. Therefore, they developed gamma-retroviral induced leukaemia. These findings validated this model further for vector genotoxicity screening (Montini et al., 2009).

The results from one study using *Cdkn2a*<sup>-/-</sup> mice revealed retroviral vectors can induce leukaemia/lymphomas depending on LTR enhancer activity in a dose dependent manner. However, lentiviral vectors seemed to have a safer profile even with higher copy numbers. Also, IS enrichment in cancer and cell cycle related genes were exhibited more in retroviral vectors but not in lentiviral vectors insertion profiles. The other objective of this study was to compare genotoxicity effects of vectors with elimination of enhancer elements (SIN-gamma-retroviral vectors). As a result, vectors mediated genotoxicity of these modified vectors were considerably reduced (Montini et al., 2009).

A current challenge for the gene therapy field is identifying stringent preclinical assay systems for vector safety assessment prior to implementation of clinical trials (Nienhuis et al., 2006, Corrigan-Curay et al., 2012). Preclinical assays for vector-induced genotoxicity have been developed. However, these models are subject to several limitations. One of this strategy is to use homologous recombination to recreate LMO-2 insertion events which were detected in some human X-SCID leukaemia cases. With this strategy and shuttling several vector elements into the endogenous LMO-2 locus they were able to evaluate transactivation potentials (Zhou et al., 2010). This technique proved to be a useful test for X-SCID gene therapy however, it was restricted by site specificity and T cell lineage only. A myeloid immortalization assay has also been established to assess vector safety and unlike the previous method was not limited to insertion events at one locus. The problem with these assays were most immortalizing events occurred in a small subset of genes including *Evi1* and *Prdm16* (Modlich et al., 2006, Ott et al., 2006).

Mouse transplanted assays are also used to assess vector safety and can detect a variety of vector mediated hematopoietic transformation events (Kustikova et al., 2005, Modlich et al., 2008). The advantage of using these assays is that they are not restricted by predetermined integration sites. Nonetheless, mouse transplant studies are very expansive, time consuming, labour intensive and they are sensitive enough to model vector mediated transformation events (Will et al., 2007). To overcome sensitivity, tumour prone mouse models were proposed which carry a targeted mutation in cancer-related genes such as a tumour suppressor gene (Montini et al., 2006, Lund et al., 2002, Montini et al., 2009). However, this method enhances intrinsic rate of tumour formation even further and reduces the specificity of the assay.

A murine serial transplant assay was also developed to assess safety of SIN lentiviral vectors. It is argued that serial transplantation assay is more sensitive to detect vector mediated hematopoietic transformation. One of the strategies to accelerate genotoxicity in the mouse is serial transplantation. This strategy facilitates clonal haematopoiesis and leukaemia in secondary and tertiary rounds of transplantation. This method selectively triggers activation of proto-oncogenes in a platform which rapid expansion of HSC are required. In the first step of the study, bone marrow cells from *Il2rg*<sup>-/-</sup> mice were infected. Transduced bone marrow cells were transplanted into sub-lethally irradiated *Rag2*<sup>-/-</sup>, *yc*<sup>-/-</sup> recipient mice. Following transplantation, mice were monitored for 20 to 28 weeks. At the end of follow up period, bone marrow cells were harvested from each primary recipient and transplanted into two to three secondary *Rag2*<sup>-/-</sup>, *yc*<sup>-/-</sup> recipients. Again, following transplantation mice were monitored for an additional 24 to 28 weeks. Full analysis was conducted in all the secondary recipients. For this study, serial transplant assays were conducted over long term which took place over 10 to 14 months. A complete downstream analysis was performed in all tumour cases detected with the X-SCID lentiviral vector. These findings exhibited the advantages and restrictions associated with long term murine transplant studies (Zhou et al., 2013). One of the advantages is that they can be performed on diseased background or other genetic background models such as tumour-prone models. However, these studies are very long, costly and they do require large number of animals. Technical procedures are severe and involves animal suffering to a great extent. Also, in these assays no human cells are used and may not resemble what truly happens in humans. To overcome sensitivity, secondary or tertiary transplants are used. However, the human relevance issue remains challenging. One other problem is tumour formation as a result of non-relevant mechanisms which questions its application in assessment of viral vector safety.

Another sensitive genotoxicity model is foetal mouse model which demonstrated lentiviral vector mediated mutagenesis resulting in liver oncogenesis (Nowrouzi et al., 2013). This model has highly expressed genes that controls cellular proliferation and differentiation which has a strong link with hepatocellular carcinoma (HCC). The MF-1 outbred mice strain was used for *in utero* injection. This strain is not genetically prone to cancer and facilitates life-long vector presence with permanent transgene expression. In this study MF-1 outbred strain received non-primate (np) and primate



LV vector injections and HCC were developed in mice receiving np LV vector. The two identified vector-inserted genes such as *Park 7* oncogene and *Uvrag* tumour suppressor gene were associated to networks specific to liver disease and HCC through Ingenuity Pathway (IP) analysis. This model is a great tool not only to assess the genotoxic potential of vectors but can also identify genes associated with liver oncogenesis (Nowrouzi et al., 2013).

Several differences exist between murine and human haematopoiesis and murine models particularly serial transplant models do require almost a year of monitoring. Despite these differences, preclinical gene therapy models can help to evaluate HSC gene therapy efficacy and provide useful insights into safety (Wu and Dunbar, 2011). Humanised mouse models have also been used to transplant transduced human long-term repopulating cells (Joseph et al., 2010, Frecha et al., 2011). These models have been used to optimise gene therapy technology and to assess gene therapy vector safety and efficiency. It is argued large animal models are an invaluable tool to evaluate genotoxicity in preclinical gene therapy studies. Dogs and non-human primates have also been used as predictive models of genotoxicity (Trobridge and Kiem, 2010). Due to easy accessibility, dogs have been used as diseased models of human inherited genetic diseases such as SCID-X1 (Felsburg et al., 1992, Beard and Kiem, 2009). In one study a dog model was used to compare IS of long-term repopulating cells transduced with gamma-retroviral, lentiviral, and foamy viral vectors. Findings revealed that gamma-retroviral vectors exhibited a high frequency of IS in proximity to transcription start sites. Also, gamma-retroviral vector insertions were more frequently located withing and proximity to proto-oncogene transcription sites than other viral vectors. It was concluded that this retroviral vector can trigger risky gene activation (Beard et al., 2007).

In comparison to dog models, non-human primates have also been used as model for genotoxicity screening (Trobridge and Kiem, 2010). They have a closer genetic resemblance to humans and results from these studies may predict better outcome for human gene therapy clinical trials. It is argued that genotoxicity data obtained from non-human primates are most relevant data available to evaluate risk of vector-mediate IM in humans. In a study 42 rhesus macaques, 23 baboons and 17 dogs with high levels of gene transfer were monitored for a period of 3.5 years, using retroviral vector transduced CD34<sup>+</sup> cells. As a result, no evidence of oligoclonal or monoclonal

haematopoiesis was detected (Kiem et al., 2004). Nonetheless, 5 years following transplantation, one rhesus macaque developed serious complication of myeloid sarcoma which is a type of acute myeloid leukaemia. Downstream analysis from the tumours confirmed two clonal vector insertions which one of them was in the anti-apoptotic gene *BCL2-A1* (Seggewiss et al., 2006). Therefore, using non-primate models over long period of follow up can be a valuable tool to assess vector safety.

### **1.1.18. Challenges for genotoxicity screening**

The current state of genotoxicity screening strategy depends on monitoring effects on DNA damage or mutation. Results can be obtained after a relatively short exposure and allowing some time for detection of any mutation. These are disadvantages if particularly vector-mediated IM event is aimed for. It is known that IM can take weeks, months or even years to occur in patients. Nonetheless, it was suggested that viral-vector mediated genotoxicity screening assay for IM should be quantified and reproducible (Bokhoven et al., 2009). Over the years, assessment of efficacy and safety of the therapeutic vector was mainly conducted through *in vivo* studies in wild type rodents and larger animal models. For *ex vivo* genetic modification of autologous or heterologous cells immunocompromised animals, diseased models and tumour prone models are required (Misra, 2013, Modlich et al., 2009). For genotoxicity screening platforms, specific animal strains such as immune-deficient NSG™ mice are employed which are highly expensive. Furthermore, for *ex-vivo* genotoxicity screening transplantation procedures are extremely difficult. These procedures usually require collection of HSC samples, cell purification followed by *ex-vivo* transduction which are complex. Also, several animal “donors” are used to obtain HSCs and purify the collected samples. Occasionally serial transplanted assays are required by regulatory bodies. Therefore, more recipient animals are needed to conduct the studies. Information regarding vector design and underlying biology of the disease environment can be obtained from preclinical animal studies. However, other factors such as target tissue, cell type and differentiation on potential genotoxicity have partially studied (Aiuti et al., 2013).

### **1.1.19. Replacement, Reduction, Refinement (3Rs) benefits**

For gene therapy product approval, animal testing is usually required by regulatory bodies. These studies are compulsory prior to initiation of human gene therapy trials

and will be conducted on each individual product. Each integrating vectors have unique properties and therefore efforts are made to design robust preclinical studies for successful prediction of clinical outcome. It is argued that a specific, fit for purpose preclinical study assessment is required by regulatory bodies based on sound science and risk assessment (Narayanan et al., 2014). For some gene therapy product, it is possible that enough information from preclinical result and/or clinical study is available. Therefore, relevant *in vivo* studies can be undermined. However, it is also possible for novel vectors and also new *in vitro* methods new data to be validated against existing previous data and to reduce number of animals.

In addition to animal models, *in vitro* culture assays such as *in vitro* immortalisation assays have also been explored. For example, primary mouse HSCs were cultured *in vitro* for an extended period and gamma-retroviral vector integration and upregulation of *Evi1* gene was measured. High-throughput IM screening has also been developed, using mice models (Modlich et al., 2006). These models were used to identify cancer genes and to assess viral vector safety. One caveat with current *in vitro* and *in vivo* models is that all employed mouse-derived stem cells or mice, respectively. Furthermore, human genetic relevance was also questioned and whether the viral vector insertion is similar between humans and mice. Notably, there is a big overlap in majority of mice integration sites with human oncogenes and tumour suppressor genes. This confirms the feasibility of considering mouse cell culture assays for monitoring viral vector-mediated IM. However, it is suggested that these assays should be limited to less common integration sites (Kool and Berns, 2009).

For successful vector-mediated genotoxicity screening, exposure and incubation time with the vectors must also be considered. It is known that vectors have a slow incubation time. Therefore, expected results might be delayed and occasionally no possible result can be observed within the standard toxicology assay timeframe. A minimum time period of 4-5 weeks are required for detection of IM event in murine HSCs *in vitro* (Modlich et al., 2006). In tumour prone *Cdkn2a*<sup>-/-</sup> mice, identification of integration sites and tumour development were assessed 6 weeks following transplantation (Montini et al., 2006). Development of new assays are required to assess non-viral factors such as age, different disease background for IM. Using human derived cells such as hiPSCs and novel 3D models may overcome the current issue of *in vivo* animal models and support extended culture time exposure,

respectively. Further, current *in vitro* assays are sensitive. However, they are restricted to specific oncogenes, specific cell lineages and one mechanisms of genotoxicity. Therefore, these assays are more applicable to screen new vectors for inherent genotoxicity via specific mechanisms rather than estimation of overall risk of oncogenicity (David and Doherty, 2017).

Based on aforementioned information, there is an urge for robust non-clinical screening to assess the safety of the therapeutic strategy for the patient. There is no “gold standard” *in vitro* or *in vivo* genotoxicity screening model is available to address the current challenges in the field. Therefore, development of a robust model is required to assess safety with high sensitivity for each individual vectors or diseases and to identify the relevant risk factors promoting the induction of tumour formation. A robust *in vitro/in silico* tool is required to evaluate the risk of IM/oncogenesis while limiting the use of animals 3Rs.

Development of new *in vitro/in silico* platform enable scientists to evaluate early safety profile of viral vectors for IM. These models facilitate vector screening for early positive/negative decisions while minimising the need for *in vivo* tumorigenicity and genotoxicity studies. Further, these models can be extensively used for future novel vectors and a good replacement for *ex vivo* and *in vivo* gene therapy approaches. Development of such technique can be extended to screen the safety profile of other emerging technologies, such as gene editing which can a risk factor of IM. Using stable unlimited self-renewal cell line such as hiPSCs and *in vitro* 3D platforms may enhance the *in vivo* resemblance and support longer-term culture for genotoxicity assays.

Human pluripotent stem cells (hPSCs) such as hiPSCs can be a great alternative to develop human based *in vitro* models to replace animal model or animal-derived stem cells-based *in vitro* models. Further, this cell line can be used for disease modelling and evaluate the risk of IM events in any desired tissue type prior to clinical gene therapy. In the next section of this thesis the focus will be on liver the tissue of interest and PSCs as an alternative source to generate mature Hepatocyte-like Cells (HLCs).

## **1.2. The Liver**

The liver is considered a vital organ due to its wide range of functions. The array of functions includes metabolism, immunity and detoxification. The functional unit of the liver is known as lobule and is mainly composed of hepatocytes. During embryonic

development, liver emerges as a section of the foregut. It derives from endodermal cells and forms as part of the hepatic diverticulum during the fourth week of development. The formation of diverticulum is induced by several signalling pathways, particularly the wnt/beta-catenin pathway and fibroblast growth factor (FGF). The freshly isolated primary liver hepatocytes are a valuable tool for *in vitro* drug screening and toxicological studies (Saxena et al., 1999).

### **1.2.1. Human liver development**

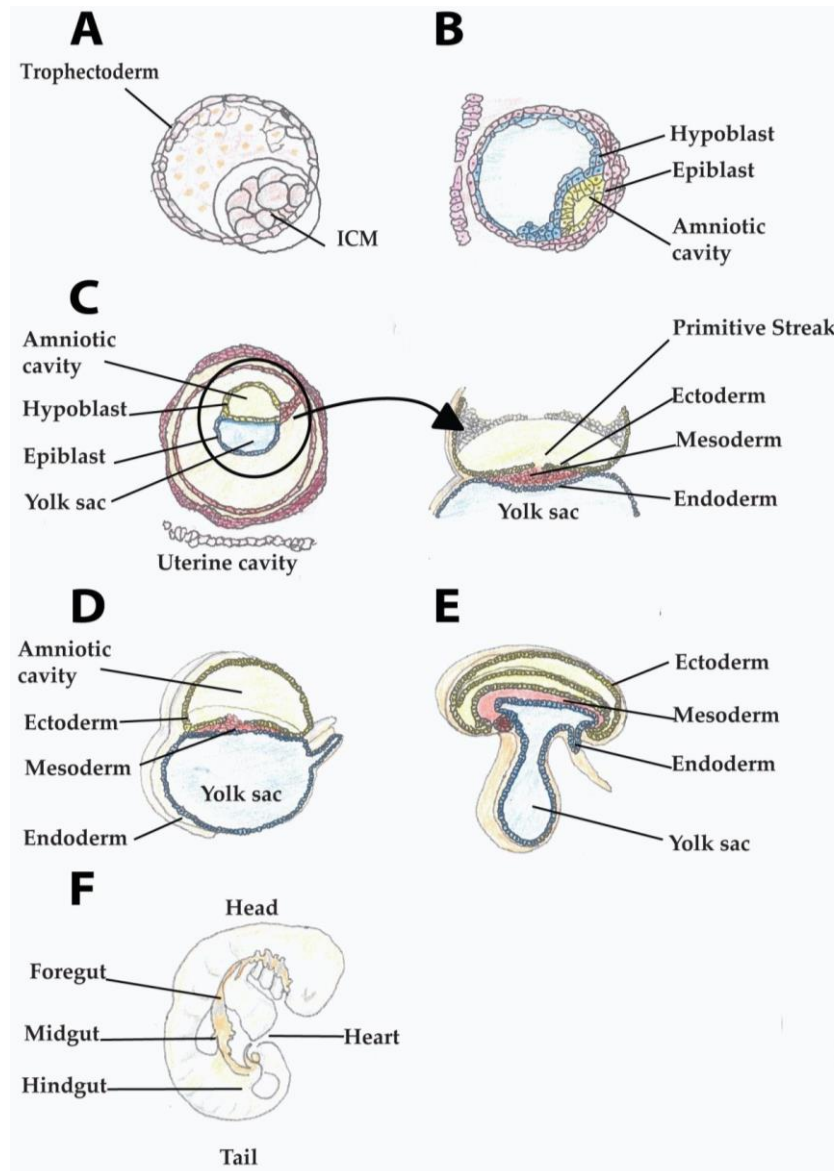
Human liver development is a multistep process which requires an extensive network of different signals from the extracellular matrix and adjacent mesoderm. These signals work in synergy in a time and/or dose-dependent manner. To generate *in vitro* hepatocytes from hPSCs and to appreciate their potential in downstream applications, it is crucial to comprehend the origin and properties of these cells during embryonic development.

#### **1.2.1.1. Mammalian Embryonic Development**

Following fertilisation of an egg from zygote, this single cell undergoes multiple rounds of divisions forming a cluster of 4-16 smaller cells called the morula (Hardy et al., 1989). These cells are surrounded by a glycoprotein membrane called the zona pellucida (Biswas and Hutchins, 2007). Following this stage, the blastomere develops and forms a central fluid cavity which is called the blastocoel (Hardy and Spanos, 2002). Blastocoel contains a cell mass known as blastocyst. During early embryonic development, the blastocysts composed of an outer layer of cells known as trophoblast and an inner cell mass (ICM) originated from the primitive blastocyst (Johnson et al., 1986). Embryonic stem cells are derived from the ICM at this embryonic developmental stage. During implantation stage, the ICM separate itself from the trophoblast forming a cavity structure known as amniotic cavity (Gardner and Rossant, 1979). Two specialised cell types form a layer within the amniotic cavity. These cells are called hypoblast and epiblast cells which are originated from the differentiation of the ICM (Rossant, 2008).

Following this stage of development, gastrulation occurs. This step involves with the formation of a thick ridge of cells called the primitive streak (PS). These cells are derived from the epiblast cells. Exposure of PS cells to specific signalling pathways such as Activin/Nodal, BMPs, FGFs and Wnt will drive the cells to migrate towards or

inwards the PS. These movements and rotations result in the formation of the three main embryonic germ layers. These three important layers of cells are called ectoderm, mesoderm and endoderm which ultimately form all of the cell types identified in the adult body (Figure 1.13)(Lawson et al., 1991).



**Figure 1.13: Early embryonic gut development. A) As a result of egg fertilisation and subsequent cell divisions, blastocyst is formed. Blastocysts composed of an outer layer of trophectoderm and an inner cell mass. B) The cells in inner cell mass form two other cell layers: the epiblast and the hypoblast which engulf the amniotic cavity. C) Following implantation, gastrulation initiates which lead to the formation of the primitive streak. Also, epiblast cells are rearranged and form three main embryonic germ layers. D-E) These cells are differentiated into the ectoderm, mesoderm, and endoderm. F)As development progresses a primitive gut tube is formed with three distinct foregut, midgut and hindgut domains (Giasecke, 2001).**

### 1.2.1.2 Hepatic Endoderm Specification

Following the movement and migration of the newly specialised endoderm cells, the embryo undergoes further rotations. These movements promote the endoderm to become the innermost layer (Lawson and Pedersen, 1987). In addition to the migration, these cells expand and ultimately generate the anterior, posterior and lateral domains to form a closed primitive gut tube. The primitive tube is further subdivided into three domains: foregut, midgut and hindgut. As the development progresses, the hindgut domain forms the large intestine, the midgut develops into the small intestine and the foregut generates the thyroid, lungs, liver, oesophagus, pancreas, the biliary tree and stomach.

Three distinct domains of hepatic progenitor cells were identified, using fate mapping experiments in mouse embryos. These experiments were performed at embryonic day 8.0 of gestation (E8.0). These cells were two lateral paired domains and a third ventral midline domain of the foregut which the liver originates (Zaret and Grompe, 2008). During this step, from the ventral endoderm segment, ALB, AFP and transthyretin are expressed. This is the first molecular evidence of the liver development stage (Hay et al., 2008a). Further, the closure of foregut occurs by the migration of the lateral domains towards the midline and fusion of this section with the ventral domain. This results in generation of a single prehepatic domain lying closely to the newly developed heart at E9.0. These events will lead to formation of the mesothelial cells of the proepicardium and septum transversum (Figure 1.14).

From previous studies in chick, frog, mouse and zebrafish embryo models, it has been reported that a surge of fibroblast growth factor FGF1 and 2 from the developing cardiac mesoderm and bone morphogenetic proteins BMP2 and 4 from the septum transversum mesenchyme (STM) play a key role in hepatic induction. It is well argued that the developing cardiac mesoderm activates FGF signalling which is crucial for hepatic induction (Calmont et al., 2006). The hepatic cell fate specification by FGF signalling is evolutionary conserved and acts in a concentration dependent manner (Cheng et al., 2003). During liver development, certain morphogenetic changes occur. These changes are induced because of distances between the hepatic endoderm and the cardiac mesoderm. Therefore, the concentration of FGF signalling decreases, prohibiting the cells to differentiate into an anterior fate such as the lungs.

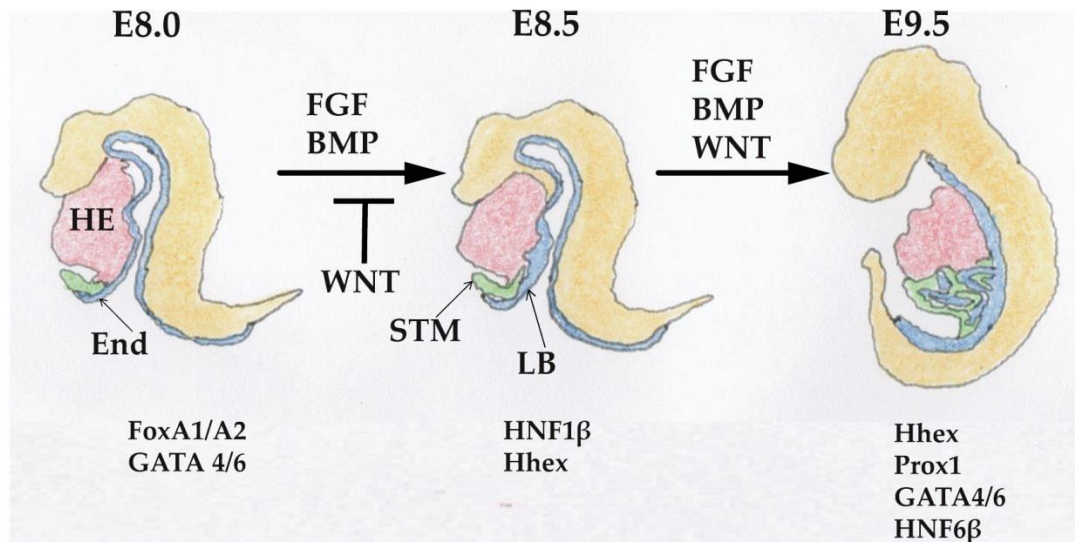
Concentration gradients of FGF signalling initiates key responses by synchronising developmental events (Serls et al., 2005). Experimental studies of ventral endoderm in the presence or absence of different members of the FGF family have revealed that FGF1 and FGF2 are crucial for the expression of ALB which is an early marker of hepatic cell fate specification (Gualdi et al., 1996). Further, it is argued that the FGF signalling activates other pathways such as the RAS/MAP kinase pathway (MAPK) ERK1 and ERK2 and the phosphoinositide 3-kinase pathway (PI3K/AKT). ERK signalling pathways has a role in enhancing hepatic gene expression and the stability of newly formed hepatocytes while PI3K/AKT pathway is more involved with the hepatic growth (Calmont et al., 2006).

For proliferation and robust differentiation of newly emerged hepatic endoderm cells different signalling molecules from members of the transforming growth factor beta (TGF $\beta$ ) superfamily and the bone morphogenic protein BMP2 and 4 are required. The former is secreted by the STM and the latter in synergy with FGF signalling enhances hepatic stability and drives the specification of the primitive endoderm through zinc finger transcription factor such as GATA4 (Huang et al., 2008). It is reported that synergism between GATA4 and the transcription factor Forkhead box (FOX) A at the gene promotor, activates other transcription factors such as hepatocyte nuclear factor 1 (HNF1) and CCAAT-enhancer binding protein beta (C/EBP $\beta$ ) (Bossard and Zaret, 1998). This will affect the transcriptional activation of ALB and controls the level of ALB expression during hepatic development. The effect of BMP2 and 4 binding to the constitutively active serine/threonine kinase type-II receptors is also crucial. This binding induces the phosphorylation of Gly-Ser (GS) domains in the kinase type-I receptors. This activation further triggers series of reactions by phosphorylation of Smad1, Smad5 and Smad8 which in coordination with Smad4 regulates the gene expression in the nucleus (Shi and Massague, 2003).

The WNT signalling pathway plays a key role during the onset of the hepatic development. However, the exact contribution of WNT signalling pathway is distinctly complex and it is varied depending on the embryonic developmental stage (McLin et al., 2007). During early somite stages, WNT signalling has a repressive role in the posterior endoderm. This will inhibit the expression of Hhex which is an important transcriptional regulator of hepatic development. On contrary, in the anterior endoderm, the expression of WNT antagonists activates the expression of Hhex and



drive the endoderm cells to a hepatic fate. Following these crucial steps, WNT signalling promote liver bud growth and further differentiation in different systems such as *Xenopus* and zebrafish (Goessling et al., 2008).



**Figure 1.14: Hepatic endoderm specification.** This figure depicts the onset of liver parenchyma cells, key signalling molecules and transcription factors. Wnt signalling pathway promotes posterior endoderm fate and must be inhibited by local Wnt antagonist expression. In later stages of differentiation, Wnt signalling in combination with FGF and BMP signalling drive hepatic specification, expansion and mature hepatocyte differentiation. In this figure E refers to days during embryonic development. Differentiated cells are indicated with colours. Foregut endoderm (End; Blue), Heart (HE; Red), Liver Bud (LB; Blue), Septum Transversum Parenchyma (STM; Green) (Duncan et al., 2005).

## 1.2.2. Formation of the Hepatic Bud from Hepatic Endoderm

### 1.2.2.1. Proliferation and Migration

The newly specified ventral endodermal cells will form the liver bud. At this stage, these cells are called hepatoblasts. At E9.5, liver bud thickens as the cells modify from a simple cuboidal to a pseudostratified columnar epithelium (Bort et al., 2006). Different hepatocyte genes are expressed including ALB, AFP and hepatocyte nuclear factor 4 alpha (HNF4 $\alpha$ ). These markers are an early indicator of early hepatic fate (Zaret and Grompe, 2008). Hepatic endoderm cells are surrounded by laminin and

collagen IV-rich basal layer which are broken down between embryonic days E9.0 and E9.5 (Zhao and Duncan, 2005). In addition, the hepatoblasts delaminate and migrate into the STM segment and the newly emerged liver bud (Bort et al., 2006). This movement and migration is controlled by an extensive network of specific transcription factors including Hhex, Prox-1, the Onecut factor 1 (OC-1 or HNF6 $\alpha$ ) and Onecut factor 2 (HNF6  $\beta$ ).

The homeobox transcription factor Hhex has a vital role in proliferation and migration of endoderm cells and is important for cell pseudostratification (Bort et al., 2006). This transcription factor is activated by GATA4 and/or GATA6 and also it is crucial to maintain the differentiation state and stability of the hepatoblasts (Zhao and Duncan, 2005). The hepatoblast delamination process is governed by the homeodomain transcription factor the Prospero-related homeobox 1 (Prox1). The Prox1 has pivotal role in degradation of basal matrix surrounding the liver bud, hepatoblast delamination and migration of hepatoblasts from the adjacent basal layer into the STM (Sosa-Pineda et al., 2000). Onecut factors OC-1 and OC-2 are responsible for the morphogenesis and expansion of the liver primordium. These factors regulate a specific gene network which involves in cell adhesion and hepatoblast migration (Margagliotti et al., 2007). Furthermore, they regulate the expression of ECM proteins and remodelling enzymes including matrix metalloproteinases (MMPs)-14 and MMP-2 which are highly expressed in the hepatic progenitors and the surrounding mesenchyme, respectively (Margagliotti et al., 2007). Endothelial progenitor cells are embraced between hepatic epithelium and the STM. These cells are in close contact with blood vessels while the hepatoblasts migrate into the stroma. It is argued that inhibition of angiogenesis prohibits liver bud growth formation in culture (Matsumoto et al., 2001). Therefore, it is concluded that endothelial cells provide crucial paracrine and physical signals which lead to hepatoblasts migration and further proliferation.

#### **1.2.2.2. Liver Bud Growth**

The progression of liver bud growth is induced by paracrine signals from surrounding and hepatic mesenchyme. Specific signalling pathways and the synergy between FGF8 via PI3K pathway, BMP4 (Sekhon et al., 2004) and Wnt/  $\beta$ -catenin signalling drive hepatic growth (McLin et al., 2007). Hepatocyte growth factor (HGF) is a potent mitogen which is highly expressed by mesenchymal cells and is required for hepatoblast migration (Deutsch et al., 2001). In addition, for hepatoblast proliferation,

HGF binds to the tyrosine kinase receptor c-Met and TGF $\beta$  signalling activation via Smad2/Smad3 pathway (Weinstein et al., 1994). It is reported that HGF and TGF $\beta$  signalling together have a key role in regulation and controlling hepatic architecture (Weinstein et al., 1994). Furthermore, FGF and HGF signalling triggers several intracellular kinase cascades such as MAPK and PI3K pathways. These pathways stimulate the activity of  $\beta$ -catenin in the liver bud, suggesting a possible crosstalk with the Wnt signalling pathway (Berg et al., 2007).  $\beta$ -catenin is so called a mediator of the canonical Wnt signalling pathway. It is widely expressed at the periphery of developing liver lobes. Further, it interacts closely with tyrosine-protein kinase Met or hepatocyte growth factor receptor (c-Met) in hepatocytes and upon HGF binding it activates in the nucleus (Monga et al., 2003).

### **1.2.2.3 Hepatocyte Specification**

Following liver bud formation and after mesenchyme invasion during mouse development (E13), the newly formed hepatoblasts are bi-potent and have ability to differentiate into either hepatocytes or cholangiocytes. These cells are also known as biliary epithelial cells. It is argued that differentiation into hepatocytes (AFP+/ALB+) or cholangiocytes (cytokeratin-19+) requires a synergy between cell signalling networks acting in a gradient manner (Jung et al., 1999). Further, the localisation of the hepatoblast also plays a key role which determines the fate of the cells. The hepatoblasts surrounding the liver parenchyma eventually differentiate into hepatocytes. However, cholangiocytes derived from hepatoblasts which are located at the portal mesenchyme (Lemaigre and Zaret, 2004).

It was reported that a gradient of Activin and TGF $\beta$  has a key role in the hepatocyte specification. The finely tuned gradient of both factors is negatively modulated by the major transcription factors Onecut-1 and Onecut-2 which drive forward the hepatocyte differentiation. Further, this pathway in coordination with Jagged-Notch pathway supports biliary differentiation (McCright et al., 2002).

### **1.2.2.4. Hepatocyte Maturation**

Following differentiation of hepatoblasts into hepatocytes, key signals from haematopoietic cells such as the cytokine oncostatin M (OSM) in addition to the HGF, regulate hepatocyte maturation which continues even following birth.

OSM is a member of the interleukin-6 (IL-6) family. It binds to the gp130 membrane receptor promoting morphological maturation into polarized epithelium via K-ras and E-cadherin via the JAK/Stat3 signalling pathway (Tan et al., 2008, Ito et al., 2000). The role of HGF is to promote the organisation of hepatocytes into cord-like structures. It also regulates the expression of diphosphate-ribosyltransferase 6 (ARF6). This is an important enzyme which involves in actin cytoskeleton remodelling. Furthermore, some studies suggest that TNF $\alpha$ , through activation of the transcription nuclear factor  $\kappa$ B (NF- $\kappa$ B), provides a synergy between HGF and OSM activity. This in turn inhibit maturation and maintains the proliferative capacity of foetal hepatocytes. Therefore, this will enable liver to grow to the appropriate size before differentiation (Kamiya and Gonzalez, 2004). It is argued that TNF $\alpha$  repression following birth is crucial to maintain fully mature hepatocytes expressing cytochrome P450 genes (Hart et al., 2009).

During maturation stage, an interaction between liver-enriched transcription factors including C/EBP $\alpha$ , HNF1 $\alpha$ , HNF3 $\alpha$ -gamma, nuclear hormone receptors and HNF4 $\alpha$  is required to control hepatocyte gene expression (Cheng et al., 2006). It is argued that HNF4 $\alpha$  expression strongly promotes hepatocyte cell phenotype (Bulla, 1997). This activation is important for hepatocyte function by regulating a chain of essential transcription factors which ultimately lead to the correct development of the foetal liver architecture (Li et al., 2000). Previous studies reported that HNF4 $\alpha$  binds to the promoters of most of the genes expressed in the mouse liver. These genes encode cell adhesion and very important functional proteins which has a role in hepatocyte epithelial structures (Konopka et al., 2007). Furthermore, some studies suggest that HNF4 $\alpha$  promote the expression of HNF1 $\alpha$  (Taraviras et al., 1994). HGF is induced by HNF1 $\alpha$  and C/EBP $\alpha$  expression which are crucial for complete hepatocyte maturation and metabolic hepatic functions (Pontoglio et al., 1996).

Furthermore, HNF4 $\alpha$  and Wnt signalling with some other factors promote the liver metabolic zonation which initiates after birth (Jungermann and Katz, 1989). Depending on which metabolism-regulating genes are expressed with the hepatic lobule, a periportal zone can be distinguished from a pericentral zone. It is reported that HNF4 $\alpha$  activates the expression of certain metabolic related genes in the pericentral zone (Stanulovic et al., 2007). The HNF4 $\alpha$  knock out livers display the same metabolic genes pattern in both periportal and pericentral zones. Additionally, Wnt signalling regulate the metabolic zonation. However, its receptor beta-catenin and the beta-

catenin negative regulator APC are expressed in the perivenous and in periportal hepatocytes, respectively (Benhamouche et al., 2006).

### **1.2.3. Liver Architecture**

The liver is considered as the largest internal organ, comprising one fiftieth of the total weight of the adult body. It has an important endocrine and exocrine properties. Endocrine functions include secretion of hormones, serum proteins such as ALB and clotting factors. The generation and secretion of bile and drug metabolism are considered as the liver's major exocrine functions.

The fundamental structural unit of the liver is the liver lobule, which exhibits a polygonal structure with a central vein. The central vein is localised in the centre of the lobule and a portal triad comprising the hepatic artery, bile duct and the hepatic portal vein. The key unit of the liver is called acinus. The acinus structure is composed of plates or cords of hepatocytes which generally extending from the central vein and lined by sinusoidal capillaries. Blood is passed the lobule via the portal triad of vessel which flows through the parenchyma and exits the structure through the central vein. This flow of blood creates crucial chemical gradient and sensitive microenvironments which divides the acinus into three key zones based on their functional properties. The periportal zone or zone 1 which has a key role in oxidative metabolism, gluconeogenesis and ureagenesis; the pericentral zone or zone 3 which promote glycolysis, liponeogenesis and metabolism of xenobiotics; and a mid-lobular region or zone 2 which exhibits a mixture of zone 1 and zone 3 functional properties (Turner et al., 2011).

The liver consists of different type of cells which are divided into two main categories: parenchymal cells and non-parenchymal cells. The parenchymal cells consisting of hepatocytes which are the major cell type of the liver. Functional hepatocytes are comprising 80% of the total liver volume and 60% of the total cell population in the liver. The non-parenchymal segment consists of bile duct epithelial cells, liver sinusoidal endothelial cells, Kupfer cells and hepatic stellate cells.

Hepatocytes are highly mature and functional epithelial cells that provide fundamental functional properties. These functions are including protein, fat, xenobiotics and steroids metabolism. The distinct hepatocyte morphology is presented by the possession of a cuboidal shape consisting of one or more nuclei. These polyploidy

cells (4N or 8N) are gradually increasing in number from zone 1 to zone 3 with distinct nuclei (Deutsch et al., 2001). Hepatocytes exhibit abundant mitochondria and Golgi complexes in close proximity to the bile canaliculi. Also, their cytoplasm is rich in rough and smooth endoplasmic reticulum. These distinctive characteristics resemble the hepatocyte secretory nature and enzymes of the phase I and phase II drug metabolism, respectively.

The non-parenchymal cells have crucial roles in the regulation of the hepatic growth and maintenance of functional properties such as the production of growth factors and other key chemical mediators of cellular function. Liver sinusoidal endothelial cells (LESC) are characterised as very thin and elongated cells. These cells are located in the hepatocyte basolateral surface within a space known as Space of Disse or perisinusoidal space. The plasma membrane of these cells exhibits several fenestrations to enhance the contact between the hepatocytes and the circulating blood. This further allows the diffusion of small molecules to the hepatocytes (Cogger et al., 2010). Hepatic stellate cells (HSC) are engulfing the endothelial cells in the Space of Disse. These cells have essential roles including sinusoid contractility, matrix remodelling and also in hepatocyte proliferation. During liver regeneration, these cells secrete certain cytokines such as HGF, TGF $\alpha$  and EGF (Friedman, 2008). Furthermore, HSC has a crucial role in deposition of extracellular matrix which is commonly seen during the progression of liver cirrhosis. This can be reversed by the role of certain cytokines secreted by the resident macrophages of the liver. These macrophages, called Kupfer cells, have a mesenchymal origin (Jaeschke, 2007).

The hepatic progenitor cells (HPC) are considered as the epithelial cells. Hepatoblasts are localised in a special compartment within the canals of Hering. These canals are very small with abundant branches which widely connect the bile canalicular system with the interlobular ducts (Gaudio et al., 2009). Under healthy conditions, HPC has a relatively low proliferation rate. However, when the liver is compromised, HPC are activated and expanded from the canals of Hering to the pericentral zone, leading to formation of mature hepatocytes or cholangiocytes (Forbes et al., 2002).

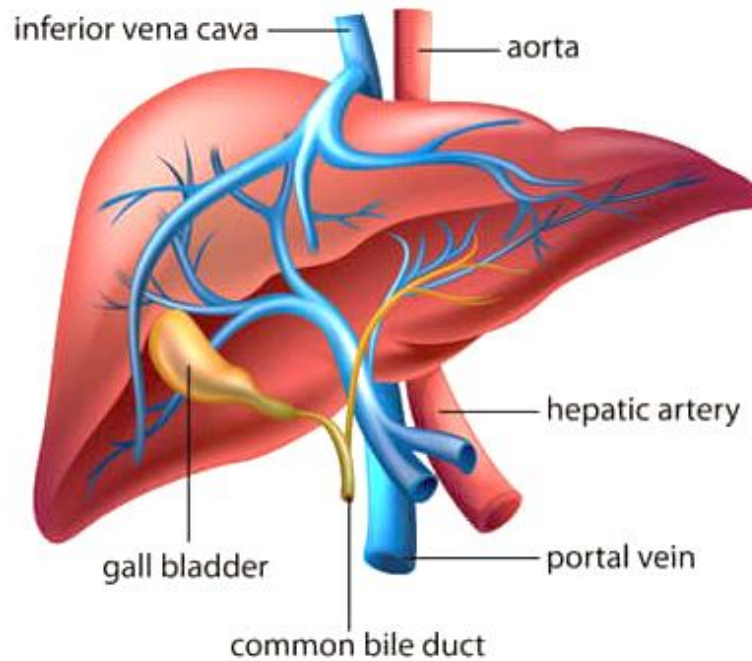
Epithelial organs generally exhibit two main domains which are called as the apical (luminal) and basolateral (blood-facing) domains. However, hepatocytes exhibit two main basolateral domains. These domains interact with the sinusoidal microvasculature on different sides of the cell layer which is known as sinusoidal

domain and an apical domain. These domains collectively form the canalicular domain. The canalicular domain generates a belt like structure surrounding each hepatocyte. This ultimately forms a network of small tubular structures terminating at the portal triad. These tubes further connect and merge with bile ducts through the canals of Hering which lead to a communal draining into the common bile duct and the gall bladder. In addition, the basolateral domains of the hepatocytes are responsible for the exchange of metabolites with the blood and providing necessary interactions with the ECM, and the apical domain has a key role in the transport of bile acids and certain products into the canaliculus which embraces the hepatocyte (Chapman and Eddy, 1989, Decaens et al., 2008).

The ECM of the liver is located in the Space of Disse which is between the hepatocytes and the liver sinusoidal endothelial cells. It has important functional properties which promote healthy architecture and phenotypic gene expression of liver cells. It is composed of a thin layer with variable composition depending on gradient. Generally, the ECM in the liver composed of fibronectin, proteoglycans and collagens type (I, III, IV, V and VI) (Martinez and Pascual, 1995). This enables a rapid bidirectional exchange of certain molecules between plasma and hepatocytes. Furthermore, it is argued that the niche of the hepatic progenitor cells is rich in laminins (Lorenzini et al., 2010).

### **1.3. Liver cells as alternative sources for *in vitro* for toxicological and pharmacological screening**

For years animal models were used for chemical toxicity screening which are both time consuming and expensive. The liver is considered as an important organ for metabolism and biotransformation of drugs. Therefore, liver cells are “gold” standard for *in vitro* toxicological and pharmacological screening. Novel and high throughput *in vitro* liver models enable scientists to screen novel compounds efficiently while minimising number of animal models (Figure 1.15).



**Figure 1.15: The liver. The functional unit of the liver is known as lobule and is composed mainly of hepatocytes.**

### **1.3.1. Primary liver hepatocytes**

Primary liver hepatocytes have been considered the “gold standard” *in vitro* tool for pharmaceutical drug testing and for other clinical applications. One caveat with primary hepatocytes is the maintenance of an *in vivo* like phenotype. Following isolation and *in vitro* culture, functional phase-I and -II metabolic activities rapidly declined 72 hours post culture (Rodriguez-Antona et al., 2002). For different cell assays and clinical applications, a prolonged functional life of hepatocytes is important.

A very recent study published that has significantly improved the viability of isolated hepatocytes to more than 90% with a stable phenotype for up to 7 days using three-dimensional printing technique. In this study alginate hydrogels have been mixed with primary hepatocytes and MSCs to improve viability of the cells (Kim et al., 2018). Although primary hepatocytes are valuable, their limitations are increasingly recognised due to their short post culture *in vitro* stability and functional activity. Thus, alternative sources such as immortalised liver cell lines namely HepG2 and HepaRG have been considered for long-term cell culture applications.



### **1.3.2. HepG2 cell line**

The HepG2 cell line was originally derived from a hepatocellular carcinoma of a 15-year old Caucasian male. Unlike primary hepatocytes, because of the low endogenous expression of cytochromes, these cells are not a great choice for the detection of hepatotoxicity (Wilkening et al., 2003). They do not resemble *in vitro* models effectively as they lack several drug transporters and a range of non-cytochrome Phase-II enzymes (Gripon et al., 2002, Guillouzo et al., 2007). Attempts have been made to improve the overall performance of the HepG2 cell line by introducing three dimensional models such as co-culturing (He et al., 2018) and the hanging drop method (Hurrell et al., 2018, Shah et al., 2018).

### **1.3.3. HepaRG cell line**

The HepaRG cell line was primarily derived from a hepatoma of a female patient with cirrhosis following hepatitis C virus infection (Gripon et al., 2002). Unlike the HepG2 cell line, these cells express several drug metabolising genes at a level similar to freshly isolated primary hepatocytes. The HepaRG cell line expresses many nuclear receptors, drug transporters and specific markers of liver hepatocytes such as ALB and aldolase B (Guillouzo et al., 2007). It is reported that HepaRG cells can be differentiated into mature hepatocytes and primitive biliary cells and have the ability to retain a relatively stable hepatic activity for several weeks (Josse et al., 2008). A few studies have demonstrated that differentiated and undifferentiated HepaRG cells have a much higher resemblance to primary human hepatocytes compared to the HepG2 cell line (Hart et al., 2010). Based on molecular profiling data, these cells are considered for studies of hepatotoxicology, xenobiotic metabolism and hepatocyte differentiation. Having said that, the rate of ALB production is much higher than in primary hepatocytes (Lubberstedt et al., 2011). Further, the levels of cytochrome activities and their relative inducibility is very similar to cultured primary hepatocytes (Lubberstedt et al., 2011). Although this cell line is similar to primary liver cells, it does not sustain long-term *in vitro* functional activity. Therefore, alternative sources such as PSCs are ideal to produce high quality human liver tissue for clinical applications. In the next section, I will focus on the properties of PSCs which is an important part of this study.

## **1.4. Pluripotent stem cells (PSCs)**

### **1.4.1. Overview**

Over the past 20 years, significant advances have been made in the field of stem cell and regenerative medicine. In this section, I provide a systemic approach to PSCs in general and recent discoveries in this rapidly developing field (Liu et al., 2020). An overview of the sections are as follows: Firstly, the experimental advances in the generation and differentiation of PSCs, secondly, maintenance of stem cells in different tissue culture settings and finally the recent advances in three-dimensional (3D) cell technology. The main aim of this section is to provide a synopsis of key historical milestones and recent research advancements in stem cell field and regenerative medicine.

### **1.4.2. Key milestones in PSCs history and classifications**

In the field of stem cell research, there are several key milestones. In 1961, more than half a century ago, the first stem cells were demonstrated by Drs Ernest A. McCulloch and James A. Till at the University of Toronto in Canada (Till and Mc, 1961). They discovered that the derived stem cells from mouse bone marrow had the ability to differentiate into different cell types and were therefore called PSCs (Till and Mc, 1961). In 1981, mouse embryonic stem cells were isolated from the blastocysts simultaneously by two different groups (Evans and Kaufman, 1981, Martin, 1981). Later, Keith Campbell, Ian Wilmut, and colleagues, successfully cloned Dolly the sheep at the Roslin Institute in 1996 and demonstrating the feasibility of somatic cell nuclear transfer (SCNT) technique (Wilmut et al., 1997). Following these advances, in 1998, the first hESCs were isolated by James Thomson in the USA (Thomson et al., 1998). In 2006, induced pluripotent stem cells (iPSCs) were generated, using adult somatic cells (ASCs). These cells were reprogrammed by four basic transcription factors (Takahashi and Yamanaka, 2006, Takahashi et al., 2007). In 2012, John Gurdon (Gurdon Institute, Cambridge, UK) and Shinya Yamanaka (Kyoto University, Japan and Gladstone Institutes, USA) received the Noble Prize for their outstanding discovery in reprogramming mature cells into pluripotent state.

In general, there are five basic categories of stem cells encompasses nuclear transfer stem cells (NTSCs), very small embryonic-like stem cells (VSELs) Embryonic Stem Cells (ESCs), reprogrammed stem cells (RSCs), and ASCs (Liu et al., 2020). It is worth

mentioning that in 2018, NTSCs have been employed to clone Macaque monkeys in China (Liu et al., 2018). However, ESCs and iPSCs have only been used to derive tissues and organs. In recent years, stem cells such as ESCs and iPSCs have shown great promise in four different fields: regenerative and transplant medicine; disease modelling; drug discovery screening and developmental biology. Therefore, the expansion of regenerative medicine field with the latest advancements in stem cells will continue forward for future translational medicine (Liu et al., 2020).

The iPSC reprogramming technology is relatively new compared to ESCs and challenges remain particularly in terms of cell proliferation and differentiation.

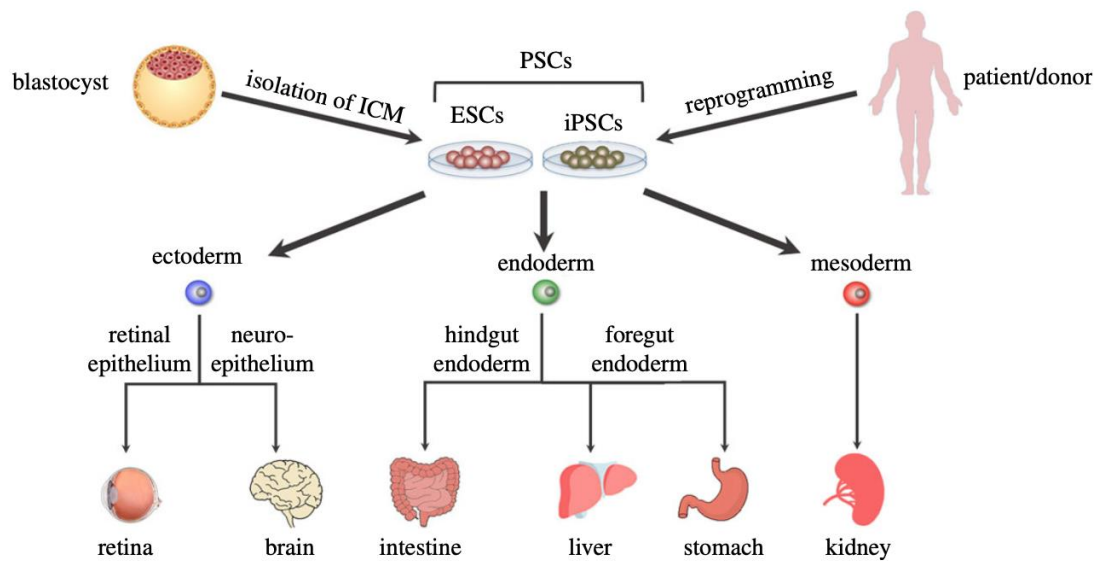
### **1.4.3. PSCs definition & properties**

The term PSCs is defined by the key properties of the cells in relation to “self-renewal” and “potency”. Self-renewal and potency refer to a cell’s ability to proliferate and differentiate into one of three primary germ layers, respectively. Three primary germ layers are classified into: ectoderm; endoderm and mesoderm (Wobus and Boheler, 2005). Three *in vivo* assays have been developed to assess the potency of PSCs in mouse models (Aoi, 2016). The “teratoma formation” assay is performed by transplantation of cells into immunocompromised mice to evaluate spontaneous generation of differentiated tissues containing cells representing all three germ layers. The “chimera formation” assay is performed by injecting stem cells into diploid early embryos (2N blastocysts) to evaluate their contribution in the developing embryo. Several endpoint assays were conducted to see if the chimeras generate functional gametes and retain chromosomal integrity whilst keeping pluripotency state. The “tetraploid (4N) complementation” assay evaluates the capacity of the tested PSCs within an entire organism. The PSCs can be injected into 4N embryos (4N blastocysts) and growth stages can be monitored for extra-embryonic lineages. These lineages are expected to originate from the transplanted stem cells and not the embryo itself (Aoi, 2016). Here, I will focus on ESCs and iPSCs and the characteristics of each cell types are described in more details.

### **1.4.4. Embryonic Stem Cells (ESCs)**

Human ESCs (hESCs) are harvested from the inner cell mass of early-stage blastocysts around 4-5 days post-fertilization (Thomson et al., 1998, Reubinoff et al., 2000). Human ESCs are the first stem cells applied in research applications,

particularly they are commonly used in different clinical trials (Figure 1.16) (Liu et al., 2020).



**Figure 1.16: Schematic of different types of PSCs. The Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of an embryo at the blastocyst stage. The hiPSCs are generated from somatic cells following reprogramming by main four Yamanaka factors (OSKM factors). These cells have the ability to form three main embryonic layers (ectoderm, endoderm and mesoderm) and differentiate to any tissue type. This picture is adapted from (Alhaque et al., 2018).**

### 1.4.5. Limitations of ESCs

The ethical controversy surrounding ESC research has restricted its use in research and clinical applications. This is because of the methods used to generate embryonic stem cell lines require destruction of living human embryos (Nordin et al., 2011). The current existing ESC lines for research studies are derived from surplus embryos generated for *in vitro* fertilisation (IVF) at blastocyst stage. This raises an ethical concern regarding the morality of destroying embryos based on the statement that human life commences when an egg is fertilised. Although only surplus embryos from *in vitro* IVF used, the moral objection still stands firmly (Nordin et al., 2011).

Another issue is the safety concerning ESC lines which pose a risk for therapeutic applications. The majority of ESCs are maintained in heterogenous cultures with irradiated mouse fibroblasts as a feeder cell layer. Therefore, such mixed cultures cannot be widely used for clinical studies (Martin et al., 2005). It is reported that most

human ESC lines undergo genetic and epigenetic changes. Thus, the genetic stability of ESCs is an important issue yet to be addressed. Additionally, due to the genetic instability of undifferentiated ESCs and possible malignant transformation, their application for therapeutic purposes are limited (Jensen et al., 2009). Research studies have shown that upon injection of a differentiated population of cells derived from ESCs, tumour formation may occur in normal or immunocompromised animals (Fong et al., 2010).

Finally, another important drawback of ESCs is the immunorejection complication. Immunorejection occurs when there is a mismatch between the donor and the recipient cells following transplantation. To overcome this problem, a global cell bank of different HLA typed stem cell lines has been established (Nakajima et al., 2007, Condic and Rao, 2008, Osafune et al., 2008). However, this approach still faces obstacles due to the limited number of available ESC lines. Having all the aforementioned reasons in mind, there is an urge to find alternative methods to generate ESC lines or non-embryonic sources of PSCs. As mentioned earlier, the outstanding research conducted by Japanese researchers to generate PSCs has received a lot of attention. In their study, adult somatic cells were used to induce four specific genes that are normally expressed in ESCs (Takahashi and Yamanaka, 2006). These stem cells were called induced pluripotent stem cells (iPSCs) and will be explained further in the following sections.

#### **1.4.6. Historical background of iPSCs research technology**

The historical background of iPSCs technology is classified into three main sections: first generation; second generation and human iPSCs which are explained in turn.

##### **1.4.6.1. First generation**

For the first time, in 2006, Yamanaka and his team at Kyoto University reported the successful reprogramming of adult mouse fibroblasts into iPSCs. This exciting and promising result set the foundation for this revolutionary technology. They employed a retrovirus to transduce mouse fibroblasts with the selected genes. The cells were collected, using antibiotic selection of Fbx15 positive cells (Takahashi and Yamanaka, 2006). Researchers proved that this combination of reprogramming factors work, albeit at modest level. However, this cell line exhibited DNA methylation errors and

failed to generate viable chimeras following injection into developing embryos (Takahashi and Yamanaka, 2006).

#### **1.4.6.2. Second generation**

In 2007, Yamanaka's group with two other independent researchers from Harvard, MIT and UCLA reported viable chimeras by reprogramming mouse fibroblasts into iPSCs (Wernig et al., 2007, Maherali et al., 2007). Similar to previous study (first generation), these cell lines were generated from mouse fibroblasts by retroviral mediation of the same four pluripotent factors (*Oct4, Sox2, Klf4* and *c-Myc*; abbreviated as; *OSKM factors*). However, a different marker was used for detection. Instead of *Fbx* marker, *Nanog* was used as a major element of cellular pluripotency which is a crucial gene in ESCs (Wernig et al., 2007).

#### **1.4.6.3. Human iPSCs**

Considering the important studies conducted in 2006 and 2007, another milestone for iPSCs research was set later in 2007. Yamanaka research group at Kyoto University and Thomson and Yu at the University of Wisconsin-Madison independently reported the successful generation of iPSCs in adult human cells (Yu et al., 2007, Takahashi et al., 2007). Yamanaka and colleagues further refined their protocols and successfully employed the same four genes with a retroviral mediated system. On the other hand, Thomson and Yu reported the use of two other alternative factors (*NANOG* and *LIN28*) for reprogramming, using a lentiviral system. These reported human ES cells expressed markers specific to ESCs and were able to differentiate into cell types of all three embryonic germ layers (Yu et al., 2007). In the next section, reprogramming strategies used to generate iPSCs from adult somatic cells will be discussed.

#### **1.4.7. Reprogramming strategies and pluripotency induction**

To generate iPSCs from adult somatic cells, reprogramming strategies are used. These methods reprogram the cells into an embryonic stem cell-like state (pluripotent state) and can be differentiated to any cell type through reversing epigenetic changes (Liu et al., 2020). There are two main methods for delivering reprogramming factors which are explained in turn.

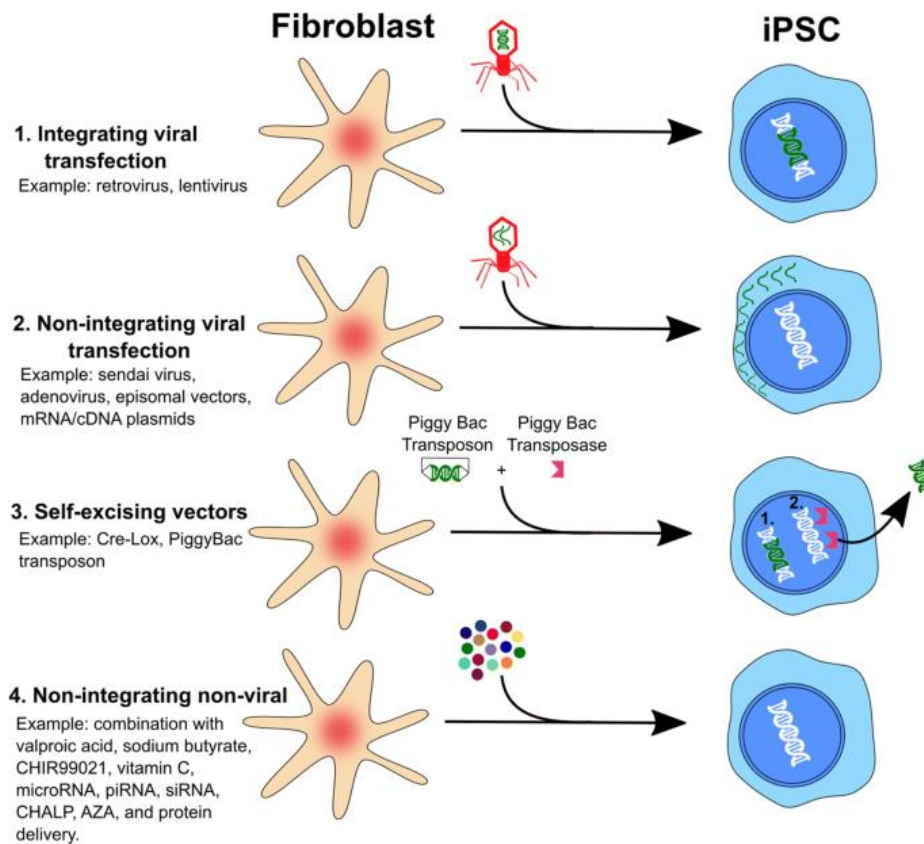
#### **1.4.7.1. Integrating viral systems**

Integrating viral systems such as ADV (Park et al., 2008), retrovirus (Wernig et al., 2007) and lentiviruses (Stadtfield et al., 2008) have been used traditionally to deliver transcription factors. However, as a result of incorporation of the genetic material and risk of teratoma formation, their applications are not favoured (Liu et al., 2020). Therefore, novel reprogramming methods have been developed to utilise non-integrating viral vectors for generation of iPSCs (Figure 1.15).

#### **1.4.7.2. Non-integrating viral systems**

Non-integrating viral systems were considered as an important step for advancing iPSC technology and for introduction of reprogramming factors into somatic cells. These systems are the most prominent non-integrating methods and are as follows: Sendai Virus (SeV), Episomal reprogramming (Epi) and mRNA transfection (Schlaeger et al., 2015). In the SeV reprogramming system, SeV particles are used to infect mature cells, using replication-competent RNA molecules encoding the main four OSKM reprogramming factors (Fusaki et al., 2009). In the Epi reprogramming system, there is an extended reprogramming factor expression, using Epstein-Barr virus-derived sequences. This allows the episomal plasmid DNA replication in dividing cells (Okita et al., 2011). In the mRNA reprogramming system, cells are transduced with *in vitro*-transcribed mRNAs encoding four OSKM reprogramming genes in addition to Lin28A mRNAs (Warren et al., 2010). It is reported that all three methods produced high quality hiPSCs; however, reprogramming efficiency and the suitability for clinical translational work remains challenging. In comparison to other systems, SeV reprogramming is highly effective, less labour intensive and most importantly, viral sequences do not appear at higher passages (Liu et al., 2020). In terms of safety, Epi reprogramming is considered as the most suitable choice for generation of iPSCs. This system is integration-free, highly efficient and requires less reagents such as plasmid DNA to generate iPSC line. Additionally, plasmid DNA can be produced readily with Current Good Manufacturing Practice (cGMP) protocol and therefore it is an ideal choice for clinical purposes (Schlaeger et al., 2015). Advances have been made and new methods were developed which are including non-integrating viral vectors, self-excising vectors and non-integrating non-viral vectors (Figure 1.17). These new methods provided significant advances in the safety and efficacy of stem

cell generation which can be applied for downstream scientific and clinical applications. In the next section, the main properties of iPSCs will be discussed.



**Figure 1.17: The four different key techniques for delivering reprogramming factors. Integrating viral transfer systems were the first applied to deliver transcription factors to generate induced pluripotent stem cells. This method was not ideal due to the incorporation of viral genetic material and contributing to teratoma formation. New methods were developed including non-integrating viral vectors, self-excising vectors and non-integrating non-viral vectors. These new methods provided significant advances in the safety and efficacy of stem cell generation which can be applied for downstream scientific and clinical applications. (Liu et al., 2020).**

### 1.4.8. Properties of iPSCs

As mentioned earlier, different methods are used to generate iPSCs from various cellular sources. The characteristics of the established iPSC lines are very much similar to naturally derived PSCs, such as ESCs (Nordin et al., 2011). Here, I will discuss the similarities between these cell lines in terms of cellular morphology, genetic profiles and differentiation potential (pluripotency state).



#### **1.4.8.1. Morphological features**

The morphology of human and mouse iPSC lines are very similar to ESCs regardless of any reprogramming methods used. The cells display compact structures, defined borders with high nucleus to cytoplasmic ratios and also a prominent nucleoli (Nordin et al., 2011). Additionally, iPSCs derived from patients with diseases such as Down syndrome, Parkinson disease and type 1 diabetes mellitus also exhibited similar morphology (Park et al., 2008).

#### **1.4.8.2. Genetic profiles**

One aspect in generating high quality iPSCs is the stable genomic integrity of mouse and human iPSC lines. Genomic stability is particularly important for any clinical applications and any changes may lead to development of certain diseases (Nordin et al., 2011). Despite karyotype stability for established mouse (Wernig et al., 2007) and human iPSC lines (Yu et al., 2007, Takahashi et al., 2007), few abnormalities were also detected (Amabile and Meissner, 2009). Nonetheless, a normal karyotype for mouse (40) and human (46) chromosomes were reported in majority of the iPSC lines (Nordin et al., 2011). A group of researchers in Spain reported that prolonged culture of human iPSCs affected the genetic stability and subsequently resulted in chromosomal abnormalities (Aasen et al., 2008). Therefore, it is of utmost importance to monitor the cells regularly for any modifications in genetic stability in culture condition.

The reactivation of telomerase reverse transcriptase (Tert) and also telomerase activity is an important similarity that exists between iPSCs and ESCs. Few studies suggested that iPSC lines maintain telomerase length. This feature enables the cells to undergo unlimited cell proliferation *in vitro* (Takahashi et al., 2007, Stadtfeld et al., 2008). It is also reported that population doubling time of human iPSCs were similar to the reported doubling time of human ESC lines (Takahashi et al., 2007). A single-cell survival assay by Zhou and colleagues, demonstrated that iPSCs can be clonally expanded. The results of this study confirmed that iPSC lines are clonogenic, similar to ESC (Zhou et al., 2009).

PSC lines express unique surface marker genes for pluripotency. By employing different techniques such as RT-PCR and immunofluorescence, established iPSCs important stage-specific embryonic antigens such as (SSEA-1 for mouse; SSEA-3 and

SSEA-4 for human), tumour-related antigen (TRA-1-60; TRA-1-81) and alkaline phosphatase (AP) can be detected. Moreover, these cell lines express several undifferentiated state markers such as Octamer-binding transcription factor- 4 (Oct4), Sex determining region Y-box 2 (Sox2), Nanog, growth and differentiation factor 3 (GDF3) and TERT (Takahashi et al., 2007, Nakagawa et al., 2008, Huangfu et al., 2008, Chambers et al., 2009).

To date, not so many studies have studied the global gene expression patterns in iPSCs and ESCs. The results from DNA microarrays revealed that mouse and human iPSC lines have similar, but not identical, gene-expression patterns to their corresponding ESCs (Wernig et al., 2007, Maherali et al., 2007). Therefore, it is important to have robust assessments in place to test the genetic stability of newly established iPSC lines. Further, efficient culture conditions are required to enhance the genetic stability and maintain the pluripotent state of the cells prior to differentiation.

#### **1.4.8.3. Differentiation potential**

The pluripotent state of the iPSCs is also an important factor which enable the cells to differentiate into the cells/tissues representative of the three embryonic germ layers, the ectoderm, endoderm and mesoderm. To determine the pluripotency state of the cells, established human and mouse iPSCs lines are differentiated via formation of 3D aggregates called embryoid bodies (EBs) or EB assay. Different types of cells can be identified within the EBs with a close morphological resemblance to neuronal cells, cobblestone-like cells and in some cases epithelial cells. Immunofluorescent results revealed that cells were positive for beta-tubulin class III (a marker for ectoderm), alpha-smooth muscle actin (alpha-SMA for mesoderm). Therefore, this assay is a safe test to assess the pluripotent state of the cells and the feasibility of formation of the three primary germ layers. Further, once the cells are differentiated, the expression of pluripotency associated genes will be decreased (Takahashi et al., 2007, Nordin et al., 2011).

There is no doubt that established mouse and human iPSCs are very much similar to their respective ESCs as previously explained. It is reported that iPSCs have been successfully differentiated into a number of different cell types including neurons, hepatocyte-like cells, and hematopoietic cells (Chambers et al., 2009, Song et al.,

2009, Choi et al., 2009). This will further confirm that iPSCs has a great potential in research and therapeutic applications without the controversial ethical issue of ESCs. To generate pluripotent iPSC lines, adult somatic sources are required which is explained further in the next section.

#### **1.4.9. Available sources of adult somatic cells for reprogramming**

Adult somatic tissues are used to generate iPSCs and hiPSC sources, such as, blood, skin and urine. Moreover, as hiPSCs are harvested from individual patients, risk of immune rejection can be diminished when the cells are transplanted autologously (self-donor). Therefore, hiPSCs have a great potential for personalised medicine and clinical applications. Various sources of iPSCs are available. Generally, any mature cell type in the human body such as umbilical cord blood cells, bone marrow cells, fibroblasts, peripheral blood cells, keratinocytes and cells in urine can be reprogrammed into iPSCs and differentiated into desired tissue types (Felfly and Haddad, 2014, Singh et al., 2015, Park et al., 2015). It is worth pointing out that umbilical cord blood or bone marrow stem cells are considered as “ready-to-use” sources. They can be used directly for transplantation without the need for reprogramming step. Non-autologous (i.e., non-self) stem cells is associated with an inherent risk of immune rejection which is not desirable. Thus, available tissues for autologous stem cell harvesting such as hair, skin and urine are used. Urine is a non-invasive stem cell source which is highly beneficial in fragile patients or in patients who have suffered from traumatic events such as heart attack or spinal cord injury (SCI). Although urine is considered as a stem cell source, it has not yet received substantial research and attention. However, it provides a promising source to generate iPSCs (Schosserer et al., 2015). Therefore, for future generation of iPSCs and clinical applications, non-invasive, simple, accessible mature somatic cell sources are required.

#### **1.4.10. Maintenance of iPSCs in feeder-free and *xeno*-free culture environments**

##### **1.4.10.1. Feeder-free cell culture**

For clinical translation and maintenance of pluripotency, it is imperative to culture the cells in feeder-free conditions. Traditionally, Thomson developed the gold-standard self-renewal culture technique for maintenance of PSCs (Thomson et al., 1998). In

this technique, the established mouse or human iPSCs were placed on a monolayer of feeder-cells, such as primary mitotically inactivated mouse embryonic fibroblasts (MEFs) (Thomson et al., 1998). Contaminating cultures with unknown proteins or zoonotic viruses are the main risks of using animal products with human cells (Llames et al., 2015, Jung and Kim, 2015). Also, secretion of growth factors from MEF feeder layers are quite inconsistent which results in batch-to-batch variation. Moreover, the anti-proliferation treatments may affect PSCs in long-term and may lead to apoptosis (Villa-Diaz et al., 2009). Over the years, using animal products for cell culture maintenance phased out and replaced with new techniques for long-term culture of human PSCs. The scalable, cost-effective and stable poly (acrylamide-co-propargyl acrylamide)-coated polystyrene flasks mixed with other polymers were used (Lambhead et al., 2018). Further, laminin-rich extracellular matrix (ECM) proteins such as Matrigel<sup>®</sup> which is derived from mouse osteocarcinoma cell line has been used to improve cell attachment and maintenance in culture (Stover and Schwartz, 2011). The novel matrices, such as CellStart, recombinant proteins including various isoforms of laminins and synthetic polymers are currently used to culture human PSCs (Ausubel et al., 2011, Chen et al., 2011). It is reported that recombinant human Laminin-521 stabilises the pluripotent state of hESCs (Albalushi et al., 2018). Apart from Matrigel<sup>®</sup>, other matrices are free from animal-derived products, and they are preferred ECM for clinical-grade cell culture applications.

#### **1.4.10.2. xeno-free culture medium**

The iPSC culture media has an important role in maintaining pluripotency state and improving stem cell differentiation capacity. Thus, culture media should be very well defined, xeno-free (i.e. free from animal products) and serum-free (Jung et al., 2012). Several studies demonstrated successful PSC culturing under xeno-free conditions. It was reported that for induction of rostral hypothalamic-like progenitor cells from neuroectoderm-derived mouse ESCs, a chemically defined media with no growth-factors were essential (Suga, 2016). Therefore, for propagation, pluripotency and improving differentiation of PSCs, it is crucial to culture the cells in feeder-free and xeno-free culture environments.

#### **1.4.10.3. *In vitro* growth and maintenance of PSCs**

For successful pluripotent stem cell culture, the indefinite self-renewal properties should be maintained. Originally MEFs were employed to support propagation of

hESCs in the undifferentiated stage (Reubinoff et al., 2000). Efforts have been made to develop a xeno-free hESC culture system, using human-derived cell types including feeder cells derived from muscle (Richards et al., 2002), foetal foreskin (Richards et al., 2002) and bone marrow (Cheng et al., 2003).

Recent advances have been made to develop fully defined culture systems for clinical and therapeutic applications. Xu and colleagues established a feeder-free culture condition using Matrigel™ (Xu et al., 2001). This matrix is derived from Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells. This matrix can be used in combination with a culture media conditioned by MEF and supplemented with basic fibroblast growth factor (bFGF or FGF2). Since then, efforts have been made to maintain hESCs pluripotency on recombinant proteins including vitronectin (Braam et al., 2008), laminin (Rodin et al., 2010), laminin 521 and E-cadherin (Rodin et al., 2014). This resulted in significant improvement in pluripotent cultures.

Following culture substrate optimisation, few research groups have focused on identifying the essential factors secreted by the feeder layers. These factors were responsible for the maintenance of the hESC pluripotent state and self-renewal properties. Overall, it was shown that the following proteins are crucial for the maintenance and undifferentiated proliferation of hESCs in serum-free media. These were including, high concentrations of bFGF, Activin A, TGFβ and the BMP repressor Noggin (Pera et al., 2004). Further, companies have released to the market defined media to standardise the hESC culture techniques. The StemPro® (Invitrogen) and mTeSR™ (Stem Cell Technologies) culture media are proved to be the most effective to support PSCs (Hannoun et al., 2010). Also, the new fully-defined E8™ media has been shown to support efficient self-renewal and differentiation of PSCs (Chen et al., 2011).

A research study conducted by the International Stem Cell Initiative Consortium which different defined culture systems for feeder cell-free propagation of embryonic stem cells were compared. For this study, eight different cell culture methods were used, and cells were propagated in the presence of Knockout Serum Replacement, FGF2. Further, MEF was used as a positive control. The cultures were analysed for up to 10 passages for attachment, death, and pluripotency (International Stem Cell Initiative et al., 2010). Of the eight culture systems, only the positive control and the commercial culture media mTeSR1 (Ludwig et al., 2006) and STEMPRO (Wang et al., 2007),

supported the long-term maintenance of most cell lines. Cultures which maintained on other culture media failed to support cells due to lack of attachment, cell death or presence of differentiation. It was concluded that the complex combination of growth factors, better development and quality control were the reasons for success of commercial formulations. A summary of defined culture media and their compositions are summarised in Table 1.1.

#### **1.4.11. Applications and future directions of iPSCs**

Development of iPSC technology has greatly deepened our understanding and also paved the way further for numerous technological advancements in stem cell field. In previous subsections, history and properties of these cells were explained. In this section, I discuss different applications, challenges and future directions of iPSC lines. For generation of patient-specific iPSC line, patient's own stem cells can be used. This unique feature has revolutionised regenerative medicine for clinical therapeutics. This appears to be a crucial step forward in developing personalised medicine by generating patient's own tissues. The main applications of iPSC technology are in pharmaco-toxicological screening, disease modelling, autologous cell transplantation and gene therapy (Nordin et al., 2011). Screening tests based on human cells can eliminate existing variations among different animal and animal-based *in vitro* models and can predict adverse effects more precisely. Patient-specific iPSC lines are also a great alternative for drug toxicity screening, personalised medicine and gene therapy applications.

To conclude, cell lines such as ESCs and iPSCs can be collected and generated from different sources. In the next section, development of 3D cell technology systems and their advantages over two-dimensional (2D) monolayer conventional methods are explained.

**Table 1.1: Summary of defined culture media and compositions for maintenance of pluripotent stem cells.**

<b>Medium</b>	<b>Composition</b>	<b>Advantages</b>	<b>Disadvantages</b>	<b>Ref</b>
<b>1</b>	XVIV0-10, NEAA, L-Glutamine-Beta-mercaptoethanol-hbFGF-hFLT3	No cell death	Low attachment, Spontaneous differentiation	(Li et al., 2005)
<b>2</b>	DMEM/F12-N2-B27-L-Glutamine-Beta-mercaptoethanol-hbFGF	Low spontaneous differentiation	Arrest/Death Low attachment	(Liu et al., 2006)
<b>3</b>	IMDM-F12-L-Glutamine-Beta-mercaptoethanol-Insulin-Transferrin-Monothioglycerol-BSA-Activin A-hbFGF		Arrest/Death Low attachment spontaneous differentiation	(Vallier et al., 2005)
<b>4</b>	DMEM/F12-L-Glutamine-Beta-mercaptoethanol-Insulin-Transferrin-Cholesterol-ALB-hbFGF-WNT3A-BAFF	Low spontaneous differentiation	Death Low attachment Variable factor concentrations	(Lu et al., 2006)
<b>5</b>	DMEM/F12-N2-B27-NEAA-L-Glutamine-Beta-mercaptoethanol-BSA Fraction V- hbFGF	Limited support	Arrest/Death, low attachment, spontaneous differentiation	(Yao et al., 2006)
<b>hESF9</b>	N/A	Very limited support	No attachment Spontaneous Differentiation	(Furue et al., 2008)
<b>mTesr1™</b>	N/A	Cell Growth Attachment Stimulation of the FGF and TGF-beta pathway via FGF2 Maintenance of pluripotency Stable karyotype Robust differentiation	Low cell death High ALB level	(Ludwig et al., 2006)

<b>Medium</b>	<b>Composition</b>	<b>Advantages</b>	<b>Disadvantages</b>	<b>Ref</b>
<b>STEMPRO®</b>	N/A	Cell Growth, attachment, stimulation of the FGF and TGF-beta pathway via FGF and Activin A, maintenance of pluripotency	Low cell death	(Wang et al., 2007)
<b>TeSR-E8™</b>	N/A	Cell Growth, attachment, maintenance of pluripotency, lack ALB, simple media, low protein	Low cell death Spontaneous differentiation	(Chen et al., 2011)
<b>Essential 8™</b>	N/A	Cell Growth, maintenance of pluripotency, long-term propagation, stable Karyotype, robust differentiation	Low cell death	(Badenes et al., 2016)



#### **1.4.12. Generation of *in vitro* hepatocyte-like cells from PSCs**

Different sources of hepatocytes are available for research studies. These are including primary hepatocytes, hepatocyte cell lines and non-human-derived hepatocytes. Despite the extensive use of these sources in research, there are some disadvantages associated which limited their use in cell applications (Hewitt et al., 2007). Primary hepatocytes are considered as the current “gold standard”. However, it displays an instable phenotype. Further, the scarcity and the quality variation of the available primary human hepatocytes have limited their application in research. Hepatocarcinoma derived cell lines and immortalised human hepatocytes have distinct functional properties. These cells are highly instable and possess an incomplete expression of hepatocyte-specific functions (Wong et al., 2000). Non-human hepatocytes are phenotypically variable. Further, the risk of xeno-contaminants and limited resemblance to the human hepatocyte biology restricted their use as reliable replacements (Behnia et al., 2000).

Therefore, it is of utmost importance to generate reliable *in vitro* hepatocyte models. For this purpose, the self-renewal capacity and pluripotent nature of PSCs in combination with robust hepatocyte differentiation protocols make them an ideal resource of human hepatocytes. In addition, our understanding from reprogramming somatic cells into PSCs enable scientists to directly transdifferentiate somatic cells into another cell type. Thus, somatic cells can become a potential source of hepatocytes for downstream applications.

Research studies in the mechanism of the liver development in addition to the recent advances in hepatocyte differentiation protocols have led to generation of hepatocyte-like cells (HLCs) from PSCs and somatic cells. These protocols mimic the patterns and key steps observed during embryonic development. The protocols are mainly based on three main strategies, either via cellular aggregation or differentiation in a 2D monolayer or 3D systems.

##### **1.4.12.1. Direct differentiation of hESC into hepatocyte-like cells (HLC)**

There are existing number of differentiation protocols for generation of functional HLCs from hESCs. These protocols are conducted in a stepwise manner, using animal-derived matrices including collagen, gelatin or Matrigel™. As mentioned earlier, majority of these protocols mimic the hepatogenesis process during embryonic

development. This is composed of different steps, including hepatic induction, priming of the hESC into definitive endoderm and a final stage of hepatocyte maturation. The resulting mature HLCs can be characterised using a number of different assays, including gene and protein expression, cytochrome P450 activity, serum protein production, urea production and glycogen storage.

#### **1.4.12.2. Endoderm Specification**

For endoderm differentiation, it is crucial to recapitulate the processes that occur during embryonic development *in vivo*. An efficient differentiation protocol was developed to differentiate hESCs into definitive endoderm, using high concentrations of Activin A (D'Amour et al., 2005). It is argued that Activin A mimics nodal signalling pathway which is a key member of the TGF $\beta$  superfamily. Activin A plays an important role in the initiation of the formation of endoderm during gastrulation stage (Green and Smith, 1990). Exposure to Activin A, resulted in over 80% of the cells expressing the endoderm marker Sox17 and FoxA2 (Cirillo et al., 2002). Another group studied the exposure to other factors including HGF, BMP4, FGF4 and all-trans-retinoic acid (ATRA) on several ECM components. It was concluded that sequential addition of Activin A and HGF on Matrigel coated plates resulted in higher level of AFP positive cells (Ishii et al., 2008).

It is well studied that WNT signalling pathway is critical during the onset of the hepatic development. Few groups studied the role of Wnt and beta-catenin signalling during the differentiation and proliferation of endoderm cells. These studies have led to significant advances in the endoderm specification from hESCs (Burke et al., 2006, Fletcher et al., 2008). Hay and colleagues extensively studied the expression of Wnt signalling during liver development. They reported the significance of Wnt3a signalling in Primitive Streak (PS) and endoderm development. During the first trimester and the expression of Wnt3a during liver development is limited to the portal system. However, in the second trimester Wnt3a expression can be detected in the liver parenchyma, suggesting its crucial role in early hepatogenesis. Therefore, they concluded that treatment with Activin A in combination with Wnt3a have led to a rapid and robust endodermal differentiation which ultimately generate more functional HLCs (Hay et al., 2008a).

#### **1.4.12.3. Hepatic Specification and differentiation**

Several groups developed differentiation protocols for hepatic specification. One group reported that following priming hESCs into definitive endoderm with Activin A, the cells were treated with FGF4 and BMP2 to generate hepatic endoderm. Following this step, the resulting endoderm cells were differentiated into mature HLCs. The hepatic endoderm was primed to hepatocyte fate by using a specific media supplemented with HGF, followed by a hepatocyte maturation step induced by OSM and dexamethasone. The resulting cells exhibited mature adult liver markers including CYP7A1, CYP3A4 and CYP2B6, tryptophan oxygenase 2, tyrosine aminotransferase and phosphoenolpyruvate-carboxykinase (PEPCK). Further, functional properties of the cells such as ALB secretion, glycogen storage and cytochrome P450 activity were also analysed. They also reported that these cells were prone to hepatitis C virus HCV infection (Cai et al., 2007).

Agarwal and colleagues suggested FGF4 and HGF growth factors to promote hepatic specification of hESCs-derived endoderm cells. Hepatocyte differentiation and maturation of the endoderm was conducted using a mixture of BSA, FGF4, HGF, OSM and dexamethasone. This resulted in the HLCs expressing a number of hepatocyte markers including ALB, AFP, CYP7A1 and CYP3A4. Other hepatocyte functions including glycogen storage and ALB secretion were also detected. Interestingly, these cells were injected in a mouse with liver injury. Following transplantation, expansion and repopulation of the liver was observed (Agarwal et al., 2008).

Brolen and colleagues applied Activin A and FGF2 for induction of definitive endoderm. For hepatic specification a cocktail mixture of BMP2 and 4 with FGF1, 2 and 4 and for hepatocyte maturation a mixture of different factors including EGF, insulin, transferrin, ascorbic acid, FGF4, HGF, dexamethasone, DMSO and OSM were applied, respectively to differentiate three different hESC lines. This resulted in generation of HLCs exhibiting hepatic functional characteristics including, urea secretion, cytochrome P450 activity and glycogen storage (Brolen et al., 2010).

Hay and colleagues conducted hepatic specification from hESC-derived endodermal cells, using a specific media supplemented with DMSO. This reagent is an organosulfur compound that promotes the acetylation of the histones. Hepatocyte differentiation and maturation was performed using a serum containing media

supplemented with HGF and OSM. The resulting HLCs possessing biliary duct-like structures, parenchymal hepatocyte markers and myofibroblast markers. Following transplantation in a mouse with damaged liver, cells were repopulated the damaged liver (Hay et al., 2008a). Further, the transplanted cells secreted human serum ALB three months following transplantation (Payne et al., 2011).

#### **1.4.12.4. Hepatic differentiation of iPSCs**

The self-renewal capacity and pluripotent nature of iPSCs make them an ideal source for generation of HLCs. Further, it is feasible to obtain autologous iPSCs which make them a desirable choice for disease modelling and transplantation. Several hepatocyte differentiation protocols of hESCs have been explored and successfully fine-tuned to obtain HLCs from iPSCs.

Hay and colleagues established a robust differentiation protocol which successfully translated into iPSC technology. The resulted HLCs were characterised and exhibited cytochrome P450 activity and expressed liver-specific proteins including ALB, fibronectin and AFP (Sullivan et al., 2010).

Song and colleagues employed a multiphase differentiation protocol to obtain HLCs from iPSCs. The resulted HLCs displayed liver-specific markers and functional proteins including urea production, ALB secretion and cytochrome P450 activity (Song et al., 2009).

Despite recent advances in 2D hepatocyte differentiation protocols, lack of phenotypic stability and hepatic functions have remained the major limitations. To overcome this hurdle, researchers have developed strategies to generate 3D *in vitro* platforms including spheroids and liver organoid models in static culture or bioreactors.

## **1.5. 3D cell culture system**

### **1.5.1. Overview**

Over the years, culturing cells in 2D format has remained the popular method of choice for *in vitro* cell growth and expansion. However, these systems fail to effectively recapitulate the spatial requirement which are critical for cellular organisation similar to the *in vivo* environment (Prior et al., 2019). In this section, I will discuss the advantages of 3D cell technologies in comparison to conventional 2D monolayer

cultures. Secondly, in general organoids and particularly 3D liver organoids will be explained in greater detail.

### **1.5.2. Two dimensional (2D) versus Three dimensional (3D) systems**

The method of culturing cells and testing their functionalities has remained a debatable discussion in the field of cell biology. A recent study shows that murine ESCs can be cultured in stirred microcarrier cultures, allowing large and robust stem cell expansion (Marinho et al., 2010). It was speculated that due to the uniform distribution of nutrition through medium, the cells had higher chance of proliferation and survival (Marinho et al., 2010). Contrary to 2D monolayer, superior cell-to-cell contact and cell-matrix interaction can be achieved under 3D culture system. These improved interactions are crucial for cell survival, proliferation and functioning (Kraehenbuehl et al., 2011, Han et al., 2014). Therefore, 3D cell culture has been more favoured than conventional 2D monolayer cultures.

