Generation of an HVS-based Episomally Maintained Gene Delivery System for Reprogramming Adult Somatic Cells

Hannah Frances Brown

Submitted in accordance with the requirements for the degree of PhD

The University of Leeds
School of Molecular and Cellular Biology
Faculty of Biological Sciences

September 2012

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

First and foremost, I would like express my deepest gratitude to my supervisor, Professor Adrian Whitehouse. He accepted me on to this project and gave me renewed hope and a new direction for my PhD, without which I could not have hoped to succeed. Further to this, he has provided immense support and guidance throughout my time in this lab, and displayed remarkable patience in attempting to shape me into a true scientist.

Secondly, I would like to thank all the members of the Whitehouse and Hewitt labs, both past and present, for all their guidance and support. Thanks go to Brian, Dave H, Marko, Dave G, Jen, Laura, Sophie, Chris, Suhana, Toral, Andy, Jim, Stu, Susie and Adam. I have truly been very lucky in my colleagues, who have provided excellent banter throughout my PhD, and made the prospect of work in the morning bearable even when science appeared to have other ideas. In addition to my fantastic colleagues, I would also like to thank all the friends I have made during my PhD for their support. I have come to look upon this excellent bunch of people as an extended family, and they will all be truly missed.

In addition, I would like to thank the other members of the Virology group here at Leeds for providing advice and reagents. Thanks also go to Dr. Gareth Howell for his time and help with bio-imaging which was always given with incredible patience and an infallibly happy demeanour.

Thanks also go to my external collaborator, Dr. Christian Unger. His valuable advice and guidance, in addition to his very generous provision of reagents, was unselfishly given, and allowed this work to be possible. For his help, I am truly grateful.

Finally I would like to thank my family, particularly my parents, who have always supported and encouraged me in my every endeavour. This has, at times, meant bearing the brunt of my frustrations or often lifting my spirits when I could see no light at the end of the tunnel. For this, I am truly grateful.

Financial support for this project was provided by the BBSRC.

Abstract

iPSC technology has the potential to generate patient specific pluripotent cells for use in stem cell therapies and disease modelling. However, current reprogramming methods utilise retroviral vectors, which integrate into the host cell genome disrupting normal gene function. Transient gene delivery methods have been investigated as safer alternatives but demonstrate poor reprogramming efficiency. Therefore, there is a requirement for iPSC gene-delivery vectors which are capable of providing prolonged transgene delivery without integrating into the host cell genome.

Herpesvirus saimiri (HVS) is a prototype member of the *gamma-2 Herpesviridae*, and is capable of persisting as a non-integrated episome in dividing and differentiating cell populations, providing sustained transgene expression. Therefore, this thesis has explored the potential of HVS-based vectors for iPSC generation.

This thesis focuses on the generation of three HVS-based vectors expressing the iPSC reprogramming transgenes, Oct4, Lin28 and Nanog.

The potential of these HVS-iPSC vectors to reprogram both primary and cancerous cells has been investigated. Human primary Neural Stem Cells demonstrated cytopathology upon transduction with HVS-based vectors, indicating the need for further improvements to the biosafety of these vectors before they are suitable for the generation of clinical grade iPSCs. However, reprogramming attempts utilising the Ewing's sarcoma cell line, A673 cells, successfully generated induced pluripotent cancer stem cell (iPC)-like colonies upon transduction with all three recombinant vectors.

Results from detailed analysis of these colonies suggest some form of reprogramming has taken place, albeit incomplete, as indicated by elevated expression of Oct4, Rex1 and Klf4; in addition to positive alkaline phosphatase staining and SSEA4 expression. Furthermore, these iPC-like colonies were capable of differentiation down the ectodermal lineage, as evidenced by upregulation of MSX1, MAP2, and Nestin.

In conclusion, this thesis has demonstrated the potential of HVS-based reprogramming vectors through the generation of A673-iPCs.

Table of Contents

List of Figures

List of Tables

Abbreviations List

α	alpha
β	beta
γ	gamma
Δ	delta
κ	kappa
λ	lamda
Ω	ohm
+ve	positive
-ve	negative
%	percent
°C	degrees Celsius
a.a.	amino acid
Ab	antibody
ABC	ATP-binding cassette transporter protein
AML	acute myeloid leukaemia
Amp	ampicillin
BAC	bacterial artificial chromosome
BF	bright field microscopy
bFGF	basic fibroblast growth factor
bp	base pairs
$CaCl_2$	calcium chloride
CAS	chromosome association site
CBS	CTCF binding site
CmR	chloramphenicol resistance
CMV	cytomegalovirus

CML	chronic myeloid leukaemia
CO_2	carbon dioxide
CSC	cancer stem cell
CTCF	cellular insulator protein
dATP	deoxyadenosine 5′-triphosphate
dCTP	deoxycytidine 5′-triphosphate
dGTP	deoxyguanosine 5′-triphosphate
dTTP	deoxythymidine 5′-triphosphate
dNTP	deoxynucleoside 5′-triphosphate
DAPI	4′6′-diamidino-2-phenylindole
dd	double distilled
DE	delayed early
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
cDNA	complementary deoxyribonucleic acid
DNAse	deoxyribonuclease
DPBS	Dulbecco's phosphate buffered saline
ds DNA	double-stranded DNA
DTT	dithiothreitol
E	early
EB	embryoid body
EBV	Epstein Barr Virus
EBNA	Epstein Barr Virus nuclear antigen
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetra-acetic acid sodium salt

EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EMT	epithelial to mesenchymal transition
ESC	embryonic stem cell
ESFT	Ewing's sarcoma family tumours
EWS	Ewing's sarcoma
F-plasmid	fertility plasmid
FACS	fluorescence activated cellular sorting
FCS	foetal calf serum
FLI	Friend's leukaemia integration transcription factor
FLIP	FLICE inhibitory protein
g	gram
GAPDH	glyceraldehyde 3-phosohpate dehydrogenase
GFP	green fluorescent protein
gp	glycoprotein
H_2O	water
H-DNA	high-density DNA
HCl	hydrochloric acid
HDAC	histone deacetylase
HEK	human embryonic kidney cells
HIV	human immunodeficiency virus
HLH	helix-loop-helix
hr	hour
HRE	hormone response element
HRP	horseradish peroxidase
HSV	Herpes simplex virus

hTERT	human telomerase reverse transcriptase
HVS	Herpesvirus saimiri
HygR	hygromycin resistance
ICM	inner cell mass
IE	Immediate Early
Ig	Immunoglobulin
iPC	induced pluripotent cancer stem cell
iPSC	induced pluripotent stem cell
KanR	kanamycin resistance
kb	kilobase
kDa	kiloDalton
Klf4	Krueppel-like factor 4
KO	Knock-Out DMEM
KO RA	Knock-Out DMEM with 1μM all-trans retinoic acid
KSHV	Kaposi's sarcoma-associated Herpesvirus
kV	kilovolt
l	litre
L	Lin28
LB	Luria broth
L-DNA	low density DNA
LIF	leukaemia inhibitor factor
Lin28	lin28 homologue A
LMP	latent membrane protein
LTR	long terminal repeat
M	Molar
MAPK	mitogen activated protein kinase

MAP2	microtubule-associated protein 2
MCS	multiple cloning site
MEF	mouse embryonic fibroblasts
mg	milligram
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
min	minute
ml	millilitre
mmol	millimole
MOI	multiplicity of infection
mRNA	messenger RNA
MSC	mesenchymal stem cell
MSX1	msh homeobox 1
MuMLV	Moloney's murine leukaemia virus
Myc	myelocytomatosis protein
µF	microfarad
µg	microgram
µl	microlitre
µmol	micromole
N	Nanog
NaOH	sodium hydroxide
NaCl	sodium chloride
NFκB	nuclear factor κB
ng	nanogram
nm	nanometre
NODE	Nanog and Oct4 deacetylase-associated complex

NPC	neuroprogenitor cell
NSC	neural stem cell
NuRD	nucleosome remodelling and histone deacetylase complex
O	Oct4
Oct4	octamer-binding transcription factor 4
OD	optical density
OMK	owl monkey kidney cells
ORF	open reading frame
Ori	origin of replication
PARP	poly ADP ribose polymerase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFG	pulsed field gel
PFGE	pulsed field gel electrophoresis
pH	$-\log_{10}[H^+]$
PNET	primitive neuroectodermal tumour
pol	polymerase
p.i.	post-infection
PIP	phosphatidyl inositol phosphate
qRT-PCR	quantitative RT-PCR
Rb	retinoblastoma
RFP	red fluorescent protein
RNA	ribonucleic acid
RNAse	ribonuclease
r.p.m	revolutions per minute
RPMI	Roswell Park Memorial Institute medium

RRM	RNA recognition motif
RTA	replication and transcription activator
RT	reverse transcriptase
rt	room temperature
s	second
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SOC	super-optimal broth with catabolite repression
Sox2	SRY (sex-determining region Y)-box 2
SSEA4	stage-specific embryonic antigen 4
ssRNA	single-stranded RNA
STAT	signal transducer and activation of transcription
STP	saimiri transforming protein
TBE	tris-borate-EDTA buffer
TBS	tris-buffered saline
Tip	tyrosine kinase interacting protein
TNFβ	tumour necrosis factor β
TR	terminal repeat
TRAIL	TNF-related apoptosis-inducing Ligand
TRAF	tumour necrosis factor receptor-associated factors
Tris	tris (hydroxymethyl)- aminoethane
U	units
Utf1	Undifferentiated embryonic cell transcription factor 1
UTR	untranslated region
UV	ultra-violet
V	volt

VEGF	vascular endothelial growth factor
VLP	virus-like particle
VPA	valproic acid
v/v	volume per volume
VZV	Varicella Zoster virus
w/v	weight per volume

Bases

adenine

cytosine

guanine

thymine

Amino Acids

G	Gly (glycine)	P	Pro (proline)
A	Ala (alanine)	D	Asp (aspartate)
V	Val (valine)	E	Glu (glutamate)
L	Leu (leucine)	N	Asn (asparagine)
I	Ile (isoleucine)	Q	Gln (glutamine)
S	Ser (serine)	K	Lys (lysine)
T	Thr (threonine)	R	Arg (arginine)
C	Cys (cysteine)	H	His (histidine)
M	Met (methionine)	F	Phe (phenylalanine)
Y	Tyr (tyrosine)	W	Trp (tryptophan)

Chapter 1

Introduction

1 Introduction

1.1 Stem Cells

Embryonic stem cells (ESCs) are pluripotent in nature, possessing the ability to self-renew, and divide indefinitely. Furthermore, they are capable of differentiating into almost any cell type within the adult organism, including the extra-embryonic tissues (Macfarlan *et al.*, 2012).

Evidence for the existence of stem cells was originally identified by investigations using teratomas; tumours comprising various different tissue types. Experiments demonstrated that a single cell from a teratoma was capable of forming new tumours with all the different cell types identified in the original teratoma (Kleinsmith & Pierce, 1964). Furthermore, Evans et al isolated a subpopulation of cells from these tumours that possessed large nuclei and scant cytoplasm and were capable of differentiation *in vitro* and *in vivo* (Martin & Evans, 1974). These cells, termed Embryonal Carcinoma cells, paved the way for the identification of Embryonic Stem Cells (ESCs) from mice in 1981 (Evans & Kaufman, 1981; Martin, 1981) and later, from humans in 1998 (Thomson *et al.*, 1998).

ESCs are isolated from the Inner Cell Mass (ICM) of the pre-implantation blastocyst. The blastocyst, which forms 3-5 days post-fertilization, consists of the ICM in addition to a fluid-filled blastocyst cavity, termed the blastocoel, and the trophoblast which comprises a layer of cells surrounding the blastocoels and the ICM, which later forms part of the placenta (Figure 1.1).

ESCs self-renew their populations by asymmetric cell division, whereby a stem cell is capable of dividing into a two different daughter cells, one being pluripotent and the other differentiated. During differentiation ESCs respond to complex interacting signalling cascades which determine the fate of these cells, thereby allowing the ordered determination of different cell fates. ESCs then differentiate to produce progenitor cells, which produce differentiated progeny forming various organ tissues of the ectodermal, endodermal and mesodermal lineages. These progenitor cells, or transit-amplifying cells, divide more rapidly than ESCs, and proliferate before

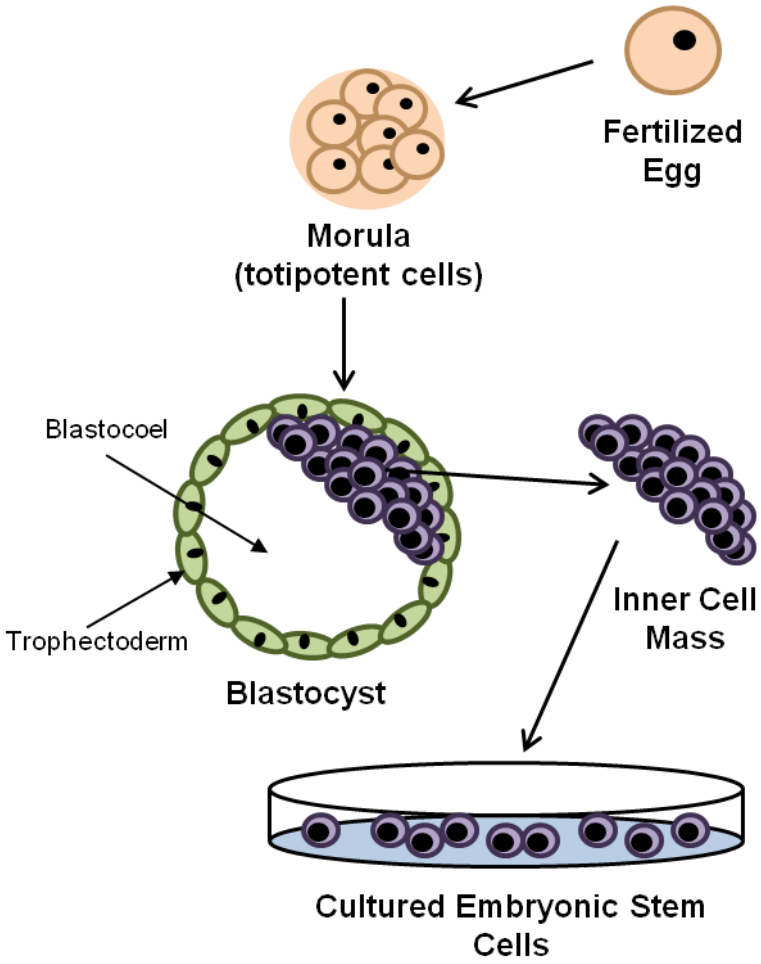


Figure 1.1: Isolation of ES cells from the Inner Cell Mass of the Blastocyst
5-8 days post-fertilization, the blastocyst forms. At the blastocyst stage of embryonic development, the embryo consists of a fluid-filled blastocoel, and an inner cell mass containing pluripotent stem cells. These are in turn surrounded by the trophoblast. Embryonic stem cells are obtained by removing the inner cell mass from the blastocyst and culturing these cells on mouse embryonic fibroblast feeder cell layers.

differentiation occurs, thereby amplifying the population of terminally differentiated cells

In addition to ESCs, populations of adult stem cells remain in most tissues. These adult stem cells are derived from ESCs, but their differentiation potential is limited to producing the various specialised cell types within the organ in which they reside, thus making these cells multipotent rather than pluripotent.

1.2 Stem Cell Promise in Therapy and Disease Modelling

Since their discovery more than 20 years ago, the unique ability of stem cells to generate differentiated tissues and whole organs has generated much

interest in their potential use in stem cell-based therapies. Indeed, there are already a number of adult stem-cell based therapies currently in use, and many registered clinical trials based on both adult and embryonic stem cells.

1.2.1 Adult Stem cell-based therapies

Most stem cell-based therapies which are currently in use or registered for clinical trials are predominantly based on adult stem cells. The first successful adult stem cell therapies were bone marrow transplants performed in 1968. This is still the most widely used adult stem cell-based therapy (Bach *et al.*, 1968; Gatti *et al.*, 1968). Since then, adult stem cells have been investigated for their therapeutic potential in a wide variety of disorders, and there are currently many adult stem cell-based clinical trials registered, including Mesenchymal Stem Cell (MSC)-based therapies for Type 1 diabetes, myocardial ischemia, liver cirrhosis and graft-versus-host disease (clinicaltrials.gov date accessed 10/08/2012). Bone marrow-derived stem cells are also currently being investigated for the treatment of cerebral ischemia (Abe *et al.*, 2012; Aurich *et al.*, 2009).

1.2.2 ESC-based Therapy

Due to their pluripotent nature, ESCs have an even greater number of potential uses in stem cell-based therapies. However, the use of ESCs in stem cell-based therapies is tightly regulated due to their highly dynamic nature and the ethical issues raised by their derivation from human embryos. These issues have limited the progress of ESC-based stem cell therapy clinical trials compared with those of adult stem cell-based therapies. However, there are many pre-clinical trials which have demonstrated the potential of ESCs in the treatment of neurodegenerative disorders. For instance, transplantation of ESC-derived dopaminergic neurones successfully restored motor function in rats and monkeys (Redmond *et al.*, 2007; Studer *et al.*, 1998). Further evidence of the potential of ESCs in the treatment of neurodegenerative diseases was highlighted in a Phase I/II clinical trial of ESC-derived motor neurones for the treatment of spinal muscular atrophy (Lunn *et al.*, 2011).

1.3 Safety and Ethical Issues with ESC-based therapies.

Despite the great potential of ESCs in stem cell-based therapies, there are a number of issues which currently make ESCs unsuitable for therapeutic use. Firstly, there are the ethical and religious implications of using cells derived from human embryos. Secondly, due to their highly dynamic nature, there is significant variability within and between master banks of stem cell lines, including differentiation potential and activity of the cells upon transplantation (Carpenter *et al.*, 2009). Thirdly, despite most stem cells being immune-privileged, there have been certain cases where tissue rejection has been reported (Swijnenburg *et al.*, 2008a; Swijnenburg *et al.*, 2008b). Finally, there are several issues with the biosafety of transplanted ESCs. Prolonged *in vitro* culture of ESCs results in certain genetic and epigenetic alterations, including for example, trisomy of chromosomes 12 and 17 (Buzzard *et al.*, 2004), and significant variability in the X-inactivation (Dhara & Benvenisty, 2004; Hoffman *et al.*, 2005). Furthermore, late-passage ESCs have been demonstrated to possess many mutations commonly observed in cancers (Maitra *et al.*, 2005). In addition to these acquired mutations, ESCs have the ability to form teratoma tumours upon transplantation into a host, and any stem cell-based therapy using differentiated cells derived from ESCs may contain contamination populations of undifferentiated ESCs which may cause tumour formation upon implantation.

However, despite these issues, there is still a vast amount of interest in the eventual use of these cells for both therapeutic uses and as models for development processes in health and disease.

1.4 Induced Pluripotent Stem Cell Technology

The advent of Induced Pluripotent Stem Cell (iPSC) technology gave renewed hope for stem cell-based therapies and disease modelling. This technology involves the reprogramming of somatic cells into a pluripotent, stem cell-like state. The cells produced from this method are not only ethically sourced, but can also be derived from individual patients, thereby avoiding tissue rejection upon implantation.

iPSCs were first generated in 2006 by Takahashi et al., who detailed in a seminal paper the minimal combination of factors required for reprogramming somatic mouse embryonic fibroblasts (MEFs) to ESC-like cells. Initially, the authors stably transfected MEF cells with a β-geo cassette, (a fusion of the β-galactosidase gene and the neomycin resistance gene), under the promoter of the Fbx15. Fbx15 is a gene that is expressed in ESCs but is dispensable for maintaining pluripotency. These cells were then transfected with 24 factors previously identified to be key regulators of pluripotency. Leaving out individual genes systematically revealed just four factors that were indispensible for reprogramming; namely Oct4, Sox2, Klf4 and c-Myc (Takahashi & Yamanaka, 2006). After this initial experiment, two additional papers were published simultaneously, detailing the reprogramming of human somatic fibroblasts to iPSCs using exogenous expression of reprogramming factors (Takahashi *et al.*, 2007; Yu *et al.*, 2007).

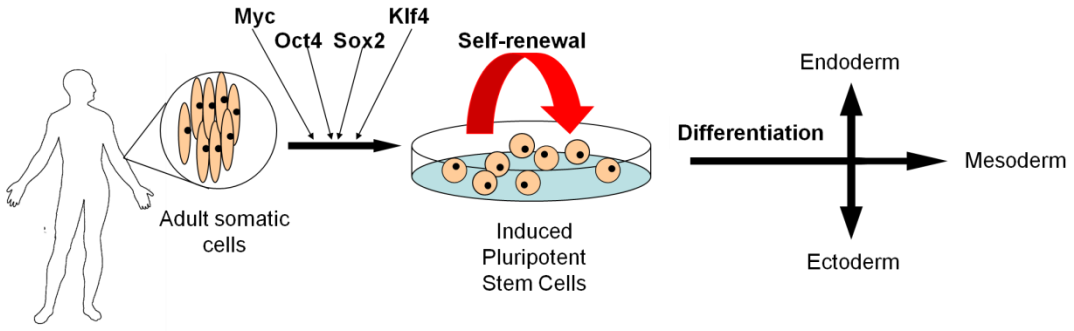


Figure 1.2: Overview of iPSC technology

Adult somatic cells are transduced with four reprogramming factors. This eventually reprograms these somatic cells to pluipotent, stem cell-like cells, which are capable of indefinite self-renewal, and can differentiate into cells comprising all three germlayers.

1.5 iPSCs in stem cell-based therapies

iPSCs are of immense interest for stem cell-based therapies as they provide the potential to generate large numbers of ethically sourced, patient specific pluripotent cells. This would therefore circumvent the issues of tissue-typing and cross-matching donors with recipients, and eliminate the possibility of tissue rejection or graft-versus-host disease.

Although there are significant hurdles to overcome before iPSCs are used in clinical therapies, a number of very promising pre-clinical trials have demonstrated the potential of these cells. For example, iPSC-derived dopaminergic neurones can integrate into the existing host brain neural networks and improve the motor function of Parkinson's disease model rats (Wernig *et al.*, 2008). Moreover iPSCs derived from both murine and human cells successfully engrafted and repaired motor function in Spinal Cord Injury mice models (Nori *et al.*, 2011; Tsuji *et al.*, 2010). In addition, Terzic et al. demonstrated that intramyocardially injected iPSCs were able to functionally and structurally repair infarcted hearts in mice (Nelson *et al.*, 2009), and Ma and colleagues utilised iPSC-derived endothelial cells expressing Factor VIII to rescue the Haemophilia Type A phenotype in mice models of the disease (Xu *et al.*, 2009).

A further potential therapeutic application is to combine iPSCs with gene therapy to repair genetic defects in patient-specific iPSCs, and transplant these cells back into the donor patient. There are number of ways utilised to repair the genomes of these iPSCs, including Zn-finger nucleases to target single point mutations, or homologous recombination. Homologous recombination can either be mediated by plasmid, virus or bacterial artificial chromosome-based vectors to repair homozygous mutations, or alternatively, the cells own mitotic-mediated recombination events can be used to repair heterozygous genetic disorders (Cheng *et al.*, 2012). This type of genetic correction of iPSCs has successfully been demonstrated for sickle-cell anaemia, whereby a humanized mouse model of sickle cell anaemia was rescued upon transplantation of iPSC-derived haematopoietic progenitor cells, previously repaired by homologous recombination (Hanna *et al.*, 2007).

1.6 Retroviruses and Lentiviruses are the predominant iPSC-gene delivery vectors

Viral vectors are ideal for efficient gene delivery. They are capable of receptor binding and internalisation, followed by transfer of transgene-encoding DNA to the nucleus and regulation of transcription (Wasungu & Hoekstra, 2006). Currently, the most commonly utilised gene delivery systems for iPSC generation are retrovirus-based vectors.

Retroviruses contain two copies of their positive sense ssRNA genome within enveloped particles, 80-100 nm in diameter (Dimmock *et al.*, 2001). There are seven different genera; *Alpharetrovirus, Betaretrovirus, Gammaretrovirus, Deltaretrovirus, Epsilonretrovirus, Lentivirus* and *Spumavirus*, although originally all genera, except *Lentivirus* and *Spumavirus*, belonged to the oncoretrovirus group due to their oncogenic nature. The most common retroviruses utilised for iPSC gene-delivery vectors are Murine Leukaemia Virus, (a *Gammaretrovirus*) and Lentiviruses.

The defining feature of retroviruses is their method of replication. Uniquely, their RNA genome is reverse-transcribed into a DNA pro-virus which then integrates into the host cell genome, mediated by the virally encoded enzymes, reverse transcriptase and integrase. Both of these are contained in the retroviral virions.

All oncogenic retroviruses comprise three genes, *gag, pol* and *env*, flanked by long terminal repeats (LTRs). The upstream LTR contains the U3 promoter, which drives transcription of these genes. Two differentially spliced mRNAs are produced, which in turn produce the Gag and Gag-Pol poly proteins, and the Env protein. The Gag and Gag-Pol polyproteins are self-proteolytically cleaved to produce the capsid proteins (produced by Gag) and the non-structural protease, reverse transcriptase and integrase proteins (produced by Pol). The env gene encodes the envelope glycoproteins that mediate viral entry (Figure 1.3). Lentiviruses and Spumaviruses contain six

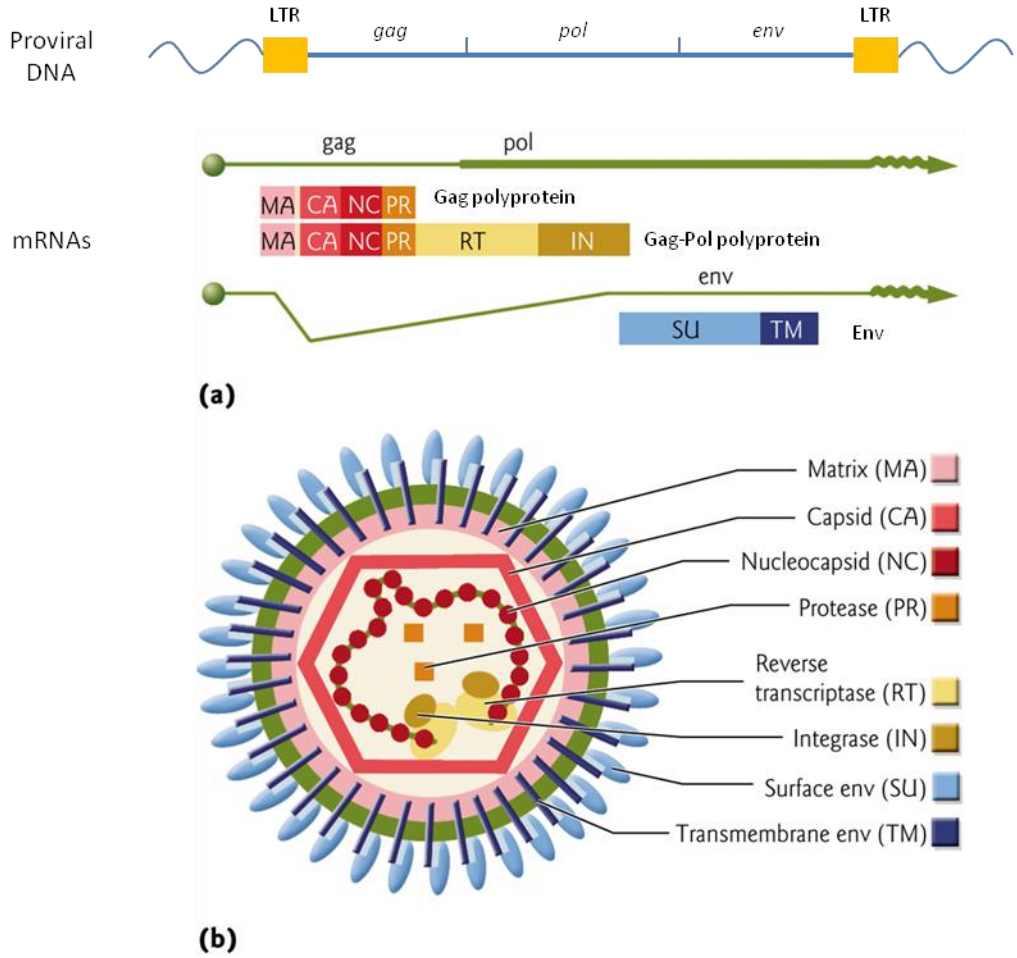


Figure 1.3: The structure of the retrovirus capsid and genome organisation.
a) After the retroviral genome has been reverse transcribed, it integrates into the host genome using the virally encoded protein, Integrase. This proviral DNA, flanked by the Long Terminal Repeat (LTR) sequences, produces two distinct mRNAs by means of alternate splicing events. The gag-pol mRNA produces two poly proteins, the Gag-polyprotein, which encoded the capsid proteins, and the Gag-Pol protein. The Gag-Pol protein encoded a viral protease, reverse transcriptase and integrase proteins, in addition to the structural proteins encoded by Gag. The env mRNA produces two proteins which are both glycoproteins and are expressed on the surface of the viral envelope. b) The completed retroviral virion structure. Adapted from (Dimmock *et al.*, 2006).

additional accessory genes (*tat, rev, nef, vif, vpr, vpu*) and three extra genes (*bel-1, bel-2 bel-3*) respectively.

Retroviruses were considered very attractive gene therapy vectors due to their ability to integrate into the host genome, thereby enabling long-term transgene expression. The most common retroviral vectors are based on the Gammaretrovirus, Moloney Murine Leukaemia Virus (MuMLV) or on the Lentivirus, Human Immunodeficiency Virus (HIV) (McTaggart & Al-Rubeai, 2002). Although MuMLV was the first virus to undergo clinical trials (Blaese

et al., 1995), these viral vectors had the disadvantage that they, like all oncoretroviruses, were only capable of integrating and expressing their transgene in dividing cells, thereby making them unsuitable for therapies involving certain tissues such as neurones. For this reason, lentiviruses, which can infect both dividing and non-dividing cells, may be a more practical alternative.

Retroviral vectors are produced by replacing the genes required for viral replication with the transgene of interest. Only the LTR regions and the *psi* (Ψ) packaging sequences remain, as these are essential for virus encapsidation. Because these modified viruses are rendered replication incompetent, their production requires the use of special packaging cell lines which contain the viral structural and functional genes on complementing expression cassettes (Figure 1.4). To further improve the biosafety of these vectors, the gag-pol and env genes are encoded from separate expression cassettes. This reduces the risk of producing replication competent retrovirus particles by homologous recombination (McTaggart & Al-Rubeai, 2002).

However, viral vectors possess certain drawbacks compared to non-viral gene delivery vectors. These include toxicity, immune recognition of virally expressed proteins, recombination with wild-type viruses and oncogenicity of virally encoded proteins involved in cell cycle regulation. A further disadvantage of retroviral vectors is the risk of mutagenesis caused by insertion of the viral genome into transcriptionally active areas of the host genome (Scherdin *et al.*, 1990). Therefore, a number of alternative methods to derived iPSCs have been assessed.

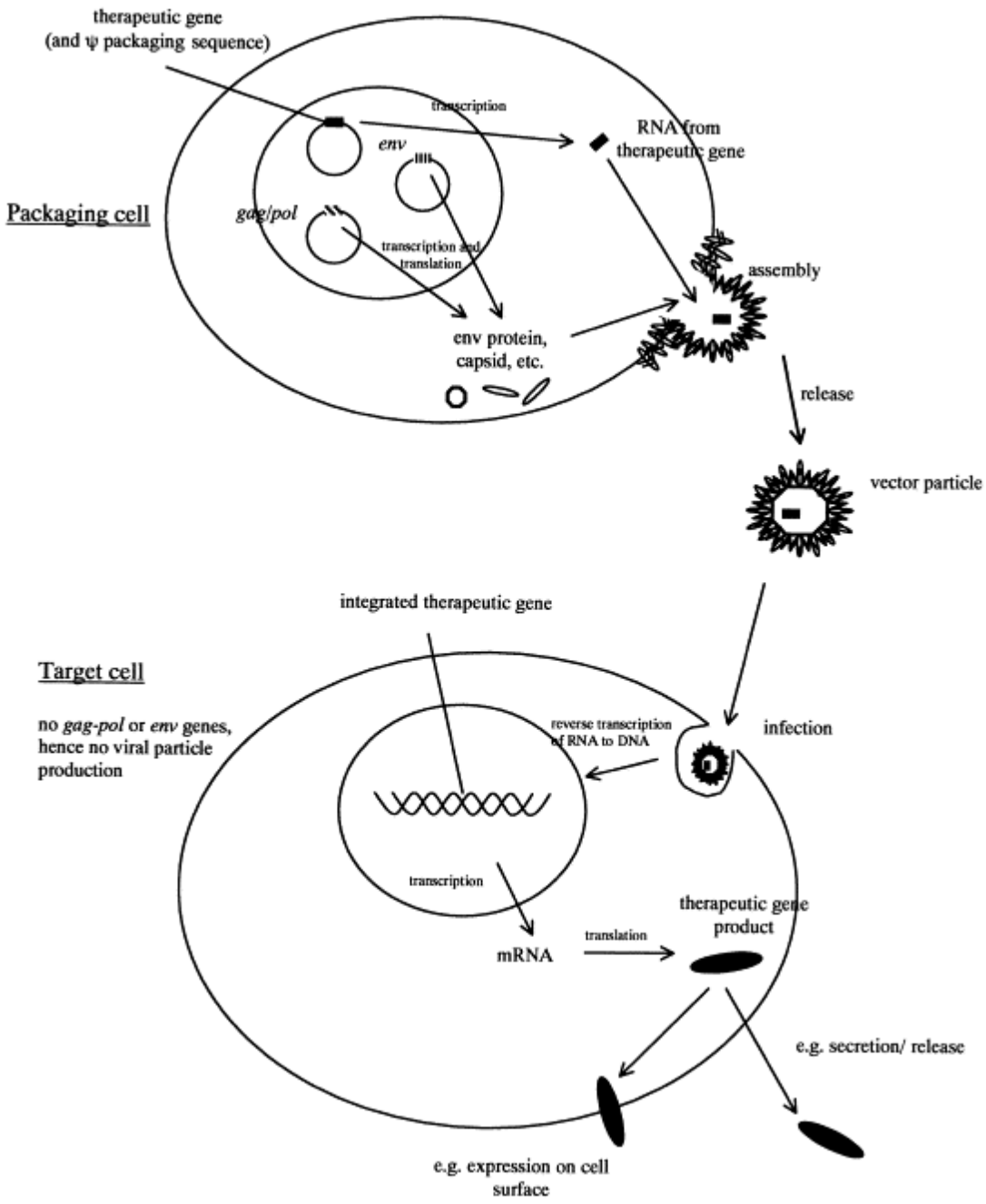


Figure 1.4: Production of retroviral gene delivery vectors.

Packaging cell lines are transfected with expression cassettes containing the gag-pol and env coding sequences, in addition to the vector cassette. Vector cassettes contain the transgene of interest, flanked by the LTR regions and the Ψ packaging sequences, and are the only cassette to be encapsidated. Viral structural and functional proteins required for infectious virus production are therefore provided by the packaging cell. The infectious virus particles are capable of infecting target cells, integration and transgene delivery, but are incapable of viral replication. Taken from (McTaggart & Al-Rubeai, 2002).

1.6.1 Alternative Gene Delivery Approaches

There have been a number of attempts to generate safer reprogramming methods, which are detailed in Table 1.1. Initially, systems were developed to excise the retroviral genome using Cre-LoxP recombination, but this system left a viral 'footprint' at the site of integration after excision, which could still potentially disrupt normal host gene function (Soldner *et al.*, 2009). Reversibly integrating piggy-Bac transposons

		Reprogrammed Cells	Factors used	Efficiency	Advantages	Disadvantages
Excisable	Transposon	Fibroblasts	OSKM	0.1	Efficient, no integration	Labour-intensive screening of colonies
	LoxP-flanked lentiviral genome	Fibroblasts	OSK	0.1 - 1	Reasonably efficient and no genomic integration	Labour-intensive screening of colonies, LoxP sites retained in the genome
Non-integrating	Adenoviral	Fibroblasts, liver cells	OSKM	0.001	no genomic integration	Low efficiency
	Plasmid	Fibroblasts	OSNL	0.001	Only occasional genomic integration	Low efficiency, requirement for repeated transfections, occasional genomic integration
Episomal	EBV EBNA1		OSKMN + SV40 LT	0.000003	no genomic integration	extremely low efficiency, EBNA1 potentially oncogenic
DNA free	Sendai Virus	Fibroblasts	OSKM	1	no genomic integration	sequence sensitive RNA-replicase, difficulty in removing virus after reprogramming
	Recombinant protein	Fibroblasts	OS	0.001		low efficiency, short half-life, requirement for multiple applications of the proteins
	Modified mRNA		OSKM or OSKML + VPA	1-4.4	No genomic integration, bypasses innate antiviral response, faster reprogramming kinetics, controllable and high efficiency	Requirement for multiple rounds of transfection
	MicroRNA		miR-200c, miR-302s or miR-396	0.1	Efficient, fast reprogramming, no exogenous transcription factors, no genomic integration	Low efficiency

Table 1.1: Safer gene delivery methods used in iPSC generation.

Adapted from (Robinton & Daley, 2012). Factors used: O – Oct4; S – Sox2; K – Klf4; M – Myc; N – Nanog; L – Lin28; VPA – Valproic Acid.

have also been used to deliver iPSC genes. This system can subsequently be excised from the genome without leaving any remaining exogenous DNA sequences. However, this system requires stringent genome-wide screening to ensure all copies of the transposon have been removed upon excision (Kaji *et al.*, 2009; Woltjen *et al.*, 2009; Yusa *et al.*, 2009).

Alternatively, transient methods have also been used, including repeated plasmid transfection, (Okita *et al.*, 2008) non-integrating adenoviral vectors, (Stadtfeld *et al.*, 2008) cell-permeable recombinant protein transduction, (Zhou *et al.*, 2009b) and recently, transfection with mRNAs (Warren *et al.*, 2010), or mature miRNAs (Miyoshi *et al.*, 2011).

However, these transient reprogramming methods have poor efficiencies, which further highlight the need for prolonged transgene expression in iPSC generation. Therefore, the ideal reprogramming vector would be capable of providing prolonged expression of the reprogramming transgenes without integrating into the host genome.

To provide sustained transgene expression in a dividing cell population, the Epstein-Barr virus latency associated nuclear antigen (EBNA-1) has also been used in a plasmid-based episomally maintained system to successfully reprogramme fibroblasts (Yu *et al.*, 2009). However, this study chose to use inefficient plasmid transfection to deliver the reprogramming factors rather than utilising the natural infectivity of EBV. Furthermore, EBNA-1 has been shown to immortalise B cells *in vitro*, thus questioning the safety of this method (Young *et al.*, 1988). However, this study did highlight the advantages of episomally maintained reprogramming systems, therefore warranting further investigation of such vectors for iPSC generation.

1.6.2 Replacing the Exogenous Reprogramming Factors

One of the intrinsic safety issues of iPSC technology is the ectopic expression of the specific factors required for reprogramming. All of these factors have been identified to be upregulated in a variety of different cancers, with Klf4 and c-Myc being well described as potent oncogenes (Sun & Liu, 2011). Therefore alternative reprogramming factors to replace Klf4 and c-Myc were sought.

In 2007, Thompson and colleagues discovered that c-Myc and Klf4 could be replaced with the non-oncogenic pluripotency factors, Nanog and Lin28 (Yu *et al.*, 2007). However, although this process improved the biosafety of the resulting iPSCs, reprogramming efficiency was reduced. Furthermore, this method still required transduction of somatic cells with four reprogramming factors, which further reduces the efficiency of iPSC generation. Therefore, researchers attempted to eliminate the requirement of multiple reprogramming factors. Kim and colleagues investigated the reprogramming potential of other cell lines. Specifically, they attempted to reprogramme Neural Stem Cells (NSCs) using fewer factors. Interestingly, NSCs can be reprogrammed to iPSCs using only exogenous Oct4 expression (Kim *et al.*, 2009b).

Furthermore, several groups have identified various small molecules capable of replacing certain reprogramming factors during iPSC generation. For instance, Huangfu and colleagues demonstrated that the histone deacetylase inhibitor, Valproic Acid (VPA), enabled the reprogramming of human fibroblasts using exogenous expression of only Oct4 and Sox2 (Huangfu *et al.*, 2008). In addition, iPSCs were successfully generated from keratinocytes using exogenous expression of Oct4 and Klf4 in conjunction with parnate, (an inhibitor of lysine specific demethylase 1) and an inhibitor of GSK3 signalling (Li *et al.*, 2009). Further work detailed the generation of human iPSCs from keratinocytes with exogenous Oct4 expression combined with a cocktail of small molecule inhibitors (Zhu *et al.*, 2010).

Interestingly, another small molecule, Nr5a2, has been demonstrated to be capable of replacing the activity of Oct4 in the generation of iPSCs (Heng *et al.*, 2010). However, the combination of Nr5a2 and the molecules defined by Zhu et al. is unable to generate iPSCs.

1.7 Oct4, Sox2 and Nanog are Key Regulators of Pluripotency

The protein-protein networks and signalling pathways governing pluripotency and self-renewal in stem cells are diverse and highly complex. As such, these mechanisms are still poorly understood despite their immense

importance in the maintenance of stem cells. However, it is clear that there are three proteins which are central to the complex interaction networks governing pluripotency. These key regulators are the transcription factors, Oct4, Sox2 and Nanog, and a simplified diagram of the key protein-protein interactions governing pluripotency is shown in Figure 1.5.