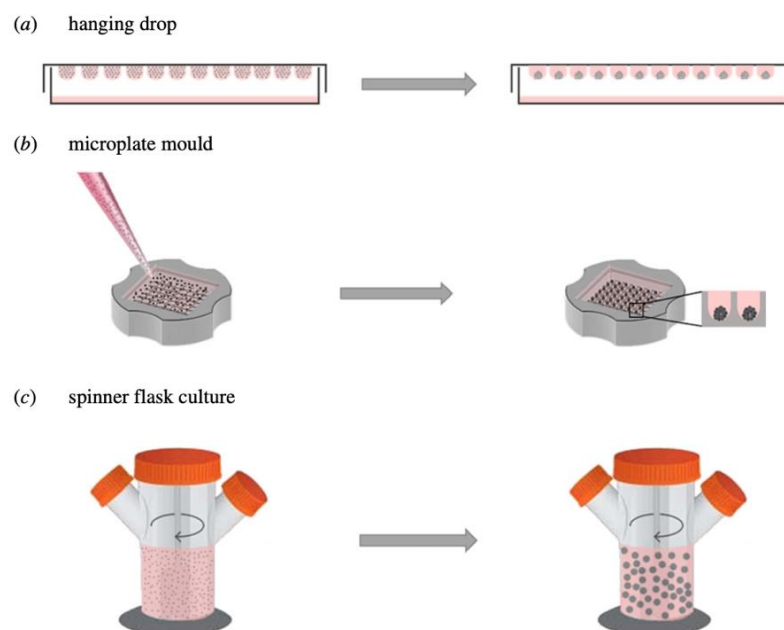
### **1.5.3. Scaffold-free 3D methods**

To overcome 2D cell culture limitations, novel approaches have been developed to recapitulate physiological conditions more closely. Efforts have been made to develop scaffold-based and scaffold-free 3D platforms. There are several techniques to generate scaffold-free 3D cellular aggregates. The “Hanging-drop” technique was the first method developed to generate 3D structures, using suspended droplets of dissociated cells (Alhaque et al., 2018). The “liquid overlay” method (Carlsson and Yuhas, 1984) was developed to produce 3D microtissues on non-adherent surfaces, resulting in formation of larger spherical-shaped cell masses. Other approaches such as “micromoulds” partially overcame the limitations commonly seen with the previous models (Napolitano et al., 2007). As the field progressed, 3D “organoid” cultures (Sato et al., 2009) have evolved as an alternative *in vitro* system. This system enables scientists recapitulating the differentiation of target tissues *in vitro* in a dish (Figure 1.16). In the next section, the key aspects of organoid cultures and recent advances in the development of 3D liver organoids will be discussed further.

### 1.5.4. What are 3D organoids?

The term “organoid” is defined as an *in vitro* 3D cell clusters which are generated from tissue-resident stem cells. The 3D cellular clusters have the ability for self-renewal and self-organisation and recapitulates the functionality of the tissue of origin (Lancaster and Knoblich, 2014). In the body, cells are located in complex microenvironments and are exposed to different factors including several signalling interactions, mechanical cues and the extra cellular matrix (ECM). These interactions play a key role in maintaining and regulating cellular phenotypes and functions. It is now strongly argued that cells cultured in 3D platforms closely resemble physiologically relevant *in vivo* tissues and cellular responses (Figure 1.18) (Baker and Chen, 2012).

The current understanding of liver biology and related diseases has been obtained, using animal studies and the established cell lines. However, existing 2D models lack the complex architecture and metabolic functions of the liver and the limited availability of human sample resources. The recent developments in organoid culture systems have advanced our understanding in human biology (Prior et al., 2019). Advancements in 3D cell cultures has led to the development of organoid systems. An understanding of ECM biology and the interaction of different signalling pathways at pluripotent and differentiation stages, is crucial to mimic the natural *in vivo* environment.



**Figure 1.18: Techniques for generating 3D microtissues. (A) The hanging drop method was the first technique used for generating and maintaining 3D**

structures. (B) The 3D micromoulds overcame limitations such as culture media restriction and size heterogeneity of spheroids seen with previous methods. (C) The spinner flask technology developed for large-scale production and maintenance of 3D structures (Alhaque et al., 2018).

### **1.5.5. iPSC-derived 3D Liver Organoid Models**

#### **1.5.5.1. Comparison of different protocols for generation of 3D iPSC-derived hepatic organoids**

Takebe and colleagues were the first to establish hepatic organoids from iPSCs, using an elegant co-culture methodology. In this model, hepatic progenitors were generated by stepwise differentiation in a 2D monolayer culture setting. Hepatic progenitors were then co-cultured with human mesenchymal stem cell (MSCs) and human umbilical vein endothelial cells (HUVECs). Following this step, 3D aggregates were called iPSC-liver buds (LB) and were continuously generated in a Matrigel-embedded culture. Upon transplantation of iPSC-derived LBs, human vasculature surrounding the LBs became functional. Some level of liver-specific proteins such as ALB was detected from the hepatic cells in the engrafted liver buds into the bloodstream of the recipient mouse from day 10 till day 45 post-transplantation (Takebe et al., 2013). Since then, efforts have been made by different research groups to establish robust protocols to generate liver organoids derived from pluripotent stem cells.

Guan and colleagues have successfully established a protocol to generate iPSC-derived hepatic organoids (Hep-org). These Hep-org were surrounded by cholangiocyte ductal structures, within about 50-60 days. Although Hep-org exhibited extended stable phenotype in culture, limited diffusion of oxygen and nutrients in large organoids proved to be the main challenge leading to reduction in proliferation and regenerative capacity of Hep-orgs. The issue was resolved by dissociating the organoids into single cells and subsequent replating and embedding them in Matrigel (Guan et al., 2017).

Rashidi and colleagues employed a robust protocol to generate self-aggregated 3D cellular spheroids from iPSCs, using the 3D Petri Dish® mould. The advantage of using this method compared to previous protocols was to exclude Matrigel which is commonly used to embed organoids. The resulted 3D spheroids exhibited hepatic cellular organisation correlating with stable phenotype, reduced proliferation and loss

of foetal protein secretion. Interestingly, 3D hepatic spheroids were maintained with modest liver function for more than a year in culture. Further, hepatic spheroids were transplanted and supported failing liver function *in vivo* (Rashidi et al., 2018).

Another study was conducted by Wu and colleagues for generation of iPSC-derived organoids with a hepato-biliary structure. Following characterisation, these organoids exhibited hepatic gene expression signatures and other key functional properties of cholangiocytes *in vitro*. Once transplanted the cells were survived for more than 8 weeks in immune-deficient mice (Wu et al., 2019).

A rapid and robust protocol was developed by Akbari and colleagues to generate human hepatic organoids from iPSCs. In this study, Epithelial Cell Adhesion Molecule (EpCAM)-positive cells were enriched and resulted in a homogenous population of endodermal cells which led to functional hepatocytes. Further, the recombinant human R-spondin 1 (RSPO1) was used to improve specification of EpCAM-positive endoderm cells. One of the advantages of this protocol was the endoderm-derived hepatic organoids (eHEPOs) can be produced in 14 days and can be maintained in the culture for more than 1 year without significant loss in culturing efficiency. Further characterisation exhibited eHEPOs can adapt epithelial morphology and a pseudostratified structure. Also, the cellular composition and the morphological structure of eHEPOs can be preserved and maintained in young and old organoids. In addition, organoids displayed functional characteristics of mature hepatocytes *in vitro*. Following transplantation in the mouse liver, human ALB was secreted at day 32 post intrasplenic injection (Akbari et al., 2019b).

More recently, a breakthrough method was established to develop an integral multi-organ structure. For this protocol, iPSCs were separately differentiated into anterior and the posterior gut spheroids. Following this step, these spheroids were allowed to merge together as an interconnected multi organ. This approach was very promising and demonstrated the possibility to recapitulate the foregut-midgut boundary *in vitro*, regardless of extrinsic factors (Koike et al., 2019).

In summary, organoids are considered as one of the most optimised 3D cultures of cells and organ fragments. It is possible to establish physiologically relevant models of different tissue types *in vitro*. These rapidly evolving models can provide different research and clinical translational applications such as disease modelling and drug



screening. Further, with a more defined ECM, it is predicted that a highly robust and reproducible culture models can be achieved and pave the way from bench to bedside. Liver organoids have shown to be the most versatile next generation cell culture system in modelling various liver diseases and optimising patient specific therapeutic strategies (Figure 1.19). The advantages and disadvantages of key iPSC derived 3D hepatic organoid protocols are summarised in Table 1.2.

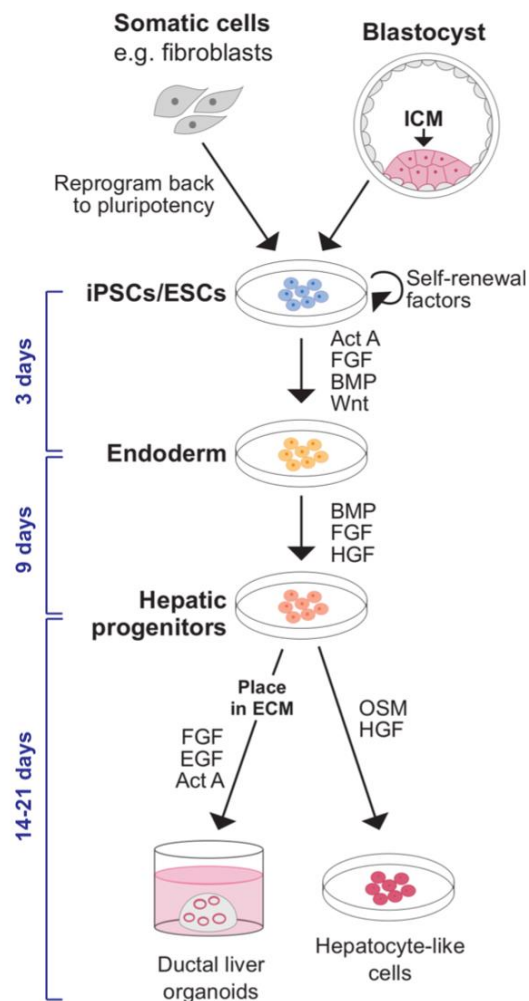
#### **1.5.5.2. Limitations of 3D *in vitro* models**

Pluripotent stem cell-derived 3D models have been developed as more physiologically relevant systems to human than animal models. With this approach personalised medicine can be applied, using patient-derived organoids. Despite holding unique advantages, they have remained underdeveloped. Efforts have been made to advance the technology even further to meet the necessary requirements. One of the most pressing obstacles in 3D technology is the variability between research groups. Each research laboratory has demonstrated contrasting protocols to generate 3D organoid systems from stem cells. However, a unique, widely accepted, and standardised approach is yet to be established. This issue is very important to reduce the variability in protocols. A joint effort must be made to set clear strategies and guidelines to evaluate the quality of final product. Single-cell profiling technologies such as transcriptome and epigenome analysis are very promising. These are considered as highly accurate assays which can be used to evaluate every cell type present in the organoids and compare it to *in vivo* data. Some factors such as age and patient genetic background might lead to further variation in organoid systems (Kim et al., 2020).

Other factors such as modelling of cell-to-cell communication with stromal cell and the development of vasculature in organoid systems yet to be fully established. Despite limited success in this area (Takebe et al., 2013, Takebe et al., 2015), vascularisation remains a difficult hurdle to overcome. Organoid co-culture systems such as mesenchymal and/or immune cell populations are exciting approaches. However, they often suffer from complications. It is reported that these systems contain pathogens such as viruses, bacteria, and parasites (Kim et al., 2019, Bar-Ephraim et al., 2020). Overall, organoid systems exhibit a large degree of complexity. One can appreciate the challenge of making this system more complex by adding additional components. Therefore, the best approach would be to select the most appropriate level of

complexity for a given study. For example, cancer organoids with epithelial cells are beneficial to assess the efficacy of most cancer drugs (Kim et al., 2020).

**Figure 1.19: Liver organoids can be produced from PSCs by regulating signalling pathways during differentiation. Liver organoids can be generated from both hESCs and hiPSCs, using three-stage differentiation protocols which recapitulate the key stages of in vivo embryonic development. The endodermal fate is initiated by exposing cells to Act A and Wnt3a growth factors. The endoderm cells gradually transit to a hepatic fate by induction of HGF and FGF signalling. At the end stage of differentiation, HLCs are formed in response to OSM signalling. Abbreviation: Act A; Activin A, HGF; hepatocyte growth factor, FGF; fibroblast growth factor, OSM; Oncostatin M. This diagram is adapted from (Prior et al., 2019).**

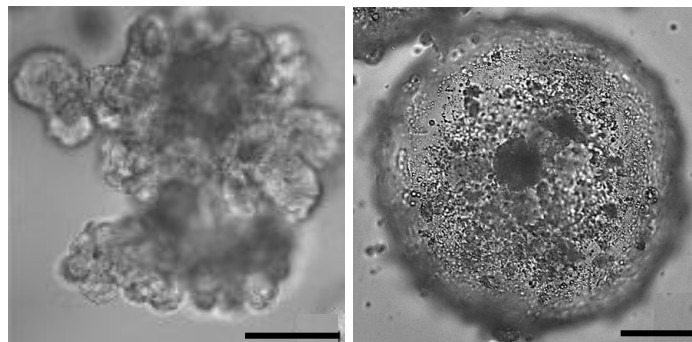


Another clear downside with 3D organoid systems is the lack of interorgan connection. 3D human organoids recapitulate part of the human body and not the whole body. Therefore, it is difficult to recapitulate organ-specific or tissue-specific microenvironment. Over the years, efforts have been made to overcome this issue. For instance, multiple organoids have been fused together in order to study the possible communication between the liver, gastrointestinal tract and pancreas (Xiang et al., 2017, Bagley et al., 2017, Koike et al., 2019). Further, few research groups tried to develop a joint model between organoid and organ-on-a-chip research resulting in an organoid-on-a-chip technology. More recently, chamber devices have been developed. This system enables separate culture of distinct organoid types while prohibiting the uncontrolled merging of organoids. This system can be very useful to evaluate organoid-organoid communication (Zhang et al., 2009, Zhang et al., 2017).

One other drawback of organoid systems is the ECM composition and its effect on organoid culture. Composition of the ECM can affect the final outcomes in genetic screening or chemical screening of human 3D organoids (Broguiere et al., 2018, Giobbe et al., 2019). Further works needs to be done in this area and as part of 'good manufacturing practice' all the raw materials must be fully defined. This can generate robust human model systems and the feasibility of clinical work. Considering all the remaining limitations, human 3D organoids hold great potential in regenerative medicine and clinical translational research.

### 1.5.5.3. 3D organoids versus 3D spheroids

The organoids and spheroids are defined collectively as 3D cellular structures. Spheroids are usually derived from PSCs as freely floating aggregates in ultra-low attachment plates. On contrary, 3D organoids are made using tissue-resident stem cells or PSCs, which are embedded in the ECM such as Matrigel (Sato et al., 2009). The 3D organoids are highly complex structures, whilst spheroids can be differentiated to any tissue type which can recapitulate some aspects of *in vivo* environment (Figure 1.20). In this study, the focus will be on generation of 3D liver spheroids or 3D hepatospheres (3D heps).



**Figure 1.20: Morphology of organoids versus spheroids. The organoids (left) are highly complex structures and are made by embedding tissue-resident stem cells or PSCs in Matrigel. The spheroids (right) are derived from PSCs and formed as floating aggregates in ultra-low attachment plates. These 3D structures mimic some aspects of the *in vivo* environment. The organoid image is adapted from (Hu et al., 2018). Scale bars 100µm.**

**Table 1.2: Comparison of key hepatic differentiation protocols for iPSC derived 3D liver organoids and hepatospheres.**

<b>Protocol</b>	<b>Protocol steps</b>	<b>3D name</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>Takebe et al, 2013</b>	Co-culture (iPSCs+MSCs+HUVECs) Matrigel Embedding Multistep protocol using GFs and FBs	iPSC-liver buds (LB)	Liver support <i>in vivo</i> for 45 days	Multistep process, use of undefined animal-derived materials like Matrigel and FBS, requires regular passaging, size variation
<b>Guan et al, 2017</b>	iPSC-derived Three-step process Combination of GFs and SMs Dissociation and Matrigel embedding	iPSC-derived Hep-Org	Long term culture (50-60 days)	Size variation, Lack of nutrient and oxygen supply, requires dissociation and re-embedding, size variation
<b>Rashidi et al, 2018</b>	iPSC-derived Four-step process Recombinant GFs Self-aggregated spheroids	PSC-derived 3D Heps	Low proliferation, long-term culture (over a year), self-aggregation, Liver support <i>in vivo</i> , No use of Matrigel	Modest liver function in long-term culture, heterogenous containing non-parenchymal cells
<b>Wu et al, 2018</b>	Four-step process Recombinant GFs	iPSC-derived Hepato-biliary organoids	Hepatic gene expression signatures, liver support <i>in vivo</i> (for 8-weeks)	Long differentiation protocol
<b>Akbari et al, 2019</b>	Three-step process Combination of GFs and SMs Enrichment of EpCAM+ cells Use of Matrigel	iPSC-derived Endoderm-derived hepatic organoids	Rapid and highly efficient (14 days), functional hepatocytes, long-term culture for over a year, liver support <i>in vivo</i>	Complex mixture of SMs and GF, requires re-embedding in Matrigel, requires cell sorting prior to differentiation
<b>Koike et al, 2019</b>	Two-phase differentiation Recombinant GFs Use of Matrigel	iPSC-derived Hepato-biliary-pancreatic organoids	Multi-organ structure 3D model Recapitulate the foregut-midgut boundary	Requires separate differentiation and fusion Size variation

**Abbreviations:** FBS, Foetal Bovine Serum; GFs, Growth Factors; SMs, Small molecules; Hep-Org, Hepatic Organoid; iPSC, induced pluripotent stem cells; MSCs, Mesenchymal Stem Cells; HUVECs, Human Vascular Endothelial Cells; EpCAM, Epithelial Cell Adhesion Molecule; 3D Heps, 3D hepatospheres.

## 1.6. Research study hypothesis

In this study, it was hypothesised that phenotypically stable hiPSC-derived 3D heps can be used as an *in vitro* platform to evaluate genotoxicity of viral vectors and can faithfully predict the risk of IM events. To further test the hypothesis, IM events was also evaluated in the parental hiPSC line which had a different gene expression pattern and would allow targeting a wider array of genes rather than liver-specific ones.

## 1.7. Main objectives:

The main objectives include:

- Developing a 3D heps differentiation platform and characterisation of the generated 3D heps functionalities discussed in chapter 3.
- Optimising transduction methodology and assessing gene viral vector efficiency in 3D heps and hiPSCs using viral vectors expressing GFP under tissue-specific (ApoE), ubiquitous (CMV & CB7) and viral (SFFV) promoters which is discussed in chapter 4.
- Performing insertional site analysis in transduced single cloned hiPSCs and genomic materials obtained from transduced hiPSCs and 3D heps using Solid-phase ligation-mediated PCR (EPTS/LM-PCR) which is discussed in chapter 5.

## 2. Materials & Methods

### 2.1. Chemicals, Reagents, Supplements and Buffers

The chemicals, reagents, buffers and supplements used to perform experiments are listed below.

#### 2.1.1. Chemicals

**Table 2.1: List of chemicals.**

Chemicals	Supplier name
Acetic acid	BDH International, UK
Ammonium chloride	Sigma-Aldrich, UK
Chloroform	Sigma-Aldrich, UK
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich, UK
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, UK
Ethanol	Fisher Scientific, UK
Ethidium Bromide	Sigma-Aldrich, UK
Formaldehyde	BDH, UK
Glycerol	BDH, UK
Industrial methylated spirit (IMS)	Fisher Scientific, UK
Isopropanol	BDH, UK
Magnesium chloride	BDH, UK
Methanol	Fisher Scientific, UK
Paraformaldehyde	Sigma-Aldrich, UK
Phenol	Sigma-Aldrich, UK
Potassium chloride	Sigma-Aldrich, UK
Sodium chloride	Fisher Scientific, UK

#### 2.1.2. Buffers

**Table 2.2: List of buffers.**

Buffers	Supplier name
Tris (tris hydroxymethyl methylamine)	BDH, UK
Opti-MEM (reduced serum medium, with L-glutamine, sodium bicarbonate, HEPES, sodium pyruvate, hypoxanthine, thymidine, trace elements, growth factors and phenol red)	Invitrogen, UK
Trypsin-EDTA, supplied as 1 X solution containing 0.05% trypsin (w/v) in 0.53 mM EDTA.	Invitrogen, UK
Phosphate buffered saline (PBS), formulated without calcium, magnesium or phenol red. (sterile tissue-culture grade)	Sigma-Aldrich, UK
Phosphate buffered saline (PBS) formulated with calcium, magnesium and no phenol red. (sterile tissue-culture grade).	Sigma-Aldrich, UK
RIPA buffer	Sigma-Aldrich, UK

### 2.1.3. Consumables

**Table 2.3: List of Consumables.**

Product	Supplier name
Glass coverslips	BDH, UK
Cryovials	Nunc, Denmark
Eppendorf (0.5 and 1.5ml)	Fisher Scientific, UK
FACS tube	Elkay Laboratory products, UK
Falcon Tubes (15ml)	Fisher Scientific, UK
Falcon tubes (50 ml)	Fisher Scientific, UK
Tissue culture flasks (25, 75,175)	Fisher Scientific, UK
Tissue culture plates (6,12,24,48)	Corning, UK
Glass pipettes	Poulsen & Graf, Germany
Graduated serological pipettes	Fisher Scientific, UK
2ml sterile aspirating pipettes	SLS, UK
Petri dishes	Fisher Scientific, UK
Pipette tips and filter tips	SLS, UK
Syringes, assorted sizes	Fisher Scientific, UK
Ultracentrifuge tubes, 12 ml	Beckman, USA
Cell culture multi-well plate	CellStar, UK
Sterile microwtube® with O-ring	Simport, UK
Ethyl alcohol, molecular grade (Ultra-pure)	Sigma, UK
Millex® filter unit	Merck, Ireland
Stericup® & Steritop®	Millipore, UK
Tissue culture sterile cell scrappers	Corning, UK
96-well optical plate	Thermo Fisher, UK
Sterile reagent reservoir	Corning, UK

## 2.1.4. Commercial Kits

Table 2.4: List of Commercial kits.

Product name	Product code	Supplier
Human liver ALB ELISA kit	1190	Alpha diagnostic
Human liver alpha fetoprotein (AFP) kit	500	Alpha diagnostic
P450-p-Glo Luminescence kit (CYP3A)	V8902	Promega
P450-p-Glo Luminescence kit (CYP1A)	V8772	Promega
BCA protein assay	23227	ThermoFisher
Endo-free plasmid purification (megaprep)	12381	Qiagen
QIAprep® plasmid purification kit (maxiprep)	12162	Qiagen
RNAeasy® Kit	74104	Qiagen
DNeasy® blood & tissue kit	69504	Qiagen
RNase free DNase set	79254	Qiagen
QuantiTect® reverse transcription kit	205311	Qiagen
Taqman® fast advanced master mix	4444556	ThermoFisher

## 2.2. Human Pluripotent Stem Cell culture and hepatic differentiation

List of culture reagents is summarised in table 2.5.

Table 2.5: List of cell culture supplements.

Cell culture supplements	Product Code	Supplier name
Agarose low gelling temperature	A9045	Sigma-Aldrich, UK
BSA	A4919	Sigma-Aldrich, UK
Penicillin-Streptomycin	15140122	Thermo Fisher, UK
Luria-Bertani (LB) broth	L3522	Sigma-Aldrich,UK
Lactose	17814	Sigma-Aldrich,UK
Water, tissue culture grade	W3500	Sigma-Aldrich, UK

### 2.2.1. Induced pluripotent stem cell lines

Two lines of hiPSCs were used in this study.

**1- JHU106i (P106) cell line:** JHU106i (P106) cell line was obtained from Johns Hopkins University - Laboratory of Dr Lewis Becker, WiCell stem cell bank. The cell



line was derived from a healthy Caucasian male and reprogrammed by non-integrating episomal plasmid method using OCT4, Klf4, Sox2, cMyC, Bcl, xl. Upon receiving the cell line from the cell bank, WiCell feeder independent E8 medium protocol used for thawing and passaging prior to expansion. Following expansion of the line, the cells were fully characterised.

**2- 33D6 cell line:** This hiPSC line was previously generated and characterised in Professor David Hay laboratory (Hay et al., 2008a, Sullivan et al., 2010). This cell line was derived from skin fibroblasts and reprogrammed using five integrating viral vectors.

### 2.2.2. Matrix Coating using BD Matrigel™ (MG)

BD Matrigel™ (Bioscience, UK) stock solution was prepared prior to coating. To prepare the stock, Matrigel stock from the freezer was left at 4°C overnight to be thawed. Next day, the Matrigel stock was resuspended in 10 ml ice cold Knockout DMEM (Life Technologies, UK). The mixture was resuspended several times and aliquoted 1 ml of the solution per sterile tube. Matrigel aliquots were stored in the freezer until further use. 1ml of Matrigel aliquot was thawed overnight or for two hours prior to coating at 4 °C. The thawed aliquot was resuspended in 18 ml of ice-cold KO-DMEM and was added to cell culture plates as summarised in Table 2.6.

**Table 2.6: Volume of Matrigel required for each specific plate formats.**

Plate Format	Volume of Matrigel per well
6-well	1ml
12-well	0.5ml
24-well	0.5ml
48-well	0.2ml
96-well	0.1ml

### 2.2.3. Vitronectin (VT) coating

Human recombinant vitronectin (Life Technologies, UK) was resuspended according to the manufacturer's instructions. The vial of vitronectin was thawed at room temperature and aliquoted into 60 µl sterile tubes. The aliquots were kept frozen for future use. To coat a 6-well plate, an aliquot of vitronectin was removed from -80°C

storage and thawed at room temperature. One aliquot was sufficient to coat a 6-well plate format. The thawed vitronectin was added to a sterile 15 ml falcon tube containing 6 ml of sterile DPBS (Sigma-Aldrich) without calcium and magnesium at room temperature. Gently the diluted vitronectin solution was resuspended and 1 ml added to each well of the plate. The plate was incubated at room temperature for 1 hour prior to use. For other plate formats and their required volumes refer to table 2.6.

#### **2.2.4. Coating of the plate with Recombinant Laminin 521 (LN-521)**

For this study, recombinant Laminin 521 matrix was purchased from BioLamina™, Sweden. To coat the cell culture plates, manufacturer's instruction was followed. The concentration in each vial was 100 µg/ml and 5 µg/cm<sup>2</sup> was used to coat the plates. Briefly, 1 ml of laminin solution was diluted in 19ml of ice-cold tissue culture grade PBS with Calcium and Magnesium (Gibco). The solution was gently mixed and 1ml of diluted laminin solution was used to coat a well of a 6-well plate. Once coated, plates were incubated at 37°C for 2 hours or alternatively stored at 4 °C overnight till further use. The plates should be warmed up prior to use for at least one hour at room temperature. For other plate formats and their required volumes refer to table 2.6.

#### **2.2.5. Culturing hiPSCs**

List of culture media and reagents is summarised in table 2-7. The 33D6 hiPSC line cultured for over 30 passages on Matrigel™ coated 6-well plates and were fed with 3ml serum free mTeSR1™ medium (StemCell Technologies) .The P106 cell line was routinely cultured on pre-coated LN-521 6-well plates and were fed with 3ml of mTeSR1™ medium without any antibiotics as previously described (Cameron et al., 2015). The cells were incubated at 37 °C in 5% (v/v) CO<sub>2</sub>, 95% (v/v) air for optimal growth with daily medium change. Further, the cell lines were monitored regularly for bacterial infection and differentiated colonies, using morphological analysis.

#### **2.2.6. Passaging hiPSCs**

hiPSCs were split at a ratio of 1:3 using Gentle Cell Dissociation Reagent (StemCell Technologies, UK). The media was removed, the cells were washed once with PBS (Sigma-Aldrich,UK). The dissociation reagent (1ml) was added and cells were incubated at 37 °C for 5 minutes until the edges of the colonies rounded up. The

dissociation reagent was aspirated, and the cells were washed once with PBS. Fresh media (3ml) was added to the cells and they were subsequently scrapped off and resuspended 1-2 times before transferring 1ml of the cell suspension into a new laminin-coated plate containing fresh media. The split ratio was maintained at 1:3 throughout the study.

**Table 2.7: Summary of essential media and reagents used for stem cell culture maintenance.**

Product	Catalogue Number	Supplier
Matrigel™ basement membrane	354234	BD, Bioscience. UK
Vitronectin	A14700	Thermo Fisher Scientific, UK
Recombinant Laminin-521	LN521-02	BioLamina™, Sweden
Essential 8™ Medium	A1517001	Thermo Fisher Scientific, UK
mTeSR™1 medium	85857	STEMCELL Technologies, UK
Gentle Cell Dissociation Reagent	07174	STEMCELL Technologies, UK
Rock inhibitor (Y-27632)	07171	STEMCELL Technologies, UK
Phosphate Buffered Saline	14190-094	Life Technologies, UK
0.4% Trypan Blue solution	15250-061	Life Technologies, UK

**Table 2.8: Volume of mTeSR™1 required for passaging hiPSCs in different plate formats.**

Plate format	Volume of mTeSR™ 1 per well
6-well	2 ml
12-well	1 ml
24-well	0.5 ml
48-well	0.25 ml
96-well	0.1 ml

### 2.2.7. Freezing and thawing hiPSCs

The freezing medium was made by mixing Knock-Out Serum Replacement (KOSR) with 10% dimethyl sulfoxide (DMSO) (Sigma-Aldrich, UK). At 80-90% confluency, hiPSCs cultured on Matrigel™ and LN-521 were scrapped off, placed into a 15ml falcon tube, and centrifuged at 0.3 RCF for 5 minutes. After centrifugation, the supernatant was aspirated and the cells were resuspended in 1ml of the ice-cold freezing mixture, stored in -80 overnight and transferred to liquid nitrogen for long-term storage.

The cells were routinely thawed by warming the cryotube gently. 10 ml of fresh MT and 10 $\mu$ M rock inhibitor molecule (Y-27,632, Calbiochem) was pre warmed in 50 ml falcon tube. The cells were transferred and resuspended in 5 ml of prepared medium and centrifuged at 0.3 RCF for 5 minutes. Following centrifugation, the supernatant was aspirated, and the cells were resuspended in 1ml of MT on Matrigel™ or LN-521 pre coated plates.

## **2.3. Hepatic differentiation of hiPSCs in 2D monolayer and 3D culture**

### **2.3.1. 2D monolayer differentiation**

In this study, it was important to accurately calculate the number of cells required for specific plate formats in order to establish a successful 2D monolayer and 3D heps. Initially, hiPSCs were expanded prior to differentiation. It is estimated that one well of a 6-well plate at 90% confluency contains 3 $\times$ 10<sup>6</sup> viable cells. The number of hiPSCs required was depended on the surface area of each tissue culture plate format. (Table 2.9).

Below is an example of how the hiPSCs numbers were calculated in order to seed four 6-well plates.

- 1) The surface area per well of a 6-well plate = 9.5 cm<sup>2</sup>
- 2) The total number of viable cells to seed per well of a 6-well plate at 1 $\times$ 1e5 per cm<sup>2</sup> =  
(1 $\times$ 1e5) $\times$ 9.5 = 9.5 $\times$ 1e5.
- 3) Total number of wells which is required to seed hiPSCs for differentiation are 24 wells of 6-well plate and therefore the total number of viable cells are roughly around 2.24 $\times$ 10<sup>7</sup>.
- 4) The number of wells of 6-well plate required before seeding the cells for differentiation can be calculated by dividing the total number of cells required by the number of viable cells in one well of a 6-well plate which:  
2.24 $\times$ 10<sup>7</sup> / 3 $\times$ 10<sup>6</sup> = 7-8 wells.

**Table 2.9: The surface area (SA) based on different plate formats.**

Plate format	Surface area per well
6-well	9.5 cm <sup>2</sup>
12-well	3.8 cm <sup>2</sup>
24-well	1.9 cm <sup>2</sup>
48-well	0.95 cm <sup>2</sup>
96-well	0.32 cm <sup>2</sup>

After seeding hiPSCs on LN-521 pre-coated plates, cells were maintained overnight with MT supplemented with Ri. Endoderm differentiation was initiated next day or when the cells were approximately 30-40 percent (%) confluent. The culture media was replaced with endoderm induction medium (EIM) RPMI 1640 containing 1 x B27 (Life Technologies) supplemented with essential growth factors, 10 ng/ml Activin A (PeproTech) and 50 ng/ml Wnt3a (R&D Systems). The medium was changed every 24 hours for 72 hours and then continued with 10ng/ml Activin A without Wnt3a for another extra two days (Table 2.10).

**Table 2.10: Endoderm induction medium (EIM) composition.**

Material	Cat. Number	Supplier	Vol.
RPMI 1640 medium	21875	Life Technologies	485 ml
B27 supplement 50x stock	17504	Life Technologies	10 ml
Pen/Strep	15140	Life Technologies	5ml
rh AA (Stock Conc. 100 µg/ml)	120-14E	PeproTech	10 ng/ml
rm Wnt3a (Stock Conc. 10 µg/ml)	1324-WN-500/CF	R&D Systems	50 ng/ml

On day 5, endoderm differentiation medium was replaced with hepatoblast differentiation medium (HDM), and this was changed every second day for a further 5 days. The medium consisted of Knockout (KO-DMEM; Life Technologies), knockout serum replacement (KOSR; Life Technologies), 0.5% Glutamax (Life Technologies), 1% non-essential amino acids (NEAA; Life-Technologies), 0.2% b-mercaptoethanol (Life technologies), and 1% DMSO (Table 2.11).

**Table 2.11: Hepatoblast differentiation medium (HDM) composition.**

Material	Cat. Number	Supplier	Vol.
KO-DMEM	10829	Life Technologies	379 ml
KO-SR	10828	Life Technologies	100 ml
GlutaMAX 100x stock	35050	Life Technologies	5 ml
NEAA 100x stock	11140	Life Technologies	5 ml
2-mercaptoethanol	31350	Life Technologies	1 ml
1% (v/v) DMSO	D5879	Sigma-Aldrich	5 ml
Pen/Strep	15140	Life Technologies	5ml

Hepatocyte maturation of the hiPSCs-derived hepatoblasts was induced at day 10 of differentiation. Cells were cultured using serum-free HepatoZYME™ medium (Life Technologies) containing 1% Glutamax (Life Technologies) , supplemented with 10 ng/ml hepatocyte growth factor (HGF, PeproTech) and 20 ng/ml OSM, PeproTech) as described previously (Cameron et al., 2015, Wang et al., 2017) for 13 days. The hepatocyte maturation medium (HMM) was replaced every 48 hours. The maturation medium with essential growth factors were prepared freshly each time (Table 2.12).

**Table 2.12: Hepatocyte maturation medium (HMM) composition.**

Material	Cat. Number	Supplier	Vol.
HepatoZYME™ medium	17705	Life Technologies	490 ml
Hydrocortisone 21-hemisuccinate sodium salt (stock solution 1 mM)	H4881	Sigma-Aldrich	5ml
Pen/Strep	15140	Life Technologies	5ml
Rh HGF (Stock Conc. 10 µg/ml)	100-39	PeproTech	10 ng/ml
Rh OSM (Stock Conc. 20 µg/ml)	300-10	PeproTech	20 ng/ml

Full list of medium change requirement for 2D hepatic differentiation is summarised in table 2.13 to 2.15.

**Table 2.13: The required volume of RPMI medium with essential growth factors during the first phase of hepatic differentiation for different plate formats.**

Plate format	Volume of EIM per well
6 well	2 ml
12-well	1 ml
24-well	0.5 ml
48-well	0.2 ml
96-well	0.1 ml

**Table 2.14: The required volume of HDM and HMM during the second phase of hepatic differentiation for different plate formats.**

Plate format	Volume of HDM & HMM
6 well	4 ml
12-well	2 ml
24-well	1 ml
48-well	0.5 ml
96-well	0.2 ml

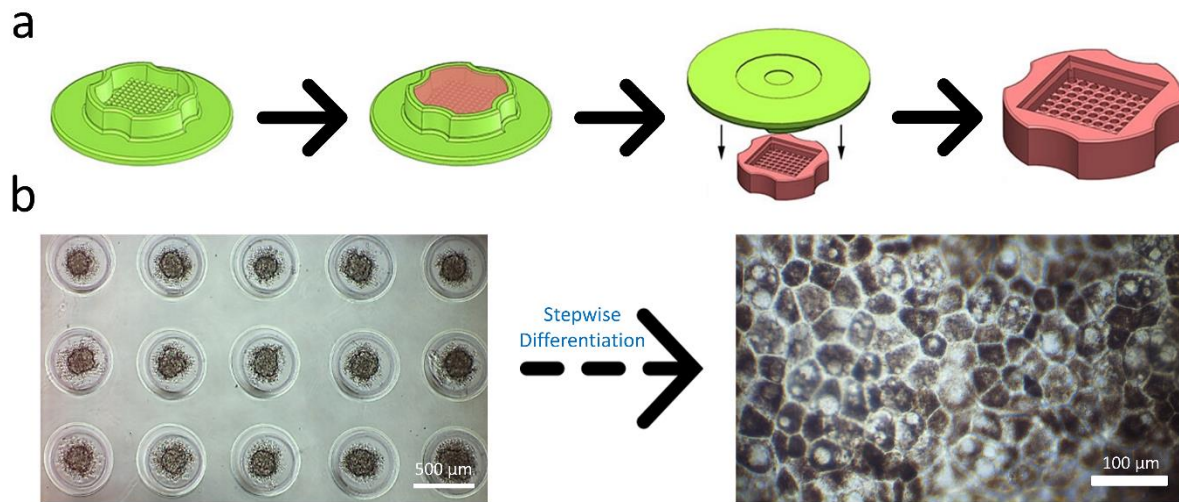
**Table 2.15: Summary of 2D hepatic differentiation stages and the days that require medium changing or medium switching.**

Days of differentiation	Hepatic differentiation stages
Days 0, 1, 3, and 5	HIM
Days 6, 8, and 10	HDM
Days 11, 13, 15, 17, 19, .....	HMM

### **2.3.2 Formation of self-aggregated hiPSCs spheroids (3D-Spheroids)**

Agarose microplates were manufactured in 256-well format using the 3D Petri Dish<sup>®</sup> mould (Sigma Aldrich) following the manufacturer instructions. The formed moulds were transferred to 12-well plates (Corning). hiPSCs cells were scaled up on LN-521 coated plates and incubated with 1ml of Gentle Dissociation Reagent (StemCell Technologies) for 7-10 minutes at 37 °C. Following this step, single cell suspensions were prepared by pipetting the buffer up and down gently. The cell suspension was centrifuged at 0.2 rcf for 5 minute and resuspended in MT supplemented with 10 µM Y-27,632 (Calbiochem) in a density of  $2.0 \times 10^6$  live cells/ml. The agarose microplates

were seeded by transferring 190  $\mu$ l of cell suspension. After 2-3 hours, 1 ml of MT supplemented with 10  $\mu$ M Y-27632 was gently added to each well of 12-well plate and incubated overnight (Figure 2.1).



**Figure 2.1: Fabrication of agarose mold using commercially available Microtissue mold. (a). Following seeding mold with dissociated hiPSC, self-aggregated 3D spheres will form overnight which can be subsequently differentiated into phenotypically stable 3D heps using a stepwise differentiation protocol (b) (Rashidi et al., 2018).**

### 2.3.3. 3D Hepatic Differentiation

Following overnight incubation and formation of 3D aggregated, endoderm differentiation was initiated by replacing the culture media with supplemented EIM (Table 2.10). The medium was changed every 24 hours for 72 hours and then continued with 10ng/ml Activin A without Wnt3a for another extra two days. On day 5 of differentiation, EIM was replaced with HDM (Table 2.11). The medium was changed every other day for a further 5 days. Hepatocyte maturation was induced at day 10 of differentiation using HMM as described earlier (Table 2.12). The medium was replenished every 48 hours for 10 days. On day 21, cells were cultured in 3DM as described elsewhere (Rashidi et al., 2018) until termination of the experiments (Table 2.16).



**Table 2.16: 3D medium composition.**

Material	Cat. Number	Supplier	Vol.
William's-E Medium, no phenol red	A1217601	Life Technologies	445 ml
KO-SR	10828	Life Technologies	50 ml
Pen/Strep	15140	Life Technologies	5ml
Rh HGF (Stock Conc. 10 µg/ml)	100-39	PeptoTech	10 ng/ml
rh VEGF (Stock Conc. 10 µg/ml)	293-VE	R&D Systems	10 ng/ml
rh EGF (Stock Conc. 10 µg/ml)	236-EG-200	PeptoTech	10 ng/ml
rh FGF2 (Stock Conc. 10 µg/ml)	100-18B	PeptoTech	10 ng/ml

## 2.4. Characterisation

### 2.4.1. Immunofluorescent staining of 2D monolayer hiPSC and hiPSC-derived HLCs

hiPSCs, hiPSCs-derived hepatic endoderm and hiPSCs-derived HLCs were fixed in ice-cold 100% methanol (Sigma-Aldrich) at -20°C for 30 minutes. Also, cells can be stored in PBS at 4°C and stained for later analysis. Following fixation, cells were washed three times with PBS, 5 minutes interval between each wash at room temperature. For blocking stage, PBS-0.1% Tween (PBST) containing 10% BSA (Sigma Aldrich) was used for one hour at room temperature. The blocking solution was removed and diluted primary antibody in 1% BSA (made up in PBST) was added to the fixed cells and incubated at 4°C overnight with shaker (Table 2.17). Next day, the cells were washed three times with PBST, 5 minutes interval between each wash. Next, the cells were incubated with the appropriate secondary antibody diluted in PBST, 1% BSA for 1 hour at room temperature. The plate was wrapped in aluminium foil to avoid light (Table 2.18). Following this step, wells were incubated with DAPI (1:1000, Sigma-Aldrich) for 20 minutes at room temperature following manufacturer's instructions. DAPI solution was removed and wells mounted with 50 µl of PermaFluor aqueous mounting medium (Thermo Scientific). For long term storage plates were wrapped with aluminium foil and kept at 4°C.

**Table 2.17. Summary of primary antibodies used for characterisation of hiPSCs and mature hepatocytes in this study.**

Primary Antibody	Host	Dilution	Supplier
AFP	Mouse Mono	1/500	Abcam
$\beta$ -tubulin III	Mouse Mono	1/1000	Sigma-Aldrich
Octamer 4	Rabbit poly	1/200	Abcam
Nanog	Rabbit poly	1/200	Abcam
Sox 17	Goat poly	1/500	R&D systems
HNF4 $\alpha$	Rabbit poly	1/100	Santa Cruz
ALB	Mouse mono	1/200	Sigma
Ki67	Mouse mono	1/400	DAKO
E-Cadherin	Mouse mono	1/200	Abcam

**Table 2.18. Summary of secondary antibodies used in this study.**

Secondary Antibodies	Host	Dilution	Supplier
Anti-Rabbit 488	Donkey	1/400	Life Technologies
Anti-Rabbit 568	Donkey	1/400	Life Technologies
Anti-Mouse 488	Rabbit	1/400	Life Technologies
Anti-Mouse 568	Goat	1/400	Life Technologies
Anti-Goat 488	Rabbit	1/400	Life Technologies
Anti-Goat 568	Rabbit	1/400	Life Technologies
Anti-Sheep 488	Donkey	1/400	Life Technologies

#### 2.4.2. Immunofluorescence (3D Spheroids)

3D spheroids were fixed in ice-cold 100% methanol for 30 minutes. Spheroids were washed in PBS and embedded in agarose. Agarose-embedded spheroids were fixed in paraffin and 4 $\mu$ m sections were obtained. Antigen retrieval was performed by heating dewaxed and rehydrated sections using 1 $\times$  Tris-EDTA buffer solution for 15 minutes at full power in microwave. Following this step slides were used for subsequent staining. To stain sectioned spheroids, hiPSCs-derived HLC spheroids were blocked with 10% BSA in PBST and incubated with primary antibody overnight at 4°C and detected using species-specific fluorescent-conjugated secondary antibody. Sections were incubated with 4,6-diamidino-2-phenylindole (DAPI) and

mounted with Fluomount-G (SouthernBiotech) before microscopy. An extended staining protocol was developed in order to stain whole mount 3D Heps, as previously described (Rashidi et al., 2018).

#### **2.4.3. Histology (3D spheroids)**

Sections with 4µm in thickness were obtained from paraffin-embedded 3D heps and were stained with Hematoxylin and Eosin for histological analysis. The slides were mounted in Pertex before imaging.

#### **2.4.4. Imaging and acquisition**

All images were captured and analysed using a Zeiss Axio Observed Z1 microscope, with LD plan-Neofluar objectives lenses (Carl Zeiss Ltd, Welwyn garden City, UK). The microscope was attached to a Zeiss AxioCamMR3 camera. The image acquisition and processing software was Zeiss Axiovision Rel 4.8 and Axiovision version 4.7.1.0, respectively. For hiPSCs-derived 3D HLC spheroid slides, brightfield images were obtained using a Nikon Eclipse e600 microscope equipped with a Retiga 2000R camera (Q-imaging) and Image-Pro Premier software was used to analyse the images.

#### **2.4.5. Flow cytometry**

Flow cytometry was employed to detect expression of cells surface markers of hiPSCs cultured in Laminin-521. Cells were washed with 2ml PBS without calcium chloride and magnesium chloride. The PBS was aspirated and 1ml of TrypLE™ was added to the cells for 7 minutes until the cells transformed into single cells. Single hiPSCs were collected and resuspended in FACS-PBS (PBS supplemented with 0.1% BSA and 0.1% sodium azide), counted and resuspended at  $1 \times 10^6$  cells/ml for use. Tubes containing 100,000 cells were incubated for 30 minutes at 4° C with the fluorochrome conjugated antibodies (Table 2.19). Following incubation, cells were then washed once with PBS, removing any unbound antibody and centrifuged at 1500 rpm for 5 minutes. After centrifugation, cells were resuspended in 100 µl of FACS-PBS. Antibody binding to the surface of the cells was measured using the optimum concentration of an appropriate fluorochrome conjugated isotype specific antibody. In this study, unstained cells were used as controls. In terms of analysis, Dead cells and remaining debris were not included. Measurement was carried out by using an electronic live gate on forward scatter and side scatter parameters. Data was acquired

for 20,000-50,000 “live” events for each sample using a BD FACS Calibur Flow Cytometry System (Becton, Dickinson and Company, Biosciences, San Diego, CA) equipped with a 488 nm laser and analysed using FlowJo software (Treestar Inc., San Carlos, CA).

**Table 2.19. Summary of antibodies used flow cytometry analysis.**

Flow Cytometry Antibodies	Host	Dilution	Supplier
SSEA4	Mouse		Biolegend
TRA-1-60	Mouse		Biolegend
TRA-1-81	Mouse		Biolegend
SSEA1	Mouse		Biolegend

#### **2.4.6. Enzyme-Linked Immunosorbent Assay (ELISA)**

Two dimensional and 3D mature hepatocytes derived from hiPSCs were incubated with HepatoZYME™ medium for 24 hours at different time points at 37°C in 5% (v/v) CO<sub>2</sub>, 95% (v/v) O<sub>2</sub>. The supernatants were collected after 24 hours and could be stored in -80 °C for later analysis. The ALB and AFP levels were measured, using commercially available microwell plates pre-coated with immobilized human anti-ALB and AFP antibodies (Alpha Diagnostic Intl. Inc., San Antonio, USA). The supernatant was diluted into 1:3 or 1:10 on the working sample diluent and transferred into the wells in duplicate followed by 1-hour incubation at room temperature as per manufacturer’s instructions. Following this step, microwells were washed with working wash solution for four times and diluted anti-human ALB HRP conjugated was added to the microwells. The wells were incubated at room temperature for 30 minutes. Again, microwells were washed with working wash solution five times, and the substrate for the HRP enzyme TMB, was added and incubated for 15 minutes in the dark. Following this step, the stop solution was added to each well and the plates luminous activity was measured at 450 nm with a reference wavelength of 630 nm using a FLUOstart Omega plate reader (BMG LabTech, Germany). Plain tissue culture media which was incubated for 24 hours at 37°C was employed as a negative control. For data analysis, the collected data was normalised per ml per 24 mg protein as measured by the BCA Assay (Pierce, UK).

#### **2.4.7. Cytochrome P450 Assay**

2D and 3D hiPSC-derived HLCs were incubated for 5 hours, with the luciferin conjugated specific CYP3A (1:40) and CYP1A2 (1:50) substrate (P450 P-Glo Luminescence Kit, Promega, UK) at 37°C. The incubated supernatants were then collected and also could be stored at -80 C for later analysis. The Luciferin detection reagent was reconstituted by mixing the buffer into the bottle containing the lyophilised Luciferin detection reagent. For measurement, in a white 96 well plate 50 µl of the supernatant sample mixed with 50 µl of the detection reagent was added and incubated at room temperature in the dark for 20 minutes. The data was collected using a luminometer (POLARstar optima). For data analysis, units of activity were measured as relative light units per ml per mg protein (RLU/ml/mg) as determined by the BCA assay.

#### **2.4.8. Cytochrome P450 drug inducibility**

Two dimensional and 3D hiPSCs derived HLCs were incubated at day 18 and day 30 of differentiation for 48 hours with specific chemical compounds which promote the activity of liver enzymes such as P450s, CYP1A2 and CYP3A. Stock solutions of phenobarbital (Sigma-Aldrich) and Dexamethasone (Sigma-Aldrich) were prepared. The concentration of phenobarbital stock solution was 1 M in PBS and diluted to a final concentration of 1 mM or 2mM. The concentration of Dexamethasone stock solution was 1 mM which diluted in fresh media to a final concentration of 10 µM. Media containing drug inducing compounds were refreshed daily. For control samples, plain media was used with an appropriate vehicle such as PBS or DMSO and the media changed daily. The level activity of CYP3A and CYP1A2 was measured using luciferin reagents CYP 3A (1:40) and CYP 1A2 (1:50) substrate (P450 P-Glo ® Luminescent Kit, Promega. UK) for 5 hours at 37° C. The luminous activity was measured using luminometer (POLARstar Optima) and the units of activity was expressed as relative light units/ml/mg protein (RLU/ml/mg), as confirmed by the BCA Assay.

## **2.5. Molecular Biology Techniques**

### **2.5.1. RNA extraction**

For RNA extraction RNeasy® Kit (Qiagen, UK) was used, and the protocol was followed according to the manufacturer's instructions. The cell samples were isolated and washed with PBS and resuspended in 350 µl of lysis buffer (Buffer RTL, Qiagen) containing β-mercapthoethanol (Gibco, UK). The cells were collected and transferred into 1.5ml Eppendorf tubes. Samples can be stored in -80 °C for later analysis. Tubes were mixed gently, and an equal volume of 70% ethanol was added to the mixture. The resulted suspension mixture was transferred to a RNeasy Spin Column with collection tube underneath. The samples were centrifuged at 10,000 rpm for 20 seconds and the resulted flow through was discarded. Next step, 700 µl buffer RW1 was added and centrifuged at 10,000 rpm for 20 seconds. Following this step, 500 µl of buffer RPE was added and centrifuged at 10,000 rpm for 30 seconds. This step was repeated twice for 2 minutes in the last step. The RNeasy spin column was added to the new collection tube and further centrifuged at 16,000 rpm for 1 minute. Using a new RNAase free 1.5ml collection tube, 100 µl RNAase free was added to the spin column membrane. After 3 minutes incubation, the tubes were centrifuged at 10,000 rpm for 1 minute. The eluted RNA can be stored at -80 °C for further analysis. The RNA samples were measured using the Nanodrop machine for the concentration and purity.

### **2.5.2.DNA extraction**

For DNA isolation in this study DNeasy Blood & Tissue Kits (Qiagen, UK) was used and the protocol was followed according to the manufacturer's instructions. A maximum of  $5 \times 10^6$  cells was centrifuged at 190 rpm and the pellet was resuspended in 200 µl PBS. To the mixture, 20 µl proteinase K and 200 µl Buffer AL was also added. The cell suspension was mixed thoroughly by vortexing. To the mixture, 200 µl ethanol (96-100%) was added. The resulted mixture transferred into a DNeasy Mini spin column placed in a 2ml collection tube. The column centrifuged at 8000 rpm for 1 minute. The flow through and collection tube were discarded. The spin column was placed in a new 2 ml collection tube. 500 µl of Buffer AW1 was added, centrifuged and the flow through was discarded. This step was repeated again using 500 µl of Buffer

AW2, the column was centrifuged at 14,000 rpm for 3 minutes. The flow through and collection tube were discarded. The spin column was transferred to a new 1.5ml or 2 ml microcentrifuge tube. The DNA was eluted by adding 200  $\mu$ l Buffer AE to the centre of the spin column membrane. The column was incubated for 1 minute at room temperature and centrifuged for 1 minute at 6000 X *g*.

### **2.5.3. RNA purification**

For RNA purification, genomic DNA was eliminated by using RNase-free DNase set kit (Qiagen, UK), according to the manufacturer's instructions. The DNase I stock solution was prepared using the RNase-Free DNase set. The lyophilized DNase I was dissolved in 550  $\mu$ l of the RNase-free water. RNase-free water was injected into the vial using an RNase-free needle and syringe. The vial was gently mixed. For long term storage of DNase-I, stock solution was removed from the glass vial and divided into single-use aliquots. The aliquots were stored at -20 °C for later use.

### **2.5.4. Reverse transcription (RT)**

To synthesis complementary DNA (cDNA) from RNA, reverse transcription (RT) was performed using the QuantiTect Reverse Transcription Kit (Qiagen), according to the manufacturer's instructions. Following nanodrop measurement of RNA samples, 1  $\mu$ g of RNA from each sample was reverse transcribed. To reverse transcribe the RNA, 1  $\mu$ l Quantiscript Reverse Transcriptase, 4  $\mu$ l 5X Quantiscript RT Buffer and 1  $\mu$ l RT primer mix were added to the purified RNA reaction and the mixture was incubated at 42°C for 30 minutes, followed by a 95 °C heat treatment for 5 minutes and a cooling step (4°C ) for 5 minutes. The resulting cDNA was employed for qPCR analysis.

### **2.5.5. Quantitative polymerase chain reaction (qPCR)**

Quantitative real-time PCR (qPCR) was performed using the TaqMan®Fast Advanced Master Mix (Thermo-Fisher Scientific, UK) and the relevant primers (Applied Biosystems). Each qPCR reaction was set up using 0.5  $\mu$ l of specific primers, (for primer details, see table 2.20). 5.5  $\mu$ l of TaqMan®Fast Advanced Master Mix and 5.5  $\mu$ l of nuclease free water containing 12 ng of cDNA per reaction. Each sample was run in triplicates. In this study, qPCR technique was used to investigate the expression of specific genes within the hiPSCs and mature HLCs. The synthesised cDNAs from

these samples are utilised as the template. The qPCR reaction consisted of three different phases:

- 1) An initial denaturation step at 95°C for 10 minutes followed by another 40 cycles of denaturation.
- 2) Annealing and extension step at 60°C for 1 minute.
- 3) Cooling at 4°C

The samples were analysed using Roche LightCycler 480 Real-Time PCR system and data analysis was carried out using Roche LightCycler 480 software (version 1.5) in the form of cycle threshold (CT) values. This value marks the point at which fluorescence light intensity reaches a set threshold above the background level. The relative expression was calculated by  $\Delta\Delta$ CT method (Schmittgen and Livak, 2008) and the results were normalised in relation to the housekeeping genes such as GAPDH and B2M. Quantitative PCR reactions were run in triplicate format and the levels of significance were measured using students' *t-test* statistical analysis.

This equation signifies the expression of the target gene in the test samples relative to the control samples. The  $(E)_{\text{Target}}$  and  $(E)_{\text{Housekeeping}}$  are the PCR amplification efficiencies which are calculated from the real-time PCR reactions. The  $\Delta\text{Target Ct}$  and  $\Delta\text{Housekeeping Ct}$  are defined as the difference between the CT values of the target gene in the control versus samples.

$$\text{Ratio} = \frac{(E)_{\text{Target}}^{\Delta\text{Target Ct (Ct control-Ct sample)}}}{(E)_{\text{Housekeeping}}^{\Delta\text{Housekeeping Ct (Ct control-Ct sample)}}$$

**Table 2.20. Summary of the oligonucleotides and their sequences used in this study.**

Gene	Primer	Supplier
Octamer 4	Hs00742896-s1	Applied Biosystems
Nanog	Hs02387400-g1	Applied Biosystems
Sox 17	Hs00751752-s1	Applied Biosystems
AFP	Hs01040607-m1	Applied Biosystems
ALB	Hs00910225-m1	Applied Biosystems
HNF4 $\alpha$	Hs01023298-m1	Applied Biosystems
E-Cadherin	Hs01023298-m1	Applied Biosystems
GAPDH	Hs02758991-g1	Applied Biosystems



## **2.6. Protein Biochemistry Techniques**

### **2.6.1. Cellular protein extraction**

The cell samples for protein extraction were lysed in 150  $\mu$ l of RIPA buffer (Millipore, UK) containing proteinase and phosphatase inhibitors with final concentration of 1% (Sigma-Aldrich). The lysed extracts were collected and centrifuged at 10,000 rpm for 15 minutes at 4°C. The Supernatant was transferred to a new 1.5 ml Eppendorf tube.

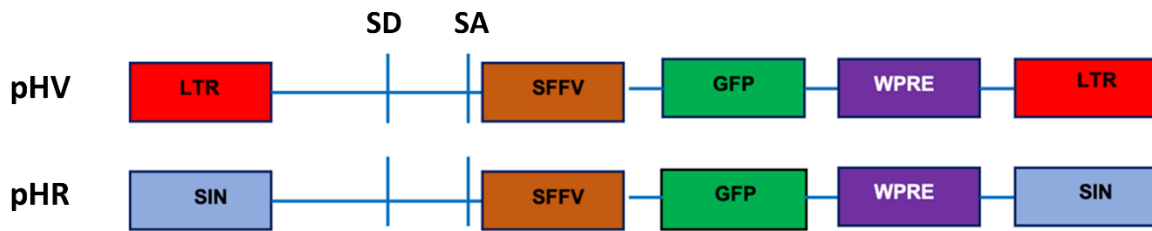
### **2.6.2. Measuring protein concentration**

The Pierce BCA (bicinchoninic acid) protein assay kit (Thermo Fisher Scientific, UK) was employed to measure the protein concentration in the lysed cell extract samples. Protein extracts were diluted 1:2 using nuclease free water. The samples were run in a 96 well plate and in triplicates format. The protocol was followed according to the manufacturer's instructions. Reagents A and B were mixed at a 1:50 ratio and 200  $\mu$ l was added to each well containing the samples. BSA standards were used as control wells ranging from 20 to 2,000  $\mu$ g/ml. Following samples loading, the plate was incubated at room temperature for 10 minutes and read at the absorbance of 562 nm. The protein concentration was quantified by linear extrapolation generating the standard curve from the protein control standards.

## **2.7. Vector production**

### **2.7.1. Lentiviral vector constructs**

In this study, the HIV type I (HIV-1) derived lentiviral constructs as reference vectors were used. They were classified into 'Safe' and 'Unsafe' vectors. The pHV vector contained the U3 region of the LTR from the wild type and considered as an unsafe vector and driving the reporter GFP gene expression under spleen focus-forming virus (SFFV) and an internal phosphoglycerate kinase (PGK) promoter. Deletion of the U3 region of the LTR of pHV vector forms self-inactivating (SIN) lentiviral vector. The pHR lentiviral vector is considered 'safe' and activates the reporter gene expression by SFFV and PGK promoters respectively. Details of lentiviral vectors are shown in Figure 2.2. Previously designed plasmids were amplified in house and used to transduce HEK293 cell line to generate two different lentiviral vectors.



**Figure 2.2: The HIV type 1 (HIV-1) derived lentiviral vectors. (I) Lentiviral vector containing wild type HIV LTR (pHV, unsafe vector). (II) Self-inactivating (SIN) lentiviral vectors with no U3 region of the LTR (pHR, safe vector).**

### 2.7.2. Adeno associated virus (AAV) viral vector constructs

Recombinant adeno-associated virus (rAAV) are commonly used as a viral vector in gene therapy. It consists of the packaging size of the expression cassette which will be located between two ITR segments. The AAV2 vectors used in this study were not constructed in house and kindly provided in collaboration with Dr Ian Alexander, Dr Sharon Cunningham and Dr Leszek Lisowski laboratories in Australia. The details of provided AAV vector constructs are as follows:

- 1) **rAAV/LK03-LSP1.EGFP (2X ApoE/1xhAAT)**. This rAAV vectors were initially produced in order to test the best hepatic transduction efficiency in 3D mature HLC model. These vectors encoding eGFP protein under the transcriptional control of a liver specific promoter in the AAV2 capsid and highly efficient liver tropic synthetic capsids. Recombinant vectors were produced in Dr Leszek Lisowski laboratory. Summary of all the AAV2 constructs are provided in Table 2.21.

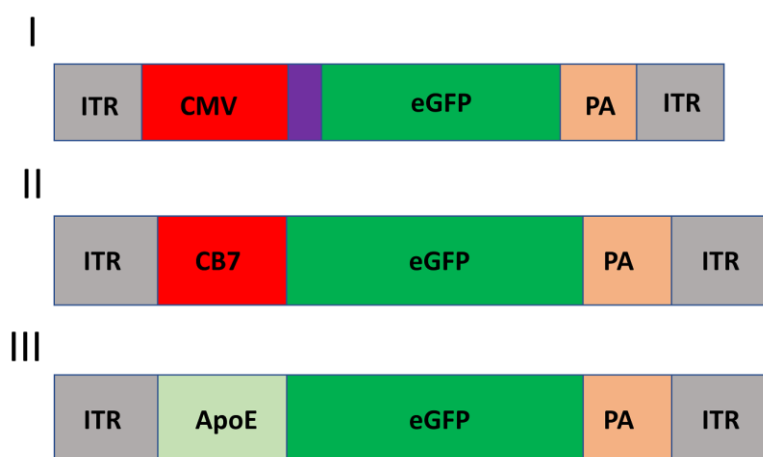
for the rest of the study, two other main rAAV vector constructs were designed which the reporter gene is driven by strong (unsafe) and weak (safe) promoters. Details of each vector constructs are shown in Figure 2.3.

- 2) **rAAV/LK03-CB7.EGFP**. This rAAV vector constructed with “clean ITRs” which the reporter gene is driven by hAAT promoter as previously described in (Chandler et al., 2015). (Unsafe promoter, Figure 2.3 (II)).
- 3) **rAAV/LK03-LSP.EGFP (Construct V) (1X ApoE/1xhAAT)**. This rAAV vector constructed with “clean ITRs” which the reporter gene is driven by the CBA

promoter as previously described in (Chandler et al., 2015). (Safe promoter, Figure 2.3 (III)).

Further, another rAAV vector construct with the reporter GFP gene driven by the CMV and wild type AAV2 3'UTR promoter enhancer element was kindly provided by the King's University to study hepatic transduction (Figure 2.3 (I)).

The map of provided constructs is shown in figure 2.3 and a list of all used AAV vectors in this study is summarised in Table 2.21.



**Figure 2.3: Toxicology details of rAAV vector design constructs employed in this study. (I) pSUB201-derived vector construct with the reporter GFP gene driven by the CMV and wild type AAV2 3'UTR promoter enhancer element (strong promoter, provided by King's College London). (II) pSUB201-derived vector construct with the reporter GFP gene driven by the CB7 promoter (strong promoter, provided by Prof Ian Alexander, University of Sydney, Australia). (III) Vector construct with the reporter GFP gene driven by the APOE promoter (weak promoter, provided by Prof Ian Alexander, University of Sydney, Australia).**

**Table 2.21. Summary of AAV vectors used in this study.**

rAAV vector	Capsid	Promoter	Titre (Vg/ml)
rAAV/LK03-LSP1.EGFP	AAV2	2 X ApoE	$1.03 \times 10^{13}$
rAAV/LK03-LSP1.EGFP	LK03	2 X ApoE	$3.07 \times 10^{13}$
rAAV/LK03-LSP1.EGFP	AAV8	2 X ApoE	$3.2 \times 10^{14}$
rAAV/LK03-LSP1.EGFP	NP59	2 X ApoE	$9.37 \times 10^{12}$
rAAV/LK03-LSP1.EGFP	AAV DJ	2 X ApoE	$7.39 \times 10^{13}$
rAAV/AAV2/2-CMV.EGFP	AAV 2/2	CMV	$8.45 \times 10^{12}$
rAAV/LK03-CB7.EGFP	LK03	CB7	$2.17 \times 10^{13}$
rAAV/LK03-LSP.EGFP	LK03	ApoE	$8.57 \times 10^{12}$

### **2.7.3. Bacterial culture and plasmid amplification**

In this study, high efficiency chemically competent *E.Coli* cells DH5 alpha Max Efficiency® Stabl2™ competent cells and DH10B were obtained from Invitrogen, UK. In this study, three different plasmids were routinely used to produce HIV-1 derived lentiviral vectors.

- 1) Packaging plasmids which contain structural (Gag) and replication (Pol) components.
- 2) Envelop plasmid which contains the G protein of the vesicular stomatitis virus envelop gene (VSV-G).
- 3) Green fluorescent protein (GFP) based plasmid.

### **2.7.4. Restriction enzyme digest**

All restriction enzymes and their reaction buffers were obtained from New England Biolabs (NEB, MA, USA) or Invitrogen, UK. To verify the purified DNA plasmids from bacterial cultures, digestion reaction was set up using 1 µg of DNA of interest, 1 X reaction buffer (New England Biolabs, UK), restriction enzyme (New England Biolabs, UK) in a total 20 µl reaction volume. Reactions were incubated at 37°C for 1 hour and the results were verified using agarose gel electrophoresis.

### **2.7.5. Polymerase Chain Reaction (PCR)**

In this study, polymerase chain reaction (PCR) technique was used to amplify a segment of LTR specific region of lentiviral vectors (pHR and pHV). A general protocol was set up and followed. A master mix was prepared composed of PCR reaction buffer at a final concentration of 1x, Magnesium Chloride (MgCl) at a final concentration of 1.0- 3.0 mM, forward (GAGCTCTCTGGCTAACTAGG, TM:60.2) and reverse (GCTAGAGATTTTCCCACTG, TM:55.9) lentiviral primers, dNTPs (0.2 mM) and *Taq* DNA polymerase (0.5 units per 20 µl reaction). The master mix can be aliquoted for future use. For each reaction, 1-2 µl of template DNA is required. Tubes were transferred to a thermal cycler (Perkin Elmer, USA).

**Table 2.22: List of RT-PCR reagents.**

Product	Supplier name
DNA Ladders (100 or 1 KB) base pairs	Invitrogen, UK
Deoxyribonucleotide triphosphates (dNTPs)	Invitrogen, UK
PCR Buffer/MgCl	Invitrogen, UK
<i>Taq</i> DNA polymerase	Invitrogen, UK

### **2.7.6. Growth and storage of bacterial cultures**

Cultures of DH5 $\alpha$  and DH10B *Escherichia coli* cells were expanded in Luria-Bertani (LB) media in 37°C shaking at 200 rpm. Cultures of other strain of *E. Coli* such as STBL2 were grown at 30°C. For miniprep plasmid extraction 5 ml and for production of large amount of plasmid 1-2 litre culture volumes were used. In this study, bacterial cultures of Gag/pol, VSV-G and GFP plasmids were routinely prepared for lentiviral vector production. In this preparation, 100  $\mu$ l of bacterial glycerol stocks were mixed with 10 ml of PDM media. Bacterial cultures incubated at 37°C on a shaker at 200 rpm overnight. Following day, 5ml of the inoculated bacterial cultures were transferred into two separate 1L flasks, containing 245 ml of PDM. The flasks were incubated on a shaker at 37°C at 140 RPM for two nights.

For long-term storage of stationary phase cultures of *E. Coli* containing the plasmids were stored by adding the glycerol solution to the cultures at a final concentration of 20% (v/v) and kept at -80°C.

### **2.7.7. Isolation of plasmid DNA by Miniprep**

Following the growth of 5ml liquid culture *E. Coli* containing the plasmid of interest with the appropriate antibiotic, the plasmid DNA was isolated using the Miniprep spin column kit (Qiagen,UK) as per manufacturer's instructions. Bacteria were centrifuged at 4,000 rpm for 5 minutes in a benchtop microcentrifuge (MSE UK Ltd, Beckenham, UK). The bacterial pellets were resuspended in an RNase A containing resuspension buffer and lysed by lysis buffer. The solution was neutralised using the kit's neutralisation buffer containing acetic acid. The lysed cells were centrifuged at 13,000 rpm for 10 minutes and the resulting supernatant was transferred to spin columns

allowing the plasmids to bind to the matrix. Following protocol wash steps, the plasmid DNA was eluted in 50  $\mu$ l water. This protocol allows the efficient isolation of approximately 10  $\mu$ g of DNA.

#### **2.7.8. Isolation of plasmid DNA by Megaprep**

For the large-scale production of plasmid DNA Megaprep kit (Qiagen, UK) was used. The principle of this technique is very similar to miniprep though in larger volume. For high copy number plasmids 500 ml and for low copy number plasmids 2 x 500 ml liquid cultures of *E. Coli* were processed in one column. A proprietary endotoxin-removal buffer was included in the kit which was added to the bacterial cell lysates prior to the binding to the column. Following elution step, isopropanol used to precipitate DNA and centrifuged at 15,000 g for 45 minutes. Wash steps were following as per protocol instructions with 70% ethanol and the purified DNA was resuspended in dH<sub>2</sub>O. It was estimated that 1.5-2.5mg of endotoxin-free plasmid DNA to be obtained using megaprep kit.

#### **2.7.9. Vector production in mammalian cell line**

Human embryonic kidney 293 (HEK293) which is originally derived from kidney cells of human embryo and are easy to grow in culture have been used in transfection studies and also to produce recombinant DNA or gene products. In this study, a derivative of HEK293 cells with high transfection efficiency was used which contains the SV40 large T-antigen. This characteristic allows the high replication of transfected plasmids containing the SV40 promoter by the T-antigen (Ooi et al., 2016).

HEK 293 cell line was maintained in growth medium: DMEM containing 10% Foetal Calf Serum (FCS), supplemented with 100 U/ml of penicillin and 100  $\mu$ g /ml streptomycin. Cells were grown as 2D monolayers in 25, 75 or 175 cm<sup>2</sup> sterile tissue culture flasks in a CO<sub>2</sub> incubator (New Brunswick Scientific, Edison, NJ, USA) maintained at 37°C.

Once the cells reached optimal confluency, medium was aspirated, and cells were washed with PBS. Trypsin-EDTA solution as a dissociation reagent was added and the flasks were incubated at 37°C until the cells were dislodged from the surface. Trypsin activity was neutralised by the addition of more than five volumes of culture medium. The cells were resuspended using serological pipettes to form a

homogenous single cell culture suspension. An aliquot of cell suspension was transferred to a new flask containing fresh culture medium and transferred to the incubator. Cells were maintained and passaged two or three times per week, using a passage ratio of 1:5-1:12 depending on different culture conditions.

Following trypsinisation and single-cell suspension, cells were centrifuged at 2,000 rpm for 10 minutes. The pellets are resuspended in freezing medium containing DMEM, 20% FCS and 10% DMSO) to the final cell density of  $5.0 \times 10^6$  to  $1 \times 10^7$  cells/ml. 1 ml of cells were transferred to the sterile cryovials and kept in to  $-80^\circ\text{C}$  for 24 hours. Next day, the vials were transferred to liquid nitrogen for long term storage until further use.

Cells were dissociated using trypsin reagent and resuspended in culture medium as described as described before. Following dissociation, cells were counted using a haemocytometer to determine the cell density. Cells were seeded at different density based on the experiments. For instance, for transfection the cell suspension adjusted to  $1.0 \times 10^5$  cells/ml. From this cell suspension, 2ml added per well of a 6 well plate or 1 ml added per well of a 12 well plate. Plates or flasks were shaken to distribute the cells evenly. Cells were transferred into incubator till the cells were attached. It usually takes 6-7 hours for HEK293 cells to attach properly.

#### **2.7.10. Lentiviral titration using flow cytometry**

This method was routinely used to quantify virus titre using flow cytometry. It is only applicable in vectors which carry a fluorescent reporter gene such as GFP. For lentiviral titration,  $5 \times 10^5$  HEK 293 cells/well were seeded in a 12-well plate containing DMEM media. The cells were incubated at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$ . Following 24 hours incubation, the cells from one well were dissociated and counted using cell counting machine. Titrations were set up for non-concentrated and concentrated vector batches. For non-concentrated vector the cells were transduced with different dilutions of crude supernatant ranging (500  $\mu\text{l}$ , 100  $\mu\text{l}$ , 50  $\mu\text{l}$ , 20  $\mu\text{l}$  and 10  $\mu\text{l}$ ) and volume of DMEM medium was adjusted to 500  $\mu\text{l}$ . For concentrated vector batches the cells were transduced with different dilutions of concentrated vector ranging (1  $\mu\text{l}$ ,  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ , and  $10^{-4}$   $\mu\text{l}$ ) and volume of DMEM medium was adjusted to 500  $\mu\text{l}$ . One well of non-transduced cells were used as a control. After 24 hours, post transduction medium was replaced with fresh complete DMEM. Following 72 hours post transduction, cells

were washed with PBS, trypsinised and cells were collected for flow cytometry analysis. Data was collected using flow cytometry software and the percentage of positive GFP cells were quantified. Using a specific formula, the titre of lentiviral vector batches was calculated using the equation below:

$$\left( \frac{\text{Number of cells counted on Day 1} \times (\text{percentage of GFP positive cells} / 100)}{\text{Volume of virus solution added expressed in ml}} \right) \times \text{dilution factor}$$

### 2.7.11. Preparation of lentivirus vector using GeneJuice™ transfection reagent

For HIV-1 derived lentivirus vector productions such as pHR and pHV (see lentiviral vector section),  $1.5 \times 10^7$  HEK293 cells were seeded in one T-175 flask one day prior to transfection. This ensures the cells were 80-90% confluent on the day of transfection. Following day, 25 ml of Opti-MEM medium was filtered using 0.2 µm filter and 66 µl of Gene Juice (Merk Millipore, UK) was added to the 12 ml of filtered Opti-Mem. Plasmid DNAs were added to the Gene Juice/Opti-MEM solution and allowed to bind at room temperature for 20 minutes. The amount of plasmid DNAs added to the solution were as follows: 1) GFP plasmid (16 µg), pMD.G2 (VSV-G, 12 µg) and p8.74 (GAG-POL, 4 µg). After 20 minutes incubation time, the cells were washed in Opti-MEM and the 12 ml of the mixed solution containing DNA complexes were added to the seeded flasks. The cells were incubated at 37 °C, 5 % CO<sub>2</sub> for 18-24 hours. After incubation time, the medium was replaced with complete DMEM. The next day, supernatant containing the virus was collected and replaced with fresh medium. The collected viral supernatant filtered using 0.45 µm filter. The collected supernatant can be used as a crude form or can be concentrated using ultracentrifugation. For ultracentrifugation, small rotor (11.5 ml) or large rotor (33 ml) can be used using ultracentrifuge tubes. The virus was ultracentrifuged at 4 °C for 2-2.5 hours at 23,000 rpm (~90,000 xg). The supernatant was carefully removed and 150 µl of Opti-MEM (or serum free media / PBS) was added to the tubes and resuspended multiple times. The concentrated virus was incubated on ice for one hour, aliquoted in Eppendorf tubes and transferred to – 80 °C for long term storage till further use.

**Table 2.23: List of transfection reagents.**

Transfection/Transduction Reagents	Supplier name
GeneJuice™ (Transfection)	Merk Millipore (Cat# 70967)
Polybrene (Transfection/Infection)	Sigma-Aldrich (Cat# TR-1003-G)



### 2.7.12. H1V-1 -derived lentiviral production using HEK293 cells and GeneJuice™

For transfection of HEK293 cells using GeneJuice™ reagent,  $1.5 \times 10^7$  cells were seeded in one T175 flask and incubated at 37°C and 5% CO<sub>2</sub> overnight. One day following seeding, cells were 80% confluent and transfected with 66 µl of GeneJuice™ in 12 ml of filtered Opti-MEM containing 32 µg of plasmid DNA (pHR'SINcPPT-SFFV-eGFP-WPRE, pCMVΔR8.74 and pMD.G2 at a ratio of 4:3:1). The mixture was incubated at room temperature for 20 minutes before adding to the cells. The cells were incubated with the virus for 18-24 hours. Following incubation, transfection medium was replaced with fresh complete DMEM medium. The collected supernatant was filtered using 0.45 µm filter and stored at 4°C. This step continued for three days (72 hours). At the final step, all the collected filtered viral supernatant can be used as crude or further concentrated using ultracentrifugation.

## 2.8. Transduction

### 2.8.1. Transduction of JHP106i hiPSCs

In this study, hiPSCs were transduced using H1V-1 derived lentiviral and recombinant (rAAV2) viral vectors (for details of each construct, refer to figure 2 and table 14). One day prior to transduction,  $3 \times 10^5$  iPSCs were seeded in pre-coated laminin-521 6-well plate. The following days, one well of the cells were dissociated using Gentle Cell Dissociation reagent (StemCell Technology, UK) and counted using haemocytometer. Based on the cell numbers and virus titre (vg/ml), multiplicity of infection (MOI) can be calculated using the following equation:

$$\left( \frac{\text{Number of virus particles per cells} \times \text{number of the cells}}{\text{virus titre}} \right) \times \text{volume of media}$$

For lentiviral transduction, mTeSR1 medium containing 10 µM Y-27,632 (Calbiochem) and polybrene reagent (Sigma-Aldrich, UK) was prepared. The virus added to the medium and the mixture was incubated for 20 minutes at room temperature. The medium from the cells removed and replaced with the medium containing the virus and rock inhibitor. The plate was gently shaken to allow even distribution of the medium and maximum viral transfer efficiency. The plate was incubated at 37°C for 24 hours. Following day, the transduction medium was replaced with fresh complete mTesr1 medium. This step was continued for three days and medium refreshed daily.

Next, the cells were extracted for flow cytometry analysis to determine the number of GFP positive cells. The remaining wells were imaged using fluorescent microscopy and harvested for DNA/RNA extraction. One well of un-transduced cells were collected for control sample. For recombinant AAV<sub>2</sub> (rAAV<sub>2</sub>) vectors, the same protocol followed, however no transduction agent added to the mTeSR1 medium with no incubation time required.

### **2.8.2. Transduction of 3D heps**

In this study, on day 30 of differentiation 3D heps were transduced using HIV-1 derived lentiviral and recombinant (rAAV<sub>2</sub>) viral vectors (for details of each construct, refer to figure 2 and table 14). One day of transduction, William's E medium (Life Technology, UK) supplemented with 10  $\mu$ M Y-27,632 (Calbiochem) , 10 ng/ml EGF (R&D Sytems), 10 ng/ml VEGF (R&D Systems), 10 ng/ml HGF (PeproTech), 10 ng/ml bFGF (PeproTech) , polybrene transduction reagent (Sigma-Aldrich) and the virus was prepared. The mixture was incubated for 20 minutes at room temperature before adding to the cells. Based on the cell numbers and virus titre (vg/ml), multiplicity of infection (MOI) can be calculated using the equation as described earlier.

The 12-well plate containing 3D heps were gently tilted to remove the old medium. The transduction medium added to the cells and transferred to the incubator while tilted to enhance maximum viral transfer efficiency. Following 24 hours incubation, transduction medium was removed and replaced with complete William's E medium with essential growth factor supplements. Following transduction, transduced 3D heps were kept 3 to 7 days until green cells were appeared. Cells were imaged using fluorescent microscopy and collected for DNA/RNA extraction. One well of un-transduced cells were chosen as control. For recombinant AAV<sub>2</sub> (rAAV<sub>2</sub>) vectors, the same protocol followed, however no transduction agent added to the William's E medium with no incubation time required.

### **2.8.3. Long-term outgrowth 3D culture for genotoxicity**

For long-term culture genotoxicity assay, 3D heps were transduced on day 30 of differentiation using HIV-1 derived lentiviral and rAAV<sub>2</sub> vectors as previously described. Transduced 3D heps were maintained in the incubator at 37 °C, 5% CO<sub>2</sub> using 3DM supplemented with growth factors (table 2.16) for six months. The medium was changed with freshly prepared supplemented medium every other day.

## **2.9. Single-Cell Cloning (SCC) of the transduced hiPSCs**

Considerations prior to SCC procedure:

- 1) Stem cell lines tend not to grow efficiently in sparse or individual cultures due to the lack of essential growth factors and secreted proteins.
- 2) Conditioned medium (CM) is a key to have successful SCC.
- 3) Stem Cell pool should be checked properly for positive GFP expression to ensure every cell in culture, in theory, has been transduced.
- 4) This protocol has been optimised for adherent 2D monolayer stem cells only.

### **2.9.1. Cloning efficiency**

Cloning efficiency was calculated using uninfected hiPSCs and transduced iPSCs with lentiviral vectors prior to SCC. Ten clones from each condition was seeded and monitored for survival and cloning efficiency.

### **2.9.2. Preparation of conditioned medium**

Un-transduced hiPSCs with healthy colonies were seeded routinely in pre coated 6-well plates for three consecutive weeks as previously described. Cells supernatant was harvested daily from the cells and replaced with fresh complete mTeSR1 (StemCell Technology). The collected supernatants were filtered using 0.45 µm filter, aliquoted in 15 ml sterile falcon tubes and stored in -20 °C for 3 months.

### **2.9.3. Single cell cloning**

Prior to SCC, hiPSCs were transduced with lentiviral vectors as previously described. On day of SCC complete mTeSR1 (StemCell Technology) medium with CM at a ratio of 1:1 was prepared. The medium was supplemented with 10 µM Y-27,632 (Calbiochem) to enhance the cell survival. A pre-coated 24 well plate was incubated at 37°C for 30 minutes. Transduced positive GFP cells were washed with PBS (Sigma-Aldrich) once and gently dissociated using Gentle Cell Dissociation reagent (StemCell Technology) for 15 minutes. The single cells were resuspended in mTeSR1 and Y-27632 and counted by haemocytometer. To achieve low volume of homogenized cell solution, a relative dilute solution was required.

This is shown in following steps as an example:

Homogenised cell solution concentration count:  $5 \times 10^3$  cells /ml

The volume required to seed a 24 well plate: 12 ml of 2 cells/ml

Total cells needed: 12 ml×2 cells/ml= 24 cells

Volume of homogenised cell solution required which corresponds to 48 cells in 12 ml is as follow:

$$\{(24 \text{ cells})/ 5 \times 10^3 \text{ cells/ml}\} = 4.8 \mu\text{l}$$

To make the final 2 cells/ml, 4.8  $\mu\text{l}$  of cell solution was transferred to 12 ml of complete/conditioned medium and 500  $\mu\text{l}$  from the cell suspension was dispensed per well of a 24-well plate. This was to ensure the plate was seeded at a density of 1 cell/well. Following seeding, the cells were undisturbed for 7 to 10 days. After 7 days, the plate was scanned for colonies. The cells from each colony were expanded and harvested for DNA/RNA extraction.

## 2.10. Statistics and Bioinformatic

Data were analysed by GraphPad Prism (version 7). I did not apply any statistical methods to predetermine sample size. The results represent the mean +/- SD of three individual biological samples per time point. Unpaired t-test is a statistical technique that compares the means of two different groups to determine level of significance. The Mann-Whitney U test was applied to evaluate degree of difference in the dependent factor for two unrelated set of data which was set at  $P < 0.05$ . One-Way ANOVA was used to compare the means of two or more independent sets to measure the level of significance. The levels of significance were measured by One-Way ANOVA whereby  $p < 0.05$  denoted as \*,  $p < 0.01$  denoted as \*\* and  $p < 0.001$  denoted as \*\*\*.

Bioinformatic analysis was performed in collaboration with GeneWerk (Germany) and a colleague at University of Brunel. Insertion Site (IS) analysis was retrieved, using solid-phase ligation-mediated PCR (EPTS/LM-PCR). For IS site analysis on bulk P106i and 3D heps transduced with lentiviral vectors, the quantity of IS has been normalised per 1 million sorted reads that the machine processed. This is how GeneWerk normalised the data and it does not alter the data. The level of significance was measured by unpaired T-test. An Unpaired T-test is a statistical procedure that compares the means of two unrelated groups and to determine if there is a significant difference between those groups.

The top 10 IS analysis in P106i cells transduced with lentiviral vectors and overtime retrieved by EPTS/LM-PCR with primers specific to HIV-1 LTR viral vectors. This processed bioinformatic data was provided by GeneWerk. To analyse, they mapped the IS within 100KB of that IS and relative sequence count (RSQ) analysis was used to quantify the number of the counts of integrated lentiviral vectors. The RSQ is the division of absolute sequence count (ASQ, i.e. The raw number of viral sequences detected per IS) by the total number of viral sequences detected. Therefore:  $RSQ = [ASQ / \text{Total number of viral sequences}]$ . For each sample number, two replicates were analysed and are presented in parallel to each other.

For genotoxicity, it is crucial to assess the total IS in genes found in P106i and 3D heps transduced with lentiviral vectors and compare them. This data was a comparison of top 10 genes between lentiviral vectors, time points and tissue types. This processed bioinformatic data was provided by GeneWerk, using RSQ analysis.

Gene Ontology was applied to performed gene enrichment analysis on gene sets, using the raw data provided by GeneWerk. The Uniprot software was used to identify the total number of cancer-related genes CIS in P106i and 3D heps transduced with lentiviral vectors. Following identification, samples were compared using Venn diagram. Venn diagram is a widely used tool which represents graphical depiction of the unions and intersections among multiple data sets. Further, the functional enrichment analysis tool (FunRich) was applied to identify biological pathways for each tissue type. All the insertions identified were significant and from three biological replicates.

## Chapter 3

### 3.1. Introduction

Gene therapy (GT) has evolved as a medical reality in the last two decades. This technology is a powerful method to introduce recombinant genetic elements into human cells. This method holds a considerable therapeutic potential for the treatment of wide range of pathological conditions, such as genetic disorders, neurological complications, diabetes, cancer and infectious diseases. To date, twenty gene therapy products are successfully approved and over two thousand human clinical trials have been reported worldwide. These exceptional results provided confidence to treat inherited rare diseases. The first set of gene therapy initiated with hematopoietic stem cells (HSCs) and the aim was to treat a broad range of human diseases such as inherited haematological disorders (Romano, 2003, Edelstein et al., 2007).

Over the past two decades, hematopoietic stem cells (HSC) have been used for gene therapy based clinical trials. They were employed for the treatment of severe inherited diseases such as X-linked severe combined immunodeficiency (SCID-X1), adenosine deaminase deficiency (ADA-SCID), X-linked chronic granulomatous disease (X-CGD), X-linked adrenoleukodystrophy (X-ALD) and Wiskott-Aldrich syndrome (WAS). Details of each clinical trial was covered in Chapter 1, Section 1.1.16. Gene therapy field faced a major setback in 2003 when clonal vector mediated leukaemias were reported in several patients during SCID-X1 trial. Therefore, it is important to develop a sensitive model to evaluate viral vector mediated genotoxicity prior to gene therapy treatment.

#### 3.1.1. Integrating viral vectors and Insertional Mutagenesis

For gene therapy treatment, different types of gene transfer systems were explored. The most common gene transfer systems were derived from viruses such as retroviruses (RV), ADVs, AAVs and other non-viral vector systems. For permanent transgene expression, shuttle vector is integrated within the target cell's chromosomal DNA. However, integrative gene transfer systems have few pitfalls such as risk of IM. This may eventually lead to development of cellular transformation and malignancies (Romano, 2012). Therefore, it is extremely important to develop safe viral vectors to minimise the risk of vector mediated genotoxicity such as the IM event. The IM event is a well-established safety concern of viral vector-based gene therapy and it is more

prevailed in integrating viral vectors such as RVs. By altering the expression of host genes in the proximity of the Insertion Site (IS), the resulting retroviral mediated IM can lead to oncogenesis (David and Doherty, 2017). The mechanisms of IM are including promoter insertion, promoter activation and formation of truncated proteins as a result of virus integration. These mechanisms were explained in greater details in Chapter 1, Section 1.1.13.

### **3.1.2. *in vivo* & *in vitro* genotoxicity models**

Over the years, a number of pre-clinical assays for the assessment of oncogenicity have been developed. Although these systems are valuable, they are applicable only to specific cell lineages, oncogenes and mechanisms of genotoxicity. Therefore, they can be used to detect very specific mechanisms for screening novel vectors rather than estimating the risk for oncogenicity and cancer diseases (Zhou et al., 2013, Rahman et al., 2017). As explained in Chapter 1, Section 1.1.17, IM events were detected in several animal and human cell culture models. The main issues associated with these systems are sensitivity and use of animal models which does not faithfully resemble the human biology. In addition, these tests are costly and time consuming. As GT progresses, there is an urge for more pre-clinical tests to assess the safety of developed therapies and therapeutic strategies. mice models are commonly used for *in vivo* genotoxicity research studies. However, these studies are laborious, costly and do require large number of animals. Therefore, developing alternative *in vitro* models are required to evaluate the risk of IM. A robust and efficient *in vitro/in silico* system is of utmost importance to evaluate the risk of IM/oncogenesis while restricting the usage of animal models.

Due to the disadvantages of the *in vivo* mice models, alternative *in vitro* models were explored. However, these models do still require mouse stem and progenitor cells. There is not an *in vitro* or *in vivo* “gold standard” genotoxicity model available to overcome the current challenges in the field of gene therapy. Therefore, development of a robust model is required to evaluate safety with high sensitivity for each individual vectors or diseases and to identify the relevant risk factors. In addition, this model should facilitate the risk benefit decisions while considering the measurable impact on 3Rs. The model should be applicable to a wide range of vector types and also different target tissues. Using the stable unlimited self-renewal cell lines such as hiPSCs and *in vitro* 3D models are great alternative to mice models or mouse stem cells. These

models may further promote the *in vivo* resemblance and support longer-term culture for genotoxicity assays.

### **3.1.3. Liver and mature hepatocytes**

The liver is one of the largest internal organs which mainly consist of parenchymal cells known as hepatocytes (Rodriguez-Antona et al., 2002). The liver has unique characteristics which makes it a great target for *in vivo* and *ex vivo* gene transfer (Nguyen and Ferry, 2004). One of the reasons for being an attractive target for *in vivo* gene transfer is hepatocytes are readily available through the blood stream. Moreover, the endothelium of hepatic sinusoids composed of 100 nm fenestrations that allow entry of viral particles into hepatocytes. For this reason, over the past decade many studies have focused on developing gene transfer methods to deliver a wide range of viral and non-viral vectors. One of the main focus of this study is to produce functional mature hepatocytes *in vitro* for gene transfer, using different viral vectors. Hepatocytes play a key role in maintaining liver functions and can be isolated from the human liver using collagenase perfusion.

Primary liver hepatocytes are still considered as the “gold standard” *in vitro* model for pharmaceutical drug testing and for other clinical applications. One caveat with primary hepatocytes is the maintenance of an *in vivo* like phenotype. Following isolation and *in vitro* culture, functional phase I and II metabolic activities rapidly declined 72 hours post culture (Rodriguez-Antona et al., 2002). Further, cultured primary hepatocytes may also lose their polarized morphology which affect their ability to excrete biotransformed compounds (Noel et al., 2013). For different cell assays and clinical applications, a prolonged functional life of hepatocytes is important. One promising approach is to transform these cells from 2D monolayer into 3D spheroids. Other cell lines such as immortalised liver cell lines namely HepG2 and HepaRG have been considered for long-term culture applications. The properties of these cell lines were previously explained in Chapter 1, Sections 1.2.2.2 & 1.2.2.3. Although these cell lines are a good alternative to freshly isolated primary hepatocytes, they do not sustain long term *in vitro* functional activity. For these reasons, other cell sources such as PSCs are great alternative to generate high quality human liver tissue for clinical applications.



### 3.1.4. 3D hepatic differentiation

To generate hepatocytes from human PSCs, several protocols have been established to efficiently produce hepatocyte-like cells (HLCs), utilizing 2D differentiation platform systems (Hay et al., 2007, Hay et al., 2008a, Hay et al., 2008b, Agarwal et al., 2008, Sullivan et al., 2010, Hannan et al., 2013, Loh et al., 2014). As previously mentioned in Chapter 1, Section 1.4.6, despite recent advances (Cameron et al., 2015, Wang et al., 2017), 2D-derived HLCs display foetal features such as high AFP and low ALB secretion and have a transient phenotype *in vitro*, limiting their clinical and industrial applications (Rashidi et al., 2018).

To generate functional and stable *in vitro* human liver tissue, 3D approaches have been explored (Szkolnicka and Hay, 2016). These systems utilise scaffold-driven formation of 3D aggregates or different matrixes with or without addition of other cell types (Takebe et al., 2013, Gieseck et al., 2014, Camp et al., 2017). The scaffold-free 3D methods were covered in detail in Chapter 1, Section 1.4.3. The 3D spheroids can be differentiated to any tissue type which can recapitulate some aspects of *in vivo* environment.

In this study, hiPSC-derived 3D human liver tissue was generated using an optimised and stepwise differentiation protocol, as previously described (Rashidi et al., 2018, Lucendo-Villarin et al., 2019). This protocol addresses issues surrounding previous 3D protocols such as scalability and long-term *in vitro* phenotypic stability. Notably, hiPSC-derived 3D liver cells displayed mature liver functions for an extended period of over one year in culture.

## 3.2. Objectives

The objectives of this chapter are as follows:

- Expansion, banking and full characterisation of an integration-free hiPSC cell line
- Optimising of the novel 3D *in vitro* culture system to generate hepatic cells from hiPSCs
- 2D and 3D hepatic differentiation using the hiPSC cell line and stepwise differentiation to generate hepatocytes
- Full characterisation of hiPSCs-derived 3D heps

## **3.3. Results**

### **3.3.1. Characterisation of the hiPSCs population**

For successful hepatic differentiation, it is crucial to maintain the proliferation and pluripotent nature of the stem hiPSCs. The latest advances in a fully defined culture system technology have enabled standardised, scalable and highly efficient methods to culture cells. These improvements have supported proliferation, self-renewal and pluripotency of the stem cells. Therefore, this method was used to culture and maintain iPSCs followed by full characterisation prior to hepatic differentiation.

### **3.3.2. Culture and characterisation of hiPSCs maintained in mTeSR1™**

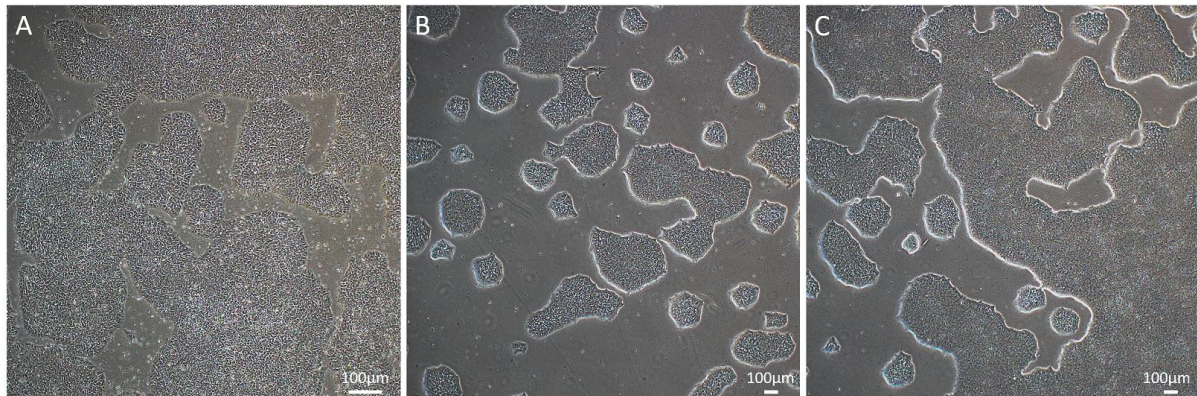
To characterise hiPSC cell lines such as 33D6 and P106, these lines were routinely thawed and cultured using laminin-521 coated plates. The cells were sustained using standardised, feeder-free mTeSR1™ medium. Further, this medium is serum-free and highly specialised which overcame the necessity to use feeder layers. It is argued that commercial mTeSR1™ medium provides more consistent cultures and stable undifferentiated phenotypes (International Stem Cell Initiative et al., 2010).

In this study, frozen hiPSCs were initially retrieved from LN and maintained in mTeSR1™ medium for minimum of five passages prior to characterisation such as flow cytometry and pluripotency capacity. The identity of iPSCs was evaluated using morphological observation, detection of the expression of transcription factors ascribed to pluripotency; SOX2, Octamer 4 (OCT3/4), and expression of cell surface markers such as SSEA-1, SSEA-4, TRA-1-60 and TRA-1-81.

### **3.3.3. Morphological analysis of healthy and unhealthy colonies**

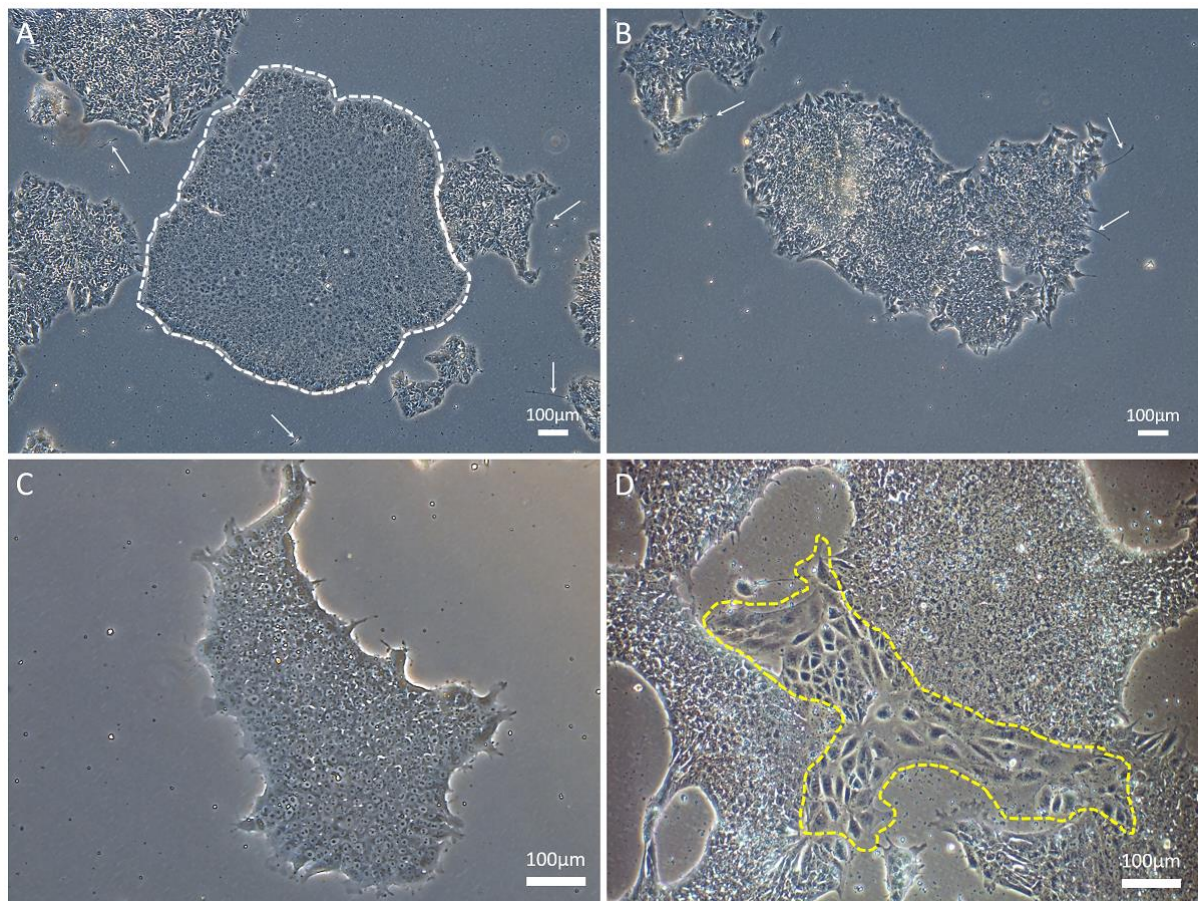
To determine morphology of healthy and unhealthy colonies of iPSC lines, morphological analysis was required. Morphological observation of the hiPSCs is a great tool to assess culture homogeneity and detect spontaneous differentiation. For this purpose, PSCs were grown routinely in laminin-521 coated plates and maintained in mTeSR1™ complete media. Stem cell colonies were maintained using this complete medium. Healthy colonies were distinguished by having a very tightly packed dome-

like structure, very well-defined edges and no spontaneous cellular differentiation (Figure 3.1).



**Figure 3.1: Morphology of hiPSC colonies.** hiPSC cell lines were routinely thawed and cultured using laminin-521 coated plates and maintained in mTeS1™ medium. A) The morphology of hiPSC P106i line 24 hours following thawing from liquid nitrogen. The colonies displayed loose structures with no clear and tight borders. B) Following a few passages, the colonies formed dense and compact structures with clear borders. The colonies were cultured using laminin-521 matrix and maintained in mTeS1™ medium. C) Morphology of healthy colonies reaching 75-80% confluency before dissociation and subculturing. Abbreviations: Human induced pluripotent stem cells; hiPSCs. Scale bars: 100 µm.

Using high magnification microscopy, cells displayed the characteristic of pluripotent stem cell morphology such as a large nucleus to cytoplasm ratio, well defined borders and pronounced nucleoli (Figures 3.2A & C). Cultures were routinely cleaned to remove spontaneously differentiated cells, allowing healthy colonies to expand. The routine procedure was to mark each unhealthy colony and aspirate during culture medium change. For optimal hepatic differentiation, it is crucial to maintain PSCs and to remove unwanted colonies. Failing to maintain this state allows the promotion of unhealthy colonies to gradually overtake the culture (Figures 3.2B & D).



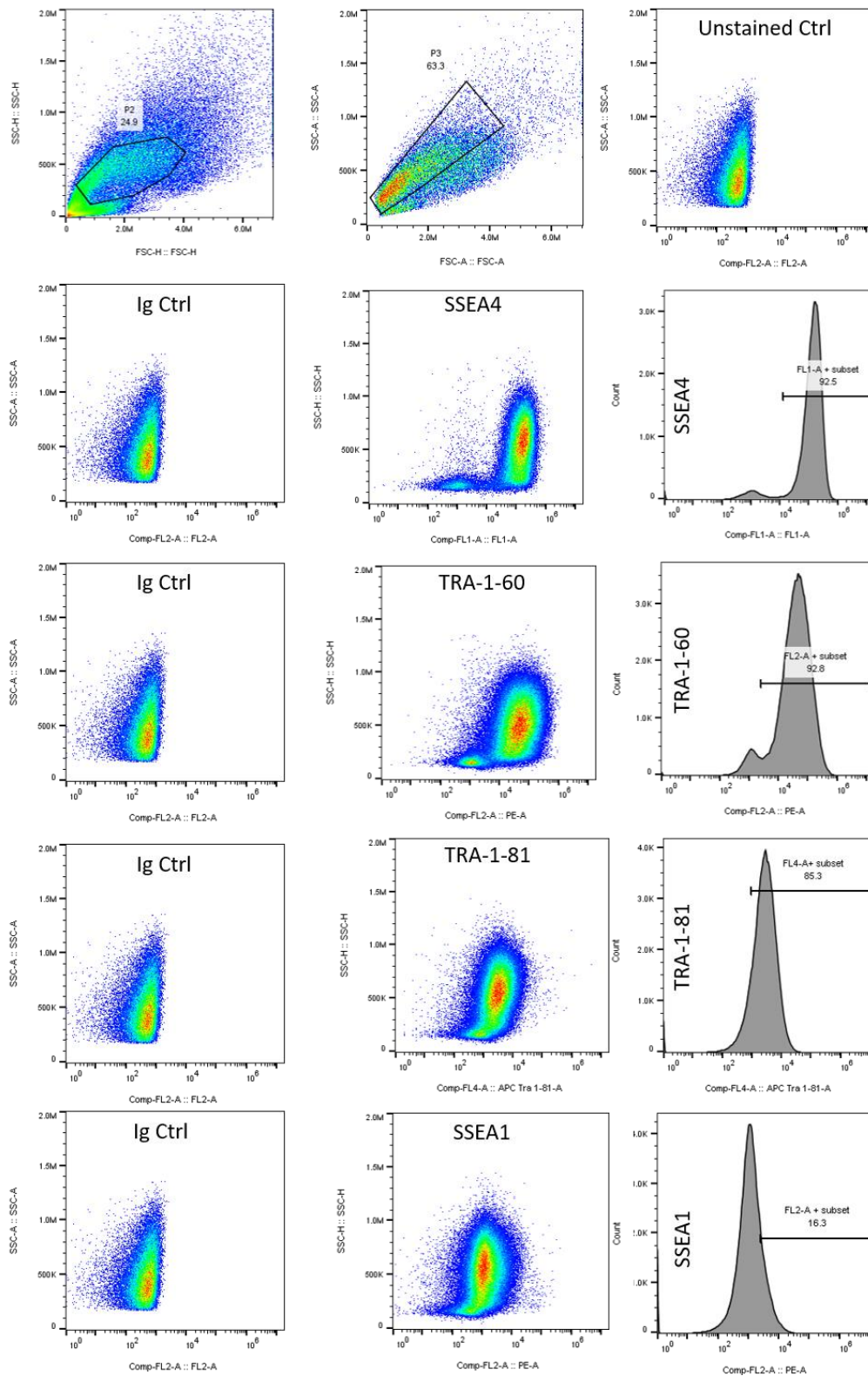
**Figure 3.2: Morphology analysis of different types of hiPSCs colonies. Morphological observation of the hiPSCs is crucial to assess culture homogeneity and detection of spontaneous differentiation. A) Morphology of a healthy colony indicated by white dashed line surrounded by colonies with spontaneous differentiation, indicated by white arrow. Unhealthy colonies were routinely marked and aspirated from the culture to preserve the pluripotent healthy colonies. B) Morphology of an unhealthy colony with full spontaneous differentiation. As it is seen in the figure, different types of cells reside in the colony and makes them unsuitable for differentiation purposes. Neuronal axons are indicated by white arrow and these cells are no longer considered as pluripotent. C) Healthy colonies were distinguished by having a packed dome-like structure, very well-defined edges and no spontaneous cellular differentiation. D) Morphology of a partially unhealthy colony with spontaneous differentiation. The presence of such colonies promotes gradual destruction of the structure of healthy colonies, breaking the tight pack borders and complete change of the pluripotent phenotype. Scale bar: 100 $\mu$ m.**

### **3.3.4. Flow cytometry analysis**

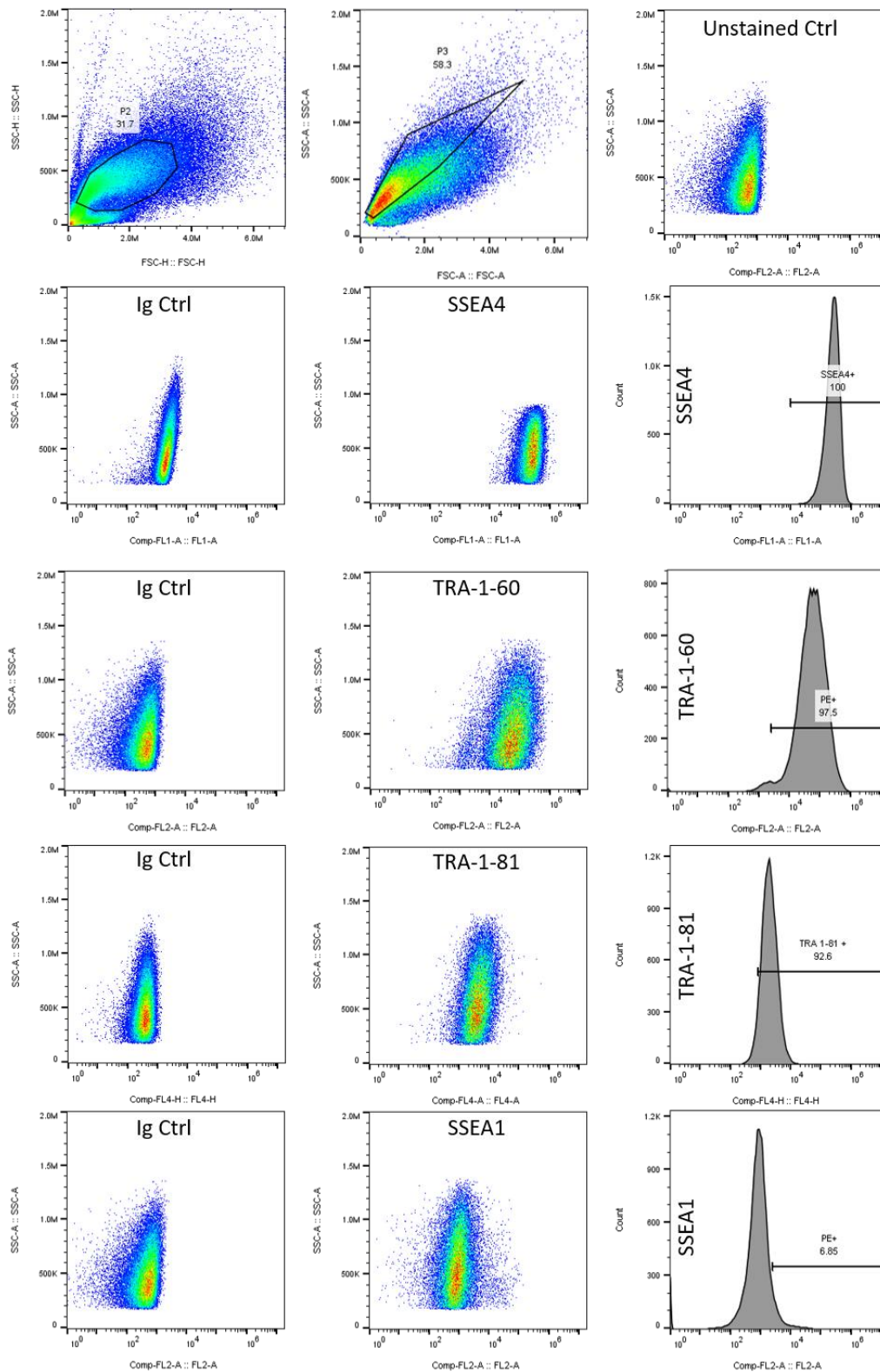
In addition to the morphological observations, flow cytometry analysis was employed as a routine procedure to confirm pluripotency and stem cell identity of the cells. In this method, expression of cell surface markers such as SSEA-4, TRA-1-60 and TRA-1-81 were quantified by flow cytometry. The SSEA-1 was used as a negative control as it uniquely marked stem cell differentiation. For each flow cytometry analysis, cells were cultured and maintained in mTeS1™ complete media for a minimum of 5 passages and at different passage ranges (p5-10, p20-30). The cells were gently dissociated and incubated with an appropriate antibody at room temperature before flow cytometry analysis. Unstained cells were used for gating side scatter in relation to forward scatter. Additionally, dead, differentiated and cell aggregates were eliminated. In this study, the expression of hiPSCs surface markers in passage numbers 15 (P15) and 17 (P17) were measured. In P15, the percentage of cell expressing SSEA-4 (92.5%), TRA1-60 (92.8%), TRA1-81 (85.3%) and SSEA-1 (16.3%) and in P17 the percentage of cell expressing SSEA-4 (100%), TRA1-60 (97.5%), TRA1-81 (92.5%) and SSEA-1 (6.85%) confirming the undifferentiated and pluripotent state of the cells (Figures 3.3, 3.4).

### **3.3.5. Stemness nature of hiPSCs**

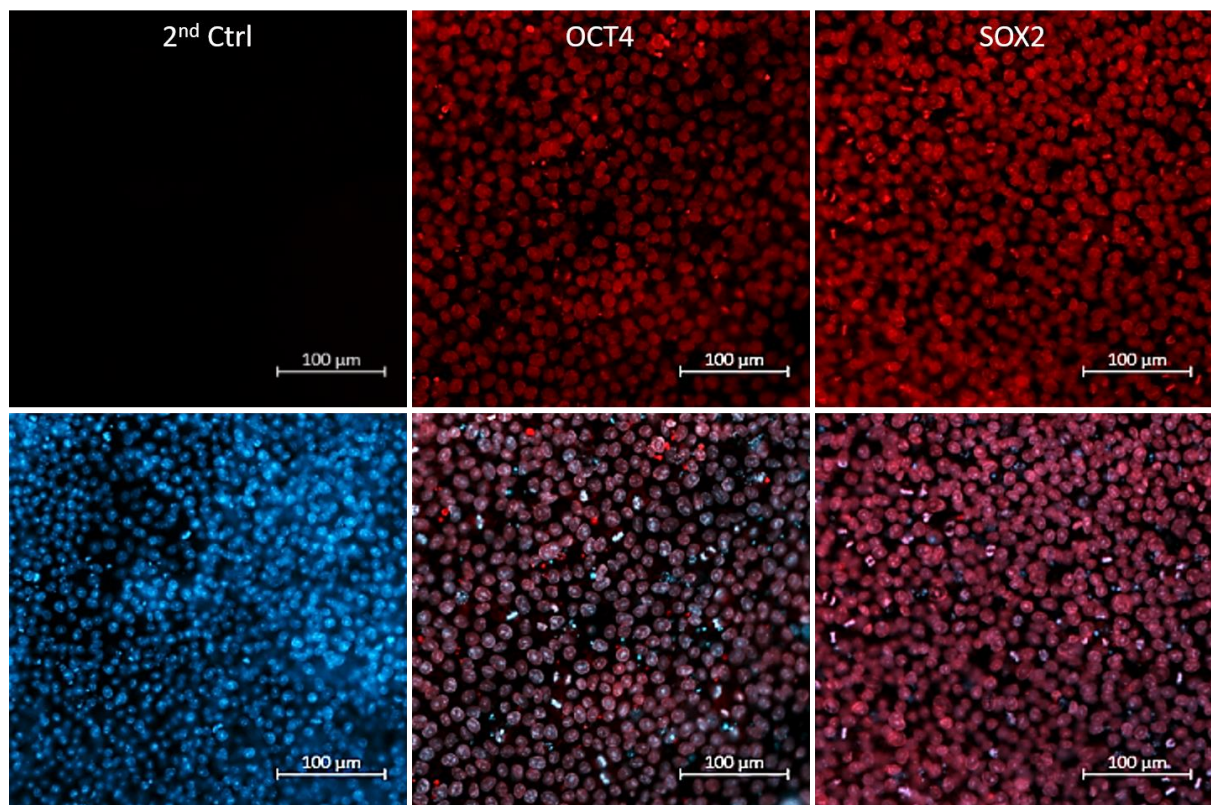
To determine the stemness, the protein expression of two well-known transcription factors OCT4 and SOX2 were analysed. SOX2 is a member of the SoxB1 transcription factor family which is a key transcriptional regulator in pluripotent stem cells. Also, together with OCT4, has a key role in gene expression and maintaining pluripotency of the cells. The protein expression, as depicted by the immunofluorescence analysis, exhibited that hiPSCs, were stained positive for both pluripotent markers. The high level of expression of OCT4 and SOX2 in addition to the flow cytometry analysis confirmed the pluripotent state and identity of hiPSCs (Figure 3.5).



**Figure 3.3: Expression of stem cell surface markers in hiPSCs passage 15.** The graph shows the percentage of cells expressing pluripotent surface markers, including SSEA-4, TRA1-60 and TRA1-81. A low percentage of cells expressed the differentiation marker SSEA-1, confirming the pluripotent nature of hiPSCs. Unstained cells were used to exclude the background noise.



**Figure 3.4: Expression of stem cell surface markers in hiPSCs passage 17. The graph shows the percentage of cells expressing pluripotent surface markers, including SSEA-4, TRA1-60 and TRA1-81. A low percentage of cells expressed the differentiation marker SSEA-1, confirming the pluripotent nature of hiPSCs. Unstained cells were used to exclude the background.**



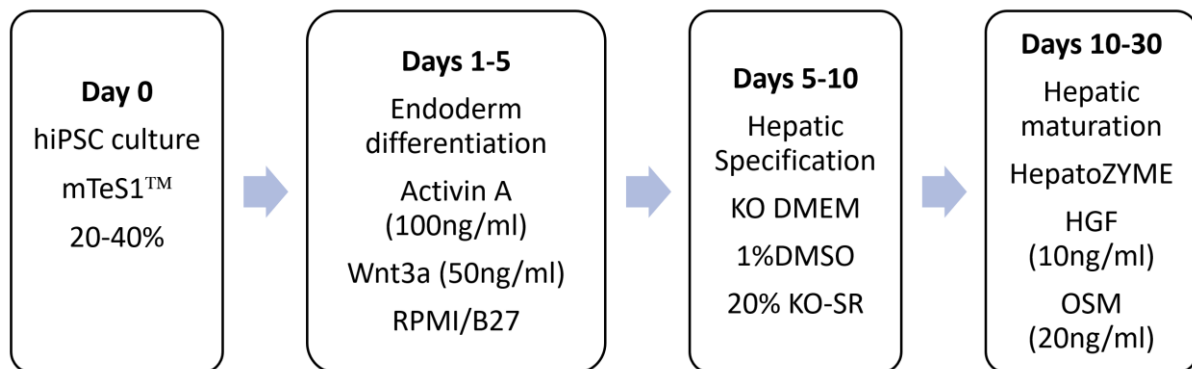
**Figure 3.5: Stemness analysis of hiPSCs population.** To determine the pluripotent identity of hiPSCs, the protein expression of two major pluripotent transcription factors OCT4 and SOX2 were analysed. Immunofluorescence analysis exhibited that all the cells stained positive for both markers, confirming the pluripotent identity of hiPSCs. The nucleus of the cells was counterstained with DAPI. Top panel (left-right); Secondary IgG control, OCT4 (Red), SOX2 (RED). Bottom panel (left-right); DAPI (Blue), merged with DAPI, merged with DAPI. Abbreviations: Sex determining region Y-box 2; SOX2, octamer-binding transcription factor 4; OCT4. Scale bar: 100  $\mu\text{m}$ .

### **3.3.6. 2D hepatic differentiation of human induced pluripotent stem cells, using a serum-free approach**

To generate pluripotent stem cell-derived hepatocytes, development of robust and reliable differentiation protocols is essential. Undifferentiated hiPSCs were maintained in mTeS1™ medium, scaled up and differentiated using an efficient and robust stepwise differentiation protocol which is described previously (Cameron et al., 2015). To initiate definitive endoderm (DE) stage, hiPSCs were primed using RPMI/B27 medium, Wnt3a and Activin A. For induction of hepatic specification, a media containing Knock-Out Serum Replacement (KOSR) and DMSO (SR/DMSO) was used, resulting in formation of liver progenitor cells called hepatoblasts. Finally, for

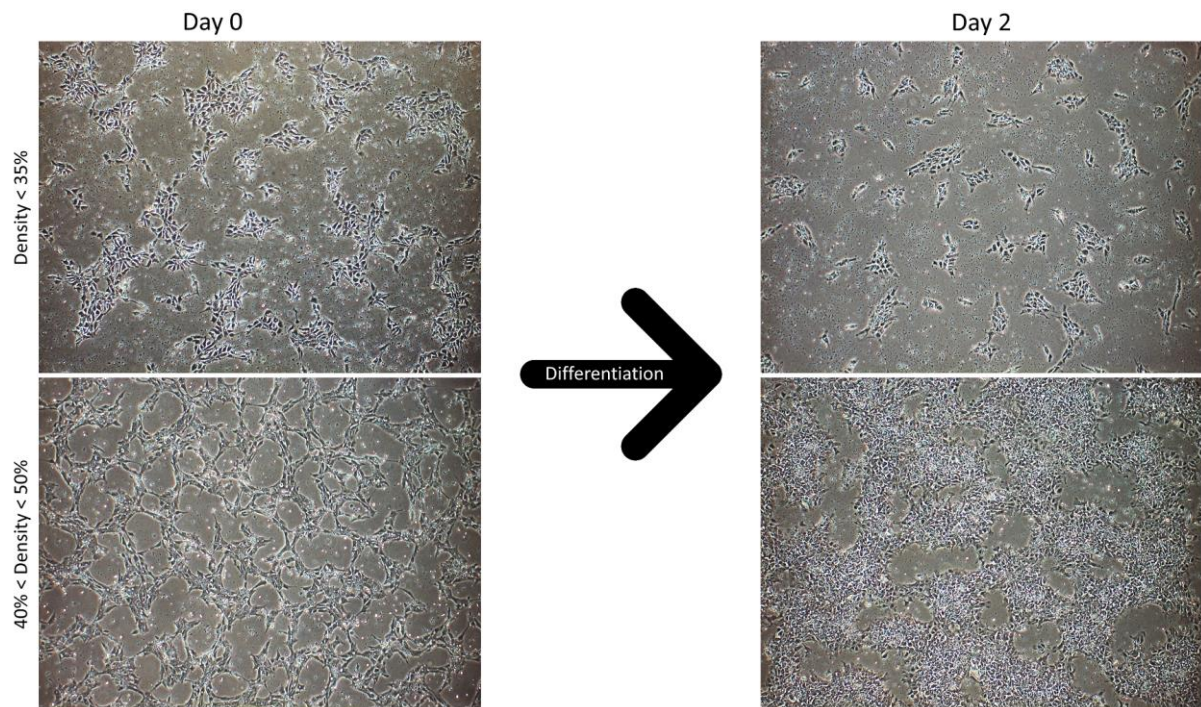


hepatocyte differentiation and maturation the hepatic specification media was replaced with a defined serum-free media containing OSM and HGF (Figure 3.6). In this study, morphological analysis, transcriptional and functional analysis were carried out at different time points. These experiments were necessary to determine the efficiency of hepatic differentiation at each step and to ensure the end products were obtained accurately.



**Figure 3.6: Flow diagram of the hepatocyte differentiation protocol in serum-free media. 33D6 hiPSCs were differentiated to mature HLCs using an efficient and stepwise differentiation protocol for up to 27 days. Abbreviations: hiPSCs- human induced pluripotent stem cells; Knock Out Dulbecco's Modified Eagle Medium KO DMEM; Dimethyl sulfoxide (DMSO); Knockout Serum Replacement (KOSR); Hepatocyte Growth Factor (HGF); Oncostatin M (OSM).**

In this study, initially 33D6 iPSC line was employed for 2D hepatic differentiation. To improve 2D hepatic differentiation of hiPSCs, the protocol was further optimised by an additional two days exposure to media containing Activin A only during endoderm differentiation. For successful hepatocyte differentiation, correct size density plays a key role. Following seeding, if the cells are too sparse (low density) they do not resist the first stage of differentiation. On the other hand, lower differentiation efficiency was observed if the cells are densely seeded. Thus, controlling the seeding density for consistent hepatic differentiation monolayer is critical (Figure 3.7). The most efficient hepatic differentiation was achieved using an initial seeding density of  $1 \times 10^5$  per  $\text{cm}^2$ .

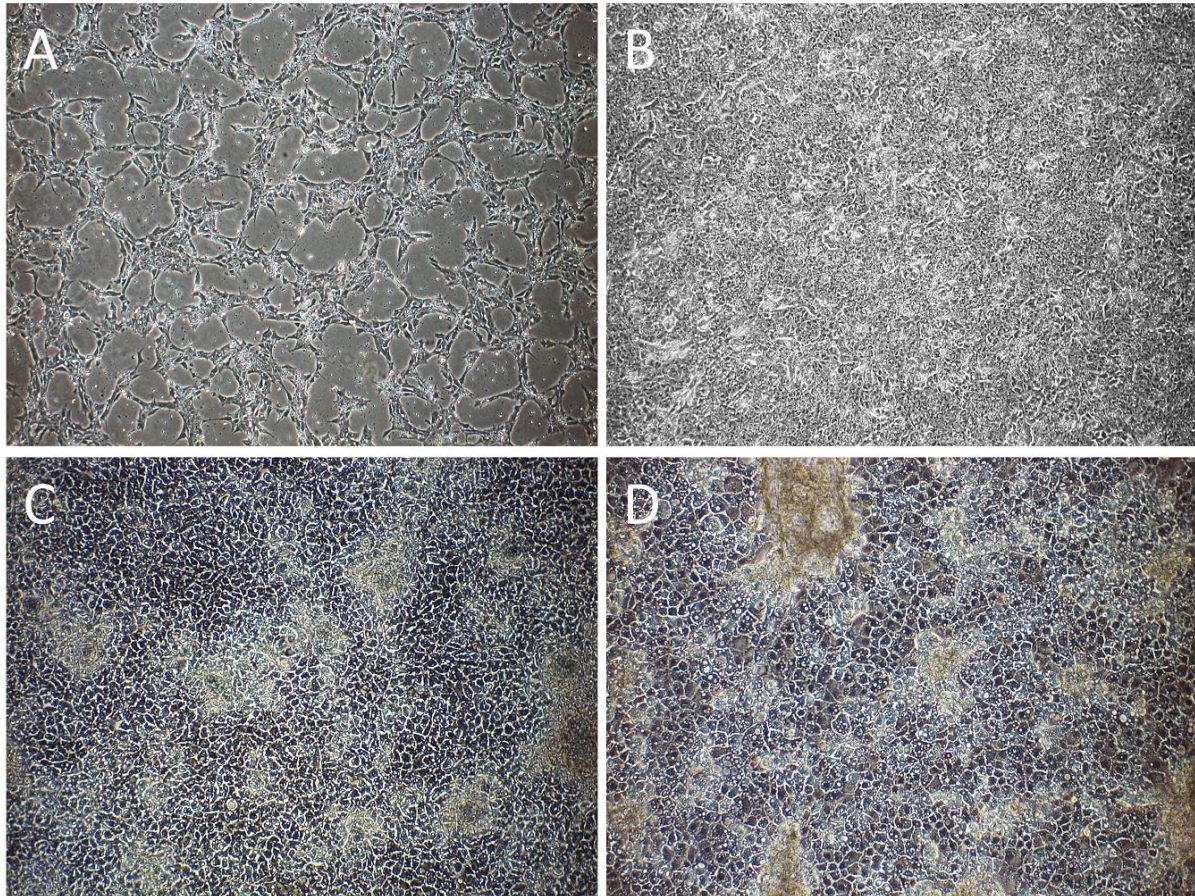


**Figure 3.7: Morphology analysis of different cell density on 33D6 hepatic differentiation. Top panel: On Day 0 of differentiation, seeding the cells with less than 35% resulted in a very sparse culture. The cells did not resist the first phase of differentiation and ultimately were lost. Bottom Panel: Seeding the cells with 40-50% density displayed a uniform culture and successful continuation of hepatocyte differentiation.**

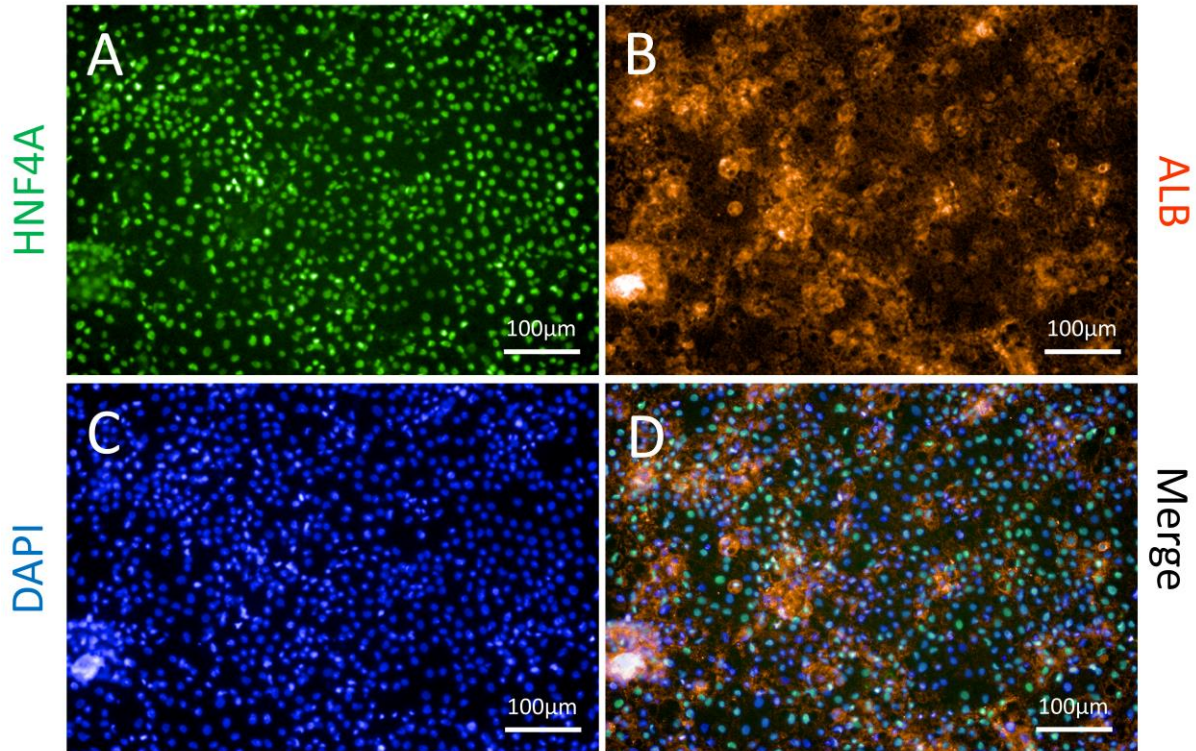
Morphology analysis of hepatic differentiation of 33D6 iPSC line is divided into three main phases including DE, hepatoblast and mature HLC stages. During the DE stage (Figure 3.8A), cells were dividing rapidly and following 72 hours in culture triangular shaped endodermal cells were observed (Figure 3.8B). For the next phase, the 33D6-derived hepatoblast cells displayed cobblestone-like morphology with clear and distinct edges (Figure 3.8C). Hepatocyte specification of mature HLCs at day 20 of differentiation indicates morphological changes which is more distinct compared to hepatoblast stage. A homogenous population of cells with hexagonal shape, clear borders and very defined large nuclei, displaying *in vivo* mature hepatocyte features (Figure 3.8D).

Differentiated cells were characterised using immunostaining and liver functional assays including quantification of ALB and AFP secretion and cytochrome P-450 metabolic activity (Wang et al., 2017). The immunostaining data for protein expression of HNF4-alpha (HNF4A) and ALB demonstrated that the majority of the cells were stained positive for these markers on day 20 of differentiation (Figure 3.9). The

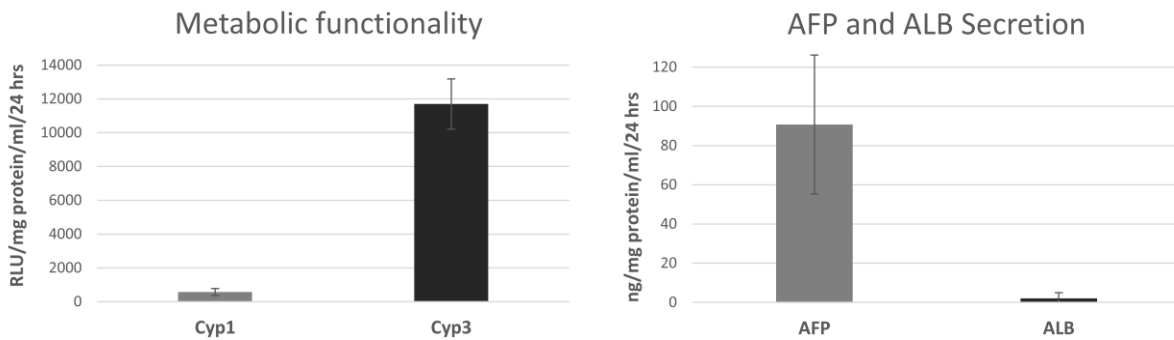
metabolic functionality of HLCs were measured, using the cytochrome P450 (CYP) activity assay. The level of CYP3A was much higher than CYP1A with  $11 \times 10^3$  and  $1 \times 10^3$  RLU/mg/ml/24hrs, respectively. To quantitatively measure the level of protein secretion, the amount of ALB and AFP were quantified using ELISA assay. The result showed a high level of AFP and a very low level of ALB production (Figure 3.10).



**Figure 3.8: Morphology of 2D hepatic differentiation of 33D6-iPSC line. A) Day 0, cells were seeded with a density of 40% prior to differentiation. B) DE stage, cells were dividing rapidly and following 72 hours in culture triangular endoderm shape was observed. C) Hepatoblast stage, displayed cobblestone-like morphology with clear and distinct edges. D) Maturation of HLCs, a homogenous population of cells with hexagonal shape, clear borders and very defined large nuclei, displaying in vivo mature hepatocyte features. Abbreviations: Definitive Endoderm (DE), Hepatocyte-like Cells (HLCs).**



**Figure 3.9:** Standard characterisation tests were performed on day 20 mature hiPSC-derived HLCs. Immunofluorescence analysis of hiPSC 33D6 derived HLCs displayed that positive expression of HNF4A (A) and ALB (B). Nucleus counterstained with DAPI (C). Scale bar: 100 µm.



**Figure 3.10:** Liver functionality tests were performed on 33D6-derived mature HLCs. On Day 20 of differentiation, metabolic functionality of mature HLCs were measured, using the cytochrome P450 (CYP) activity assay. The level of CYP3A was much higher than CYP1A with  $11 \times 10^3$  and 1000 RLU/mg/ml/24hrs respectively. The secretion of serum ALB and AFP was measured using ELISA assay. The result showed a high level of AFP and a very low level of ALB production indicating a more foetal phenotype. The data was normalised versus mg protein per ml. Abbreviation: Albumin (ALB), Alpha fetoprotein (AFP).

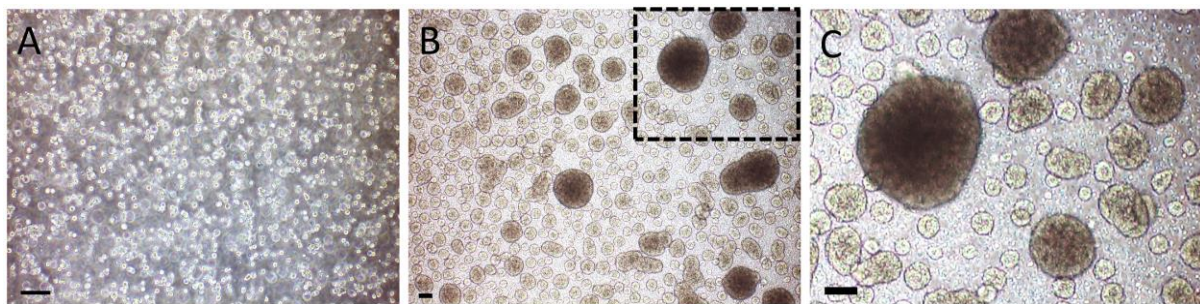
Following optimisation of the step-wise differentiation protocol using the 33D6 hiPSC line and generation of functional human mature HLCs from hiPSCs (Wang et al., 2017), an integration-free iPSC line (P106i) was purchased to study viral vectors insertion sites at the pluripotent stage and in hiPSC-derived HLCs. The 33D6 cell line was originally reprogrammed with five retroviral vectors and was not fit for this purpose. Further, for genotoxicity assays and to see the effects of viral insertions on gene expression a long-term culture with stable phenotype is of utmost importance. From 2D differentiation data, mature HLCs were functional in culture for a maximum of 10 days before the cells trans-differentiated and lost their hepatic function. For this reason, I was involved in developing and optimising a 3D platform to generate hESC-derived 3D heps, which maintained their hepatic phenotype and function for more than a year in culture (Rashidi et al., 2018, Lucendo-Villarín et al., 2019). The 2D hepatic differentiation protocol was used to differentiate hiPSCs-derived spheroids into mature 3D heps. The 3D heps were extensively characterised at different time points of differentiation stages. Morphological analysis of the cells, gene expression, protein expression and liver functional analysis were performed and will be explained in the following sections.

### **3.3.7. Development and optimisation of 3D heps for genotoxicity assay**

The majority of the cell-based assays have employed 2D monolayer culture systems in plates or flasks. Traditionally, these systems proved to be valuable with some existing limitations. Cells in the *in vivo* environment are embedded in extracellular matrix (ECM) and in close contact with other cells in a 3D configuration. In comparison, the 2D cell culture system is lacking the natural 3D microenvironment architecture with limited ECM and contact to neighbouring cells, therefore, does not truly represent physiological conditions and occasionally might provide non-predictive data. Thus, 3D cell culture systems may better represent the physiology of target tissue and more closely mimic the actual microenvironment. Over the recent years, efforts have been made to develop a variety of 3D culture systems for different purposes such as drug discovery, cell-based therapies and other cell-based analysis. Such 3D systems are excellent *in vitro* models to study cellular responses in a more physiologically relevant setting. In the next section, methods for the generation of hiPSCs 3D spheroids will be explained.

### 3.3.8. Development of the 3D platform to generate hPSC-derived 3D heps

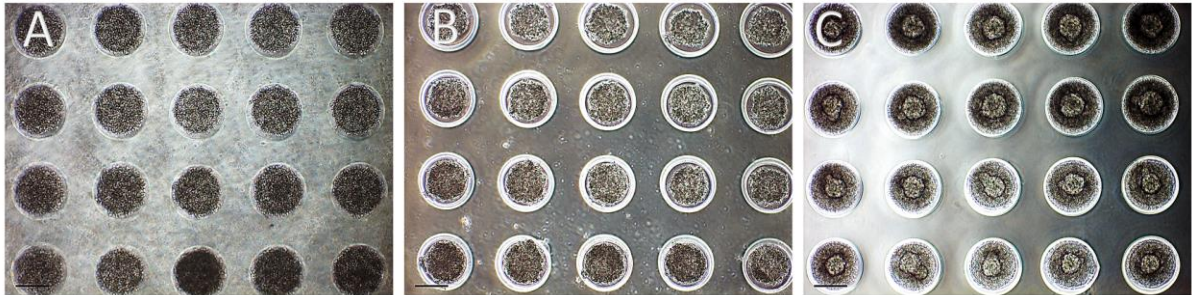
To develop the 3D platform, two different culture methods of cell suspension and agarose micromold was explored in parallel. For the 3D cell suspension methodology, hiPSCs were dissociated into a single cell suspension and transferred into an ultra-low adherent culture dish to facilitate self-aggregation. Although this method was successful in the generation of 3D hiPSC spheroids and subsequently differentiated into hepatocyte-like cells (HLC), some limitations were increasingly recognised including lack of control over the initial size of formed spheroids and the fusion of spheroids post differentiation. These limitations resulted in heterogeneity in culture and in the formation of very large and dense 3D microtissues with a necrotic core (Figure. 3.11).



**Figure 3.11: Static suspension culture. A) Single hiPSCs were dissociated into a single cell suspension and transferred into ultra-low adherent culture dish to facilitate self-aggregation (seeding). B) Formation of 3D hiPSC spheroids with different sizes resulted in a heterogenous culture. The spheroids became bigger as they were maintained in the culture. C) This picture is the magnified version of the field that has been highlighted with a black dash line border and shows the formation of different spheroid sizes. Scale bar: 100  $\mu$ m.**

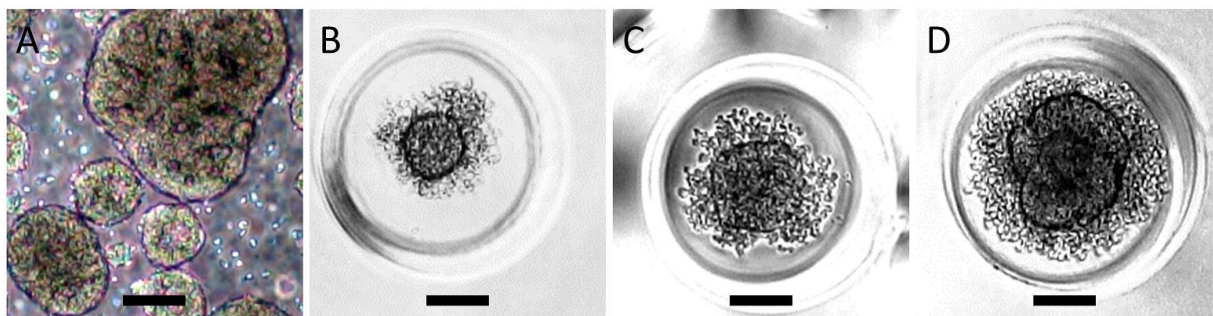
To overcome these limitations, an agarose micro mould system was used to minimise size variation of spheroids which remained separately in microwells at initial stages of differentiation. Agarose microplates were manufactured in a 256-well format using the 3D Petri Dish<sup>®</sup> Microtissue mould. The advantages of using this system was to maximise cell-cell interaction and consistent batches with controlled cell density and size ranges. By using this method, controlling spheroid sizes were improved considerably, resulting in formation of individual 3D spheroids with consistent sizes

throughout the plate (Figure 3.12). This technique was used to differentiate P106 iPSC line into mature HLCs in 3D cell culture setting and will be explained in later sections.



**Figure 3.12: Formation of spheres in agarose microplates using the 3D Petri Dish® mould. A) P106-iPSCs were seeded in agarose microplates. Cells were floating and gradually sedimented inside the microwells. B) Two hours following seeding, the cells are located inside the microwells and spheroids are forming. C) One day following seeding, tiny 3D spheroids were formed with a clear border inside the microwells following self-aggregation of live cells. Scale bar: 200  $\mu\text{m}$ .**

In this study, the aim was to maintain a long-term *in vitro* 3D culture and therefore maintaining the right size for consistent hepatic differentiation was crucial. Each spheroid size consisted of certain number of cells on seeding day. The initial seeding for each spheroid was around 800, 1500 and 2200 cells per spheroid for 50-100, 100-150 and 150-200  $\mu\text{m}$  in diameter sizes, respectively. Since 3D spheroids of lower than 100  $\mu\text{m}$  in diameter (Figure 3.13B) did not survive the first phase of differentiation, spheroids in the range of 100-150  $\mu\text{m}$  in diameters (Figure 3.13C) were selected for hepatic differentiation and production of mature hepatic 3D HLCs (Figure 3.13D).

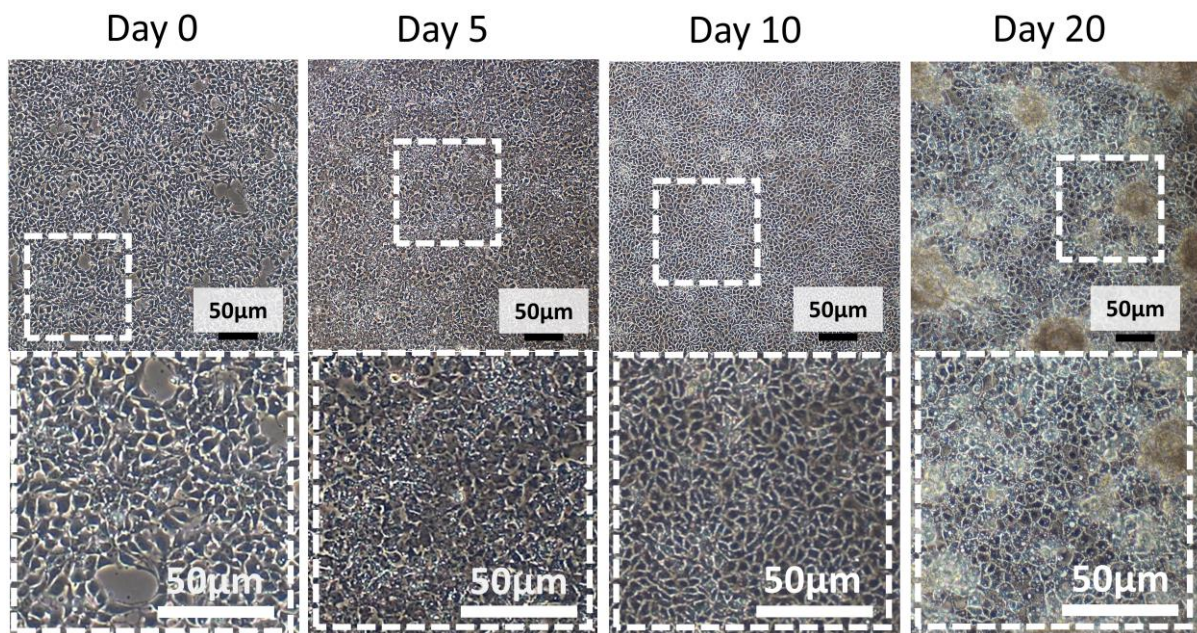


**Figure 3.13: Three different size ranges for 3D spheroids (50-100  $\mu\text{m}$ ), (100-150  $\mu\text{m}$ ) and (150-200  $\mu\text{m}$ ) were generated. A) Different spheroid sizes in static suspension culture. B) 50-100  $\mu\text{m}$  spheroid, these spheroids do not resist the first stage of differentiation. C) 100-150  $\mu\text{m}$  spheroid, demonstrated to be effective for production of**

mature hepatic 3D HLCs. D) Differentiation of 150-200  $\mu\text{m}$  spheroids resulted in more heterogenous population. Scale bar: 100  $\mu\text{m}$ .

### 3.3.9. Morphological analysis of 2D and 3D hepatic differentiation of P106-iPSC line.

The hepatic differentiation protocol was used to differentiate P106 iPSCs and P106-derived spheroids into mature 2D and 3D HLCS, respectively. There was no morphological difference between 33D6 and P106i cell lines. During the DE stage, cells were formed triangular endoderm cells (Figure 3.14). During the hepatoblast stage, cells displayed a cobblestone-like morphology with clear and distinct edges (Figure 3.14). At the hepatocyte maturation stage, cells exhibited a hexagonal shape and clear borders (Figure 3.14).

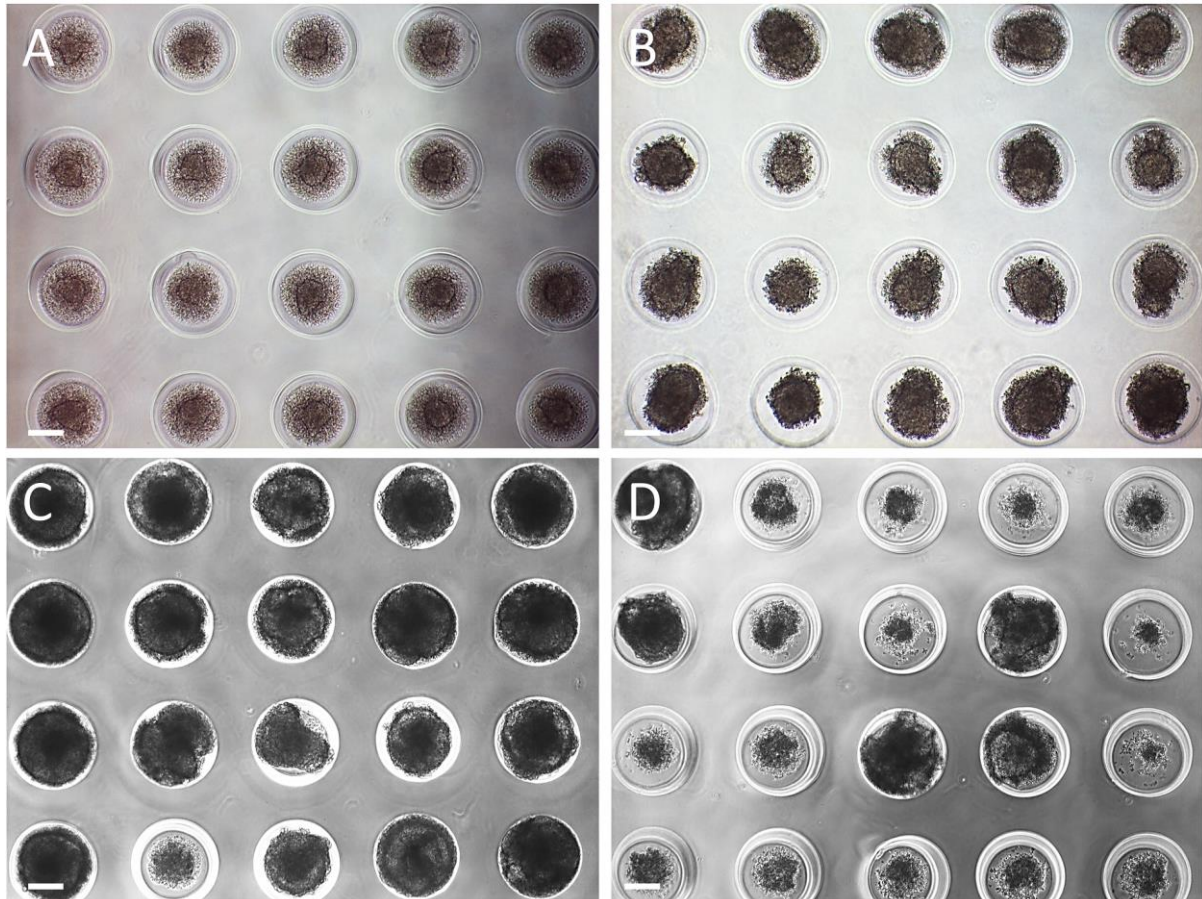


**Figure 3.14: Morphology of 2D hepatic differentiation of P106i cell line. Top panel: A) Seeding of the cells and priming to DE (Day 0). B) DE phase (Day 5). Hepatoblast stage (Day 10), and HLC (Day 20). Bottom panel: Is the magnified version of the fields highlighted in white dash border lines. Abbreviation: Definitive Endoderm (DE) Scale bar: 50  $\mu\text{m}$ .**

The stepwise differentiation protocol was used to differentiate P106 iPSCs and P106-derived spheroids into mature 3D haptospheres as described before. On day 0 of differentiation, the P106-derived spheroids were formed with a clear edge around each spheroid. As the differentiation progressed, the 3D spheroids became bigger and on



day 5 of differentiation formed the DE stage. Due to the 3D nature of the spheroids, morphological analysis was not very clear (Figure 3.15). From day 10 of differentiation onward, differentiated spheroids formed two different types of solid and hollow spheroids. The solid hepatospheres stayed inside the mold while the hollow spheroids floated out of microwells (Figures 3.15D & 3.16A, respectively).



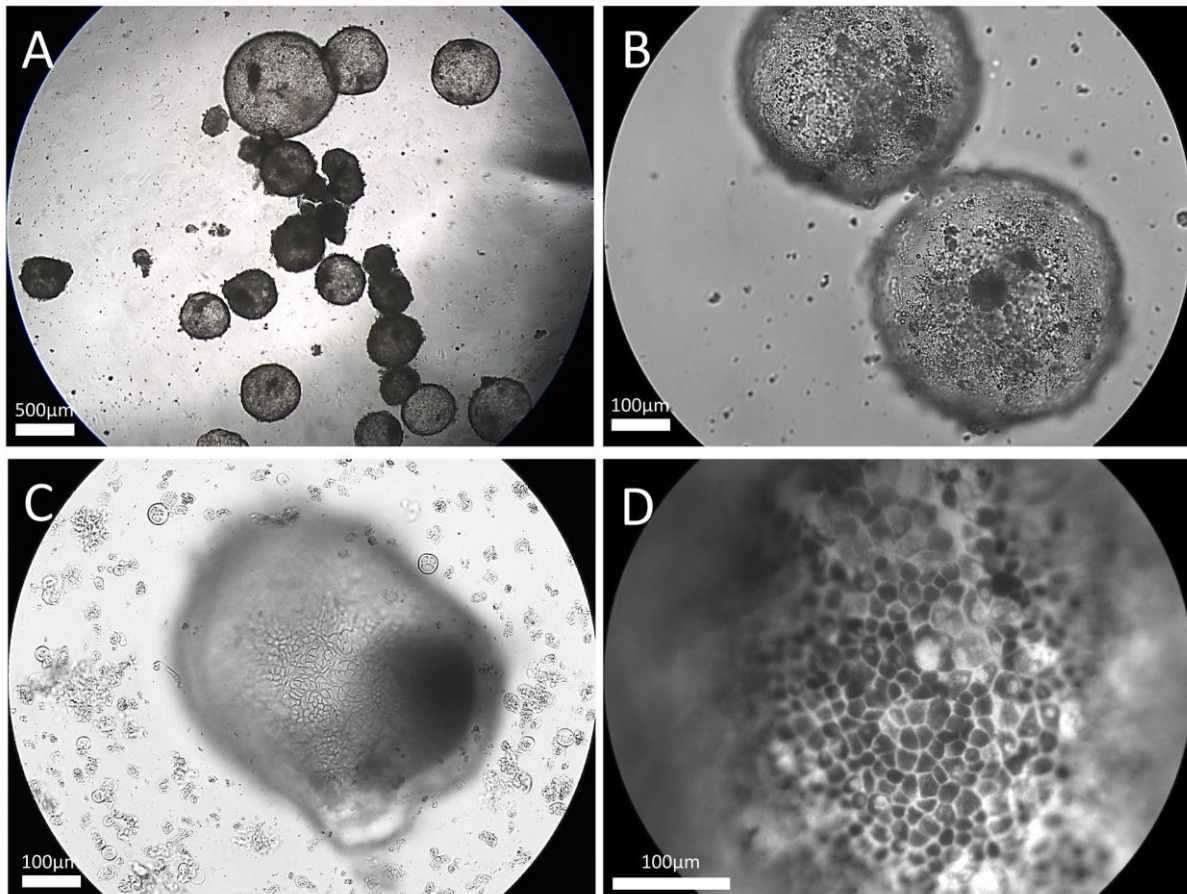
**Figure 3.15: Morphology of 3D hepatic differentiation of P106-iPSC hepatospheres. A) Day 0 of differentiation, DE stage B) Day 5 of differentiation, hepatoblasts stage. C, D) Day 15 of differentiation, spheroids become bigger and form hollow structures which ultimately leave the microwell space and float inside the wells. Scale bar: 100  $\mu$ m.**

From day 20 onward, the content of the molds was transferred into 12-well plates coated with poly-HEMA (Figure 3.16). 3D heps gradually exhibited polygonal morphology and accumulated dark deposit by elongation of the culture (Figure 3.16D).

### **3.3.10. Functional characterisation of 2D P106 and 3D heps.**

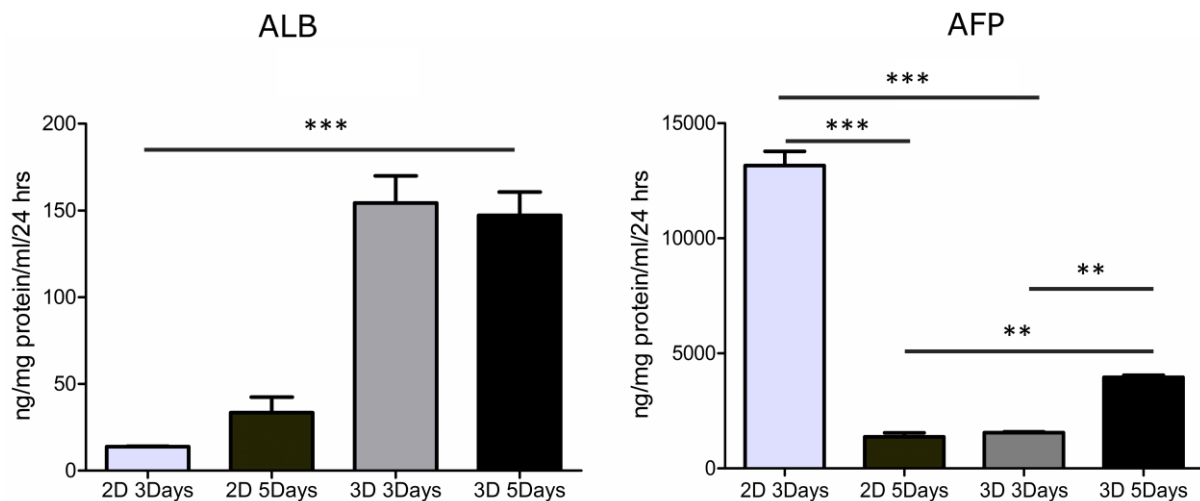
Serum protein production of mature hepatocytes is crucial for maintenance of healthy liver metabolism and homeostasis. The secretion of ALB and AFP from the collected

supernatant was measured on Days 17 and 20 of differentiation using ELISA technique. The 2D P106-derived HLCs were compared with P106 3D heps, employing 3 days protocol (Wnt 3a, Activin A for three days) and 5 days protocol (Wnt 3a, Activin A and two more additional days of only Activin A). The results are shown in (Figure 3.17). The level of ALB production in 2D (5 days) was higher compared with 2D (3 days), confirming the efficiency of the 5 days protocol for the first stage of differentiation. In comparison, the level of ALB production in 3D heps was high in the 3 days and the 5 days protocols, although the level of ALB was slightly higher in the 3 days protocol. Overall, the rate of ALB protein production was significantly greater in 3D heps compared to 2D hepatic monolayer, a 5 folds increase from 30 ng/mg protein/ml/24hours to 150 ng/mg protein/ml/24hrs. This further confirms the efficiency of the 3D protocol and the maturity of 3D heps.



**Figure 3.16: Morphology of 3D heps and maturation phase. A) Agarose microplates were removed and remained in suspension (scale bar 500 μm). B) Day 30 of differentiation. C) Day 45 of differentiation. D) Day 150 of differentiation, as the differentiation progressed, cells became more polarised with a distinct hexagonal shape while some deposited dark materials (B, C & D scale bars 100 μm).**

In addition, the level of AFP production was significantly higher in 2D hepatocytes (2 days) compared to 2D hepatocytes (5 days). This further indicates the efficiency of the 5 days hepatic differentiation protocol. However, the level of AFP production was significantly lower in 3D heps in both the 3 days and the 5 days protocol. A high level of ALB production and a low level of AFP secretion indicated superiority of 3D heps over HLC derived under 2D culture condition.



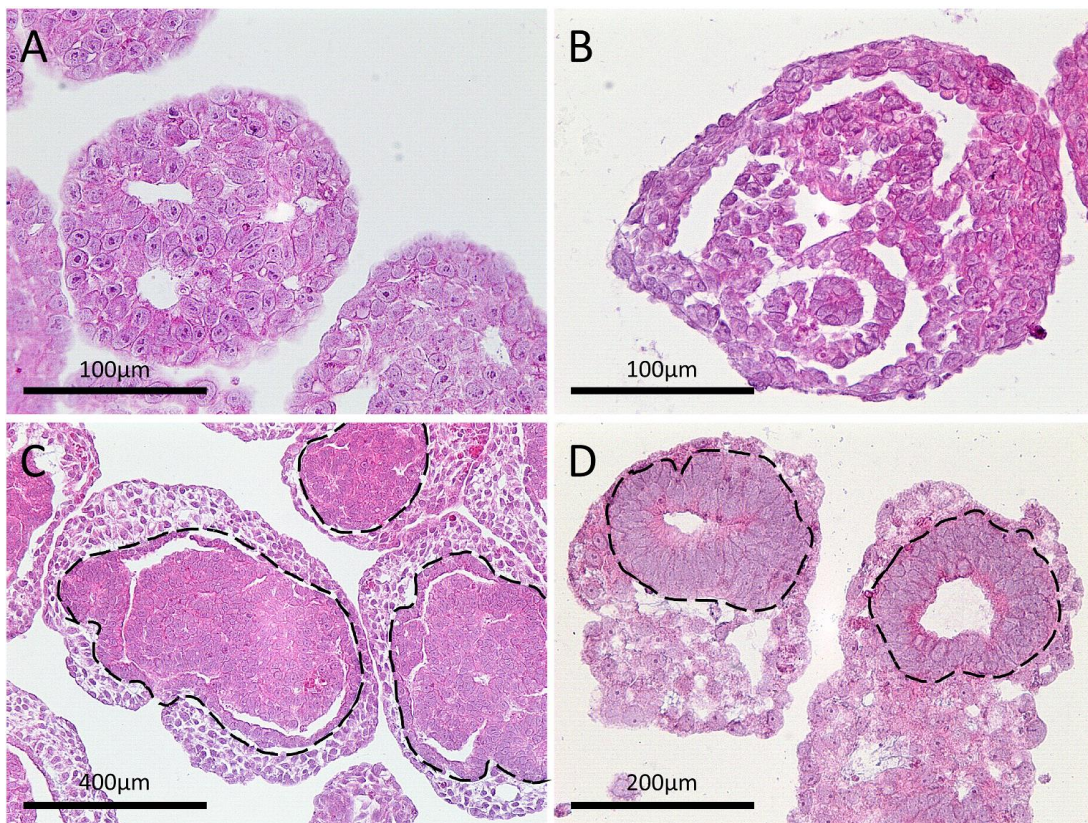
**Figure 3.17: ALB protein production of 2D HLCs and 3D heps.** The graph displays ALB protein production of 2D mature HLCs and 3D heps and comparing 2 days versus 5 days in vitro culture. The level of ALB production in 2D (5days) was higher compared with 2D (3 days), confirming the efficiency of 5 days protocol for the first stage of differentiation. the level of ALB production in 3D heps was high in 3 days and 5 days protocols, although the level of ALB was slightly higher in 3 days protocol. Overall, the rate of ALB protein production was significantly increased in 3D heps compared to 2D hepatic monolayer, a 5 folds increase from 30 ng/mg protein/ml/24hours to 150 ng/mg protein/ml/24hrs. The results represent the mean +/- SD of three individual biological samples per time point. The levels of significance were measured by one-way ANOVA where  $p < 0.05$  is denoted as \*,  $p < 0.01$  is denoted as \*\* and  $p < 0.001$  is denoted as \*\*\*.

Following successful hepatic differentiation, protein expression, gene expression and metabolic functionality were studied to further characterise the 3D Heps.

### 3.3.11. Characterisation of P106-derived 3D heps

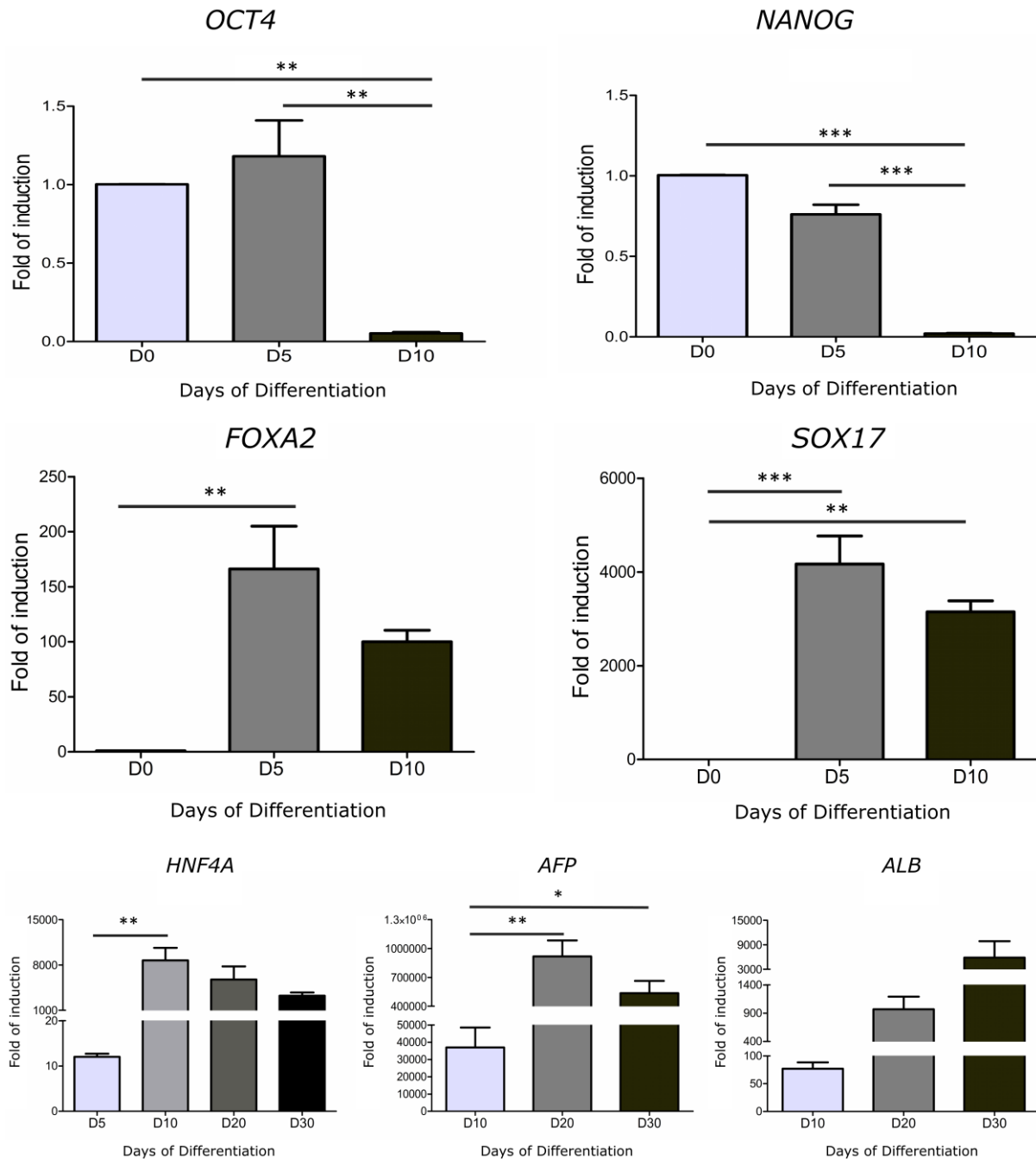
In order to further analyse the structure of P106 3D Heps, hematoxylin & eosin staining (H&E) was performed following sectioning of the 3D Heps at Days 0, 5, 10 and 20 of hepatic differentiation. Data from histology analysis showed that on Day 0 of

differentiation, cells formed a compact structure while exhibiting high nucleus to cytoplasm ratio (Figure 3.18A). On Day 5 of differentiation (DE induction), cells were rearranged to form a triangular endoderm shape structure and the ratio of nucleus/cytoplasm was reduced (Figure 3.18B). As the differentiation progressed into hepatoblasts and maturation stage, layers of mature hexagonal hepatocytes formed in outer layers while dense mesenchymal structures formed in the middle (Figures 3.18C , 3.18D), as previously observed in 3D Heps generated from the H9 embryonic stem cell line (Rashidi et al., 2018, Lucendo-Villarín et al., 2019).



**Figure 3.18: Representative histology of 3D P106-derived hepatospheres. This figure represents sectioned areas of 3D heps at different days of differentiation, using H&E staining method. A) Day 0, the cells formed compact structures with high nucleus to cytoplasmic ratio (scale bar 100 µm). B) Day 5 (DE stage), cells were rearranged to form triangular DE structures (scale bar 100 µm). C & D) Day 10 & 20 of differentiation (hepatoblast and hepatic specification stage), mature hepatocyte cells were formed in outer segment layers while containing a very dense structure of mesenchymal cells in the middle section (scale bars 400 µm and 200 µm, respectively). In this figure, the mesenchymal segment is outlined with a black dash line border in the centre of each 3D heps. Abbreviation: Haematoxylin & eosin staining; (H&E).**

During the first few days of hepatic differentiation and in line with significant changes in cell morphology, gene expression analysis using qPCR indicated a decrease in the expression of the pluripotent stem cell markers such as *OCT4* and *NANOG* at day 10 of differentiation compared to undifferentiated hiPSCs (Figure 3.19).



**Figure 3.19: Gene expression analysis of 3D P106-derived spheroids at different stages of differentiation. Gene expression analysis of 3D P106-derived spheroids were measured using quantitative PCR. Results exhibited a significant decrease in the expression of pluripotency markers; *OCT4* and *NANOG* from Day 0 to Day 10 of differentiation. The first stage of differentiation is DE. Results exhibited a significant**

increase in the expression of definitive endoderm markers; *FOXA2* and *SOX17* from Day 5 to Day 10 of differentiation, confirming the successful transition into definitive endoderm. Maturation phase involves two stages namely hepatoblasts and hepatic specification. The gene expression of *ALB*, *AFP* and *HNF4A* was detected on Days 10, 20 and 30 of differentiation. The level of *HNF4A* was significantly increased on Day 20 whilst there was a reduction in expression on Days 20 and 30. The level of *AFP* expression was significantly increased on Day 20 of differentiation. However, as the hepatospheres became mature the expression was reduced. The level of *ALB* gene expression showed a general rise from Day 10 to Day 30 of differentiation. However, this was not significant. The results were normalised to the house-keeping gene *GAPDH*. The results represent the mean  $\pm$  SD of three individual biological samples per time point. The levels of significance were measured by one-way ANOVA where  $p < 0.05$  is denoted as \*,  $p < 0.01$  is denoted as \*\* and  $p < 0.001$  is denoted as \*\*\*.

The result of gene expression analysis is further confirmed with immunofluorescence analysis. The P106-derived 3D Heps were fixed in ice-cold methanol, wax embedded, sectioned and stained with OCT4 and NANOG as pluripotency markers on Day 0 of hepatic differentiation. This data suggests the high expression levels of pluripotent markers confirm the pluripotent state and identity of the cells prior to hepatic differentiation (Figure 3.20, Day 0).

The first stage of hepatic differentiation was priming P106-derived 3D heps into DE using RPMI/B27 media supplemented with recombinant mouse Wnt3a and Activin A growth factors. As the differentiation progressed, a significant increase in the expression of endodermal markers of DE, namely Sry-related HMG box factor (*SOX17*) and Forkhead box A2 (*FOXA2*), was observed at day 5 of differentiation while the expression was slightly down-regulated at day 10 of differentiation. The *FOXA2* is a hepatocyte nuclear factor 3-beta (*HNF3B*) which has a role in activating liver-specific genes. The *SOX17* was mainly expressed during endoderm differentiation.

Further, immunofluorescence analysis was performed on the DE and hepatoblast stages and confirmed with qPCR gene expression analysis. Immunostaining data revealed that the majority of the cells were stained positive for *SOX17* and *FOXA2* markers, confirming endoderm differentiation. During DE and early endoderm differentiation, these markers were highly expressed in the nucleus (Figure 3.20, Days 5 & 10).

For induction of hepatic specification, a media containing Knock-Out Serum Replacement (KOSR) and DMSO (SR/DMSO) was used, resulting in formation of liver progenitor cells called hepatoblasts. Finally, for hepatocyte differentiation and maturation a defined serum-free media containing OSM and hepatocyte growth factor (HGF) was used.