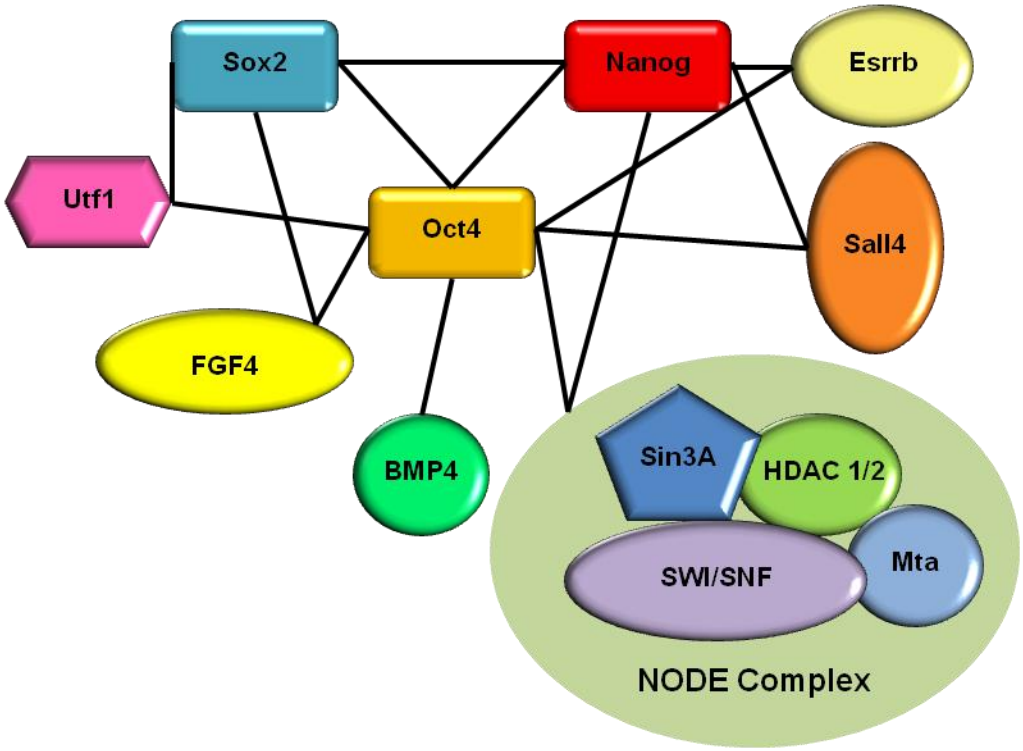


Figure 1.5: The protein-protein networks governing pluripotency.
The network of protein-protein interactions involved in stem cell self-renewal and pluripotency are governed by three proteins, Oct4, Sox2 and Nanog. These proteins are capable of regulating the expression of a myriad of different genes involved in differentiation and proliferation. In addition, these factors are capable of regulating their own expression to fine tune the mechanisms involved in pluripotency. Oct4 and Nanog also form a histone remodelling complex with several members of the NuRD complex.

1.7.1 Oct4

Oct4 is a homeodomain transcription factor of the POU family of proteins. It recognises the consensus octamer binding sequence of 5'-AGTCAAAT (Herr & Cleary, 1995). The role of Oct4 in stem cell biology was highlighted using mouse embryos containing Oct4 null mutations which were capable of developing into blastocysts, however the cells within the inner cell mass

were not pluripotent, and differentiated down the extraembryonic trophoblast lineage (Nichols *et al.*, 1998). Oct4 remains the only irreplaceable factor for reprogramming, with one exceptional case of its replacement using Nr5a2. Nr5a2 is an activator of Oct4 and enhances reprogramming efficiency partly through activation of Nanog (Heng *et al.*, 2010).

There are four Conserved Regions identified in the upstream regulatory region of Oct4, CR1 – CR4, which have implications in Oct4 regulation. For instance, CR1 has a putative Sp1 binding site which is overlapped by a Hormone Responsive Element (HRE), which allows Retinoic Acid to shut off transcription of Oct4, and CR4 has a ATTA sequence, which is recognised by many transcription factors, including Nanog (Loh *et al.*, 2006; Nordhoff *et al.*, 2001).

1.7.2 Sox2

Sox2 also plays an important role in reprogramming and regulating pluripotency. It is a member of the High Mobility Group (HMG) of protein transcription factors. Sox2 null embryos develop an inner cell mass, but fail to maintain the epiblast (Avilion *et al.*, 2003), although Sox2 null embryos are still capable of maintaining some degree of pluripotency. This may be due to residual maternal Sox2 and the partial compensation for Sox2 activity by Sox3 (Avilion *et al.*, 2003; Wang *et al.*, 2012).

Sox2 forms a heterodimer with Oct4 (Ambrosetti *et al.*, 1997), which is then capable of synergistically binding to, and regulating transcription from a highly conserved Oct-Sox binding element, CTTTTGCATTACAATG. This Sox2-Oct4 complex can then upregulate FGF4 and Utf1 expression by differential binding to enhancer elements within the promoter regions of these genes (Nishimoto *et al.*, 1999; Yuan *et al.*, 1995). In addition, this complex can also regulate expression of other genes associated with pluripotency, including Nanog (Rodda *et al.*, 2005), and Oct4 and Sox2 themselves (Okumura-Nakanishi *et al.*, 2005; Tomioka *et al.*, 2002). Sox2 is also capable of fine-tuning its own expression via an interaction with the poly ADP-ribose polymerase 1 protein (PARP1), which occurs in response to signals from FGF4. The PARP1-Sox2 complex prevents binding of Sox2 to the Oct-Sox element, thereby repressing Sox2 expression (Lai *et al.*, 2012).

1.7.3 Nanog

The final protein involved in governing the pluripotency networks which will be discussed herein is Nanog. Nanog is another homeodomain-containing protein which is crucial for maintaining pluripotency. It is capable of maintaining stem cell self-renewal even in the absence of Leukaemia Inhibitor Factor (Chambers *et al.*, 2003). Nanog null embryos develop an inner cell mass but this fails to become an epiblast. Instead, these cells differentiate into parietal endoderm-like cells (Mitsui *et al.*, 2003). Furthermore, Nanog is responsible for repressing differentiation into neuroectoderm and neural crest cells, but has no effect on repressing differentiation into other lineages (Wang *et al.*, 2012).

Like Oct4 and Sox2, Nanog also regulates pluripotency via a number of protein-protein interactions. For example, Nanog and Oct4 can recruit other proteins to form the Nanog and Oct4-associated Deacetylase (NODE) complex. This complex contains certain members of the NuRD and Sin3A complexes; Mta1, HDAC1/2 and Sin3a, and is capable of repressing genes involved in differentiation, thereby promoting self-renewal and maintaining pluripotency (Liang *et al.*, 2008). In addition to the NODE complex, Nanog mediates pluripotency by regulating the expression of Oct4 and Sox2 (Loh *et al.*, 2006).

1.7.4 Pluripotency Interactomes

There have been several studies aimed at identifying the various interaction partners for Oct4, Sox2 and Nanog. These methods predominantly rely on tagging the protein of interest and affinity purification of any complexes formed (Pardo *et al.*, 2010; van den Berg *et al.*, 2010; Wang *et al.*, 2006). However, these methods have several drawbacks, including the detection of ‘false-positives’ caused by optimisation to detect protein-protein interactions and certain missing interactions. This is highlighted by Nanog, which is frequently not detected in Oct4 interactome lists, but has previously been shown to form a complex with Oct4 (Zhang *et al.*, 2007).

These missing links may be caused by the lower levels of expression of certain interaction partners, but as many transcription factors are expressed

at low levels, this may cause many of these potential interaction partners to be missed.

1.7.5 Post-translational modifications of Pluripotency Factors

The complex protein-protein interactions governing pluripotency are made even more intricate by post-translational modifications. The master regulator of pluripotency, Oct4, is phosphorylated at two residues, T234 and S235. Mutational analysis of these residues revealed that these phosphorylation sites are important for sequence-specific DNA binding by Oct4 (Brumbaugh *et al.*, 2012). Furthermore, O-linked β-D-N-acetylglucoamine modification of Oct4 and Sox2 has been shown to be important in maintaining pluripotency, and is rapidly lost upon differentiation (Jang *et al.*, 2012).

1.8 Epigenetic Gene Regulation

The proteins involved in pluripotency regulation, such as those discussed in section 1.8, are all capable of causing epigenetic alterations in cellular gene expression. Gene expression can be regulated by several of epigenetic mechanisms which are discussed here.

1.8.1 Histone Modification

Within the nucleus, DNA is condensed by through winding around nucleosomes; octamer complex of four histone proteins (two H2A-H2B heterodimers, and two H3, H4 heterodimers). DNA winds around these nucleosomes 1.65 times (147 bps), and a final histone, H1, binds to the linking DNA between nucleosomes, thereby allowing higher order chromatin structures to form. The histones within the nucleosome have extended N-terminal tails which can be modified in a number of ways, including acetylation, methylation, phosphorylation and ubiquitination. These histone tail modifications have varying effects on gene expression, and also depend on the specific lysine or arginine they affect. Generally, acetylation of histone tails results in disruption of the wound DNA, leading to less-densely packed chromatin (euchromatin) and results in gene activation. Acetylation of histone tails is performed by Histone Acetyl-Transferases (HATs). Conversely, Histone Deacetylase enzymes (HDACs) remove acetyl groups

from histone tails, resulting in more densely packed chromatin (heterochromatin) and gene repression (Alberts, 2002).

1.8.2 DNA Methylation

DNA methylation occurs on cytosine residues located next to G residues (CpG sites), forming 5-methylcytosine. CpG sites occur within promoters of various genes, and are maintained by maintenance methyltransferases which methylate CpG sites on newly synthesised DNA strands according the methylation pattern of the parental strand. Shortly after fertilisation, de novo DNA-methyltransferases begin to methylate specific CpG sites, and during development, increasing numbers of promoter regions become more heavily methylated. Proteins containing a Methyl CpG Binding Domain, (such as MeCP2), are capable of binding to methylated DNA and recruit histone remodelling complexes, thereby resulting in heterochromatin formation and gene silencing (Alberts, 2002).

1.8.3 MicroRNAs (miRNAs)

The role of microRNAs in gene regulation is widely varied between cell types. microRNAs are transcribed by either RNA polymerase II or III and are initially formed as large primary miRNA transcripts (pri-miRNAs which contain a number of hairpin loops). These hairpin loops of the pri-miRNAs are cleaved in the nucleus by Drosha to produce 70 bp pre-miRNAs stem-loops which are exported by Exportin-5 into the cytoplasm. Within the cytoplasm, these pre-miRNAs are then further cleaved by Dicer to produce a 22bp RNA duplex of which one strand is incorporated into the Argonaut 2 (Ago2) subunit of the RNA-Induced Silencing Complex (RISC). This single-stranded miRNA can then base-pair with imperfect complementarity to a number of target mRNAs and cause transcriptional repression either by mRNA cleavage, mRNA decay by deadenylation, or by transcriptional repression caused by interfering with ribosomal binding to target mRNAs (Kim, 2005; Winter *et al.*, 2009). An overview of miRNA processing is shown in Figure 1.6.

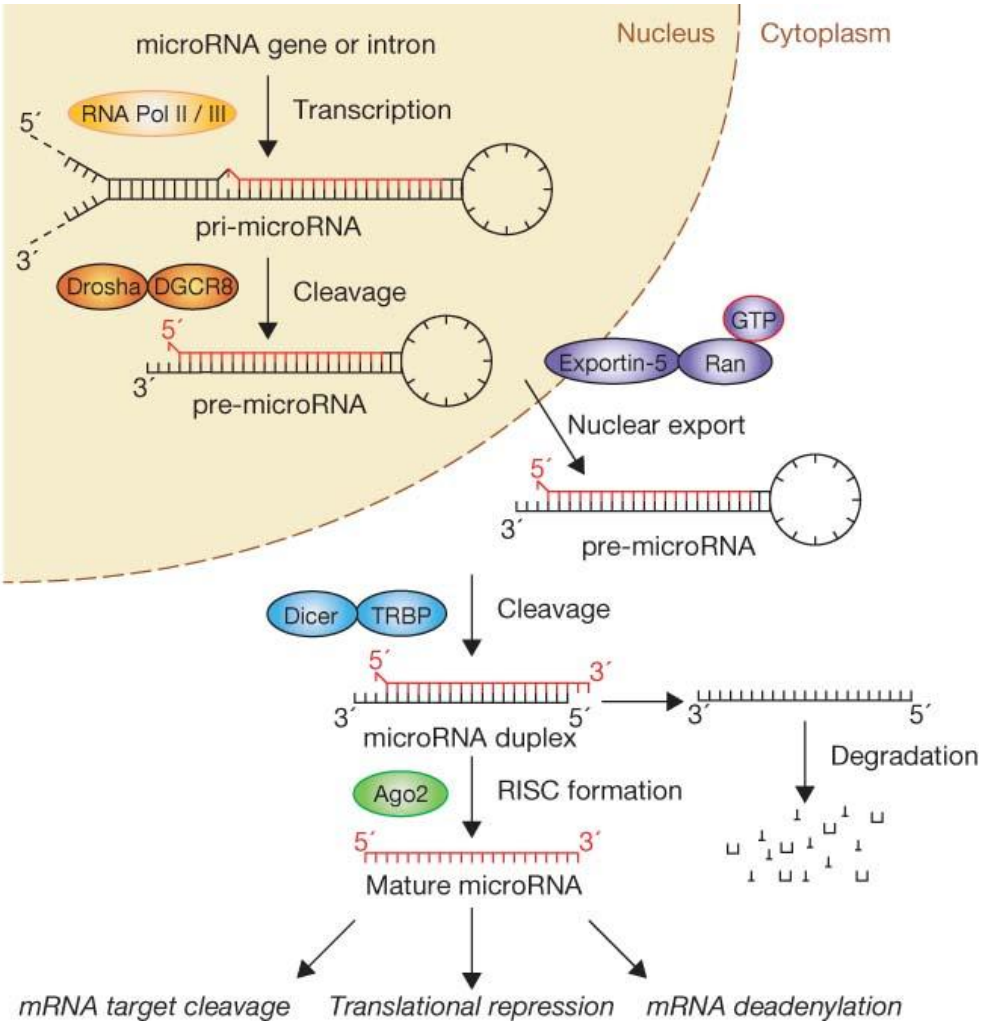


Figure 1.6: miRNA processing pathway (taken directly from (Winter *et al.*, 2009)).
miRNAs are expressed in the nucleus as large pri-miRNAs containing several hairpins. Pri-miRNAs are then cleaved in the nucleus by Drosha to produce 70bp hairpin stem-loops termed pre-miRNAs. These pre-miRNAs are exported into the cytoplasm by Exportin-5, where they are further cleaved by Dicer to form 22bp RNA duplexes. One strand of this RNA duplex is incorporated into the RNA-Induced Silencing complex and can therefore repress expression of a number of target mRNAs.

1.9 miRNA profiles in ESCs

The importance of miRNAs in pluripotency and differentiation regulation was demonstrated in Dicer knock-out mice. These mice were embryonically lethal yet produced viable ESCs. However, these ESCs possessed limited differentiation capabilities, thereby highlighting a role of miRNAs in pluripotency and regulation of differentiation (Bernstein *et al.*, 2003; Kanellopoulou *et al.*, 2005).

Since this initial observation a number of miRNA clusters have been identified as being upregulated in ESCs and rapidly downregulated upon differentiation. One such cluster is the miR-302-367 cluster. These miRNAs are activated by the pluripotency factors, Oct4, Sox2 and Nanog and have been shown to promote self-renewal by repressing Cyclin D and promoting entry into the S-phase of the cell cycle, causing genome-wide demethylation, and maintaining pluripotency (Barroso-del Jesus *et al.*, 2009; Card *et al.*, 2008; Lin *et al.*, 2008).

1.10 The Flaws of iPSC Technology

Initially, iPSC technology held great promise for stem-cell based therapies. However, there is increasing evidence to demonstrate that iPSC-based therapies are not currently feasible due to the biosafety concerns of iPSCs. Aside from the risks associated with insertional mutagenesis from retroviral gene delivery vectors and the ectopic expression of the oncogenic reprogramming factors, the reprogramming process itself can be oncogenic, inducing small nucleotide polymorphisms and copy number variations in resulting iPSCs (Laurent *et al.*, 2011; Pasi *et al.*, 2011). In addition, the starting somatic cells used for iPSC generation often have mutations which are harmless due to the limited proliferative potential of these differentiated cells. However when these cells are reprogrammed to a pluripotent cell with unlimited self-renewal, these mutations have the potential to cause abnormal gene expression and potentially aberrant proliferation.

iPSCs are also epigenetically different from ESCs, and possess aberrant methylation patterns, indicative of a somatic cell memory. One study analysed four iPSC cell lines derived by separate groups using various reprogramming protocols. These protocols included the use of either inducible or constitutive iPSC transgene expression in addition to utilising either permanently integrated, Cre-excisable or episomal reprogramming vectors. This study identified 15 genes which all four iPSC lines expressed at different levels to ESCs, indicating some functional similarity and potentially indicating an iPSC-specific gene expression profile distinct from that of ESCs or parental fibroblasts (Chin *et al.*, 2009). Furthermore, although comparisons of ESC and iPSC gene expression also revealed that

most miRNAs expression levels were identical between these cells, there were a number of these non-coding RNAs which were expressed at different levels, including miR-302b, which has important implications in the regulation of self-renewal and pluripotency (Chin *et al.*, 2009; Wilson *et al.*, 2009).

iPSCs derived from female donors can also retain a partially or completely silenced X-chromosome (Lister *et al.*, 2011; Pick *et al.*, 2009; Tchieu *et al.*, 2010). Interestingly, the aberrant methylation patterns found in iPSCs are maintained upon differentiation of these cells (Lister *et al.*, 2011). This may add to the oncogenic potential of iPSCs upon implantation, as it is speculated that some cancers may be caused by changes at the epigenetic level (Feinberg *et al.*, 2006).

Perhaps the biggest disappointment in iPSC technology was the immune rejection observed in a strain of in-bred mice. Specifically, iPSCs derived from these mice were transplanted back into the same inbred line, However, the iPSCs had very poor engraftment success and tissue rejection was also observed (Zhao *et al.*, 2011)

1.11 Disease-in-a dish models

Because of the issues detailed above in section 1.9 of iPSCs, their application in stem cell-based therapy may be a very distant promise. Indeed, it is now unclear whether these cells will ever be suitable for therapeutic applications. However, iPSC technology does provide an abundant source of pluripotent cells which can be derived from specific patients. These cells are therefore ideal to provide *in vitro* models for a myriad of human genetic disorders which were previously modelled in animals. Animal models have many limitations, primarily because of interspecies variations which can affect the signalling interactions involved in disease progression. Furthermore, many genetic disorders do not naturally occur in these animal models, and thus have to be recreated artificially.

iPSCs have therefore been used to model a wide variety of genetic disorders (Table 1.2). This ‘disease-in-a-dish’ modelling is perhaps particularly

Disease	Cell type differentiated from iPSCs	Disease phenotype demonstrated in Differentiated cells
Neurological		
Alzheimer's disease	Heterogeneous population of neurones	Yes
Amyotrophic lateral sclerosis (ALS)	Motor neurons and glial cells	ND
Spinal muscular atrophy (SMA)	Neurons and astrocytes, and mature motor neurons	Yes
Parkinson's disease	Dopaminergic neurons	No
Huntington's disease	None	NA
Down's syndrome	Teratoma with tissue from each of the three germ layers	Yes
Familial dysautonomia	Central nervous-system lineage, peripheral neurons, haematopoietic cells, endothelial cells and endodermal cells	Yes
Schizophrenia	Neurons	Yes
Haematological		
ADA SCID	None	ND
Sickle-cell anaemia	None	NA
Metabolic		
Type 1 diabetes	β-Cell-like cells (express somatostatin, glucagon and insulin; glucose-responsive)	ND
Glycogen storage disease1a (GSD1a)	Hepatocyte-like cells (foetal)	Yes
Familial hypercholesterolaemia	Hepatocyte-like cells (foetal)	Yes
Cardiovascular		
LEOPARD syndrome	Cardiomyocytes	Yes
Type 1 long QT syndrome	Cardiomyocytes	Yes
Type 2 long QT syndrome	Cardiomyocytes	Yes
Other category		
Duchenne muscular dystrophy	None	NA
Cystic fibrosis	None	NA
Retinitis pigmentosa	Retinal progenitors, photoreceptor precursors, retinal-pigment epithelial cells and rod photoreceptor cells	Yes
Osteogenesis imperfect	None	NA

Table 1.2: iPSC models of genetic diseases.

Adapted from (Robinton & Daley, 2012).

important for neurodegenerative diseases, where isolating tissue samples for research purposes is generally not feasible from living patients, and there are limited suitable animal models. Neurological diseases whose progression has been modelled with iPSCs include Parkinson's disease (Nguyen *et al.*, 2011; Soldner *et al.*, 2009), Alzheimer's disease (Israel *et al.*, 2012). Smooth muscle atrophy (Ebert *et al.*, 2009) and familial dysautonomia (Lee *et al.*, 2009). Often, the resulting differentiated cells produced from these disease-specific iPSCs possess the disease phenotype. This therefore gives vital insights into the progression of these genetic diseases.

Furthermore, these 'disease-in-a-dish' models allow safer drug screening *in vitro* in human models of the genetic disease, rather than less physiologically relevant animal models. Indeed, iPSC cardiomyocyte models of the genetic heart defect, Type 2 long QT syndrome, demonstrated that certain drugs aggravated the condition, highlighting a complex cardiotoxic effect of these drugs (Itzhaki *et al.*, 2011). Screening drugs in these functional *in vitro* models could therefore reduce the risk of drugs inducing arrhythmias prior to market release.

1.12 Pluripotency Mechanisms and Cancer

The pluripotency networks of stem cells enable these cells to maintain their undifferentiated state, and allow these cells an unlimited capacity for cellular division. As such, many of the pluripotency mechanisms are also found to be deregulated in cancers, and the disruption of various tumour suppressor genes has also been demonstrated to enhance reprogramming efficiency (Bernhardt *et al.*, 2012). An overview of the links between pluripotency and cancer are shown in Figure 1.7.

For example, ectopic expression of Oct4 has been shown to increase the malignant phenotype of ESCs. Specifically, mice injected with an ES-cell line carrying an inducible Oct4 expression cassette demonstrated that Oct4 had the potential to cause a malignant phenotype. Moreover, repression of Oct4 resulted in a 10-fold decrease in tumour formation of these cells (Gidekel *et al.*, 2003). Furthermore, Oct4 and Sox2 have been found to be upregulated in a number of cancer tissues, including bladder cancer, prostate cancer,

leukaemia, glioblastoma, primary colon cancer, hepatocellular carcinoma, bladder carcinoma, pancreatic cancer and breast cancer (Sun & Liu, 2011).

Signalling pathways which are involved in self-renewal and development are often upregulated in many cancers. Sonic Hedgehog, which is involved in cortical development, induces basal cell carcinoma of the skin if overexpressed (Mullor *et al.*, 2002), and Notch signalling, which controls the inhibition of many differentiation pathways, has the potential to immortalise haematopoietic cells if constitutively activated (Varnum-Finney *et al.*, 2000) and is also often upregulated in adenoma cells (van Es *et al.*, 2005). Wnt also plays an important role in cancer progression. This pathway is involved in a number of different processes in animal development and controls self-renewal in a variety of adult stem cells. Wnt signalling is positively regulated

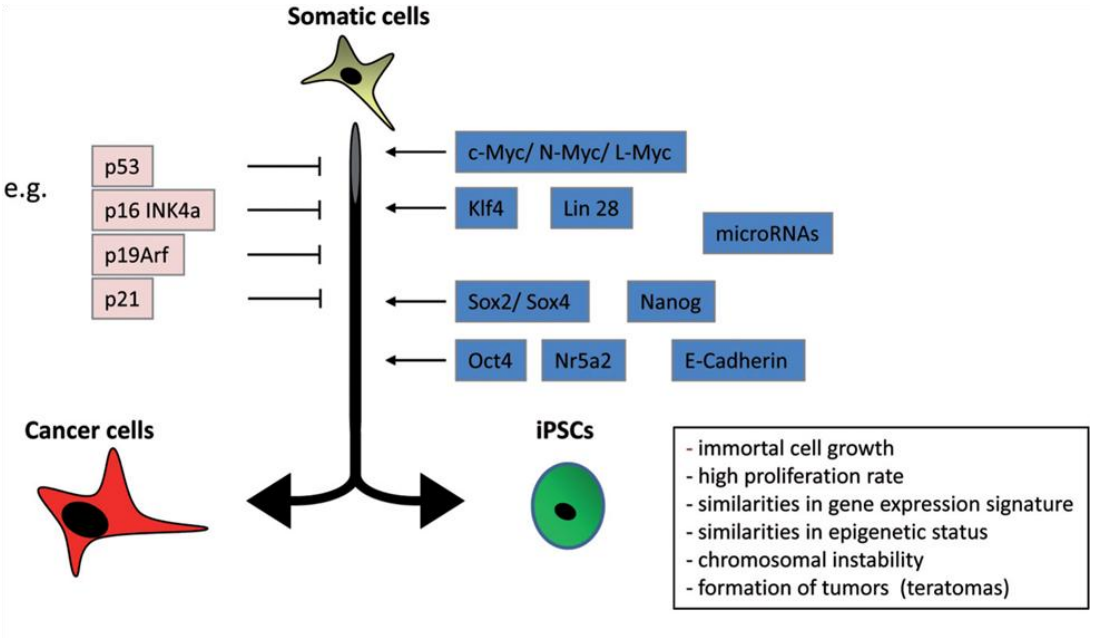


Figure 1.7: The links between pluripotency and cancer.

Many of the genes which promote reprogramming also enhance proliferation and are upregulated in many cancers. Furthermore, disruption of tumour suppressor pathways can also enhance the efficiency of reprogramming. Adapted from (Bernhardt *et al.*, 2012).

by Nanog and Oct4, and in turn, the downstream effector protein of the Wnt pathway, β-catenin, is capable of upregulating Nanog expression by interacting with Oct4 (Knoop & Baker, 2001; Loh *et al.*, 2006). One of the downstream affects of Wnt is the stabilisation and subsequent cytoplasmic accumulation of β-catenin. After accumulating in the cytoplasm, β-catenin can translocate to the nucleus, and regulated target gene expression. Therefore, aberrant Wnt signalling can cause deregulation of these target genes and as such is involved in the formation of a wide variety of cancers, including colorectal cancer and leukaemias (Reya & Clevers, 2005).

1.13 Tumour Heterogeneity and the Tumour Microenvironment

Cancerous tumours are complex heterogeneous populations of cells, each with their own role in cancer formation and progression. This complex structure of different cells, and the niche environment created by various cytokines and growth factors, is termed the Tumour Microenvironment (Figure 1.8). Cells within the Tumour Microenvironment include endothelial cells and pericytes which provide the tumour associated vascular structure, cancer-associated fibroblasts, which can serve a structural function in the tumour, and immune inflammatory cells such as macrophages, neutrophils, and T and B lymphocytes. These immune inflammatory cells provide cytokines which promote tumour growth, including EGF, VEGF, FGF2 and TGF-β (Hanahan & Weinberg, 2011). However, perhaps the most important cell type identified in the Tumour Microenvironment is the Tumour-Initiating Cancer cell, or Cancer Stem Cell (CSC). These cells are capable of differentiating into many of the cells found within tumours and are thought to have the ability to seed new tumours.

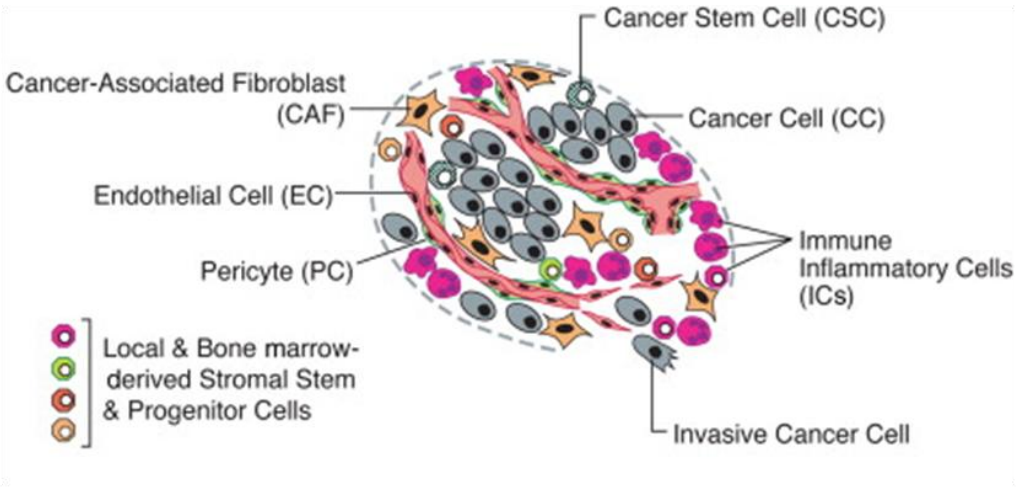


Figure 1.8: The Tumour Microenvironment.

Tumours are heterogeneous populations of cells, each with their own role in tumour formation. Cancer-associated fibroblasts serve a structural function, endothelial and pericytes cells provide the vascular system of the tumour, immune inflammatory cells produce cytokines which enhance tumour growth and CSCs are able to self-renew their populations and differentiate into many of the tumour stromal cells. Adapted from (Hanahan & Weinberg, 2011).

1.14 Cancer Stem Cells

CSCs are defined by their ability to self-renew and to generate new tumours upon transplantation. The first evidence for CSCs was shown in 1994 (Lapidot *et al.*, 1994). A sub-population of leukaemia cells possessing CD34+/CD38- antigenic properties were identified; a phenotype which is associated with haematopoietic and pluripotent stem cells. This led to the discovery in 1997 by Bonnet and Dick that this population of cells were capable of self-renewal, and generated differentiated progeny with aberrant phenotypes upon transplantation into mice (Bonnet & Dick, 1997). Since their initial discovery, CSCs have been identified in a wide variety of tumours, including breast (Al-Hajj *et al.*, 2003), brain (Singh *et al.*, 2004), prostate (Collins *et al.*, 2005) and many others (Table 1.3).

CSCs are identified by their ability to form tumours in mice xenograft models and by certain marker proteins which were originally identified in adult stem cells such as $CD34^{+}/CD38^{low/-}$ for lymphomas, $CD133^{+}$ for brain tumours, and $CD44^{+}/CD24^{low/-}$ for breast cancers. Other markers include drug

Cancers with Identified CSC populations	Reference
Breast Cancer	(Al-Hajj *et al.*, 2003)
CNS	(Singh *et al.*, 2004)
Multiple Myeloma	(Matsui *et al.*, 2008)
Melanoma	(Fang *et al.*, 2005)
Prostate	(Collins *et al.*, 2005)
HNSCC	(Prince *et al.*, 2007)
Colorectal cancer	(O'Brien *et al.*, 2007; Ricci-Vitiani *et al.*, 2007)
Pancreas	(Li *et al.*, 2007)
Lung	(Kim *et al.*, 2005)
Ovarian Cancer	(Zhang *et al.*, 2008)
Cervical Cancer	(Feng *et al.*, 2009)
Bladder Cancer	(Chan *et al.*, 2009)
Liver Cancer	(Yamashita *et al.*, 2009)
Renal Cancer	(Bussolati *et al.*, 2008)
Gastric Cancer	(Wakamatsu *et al.*, 2012; Xue *et al.*, 2012)
Osteosarcoma	(Tirino *et al.*, 2011)
Ewings sarcoma	(Suvà *et al.*, 2009)
Rhabdomyosarcoma	(Walter *et al.*, 2011)
Synovial Sarcoma	(Terry & Nielsen, 2010)

Table 1.3: Cancer Stem Cell Identification

Marker	Cancer	Marker	Cancer
CD133	Brain	**CD90**	Leukaemia
	Prostate		Glioma
	Pancreas		Liver
	Melanoma	**ABCB5**	Melanoma
	Colon	**ALDH**	Breast
	Liver		Lung
	Lung		Head and Neck
	Ovary		Colon
	Gastric		Liver
CD44	Colon		Pancreas
	Breast		Gastric
	Prostate		Prostate
	Pancreatic		Renal
	Gastric		
CD34	Leukaemia		

Table 1.4: Markers used to Identify Cancer Stem Cells.
Adapted from (Hu & Fu, 2012).

efflux pumps, such as the ABCG2 multidrug resistance pump and ALDH. A full list of the different marker types in various cancers is shown in Table 1.4 (Hu & Fu, 2012).

1.15 Roles of CSCs in Tumour Resistance to Anti-cancer Therapies

CSCs have been implicated in the acquisition of drug resistance in many tumours. Firstly, like normal stem cells, they have a quiescent or slowly proliferating nature. Since most anti-cancer therapies target DNA replication or protein synthesis, this makes CSCs more resistant to drug treatments than their more proliferative differentiated progeny. As such, CD133+ glioma cells have increased resistance to ionizing radiation in comparison to their CD133- counterparts (Bao *et al.*, 2006). Similar results have also been shown in breast cancer (Diehn *et al.*, 2009).

A second mechanism of drug resistance which CSCs possess is the ability to efflux anti-cancer drugs out of the cell. This is achieved through increased expression of ABC transporters, which are membrane proteins involved in the transport of various hydrophobic and hydrophilic compounds across cell membranes by ATP hydrolysis. ESCs and CSCs possess higher levels of ABC transporter protein expression than their differentiated progeny. This property confers increased drug resistance to CSCs, by efficiently effluxing anti-cancer drugs (Hirschmann-Jax *et al.*, 2004; Scharenberg *et al.*, 2002). The primary drug efflux pumps upregulated on CSCs are ABCB1, ABCC1 and ABCG2, which are all capable of effluxing a wide range of anti-cancer agents (Dean *et al.*, 2005).

Because of their role in resistance of tumours to anti-cancer therapies, there have been a number of efforts to develop drugs which specifically target CSCs, summarised in Table 1.5. These therapies work by either inhibiting the self-renewal signalling pathways which become deregulated in cancers, such as Wnt, Notch and Hedgehog, or by using monoclonal antibodies to target markers upregulated on CSCs (Hu & Fu, 2012).

Target	Drug	Cancer	Method of Action	Phase
Wnt	vitamin D3	Basal Cell Carcinoma	β-catenin	III
	PRI-724	Advanced solid tumours	CBP/β-catenin	I
	CWP232291	AML	β-catenin	I
Notch	MK0752	Advanced Breast Cancer	γ-secretase inhibitor	I
	RO4929097	Lung Cancer	γ-secretase inhibitor	II
	PF-03084014	Leukaemia	γ-secretase inhibitor	I
	OMP-21M18	Pancreatic Cancer	anti-DLL4 specific mAbs	I
Hedgehog	GDC-0449	Solid tumours, Colorectal	PTCH and/or SMO antagonist	I
			SMO antagonist	II
	BMS-833923	Basal cell	SMO antagonist	I
	IPI-926	Primary Myelofibrosis, Fibrosis, Bone Marrow	SMO antagonist	II
ABCB5	3C2-1D12	Malignant Melanoma	An ABCB5-specific mAb	target validation
CD44	H90	AML	A CD44-specific mAb	pre-clinical
	p245	Breast Cancer	A CD44-specific mAb	pre-clinical
EPCAM	Adecatumumab	Metastatic Breast Cancer	an EPCAM-specific mAb	II
	Edrecolomab	Colon Cancer	an EPCAM-specific mAb	II-III
	Catumaxomab	Ovarian and Gastric Cancer	Ab against EPCAM and CD3	II-III
	MT110	Lung and Gastrointestinal Cancer	Ab against EPCAM and CD3	I

Table 1.5: Cancer stem cell specific drugs and their targets.
Adapted from (Hu & Fu, 2012; Zhou *et al.*, 2009a).

1.16 The Role of the Epithelial – Mesenchymal Transition in Cancer

The Epithelial – Mesenchymal Transition (EMT) is an evolutionarily conserved process which occurs both in normal development and during tissue injury (Thiery, 2003). However, there is increasing evidence to suggest that this developmental process is involved in tumour formation and metastasis.

EMT is primarily initiated by transforming growth factor-β (TGF-β), which is also associated with cancer formation and progression. TGF-β receptors phosphorylate SMAD 2 and 3, which in turn form a heterodimer with SMAD 4 (Comijn *et al.*, 2001; Postigo *et al.*, 2003; Verschueren *et al.*, 1999). These SMAD complexes associate with transcription factors and upregulate expression of various EMT associated genes, including C2H2-type zinc finger proteins, Snail, Slug and Zeb; and the bHLH factor, Twist. Expression of these EMT-related proteins results in the repression of the cell-adhesion molecule, E-cadherin, by the recruitment of histone deacetylases. This increases cell motility (Peinado *et al.*, 2007), and as such, the EMT is implicated in cancer metastasis (Figure 1.9).

Interestingly, in addition to TGF-β, there are a number of pathways involved in EMT induction which are deregulated in cancer, including Wnt signalling, Notch signalling and NF-κB (Chua *et al.*, 2007; Kim *et al.*, 2002; Min *et al.*, 2008; Noseda *et al.*, 2004; Sahlgren *et al.*, 2008). Furthermore, during EMT the normally pro-apoptotic signalling protein, TGF-β, can become anti-apoptotic through TGFβ-mediated induction of the PI3K signalling pathway. This TGF-β initiation of PIK/AKT signalling has been shown to block apoptosis in mammary epithelial cells (Gal *et al.*, 2008).

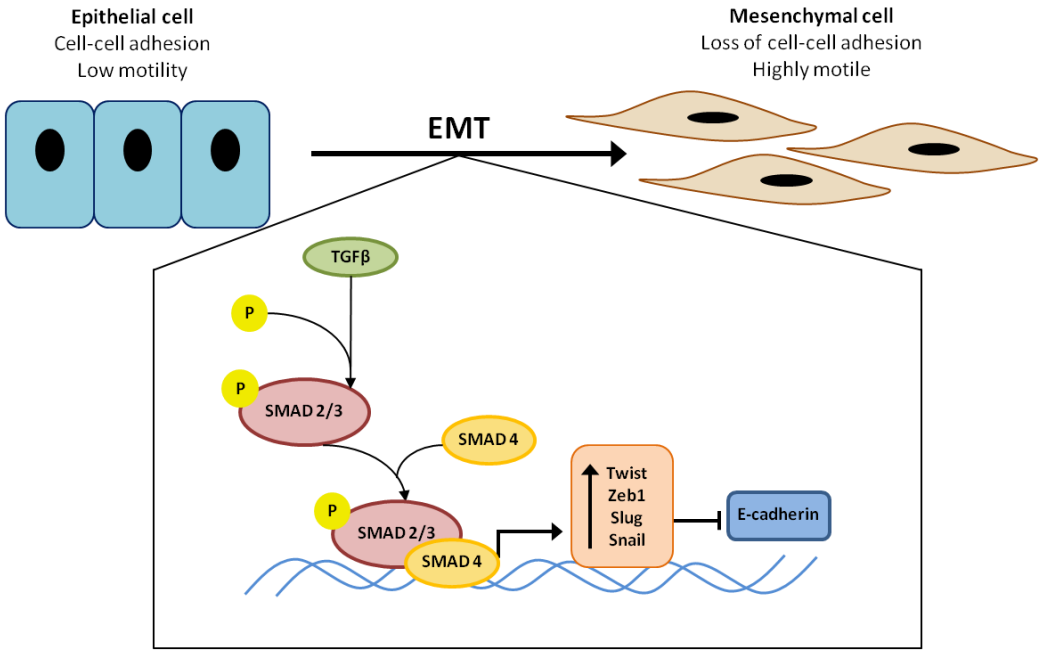


Figure 1.9: The Epithelial-to-Mesenchymal Transition

TGFβ signalling causes SMADs 2 and 3 to become phosphorylated. SMADs2 or 3 can then form a heterodimer with SMAD4 which then translocates to the nucleus and activates transcription of the EMT-associated genes, Twist, Zeb1, Slug and Snail. These proteins repress expression of the cell-adhesion molecule, E-cadherin, resulting in reduced cell-cell adhesion and increased cellular motility.

1.16.1 The EMT can generate Cancer Stem Cells

Recently, the EMT has been shown to be capable of generating populations of CSCs. This was initially highlighted by the observation that disseminated secondary breast cancer possesses a high percentage of $CD44^{high}/CD24^{low}$ cells (Al-Hajj *et al.*, 2003). From this observation, Mani and colleagues provided the first evidence for the role of EMT in the generation of CSCs. They induced EMT in both primary mammary carcinoma samples and non-tumorigenic immortalized human mammary epithelial cells, either by ectopic expression of Snail or Twist, or by treatment with TGF-β. These cells possessed a $CD44^{high}/CD24^{low}$ antigenic phenotype, an increased ability to form mammospheres, and were capable of self-renewal and forming differentiated progeny (Mani *et al.*, 2008). Further evidence supporting this hypothesis has been demonstrated by showing that several of the EMT-associated proteins promote cell de-differentiation. These include Zeb-1 which modulates KLF4 and Sox2 directly, Snail cooperating with SOX9 in

forming the mammary stem cell state and Twist acting downstream of BMI-1 to silence target genes involved in differentiation (Ansieau, 2012).

1.16.2 EMT-mediated Cancer Cell drug resistance

In addition to roles in metastasis and the enrichment of CSC populations in tumours, the EMT has also been implicated in cancer drug-resistance. Indeed, many of the epigenetic changes observed in cancer cells which have acquired drug resistance are associated with EMT, such as the increased expression of the EMT-associated protein Zeb-1, which protects cells from damaged induced apoptosis of cancers (Sayan *et al.*, 2009). Moreover, the resulting cells produced after EMT possess phenotypes and gene expression profiles which more closely resemble MSCs. MSCs are capable of enhancing anti-cancer drug resistance of a number of cancer cells in co-culture experiments (Houthuijzen *et al.*, 2012; Singh & Settleman, 2010).

Importantly, the ability of EMT to promote cell de-differentiation and generate new populations of CSCs from differentiated progeny complicates the design of new anti-cancer therapies. If therapies target the more proliferative differentiated cells within a tumour, the tumour-initiated CSCs are likely to remain, thereby seeding new tumours. However, if CSCs alone are targeted, the process of EMT could cause the dedifferentiation of other tumour stromal cells to form new CSCs, once again enabling the formation of new tumours.