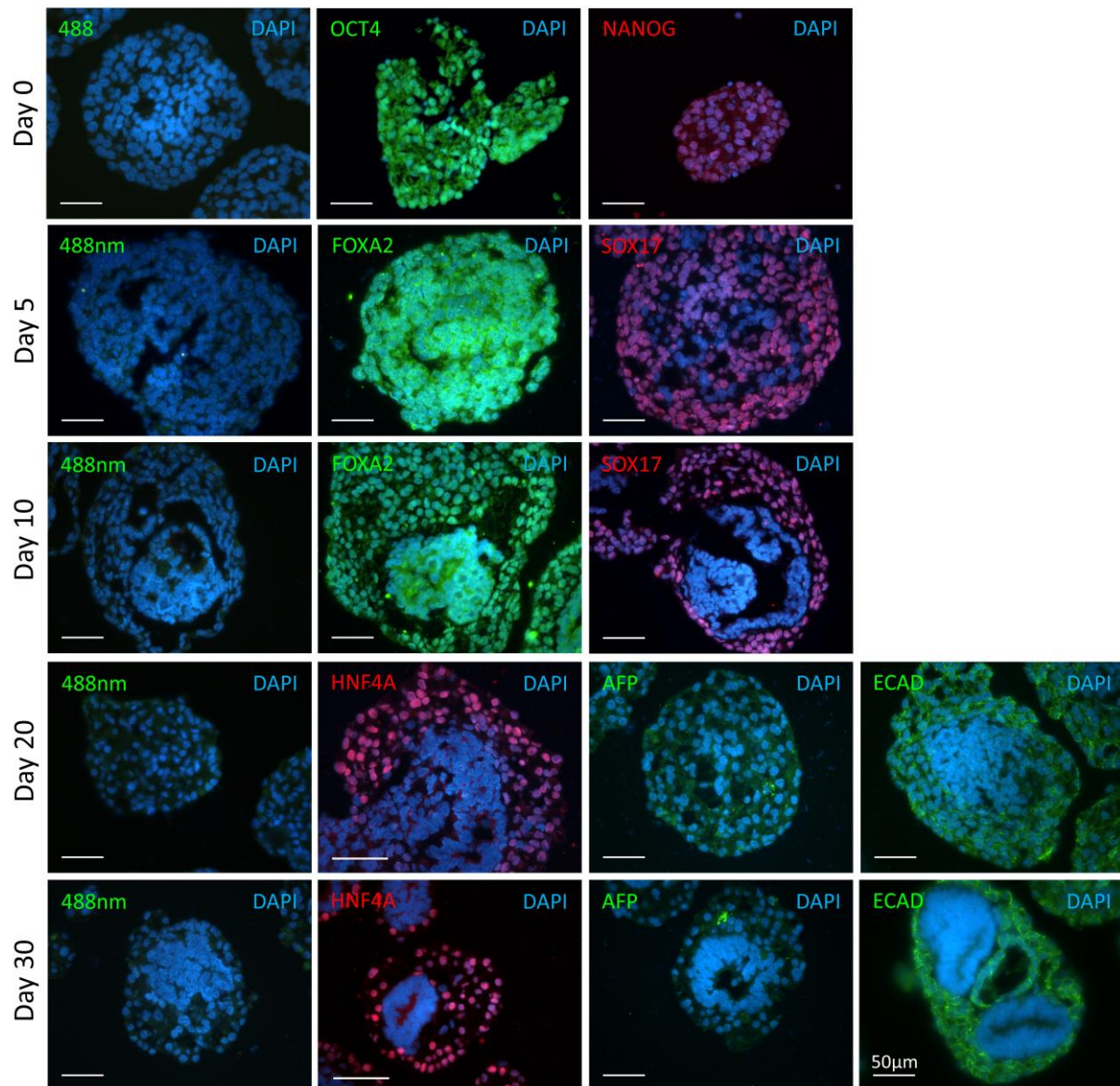
The gene expression of the foetal hepatic marker, *AFP* was detected in P106-derived 3D Heps at Day 10, 20 and 30 of differentiation. At Day 20 of differentiation, there was a significant increase in the level of *AFP* expression. However, as the differentiation and maturation of the cells continued the level of expression was significantly reduced.

*ALB* gene expression was detected at Day 10, 20 and 30 of differentiation in P106-derived 3D Heps. Although the increasing level of *ALB* gene expression was not significant, it was showing a trend of upregulation which continued throughout differentiation.

HNF4-alpha (*HNF4A*) is a master liver transcription factor which is expressed by hepatoblasts and mature hepatocytes. The expression of *HNF4A* in 3D Heps were detected at Days 5, 10, 20 and 30 of differentiation. On Day 10, there was a significant increase in the level of expression. As the hepatospheres became more mature, the level of *HNF4A* expression was decreased.

To characterise mature 3D HLC hepatospheres, the protein expression of different hepatic markers such as Epithelial cell adhesion molecule (E-cadherin; ECAD), HNF4A and AFP were investigated. On Day 30 of differentiation, 3D heps were fixed, sectioned and stained with appropriate antibodies. Expression of HNF4A was detected in cells forming the outer ring. Similarly, cells forming the outer shell of the 3D Heps were stained positive for ECAD on Days 20 and 30 of differentiation, while the level of expression seemed to be more pronounced on Day 30 of differentiation compared to Day 20, suggesting further maturation by elongation of differentiation. The ECAD is a cell-cell adhesion molecule which has a role in epithelial cell behaviour and tissue formation. It is mainly expressed by mature hepatocytes and biliary epithelial cells. In addition, immunostaining was performed to detect AFP expression. AFP is considered a foetal marker during hepatic development. Expression of AFP was detected only in few cells on Days 20 and 30 of differentiation suggesting that 3D

Heps exhibited a more mature phenotype compared to 2D counterparts (Figure 3.20, Days 20 & 30).



**Figure 3.20: Immunofluorescence analysis of 3D P106-derived spheroids. On Day 0 of differentiation the spheroids were fixed, sectioned and stained with OCT4 and NANOG antibodies. Immunostaining results showed high expression levels of OCT4 (Green) and NANONG (PINK), confirming the pluripotent state of the cells prior to differentiation. On Days 5 and 10 of differentiation the spheroids were stained with SOX17 and FOXA2 antibodies showing high levels of SOX17 (pink) and FOXA2 (green) expression, confirming the endoderm stage. These endodermal markers were highly expressed in nucleus. The E-cadherin, HNF4A and AFP antibodies were used for day 20 and 30 of differentiation. E-cadherin (green) expression in 3D heps were detected at cellular boundaries on Days 20 and 30 of differentiation with higher level of expression**



on Day 30 compared to Day 20 of differentiation. HNF4A (pink) expression in 3D heps were detected on Days 20 and 30 of differentiation. It is a master regulator which is highly expressed in the nucleus. Immunostaining was also performed using an AFP antibody. Only a few cells expressed AFP (green) marker in cytoplasm of 3D heps. No non-specific signal was detected in secondary antibody control (488 nm). The nuclei were counterstained with DAPI (blue). Scale bar: 50  $\mu$ m. Abbreviation: Epithelial adhesion molecule marker; E-cadherin. Hepatocyte nuclear factor alpha; HNF4A. Alpha fetoprotein; AFP.

### 3.4. Discussion

The current “gold standard” model to study *in vitro* liver hepatocyte biology and to predict *in vivo* functionality is using freshly isolated human primary hepatocytes (Hewitt et al., 2007). However, scarcity, short life span, low proliferation rate and batch to batch variation among human tissues have restricted their use in research and clinical applications (Cheng et al., 2008, Fox and Strom, 2008). Other sources such as hepatocarcinoma and immortalised human hepatocyte cell lines such as HepG2 (Wilkening et al., 2003) and HepaRG (Gripon et al., 2002) are also available. The problems such as low metabolic activity and long-term stable functionality has limited their application as an alternative source of hepatocytes (Vessey and de la Hall, 2001, Alison et al., 2007). Therefore, a credible source such as PSCs can produce reliable and functional *in vitro* hepatocytes. PSCs such as hESCs and hiPSCs have unique characteristics which the previous predecessor cell lines do not hold. Their unlimited self-renewal ability and the pluripotency capacity enabled them to differentiate into any desired tissue type (Cai et al., 2006, Vazin and Freed, 2010). Further, these cell lines can provide inexhaustible source of mature and functional hepatocytes. To this end, various research groups have focused on developing reliable and fully-defined protocols to obtain mature and functional mature hepatocyte-like cells (HLCs) from PSCs (Hay et al., 2007, Hay et al., 2008a, Hay et al., 2008b, Agarwal et al., 2008, Sullivan et al., 2010, Hannan et al., 2013, Loh et al., 2014).

From the developmental biology perspective, liver development and maturation occur in distinct stages. While the existing protocols for derivation of hepatocytes from hPSCs mimic aspects of *in vivo* development, generated HLC generally exhibited more foetal-like phenotype and have a transient phenotype. In this study, an efficient stepwise hepatic differentiation protocol was employed which recapitulates *in vivo*

developmental stages of definitive endoderm induction (DE), hepatoblasts and hepatocyte specification and maturation. In addition, generated 3D Heps exhibited stable phenotype with prolonged hepatic functionality of over a year in culture (Rashidi et al., 2018).

During embryonic gastrulation stage, induction of DE highly depends on the network of coordinated signals Wnt/ $\beta$ -catening derived from emerging cardiac mesoderm. BMP/Nodal signalling is released from septum transversum mesenchyme which maintains the complete hepatic induction by activation of certain transcription factors. These transcription factors are including Sox17, Forkhead box (Fox) A1 and A2 (HNF3A and B) (Clements et al., 2003). Wnt/  $\beta$ -catening signalling pathway has a key role in differentiation and proliferation of DE cells (Burke et al., 2006, Fletcher et al., 2008). Further, BMP signalling pathway plays an important role in coordination with the above-mentioned signalling pathways in maintenance of hepatic induction and liver development. In this study, for initiation of hepatic induction and BMP signalling, the cells were treated with Activin A (Hay et al., 2007, Hay et al., 2008a) and Wnt3a (Hay et al., 2008a). Using these growth factors activate the transcription factors such as Sox17 and FoxA2, confirming the generation of efficient DE cells (Costa et al., 2003).

For hepatic specification and formation of hepatoblasts, expression of certain factors such as FoxA and GATA and chromatin changes are required (Cirillo et al., 2002). These factors promote expression of hepatic genes such as ALB, AFP and HNF4A from the liver bud. The process of *in vitro* hepatic specification can be mimicked using a cocktail of growth factors such as BMPs (Cai et al., 2007, Duan et al., 2010, Touboul et al., 2010) or supplementing the media with DMSO (Hay et al., 2008a, Duan et al., 2010) or using with other factors (Duan et al., 2010). In this study, the medium was supplemented with DMSO to induce hepatic induction. Similar to previously published papers (Cameron et al., 2015, Wang et al., 2017, Rashidi et al., 2018), upregulation of hepatoblast markers such as HNF4A, AFP and ALB were detected at both gene and protein levels on Day 10 of differentiation.

The hepatic maturation stage involves a close synergy between gradient cell signalling and combination of growth factors; OSM and HGF. OSM promotes maturation and polarisation of hepatocytes, whilst HGF has a role in cell proliferation and organisation of hepatocytes into cord-like structures (Kamiya and Gonzalez, 2004). Current hepatic

differentiation protocols employ growth factors combination with glucocorticoids to promote specification and maturation of derived hepatoblasts from PSCs (Cai et al., 2007, Hay et al., 2008a, Basma et al., 2009, Duan et al., 2010, Rashidi et al., 2018, Lucendo-Villarin et al., 2019). HNF4A is a liver master regulator which has a key role in hepatocytes functionality by regulating the essential transcription factors involved in the liver development (Li et al., 2000). Mature and functional hepatocytes express different genes involved in cytochrome P450 activities and secretion of serum ALB and AFP. Consistent with previous published studies (Hay et al., 2008a, Wang et al., 2017, Rashidi et al., 2018, Lucendo-Villarin et al., 2019), for hepatic maturation stage cells were treated with HGF and OSM in addition to glucocorticoids. Results showed that P106 derived 3D heps formed from two distinct populations of cells with mesenchymal identity at the core and a homogenous population of polar and hexagonal shaped cells resembling *in vivo* mature hepatocytes. The gene and protein expression on Day 20 and 30 of differentiation revealed expression of high levels of mature hepatic markers. Further, functional characterisation of 3D heps showed high level of serum ALB protein secretion whilst detecting a very low level of AFP, confirming the functionality and maturity of 3D heps.

For *in vitro* culture of primary and PSCs-derived hepatocytes, physiological reagents such as insulin or glucocorticoids are added to the culture media (Godoy et al., 2010, Szkolnicka et al., 2014). It has been reported that complementing media with glucocorticoids including dexamethasone or hydrocortisone promotes cell attachment, maturation of hepatocytes and secretion of liver specific proteins (Kim et al., 2001). Also, HepatoZYME serum free medium was considered to be the most suitable media for hepatocyte maturation (Szkolnicka et al., 2014). Using this media promoted mature hepatocyte morphological features, hepatic gene and protein expression patterns and a more stable phenotype throughout the differentiation. One other non-physiological component that was added to the differentiation media was DMSO reagent. It acts as an anti-dedifferentiation reagent by preserving hepatocellular phenotype, improving polarization of the cells and an increase in the expression of hepatic transcription factors such HNF3A and B and HNF4A (Su and Waxman, 2004, Vinken et al., 2006). Based on this information, in this study DMSO was used in the second stage of differentiation and HepatoZYME serum free medium with essential growth factors was employed for maturation of hepatocytes. However, it was shown that long-term culture

with HepatoZYME culture medium resulted in disintegration of 3D heps and loss of hepatic phenotype and an alternative recipe was developed to support long-term culture of 3D Heps *in vitro* (Rashidi et al, 2018). The same medium composed of Williams E supplemented with FGF2, EGF, VEGF and HGF was employed in this study from day 30 of differentiation onward.

In addition to the reagents used for hepatocyte differentiation, the microenvironment surrounding the cells is also equally important. Traditionally for 2D hepatic differentiation, biological matrices such as Matrigel™ or collagen I have been extensively used to preserve primary hepatocytes *in vitro* culture. Using matrices restores the cell-to-extracellular matrix (ECM) and cell-to-cell interactions. For instance, Matrigel™ promotes the expression of hepatic transcription factors (Elaut et al., 2006) whilst collagen-I maintains the liver mature phenotype by improving hepatocyte attachment and survival (Skett and Bayliss, 1996). Very recently, a recombinant isoform of laminin (LN521) has been developed which has been used extensively for PSCs culture. This natural laminin provides efficient self-renewal and differentiation of PSCs in a feeder-free, animal-free and chemically defined culture system. Using this natural laminin promoted functional mature hepatocytes and a much more stable phenotype compared to their predecessors (Cameron et al., 2015, Wang et al., 2017).

Over the years, efforts have been made to develop 3D culture systems for hepatocyte differentiation using biological matrices (Tuschl and Mueller, 2006, Page et al., 2007, Rowe et al., 2010). However, using 3D Matrigel structures had a negative impact on the cellular stress response signalling by decreasing the mRNA expression level of interleukins and other protein kinases (Page et al., 2007). Therefore, using these matrices for 3D systems compromises the functionality and maturity of hepatocytes. Further, liver sinusoidal endothelial cells (LSEC) have been extensively employed for primary hepatocytes and PSC derived HLCs. These cells in combination with other factors secrete ECM components which has a key role in the regulation of liver specific genes (McCuskey, 2008, Kim and Rajagopalan, 2010).

Hepatic organoids are 3D mini-organ structures in a 3D matrix which recapitulates the morphology and functional activities of real organs. The presence of multiple signalling pathways is required to promote the proliferation and differentiation of the cells and ultimately organogenesis. Newly established methods suggest that PSCs or liver

specific progenitor cells can generate liver-specific organoids. Several studies have focused on developing robust and long-term cultures of 3D organoids which mimics *in vivo* physiology of the liver (Rashidi et al., 2018, Akbari et al., 2019a). Takebe and colleagues were the first to propose the generation of hepatic organoids from iPSCs, using a sophisticated co-culture method. In this study, human MSCs and human umbilical vein endothelial cells (HUVECs) were combined with iPSC-derived HLCs in a Matrigel-embedded culture to generate *in vitro* vascularised and mature liver buds (LBs). In terms of functionality and following transplantation, ALB was secreted into the bloodstream from days 10 to 45 post-transplantation (Takebe et al., 2013). In a follow-up study, single-cell RNA sequencing displayed unique patterns of gene expression and identification of different cell populations from the pluripotency stage to liver bud formation (Camp et al., 2017, Potter, 2018). Although the development of such systems was crucial from the developmental biology point of view, having batch to batch variation restricted widespread application of this technology. Therefore, many research groups were focused on developing robust protocols to produce liver organoids derived from PSCs.

Using mouse-derived fibroblasts in combination with stem cell-derived hepatocytes have enhanced cell-to-cell contacts in hepatocyte cultures. It is argued that fibroblasts secrete essential ECM components and growth factors such as HGF which affects the survival of mature hepatocytes (Sugimachi et al., 2004, Leite et al., 2011). One study suggested that a 3D culture system comprising of MEFs and PSC derived HLCs embedded in two different ECM such as Matrigel and collagen have improved hepatocyte phenotype and metabolic functionality (Berger et al., 2015). Guan and colleagues generated iPSC-derived hepatic organoids, embedded in cholangiocyte ductal structures for 50-60 days. At later stages of differentiation, due to the nutrient deficiency and lack of oxygen supply to the organoids, their proliferation and regenerative capacity decreased significantly (Guan et al., 2017). Another protocol was established to generate organoids from PSCs with a hepato-biliary structure. The PSC-derived organoids displayed hepatic gene expression and *in vitro* functional characteristics of cholangiocytes. Following transplantation, the organoids remained stable for more than 8 weeks in immune-deficient mice (Wu et al., 2019). A recent protocol further suggested that enrichment of Epithelial Cell Adhesion Molecule (EpCAM)-positive cells promoted homogenous population of endodermal cells and

differentiation of functional hepatocytes. Further, the addition of R-spondin and combination of growth factors has promoted the specification of EpCAM-positive endoderm cells in this unique culture system. These endoderm-derived hepatic organoids (eHEPSs) were mature in 14 days and successfully expanded for more than one year in culture without significant loss of phenotype and culturing efficiency (Akbari et al., 2019b).

In this study, initially iPSC line from WiCell was characterised using a range of assays including immunostaining and flow cytometry to identify pluripotent markers. As the ability of this line to differentiate into all three germ layers of endoderm, ectoderm and mesoderm was confirmed earlier by other colleagues in the lab, here spontaneous differentiation into all three germ layers was not performed; however, this is part of the standard procedure to fully characterise pluripotent stem cells. Then, the iPSCs-derived 3D human liver tissue was generated using an optimised and step wise differentiation protocol, as previously described (Rashidi et al., 2018, Lucendo-Villarin et al., 2019). This protocol addresses issues surrounding previous 3D protocols such as scalability and long-term *in vitro* phenotypic stability. The use of a 3D culture system with a defined hepatocyte maturation media and eliminating animal-derived components such as Matrigel and fibroblasts have improved the life span of the cells up to 400 days in culture while promoting overall long-term mature HLC functionality (Rashidi et al., 2018, Lucendo-Villarin et al., 2019). In addition, this protocol surpasses the regular need for mechanical dissociation which is a common feature for previously published 3D hepatic organoid protocols (Huch et al., 2013, Huch et al., 2015, Hu et al., 2018).

Functional analysis of the 2D HLCs and 3D heps were performed in Days 20 and 30 of differentiation, respectively. It is crucial to quantify and compare cytochrome p450 3A (CYP3A) and cytochrome p450 1A2 (CYP1A2) in 2D and 3D culture settings. These are important enzymes of functional HLCs and their activity was assessed using established assays. It is reported that high levels of ALB and CYP3A4 activity are associated with hepatocytes from the adult liver while secretion of AFP and CYP3A7 activity are detected in the liver during foetal life (Stevens et al., 2003). In a study published by Balta Lucendo-Villarin and colleagues, H9 and P106i-derived 3D heps were compared for CYP activities. H9-derived hepatospheres exhibited CYP3A activity of 220,375 RLU/ml/mg protein and CYP1A2 activity of 732,440 RLU/ml/mg

protein. However, in P106-derived 3D heps levels of CYP3A and CYP1A2 activities were 132,117 and 409,907 RLU/ml/mg protein, respectively. These results were compared with two batches of primary human hepatocytes and 3D heps which exhibited modest levels of CYP activity compared to primary hepatocytes (Lucendo-Villarin et al., 2019). In a similar study by Freyer and colleagues, CYP activities of undifferentiated iPSCs, iPSC-derived 2D and 3D heps were measured and compared to primary hepatocytes. Results demonstrated a high-level of CYP1A2 activity in 3D bioreactors compared to 2D setting. However, this result was much lower in undifferentiated iPSC and primary hepatocytes. In contrast, CYP3A activity was higher in undifferentiated iPSC and 2D-derived HLCs compared to 3D bioreactors. Overall, all enzymes were significantly higher in primary hepatocytes compared to 2D and 3D HLCs derived from iPSCs (Freyer et al., 2016). The existing dynamics of 3D culture systems promote an efficient microenvironment, maturation, and better formation of *in vitro* tissue structures. Furthermore, cell populations within 3D heps have better cell-cell interactions. It is speculated that these characteristics generate a natural gradient of nutrients and growth factors. This ultimately leads to better cell proliferation, differentiation, and maturation. One clear drawback is heterogeneity of cell populations within the 3D heps. Further works are required to standardise differentiation protocols and reduce heterogenous cell population within 3D culture systems.

In conclusion, the use of a 3D system with no ECM, an optimised stepwise differentiation protocol and a fully defined hepatocyte differentiation media supports the generation of consistent and stable mature HLCs for long-term *in vitro* culture. This resulted in mature and functional 3D heps as evidenced by ALB and AFP secretion and P450 metabolic activities. One of the limitations of the characterisation is lack of primary foetal and adult hepatocytes as a comparison. Without these controls the significance of ALB and AFP expression cannot be appreciated. To evaluate *in vitro* genotoxicity testing of viral vectors, hiPSCs and mature hiPSC-derived 3D heps were used. This model was applied as an individualised genotoxicity testing (In GeTox) 3D platform to assess genotoxicity of viral vectors *in vitro*. In the next chapter, lentiviral and adeno-associated viral vectors were used to transduce hiPSCs and 3D heps.

## Chapter 4

### 4.1. Introduction

Viral vectors have been used to deliver transgenes to treat inherited monogenic disorders. Numerous studies have been conducted to provide *ex vivo* (Montini and Cesana, 2012) or *in vivo* (Follenzi et al., 2002) gene correction (addition/correction) which demonstrates a promising approach to achieve therapeutic efficacy.

Despite the growing demand for the development of gene therapy regimen for clinical applications, there are concerns regarding the biosafety of viral vectors. Oncogenesis resulting from IM is a well-known gene therapy safety concern. IM has been shown in a number of trials such as *ex vivo* transduction of haematopoietic stem cells with gamma-retroviral vectors (Howe et al., 2008, Hacein-Bey-Abina et al., 2008, Braun et al., 2014).

Existing genotoxicity models play a key role in assessing the biosafety of viral vectors. To evaluate the biosafety profile of viral vectors, direct and systemic administration of the transgene vectors are predominantly based on animal models such as, *in vivo* studies in wildtype rodents or sometimes in non-rodent species (Cantore et al., 2015). In some studies, immunocompromised animals (Montini et al., 2006), disease models (Modlich and Baum, 2009) or tumour prone animals (Montini and Cesana, 2012, Nowrouzi et al., 2013) are used for *ex vivo* genetic modification of autologous or heterologous cells. Notably, efficacy of the gene therapy product is also important. In order to see an optimum therapeutic effect, the level of gene-corrected cells following treatment must be high. However, information regarding the clonal dominance between the gene-corrected cells in non-clinical models is not readily accessible. For this reason, a detailed evaluation of viral insertions in the target cells and the clonal dominance of the corrected cells is required for safety purposes and for an improved clinical prediction outcome.

In recent years, the use of recombinant AAV (rAAV) vectors for *in vivo* gene transfer have steadily increased. The safety profile and the high efficiency of transduction of a broad range of tissues have made them an excellent candidate for *in vivo* gene therapy (Colella et al., 2018). One of the main advantages of using AAV vectors in gene therapy is the low frequency of vector genome integration in the host DNA and



subsequent reduction in genotoxicity (Balakrishnan and Jayandharan, 2014). Despite this assertion, the issue of AAV-related genotoxicity remains crucial particularly in the clinical gene therapy setting (Colella et al., 2018).

Similarly, lentiviral vectors are the gene transfer tool of choice for gene or cell therapies. The safety and efficacy of these vectors have been reported in several clinical studies. Notably, the third-generation vector system is the most well characterised among all the generations and maintains long term stable expression in target cells. Reports published state that lentiviral-mediated gene delivery does not provoke IM in oncogenic hotspots thus decreasing the risk of cellular transformation compared to other viral vectors. However, random insertions still pose a risk as oncogenes and tumour suppressor genes are present through the whole genome (White et al., 2017). So far, lentiviral vectors have been extensively used for the genetic manipulation of PSCs. Also, they are used to generate hiPSCs from skin fibroblasts using pluripotency markers (Takahashi and Yamanaka, 2006, Takahashi et al., 2007). In a study by Nowrouzi et al, they developed *in utero* model that demonstrates the formation of HCC in mice receiving lentiviral vectors. Analysis of lentiviral integrations in the liver genomes of treated mice exhibited clear differences between vector insertions in gene dense regions and the genes that are highly expressed. This further suggests that vector has a preference for insertion or possible clonal outgrowth. Using the insertion site analysis, this genotoxicity model revealed not only the genotoxic potential of lentiviral vectors but also the genes associated with liver oncogenesis (Nowrouzi et al., 2013).

Despite vital information provided from animal genotoxicity studies, these models are time consuming and expensive. Results from *in vivo* and *in vitro* assays may not always extrapolate to that in the liver of the large animals or human liver. Efforts have been made to develop a “humanised” liver model (Lisowski et al., 2014). In this study, a chimeric human-mouse liver model was developed, using transplanted human hepatocytes in immune deficient *Fah*<sup>-/-</sup> mice. Mature human hepatocytes were repopulated in the mouse liver with variable efficiency ranges (5%-40%) (Lisowski et al., 2014). Despite successful screening of novel AAV capsid variants, this model still utilised the mouse model which may not resemble the true population of human liver hepatocytes.

Having all the aforementioned information in mind, there is a growing demand to develop *in vitro* non-clinical tests to assess the biosafety of the viral vectors for the patients. A robust humanised *in vitro/in silico* tool is required to evaluate the risk of vector mediated genotoxicity (IM/oncogenesis) which supports the risk assessment set by regulatory bodies whilst limiting the use of animals.

Human hepatocytes are considered as “gold standard” tool to assess *in vitro* hepatic metabolism and toxicity of pharmacological compounds and other xenobiotics (LeCluyse et al., 2005). However, there are some limitations while using liver primary hepatocytes such as low throughput, loss of viability and a reduction in liver-specific functionality and gene expression (Soldatow et al., 2013). PSCs such as hiPSCs is a great alternative and are able to differentiate into any desired tissue type such as mature hepatocytes.

As mentioned in Chapter 3, many *in vitro* cell culture-based assays use conventional 2D systems. However, there are limitations such as morphology changes, loss of phenotype and polarity. In addition, they do not mimic the natural *in vivo* environment. 3D systems can faithfully recapitulate the *in vivo* environment (24). Notably, it can support long-term culture for genotoxicity studies.

Over the years, four generations of LVs were developed which are different from each other based on their genetic constructs to drive the expression of the viral components and the number of wild-type genes. This ensures an increase in vector titres and also enhances vector safety. Third generation LVs present a higher level of biosafety and are commonly used. In this study, third generation Vesicular stomatitis virus G (VSVG) Pseudotyped HIV-1 based LVs were produced. These vectors were [pHR'SIN-cPPT-GFP-WPRE(SEW)] (safe) and [pHV-cPPT-SEW] (unsafe) which is the full long terminal repeat (LTR) equivalent to the pHR vector. The self-inactivating (SIN) pHR vector has deletions/mutations in 3'LTR of the viral genome. This will disrupt the promoter/enhancer activity of the LTR further improving the biosafety (27). This alteration did not affect the vector titer or transgene expression. In the SEW the internal promoter spleen focus forming virus (SFFV) was incorporated to drive the green fluorescent protein (GFP) expression. The woodchuck post-transcriptional regulatory element (WPRE) and central polypurine tract element (cPPT) have been used to improve the overall vector performance.

Recombinant Adeno associated (rAAV)/GFP serotype 2 vectors were kindly provided by our collaborators in Australia. Two main rAAV constructs were generated for this study:

1. The first vector construct carries “Clean ITRs” and a reporter gene that is driven by a strong ubiquitous CB7 promoter. This promoter was previously reported in Chandler et al (Chandler et al., 2015) [rAAV/CB7/GFP].
2. The second vector construct carries “Clean ITRs” and a reporter gene that is driven by a weaker hepatocyte specific combination of Apolipoprotein E (ApoE) enhancer and a human alpha-1antitrypsin (hAAT) promoter [rAAV/ApoE/GFP].

Both vectors do not have the 3'UTR enhancer element from wild type AAV. To enhance transduction efficiency, the LK03 (liver tropic) capsid has been used.

In general terms, transduction is the process by which viruses encounter the host cell and introduces its viral material into the cell (Subramanian and Geraghty, 2007). Lentiviral vectors are extensively used for gene transfer and have the ability to integrate different numbers of proviral DNA copies in variable segments of cells. To enhance transduction efficiency of lentiviral vectors a few points were considered. Vesicular stomatitis virus G (VSVG) pseudotyped LVs transduce a wide range of cells such as hiPSCs (dividing) and hepatocytes (non-dividing). The role of the VSVG protein is to recognise a ubiquitously expressed receptor called low density lipoprotein (LDL) receptor which helps the vector to transduce a wide range of cells. Other problems such as electrostatic repulsion forces between negatively charged cells and the viral vector may affect the enveloped viral vector interaction with the target cell (Zare et al., 2016). Hence, to overcome this problem polybrene as a polycationic transduction reagent was used to neutralise the repulsion forces. Further, to enhance transgene expression and successful transduction, the WPRE and cPPT were incorporated in the vector constructs.

On the other hand, the genome of the recombinant AAV vectors is not capable of site-specific integration within the host DNA. However, it remains episomal in the nucleus of transduced cells with reported random integration events with a low frequency - (0.1-1%) of transduction events (Colella et al., 2018). Collective data from different studies suggests the transduction efficiency of AAV vectors in PSCs is highly variable but the reason for this is unclear. In recent years, different methods of gene transfer

into human PCs, such as transfection, electroporation and viral transduction were established. However, these initial reports demonstrated that transduction efficiency of natural AAV serotypes are quite incompetent in gene delivery to human PCs (Brown et al., 2017). The first demonstration of AAV gene editing in hESCs was reported by Mitsui et al. In this study, AAV vector-induced chromosomal editing was performed only in the presence of ROCK inhibitor to support the cells' inherent tendency toward apoptosis (Mitsui et al., 2009). Thereafter, another report demonstrated successful gene modification in hESCs and iPSCs without ROCK inhibitor following transduction with AAV gene targeting vectors (Khan et al., 2011). On contrary, it has been shown that hESCs are extremely sensitive to AAV vector transduction (Hirsch et al., 2011). As a result of ITR-induced cellular stress, p53-dependent apoptotic response is activated. Interestingly, another study confirmed these observations and mediation of apoptosis by AAV vector was extended to iPSCs with a similar proposed mechanism (Hirsch et al., 2011, Rapti et al., 2015). Factors such as different stem cell culture conditions set out in each lab, vector production, titre and sensitivity of the iPSCs, can have a profound impact on overall viral transduction.

Lentiviral vectors are extensively used for gene transfer and have the ability to integrate different numbers of proviral DNA copies in variable segments of cells. The efficiency levels of transduction of a cellular population relies on some experimental parameters affecting the frequency and/or distribution of vector integrations within that population. Hence, measuring VCN in individual cells are required to perform such analysis (Charrier et al., 2011). There is a strong relation between the biological potency of the vector and the frequency of transduced cells as well as the number of integrations per cell (Liu et al., 2008). From a genotoxicity point of view, the number of vector insertions per cell can result in inappropriate transgene expression in cells (Chang and Sadelain, 2007) or from the effects of the integrated cassette (Modlich et al., 2006). Therefore, controlling the number of gene insertions per cell is required for safety considerations. To assess safety and efficacy of the viral vectors, it is vital to determine the distribution of vector copies in the transduced cell population at the single cell level (Charrier et al., 2011). In this study, to determine the number of vector insertions per cell VCN was quantified using quantitative PCR (qPCR) method.

Additionally, standard PCR technique was used to validate the viral insertions within the transduced cell population.

In the previous chapter, P106i hiPSC line was expanded and fully characterised using pluripotency markers. We demonstrated successful differentiation of P106i cells into 3D mature hepatospheres which was confirmed by hepatic functional assays. Hereby, we proposed a robust, xeno-free *in vitro* human hepatic genotoxicity model to study the effects of different viral vectors. In this chapter, the viral vectors were used to transduce *in vitro* 3D heps and P106i cells.

## 4.2. Objectives

The objectives of this chapter are as follows:

- High titre lentiviral production
- Transduction of hiPSC P106i cell line at two time points for proliferation assay
- Transduction of 3D mature hepatospheres (Day 30 of differentiation)
- DNA/RNA extraction of samples and quantification
- Vector copy number (VCN) of transduced cell population
- Detection of lentiviral and rAAV insertions, using PCR method
- Single-cell cloning of transduced hiPSCs using lentiviral vectors
- Detection of positive GFP expression in clones by flow cytometry analysis, fluorescent microscopy and PCR method.

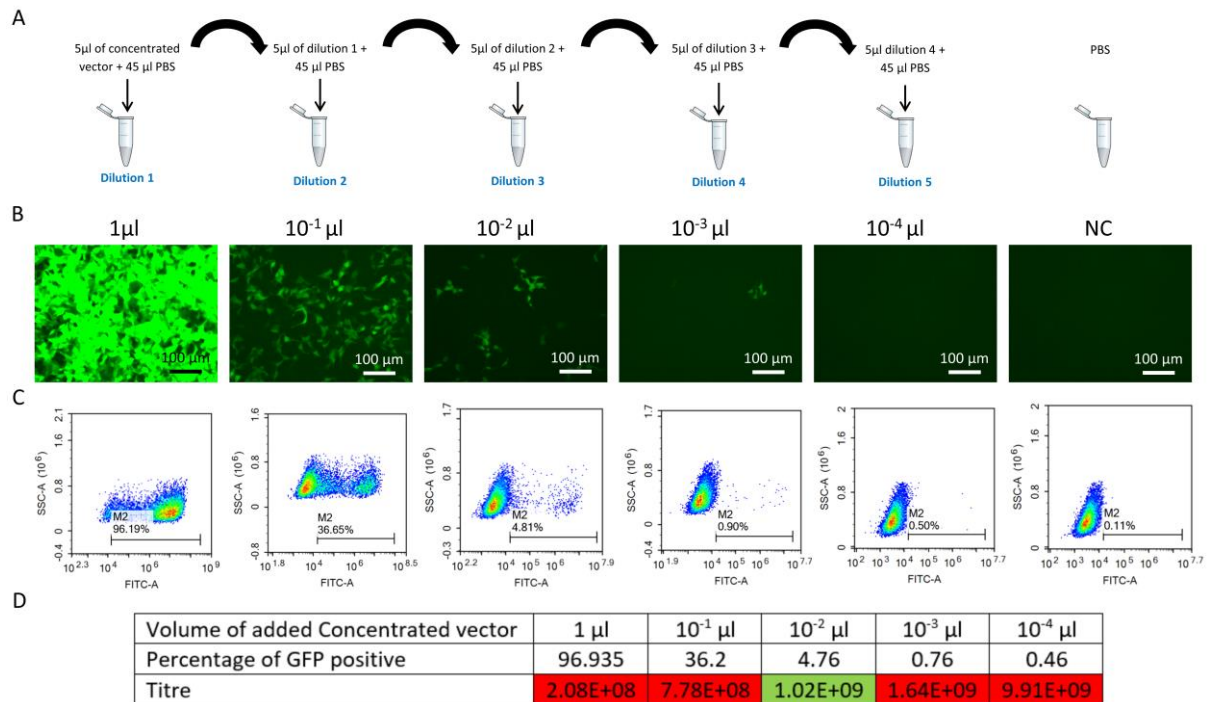
## 4.3. Results

### 4.3.1. High titer GFP expressing lentiviral vector preparation

The lentiviral vectors (pHR, pHV) were regularly generated in house at high titers by transient transfection of adherent human embryonic kidney (HEK) 293T cells. To measure the titer of the viral vector, the concentrated viral vector was added to the HEK293T cells in different dilutions. The relative intensity and percentage of the GFP-positive cells were measured using flow cytometry analysis, as previously described in the Materials and Methods.

From flow cytometry analysis for pHR vector, 1µl of undiluted concentrated virus showed 96.16% GFP-positive cells. To calculate the titre, the concentrated virus was serially diluted (1:10) and the results were as follows: For  $10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$  µl of

diluted samples, the percentage of GFP-positive cells were 35.75%, 4.72%, 0.63% and 0.5%, respectively. Using the flow cytometry result, between (1-30%) of GFP expression percentages were used to calculate the final titer of pHR virus. The titer was calculated as  $1.02 \times 10^9$  TU/ml (Figure 4.1). The range of pHR titer in our lab is between ( $1.04 \times 10^9 - 3.98 \times 10^9$ ). The pHV vector was produced and titered in a similar manner, using flow cytometry analysis and the range of pHV titer in our lab is between ( $3.8 \times 10^9 - 5.8 \times 10^9$ ).



**Figure 4.1: High titre pHR lentiviral preparation and titration, using flow cytometry analysis. A) To measure the titre of the functional pHR virus, the concentrated virus was serially diluted at a ratio of (1:10). B) Representative fluorescent images showing the HEK293T cells transduced with concentrated and diluted GFP-expressing lentivirus. Non-transduced HEK293T cells were used as negative control (NC) (Scale bars 100µM). C) Representative images of the flow cytometry analysis of HEK293T cells transduced with the GFP-expressing pHR lentivirus. The HEK293T cells were trypsinised 36 hours following transduction for the flow cytometry analysis. D) A summary table of the percentages GFP-positive cells analysed by flow cytometry, different volumes of concentrated virus and the final titre of the pHR lentivirus vector (highlighted in yellow). Using the flow cytometry result, between (1-30%) of GFP expression percentages were used to calculate the final titre of pHR virus (highlighted in green). The GFP expression percentages highlighted in red were excluded.**

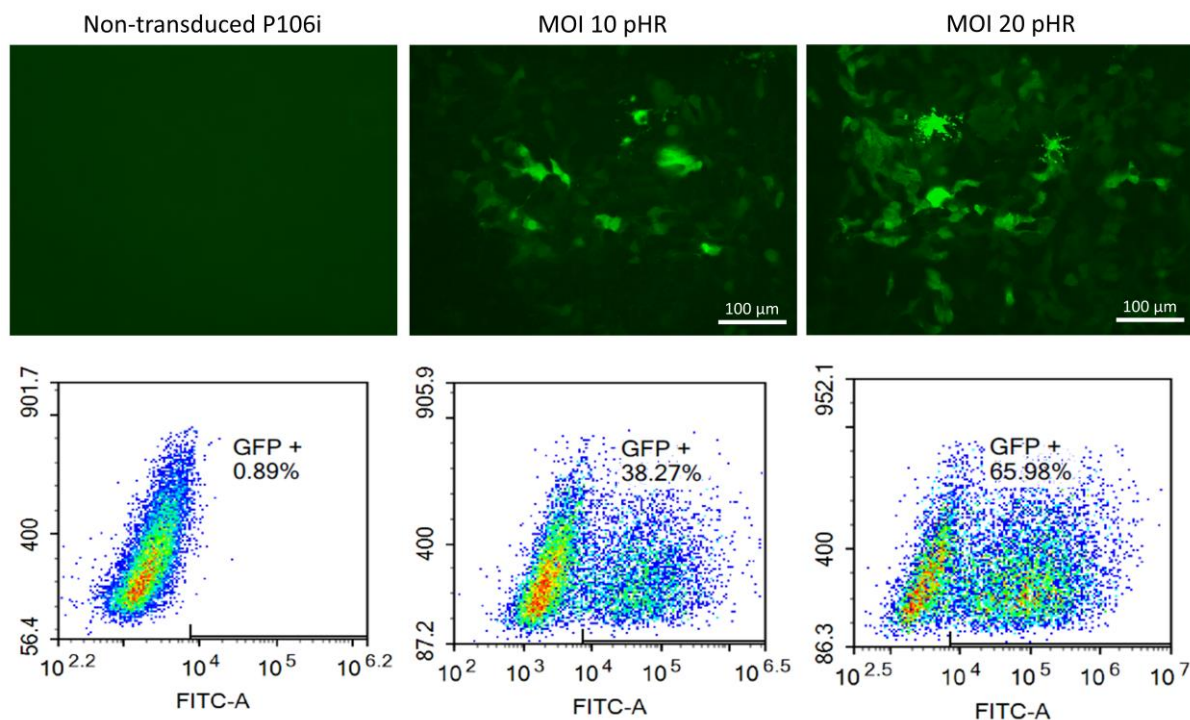
### **4.3.2. Transduction of P106i cell line with pHR lentiviral vector at MOIs 10, 20**

As mentioned earlier, for transduction of P106i cell line, using viral vectors the following points were adapted:

- To accurately calculate the multiplicity of infection (MOI) of the lentiviral and rAAV vectors, healthy undifferentiated P106i cells were single seeded one day prior to transduction.
- For lentiviral transduction, the vector and polybrene reagent were added to the pre-warmed mTsr1 medium supplemented with 10  $\mu$ M Rho-associated Kinase (ROCK) inhibitor (Y27632) to support cells for the first 24 hours.
- However, for rAAV viral transduction, ROCK inhibitor was only added for cell survival. To enhance transduction, the medium containing the virus was kept to the minimum level. Following this step, viral insertions were validated, using vector copy number (VCN) and standard PCR method.

Initially pHR lentiviral vector was used to transduce P106i cell line at different MOIs. Few studies have employed different MOIs ranging from (10 to 35) to transduce PSCs, using lentiviral vectors (Geis et al., 2017, Zhang et al., 2007). Therefore, for this study three different MOIs (10, 20) were considered. This was to determine the best MOI for optimum transduction efficiency while reducing toxicity to the cells. To accurately calculate the multiplicity of infection (MOI), P106i cells were single seeded one day prior to transduction at  $3 \times 10^5$  cells/ml. For the transduction medium, the vector and polybrene reagent were added to pre-warmed mTeSR1 medium supplemented with ROCK inhibitor to support cells for the first 24 hours.

Fluorescent images exhibited positive GFP expression at 24 hours post transduction. GFP positive cells were determined by flow cytometry analysis following transduction. The level of GFP expression for MOIs 10 and 20 were 38.27% and 65.98%, respectively. At MOI 20, the level of GFP expression was higher and stronger compared to MOI 10 at 24 hours post transduction. Further, normal morphology of the cells was retained with no major cell death throughout the well. Rock inhibitor was added to the medium to support the cells during transduction (Figure 4.2).



**Figure 4.2: Transduction of P106i cell line using pHR lentiviral vector at MOIs 10 and 20. Representative fluorescent images showing P106i cells infected with GFP expressing lentivirus (pHR) at MOI 10 and 20. The cells were single seeded one day prior to transduction at density of  $3 \times 10^5$  cells/ml. The transduction medium contained pHR vector, polybrene reagent and ROCK inhibitor to support cells for the first 24 hours. Non-transduced P106i cells were used as control. GFP positive cells were measured by flow cytometry analysis following transduction. The level of GFP expression for MOIs 10 and 20 were 38.27% and 65.98%, respectively. Scale bar: 100 μm.**

#### **4.3.3. Cell viability analysis for P106i cells transduced with pHR lentiviral vector**

Following transduction, cell viability was quantified for the P106i cells transduced with pHR vector at MOIs 10 and 20, using an automated cell counter. Using Trypan Blue staining which penetrates cells with disrupted cell membrane, the number of viable cells were quantified at different timepoints. The cell viability were 79% and 76% 6 hours post-transduction at MOIs 10 and 20, respectively. The cell viability were 94% and 93% 24 hours post-transduction at MOIs 10 and 20, respectively, suggesting successful recovery following transduction (Table 4.1). Based on morphological analysis and cell viability results, MOI 20 was shown to provide the best transduction



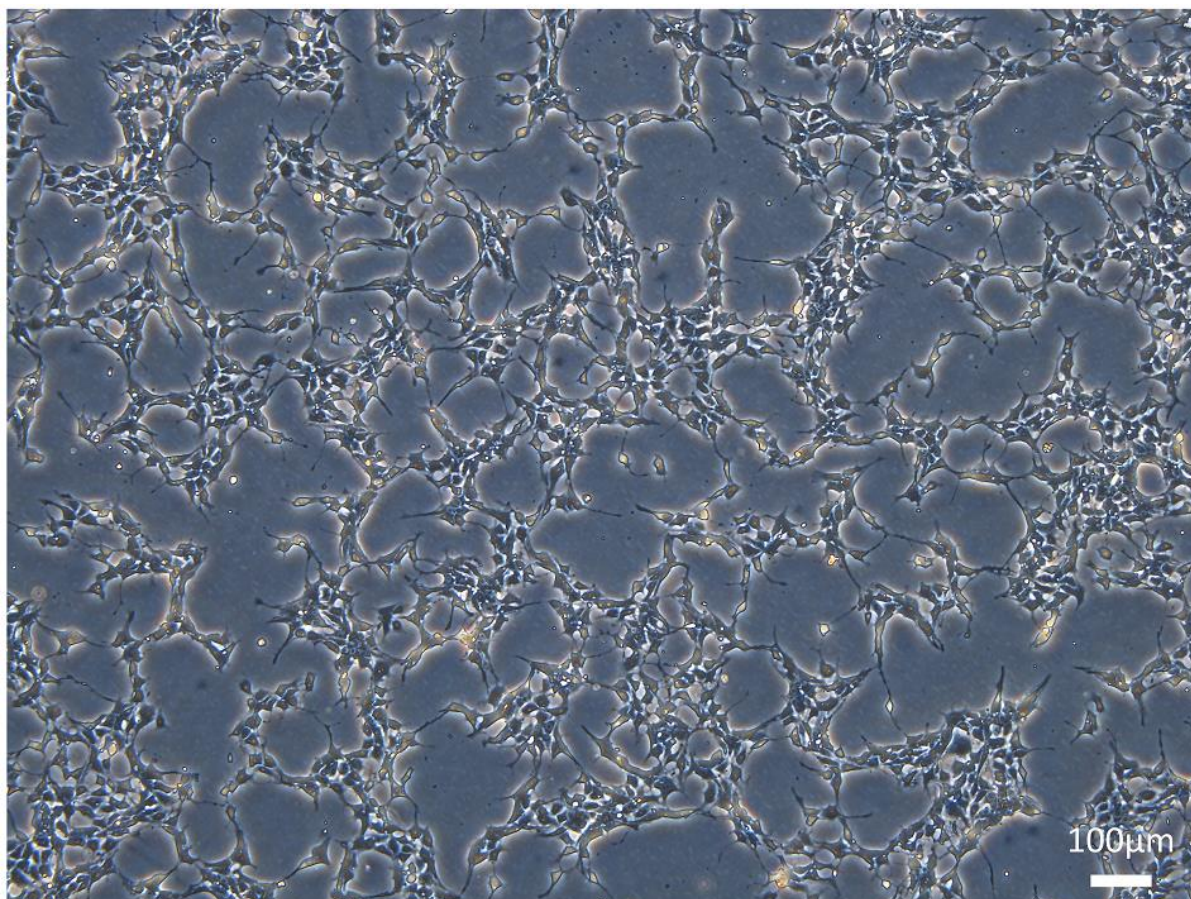
with high cell viability. Therefore, this MOI was applied to transduce P106i cells and 3D heps for the study of viral vectors and downstream analysis.

**Table 4.1. Summary of the average number of cells, the cell viability and the level of GFP expression in P106 hiPSC line transduced with pHR vector at different timepoints, using MOIs 10 and 20.**

Time	Average Number of cells	Cell viability (MOI 10)	Cell viability (MOI 20)
6 hrs	3.2 X 1e5	79%	76%
9 hrs	2.0 X 1e5	78%	75%
12 hrs	3.0 X 1e5	85%	80%
24 hrs	4.8 X 1e5	94%	93%

#### **4.3.4. Brightfield microscopy analysis of the P106i cells transduced with pHR lentiviral vector**

One day following transduction of P106i cells with pHR lentiviral vector at MOI 10, morphology of the cells was monitored using brightfield microscopy. The cells adhered consistently to the wells with minimum number of floating cells. They retained the single-cell morphology clusters without forming dense and compact colonies. Cluster of cells were interconnected (spiky appearance) and no significant stress or cell death was observed. It was speculated that due to the addition of virus and the polycationic reagent (polybrene) cells exhibited such morphology (Figure 4.3). As mentioned earlier, ROCK inhibitor was added to transduction medium to support cell survival for the first 24 hours.

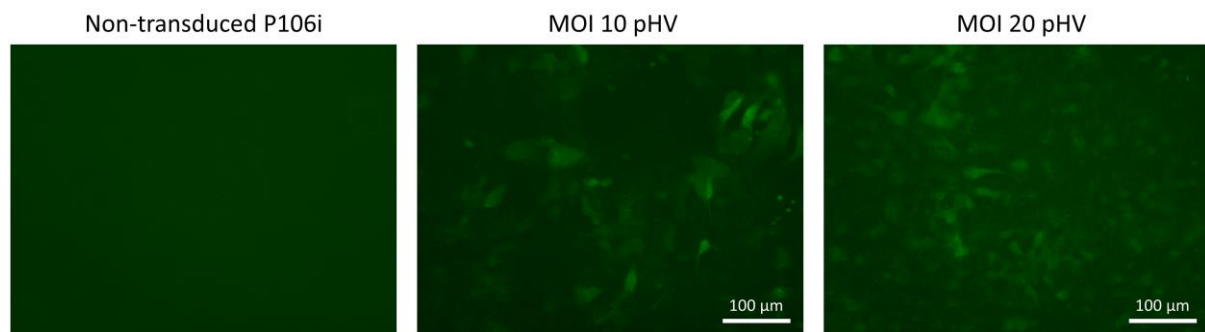


**Figure 4.3: Representative brightfield image of P106i cells transduced with pHR lentiviral vector. One day following transduction, morphology of P106i cells was monitored using brightfield microscopy. The cells retained the single cluster morphology without formation of compact colonies. Cluster of cells were interconnected with minimum level of dead cells. Scale bar: 100μm.**

#### **4.3.5. Transduction of P106i cell line with pHV lentiviral vector, using MOIs 10 and 20.**

Similarly, the pHV lentiviral vector was used to transduce P106i cells at MOIs 10 and 20. To calculate the multiplicity of infection (MOI), P106i cells were single seeded one day prior to transduction at  $3 \times 10^5$  cells/ml. For the transduction medium, the vector and polybrene reagent were added to pre-warmed mTeSR1 medium supplemented with ROCK inhibitor to support cells for the first 24 hours. Fluorescent images exhibited positive GFP expression in a limited number of cells following 24 hours post transduction.

Overall, a low level of GFP expression was observed at MOIs 10 and 20. However, at MOI 20 the level of expression was slightly higher at 24 hours post transduction. The cells retained normal 2D monolayer morphology, as previously described. Although pHR and pHV lentiviral vectors share the same internal promoter to drive transgene expression, the level of GFP expression was much higher and stronger in cells transduced with pHR vector (Figure 4.4).



**Figure 4.4: Transduction of P106i cell line, using pHV lentiviral vector. Representative fluorescent images showing the P106i cells infected with GFP expressing lentivirus vector (pHV) at MOIs 10 and 20. The cells were single seeded one day prior to transduction at a density of  $3 \times 10^5$  cells/ml. Transduction medium contained pHR vector, polybrene reagent and ROCK inhibitor to support cells for the first 24 hours. Non-transduced P106i cells were used as control. From fluorescent images, a low level of GFP expression was detected in P106i infected with pHR vector at MOIs 10 and 20. However, at MOI 20, the cells were expressing a higher level of GFP expression 24 hours post transduction. Scale bar: 100  $\mu$ m.**

#### **4.3.6. Recombinant AAV/GFP serotype 2 transgene (GFP) expression in P106i cells.**

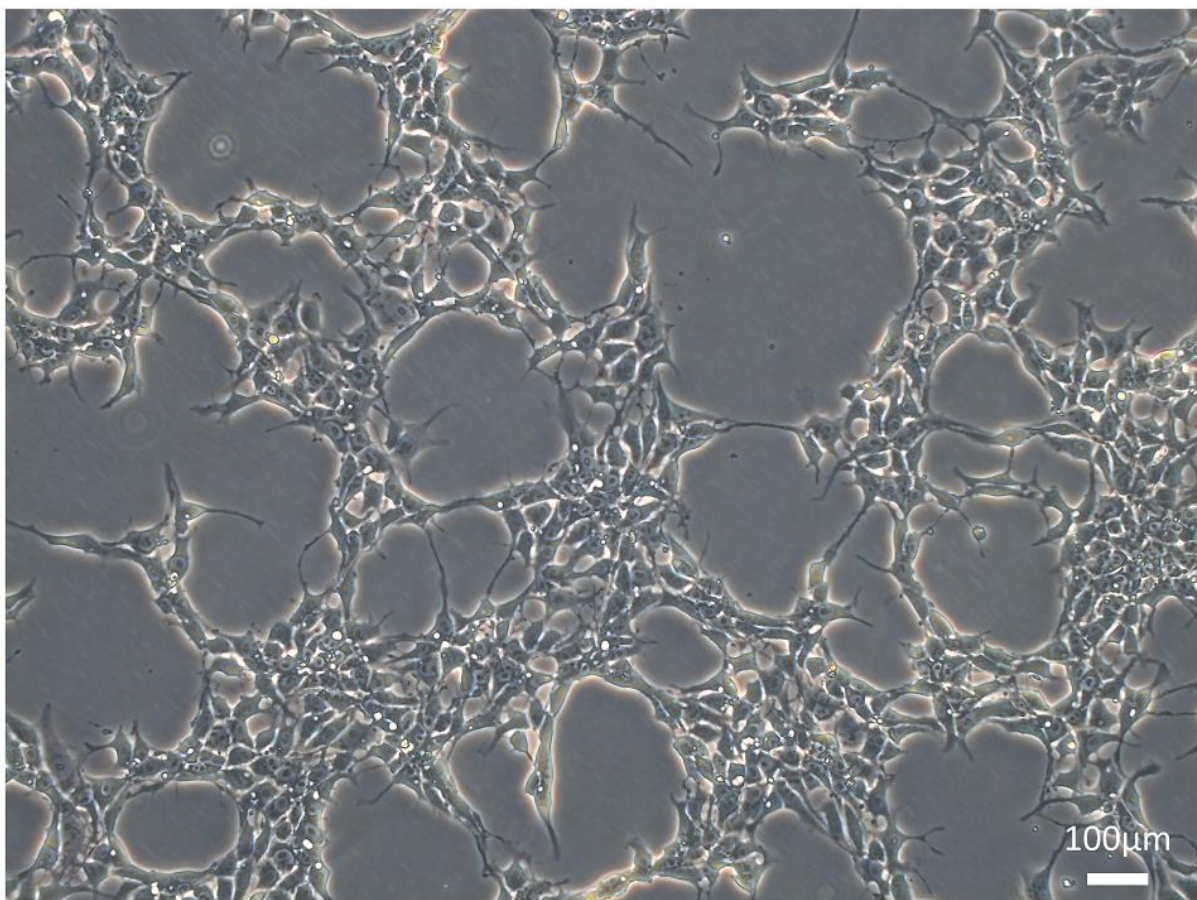
Transduction and transgene expression of recombinant adeno-associated virus (rAAV) with different serotypes (1, 2, 6 and 9) in hiPSC line have been reported. It was reported that AAV2 and AAV6 serotypes have high tropism for PSCs (Rapti et al., 2015). A wide selection of vector genomes per cell ranging from (10,000 to 100,000) have been used (Ellis et al., 2013, Rapti et al., 2015, Duong et al., 2019).

For this study, initially rAAV/GFP serotype 2 vectors at high concentration 100,000 (1E5) vector genomes per cell (vg/cell) was used to transduce P106i cell line. The amount of virus required to achieve an MOI (1E5) was calculated based on cell numbers per well for that particular serotype. Similar to lentiviral transduction, to

enhance cell survival the pre-warmed mTsr1 medium supplemented with ROCK inhibitor for the first 24 hours.

#### **4.3.7. Brightfield microscopy analysis of the P106i cell line following transduction with rAAV/GFP serotype 2 vector.**

One day following transduction, the morphology of the cells was monitored using brightfield microscopy. The cells were adhered consistently to the wells with minimum number of floating cells. The tiny cluster of cells were interconnected, and no significant stress or cell death was observed (Figure 4.5).



**Figure 4.5: Representative brightfield image of P106i cells transduced with rAAV/GFP serotype 2 AAV vector. One day following transduction, the morphology of P106i cells was monitored using brightfield microscopy. The cells retained 2D monolayer morphology with minimum number of dead cells. Scale bar: 100 μm.**

#### 4.3.8. Cell viability and fluorescent microscopy analysis for P106i cells transduced with rAAV/GFP serotype-2 vector.

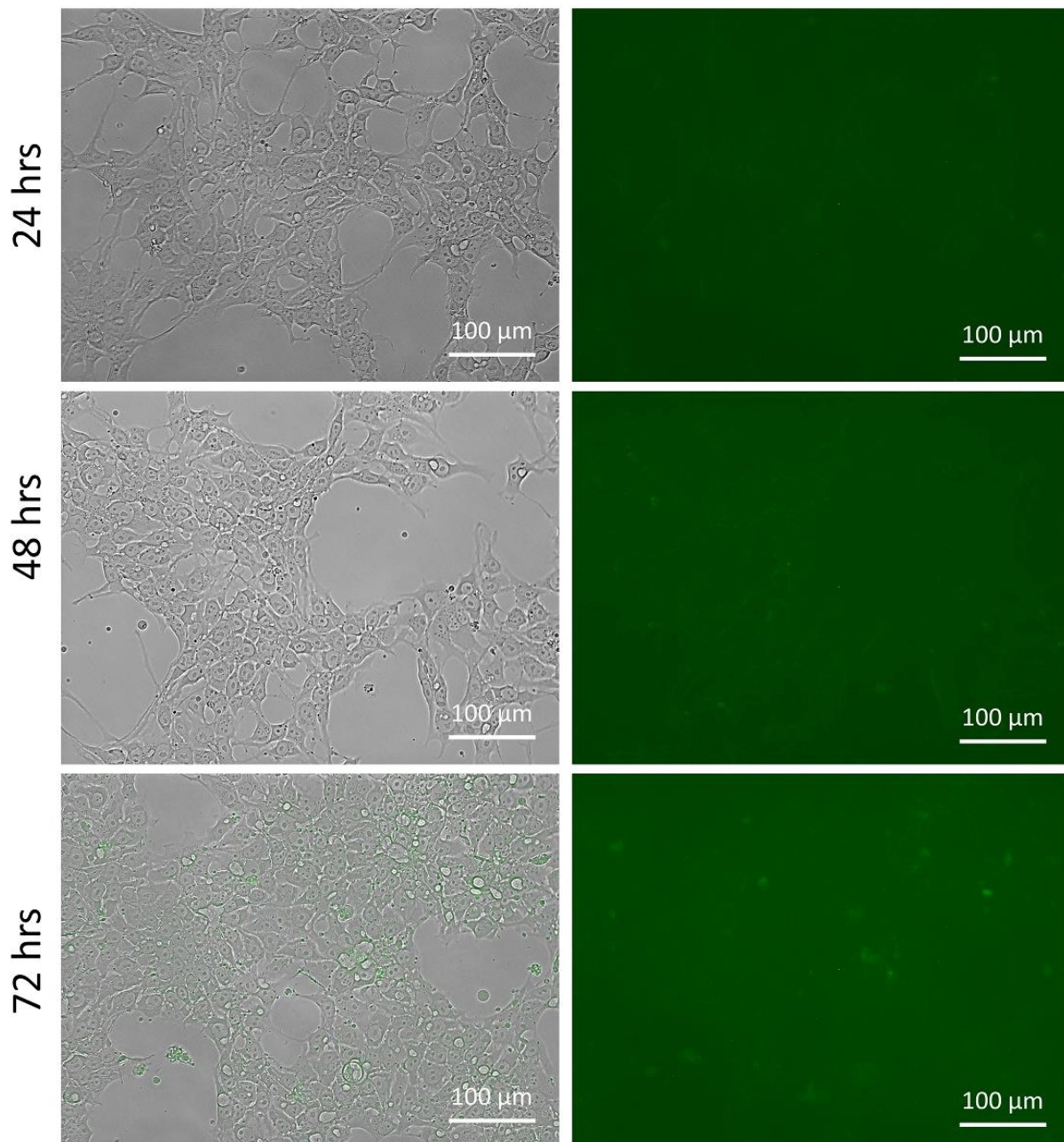
Following transduction with rAAV/GFP serotype 2 at MOI 1E5, viability of P106i cells was analysed using an automated cell counter. Trypan Blue staining was used to differentiate viable and non-viable cells based on dye penetration into the cells with disrupted membrane. Over 75% of the transduced P106i remained viable 6 hours post transduction with no drastic changes up to 72 hrs post transduction (Table 4.2).

The level of GFP expression was monitored using fluorescent microscopy which was also time dependent. Fluorescent microscopy images revealed no positive GFP expression at 6- and 9-hours post-transduction. However, at 72 hours post transduction a low level of GFP expression was detected (Figure 4.6).

**Table 4.2. The average number of cells, the cell viability and the level of GFP expression in P106i cell line transduced with rAAV/GFP serotype-2 vector at different time points using MOI (1E5).**

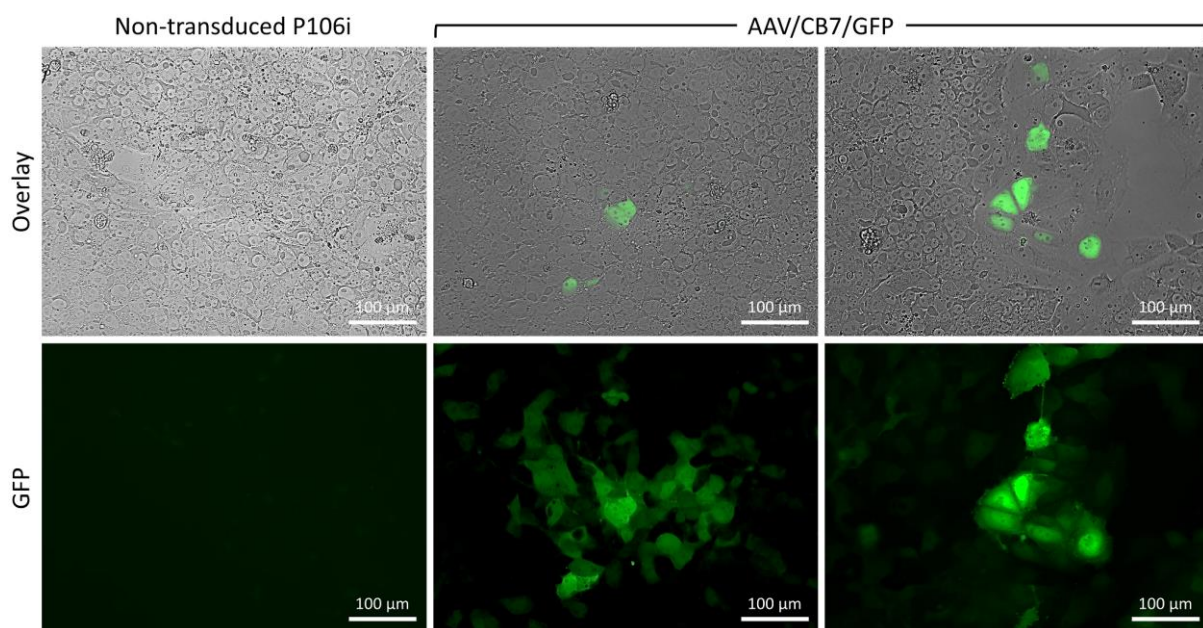
Time	Average Number of cells	Cell viability [MOI (1E5)]
6 hrs	3.5 X 1e5	77%
9 hrs	2.3 X 1e5	80%
12 hrs	2.8 X 1e5	78%
24 hrs	4.6 X 1e4	90%
72 hrs	4.4 X 1e4	91%

## rAAV2/2/ApoE/GFP

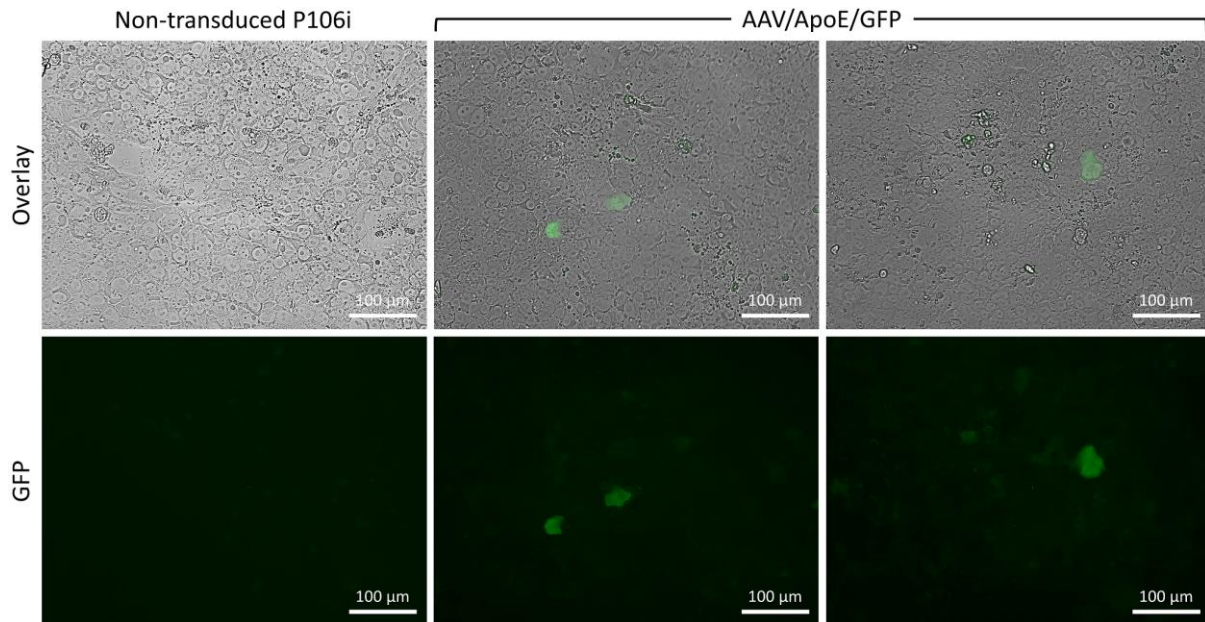


**Figure 4.6: Transduction of P106i cell line, using rAAV/GFP serotype-2 vector. Representative fluorescent images showing the P106i cells infected with GFP expressing rAAV at MOI 1e5. The cells were single seeded one day prior to transduction at a density of  $3 \times 10^5$  cells/ml. Non-transduced P106i cells were used as control. Fluorescent microscopy images revealed no positive GFP expression at 6- and 9-hours post-transduction. However, at 72 hours post transduction a low level of GFP expression was detected. Scale bar: 100  $\mu$ m.**

Following initial testing of rAAV/GFP serotype 2 vectors, the main rAAV vectors (rAAV/CB7/GFP and rAAV/ApoE/GFP) were used to transduce P106i cell line at MOI (1E5). The GFP was expressed by CB7 as a putative and ApoE as a liver-specific promoter in rAAV/CB7/GFP and rAAV/ApoE/GFP, respectively. The level of GFP signal was intensified by time post-transduction and was considerably higher in rAAV/CB7/GFP vector compared to the rAAV/ApoE/GFP. The rAAV/CB7/GFP led to the highest level of GFP expression in P106i cell line after 72 hours post-transduction (Figure 4.7). Whereas the cells transduced with the ApoE promoter exhibited a lower level of GFP expression and only in a few numbers of cells (Figure 4.8).



**Figure 4.7: Transduction of P106i cell line, using the rAAV/CB7/GFP serotype-2 vector. Representative fluorescent images showing P106i cells infected with a GFP-expressing rAAV at MOI 1e5. Non-transduced P106i cells were used as control. The level of GFP expression was promoter- and time-dependent. The rAAV/CB7GFP serotype-2 vector led to the highest level of GFP expression after 72 hours post transduction. Scale bar: 100 µm.**



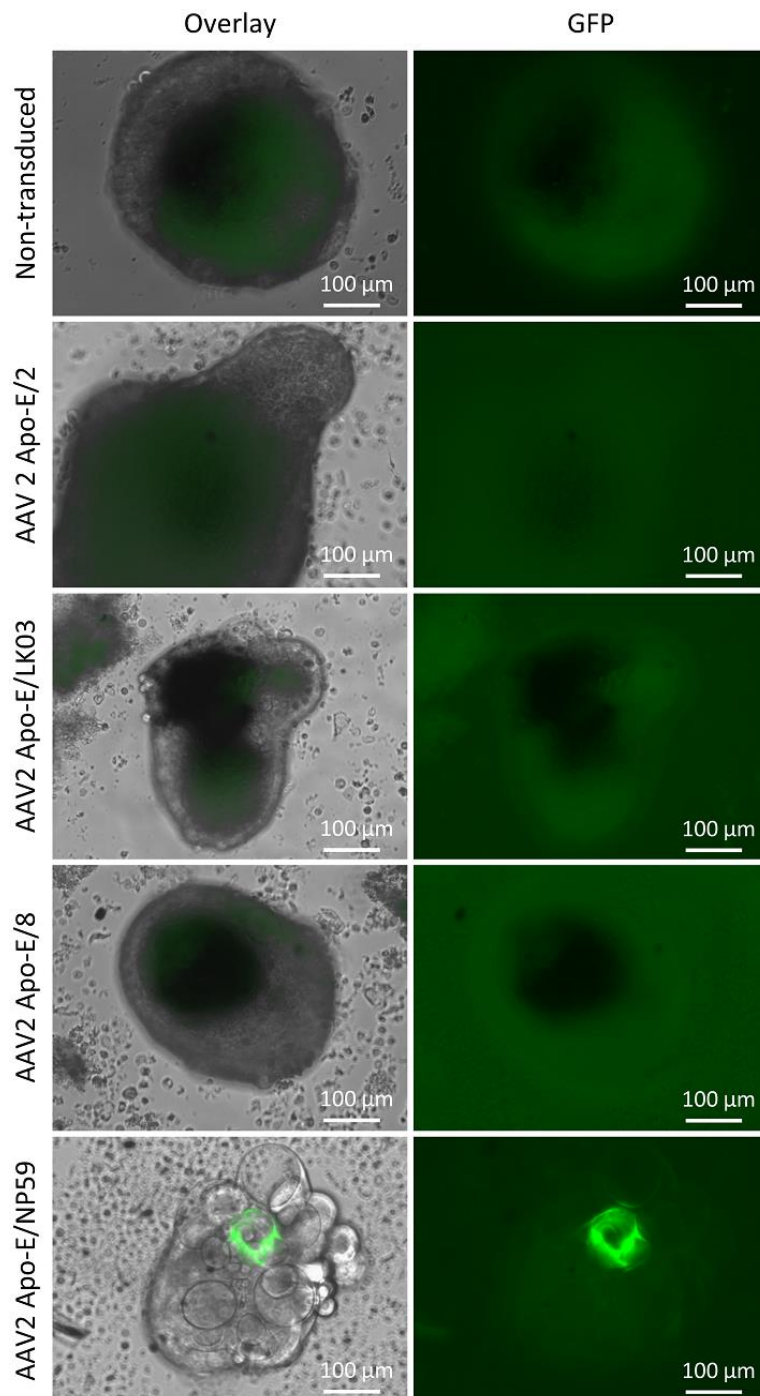
**Figure 4.8: Transduction of P106i cell line, using the rAAV/ApoE/GFP serotype-2 vector. Representative fluorescent images showing P106i cells infected with GFP-expressing rAAV at MOI 1e5. Non-transduced P106i cells were used as control. The level of GFP expression was promoter- and time-dependent. The rAAV/ApoE/GFP serotype 2 vector exhibited lower level of GFP expression. The positive GFP expression was detected only in a few of the cells. Scale bar: 100  $\mu$ m.**

#### **4.3.9. Transduction of *in vitro* 3D heps, using the rAAV/GFP serotype 2 vectors with different liver specific capsids.**

Initially for our *in vitro* 3D hepatic model, rAAV/GFP serotype 2 vectors with different liver specific capsids (AAV8-LK03-NP59 and AAV2) were used to test the transduction efficiency. The transgene expression was driven under the ApoE promoter. It has been reported that LK03 capsid has a preference to transduce human cells and particularly primary human hepatocytes 100-fold better than AAV8, while AAV2 shows the least preference to transduce human hepatocytes (Lisowski et al., 2014).

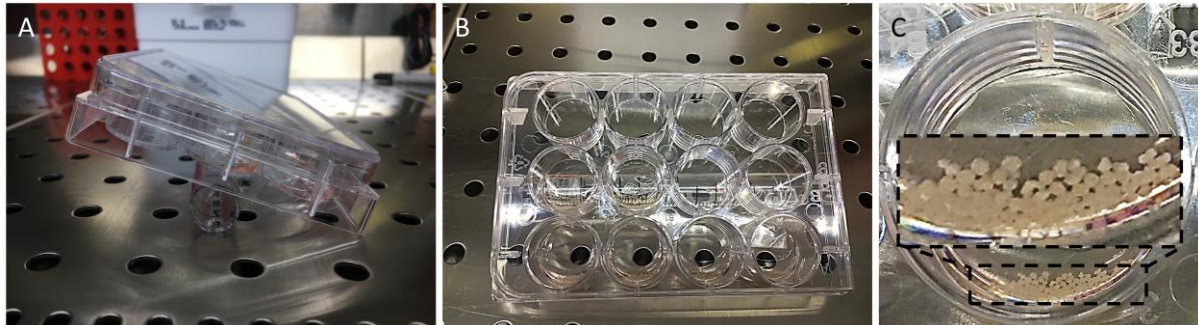
Fluorescent microscopy images revealed positive GFP expression in a few 3D heps transduced with NP59 capsid. However, there was no GFP expression in 3D heps transduced with LK03 and AAV8 capsids. Although there was some level of GFP expression, transduction efficiency was quite low (Figure 4.9).





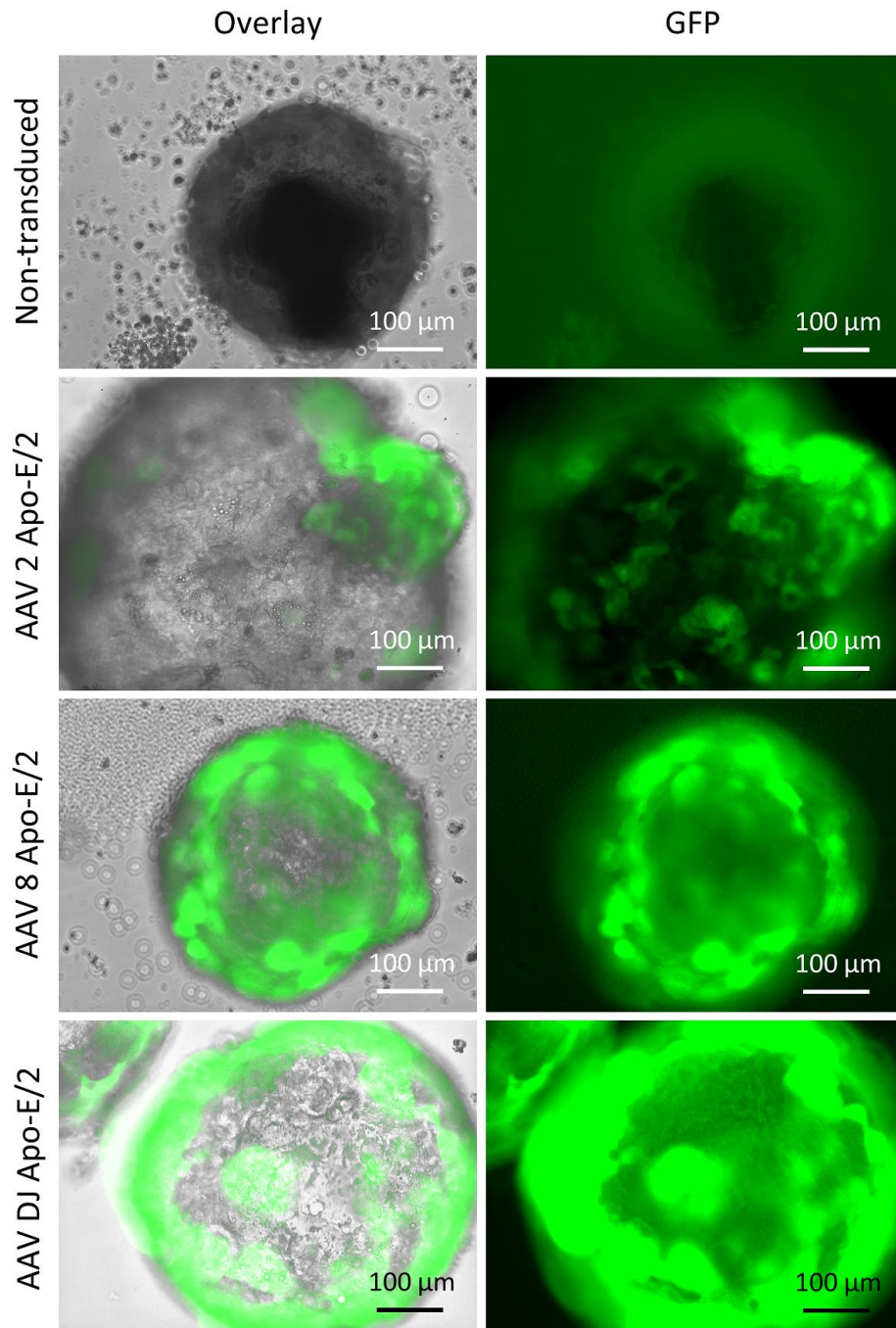
**Figure 4.9: Transduction of in vitro 3D heps, using rAAV/ApoE/GFP serotype-2 vector with AAV2, LK03, AAV8 and NP59 liver specific capsids. Representative fluorescent images showing 3D heps infected with GFP expressing rAAV/GFP vectors at MOI 1e5. Non-transduced 3D heps were used as control. From fluorescent microscopy images no positive GFP expression detected in 3D heps transduced with rAAV/GFP vectors with AAV2, LK03 and AAV8 capsids. However, only a few GFP-positive cells were observed in 3D heps transduced with rAAV/GFP vector the NP59 capsid. Scale bar: 100  $\mu$ m.**

To improve transduction efficiency in 3D heps, the plate was modified from a flat to a tilted position during the transduction period. The tilted position allows removal of the supernatant medium to avoid further dilution of vectors. Then the tilted plates containing 3D heps were transduced and transferred into the incubator for the first 24 hours. The plates were occasionally shaken to ensure the hepatospheres were not attaching together while the whole 3D surface was exposed to the viral vectors (Figure 4.10).

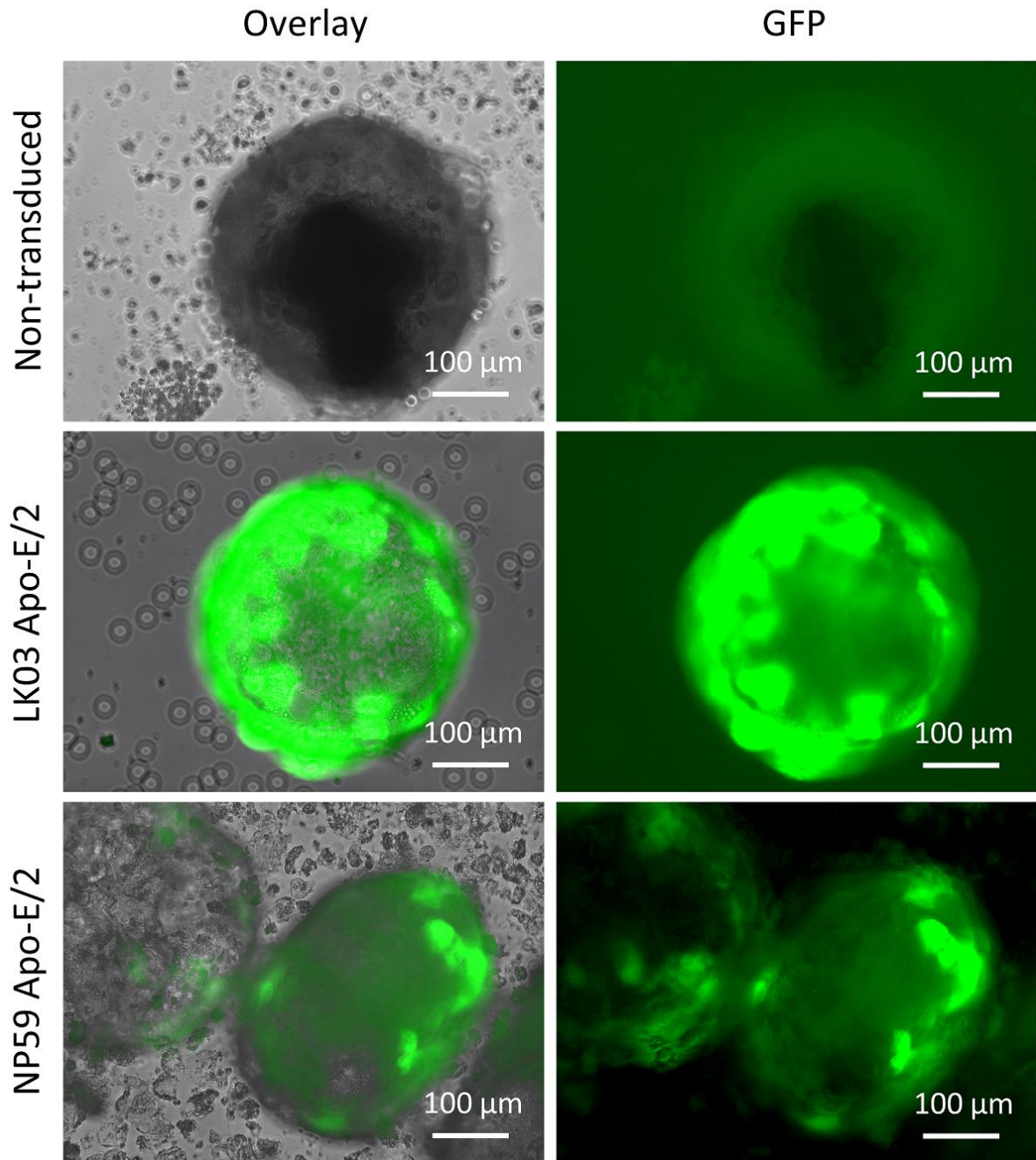


**Figure 4.10: Modifications to enhance transduction efficiency in 3D heps. A) The position of the plate was altered from a flat to a tilted position inside the incubator for the first 24 hours. B) An overall view of the plate showing the collection of 3D heps in one corner of the wells with a low volume of transduction medium. C) A magnified version of one well containing 3D heps collectively on bottom of tilted plate.**

Following implementing the above-mentioned amendments, the second set of transductions was performed successfully. On day 30 of differentiation, rAAV/GFP serotype 2 vectors with different liver specific capsids (AAV8, LK03, NP59, DJ, AAV2 and AAV2/2) were used to transduce the P106i-derived 3D heps at MOI (1E5). Fluorescent microscopy images revealed positive GFP expression in 3D heps confirming a high efficiency of transduction. The level of GFP expression was much higher compared to the first trial, showing a strong GFP expression throughout the 3D heps architecture. The outer layer of the 3D heps exhibited high level of GFP expression 96 hours post transduction (Figures 4.11 & 4.12). Furthermore, the 3D heps transduced with rAAV/GFP vectors with LK03 and DJ capsids displayed the highest level of GFP expression compared to other capsids.