One possible solution to this paradox could be to identify anti-cancer drugs which selectively target cells which have undergone EMT. Gupta and colleagues successfully identified four such compounds from an initial screen of 16,000. These compounds were capable of selectively targeting cells in which EMT had been induced by shRNA knockdown of E-cadherin with IC_{50} values around 7-10 fold lower than for control cells. Further investigation of one of these drugs, salinomycin, demonstrated it was capable of reducing CD44+/CD24- CSC populations in breast cancer cells by 20-fold and selected for cells with morphological features which more closely resembled epithelial cells. This was in contrast to the chemotherapeutic drug, paclitaxel, which selected for cells exhibiting a mesenchymal and migratory phenotype. Finally, salinomycin pre-treatment

of breast cancer cells prior to implantation into mice resulted in a >100 fold reduction in tumour seeding capacity relative to paclitaxel treated cells (Gupta *et al.*, 2009b).

1.17 Controversy surrounding the existence of CSCs

CSCs, like other cancer cells, contain a number of genetic and epigenetic abnormalities which give these cells a highly dynamic and unstable nature, thereby complicating their detection and isolation. Therefore, the presence of CSCs has been extremely difficult to prove. This has resulted in a great deal of controversy surrounding whether these cells actually exist. Some of the difficulties in identifying and isolating these cells are discussed herein.

One of the 'gold-standard' detection methods for CSCs is their ability to form new tumours upon transplantation into mice. However, these tumour xenograft models are less than ideal as there is significant variation in the physiology of the various different animal host models. These differences in physiology present several problems, including the potential lack of cross-species functionality of the growth factors and cytokines, which provide the CSC niche, thereby promoting tumourigenesis. Indeed, to prevent immune rejection of human tumour cells in mice models, immunocompromised mice are often used for tumour xenograft experiments. However, these mice are unable to produce many of the cytokines which promote tumour growth. The importance of a cytokine-rich environment in CSC tumour formation was first demonstrated by Kelly and colleagues in 2007. Transgenic mice were used which contained the N-ras or Myc/immunoglobulin chromosomal translocations used to model lymphomas. Experiments with these transgenic mice demonstrated that as many as 10% of cells from primary lymphomas were capable of initiating lymphomas in this histocompatible xenograft animal model (Kelly *et al.*, 2007). This contradicted the earlier belief that CSCs represent a very small population within cancers.

A further issue complicating the identification of CSCs is the current lack of stringent CSC markers which are consistently present in tumour-initiating cells, but absent in the remaining cells of the tumour stroma. For example, cells isolated from Acute Myeloid Leukaemia (AML) in patients possessing a mutation in the nucleophosmin gene were $CD34^-$, but were still capable of

generating malignant progeny (Taussig *et al.*, 2010). This may lead to false identification of CSCs due to the isolated populations of CSCs being impure.

Solid tumours present a further hurdle to the isolation of CSCs. During CSC isolation from solid tumours, the tumours must first be dissociated into single cells before fluorescence activated cellular sorting of cells expressing certain markers, and seeding into artificial niches in immunocompromised mice. The mechanical stresses of disseminating the cells may select for cells that are more capable of surviving the various insults of the isolation process rather than CSCs.

The EMT also adds further complexity to the CSC hypothesis, as it is capable of generating CSC populations from differentiated progeny (Mani *et al.*, 2008). This *de novo* generation of CSCs adds to the already highly dynamic nature of these cells, further complicating any attempts to isolate these cells.

1.18 Tracking CSC differentiation provides strong evidence for their existence

The controversy surrounding the existence of CSCs may have finally been put to rest after the recent publication of three reports which, for the first time, provide strong evidence for the presence of CSCs. The first of these reports, by Parada et al., utilised a nestin-responsive reporter gene expressing GFP to specifically label neural stem cells. They demonstrated the presence of GFP positive 'stem-like cells' within glioblastoma, which when ablated, significantly arrested tumour growth (Chen *et al.*, 2012). Secondly, Blanplain and colleagues non-specifically labelled cells within squamous skin tumours, and identified two distinct proliferative cell compartments within the tumour; one which had limited potential to divide and produced terminally differentiated cells, and another which was capable of unlimited cell divisions and possessed stem-like characteristics. In addition, more aggressive tumours were shown to possess a higher proportion of these stem-like cells (Driessens *et al.*, 2012). Finally, and perhaps in the most elegant of these studies, Clevers et al. used a dual labelling strategy to track a population of gut cancer stem cells throughout

their differentiation. Firstly, they utilised a GFP reporter gene under the control of the Lgr5 promoter. Lgr5 is expressed in pluripotent cells from both healthy crypt stem cells, and intestinal adenocarcinoma CSCs (Barker *et al.*, 2009). This therefore enabled the identification of undifferentiated CSCs within adenocarcinoma samples, but GFP expression was lost upon differentiation of these cells. Lgr5 labelling was then combined with stable transfection of these cells with a fluorescent protein expression cassette. Expression from this cassette was induced by addition of the drug tamoxifen, and constitutive expression of the fluorescent marker protein was maintained throughout differentiation. Furthermore, a second treatment with tamoxifen induced a colour switch in the fluorescent marker expression cassette. This allowed the investigators to trace the differentiation and expansion of individual CSCs. The results of this dual labelling system,

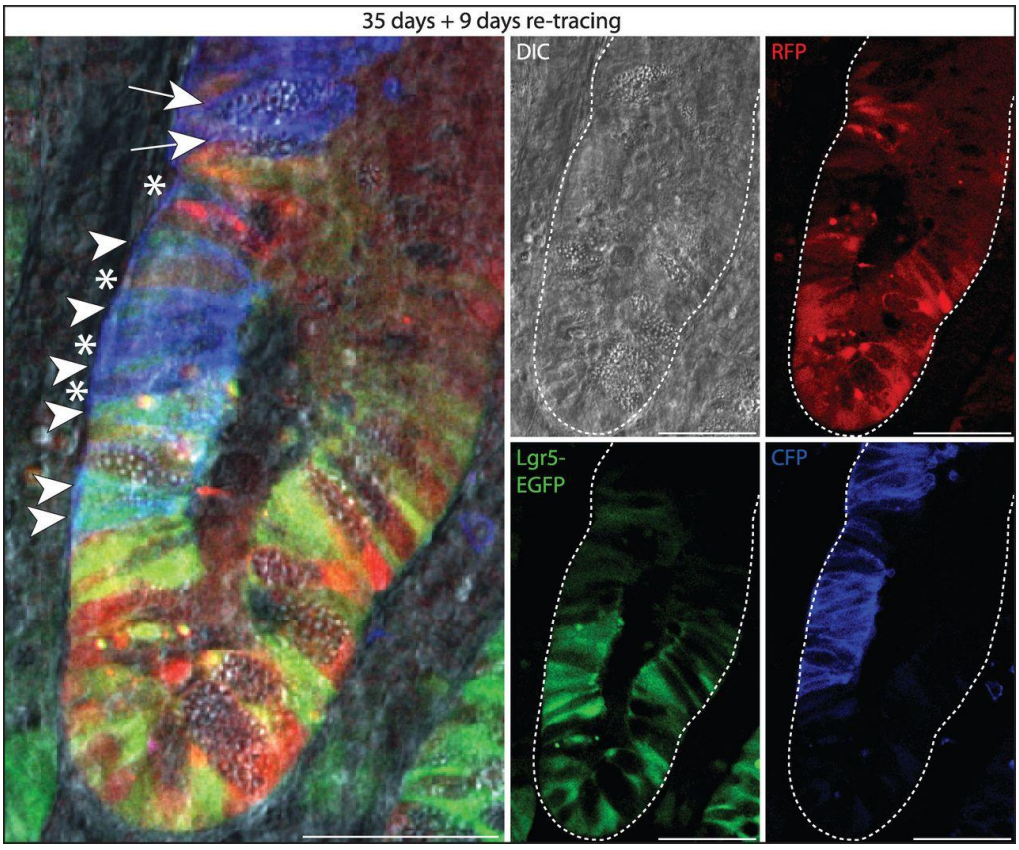


Figure 1.10: Tracing CSCs within an adenocarcinoma segment.
Differential Interference Contrast (DIC) is shown in grey, Lgr+ CSCs express EGFP. The cells which were originally labelled and produced from Lgr+ cells are displayed in red (RFP). A second addition of tamoxifen induced a colour change to blue (CFP) in one CSCs. This CSC could be traced as it expanded and produced differentiated progeny which subsequently lost GFP expression. Scale bars indicate 50 µm. Taken from (Schepers *et al.*, 2012).

demonstrated in Figure 1.10, showed that the Lgr5+ CSCs were not only capable of self-renewing, but they could form Lgr5- differentiated progeny in a hierarchical manner closely resembling normal intestinal development (Schepers *et al.*, 2012). Together, these reports demonstrate that cancers possess a subpopulation of cells which are undifferentiated and are capable of self-renewing their populations, in addition to producing differentiated progeny which form the bulk of the tumour.

1.19 Reprogramming Cancer Cells

Tumour-initiating CSCs have been identified in a number of cancers. These cells are capable of differentiating into various cells within tumours, and are believed to be the main cause of metastasis and secondary cancers. However, CSCs are difficult to isolate from primary tumour samples due to their rarity and highly dynamic nature. Therefore, utilising iPSC technology to reprogramme somatic cancer cells to these tumour-initiating CSCs would provide an abundant source of these cells for studying their role in cancer progression. To date, there have been a number of attempts to generate induced Pluripotent Cancer Stem Cells (iPCs) which are detailed in Table 1.6.

	Reprogramming Factors	Vector Delivery	Reference
Colorectal cancer, pancreatic cancer, hepatocellular carcinoma (cell lines)	OSKM	Retroviral /lentiviral	(Miyoshi *et al.*, 2010)
Chronic Myeloid Leukaemia (cell line)	OSKM	retroviral	(Carette *et al.*, 2010)
Melanoma, pancreatic cancer (cell lines)	miRNA-302	retroviral	(Lin *et al.*, 2008)
Melanoma (cell line)	OKM	lentiviral	(Utikal *et al.*, 2009)
Polycythemia vera, primary myelofibrosis (primary samples)	OSKM	retroviral	(Hu *et al.*, 2011)
Chronic Myeloid Leukaemia (primary samples)	OSKM	retroviral	(Kumano *et al.*, 2012)
EBV-immortalised B-cells	OSKM	episomal EBNA1/OriP based	(Choi *et al.*, 2011)

Table 1.6: Generation of Induced Pluripotent Cancer Stem Cells from somatic cancer cells. Adapted from (Sun & Liu, 2011). O – Oct4, S – Sox2, K – Klf4, M – Myc.

Perhaps the most interesting of these studies was performed by Hu and colleagues, whereby primary human lymphoblasts from a patient with Chronic Myeloid Leukaemia containing the JAK2-V617F mutation (+) polycythemia vera (PV) mutation were successfully reprogrammed to iPC cells (Hu *et al.*, 2011). This study demonstrated it was possible to reprogramme leukemic cells containing a complex genetic mutation which, as yet, cannot be recreated in experimental models by genetic engineering. Furthermore, a non-integrating, episomally maintained reprogramming vector based on the Epstein-Barr virus nuclear protein, EBNA1, was used for this study. The episomal nature of the reprogramming vector would prevent any perturbations to the gene expression profiles of these cells which could occur with integrating retroviral gene delivery vectors. Therefore, iPCs generated by this method would represent a more accurate model of leukemic stem cells. Although, as previously mentioned, EBV EBNA-1 is capable of immortalising B-cells *in vitro* (Young *et al.*, 1988).

1.20 Ewings Sarcoma Family Tumours

First identified in 1921 by Professor James Ewing (Ewing, 1921), Ewings sarcoma is the second most common bone malignancy after osteosarcoma. It predominantly affects children and young adults, with a peak incidence at the age of 15, and most tumours found in the long bones of the arms and legs or in the ribs (Riggi & Stamenkovic, 2007). However, Ewings sarcoma is only part of a larger family, termed the Ewing's Sarcoma Family of Tumours (ESFTs). These incorporate a family of morphologically similar cancers, including the 'classic' Ewing's sarcoma of bone, extraskeletal Ewing's sarcoma (Angervall & Enzinger, 1975), small cell tumours of the thoracopulmonary region (Askin *et al.*, 1979), and soft tissue-based primitive neuroectodermal tumours (PNET) (Jaffe *et al.*, 1984).

ESFTs are treated with a combination of surgery, radiation, and chemotherapy. However the survival rates remain poor, with 50% survival after 5 years, and less than 30% after 10 years. Survival rates can even be as little as 25% if metastases are present at diagnosis (Riggi & Stamenkovic, 2007).

1.21 ESFTs are caused by a chromosomal translocation

ESFTs are caused by a chromosomal translocation which fuses the N-terminal transactivation domain of the Ewing's Sarcoma (EWS) gene with the C-terminal DNA-binding domains of various ETS family transcription factors. The most commonly occurring chromosomal translocation in ESFTs is t(11;22)(q24:q12) (Figure 1.11a), present in 85% of ESFTs. This translocation causes a fusion of the EWS gene with the Friend Leukaemia Integration 1 transcription factor (FLI1) gene (Delattre *et al.*, 1992). Other, less common translocations produce different oncogenic fusion proteins, including the EWS-ERG fusion protein which occurs in 10% of ESFTs, and various other EWS-ETS protein fusions account for the remaining cases.

1.21.1 The EWS gene

EWS is a member of the TET family of proteins, which also include the Translocated in Liposarcoma (TLS), and TATA-binding protein-associated factor 15 (TAF15). These proteins are believed to have a role in RNA transcription and processing. EWS contains an N-terminal transcriptional activation domain and a C-terminal RNA-binding domain. The N-terminal domain is capable of interacting with RNA pol II and TFIID (Bertolotti *et al.*, 1998) and thus is implicated as a transcriptional co-activator (Law *et al.*, 2006; Lee *et al.*, 2005; Thomas *et al.*, 2004). The C-terminal domain of EWS contains an 86-amino acid RNA Recognition Motif (RRM) implicated in RNA binding (Ohno *et al.*, 1994), and has three RGG boxes, which repress the activity of the N-terminus (Alex & Lee, 2005).

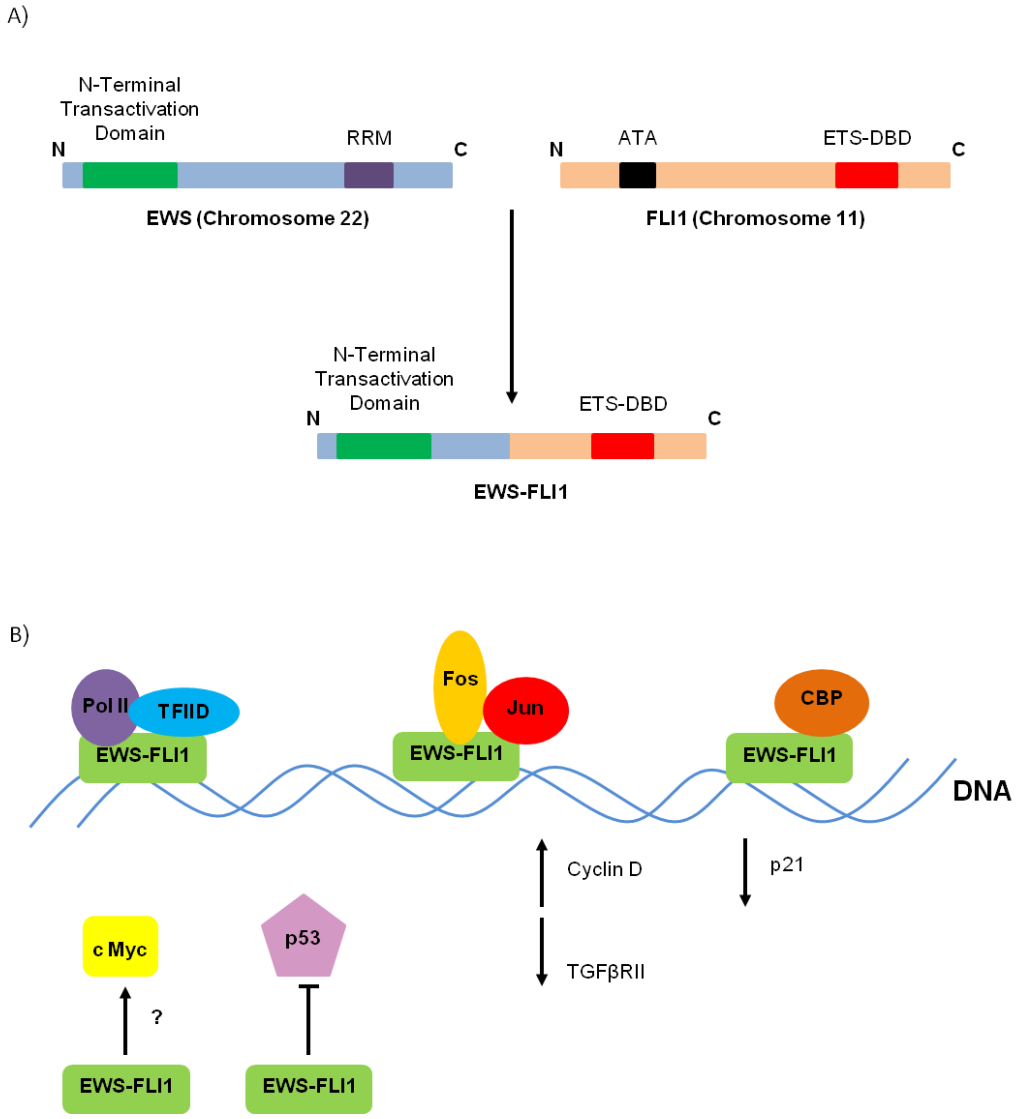


Figure 1.11 The EWS-FLI1 fusion protein

A) A chromosomal translocation occurs between chromosome 22 and chromosome 11. This fuses the N-terminal transactivation domain of the EWS gene to the DNA-binding domain (DBD) located at the C-terminus of the FLI1 gene. B) The EWS-FLI1 gene product is a potent oncoprotein, capable of interacting with various proteins involved in transcription, thereby regulating the expression of a number genes involved in cell cycle regulation and proliferation.

1.21.2 The FLI1 gene

The FLI1 gene was first identified as a proto-oncogene which is a preferred site of entry for the Friend Murine Leukaemia Virus (Ben-David *et al.*, 1991). It is a member of the ETS family of transcription factor proteins, classified as such due to a characteristic Helix-Loop-Helix N-terminal activation domain (ATA) (Rao *et al.*, 1993). The FLI1 C-terminus possesses an ETS-DNA Binding Domain, containing a Helix-Turn-Helix motif capable of binding to the consensus sequence, GAAA/T. Additional adjacent sequences determine the specificity of FLI1 recognition (Liang *et al.*, 1994; Mao *et al.*, 1994).

1.21.3 The EWS-FLI1 fusion protein

The EWS-FLI1 fusion protein is therefore a potent oncoprotein capable of recognising the same specific DNA motifs as the wild-type FLI1 protein (Mao *et al.*, 1994), but possessing enhanced transcription activation capabilities, due to the EWS N-terminal domain and the loss of the regulatory EWS C-terminus.

The disordered nature of the EWS-FLI1 protein also enables it to interact with multiple proteins to initiate oncogenesis. As such, it is capable of interacting with proteins involved in transcriptional regulation, such as the Fos-Jun dimer, RNA pol II, TFIID, and CBP (Figure 1.11b). An excellent review of these protein interactions is provided in (Erkizan *et al.*, 2010).

EWS-FLI1 can also affect pre-mRNA processing. Specifically, it can alter splice site selection by counteracting heterogeneous nuclear ribonuclear protein A1 (hnRNPA1). It interacts with Splicing Factor 2 (SF2) and the U1 small nuclear ribonuclear protein U1C to redirect the splicing machinery to more proximal 5′ splicing sites (Knoop & Baker, 2000; Knoop & Baker, 2001). This alternative splicing can lead to oncogenesis by altering the relative expressions of Cyclin D splice variants. Cyclin D has 2 possible splice variants; Cyclin Da and Cyclin Db. Of these, Cyclin Db is more oncogenic and usually expressed at lower levels. However, EWS-FLI1-mediated alternative splicing causes upregulation of Cyclin Db, thereby

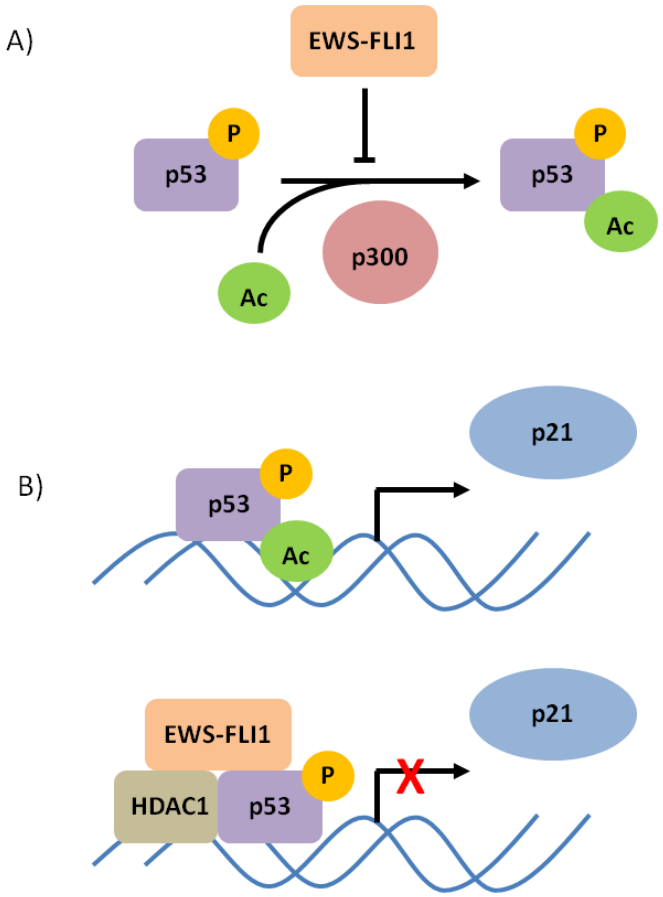


Figure 1.12: EWS-FLI1 represses p53 activation of p21

A) EWS-FLI1 binds to p53 and represses p300-mediated acetylation. This inhibits the binding of p53 to its target sequences. B) Phosphorylated and acetylated p53 is capable of promoting p21 transcription. However, in the presence of EWS-FLI1, p21 transcription is repressed. This is caused by EWS-FLI1 repressing p53 acetylation, and by recruiting HDAC1 into a complex with p53, thereby further repressing p21 transcription.

enhancing the G1-S phase progression of the cell cycle (Sanchez *et al.*, 2008).

In addition, EWS-FLI1 can deregulate the activities of the tumour suppressor protein, p53 (Figure 1.12). EWS-FLI1 is capable of forming a complex with p53 via its N-terminus, at residues 65-109. This interaction represses p53 activity in two ways. Firstly, it represses p300-mediated acetylation of p53, reducing the ability of p53 to bind to its target promoter sequences. Secondly, the EWS-FLI1 protein recruits HDAC1 to the EWS-FLI1/p53 complex, which further represses the transcription of p53 target genes, such as p21. p21 prevents cell cycle progression from G1-S phase by binding to cyclin dependent kinase 4. Therefore, by repressing expression of p21,

EWS-FLI1 drives cell cycle progression, enhancing proliferation (Li *et al.*, 2012; Li *et al.*, 2010).

1.22 ESFTs Originate from Mesenchymal Cells

Although ESFT cells express various genes found in neural differentiation and ectodermal lineages (Hu-Lieskovan *et al.*, 2005), there is evidence to suggest they may originate from Mesenchymal Stem Cells (MSCs). This is exemplified by results showing that the ectopic expression of the EWS-FLI1 protein in fibroblast cells results in growth arrest and cell death; whereas in contrast, ectopic expression of the EWS-FLI1 protein in both murine and human MSCs produced cells with ESFT-like characteristics (Lessnick *et al.*, 2002). These include the expression of the ESFT-associated marker CD99, morphological changes from elongated spindle shaped cells to small, rounded cells with scant cytoplasm, and the upregulation of genes implicated in neural differentiation (Riggi *et al.*, 2005; Riggi *et al.*, 2008). Further evidence for MSCs being the cells of origin for ESFTs is the ability of ESFT CSCs to differentiate down mesenchymal lineages, suggesting these cells have maintained some of their parental MSC characteristics (Suvà *et al.*, 2009).

1.23 ESFT CSCs

Recently a small subpopulation of ESFT cancer stem cells has been identified which are CD133+. These CSCs account for between 4% - 8% of ESFT cell populations in primary tumours and are capable of forming multicellular spheroids *in vitro*. Moreover, only 2,300 of these CSCs were required to form tumours in mice. Interestingly, these cells were also capable of differentiating down three mesenchymal lineages, namely adipogenic, chondrogenic and osteogenic. This further indicates these cells may have been derived from mesenchymal stem cell precursors (Suvà *et al.*, 2009).

1.24 Herpesviruses

Herpesviruses are large DNA viruses that are prevalent throughout the animal kingdom, spanning both vertebrates and invertebrates. These viruses belong to the order *herpesvirales*, which contains three families; *Alloherpesviridae*, which infect fish and amphibians, *Malacoherpesviridae*, which infect molluscs, and *Herpesviridae* which infects mammals, birds and reptiles. The *Herpesviridae* encompasses 90% of herpesviruses (McGeoch *et al.*, 2006) and contains 3 sub-families; α, β, and γ, which are grouped on their genome sequence and organization.

Alphaherpesvirinae primarily form latent infections in the host sensory ganglia. This subfamily includes the two mammalian genera (*Simplexvirus* and *Varicellovirus*), two avian genera (*Mardivirus* and *Iltovirus*), and several reptilian viruses which remain unclassified. The human herpesviruses, Herpes simplex virus type-1 and 2, (HSV-1 and -2), which form cold sores and genital herpes respectively, and Varicella Zoster virus (VSV), which causes chicken pox and shingles, belong to this sub-family.

The *Betaherpesvirinae* typically has a restricted host range and a long reproductive cycle. This family consists of four genera; *Cytomegalovirus*, *Muromegalovirus*, *Roseolovirus* and *Provoscivirus*. The most notable of human viruses in this family is human Cytomegalovirus (HCMV), which typically causes mild diseases or asymptomatic infections in up to 90% of the population worldwide. However, HCMV can cause serious pneumonia-like symptoms in immune-suppressed patients.

Gammaherpesvirinae contains four genera; *Lymphocryptovirus* (gamma-1), *Rhadinovirus* (gamma-2), *Macavirus* and *Percavirus*. Viruses within this family typically replicate in T- and B-lymphocytes, with gamma-1 herpesviruses typically infecting B-cells, and gamma-2 herpesviruses persistently infecting T-cells.

Herpesvirinae possess ds DNA genomes of 124-241kb in size (Davison *et al.*, 2003; Gray *et al.*, 2001), which are packaged within an icosohedral (T=16) capsid composed of 162 capsomeres (McGeoch *et al.*, 2006). 150 of these capsomeres are homo-hexamers, 11 are homo-pentameric multimers,

and the final pentameric position is occupied by a portal complex which allows DNA entry into the completed capsid. Surrounding the capsid is the tegument layer which contains a number of pre-synthesised viral and cellular proteins required for infection and initiation of replication (Varnum *et al.*, 2004; Zhu *et al.*, 2005). Finally, the entire virus particle is enveloped by a lipid bi-layer membrane which contains virally encoded glycoproteins required for viral attachment and entry into infected cells. The size of these entire herpesvirus particles varies between 120-260 nm, depending on the size of the tegument layer (Dimmock *et al.*, 2001). The typical structure of a herpesvirus virion is shown in Figure 1.13.

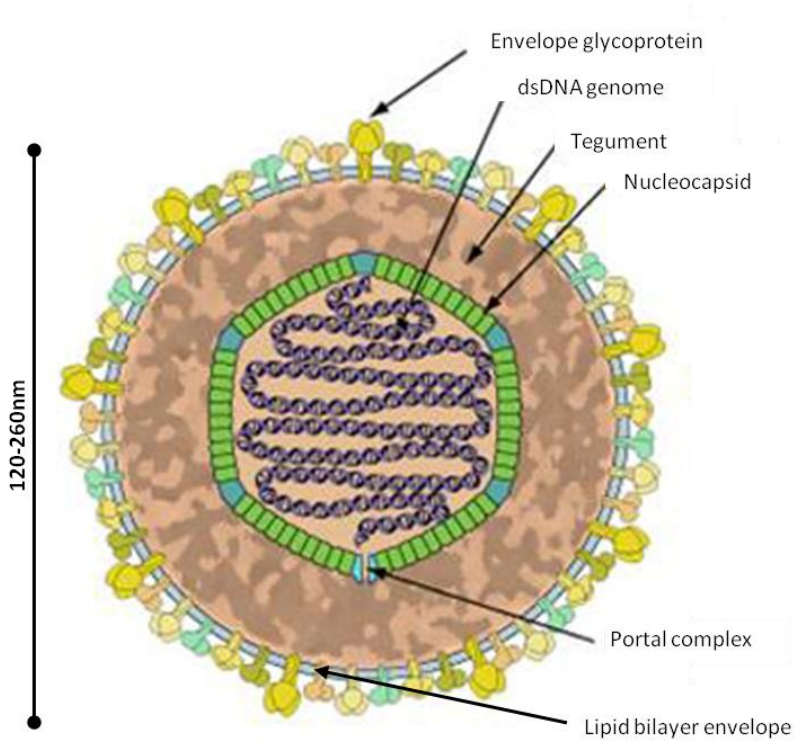


Figure 1.13: The Herpesvirus capsid structure.

Herpesvisues possess a dsDNA genome which is encased within a T=16 icosohedral nucleocapsid. This capsid contains a portal complex which enables the dsDNA genome to be inserted into the completed capsid after its assembly. Surrounding the nucleocapsid is the protein-filled Tegument layer, which is in turn surrounded by a lipid-bilayer envelope. Adapted from abcam.com, date accessed 04/09/2012.

1.25 Gammaherpesviruses

Like all herpesviruses, the *Gammaherpesvirinae* are able to form latently persistent infections in their host. Moreover, unlike the alpha- and betaherpesviruses, gammaherpesviruses can replicate their genome in dividing cell populations due to the presence of a virally encoded latent origin of replication (OriP). These viruses are lymphotrophic and predominantly infect T- and B-lymphocytes. However, they have been shown to infect a number of other cell types, to form latent infections or undergo lytic replication.

Gammaherpesviruses show remarkable similarity within their genomes, with more shared genes compared to members of the alpha- or betaherpesviruses. Furthermore, their genomes are highly co-linear, containing conserved blocks of highly similar genes with only a few genes unique to each virus. These coding regions are, in turn, flanked by terminal repeat regions (Pellet & Roizman, 2007). Figure 1.14 demonstrates the co-linearity of three gammaherpesviruses, Herpesvirus saimiri (HVS), Epstein - Barr virus (EBV), and Kaposi's Sarcoma-associated Herpesvirus (KSHV).

A further property of the gammaherpesviruses is their ability to induce neoplastic disease, either in their natural host, in the case of EBV or KSHV, or in a foreign host, as is observed for HVS. All three of these viruses are associated with lymphoproliferative disorders. However, EBV has also been linked with the formation of nasopharyngeal carcinomas, and KSHV linked with endotheliosarcomas, such as Kaposi's sarcoma (Damania, 2004). This oncogenic potential is due to virally encoded genes involved in lytic and latent stages of the virus life cycle. These genes subvert the host immune response, regulate the cell cycle and inhibit apoptosis. The two gammaherpesviruses which will be discussed in further detail here are EBV and HVS.

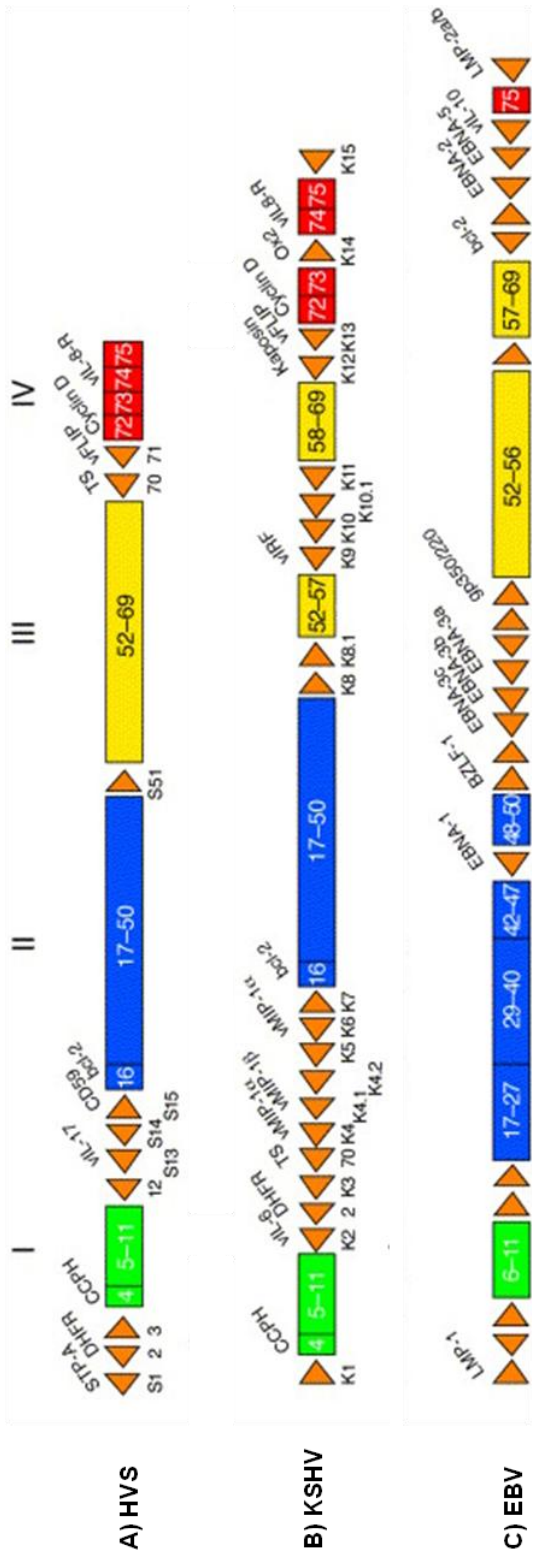


Figure 1.14: The genome structures of HVS and related gamma-herpesviruses.

A) Herpesvirus saimiri (HVS) genome arrangement, B) Kaposi's Sarcoma-associated Herpesvirus (KSHV) genome structure, C) The genome structure of Epstein-Barr Virus (EBV). There is significant co-linearity between the gammaherpesvirus genomes, with four conserved gene blocks (shown here in different colours). These gene blocks contain proteins with conserved functions between gammaherpesviruses, with unique proteins found between these conserved blocks. Adapted from (Pedro Simas & Efstathiou, 1998).

1.26 Epstein - Barr virus

EBV is the prototype member of the type-1 gammaherpesviruses, and can be found in over 95% of the world's population. It forms asymptomatic infections in children, but can cause infectious mononucleosis if acquired during adolescence (Henle *et al.*, 1968). Following initial infection, EBV develops life-long asymptomatic latent infections in the host. Typically, EBV infections are non-carcinogenic; however, it has been associated with a number of malignancies, including Hodgkin's lymphoma, nasopharyngeal carcinoma and T-cell lymphomas (Jones *et al.*, 1988; Levine *et al.*, 1971; Pathmanathan *et al.*, 1995).

EBV primarily infects B-lymphocytes via an interaction with the viral glycoprotein, gp350/220, and the cellular complement receptor, CD21 (Fingeroth *et al.*, 1984; Kawaguchi *et al.*, 2009). Upon initial infection, a short lytic reproductive cycle occurs, activating the naïve infected B-cells, causing them to proliferate. These activated B-cells then further differentiate into long-lived memory B-cells in which the EBV can reside in a latent state (Babcock *et al.*, 1998).

Only a limited number of genes are transcribed in the latent phase of the EBV lifecycle. These comprise of six EBV nuclear antigens (EBNAs 1-6), three latent membrane proteins, (LMP-1, LMP-2A and LMP-2B), and two small RNAs (EBERs 1 and 2). There are three distinct forms of EBV latency described, each possessing different expression patterns of the latency-associated genes. During Type I latency, only the EBERs and the episomal maintenance protein, EBNA-1, are expressed. Type I latency is frequently associated with Burkitt's Lymphoma, but the expression of these genes alone is not sufficient for transformation (Klein *et al.*, 2007). Type II latency involves the expression of the LMPs in addition to the genes transcribed during Type I latency. Once again, this expression profile is not capable of inducing cell proliferation, but is, however, associated with Hodgkin's lymphoma and nasopharyngeal carcinoma (Chiang *et al.*, 1996; Deacon *et al.*, 1993). Finally, Type III latency involves the expression of the complete set of latency-associated genes, and this expression profile is capable of inducing autonomous proliferation.

Although EBV usually remains latent, it can reactivate, initiating lytic gene expression in a temporal cascade. Like other herpesviruses, the genes involved in lytic replication are expressed in a temporal cascade, with the Immediate Early (IE) genes being primarily expressed. This in turn enables the expression of the Delayed Early (DE) genes, and the structural late (L) genes. EBV has two IE genes which control the lytic replication cascade, termed BRLF1 and BZLF1, which produce the transcription factors Rta and Zta, respectively (Countryman *et al.*, 1987; Countryman & Miller, 1985; Hardwick *et al.*, 1988). These proteins activate transcription of the other genes involved in the lytic cascade by binding directly to the viral promoters (Liu & Speck, 2003; Ragoczy & Miller, 1999). Consequently, lytic replication leads to the production of infectious virions.

1.27 Herpesvirus saimiri (HVS)

Herpesvirus saimiri (HVS) is a member of the *Rhadinovirus* genera of *Gammaherpesvirinae*, first identified in 1968. HVS establishes a persistent asymptomatic infection in T-lymphocytes of their natural host, the squirrel monkey (Saimiri sciureus) (Falk *et al.*, 1972; Melendez *et al.*, 1968). However, HVS is capable of forming acute malignant T-cell lymphomas upon experimental infection of New World primates, including owl monkeys, common marmosets, and cottontop tamarins (Fleckenstein & Desrosiers, 1982). Moreover, HVS can transform a variety of simian and human cells to continuous growth in vitro (Desrosiers *et al.*, 1985a; Schirm *et al.*, 1984). HVS also shares considerable sequence homology with the human herpesvirus, Kaposi's sarcoma-associated Herpesvirus (KSHV), and is used as a model to study the KSHV lytic lifecycle.

1.28 Genome Structure

The HVS ds DNA genome consists of a central region, 112,930 bp in length, which has a low G+C content (36%) termed L-DNA, flanked by a variable number of G+C rich (71%) H-DNA repeats, 1,444bp in length (Albrecht *et al.*, 1992; Murthy *et al.*, 1986). The L-DNA region contains the unique region of the genome, and encodes all 75 open reading frames (ORFs). In addition, the L-DNA also contains several virally encoded small nuclear viral RNAs; 5

in HVS strain C288, and 7 in HVS strain A11 (Albrecht *et al.*, 1992; Biesinger *et al.*, 1990). Among the 75 ORFs in HVS, there are 15 cellular homologues, including ORF4 and ORF15, which encode glycoproteins related to complement control proteins (Albrecht *et al.*, 1992), and ORF74, a G-protein coupled receptor homologue (Nicholas *et al.*, 1992). It is believed these cellular homologues enhance immune evasion, transformation and pathogenesis.

1.29 Viral Entry

HVS entry has not been extensively studied. However, HVS possesses many of the envelope glycoproteins conserved in other herpesviruses; namely gB, gD, gH, gL, gM and gN (Ensser *et al.*, 2003; Neipel *et al.*, 1997). Therefore the entry mechanisms of HVS can be elucidated by studying the mechanisms of other herpesviruses. For example, the viral entry mechanisms of Herpes Simplex Virus Type-1 (HVS-1) have been extensively studied, and will therefore be discussed here.

HSV-1 primary attachment occurs by binding of the viral gB and gC glycoproteins to heparan sulphate-containing proteoglycans on the cell surface (Herold *et al.*, 1994; Laquerre *et al.*, 1998; Shieh *et al.*, 1992). Following this binding, gD interacts with either Herpes Virus Mediator of Entry (HVEM) (Whitbeck *et al.*, 1997), nectin-1 (Richart *et al.*, 2003), or 3-O-Sulphate (Shukla *et al.*, 1999), which causes a conformational change in the C-terminal domain of gD. This conformational change has been implicated in transmitting a signal to the gB and gH/gL glycoproteins which subsequently mediate viral entry (Reske *et al.*, 2007). Because the glycoproteins, gB, gH and gL, are highly conserved, a similar mechanism for viral entry is believed to occur upon HVS infection. However, the exact mechanism involving these proteins is, as yet, unknown (Reske *et al.*, 2007).

Although gB, gM and gL are conserved, there is some variation between species. The HVS glycoprotein, ORF51, has also been implicated in viral entry into host cells, as it is present in the HVS virion and is capable of binding to heparan sulphate. Furthermore, ORF51 also shares significant homology to the KSHV protein, K8.1, which has also been shown to bind to heparan sulphate glycosaminoglycans and is believed to be involved in initial

binding in a similar way to gC of HSV-1 (Birkmann *et al.*, 2001; Means, 2004; Wang *et al.*, 2001).

1.30 HVS lifecycle

Upon viral entry, the linear HVS genome is transported to the nucleus. This process is believed to be similar to that of KSHV, whereby genome transport is mediated through dynein motors processing along microtubules (Naranatt *et al.*, 2005). Like other herpesviruses HVS is then capable of undergoing two distinct phases of infection; lytic replication, or a latent state. During latent infections many of the viral genes are repressed. In contrast, during lytic infections the virus is capable of replicating its genome and forming new infectious virions.

1.30.1 Lytic replication

Lytic infection allows HVS to reproduce its genome and eventually produce viral progeny within the host cell. Like all Herpesviruses, HVS genes are categorised into three distinct groups based on their temporal expression. These categories are Immediate Early (IE), Delayed Early (DE), and Late (L). IE genes are expressed first, and some are even present in the HVS virions. These proteins are responsible for initiation of DE gene expression. DE proteins are involved in DNA replication and L gene products encode the structural proteins involved in capsid formation and the envelope glycoproteins.