**Figure 4.11: Transduction of in vitro 3D heps, using the rAAV/ApoE/GFP serotype 2 vectors with AAV2, AAV8 and DJ liver specific capsids. Representative fluorescent images showing 3D heps infected with GFP expressing rAAV/GFP vectors at MOI 1e5. Non-transduced 3D heps were used as control. Fluorescent microscopy images revealed positive GFP expression in 3D heps. The outer layer of the 3D heps exhibited the highest level of GFP expression 96 hours following post transduction. Scale bar: 100  $\mu$ m.**



**Figure 4.12: Transduction of in vitro 3D heps, using rAAV/ApoE/GFP serotype 2 vectors with LK03 and NP59 liver specific capsids. Representative fluorescent images showing 3D heps infected with GFP expressing rAAV vectors at MOI 1e5. Non-transduced 3D heps were used as control. Fluorescent microscopy images revealed positive GFP expression in 3D heps. The outer layer of the 3D heps exhibited the highest level of GFP expression 96 hours following post transduction. Scale bar: 100  $\mu$ m.**

Following testing mature 3D heps with rAAV/GFP serotype 2 vectors, two others main rAAV vectors [(rAAV/CB7/GFP)/strong promoter] and [(rAAV/ApoE/GFP)/ weak

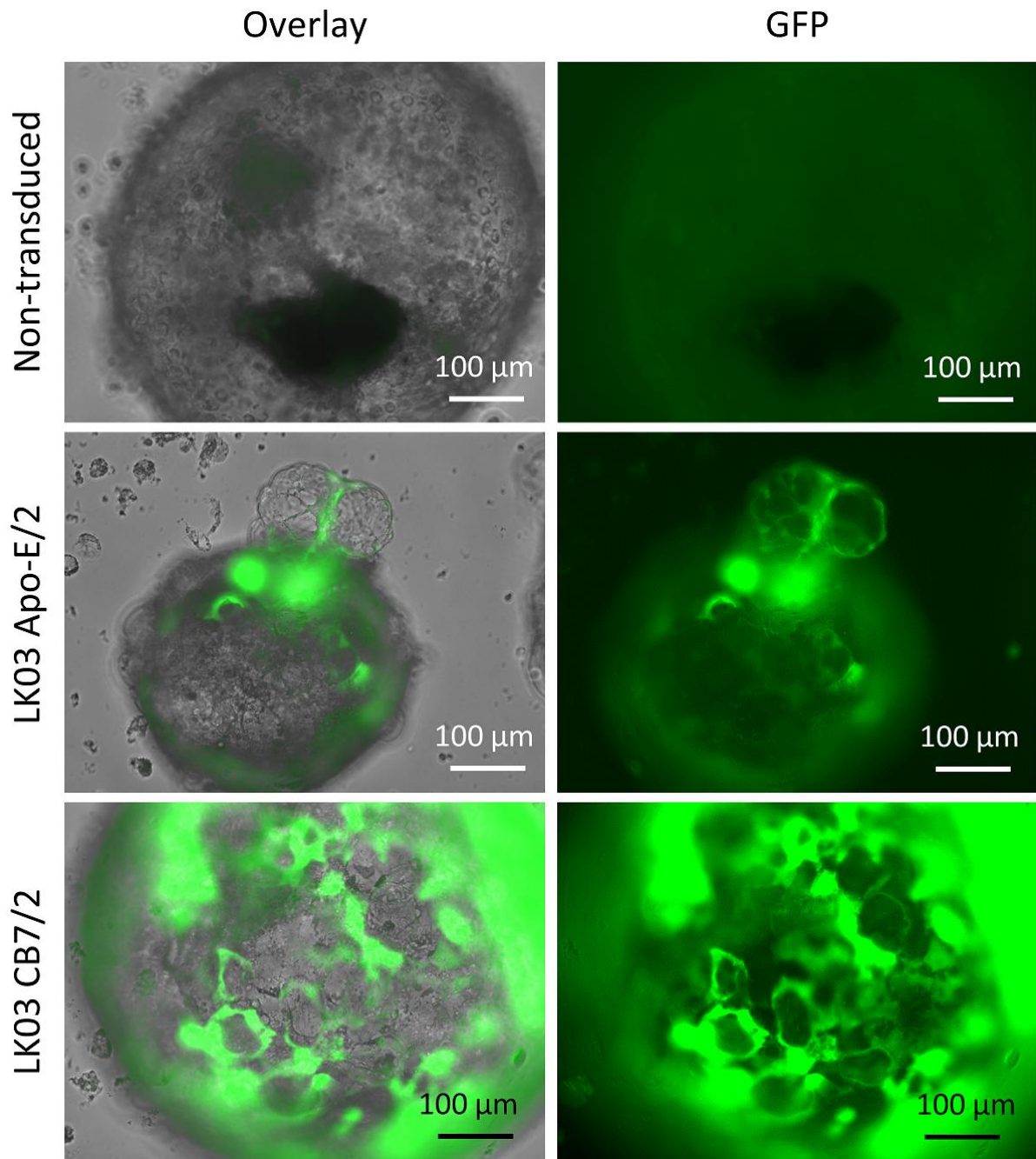
promoter] were employed to transduce 3D heps on day 30 of differentiation at MOI (1E5). The LK03 capsids were incorporated in these vectors and demonstrated high efficacy for mature hepatocytes. Fluorescent images revealed positive GFP expression in 3D heps transduced with rAAV/GFP with strong and weak promoters. However, the level of GFP expression was more prominent in hepatospheres transduced with rAAV/CB7 compared to rAAV/ApoE at 96 hours post transduction (Figure 4.13).

#### **4.3.10. Transduction of *in vitro* 3D heps with pHR and pHV lentiviral vectors at MOI 20.**

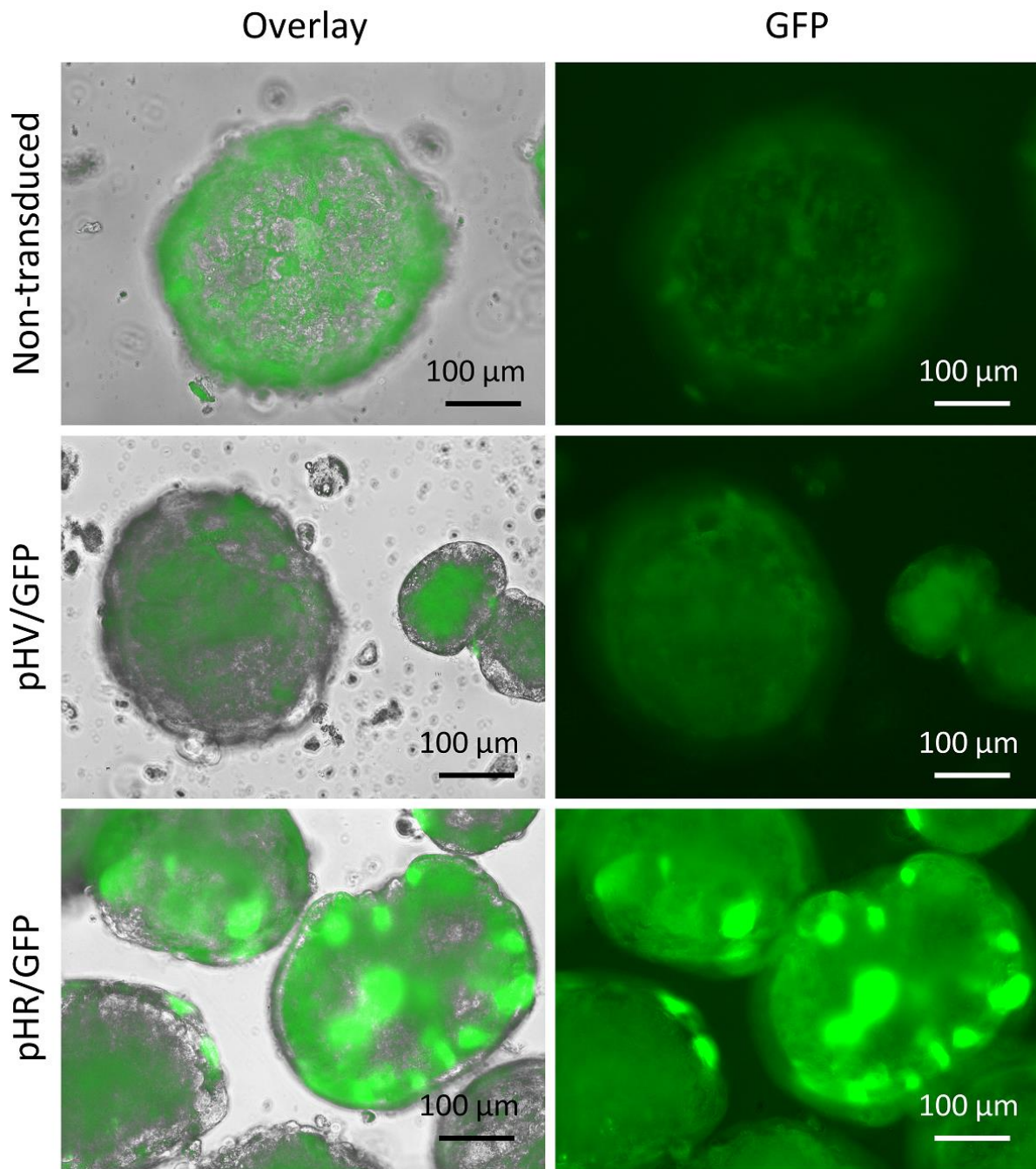
The pHR and pHV lentiviral vectors were employed to transduce 3D heps at MOI 20. The transduction medium contained vectors and polybrene reagent which added to pre-warmed Williams's E medium. Fluorescent images exhibited very low GFP expression in 3D heps transduced with pHV lentiviral vectors 48 hours following transduction. However, a strong level of GFP expression was detected in 3D heps transduced with pHR lentiviral vector. The same pattern was observed in transduction of P106i cells which showed GFP expression was much higher and stronger in cells transduced with pHR vector (Figure 4.14).

#### **4.3.11. DNA/RNA extraction of P106i hiPSCs and 3D heps transduced with rAAV/GFP and lentiviral vectors**

Following successful transduction of P106i hiPSCs and P106i-derived 3D heps with lentiviral vectors (pHR, pHV) and rAAV/GFP serotype 2 vectors (rAAV/ApoE/GFP, rAAV/CB7/GFP), the samples were harvested. The DNA/RNA of the collected samples were extracted using extraction kits, as previously described. The purity and quantity of the samples were quantified using the Nanodrop machine. These samples were sent to our collaborator in GeneWerk (Germany) for downstream analysis of vector copy number (VCN) and PCR tests. The DNA and RNA samples for 3D heps and P106i cells transduced with lentiviral and rAAV/GFP vectors are summarised in (Tables 4.3, 4.4, 4.5 & 4.6).



**Figure 4.13: Transduction of in vitro 3D heps, using rAAV/ApoE/GFP and rAAV/CB7/GFP serotype 2 vectors. Representative fluorescent images showing 3D heps infected with GFP expressing rAAV vectors at MOI 1e5. Non-transduced 3D heps were used as control. Fluorescent microscopy images revealed positive GFP expression in 3D heps transduced with rAAV/ApoE/GFP and rAAV/CB7/GFP vectors. The level of GFP expression was much higher in hepatospheres transduced with rAAV/CB7 compared to rAAV/ApoE at 96 hours post transduction. Scale bar: 100  $\mu$ m.**



**Figure 4.14: Transduction of in vitro 3D heps, using pHR and pHV lentiviral vectors. Representative fluorescent images showing 3D heps infected with GFP expressing pHR and pHV lentiviral vectors at MOI 20. Non-transduced P106i cells were used as control. Fluorescent microscopy images revealed positive GFP expression in 3D heps transduced with pHR lentiviral vector. However, no positive GFP expression was detected in 3D heps transduced with pHV lentiviral vector. The level of GFP expression was greater in 3D heps transduced with pHR vector at 48 hours post transduction. Scale bar: 100  $\mu$ m.**

**Table 4.3: The DNA samples for P106i samples transduced with the lentiviral and rAAV/GFP vectors.**

P106i samples	ng/ul	Abs 260/280	Abs 260/230	Quantity (μl)
Control (non-transduced P106i)	334.7	1.96	2.19	120
pHR transduced	512.9	1.96	2.25	120
pHV transduced	411.8	1.98	2.22	120
rAAV/ApoE/GFP transduced	722.2	2.03	2.05	120
rAAV/CB7/GFP transduced	515.2	2.04	2.13	120

The RNA samples for P106i hiPSCs transduced with the lentiviral and rAAV/GFP vectors are summarised in (Table 4.4).

**Table 4.4: The RNA samples for P106i samples transduced with the lentiviral and rAAV/GFP vectors.**

P106i samples	ng/ul	Abs 260/280	Abs 260/230	Quantity (μl)
Control (Uninfected P106i)	738.6	1.96	2.19	30
pHR transduced	318.9	1.96	2.25	30
pHV transduced	524.7	2.08	2.02	30
rAAV/ApoE/GFP transduced	2510.5	2.08	2.07	30
rAAV/CB7/GFP transduced	2046.2	2.10	2.06	30

**Table 4.5: The DNA samples for 3D heps transduced with lentiviral and rAAV/GFP vectors.**

3D HLC samples	ng/ul	Abs 260/280	Abs 260/230	Quantity (μl)
CONTROL-1 (Uninfected HLC)	79.5	1.88	1.51	160
CONTROL-2 (Uninfected HLC)	88.5	1.91	1.51	160
CONTROL-3 (Uninfected HLC)	116.6	1.93	1.53	160
HLC-infected with PHV-1	107.9	1.93	1.60	160
HLC-Infected with PHV-2	125.6	1.93	1.74	160
HLC-Infected with PHV-3	85.7	1.90	1.50	160
HLC-Infected with PHR-1	81.6	1.91	1.57	160
HLC-Infected with PHR-2	75.5	1.90	1.52	160
HLC-Infected with PHR-3	71.3	1.98	1.39	160
HLC-infected with AAV-SAFE-1	70.5	1.89	1.42	160
HLC-infected with AAV-SAFE-2	77.0	1.93	1.50	160
HLC-Infected with AAV-SAFE-3	59.1	1.88	1.50	160
HLC-Infected with AAV-UNSAFE-1	92.6	1.93	1.55	160
HLC-Infected with AAV-UNSAFE-2	106.6	1.91	1.65	160
HLC-Infected with AAV-UNSAFE-3	78.1	1.98	1.54	160



The RNA samples for 3D heps transduced with lentiviral and rAAV/GFP vectors are summarised in (Table 4.6).

**Table 4.6: The RNA samples for 3D heps transduced with lentiviral and rAAV/GFP vectors.**

3D HLC samples	ng/ul	Abs 260/280	Abs 260/230	Quantity (µl)
CONTROL-1 (Uninfected HLC)	380	1.83	1.52	70
CONTROL-2 (Uninfected HLC)	208.1	1.92	2.05	70
CONTROL-3 (Uninfected HLC)	188.3	1.95	2.22	70
HLC-infected with PHV-1	162.8	1.90	2.12	60
HLC-Infected with PHV-2	175.6	1.95	2.25	60
HLC-Infected with PHV-3	187.7	1.85	1.83	60
HLC-Infected with PHR-1	183.1	1.92	2.22	60
HLC-Infected with PHR-2	146	1.95	2.27	60
HLC-Infected with PHR-3	109	1.94	2.28	60
HLC-infected with AAV-SAFE-1	73.6	1.95	2.24	60
HLC-infected with AAV-SAFE-2	133.1	1.93	2.19	60
HLC-Infected with AAV-SAFE-3	118	1.94	2.20	60
HLC-Infected with AAV-UNSAFE-1	106.4	1.90	2.03	60
HLC-Infected with AAV-UNSAFE-2	173.1	1.89	2.05	60
HLC-Infected with AAV-UNSAFE-3	204.9	1.94	2.19	60

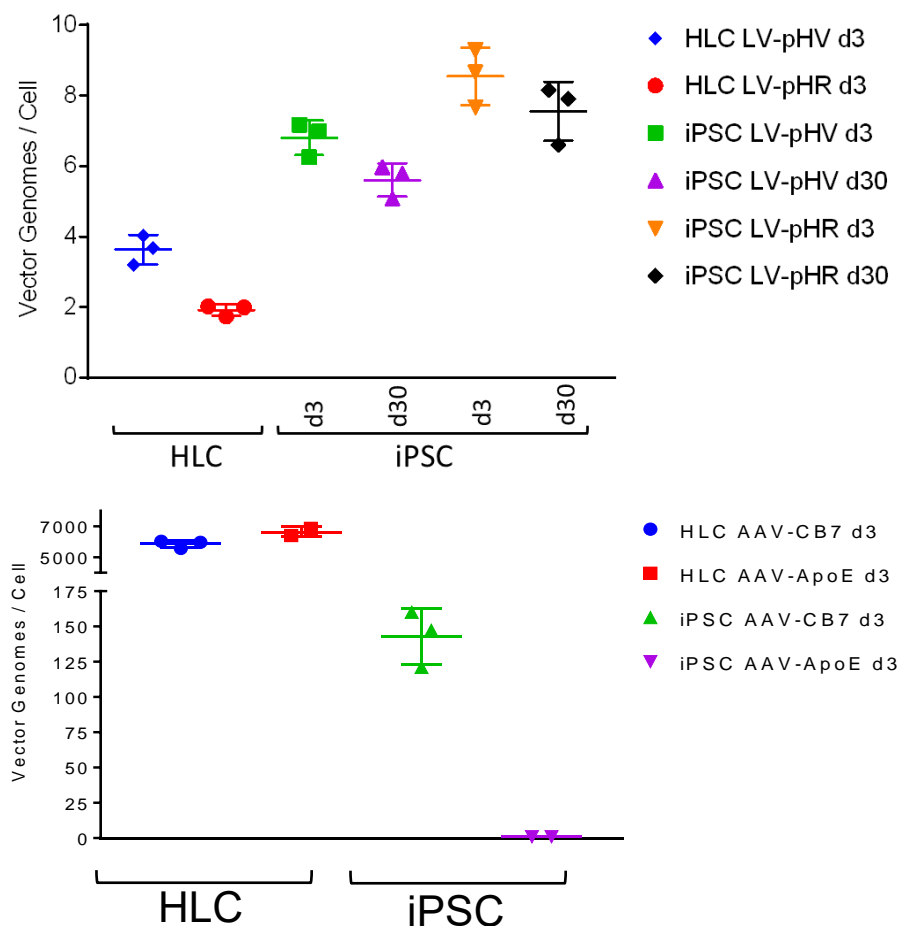
#### 4.3.12. Vector copy number

To evaluate the transduction of P106i hiPSCs and P106i-derived 3D heps with pHR and pHV, the VCN per cell was measured using a standardised quantitative PCR (q-PCR) in collaboration with GeneWerk (Germany). To determine average copy number of cell population transduced with viral vectors, various dilutions of a plasmid containing both vector and genomic sequences were prepared using the q-PCR method. By measuring VCN, it was possible to compare the number of viral insertions per cell between hiPSC and 3D heps. The result showed that the P106i transduced with the pHV lentiviral vector has a lower VCN of 7 compared to P106i transduced with the pHR lentiviral vector which has a VCN of 9.

There was an overall reduction in clonality and VCN number in P106i transduced with both vectors at day 30 time point, validating the necessity to study two time points. However, in P106i-derived 3D heps transduced with the pHR and pHV lentiviral vectors lower VCNs were obtained compared to P106i cells thus confirming the lower

efficiency of lentiviral vectors in non-dividing cells. The VCN in 3D heps transduced with pHV was higher compared to the VCN in 3D heps transduced with pHR vector (Figure 4.15). In contrast, in 3D heps transduced with rAAV/CB7/GFP and rAAV/ApoE/GFP serotype 2 vectors the VCN was profoundly higher 6000 and 7000 VG/Cell, respectively. However, in P106i transduced with rAAV vectors the VCN was much lower compared to the mature hepatocytes. The VCN for P106i transduced with rAAV/CB7/GFP was measured as 130 VG/Cell whilst for P106i transduced with rAAV/ApoE/GFP was quantified as nearly 0. This result demonstrates higher transduction efficiency and GFP expression in the P106i cell line, mediated by rAAV/CB7/GFP vector.

From this result, it can be concluded that distribution of VCN in hiPSCs and 3D mature hepatocytes may depend upon experimental variables such as vector construct designs, different cell types and viral titres.



**Figure 4.15: Quantification of vector copy number (VCN) in P106i and 3D heps, using rAAV/GFP and lentiviral vectors. Vector genomes per cell were determined by Taqman quantitative PCR (qPCR) by absolute quantification and the data was normalised**

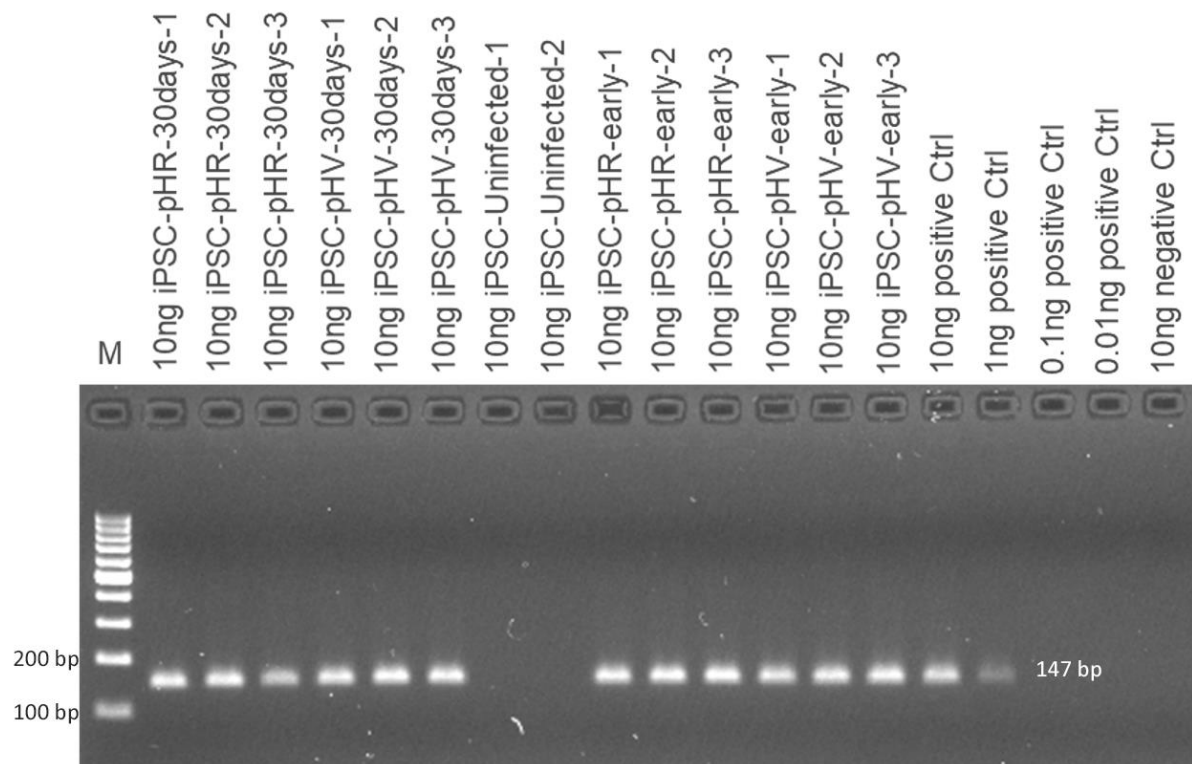
considering that one cell contains 6.6 pg of DNA. The result has shown that the P106i transduced with pHV lentiviral vector has a lower VCN compared to P106i transduced with pHR lentiviral vector. There was an overall reduction in clonality and VCN number in P106i transduced with both vectors at day 30 time point. In 3D heps transduced with pHR and pHV lentiviral vectors VCNs was much lower compared to P106i stage. In contrast, in 3D heps transduced with rAAV/CB7/GFP and rAAV/ApoE/GFP serotype 2 vectors the VCN was profoundly higher. Abbreviations: HLC; mature 3D heps, iPSC; P106i cells, d3; day 3 post transduction, d30; day 30 post transduction.

#### **4.3.13. Detection of pHR and pHV lentiviral and rAAV/GFP insertions in P106i and 3D heps**

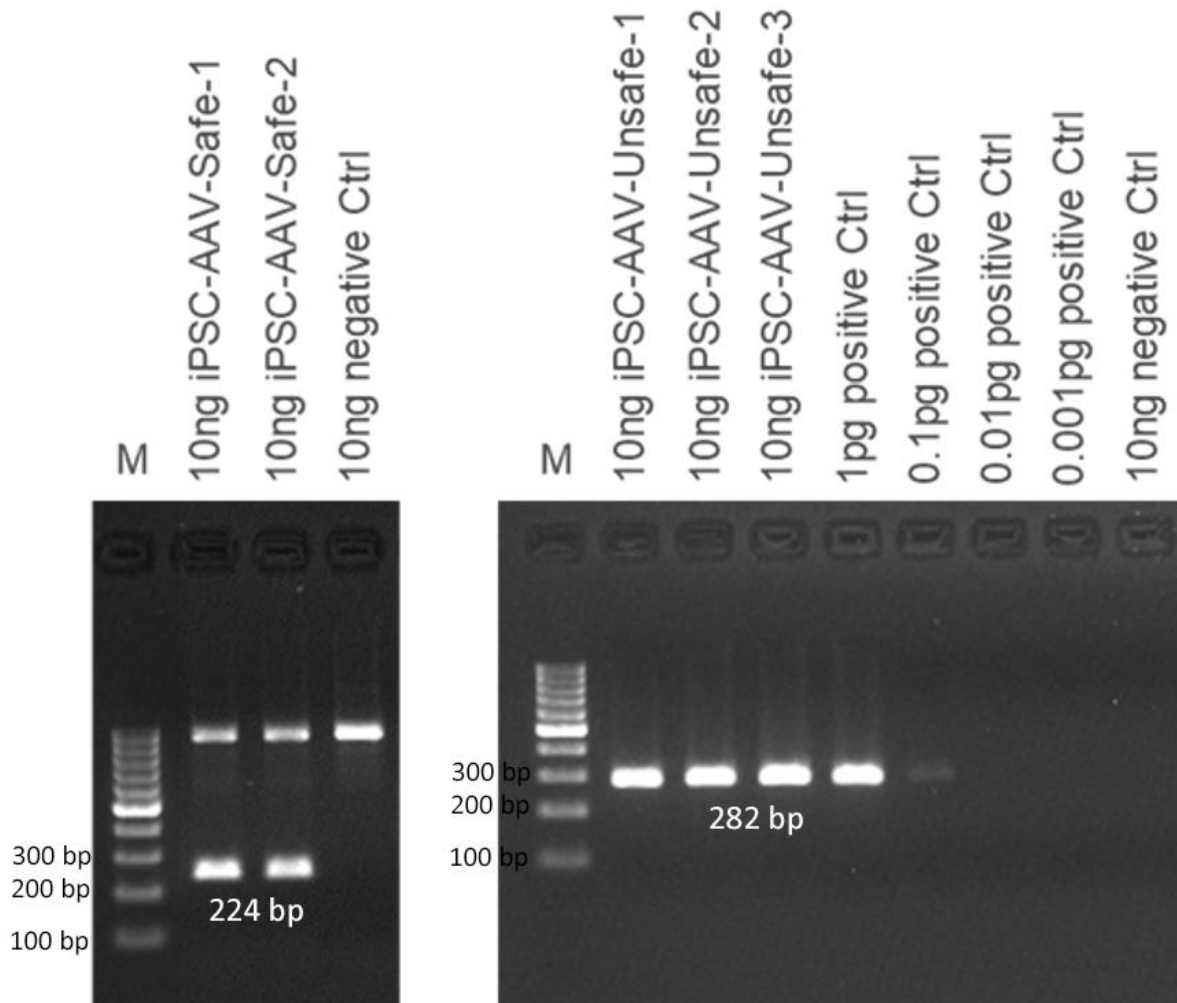
Validation of pHR and pHV lentiviral insertions in P106i hiPSC and 3D heps was performed using a standard PCR method. This result was carried out in collaboration with GeneWerk (Germany). From extracted DNA of transduced samples with lentiviral vectors, PCR was performed to amplify a segment of LTR specific region which is unique to HIV-1 based lentiviral vectors, using specific primer sequences. The primers designed for pHR and pHV lentiviral vectors have an annealing temperature of 51°C and produced 147 bp fragment as seen in the gel images. Similarly, the PCR was performed to detect insertions in cells transduced with rAAV/GFP vectors. The primers for rAAV/ApoE/GFP and rAAV/CB7/GFP vectors were specifically designed for their promoters and generated 224 bp and 282 bp bands, respectively. Uninfected P106i cells and 3D heps were used as controls.

From the gel image results, positive LTR bands (147 bp) were detected in P106i samples transduced with pHR and pHV lentiviral vectors for early and late time points. Similarly, positive bands were observed in P106i samples transduced with rAAV/ApoE/GFP and rAAV/CB7/GFP vectors at 224 and 282 bp, respectively. No bands were detected in uninfected P106i and negative control samples. (Figures 4.16, Figure 4.17).

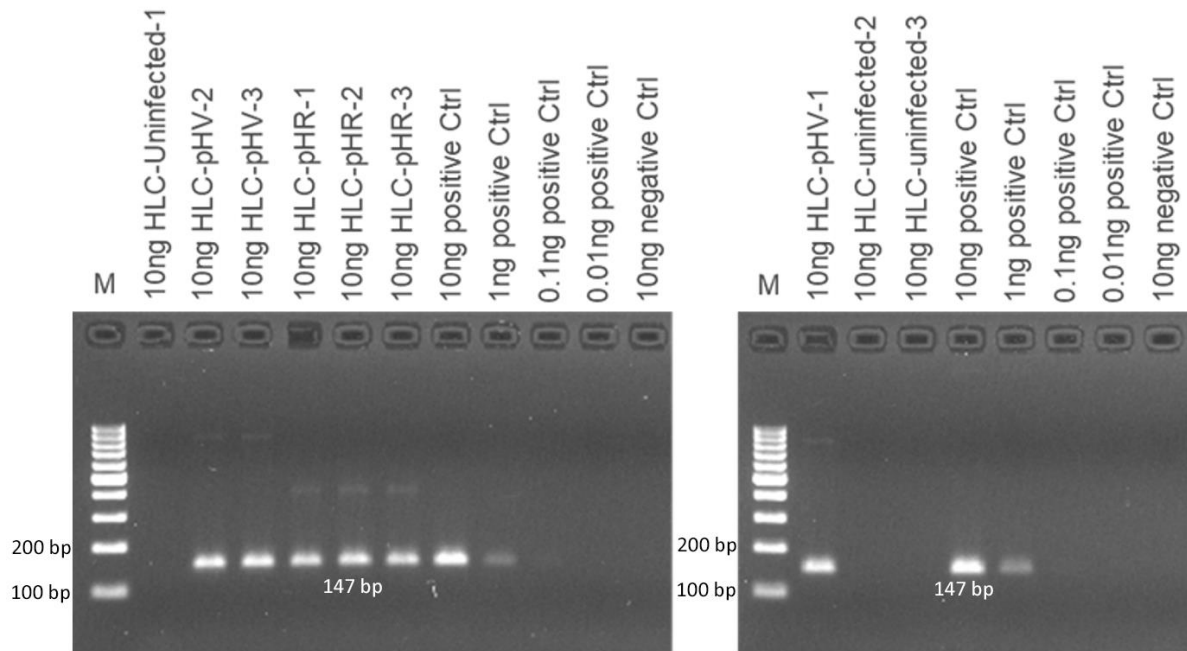
Further, results revealed positive LTR bands (147 bp) in 3D heps transduced with pHR and pHV lentiviral vectors. Also, positive bands were detected in 3D hepatosphere samples transduced with rAAV/ApoE/GFP and rAAV/CB7/GFP vectors at 224 and 282 bp, respectively (Figures 4.18, 4.19).



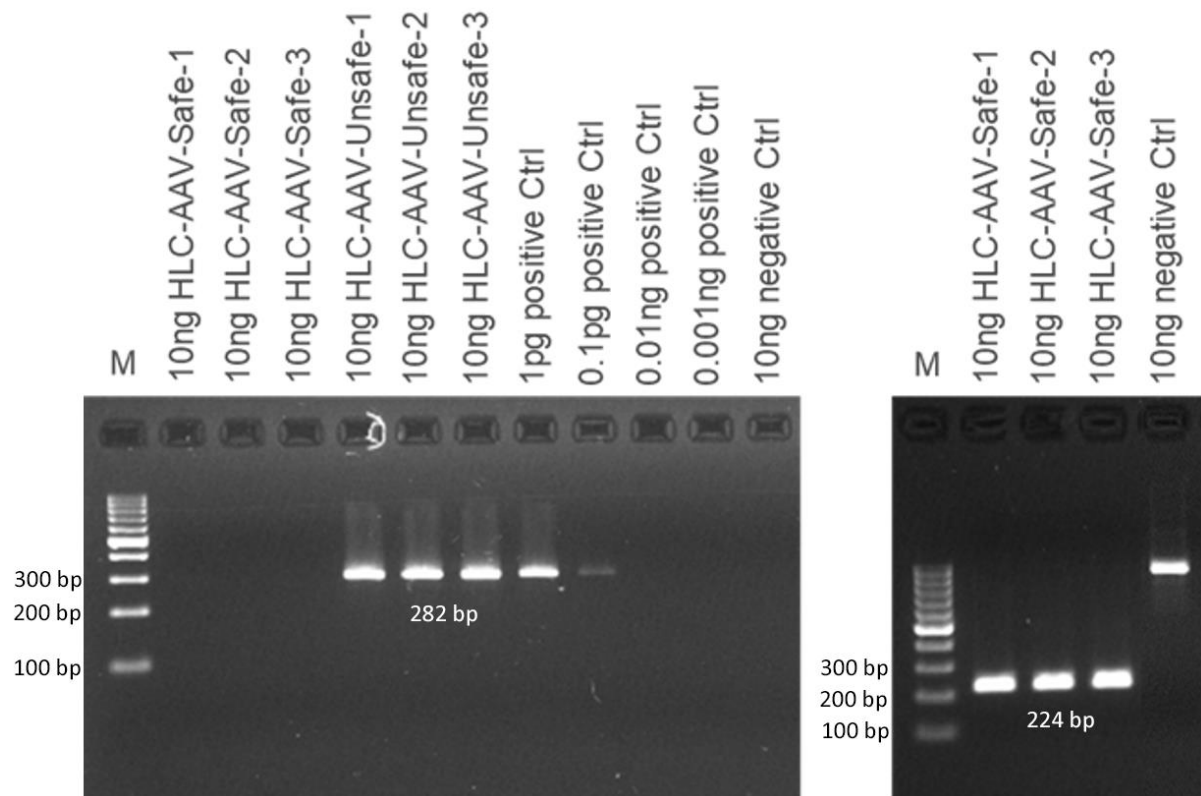
**Figure 4.16: Detection of lentiviral insertions in P106i hiPSCs transduced with pHR and pHV lentiviral vectors. To validate insertions, a segment of LTR specific region was amplified using PCR. The result revealed positive LTR bands in P106i samples transduced with pHR and pHV lentiviral vectors at early and late time points. The positive LTR band was detected at 147 bp. No positive LTR band detected in uninfected P106i and negative control samples. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V.**



**Figure 4.17: Detection of rAAV viral insertions in P106i hiPSCs transduced with rAAV/ApoE and rAAV/CB7 vectors, using PCR method. The result revealed positive bands in P106i samples transduced with rAAV/GFP vectors. The positive bands were detected at 224 bp and 282 bp in P106i samples transduced with rAAV/ApoE/GFP and rAAV/CB7/GFP, respectively. No positive bands were detected in the uninfected P106i and the negative control samples. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V. Abbreviation: Safe vector; rAAV/ApoE/GFP (weak promoter), Unsafe vector; rAAV/CB7/GFP (strong promoter).**



**Figure 4.18: Detection of lentiviral insertions in 3D heps transduced with pHR and pHV lentiviral vectors. To validate insertions, a segment of LTR specific region was amplified, using PCR method. The result revealed positive LTR bands in 3D heps transduced with pHR and pHV lentiviral vectors. The positive LTR band was detected at 147 bp. No positive band was observed in the uninfected P106i and the negative control samples. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V.**



**Figure 4.19: Detection of rAAV viral insertions in 3D heps transduced with rAAV/ApoE and rAAV/CB7 vectors using PCR. The result revealed positive bands in 3D heps transduced with rAAV/GFP vectors. The positive bands were detected at 224 bp and 282 bp in 3D heps transduced with rAAV/ApoE/GFP and rAAV/CB7/GFP, respectively. No positive bands detected in uninfected P106i samples. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V. Abbreviation: Safe vector; rAAV/ApoE/GFP (weak promoter), Unsafe vector; rAAV/CB7/GFP (strong promoter), HLC;3D heps.**

#### **4.3.14. Single-Cell Cloning (SCC) of P106i hiPSCs transduced with pHR and pHV lentiviral vectors**

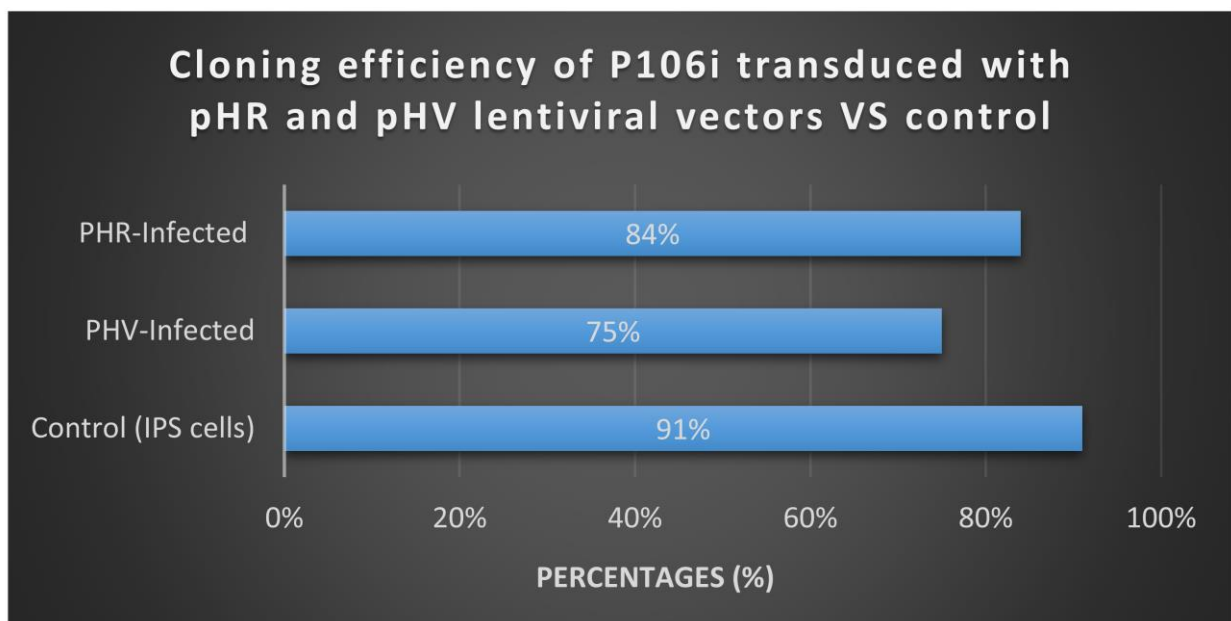
In order to see the effect of gene expression in the proximity of viral insertions at individual cell level, SCC was performed. Prior to SCC, hiPSCs were transduced with lentiviral vectors as previously described. To perform SCC, mTeSR1 medium mixed with CM at a ratio of 1:1 was used to feed clones until healthy colonies established in culture. In addition, the medium was supplemented with 10  $\mu$ M ROCK inhibitor to enhance the cell survival.

Initially the concentration of 2 cells/ml was used to seed two 12-well plates (500  $\mu$ l/well) for each lentiviral vector. After 7 days, the plate was scanned for colonies.

Establishment of 6 colonies for each vector was achieved. Individual colonies were picked and expanded in 6-well plate to obtain enough genomic material for downstream analysis. Following expansion, the DNA/RNA was extracted, and the samples were sent to our collaborator (GeneWerk, Germany) for downstream analysis.

#### 4.3.15. Cloning efficiency of P106i transduced with pHR and pHV lentiviral vectors

Prior to SCC, the cloning efficiency of P106i transduced with pHR and pHV lentiviral vectors were quantified. The cloning efficiency of P106i transduced with pHR vector was much higher (84%) compared to the P106i transduced with pHV vector (75%). However, uninfected P106i cells exhibited the highest level of cloning efficiency. This result demonstrates that the SCC is feasible, using the P106i cell line and the protocol (Figure 4.20).

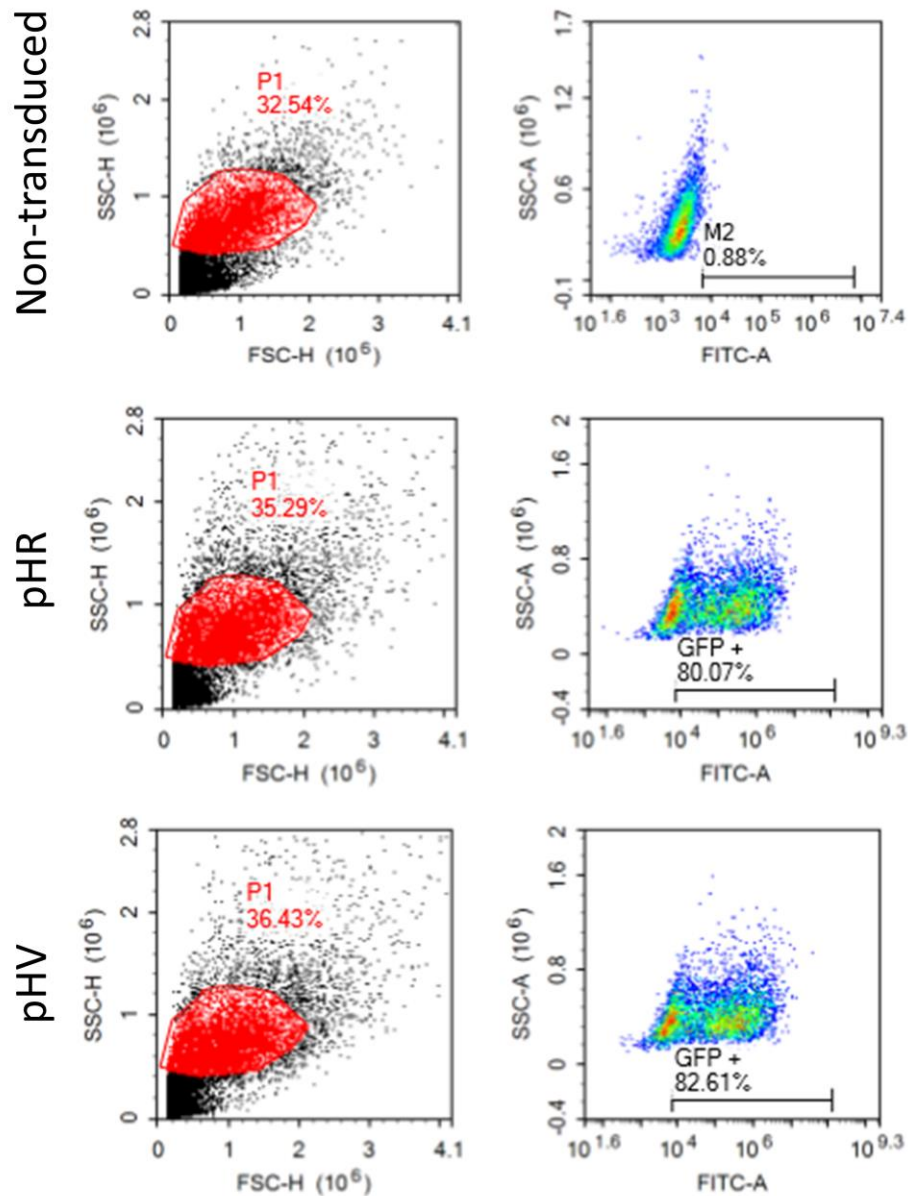


**Figure 4.20: Cloning efficiency of P106i hiPSCs transduced with pHR and pHV lentiviral vectors. The P106i cells transduced with pHR and pHV lentiviral vectors exhibited high cloning efficiency, confirming the feasibility of the SCC study.**

In the first step, the P106i cell line was transduced with pHR and pHV lentiviral vectors. After 72 hrs, the cells were analysed for positive GFP expression using flow cytometry analysis. The result showed positive GFP expression in P106i cells transduced with pHR (80.7%) and pHV (82.6%) lentiviral vectors, confirming the transduction. The



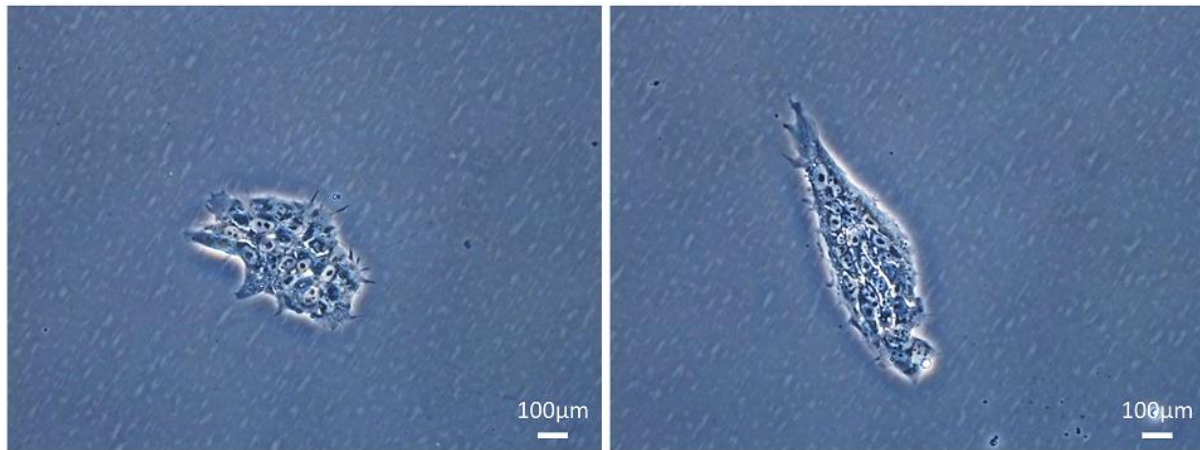
level of GFP expression was higher compared to the first transduction at MOI 20. No GFP expression was detected in the control sample (Figure 4.21).



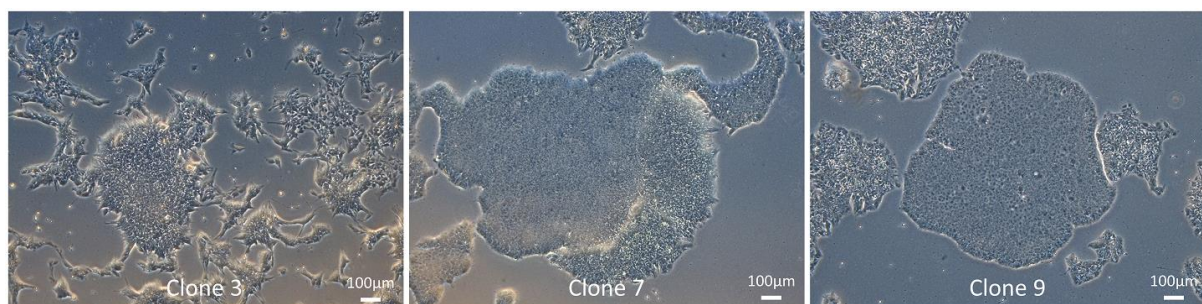
**Figure 4.21: Flow cytometry analysis of P106i cells transduced with pHR and pHV lentiviral vectors prior to SCC. P106i cells were analysed for positive GFP expression using flow cytometry analysis. The results exhibited high positive GFP expression in P106i cells transduced with pHR (80.7%) and pHV (82.6%), confirming the transduction.**

Following successful transduction of P106i cells with lentiviral vectors, SCC was performed for each transduced pool of cells, as previously described. After a week, 6 clones for each vector was formed and carefully expanded. Morphology of clones were analysed using brightfield microscopy. Each clone formed a compact structure with a

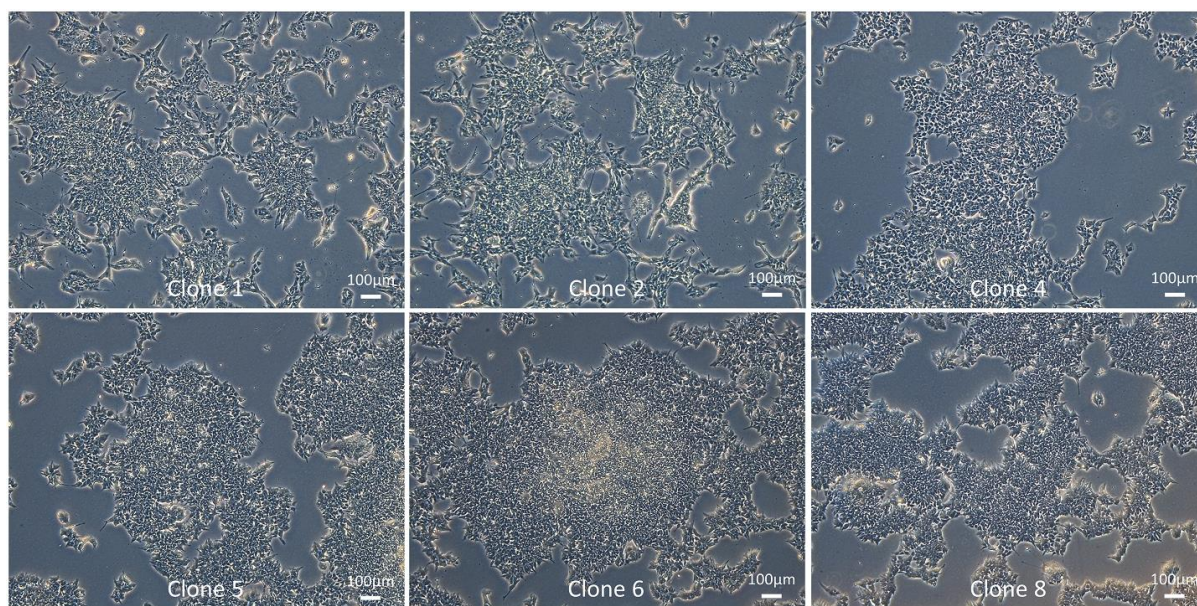
distinct border, confirming the stability of the colonies for expansion. As the clones became larger, each clone was transferred to one well of a 6 well plate for expansion. Brightfield microscopy images revealed the successful expansion of the clones. Some clones formed a tight and dense structure while others formed loose colonies which interconnected to each other. Clearly, the morphology of the cells varies differently following viral transduction in comparison with the cells which have not been infected. (Figures 4.22, 4.23, 4.24).



**Figure 4.22: Morphology of single clone following SCC. Representative brightfield images of single clones following SCC. Following 7 days after SCC, single clones were formed. Each clone formed a compact and dense structure with distinct borders. The cells retained 2D monolayer morphology prior to expansion, confirming the stability of each single clones. Scale bar: 100 µm.**

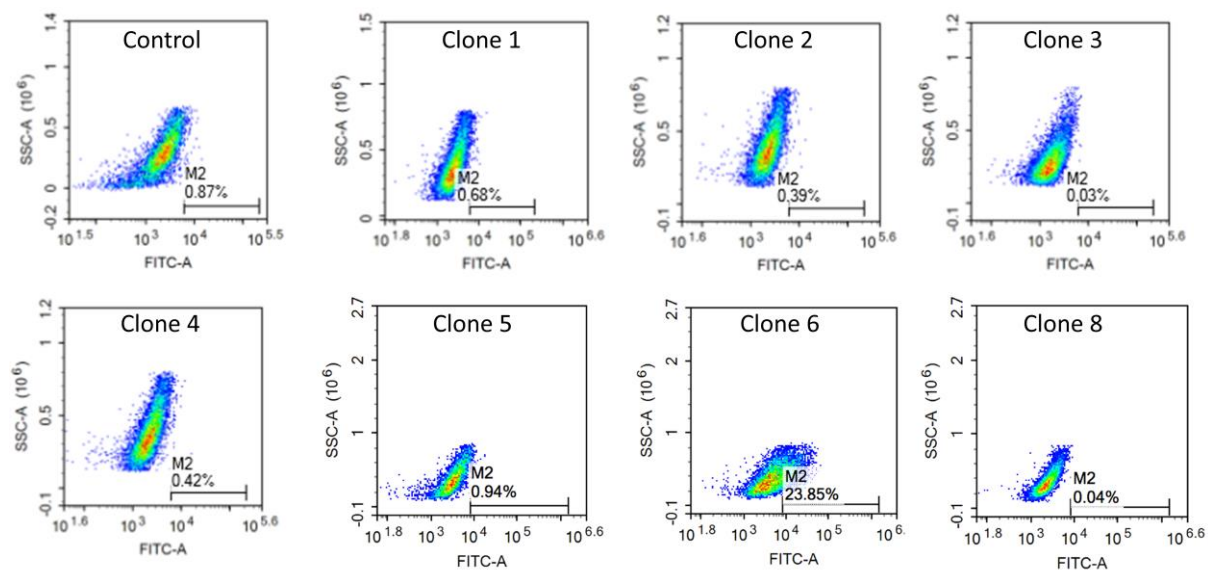


**Figure 4.23: Morphology of single clones during expansion. Representative brightfield images of expanded single clones. The cells retained a 2D monolayer morphology, confirming the successful expansion of clones. Healthy colonies with compact structures and tight borders were observed in a few single clones. Scale bar: 100 µm.**



**Figure 4.24: Morphology of single clones during expansion. Representative brightfield images of expanded single clones. In some of the clones, the structure of colonies was loose while interconnected. Scale bar: 100  $\mu$ m.**

Following expansion of the clones, flow cytometry was conducted for 7 clones to measure the GFP expression level. The level of GFP expression was lower in clones transduced with pHR and pHV vectors compare to original P106i population initially transduced with these vectors. Only in one clone of pHR, a low level of GFP expression (23.8%) was detected (Figure 4.25). The level of GFP expression in the transduced pool population prior to SCC was much higher. This result may suggest that the colonies with positive GFP expression have been lost during expansion or the level of GFP expression was reduced. Despite this result, the DNA/RNA of each clone was extracted. The samples were sent to GeneWerk (Germany) for validation of lentiviral insertions in single clones.



**Figure 4.25: Flow cytometry analysis of single clones transduced with pHR and pHV lentiviral vectors. Following expansion, three clones of pHV (top panel) and four clones of pHR (bottom panel) were analysed for GFP expression, using flow cytometry analysis. The result revealed no positive GFP expression in pHV and pHR clones except in one clone (clone 6 pHR) which positive GFP expression (23.85%) was detected.**

To further analyse the samples, DNA and RNA was extracted from expanded clones and was sent to GeneWerk for downstream analysis (Tables 4.7 & 4.8).

**Table 4.7: The DNA samples for single clones transduced with lentiviral vectors**

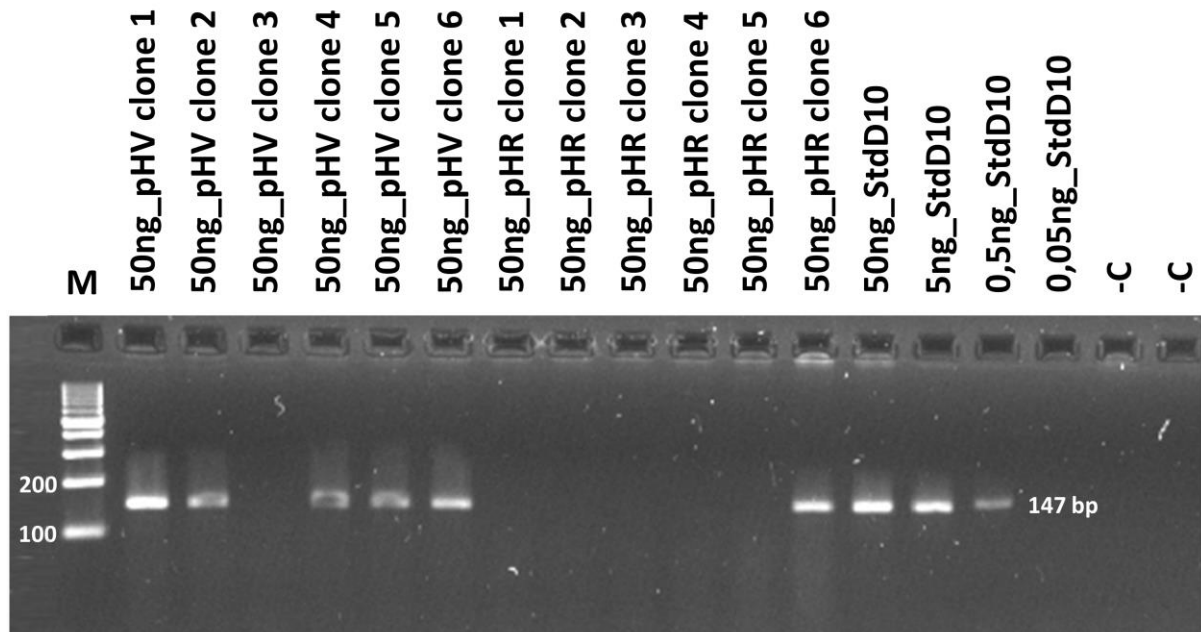
<b>P106i transduced samples</b>	<b>ng/ul</b>	<b>260/280</b>	<b>260/230</b>	<b>Quantity (µl)</b>
pHV/CLONE 1	237.3	2.02	2.04	120
pHV/CLONE 2	283.6	2.03	2.05	120
pHV/CLONE 3	336.5	2.04	2.04	120
pHV/CLONE 4	341.4	2.04	2.04	120
pHV/CLONE 5	324.0	2.05	2.04	120
pHV/CLONE 6	310.8	2.04	2.05	120
pHR/CLONE 1	184.4	2.02	2.02	280
pHR/CLONE 2	204.9	2.02	2.02	280
pHR/CLONE 3	234.7	2.01	2.02	280
pHR/CLONE 4	236.02	2.03	2.02	280
pHR/CLONE 5	284.0	2.02	2.02	280
pHR/CLONE 6	241.4	2.01	2.1	280

**Table 4.8: The RNA samples for single clones transduced with lentiviral vectors**

<b>P106i transduced samples</b>	<b>ng/ul</b>	<b>260/280</b>	<b>260/230</b>	<b>Quantity (µl)</b>
pHV/CLONE 1	1114.8	2.11	2.13	30
pHV/CLONE 2	1085.0	2.10	2.11	30
pHV/CLONE 3	1486.5	2.10	1.94	30
pHV/CLONE 4	785.8	2.09	2.08	30
pHV/CLONE 5	745.9	2.09	2.18	30
pHV/CLONE 6	1338.5	2.09	1.91	30
pHR/CLONE 1	816.0	2.09	2.17	30
pHR/CLONE 2	1668.7	2.10	2.06	30
pHR/CLONE 3	1862.4	2.11	2.26	30
pHR/CLONE 4	1631.2	2.10	1.94	30
pHR/CLONE 5	1328.7	2.09	1.96	30
pHR/CLONE 6	1550.4	2.09	2.18	30

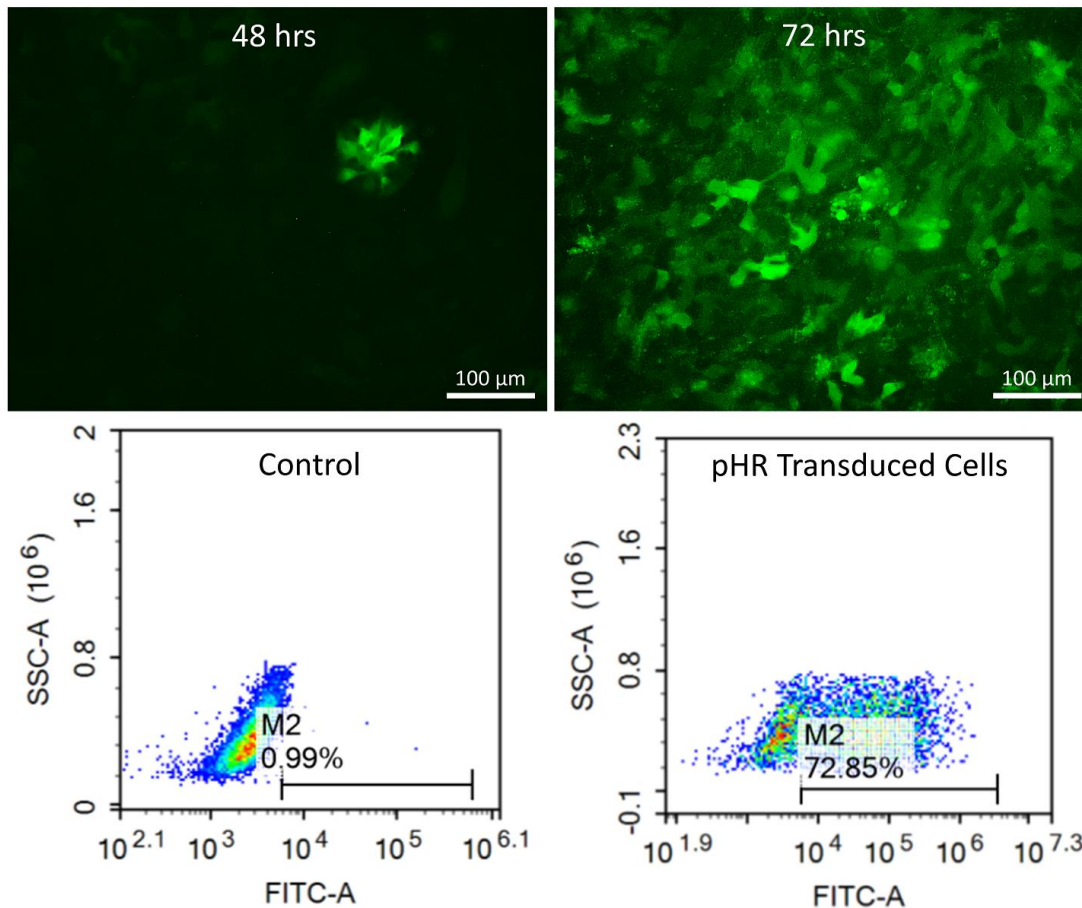
#### **4.3.16. Detection of pHR and pHV lentiviral insertions in P106i transduced clones**

Validation of pHR and pHV lentiviral insertions in P106i transduced clones was performed, using a standard PCR method (Figure 4.26). From the extracted DNA of transduced samples, PCR was performed to amplify a segment of LTR specific region, using specific primer sequences. Following PCR, LTR band was detected in 5 out of 6 pHV-transduced P106i clones. In contrast, for pHR clones no bands were detected in 5 out of 6 clones and only 1 clone confirmed to be positive for LTR insertion. The PCR result, further reinforced the low level of GFP expression in pHR clones, as shown in flow cytometry analysis (Figure 4.25).



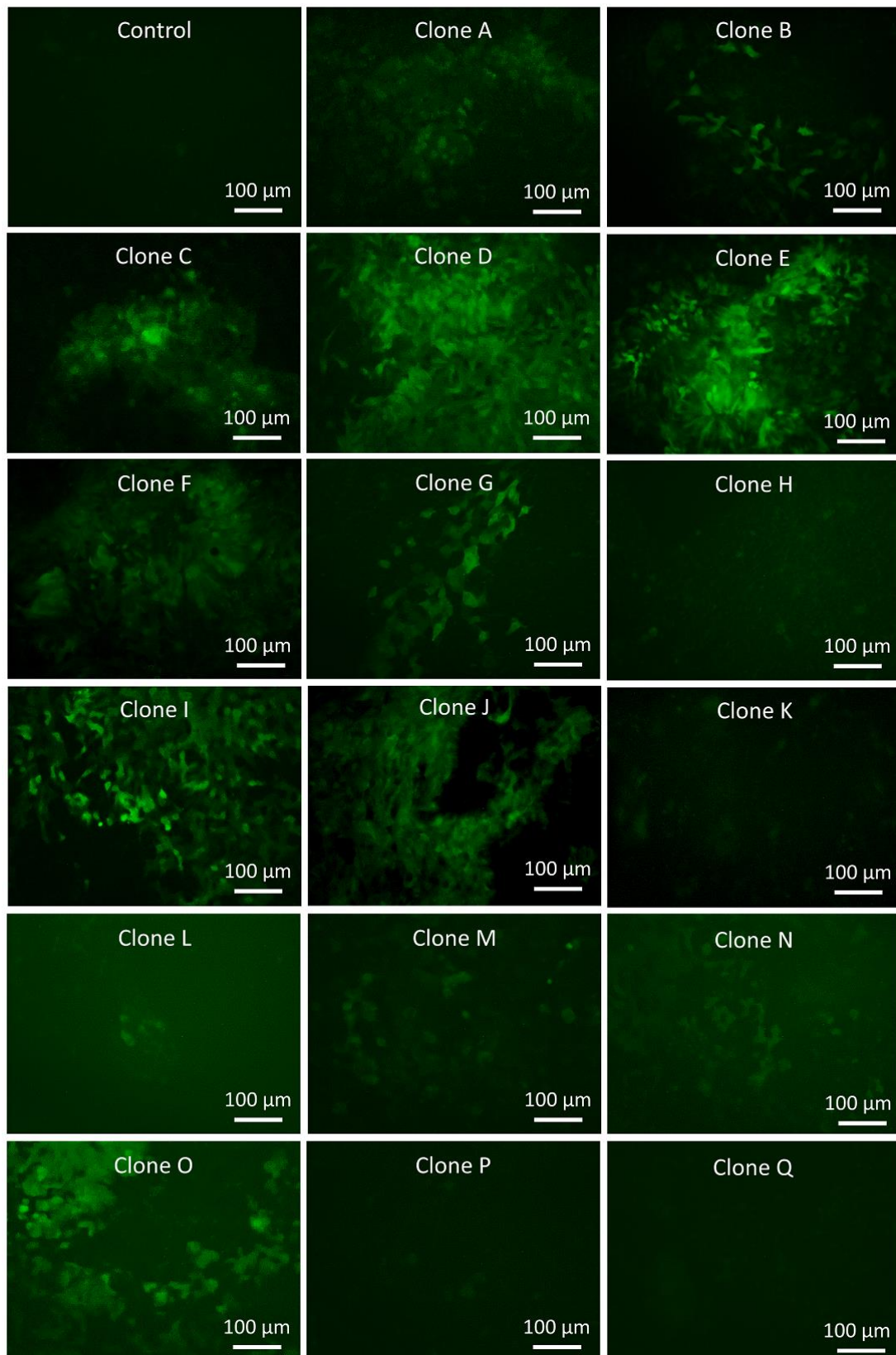
**Figure 4.26: Detection of lentiviral insertions in single clones transduced with pHR and pHV lentiviral vectors. To validate insertions, a segment of LTR specific region was amplified using PCR. The result revealed positive LTR bands in clones transduced with pHV lentiviral vector. However, no positive band was detected for pHV clone 3. Also, a positive band was detected only in one clone of pHR. The positive LTR band was detected at 147 bp. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V.**

Based on this result, SCC was repeated to obtain positive pHR clones. SCC was performed using transduced P106i cells with pHR lentiviral vector and the cells were analysed for positive GFP expression 72 hrs post transduction using flow cytometry analysis. The result has shown the positive GFP expression in P106i transduced with pHR (80.7%), confirming the transduction. (Figure 4.27).



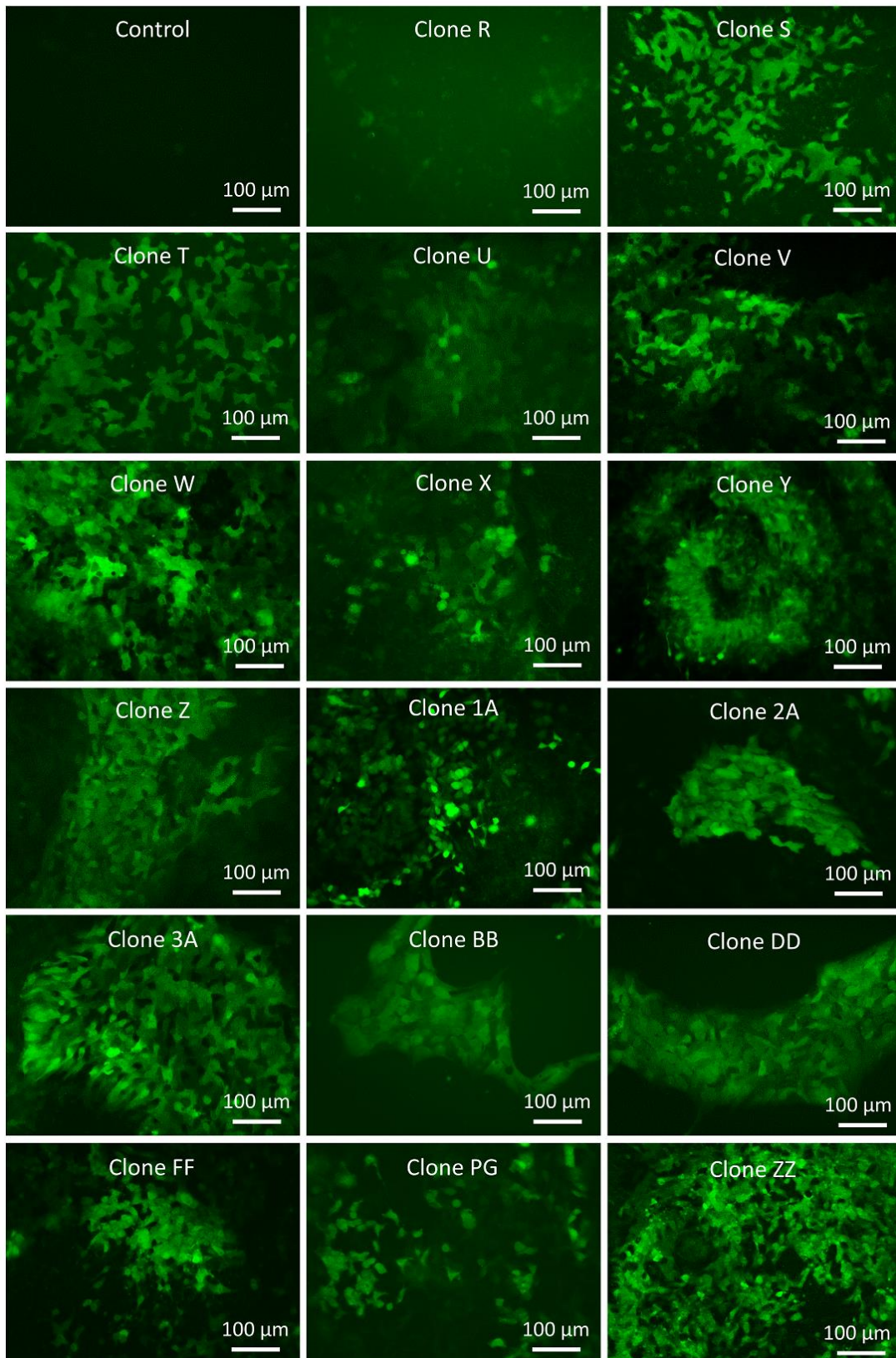
**Figure 4.27: Fluorescent microscopy and flow cytometry analysis of P106i cells transduced with pHR lentiviral vector. Representative fluorescent images showing P106i cells infected with GFP expressing pHR lentiviral vectors at MOI 20. Fluorescent microscopy images revealed positive GFP expression in P106i cells transduced with pHR lentiviral vector at 72 hours post transduction (Scale bar 100  $\mu\text{m}$ ). Further, the cells were analysed for positive GFP expression, using flow cytometry analysis. The result revealed positive GFP expression (72.85%) in P106i cells transduced with pHR lentiviral vector.**

Following successful transduction of P106i cells with lentiviral vectors, SCC was performed for the pHR transduced pool of cells, as previously described. Then 34 clones were picked 7 days post seeding and expanded. Unlike the previous set, efforts were made to select and expand the clones that exhibited positive GFP expression, confirmed by fluorescent microscopy. Different levels of GFP expression was detected throughout the whole batch of clones (Figures 4.28 & 4.29). Each of the clones were carefully dissociated and collected as cell pellets. The pellets were sent to GeneWerk for lentiviral insertion detection and downstream analysis.



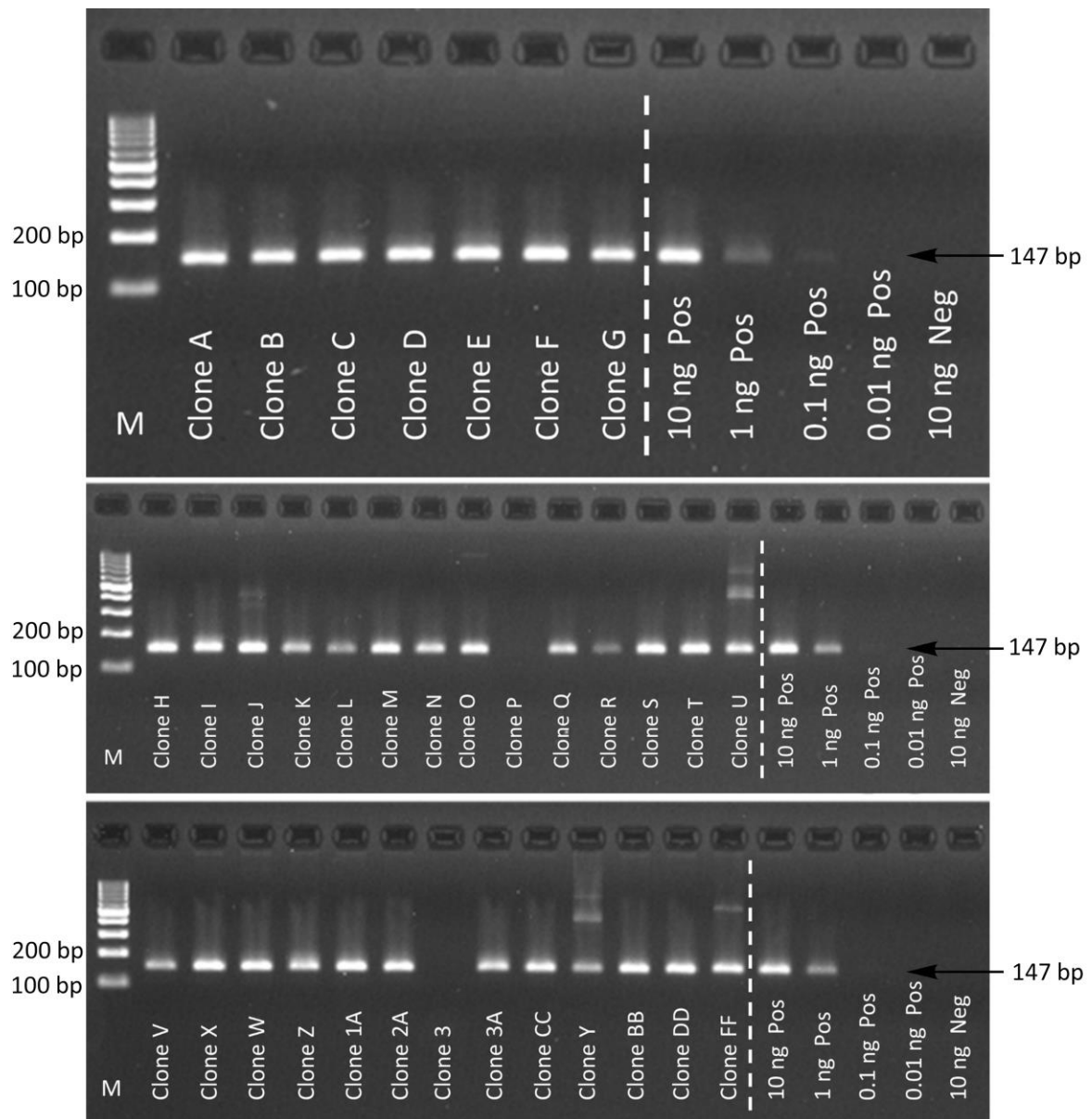
**Figure 4.28: Representative fluorescent images of single clones A-Q transduced with pHR lentiviral vector showing positive GFP expression. The level of GFP expression was lower in clones H, K, P and Q. Scale bar:100 μm.**





**Figure 4.29: Fluorescent microscopy imaging of expanded single clones R-ZZ transduced with pHR lentiviral vector. The level of GFP expression was lower in clones R and BB. Scale bar:100 μm.**

Validation of pHR lentiviral insertions in P106i transduced clones was performed, using standard PCR method. The result revealed positive LTR bands in 32 out of 34 pHR single cell clones. However, no positive band was detected in 2 out of 34 pHR clones. Despite low GFP expression in some clones, lentiviral insertion was confirmed in 32 clones (Figure 4.30).



**Figure 4.30: Detection of lentiviral insertions in single clones transduced with pHR lentiviral vectors. The result revealed positive LTR bands (147 bp) in all clones except two clones (P and 3) using 10 ng DNA for PCR. The ladder is GeneRuler (100 bp). The gel is 2% and ran for 35 minutes at 130V.**

## 4.4 Discussion

Existing *in vivo/ex vivo* animal genotoxicity models suffer from several limitations. Therefore, there is an urge to develop *in vitro* non-clinical tests to study the biosafety of the viral vectors for the patients. A robust humanised *in vitro* model is required to evaluate the risk of vector mediated genotoxicity (IM/oncogenesis) which supports the risk assessment set by regulatory bodies whilst limiting the use of animals.

A major disadvantage of conventional *in vivo/ex vivo* genotoxicity models is that animal physiology does not recapitulate the natural *in vivo* human cellular response. As mentioned earlier, *in vitro* 3D models are valuable as they can faithfully mimic natural human physiology. Further, existence of such models is also beneficial to animal welfare. To date, *ex vivo* “reconstituted” 3D liver (Wagner et al., 2015) and *in vivo* “humanised” liver models (Lisowski et al., 2014) have been developed to study viral gene vectors. Reconstituted 3D liver models are formed by engulfing hepatic cells into an animal extra cellular matrix (ECM) (Wagner et al., 2015). In this study, they constructed the 3D liver model, using human hepatocarcinoma cell line (HepG2) on a rat ECM. The AAV vector expressing Emerald green fluorescent protein (EmGFP) was injected into the portal vein of 3D liver model. Authors concluded the widespread distribution of the AAV vectors throughout the liver model followed by high internalisation of vector genomes per cell (Wagner et al., 2015). Another study developed a chimeric human-mouse liver model, called “humanised” liver model, to study novel AAV capsid variants. They applied transplanted human hepatocytes in immune deficient *Fah*<sup>-/-</sup> mice. Mature human hepatocytes were repopulated in the mouse liver with variable efficiency ranges (5%-40%) (Lisowski et al., 2014). Although it is claimed that recellularized biological ECM or chimeric human-mouse models are valuable to study viral vectors, these models are not ideal. The HepG2 cell line used in the recellularized model is a carcinoma cell line which has high proliferation with low ALB secretion. However, this cell line is cancerous and lacking important metabolic activities of mature hepatocytes (Kammerer and Kupper, 2018). Further, for production of recellularized biological ECM or chimeric human-mouse model, they relied on animals. In our study, P106i hiPSCs were used as an alternative to HepG2 cell line and primary hepatocytes. They were successfully differentiated into 3D heps which are xeno-free and robust. We demonstrated that the proposed model benefits

from mature hepatic functional activities and long-term culture stability which eliminates the need for conventional 2D monolayer cultures.

For this study, recombinant AAV/GFP serotype 2 vectors were kindly provided by our collaborator in Australia. These vectors were used to transduce P106i cells and our proposed *in vitro* 3D hepatospheres model. Further, third generation VSVG pseudotyped HIV-1 based lentiviral vectors were generated in house at high titres by transient transfection of adherent embryonic kidney (HEK) 293T cells. The safe vector (pHR) and the unsafe vector (pHV) is the full LTR equivalent to the pHR vector. The self-inactivating (SIN) pHR vector has deletions in 3'LTR of the viral genome. This will disrupt the promoter/enhancer activity of the LTR further improving the biosafety. This modification did not affect the vector titer or transgene expression. In pHR and pHV lentiviral vectors the internal promoter SFFV was incorporated to drive the GFP expression. The WPRE and cPPT have been applied to enhance the overall vector performance.

To enhance the transduction of lentiviral vectors in P106i cells polybrene reagent was added. The effect of ROCK inhibitor on viral vectors transduction in PSCs has been reported (Mitsui et al., 2009). Hence, for rAAV/GFP and lentiviral vectors transduction of P106i, medium supplemented with ROCK inhibitor to support cells for the first 24 hours. It is reported that the transduction efficiency of AAV vectors in PSCs is highly variable and that transduction efficiency of natural AAV serotypes are quite incompetent in gene delivery to human PSCs (Brown et al., 2017). From our results, it can be concluded that the level of transduction in P106i cells was highly promoter dependent and the cells transduced with rAAV/CB7/GFP vector exhibited high level of GFP expression. Further, P106i samples transduced with pHR lentiviral vector have shown high level of GFP expression. Considering the culture condition and sensitivity of P106i cells, these cells can be transduced with rAAV/GFP serotype 2 and lentiviral vectors.

Previously in our lab viral vectors were used to transduce 2D mature hepatocytes which was not very successful (Data unavailable). One can speculate that due to the limited life span of 2D hepatocytes and the transient phenotype, the rate of transduction varies. Transduction efficiency in conventional 2D culture systems is variable and does not support long-term screening. For optimum genotoxicity assays,

long-term culture, stable phenotype and efficient transduction is of utmost importance. Therefore, *in vitro* 3D cell models can be used as a great alternative to fit the purpose. The 3D dimensional architecture with abundant ECM and more cell-cell and cell-ECM contacts compared to conventional 2D culture systems can greatly enhance transduction efficiency. To further improve the transduction of 3D heps, plates were modified for the first 24 hours. In this study, lentiviral and rAAV/GFP serotype 2 vectors were used to transduce 3D heps on day 30 of differentiation. The level of GFP expression was also promoter dependent. The 3D heps were successfully transduced with lentiviral and rAAV/GFP vectors. Fluorescent microscopy images revealed a strong level of GFP expression in 3D heps transduced with rAAV/CB7/GFP and pHR lentiviral vector. However, a low level of GFP expression was detected in 3D heps infected with rAAV/ApoE/GFP and pHV lentiviral vector. The efficiency of transduction was confirmed by PCR and VCN quantification.

The PSCs and iPSC-derived cell models are great alternative to bridge the remaining gaps between animal models and humans. Human cell models such as iPSCs and differentiated cells derived from iPSCs can be used as a viable alternative for testing, efficacy and genotoxicity screening. The foetal mouse model has been explored as a sensitive genotoxicity model that can predict lentiviral-associated mutagenesis resulting in liver oncogenesis. However, this model targets immature cell type with a larger readout. Therefore, human liver cell models such as differentiated HLCs resemble the *in vivo* situation more closely. In a study, *in vitro* human liver models were compared to *in vivo* human liver, using RNA sequencing. It was reported that 3D iPSC-derived liver models exhibited a high degree of similarity with *in vivo* liver (Gupta et al., 2021). Therefore, it is a great tool to transduce iPSCs and iPSC-derived HLCs with viral vectors for genotoxicity screening. This also reduce the number of animals required for preliminary screening and eliminates the need for long-term studies with animals. The 3D *in vitro* models support targeting specific cells (differentiated cell types), long-term culture and longer incubation times.

Another advantage of transducing at two stages is targeting different gene expression profiles. Transduction of PSCs with viral vectors can target key stem cells markers and the effect of viral vectors at early stages. When the cells are differentiated, there will be distinct tissue-specific markers and transcriptions factors which viral vectors can target. In addition, due to the differences between pluripotent and differentiated

states, downstream analysis such as IS readout would also be different. In terms of testing vector safety, targeting two stages will help researchers to target wide array of genes and risk of IM can be evaluated more effectively.

For genotoxicity screening, tumour prone mice models are often used. Therefore, IS mutagenesis might be influenced by the effect of viral vectors and animal genetic cancer predisposition. Further, certain *in vitro* assays such as *in vitro* immortalisation assay (IVIM) are not sensitive enough and suffers from considerable inter-assay variation (Modlich et al., 2009). Using PSCs and stem cell-derived differentiated cells eliminates the carcinogenic background effect and variability to a great extent.

Despite significant advances for the development of *in vitro* cell-based platforms, challenges are remained. The conventional 2D monolayer culture systems suffer from transient phenotype, short-term culture, and variable transduction efficiency. These qualities are not desirable for long-term genotoxicity screening or very long incubation times. The stem cell-derived 3D culture platforms have been established to improve the *in vitro* systems. However, it is extremely difficult to recapitulate interorgan connections and tissue-specific microenvironments. As discussed earlier, one of the problems in 3D technology is the variability between 3D system platforms. For genotoxicity screening, existing variability between protocols can affect downstream analysis. Sample variations promote different IS targeting. This is particularly of importance if patient-derived iPSCs are used to evaluate vector safety. Despite the great potential that the *in vitro* models hold for genotoxicity screening, it is still difficult to fully recapitulate *in vivo* events following viral vectors administration. In addition, for genotoxicity screening, clonal outgrowth studies are performed. Clonal outgrowth transformation can be expected for *in vitro* studies. However, our data doesn't prove the adverse effects of outgrowth. In addition, further analysis can be performed including ATP production, cell viability and oxidative stress assays to study the effect of viral transduction on hiPSCs and 3D heps following transduction.

Viral insertions in P106i cells and 3D heps were confirmed, using standard PCR. Further, VCN per cell was measured using a standardised quantitative PCR (q-PCR) in collaboration with GeneWerk (Germany). There was an overall reduction in clonality and VCN number in P106i transduced with both vectors at day 30 time point. However, 3D heps transduced with pHR and pHV lentiviral vectors had lower VCNs compared

to the hiPSC stage, confirming the low transduction of lentiviral vectors in non-dividing cells with low copy numbers.

In conclusion, P106i cells infected using lentiviral and rAAV viral vectors. The *in vitro* 3D human hepatosphere system is a xeno-free alternative to conventional *in vivo/ex vivo* animal genotoxicity models. This model has shown a high rate of transduction efficiency for lentiviral and rAAV/GFP serotype 2 vectors which can be extended to other viral vectors. The phenotypic stability and lack of proliferation are valuable characteristics of *in vitro* 3D hepatosphere model which allows a long-term assessment for vector mediated genotoxicity and risk of IM.

## Chapter 5

### 5.1. Introduction

#### 5.1.1. Gene Transfer Technology in Gene Therapy

Gene transfer technology is a powerful method to introduce recombinant genetic elements into human cells. This method holds a considerable therapeutic potential for the treatment of wide range of pathological conditions, such as genetic disorders, neurological complications, diabetes, cancer and infectious diseases (Edelstein et al., 2004, Romano, 2003, Edelstein et al., 2007). Cancer and infectious diseases gene therapy may require only a transient expression of the therapeutic transgene in order to destroy targeted neoplastic tissues or infectious agent (Romano et al., 2000, Romano, 2003). Conversely, inherited genetic disorders or cardiovascular diseases may require long-term transgene expression. With successful introduction of functional copies of the corrected genes, healthy disease phenotypes can be restored. Gene transfer systems are able to integrate into the target cell's chromosomal DNA effectively and allow for a more stable and long-term transgene expression (Romano et al., 2000). To date, several types of gene transfer systems were developed. The most common gene transfer technology types were derived from viruses such as retroviruses (RV and LV), adeno-associated viruses (AAV) and other non-viral vector systems (Rosenberg et al., 1990, Blaese et al., 1995, Muul et al., 2003, Romano et al., 2003, Romano, 2005a, Romano, 2005b, Romano, 2006, Romano, 2007a). AD and majority of non-viral derived vectors are considered as episomal gene transfer systems (Romano, 2006, Romano, 2007a). However, vectors derived from RV, LV, AAV, Sleeping Beauty (SB) DNA transposon systems hold integration properties. SB is a non-viral based gene transfer system and for delivery of therapeutic transgene do require a plasmid DNA. Following transfection, it is permanently integrated within the target cell genome (Romano, 2006).