The exact mechanisms of HVS DNA replication and assembly are yet to be fully elucidated. However, in HSV-1 it is believed that once the viral genome has entered the cell, it is ligated into a circular genome, initiating rolling circle replication (Muylaert & Elias, 2007). Once DNA concatemers have been produced and Late structural proteins have formed capsids in the nucleus, DNA is incorporated into the capsids via a capsid pore complex. The exact packaging mechanism of HVS is again unknown, but the putative packaging protein encoded by ORF29 is believed to be involved in this process (Ensser *et al.*, 2003).

1.30.2 HVS Lytic replication is mediated by the IE genes, ORF50 and ORF57

Two IE genes regulate expression during HVS lytic replication, termed ORF50, (also known as the replication and transcription activator), and ORF57. ORF50 is the first protein to be expressed in the lytic gene temporal cascade. This transcriptional activator gene, encodes two transcripts, ORF50a and ORF50b. ORF50a is a spliced transcript containing a single intron, 960bp in length, and is detected early in viral replication. ORF50a is the major transcriptional activator, whereas ORF50b is expressed later in replication from a promoter that lies within the second exon (Nicholas *et al.*, 1991; Whitehouse *et al.*, 1997). ORF50a upregulates expression of many HVS genes, including ORF57, via ORF50 response elements which have the consensus sequence, CCN9GG (Walters *et al.*, 2004). ORF50 is capable of enhancing transcription by interacting with cellular proteins, such as the TATA-binding protein and components of the TFIID complex (Hall *et al.*, 1999; Smale & Kadonaga, 2003).

The second IE gene crucial for HVS replication is the multifunction protein, ORF57. ORF57 enhances the translation of viral intron-less mRNAs by transporting them out of the nucleus (Goodwin *et al.*, 1999). During normal cellular mRNA processing, unspliced pre-mRNA transcripts acquire distinct protein complexes required for their nuclear export during splicing. However, as most of the transcripts produced by HVS are intron-less, they require recruitment of these protein complexes to enable efficient transport of these mRNAs out of the nucleus and their subsequent translation. ORF57 has been shown to bind these intronless viral RNAs and recruit cellular export proteins to enhance translocation through the nuclear pore, thereby allowing subsequent translation (Jackson *et al.*, 2012). Specifically, the C-terminus binds to the consensus binding motif GAAGRG, which is found in the majority of HVS viral mRNAs (Colgan *et al.*, 2009a). ORF57 then shuttles these viral mRNAs out of the nucleus by forming an export-competent ribonucleoprotein particle by binding to the human TREX component, Aly, at residues 103-120 of ORF57 (Tunnicliffe *et al.*, 2011). This interaction then

mediates HVS mRNA nuclear export by the TAP-mediated pathway (Colgan *et al.*, 2009b).

In addition to exporting viral mRNAs, ORF57 is also capable of downregulating expression of genes containing at least 1 intron. This is achieved through interactions with the host cell splicing machinery, including SC-35 (a member of the spliceosome complex), thereby causing these proteins to aggregate in nuclear bodies (Cooper *et al.*, 1999; Fu & Maniatis, 1990). This disrupts the splicing and export of cellular mRNAs, a process which is intrinsically linked to nuclear export of these mRNAs. Furthermore, this function of ORF57 allows it to regulate the activity of ORF50, whereby the more potent transactivator of expression, ORF50a, is repressed by ORF57 later in infection as ORF50a is spliced. In contrast, the unspliced ORF50b is upregulated by ORF57 later in infection (Whitehouse *et al.*, 1998; Williams *et al.*, 2005).

1.30.3 Latency

One of the unique features of *gammaherpesvirinae* is their ability to persist as a high-copy number episome without viral integration in a dividing cell population during the latent phase of their infection. To this end, the viral genome is replicated during host cell mitosis and segregated to daughter cells (Fickenscher & Fleckenstein, 2001). During latency, only three genes are expressed from one polycistronic mRNA; ORF71, ORF72 and ORF73. ORF71 encodes an anti-apoptotic gene, termed viral FLICE protein (vFLIP), and ORF72 encodes a cyclin D homologue (Chang *et al.*, 1996; Thome *et al.*, 1997). These proteins protect latently infected cells from apoptosis induced by the extrinsic death receptor associated pathway, and promote progression from G_1 to S phase of the cell cycle, respectively.

ORF73 however, is responsible for HVS episomal maintenance by attaching the HVS episome to the host mitotic chromosome through its C-terminal region. The C-terminus of ORF73 possesses 2 chromosome association sites, termed CAS1 and CAS2, at amino acid positions 285-303 and 396-407, respectively (Calderwood *et al.*, 2004). ORF73 is also capable of interacting with the abundant cellular chromosome-associated protein, MeCP2, independently of these two CAS sites (a.a. 324-396), and Histone

H1 at a similar site (a.a. 342-379) (Griffiths *et al.*, 2008; Griffiths & Whitehouse, 2007). Furthermore, the cellular insulator protein (CTCF) is also capable of binding the HVS-episome at one of two CTCF Binding Sites (CBS) within the *orf73* promoter region (Zielke *et al.*, 2012). The features of the ORF73 protein product are demonstrated in Figure 1.15.

ORF73 tethering of the HVS episome is mediated by its interaction with TR domains within the H-DNA at two sites termed 73 binding region (BR)-1 and 73BR2 (Verma & Robertson, 2003). These regions are therefore essential for HVS latent persistence, as demonstrated by White et al. whereby, upon removal of the TR regions by RecA mediated recombination, viral genomes were produced which were incapable of latent persistence (White *et al.*, 2003).

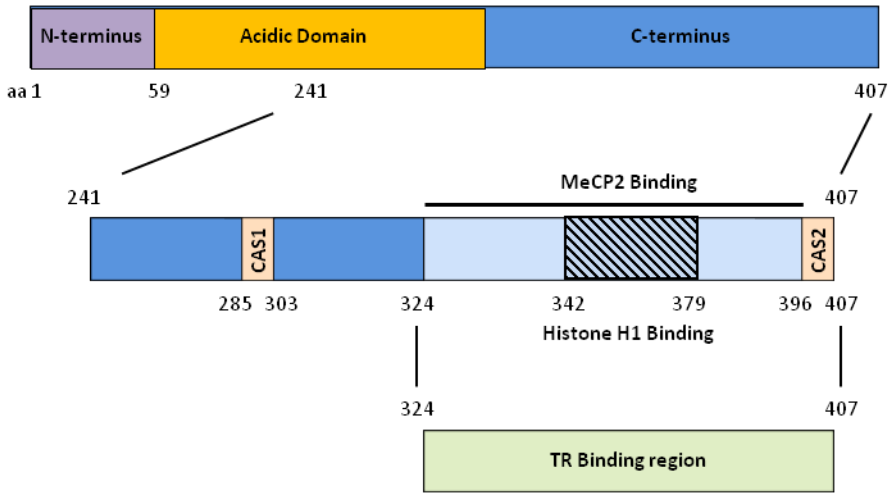


Figure 1.15: Features of the HVS ORF73 protein.

The ORF73 protein of HVS consists of an N-terminus which contains 2 nuclear localisation signals, a central acidic repeat domain, and the C-terminus, which is responsible for tethering the viral genome to host cell chromosomes. The C-terminus is capable of binding to the chromosomes either directly through two chromosome association sites (CAS 1 and 2), or indirectly through interactions with the chromosome associated proteins, MeCP2 and Histone H1. This region is also capable of binding the Terminal Repeat (TR) regions within the HVS genome, thereby allowing HVS to persist as a non-integrated episome. Adapted from (Turrell, 2010).

In addition to maintaining the HVS episome during latent infections, ORF73 negatively regulates the transactivator protein gene product of ORF50, thereby suppressing the lytic gene cascade (Schafer *et al.*, 2003).

1.31 Strain Variation and Oncogenicity

There are three HVS subgroups, termed A, B and C. These are classified on their ability to transform a variety of cells and on their homology within a highly variable 7kb region at the left-most region of the L-DNA (Medveczky *et al.*, 1984). These subgroups vary in their transforming capability, with strains A and C being capable of transforming common marmoset T-lymphocytes, and strain B being non-oncogenic (Brinkmann & Schulz, 2006).

1.31.1 STP

The transforming capabilities of HVS are associated with the product of the ORF1 gene which encodes the Saimiri Transformation-Associated Protein (STP). Both STP-A and STP-C contain a 17 a.a. cytoplasmic acidic N-terminus, 18 collagen-like repeats in the central region and a 16 a.a. hydrophobic, plasma spanning C-terminal domain. STP-B however, is a truncated version of the protein and does not possess the acidic N-terminal region or the collagen-like repeats (Choi *et al.*, 2000).

The collagen-like repeats appear to be essential for the transforming capabilities of STP, as insertion of this domain into the non-transforming *stp* gene of strain B confers transforming ability (Choi *et al.*, 2000). STP-C can associate with cellular Ras which stimulates the MAPK pathway (Jung & Desrosiers, 1995), whereas STP-A interacts with the SH2 motif of Src kinase and Stat3 (Chung *et al.*, 2004; Lee *et al.*, 1997). Furthermore, all STP genotypes are capable of interacting with the Tumour Necrosis Factor Receptor-Associated Factors (TRAFs), with STP-A and STP-C being capable of activating NF-κB via these interactions (Brinkmann & Schulz, 2006; Garcia *et al.*, 2007).

1.31.2 Tip

HVS strain C viruses encode a further protein with transforming potential adjacent to ORF1, termed the Tyrosine kinase Interacting protein (Tip). This

protein is membrane located and is largely unstructured, possessing a single C-terminal transmembrane domain, an N-terminal glutamate-rich region, serine-rich regions and three conserved tyrosine residues (Fickenscher & Fleckenstein, 2001; Mitchell *et al.*, 2007). Tip-mediated transformation is mediated through a strong interaction with the Src family protein tyrosine kinase, Lck, which is a key regulator of T-cell activation (Isakov & Biesinger, 2000). This interaction is mediated by the Lck Binding Domain of Tip, which consists of a 9 aa motif with homology to the C-termini of various Src family kinase domains (CSKH), a 20 aa linker and a 9 aa proline-rich consensus sequence for SH3 domain binding (Biesinger *et al.*, 1995; Jung *et al.*, 1995a). These domains work cooperatively to bind Lck, enabling transformation of T-lymphocytes by the T-cell receptor (TCR) signalling pathway (Heck *et al.*, 2006). However, there have been conflicting reports on this activation mechanism, as Tip-C488 overexpression in T-cells partially reversed the activities of constitutive Lck expression (Jung *et al.*, 1995b), making the activity of Tip in HVS-infection unclear. One explanation for this downregulation could be to evade host cell immune control. Furthermore, Tip has been shown to interact with the lysosomal protein, p80, enhancing lysosome formation and subsequently recruiting Lck into the lysosome for degradation (Park *et al.*, 2003). This in turn disrupts cellular trafficking by interacting with the Vps35 subunit of the retromer complex, leading to the downregulation of CD4 expression on the cell surface, thereby further reducing TCR signalling (Kingston *et al.*, 2011).

Tip also associates with the SH2 domains of STAT1 and 3 via a phosphotyrosine residue at position Y114 in Tip-C488, or Y72 in Tip-C484, which is also believed to mediate T-cell transformation (Hartley & Cooper, 2000; Lund *et al.*, 1997).

1.32 HVS miRNAs

Upon infection, viruses are able to exploit host cell miRNAs to achieve regulation of a number of viral constructs and to redirect host cellular machinery. This can either enable virus replication or shut off virally encoded proteins during latent infection (Ghosh *et al.*, 2009).

Several dsDNA viruses, including Herpesviruses, also encode their own set of miRNAs which allow the virus to regulate viral and host cell transcripts, thereby controlling latency or lytic replication. For instance, EBV encodes 25 miRNAs, some of which regulate latent infection of the virus. Firstly, miR-BART2 represses lytic infection through downregulation of the EBV transcription factor BALF5 (Barth *et al.*, 2008). Secondly, three EBV-encoded miRNAs tightly regulate expression of LMP1, which promotes cell proliferation during latent infection, but can result in cell cycle arrest and apoptosis if overexpressed (Lo *et al.*, 2007).

Additionally, Herpesvirus miRNAs can bear homology to cellular miRNAs, as observed for the miR-K12-11 of KSHV. This miRNA shares the same first 8 nucleotides as the cellular miRNA, miR-155, an miRNA involved in cell-growth regulation. Overexpression of miR-155 in B-cells is associated with lymphomas, thereby highlighting a role of virally encoded miRNAs in cancer progression (Eis *et al.*, 2005; Skalsky *et al.*, 2007).

Herpesvirus saimiri also encodes its own microRNAs, termed Herpesvirus saimiri U-rich RNAs (HSURs); 5 in HVS strain C288, and 7 in HVS strain A11 (Albrecht *et al.*, 1992; Biesinger *et al.*, 1990). Two of these virally encoded miRNAs (HSUR 1 and 2) have been shown to rapidly degrade the host cell miRNA, miR-27, which is an activator of Wnt and promotes proliferation (Cazalla *et al.*, 2010). The rapid degradation of miR-27 has also been observed upon murine CMV infection and, interestingly, appeared to be important for viral replication (Marcinowski *et al.*, 2012).

1.33 HVS as a Gene Therapy Vector

HVS possesses many features which make it attractive as a gene therapy vector. Firstly, like all herpesviruses, HVS possesses a large dsDNA genome (150kb), which gives it a large packaging capacity for heterologous DNA. Secondly, HVS is capable of forming a persistent latent infection, maintaining its genome as a non-integrated circular episome in host cells for up to 150 cell doublings (Grassmann & Fleckenstein, 1989; Simmer *et al.*, 1991). This allows prolonged transgene expression without the risk of insertional mutagenesis. Finally, HVS has a broad tropism and can infect a variety of cell types.

1.33.1 Tropism

HVS is capable of forming persistent latent infections in a wide variety of primary and cancerous cell lines. Indeed, intravenous and intraperitoneal injection of HVS expressing luciferase demonstrated, using non-invasive optical imaging, that HVS has a broad biodistribution. Viral DNA was detected in a number of organs, including the liver, spleen, lung, kidney, heart and brain. HVS infection was well tolerated in this mouse model, with predominantly ORF73 transcripts observed in infected organs, indicating the establishment of latent infection. In addition, no toxicity was observed in the liver, where the highest levels of infection were seen (Smith *et al.*, 2005).

HVS has a natural tropism for cells of the haematopoietic lineage, including T, B, Pre-B, myeloid and monocytoid cell lines and bone marrow stromal cells (Frolova-Jones *et al.*, 2000; Stevenson *et al.*, 1999; Stevenson *et al.*, 2000b). In addition to terminally differentiated cells, HVS is capable of infecting more primitive cells, as demonstrated by Doody and colleagues, who showed that HVS could infect CD34+ haematopoietic progenitor cells isolated from human umbilical cord blood and persist as a non-integrated episome through differentiation into erythroid cells (Doody *et al.*, 2005).

In addition to infecting primary pluripotent and haematopoietic cells, HVS is also capable of stably infecting a wide variety of cancer cell lines, including pancreatic, lung, and colorectal cells (Griffiths *et al.*, 2006), and is capable of infecting ESFT cell lines in multicellular spheroidal cultures (Smith *et al.*, 2004). HVS is also capable of maintaining latent infections in tumours *in vivo*, as demonstrated with MiaPaCa and SW480 tumour xenografts in Balb/C nude mice, without infecting surrounding tissues (Smith *et al.*, 2001) . A complete list of cell lines demonstrated to be stably infected with HVS is shown in Table 1.7. The ability of HVS to infect a broad range of cancer cells demonstrates its potential as a gene delivery vector for reprogramming cancer cells to their tumour-initiating stem cell-like state.

Cell Line	Status of DNA	Virus Production
Carcinoma Cells		
Lung- A549	Episomal	yes
Colorectal- SW480, HCT116	Episomal	no
Pancreatic- MiaPaCa, Panc1	Episomal/Linear	yes
Breast- MCF7	Episomal	no
Gastric- MKN45	Episomal	no
Oesophageal- Kyse 30, Oe19	Episomal	no
Ewings Family Tumours- A673, SK-ES	Episomal	yes
Hepatocellular Carcinoma- Huh7, Huh7.5, HepG2	Episomal	no
Human Haematopoietic Cells		
T cells- Jurkat, Molt 4	Episomal	no
Pre-B cell- LAS-221	Episomal	no
B Cells- Raji, Daudi	Episomal	yes
Myeloid- K56, HEL 92	Episomal	no
Monocytoid- THP-1, U937	Episomal	no
Pluripotent Cells		
Mouse Embryonic Fibroblasts- CCE	Episomal	no

Table 1.7: HVS infections and episomal persistence in various cell types. Adapted from (Griffiths *et al.*, 2006).

1.33.2 Engineering HVS-based Gene Delivery Vectors

Traditionally, modifying the HVS genome was very technically demanding due to its large size and fragility, a typical feature of all gamma-2 herpesviruses (Longnecker & Neipel, 2007). Initially, one approach was to clone the HVS C488 strain into an overlapping cosmid library and co-transfect these cosmids into the permissive owl monkey kidney cells. This resulted in reconstitution of recombinant HVS-based viruses, but was very time consuming, and required replication competent viruses (Ensser *et al.*, 1999; Wieser *et al.*, 2005).

Engineering HVS-based vectors was made easier upon the use of the bacterial F-factor-based bacterial artificial chromosome (BAC) system. This system can stably maintain DNA fragments of up to 300 kb in length and can be easily genetically manipulated in *E.coli* (Shizuya *et al.*, 1992). The first HVS-BAC was based on the HVS-C484-77 strain, but was incapable of forming persistent latent infections, as the bacterial replication elements

disrupted the TR regions of the virus. However, episomal maintenance could be restored by incorporation of four copies of the TR unit into the construct (Collins *et al.*, 2002). The second HVS-BAC was created by White et al. in 2003 and was based on the non-transforming HVS strain A11. Here, the bacterial replication elements were inserted in ORF15 of the HVS genome to prevent disruption of the TR regions and episomal maintenance. ORF15 encodes a viral homologue of the complement control protein, CD59, and is not required for HVS lytic or latent infections. The BAC element also contains an *I-PpoI* restriction site, enabling simple and fast cloning of desired genes into the HVS genome. *I-PpoI* is an extremely rare cutting enzyme, with only three restriction sites in the entire human genome, so there is little chance of a suitable restriction site being present in any transgenes cloned into the HVS-BAC. Genes can therefore be quickly cloned into the HVS-BAC by *I-PpoI* restriction digestion of the pShuttle Link1 construct containing the desired transgene, and subsequent ligation of the excised fragment into the HVS-BAC. The BAC element also contains the chloramphenicol resistance marker for easy selection of successful clones and either a GFP or RFP mammalian expression cassette, enabling quick identification of cells successfully infected with recombinant HVS virus. The features of the HVS-BAC cloning system are detailed in Figure 1.16.

The generation of the HVS-BAC allowed for improvement of the biosafety of HVS-based vectors by creating the HVS-based amplicon system (Macnab *et al.*, 2008). An amplicon is specifically a defective vector derived from a viral genome. It contains a therapeutic transgene expression cassette and the cis-acting sequences required for replication, cleavage and packaging into the Virus-Like Particle (VLP). Amplicons carry no, or few, trans-acting viral genes, and do not induce synthesis of viral proteins. Therefore, these vectors are non-toxic and non-pathogenic to the transduced cell or tissue.

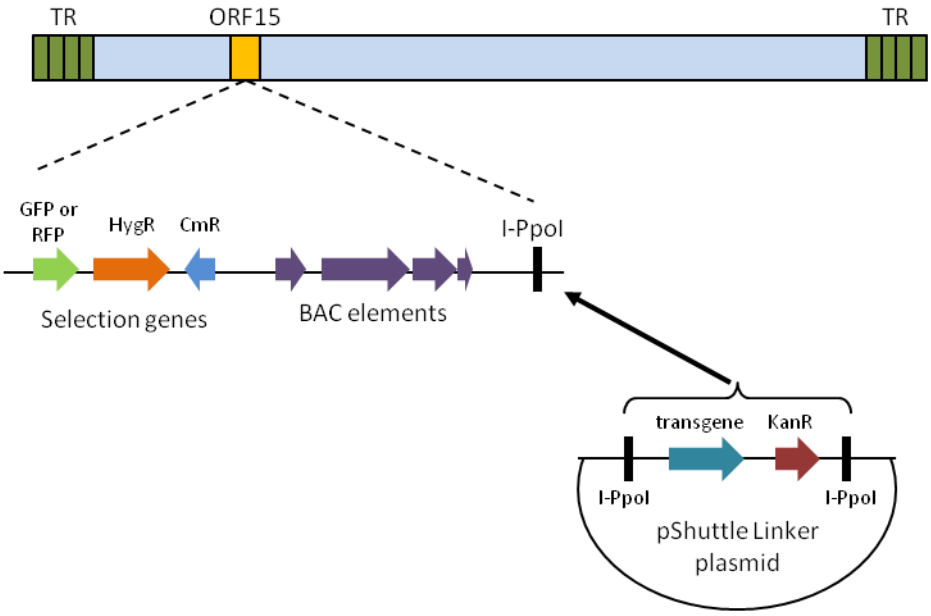


Figure 1.16: Features of the HVS-BAC system

The BAC elements were inserted within ORF15 of the HVS strain A11 genome via RecA mediated recombination, in addition to three selective marker genes. These genes comprised of a GFP or RFP fluorescent marker, and the hygromycin and chloramphenicol resistance genes. Transgenes of interest can be cloned into the BAC by utilising the pShuttle Linker plasmid which contains a multiple cloning site and a kanamycin selection marker, flanked by two *I-PpoI* restriction sites.

The amplicon-like system which has been generated for HVS was developed by digesting a 50kb fragment of HVS genomic DNA spanning ORFs 19-62. Although this is not a true amplicon as it still contains some viral gene sequences, it did remove the two genes essential for controlling lytic replication; namely ORF50 and ORF57. Furthermore, this HVS amplicon still contained ORF73, and was therefore capable of persisting as a non-integrated episome through repeated cellular divisions (Macnab *et al.*, 2008).

Further engineering of HVS-based vectors for gene therapy applications has involved re-targeting the virus to preferentially infect hepatocellular carcinoma cell lines. The naturally broad tropism of HVS is useful in gene therapy applications; however, there are certain cases where a more targeted delivery of the vector into specific diseased cells is required. To this end, HVS was genetically altered by RecA mediated recombination to modify the ORF51 glycoprotein; replacing the heparan sulphate binding

domain with a somatostatin receptor binding motif (Turrell & Whitehouse, 2011). Hepatocellular carcinoma cell lines often have increased levels of somatostatin receptor expressed on their cell surface. Therefore, this recombination aimed to re-target these modified HVS vectors to preferentially infect hepatocellular carcinoma. This recombination step swapped only 15 a.a. in the ORF51 region at positions 214-228 with 17 a.a. from the somatostatin binding motif. Interestingly, the resulting virus, termed $HVSmORF_{51,}$ did not display improved binding to somatostatin receptors. Perhaps this is not so surprising, as there is no way of ensuring the 17 a.a. somatostatin binding motif will be in the correct conformation or even exposed when it is translated as part of the ORF51 gene product. However, its binding to non-hepatocellular carcinoma human cell lines was perturbed, whereas binding to hepatocellular carcinoma lines remained high. Therefore, the $HVSmORF_{51}$ virus was successfully engineered to possess enhanced specificity for hepatocellular carcinoma cell lines (Turrell & Whitehouse, 2011).

1.33.3 HVS-based therapies

The potential of HVS as a gene-delivery vector was first realised in 1985. Recombinant HVS expressing the bovine growth hormone gene was able to successfully produce circulating bovine growth hormone upon infection of new world primates (Desrosiers *et al.*, 1985b).

Since this initial pre-clinical trial, HVS has been investigated as a gene therapy vector for the treatment of rheumatoid arthritis. HVS vectors encoding a combination of the anti-inflammatory cytokines IL-1RA and IL-10 were demonstrated capable infecting rheumatoid arthritis fibroblasts-like cells. Furthermore, these vectors conferred a ‘protective’ effect on cells treated with pro-inflammatory cytokines (Wieser *et al.*, 2005).

In addition to these initial studies, HVS has since been successfully engineered as an anti-cancer gene therapy vector. The potential of HVS as an oncolytic virus was first suggested by infection of the pancreatic cancer cell lines, MiaPaCa and Panc-1. HVS infection of these cells resulted in extensive cell death. In contrast, infection of all other carcinoma cell lines

tested showed no cytopathic effect, wherein the virus established a latent persistent infection (Stevenson *et al.*, 2000c).

The first HVS-based anti-cancer gene delivery vector generated expressed a 64 nt synthetic shRNA against the Endothelin-Converting Enzyme-1. Infection of prostate cancer cells with this vector resulted in a 32% reduction in endogenous expression of the mitogenic protein Endothelin-1. Moreover, this vector significantly reduced prostate cancer cell invasion in scratch assays and Matrigel invasion assays (Hong *et al.*, 2011). A further anti-cancer HVS-based therapy involved generating a recombinant HVS-based vector expressing a negative regulator of Wnt signalling, the Adenomatous Polyposis Coli (APC) protein. Upon infection of two colorectal cancer cell lines, SW480 and SW620, these cells displayed reduced proliferation and migration in scratch assays (Macnab *et al.*, 2011).

Most recently, HVS has been engineered to express the anti-cancer gene, TNF-Related Apoptosis Inducing Ligand (TRAIL). Here, the TRAIL gene was inserted under the control of the α-survivin promoter into a deletion mutant version of the HVS-BAC, HVSΔ71-73, as ORF71 of HVS encodes vFLIP (Thome *et al.*, 1997). The HVSΔ71-73-TRAIL virus was capable of inducing apoptosis in a number of human cancer cell lines, including colorectal (SW480), Hepatocellular (Huh7, Huh7.5, HepG2), Prostate (PC3), ESFTs (TC32), and lung (A549), among others. In addition to inducing apoptosis in these monolayer cultures, HVSΔ71-73-TRAIL infection was used in conjunction with a JAK-inhibitor to successfully disrupt spheroid cultures of the melanoma cancer cell line, Mel888, after 4 days post-infection (Turrell *et al.*, 2012).

1.34 HVS as a potential iPSC gene-delivery vector

As previously stated, there is great potential for episomally maintained viral vectors in iPSC generation. Episomally maintained reprogramming vectors would be able to provide prolonged expression of the reprogramming factors without the risks associated with retroviral insertional mutagenesis. Indeed, episomal persistence has already been utilised in a plasmid-based reprogramming method which utilised the EBV episomal maintenance protein EBNA1 and OriP origin of replication (Yu *et al.*, 2009). This system,

however, was highly inefficient, likely due to the lack of virus-mediated entry. Nevertheless, this study demonstrated the potential of a non-integrating, episomally maintained system in iPSC generation. Moreover, an additional advantage is that episomal vectors could be removed from the reprogrammed cell populations by repeated cellular division without selection. This is an extremely desirable trait in any vector used for iPSC technology, as withdrawal of transgene expression upon successful reprogramming of cells has been demonstrated to produce iPSCs which more closely resemble ESCs in their gene expression profiles (Soldner *et al.*, 2009; Sommer *et al.*, 2010). Because HVS episomal persistence is reliant on a single protein, ORF73, HVS-based iPSC reprogramming vectors could, in theory, be removed from reprogrammed cells by regulating expression of this viral gene.

Perhaps the most interesting feature of HVS as a potential gene delivery vector for iPSC reprogramming is the ability of HVS to infect undifferentiated ESCs and persist through their differentiation. This was demonstrated by Stevenson et al. whereby Mouse Embryonic Stem Cells (mESCs) were infected with HVS carrying a neomycin resistance gene and a GFP reporter gene. The initial infection rate in these cells was 20%, but infected cell populations were enriched by G418 selection. The virus remained as a non-integrated episome in these cells. Infected mESCs appeared to have no morphological changes, and no infectious virus was produced. Moreover, the episome was maintained in these cells during differentiation into macrophages, and was capable of maintaining GFP expression in these cells (Figure 1.17). Similar studies were performed in human CD34+ haematopoietic progenitors selected from cord blood. HVS-delivered expression of GFP was readily detected in cells of the erythroid lineage upon differentiation (Doody *et al.*, 2005).

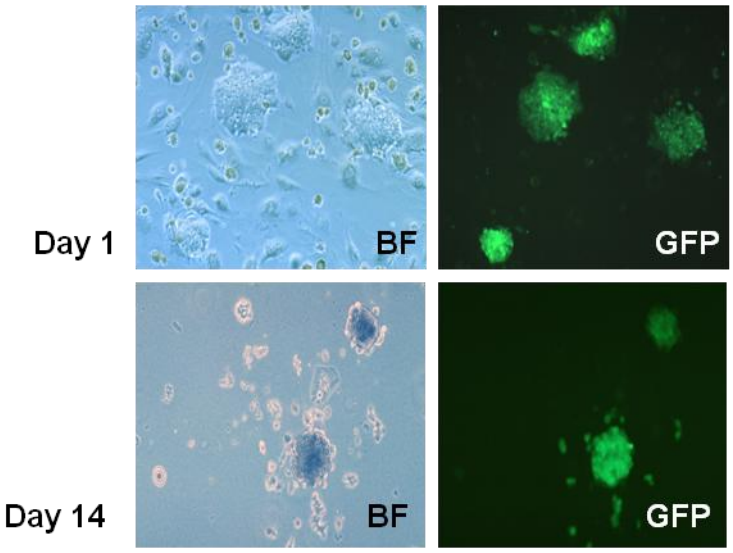


Figure 1.17: HVS episomal persistence in differentiating cells.
Mouse Embryonic Stem Cells (mESCs) were infected with HVS-GFP (Day 1), with an efficiency of 20%. After G418 selection to enrich the infected cell population, the mESCs were differentiated for a period of 14 days to terminally differentiated macrophages (Day 14). These cells displayed expression of the GFP reporter transgene from the HVS-GFP genome, demonstrating the ability of HVS-GFP to persist and maintain transgene expression within a differentiating cell population. These images were taken at 100x magnification. Adapted from (Stevenson *et al.*, 2000a).

These results indicate that HVS may be a suitable viral vector for iPSC generation, and is not subject to the same positional effect silencing which is observed in retroviral vectors (Hotta & Ellis, 2008). Moreover, A549 lung carcinoma cells infected with HVS carrying a GFP reporter gene demonstrated GFP transgene expression up to 12 months post-infection, indicating HVS is capable of prolonged transgene expression. This would be a valuable trait in iPSC reprogramming where prolonged expression of the reprogramming factors enhances the efficiency of iPSC generation (Hanna *et al.*, 2009).

Therefore, HVS possesses many desirable traits of an episomally maintained iPSC reprogramming vector, which could be used for safer iPSC generation. Furthermore, HVS-based vectors would not perturb the host cell gene functions through insertion of its genome, unlike their retroviral counterparts. Therefore, this virus presents a useful tool to reprogramme somatic cancer cells to tumour-initiating iPCs, which can then be used as an accurate model for studying tumour formation and progression.

1.35 Thesis Aims

This thesis will explore the potential of HVS-based reprogramming vectors to generate iPSCs and iPCs in a proof-of-principle study.

Firstly, HVS-based recombinant vectors will be produced expressing the iPSC transgenes, Oct4, Sox2, Lin28 and Nanog under the control of the immediate early CMV promoter.

Secondly, the potential of these vectors to reprogramme primary cells and cancerous cell lines will be investigated

Finally, resulting iPSC and iPC colonies will be characterised to determine the extent of reprogramming achieved within these cells. Furthermore, the potential of resulting iPCs to differentiate into various cell types will be determined.

.

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and reagents

All chemicals and reagents were purchased from Sigma-Aldrich and Melford unless otherwise stated. Primers, T4 DNA ligase, DNA ladders and protein ladders were supplied by Invitrogen. Restriction enzymes were purchased from New England Biolabs or Promega. Gel extraction and Maxi-prep kits were obtained from QIAGEN. Reagents for qRT-PCR were supplied by Bioline.

2.1.2 Plasmid constructs

Lentiviral constructs containing cDNAs of each iPSC gene, pSin-E2F-Oct4-Pur, pSin-E2F-Sox2-Pur, pSin-E2F-Lin28-Pur and pSin-E2F-Nanog-Pur, were obtained from Addgene. pEGFP-C1 was purchased from Clontech. The pShuttle Link 1 vector and the HVS-BAC construct were created by Dr. Robert White (White *et al.*, 2003).

2.1.3 Oligonucleotides

Oligonucleotide primers used for sequencing and Polymerase Chain Amplification Reactions (PCR) were purchased from Sigma-Aldrich. A full list of these primers is given in Table 2.1. The Oligo(dT) primer used for Reverse Transcription reactions was purchased from Promega.

2.1.4 Antibodies

Primary antibodies were supplied from a range of companies, or provided by Dr. Christian Unger (Centre for Stem Cell Biology, University of Sheffield), and are detailed in Table 2.2. Horseradish peroxidise (HRP) conjugated secondary antibodies were supplied by Dako and used for western blotting. Alexa-fluor conjugated anti-mouse Immunoglobulin G (IgG) antibodies, used for immunofluorescence microscopy, were purchased from Invitrogen.

Primer	Use		Sequence
***AgeI/EcoRI* Oct4**	PCR, Cloning, RT-PCR	FWD	GCGACCGGTATGGCGGGACACCTGGCTTCGGAT
		REV	GCGGAATTCTCAGTTTGAATGCATGGGAGAGCC
***AgeI/EcoRI* Sox2**	PCR, Cloning, RT-PCR	FWD	GCGACCGGTATGTACAACATGATGGAGACGGAG
		REV	GCGGAATTCTCACATGTGTGAGAGGGGCAGTGT
***AgeI/EcoRI* Lin28**	PCR, Cloning, RT-PCR	FWD	GCGACCGGTATGGGCTCCGTGTCCAACCAGCAG
		REV	GCGGAATTCTCAATTCTGTGCCTCCGGGAGCAG
***AgeI/EcoRI* Nanog**	PCR, Cloning, RT-PCR	FWD	GCGACCGGTATGAGTGTGGATCCAGCTTGTCCC
		REV	GCGGAATTCTCACACGTCTTCAGGTTGCATGTT
pEGFP *NotI/MluI*	PCR of expression cassette	FWD	ATAAGAATGCGGCCGCAACCGTATTACCGCCATGCAT
		REV	GCG ACGCGT GGACAAACCACAACTAGAATGC
qPCR Oct4	qRT-PCR	FWD	AAGCTGGAGAAGGAGAAGCTG
		REV	TCAAAGCGGCAGATGGTCGTT
qPCR Sox2	qRT-PCR	FWD	TGAACCAGCGCATGGACAGTT
		REV	GTTCATGTAGGTCTGCGAGCT
qPCR Lin28	qRT-PCR	FWD	AAGTGGTTCAACGTGCGCATG
		REV	GTGACACGGATGGATTCCAGA
qPCR Nanog	qRT-PCR	FWD	TGAGATGCCTCACACGGAGACT
		REV	TACACACAGCTGGGTGGAAGA
qPCR GFP	qRT-PCR	FWD	GCTGACCCTGAAGTTCATC
		REV	GATGCGGTTCACCAGGGTG
pShuttle REV	Sequencing	FWD	-
		REV	CTTGTGCAATGTAACATCAGAG
Endogenous Oct 4	qRT-PCR ESC marker	FWD	CGCAAGCCCTCATTTCACC
		REV	CACAGAACTCATACGGCGG
hTERT	qRT-PCR ESC marker	FWD	CCTGCTCAAGCTGACTCGACACCGTG
		REV	GGAAAAGCTGGCCCTGGGGTGGAGC
Klf4	qRT-PCR ESC marker	FWD	GCTGCTGAGTGGAAGAGAG
		REV	GAGAAAGATGGGAGCAGCG
Rex1	qRT-PCR ESC marker	FWD	CAGATCCTAAACAGCTCGCAGAAT
		REV	GCGTACGCAAATTAAAGTCCAGA
FLK1	qRT-PCR Differentiation marker	FWD	TGATCGGAAATGACACTGGA
		REV	CACGACTCCATGTTGGTCAC
MAP2	qRT-PCR Differentiation marker	FWD	CAGGAGACAGAGATGAGAATTCCTT
		REV	GTAGTGGGTGTTGAGGTACCACTCTT
Sox17	qRT-PCR Differentiation marker	FWD	CTCTGCCTCCTCCACGAA
		REV	CAGAATCCAGACCTGCACAA
MSX 1	qRT-PCR Differentiation marker	FWD	CGAGAGGACCCCGTGGATGCAGAG
		REV	GGCGGCCATCTTCAGCTTCTCCAG
Nestin	qRT-PCR Differentiation marker	FWD	CCGGGTCAAGACGCTAGAAGA
		REV	CTCCAGCTCTTCCGCAAGGTTGT
β-Tubulin III	qRT-PCR Differentiation marker	FWD	GTCCGCCTGCCTCTTCGTCTCTA
		REV	GGCCCCTATCTGGTTGCCGCACT

Table 2.1: Primer sequences and their applications

Target	Cat. No.	Working Dilution	Type	Provider
Oct4	SC-5279	1:200	Mouse Monoclonal	Santa Cruz
Sox2	Ab75485	1:200	Mouse Monoclonal	Abcam
Nanog	Ab76586	1:500	Mouse Monoclonal	Abcam
Lin28	Ab76369	1:500	Mouse Monoclonal	Abcam
GAPDH	Ab8245	1:5000	Mouse Monoclonal	Abcam
SSEA4		1:10	Mouse Serum	Dr. Christian Unger

Table 2.2: Primary Antibodies used for various applications

2.1.5 Bacterial Growth Media

LB medium (Luria Bertani) – Tryptone (1% w/v), yeast extract (0.5% w/v), NaCl (1% w/v). For LB agar plates, agar (1.5% w/v) was added to LB medium. Media was autoclaved at 15 lb/sq. Inch on a liquid cycle. For selective media, antibiotic was added at either 50 µg/ml or 12.5 µg/ml once the media had cooled to 50°C.

SOC medium (Super optimal broth with catabolite repression) – Tryptone (2% w/v), yeast extract (0.5% w/v), NaCl (0.05% w/v), 2.5 mM KCl. Media was autoclaved and, once cooled, sterile glucose and sterile $MgCl_2$ were added to a final concentration of 20 mM and 10 mM respectively.

2.1.6 Mammalian culture reagents

Media and supplements were supplied by Invitrogen, Lonza and Stem Cell Technologies. A full list of cell culture reagents and their suppliers is given in Table 2.3.

Medium	Supplements and Coating Agents	Provider
DMEM		Lonza
RPMI		Invitrogen
Knock-Out DMEM		Invitrogen
Neural Stem Cell Media Kit		Invitrogen
mTeSR		Stem Cell Technologies
	FCS	Invitrogen
	Knock-Out Serum Replacement	Invitrogen
	mTeSR Supplement	Stem Cell Technologies
	Penicillin/Streptomycin	Invitrogen
	hygromycin	Invitrogen
	β-mercaptoethanol	Sigma-Aldrich
	Non-Essential Amino Acids	Invitrogen
	Glutamax	Invitrogen
	Matrigel	BD Biosciences
	Gelatin	Sigma-Aldrich

Table 2.3: Mammalian cell culture media and supplements

2.2 Cloning

2.2.1 Restriction Enzyme Digest

Restriction enzymes were obtained from New England Biolabs. *I-PpoI* was purchased from Promega. Typically, restriction digests were performed in a 20 µl reaction, with 2 µl of the appropriate 10x buffer, 1 µl enzyme, 0.1-2 µg DNA, 17 µl ddH_2O, and were incubated for 1-16 hours at 37°C. Digested fragments were purified by agarose gel electrophoresis and extracted as described in section 2.2.6. For *I-PpoI* digests, 0.1 µl of enzyme was used and reactions incubated at 37°C for 50 minutes before heat inactivating the enzyme at 65°C for 20 minutes to prevent over-digestion.

2.2.2 Cloning Strategies

Restricted DNA fragments were separated by agarose gel electrophoresis and the bands visualised and excised using a UV transilluminator. DNA was extracted using the QIAquick Gel Extraction Kits as described in section 2.2.6. DNA was eluted off the columns in 30-50µl of elution buffer (10 mM Tris-HCl, pH 8.5).

Purified products were ligated into pre-digested parent vectors in a 3:1 molar ration of insert: vector. The ligation reaction consisted of 0.5 µl (200 Units) of T4 DNA ligase and 2 µl of 10x ligase buffer (50mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) made to a total volume of 20 µl with ddH_20 and samples incubated for 16 hours at 16ºC prior to transformation of chemically competent cells (section 2.2.7). A list of the recombinant plasmid constructs generated and their parent vectors is given in Table 2.4.

Recombinant Vector	Parent Vector	Primers Used	Restriction Enzymes Used For Cloning
pCMV-Oct4	pEGFP-C1	*Age/EcoRI* Oct4 FWD and REV	*AgeI/EcoRI*
pCMV-Sox2	pEGFP-C1	*Age/EcoRI* Sox2 FWD and REV	*AgeI/EcoRI*
pCMV-Lin28	pEGFP-C1	*Age/EcoRI* Lin28 FWD and REV	*AgeI/EcoRI*
pCMV-Nanog	pEGFP-C1	*Age/EcoRI* Nanog FWD and REV	*AgeI/EcoRI*
pShuttle-Oct4	pShuttle Link 1	pEGFP *NotI/MluI*	*NotI/MluI*
pShuttle-Sox2	pShuttle Link 1	pEGFP *NotI/MluI*	*NotI/MluI*
pShuttle-Lin28	pShuttle Link 1	pEGFP *NotI/MluI*	*NotI/MluI*
pShuttle-Nanog	pShuttle Link 1	pEGFP *NotI/MluI*	*NotI/MluI*
HVS-BAC-Oct4	HVS-BAC-GFP	-	*I-PpoI*
HVS-BAC-Sox2	HVS-BAC-GFP	-	*I-PpoI*
HVS-BAC-Lin28	HVS-BAC-GFP	-	*I-PpoI*
HVS-BAC-Nanog	HVS-BAC-GFP	-	*I-PpoI*

Table 2.4: List of recombinant vectors produced and their parental plasmids.