Advanced gene transfer methods were also employed in the field of stem cells. These methods were used to genetically manipulate human and mouse embryonic (ES) cells, different types of adult stem cells, and, more recently, for the generation of induced pluripotent stem cells (iPSCs), which is derived from human or animal somatic cells. Over the last decade, gene transfer technology had a great impact on stem cell research and contributed to the progress of regenerative medicine (Romano, 2008b,



Romano, 2008a, Narsinh and Wu, 2010, Wu and Dunbar, 2011). The co-interaction between stem cell research and gene therapy has led to major advances in the treatment of genetic haematological disorders such as severe combined immunodeficiency (SCID-X1) (Edelstein et al., 2004, Edelstein et al., 2007, Romano et al., 2000, Romano, 2003).

One of the requirements for a permanent transgene expression, is the integration of the shuttle vector within the target cell's chromosomal DNA. However, integrative gene transfer systems have few pitfalls such as risk of IM. This may eventually lead to development of malignancies (Romano et al., 2000, Romano, 2003, Romano, 2005a, Romano, 2005b, Romano, 2006, Romano, 2007a). The integration of shuttle vector may interfere with the organisation of the chromosomes. This will induce a series of events which may ultimately lead to cellular transformation and change of cell phenotype (Romano et al., 2000, Romano, 2003). Unfortunately, following administration of retroviral-mediated gene transfer five patients who underwent French and England clinical trials developed leukaemia (Hacein-Bey-Abina et al., 2003, Baum, 2007, Cavazzana-Calvo and Fischer, 2007). The mechanisms underlying IM-induced carcinogenesis in SCID-X1 trials will be discussed in greater detail in the next section. Gene therapy clinical trials, definition of IM and important mechanisms of IM were fully covered in Chapter 1, Section 1.1.13.

### **5.1.2. Key mechanisms of IM-induced Oncogenesis in the SCID-X1 trials**

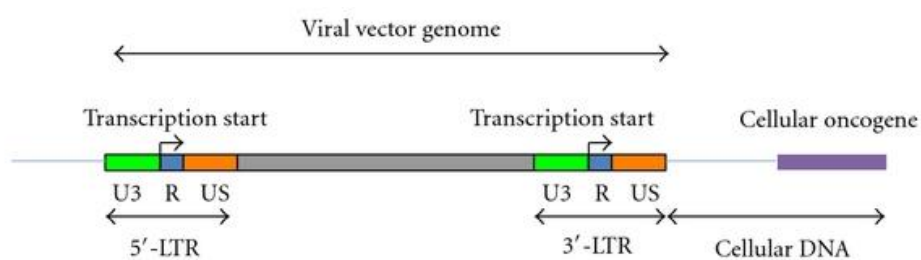
For *retroviridae*-derived gene transfer methods (including retroviral and lentiviral vectors) two important mechanisms of vector-induced IM were proposed. The genome structure of *retroviridae* viral vector has two long terminal repeats (LTRs) which are located at the 5' and 3' end of their genome. Each of the LTR segment contains an enhancer and promoter elements. As a result of integration into the target cell chromosomal DNA, the 3' LTR might induce the expression of an endogenous oncogene. This is particularly important if an endogenous oncogene is located in proximity of the insertion site. Additionally, the expression of cellular tumour suppressor genes might be inhibited. This will occur if viral vector inserted within an exon of tumour suppressor gene. The first mechanism involves with the induction of expression of cellular oncogenes and/or inhibition of tumour suppressor genes

(Figures 5.1-5.2). The second mechanism for IM which has a longer range of action involves a close interaction between the enhancer element of a retroviral LTR segment and a cellular promoter. This will promote expression of an endogenous oncogene (Romano et al., 2000, Romano, 2003, Romano et al., 2003, Romano, 2005a) (Figure 5.3). The second mechanism had a great role in onset of haematological malignancies detected in five patients of the SCID-X1 clinical trials (Hacein-Bey-Abina et al., 2003, Baum, 2007, Cavazzana-Calvo and Fischer, 2007). Downstream analysis of the tumours obtained from the first two patients revealed and overexpression of LIM Domain Only 2 (LMO2) oncogene. A detailed analysis was conducted on malignant T cells obtained from the first two patients who developed leukaemia in the French clinical trial. In one patient, the site of viral vector integration was detected to be at 3kb upstream of the LMO2 transcription start site. Other patient, viral vector insertion was located in antisense orientation to the LMO2 promoter and 5kb downstream of the LMO2 transcription start site. Further, no replication-competent retrovirus was reported in the study, and it was concluded that overexpression of recombinant gamma C subunit was responsible for the onset of the haematological malignancy (McCormack and Rabbitts, 2004, Baum, 2007). In total five patients developed leukaemia as a result vector-mediated IM event. Few animal studies were also reported that overexpression of recombinant gamma C subunit was induced tumours in animal models (Dave et al., 2004, Woods et al., 2006). Nonetheless, few studies argued that overexpressed recombinant gamma C subunit had no effect on oncogenesis (Pike-Overzet et al., 2007, Modlich et al., 2008). These findings suggest that animal models do not faithfully recapitulate human diseases particularly diseases concerning oncogenesis.

*LMO2* overexpression, *per se*, was not the only factor to promote oncogenesis. This was a multiplex mechanism which required different genetic alterations, epigenetic, induction of cellular oncogenes and/or silencing of tumour suppressor genes (Fucito et al., 2008). A follow up study analysed the phenotype of tumours of the other two leukemic patients. These patients were also participated in the French SCID-X1 clinical trials. *LMO2* overexpression was also detected in these two cases. In addition, a second retroviral vector insertion site was also detected in the proximity of the proto-oncogene *CCDN2* in tumour cells of one patient. However, for other patient the retroviral vector insertion site in tumour cells was close to the proto-oncogene *BMI1*.

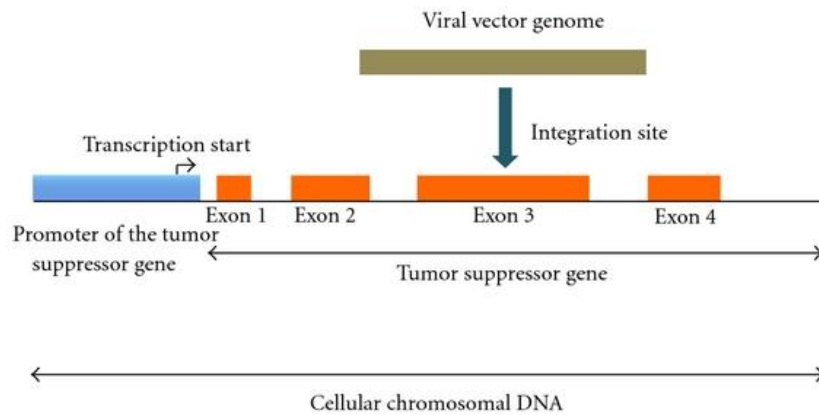
It was concluded that both leukemic patients exhibited several genetic alterations such as deletion of tumour suppressor gene cyclin-dependent kinase 2A (*CDKN2A*) (Hacein-Bey-Abina et al., 2008). A similar study was conducted for the analysis of tumour cells of the leukemic patient who was participated in the British SCID-X1 gene therapy trials. Similarly, overexpression of LIM Domain Only 2 (*LMO2*) was also detected in this case along with several genetic alterations such as deletion of tumour suppressor gene *CDKN2A* and gain of functions (Chinen and Puck, 2004, Howe et al., 2008).

For detailed analysis of the retroviral-derived vector insertion site, a protocol based on linear amplification-mediated PCR (LAM-PCR) was developed. This method was employed to determine the insertion site profile of retroviral-derived vectors within the genome of CD34<sup>+</sup> hematopoietic stem cells (HSC) of two patients (Metais and Dunbar, 2008, Gabriel et al., 2009). These patients had Wiskott-Aldrich syndrome (WAS) syndrome and underwent gene therapy treatment. Following 892- and 891-days post gene therapy treatment, LAM-PCR analysis was conducted in patients 1 and 2, respectively. Results exhibited 5,709 and 9,538 unique insertion retroviral vector insertion sites in patients 1 and 2, respectively. The most common insertion sites (CIS) were as follows: *LMO2*, *MDS-EVI1*, *CCDN2* and *PRDM16*. It was concluded that retroviral vector insertions were either close or within the *LMO2* and *CCDN2* gene loci. These insertion sites were frequent in lymphoid cells. However, retroviral vector insertions associated with *PRDM16* and *MDS1-EVI1* gene loci were more pronounced in myeloid cells. Following gene therapy treatment, polyclonal blood cell populations were detected in both patients. Nonetheless, there was no sign of haematological malignancies (Deichmann et al., 2011). Therefore, characterisation of retroviral integration sites might reveal important clues on how viral vectors integrate and also enable us to evaluate potential risk of oncogenesis.

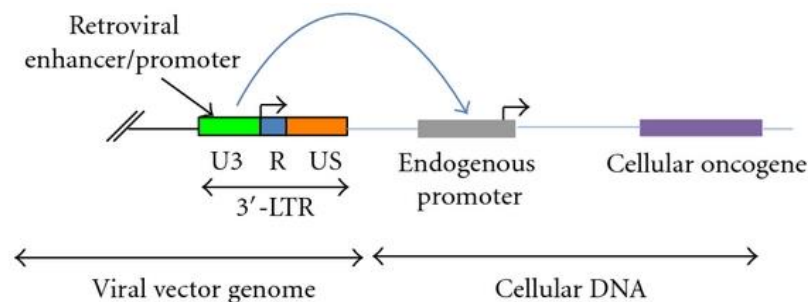


**Figure 5.1: Schematic representation of first mechanism of *retroviridae*-based gene transfer vectors. This is first mechanism for cellular oncogene expression. The**

retroviral 3' long terminal repeat (LTR) promote transcriptional activation promoter which may lead to cellular oncogene expression. The bent arrows exhibit the transcription start site (TSS) of the 5'- and 3'-LTR of the genome (Romano, 2012).



**Figure 5.2: Schematic representation of first mechanism of *retroviridae*-based gene transfer vector induced insertional mutagenesis. This is the mechanism for tumour suppressor gene silencing. This mechanism is triggered by the retroviral vector integration within an exon of the tumour suppressor gene (Romano, 2012).**



**Figure 5.3: Schematic representation of the second mechanism of *retroviridae*-based gene transfer vectors induced insertional mutagenesis. This is the mechanism for expression of cellular oncogene. It is mediated by close interaction and prolonged effect of the retroviral enhancer and the endogenous promoter of the cellular oncogene. The blue block arrow depicts the close interaction between the retroviral enhancer and the endogenous promoter of the cellular oncogene (Romano, 2012).**

### 5.1.3. Preclinical Studies (*in vivo* & *in vitro*) for IM

The use of *retroviridae*-derived vector systems in gene therapy clinical trials can cause a serious safety concern. Concerns were raised due to the onset of haematological malignancies that can be mediated by IM events (Edelstein et al., 2004, Edelstein et al., 2007, Romano, 2003). As already mentioned in Chapter 1, Section 1.1.16.1, IM-induced malignancies were detected in five leukemic patients of the SCID-X1 gene

therapy clinical trials. IM events were also demonstrated both in animal models (Modlich et al., 2008, Cavazzana-Calvo et al., 2000, Modlich et al., 2005, Kustikova et al., 2005, Maetzig et al., 2011a) and also *in vitro* human cell culture models (Maetzig et al., 2011a, Mitchell et al., 2004). Most of these studies were focused on the integrating characteristics of *retroviridae*-based gene transfer systems based on murine leukaemia virus (MLV), avian sarcoma-leukosis virus (ASLV), and human immunodeficiency virus type 1 (HIV-1). Results from the aforementioned *retroviridae*-based gene transfer systems suggested that the integration within the human genome was not completely a random process. In fact, viral vectors were chosen preferential sites of integration in the proximity of chromosomal regions that contained transcriptionally active genes. For instance, MLV-based gene transfer systems exhibited a great preference for integration points near TSS site (Mitchell et al., 2004, Lewinski et al., 2006, Ciuffi et al., 2006). However, HIV-1 derived lentiviral vectors prefer specific regions which are rich in active gene and intercalated with genomic regions containing methylated CpG islands. These regions are not permissive for gene expression (Mitchell et al., 2004).

It is important to know all integrating gene transfer systems may pose a serious risk of IM in target cells. Interestingly, results from animal models and *in vitro* cell culture models demonstrated that IM might also be induced by AAV-derived viral vectors and also integrating non-viral gene transfer systems, such as SB DNA transposon-based vectors (Romano, 2005a, Romano, 2007b). Following *in vivo* administration of AAV-derived vectors, IM-induced angiosarcomas or hepatocellular carcinomas were detected in mice. Many studies reported that retroviral and AAV-derived viral vectors have remarkable predisposition for integration in common fragile sites. Fragile sites composed of specific regions that is linked with chromosomal breakpoints which has a relevant role in the induction of early stages of malignant cellular transformation (Romano, 2005a, Romano, 2007b). Based on aforementioned points, a detailed analysis of so-called hotspots for viral based gene transfer system insertion sites are crucial. This analysis enables scientists to determine the mechanism of IM-induced onset of malignancies in patients.

#### **5.1.4. Strategies to minimise the incidence of IM**

The first approach to minimise the incidence of IM is to produce safer integrating gene delivery systems. The aim is to develop a gene transfer shuttle vector that do not

interfere with the physiological genomic organisation of transduced cells. Hence, it is crucial to prohibit interactions between enhancer and/or promoters of the gene transfer vector system and the target cell's genome. The engineering of self-inactivating (SIN) *retroviridae*-derived vector systems may reduce the incidence of IM mediated events in transduced cells. SIN *retroviridae*-derived vector systems are generated by a deletion of the U3 region in the 3' LTR segment. In this strategy, the 5' LTR initiates the transcription of a viral mRNA. Ultimately a proviral DNA is generated without transcriptional activity at both LTRs. Nonetheless, SIN *retroviridae*-derived vector systems require an internal promoter to express the transgene.

Generally, SIN *retroviridae*-derived vector systems are considered safer. Particularly when they are compared with their original counterparts such as gamma-retroviral or gamma-lentiviral derived vectors. The probability of interactions between the SIN vectors and the host genome is minimised by deletion of retroviral enhancers and promoter regions in the LTRs. Notably, the non-transcribing 3' LTR is not capable of inducing the expression of cellular oncogenes if it is in close proximity of the insertion site. In addition, the long-range interactions of cellular promoters that modulate the expression of endogenous oncogenes is prohibited. This is due to the removal of *retroviridae* enhancer regions. Nevertheless, presence of an enhancer within the internal promoter of a SIN vector can be problematic. This can trigger cellular promoters of oncogenic factors and may increase the risk of IM (Romano, 2003, Romano, 2005a, Romano, 2005b). Some studies reported development of SIN-lentiviral vectors with enhancer-less internal promoters to improve their safety profile. One major drawback with SIN *retroviridae*-derived vector systems is elevation of polyadenylation (polyA) signal read-through in the 3' LTR region. Activation of muted cellular oncogenes in transduced cells can be triggered by leaky transcriptional termination signals. New strategies such as incorporation of seven SV-40-derived upstream polyA enhancer elements might improve transcription termination efficiency in these vectors (Schambach et al., 2007).

The second approach is to engineer episomal vector systems. These systems facilitate a more stable and long-lasting transgene expression in transduced cells (Romano, 2008a). The third approach is to employ engineered zinc-finger proteins. These proteins enable scientists to site-specific correct defected genes or insert genetic factors into specific *loci* of the genome of transduced cells. However, resolving

IM issue can have a huge impact on successful application of gene therapy treatments (Romano, 2012). In the next sections, methods to analyse insertion sites and assess safety of viral vectors will be discussed.

### **5.1.5. Methods for tracking virus Integration Sites**

Different classes of integrating viral vectors have been employed in research studies and clinical gene therapy trials. The aim of these vectors was to establish stable expression of corrected genes in transduced HSCs and their progeny. However, it is known that an integrated vector provirus can affect the expression of neighbouring genes. Integration can occur up to 50-100 kb upstream or downstream of neighbouring genes and promote phenotypic changes on transduced cells. As mentioned in Chapter 1, Section 11.5.1, integrating viral vectors utilised in clinical gene therapy have exhibited clinical improvement and also insertional activation of proto-oncogenes. This has led to malignant or uncontrolled clonal proliferation in several patients with SCID-X1, WAS and X-CGD (Ott et al., 2006, Hacein-Bey-Abina et al., 2003, Howe et al., 2008). Therefore, it is extremely important to be able to identify and track proviral insertion sites following treatment. It is also crucial to detect IS in the setting of clonal dominance. This information is vital for better understanding of vector-host cell interactions and develop strategies to avoid genotoxicity. Prior to the full sequencing of the human genome, it was not feasible to identify integration sites. This is because integrating retroviruses do not integrate into particular motif or small set of genes. However, since 2001 it is possible to perform detailed IS analysis. In the past decades several methods for tracking IS have been developed. The general principle of these methods is to identify known sequences in the proviral integrated into the unknown region of genomic DNA. Then the identified junction fragment is isolated, amplified and sequenced (Aiuti et al., 2007). Methods for tracking IS are including Ligation-Mediated PCR (LM-PCR), Linear Amplification-Mediated polymerase chain reaction (LAM-PCR), Non-Restrictive LAM PCR (nrLAM-PCR) and Solid-phase ligation-mediated PCR (EPTS/LM-PCR). For the purpose of this thesis, LAM-PCR and nrLAM-PCR and EPTS/LM-PCR methods will be explained in greater details.

### 5.1.6. LM-PCR & LAM-PCR methods

Ligation-mediated PCR (LM-PCR) was first described in 1989 (Mueller and Wold, 1989). In this method genomic DNA was digested with restriction enzymes. Following this step, a linker was ligated to the genomic end of the cleaved DNA. For amplification step one primer was attached to the linker and another primer to the vector LTR resulting in amplification of vector-genome junction. Sequencing step was performed at the end. This method was further refined in 2008 and was used for analysing retroviral IS in HSCs (Kustikova et al., 2008). This method has certain drawbacks such as restriction enzyme bias. Also, the linker ligation step was quite inefficient. However, this method proved to be useful for identification of IS in clonal or oligoclonal samples containing vector copy numbers of 0.10-0.20 or even higher (Kustikova et al., 2005). Most clinical samples had much lower copy numbers and therefore this method was not sensitive enough to retrieve IS effectively.

One decade ago, LM-PCR was optimised to enhance sensitivity and efficiency and so-called LAM-PCR. By inclusion of a linear amplification step, efficiency and sensitivity was greatly enhanced. The linear amplification step was performed, using a biotinylated primer. This primer was annealed to the LTR in the orientation out toward the vector-genome junction. Further, those fragments with no vector containing DNA were filtered using streptavidin-coated magnetic beads. The obtained purified fragments were converted to double stranded DNA (dsDNA) by random hexamer priming. These fragments were digested with a restriction enzyme(s). Then a double stranded linker was ligated to the ends. Exponential nested PCR followed by direct sequencing was performed as final steps. This method is technically challenging and also it is labour intensive (Schmidt et al., 2002, Schmidt et al., 2009).

LAM-PCR enable scientists to retrieve and identify large numbers of IS in highly polyclonal samples. This technique has been reported in numerous publications which IS data from gene therapy trials and preclinical animal studies was studied (Ott et al., 2006, Howe et al., 2008). This technique is very sensitive and efficient even if level of vector copy number quite low. Further, small amounts of DNA (100-500ng) is sufficient to conduct IS analysis of sorted cell populations (Harkey et al., 2007, Gabriel et al., 2009). Although this technique is sensitive and efficient it also has limitations. It is difficult to do full IS analysis particularly if it is based on restriction enzyme cutting. The reason for this is some IS can either be too close or too distant from any specific



restriction enzyme site. As a result, some fragments are too small or too long to be amplified. Hence, the whole analysis is restricted to a subset of clones (Harkey et al., 2007, Gabriel et al., 2009).

### **5.1.7. Non-restrictive linear-amplification-mediated PCR (nrLAM-PCR)**

To overcome limitations related to restriction enzymes cutting in LM-PCR and LAM-PCR, an alternative method was explored. This method was called non-restrictive linear-amplification-mediated PCR (nrLAM-PCR). It is very similar to LAM-PCR and a biotinylated LTR-specific primer is used for the linear amplification step. However, there is no restriction enzyme cutting following bead enrichment step. Instead, the single-stranded DNA fragments containing the provirus-genome junctions are attached to a linker oligonucleotide with no restriction enzyme cutting. Then an exponential nested PCR is performed followed by high-throughput pyrosequencing. The single-stranded attachment step in nrLAM-PCR is less efficient compared to LAM-PCR (Schmidt et al., 2002, Schmidt et al., 2007). Hence, the sensitivity of nrLAM-PCR is much lower and larger DNA samples are required.

### **5.1.8. Solid-phase ligation-mediated PCR (EPTS/LM-PCR) technique**

A newer method was developed which is more reliable than existing predecessors. This method also removes nontarget DNA and avoids restriction bias commonly seen with LM-PCR and LAM-PCR methods. The magnetic extension primer tag selection (EPTS) is used prior to solid-phase ligation-mediated PCR (EPTS/LM-PCR). The advantage of this technique is that multiple proviral LTR-flanking sequences of retroviral and lentiviral vectors can be identified even if it is only present in 1 per 100 or 1000 cells containing the provirus (Schmidt et al., 2001). Furthermore, this method is fast, sensitive and inexpensive using minimum number of samples.

In chapter 4, [pHR'SIN-cPPT-GFP-WPRE(SEW)] (pHR) and [pHV-cPPT-SEW] (pHV) which is the full long terminal repeat (LTR) equivalent to the pHR lentiviral vector and recombinant AAV/GFP serotype 2 vectors were used to transduce P106i cell line and *in vitro* 3D heps. It was shown that the 3D hepatic model can be transduced using lentiviral vectors (pHR and pHV) and adeno-associated serotype 2 viral vectors [(rAAV/CB7/GFP and rAAV/ApoE/GFP)]. In this chapter, to evaluate the genotoxicity

risk of pHR and pHV lentiviral vectors, IS analysis, cancer gene identification and gene enrichment analysis were performed. Further, to identify viral insertions and their effect on gene expression in the proximity of vector IS, single cell cloning (SCC) was conducted. From collected samples, DNA/RNA was extracted and provided to our collaborator GeneWerk (Germany) for EPTS/LM-PCR analysis. Some of the processed bioinformatic data was provided by GeneWerk as part of this collaboration. Additional bioinformatic analysis was also performed in-house by Saqlain Suleman carried out a PhD in the Themis laboratory using EPTS/LM-PCR raw data provided by GeneWerk. In this chapter, the bioinformatics analysis of the P106i cell line, 3D heps and single cell clones transduced with pHR and pHV lentiviral vectors are shown.

## **5.2. Objectives**

The objectives of this chapter are as follows:

- To compare the frequency of integration sites for pHR and pHV lentiviral vectors at two time points in transduced P106i cells and P106i derived 3D heps.
- To identify and compare the common insertion site (CIS) in transduced cells
- To assign gene ontology to each gene IS and its association with cancer in transduced P106i cells and 3D heps
- To identify IS in iPSC clones and relative gene expression in proximity to IS

## **5.3. Results**

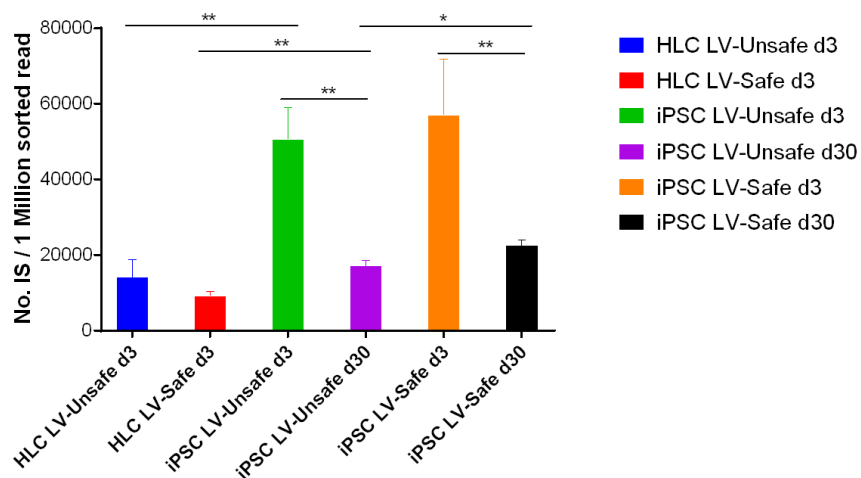
To identify potential cancer gene candidates, genomic regions were identified using EPTS/LM-PCR with primers specific to HIV-1 LTR viral vectors. P106i cells lines were transduced with pHR and pHV lentiviral vectors and propagated for 30 days. DNA samples were collected at days 3 and 30 timepoints following post transduction. The frequency of insertions by pHV and pHR lentiviral vectors were then compared between two timepoints. The frequency of insertions was also analysed for 3D heps which were transduced with pHR and pHV lentiviral vectors.

### **5.3.1. Integration site analysis (IS) on bulk P106i iPSC line and P106i-derived 3D heps transduced with pHR and pHV lentiviral vectors.**

Using the integration sites (IS) provided by GeneWerk, Germany, on P106i cells transduced with pHR and pHV lentiviral vectors was initially conducted to determine

the levels of lentiviral vectors insertions between days 3 and 30 timepoints. These data are shown in Figure 5.4. This result revealed an overall decrease in IS overtime in transduced cells for both lentiviral vectors by comparison of the day 3 and day 30 timepoints. The number of IS identified was also lower in the 3D heps infected compared to iPSCs for each lentiviral vectors. It was shown that 3D heps transduced with pHV and pHR lentiviral vectors had low number of vector copy numbers (VCN) per cell compared to P106i cells, hence a lower transduction efficiency. This data was previously presented in Chapter 4, Figure 4.12. There was also an apparent reduction in IS in iPSCs, over time suggesting cell death during propagation or the possibility of certain subpopulations of cells with low copy numbers expanding and outgrowing cells with high copy numbers in the bulk subpopulations. Figure 5.4.

Since lentiviral vectors can transduce both dividing and non-dividing cells, they are considered as a great tool to evaluate combinatorial gene transfer into different cell types. P106i cells are characterised by a high proliferation rate while 3D heps applied for this study had a low proliferation rate. The low rate of proliferation in 3D heps was previously reported (Rashidi et al., 2018). In our study, 3D heps transduced with pHV and pHR lentiviral vectors had a low number of VCN compared to highly proliferative p106i cells. It was speculated that the low proliferation rate of hepatospheres and specificity of HLCs affected VCN and lower transduction efficiency compared to P106i cells. Further, permissivity of lentiviral to transduce target cells highly depend on the receptors and on the viruses themselves. Different pseudotyping strategies will enhance transduction outcome.



**Figure 5.4: Integration site analysis on bulk P106i iPSCs and 3D heps. Integration site analysis of P106i and 3D heps transduced with pHR and pHV lentiviral vectors. The**

result exhibited a significant reduction in IS in iPSCs overtime for both lentiviral vectors. Also, an overall decrease in IS was observed in 3D heps transduced with pHR and pHV lentiviral vectors compared to P106i cells. This analysis was normalised based on one million sorted read. The levels of significance were measured by unpaired T-test and \* and \*\* is denoted as  $P < 0.05$  and  $P < 0.005$ , respectively. Abbreviation: d3: Day 3; d30: Day 30; Safe: pHR lentiviral vector; Unsafe: pHV lentiviral vector; HLC: 3D heps; iPSC: P106i cells.

### **5.3.2. Insertion site analysis on P106i samples transduced with pHR and pHV lentiviral vectors at two different time points**

Due to reduction in IS identified over time (Figure 5.4), a comparison of the IS was carried out. To do this, the top 10 IS were examined for each time point for P106i cells transduced with pHR and pHV lentiviral vectors. Using the relative sequence count (RSQ) analysis, the number of the counts of integrated lentiviral vectors was assigned. The RSQ is the division of ASQ by the total sequence, resulting in normalised sequence count per gene. Three biological replicates were analysed for each transduced sample and the top 10 IS identified were also significant [ $P < 0.05$ ]. As a result, this analysis provided the top 10 insertions with the highest sequence counts which are shown in (Figures 5.5-5.8) accordingly.

Sample	iPSCxPHRx1			iPSCxPHRx1		
Replicate	CL0207u03			CL0207u04		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	ZIC3	0.062	X+137724375	GRB2	0.071	17+75365752
Top2	TELO2	0.057	16+1498343	PPP3CA	0.062	4+100859010
Top3	GRB2	0.056	17+75365752	ITGBL1	0.051	13+101621168
Top4	FBN3	0.049	19+8082359	RPS6KA2	0.049	6+166509445
Top5	PRKAR2A	0.049	3-48762504	MAOB	0.046	X+43862190
Top6	LINC00113	0.048	21-27539716	SERHL	0.043	22-42477456
Top7	CNTN5	0.046	11+100131925	LOC101929497	0.041	11+127657140
Top8	TCEB3	0.046	1+23752717	UIMC1	0.041	5-176956476
Top9	CIRBP	0.045	19-1272965	CDCA4	0.04	14+105010821
Top10	LINC01478	0.045	18-44202067	PPP1R3C	0.04	10-91686277
#All Other mapp. IS	19797	99.497		25009	99.516	
Seq Count 10 Strongest	597			678		
Seq Count all other mapp. IS	118038			139465		
Total Seq Counts Used	118635			140143		
Sample	iPSCxPHRx2			iPSCxPHRx2		
Replicate	CL0207u05			CL0207u06		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	LOC101928911	0.068	6-87863209	DIP2B	0.077	12-50537871
Top2	TBC1D22A	0.046	22-47004981	LOC1720	0.06	2+83128242
Top3	ARFGEF1	0.045	8-67273941	NLGN1	0.06	3-174251674
Top4	ANKS1B	0.044	12+99427461	P4HTM	0.06	3-48999447
Top5	CCDC57	0.044	17-82114312	EFNA2	0.054	19+1305818
Top6	EPC1	0.044	10-32327073	EMB	0.054	5-50409689
Top7	PXN	0.044	12-120252743	EPHA4	0.054	2-220762658
Top8	SLCO3A1	0.044	15-91968807	TCF12	0.054	15+56922333
Top9	SPOP	0.044	17+49635480	C9orf62	0.048	9-135374602
Top10	SPATC1	0.043	8+144009360	GPATCH1	0.048	19+33114926
#All Other mapp. IS	21549	99.534		10571	99.431	
Seq Count 10 Strongest	503			95		
Seq Count all other mapp. IS	107485			16682		
Total Seq Counts Used	107988			16777		
Sample	iPSCxPHRx3			iPSCxPHRx3		
Replicate	CL0207u07			CL0207u08		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	LRP1B	0.053	2-141525084	LINC01021	0.074	5-27660390
Top2	PPFIA1	0.053	11+70346466	SMCO4	0.059	11-93471094
Top3	LUZP2	0.051	11-24293924	STPG2	0.059	4-98182209
Top4	C11orf80	0.047	11-66768845	KCNH8	0.056	3+19533606
Top5	CXCL5	0.045	4+74003161	VPS8	0.051	3+184933538
Top6	OR51B5	0.045	11-5343225	MOV10L1	0.05	22+50098458
Top7	PCDH18	0.041	4-137081056	MYCBP2	0.05	13-77103934
Top8	DIS3L2	0.04	2+232104293	LINC01300	0.046	8-141301038
Top9	ACHE	0.038	7+100917818	HAS2	0.044	8-121238626
Top10	UBE2J2	0.038	1-1263284	LUZP2	0.044	11+25548016
#All Other mapp. IS	22497	99.549		19218	99.467	
Seq Count 10 Strongest	598			590		
Seq Count all other mapp. IS	131956			110104		
Total Seq Counts Used	132554			110694		

**Figure 5.5: Top 10 insertions sites (IS) in P106i iPSCs transduced with pHR lentiviral vector at Day 3. Three biological replicates of the top 10 IS in P106i cells**

transduced with pHR lentiviral vector (P106i pHR 1-2-3) at Day 3 time point, using relative sequence count (RSQ) analysis, are shown. The RSQ is the division of ASQ by the total sequence, resulting in the average sequence count per gene. In the left side of the figure, based on percentage of integration frequency for each gene Top 1 to Top 10 insertion sites were ranked sequentially, and colour coded. The frequency percentage is the percentage of that particular viral insertion in the total bulk samples. These percentages are the most frequent insertion for each sample replicate. The highest sequence count (Top 1) is the first row of the gene column, and the lowest sequence count (Top 10) is the last row which are shown in maroon and purple colours, respectively. For each sample number, two replicates were analysed which are presented in parallel to each other. Further, gene names that are highlighted in grey are repeated twice in each replicate.

Sample	iPSCxPHRx1			iPSCxPHRx1		
Replicate	CL0210u01			CL0210u02		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	DENND1B	0.158	1+197762619	ANKFN1	0.173	17-56435061
Top2	ARMCX2	0.144	X-101683631	IKZF3	0.128	17+39773186
Top3	IKZF3	0.144	17+39773186	COL18A1	0.114	21-45446260
Top4	NEGR1	0.127	1+71386274	STK38L	0.108	12+27217186
Top5	STK38L	0.126	12+27217186	ARHGAP32	0.106	11-129150113
Top6	C3orf67-AS1	0.122	3+58867823	DLEU2	0.104	13-50042982
Top7	MIR7977	0.121	3+176346768	C9orf62	0.096	9-135310498
Top8	SPATS2	0.12	12-49411255	KIAA1598	0.095	10-116891268
Top9	KCTD3	0.118	1-215503066	MIR4679-2	0.095	10+89041159
Top10	LOC100133669	0.118	8+143002161	SHROOM3	0.095	4+76515590
#All Other mapp. IS	8946	98.702		10848	98.886	
Seq Count 10 Strongest	1234			1126		
Seq Count all other mapp. IS	93823			99961		
Total Seq Counts Used	95057			101087		

---

Sample	iPSCxPHRx2			iPSCxPHRx2		
Replicate	CL0210u03			CL0210u04		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	IKZF3	0.163	17+39773186	NMS	0.147	2+100510429
Top2	ANKFN1	0.154	17-56435061	KCTD3	0.139	1-215503066
Top3	STK38L	0.124	12+27217186	STK38L	0.123	12+27217186
Top4	GHITM	0.122	10+83574656	CHST12	0.121	7+2414238
Top5	ERI3	0.12	1-44341910	MIR4500HG	0.119	13-86646644
Top6	C3orf67-AS1	0.119	3+58867823	SNTG1	0.118	8-49610898
Top7	ESYT2	0.113	7-158758861	PDLIM7	0.112	5+177479255
Top8	KCTD3	0.113	1-215503066	GRIK2	0.11	6+102713082
Top9	RARRES2	0.112	7+150348320	TLK1	0.108	2-171018870
Top10	GPD2	0.111	2-156654771	PXDNL	0.106	8+51235727
#All Other mapp. IS	8810	98.749		9492	98.797	
Seq Count 10 Strongest	1280			1123		
Seq Count all other mapp. IS	101044			92258		
Total Seq Counts Used	102324			93381		

---

Sample	iPSCxPHRx3			iPSCxPHRx3		
Replicate	CL0210u05			CL0210u06		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	IKZF3	0.169	17+39773186	NEGR1	0.154	1+71386274
Top2	STK38L	0.148	12+27217186	ANKFN1	0.149	17-56435061
Top3	COL18A1	0.131	21-45446260	STK38L	0.144	12+27217186
Top4	ANKFN1	0.122	17-56435061	IKZF3	0.127	17+39773186
Top5	SPATS2	0.113	12-49411255	COL18A1	0.122	21-45446260
Top6	CLK4	0.106	5+178612092	CDH4	0.121	20+61216560
Top7	MIR4500HG	0.104	13-86646644	NMS	0.116	2+100510429
Top8	GRIK2	0.102	6+102713082	LCLAT1	0.107	2-30582954
Top9	GID8	0.101	20+62944174	CTB-7E3.1	0.105	5+165998349
Top10	LINC01195	0.101	16+9509905	SPATS2	0.1	12-49411255
#All Other mapp. IS	10949	98.803		11052	98.755	
Seq Count 10 Strongest	1193			1145		
Seq Count all other mapp. IS	98451			90806		
Total Seq Counts Used	99644			91951		

**Figure 5.6: Top 10 insertions sites (IS) in P106i iPSCs transduced with pHR lentiviral vector at Day 30. Three biological replicates of the top 10 IS in P106i cells transduced with pHR lentiviral vector (P106i pHR 1-2-3) at day 30 time point, using relative**

sequence count (RSQ) analysis, are shown. The highest sequence count (Top 1) is the first row of gene column, and the lowest sequence count (Top 10) is the last row which are shown in maroon and purple colours, respectively. 10 genes were repeatedly detected among replicates and are highlighted in grey. 14 genes were repeatedly detected among P106i pHV sample replicates. These genes are highlighted in grey. The IKZF3, SPATS2, ANKFN1, COLI8AI and STK28L, were the most frequent insertions detected in the top 10 IS between biological replicates.



Sample	iPSCxPHVx1			iPSCxPHVx1		
Replicate	CL0207u09			CL0207u10		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	LOC101928861	0.06	7-133821079	RNF169	0.046	11-74807230
Top2	GALNT12	0.057	9+98895529	ACOX1	0.045	17+75957965
Top3	PLSCR2	0.056	3-146380145	NCLN	0.045	19+3201028
Top4	CAD	0.05	2+27225332	ARHGAP15	0.044	2-143758058
Top5	ELN	0.049	7-74006047	FRMD8	0.043	11-65409304
Top6	MTM1	0.049	X+150580519	FETUB	0.042	3+186640712
Top7	SLC1A3	0.047	5-36665173	HNRNPR	0.042	1+23310529
Top8	NPLOC4	0.046	17-81593613	PCSK4	0.041	19+1483155
Top9	ARID1B	0.044	6-156975770	LOC283585	0.04	14-87373613
Top10	NOX3	0.044	6+155658082	MED13L	0.04	12+115461930
#All Other mapp. IS	16441	99.498		19132	99.572	
Seq Count 10 Strongest	649			584		
Seq Count all other mapp. IS	128540			135818		
Total Seq Counts Used	129189			136402		

---

Sample	iPSCxPHVx2			iPSCxPHVx2		
Replicate	CL0207u11			CL0207u12		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	CACNG8	0.058	19-53968448	LOC100507477	0.045	6+139969600
Top2	PKN2-AS1	0.044	1-88254018	GYS1	0.039	19+48988919
Top3	MACF1	0.041	1-39361554	MN1	0.038	22-27778940
Top4	SGCZ	0.04	8+14678461	EPHA7	0.037	6+93134365
Top5	ZFR2	0.038	19-3864358	MYCNUT	0.036	2-15901648
Top6	DIS3L2	0.037	2-232114394	CMPK1	0.035	1+47354067
Top7	KCNK2	0.037	1-215001717	FANCI	0.035	15-89263748
Top8	MIR5007	0.037	13+55067108	UBE2I	0.035	16+1318100
Top9	AURKAIP1	0.036	1+1379033	MYCBP2	0.034	13+77260664
Top10	CRAMP1L	0.036	16-1664252	TRPS1	0.033	8+115657241
#All Other mapp. IS	19480	99.596		24861	99.633	
Seq Count 10 Strongest	364			584		
Seq Count all other mapp. IS	89706			158610		
Total Seq Counts Used	90070			159194		

---

Sample	iPSCxPHVx3			iPSCxPHVx3		
Replicate	CL0207u13			CL0207u14		
DNA Input	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	LINGO2	0.058	9-28567418	PKP4	0.052	2-158485457
Top2	DOCK10	0.048	2+224967570	LIPE-AS1	0.051	19-42516930
Top3	LOC101927924	0.047	2+129942557	LRRC2	0.046	3-46511835
Top4	CNTN5	0.045	11-98608643	TSC2	0.046	16+2050487
Top5	GRIP1	0.042	12+66607451	XPNPEP1	0.043	10-109918840
Top6	LOC339166	0.039	17+5830747	CPSF3L	0.04	1+1314697
Top7	IL1RAPL1	0.038	X+28683843	SPATS2	0.04	12+49477582
Top8	ATP10B	0.037	5-160807941	PSMD13	0.039	11+249273
Top9	HDAC7	0.037	12+47815436	LRRIQ3	0.038	1-73947561
Top10	KDM2A	0.037	11+67245089	MIR4495	0.038	12+97852995
#All Other mapp. IS	21811	99.572		19375	99.567	
Seq Count 10 Strongest	594			612		
Seq Count all other mapp. IS	138153			140671		
Total Seq Counts Used	138747			141283		

**Figure 5.7: Top 10 insertions sites (IS) in P106i iPSCs transduced with pHV lentiviral vector at Day 3. Three biological replicates of the top 10 IS in P106i cells transduced with pHV lentiviral vector (P106i pHV 1-2-3) at Day 3 time point, using**

relative sequence count (RSQ) analysis, are shown. The highest sequence count (Top 1) is the first row of gene column, and the lowest sequence count (Top 10) is the last row which are shown in maroon and purple colours, respectively. No common genes were detected in top 10 IS in P106i cells transduced with pHV lentiviral vector at Day 3 time point.

Sample Replicate DNA Input	iPSCxPHVx1			iPSCxPHVx1		
	CL0210u07			CL0210u08		
	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	YTHDC2	0.217	5+113525064	CD164	0.22	6-109357042
Top2	GNA12	0.215	7-2776673	OBSCN	0.201	1+228328847
Top3	CD244	0.189	1+160861365	PPARD	0.197	6+35355871
Top4	OBSCN	0.187	1+228328847	YTHDC2	0.196	5+113525064
Top5	CPVL	0.183	7+29027781	MOB1A	0.181	2+74164615
Top6	RPS2	0.162	16-1963441	CPVL	0.154	7+29027781
Top7	ZZZ3	0.157	1+77609014	GNA12	0.149	7-2776673
Top8	CD27-AS1	0.155	12-6431713	ZMYM4	0.145	1+35384465
Top9	CD164	0.15	6-109357042	ACIN1	0.144	14-23087545
Top10	CDK14	0.15	7+90606416	ASH1L	0.144	1+155499196
#All Other mapp. IS	9514	98.235		9570	98.269	
Seq Count 10 Strongest	2017			2289		
Seq Count all other mapp. IS	112278			129953		
Total Seq Counts Used	114295			132242		

Sample Replicate DNA Input	iPSCxPHVx2			iPSCxPHVx2		
	CL0210u09			CL0210u10		
	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	CD164	0.232	6-109357042	GNA12	0.262	7-2776673
Top2	RPS2	0.222	16-1963441	CD164	0.238	6-109357042
Top3	AK4	0.191	1+65179261	YTHDC2	0.2	5+113525064
Top4	GNA12	0.189	7-2776673	PPP1R12B	0.187	1-202424153
Top5	GRIN2A	0.187	16+9920297	CPVL	0.183	7+29027781
Top6	EFNA5	0.174	5-107612169	MOB1A	0.181	2+74164615
Top7	LIMCH1	0.167	4-41368336	ACIN1	0.18	14-23087545
Top8	ZZZ3	0.162	1+77609014	RPS2	0.177	16-1963441
Top9	SKI	0.155	1+2265194	SKI	0.161	1+2265194
Top10	YTHDC2	0.152	5+113525064	ZC3H14	0.158	14+88583052
#All Other mapp. IS	8997	98.169		9913	98.073	
Seq Count 10 Strongest	2775			2517		
Seq Count all other mapp. IS	148799			128091		
Total Seq Counts Used	151574			130608		

Sample Replicate DNA Input	iPSCxPHVx3			iPSCxPHVx3		
	CL0210u11			CL0210u12		
	1000			1000		
Rank	Gene Name	Frequency [%]	Location	Gene Name	Frequency [%]	Location
Top1	GNA12	0.252	7-2776673	RPS2	0.223	16-1963441
Top2	ZC3H14	0.229	14+88583052	CD164	0.222	6-109357042
Top3	YTHDC2	0.224	5+113525064	GNA12	0.217	7-2776673
Top4	CPVL	0.21	7+29027781	YTHDC2	0.21	5+113525064
Top5	SKI	0.207	1+2265194	PPP1R12B	0.208	1-202424153
Top6	RPS2	0.198	16-1963441	ACIN1	0.202	14-23087545
Top7	CS	0.19	12+56283907	CPVL	0.17	7+29027781
Top8	MOB1A	0.175	2+74164615	SKI	0.169	1+2265194
Top9	TENM3	0.172	4+182485261	MOB1A	0.166	2+74164615
Top10	AK4	0.15	1+65179261	CCK	0.164	3-42330022
#All Other mapp. IS	8777	97.993		8922	98.049	
Seq Count 10 Strongest	2414			2416		
Seq Count all other mapp. IS	117851			121414		
Total Seq Counts Used	120265			123830		

**Figure 5.8: Top 10 insertions sites (IS) in P106i iPSCs transduced with pHV lentiviral vector at Day 30. Three biological replicates of the top 10 IS in P106i cells transduced with pHV lentiviral vector (P106i pHV 1-2-3) at Day 30 time point,**

using relative sequence count (RSQ) analysis, are shown. The highest sequence count (Top 1) is the first row of gene column, and the lowest sequence count (Top 10) is the last row which are shown in maroon and purple colours, respectively. 13 genes were repeatedly detected among P106i pHV sample replicates. These genes are highlighted in grey. The YTHDC2, GNA12, CPVL, RPS2, CD164, MOB1A and SKI were the most frequent insertions detected in the top 10 IS between biological replicates.

Overall, the IS data shown in figures 5.6-5.8 exhibited minimum similarity and/or overlap in top 10 IS between P106i samples at day 3 and day 30 post transduction timepoints for each vector. The number of genes repeatedly detected among replicates in P106i cells transduced with pHR and pHV lentiviral vectors at Day 30 time points were 14 and 13 genes, respectively. The number of frequent insertions between replicates were much higher for both vectors at day 30 compared to day 3 timepoint. This result suggests despite similarity in pHV and pHR lentiviral vectors, both vectors have different IS choice selection. It appears lentiviral vectors were chosen different IS, regardless of similar cell type. In addition, pHV with full LTR possibly affect the RSQ level and therefore conferring more proliferation in P106i cells.

### **5.3.3. Identification of total insertion site in genes found in P106i cells transduced with pHR and pHV lentiviral vectors at two time points**

Lentiviral integration sites were retrieved by EPTS/LM-PCR and IS were identified in P106i samples and 3D heps transduced with pHR and pHV lentiviral vectors. This is comparison of the top 10 genes identified in P106i cells transduced with pHV lentiviral vectors. These are the genes that were identified in replicates at the different time points and then were analysed to see if the same locus/gene detected at each time point. The same comparison was performed for P106i cells transduced with pHR lentiviral vectors between two time points. Results suggest the top 10 identified genes did not persist overtime and only six and twelve genes were common between P106i cells transduced with pHV and pHR lentiviral vectors at day 3 and 30 timepoints, respectively (Figures 5.9, 5.10).

Sample	Top10 CIS	CIS Order	Chromosome	Integration Locus	Gene
IPS pHV D3	Top 1	207	8	144153434	MROH1
	Top 2	169	11	66054651	PACS1, SF3B2
	Top 3	154	17	81557449	NPLOC4
	Top 4	146	8	144268041	BOP1, HSF1
	Top 5	131	16	2359813	ABCA17P
	Top 6	130	11	67124156	KDM2A
	Top 7	124	16	1669622	CRAMP1L, HN1L
	Top 8	117	3	48922554	ARIH2
	Top 9	113	3	49957760	RBM6
	Top 10	112	9	136889215	FBXW5, TRAF2
IPS pHV D30	Top 1	37	9	136892479	FBXW5, TRAF2
	Top 2	31	9	136810305	RABL6
	Top 3	31	1	1814392	GNB1
	Top 4	30	11	66097286	PACS1
	Top 5	24	5	176908026	UIMC1
	Top 6	22	19	1350779	MUM1
	Top 7	21	1	17026689	SDHB
	Top 8	21	16	1681390	HN1L
	Top 9	20	1	1388801	CCNL2
	Top 10	20	11	66786609	C11orf80

**Figure 5.9: Comparison of top 10 genes in P106i cells transduced with pHV lentiviral vector at days 3 and 30 timepoints. The genes that are highlighted in colours are detected in both time points. Integration locus is the location within the genome which pHV lentiviral vector is integrated. IS order indicating the number of times that the integration of pHV vector was detected.**

Sample	Top10 CIS	CIS Order	Chromosome	Integration Locus	Gene
IPS pHR D3	Top 1	1019	7	79048142	A2ML2, ABCA17P, ABL2, ACOX2, ACOXL, ACSF4, ACTN4, ADAMTS20, ADARB2, ADCK6, ADK, AFG3L1P,
	Top 2	714	6	78274499	A2ML2, ABCA17P, ACAN, ACER4, ACRBP, ACSL4, ACSS4, ACTN5, ADAMTS19, AGBL2, AGFG2, AKAP14,
	Top 3	464	5	78566246	ABCC6, ADAMTS19, ADAMTS3, ADCK6, AGMO, AHCYL3, ANK3, ANKHD1-EIF4EBP3, ANKRD8, ANXA2,
	Top 4	379	5	78205053	AADACL2-AS2, ACAP3, ACHE, ACN10, ACTL7B, ADAMTS20, ADAMTS4, ADAMTS7, ADAMTSL2, ADCY3,
	Top 5	339	6	78613645	AAK2, ABCA17P, ACOX2, AGPS, AKAP8L, AKAP9, ALS3, AMFR, ANKRD29, AP3S3, ARHGAP35, ATCAY, AXIN2,
	Top 6	315	7	78343945	ADAR, ADK, AJAP2, AK6, ANP32E, AP2S2, ARHGAP22, ARHGAP43, ASH1L, ATE1, BRINP4, C11orf54, C1orf211,
	Top 7	297	7	78661045	ADK, AGO1, AGO2, AMER3, ANKRD13C, ANKS1B, ANXA2P4, ARHGAP30, ARHGAP33, ARHGEF12, ARID1A,
	Top 8	284	8	78413575	ACOX2, ANGPT1, ANKRD8, ANKS4, ANXA2, APOOP6, ARFIP2, BCHE, BET2, BTBD2, BTRC, C19orf45, C5orf52,
	Top 9	272	6	78860197	ABCA17P, ACAP3, ADAM21P2, ADCY10, AGBL2, AGPAT4, ALPPL3, ALYREF, ANKS4, AP3D2, ARIH3,
	Top 10	270	8	79943591	ACSM7, AGBL5, ANKRD13D, ANKRD56, ANTXR2, ARHGAP21, ARL4C, ARPP22, ATCAY, ATG11, ATP2A3,
IPS pHR D30	Top 1	1170	7	79048142	A2ML2, ABCA17P, ABL2, ACOX2, ACOXL, ACSF4, ACTN4, ADAMTS20, ADARB2, ADK, AFG3L1P, AGAP2,
	Top 2	284	6	78625587	ABCA17P, ACOX2, AGPS, AKAP8L, AKAP9, ALS3, AMFR, AP3S3, ARHGAP35, ATCAY, AXIN2, BANP, BAZ1A,
	Top 3	110	8	79022605	AASS, AMDHD2, ANKRD29, B4GALNT5, C11orf96, CCDC138, CCDC92, CHST13, CNTNAP6, CTB-7E3.2,
	Top 4	64	9	78973736	C7orf51, CEP85L, CLEC2D, COL6A2, CTNNA2, DCP1A, DEFB114, FPGT-TNNI3K, IKZF3, KCNA6, KIAA1110,
	Top 5	58	1	1378986	AURKAIP1, CCNL2
	Top 6	53	5	78603908	ABCC6, ANK3, ARIH3, ATP11B, BANK2, CADM3, CHMP2B, CLDN2, CLNK, CNTN5, CRIPAK, DGKG, DOCK4,
	Top 7	45	19	49602947	PRR12
	Top 8	44	5	78984208	ABCA17P, ABI2, C14orf181, CNOT3, DCAF7, EPHA3, ETS2, KND2, LINC00669, LINC01117, LOC101928477,
	Top 9	38	6	78613645	ANKRD29, BCHE, CACNG3, CBLB, COL6A2, CPNE5, CSNK2A2, CTCFL, DIP2A, DYRK1A, GABPA, GMEB3,
	Top 10	38	7	79014889	ACOT8, ADCYAP2, AMICA2, AMOTL2, ANXA2P4, BAHCC2, FER, GADL2, GMCL2, LINC01193, LINC01371,

**Figure 5.10: Comparison of top 10 genes in P106i cells transduced with pHR lentiviral vector at days 3 and 30 timepoints. The genes that are highlighted in colours are detected in both time points. Integration locus is the location within the genome which pHV lentiviral vector is integrated. IS order indicating the number of times that the integration of pHV vector was detected.**

### **5.3.4. Identification of insertion sites in genes found in P106i derived 3D heps transduced with pHR and pHV lentiviral vectors**

This is comparison of the top 10 genes identified in P106i-derived 3D heps transduced with pHR and pHV lentiviral vectors. These are the genes that were identified in replicates and then were analysed to see if the same locus/gene detected between

lentiviral vectors. Three genes (*PACS1*, *CHD3*, *CYB5D1*) were detected in 3D heps transduced with pHR and pHV lentiviral vectors which were ranked as the top 2 and top 4, respectively. Other genes were not similar between 3D heps infected with pHR and pHV lentiviral vectors. The identified genes were ranked from top 1 to top 10 genes based on there is order. This suggests that despite similarity between vectors, insertion sites and number of identified genes might be different. Hence, different vectors constructs were appeared to have specific preference for IS regardless of same tissue type (Figure 5.11).

Sample	Top10 CIS	CIS Order	Chromosome	Integration Locus	Gene
HLC pHR	Top 1	17	9	136810825	RABL6
	Top 2	12	11	66092210	<b>PACS1</b>
	Top 3	12	16	1676560	CRAMP1L, HN1L
	Top 4	11	17	7870521	<b>CHD3, CYB5D1</b>
	Top 5	11	9	136892479	TRAF2
	Top 6	10	2	2573126	MYT1L
	Top 7	10	11	64769357	SF1
	Top 8	10	11	65426091	NEAT1
	Top 9	10	16	2255950	RNPS1
	Top 10	10	9	135866918	CAMSAP1
HLC pHV	Top 1	21	16	2248352	ECI1, RNPS1
	Top 2	20	17	7867680	<b>CHD3, CYB5D1</b>
	Top 3	20	19	1357855	MUM1
	Top 4	18	11	66163385	<b>PACS1</b>
	Top 5	17	19	49603220	PRR12
	Top 6	15	22	50373041	PPP6R2
	Top 7	15	1	1381710	CCNL2, LOC148413, MRPL20
	Top 8	14	9	136449996	SEC16A
	Top 9	13	11	65190379	CAPN1
	Top 10	13	17	44457900	GPATCH8

**Figure 5.11: Comparison of top 10 genes in P106i-derived 3D heps transduced with pHV and pHR lentiviral vectors. The genes that are highlighted in colours are detected in hepatospheres transduced with both vectors. Integration locus is the location within the genome which pHV and pHR lentiviral vectors are integrated. IS order indicating the number of times that the integration of pHV vector was detected. These genes were ranked from top 1 to top 10 based on their insertion site order.**

A comparison was also performed between P106i cells and P106i derived 3D heps transduced with pHR and pHV lentiviral vectors. No similar genes/locus were detected between P106i cells and 3D heps transduced with pHR lentiviral vector. Identified

genes were ranked from top 1 to top 10 based on their order for P106i cells and 3D heps. However, two genes (*PACS1* and *SF3B2*) were identified in both P106i cells and 3D heps transduced with pHV lentiviral vector. This result suggests that pHR lentiviral vector have different IS preference for P106i cells and 3D heps. Therefore, the top IS and identified genes were different. It can be concluded that for each cell and tissue types of lentiviral vectors have different IS selection (Figures 5.12 & 5.13).

Sample	Top10 CIS	CIS Order	Chromosome	Integration Locus	Gene
HLC pHR	Top 1	17	9	136810825	RABL6
	Top 2	12	11	66092210	PACS1
	Top 3	12	16	1676560	CRAMP1L, HN1L
	Top 4	11	17	7870521	CHD3, CYB5D1
	Top 5	11	9	136892479	TRAF2
	Top 6	10	2	2573126	MYT1L
	Top 7	10	11	64769357	SF1
	Top 8	10	11	65426091	NEAT1
	Top 9	10	16	2255950	RNPS1
	Top 10	10	9	135866918	CAMSAP1
IPS pHR D3	Top 1	1019	7	79048142	A2ML2, ABCA17P, ABL2, ACOX2, ACOXL, ACSF4, ACTN4, ADAMTS20, ADARB2, ADCK6,
	Top 2	714	6	78274499	A2ML2, ABCA17P, ACAN, ACER4, ACRBP, ACSL4, ACSS4, ACTN5, ADAMTS19, AGBL2,
	Top 3	464	5	78566246	ABCC6, ADAMTS19, ADAMTS3, ADCK6, AGMO, AHCYL3, ANK3, ANKHD1-EIF4EBP3,
	Top 4	379	5	78205053	AADACL2-AS2, ACAP3, ACHE, ACN10, ACTL7B, ADAMTS20, ADAMTS4, ADAMTS7,
	Top 5	339	6	78613645	AAK2, ABCA17P, ACOX2, AGPS, AKAP8L, AKAP9, ALS3, AMFR, ANKRD29, AP3S3,
	Top 6	315	7	78343945	ADAR, ADK, AJAP2, AK6, ANP32E, AP2S2, ARHGAP22, ARHGAP43, ASH1L, ATE1,
	Top 7	297	7	78661045	ADK, AGO1, AGO2, AMER3, ANKRD13C, ANKS1B, ANXA2P4, ARHGAP30, ARHGAP33,
	Top 8	284	8	78413575	ACOX2, ANGPT1, ANKRD8, ANKS4, ANXA2, APOOP6, ARFIP2, BCHE, BET2, BTBD2, BTRC,
	Top 9	272	6	78860197	ABCA17P, ACAP3, ADAM21P2, ADCY10, AGBL2, AGPAT4, ALPPL3, ALYREF, ANKS4,
	Top 10	270	8	79943591	ACSM7, AGBL5, ANKRD13D, ANKRD56, ANTXR2, ARHGAP21, ARL4C, ARPP22, ATCAV,

**Figure 5.12: Comparison of top 10 genes in P106i cells and 3D heps. This is the comparison of top 10 genes between P106i cells and P106i derived 3D heps transduced with pHR lentiviral vectors. No similar top 10 genes were detected between P106i cells at early time point and 3D heps. These genes were ranked from top 1 to top 10 based on their insertion site order.**



Sample	Top10 CIS	CIS Order	Chromosome	Integration Locus	Gene
HLC pHV	Top 1	21	16	2248352	ECI1, RNPS1
	Top 2	20	17	7867680	CHD3, CYB5D1
	Top 3	20	19	1357855	MUM1
	Top 4	18	11	66163385	PACS1
	Top 5	17	19	49603220	PRR12
	Top 6	15	22	50373041	PPP6R2
	Top 7	15	1	1381710	CCNL2, LOC148413, MRPL20
	Top 8	14	9	136449996	SEC16A
	Top 9	13	11	65190379	CAPN1
	Top 10	13	17	44457900	GPATCH8
IPS pHV D3	Top 1	207	8	144153434	MROH1
	Top 2	169	11	66054651	PACS1, SF3B2
	Top 3	154	17	81557449	NPLOC4
	Top 4	146	8	144268041	BOP1, HSF1
	Top 5	131	16	2359813	ABCA17P
	Top 6	130	11	67124156	KDM2A
	Top 7	124	16	1669622	CRAMP1L, HN1L
	Top 8	117	3	48922554	ARIH2
	Top 9	113	3	49957760	RBM6
	Top 10	112	9	136889215	FBXW5, TRAF2

**Figure 5.13: Comparison of top 10 genes in P106i cells and 3D heps. This is the comparison of top 10 genes between P106i cells and P106i derived 3D heps transduced with pHV lentiviral vector. Two genes were identified which were similar in top 10 genes between P106i cells and 3D heps transduced with pHV lentiviral vector. These genes are highlighted in colours. Identified genes were ranked from top 1 to top 10 based on their insertion site order.**

### **5.3.5. Identification of cancer related genes IS and common IS in P106i cells and 3D heps**

The total number of cancer related genes; Proto-Oncogenes (PO), Tumour Suppressor genes (TSGs) & oncogenes were identified in P106i samples and 3D heps transduced with pHR and pHV lentiviral vectors. This data was used to find common insertion sites (CIS) within these samples. The CIS are the all the genes that the virus has been inserted into at least once in the three replicate samples that were sequenced.

For the proliferation assay, in P106i samples (n=3) transduced with pHR and pHV lentiviral vectors, 154 and 169 POs were identified at early timepoint, respectively. As the time progressed following transduction, number of POs were decreased. Similarly, in P106i samples (n=3) transduced with pHR and pHV lentiviral vectors 122 and 126 TSGs were identified at late time points, respectively. However, at late time points the

number of TSGs were decreased in both vectors. Oncogenes comprised a small portion of cancer-related genes and in P106i samples infected with pHR and pHV lentiviral vectors number of oncogenes were found to be 6 and 9, respectively. The number of oncogenes decreased in P106i samples infected with pHV lentiviral vector at late time points. In 3D heps infected with pHR lentiviral vectors, the number of POs, TSGs and oncogenes were 119, 93 and 6 respectively. These numbers were increased in 3D heps transduced with pHV lentiviral vector. This result suggests the difference between SIN and full LTR lentiviral vectors (Table 5.1).

**Table 5.1. Identification of number of cancer related genes in P106i and 3D hepatosphere samples transduced with pHR and pHV lentiviral vectors. Abbreviations: Proto-Oncogenes (PO), Tumour Suppressor Genes (TSG).**

Samples	Vector Name	PO		TSG		Oncogenes	
		D3	D30	D3	D30	D3	D30
P106i	pHR	154	108	122	80	6	6
P106i	pHV	169	141	126	1e5	9	6
3D heps	pHR	119		93		6	
3D heps	pHV	129		101		8	

Cancer-related genes CIS identified in P106i and 3D hepatosphere samples and compared. Comparison of cancer related genes IS were identified in P106i samples transduced with pHR and pHV lentiviral vectors and at two early and late time points. In P106i samples transduced with pHR and pHV lentiviral vectors, the total number of cancer genes 64% (183) and 69.4% (222) were shared between early and late time points respectively. In P106i samples infected with pHR lentiviral vector the number of oncogenes 71.4% (5) were shared between early and late time points. However, in samples infected with pHV vector 66.7% (6) were shared between early and late time points. In terms of POs, in P106i samples transduced with pHR and pHV lentiviral vectors 66.7% (104) and 74.0% (131) were shared between early and late time points, respectively. For TSG, 60.8% (78) and 64.7% (90) were shared between the early and late time points. In 3D heps transduced pHR and pHV lentiviral vectors, the total number of cancer genes 67.4% (182) were shared between two vectors. The number of oncogenes, POs and TSGs, 75.0% (6), 72.2% (104) and 61.3% (73) were shared between pHR and pHV lentiviral vectors respectively. These results suggest a

possible clonal dominance in P106i cells transduced with pHR and pHV lentiviral vectors and between two time points. In P106i cells transduced with pHV lentiviral vectors numbers of cancer related genes were higher.

Overall, this finding exhibited pHV lentiviral vector might be conferring more proliferation and a possible clonal dominance in P106i samples overtime. This was possibly triggered by the LTR effect of the pHV lentiviral vector. Similarly, number of cancer related genes were high in 3D heps transduced with pHV lentiviral vector. This shows a similar pattern in samples transduced with pHV vector, regardless of cell type. The pHR lentiviral vector was considered a safer vector, nevertheless levels of identified cancer related genes were also high. This signifies the importance.

### **5.3.6. Gene enrichment analysis**

The enrichment of IS genes for biological pathways was used to interpret sets of genes and related functional properties. One of the gene enrichment analyses used for this study was to find the biological pathways and comparing these pathways between different samples. From this analysis, the top 10 biological pathways from CIS retrieved by EPTS/LM-PCR in iPSC and 3D heps transduced with pHR and pHV lentiviral vectors by the ASQ analysis were identified. All the identified top 10 biological pathways were significant [ $P < 0.05$ ].

To compare the biological pathways in P106i samples transduced with pHR and pHV lentiviral vectors at the early and late time points, the shared biological pathways were also analysed. The shared biological pathways between pHR and pHV lentiviral vectors were 80% and 30% respectively. However, in 3D heps, 80% of the common biological pathways were shared between two lentiviral vectors [ $p < 0.05$ ]. This finding suggests that the similar biological pathways were targeted by two vectors. As explained earlier there was a reduction in IS in P106i samples overtime and therefore biological pathways might have been affected.

The comparison between P106i and 3D heps, the cells transduced with pHR and pHV vectors shared 60% and 80% of the shared biological pathways [ $p < 0.05$ ]. This result suggests that similar biological pathways were shared between different two different tissue types. Also, there was a slight decrease in the number of biological pathways, particularly in P106i transduced with pHR lentiviral vector.

The top 10 biological pathways identified in P106i samples transduced with pHR lentiviral vectors at days 3 and 30 time points were mostly involved with signal transduction, neuronal development, cell cycle and proliferation. However, biological pathways identified in P106i cells at day 30 timepoint was more involved with cell signalling. In 3D heps transduced with pHR lentiviral vector, biological pathways were associated with metabolism, cell signalling and cell cycle. This finding suggests that biological pathways were shared between P106i cells and 3D heps transduced with pHR lentiviral vectors (Tables 5.2 & 5.3).

**Table 5.2. The top 10 biological pathways identified in P106i samples transduced with pHR lentiviral vector at Day 3 and Day 30 time points.**

Time-point	Biological Pathway	Function
Day 3	BCR signalling pathway	B-cell development and activation
	NGF signalling via TRKA from the plasma membrane	Neuronal survival and differentiation
	Signalling by NGF	Survival and apoptosis in neurons
	Regulation of signalling by CBL	Regulation of cell function
	CREB phosphorylation	Neuronal development
	Interleukin receptor SHC signalling	Regulation of different signalling pathways
	RSK activation	Role in cancer development
	Axon guidance	Neuronal development
	GRB2: SOS provides linkage to MAPK signalling for Integrins	Signal transduction/cell communication
	Day 30	IL2-mediated signalling events
Peptide ligand-binding receptors		
Class A/1 (Rhodopsin-like receptors)		Signalling and extracellular signals
GPCR ligand binding		Signal transduction
Signalling by GPCR		Signal transduction
Signal Transduction		Signal transduction

**Table 5.3. The top 10 biological pathways identified in 3D heps transduced with pHR lentiviral vector at Day 3 post transduction.**

Biological Pathway	Function
NOSIP mediated eNOS trafficking	Cell cycle
BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members	Apoptosis
D-myo-inositol (1,4,5)-trisphosphate degradation	Signal transduction and cardiac problems
SEMA3A-Plexin repulsion signalling by inhibiting Integrin adhesion	Neuronal structure
1D-myo-inositol hexakisphosphate biosynthesis II (mammalian)	Production of phytate or phytic acid
D-myo-inositol (1,3,4)-trisphosphate biosynthesis	Metabolism
eNOS activation and regulation	Protein phosphorylation
Metabolism of nitric oxide	Metabolism
3-phosphoinositide degradation	Cell signalling
super pathway of D-myo-inositol (1,4,5)-trisphosphate metabolism	Metabolism

The top 10 biological pathways identified in P106i samples transduced with pHV lentiviral vectors at day 3 time point were mainly associated with lipid and glucose homeostasis and metabolism. However, biological pathways identified in P106i cells at day 30 timepoint was more involved with S1P family pathways such as S1P4, S1P5, S1P2 and S1P3 pathways. These pathways have a role in angiogenesis, fibrosis proliferation and cancer. Also, reelin signalling pathways has a role in central nervous development. In 3D heps transduced with pHV lentiviral vector, biological pathways were associated with mainly associated with cell cycle, DNA replication and DNA repair.

From previous IS result analysis, it was shown that viral insertions were different between vectors, tissue types and also at different time points. Interestingly, biological pathways analysis revealed both lentiviral vectors target same biological pathways. Further, similar biological pathways persisted over time in P106i cells at different time points. It appears that pHV lentiviral vector was targeting similar biological pathways regardless of cell type. In addition, biological pathways did not decrease overtime. Hence, it is imperative to study IS as well as biological pathways for genotoxicity studies (Tables 5.4 & 5.5).

**Table 5.4. The top 10 biological pathways identified in P106i samples transduced with pHV lentiviral vector at Day 3 and Day 30 time points.**

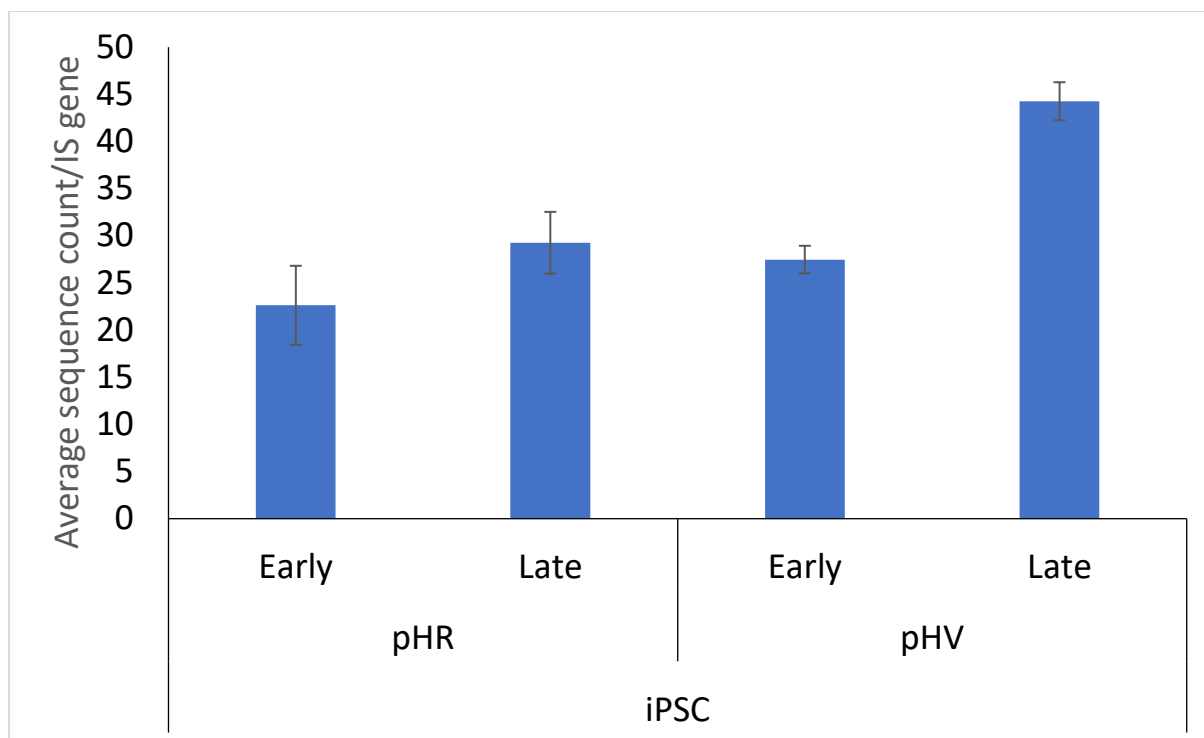
Time-point	Biological Pathway	Function
Day 3	Inhibition of HSL	Lipid and glucose homeostasis
	PDE3B signalling	Regulation of heart muscle,vascular smooth muscle
	Inhibition of TSC complex formation by PKB	Pigmentation and hyperactivation of glycogen synthase
	AKT phosphorylates targets in the cytosol	
	Regulation of Rheb GTPase activity by AMPK	Differential regulation of mTORC1 signalling in response to alcohol and leucine
	Regulation of AMPK activity via LKB1	Metabolism and growth control in tumour
	Energy dependent regulation of mTOR by LKB1-AMPK	
	PIP3 activates AKT signaling	
	p38 signalling mediated by MAPKAP kinases	Signal transduction mediator
	mTOR signalling	Cell cycle, proliferation, cancer
Day 30	S1P5 pathway	Fibrosis and angiogenesis
	Unblocking of NMDA receptor, glutamate binding and activation	Signal transduction
	S1P4 pathway	Cell migration particularly for lymphocytes
	CREB phosphorylation through the activation of CaMKII	Signal transduction
	Ras activation upon Ca <sup>2+</sup> influx through NMDA receptor	Signal transduction and their role in cancer
	S1P2 pathway	Proliferation, cell migration
	CREB phosphorylation through the activation of Ras	Activation of dopaminergic neurons
	Reelin signalling pathway	Central nervous system development
	S1P3 pathway	Cancer, angiogenesis
Post NMDA receptor activation events	Cell signalling	

**Table 5.5. The top 10 biological pathways identified in 3D heps transduced with pHV lentiviral vector at Day 3 post transduction.**

Time-point	Biological Pathway	Function
Day 3	Assembly of the ORC complex at the origin of replication	DNA replication and cell cycle
	E2F-enabled inhibition of pre-replication complex formation	DNA replication and cell cycle
	CDC6 association with the ORC:origin complex	Cell cycle
	G1/S-Specific Transcription	Cell cycle
	3-phosphoinositide biosynthesis	Cell proliferation and oncogenic transformation
	ATM pathway	Cell cycle arrest, DNA repair and apoptosis
	E2F mediated regulation of DNA replication	DNA replication and cell cycle
	G2/M DNA damage checkpoint	Cell cycle
	Activation of the pre-replicative complex	DNA replication

In addition to the gene enrichment analysis, we tracked the viral sequence count over time (30 days) with P106i cell line, using proliferation assay. The hiPSC cell line has a profound proliferation ability and can be maintained for a long time in culture. The average ASQ per IS gene in P106i samples transduced with pHR and pHV lentiviral vectors at the early and late time points was quantified.

The result exhibited a slight increase in the average sequence count in the P106i samples transduced with pHR lentiviral vector at the late time point which was not significant. However, in P106i samples transduced with pHV lentiviral vector there is a significant increase [ $P < 0.05$ ] in the average ASQ at the late time point. This further confirms the effect of full LTR vector on the average sequence count and also over time (Figure 5.14). This result suggests that pHV lentiviral vector with full LTR possibly conferring more proliferation compared to pHR SIN vector. Hence, higher number of sequence count. The average sequence count was also high in pHR lentiviral vector which is still not desirable.



**Figure 5.14: The quantification of average ASQ per IS gene in P106i samples transduced with pHR and pHV lentiviral vectors at two time points. In P106i samples transduced with pHV lentiviral vectors there is a significant increase in the average ASQ at a late time point. The level of significance was measured, using two tailed Mann-Whitney U test which \* is denoted as  $P < 0.05$ .**

Following the sequence count analysis, the number of cancer genes which were increased in sequence count analysis were also identified. In total, nine cancer genes were identified for 8.74% and 8.57% of all the cancer gene IS retrieved in P106i samples transduced with pHR and pHV lentiviral vectors, respectively. The nine POs were increased by >2 fold in average ASQ in P106i samples transduced with pHR and pHV lentiviral vectors at the early and late time points. There was no ASQ in oncogenes identified which had an increase by >2 fold. Further, no TSGs were identified which showed a reduction by >2 fold. The nine identified POs in P106i samples transduced with pHR and pHV lentiviral vectors were shown in (Tables 5.6 & 5.7). In addition, there is no common PO between pHR and pHV lentiviral vectors which suggests the possible effect of two vector constructs on different PO.



**Table 5.6. The nine identified Proto-Oncogenes in P106i samples transduced with pHR lentiviral vector.**

Gene symbol	D3	D30	Fold change
PLAG1	1.00	11.00	10.00
GLI1	1.00	28.67	27.67
TPM3	1.00	8.00	7.00
RUNX1	10.00	32.50	2.25
MDS2	4.00	12.00	2.00
OLIG2	2.67	25.00	8.38
GFI1B	7.00	94.67	12.52
FGF6	5.50	17.33	2.15
FGR	5.50	21.67	2.94

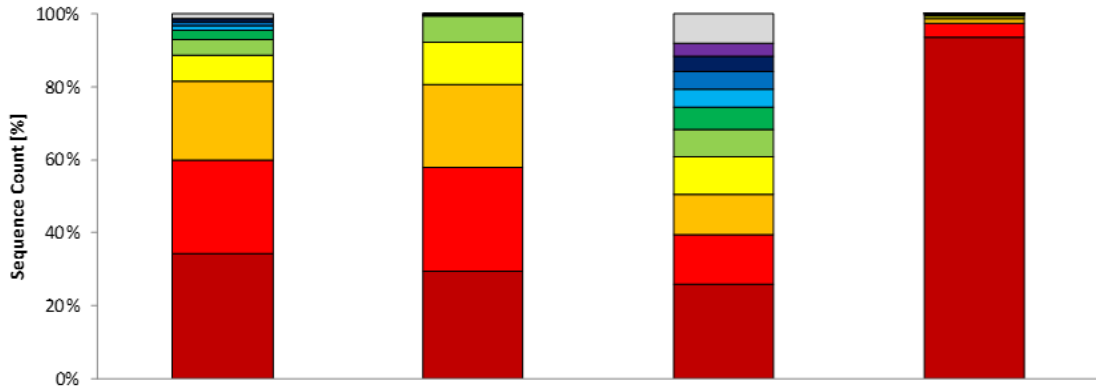
**Table 5.7. The nine identified Proto-Oncogenes in P106i samples transduced with pHV lentiviral vector.**

Gene symbol	D3	D30	Fold change
PML	4.00	24.00	5.00
CCND1	37.33	135.00	2.62
WDR11	10.00	56.67	4.67
NFKB2	11.00	89.00	7.09
SKI	125.00	594.00	3.75
JAK2	9.67	46.33	3.79
PRCC	30.00	132.33	3.41
PBX1	50.33	156.67	2.11
AFF1	22.33	159.67	6.15

### 5.3.7. Single-Cell clones Insertion Site analysis

To determine the effect of viral insertion at single cell level, P106i cells were transduced with pHR and pHV lentiviral vectors. These cells were single cell cloned as previously described in Chapter 2. Insertion sites were recovered by EPTS/LM-

PCR from fragmented DNA using primers specific to the HIV-1 LTR. All the insertions for pHV and pHR clones were significant and the top 10 IS for both lentiviral vectors and are listed in (Figures 5.15, 5.16, 5.17, 5.18, 5.19 & 5.20). The result exhibited some polyclonality in pHV clones. However, only pHV clone 6 was identified to be a single cell clone with one prominent insertion site. Similar analysis was performed for pHR clones. From result analysis, variety of polyclonality was detected in pHR clones. Nevertheless, pHR clone R was identified to be a single cell clone with one dominant insertion site (*GR1D2*). In addition, pHR clone Q had two prominent insertion sites.



Vector	pHV	
Sample	Clone1	Clone2
DNA Input [ng]	400	400

Rank	Gene	[%]	Gene	[%]
Top1	ATRNL1	34.452	IRAK1	29.472
Top2	IP6K1	25.284	RPS6KA4	28.589
Top3	ANGPT1	21.836	FRYL	22.486
Top4	ZMYM4	7.199	LEF1	11.839
Top5	LOC100129520	4.151	IL26	7.11
Top6	LOC101927768	2.456	FRYL	0.074
Top7	VBP1	1.314	IRAK1	0.057
Top8	FUT9	0.98	FRYL	0.047
Top9	MIR4431	0.947	RPS6KA4	0.044
Top10	NLGN4X	0.211	FRYL	0.043
#All Other mapp. IS	47	1.17	63	0.239

Seq Count 10 Strongest	382331	223992
Seq Count all other mapp. IS	4526	537
Total Seq Counts Used	386857	224529

Vector	pHV	
Sample	Clone4	Clone6
DNA Input [ng]	400	400

Rank	Gene	[%]	Gene	[%]
Top1	AXIN1	25.903	LINC01249	93.454
Top2	LINC01571	13.498	PLCG1	4.119
Top3	MED13L	11.187	LOC101928223	1.13
Top4	SOX8	10.163	ADAM21P1	0.569
Top5	THTPA	7.512	LINC01249	0.319
Top6	CASC6	6.114	LINC01249	0.272
Top7	CSK	5.008	MGAT4C	0.135
Top8	RBFOX3	4.818	MGAT4C	0.002
Top9	VPS13C	4.152	LINC01249	0.001
Top10	ANKRD11	3.65	LINC01249	0.001
#All Other mapp. IS	68	7.995	4	-0.002

Seq Count 10 Strongest	231942	301094
Seq Count all other mapp. IS	20155	4
Total Seq Counts Used	252097	301098

**Figure 5.15: Top 10 insertion sites in P106i iPSCs cells which were transduced with pHV lentiviral vector and single cell cloned manually. The top 10 IS in pHV clones (1, 2, 4 and 6), using relative sequence count (RSQ) analysis, are shown. In the left side of the figure, based on integration frequency for each gene Top 1 to Top 10 insertion sites were ranked sequentially, and colour coded. The highest sequence count (Top 1) is the first row of the gene column, and the lowest sequence count (Top 10) is the last row which are in maroon and purple colours, respectively. The result exhibited polyclonality in all pHV clones. However, only pHV clone 6 was identified to be a single cell clone with one prominent insertion site.**

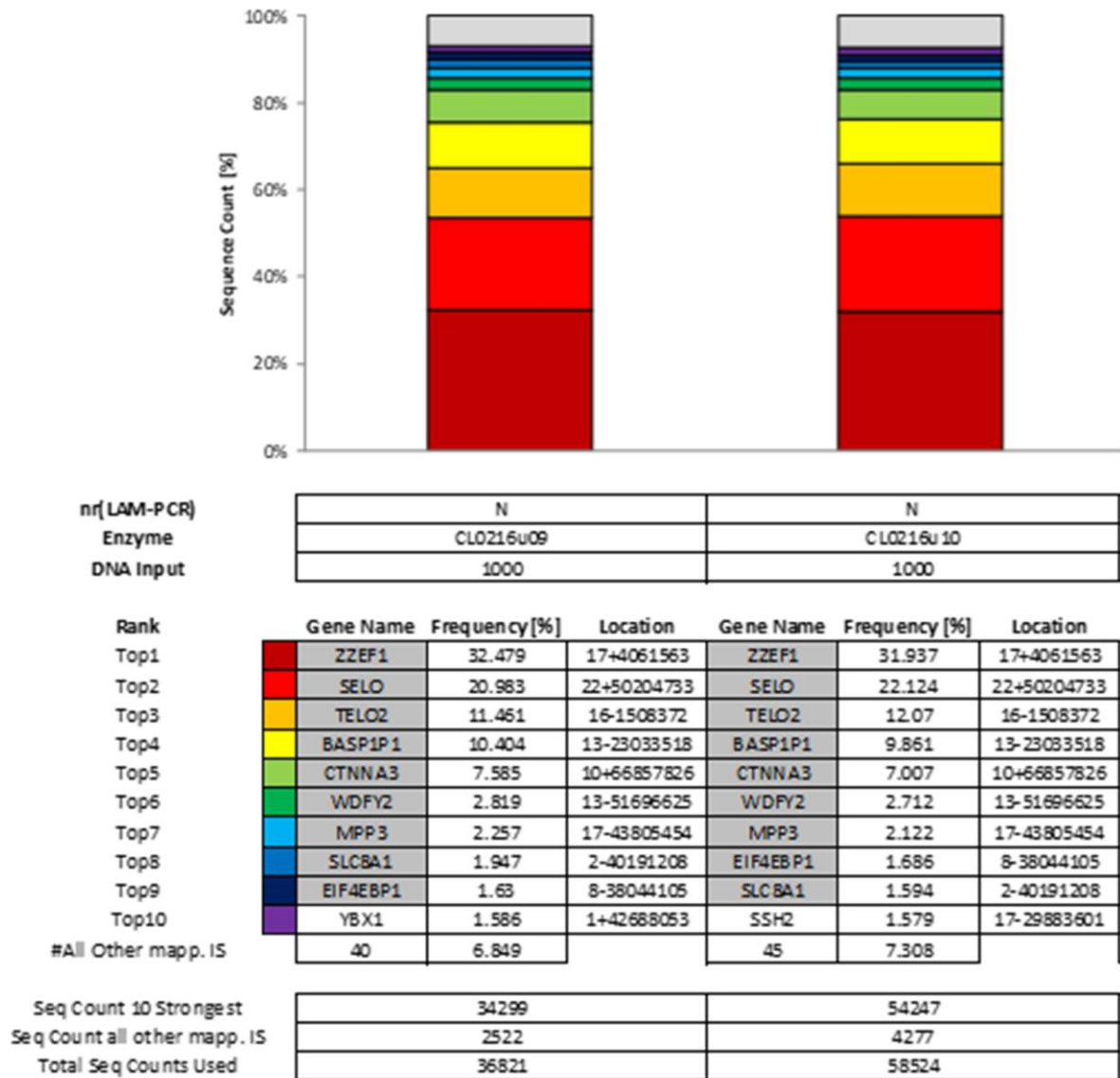


Figure 5.16: Top 10 insertion sites in P106i iPSCs cells which were transduced with pHR lentiviral vector and single cell cloned manually. The top 10 IS in pHR clone N, using RSQ analysis, are shown. In the left side of the figure, based on integration frequency for each Top 1 to Top 10 insertion sites were ranked sequentially, and colour coded. The highest sequence count (Top 1) is the first row of the gene column, and the lowest sequence count (Top 10) is the last row which are in maroon and purple colours, respectively. The genes that are frequently repeated in pHR clone N are highlighted in grey. The result exhibited variety of polyclonality in clone N.