2.2.3 Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was performed with Platinum *Pfx* DNA Polymerase or *Taq* polymerase. Conditions for *Pfx* reactions consisted of denaturation for 5 min at 95°C followed by 25 cycles of 95°C for 1 minute, an annealing step of 55°C for 1 minute, and an extension step for 1 minute/1000 bp at 68°C. A final extension of 68°C for 10 minutes was used to ensure full extension of any remaining incomplete fragments. For reactions using *Taq* polymerase, a 5 minute incubation at 95°C was followed by 25-35 cycles of denaturation at 95°C for 1 minute, annealing of primers at 55°C, and extension at 72°C for 1 minute/1000 bp. Annealing temperatures were optimised for each primer pair. A list of all the primers used is given in Table 2.1.

2.2.4 Agarose Gel Electrophoresis

1% w/v agarose gels were prepared using electrophoresis grade agarose and 1x TBE buffer (0.045 M TRIS-borate, 0.002 M EDTA) and 0.1 µg/ml ethidium bromide was added prior to setting the gel. 5-10 µl of DNA loading buffer (1 mg/ml Orange G, 30% glycerol) was added to each sample before loading, along with either 1 kb or 100 bp DNA ladders. Electrophoresis was performed in 1x TBE buffer at 90 V or 120 V for typically 1 hour.

2.2.5 Pulsed Field Gel Electrophoresis

1.2% pulsed-field grade agarose gels were made 0.5x TBE buffer. Midrange I PFG Marker (New England Biolabs) was used along with λ DNA *HindIII* Digest (New England Biolabs) and 1 kb DNA ladder (Invitrogen) as size comparisons, and 10 µl of loading dye was mixed with DNA samples prior to loading. Electrophoresis was performed at 6 V for 9-16 hours, with an initial switch of 2 seconds and a final switch of 16 seconds, and the buffer chilled to 15.5°C using a Bio-Rad Model 1000 Mini chiller. Bands were visualised by staining the gels for 1 hour using 200 ml of 0.5x TBE containing 0.1 µg/ml of Ethidium Bromide.

2.2.6 Gel Purification of DNA

Following agarose gel electrophoresis, DNA bands were excised and purified using the QIAquick gel extraction kit according to the manufacturer's protocol. The gel slice was weighed and incubated with 3 times the gel volume of QG buffer (patented by manufacturer) at 50°C for 10 minutes to melt the agarose. Once melted, 1 volume of isopropanol was added and the sample was mixed before applying to a QIAquick column. The samples were centrifuged for 1 minute at 13,000 r.p.m and the flow-through discarded. The DNA bound to the column was washed once with 750 µl of buffer PE prior to centrifugation at 13,000 r.p.m for 1 minute to remove impurities. The flow-through was once more discarded and a further centrifugation step of 13,000 r.p.m for 1 minute was performed to remove the remaining PE buffer from the column. Finally, the DNA was eluted using 30-50 µl of ddH_2O, by centrifuging the sample at 13,000 r.p.m for 1 minute. Pure DNA was stored at -20°C until use.

2.2.7 Transformation

Chemically competent DH5α *E.coli* cells (kindly provided by Dr. Brian Jackson) were thawed on ice for 10 minutes prior to use. 1-5 µl of plasmid DNA was added to 50 µl aliquots of cells, and incubated on ice for 15-30 minutes before heat shock treatment in a 42°C water bath for 30 seconds. Cells were chilled on ice for 2 minutes prior to addition of 250 µl of LB, and incubated at 37°C at 200 r.p.m for 1 hour before plating on selective LB agar.

The HVS BAC DNA required the use of the E-Shot Electromax DH10β T1R Electrocompetent cells (Invitrogen) which are RecA- and more efficient at taking up large DNA fragments. 5-10 µl of HVS-BAC DNA was prepared by drop dialysis on 0.025 µm nitrocellulose membrane filters (Millipore) for 2 hours. Electroporation was performed as detailed in the product protocols. Briefly, cells were thawed for 10 minutes on ice before adding the dialysed viral DNA. Electroporation was performed using the following conditions: 2.0 kV, 200 Ω and 25 µF. Immediately after electroporation, 250 µl of SOC medium was added to the cuvettes and cells were incubated at 37°C, with shaking at 200 r.p.m for 1 hour before plating onto selective LB agar plates.

2.3 Purification of plasmid DNA and BAC DNA

2.3.1 Small scale plasmid purification (mini-preps)

The protocol was adapted from 'Molecular Cloning – A Laboratory Manual' (Sambrook *et al.*, 1989). Colonies were picked into 2 ml of selective LB medium and incubated overnight at 37°C, 200 r.p.m. 1ml of each culture was centrifuged for 1 min at 13,000 r.p.m in a bench-top centrifuge and the bacterial cell pellets resuspended in 100 µl of Solution 1 (50 mM glucose, 25 mM TRIS-HCl (pH 8.0), 100 mM EDTA (pH 8.0)). To lyse the cells, 200 µl of Solution 2 (0.2 M NaOH, 1% SDS) was added and samples mixed by inverting 4-6 times before incubating at room temperature for 5 minutes. Following lysis, 150 µl of ice-cold Solution 3 (60 ml 5 M potassium acetate, 11.5 ml glacial acetic acid, 28.5 ml dH_2O) was added to each sample to neutralise the pH and precipitate cellular protein and chromosomal DNA. Samples were mixed by inverting 4-6 times and incubated on ice for 10

minutes prior to centrifugation at 13,000 r.p.m for 10 minutes to pellet the cellular debris. The supernatant was transferred to a fresh eppendorf containing 450 µl phenol:chlorophorm (1:1, pH 8.0), mixed by vortexing, and centrifuged for 3 minutes at 13,000 r.p.m. The aqueous top fraction was collected and mixed with 1 ml of 100% ice-cold ethanol to precipitate plasmid DNA. Samples were centrifuged at 13,000 r.p.m for 10 minutes to pellet the DNA and washed with 1 ml 70% ethanol. DNA pellets were allowed to air dry for 5 minutes before resuspension in 50 µl ddH_2O with 0.2 µl RNAse A.

2.3.2 Large Scale plasmid purification (maxi-preps)

The QIAGEN Plasmid Purification Maxi Kit was used according to the manufacturer's protocols. 50 ml selective LB was inoculated with a 2 ml starter culture, and cells incubated at 37°C shaking at 200 r.p.m. Cells were pelleted by centrifugation at 5,000 r.p.m for 15 minutes at 4°C, and resuspended in 10 ml of resuspension buffer P1 (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 100 µg/ml RNAse A). Lyseblue was also added to buffer P1 to allow visualisation of complete lysis and neutralization in subsequent stages of the protocol. 10 ml of lysis buffer P2 (200 mM NaOH, 1% SDS) was added and samples inverted 4-6 times before incubating for 5 minutes at room temperature. Following lysis, the pH was neutralised using 10 ml of neutralisation buffer (3M potassium acetate pH 5.5, pre-chilled to 4°C), and samples were incubated on ice for 20 minutes to ensure complete precipitation of genomic DNA and protein. Cellular debris was removed by centrifugation at 5,000 r.p.m for 30 minutes, during which time a QIAGEN-tip 500 column was prepared using 10 ml of buffer QBT, (750 mM NaCl, 50 mM MOPS (pH 7.0), 15% (v/v) isopropanol, 0.15% (v/v) Triton X-100). The supernatant was added to the equilibrated column and passed through the resin by gravity flow. Once all the supernatant had passed through the column, it was washed twice with 30 ml of wash buffer QC (1 M NaCl, 50 mM MOPS (pH 7.0), 15% (v/v) isopropanol), and plasmid DNA eluted with 15 ml of elution buffer QF (1.25 M NaCl, 50 mM Tris-HCl (pH 8.5), 15% (v/v) isopropanol). DNA was precipitated by mixing the eluate with 10.5 ml of 100% isopropanol, and centrifugation at 5,000 r.p.m for 45 minutes at 4°C.

The supernatant was discarded and the DNA pellet washed with 70% ethanol and air dried for 5-10 minutes. Purified plasmid DNA pellets were resuspended in 400 µl of ddH_2O and quantified by nanodrop.

2.3.3 Small scale purification of low copy number plasmids and BACs

All solutions used for this method are described in section 2.3.1. Successful colonies were picked to 15 ml LB cultures containing appropriate antibiotic and cultured for 16 hours at 37°C, shaking at 200 r.p.m. Following centrifugation at 5,000 r.p.m for 10 minutes, the cell pellets were resuspended in 600 µl of Solution 1. 1.2 ml of lysis buffer (Solution 2) was added to the resuspended cells and the samples mixed by inverting 4-6 times before incubating at 18-25°C for 5 minutes. Lysis of the cells was stopped and by using 1 ml of ice cold neutralization buffer (Solution 3). Subsequently, the samples were centrifuged at 5,000 r.p.m for 20 minutes at 4°C to pellet the cellular debris. The samples were mixed with 2 ml of ice cold isopropanol and incubated on ice for 15 minutes prior to centrifugation at 5,000 r.p.m for 30 minutes to pellet the precipitated DNA, and the pellet washed in 70% ethanol. Pellets were air dried for 5-10 minutes before resuspension in 50 µl ddH_2O containing 0.2 µl RNAse A.

2.3.4 Large scale purification of low copy number plasmids and BACs

This protocol is modified from the large scale plasmid purification QIAGEN protocol, and all solutions are described in section 2.3.2. A 1 litre culture of LB containing appropriate antibiotic was inoculated with 2 ml of starter culture and incubated for 16 hours at 37°C, shaking at 200 r.p.m. The culture was centrifuged at 6,000 r.p.m for 15 minutes at 4°C using a Sorvall ultracentrifuge, SLA 1500 rotor using 4 x 250 ml containers. Cell pellets were resuspended in 5 ml of buffer P1 before combining the pellets to give 2 x 10 ml samples. 10 ml of buffer P2 was added to each of these and samples incubated for 20 minutes at 4°C on a rotating wheel to lyse samples without shearing the larger BAC DNA. Samples were then neutralized by addition of 10 ml of ice cold buffer P3 before incubating for a further 20 minutes on a

rotator wheel at 4°C. Cellular debris was removed by firstly centrifuging the samples at 12,000 r.p.m for 30 minutes before filtering the supernatant through sterile gauze to a fresh polypropylene tube and a further centrifugation at 5,000 for 15 minutes. Subsequently, 20 ml of 100% isopropanol was added to each of the two samples, and mixed gently by inversion. Following this, the samples were incubated on ice for 30 minutes and centrifuged at 5,000 r.p.m for 30 minutes for the first precipitation of the DNA.

The resulting pellet was resuspended in 500 µl of TE buffer by gently pipetting up and down with a p1000, cut-off tip and agitation in a 37°C water bath. 4.5 ml of QBT buffer was added and the two samples combined. A QIAGEN-tip 500 column was prepared using 10 ml of QBT buffer before adding the suspension and allowing this to enter the column by gravity flow. The column was then washed once using 20 ml of QC buffer, and DNA eluted using 20 ml of buffer QF. 30 ml of isopropanol was mixed with the eluate and incubated on ice for 30 minutes, prior to centrifugation at 5,000 r.p.m for 30 minutes. The resulting pellet was washed in 70% ethanol, resuspended in 100 µl TE buffer, and stored at 4°C to prevent repeated freeze-thaw which could compromise the stability of the fragile HVS-BAC DNA.

2.4 Cell Culture

2.4.1 Cell Maintenance

Foreskin Fibroblasts, Human Embryo Kidney 293T, , Owl Monkey Kidney and Huh7 cells were all maintained in DMEM containing 10% heat-inactivated Foetal Calf Serum (FCS), and 1% penicillin/streptomycin, on Corning tissue culture-grade plastic.

The murine mESC-derived Neural Stem Cell NS5 adherent cell line was kindly provided by Dr. Ian Wood (School of Biomedical Sciences, University of Leeds).

The Ewings sarcoma cell lines, TC32 and TTC466, were cultured in RPMI containing 10% FCS and 1% penicillin/streptomycin. A673 cells were grown in DMEM with 10% FCS and 1% penicillin/streptomycin. All Ewings cell lines

were cultured on Primaria tissue culture-grade plastic (BD biosciences). Induced pluripotent stem cell line (iPSC), ShiPS-FF5 was kindly provided by Dr. Christian Unger (Centre for Stem Cell Biology, University of Sheffield). They were reprogrammed through lentiviral overexpression of Oct4, Sox2, Nanog and Lin28 in human foreskin fibroblasts. ShiPSFF5 cells are karyotpically normal, have typical embryonic stem cell morphology, and express pluripotency markers SSEA4, Tra1-81, Oct4 and Nanog. iPSCs and induced pluripotent cancer stem cells (iPCs) were cultured on Matrigel ESC qualified extracellular matrix (BD Biosciences) coated tissue cultured plates, in mTeSR 1 complete medium. The cells were split 1 in 6, once every 5-7 days by treating with 1 mg/ml collagenase type IV (Stem Cell Technologies) for 5 minutes at 37°C. The collagenase was removed and cells were washed twice with 2 ml of pre-warmed serum-free DMEM-F12 and 1 ml of fresh mTeSR added to the well. The colonies were scored by scraping with a blue tip and pipetted up and down with a P1000 pipette 2-3 times to break up the colonies, prior to centrifuging the cells at 400 r.p.m to replace the mTeSR. Cells were split 1 in 6 or 1 in 10 to pre-coated 12 well or 6 well dishes.

Cells were stored in a humid atmosphere with 5% CO_2 at 37°C. All cells were tested monthly for mycoplasma and they genotype was confirmed by suppliers.

2.4.2 Reprogramming A673 cells

1 day prior to infection, 2 x 10^4 A673 cells were plated per 6 well dish. Reprogramming was initiated by infecting cells with an MOI of 0.5 of HVS-based vectors per well. Viral supernatants were added to 2 ml of DMEM containing 5% FCS. If multiple viruses were used, the total MOI of virus was maintained using HVS-GFP. After day 1 post-infection (p.i.), the medium was removed, the cells washed twice with 2 ml of 1x DPBS, and fresh 10% DMEM added to the cells. On day 2 p.i., cells were dissociated using 0.05% trypsin-EDTA and plated onto 10 cm Primaria dishes in 10% DMEM. Day 3 p.i., cells were treated with 100 µg/ml of hygromycin B and maintained in selective medium. Medium was replaced every 2-3 days until the emergence of colonies (around day 19 p.i.). Colonies were collected by washing them off the 10 cm plates in DPBS by gentle agitation. Subsequently, colonies

were centrifuged at 400 r.p.m for 3 minutes and resuspended in medium comprising ½ DMEM 10%, and ½ mTeSR complete medium, and plated into 2-4 wells of a 24 well dish, which had been pre-coated in Matrigel ESC qualified extracellular matrix. Alternatively, colonies were picked directly from the 10 cm dishes using a P20 pipette, and each colony transferred to a 24 well dish. The following day, the medium was replaced with mTeSR complete medium. The colonies were maintained for a further 1-2 weeks in 24 well dishes until large enough to split (usually 1-2 mm in size). To split colonies the medium was firstly replaced with fresh mTeSR and each colony scored with a yellow tip 4-6 times and pipetted up and down twice to break up clumps. No cell dissociation reagents were used in the first split. The cells were transferred to a 6 well dish in 2 ml mTeSR.

2.4.3 Freezing and Thawing of cells

Cells were dissociated from tissue culture plastics using 0.05% trypsin-EDTA and washed with medium containing 10% FCS. Following centrifugation at 1500 r.p.m for 3 minutes, the medium was removed and cells resuspended in freezing medium, (DMEM or RPMI with 20% FCS and 10% Dimethyl Sulphoxide (DMSO)). Cells were frozen in 1.8 ml cryovials (Nunc) in a Mr. Frosty at -80°C overnight before transferring to liquid nitrogen for long term storage. Frozen cells were recovered by rapid thawing in a 37°C water bath, washing in 10 mls of growth medium and resuspending in fresh growth medium. For iPSCs and iPCs the same procedure was followed with the substitution of freezing medium with mFreSR stem cell freezing medium (Stem Cell Technologies).

2.4.4 Transfection

Cells were cultured in Corning 6-well tissue culture dishes until 50-70% confluency. Transfections were performed in 2 ml serum free medium without antibiotics. 3 µl Lipofectamine 2000 (Invitrogen) was used per 1 µg of DNA transfected into cells, and transfections were performed in a final volume of 200 µl. 4-6 hours post-transfection, the medium was replaced with medium containing 10% FCS, 1% penicillin/streptomycin.

2.4.5 Transducing cells

Cells were transduced with an MOI of 0.5-1 of each virus in DMEM with 5% FCS for 24 hrs. Cells were washed twice with 2 ml of DPBS and 10% medium added to the cells. If multiple viruses were used, an MOI of 0.5 of each virus was used and total MOI of virus maintained using HVS-GFP.

2.5 Virus Generation and Characterisation

2.5.1 Generation of Viral Stocks

The HVS-BAC-iPSC constructs were each transfected into the permissive cell line, Owl Monkey Kidney (OMK) cells. 6-12 µl of BAC DNA was mixed with 10 µl of Lipofectamine 2000 in serum free medium, according to the manufacturer's protocols, and 200 µl of this mixture was added to 2 ml total volume of serum free medium per 3.5 cm dish. After 6 hours, the medium was replaced with DMEM containing 5% FCS, and cells were cultured for 2-3 weeks until the cell sheet was lysed. Virus containing supernatant was subsequently used to infect $1x10^7$ OMK cells to produce higher titre viral stocks. These higher titre viral supernatants were harvested when the cell sheet was lysed (typically 3-7 days p.i.).

2.5.2 Titration of Viral supernatants

Upon production of fresh viral supernatants, titres were performed in OMK and A673 cells. Serial 1 in 10 dilutions were made in 5% DMEM and 1 ml of these dilutions was added to each 6 well dish of cells. 24 hours post-infection, cells were analysed for GFP expression from the virus by flow cytometry

The GFP expressing Units/ml of each viral supernatant was calculated using the following equation:

$$\frac{\left(\left(\frac{Total\ no.of\ cells}{100} \times \%\ cells\ infected\right) \times dilution\ factor\right)}{volume\ of\ innoculum}$$

2.5.3 Flow Cytometry

Cells were prepared for flow cytometry using 0.05% trypsin-EDTA per 6 well dish, and washed in 1 ml DMEM 10% FCS to inactivate the trypsin. Cells

were centrifuged at 1,500 r.p.m for 3 minutes before washing once in 1 ml of 1x DPBS. Subsequently, cells were pelleted by centrifugation at 1,500 r.p.m for 3 minutes, resuspended in 0.5 ml of 1x DPBS, and stored on ice prior to flow cytometry analysis. Cells were live sorted using a Becton Dickinson BD-LSRFortessa and results were analysed using the FACS DiVa6 software (BD Biosciences).

2.6 Electrophoretic analysis of proteins

2.6.1 Tris-glycine SDS-PAGE

$1x10^6$ cells were washed once in 1x DPBS before lysing with 100-250 µl RIPA (50 mM Tris base, 150 mM NaCl, 1% NP40, pH 7.6) containing 2% EDTA-free protease inhibitor (Roche) for 20-30 minutes on ice. Samples were centrifuged at 13,000 r.p.m to pellet cellular debris, and the supernatants harvested. Protein concentration was determined with the DC assay (BioRad) as directed by the manufacturer's protocol. Samples were mixed with an equal volume of 2x Laemmli buffer (50 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 20% (v/v) glycerol, 50 µg/ml bromophenol blue, 10 mM DTT), and incubated at 95°C for 5 minutes. All SDS-PAGE was performed using a BioRad mini-PROTEAN 3 cell, set up according to the manufacturer's instructions. Samples were loaded onto 10% or 12% polyacrylamide Tris-glycine gels, overlaid with 5% polyacrylamide Tris-glycine stacking gel. 10-20 µg of protein was loaded onto the gels before running at 180 V for typically 45 minutes, until the bromophenol blue had reached the bottom of the gel. SeeBlue Plus 2 (Invitrogen) or Colourplus (New England Biolabs) pre-stained protein standard markers were run alongside samples to indicate molecular weight.

2.6.2 Western Blotting

10-20 µg of protein was separated by SDS-PAGE as described in section 2.6.1, prior to transferring the proteins onto Hybond C extra nitrocellulose (Amersham biosciences), using the BioRad Mini Trans-Blot Electrophoretic Transfer Cell, as stated in the manufacturer's instructions. The gel, four pieces of Whatman 3 mm filter paper, and one piece of Hybond C were arranged in a specific order and soaked in transfer buffer (25 mM Tris-base,

190 mM Glycine, 20% (v/v) methanol), before being secured in the Mini Trans-Blot, according to the manufacturer's instructions. Proteins were transferred at 100 V for 1 hr. Subsequently, membranes were blocked for 1 hr at 18-25°C in 5% non-fat milk (Marvel) dissolved in 1 x TBS (Tris-base, 500 mM NaCl) containing 0.1% TWEEN (TBST), before incubating with primary antibody diluted to the appropriate concentration (see Table 2.2) in 5% milk/TBST for 16 hr at 4°C. Membranes were washed three times for 10 minutes in 1x TBST before incubating with HRP-conjugated secondary antibody (Dako) diluted in 5% milk/TBST for 1 hour. Three further washes of 10 minutes in 1 x TBST were performed, and membranes incubated with the EZ-ECL kit (Geneflow) prior to exposure on ECL hyperfilm (Amersham).

2.7 Characterisation of iPC cells

2.7.1 Alkaline Phosphatase staining of cells

The Alkaline Phosphatase staining kit II from Stemgent was used to stain cells as described in the manufacturer's protocols. Briefly, the cells were washed with 2 ml 1x DPBS fixed for 2 minutes with Fix Solution (4% w/v paraformaldehyde) and washed using 1x DPBS containing 0.05% v/v Tween-20. The staining solution was prepared by mixing 1.5 ml of Solution A (<1% w/v Fast Red Violet solution) with 1.5 ml of Solution B (<1% w/v Napthol AS-BI Phosphate solution) and incubating at room temperature for 2 minutes before finally adding 2 ml of Solution C (<1% w/v Sodium Nitrate solution). The prepared staining solution was immediately added to the fixed cells and incubated at room temperature in the dark for 15 minutes. Following this, the staining solution was removed and the cells washed 3 times using 2 ml of 1x DPBS.

2.7.2 Reverse-Transcription Polymerase Chain Reaction

2.7.2.1 RNA extraction

Cells were harvested using 1 ml of Trizol per 6 well dish. 200 µl of chloroform was added, and the samples were mixed by vigorous shaking for 15 seconds before incubating at 18-25°C for 2-3 minutes. Samples were then centrifuged at 13,000 r.p.m, 4°C for 15 minutes and the top layer of the supernatant collected and transferred to a fresh 1.5 ml eppendorf containing

0.5 ml of 100% isopropanol. The samples were mixed and incubated at 18 - 25°C for 10 minutes to precipitate the RNA. RNA was then pelleted by centrifugation at 13,000 r.p.m for 10 minutes at 4°C, and the pellet washed in 1 ml 70% ethanol before a further centrifugation at 9,000 r.p.m for 5 minutes. The ethanol was removed and pellets allowed to air dry for 5 - 10 minutes before resuspending them in 20 µl of RNAse-free ddH_20 at 18-25°C for 30 minutes. RNA was stored at -80°C and repeated freeze-thawing avoided.

2.7.2.2 Generation of cDNA

Total cellular RNA was extracted as described in section 2.7.2.1. The concentration of each sample was determined by nanodrop and 10 µg of RNA was treated with 1 µl of 2 Units/µl DNA-free DNAse (Ambion) in a 50 µl reaction containing 5 µl of 10x DNAse buffer (100 mM Tris-HCl pH 7.5, 25 mM $MgCl_2$, 5 mM $CaCl_2$) for 30 minutes at 37°C, according to the manufacturer's protocol. The reaction was stopped using 5 µl of DNAse Inactivation Reagent, and samples centrifuged for 2 minutes at 13,000 r.p.m. The supernatant, containing DNAse-treated RNA, was transferred to a fresh eppendorf and 5 µl of treated RNA was used to produce cDNA using M-MuLV Reverse Transcriptase (New England Biolabs) according to the manufacturer's protocol. This required incubating the RNA with 1 µl of Oligo(dT) primer and 1 µl of 10 mM dNTP mix in a total reaction volume of 13 µl before incubating at 65°C for 5 minutes. The samples were chilled on ice for 1 minute before adding 1 µl M-MuLV reverse transcriptase, 2 µl of M-MuLV buffer, and 1 µl of RNAse OUT (Invitrogen), with the total volume of the reaction mix made up to 20 µl with ddH_2O. Reverse transcription reactions were performed at 42°C for 1 hour, prior to enzyme inactivation at 80°C for 5 minutes. Negative controls containing no reverse transcriptase enzyme were also prepared to demonstrate the samples were free from DNA contamination. cDNA samples were stored at -20°C.

2.7.2.3 Reverse Transcription Polymerase Chain reaction (RT-PCR)

Total RNA was extracted as described in section 2.7.2.1, and cDNA produced (section 2.7.2.2). 2.5 µl of cDNA was subsequently PCR amplified as described in section 2.2.3. A typical reaction cycle for RT-PCR involved

an initial denaturation step of 95°C for 5 minutes before 35 cycles of denaturation at 95°C for 30 seconds, primer annealing at 55°C for 30 seconds, and extension for 30 seconds at 72°C. Samples were analysed using agarose gel electrophoresis as described in section 2.2.4.

2.7.2.4 Quantitative Real Time-Polymerase Chain Reaction (qRT-PCR)

Samples were set up in triplicate for each condition and total RNA harvested as described in section 2.7.2.1. cDNA was generated, as described in section 2.7.2.2, and diluted to a final concentration of 2 ng/µl. 10 ng of cDNA (5 µl) was mixed with 5.5 µl ddH_20, 2 µl of 10 mM primer mix and 12.5 µl of 2x SensiMix SYBR Green (Bioline), and qPCR performed using a Rotor-Gene Q 2-plex Thermocycle (Qiagen). Typical reaction conditions for a qPCR run involved a denaturation of 10 minutes at 95°C followed by 45 cycles of 95°C for 15 seconds, primer binding at 55°C for 30 seconds and extension with acquisition of fluorescence levels at 72°C for 20 seconds. Rotor-Gene 2000 software was used to analyse results.

2.7.3 Immunofluorescence

Cells were grown in 24 well dishes pre-coated in Matrigel and fixed after 3 days in culture using 0.5 ml of 4% Paraformaldehyde in PBS per 24 well for 15 minutes. Cells were washed 3 times using 2 ml 1 x DPBS and blocked for non-specific antibody binding using 1 x DPBS containing 1% Bovine Serum Albumin for 1 hour at 37°C. Primary antibody was diluted to the appropriate concentration in 1 x DPBS containing 1% Bovine Serum Albumin, and samples incubated with 200 µl of primary antibody at 37°C for 1 hour. Five washes in 1 x DPBS were performed to remove the primary antibody, before incubating samples with 200 µl of fluorescently-tagged secondary antibody diluted to the appropriate concentration at 37°C for 1 hour. Secondary antibody was removed and the cells washed 5 times in 1 x DPBS before staining the DNA with 5 µg/ml Hoechst stain in 1 x DPBS for 15 minutes. Hoechst was then removed and the cells washed a further 3 times using 2 ml of 1 x DBPS.

2.7.4 Microscopy

Images of colonies were taken using a DM1L fluorescent microscope attached to a Leica DC 300 camera, or Zeiss AxioVert 200 inverted microscope with images captured using a Hamamatsu CCD camera controlled by Openlab software using 10x or 20x magnification. Immunofluorescence analysis was performed on an inverted Zeiss AxioVert 135 TV microscope using a Zeiss AxioVertCam HRC camera under 20x magnification. Images were analysed using ImageJ software.

2.8 Differentiation of Induced Pluripotent Stem Cells and Induced Pluripotent Cancer Stem Cells

2.8.1 Embryoid Body Formation

Cells were split, as described in section 2.4.1, and colonies centrifuged at 400 r.p.m for 3 minutes before resuspending in 6 ml of differentiation medium (Knock-out DMEM, 20% Knock-out Serum Replacement,1% Non-essential amino acids, 1mM Glutamax, and 0.1mM β-mercaptoethanol), or 6 ml of DMEM containing 10% FCS, and plated out onto low-adherent tissue culture dishes in a total volume of 2 ml of the chosen differentiation medium. The Embryoid Bodies (EBs) were cultured for 6-8 days as suspension cultures, with medium exchanged every 2-3 days, before transferring half of the cells onto 6 well dishes pre-coated in 2 ml of 0.1% gelatin. The remaining cells were harvested in Trizol for later analysis. EBs were further cultured on gelatin coated plates for 6-8 days and cells imaged using DM1L fluorescent microscope before being harvested in Trizol.

2.8.2 Directed Neural Differentiation

Colonies were split and EBs formed as described in section 2.8.1. The EBs were then cultured in low-adherent dishes in 2 ml of EB medium containing 1 μM All-Trans retinoic acid. Cells were cultured in this medium in suspension cultures for 2 days before transferring the EBs to gelatin coated 6 well dishes. These EBs were allowed to attach to the gelatin-coated plates, and cultured for a further 8 days in EB medium without All-Trans retinoic acid.

Alternatively, the EB were cultured in Neural Stem Cell (NSC) Medium for 2 days in suspension culture before transferring to gelatin coated dishes. These EBs were then cultured for a further 8 days in NSC medium.

Chapter 3
Generation of HVS-based iPSC-gene Delivery Vectors

3 Generation of HVS-based iPSC-gene Delivery Vectors

3.1 Introduction

Induced pluripotent stem cell (iPSC) technology involves generating stem-cell like cells from adult somatic cells by exogenous expression of certain reprogramming factors such as Oct4, Sox2, Klf4 and Myc (Takahashi & Yamanaka, 2006). This technology therefore has the potential to generate stem cells which are patient-specific and ethically sourced, and is of great interest in stem cell-based therapies. However, initially there were certain problems with the biosafety of iPSCs. Firstly, two of the original reprogramming factors used, Myc and Klf4, are oncogenic and lead to cell proliferation; although this issue was partially solved by the discovery that these factors could be replaced with Lin28 and Nanog, thereby making the reprogramming process safer (Yu *et al.*, 2007).

A further issue with iPSCs was the requirement of integrating retroviral vectors for delivery of the reprogramming factors. Retroviral vectors have the advantage of providing prolonged expression of the reprogramming factor transgenes which is essential for efficient reprogramming. However, retroviruses preferentially integrate into highly expressed regions of the genome and can disrupt normal gene function by causing overexpression of genes related to proliferation or, alternatively, silence regulatory genes. There have been many attempts to develop safer reprogramming vectors including generating excisable retroviral vectors using Cre/LoxP recombination (Kaji *et al.*, 2009) or piggyBac transposons (Woltjen *et al.*, 2009). However, both these systems still leave behind a “footprint” after excision which can still disrupt normal gene function and require very stringent screening processes to ensure all viral DNA is excised. Alternative gene delivery methods have been assessed, including adenoviral infection (Stadtfeld *et al.*, 2008), repeated plasmid transfection (Okita *et al.*, 2008), mRNA transfection, (Warren *et al.*, 2010) and cell permeable recombinant reprogramming factor proteins (Kim *et al.*, 2009a). While these transient delivery methods have had some success, their efficiencies are poor when compared to the retroviral vectors. For example, the Epstein-Barr virus latency associated nuclear antigen (EBNA-1) has previously been used in a

plasmid-based episomal system to successfully reprogramme fibroblasts. However, this study chose to use inefficient plasmid transfection to deliver the reprogramming factors rather than utilising the natural infectivity of the virus. Furthermore, EBNA-1 has been shown to immortalise B cells in vitro, thus questioning the safety of this method (Young *et al.*, 1988).

Recently, certain small molecules have been identified which are capable of replacing the activities of the reprogramming factors. Successful reprogramming of keratinocytes has been achieved with a cocktail of small molecule inhibitors and Oct4 expression (Zhu *et al.*, 2010). As yet however, iPSCs have not been generated using small molecules alone; therefore there is still the requirement for the delivery of exogenous Oct4.

The aim of this chapter was to generate reprogramming vectors capable of efficient delivery and prolonged transgene expression without integration and disruption of normal host cell gene function using Herpesvirus saimiri (HVS) based viral vectors. HVS is a prototype member of the gamma type II *herpesvirinae* and is capable of maintaining its genome as a non-integrated episome in host cells. This episome is tethered to host cell chromosomes by the virally encoded protein ORF73, which interacts with the abundant chromosome associated protein, MeCP2 (Calderwood *et al.*, 2004; Griffiths & Whitehouse, 2007), thus allowing long term persistence in a dividing cell population.

All HVS-based vectors are based on the HVS-A11 strain, which contains a natural deletion removing the transforming gene Tyrosine Kinase Interacting Protein (Tip) (Biesinger *et al.*, 1990). The biosafety of these viral vectors is further improved by deleting the STP coding region (Griffiths *et al.*, 2006).

Perhaps of particular interest in iPSC technology is the ability of HVS to persist and provide prolonged transgene expression in differentiating cell populations. This was demonstrated by Stevenson and colleagues, whereby Mouse Embryonic Stem Cells (mESCs) were infected with a HVS-based vector encoding a GFP reporter gene. These mESCs were then differentiated to terminally differentiated macrophages which remained GFP positive (Stevenson *et al.*, 2000a). Therefore, HVS could potentially be capable of maintaining its episome in reprogramming cells without transgene

silencing occurring. Furthermore, HVS has a naturally broad tropism and has been demonstrated to be capable of infecting both a wide variety of primary tissues and cultures (Doody *et al.*, 2005; Simmer *et al.*, 1991; Smith *et al.*, 2005; Stevenson *et al.*, 1999) and also cancer cells (Smith *et al.*, 2004; Smith *et al.*, 2001). These properties make HVS an attractive viral vector for use in reprogramming both primary and cancer cells to a stem cell-like state.

In order to produce HVS-based vectors expressing Oct4, Sox2, Lin28 and Nanog, a multistep cloning strategy was performed as detailed in Figure 3.1. This chapter details each stage of this cloning procedure and the generation of infectious HVS-iPSC virus particles.

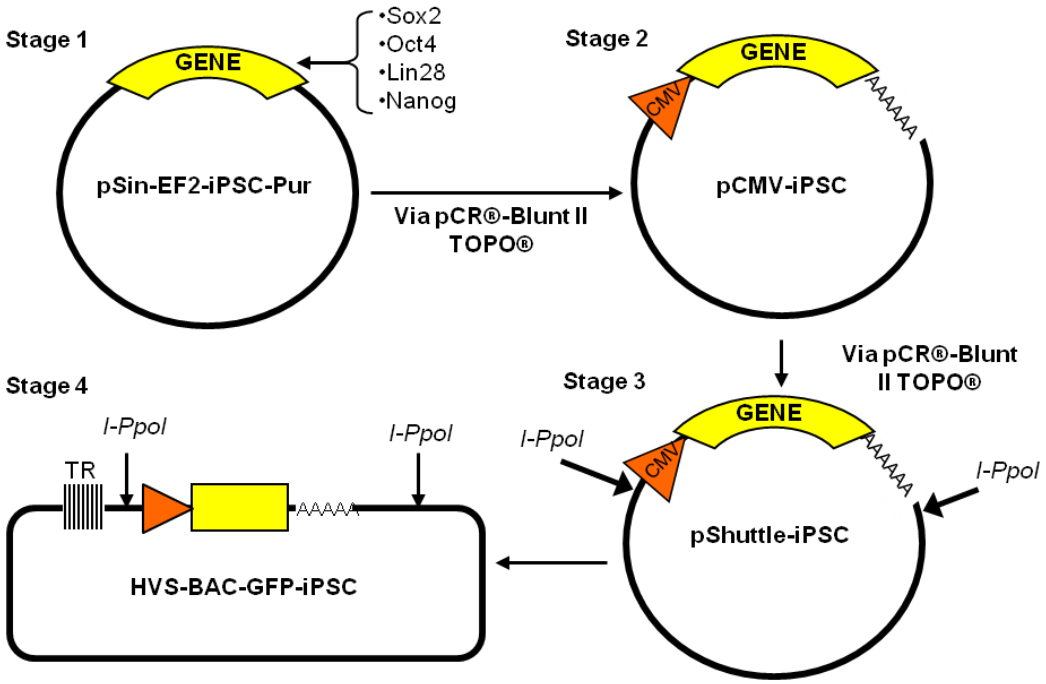


Figure 3.1: Overview of the Generation of the HVS-based iPSC vector constructs

The iPSC coding sequences were firstly PCR amplified to incorporate *AgeI/EcoRI* restriction sites, enabling the fragments to be cloned into the empty pEGFP-C1 vectors to create the pCMV-iPSC constructs. These constructs contain a specific iPSC gene under the control of the CMV promoter. Each CMV-iPSC expression cassette was then PCR amplified to incorporate *NotI/MluI* restriction sites, allowing them to be cloned into the pShuttle Link 1 vector. Finally each expression cassette was cloned into the HVS-BAC-GFP at the *I-PpoI* sites.

3.2 Stage 2: Construction of pCMV-iPSC constructs

To provide strong expression of each of the transgenes, the first stage of the cloning strategy involved placing each gene under the control of the human Cytomegalovirus (CMV) promoter. To achieve this, the EGFP cloning vector, pEGFP-C1 was utilised. This plasmid contains the EGFP coding sequence under the control of a CMV promoter and SV40 polyadenylation signal. The cloning strategy devised removed the EGFP coding region by *AgeI/EcoRI* restriction digestion and replaced this with the coding region of the iPSC genes, yielding the expression constructs, pCMV-iPSC. As no suitable restriction sites were present flanking the coding sequences for the reprogramming genes in their parent vectors, PCR amplification of the genes was required to incorporate *AgeI/EcoRI* restriction sites flanking the gene of interest from the parental plasmid, pSin-E2F-iPSC-Pur.

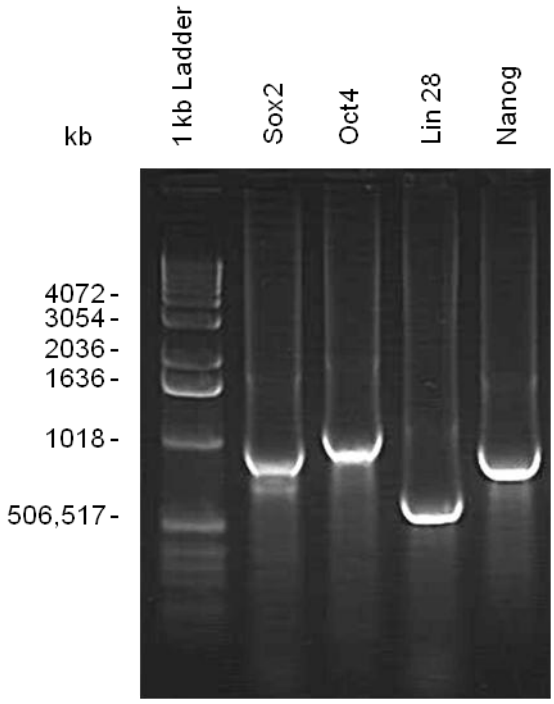


Figure 3.2: PCR amplification of iPSC genes.

Sox2, Oct4, Lin28 and Nanog coding sequences were PCR amplified from each of parent vectors, pSin-E2F-iPSC-Pur. *AgeI* and *EcoRI* restriction sites were incorporated at the 5' and 3' termini of the amplified fragments, respectively. Fragments of sizes 950 bp, 1,000 bp, 650 bp and 900 bp were produced for Sox2, Oct4, Lin28 and Nanog respectively.

Figure 3.2 shows PCR fragments were amplified comprising 950 bp, 1,000 bp, 650 bp and 900 bp for Sox2, Oct4, Lin28 and Nanog, respectively. This indicated the successful amplification of the full length coding sequences for each gene. These PCR fragments were then sub-cloned into the pre-digested pEGFP-C1 vector via the PCR cloning vector, pCR®-Blunt TOPO II, by *AgeI/EcoRI* restriction digestion. Figure 3.3 shows restriction digestion with *AgeI/EcoRI* to screen for the presence of the iPSC coding sequence inserts in picked colonies. Results are shown for pCMV-Oct4 as an example and similar results were produced for Sox2, Lin28 and Nanog (data not shown).

Inserts of 1 kb were produced after *AgeI/EcoRI* restriction digestion, indicating the presence of the Oct4 insert. Sequencing analysis was performed to confirm the integrity of the cloned inserts in successful colonies (data not shown).

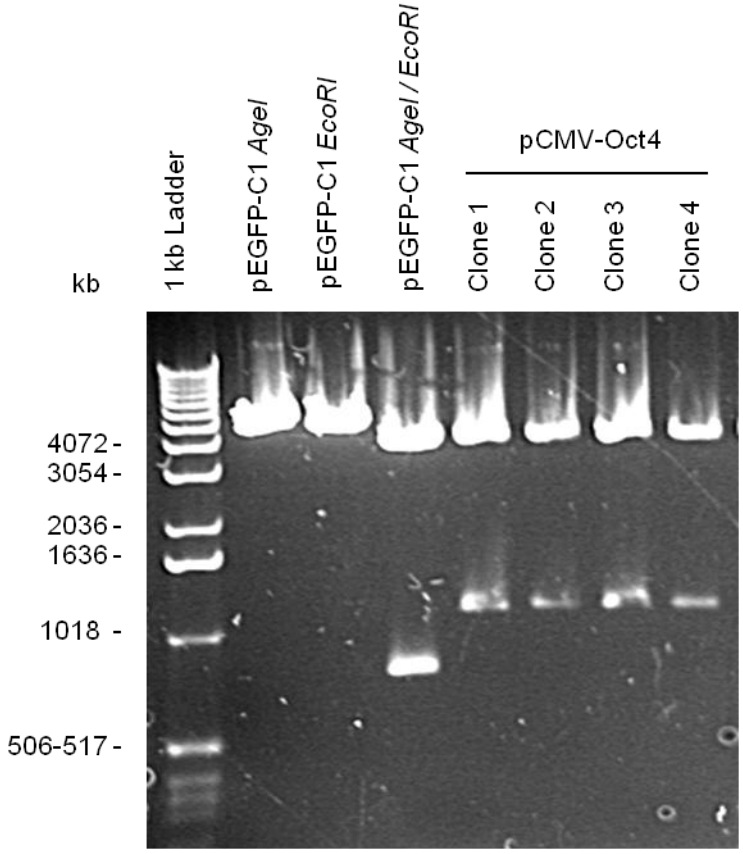


Figure 3.3: *AgeI/EcoRI* restriction digestion of pCMV-Oct4 colonies
DNA was extracted from surviving colonies and screened by *AgeI/EcoRI* restriction digestion. The presence of a 1 kb fragment demonstrates the presence of the Oct4 coding sequence has been successfully cloned into the empty pEGFP-C1 vector. Similar screening was performed for each of the pCMV-iPSC constructs.

3.3 Expression of iPSC transgenes from pCMV-iPSC constructs

Prior to producing recombinant HVS-based vectors, it was first necessary to confirm each pCMV-iPSC expression construct was capable of expressing their relative reprogramming factors. Transgene expression analysis was performed using RT-PCR and western blot analysis.

HEK 293T cells were transfected with the control pEFGP-C1 or each pCMV-iPSC construct. After 24 hours, RNA was isolated and reverse transcribed. Semi-quantitative RT-PCR was then performed and results showed the successful amplification of iPSC gene expression compared to the controls (Figure 3.4). Samples incubated without reverse transcriptase (-ve RT) were used to demonstrate no DNA contamination was present.

Further confirmation that the pCMV-iPSC constructs were able to express their respective transgenes was provided by western blot analysis. Again, 293T cells were transfected with each pCMV-iPSC construct, or the pEGFP-C1 construct as a negative control. Cell lysates were then analysed by immunoblotting with specific antibodies to each iPSC protein. Figure 3.5 shows expression of each iPSC protein, with bands detected at approximately 39 kDa, 34 kDa, 23 kDa, and 35 kDa for Oct4, Sox2, Lin28 and Nanog respectively. These data therefore demonstrate that the pCMV-iPSC constructs are capable of producing full length iPSC proteins.

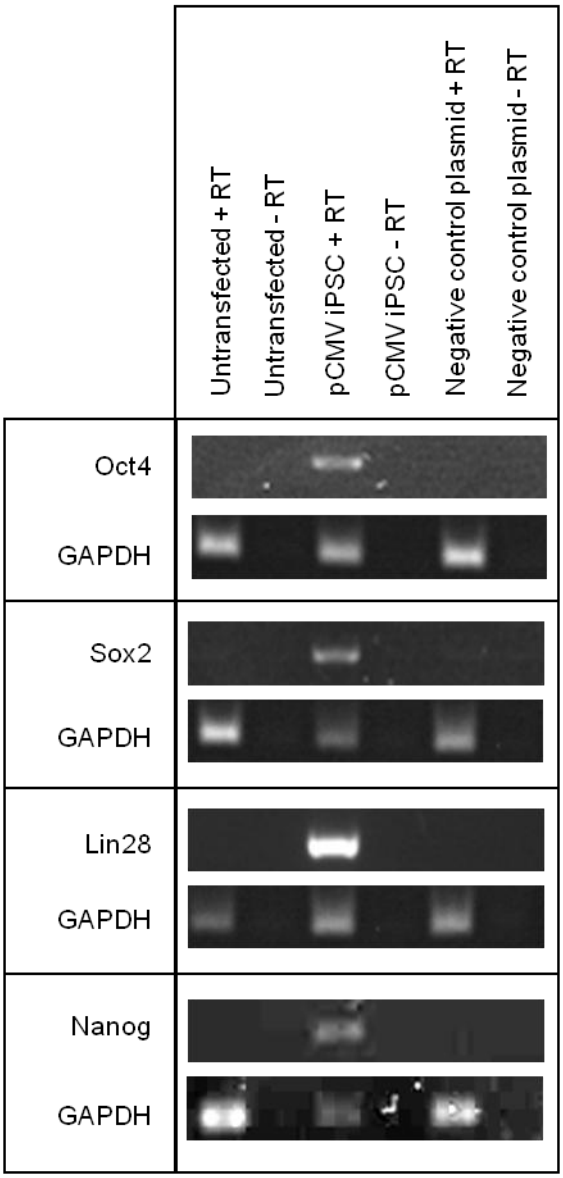


Figure 3.4: RT-PCR Analysis of pCMV-iPSC gene expression in 293T cells

Total RNA from transfected cells was extracted and reverse transcribed with Oligo(dT) primers. PCR was performed on the cDNA produced using primers specific for each iPSC transgene. Cells transfected with each pCMV-iPSC construct express their respective transgenes. Samples incubated without Reverse Transcriptase (-ve RT) controls were used to demonstrate the absence of DNA contamination.

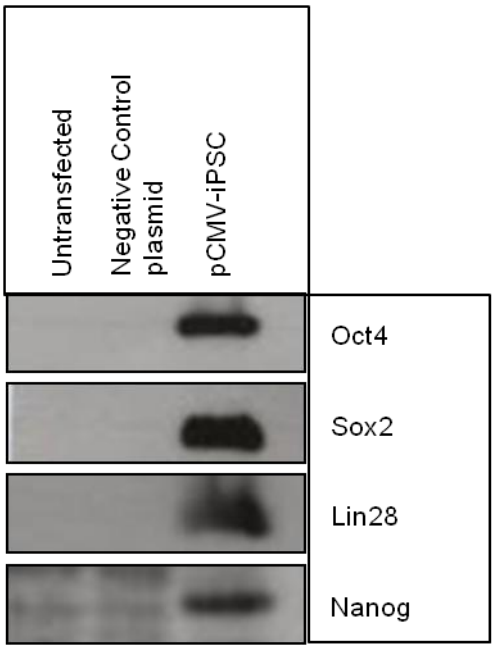


Figure 3.5: Western Blot Analysis of pCMV-iPSC expression in 293T cells
Whole cell lysates were extracted from transfected 293T cells and separated by SDS-PAGE followed by western blotting. Antibodies specific for each iPSC gene were used to detect protein expression. Expression of each iPSC gene is detected in transfected 293T cells, and is absent in cells transfected with the negative control plasmid (pEGFP-C1).

3.4 Stage 3: pShuttle-iPSC construct generation

Cloning of the CMV-iPSC expression cassettes into the HVS-BAC genome firstly required each iPSC factor to be sub-cloned into the shuttle vector, pShuttle Link1 (White *et al.*, 2003). This construct contains a multiple cloning site and kanamycin resistance gene flanked by two *I-PpoI* sites. *I-PpoI* is an extremely infrequent restriction site, with no naturally occurring restriction sites in the HVS genome. Firstly, the CMV-iPSC cassette, incorporating the CMV promoter, iPSC coding region and SV40 poly adenylation signal were PCR amplified to incorporate *NotI* and *MluI* restriction sites at the 5' and 3' ends of the expression cassette, respectively. The PCR fragments generated were 2 kb for CMV-Oct4, 1.85 kb for CMV-Sox2, 1.5 kb for CMV-Lin28 and 1.8 kb for CMV-Nanog (Figure 3.6). These sizes indicate successful amplification of each transgene expression cassette, with the 500 bp CMV promoter and the 50 bp SV40 poly adenylation signal. Each amplified PCR fragment was then cloned into pCR®-Blunt TOPO II prior to cloning into pShuttle Link1 via *NotI/MluI* restriction digest. Colonies were

then screened for the presence of the CMV-iPSC inserts by *NotI/MluI* restriction digestion. As an example, Figure 3.7 shows pShuttle-Oct4 colonies digested with *NotI/MluI* with the 2 kb CMV-Oct4 insert in all colonies except colony 6. Sequencing analysis was performed on successful clones to confirm the integrity of the cloned insert (data not shown). Similar results were observed for Sox2, Lin28 and Nanog (data not shown).

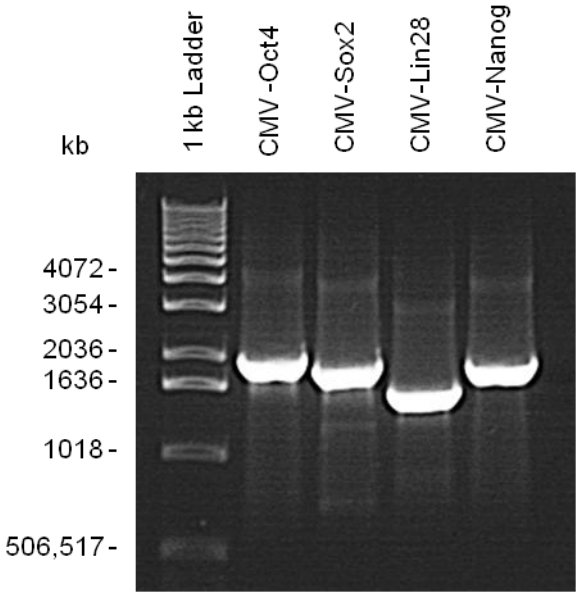


Figure 3.6: PCR amplification of the CMV-iPSC expression cassettes

The iPSC expression cassettes were PCR amplified using 1 µg of each pCMV-iPSC template. Clean PCR fragments were produced of sizes 2 kb, 1.85 kb, 1.5 kb and 1.8 kb for the CMV-Oct4, CMV-Sox2, CMV-Lin28 and CMV-Nanog expression cassettes, respectively.

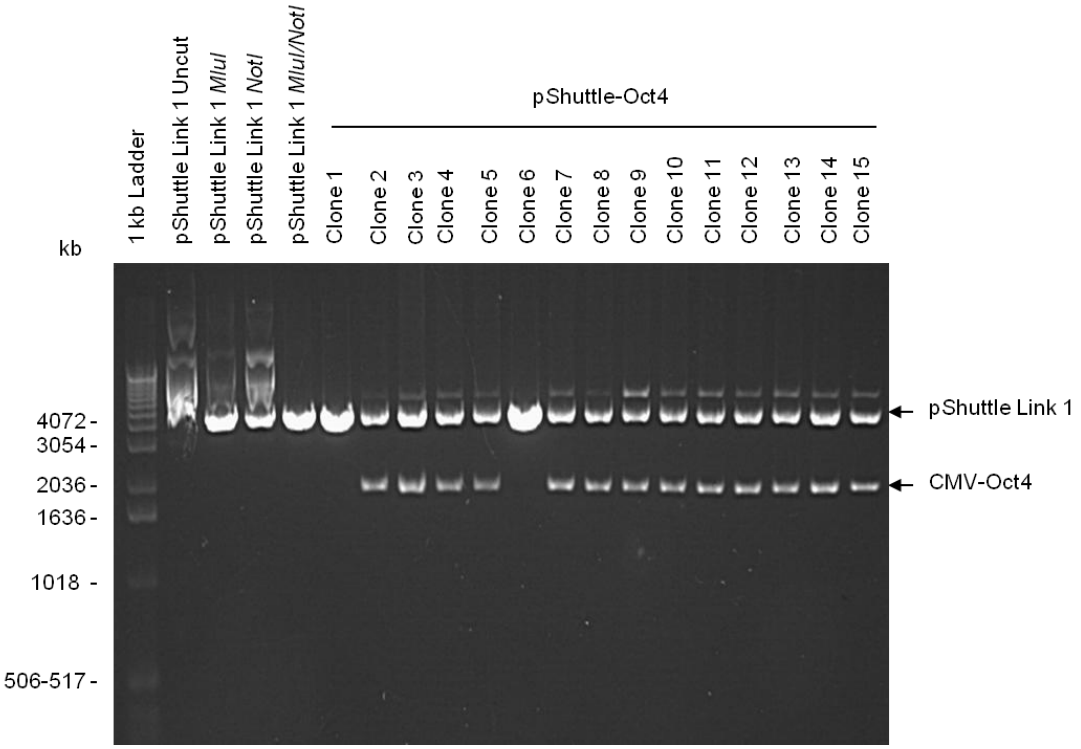


Figure 3.7: *NotI/MluI* restriction digestion of pShuttle-iPSC colonies

DNA was extracted from picked colonies and screened for correct inserts by *NotI/MluI* restriction digestion. Successful colonies were identified by presence of the CMV-Oct4 insert, shown by a 2 kb band.

3.5 Stage 4: Generation of HVS-BAC-iPSC recombinant virus

To insert the iPSC expression cassettes into the HVS genome, the HVS-BAC system was utilised (White *et al.*, 2003). The HVS-BAC has been engineered to contain a single *I-PpoI* restriction site, which enables simple and fast cloning of desired genes into the HVS genome. The presence of a chloramphenicol resistance marker allows for easy selection of successful clones and either a GFP or RFP mammalian expression cassette enables quick identification of cells which have successfully been infected with HVS-BAC recombinant virus.

To insert the iPSC expression construct, the HVS-BAC-GFP was linearised with *I-PpoI* followed by drop dialysis to purify the BAC DNA. pShuttle-iPSC constructs were also digested with *I-PpoI* which produced a fragment containing the expression cassette containing the iPSC gene of interest along with the kanamycin resistance marker of the pShuttle-Link 1 plasmid.

This kanamycin resistance gene allows selection of successful HVS-BAC-iPSC colonies under dual selection with kanamycin and chloramphenicol. DNA was extracted from successful colonies by low-copy number mini-preparations prior to screening by *I-PpoI* restriction digest and Pulsed-Field Gel Electrophoresis for the presence of the CMV-iPSC insert. Figure 3.8 shows mini-prepped analysis of putative HVS-Oct4 recombinants, with similar results obtained for Sox2, Lin28 and Nanog (data not shown). This analysis yielded fragments of sizes 3.3 kb, 3.15 kb, 2.85 kb and 3.1 kb for HVS-Oct4, HVS-Sox2, HVS-Lin28 and HVS-Nanog, respectively. These fragment sizes correspond to the size of the CMV-iPSC fragments and the remaining pShuttle Link 1 backbone containing the kanamycin resistance gene and indicate the successful cloning of each CMV-iPSC cassette into the HVS-BAC-GFP genome. The integrity of each insert was confirmed by sequencing analysis (data not shown).

After confirming the integrity of the cloned inserts, each HVS-BAC-iPSC construct was isolated by low-copy number maxi-preparations of the DNA, and their purity confirmed by Pulsed-Field Gel Electrophoresis (Figure 3.9).

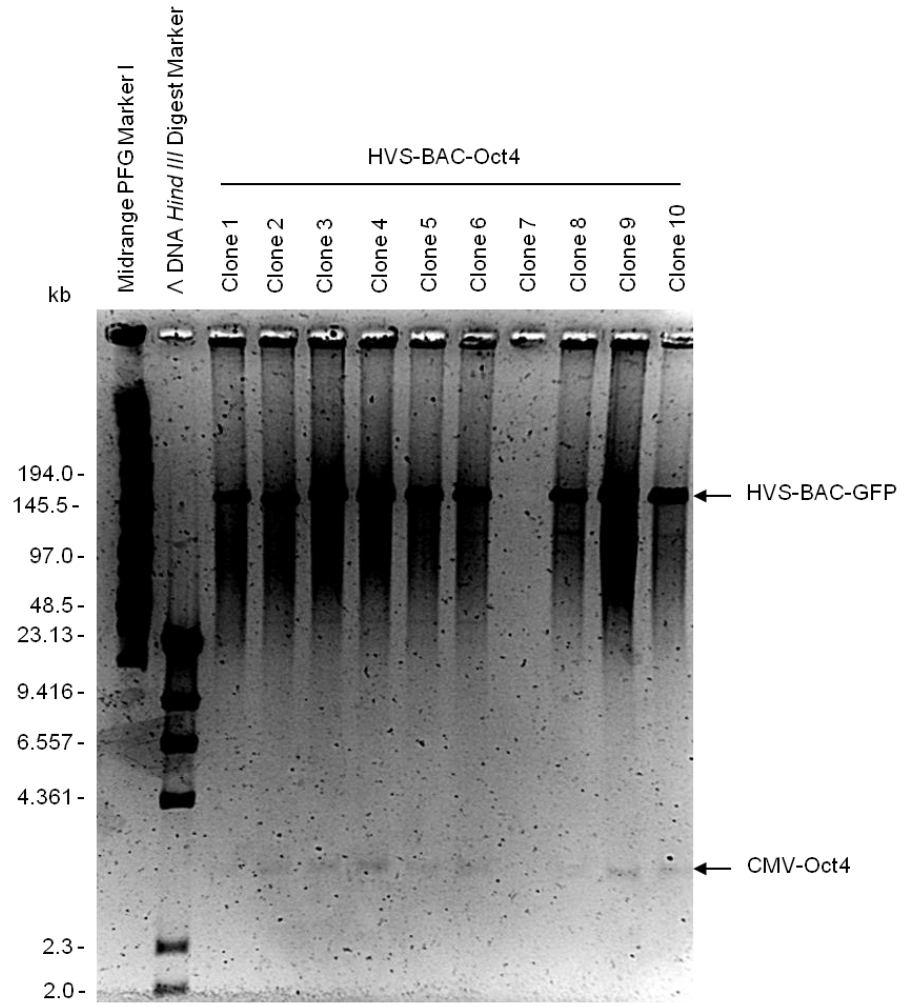


Figure 3.8: Pulsed-Field Gel Electrophoresis of HVS-BAC-Oct colony screening

DNA was extracted from colonies grown under kanamycin and chloramphenicol selection by low-copy number mini-preps. *I-PpoI* restriction digests were performed to remove the CMV-Oct4 insert from the HVS-BAC genome, thus enabling identification successful colonies. All colonies contained a 3 kb insert indicating the presence of the cloned CMV-Oct4 expression cassette.

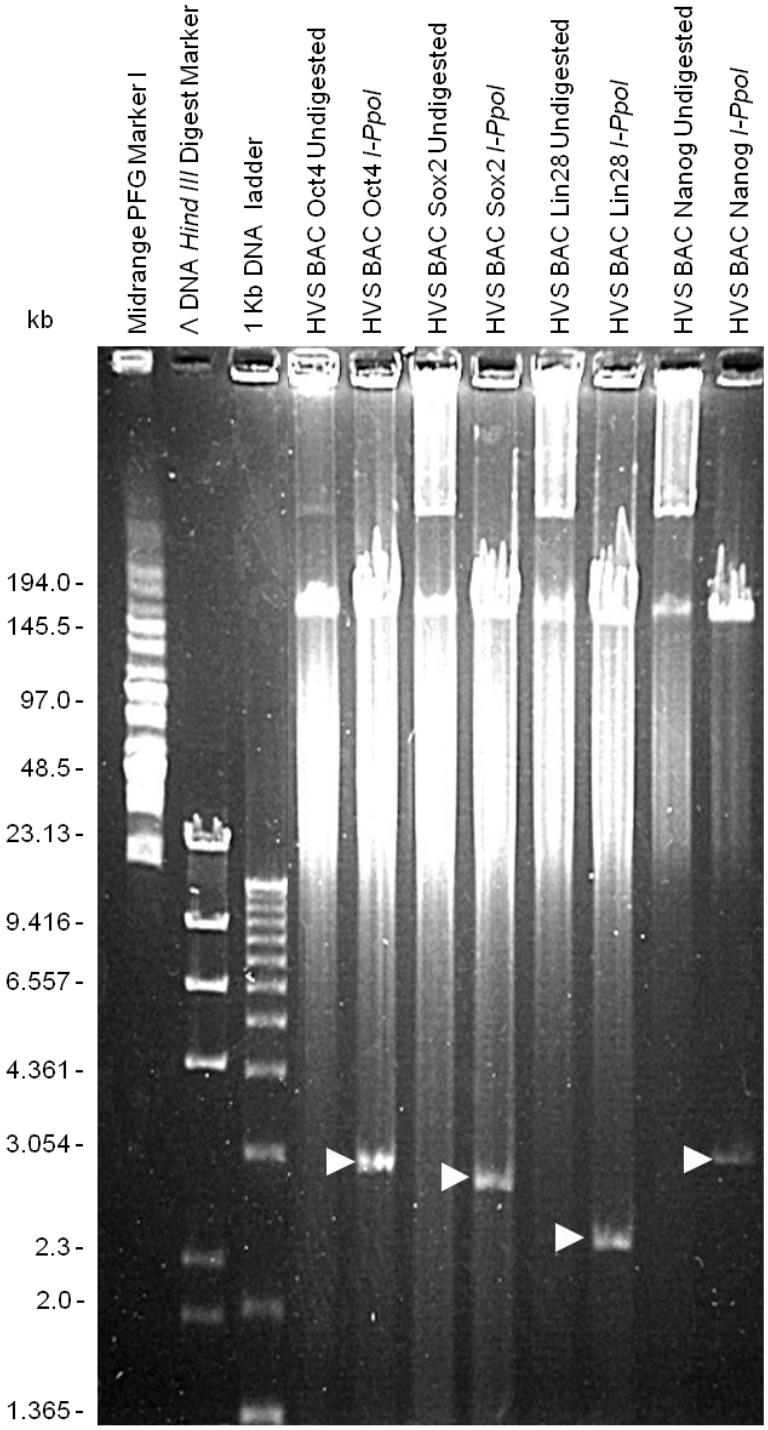


Figure 3.9: Pulse-Field Gel Electrophoresis of the HVS-BAC-iPSC constructs

Low-copy number maxi-preparations of DNA were performed to purify each HVS-BAC-iPSC construct. Each construct was digested with *I-PpoI* and analysed by PFG electrophoresis to demonstrate the presence of each cloned CMV-iPSC insert and to verify the purity and integrity of each HVS-BAC-iPSC construct. The HVS-BAC can be seen as a 175 kb band, while the CMV-iPSC inserts are seen as bands of sizes 3.3 kb, 3.15 kb, 2.85 kb and 3.1 kb for CMV-Oct4, CMV-Sox2, CMV-Lin28 and CMV-Nanog respectively.

3.6 Generation of recombinant HVS vectors expressing iPSC transgenes

Infectious HVS-iPSC virus particles were produced by transfecting each HVS-BAC-iPSC construct into the permissive Owl Monkey Kidney (OMK) cell line and culturing in DMEM with 5% foetal calf serum (FCS). Two weeks post-transfection, plaques started to form in the OMK cell sheet with classic herpesvirus comet-tail appearance. The presence of the virus in these plaques was confirmed by the expression of the GFP marker gene in the infected cells surrounding the plaque (Figure 3.10). After 3 weeks, the cell sheet was completely lysed and viral supernatant could then be used to infect more OMK cells to produce higher titre virus stocks by standard protocols.

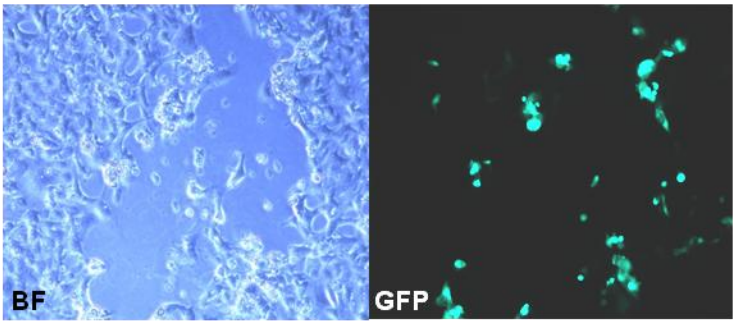


Figure 3.10: Generation of the HVS-Oct4 virus stocks in Owl Monkey Kidney Cells
OMK cells were transfected with each HVS-BAC-iPSC construct (HVS-BAC-Oct4 is shown here) and cells were cultured for 2-3 week weeks in DMEM 5% foetal bovine serum. GFP positive plaques were observed after 2 weeks in culture indicating the production of infectious HVS-iPSC particles. Image was taken at 50x magnification.

3.7 Expression of iPSC transgenes in the HVS-iPSC viral vectors

Once high titre viral stocks had been produced (1x 10^7 GFP units/ml or higher) it was necessary to assess if these viral vectors were capable of transgene delivery and expression. To assess transgene expression in infected cells viral supernatants were used to infect 1x10^6 293T cells with an MOI of 1, and infected cells were harvested 48 hours post-infection. Cell lysates were then analysed by immunoblotting with antibodies specific for each transgene. Figure 3.11 shows expression of the transgenes demonstrated by the presence of bands of sizes 39, 23 and 34 kDa for Oct4, Lin28 and Nanog, respectively. However, no band was detected for Sox2, indicating the HVS-Sox2 virus was not expressing despite the confirmation of the integrity of the Sox2 insert and the CMV promoter and poly adenylation sequences of the expression cassette. Attempts to re-clone the HVS-Sox2 construct were unsuccessful. Therefore, cell lines that endogenously expressed Sox2 were selected for reprogramming efforts, as it had previously been demonstrated that these cells were capable of being reprogrammed without exogenous Sox2 (Kim *et al.*, 2009b).

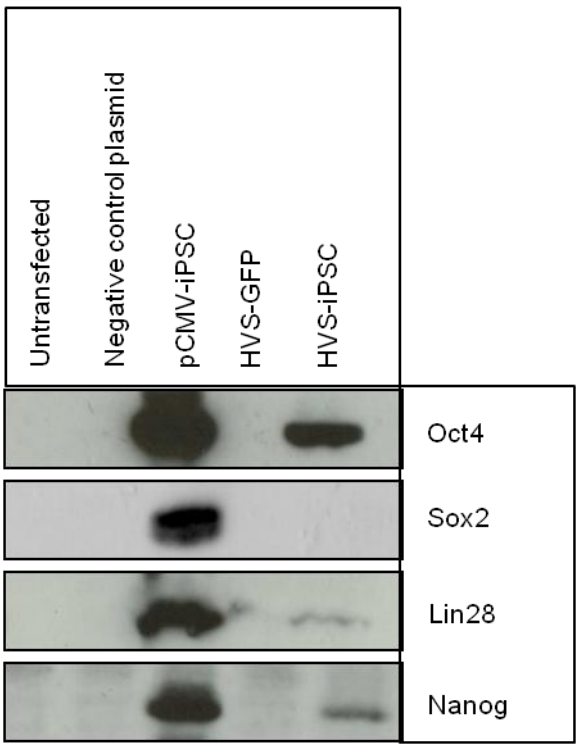


Figure 3.11: Expression of HVS-iPSC viruses in infected 293T cells
293T cells were infected with HVS viral stocks and total cell lysate extracted by RIPA lysis 48 hours post-infection. Cell lysates were analysed by western blot following separation on SDS-PAGE. Expression is demonstrated for HVS-Oct4, HVS-Lin28 and HVS-Nanog.

3.8 Discussion

In this chapter, HVS-based vectors were developed for the delivery of the reprogramming factors, Oct4, Sox2, Lin28 and Nanog. Initially, these genes were placed under the control of the CMV promoter which was chosen as it provides strong expression of transgenes in a wide variety of mammalian cell types. Furthermore, it is already commonly used in iPSC generation, indicating its suitability for vectors designed for iPSC technology and has previously been shown to be active in HVS-based vectors in a differentiation model (Stevenson *et al.*, 2000a).

Although the CMV promoter is eventually silenced, this is a useful trait in iPSC vectors as it has been demonstrated that removal of iPSC gene expression after successful reprogramming produced colonies which more closely resemble ESCs in their gene expression profiles and differentiation potential (Kim *et al.*, 2009c; Soldner *et al.*, 2009). Therefore, the CMV promoter seems an appropriate choice for these proof-of-principle investigations. Future work would ideally utilise tetracycline inducible expression systems (Forster *et al.*, 1999; Gossen & Bujard, 1992a) to provide direct, tight controlled expression of the reprogramming transgenes.

The potential of HVS as a gene delivery system for iPSC generation has already been highlighted by Stevenson et al. (Stevenson *et al.*, 2000a) wherein mESCs were shown to be readily infected with HVS carrying a GFP reporter gene. Moreover, these HVS vectors were capable of maintaining GFP expression throughout differentiation of the mES cells into macrophages, whereas transgene silencing is commonly observed in retroviral vectors upon differentiation (Hotta & Ellis, 2008; Laker *et al.*, 1998). Therefore, if HVS is capable of maintaining transgene expression through differentiation, the reverse could be applicable. This will enable more reliable determination of the optimal stoichiometry of reprogramming factors as the dynamic pattern of positional silencing observed for retroviral vectors (Ramunas *et al.*, 2007) would not be affected using a HVS-based episomally maintained vector. Dynamic positional silencing causes varying expression levels of the reprogramming transgenes depending on the integration site of the retrovirus in the genome. This therefore alters the stoichiometry of the

various reprogramming factors. (Hotta & Ellis, 2008) Furthermore, the ability of HVS to persist as a non-integrating episome enables this virus to provide prolonged transgene expression without disrupting normal gene expression. This makes HVS a safer vector for reprogramming cells and will ensure the iPSC and iPC cells produced have less disrupted gene expression profiles.

This chapter details a suitable cloning technique for the quick and efficient generation of HVS-iPSC vectors using the HVS-BAC-GFP genome. The HVS-BAC contains a single *I-PpoI* restriction site and bacterial replication elements which enables fast and efficient manipulation of the HVS genome to incorporate the reprogramming transgenes. Furthermore, the HVS-BAC also contains a hygromycin resistance gene and a GFP reporter gene. The GFP reporter gene allows visual identification of cells carrying the HVS episome, and the hygromycin resistance gene allows selection of infected cells without the need for cell sorting. Furthermore, the hygromycin resistance gene will ensure that cells grown under prolonged hygromycin selection will not lose the HVS episomes after repeated cell divisions.

Figures 3.4, 3.5 and 3.11 demonstrate that each pCMV-iPSC construct and HVS-iPSC virus highly express their respective transgenes, with the exception of the HVS-Sox2 virus. This could present an issue in future iPSC reprogramming studies as Oct4 and Sox2 are the key regulators of reprogramming to pluripotency and are capable of regulating their own expression levels as well as many other pluripotency associated genes, including Nanog (Masui *et al.*, 2007; Okumura-Nakanishi *et al.*, 2005; Ovitt & Schöler, 1998). Human fibroblasts have even been successfully reprogrammed by only using Oct4 and Sox2 (Baron *et al.*, 1997). However, Kim et al demonstrated that Oct4 alone is capable of reprogramming neuronal stem cells as these cells endogenously express Sox2 (Episkopou, 2005; Kim *et al.*, 2009b). Therefore, this issue can be circumvented by selecting cell lines that endogenously express Sox2 for reprogramming attempts.

In summary, HVS possesses many attractive qualities as a vector for iPSC generation. As demonstrated in this chapter, a cloning strategy has been devised which utilises the HVS-BAC (White *et al.*, 2003) to enable quick and

efficient generation of HVS-iPSC viruses. The efficacy of this cloning procedure is demonstrated by the production of HVS-iPSC constructs that are capable of delivering three out of the four reprogramming transgenes.

Chapter 4
Generation of iPCs from the Ewings Sarcoma A673 Cell Line

4 Generation of Induced iPCs from the Ewings Sarcoma A673 Cell Line

4.1 Introduction

iPSC technology potentially holds great promise for stem-cell based therapies. This has been demonstrated in pre-clinical studies for a variety of diseases, including neurodegenerative diseases. Specifically, iPSCs have been used in several pre-clinical trials for cell-based replacement therapies to treat Parkinson's disease (Hargus *et al.*, 2010; Wernig *et al.*, 2008), strokes (Chen *et al.*, 2010), and spinal chord injuries (Tsuji *et al.*, 2010).

Furthermore, aside from their therapeutic potential, iPSCs provide an excellent model for the progression of neurodegenerative diseases. Currently, most studies of neurodegenerative diseases are performed on tissue from dead patients or animal models as the availability of live brain tissue from humans is limited. These models are less than ideal, as tissues from dead patients only represent the end stages of the disease and animal models cannot compensate for variances caused by species, age, patient specificity and the complexities of the human brain.

There have already been a number of reports detailing the generation of patient specific iPSCs from patients suffering from a variety of genetic neurodegenerative diseases including Parkinson's disease (Soldner *et al.*, 2009), Huntington's disease (Park *et al.*, 2008), Alzheimer's disease (Israel *et al.*, 2012) and Spinal Muscular Atrophy (Ebert *et al.*, 2009) to name but a few. These disease-specific derived iPSCs enable further detailed study of the development and progression of these diseases, as well as providing appropriate models for drug-screens.

Another interesting application of iPSC technology for disease modelling includes reprogramming somatic cancer cells to induced pluripotent cancer stem cells (iPCs). This has already been achieved for chronic myeloid leukaemia (Hu *et al.*, 2011), colorectal cancer (Miyoshi *et al.*, 2010), and hepatocellular carcinoma (Miyoshi *et al.*, 2010). Cancer Stem Cells (CSCs) are defined by their ability to seed tumours in animal hosts, to self renew and to produce differentiated progeny. They were originally identified in

leukaemia (Lapidot *et al.*, 1994) and their presence has also been demonstrated in many solid tumours including breast, brain, pancreatic, colon, prostate, and various others (Al-Hajj *et al.*, 2003; Collins *et al.*, 2005; Li *et al.*, 2007; O'Brien *et al.*, 2007; Ricci-Vitiani *et al.*, 2007; Singh *et al.*, 2004).These cells are more resistant to chemotherapy than other cells within the tumour (Dean *et al.*, 2005). As a result, chemotherapy treatments are believed to remove the bulk of tumours, but leave the CSCs free to migrate into the lymph system. This, combined with their tumour seeding capabilities, has implicated cancer stem cells as having roles in metastasis and secondary cancers.

Despite their important role in cancer progression and metastasis, little is known about CSCs due to the difficulties in their identification and isolation. Therefore, utilising iPSC technology to reprogram cancer cells into these rare cancer stem cells would provide a source of easily obtainable, pure cells to use for further studies.

This chapter will focus on investigating the potential of the HVS-iPSC virus gene delivery vectors for reprogramming both normal somatic neural stem cells and somatic cancer cells to a stem cell-like state.

4.2 Neural Stem Cells

4.2.1 Isolation of primary mouse Neural Stem Cells

iPSC technology was initially designed to reprogramme primary somatic cells to an embryonic stem-cell like state; therefore the HVS-based vectors generated would ideally be tested for their ability to reprogramme primary somatic cells, such as fibroblasts. However, as discussed in Chapter 3, the HVS-Sox2 virus was incapable of expressing the Sox2 transgene which made reprogramming these cells unfeasible. Alternatively however, Kim et al. demonstrated it was possible to reprogramme Neural Stem Cells (NSCs) to iPSC cells without the need for exogenous Sox2 expression (Kim *et al.*, 2009b) as these cells endogenously express Sox2 (Episkopou, 2005). Therefore initial attempts were made to reprogramme Neural Stem Cells (NSCs) using the HVS-based vectors.

In collaboration with Dr. Ian Wood (Faculty of Biological Sciences), murine brains were harvested and cultured for four weeks according to the protocol described by Kim et al. (Kim *et al.*, 2009c). Attempts were then made to isolate live neurospheres from other neural cells by harvesting the supernatant every 7 days and centrifuging at 100 rcf to pellet the neurospheres. The supernatant was then discarded and the neurospheres gently resuspended in fresh NSC medium (detailed in (Kim *et al.*, 2009c)). Pictures were taken of the cells every 2-4 days (Figure 4.1).

As Figure 4.1 demonstrates, the repeated purification stages produced fewer and fewer floating spheroids which did not increase in size. As such it was concluded that the cells were dying under prolonged culture and the floating spheroids were either dead cells or cellular debris. Therefore an alternative source of NSCs was kindly provided by Dr Ian Wood (University of Leeds) (see section 2.4.1).

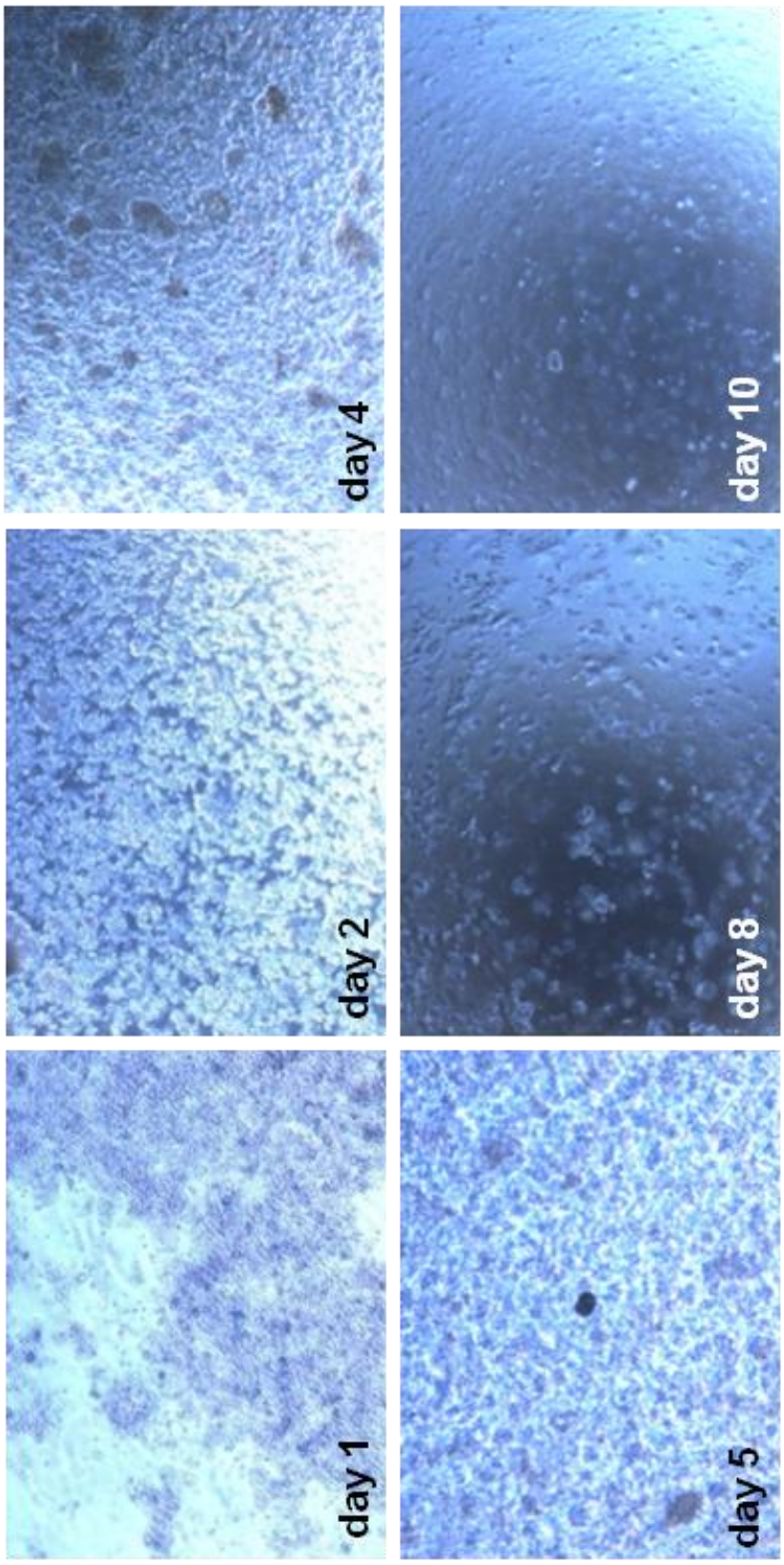


Figure 4.1: Isolation of NSCs from mouse brains

Whole murine brains were extracted and trypsinised before plating on 10 cm dishes in NSC medium. Over time extensive cell death was seen. Some spheroidal clumps of cells were observed, but these were poorly defined and resembled aggregated dead cells, and were unlike the regular, tightly formed morphology of neurospheres. All images were taken at 10 x magnification.

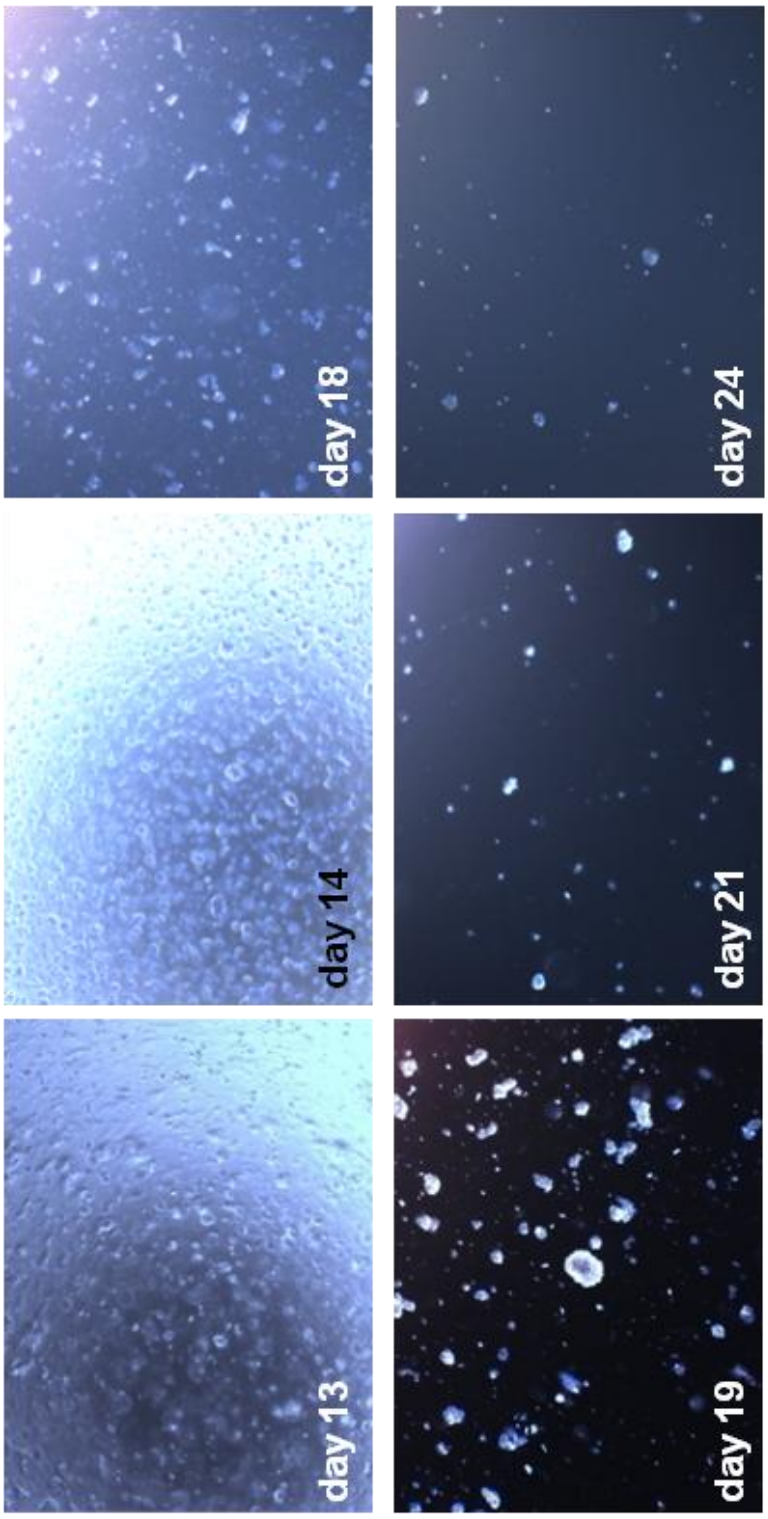


Figure 4.1 (continued): Isolation of NSCs from mouse brains.