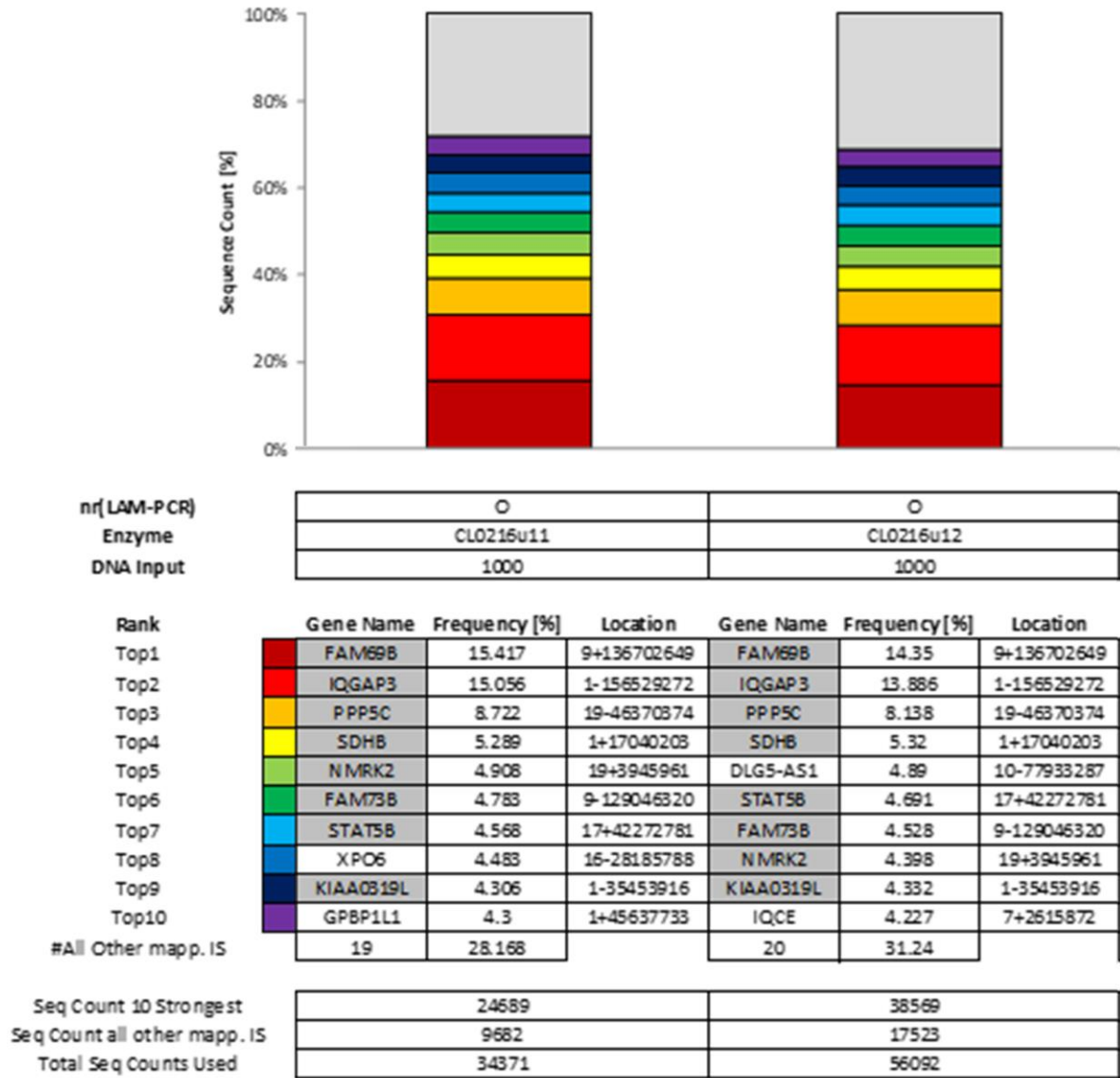


Figure 5.17: Top 10 insertion sites in P106i iPSCs cells which were transduced with pHR lentiviral vector. The top 10 IS in pHR clone O, using RSQ analysis. The frequent genes for clone N are highlighted in grey. Polyclonality was also detected in P106i cells clone N.

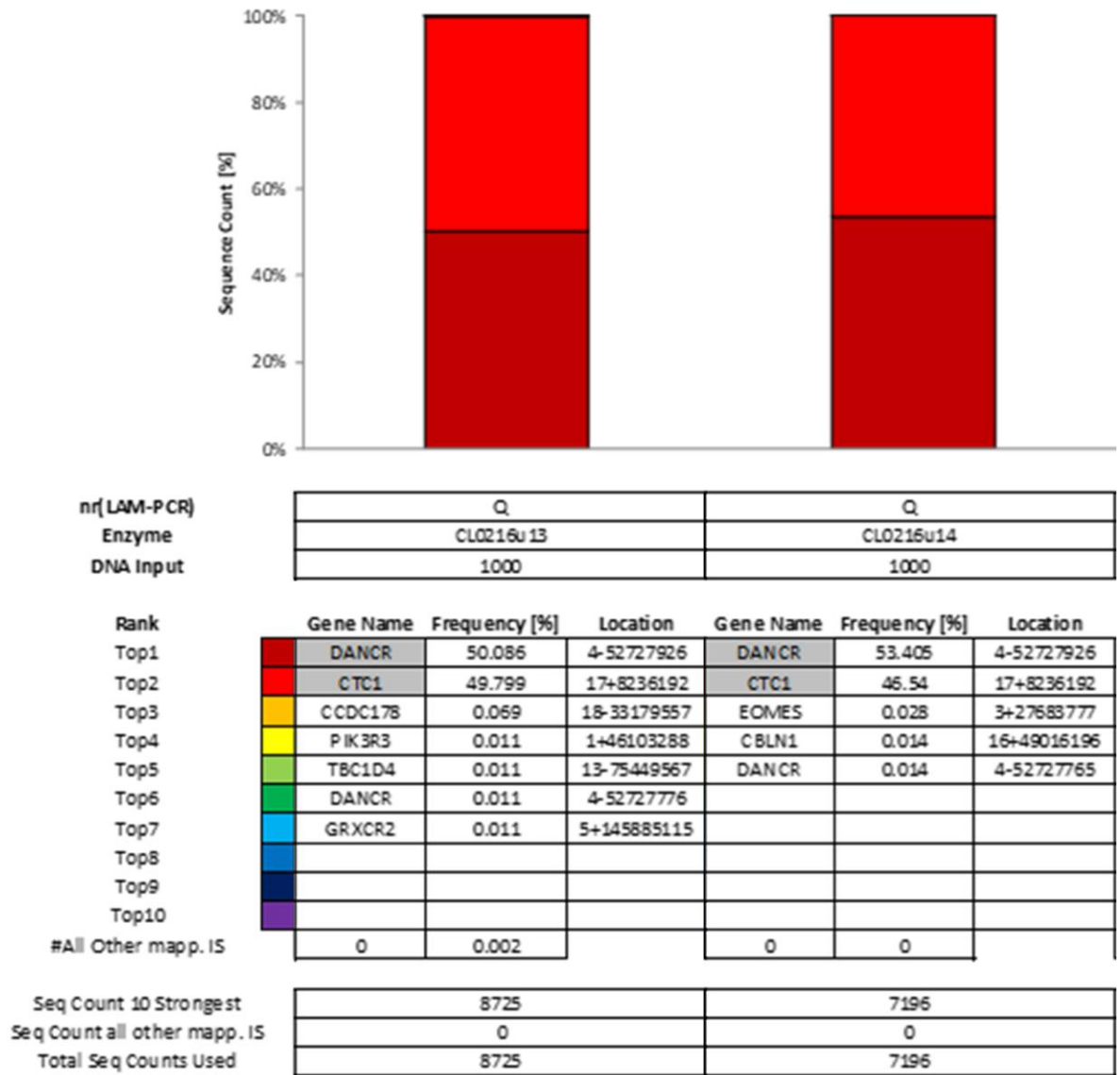


Figure 5.18: Top 10 insertion sites in P106i iPSCs transduced with pHR lentiviral vector. The top 10 IS in pHR clone Q, using RSQ analysis. Only two frequent genes for clone Q were detected which were highlighted in grey. There was less degree of polyclonality compared to other pHR clones.

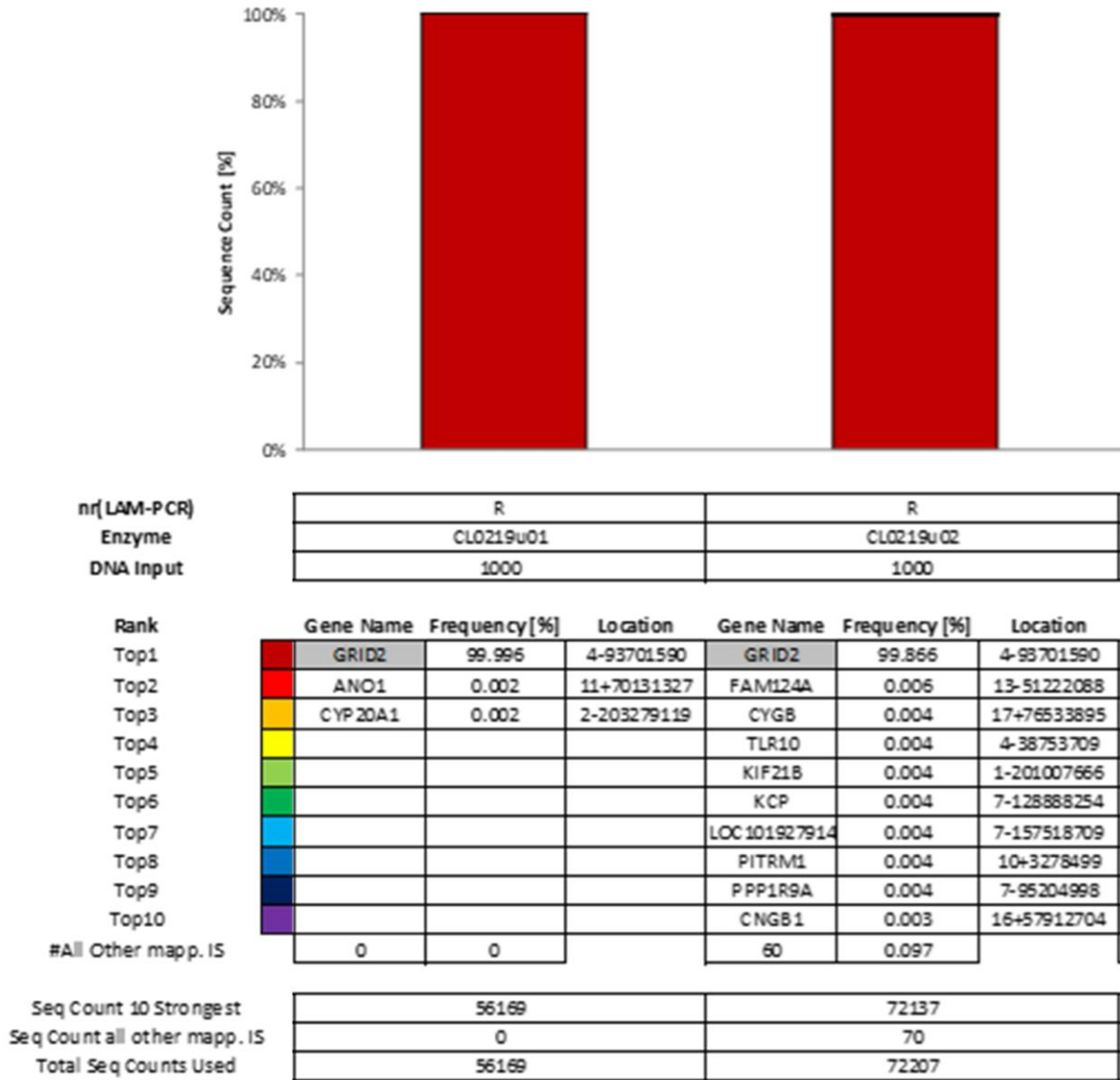
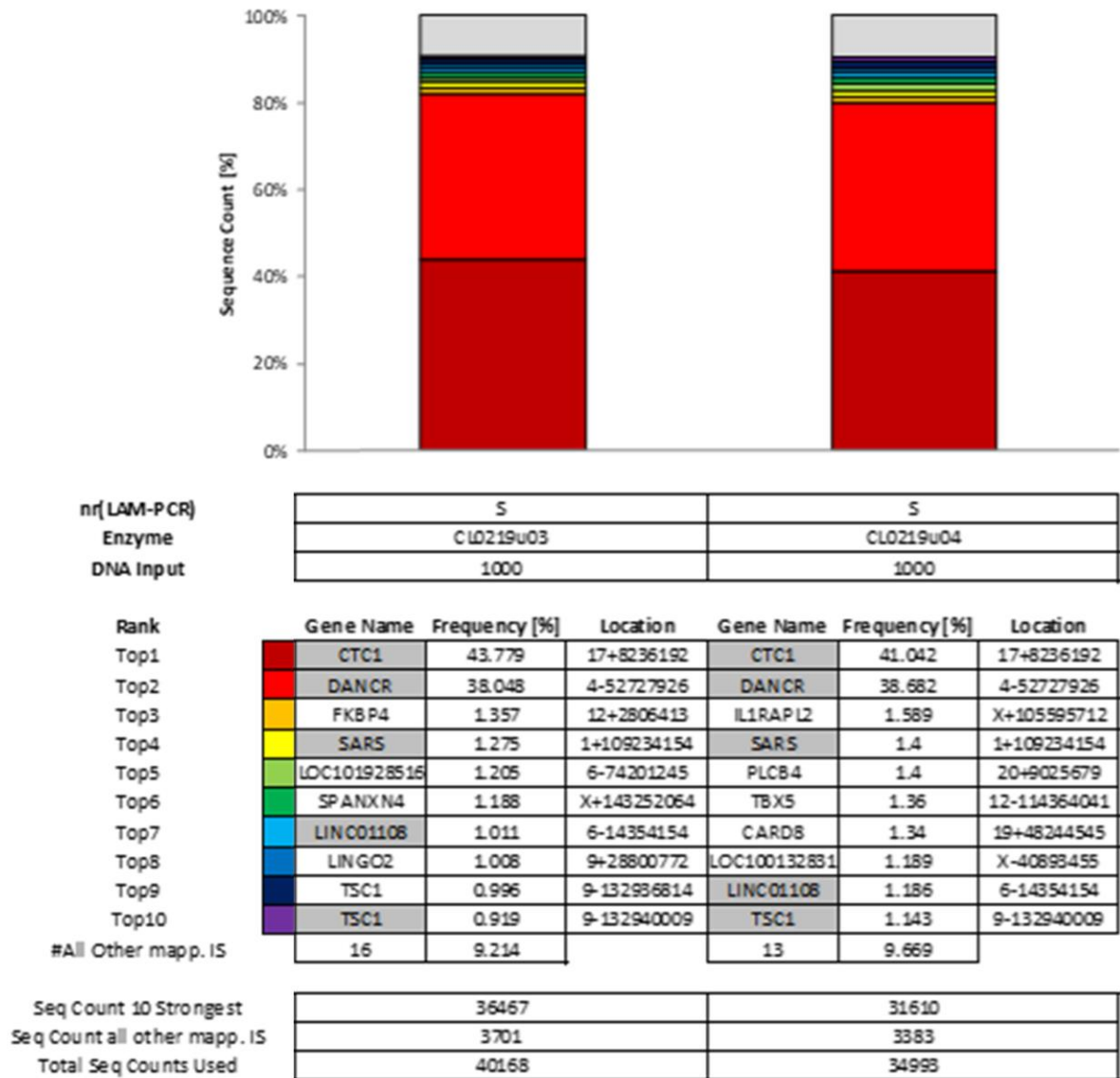


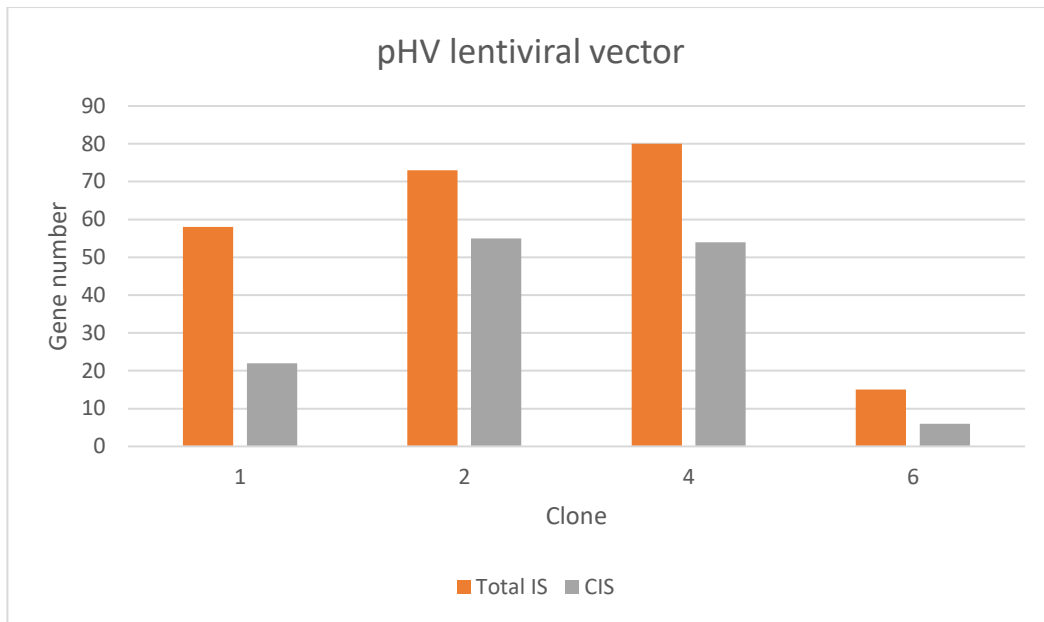
Figure 5.19: Top 10 insertion sites in P106i iPSCs transduced with pHR lentiviral vector. The top 10 IS in pHR clone R, using RSQ analysis. Only one gene was shared between top 10 insertions and was highlighted in grey. Clone Q was identified to a single clone cell clone with one prominent IS.



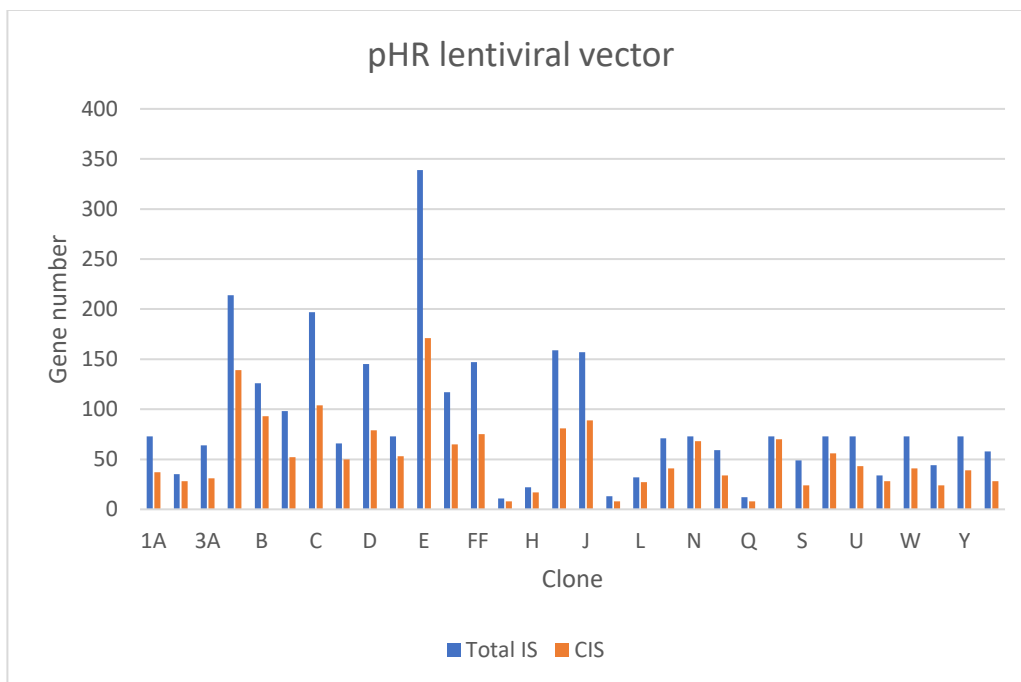


**Figure 5.20: Top 10 insertion sites in P106i iPSCs transduced with pHR lentiviral vector. The top 10 IS in pHR clone S, using RSQ analysis. Four genes were identified to be shared between top 10 IS and were highlighted in grey. The result exhibited polyclonality for iPSCs transduced with pHR lentiviral vector, clone S.**

Further, the quantity of total IS and CIS in P106 clones transduced with pHV and pHR lentiviral vectors were identified. The result exhibited that only one clone transduced with pHV lentiviral vector has <10 CIS and the total 6 CIS were identified (*LINC01249*, *PLCG1.LOC101928223*, *ADAM21P1*, *MGAT4C* and *MARS*). However, in clone G, K and Q transduced with pHR lentiviral vector, 8 CIS were identified (Figures 5.21 & 5.22).



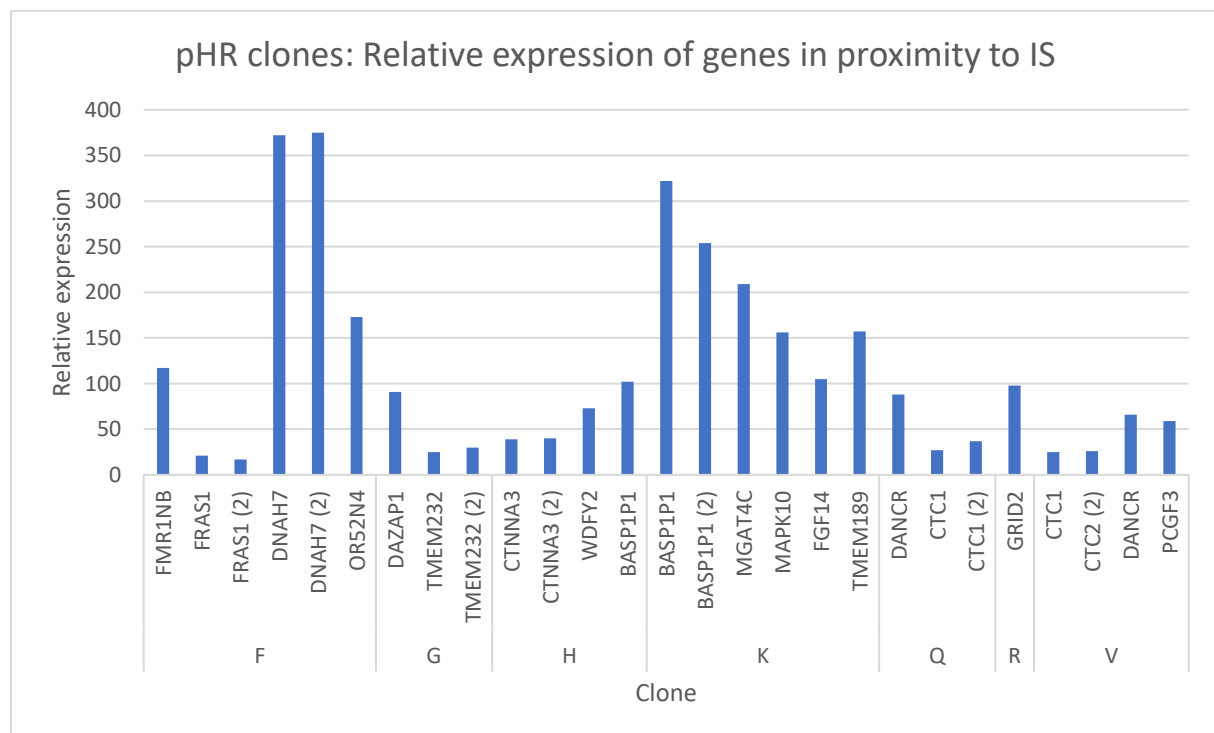
**Figure 5.21: The quantity of total IS and CIS in P106i clones transduced with pHV lentiviral vectors. The result exhibited that only clone 6 has <10 CIS.**



**Figure 5.22: The quantity of total IS and CIS in P106i clones transduced with pHR lentiviral vectors. The result revealed that only in clones G, K and Q <10 CIS were identified.**

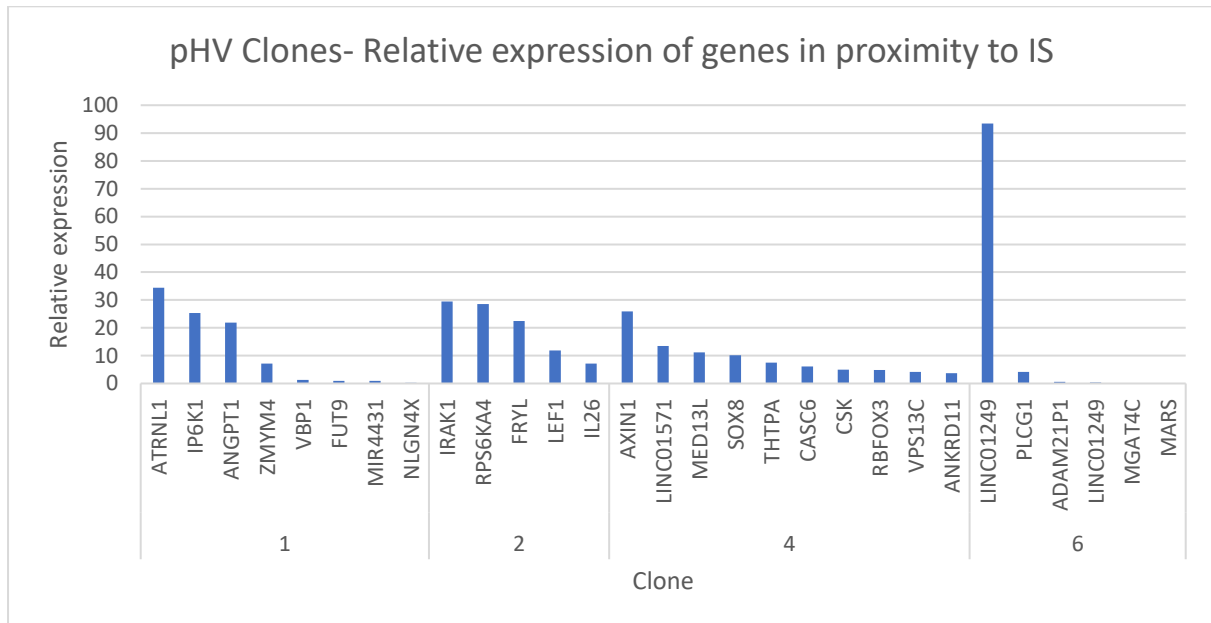
### 5.3.8. Relative gene expression of genes in proximity to the insertion sites in P106i clones

The relative gene expression of P106i clones which were transduced with pHR and pHV lentiviral vectors were shown in (Figures 5.21 & 5.22). Genes in proximity to IS were identified and analysed, using qRT-PCR technique. This technique was used to quantify gene expression. P106i cells transduced with pHR and pHV lentiviral vectors and manually single cell cloned. The gene expression of GAPDH was used as an internal reference to calculate the deltaCt values within each sample. Uninfected P106i cells was used as a reference to calculate the deltadeltaCt values, setting the expression of uninfected cells to “100”. For pHR clones, the expression of eight out of the 20 loci was increased remarkably and no reduction was detected. This result concerns six out of the seven clones analysed (with only one gene analysed for clone R, which did not show a change in expression). Hence, all genes in proximity to IS were found to be upregulated (Figures 5.23 & 5.24). For pHV clones, the expression of genes in proximity to IS in clones 1, 2 and 4 were upregulated. However, in clone 6, only one gene out of 29 loci was increased remarkably.



**Figure 5.23: The relative gene expression of P106i clones transduced with pHR lentiviral vector. Genes in proximity to IS were identified and analysed, using qRT-PCR technique to quantify gene expression. The gene expression of GAPDH was used as**

an internal reference to calculate the deltaCt values within each sample. Uninfected P106i cells was used as a reference to calculate the deltadeltaCt values, setting the expression of uninfected cells to “100”. A technical replication of the analysis was performed for seven samples to confirm the reproducibility of the analysis. All genes identified in proximity to IS were found to be upregulated. Each pHR clones (F, G, H, K, Q, R and V) and their respective genes in proximity to IS are shown. Only one gene analysed for clone R, which did not show a change in expression.



**Figure 5.24:** The relative gene expression of P106i clones transduced with pHV lentiviral vector. Genes in proximity to IS were identified and analysed, using qRT-PCR technique to quantify gene expression. All genes identified in proximity to IS were found to be upregulated. For pHV clones, the expression of genes in proximity to IS in clones 1, 2 and 4 were upregulated. However, in clone 6, only one gene out of 29 loci was increased remarkably. Uninfected P106i cells were used as controls.

## 5.4. Discussion

Gene transfer technology is a powerful method to introduce recombinant genetic elements into human cells. To this end, several types of gene transfer systems were developed. The most common gene transfer technology types were derived from viruses such as RVs, ADVs, AAVs and other non-viral vector systems. One of the requirements for a permanent transgene expression, is the integration of the shuttle vector within the target cell’s chromosomal DNA. However, integrative gene transfer systems have few pitfalls such as risk of IM. This may eventually lead to development

of malignancies. The integration of shuttle vector may interfere with the organisation of the chromosomes. This will induce a series of events which may ultimately lead to cellular transformation and change of cell phenotype. Unfortunately, following administration of retroviral-mediated gene transfer five patients who underwent French and England clinical trials developed leukaemia. For *retroviridae*-derived gene transfer methods (including retroviral and lentiviral vectors) two important mechanisms of vector-induced IM were proposed. The genome structure of *retroviridae* viral vector has two long terminal repeats (LTRs) which are located at the 5' and 3' end of their genome. Each of the LTR segment contains an enhancer and promoter elements. As a result of integration into the target cell chromosomal DNA, the 3' LTR might induce the expression of an endogenous oncogene. This is particularly important if an endogenous oncogene is located in proximity of the insertion site.

The main approach to minimise the incidence of IM is to produce safer integrating gene delivery systems. The engineering of self-inactivating (SIN) *retroviridae*-derived vector systems may reduce the incidence of IM mediated events in transduced cells. SIN *retroviridae*-derived vector systems are generated by a deletion of the U3 region in the 3' LTR segment. Upon viral vector integration, it is crucial to detect IS in the setting of clonal dominance. This information is vital for better understanding of vector-host cell interactions and develop strategies to avoid genotoxicity. Over the years, different methods were explored to detect IS. The general principle of these methods were to identify known sequences in the proviral integrated into the unknown region of genomic DNA. Then the identified junction fragment is isolated, amplified and sequenced.

In chapter 4, pHR and pHV lentiviral vectors were used to transduce P106i cell line and *in vitro* 3D heps. It was shown that the 3D hepatic model can be transduced using lentiviral vectors. In this chapter, to evaluate the genotoxicity risk of pHR and pHV lentiviral vectors, IS analysis, cancer gene identification and gene enrichment analysis were performed. Further, to identify viral insertions and their effect on gene expression in the proximity of vector IS, single cell cloning (SCC) was conducted. In order to see the effect of time on clonality and IS selection, proliferation assay was conducted on P106i cell line at two different timepoints. This was because hiPSCs has the ability to proliferate for a long period of time whilst the rate of proliferation in 3D heps are quite low.

To our knowledge, there is no standard *in vitro* human model available to evaluate the genotoxicity risks associated with lentiviral vectors. For bioinformatic analysis, lentiviral integration sites were retrieved by EPTS/LM-PCR and CIS were detected in P106i and 3D hepatosphere samples. Integration site analysis on bulk P106i and P106i-derived 3D heps were performed. This result revealed an overall decrease in IS overtime in transduced cells for both lentiviral vectors by comparison of the days 3 and 30 time points. The number of identified IS in 3D heps transduced with pHR and pHV lentiviral was low compared to P106i cells. It was previously shown that 3D heps transduced with pHV and pHR lentiviral vectors had low vector copy numbers (VCN) per cell compared to P106i cells, hence a lower transduction efficiency and possibly low number of IS. There was also an apparent reduction in IS in P106i transduced with pHR and pHV lentiviral vectors over time suggesting possible cell death during propagation and culture expansion.

IS analysis was also performed on P106i cells transduced with pHR and pHV lentiviral vectors at two different time points. A comparison of the IS was also carried out between the two timepoints. To do this, the top 10 IS were examined for each time point for P106i cells transduced with pHR and pHV lentiviral vectors. Using the relative sequence count (RSQ) analysis, the number of the counts of integrated lentiviral vectors was assigned. From the top 10 IS analysis in P106i cells transduced with pHR lentiviral vector, only one common gene repeated at day 3 among replicates. However, over time IS analysis revealed 10 genes were repeated among replicates at day 30 timepoint. Similar IS analysis comparison was performed on P106i cells transduced with pHV lentiviral vector. No common genes among replicates were detected at day 3. However, 13 genes were commonly repeated at day 30 timepoint. This result suggests that pHV lentiviral vector with full LTR possibly conferring more proliferation. This data exhibited that pHV lentiviral vector presents a higher number of overlapping IS between two time points. Hence, an increased number of overtimes persisting IS. In addition, minimum level of overlapping common ISs were detected between two vectors and at different time points. This indicates despite similarity between the two lentiviral vectors structural backbone, following transduction insertion sites can be different.

A comparison of the top 10 genes were also performed in P106i-derived 3D heps transduced with pHR and pHV lentiviral vectors. Three common genes (*PACS1*,

*CHD3*, *CYB5D1*) were detected in 3D heps transduced with pHR and pHV lentiviral vectors which were ranked as the top 2 and top 4 identified genes. Other genes were not similar between 3D heps infected with pHR and pHV lentiviral vectors. A similar comparison was also performed between P106i cells and P106i derived 3D heps transduced with pHR and pHV lentiviral vectors. No similar genes/locus were detected between P106i cells and 3D heps transduced with pHR lentiviral vector. However, two genes (*PACS1* and *SF3B2*) were identified in both P106i cells and 3D heps transduced with pHV lentiviral vector. From this data, it can be concluded that different cell and/or tissue types have different ISs.

Following IS analysis, identification of cancer related genes IS and CIS in P106i cells and 3D heps were performed. From the result analysis it can be concluded that in P106i samples transduced with pHR and pHV lentiviral vectors, the number of POs were higher in samples transduced with pHV lentiviral vector. This result suggests the effect of full LTR component of the lentiviral vector. Further, the number of POs were also higher in 3D heps transduced with pHV lentiviral vector. In addition to the IS and CIS analysis, we quantified the viral sequence count over time (30 days) with P106i cells. The result revealed a slight increase in the average sequence count in the P106i samples transduced with pHR lentiviral vectors at the late time point which was not significant. However, in P106i samples transduced with pHV lentiviral vectors there was a significant increase in the average sequence count at the late time point. This further suggests the possible effect of full LTR vector on the average sequence time and also overtime.

The enrichment of IS genes for biological pathways was used to interpret sets of genes and related functional properties. One of the gene enrichment analysis used for this study was to find the biological pathways and their comparison within unique IS. The result revealed that the similar biological pathways were targeted by two lentiviral vectors in P106i cells transduced with pHR and pHV lentiviral vectors at two different timepoints. As explained earlier there was a reduction in IS overtime for both vectors and therefore this might have an effect on biological pathways. The comparison between P106i and 3D heps, the cells transduced with pHR and pHV vectors shared 60% and 80% of the shared biological pathways [ $p < 0.05$ ]. This result exhibited a slight decrease in the number of biological pathways over time, particularly in P106i transduced with pHR lentiviral vector. From IS result analysis of this study, it was

shown that viral insertions were different between vectors, cell types and also at different timepoints. Interestingly, biological pathways analysis revealed both lentiviral vectors target same biological pathways. Further, similar biological pathways persisted over time in P106i cells at different timepoints. It also appears that pHV lentiviral vector was targeting similar biological pathways regardless of cell type. Therefore, it is important to study biological pathways as well as IS analysis for genotoxicity studies.

Finally, to determine the effect of viral insertion at single cell level, single cell cloning was performed in iPSCs transduced with pHR and pHV lentiviral vectors. Both pHV and pHR clones exhibited polyclonality. However, in clone 6 transduced with pHV lentiviral vector the gene *LINC01249* exhibited 93.4% of the viral sequence count. Further, in pHR clone R only one gene revealed the highest percentage of sequence count whilst other clones displayed variety of polyclonality. In addition, despite the polyclonality of the samples, the qPCR result revealed that the genes near to IS have increased in level of expression. To study the impact of single IS in single clones and reducing polyclonality, a lower multiplicity of infection (MOI) and using fluorescence-activated cell sorting (FACS) machine will be attempted in the future. From the result of this study, it can be concluded that *in vitro* 3D hepatic model and P106i cells provide a valuable tool to assess the genotoxicity risks associated with lentiviral vectors. Although the readouts from 3D heps and P106i cells were different, they are comparable. Moreover, the results exhibited that pHV lentiviral vector with full LTR was linked with an increased number of average sequence count, high number of POs. Hence, it is less safe than pHR SIN lentiviral vector.

The relative gene expression of P106i clones which were transduced with pHR and pHV lentiviral vectors were also quantified. Genes in proximity to IS were identified and analysed, using qRT-PCR technique. For pHR clones, the expression of eight out of the 20 loci was increased remarkably and no reduction was detected. In pHR and pHV clones nearly all genes in proximity to IS were found to be upregulated. It would be of interest to check how close the genes analysed are to the IS, and whether the increase in expression correlates inversely with the vicinity. Further, if genes close to the IS expressed more highly than those further away.



## 6. 1 Discussion & Final Remarks

The *in vivo* pre-clinical genotoxicity screening causes enormous animal suffering. Animal screening is expensive, time-consuming, labour-intensive, does not resemble human physiology, and has inadequate predictive power. Currently, no “gold standard” *in vitro* or *in vivo* genotoxicity screening models are available to address the current challenges in the field. The main objective of this study was to develop an *in vitro* genotoxicity platform to screen and map viral vector insertional sites.

Current sources of hepatocytes including primary human hepatocytes, hepatocarcinoma cell lines and animal-derived hepatocytes have potential drawbacks limiting their applications as a tool for *in vitro* assessment of genotoxicity.. The aim of this thesis was to produce 3D *in vitro* hepatospheres from a renewable and scalable source. The hiPSCs are characterised by their unlimited self-renewal capacity and ability to differentiate into three embryonic germ layers and subsequently all cell types found in human body. Therefore, hiPSCs represent a promising source to generate functional hepatocytes. Despite encouraging results from 2D monolayer, generated HLCs and exhibit foetal functions and transient phenotype in culture.

To address the above-mentioned issues, I was initially involved in development of a novel *in vitro* 3D platform for generation of functional hepatocytes.. from hiPSCs. *In vitro* generated 3D heps exhibited stable phenotype for over one year in culture, providing an ideal tool for long-term genotoxicity studies. Following characterisation and differentiation, hiPSCs and 3D heps were used to map IS of different viral vectors. This protocol addresses issues surrounding previous 3D protocols, such as scalability and long-term *in vitro* phenotypic stability.

Here, we demonstrated proof of concept that *in vitro* genotoxicity can be evaluated using a personalised human-based model. This model can be used for genotoxicity screening of viral vectors intended for human gene therapy. For this study, hiPSCs and hiPSCs-derived 3D heps were successfully transduced using lentiviral and AAV vectors. We optimised transduction of hiPSCs and 3D heps by use of “safe” and “unsafe” lentiviral and recombinant AAV serotype-2 vectors. We performed genotoxicity assays including viral vectors IS profiling to show the differences between hiPSCs and 3D heps IS selection. Further, the difference between SIN LV and full LTR safety in lentiviral infected human clones for gene activation was quantified.

There is no safe native lentiviral vector (non-SIN vector) which can be used in gene therapy clinical trials. One of the requirements for a permanent transgene expression, is the integration of the shuttle vector within the target cell's chromosomal DNA. This was evidenced during French and England clinical trials. Unfortunately, following administration of retroviral mediated gene transfer five patients who underwent clinical trials developed leukaemia (Cavazzana-Calvo and Fischer, 2007, Hacein-Bey-Abina et al., 2003). Native lentiviral vectors can integrate into the genome and promote oncogenesis by altering local gene expression. Researchers tried to develop safer vectors that can be used for clinical trials. SIN lentiviral vectors contain a deletion in 3' LTR that prohibits aberrant activation of nearby genes. These vectors have now become a standard in gene therapy and can be used safely for clinical trials. Due to improved safety feature, SIN vectors have emerged as a promising tool for clinical conditions and their application is favoured over native lentiviral vectors. Lentiviral vectors safety was improved by splitting the viral genome into separate plasmids, reducing recombinant virus generation. However, this modification was not robust enough to reduce the possible associated risks. Therefore, SIN vectors with additional safety feature became a better alternative. It is reported that lentiviral vector composed of SIN LTRs possess a 10-fold lower risk of genotoxicity than native counterparts (Montini et al., 2009). Tisagenlecleucel (CTL019, Kymriah) was the first SIN lentiviral vector which approved in the United States in August 2017 for the treatment of paediatric and young adult patients with acute lymphoblastic leukaemia. It is argued SIN vectors have improved safety for gene transfer into stem cells and T cells while one cannot undermine risk of IM of lentiviral vectors (Schuster et al., 2019). As such, SIN vectors are preferred for gene therapy clinical trials.

In this study hiPSCs and 3D heps were successfully transduced with both native and SIN viral vectors. In 3D heps transduced with pHV (native) and pHR (SIN) lentiviral vectors, a higher level of GFP expression was observed in cells transduced with pHR lentiviral vector. For bioinformatic analysis, lentiviral IS was retrieved by EPTS/LM-PCR, and CIS were detected in P106i and 3D heps samples. This result revealed an overall decrease in IS overtime in transduced cells for both lentiviral vectors by comparison to the days 3 and 30 post-transduction. While results are promising further refinements are required to permit technology translation from bench to clinic.

As discussed above the 3D generation of HLCs in a defined microenvironment is a great starting point to improve the technology for genotoxicity screening. The key areas where the current state of art would benefit from are:

- 1) **Improved *in vitro* 3D HLC biochemical characterisation.** Existing differentiation approaches from PSCs generate functional HLCs. These cells exhibit hepatocyte markers and function. However, they still express certain foetal hepatocyte features. In this study, differentiated HLCs expressed cytochrome 3A and ALB. A study conducted by Baxter and colleagues, it was reported that the functionality of these enzymes can also be detected in foetal hepatocytes (Baxter et al., 2015). Therefore, a comparative proteomic and functional analysis can be performed to identify reliable markers of hepatocyte maturity. One study reported that a set of proteins including cytochrome P450 2A6, glutathione S transferase P and alcohol dehydrogenase are regarded as key indicators of hepatocyte differentiation (Rowe et al., 2013).
- 2) ***In vitro* 3D culture optimisation.** Development of novel co-culture strategies with immune compartments will be a more realistic approach for future applications.
- 3) **Single cell cloning optimisation.** This model will benefit from further characterisation and preparing single clones using fluorescence-activated cell sorting (FACS) rather than dilution-based methodology and manual picking of the clones.
- 4) **Cell viability assays.** For this study, trypan blue was used to differentiate viable from dead cells. However, other reliable assays such as MTT and Luminescent ATP assays can be used as an indicator of cell viability and cytotoxicity following transduction with viral vectors. Further, BrdU assay can be applied to monitor cell proliferation following transduction with different MOIs.

Overall, it can be concluded that under 3D culture systems superior cell-to-cell contact and cell matrix interaction can be achieved. These improved 3D interactions are crucial for cell survival, proliferation, and function. In addition, 3D models can more closely recapitulate *in vivo* cellular responses offering physiologically-relevant information. This will bridge the gap between *in silico* hypothesis/*in vitro* results and natural *in vivo* environment.

In addition, phenotypic stability and lack of proliferation are key characteristics required in a genotoxicity assay. Current 2D cultured HLCs lack both of these features while 3D HLCs can be maintained for over a year in culture, therefore, they can be used as a platform for in vitro genotoxicity assays and preferred over conventional 2D monolayer cultures.

## 7. References

- AASEN, T., RAYA, A., BARRERO, M. J., GARRETA, E., CONSIGLIO, A., GONZALEZ, F., VASSENA, R., BILIC, J., PEKARIK, V., TISCORNIA, G., EDEL, M., BOUE, S. & IZPISUA BELMONTE, J. C. 2008. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol*, 26, 1276-84.
- AGARWAL, S., HOLTON, K. L. & LANZA, R. 2008. Efficient differentiation of functional hepatocytes from human embryonic stem cells. *Stem Cells*, 26, 1117-27.
- 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. & BORDIGNON, C. 2007. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. *J Clin Invest*, 117, 2233-40.
- AIUTI, A., CATTANEO, F., GALIMBERTI, S., BENNINGHOFF, U., CASSANI, B., CALLEGARO, L., SCARAMUZZA, S., ANDOLFI, G., MIROLO, M., BRIGIDA, I., TABUCCHI, A., CARLUCCI, F., EIBL, M., AKER, M., SLAVIN, S., AL-MOUSA, H., AL GHONAIUM, A., FERSTER, A., DUPPENTHALER, A., NOTARANGELO, L., WINTERGERST, U., BUCKLEY, R. H., BREGNI, M., MARKTEL, S., VALSECCHI, M. G., ROSSI, P., CICERI, F., MINIERO, R., BORDIGNON, C. & RONCAROLO, M. G. 2009. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N Engl J Med*, 360, 447-58.
- AIUTI, A., COSSU, G., DE FELIPE, P., GALLI, M. C., NARAYANAN, G., RENNER, M., STAHLBOM, A., SCHNEIDER, C. K. & 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. *Hum Gene Ther Clin Dev*, 24, 47-54.
- AIUTI, A., SLAVIN, S., AKER, M., FICARA, F., DEOLA, S., MORTELLARO, A., MORECKI, S., ANDOLFI, G., TABUCCHI, A., CARLUCCI, F., MARINELLO, E., CATTANEO, F., VAI, S., SERVIDA, P., MINIERO, R., RONCAROLO, M. G. & BORDIGNON, C. 2002. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science*, 296, 2410-3.
- AKBARI, S., ARSLAN, N., SENTURK, S. & ERDAL, E. 2019a. Next-Generation Liver Medicine Using Organoid Models. *Front Cell Dev Biol*, 7, 345.
- AKBARI, S., SEVINC, G. G., ERSOY, N., BASAK, O., KAPLAN, K., SEVINC, K., OZEL, E., SENGUN, B., ENUSTUN, E., OZCIMEN, B., BAGRIYANIK, A., ARSLAN, N., ONDER, T. T. & ERDAL, E. 2019b. Robust, Long-Term Culture of Endoderm-Derived Hepatic Organoids for Disease Modeling. *Stem Cell Reports*, 13, 627-641.

- ALBA, R., BOSCH, A. & CHILLON, M. 2005. Gutless adenovirus: last-generation adenovirus for gene therapy. *Gene Ther*, 12 Suppl 1, S18-27.
- ALBALUSHI, H., KUREK, M., KARLSSON, L., LANDREH, L., KJARTANSDOTTIR, K. R., SODER, O., HOVATTA, O. & STUKENBORG, J. B. 2018. Laminin 521 Stabilizes the Pluripotency Expression Pattern of Human Embryonic Stem Cells Initially Derived on Feeder Cells. *Stem Cells Int*, 2018, 7127042.
- ALHAQUE, S., THEMIS, M. & RASHIDI, H. 2018. Three-dimensional cell culture: from evolution to revolution. *Philos Trans R Soc Lond B Biol Sci*, 373.
- ALISON, M. R., CHOONG, C. & LIM, S. 2007. Application of liver stem cells for cell therapy. *Semin Cell Dev Biol*, 18, 819-26.
- AMABILE, G. & MEISSNER, A. 2009. Induced pluripotent stem cells: current progress and potential for regenerative medicine. *Trends Mol Med*, 15, 59-68.
- AMALFITANO, A., HAUSER, M. A., HU, H., SERRA, D., BEGY, C. R. & CHAMBERLAIN, J. S. 1998. Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. *J Virol*, 72, 926-33.
- AOI, T. 2016. 10th anniversary of iPS cells: the challenges that lie ahead. *J Biochem*, 160, 121-9.
- ATCHISON, R. W., CASTO, B. C. & HAMMON, W. M. 1965. Adenovirus-Associated Defective Virus Particles. *Science*, 149, 754-6.
- AUSUBEL, L. J., LOPEZ, P. M. & COUTURE, L. A. 2011. GMP scale-up and banking of pluripotent stem cells for cellular therapy applications. *Methods Mol Biol*, 767, 147-59.
- BADENES, S. M., FERNANDES, T. G., CORDEIRO, C. S., BOUCHER, S., KUNINGER, D., VEMURI, M. C., DIOGO, M. M. & CABRAL, J. M. 2016. Defined Essential 8 Medium and Vitronectin Efficiently Support Scalable Xeno-Free Expansion of Human Induced Pluripotent Stem Cells in Stirred Microcarrier Culture Systems. *PLoS One*, 11, e0151264.
- BAGLEY, J. A., REUMANN, D., BIAN, S., LEVI-STRAUSS, J. & KNOBLICH, J. A. 2017. Fused cerebral organoids model interactions between brain regions. *Nat Methods*, 14, 743-751.
- BAI, M., HARFE, B. & FREIMUTH, P. 1993. Mutations that alter an Arg-Gly-Asp (RGD) sequence in the adenovirus type 2 penton base protein abolish its cell-rounding activity and delay virus reproduction in flat cells. *J Virol*, 67, 5198-205.

- BAKER, B. M. & CHEN, C. S. 2012. Deconstructing the third dimension: how 3D culture microenvironments alter cellular cues. *J Cell Sci*, 125, 3015-24.
- BALAKRISHNAN, B. & JAYANDHARAN, G. R. 2014. Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. *Curr Gene Ther*, 14, 86-100.
- BANK, A. 1996. Human somatic cell gene therapy. *Bioessays*, 18, 999-1007.
- BAR-EPHRAIM, Y. E., KRETZSCHMAR, K. & CLEVERS, H. 2020. Organoids in immunological research. *Nat Rev Immunol*, 20, 279-293.
- BASMA, H., SOTO-GUTIERREZ, A., YANNAM, G. R., LIU, L., ITO, R., YAMAMOTO, T., ELLIS, E., CARSON, S. D., SATO, S., CHEN, Y., MUIRHEAD, D., NAVARRO-ALVAREZ, N., WONG, R. J., ROY-CHOWDHURY, J., PLATT, J. L., MERCER, D. F., MILLER, J. D., STROM, S. C., KOBAYASHI, N. & FOX, I. J. 2009. Differentiation and transplantation of human embryonic stem cell-derived hepatocytes. *Gastroenterology*, 136, 990-9.
- BAUM, C. 2007. What are the consequences of the fourth case? *Mol Ther*, 15, 1401-2.
- 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. & HANLEY, N. A. 2015. Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes. *J Hepatol*, 62, 581-9.
- BEARD, B. C., KEYSER, K. A., TROBRIDGE, G. D., PETERSON, L. J., MILLER, D. G., JACOBS, M., KAUL, R. & KIEM, H. P. 2007. Unique integration profiles in a canine model of long-term repopulating cells transduced with gammaretrovirus, lentivirus, or foamy virus. *Hum Gene Ther*, 18, 423-34.
- BEARD, B. C. & KIEM, H. P. 2009. Canine models of gene-modified hematopoiesis. *Methods Mol Biol*, 506, 341-61.
- BEHNIA, K., BHATIA, S., JASTROMB, N., BALIS, U., SULLIVAN, S., YARMUSH, M. & TONER, M. 2000. Xenobiotic metabolism by cultured primary porcine hepatocytes. *Tissue Eng*, 6, 467-79.
- BENHAMOUCHE, S., DECAENS, T., GODARD, C., CHAMBREY, R., RICKMAN, D. S., MOINARD, C., VASSEUR-COGNET, M., KUO, C. J., KAHN, A., PERRET, C. & COLNOT, S. 2006. Apc tumor suppressor gene is the "zonation-keeper" of mouse liver. *Dev Cell*, 10, 759-70.

- BERG, T., ROUNTREE, C. B., LEE, L., ESTRADA, J., SALA, F. G., CHOE, A., VELTMAAT, J. M., DE LANGHE, S., LEE, R., TSUKAMOTO, H., CROOKS, G. M., BELLUSCI, S. & WANG, K. S. 2007. Fibroblast growth factor 10 is critical for liver growth during embryogenesis and controls hepatoblast survival via beta-catenin activation. *Hepatology*, 46, 1187-97.
- BERGER, D. R., WARE, B. R., DAVIDSON, M. D., ALLSUP, S. R. & KHETANI, S. R. 2015. Enhancing the functional maturity of induced pluripotent stem cell-derived human hepatocytes by controlled presentation of cell-cell interactions in vitro. *Hepatology*, 61, 1370-81.
- BIASCO, L., AMBROSI, A., PELLIN, D., BARTHOLOMAE, C., BRIGIDA, I., RONCAROLO, M. G., DI SERIO, C., VON KALLE, C., SCHMIDT, M. & 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 Mol Med*, 3, 89-101.
- BICEROGLU, S. & MEMIS, A. 2005. Gene therapy: applications in interventional radiology. *Diagn Interv Radiol*, 11, 113-8.
- BISWAS, A. & HUTCHINS, R. 2007. Embryonic stem cells. *Stem Cells Dev*, 16, 213-22.
- 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. & ANDERSON, W. F. 1995. T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science*, 270, 475-80.
- BOKHOVEN, M., STEPHEN, S. L., KNIGHT, S., GEVERS, E. F., ROBINSON, I. C., TAKEUCHI, Y. & COLLINS, M. K. 2009. Insertional gene activation by lentiviral and gammaretroviral vectors. *J Virol*, 83, 283-94.
- BORT, R., SIGNORE, M., TREMBLAY, K., MARTINEZ BARBERA, J. P. & ZARET, K. S. 2006. Hex homeobox gene controls the transition of the endoderm to a pseudostratified, cell emergent epithelium for liver bud development. *Dev Biol*, 290, 44-56.
- BOSSARD, P. & ZARET, K. S. 1998. GATA transcription factors as potentiators of gut endoderm differentiation. *Development*, 125, 4909-17.
- BOZTUG, K., SCHMIDT, M., SCHWARZER, A., BANERJEE, P. P., DIEZ, I. A., DEWEY, R. A., BOHM, M., NOWROUZI, A., BALL, C. R., GLIMM, H., NAUNDORF, S., KUHLCHE, K., BLASCZYK, R., KONDRATENKO, I., MARODI, L., ORANGE, J. S., VON KALLE, C. & KLEIN, C. 2010. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N Engl J Med*, 363, 1918-27.



BRAAM, S. R., ZEINSTRA, L., LITJENS, S., WARD-VAN OOSTWAARD, D., VAN DEN BRINK, S., VAN LAAKE, L., LEBRIN, F., KATS, P., HOCHSTENBACH, R., PASSIER, R., SONNENBERG, A. & MUMMERY, C. L. 2008. Recombinant vitronectin is a functionally defined substrate that supports human embryonic stem cell self-renewal via alphavbeta5 integrin. *Stem Cells*, 26, 2257-65.

BRAUN, C. J., BOZTUG, K., PARUZYNSKI, A., WITZEL, M., SCHWARZER, A., ROTHE, M., MODLICH, U., BEIER, R., GOHRING, G., STEINEMANN, D., FRONZA, R., BALL, C. R., HAEMMERLE, R., NAUNDORF, S., KUHLCHE, K., ROSE, M., FRASER, C., MATHIAS, L., FERRARI, R., ABBOUD, M. R., AL-HERZ, W., KONDRATENKO, I., MARODI, L., GLIMM, H., SCHLEGELBERGER, B., SCHAMBACH, A., ALBERT, M. H., SCHMIDT, M., VON KALLE, C. & KLEIN, C. 2014. Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. *Sci Transl Med*, 6, 227ra33.

BROGUIERE, N., ISENMANN, L., HIRT, C., RINGEL, T., PLACZEK, S., CAVALLI, E., RINGNALDA, F., VILLIGER, L., ZULLIG, R., LEHMANN, R., ROGLER, G., HEIM, M. H., SCHULER, J., ZENOBI-WONG, M. & SCHWANK, G. 2018. Growth of Epithelial Organoids in a Defined Hydrogel. *Adv Mater*, 30, e1801621.

BROLEN, G., SIVERTSSON, L., BJORQUIST, P., ERIKSSON, G., EK, M., SEMB, H., JOHANSSON, I., ANDERSSON, T. B., INGELMAN-SUNDBERG, M. & HEINS, N. 2010. Hepatocyte-like cells derived from human embryonic stem cells specifically via definitive endoderm and a progenitor stage. *J Biotechnol*, 145, 284-94.

BROWN, N., SONG, L., KOLLU, N. R. & HIRSCH, M. L. 2017. Adeno-Associated Virus Vectors and Stem Cells: Friends or Foes? *Hum Gene Ther*, 28, 450-463.

BULLA, G. A. 1997. Hepatocyte nuclear factor-4 prevents silencing of hepatocyte nuclear factor-1 expression in hepatoma x fibroblast cell hybrids. *Nucleic Acids Res*, 25, 2501-8.

BURKE, Z. D., THOWFEEQU, S. & TOSH, D. 2006. Liver specification: a new role for Wnts in liver development. *Curr Biol*, 16, R688-90.

CAI, J., CHEN, J., LIU, Y., MIURA, T., LUO, Y., LORING, J. F., FREED, W. J., RAO, M. S. & ZENG, X. 2006. Assessing self-renewal and differentiation in human embryonic stem cell lines. *Stem Cells*, 24, 516-30.

CAI, J., ZHAO, Y., LIU, Y., YE, F., SONG, Z., QIN, H., MENG, S., CHEN, Y., ZHOU, R., SONG, X., GUO, Y., DING, M. & DENG, H. 2007. Directed differentiation of human embryonic stem cells into functional hepatic cells. *Hepatology*, 45, 1229-39.

CALMONT, A., WANDZIOCH, E., TREMBLAY, K. D., MINOWADA, G., KAESTNER, K. H., MARTIN, G. R. & ZARET, K. S. 2006. An FGF response pathway that mediates hepatic gene induction in embryonic endoderm cells. *Dev Cell*, 11, 339-48.

CAMERON, K., TAN, R., SCHMIDT-HECK, W., CAMPOS, G., LYALL, M. J., WANG, Y., LUCENDO-VILLARIN, B., SZKOLNICKA, D., BATES, N., KIMBER, S. J., HENGSTLER, J. G., GODOY, P., FORBES, S. J. & HAY, D. C. 2015. Recombinant Laminins Drive the Differentiation and Self-Organization of hESC-Derived Hepatocytes. *Stem Cell Reports*, 5, 1250-1262.

CAMP, J. G., SEKINE, K., GERBER, T., LOEFFLER-WIRTH, H., BINDER, H., GAC, M., KANTON, S., KAGEYAMA, J., DAMM, G., SEEHOFER, D., BELICOVA, L., BICKLE, M., BARSACCHI, R., OKUDA, R., YOSHIZAWA, E., KIMURA, M., AYABE, H., TANIGUCHI, H., TAKEBE, T. & TREUTLEIN, B. 2017. Multilineage communication regulates human liver bud development from pluripotency. *Nature*, 546, 533-538.

CANTORE, A., RANZANI, M., BARTHOLOMAE, C. C., VOLPIN, M., VALLE, P. D., SANVITO, F., SERGI, L. S., GALLINA, P., BENEDICENTI, F., BELLINGER, D., RAYMER, R., MERRICKS, E., BELLINTANI, F., MARTIN, S., DOGLIONI, C., D'ANGELO, A., VANDENDRIESSCHE, T., CHUAH, M. K., SCHMIDT, M., NICHOLS, T., MONTINI, E. & NALDINI, L. 2015. Liver-directed lentiviral gene therapy in a dog model of hemophilia B. *Sci Transl Med*, 7, 277ra28.

CAPLEN, N. J., KINRADE, E., SORGI, F., GAO, X., GRUENERT, D., GEDDES, D., COUTELLE, C., HUANG, L., ALTON, E. W. & WILLIAMSON, R. 1995. In vitro liposome-mediated DNA transfection of epithelial cell lines using the cationic liposome DC-Chol/DOPE. *Gene Ther*, 2, 603-13.

CARLSSON, J. & YUHAS, J. M. 1984. Liquid-overlay culture of cellular spheroids. *Recent Results Cancer Res*, 95, 1-23.

CARTER, B. J. 2004. Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective. *Mol Ther*, 10, 981-9.

CARTIER, N., HACEIN-BEY-ABINA, S., BARTHOLOMAE, C. C., VERES, G., SCHMIDT, M., KUTSCHERA, I., VIDAUD, M., ABEL, U., DAL-CORTIVO, L., CACCAVELLI, L., MAHLAOU, N., KIERMER, V., MITTELSTAEDT, D., BELLESME, C., LAHLOU, N., LEFRERE, F., BLANCHE, S., AUDIT, M., PAYEN, E., LÉBOULCH, P., L'HOMME, B., BOUGNERES, P., VON KALLE, C., FISCHER, A., CAVAZZANA-CALVO, M. & AUBOURG, P. 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science*, 326, 818-23.

CASSANI, B., MONTINI, E., MARUGGI, G., AMBROSI, A., MIROLO, M., SELLERI, S., BIRAL, E., FRUGNOLI, I., HERNANDEZ-TRUJILLO, V., DI SERIO, C., RONCAROLO, M. G., NALDINI, L., MAVILIO, F. & AIUTI, A. 2009. Integration of retroviral vectors induces minor changes in the transcriptional activity of T cells from ADA-SCID patients treated with gene therapy. *Blood*, 114, 3546-56.

CAVAZZANA-CALVO, M. & FISCHER, A. 2007. Gene therapy for severe combined immunodeficiency: are we there yet? *J Clin Invest*, 117, 1456-65.

CAVAZZANA-CALVO, M., HACEIN-BEY, S., DE SAINT BASILE, G., GROSS, F., YVON, E., NUSBAUM, P., SELZ, F., HUE, C., CERTAIN, S., CASANOVA, J. L., BOUSSO, P., DEIST, F. L. & FISCHER, A. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science*, 288, 669-72.

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. & MONTINI, E. 2014. Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Mol Ther*, 22, 774-85.

CESANA, D., SGUALDINO, J., RUDILOSSO, L., MERELLA, S., NALDINI, L. & MONTINI, E. 2012. Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. *J Clin Invest*, 122, 1667-76.

CHAMBERS, S. M., FASANO, C. A., PAPAPETROU, E. P., TOMISHIMA, M., SADELAIN, M. & STUDER, L. 2009. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol*, 27, 275-80.

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. & VENDITTI, C. P. 2015. Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. *J Clin Invest*, 125, 870-80.

CHANG, A. H. & SADELAIN, M. 2007. The genetic engineering of hematopoietic stem cells: the rise of lentiviral vectors, the conundrum of the Itr, and the promise of lineage-restricted vectors. *Mol Ther*, 15, 445-56.

CHAPMAN, L. M. & EDDY, E. M. 1989. A Protein Associated with the Mouse and Rat Hepatocyte Junctional Complex. *Cell and Tissue Research*, 257, 333-341.

CHARRIER, S., FERRAND, M., ZERBATO, M., PRECIGOUT, G., VIORNERY, A., BUCHER-LAURENT, S., BENKHELIFA-ZIYYAT, S., MERTEN, O. W., PEREA, J. & GALY, A. 2011. Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. *Gene Ther*, 18, 479-87.

CHEN, G., GULBRANSON, D. R., HOU, Z., BOLIN, J. M., RUOTTI, V., PROBASCO, M. D., SMUGA-OTTO, K., HOWDEN, S. E., DIOL, N. R., PROPSON, N. E., WAGNER, R., LEE, G. O., ANTOSIEWICZ-BOURGET, J., TENG, J. M. & THOMSON, J. A. 2011. Chemically defined conditions for human iPSC derivation and culture. *Nat Methods*, 8, 424-9.

CHENG, L., HAMMOND, H., YE, Z., ZHAN, X. & DRAVID, G. 2003. Human adult marrow cells support prolonged expansion of human embryonic stem cells in culture. *Stem Cells*, 21, 131-42.

- CHENG, N., WAUTHIER, E. & REID, L. M. 2008. Mature human hepatocytes from ex vivo differentiation of alginate-encapsulated hepatoblasts. *Tissue Eng Part A*, 14, 1-7.
- CHENG, W., GUO, L., ZHANG, Z., SOO, H. M., WEN, C., WU, W. & PENG, J. 2006. HNF factors form a network to regulate liver-enriched genes in zebrafish. *Dev Biol*, 294, 482-96.
- CHINEN, J. & PUCK, J. M. 2004. Successes and risks of gene therapy in primary immunodeficiencies. *J Allergy Clin Immunol*, 113, 595-603; quiz 604.
- CHOI, K. D., YU, J., SMUGA-OTTO, K., SALVAGIOTTO, G., REHRAUER, W., VODYANIK, M., THOMSON, J. & SLUKVIN, I. 2009. Hematopoietic and endothelial differentiation of human induced pluripotent stem cells. *Stem Cells*, 27, 559-67.
- CIRILLO, L. A., LIN, F. R., CUESTA, I., FRIEDMAN, D., JARNIK, M. & ZARET, K. S. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell*, 9, 279-89.
- CIUFFI, A., MITCHELL, R. S., HOFFMANN, C., LEIPZIG, J., SHINN, P., ECKER, J. R. & BUSHMAN, F. D. 2006. Integration site selection by HIV-based vectors in dividing and growth-arrested IMR-90 lung fibroblasts. *Mol Ther*, 13, 366-73.
- CLEMENTS, D., CAMELEYRE, I. & WOODLAND, H. R. 2003. Redundant early and overlapping larval roles of Xsox17 subgroup genes in Xenopus endoderm development. *Mech Dev*, 120, 337-48.
- COGGER, V. C., MCNERNEY, G. P., NYUNT, T., DELEVE, L. D., MCCOURT, P., SMEDSROD, B., LE COUTEUR, D. G. & HUSER, T. R. 2010. Three-dimensional structured illumination microscopy of liver sinusoidal endothelial cell fenestrations. *J Struct Biol*, 171, 382-8.
- COLELLA, P., RONZITTI, G. & MINGOZZI, F. 2018. Emerging Issues in AAV-Mediated In Vivo Gene Therapy. *Mol Ther Methods Clin Dev*, 8, 87-104.
- COLLINS, M. & THRASHER, A. 2015. Gene therapy: progress and predictions. *Proc Biol Sci*, 282, 20143003.
- CONDIC, M. L. & RAO, M. 2008. Regulatory issues for personalized pluripotent cells. *Stem Cells*, 26, 2753-8.
- CORRIGAN-CURAY, J., COHEN-HAGUENAUER, O., O'REILLY, M., ROSS, S. R., FAN, H., ROSENBERG, N., SOMIA, N., KING, N., FRIEDMANN, T., DUNBAR, C., AIUTI, A., NALDINI, L., BAUM, C., VON KALLE, C., KIEM, H. P., MONTINI, E., BUSHMAN, F., SORRENTINO, B. P., CARRONDO, M., MALECH, H., GAHRTON, G., SHAPIRO, R., WOLFF, L., ROSENTHAL, E., JAMBOU, R., ZAIA, J. & KOHN, D.

B. 2012. Challenges in vector and trial design using retroviral vectors for long-term gene correction in hematopoietic stem cell gene therapy. *Mol Ther*, 20, 1084-94.

COSTA, R. H., KALINICHENKO, V. V., HOLTERMAN, A. X. & WANG, X. 2003. Transcription factors in liver development, differentiation, and regeneration. *Hepatology*, 38, 1331-47.

CRYSTAL, R. G. 2014. Adenovirus: the first effective in vivo gene delivery vector. *Hum Gene Ther*, 25, 3-11.

D'AMOUR, K. A., AGULNICK, A. D., ELIAZER, S., KELLY, O. G., KROON, E. & BAETGE, E. E. 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol*, 23, 1534-41.

DAVE, U. P., JENKINS, N. A. & COPELAND, N. G. 2004. Gene therapy insertional mutagenesis insights. *Science*, 303, 333.

DAVID, R. M. & DOHERTY, A. T. 2017. Viral Vectors: The Road to Reducing Genotoxicity. *Toxicol Sci*, 155, 315-325.

DE WERT, G. & MUMMERY, C. 2003. Human embryonic stem cells: research, ethics and policy. *Hum Reprod*, 18, 672-82.

DECAENS, C., DURAND, M., GROSSE, B. & CASSIO, D. 2008. Which in vitro models could be best used to study hepatocyte polarity? *Biology of the Cell*, 100, 387-398.

DEICHMANN, A., BRUGMAN, M. H., BARTHOLOMAE, C. C., SCHWARZWAELDER, K., VERSTEGEN, M. M., HOWE, S. J., ARENS, A., OTT, M. G., HOELZER, D., SEGER, R., GREZ, M., HACEIN-BEY-ABINA, S., CAVAZZANA-CALVO, M., FISCHER, A., PARUZYSKI, A., GABRIEL, R., GLIMM, H., ABEL, U., CATTOGLIO, C., MAVILIO, F., CASSANI, B., AIUTI, A., DUNBAR, C. E., BAUM, C., GASPARI, H. B., THRASHER, A. J., VON KALLE, C., SCHMIDT, M. & WAGEMAKER, G. 2011. Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy. *Mol Ther*, 19, 2031-9.

DEUTSCH, G., JUNG, J., ZHENG, M., LORA, J. & ZARET, K. S. 2001. A bipotential precursor population for pancreas and liver within the embryonic endoderm. *Development*, 128, 871-81.

DIAGNOSTIC, C.

DRAPER, J. S., SMITH, K., GOKHALE, P., MOORE, H. D., MALTBY, E., JOHNSON, J., MEISNER, L., ZWAKA, T. P., THOMSON, J. A. & ANDREWS, P. W. 2004. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol*, 22, 53-4.

DU, Y., JENKINS, N. A. & COPELAND, N. G. 2005. Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood*, 106, 3932-9.

DUAN, Y., MA, X., ZOU, W., WANG, C., BAHBAHAN, I. S., AHUJA, T. P., TOLSTIKOV, V. & ZERN, M. A. 2010. Differentiation and characterization of metabolically functioning hepatocytes from human embryonic stem cells. *Stem Cells*, 28, 674-86.

DUFAIT, I., LIECHTENSTEIN, T., LANNA, A., BRICOGNE, C., LARANGA, R., PADELLA, A., BRECKPOT, K. & ESCORS, D. 2012. Retroviral and lentiviral vectors for the induction of immunological tolerance. *Scientifica (Cairo)*, 2012.

DUONG, T. T., LIM, J., VASIREDDY, V., PAPP, T., NGUYEN, H., LEO, L., PAN, J., ZHOU, S., CHEN, H. I., BENNETT, J. & MILLS, J. A. 2019. Comparative AAV-eGFP Transgene Expression Using Vector Serotypes 1-9, 7m8, and 8b in Human Pluripotent Stem Cells, RPEs, and Human and Rat Cortical Neurons. *Stem Cells Int*, 2019, 7281912.

EDELSTEIN, M. L., ABEDI, M. R. & WIXON, J. 2007. Gene therapy clinical trials worldwide to 2007--an update. *J Gene Med*, 9, 833-42.

EDELSTEIN, M. L., ABEDI, M. R., WIXON, J. & EDELSTEIN, R. M. 2004. Gene therapy clinical trials worldwide 1989-2004-an overview. *J Gene Med*, 6, 597-602.

ELAUT, G., HENKENS, T., PAPELEU, P., SNYKERS, S., VINKEN, M., VANHAECKE, T. & ROGIERS, V. 2006. Molecular mechanisms underlying the dedifferentiation process of isolated hepatocytes and their cultures. *Curr Drug Metab*, 7, 629-60.

ELLIS, B. L., HIRSCH, M. L., BARKER, J. C., CONNELLY, J. P., STEININGER, R. J., 3RD & PORTEUS, M. H. 2013. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virology*, 10, 74.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, 154-6.

FELFLY, H. & HADDAD, G. G. 2014. Hematopoietic stem cells: potential new applications for translational medicine. *J Stem Cells*, 9, 163-97.

FELSBURG, P. J., SOMBERG, R. L. & PERRYMAN, L. E. 1992. Domestic animal models of severe combined immunodeficiency: canine X-linked severe combined immunodeficiency and severe combined immunodeficiency in horses. *Immunodeficiency Rev*, 3, 277-303.

FISCHER, A., HACEIN-BEY-ABINA, S. & CAVAZZANA-CALVO, M. 2010. 20 years of gene therapy for SCID. *Nat Immunol*, 11, 457-60.

FLETCHER, J., CUI, W., SAMUEL, K., BLACK, J. R., HANNOUN, Z., CURRIE, I. S., TERRACE, J. D., PAYNE, C., FILIPPI, C., NEWSOME, P., FORBES, S. J., ROSS, J. A., IREDALE, J. P. & HAY, D. C. 2008. The inhibitory role of stromal cell mesenchyme on human embryonic stem cell hepatocyte differentiation is overcome by Wnt3a treatment. *Cloning Stem Cells*, 10, 331-9.

FOLLENZI, A., SABATINO, G., LOMBARDO, A., BOCCACCIO, C. & NALDINI, L. 2002. Efficient gene delivery and targeted expression to hepatocytes in vivo by improved lentiviral vectors. *Hum Gene Ther*, 13, 243-60.

FONG, C. Y., GAUTHAMAN, K. & BONGSO, A. 2010. Teratomas from pluripotent stem cells: A clinical hurdle. *J Cell Biochem*, 111, 769-81.

FORBES, S., VIG, P., POULSOM, R., THOMAS, H. & ALISON, M. 2002. Hepatic stem cells. *Journal of Pathology*, 197, 510-518.

FOX, I. J. & STROM, S. C. 2008. To be or not to be: generation of hepatocytes from cells outside the liver. *Gastroenterology*, 134, 878-81.

FRECHA, C., FUSIL, F., COSSET, F. L. & VERHOEYEN, E. 2011. In vivo gene delivery into hCD34+ cells in a humanized mouse model. *Methods Mol Biol*, 737, 367-90.

FREYER, N., KNOSPEL, F., STRAHL, N., AMINI, L., SCHRADE, P., BACHMANN, S., DAMM, G., SEEHOFER, D., JACOBS, F., MONSHOUWER, M. & ZEILINGER, K. 2016. Hepatic Differentiation of Human Induced Pluripotent Stem Cells in a Perfused Three-Dimensional Multicompartment Bioreactor. *BioResearch Open Access*, 5, 235-248.

FRIEDMAN, S. L. 2008. Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. *Physiol Rev*, 88, 125-72.

FUCITO, A., LUCCHETTI, C., GIORDANO, A. & ROMANO, G. 2008. Genetic and epigenetic alterations in breast cancer: what are the perspectives for clinical practice? *Int J Biochem Cell Biol*, 40, 565-75.

FURUE, M. K., NA, J., JACKSON, J. P., OKAMOTO, T., JONES, M., BAKER, D., HATA, R., MOORE, H. D., SATO, J. D. & ANDREWS, P. W. 2008. Heparin promotes the growth of human embryonic stem cells in a defined serum-free medium. *Proc Natl Acad Sci U S A*, 105, 13409-14.

FUSAKI, N., BAN, H., NISHIYAMA, A., SAEKI, K. & 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. *Proc Jpn Acad Ser B Phys Biol Sci*, 85, 348-62.

GABRIEL, R., ECKENBERG, R., PARUZYNSKI, A., BARTHOLOMAE, C. C., NOWROUZI, A., ARENS, A., HOWE, S. J., RECCHIA, A., CATTOGLIO, C., WANG, W., FABER, K., SCHWARZWAELDER, K., KIRSTEN, R., DEICHMANN, A., BALL, C. R., BALAGGAN, K. S., YANEZ-MUNOZ, R. J., ALI, R. R., GASPAR, H. B., BIASCO, L., AIUTI, A., CESANA, D., MONTINI, E., NALDINI, L., COHEN-HAGUENAUER, O., MAVILIO, F., THRASHER, A. J., GLIMM, H., VON KALLE, C., SAURIN, W. & SCHMIDT, M. 2009. Comprehensive genomic access to vector integration in clinical gene therapy. *Nat Med*, 15, 1431-6.

GAO, G., VANDENBERGHE, L. H., ALVIRA, M. R., LU, Y., CALCEDO, R., ZHOU, X. & WILSON, J. M. 2004. Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol*, 78, 6381-8.

GARDNER, R. L. & ROSSANT, J. 1979. Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J Embryol Exp Morphol*, 52, 141-52.

GASPAR, H. B., PARSLEY, K. L., HOWE, S., KING, D., GILMOUR, K. C., SINCLAIR, J., BROUNS, G., SCHMIDT, M., VON KALLE, C., BARRINGTON, T., JAKOBSEN, M. A., CHRISTENSEN, H. O., AL GHONAIUM, A., WHITE, H. N., SMITH, J. L., LEVINSKY, R. J., ALI, R. R., KINNON, C. & THRASHER, A. J. 2004. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet*, 364, 2181-7.

GAUDIO, E., CARPINO, G., CARDINALE, V., FRANCHITTO, A., ONORI, P. & ALVARO, D. 2009. New insights into liver stem cells. *Digestive and Liver Disease*, 41, 455-462.

GEIS, F. K., GALLA, M., HOFFMANN, D., KUEHLE, J., ZYCHLINSKI, D., MAETZIG, T., SCHOTT, J. W., SCHWARZER, A., GOFFINET, C., GOFF, S. P. & SCHAMBACH, A. 2017. Potent and reversible lentiviral vector restriction in murine induced pluripotent stem cells. *Retrovirology*, 14, 34.

GENOME.GOV.

GERECHT-NIR, S. & ITSKOVITZ-ELDOR, J. 2004. Cell therapy using human embryonic stem cells. *Transpl Immunol*, 12, 203-9.

GIESECK, R. L., 3RD, HANNAN, N. R., BORT, R., HANLEY, N. A., DRAKE, R. A., CAMERON, G. W., WYNN, T. A. & VALLIER, L. 2014. Maturation of induced pluripotent stem cell derived hepatocytes by 3D-culture. *PLoS One*, 9, e86372.

GIESECKE, A. H. 2001. Medical research in the 21st century. *JAMA*, 286, 1833; author reply 1834-5.



GINTER, E. K. 2000. [Gene therapy of hereditary diseases]. *Vopr Med Khim*, 46, 265-78.

GIOBBE, G. G., CROWLEY, C., LUNI, C., CAMPINOTI, S., KHEDR, M., KRETZSCHMAR, K., DE SANTIS, M. M., ZAMBAITI, E., MICHIELIN, F., MERAN, L., HU, Q., VAN SON, G., URBANI, L., MANFREDI, A., GIOMO, M., EATON, S., CACCHIARELLI, D., LI, V. S. W., CLEVERS, H., BONFANTI, P., ELVASSORE, N. & DE COPPI, P. 2019. Extracellular matrix hydrogel derived from decellularized tissues enables endodermal organoid culture. *Nat Commun*, 10, 5658.

GODOY, P., LAKKAPAMU, S., SCHUG, M., BAUER, A., STEWART, J. D., BEDAWI, E., HAMMAD, S., AMIN, J., MARCHAN, R., SCHORMANN, W., MACCOUX, L., VON RECKLINGHAUSEN, I., REIF, R. & HENGSTLER, J. G. 2010. Dexamethasone-dependent versus -independent markers of epithelial to mesenchymal transition in primary hepatocytes. *Biol Chem*, 391, 73-83.

GOESSLING, W., NORTH, T. E., LORD, A. M., CEOL, C., LEE, S., WEIDINGER, G., BOURQUE, C., STRIJBOSCH, R., HARAMIS, A. P., PUDER, M., CLEVERS, H., MOON, R. T. & ZON, L. I. 2008. APC mutant zebrafish uncover a changing temporal requirement for wnt signaling in liver development. *Dev Biol*, 320, 161-74.

GONCALVES, G. A. R. & PAIVA, R. M. A. 2017. Gene therapy: advances, challenges and perspectives. *Einstein (Sao Paulo)*, 15, 369-375.

GOSWAMI, R., SUBRAMANIAN, G., SILAYEVA, L., NEWKIRK, I., DOCTOR, D., CHAWLA, K., CHATTOPADHYAY, S., CHANDRA, D., CHILUKURI, N. & BETAPUDI, V. 2019. Gene Therapy Leaves a Vicious Cycle. *Front Oncol*, 9, 297.

GREEN, J. B. & SMITH, J. C. 1990. Graded changes in dose of a *Xenopus* activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature*, 347, 391-4.

GREGORY, C. A., PROCKOP, D. J. & SPEES, J. L. 2005. Non-hematopoietic bone marrow stem cells: molecular control of expansion and differentiation. *Exp Cell Res*, 306, 330-5.

GRIEGER, J. C. & SAMULSKI, R. J. 2012. Adeno-associated virus vectorology, manufacturing, and clinical applications. *Methods Enzymol*, 507, 229-54.

GRIPON, P., RUMIN, S., URBAN, S., LE SEYEC, J., GLAISE, D., CANNIE, I., GUYOMARD, C., LUCAS, J., TREPO, C. & GUGUEN-GUILLOUZO, C. 2002. Infection of a human hepatoma cell line by hepatitis B virus. *Proc Natl Acad Sci U S A*, 99, 15655-60.

GROSS, G., WAKS, T. & ESHHAR, Z. 1989. Expression of immunoglobulin-T-cell receptor chimeric molecules as functional receptors with antibody-type specificity. *Proc Natl Acad Sci U S A*, 86, 10024-8.

GUALDI, R., BOSSARD, P., ZHENG, M., HAMADA, Y., COLEMAN, J. R. & ZARET, K. S. 1996. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes Dev*, 10, 1670-82.

GUAN, Y., XU, D., GARFIN, P. M., EHMER, U., HURWITZ, M., ENNS, G., MICHIE, S., WU, M., ZHENG, M., NISHIMURA, T., SAGE, J. & PELTZ, G. 2017. Human hepatic organoids for the analysis of human genetic diseases. *JCI Insight*, 2.

GUILLOUZO, A., CORLU, A., ANINAT, C., GLAISE, D., MOREL, F. & GUGUEN-GUILLOUZO, C. 2007. The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact*, 168, 66-73.

GUPTA, R., SCHROODERS, Y., HAUSER, D., VAN HERWIJNEN, M., ALBRECHT, W., TER BRAAK, B., BRECKLINGHAUS, T., CASTELL, J. V., ELENSCHNEIDER, L., ESCHER, S., GUYE, P., HENGSTLER, J. G., GHALLAB, A., HANSEN, T., LEIST, M., MACLENNAN, R., MORITZ, W., TOLOSA, L., TRICOT, T., VERFAILLIE, C., WALKER, P., VAN DE WATER, B., KLEINJANS, J. & CAIMENT, F. 2021. Comparing in vitro human liver models to in vivo human liver using RNA-Seq. *Archives of Toxicology*, 95, 573-589.

HACEIN-BEY-ABINA, S., GARRIGUE, A., WANG, G. P., SOULIER, J., LIM, A., MORILLON, E., CLAPPIER, E., CACCAVELLI, L., DELABESSE, E., BELDJORD, K., ASNAFI, V., MACINTYRE, E., DAL CORTIVO, L., RADFORD, I., BROUSSE, N., SIGAUX, F., MOSHOUS, D., HAUER, J., BORKHARDT, A., BELOHRADSKY, B. H., WINTERGERST, U., VELEZ, M. C., LEIVA, L., SORENSEN, R., WULFFRAAT, N., BLANCHE, S., BUSHMAN, F. D., FISCHER, A. & CAVAZZANA-CALVO, M. 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest*, 118, 3132-42.

HACEIN-BEY-ABINA, S., HAUER, J., LIM, A., PICARD, C., WANG, G. P., BERRY, C. C., MARTINACHE, C., RIEUX-LAUCAT, F., LATOUR, S., BELOHRADSKY, B. H., LEIVA, L., SORENSEN, R., DEBRE, M., CASANOVA, J. L., BLANCHE, S., DURANDY, A., BUSHMAN, F. D., FISCHER, A. & CAVAZZANA-CALVO, M. 2010. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med*, 363, 355-64.

HACEIN-BEY-ABINA, S., LE DEIST, F., CARLIER, F., BOUNEAUD, C., HUE, C., DE VILLARTAY, J. P., THRASHER, A. J., WULFFRAAT, N., SORENSEN, R., DUPUIS-GIROD, S., FISCHER, A., DAVIES, E. G., KUIS, W., LEIVA, L. & CAVAZZANA-CALVO, M. 2002. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med*, 346, 1185-93.

HACEIN-BEY-ABINA, S., VON KALLE, C., SCHMIDT, M., MCCORMACK, M. P., WULFFRAAT, N., LEBOULCH, P., LIM, A., OSBORNE, C. S., PAWLIUK, R., MORILLON, E., SORENSEN, R., FORSTER, A., FRASER, P., COHEN, J. I., DE SAINT BASILE, G., ALEXANDER, I., WINTERGERST, U., FREBOURG, T., AURIAS, A., STOPPA-LYONNET, D., ROMANA, S., RADFORD-WEISS, I., GROSS,

F., VALENSI, F., DELABESSE, E., MACINTYRE, E., SIGAUX, F., SOULIER, J., LEIVA, L. E., WISSLER, M., PRINZ, C., RABBITTS, T. H., LE DEIST, F., FISCHER, A. & CAVAZZANA-CALVO, M. 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, 302, 415-9.