4.2.2 Characterisation of HVS Transduction in NSCs

Although HVS has previously been demonstrated to have a naturally broad tropism (Simmer *et al.*, 1991; Smith *et al.*, 2005; Stevenson *et al.*, 1999), to date, the transduction rate of NSCs with HVS has not been assessed. Therefore it was first necessary to ascertain how efficiently these cells could be infected with HVS. To this end, NSCs were infected with varying MOIs of HVS-GFP and the percentage of GFP expressing cells analysed using flow cytometry 24 hours post-infection (Figure 4.2a).

Results indicated that HVS-GFP was able to infect NSCs with a high efficiency; with 70% of cells transduced using HVS-GFP at an MOI of 0.5, and 85% of cells transduced using an MOI of 1. These rates of transduction were higher than those observed for the permissive cell line, Owl Monkey Kidney (OMK) cells which demonstrated transduction rates of 70% using an MOI of 1.

However, as Figure 4.2b demonstrates, HVS-GFP infection appeared to be toxic to the cells. An increase in cytotoxicity and resulting cell death was observed with increased MOI. Due to the cytotoxic effects observed and time constraints, it was decided to focus reprogramming efforts on generating iPCs from somatic cancer cells.

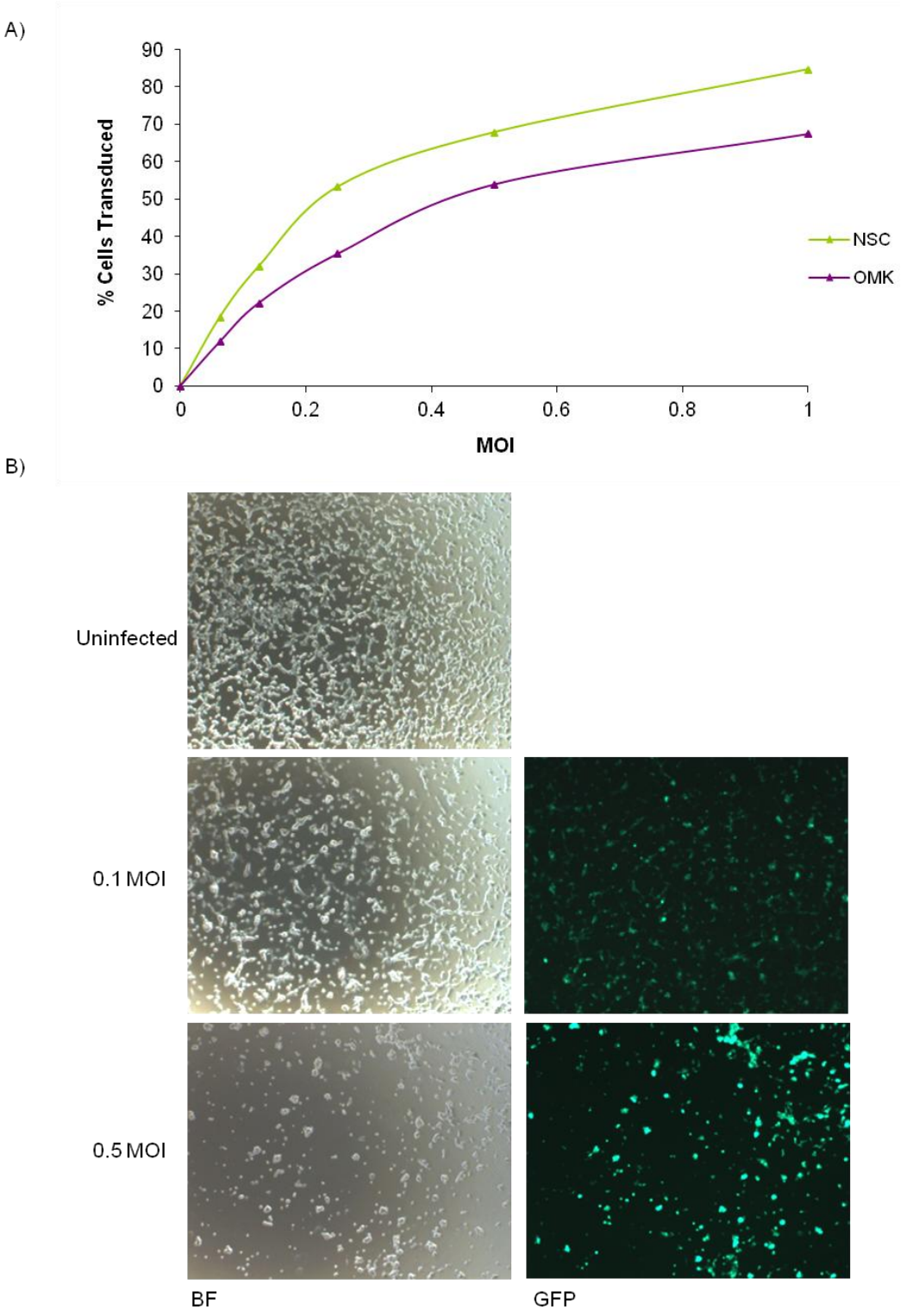


Figure 4.2: Transduction rate of Neural Stem Cells

a) NSCs displayed good transduction rates, with 85% of the cells transduced upon infection with an MOI of 1. In comparison, the OMK permissive cell line demonstrated lower transduction rates, with 70% GFP positive cells. b) Extensive cytotoxicity and cell death were observed in NSCs upon transduction with HVS-GFP. All images were taken at 20 x magnification.

4.3 Characterisation of Three Ewings Sarcoma Family Tumour Cell Lines

The second focus of this chapter was to utilise the HVS-iPSC vectors to reprogramme somatic cancer cells to tumour-initiating stem cell-like cells. Once again, cells that endogenously expressed Sox2 were required for reprogramming efforts due to the lack of a functional HVS-Sox2 delivery vector. Ewings Sarcoma Family Tumour (ESFT) cells were selected as they have previously been shown to endogenously express Sox2 (Riggi *et al.*, 2010). Furthermore, HVS has already been demonstrated to efficiently infect ESFT cell lines (Smith *et al.*, 2004).

ESFTs are the second most frequent malignant bone tumours in adolescents, with a peak incidence between the ages 14 and 20 (Soldner *et al.*, 2009). Prognosis for affected patients is poor with, survival rates of <50% at 5 years and <25% after metastasis has occurred. Recently ESFT cancer stem cells have been identified comprising of a small population of $CD133^+$ cells within the tumour. These CSCs account for between 4%-8% of the ESFT cell population in primary tumours and are capable of forming spheroids *in vitro* and forming new tumours *in vivo* (Suvà *et al.*, 2009). However, due to their low frequency in these tumours and difficulties in their isolation, there is little information on the exact role of ESFT CSCs in cancer formation and progression. Moreover, the high mortality rates associated with ESFT metastasis and the implications of CSC involvement in metastasis, chemotherapy resistance, and radiotherapy resistance make ESFT CSCs of great interest, thereby requiring further study. As such, utilising iPSC technology to reprogramme somatic ESFT cells to induced Cancer Stem Cells (iPCs) would enable the generation of increased numbers of these rare cells, thereby allowing further study into their migration and differentiation potential.

4.3.1 ESFT Cells Endogenously Express Sox2

To assess if the ESFT cell lines endogenously expressed Sox2, quantitative real-time PCR (qPCR) was performed on whole cell RNA extracted from the ESFT cell lines; A673, TC32 and TTC466, and Sox2 levels were analysed in comparison to iPSC cells which were used as a reference standard. Figure 4.3 demonstrates that all three ESFT cell lines endogenously express Sox2, with A673 and TC43 cells expressing 20% of the Sox2 levels observed in the iPSC controls and TTC466 cells showing 40%. Therefore ESFT cells were suitable for reprogramming efforts using the HVS-iPSC viral vectors.

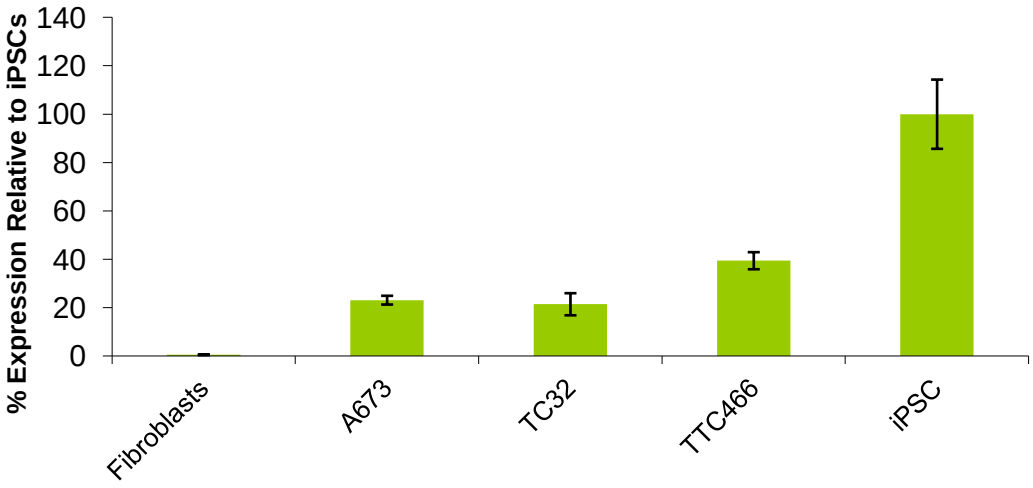


Figure 4.3: Endogenous Expression of Sox2 in the three ESFT cell lines relative to iPSCs
Expression of Sox2 was measured as a percentage of iPSC control cell expression. All three ESFT cell lines expressed Sox2, with A673 and TC32 expressing Sox2 at 20% of the levels observed in iPSCs, and TTC466 cells expressing 40% of iPSC Sox2 expression levels.

4.3.2 Transduction rates of EWS cell lines

HVS has previously been demonstrated to efficiently infect multicellular spheroids formed from two ESFT cell lines, A673 and SK-ES, supporting evidence that HVS can infect ESFT cell lines (Smith *et al.*, 2005). To confirm efficient transduction, a range of ESFT cell lines were infected with varying MOIs of HVS-GFP virus and the percentage of GFP expressing cells were analysed using flow cytometry analysis at 24 hours post-infection.

Figure 4.4a demonstrates that the A673 ESFT cell line exhibited the highest transduction rates, with 87% of the cell population becoming GFP positive upon infection with HVS-GFP at an MOI of 1. This transduction rate was much higher than the other ESFT cell lines, TC32 (32%), and TTC466 (38%), and the permissive cell line, OMK (70%). Due to the higher rates of transduction observed in the A673 cell line, these cells were selected for further reprogramming efforts. Moreover, although a small amount of cell death was observed upon infecting ESFT cells with HVS-GFP at an MOI of 1, the majority of cells appeared to be viable after HVS-GFP infection (Figure 4.4b).

4.4 Reprogramming A673 using HVS-Oct4

Kim et al. demonstrated it was possible to reprogramme somatic cells that endogenously expressed Sox2 using only exogenous Oct4 (Kim *et al.*, 2009b). Therefore, HVS-Oct4 alone was used in initial reprogramming attempts performed on A673 cells. Cells of passage 24 – 30 were used to reduce the variance between cells caused by chromosomal instability and aneuploidy. To this end, $1x10^6$ cells were infected with either mock, control HVS-GFP or HVS-Oct4 at varying MOIs between 0.06-1.0. At 4 days post-infection (p.i.) the cells were transferred to 10 cm dishes to allow them to expand and form colonies.

Results demonstrated that the optimum MOI of HVS-Oct4 was 0.5, as lower MOIs resulted in fewer colonies being formed, whereas some cytotoxicity was observed with an MOI of 1 (data not shown). Figure 4.5 shows representative colonies generated upon infection with HVS-Oct4 at an MOI

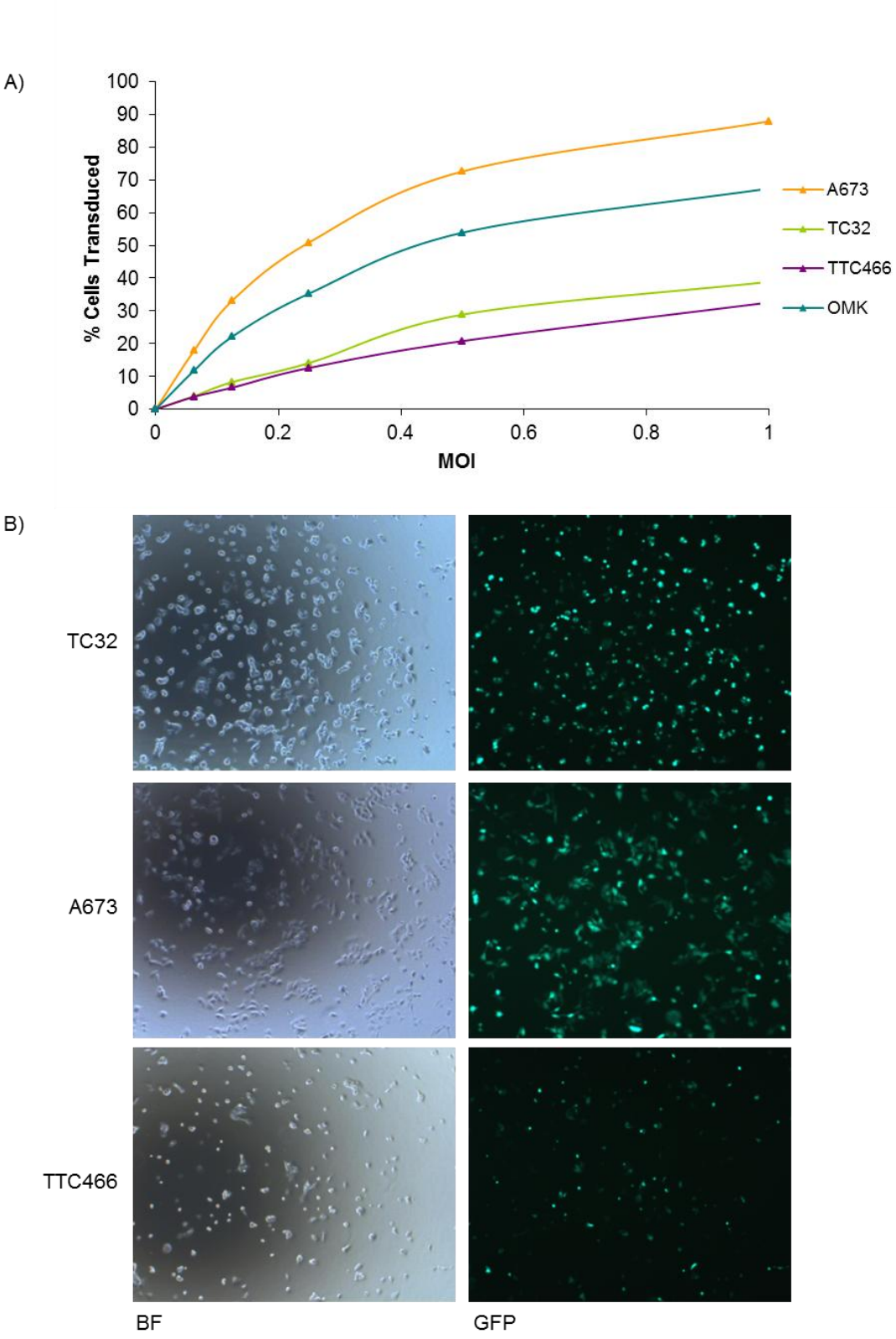


Figure 4.4: Transduction rate of three ESFT cell lines; A673, TC32 and TTC466.

A) TC32 and TTC466 cells displayed poor transduction rates, with a maximum of 32% and 38% of GFP positive cells upon infection with HVS-GFP at an MOI of 1. However, A673 cells displayed excellent transduction rates, with 87% of cells becoming GFP-positive after infecting with the same MOI of HVS-GFP. OMK cells were used as a control, and displayed a 67% transduction rate with an MOI of 1. B) Some cell death was observed in cells infected with HVS-GFP at an MOI of 1, but most of the cells remained viable. All images were taken at 20 x magnification.

of 0.5. Colonies possessing a tightly packed, spheroidal morphology, clearly distinct from the surrounding A673 cells, were observed 18 days p.i. on 10cm Primaria dishes. In contrast, no colonies were formed from mock or HVS-GFP infected cells.

The colonies were then assessed for their ability to grow in conditions which allow feeder cell-independent growth of stem cells. At day 20 p.i., the colonies were transferred to 6 well plates, pre-coated in hESC qualified Matrigel, and cultured in medium consisting of ½ DMEM, ½ mTeSR and 5% serum to allow the cells to adjust to the mTeSR medium. At day 21 p.i. (1 day after transferring to feeder-free stem cell culture conditions), results showed that the colonies were still viable and retained their densely packed spheroidal morphology. However, prolonged culturing in stem cell culture conditions (day 24 and 26 p.i., Figure 4.5) caused the colonies to deform, with cells growing out from the colony, and extensive cell death being observed. Interestingly, the uninfected and HVS-GFP infected control cells also failed to thrive under these culture conditions.

4.5 Second reprogramming attempt using A673 cells

The deformation observed in the colonies arising from HVS-Oct4 infected cells could indicate that HVS-Oct4 alone is not sufficient to reprogramme A673 cells, and that additional iPSC factors are required. Therefore, a second reprogramming attempt was undertaken, whereby A673 cells were infected with HVS-Oct4 virus in combination with either HVS-Lin28 (OL), HVS-Nanog (ON), or all three viruses (OLN). Cells were infected and cultured as described in Figure 4.6 and colonies possessing the same tightly packed spheroidal appearance were once again observed after 18-19 days p.i. (Figure 4.7). Interestingly, some of the colonies generated from the OL combination appeared to lose GFP expression despite being cultured under hygromycin selection, indicating silencing of the reporter GFP gene.

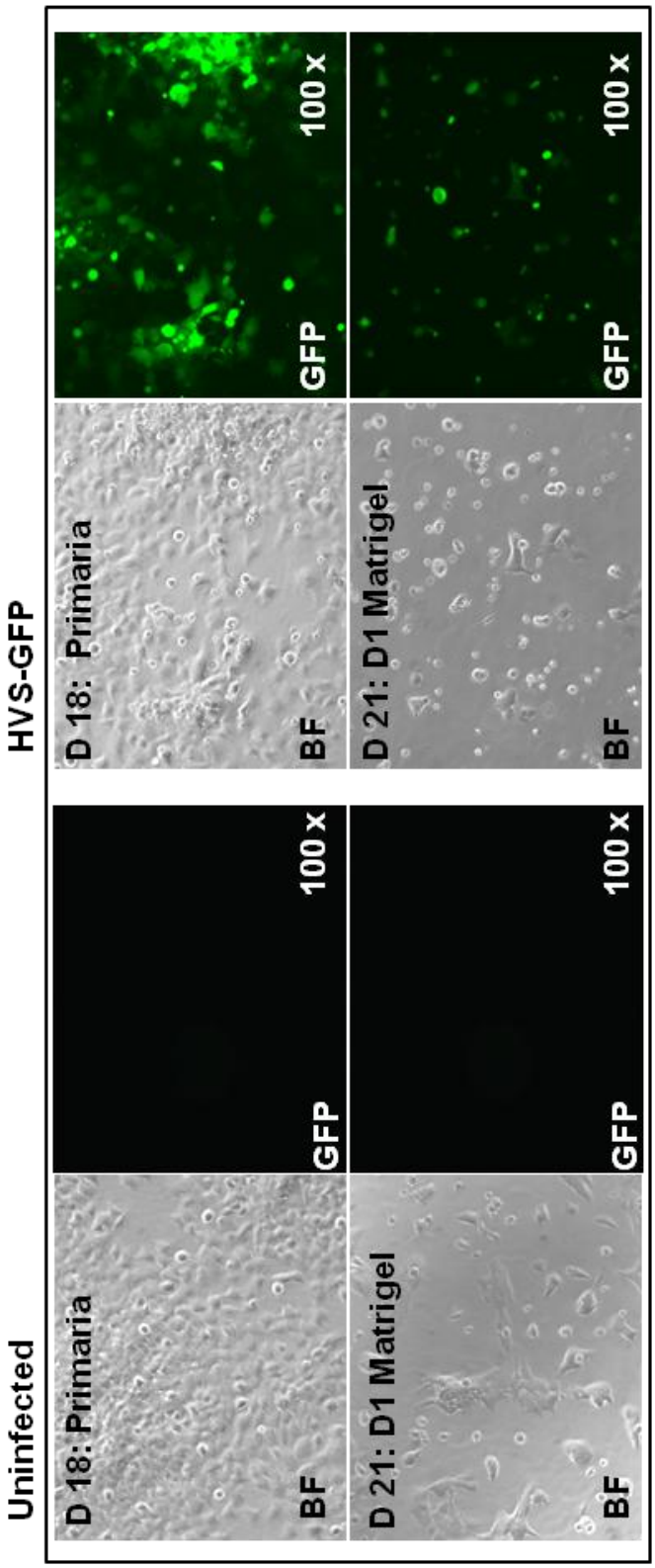


Figure 4.5: Generation of A673-iPC colonies upon transduction with HVS-Oct4.

Colonies were observed on 10 cm Primaria dishes 18 days p.i. These colonies were then transferred to feeder-free stem cell culture conditions after 21 days p.i. However, after prolonged culture in these conditions the colonies deformed and extensive cell death was observed. This cell death was also observed in HVS-GFP and uninfected A673 cultured in these conditions.

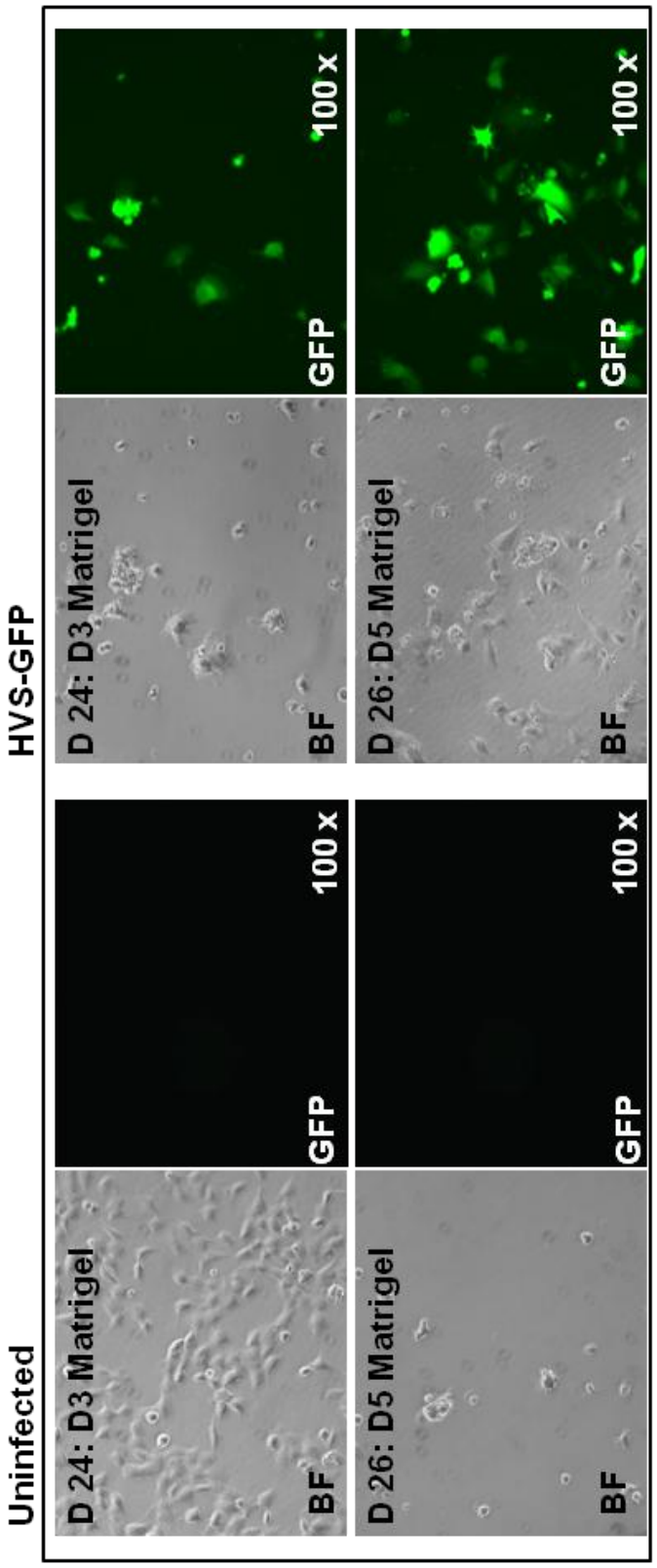


Figure 4.5 (continued): Generation of A673-iPC colonies upon transduction with HVS-Oct4.

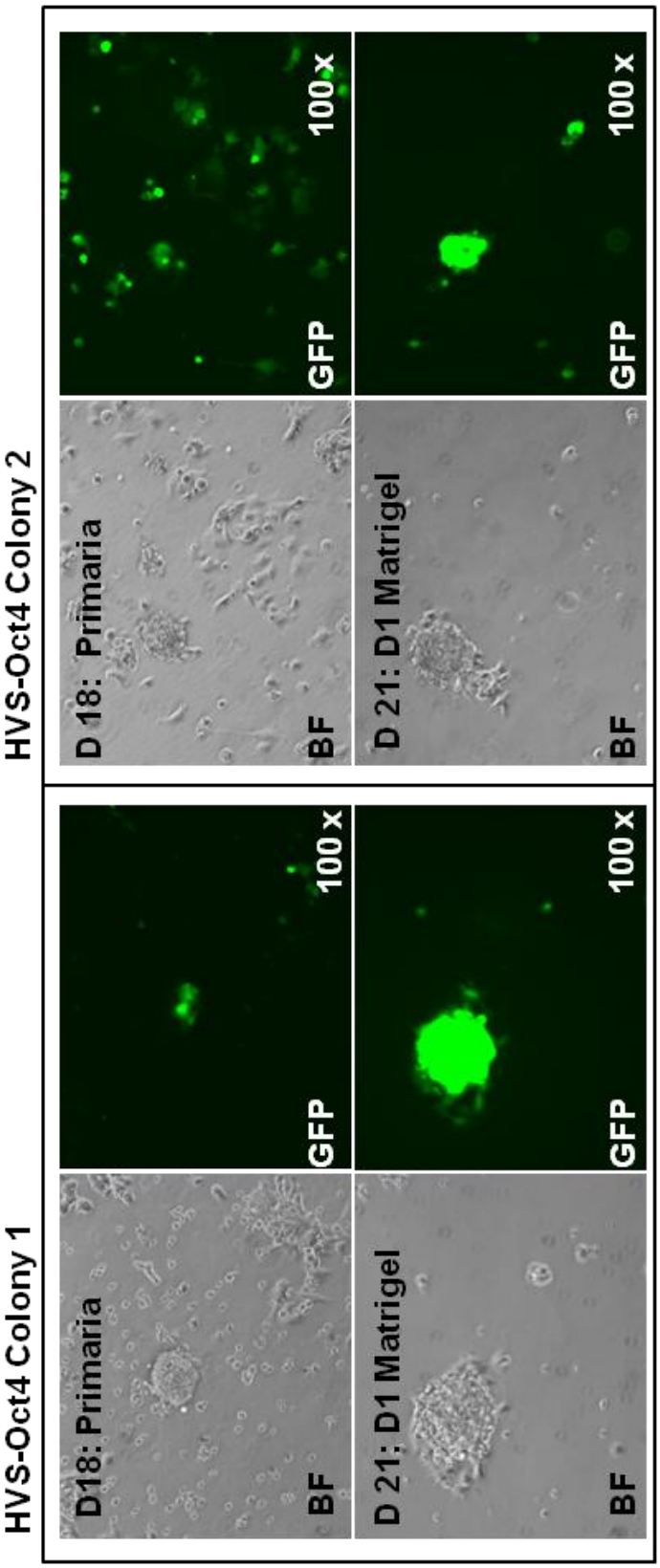


Figure 4.5 (continued): Generation of A673-iPC colonies upon transduction with HVS-Oct4.

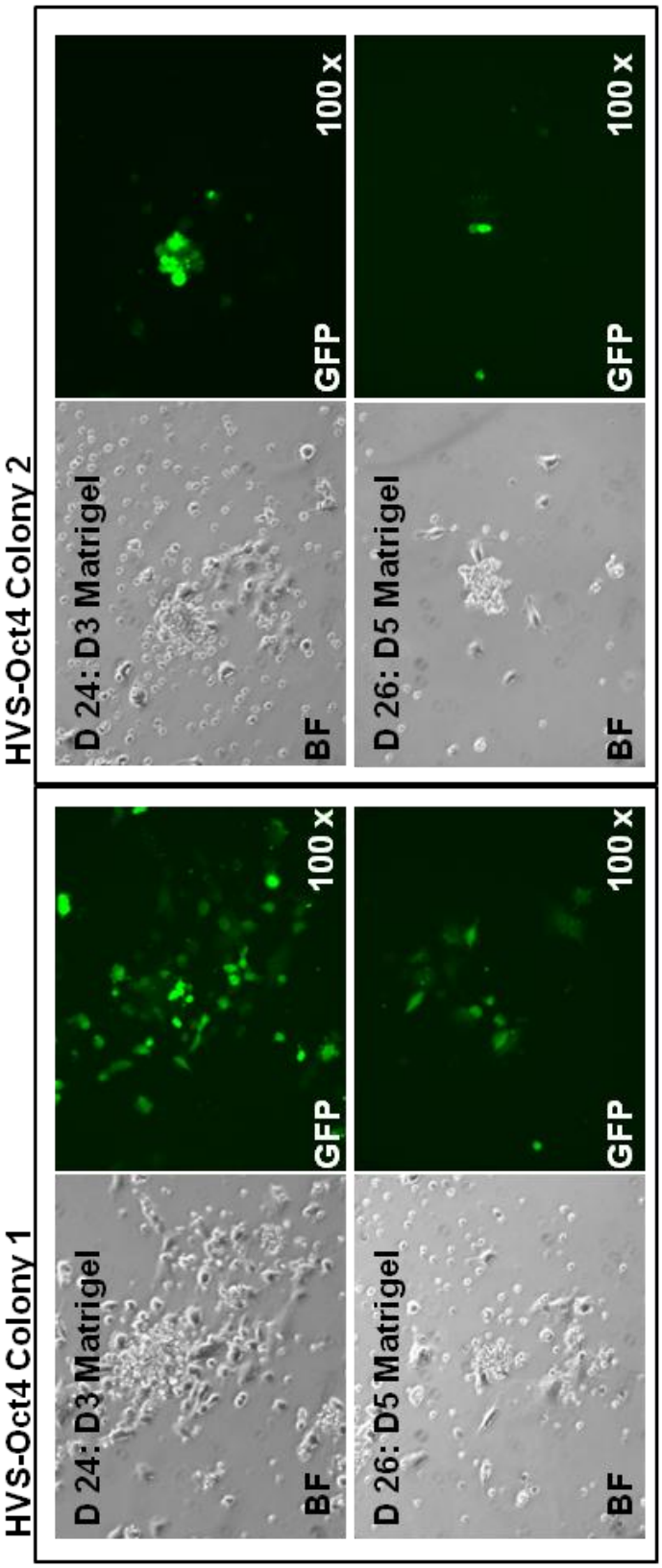


Figure 4.5 (continued): Generation of A673-iPC colonies upon transduction with HVS-Oct4.

As previously, the colonies were transferred onto Matrigel coated 24 well dishes by gently washing off the colonies using DPBS, and cultured in feeder-free stem cell culture conditions. After 1 day in these culture conditions (day 22 p.i.), colonies derived from all three virus combinations were viable, with ON and OLN colonies possessing the most tightly formed colonies (Figure 4.8). The colonies were expanded until they were large enough to be picked onto individual wells of a Matrigel coated 24-well plate (day 24 p.i.). The OL combination produced 6 viable colonies after picking, with ON and OLN producing 3 and 2 colonies respectively. However, the OL colonies possessed poorly defined edges. In contrast, ON and OLN colonies displayed more densely packed spheroidal morphologies. Figure 4.9 shows the colonies 1 day after being picked (day 25 p.i.).

As Figure 4.10 demonstrates, after prolonged culture in feeder-free stem cell culture conditions several colonies from each virus combination condition deformed, and extensive cell death was observed for the cells growing out from these colonies. This is similar to the deformation of colonies generated upon infection with HVS-Oct4 alone (Figure 4.5). However, the remaining viable colonies, after also initially shrinking in size, stabilised after 7-9 days on Matrigel.

After 38 days p.i, 4 colonies remained, which were named Colony 1 (ON), Colony 2 (OL) Colony 3 (OL) and Colony 4 (OLN) (Figure 4.11). Colonies 2, 3 and 4 were viable in feeder-free stem cell culture conditions, and were therefore expanded for further analysis. Colony 1, however, appeared much smaller and possessed poorly defined edges. Furthermore, certain cells within Colony 1 appeared to be blebbing, indicative of apoptosis, and attempts to expand this colony were unsuccessful.

In contrast, Colonies 2, 3 and 4 all possessed densely packed morphologies and phase-bright characteristics, although Colony 3 contained a dark centre. Colonies 2 and 3 (OL) also possessed well defined edges, although the edges of Colony 4 were slightly less well defined, with cells growing out from the central colony mass. Interestingly, both colonies from the OL condition, (Colonies 2 and 3) had significantly reduced GFP expression, which was not observed for Colony 4 (OLN).

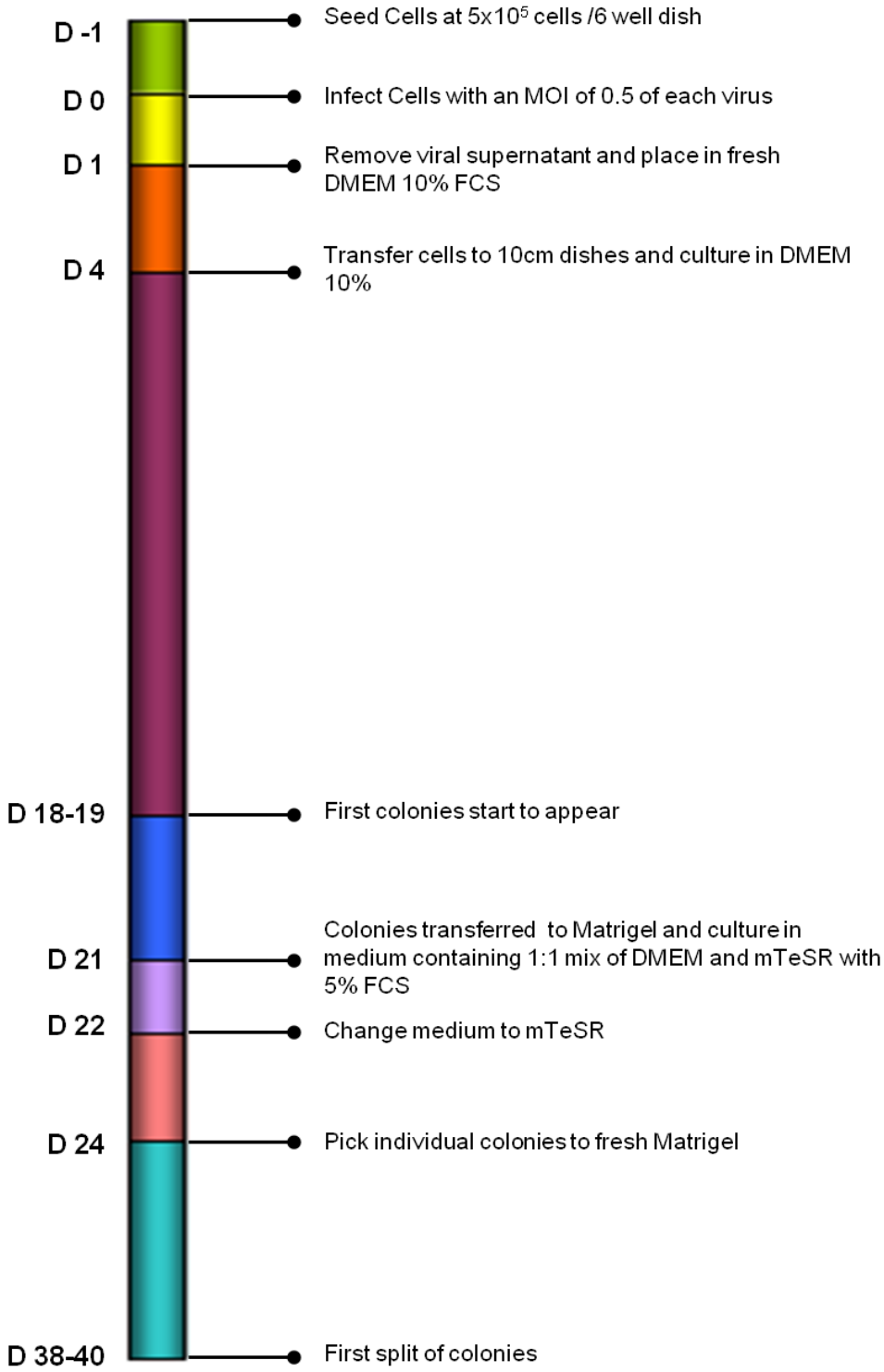


Figure 4.6: Overview of A673-iPC colony generation

Cells were infected with an MOI of 0.5 of each virus and viral supernatants were removed after 24 hours. After 4 days p.i., the cells were transferred to 10 cm dishes and cultured in DMEM with 10% FCS. Medium was replaced every 2 days until the emergence of colonies at day 18-19 p.i. After a further 2 days in DMEM 10% FCS, the cells were transferred to Matrigel and cultured for 1 day in medium consisting of a 1:1 mix of DMEM and mTeSR, with 5% FCS before changing the medium to full mTeSR. When colonies were of a sufficient size, they were picked individually to fresh Matrigel. At day 38-40 p.i., the colonies were large enough to be split and further expanded.

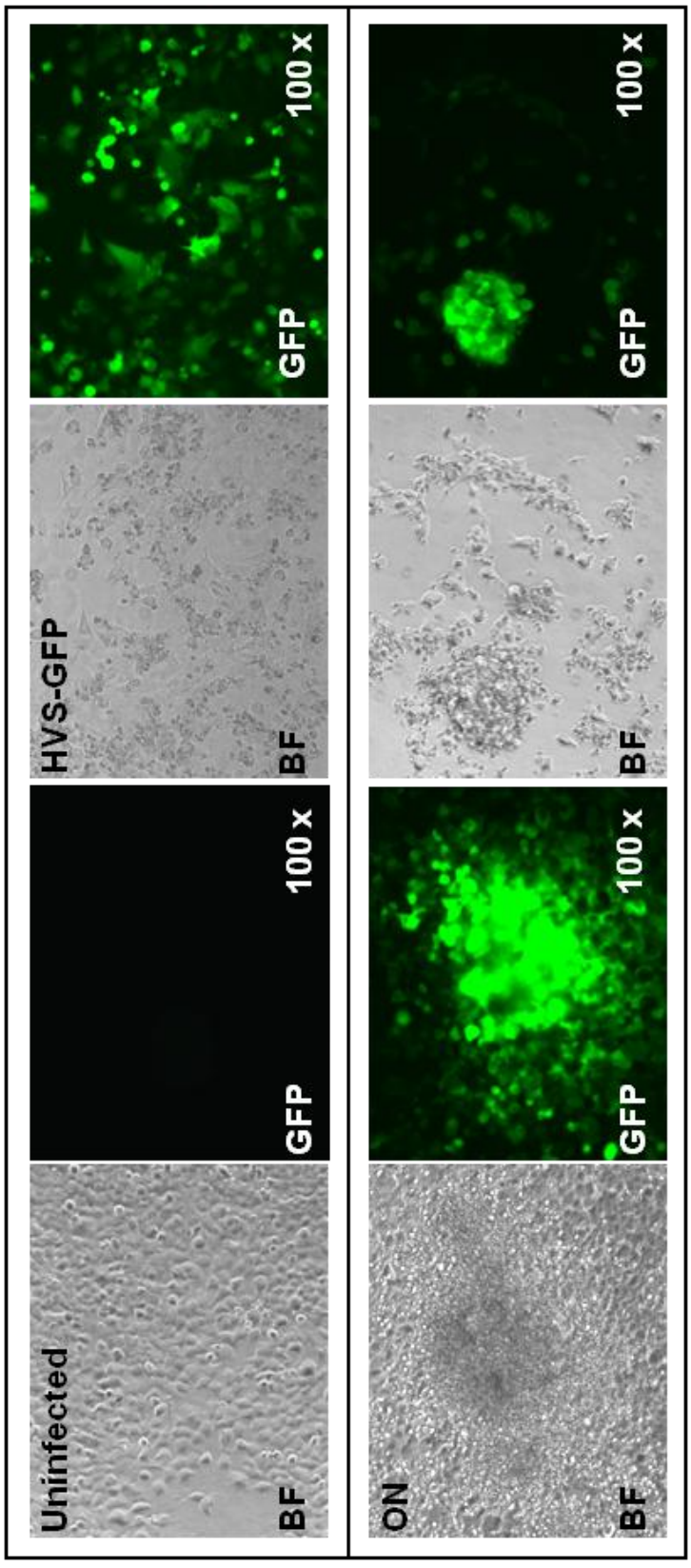


Figure 4.7: A673-iPC Colonies at day 19 post-infection on Primaria 10 cm dishes

After 19 days post-infection, colonies appeared which possessed densely packed morphologies with clearly defined edges. This morphology was distinct from that of the surrounding A673 cells. No such colonies were observed in uninfected and HVS-GFP transduced A673 cells.