HAN, Y. L., WANG, S., ZHANG, X., LI, Y., HUANG, G., QI, H., PINGGUAN-MURPHY, B., LI, Y., LU, T. J. & XU, F. 2014. Engineering physical microenvironment for stem cell based regenerative medicine. *Drug Discov Today*, 19, 763-73.

HANNAN, N. R., SEGERITZ, C. P., TOUBOUL, T. & VALLIER, L. 2013. Production of hepatocyte-like cells from human pluripotent stem cells. *Nat Protoc*, 8, 430-7.

HANNOUN, Z., FLETCHER, J., GREENHOUGH, S., MEDINE, C., SAMUEL, K., SHARMA, R., PRYDE, A., BLACK, J. R., ROSS, J. A., WILMUT, I., IREDALE, J. P. & HAY, D. C. 2010. The comparison between conditioned media and serum-free media in human embryonic stem cell culture and differentiation. *Cell Reprogram*, 12, 133-40.

HARDY, K., HANDYSIDE, A. H. & WINSTON, R. M. 1989. The human blastocyst: cell number, death and allocation during late preimplantation development in vitro. *Development*, 107, 597-604.

HARDY, K. & SPANOS, S. 2002. Growth factor expression and function in the human and mouse preimplantation embryo. *J Endocrinol*, 172, 221-36.

HARKEY, M. A., KAUL, R., JACOBS, M. A., KURRE, P., BOVEE, D., LEVY, R. & BLAU, C. A. 2007. Multiarm high-throughput integration site detection: limitations of LAM-PCR technology and optimization for clonal analysis. *Stem Cells Dev*, 16, 381-92.

HART, S. N., CUI, Y., KLAASSEN, C. D. & ZHONG, X. B. 2009. Three patterns of cytochrome P450 gene expression during liver maturation in mice. *Drug Metab Dispos*, 37, 116-21.

HART, S. N., LI, Y., NAKAMOTO, K., SUBILEAU, E. A., STEEN, D. & ZHONG, X. B. 2010. A comparison of whole genome gene expression profiles of HepaRG cells and HepG2 cells to primary human hepatocytes and human liver tissues. *Drug Metab Dispos*, 38, 988-94.

HAY, D. C., FLETCHER, J., PAYNE, C., TERRACE, J. D., GALLAGHER, R. C., SNOEYS, J., BLACK, J. R., WOJTACHA, D., SAMUEL, K., HANNOUN, Z., PRYDE, A., FILIPPI, C., CURRIE, I. S., FORBES, S. J., ROSS, J. A., NEWSOME, P. N. & IREDALE, J. P. 2008a. Highly efficient differentiation of hESCs to functional hepatic endoderm requires ActivinA and Wnt3a signaling. *Proc Natl Acad Sci U S A*, 105, 12301-6.

- HAY, D. C., ZHAO, D., FLETCHER, J., HEWITT, Z. A., MCLEAN, D., URRUTICOECHEA-URIGUEN, A., BLACK, J. R., ELCOMBE, C., ROSS, J. A., WOLF, R. & CUI, W. 2008b. Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. *Stem Cells*, 26, 894-902.
- HAY, D. C., ZHAO, D., ROSS, A., MANDALAM, R., LEBKOWSKI, J. & CUI, W. 2007. Direct differentiation of human embryonic stem cells to hepatocyte-like cells exhibiting functional activities. *Cloning Stem Cells*, 9, 51-62.
- HE, Q., OKAJIMA, T., ONOE, H., SUBAGYO, A., SUEOKA, K. & KURIBAYASHI-SHIGETOMI, K. 2018. Origami-based self-folding of co-cultured NIH/3T3 and HepG2 cells into 3D microstructures. *Sci Rep*, 8, 4556.
- HEWITT, N. J., LECHON, M. J., HOUSTON, J. B., HALLIFAX, D., BROWN, H. S., MAUREL, P., KENNA, J. G., GUSTAVSSON, L., LOHMANN, C., SKONBERG, C., GUILLOUZO, A., TUSCHL, G., LI, A. P., LECLUYSE, E., GROOTHUIS, G. M. & HENGSTLER, J. G. 2007. Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies. *Drug Metab Rev*, 39, 159-234.
- HEYWORTH, P. G., CROSS, A. R. & CURNUTTE, J. T. 2003. Chronic granulomatous disease. *Curr Opin Immunol*, 15, 578-84.
- HIRSCH, M. L., FAGAN, B. M., DUMITRU, R., BOWER, J. J., YADAV, S., PORTEUS, M. H., PEVNY, L. H. & SAMULSKI, R. J. 2011. Viral single-strand DNA induces p53-dependent apoptosis in human embryonic stem cells. *PLoS One*, 6, e27520.
- 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., GASPARD, H. B. & THRASHER, A. J. 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest*, 118, 3143-50.
- HU, H., GEHART, H., ARTEGIANI, B., C. L. O.-I., DEKKERS, F., BASAK, O., VAN ES, J., CHUVA DE SOUSA LOPES, S. M., BEGTHEL, H., KORVING, J., VAN DEN BORN, M., ZOU, C., QUIRK, C., CHIRIBOGA, L., RICE, C. M., MA, S., RIOS, A., PETERS, P. J., DE JONG, Y. P. & CLEVERS, H. 2018. Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids. *Cell*, 175, 1591-1606 e19.
- HUANG, H., RUAN, H., AW, M. Y., HUSSAIN, A., GUO, L., GAO, C., QIAN, F., LEUNG, T., SONG, H., KIMELMAN, D., WEN, Z. & PENG, J. 2008. Mypt1-mediated

spatial positioning of Bmp2-producing cells is essential for liver organogenesis. *Development*, 135, 3209-18.

HUANGFU, D., OSAFUNE, K., MAEHR, R., GUO, W., EIJKELENBOOM, A., CHEN, S., MUHLESTEIN, W. & MELTON, D. A. 2008. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol*, 26, 1269-75.

HUCH, M., DORRELL, C., BOJ, S. F., VAN ES, J. H., LI, V. S., VAN DE WETERING, M., SATO, T., HAMER, K., SASAKI, N., FINEGOLD, M. J., HAFT, A., VRIES, R. G., GROMPE, M. & CLEVERS, H. 2013. In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature*, 494, 247-50.

HUCH, M., GEHART, H., VAN BOXTEL, R., HAMER, K., BLOKZIJL, F., VERSTEGEN, M. M., ELLIS, E., VAN WENUM, M., FUCHS, S. A., DE LIGT, J., VAN DE WETERING, M., SASAKI, N., BOERS, S. J., KEMPERMAN, H., DE JONGE, J., IJZERMANS, J. N., NIEUWENHUIS, E. E., HOEKSTRA, R., STROM, S., VRIES, R. R., VAN DER LAAN, L. J., CUPPEN, E. & CLEVERS, H. 2015. Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell*, 160, 299-312.

HURRELL, T., ELLERO, A. A., MASSO, Z. F. & CROMARTY, A. D. 2018. Characterization and reproducibility of HepG2 hanging drop spheroids toxicology in vitro. *Toxicol In Vitro*, 50, 86-94.

IB.COM.AU.

IBRAHEEM, D., ELAISSARI, A. & FESSI, H. 2014. Gene therapy and DNA delivery systems. *Int J Pharm*, 459, 70-83.

INTERNATIONAL STEM CELL INITIATIVE, C., AKOPIAN, V., ANDREWS, P. W., BEIL, S., BENVENISTY, N., BREHM, J., CHRISTIE, M., FORD, A., FOX, V., GOKHALE, P. J., HEALY, L., HOLM, F., HOVATTA, O., KNOWLES, B. B., LUDWIG, T. E., MCKAY, R. D., MIYAZAKI, T., NAKATSUJI, N., OH, S. K., PERA, M. F., ROSSANT, J., STACEY, G. N. & SUEMORI, H. 2010. Comparison of defined culture systems for feeder cell free propagation of human embryonic stem cells. *In Vitro Cell Dev Biol Anim*, 46, 247-58.

ISHII, T., FUKUMITSU, K., YASUCHIKA, K., ADACHI, K., KAWASE, E., SUEMORI, H., NAKATSUJI, N., IKAI, I. & UEMOTO, S. 2008. Effects of extracellular matrixes and growth factors on the hepatic differentiation of human embryonic stem cells. *Am J Physiol Gastrointest Liver Physiol*, 295, G313-21.

ITO, Y., MATSUI, T., KAMIYA, A., KINOSHITA, T. & MIYAJIMA, A. 2000. Retroviral gene transfer of signaling molecules into murine fetal hepatocytes defines distinct roles for the STAT3 and ras pathways during hepatic development. *Hepatology*, 32, 1370-6.

JAY, F. T., LAUGHLIN, C. A. & CARTER, B. J. 1981. Eukaryotic translational control: adeno-associated virus protein synthesis is affected by a mutation in the adenovirus DNA-binding protein. *Proc Natl Acad Sci U S A*, 78, 2927-31.

JENSEN, J., HYLLNER, J. & BJORQUIST, P. 2009. Human embryonic stem cell technologies and drug discovery. *J Cell Physiol*, 219, 513-9.

JOHNSON, M. H., CHISHOLM, J. C., FLEMING, T. P. & HOULISTON, E. 1986. A role for cytoplasmic determinants in the development of the mouse early embryo? *J Embryol Exp Morphol*, 97 Suppl, 97-121.

JOSEPH, A., ZHENG, J. H., CHEN, K., DUTTA, M., CHEN, C., STIEGLER, G., KUNERT, R., FOLLENZI, A. & GOLDSTEIN, H. 2010. Inhibition of in vivo HIV infection in humanized mice by gene therapy of human hematopoietic stem cells with a lentiviral vector encoding a broadly neutralizing anti-HIV antibody. *J Virol*, 84, 6645-53.

JOSSE, R., ANINAT, C., GLAISE, D., DUMONT, J., FESSARD, V., MOREL, F., POUL, J. M., GUGUEN-GUILLOUZO, C. & GUILLOUZO, A. 2008. Long-term functional stability of human HepaRG hepatocytes and use for chronic toxicity and genotoxicity studies. *Drug Metab Dispos*, 36, 1111-8.

JUNG, J., ZHENG, M., GOLDFARB, M. & ZARET, K. S. 1999. Initiation of mammalian liver development from endoderm by fibroblast growth factors. *Science*, 284, 1998-2003.

JUNG, J. H. & KIM, B. S. 2015. A Novel Culture Model for Human Pluripotent Stem Cell Propagation on Gelatin in Placenta-conditioned Media. *J Vis Exp*, e53204.

JUNG, S., PANCHALINGAM, K. M., ROSENBERG, L. & BEHIE, L. A. 2012. Ex vivo expansion of human mesenchymal stem cells in defined serum-free media. *Stem Cells Int*, 2012, 123030.

JUNGERMANN, K. & KATZ, N. 1989. Functional specialization of different hepatocyte populations. *Physiol Rev*, 69, 708-64.

KAMIYA, A. & GONZALEZ, F. J. 2004. TNF-alpha regulates mouse fetal hepatic maturation induced by oncostatin M and extracellular matrices. *Hepatology*, 40, 527-36.

KAMMERER, S. & KUPPER, J. H. 2018. Optimized protocol for induction of cytochrome P450 enzymes 1A2 and 3A4 in human primary-like hepatocyte cell strain HepaFH3 to study in vitro toxicology. *Clin Hemorheol Microcirc*, 70, 563-571.

KANG, E. M., CHOI, U., THEOBALD, N., LINTON, G., LONG PRIEL, D. A., KUHN, D. & MALECH, H. L. 2010. Retrovirus gene therapy for X-linked chronic

granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood*, 115, 783-91.

KAUFMANN, K. B., BUNING, H., GALY, A., SCHAMBACH, A. & GREZ, M. 2013. Gene therapy on the move. *EMBO Mol Med*, 5, 1642-61.

KHAN, I. F., HIRATA, R. K. & RUSSELL, D. W. 2011. AAV-mediated gene targeting methods for human cells. *Nat Protoc*, 6, 482-501.

KIEM, H. P., SELLERS, S., THOMASSON, B., MORRIS, J. C., TISDALE, J. F., HORN, P. A., HEMATTI, P., ADLER, R., KURAMOTO, K., CALMELS, B., BONIFACINO, A., HU, J., VON KALLE, C., SCHMIDT, M., SORRENTINO, B., NIENHUIS, A., BLAU, C. A., ANDREWS, R. G., DONAHUE, R. E. & DUNBAR, C. E. 2004. Long-term clinical and molecular follow-up of large animals receiving retrovirally transduced stem and progenitor cells: no progression to clonal hematopoiesis or leukemia. *Mol Ther*, 9, 389-95.

KIM, J., KOO, B. K. & KNOBLICH, J. A. 2020. Human organoids: model systems for human biology and medicine. *Nat Rev Mol Cell Biol*, 21, 571-584.

KIM, J., KOO, B. K. & YOON, K. J. 2019. Modeling Host-Virus Interactions in Viral Infectious Diseases Using Stem-Cell-Derived Systems and CRISPR/Cas9 Technology. *Viruses*, 11.

KIM, J. J., PARK, B. C., KIDO, Y. & ACCILI, D. 2001. Mitogenic and metabolic effects of type I IGF receptor overexpression in insulin receptor-deficient hepatocytes. *Endocrinology*, 142, 3354-60.

KIM, Y., KANG, K., YOON, S., KIM, J. S., PARK, S. A., KIM, W. D., LEE, S. B., RYU, K. Y., JEONG, J. & CHOI, D. 2018. Prolongation of liver-specific function for primary hepatocytes maintenance in 3D printed architectures. *Organogenesis*, 14, 1-12.

KIM, Y. & RAJAGOPALAN, P. 2010. 3D hepatic cultures simultaneously maintain primary hepatocyte and liver sinusoidal endothelial cell phenotypes. *PLoS One*, 5, e15456.

KNOPP, Y., GEIS, F. K., HECKL, D., HORN, S., NEUMANN, T., KUEHLE, J., MEYER, J., FEHSE, B., BAUM, C., MORGAN, M., MEYER, J., SCHAMBACH, A. & GALLA, M. 2018. Transient Retrovirus-Based CRISPR/Cas9 All-in-One Particles for Efficient, Targeted Gene Knockout. *Mol Ther Nucleic Acids*, 13, 256-274.

KOIKE, H., IWASAWA, K., OUCHI, R., MAEZAWA, M., GIESBRECHT, K., SAIKI, N., FERGUSON, A., KIMURA, M., THOMPSON, W. L., WELLS, J. M., ZORN, A. M. & TAKEBE, T. 2019. Modelling human hepato-biliary-pancreatic organogenesis from the foregut-midgut boundary. *Nature*, 574, 112-116.

- KONOPKA, G., TEKIELA, J., IVERSON, M., WELLS, C. & DUNCAN, S. A. 2007. Junctional adhesion molecule-A is critical for the formation of pseudocanaliculi and modulates E-cadherin expression in hepatic cells. *J Biol Chem*, 282, 28137-48.
- KOOL, J. & BERNIS, A. 2009. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nat Rev Cancer*, 9, 389-99.
- KRAEHENBUEHL, T. P., LANGER, R. & FERREIRA, L. S. 2011. Three-dimensional biomaterials for the study of human pluripotent stem cells. *Nat Methods*, 8, 731-6.
- KUSTIKOVA, O., FEHSE, B., MODLICH, U., YANG, M., DULLMANN, J., KAMINO, K., VON NEUHOFF, N., SCHLEGELBERGER, B., LI, Z. & BAUM, C. 2005. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science*, 308, 1171-4.
- KUSTIKOVA, O. S., BAUM, C. & FEHSE, B. 2008. Retroviral integration site analysis in hematopoietic stem cells. *Methods Mol Biol*, 430, 255-67.
- LAMBSHEAD, J. W., MEAGHER, L., GOODWIN, J., LABONNE, T., NG, E., ELEFANTY, A., STANLEY, E., O'BRIEN, C. M. & LASLETT, A. L. 2018. Long-Term Maintenance of Human Pluripotent Stem Cells on cRGDfK-Presenting Synthetic Surfaces. *Sci Rep*, 8, 701.
- LANCASTER, M. A. & KNOBLICH, J. A. 2014. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science*, 345, 1247125.
- LAUGHLIN, C. A., WESTPHAL, H. & CARTER, B. J. 1979. Spliced adenovirus-associated virus RNA. *Proc Natl Acad Sci U S A*, 76, 5567-71.
- LAWSON, K. A., MENESES, J. J. & PEDERSEN, R. A. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development*, 113, 891-911.
- LAWSON, K. A. & PEDERSEN, R. A. 1987. Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development*, 101, 627-52.
- LECLUYSE, E. L., ALEXANDRE, E., HAMILTON, G. A., VIOLLON-ABADIE, C., COON, D. J., JOLLEY, S. & RICHERT, L. 2005. Isolation and culture of primary human hepatocytes. *Methods Mol Biol*, 290, 207-29.
- LEITE, S. B., TEIXEIRA, A. P., MIRANDA, J. P., TOSTOES, R. M., CLEMENTE, J. J., SOUSA, M. F., CARRONDO, M. J. & ALVES, P. M. 2011. Merging bioreactor technology with 3D hepatocyte-fibroblast culturing approaches: Improved in vitro models for toxicological applications. *Toxicol In Vitro*, 25, 825-32.



LEMAIGRE, F. & ZARET, K. S. 2004. Liver development update: new embryo models, cell lineage control, and morphogenesis. *Curr Opin Genet Dev*, 14, 582-90.

LEONARD, W. J. 1996. The molecular basis of X-linked severe combined immunodeficiency: defective cytokine receptor signaling. *Annu Rev Med*, 47, 229-39.

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. & BUSHMAN, F. D. 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog*, 2, e60.

LI, J., NING, G. & DUNCAN, S. A. 2000. Mammalian hepatocyte differentiation requires the transcription factor HNF-4alpha. *Genes Dev*, 14, 464-74.

LI, Y., POWELL, S., BRUNETTE, E., LEBKOWSKI, J. & MANDALAM, R. 2005. Expansion of human embryonic stem cells in defined serum-free medium devoid of animal-derived products. *Biotechnol Bioeng*, 91, 688-98.

LINDEN, R. Gene therapy : what it is , what it is not and what it will be. 2013.

LISOWSKI, L., DANE, A. P., CHU, K., ZHANG, Y., CUNNINGHAM, S. C., WILSON, E. M., NYGAARD, S., GROMPE, M., ALEXANDER, I. E. & KAY, M. A. 2014. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature*, 506, 382-6.

LIU, G., DAVID, B. T., TRAWCZYNSKI, M. & FESSLER, R. G. 2020. Advances in Pluripotent Stem Cells: History, Mechanisms, Technologies, and Applications. *Stem Cell Rev Rep*, 16, 3-32.

LIU, Q. & MURUVE, D. A. 2003. Molecular basis of the inflammatory response to adenovirus vectors. *Gene Ther*, 10, 935-40.

LIU, Y., HANGOC, G., CAMPBELL, T. B., GOODMAN, M., TAO, W., POLLOK, K., SROUR, E. F. & BROXMEYER, H. E. 2008. Identification of parameters required for efficient lentiviral vector transduction and engraftment of human cord blood CD34(+) NOD/SCID-repopulating cells. *Exp Hematol*, 36, 947-56.

LIU, Y., SONG, Z., ZHAO, Y., QIN, H., CAI, J., ZHANG, H., YU, T., JIANG, S., WANG, G., DING, M. & DENG, H. 2006. A novel chemical-defined medium with bFGF and N2B27 supplements supports undifferentiated growth in human embryonic stem cells. *Biochem Biophys Res Commun*, 346, 131-9.

LIU, Z., CAI, Y., WANG, Y., NIE, Y., ZHANG, C., XU, Y., ZHANG, X., LU, Y., WANG, Z., POO, M. & SUN, Q. 2018. Cloning of Macaque Monkeys by Somatic Cell Nuclear Transfer. *Cell*, 174, 245.

LLAMES, S., GARCIA-PEREZ, E., MEANA, A., LARCHER, F. & DEL RIO, M. 2015. Feeder Layer Cell Actions and Applications. *Tissue Eng Part B Rev*, 21, 345-53.

LOH, K. M., ANG, L. T., ZHANG, J., KUMAR, V., ANG, J., AUYEONG, J. Q., LEE, K. L., CHOO, S. H., LIM, C. Y., NICHANE, M., TAN, J., NOGHABI, M. S., AZZOLA, L., NG, E. S., DURRUTHY-DURRUTHY, J., SEBASTIANO, V., POELLINGER, L., ELEFANTY, A. G., STANLEY, E. G., CHEN, Q., PRABHAKAR, S., WEISSMAN, I. L. & LIM, B. 2014. Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell*, 14, 237-52.

LORENZINI, S., BIRD, T. G., BOULTER, L., BELLAMY, C., SAMUEL, K., AUCOTT, R., CLAYTON, E., ANDREONE, P., BERNARDI, M., GOLDING, M., ALISON, M. R., IREDALE, J. P. & FORBES, S. J. 2010. Characterisation of a stereotypical cellular and extracellular adult liver progenitor cell niche in rodents and diseased human liver. *Gut*, 59, 645-654.

LU, J., HOU, R., BOOTH, C. J., YANG, S. H. & SNYDER, M. 2006. Defined culture conditions of human embryonic stem cells. *Proc Natl Acad Sci U S A*, 103, 5688-93.

LUBBERSTEDT, M., MULLER-VIEIRA, U., MAYER, M., BIEMEL, K. M., KNOSPEL, F., KNOBELOCH, D., NUSSLER, A. K., GERLACH, J. C. & ZEILINGER, K. 2011. HepaRG human hepatic cell line utility as a surrogate for primary human hepatocytes in drug metabolism assessment in vitro. *J Pharmacol Toxicol Methods*, 63, 59-68.

LUCENDO-VILLARIN, B., RASHIDI, H., ALHAQUE, S., FISCHER, L., MESEGUER-RIPOLLES, J., WANG, Y., O'FARRELLY, C., THEMIS, M. & HAY, D. C. 2019. Serum Free Production of Three-dimensional Human Hepatospheres from Pluripotent Stem Cells. *J Vis Exp*.

LUDWIG, T. E., BERGENDAHL, V., LEVENSTEIN, M. E., YU, J., PROBASCO, M. D. & THOMSON, J. A. 2006. Feeder-independent culture of human embryonic stem cells. *Nat Methods*, 3, 637-46.

LUND, A. H., TURNER, G., TRUBETSKOY, A., VERHOEVEN, E., WIENTJENS, E., HULSMAN, D., RUSSELL, R., DEPINHO, R. A., LENZ, J. & VAN LOHUIZEN, M. 2002. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nat Genet*, 32, 160-5.

MAETZIG, T., BRUGMAN, M. H., BARTELS, S., HEINZ, N., KUSTIKOVA, O. S., MODLICH, U., LI, Z., GALLA, M., SCHIEDLMEIER, B., SCHAMBACH, A. & BAUM, C. 2011a. Polyclonal fluctuation of lentiviral vector-transduced and expanded murine hematopoietic stem cells. *Blood*, 117, 3053-64.

MAETZIG, T., GALLA, M., BAUM, C. & SCHAMBACH, A. 2011b. Gammaretroviral vectors: biology, technology and application. *Viruses*, 3, 677-713.

- MAHERALI, N., SRIDHARAN, R., XIE, W., UTIKAL, J., EMINLI, S., ARNOLD, K., STADTFELD, M., YACHECHKO, R., TCHIEU, J., JAENISCH, R., PLATH, K. & HOCHEDLINGER, K. 2007. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell*, 1, 55-70.
- MALI, S. 2013. Delivery systems for gene therapy. *Indian J Hum Genet*, 19, 3-8.
- MARGAGLIOTTI, S., CLOTMAN, F., PIERREUX, C. E., BEAUDRY, J. B., JACQUEMIN, P., ROUSSEAU, G. G. & LEMAIGRE, F. P. 2007. The Onecut transcription factors HNF-6/OC-1 and OC-2 regulate early liver expansion by controlling hepatoblast migration. *Developmental Biology*, 311, 579-589.
- MARINHO, P. A., FERNANDES, A. M., CRUZ, J. C., REHEN, S. K. & CASTILHO, L. R. 2010. Maintenance of pluripotency in mouse embryonic stem cells cultivated in stirred microcarrier cultures. *Biotechnol Prog*, 26, 548-55.
- MARRAFFINI, L. A. & SONTHEIMER, E. J. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*, 11, 181-90.
- MARTIN, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A*, 78, 7634-8.
- MARTIN, M. J., MUOTRI, A., GAGE, F. & VARKI, A. 2005. Human embryonic stem cells express an immunogenic nonhuman sialic acid. *Nat Med*, 11, 228-32.
- MARTINEZ, A. & PASCUAL, E. 1995. Intra and Extracellular Cppd Crystals Are Regularly Found in the Synovial-Fluid of Uninflamed Joints of Cppd Related Arthropathy. *Arthritis and Rheumatism*, 38, 557-557.
- MATSUMOTO, K., YOSHITOMI, H., ROSSANT, J. & ZARET, K. S. 2001. Liver organogenesis promoted by endothelial cells prior to vascular function. *Science*, 294, 559-63.
- MATTHEWS, Q. L. & CURIEL, D. T. 2007. Gene therapy: human germline genetic modifications--assessing the scientific, socioethical, and religious issues. *South Med J*, 100, 98-100.
- MCCORMACK, M. P. & RABBITTS, T. H. 2004. Activation of the T-cell oncogene LMO2 after gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med*, 350, 913-22.
- MCCRIGHT, B., LOZIER, J. & GRIDLEY, T. 2002. A mouse model of Alagille syndrome: Notch2 as a genetic modifier of Jag1 haploinsufficiency. *Development*, 129, 1075-82.

- MCCUSKEY, R. S. 2008. The hepatic microvascular system in health and its response to toxicants. *Anat Rec (Hoboken)*, 291, 661-71.
- MCDONNELL, W. M. & ASKARI, F. K. 1996. DNA vaccines. *N Engl J Med*, 334, 42-5.
- MCLIN, V. A., RANKIN, S. A. & ZORN, A. M. 2007. Repression of Wnt/beta-catenin signaling in the anterior endoderm is essential for liver and pancreas development. *Development*, 134, 2207-17.
- METAIS, J. Y. & DUNBAR, C. E. 2008. The MDS1-EVI1 gene complex as a retrovirus integration site: impact on behavior of hematopoietic cells and implications for gene therapy. *Mol Ther*, 16, 439-49.
- MISRA, S. 2013. Human gene therapy: a brief overview of the genetic revolution. *J Assoc Physicians India*, 61, 127-33.
- MITCHELL, R. S., BEITZEL, B. F., SCHRODER, A. R., SHINN, P., CHEN, H., BERRY, C. C., ECKER, J. R. & BUSHMAN, F. D. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol*, 2, E234.
- MITSUI, K., SUZUKI, K., AIZAWA, E., KAWASE, E., SUEMORI, H., NAKATSUJI, N. & MITANI, K. 2009. Gene targeting in human pluripotent stem cells with adeno-associated virus vectors. *Biochem Biophys Res Commun*, 388, 711-7.
- MODLICH, U. & BAUM, C. 2009. Preventing and exploiting the oncogenic potential of integrating gene vectors. *J Clin Invest*, 119, 755-8.
- MODLICH, U., BOHNE, J., SCHMIDT, M., VON KALLE, C., KNOSS, S., SCHAMBACH, A. & BAUM, C. 2006. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood*, 108, 2545-53.
- MODLICH, U., KUSTIKOVA, O. S., SCHMIDT, M., RUDOLPH, C., MEYER, J., LI, Z., KAMINO, K., VON NEUHOFF, N., SCHLEGELBERGER, B., KUEHLCKE, K., BUNTING, K. D., SCHMIDT, S., DEICHMANN, A., VON KALLE, C., FEHSE, B. & BAUM, C. 2005. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood*, 105, 4235-46.
- MODLICH, U., NAVARRO, S., ZYCHLINSKI, D., MAETZIG, T., KNOESS, S., BRUGMAN, M. H., SCHAMBACH, A., CHARRIER, S., GALY, A., THRASHER, A. J., BUEREN, J. & BAUM, C. 2009. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther*, 17, 1919-28.
- MODLICH, U., SCHAMBACH, A., BRUGMAN, M. H., WICKE, D. C., KNOESS, S., LI, Z., MAETZIG, T., RUDOLPH, C., SCHLEGELBERGER, B. & BAUM, C. 2008. Leukemia induction after a single retroviral vector insertion in Evi1 or Prdm16. *Leukemia*, 22, 1519-28.

MONGA, S. P., MONGA, H. K., TAN, X., MULE, K., PEDIADITAKIS, P. & MICHALOPOULOS, G. K. 2003. Beta-catenin antisense studies in embryonic liver cultures: role in proliferation, apoptosis, and lineage specification. *Gastroenterology*, 124, 202-16.

MONTINI, E. & CESANA, D. 2012. Genotoxicity assay for gene therapy vectors in tumor prone *Cdkn2a(-)/(-)* mice. *Methods Enzymol*, 507, 171-85.

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. & 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. *J Clin Invest*, 119, 964-75.

MONTINI, E., CESANA, D., SCHMIDT, M., SANVITO, F., PONZONI, M., BARTHOLOMAE, C., SERGI SERGI, L., BENEDICENTI, F., AMBROSI, A., DI SERIO, C., DOGLIONI, C., VON KALLE, C. & NALDINI, L. 2006. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol*, 24, 687-96.

MOSER, H. W., MAHMOOD, A. & RAYMOND, G. V. 2007. X-linked adrenoleukodystrophy. *Nat Clin Pract Neurol*, 3, 140-51.

MUELLER, P. R. & WOLD, B. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science*, 246, 780-6.

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. & 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, 2563-9.

NAKAGAWA, M., KOYANAGI, M., TANABE, K., TAKAHASHI, K., ICHISAKA, T., AOI, T., OKITA, K., MOCHIDUKI, Y., TAKIZAWA, N. & YAMANAKA, S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol*, 26, 101-6.

NAKAJIMA, F., TOKUNAGA, K. & NAKATSUJI, N. 2007. Human leukocyte antigen matching estimations in a hypothetical bank of human embryonic stem cell lines in the Japanese population for use in cell transplantation therapy. *Stem Cells*, 25, 983-5.

NAPOLITANO, A. P., CHAI, P., DEAN, D. M. & MORGAN, J. R. 2007. Dynamics of the self-assembly of complex cellular aggregates on micromolded nonadhesive hydrogels. *Tissue Eng*, 13, 2087-94.

NARAYANAN, G., COSSU, G., GALLI, M. C., FLORY, E., OVELGONNE, H., SALMIKANGAS, P., SCHNEIDER, C. K. & TROUVIN, J. H. 2014. Clinical development of gene therapy needs a tailored approach: a regulatory perspective from the European Union. *Hum Gene Ther Clin Dev*, 25, 1-6.

NARSINH, K. H. & WU, J. C. 2010. Gene correction in human embryonic and induced pluripotent stem cells: promises and challenges ahead. *Mol Ther*, 18, 1061-3.

NATHWANI, A. C., NIENHUIS, A. W. & DAVIDOFF, A. M. 2014. Our journey to successful gene therapy for hemophilia B. *Hum Gene Ther*, 25, 923-6.

NEMEROW, G. R., STEWART, P. L. & REDDY, V. S. 2012. Structure of human adenovirus. *Curr Opin Virol*, 2, 115-21.

NGUYEN, T. H. & FERRY, N. 2004. Liver gene therapy: advances and hurdles. *Gene Ther*, 11 Suppl 1, S76-84.

NIEDERER, H. A. & BANGHAM, C. R. 2014. Integration site and clonal expansion in human chronic retroviral infection and gene therapy. *Viruses*, 6, 4140-64.

NIENHUIS, A. W., DUNBAR, C. E. & SORRENTINO, B. P. 2006. Genotoxicity of retroviral integration in hematopoietic cells. *Mol Ther*, 13, 1031-49.

NOEL, G., LE VEE, M., MOREAU, A., STIEGER, B., PARMENTIER, Y. & FARDEL, O. 2013. Functional expression and regulation of drug transporters in monolayer- and sandwich-cultured mouse hepatocytes. *Eur J Pharm Sci*, 49, 39-50.

NORDIN, N., LAI, M. I., VEERAKUMARASIVAM, A., RAMASAMY, R., ABDULLAH, S., WENDY-YEO, W. Y. & ROSLI, R. 2011. Induced pluripotent stem cells: history, properties and potential applications. *Med J Malaysia*, 66, 4-9.

NOWROUZI, A., CHEUNG, W. T., LI, T., ZHANG, X., ARENS, A., PARUZYNSKI, A., WADDINGTON, S. N., OSEJINDU, E., REJA, S., VON KALLE, C., WANG, Y., AL-ALLAF, F., GREGORY, L., THEMIS, M., HOLDER, M., DIGHE, N., RUTHE, A., BUCKLEY, S. M., BIGGER, B., MONTINI, E., THRASHER, A. J., ANDREWS, R., ROBERTS, T. P., NEWBOLD, R. F., COUTELLE, C., SCHMIDT, M. & THEMIS, M. 2013. The fetal mouse is a sensitive genotoxicity model that exposes lentiviral-associated mutagenesis resulting in liver oncogenesis. *Mol Ther*, 21, 324-37.

OKITA, K., MATSUMURA, Y., SATO, Y., OKADA, A., MORIZANE, A., OKAMOTO, S., HONG, H., NAKAGAWA, M., TANABE, K., TEZUKA, K., SHIBATA, T., KUNISADA, T., TAKAHASHI, M., TAKAHASHI, J., SAJI, H. & YAMANAKA, S. 2011. A more efficient method to generate integration-free human iPS cells. *Nat Methods*, 8, 409-12.

OOI, A., WONG, A., ESAU, L., LEMTIRI-CHLIEH, F. & GEHRING, C. 2016. A Guide to Transient Expression of Membrane Proteins in HEK-293 Cells for Functional Characterization. *Front Physiol*, 7, 300.

OSAFUNE, K., CARON, L., BOROWIAK, M., MARTINEZ, R. J., FITZ-GERALD, C. S., SATO, Y., COWAN, C. A., CHIEN, K. R. & MELTON, D. A. 2008. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol*, 26, 313-5.

OTT, M. G., SCHMIDT, M., SCHWARZWAELDER, K., STEIN, S., SILER, U., KOEHL, U., GLIMM, H., KUHLCHE, 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. & GREZ, M. 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med*, 12, 401-9.

PAGE, J. L., JOHNSON, M. C., OLSAVSKY, K. M., STROM, S. C., ZARBL, H. & OMIECINSKI, C. J. 2007. Gene expression profiling of extracellular matrix as an effector of human hepatocyte phenotype in primary cell culture. *Toxicol Sci*, 97, 384-97.

PARK, B., YOO, K. H. & KIM, C. 2015. Hematopoietic stem cell expansion and generation: the ways to make a breakthrough. *Blood Res*, 50, 194-203.

PARK, I. H., ARORA, N., HUO, H., MAHERALI, N., AHFELDT, T., SHIMAMURA, A., LENSCH, M. W., COWAN, C., HOCHEDLINGER, K. & DALEY, G. Q. 2008. Disease-specific induced pluripotent stem cells. *Cell*, 134, 877-86.

PAYNE, C. M., SAMUEL, K., PRYDE, A., KING, J., BROWNSTEIN, D., SCHRADER, J., MEDINE, C. N., FORBES, S. J., IREDALE, J. P., NEWSOME, P. N. & HAY, D. C. 2011. Persistence of functional hepatocyte-like cells in immune-compromised mice. *Liver Int*, 31, 254-62.

PIKE-OVERZET, K., VAN DER BURG, M., WAGEMAKER, G., VAN DONGEN, J. J. & STAAL, F. J. 2007. New insights and unresolved issues regarding insertional mutagenesis in X-linked SCID gene therapy. *Mol Ther*, 15, 1910-6.

PILLAY, S., MEYER, N. L., PUSCHNIK, A. S., DAVULCU, O., DIEP, J., ISHIKAWA, Y., JAE, L. T., WOVEN, J. E., NAGAMINE, C. M., CHAPMAN, M. S. & CARETTE, J. E. 2016. An essential receptor for adeno-associated virus infection. *Nature*, 530, 108-12.

PLANK, C., TANG, M. X., WOLFE, A. R. & SZOKA, F. C., JR. 1999. Branched cationic peptides for gene delivery: role of type and number of cationic residues in formation and in vitro activity of DNA polyplexes. *Hum Gene Ther*, 10, 319-32.

PONTOGLIO, M., BARRA, J., HADCHOUËL, M., DOYEN, A., KRESS, C., BACH, J. P., BABINET, C. & YANIV, M. 1996. Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome. *Cell*, 84, 575-85.

POTTER, S. S. 2018. Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol*, 14, 479-492.

PRIOR, N., INACIO, P. & HUCH, M. 2019. Liver organoids: from basic research to therapeutic applications. *Gut*, 68, 2228-2237.

RAHMAN, S., MAGNUSSEN, M., LEON, T. E., FARAH, N., LI, Z., ABRAHAM, B. J., ALAPI, K. Z., MITCHELL, R. J., NAUGHTON, T., FIELDING, A. K., PIZZEY, A., BUSTRAAN, S., ALLEN, C., POPA, T., PIKE-OVERZET, K., GARCIA-PEREZ, L., GALE, R. E., LINCH, D. C., STAAL, F. J. T., YOUNG, R. A., LOOK, A. T. & MANSOUR, M. R. 2017. Activation of the LMO2 oncogene through a somatically acquired neomorphic promoter in T-cell acute lymphoblastic leukemia. *Blood*, 129, 3221-3226.

RAPTI, K., STILLITANO, F., KARAKIKES, I., NONNENMACHER, M., WEBER, T., HULOT, J. S. & HAJJAR, R. J. 2015. Effectiveness of gene delivery systems for pluripotent and differentiated cells. *Mol Ther Methods Clin Dev*, 2, 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. & HAY, D. C. 2018. 3D human liver tissue from pluripotent stem cells displays stable phenotype in vitro and supports compromised liver function in vivo. *Arch Toxicol*, 92, 3117-3129.

RAUSCHHUBER, C., NOSKE, N. & EHRHARDT, A. 2012. New insights into stability of recombinant adenovirus vector genomes in mammalian cells. *Eur J Cell Biol*, 91, 2-9.

REUBINOFF, B. E., PERA, M. F., FONG, C. Y., TROUNSON, A. & BONGSO, A. 2000. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat Biotechnol*, 18, 399-404.

RICHARDS, M., FONG, C. Y., CHAN, W. K., WONG, P. C. & BONGSO, A. 2002. Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells. *Nat Biotechnol*, 20, 933-6.

RODIN, S., ANTONSSON, L., NIAUDET, C., SIMONSON, O. E., SALMELA, E., HANSSON, E. M., DOMOGATSKAYA, A., XIAO, Z., DAMDIMOPOULOU, P., SHEIKHI, M., INZUNZA, J., NILSSON, A. S., BAKER, D., KUIPER, R., SUN, Y., BLENNOW, E., NORDENSKJOLD, M., GRINNEMO, K. H., KERE, J., BETSHOLTZ, C., HOVATTA, O. & TRYGGVASON, K. 2014. Clonal culturing of human embryonic



stem cells on laminin-521/E-cadherin matrix in defined and xeno-free environment. *Nat Commun*, 5, 3195.

RODIN, S., DOMOGATSKAYA, A., STROM, S., HANSSON, E. M., CHIEN, K. R., INZUNZA, J., HOVATTA, O. & TRYGGVASON, K. 2010. Long-term self-renewal of human pluripotent stem cells on human recombinant laminin-511. *Nat Biotechnol*, 28, 611-5.

RODRIGUEZ-ANTONA, C., DONATO, M. T., BOOBIS, A., EDWARDS, R. J., WATTS, P. S., CASTELL, J. V. & GOMEZ-LECHON, M. J. 2002. Cytochrome P450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells. *Xenobiotica*, 32, 505-20.

ROMANO, G. 2003. Gene transfer in experimental medicine. *Drug News Perspect*, 16, 267-76.

ROMANO, G. 2005a. Current development of adeno-associated viral vectors. *Drug News Perspect*, 18, 311-6.

ROMANO, G. 2005b. Current development of lentiviral-mediated gene transfer. *Drug News Perspect*, 18, 128-34.

ROMANO, G. 2006. The controversial role of adenoviral-derived vectors in gene therapy programs: where do we stand? *Drug News Perspect*, 19, 99-106.

ROMANO, G. 2007a. Current development of nonviral-mediated gene transfer. *Drug News Perspect*, 20, 227-31.

ROMANO, G. 2007b. The standpoint of gene therapy programs. *Drug News Perspect*, 20, 335-43.

ROMANO, G. 2008a. Artificial reprogramming of human somatic cells to generate pluripotent stem cells: a possible alternative to the controversial use of human embryonic stem cells. *Drug News Perspect*, 21, 440-5.

ROMANO, G. 2008b. The standpoint of stem cell research. *Drug News Perspect*, 21, 408-12.

ROMANO, G. 2012. Development of safer gene delivery systems to minimize the risk of insertional mutagenesis-related malignancies: a critical issue for the field of gene therapy. *ISRN Oncol*, 2012, 616310.

ROMANO, G. 2019. Adeno-Associated Viral-Mediated Gene Transfer. *Mater Methods*, 9, 2796.

ROMANO, G., CLAUDIO, P. P., TONINI, T. & GIORDANO, A. 2003. Human immunodeficiency virus type 1 (HIV-1) derived vectors: safety considerations and controversy over therapeutic applications. *Eur J Dermatol*, 13, 424-9.

ROMANO, G., MICHELL, P., PACILIO, C. & GIORDANO, A. 2000. Latest developments in gene transfer technology: achievements, perspectives, and controversies over therapeutic applications. *Stem Cells*, 18, 19-39.

ROSENBERG, S. A., AEBERSOLD, P., CORNETTA, K., KASID, A., MORGAN, R. A., MOEN, R., KARSON, E. M., LOTZE, M. T., YANG, J. C., TOPALIAN, S. L. & ET AL. 1990. Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N Engl J Med*, 323, 570-8.

ROSSANT, J. 2008. Stem cells and early lineage development. *Cell*, 132, 527-31.

ROWE, C., GERRARD, D. T., JENKINS, R., BERRY, A., DURKIN, K., SUNDSTROM, L., GOLDRING, C. E., PARK, B. K., KITTERINGHAM, N. R., HANLEY, K. P. & HANLEY, N. A. 2013. Proteome-wide analyses of human hepatocytes during differentiation and dedifferentiation. *Hepatology*, 58, 799-809.

ROWE, C., GOLDRING, C. E., KITTERINGHAM, N. R., JENKINS, R. E., LANE, B. S., SANDERSON, C., ELLIOTT, V., PLATT, V., METCALFE, P. & PARK, B. K. 2010. Network analysis of primary hepatocyte dedifferentiation using a shotgun proteomics approach. *J Proteome Res*, 9, 2658-68.

ROWE, W. P., HUEBNER, R. J., GILMORE, L. K., PARROTT, R. H. & WARD, T. G. 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc Soc Exp Biol Med*, 84, 570-3.

SAKHUJA, K., REDDY, P. S., GANESH, S., CANTANIAG, F., PATTISON, S., LIMBACH, P., KAYDA, D. B., KADAN, M. J., KALEKO, M. & CONNELLY, S. 2003. Optimization of the generation and propagation of gutless adenoviral vectors. *Hum Gene Ther*, 14, 243-54.

SATO, T., VRIES, R. G., SNIPPERT, H. J., VAN DE WETERING, M., BARKER, N., STANGE, D. E., VAN ES, J. H., ABO, A., KUJALA, P., PETERS, P. J. & CLEVERS, H. 2009. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459, 262-5.

SAXENA, R., THEISE, N. D. & CRAWFORD, J. M. 1999. Microanatomy of the human liver-exploring the hidden interfaces. *Hepatology*, 30, 1339-46.

SCHAMBACH, A., GALLA, M., MAETZIG, T., LOEW, R. & BAUM, C. 2007. Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors. *Mol Ther*, 15, 1167-73.

SCHAMBACH, A., SWANEY, W. P. & VAN DER LOO, J. C. 2009. Design and production of retro- and lentiviral vectors for gene expression in hematopoietic cells. *Methods Mol Biol*, 506, 191-205.

SCHLAEGER, T. M., DAHERON, L., BRICKLER, T. R., ENTWISLE, S., CHAN, K., CIANCI, A., DEVINE, A., ETTENGER, A., FITZGERALD, K., GODFREY, M., GUPTA, D., MCPHERSON, J., MALWADKAR, P., GUPTA, M., BELL, B., DOI, A., JUNG, N., LI, X., LYNES, M. S., BROOKES, E., CHERRY, A. B., DEMIRBAS, D., TSANKOV, A. M., ZON, L. I., RUBIN, L. L., FEINBERG, A. P., MEISSNER, A., COWAN, C. A. & DALEY, G. Q. 2015. A comparison of non-integrating reprogramming methods. *Nat Biotechnol*, 33, 58-63.

SCHMIDT, M., HOFFMANN, G., WISSELER, M., LEMKE, N., MUSSIG, A., GLIMM, H., WILLIAMS, D. A., RAGG, S., HESEMANN, C. U. & VON KALLE, C. 2001. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum Gene Ther*, 12, 743-9.

SCHMIDT, M., SCHWARZWAELDER, K., BARTHOLOMAE, C., ZAOUI, K., BALL, C., PILZ, I., BRAUN, S., GLIMM, H. & VON KALLE, C. 2007. High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat Methods*, 4, 1051-7.

SCHMIDT, M., SCHWARZWAELDER, K., BARTHOLOMAE, C. C., GLIMM, H. & VON KALLE, C. 2009. Detection of retroviral integration sites by linear amplification-mediated PCR and tracking of individual integration clones in different samples. *Methods Mol Biol*, 506, 363-72.

SCHMIDT, M., ZICKLER, P., HOFFMANN, G., HAAS, S., WISSELER, M., MUESSIG, A., TISDALE, J. F., KURAMOTO, K., ANDREWS, R. G., WU, T., KIEM, H. P., DUNBAR, C. E. & VON KALLE, C. 2002. Polyclonal long-term repopulating stem cell clones in a primate model. *Blood*, 100, 2737-43.

SCHOSSERER, M., REYNOSO, R., WALLY, V., JUG, B., KANTNER, V., WEILNER, S., BURIC, I., GRILLARI, J., BAUER, J. W. & GRILLARI-VOGLAUER, R. 2015. Urine is a novel source of autologous mesenchymal stem cells for patients with epidermolysis bullosa. *BMC Res Notes*, 8, 767.

SCHRODER, A. R., SHINN, P., CHEN, H., BERRY, C., ECKER, J. R. & BUSHMAN, F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, 110, 521-9.

SCHUSTER, S. J., BISHOP, M. R., TAM, C. S., WALLER, E. K., BORCHMANN, P., MCGUIRK, J. P., JAGER, U., JAGLOWSKI, S., ANDREADIS, C., WESTIN, J. R., FLEURY, I., BACHANOVA, V., FOLEY, S. R., HO, P. J., MIELKE, S., MAGENAU, J. M., HOLTE, H., PANTANO, S., PACAUD, L. B., AWASTHI, R., CHU, J. F., ANAK, O., SALLES, G., MAZIARZ, R. T. & INVESTIGATORS, J. 2019. Tisagenlecleucel in Adult Relapsed or Refractory Diffuse Large B-Cell Lymphoma. *New England Journal of Medicine*, 380, 45-56.

SEGGEWISS, R., PITTALUGA, S., ADLER, R. L., GUENAGA, F. J., FERGUSON, C., PILZ, I. H., RYU, B., SORRENTINO, B. P., YOUNG, W. S., 3RD, DONAHUE, R. E., VON KALLE, C., NIENHUIS, A. W. & DUNBAR, C. E. 2006. Acute myeloid leukemia is associated with retroviral gene transfer to hematopoietic progenitor cells in a rhesus macaque. *Blood*, 107, 3865-7.

SEKHON, S. S., TAN, X., MICSENYI, A., BOWEN, W. C. & MONGA, S. P. 2004. Fibroblast growth factor enriches the embryonic liver cultures for hepatic progenitors. *Am J Pathol*, 164, 2229-40.

SERLS, A. E., DOHERTY, S., PARVATIYAR, P., WELLS, J. M. & DEUTSCH, G. H. 2005. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development*, 132, 35-47.

SHAH, U. K., MALLIA, J. O., SINGH, N., CHAPMAN, K. E., DOAK, S. H. & JENKINS, G. J. S. 2018. A three-dimensional in vitro HepG2 cells liver spheroid model for genotoxicity studies. *Mutat Res Genet Toxicol Environ Mutagen*, 825, 51-58.

SHAYAKHMETOV, D. M. & LIEBER, A. 2000. Dependence of adenovirus infectivity on length of the fiber shaft domain. *J Virol*, 74, 10274-86.

SHI, Y. & MASSAGUE, J. 2003. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*, 113, 685-700.

SHIM, G., KIM, D., PARK, G. T., JIN, H., SUH, S. K. & OH, Y. K. 2017. Therapeutic gene editing: delivery and regulatory perspectives. *Acta Pharmacol Sin*, 38, 738-753.

SINGH, V. K., KUMAR, N., KALSAN, M., SAINI, A. & CHANDRA, R. 2015. Mechanism of Induction: Induced Pluripotent Stem Cells (iPSCs). *J Stem Cells*, 10, 43-62.

SKETT, P. & BAYLISS, M. 1996. Time for a consistent approach to preparing and culturing hepatocytes? *Xenobiotica*, 26, 1-7.

SOLDATOW, V. Y., LECLUYSE, E. L., GRIFFITH, L. G. & RUSYN, I. 2013. In vitro models for liver toxicity testing. *Toxicol Res (Camb)*, 2, 23-39.

SONG, Z., CAI, J., LIU, Y., ZHAO, D., YONG, J., DUO, S., SONG, X., GUO, Y., ZHAO, Y., QIN, H., YIN, X., WU, C., CHE, J., LU, S., DING, M. & DENG, H. 2009. Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. *Cell Res*, 19, 1233-42.

SOSA-PINEDA, B., WIGLE, J. T. & OLIVER, G. 2000. Hepatocyte migration during liver development requires Prox1. *Nature Genetics*, 25, 254-255.

STADTFELD, M., BRENNAND, K. & HOCHEDLINGER, K. 2008. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr Biol*, 18, 890-4.

STANULOVIC, V. S., KYRMIZI, I., KRUITHOF-DE JULIO, M., HOOGENKAMP, M., VERMEULEN, J. L., RUIJTER, J. M., TALIANIDIS, I., HAKVOORT, T. B. & LAMERS, W. H. 2007. Hepatic HNF4alpha deficiency induces periportal expression of glutamine synthetase and other pericentral enzymes. *Hepatology*, 45, 433-44.

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., KUHLCHE, K., SCHLEGELBERGER, B., THRASHER, A. J., HOELZER, D., SEGER, R., VON KALLE, C. & GREZ, M. 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med*, 16, 198-204.

STEVENS, J. C., HINES, R. N., GU, C. G., KOUKOURITAKI, S. B., MANRO, J. R., TANDLER, P. J. & ZAYA, M. J. 2003. Developmental expression of the major human hepatic CYP3A enzymes. *Journal of Pharmacology and Experimental Therapeutics*, 307, 573-582.

STOVER, A. E. & SCHWARTZ, P. H. 2011. Adaptation of human pluripotent stem cells to feeder-free conditions in chemically defined medium with enzymatic single-cell passaging. *Methods Mol Biol*, 767, 137-46.

SU, T. & WAXMAN, D. J. 2004. Impact of dimethyl sulfoxide on expression of nuclear receptors and drug-inducible cytochromes P450 in primary rat hepatocytes. *Arch Biochem Biophys*, 424, 226-34.

SUBRAMANIAN, R. P. & GERAGHTY, R. J. 2007. Herpes simplex virus type 1 mediates fusion through a hemifusion intermediate by sequential activity of glycoproteins D, H, L, and B. *Proc Natl Acad Sci U S A*, 104, 2903-8.

SUERTH, J. D., LABENSKI, V. & SCHAMBACH, A. 2014. Alpharetroviral vectors: from a cancer-causing agent to a useful tool for human gene therapy. *Viruses*, 6, 4811-38.

SUGA, H. 2016. Recapitulating Hypothalamus and Pituitary Development Using Embryonic Stem/Induced Pluripotent Stem Cells. *In: PFAFF, D. & CHRISTEN, Y. (eds.) Stem Cells in Neuroendocrinology*. Cham (CH).

SUGIMACHI, K., SOSEF, M. N., BAUST, J. M., FOWLER, A., TOMPKINS, R. G. & TONER, M. 2004. Long-term function of cryopreserved rat hepatocytes in a coculture system. *Cell Transplant*, 13, 187-95.

SULLIVAN, G. J., HAY, D. C., PARK, I. H., FLETCHER, J., HANNOUN, Z., PAYNE, C. M., DALGETTY, D., BLACK, J. R., ROSS, J. A., SAMUEL, K., WANG, G., DALEY, G. Q., LEE, J. H., CHURCH, G. M., FORBES, S. J., IREDALE, J. P. & WILMUT, I. 2010. Generation of functional human hepatic endoderm from human induced pluripotent stem cells. *Hepatology*, 51, 329-35.

SZKOLNICKA, D., FARNWORTH, S. L., LUCENDO-VILLARIN, B., STORCK, C., ZHOU, W., IREDALE, J. P., FLINT, O. & HAY, D. C. 2014. Accurate prediction of drug-induced liver injury using stem cell-derived populations. *Stem Cells Transl Med*, 3, 141-8.

SZKOLNICKA, D. & HAY, D. C. 2016. Concise Review: Advances in Generating Hepatocytes from Pluripotent Stem Cells for Translational Medicine. *Stem Cells*, 34, 1421-6.

TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131, 861-72.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126, 663-76.

TAKEBE, T., ENOMURA, M., YOSHIZAWA, E., KIMURA, M., KOIKE, H., UENO, Y., MATSUZAKI, T., YAMAZAKI, T., TOYOHARA, T., OSAFUNE, K., NAKAUCHI, H., YOSHIKAWA, H. Y. & TANIGUCHI, H. 2015. Vascularized and Complex Organ Buds from Diverse Tissues via Mesenchymal Cell-Driven Condensation. *Cell Stem Cell*, 16, 556-65.

TAKEBE, T., SEKINE, K., ENOMURA, M., KOIKE, H., KIMURA, M., OGAERI, T., ZHANG, R. R., UENO, Y., ZHENG, Y. W., KOIKE, N., AOYAMA, S., ADACHI, Y. & TANIGUCHI, H. 2013. Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature*, 499, 481-4.

TAN, X., YUAN, Y., ZENG, G., APTE, U., THOMPSON, M. D., CIEPLY, B., STOLZ, D. B., MICHALOPOULOS, G. K., KAESTNER, K. H. & MONGA, S. P. 2008. Beta-catenin deletion in hepatoblasts disrupts hepatic morphogenesis and survival during mouse development. *Hepatology*, 47, 1667-79.

TARAVIRAS, S., MONAGHAN, A. P., SCHUTZ, G. & KELSEY, G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mech Dev*, 48, 67-79.

THE-GIST.ORG.

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. & COUTELLE, C. 2005. Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice. *Mol Ther*, 12, 763-71.
- THOMSON, J. A., ITSKOVITZ-ELDOR, J., SHAPIRO, S. S., WAKNITZ, M. A., SWIERGIEL, J. J., MARSHALL, V. S. & JONES, J. M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science*, 282, 1145-7.
- TILL, J. E. & MC, C. E. 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res*, 14, 213-22.
- TOUBOUL, T., HANNAN, N. R., CORBINEAU, S., MARTINEZ, A., MARTINET, C., BRANCHEREAU, S., MAINOT, S., STRICK-MARCHAND, H., PEDERSEN, R., DI SANTO, J., WEBER, A. & VALLIER, L. 2010. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology*, 51, 1754-65.
- TOUW, I. P. & ERKELAND, S. J. 2007. Retroviral insertion mutagenesis in mice as a comparative oncogenomics tool to identify disease genes in human leukemia. *Mol Ther*, 15, 13-9.
- TROBRIDGE, G. D. & KIEM, H. P. 2010. Large animal models of hematopoietic stem cell gene therapy. *Gene Ther*, 17, 939-48.
- TURNER, R., LOZOYA, O., WANG, Y., CARDINALE, V., GAUDIO, E., ALPINI, G., MENDEL, G., WAUTHIER, E., BARBIER, C., ALVARO, D. & REID, L. M. 2011. Human hepatic stem cell and maturational liver lineage biology. *Hepatology*, 53, 1035-45.
- TUSCHL, G. & MUELLER, S. O. 2006. Effects of cell culture conditions on primary rat hepatocytes-cell morphology and differential gene expression. *Toxicology*, 218, 205-15.
- VALLIER, L., ALEXANDER, M. & PEDERSEN, R. A. 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci*, 118, 4495-509.
- VAZIN, T. & FREED, W. J. 2010. Human embryonic stem cells: derivation, culture, and differentiation: a review. *Restor Neurol Neurosci*, 28, 589-603.
- VESSEY, C. J. & DE LA HALL, P. M. 2001. Hepatic stem cells: a review. *Pathology*, 33, 130-41.
- VILLA-DIAZ, L. G., PACUT, C., SLAWNY, N. A., DING, J., O'SHEA, K. S. & SMITH, G. D. 2009. Analysis of the factors that limit the ability of feeder cells to maintain the undifferentiated state of human embryonic stem cells. *Stem Cells Dev*, 18, 641-51.

VINKEN, M., PAPELEU, P., SNYKERS, S., DE ROP, E., HENKENS, T., CHIPMAN, J. K., ROGIERS, V. & VANHAECKE, T. 2006. Involvement of cell junctions in hepatocyte culture functionality. *Crit Rev Toxicol*, 36, 299-318.

WAGNER, A., ROHRS, V., MATERNE, E. M., HILLER, T., KEDZIERSKI, R., FECHNER, H., LAUSTER, R. & KURRECK, J. 2015. Use of a three-dimensional humanized liver model for the study of viral gene vectors. *J Biotechnol*, 212, 134-43.

WANG, L., SCHULZ, T. C., SHERRER, E. S., DAUPHIN, D. S., SHIN, S., NELSON, A. M., WARE, C. B., ZHAN, M., SONG, C. Z., CHEN, X., BRIMBLE, S. N., MCLEAN, A., GALEANO, M. J., UHL, E. W., D'AMOUR, K. A., CHESNUT, J. D., RAO, M. S., BLAU, C. A. & ROBINS, A. J. 2007. Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood*, 110, 4111-9.

WANG, Y., ALHAQUE, S., CAMERON, K., MESEGUER-RIPOLLES, J., LUCENDO-VILLARIN, B., RASHIDI, H. & HAY, D. C. 2017. Defined and Scalable Generation of Hepatocyte-like Cells from Human Pluripotent Stem Cells. *J Vis Exp*.

WARREN, L., MANOS, P. D., AHFELDT, T., LOH, Y. H., LI, H., LAU, F., EBINA, W., MANDAL, P. K., SMITH, Z. D., MEISSNER, A., DALEY, G. Q., BRACK, A. S., COLLINS, J. J., COWAN, C., SCHLAEGER, T. M. & ROSSI, D. J. 2010. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, 7, 618-30.

WEINSTEIN, D. C., RUIZ I ALTABA, A., CHEN, W. S., HOODLESS, P., PREZIOSO, V. R., JESSELL, T. M. & DARNELL, J. E., JR. 1994. The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell*, 78, 575-88.

WERNIG, M., MEISSNER, A., FOREMAN, R., BRAMBRINK, T., KU, M., HOCHEDLINGER, K., BERNSTEIN, B. E. & JAENISCH, R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*, 448, 318-24.

WHITE, M., WHITTAKER, R., GANDARA, C. & STOLL, E. A. 2017. A Guide to Approaching Regulatory Considerations for Lentiviral-Mediated Gene Therapies. *Hum Gene Ther Methods*, 28, 163-176.

WILKENING, S., STAHL, F. & BADER, A. 2003. Comparison of primary human hepatocytes and hepatoma cell line Hepg2 with regard to their biotransformation properties. *Drug Metab Dispos*, 31, 1035-42.

WILL, E., BAILEY, J., SCHUESLER, T., MODLICH, U., BALCIK, B., BURZYNSKI, B., WITTE, D., LAYH-SCHMITT, G., RUDOLPH, C., SCHLEGELBERGER, B., VON KALLE, C., BAUM, C., SORRENTINO, B. P., WAGNER, L. M., KELLY, P., REEVES,



L. & WILLIAMS, D. A. 2007. Importance of murine study design for testing toxicity of retroviral vectors in support of phase I trials. *Mol Ther*, 15, 782-91.

WILMUT, I., SCHNIEKE, A. E., MCWHIR, J., KIND, A. J. & CAMPBELL, K. H. 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385, 810-3.

WOBUS, A. M. & BOHELER, K. R. 2005. Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiol Rev*, 85, 635-78.

WONG, N., LAI, P., PANG, E., LEUNG, T. W., LAU, J. W. & JOHNSON, P. J. 2000. A comprehensive karyotypic study on human hepatocellular carcinoma by spectral karyotyping. *Hepatology*, 32, 1060-8.

WOODS, N. B., BOTTERO, V., SCHMIDT, M., VON KALLE, C. & VERMA, I. M. 2006. Gene therapy: therapeutic gene causing lymphoma. *Nature*, 440, 1123.

WU, C. & DUNBAR, C. E. 2011. Stem cell gene therapy: the risks of insertional mutagenesis and approaches to minimize genotoxicity. *Front Med*, 5, 356-71.

WU, F., WU, D., REN, Y., HUANG, Y., FENG, B., ZHAO, N., ZHANG, T., CHEN, X., CHEN, S. & XU, A. 2019. Generation of hepatobiliary organoids from human induced pluripotent stem cells. *J Hepatol*, 70, 1145-1158.

XIANG, Y., TANAKA, Y., PATTERSON, B., KANG, Y. J., GOVINDAIAH, G., ROSELAAR, N., CAKIR, B., KIM, K. Y., LOMBROSO, A. P., HWANG, S. M., ZHONG, M., STANLEY, E. G., ELEFANTY, A. G., NAEGELE, J. R., LEE, S. H., WEISSMAN, S. M. & PARK, I. H. 2017. Fusion of Regionally Specified hPSC-Derived Organoids Models Human Brain Development and Interneuron Migration. *Cell Stem Cell*, 21, 383-398 e7.

XU, C., INOKUMA, M. S., DENHAM, J., GOLDS, K., KUNDU, P., GOLD, J. D. & CARPENTER, M. K. 2001. Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol*, 19, 971-4.

YANG, Y., NUNES, F. A., BERENCSI, K., FURTH, E. E., GONCZOL, E. & WILSON, J. M. 1994. Cellular immunity to viral antigens limits E1-deleted adenoviruses for gene therapy. *Proc Natl Acad Sci U S A*, 91, 4407-11.

YAO, S., CHEN, S., CLARK, J., HAO, E., BEATTIE, G. M., HAYEK, A. & DING, S. 2006. Long-term self-renewal and directed differentiation of human embryonic stem cells in chemically defined conditions. *Proc Natl Acad Sci U S A*, 103, 6907-12.

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, II & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318, 1917-20.

ZARE, M., SOLEIMANI, M., MOHAMMADIAN, M., AKBARZADEH, A., HAVASI, P. & ZARGHAMI, N. 2016. Efficient biotechnological approach for lentiviral transduction of induced pluripotent stem cells. *Artif Cells Nanomed Biotechnol*, 44, 743-8.

ZARET, K. S. & GROMPE, M. 2008. Generation and regeneration of cells of the liver and pancreas. *Science*, 322, 1490-4.

ZHANG, C., ZHAO, Z., ABDUL RAHIM, N. A., VAN NOORT, D. & YU, H. 2009. Towards a human-on-chip: culturing multiple cell types on a chip with compartmentalized microenvironments. *Lab Chip*, 9, 3185-92.

ZHANG, F., THORNHILL, S. I., HOWE, S. J., ULAGANATHAN, M., SCHAMBACH, A., SINCLAIR, J., KINNON, C., GASPAR, H. B., ANTONIOU, M. & THRASHER, A. J. 2007. Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. *Blood*, 110, 1448-57.

ZHANG, Y. S., ALEMAN, J., SHIN, S. R., KILIC, T., KIM, D., MOUSAVI SHAEGH, S. A., MASSA, S., RIAHI, R., CHAE, S., HU, N., AVCI, H., ZHANG, W., SILVESTRI, A., SANATI NEZHAD, A., MANBOHI, A., DE FERRARI, F., POLINI, A., CALZONE, G., SHAIKH, N., ALERASOOL, P., BUDINA, E., KANG, J., BHISE, N., RIBAS, J., POURMAND, A., SKARDAL, A., SHUPE, T., BISHOP, C. E., DOKMECI, M. R., ATALA, A. & KHADEMHOSEINI, A. 2017. Multisensor-integrated organs-on-chips platform for automated and continual in situ monitoring of organoid behaviors. *Proc Natl Acad Sci U S A*, 114, E2293-E2302.

ZHAO, R. & DUNCAN, S. A. 2005. Embryonic development of the liver. *Hepatology*, 41, 956-67.

ZHOU, H., WU, S., JOO, J. Y., ZHU, S., HAN, D. W., LIN, T., TRAUGER, S., BIEN, G., YAO, S., ZHU, Y., SIUZDAK, G., SCHOLER, H. R., DUAN, L. & DING, S. 2009. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell*, 4, 381-4.

ZHOU, S., MA, Z., LU, T., JANKE, L., GRAY, J. T. & SORRENTINO, B. P. 2013. Mouse transplant models for evaluating the oncogenic risk of a self-inactivating XSCID lentiviral vector. *PLoS One*, 8, e62333.

ZHOU, S., MODY, D., DERA VIN, S. S., HAUER, J., LU, T., MA, Z., HACEIN-BEY ABINA, S., GRAY, J. T., GREENE, M. R., CAVAZZANA-CALVO, M., MALECH, H. L. & 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, 900-8.

ZINCARELLI, C., SOLTYS, S., RENGO, G. & RABINOWITZ, J. E. 2008. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther*, 16, 1073-80.

ZWAKA, T. P. 2006.

ZYCHLINSKI, D., SCHAMBACH, A., MODLICH, U., MAETZIG, T., MEYER, J., GRASSMAN, E., MISHRA, A. & BAUM, C. 2008. Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol Ther*, 16, 718-25.

## 8. Supplementary Data

### 8.1. Publications

Arch Toxicol (2016) 90:1757–1761  
DOI 10.1007/s00204-016-1689-8



SHORT COMMUNICATION

## Fluid shear stress modulation of hepatocyte-like cell function

Hassan Rashidi<sup>1</sup> · Sharmin Alhaque<sup>1</sup> · Dagmara Szkolnicka<sup>1</sup> · Oliver Flint<sup>1</sup> · David C. Hay<sup>1</sup>

Received: 31 January 2016 / Accepted: 25 February 2016 / Published online: 15 March 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Freshly isolated human adult hepatocytes are considered to be the gold standard tool for in vitro studies. However, primary hepatocyte scarcity, cell cycle arrest and the rapid loss of cell phenotype limit their widespread deployment. Human embryonic stem cells and induced pluripotent stem cells provide renewable sources of hepatocyte-like cells (HLCs). Despite the use of various differentiation methodologies, HLCs like primary human hepatocytes exhibit unstable phenotype in culture. It has been shown that the functional capacity can be improved by adding back elements of human physiology, such as cell co-culture or through the use of natural and/or synthetic surfaces. In this study, the effect of fluid shear stress on HLC performance was investigated. We studied two important liver functions, cytochrome P450 drug metabolism and serum protein secretion, in static cultures and those exposed to fluid shear stress. Our study demonstrates that fluid shear stress improved Cyp1A2 activity by approximately fivefold. This was paralleled by an approximate ninefold increase in sensitivity to a drug, primarily metabolised by Cyp2D6. In addition to metabolic capacity, fluid shear stress also improved hepatocyte phenotype with an approximate fourfold reduction in the secretion of a foetal marker, alpha-fetoprotein. We believe these studies highlight the importance of introducing physiologic cues in cell-based models to improve somatic cell phenotype.

**Keywords** Fluid shear stress · Hepatocyte-like cell · Embryonic stem cell · Cytochrome P450 Metabolism · Albumin secretion · Alpha-fetoprotein secretion

### Introduction

Static and two-dimensional (2D) culture systems have been used extensively to study human biology. Although those systems are facile and cost-effective to use, they lack the complexity of the three-dimensional (3D) tissue microenvironment. As a result, those systems do not accurately model tissue physiology and can generate inaccurate datasets. It is a well-known phenomenon that cells respond to physical and chemical stimuli provided by the tissue niche (Rashidi et al. 2014). Therefore, the add-back of human physiology to cell based models is of the utmost importance. This will likely lead to an improvement in cell phenotype and more informative biological readouts from those systems (Godoy et al. 2015).

Fluid shear stress is one mechanical stimulus that is absent in static culture systems. The role of fluid transport is fundamental for organogenesis (Freund et al. 2012), cell signalling (Mammoto and Ingber 2010) and normal patterns of organ function (Hahn and Schwartz 2009; Hildebrandt et al. 2011). Living cells possess the ability to sense mechanical forces and transduce those into biological responses (Bao and Suresh 2003; Freund et al. 2012). The mechanisms that modulate cell behaviour by fluid shear stress is diverse, but is mediated primarily through cell surface receptors (Tzima et al. 2001), cell adhesion molecules (Tzima et al. 2005) and heterodimeric G proteins (White and Frangos 2007).

The effects of fluid shear stress have also been studied in primary hepatocytes and transformed hepatocyte cell lines.

✉ David C. Hay  
davehay@talktalk.net

<sup>1</sup> MRC Centre for Regenerative Medicine, University of Edinburgh, Edinburgh EH16 4UU, UK

## Video Article

## Defined and Scalable Generation of Hepatocyte-like Cells from Human Pluripotent Stem Cells

Yu Wang<sup>1</sup>, Sharmin Alhaque<sup>1</sup>, Kate Cameron<sup>1</sup>, Jose Meseguer-Ripolles<sup>1</sup>, Baltasar Lucendo-Villarin<sup>1</sup>, Hassan Rashidi<sup>1</sup>, David C. Hay<sup>1</sup><sup>1</sup>MRC Centre for Regenerative Medicine, University of EdinburghCorrespondence to: David C. Hay at [davehay@talktalk.net](mailto:davehay@talktalk.net)URL: <https://www.jove.com/video/55355>DOI: [doi:10.3791/55355](https://doi.org/10.3791/55355)

Keywords: Developmental Biology, Issue 121, hepatocyte-like cells, recombinant laminins, human pluripotent stem cells, GMP-ready system, scale-up, drug screening, cell therapy

Date Published: 3/2/2017

Citation: Wang, Y., Alhaque, S., Cameron, K., Meseguer-Ripolles, J., Lucendo-Villarin, B., Rashidi, H., Hay, D.C. Defined and Scalable Generation of Hepatocyte-like Cells from Human Pluripotent Stem Cells. *J. Vis. Exp.* (121), e55355, doi:10.3791/55355 (2017).

### Abstract

Human pluripotent stem cells (hPSCs) possess great value for biomedical research. hPSCs can be scaled and differentiated to all cell types found in the human body. The differentiation of hPSCs to human hepatocyte-like cells (HLCs) has been extensively studied, and efficient differentiation protocols have been established. The combination of extracellular matrix and biological stimuli, including growth factors, cytokines, and small molecules, have made it possible to generate HLCs that resemble primary human hepatocytes. However, the majority of procedures still employ undefined components, giving rise to batch-to-batch variation. This serves as a significant barrier to the application of the technology. To tackle this issue, we developed a defined system for hepatocyte differentiation using human recombinant laminins as extracellular matrices in combination with a serum-free differentiation process. Highly efficient hepatocyte specification was achieved, with demonstrated improvements in both HLC function and phenotype. Importantly, this system is easy to scale up using research and GMP-grade hPSC lines promising advances in cell-based modelling and therapies.

### Video Link

The video component of this article can be found at <https://www.jove.com/video/55355/>

### Introduction

Primary human tissue and the derivative cell types are regularly used, both for cell-based screening and in the clinic. However, access to these cells is severely limited due to insufficient organ donation and loss of cell phenotypes post-isolation<sup>1</sup>. hPSCs represent a promising alternative to primary tissue and facilitate the generation of genetically defined and renewable human somatic cells. Hepatocyte-like cells (HLCs) derived from hPSCs have already shown promise in this field. HLCs resemble primary human hepatocytes in various aspects, including cell morphology, hepatocyte gene expression, metabolic function, and sensitivity to drugs and viruses<sup>2,3,4,5,6,7,8</sup>. In addition, the unlimited proliferation and self-renewal capacity of both research- and GMP-grade hPSCs facilitates their application<sup>9,10</sup>.

Over a decade of research has produced a number of efficient hepatocyte differentiation procedures<sup>2,3,5,11,12,13,14,15,16,17,18,19,20</sup>. However, most of these systems use undefined components and/or viral transduction to drive hepatocellular specification. To improve the reliability of the technology at scale, it is important to develop a robust hepatocyte differentiation system that is truly defined, xeno-free, and GMP-compatible.

Laminins (LNs) are important extracellular matrix proteins that can influence cell adhesion, proliferation, migration, and differentiation. Laminins are heterotrimeric glycoproteins composed of one  $\alpha$ , one  $\beta$ , and one  $\gamma$  chain. Recently, recombinant human laminins have been produced and used in cell biology. LN-511 has been shown to support the maintenance of hPSCs<sup>21</sup>, while a mixture of LN-521 and E-cadherin allows clonal derivation and the expansion of human embryonic stem cells<sup>22</sup>. LN-111, on the other hand, supports the maintenance of hepatoblast-like cells derived from hPSCs<sup>23</sup>. However, before our report, laminins 521 and 111 had not been utilized to generate HLCs with mature characteristics from hPSCs<sup>10</sup>.

Here, we detail procedures for culturing hPSCs on LN-521 and differentiating them on either LN-521 or a blend of LN-521 and LN-111 (LN-521/LN-111). We optimized the differentiation protocol using single-cell seeding to generate a highly reproducible and homogenous monolayer of HLCs in many formats<sup>14</sup>. We believe that our defined differentiation system represents a simple and cost-effective method to manufacture active HLCs for application, representing a significant step forward in the field.

### Protocol

NOTE: Vendor information for all reagents used in this protocol has been listed in **Table 1**. All media/plates should be sterile and at least at room temperature when cells are to have direct contact with them.



## 3D human liver tissue from pluripotent stem cells displays stable phenotype in vitro and supports compromised liver function in vivo

Hassan Rashidi<sup>1</sup> · Nguyet-Thin Luu<sup>2</sup> · Salamah M. Alwahsh<sup>1</sup> · Maaria Ginai<sup>4</sup> · Sharmin Alhaque<sup>1</sup> · Hua Dong<sup>5</sup> · Rute A. Tomaz<sup>6</sup> · Bertrand Vernay<sup>1</sup> · Vasanthy Vigneswara<sup>2</sup> · John M. Hallett<sup>1</sup> · Anil Chandrashekran<sup>7</sup> · Anil Dhawan<sup>7</sup> · Ludovic Vallier<sup>6</sup> · Mark Bradley<sup>5</sup> · Anthony Callanan<sup>4</sup> · Stuart J. Forbes<sup>1</sup> · Philip N. Newsome<sup>2,3</sup> · David C. Hay<sup>1</sup>

Received: 23 June 2018 / Accepted: 31 July 2018 / Published online: 28 August 2018  
© The Author(s) 2018

### Abstract

Liver disease is an escalating global health issue. While liver transplantation is an effective mode of therapy, patient mortality has increased due to the shortage of donor organs. Developing renewable sources of human liver tissue is therefore attractive. Pluripotent stem cell-derived liver tissue represents a potential alternative to cadaver derived hepatocytes and whole organ transplant. At present, two-dimensional differentiation procedures deliver tissue lacking certain functions and long-term stability. Efforts to overcome these limiting factors have led to the building of three-dimensional (3D) cellular aggregates. Although enabling for the field, their widespread application is limited due to their reliance on variable biological components. Our studies focused on the development of 3D liver tissue under defined conditions. In vitro generated 3D tissues exhibited stable phenotype for over 1 year in culture, providing an attractive resource for long-term in vitro studies. Moreover, 3D derived tissue provided critical liver support in two animal models, including immunocompetent recipients. Therefore, we believe that our study provides stable human tissue to better model liver biology ‘in the dish’, and in the future may permit the support of compromised liver function in humans.

**Keywords** Liver tissue · Pluripotent stem cell · Stable cell phenotype · Implantable liver graft · Interdisciplinary research

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00204-018-2280-2>) contains supplementary material, which is available to authorized users.

✉ David C. Hay  
davehay@talktalk.net

- <sup>1</sup> MRC Centre for Regenerative Medicine, University of Edinburgh, Edinburgh EH16 4UU, UK
- <sup>2</sup> Centre for Liver Research, Institute of Immunology and Immunotherapy and National Institute for Health Research Biomedical Research Centre at University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham, Birmingham, UK
- <sup>3</sup> Liver Unit, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
- <sup>4</sup> Institute of Bioengineering, The University of Edinburgh, King’s Buildings, Edinburgh EH9 3DW, UK
- <sup>5</sup> School of Chemistry, University of Edinburgh, Kings Buildings, EH9 3FJ Edinburgh, UK
- <sup>6</sup> Anne McLaren Laboratory, Wellcome Trust-MRC Stem Cell Institute, University of Cambridge, Cambridge CB2 0SZ, UK
- <sup>7</sup> Child Health Clinical Academic Group, MRC Centre for Transplantation, King’s College London, London, UK

### Introduction

The liver performs a wide range of functions essential for body function. Significant loss of liver function therefore has serious consequences for human health (Ebrahimkhani et al. 2014). In the UK, liver disease is the fifth biggest killer, and unlike the top four, the incidence is increasing. Orthotopic liver transplantation (OLT) is the most effective treatment for acute or end-stage liver failure. However, due to the shortage of organ donors, and complications associated with lifelong immunosuppression, there is an impetus to develop alternative therapies.

While donor hepatocytes have been used to successfully treat metabolic liver disease (Alwahsh et al. 2018), the graft is eventually lost. This is further compounded by the scarcity of human hepatocytes. Therefore, we and others have proposed that pluripotent stem cells (PSCs) are a credible alternative with which to generate a stable source of quality assured human liver tissue for applied medicine. PSCs can self-renew and differentiate to all cells types in the body, promising an unlimited supply of human tissue for



**Cite this article:** Alhaque S, Themis M, Rashidi H. 2018 Three-dimensional cell culture: from evolution to revolution. *Phil. Trans. R. Soc. B* **373**: 20170216. <http://dx.doi.org/10.1098/rstb.2017.0216>

Accepted: 19 March 2018

One contribution of 18 to a theme issue 'Designer human tissue: coming to a lab near you'.

**Subject Areas:**

bioengineering, biomaterials, cellular biology, developmental biology

**Keywords:**

pluripotent stem cells, organoids, three-dimensional culture

**Author for correspondence:**

Hassan Rashidi  
e-mail: h.rashidi@ucl.ac.uk

<sup>†</sup>Present address: Great Ormond Street Institute of Child Health, University College London, London, UK.

# Three-dimensional cell culture: from evolution to revolution

Sharmin Alhaque<sup>1,2</sup>, Michael Themis<sup>2</sup> and Hassan Rashidi<sup>1,†</sup>

<sup>1</sup>Scottish Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK

<sup>2</sup>Division of Biosciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Uxbridge, Middlesex, UK

HR, 0000-0001-8078-6688

Recent advances in the isolation of tissue-resident adult stem cells and the identification of inductive factors that efficiently direct differentiation of human pluripotent stem cells along specific lineages have facilitated the development of high-fidelity modelling of several tissues *in vitro*. Many of the novel approaches have employed self-organizing three-dimensional (3D) culturing of organoids, which offer several advantages over conventional two-dimensional platforms. Organoid technologies hold great promise for modelling diseases and predicting the outcome of drug responses *in vitro*. Here, we outline the historical background and some of the recent advances in the field of three-dimensional organoids. We also highlight some of the current limitations of these systems and discuss potential avenues to further benefit biological research using three-dimensional modelling technologies.

This article is part of the theme issue 'Designer human tissue: coming to a lab near you'.

## 1. Introduction

The isolation and maintenance of mammalian cells have significantly advanced scientific research into cellular processes and mechanisms of disease, including stem cell development and differentiation, the production of monoclonal antibodies, and therapeutic proteins and for modelling cancer *in vitro* [1]. Although culturing tissues dates back to the late nineteenth century, present cell culture systems draw from studies on the action of serum on fibroblast cells [2] and the development of novel synthetic cell culture media [3,4]. A classic example of this was the isolation and expansion of HeLa cells from a cervical tumour on a two-dimensional monolayer culture [5].

Since then, culturing cells in two-dimensional format has remained the predominant methodology of *in vitro* cell growth and expansion. However, the two-dimensional platforms do not effectively recapitulate the spatial requirements that are essential for the organization and cellular interactions that occur *in vivo*. In addition, it is suspected that limited cell–cell contact and altered *in vitro* cell signalling networks can result in major discrepancies between the data acquired from two-dimensional *in vitro* versus *in vivo* research.

## 2. Historical background

To overcome two-dimensional platform limitations, efforts have led to the development of novel approaches to recreate a more physiologically relevant environment in the form of three-dimensional cell culture [1]. To successfully construct and maintain a three-dimensional structure, much research has been devoted to the development of synthetic or natural polymeric three-dimensional scaffolds to facilitate cell growth. These efforts have resulted in the fabrication and characterization of several non-degradable or biodegradable synthetic polymers, such as poly-lactic acid, poly-glycolic acid, poly-lactic-co-glycolic acid and

## Video Article

# Serum Free Production of Three-dimensional Human Hepatospheres from Pluripotent Stem Cells

Balta Lucendo-Villarin<sup>1</sup>, Hassan Rashidi<sup>1,2</sup>, Sharmin Alhaque<sup>1,3</sup>, Lena Fischer<sup>1,4</sup>, Jose Meseguer-Ripolles<sup>1</sup>, Yu Wang<sup>1</sup>, Cliona O'Farrelly<sup>4</sup>, Michael Themis<sup>3</sup>, David C. Hay<sup>1</sup>

<sup>1</sup>MRC Centre for Regenerative Medicine, University of Edinburgh

<sup>2</sup>UCL Great Ormond Street Institute of Child Health, University College London

<sup>3</sup>Division of Biosciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University London

<sup>4</sup>School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin

Correspondence to: David C. Hay at [davehay@talktalk.net](mailto:davehay@talktalk.net)

URL: <https://www.jove.com/video/59965>

DOI: [doi:10.3791/59965](https://doi.org/10.3791/59965)

Keywords: Developmental Biology, Issue 149, pluripotent stem cells, three-dimensional culture, hepatospheres, drug metabolism and protein secretion, stable cell phenotype, defined culture system

Date Published: 7/20/2019

Citation: Lucendo-Villarin, B., Rashidi, H., Alhaque, S., Fischer, L., Meseguer-Ripolles, J., Wang, Y., O'Farrelly, C., Themis, M., Hay, D.C. Serum Free Production of Three-dimensional Human Hepatospheres from Pluripotent Stem Cells. *J. Vis. Exp.* (149), e59965, doi:10.3791/59965 (2019).

## Abstract

The development of renewable sources of liver tissue is required to improve cell-based modelling, and develop human tissue for transplantation. Human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs) represent promising sources of human liver spheres. We have developed a serum free and defined method of cellular differentiation to generate three-dimensional human liver spheres formed from human pluripotent stem cells. A potential limitation of the technology is the production of dense spheres with dead material inside. In order to circumvent this, we have employed agarose microwell technology at defined cell densities to control the size of the 3D spheres, preventing the generation of apoptotic and/or necrotic cores. Notably, the spheres generated by our approach display liver function and stable phenotype, representing a valuable resource for basic and applied scientific research. We believe that our approach could be used as a platform technology to develop further tissues to model and treat human disease and in the future may permit the generation of human tissue with complex tissue architecture.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/59965/>

## Introduction

The ability of human pluripotent stem cells (hPSCs) to self-renew, whilst retaining pluripotency, provides an opportunity to produce human cell types and tissues on demand. hPSCs have been efficiently differentiated into hepatocyte-like cells (HLCs) using two-dimensional (2D) adherent culture systems<sup>1,2,3,4,5,6,7,8,9,10</sup>. These systems have been used to successfully model monogenic disease, virus lifecycle, drug induced liver injury (DILI), fetal exposure to toxins and non-alcoholic fatty liver disease (NAFLD)<sup>11,12,13,14,15</sup>. However, these models do possess some drawbacks, which limit their routine use. Those include fetal marker expression, unstable phenotype and poor tissue architecture<sup>16,17,18,19</sup>, which could also limit extrapolation to organ function in vivo.

To overcome these limitations, three-dimensional (3D) differentiation platforms have been developed to mimic in vivo tissue architecture. Although enabling, those approaches rely on the use of animal derived products and matrices to drive tissue genesis<sup>20,21,22</sup>, limiting scale-up and widespread application.

Here, we detail procedures to generate large quantities of 3D hepatospheres from hPSCs using defined materials and cell self-assembly. Notably, the tissue generated by our procedure remains functional for more than one year in cell culture and is capable of supporting liver function in vivo<sup>23</sup>.

In summary, our defined differentiation approach allows the generation of stable human hepatospheres from both human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs). We believe the described procedure represents a significant breakthrough in the generation of 3D hepatospheres for basic and applied scientific research.