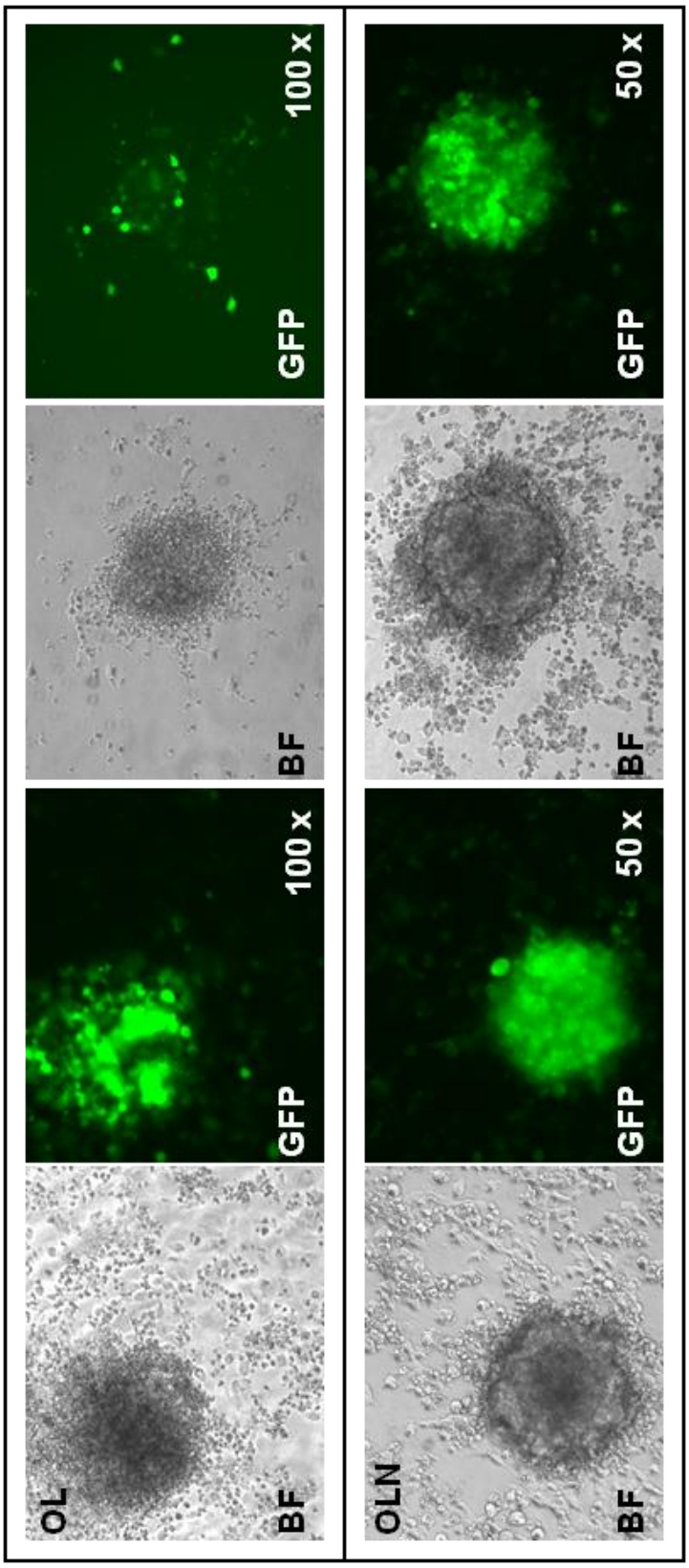


Figure 4.7 (continued): A673-iPC Colonies at day 19 post-infection on Primaria 10 cm dishes.

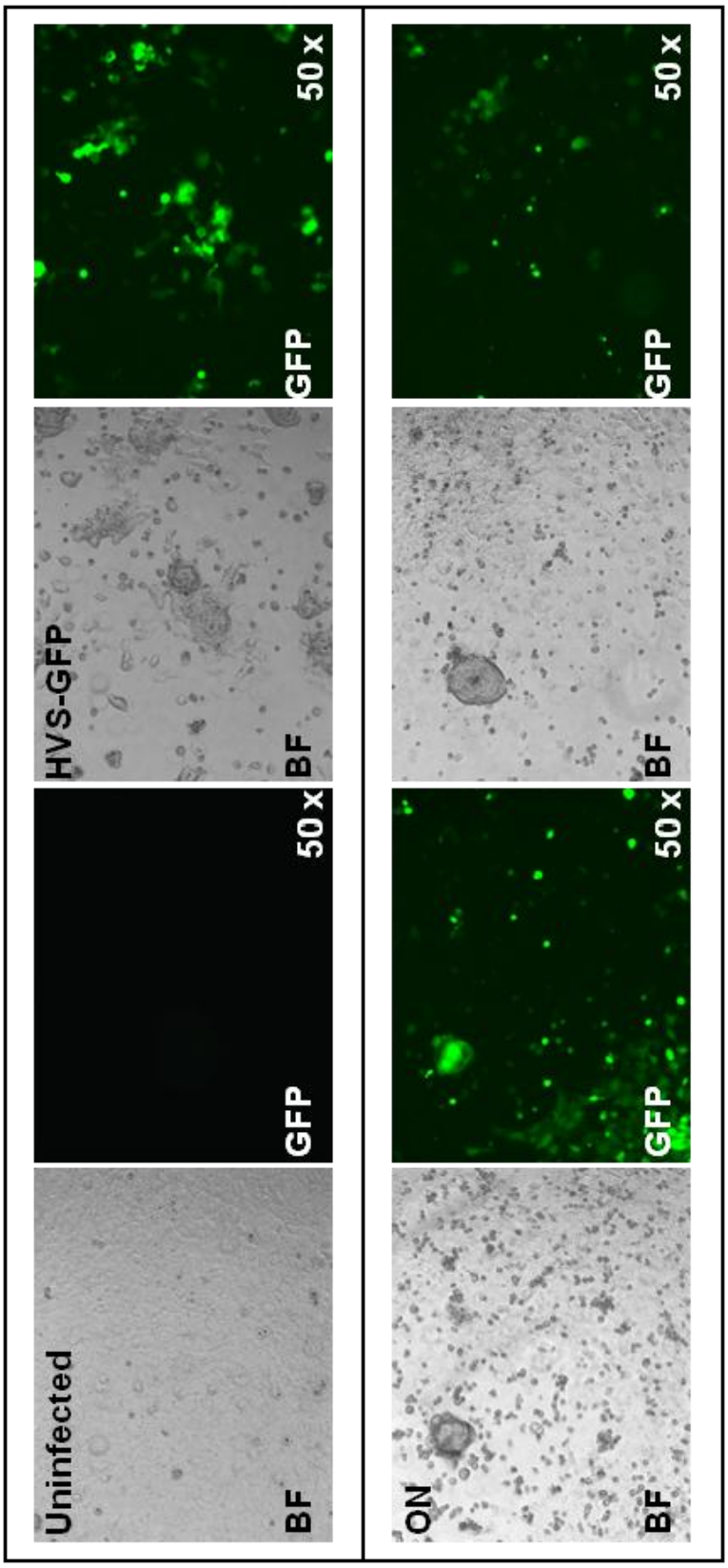


Figure 4.8: A673-iPC colonies at day 22 post-infection; 1 day on Matrigel.

Colonies were imaged 1 day after transferring onto Matrigel coated plates. Uninfected cells were trypsinised prior to transfer onto Matrigel coated plates. Colonies from the ON and OLN conditions appeared to have densely packed spheroidal morphologies, but colonies from the OL condition appeared to be have less defined edges.

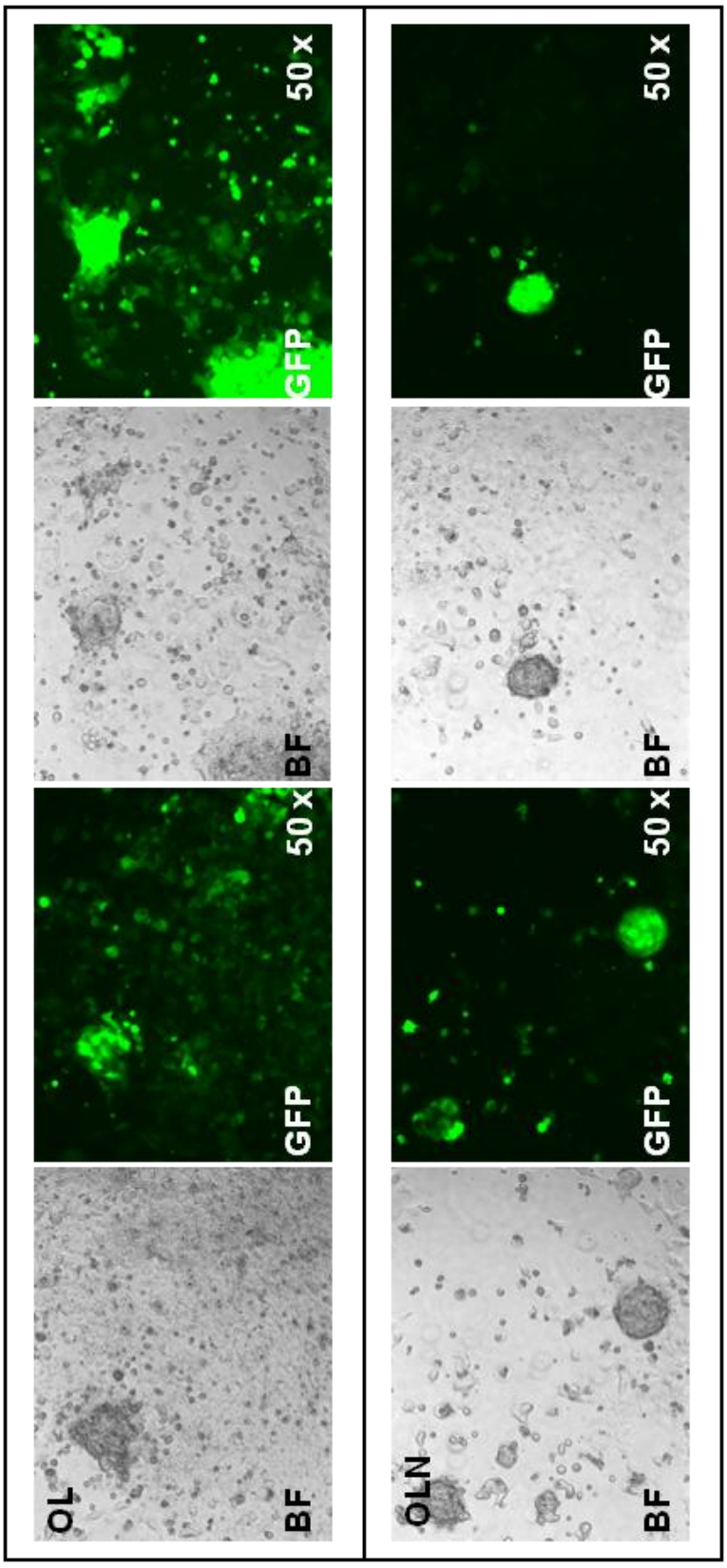


Figure 4.8 (continued): A673-iPC colonies at day 22 post-infection; 1 day on Matrigel.

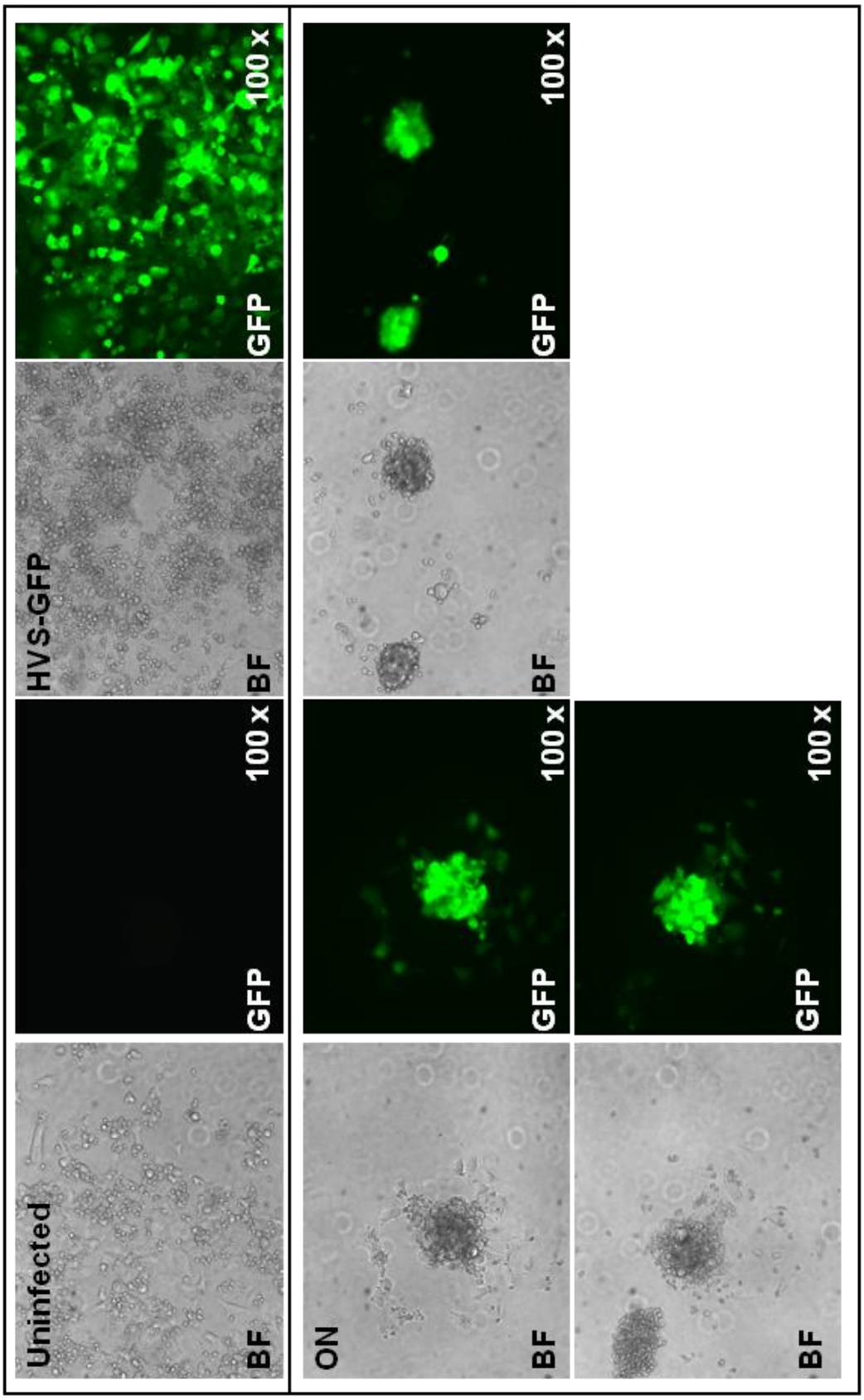


Figure 4.9: Individually picked A673-iPC at day 25 post-infection; 4 days on Matrigel, 1 day after picking. OL and OLN colonies retain their densely packed, spheroidal morphologies, with few cells growing out from the colony. ON colonies, however, are less well formed with more cells growing out from the colony.

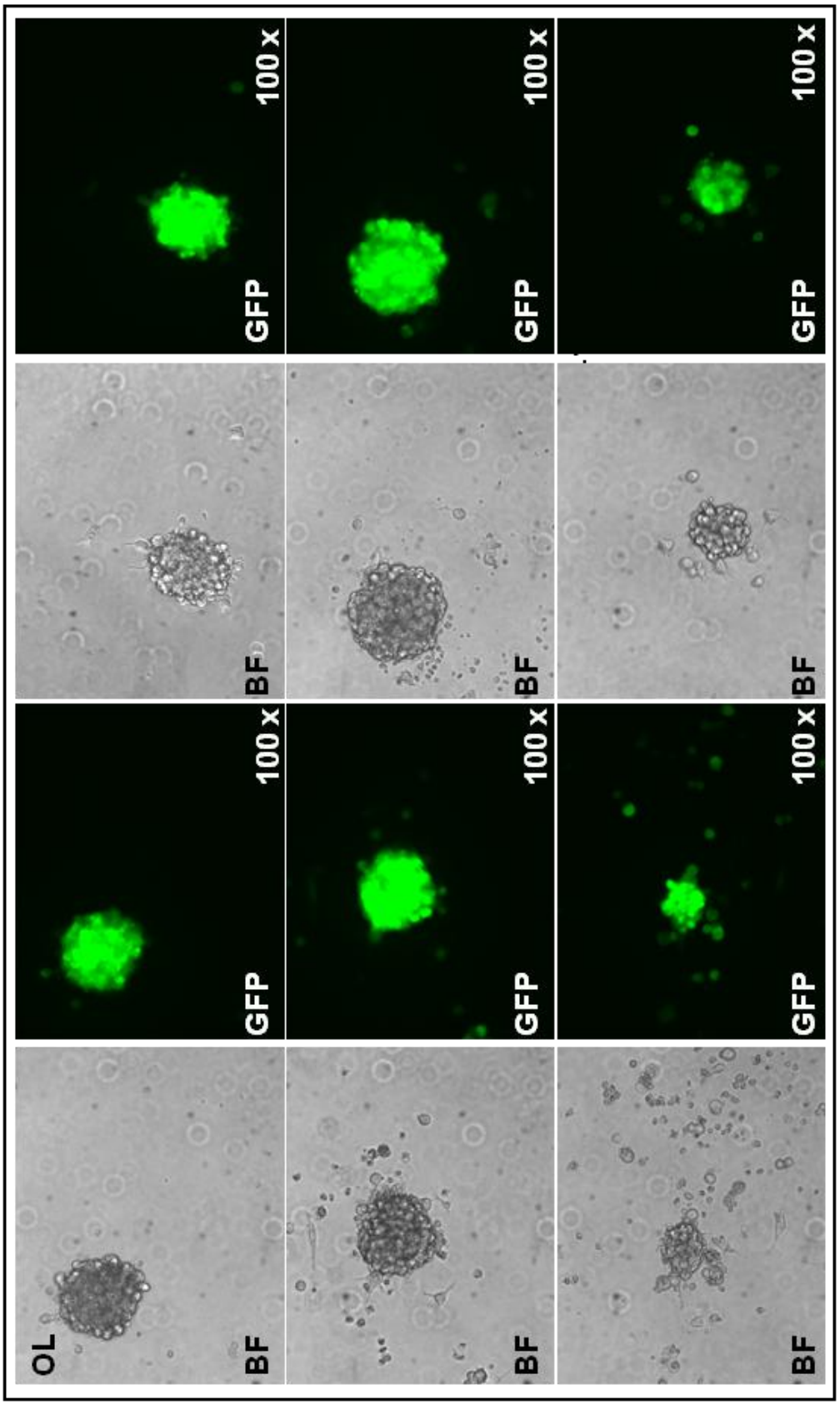


Figure 4.9 (continued): Individually picked A673-iPC at day 25 post-infection; 4 days on Matrigel, 1 day after picking.

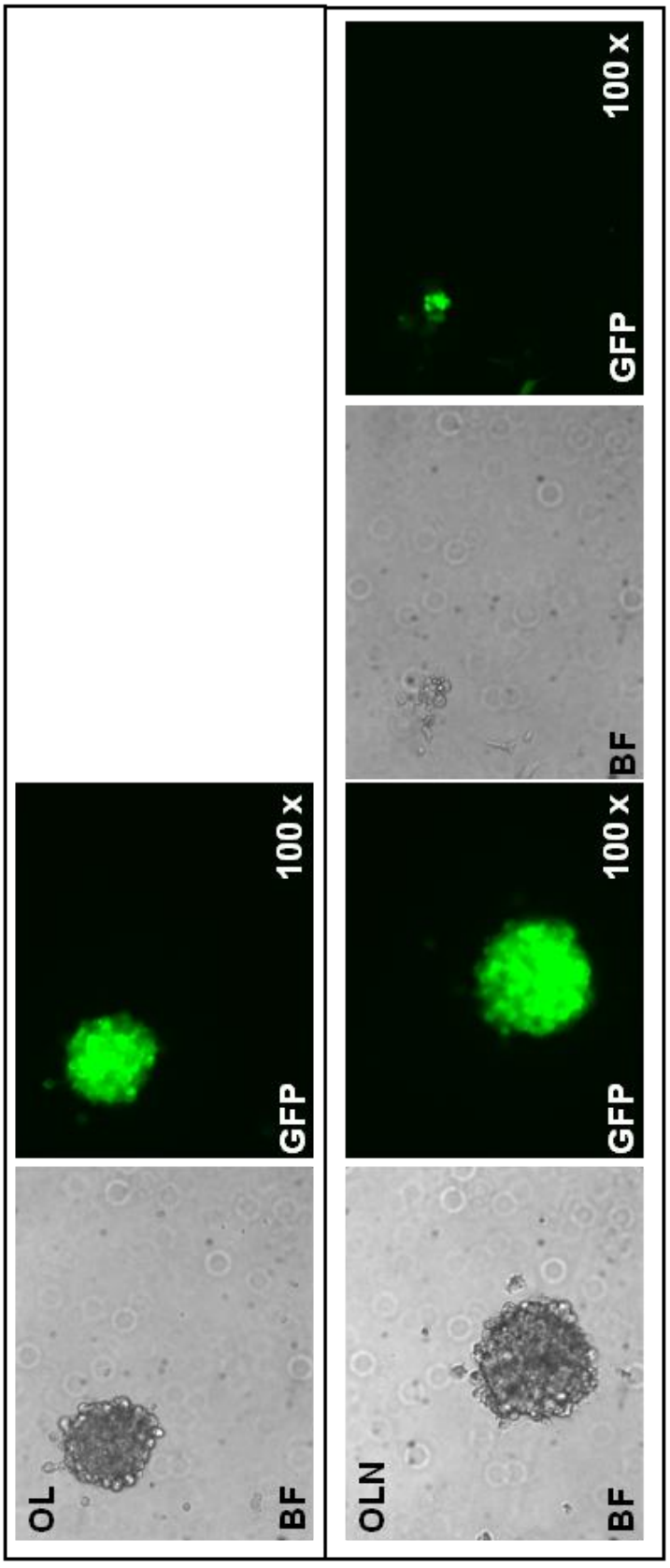


Figure 4.9 (continued): Individually picked A673-iPC at day 25 post-infection; 4 days on Matrigel, 1 day after picking.

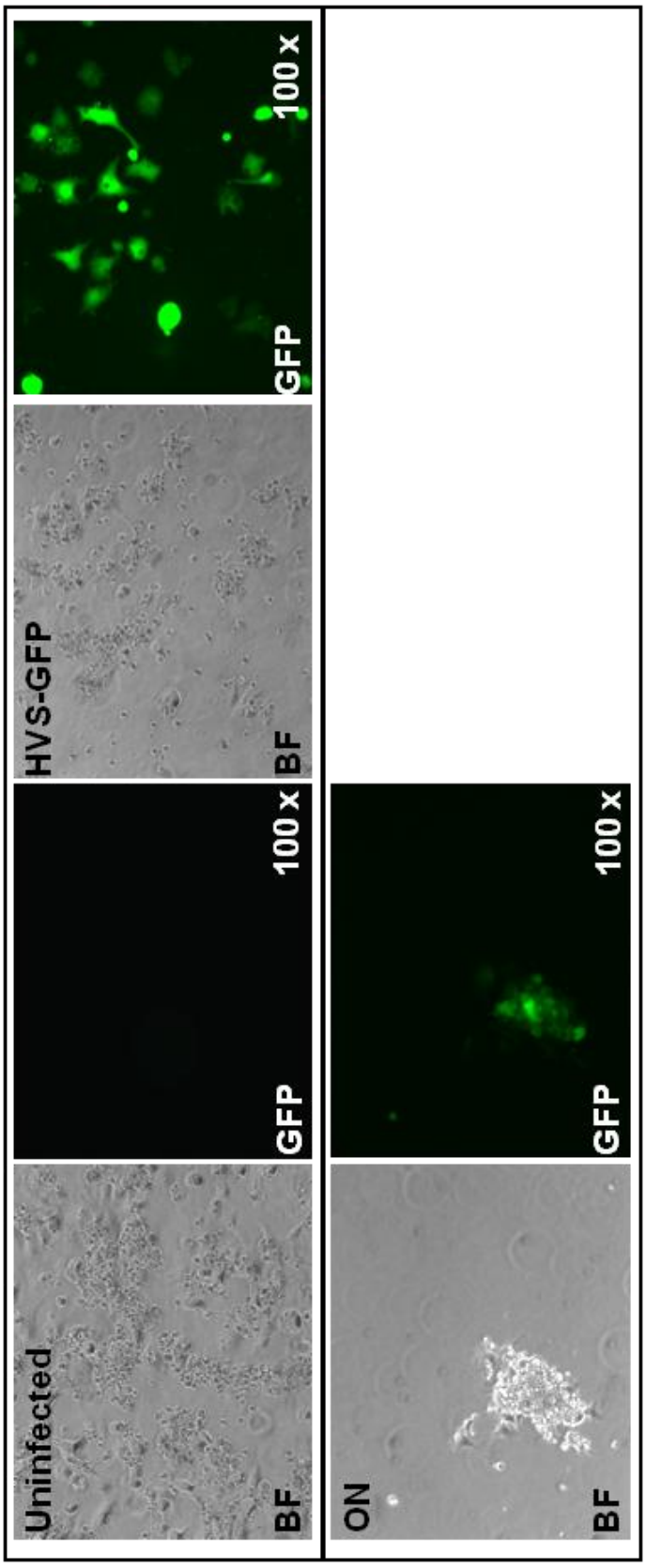


Figure 4.10: A673-iPC colonies at day 28; 7 days on Matrigel.

Many of the picked A673-iPC colonies deformed after picking. The remaining colonies reduced in size after being picked and their edges became less defined. OL produced the largest number of viable colonies, with 6 colonies. ON and OLN both produced a single viable colony each.

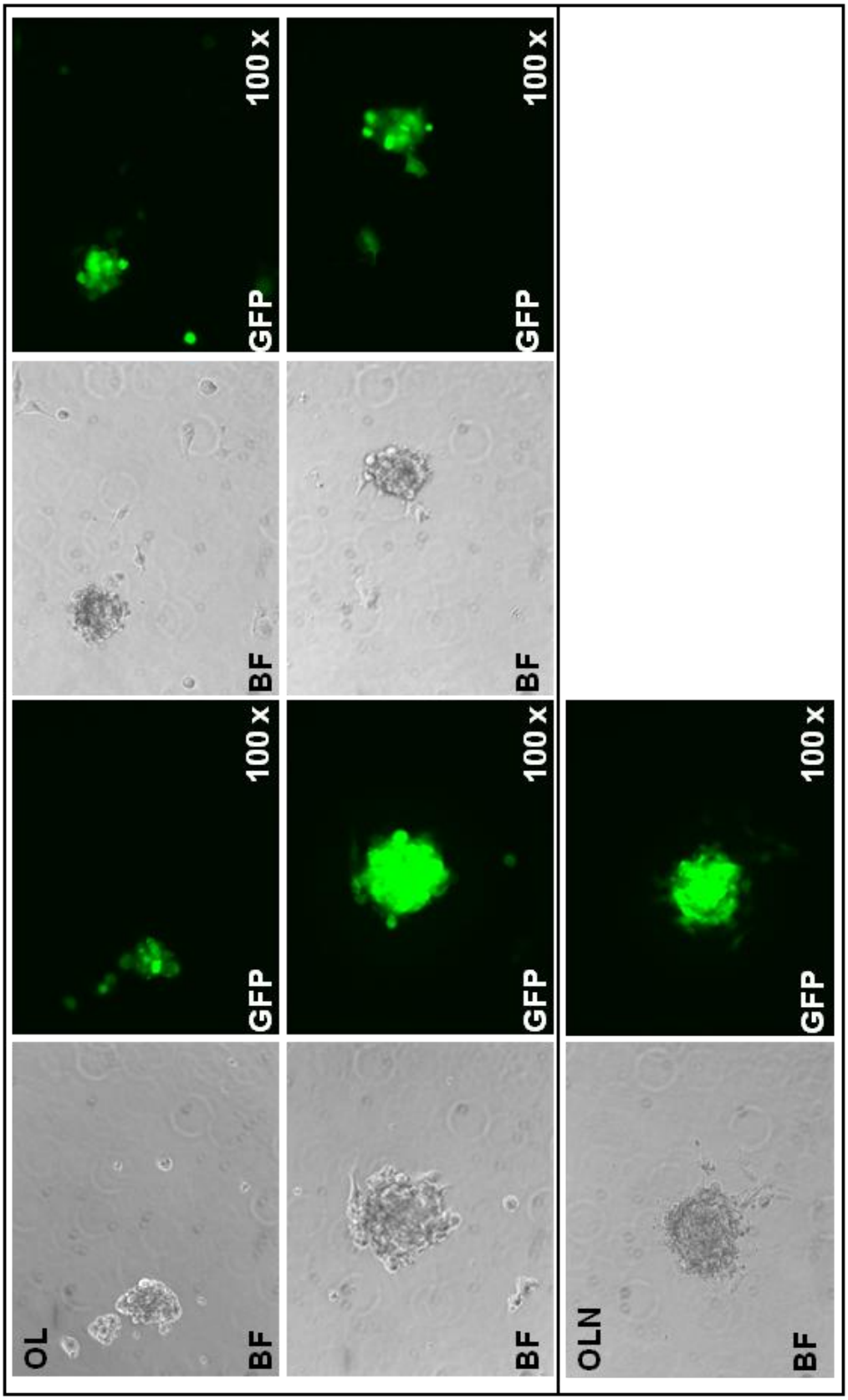


Figure 4.10 (continued): A673-iPC colonies at day 28; 7 days on Matrigel.

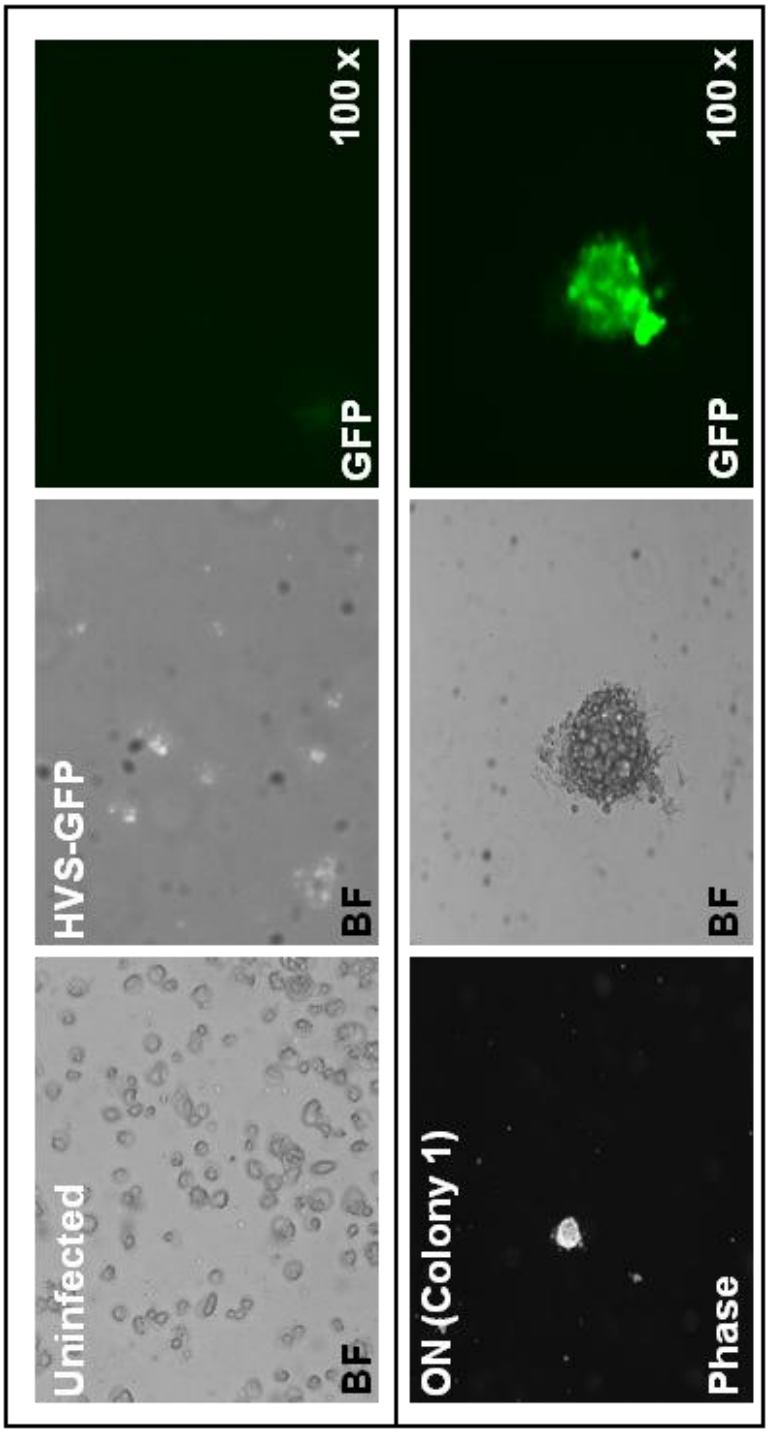


Figure 4.11: Surviving colonies 38 days post-infection (17 days on Matrigel).

Uninfected and HVS-GFP infected A673 cells did not survive after 17 days on Matrigel. Four colonies survived, although Colony 1 (ON) could not be expanded, and cells within this colony appeared to be apoptosing. Colonies 2, 3 and 4 produced densely packed, phase-bright colonies, with clearly defined edges. However, Colony 4 had some cells outgrowing from the central cell mass. Colony 3 also appeared to have a dark centre under phase-contrast.

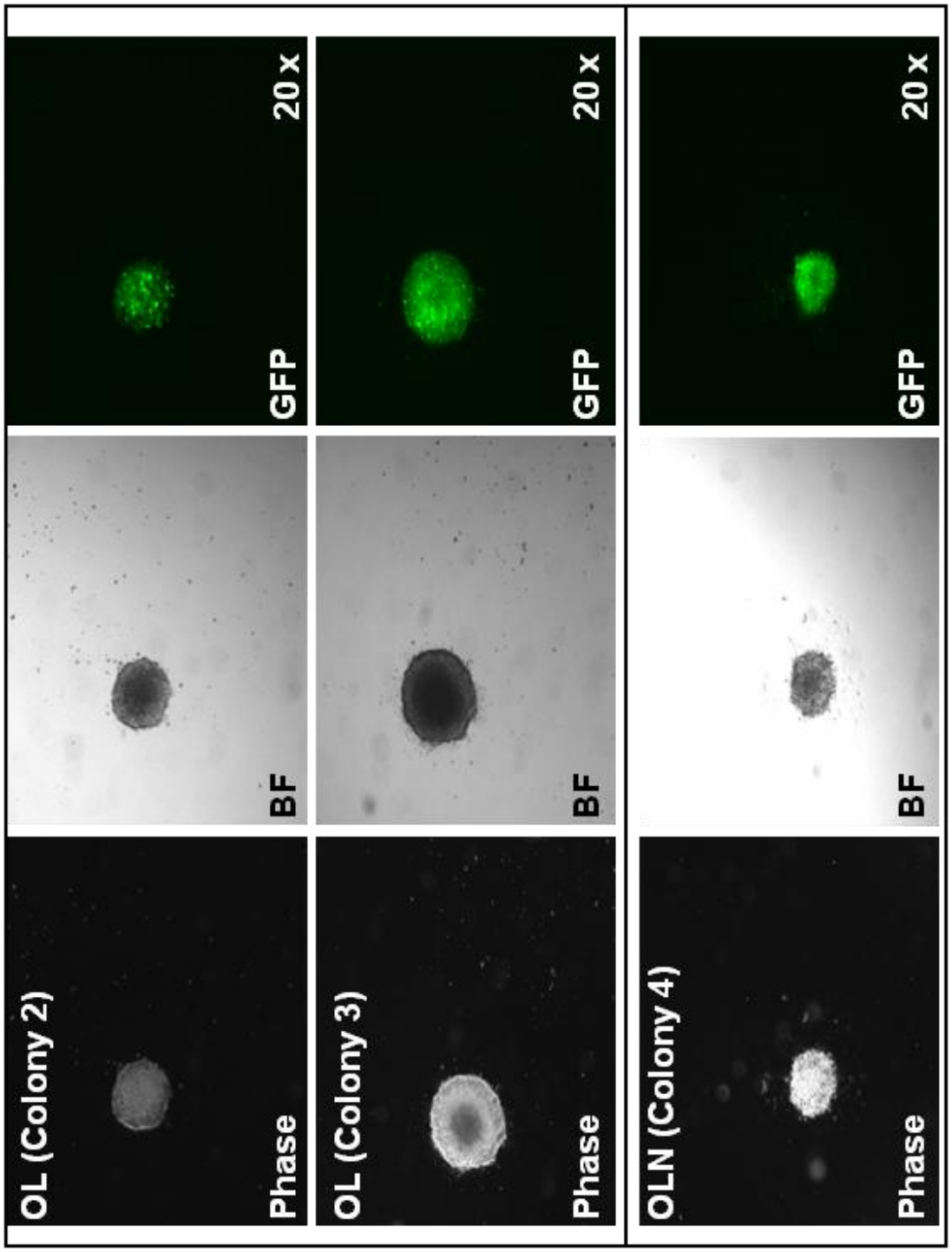


Figure 4.11 (continued): Surviving colonies 38 days post-infection (17 days on Matrigel)

4.6 Optimising Reprogramming Efficiency in A673

A key interest in iPSC technology is to efficiently reprogramme cells using as few exogenous reprogramming factors as possible. As many of the pathways involved in pluripotency are also associated with oncogenesis, it is plausible to assume that A673 cells also express some of the genes associated with pluripotency. Evidence from Figure 4.3 has already demonstrated that A673 cells express Sox2. In addition, the clearly marked differences in colony morphology and survival rates observed in section 4.5 indicate that the different stoichiometries of the pluripotency genes had varying effects on reprogramming efficiency of A673 cells. Therefore, the optimum combination of HVS-iPSC viral vectors required for efficient generation of A673-iPC colonies was determined by infecting A673 cells with various combinations of each of the HVS-iPSC viruses, as shown in Table 4.1. HVS-GFP virus used to ensure each infection condition had the same total MOI of 1.5. Low MOIs were used due to the cytotoxicity issues observed for previous reprogramming attempts, but hygromycin selection ensured that all cells cultured contained HVS episome. The cells were cultured on Primaria 10cm dishes, and the resulting colonies were counted after 19 days p.i.

Virus combination	MOI of individual viruses			
	HVS-GFP	HVS-Oct4	HVS-Lin28	HVS-Nanog
Uninfected	0	0	0	0
HVS-GFP	1.5	0	0	0
O	1	0.5	0	0
L	1	0	0.5	0
N	1	0	0	0.5
OL	0.5	0.5	0.5	0
ON	0.5	0.5	0	0.5
LN	0.5	0	0.5	0.5
OLN	0	0.5	0.5	0.5

Table 4.1: MOIs of each HVS-iPSC virus used to determine the optimum reprogramming virus combination
An MOI of 0.5 of each HVS-iPSC virus was used with cells ultimately infected with a total MOI of 1.5. HVS-iPSC viruses were used alone or in various combinations, with HVS-GFP co-infected to ensure the total MOI used to transduce the cells remained constant.

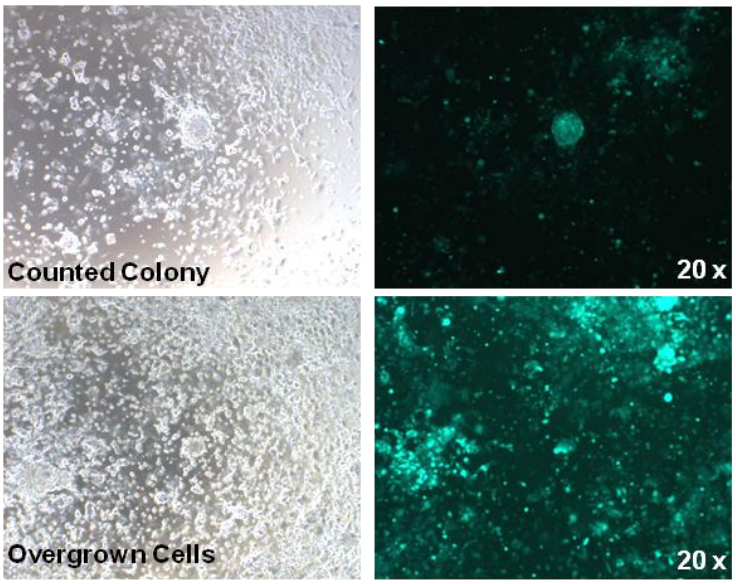


Figure 4.12: Morphology of counted colonies

Colonies that possessed densely packed, spheroidal morphologies with well defined edges were counted. Cells that were roughly clustered, with poorly defined edges to the clusters were identified as overgrown A673 cells, and were not included in the colony count. The examples shown here are a counted OLN colony and a HVS-GFP clustered cell mass.

Virus Combination	No. of Colonies
Uninfected	0
HVS-GFP	0
O	8
L	6
N	7
OL	23
ON	12
LN	6
OLN	29

Table 4.2: Number of counted colonies for each virus combination

Only overgrown clusters of cells were observed in the Uninfected and HVS-GFP infected cells. The greatest numbers of colonies were observed for the cells infected with the OLN and OL virus combinations, with 29 and 23 colonies counted, respectively. ON produced 12 colonies, and low numbers of colonies were also observed for O, L, N and LN.

Figure 4.12 shows representative morphologies of the colonies which were included in the colony count. They have a tightly packed, regular spheroidal appearance with clearly defined edges and the cells within are small with scant cytoplasm. Cells infected with HVS-GFP virus also produced clusters of cells; however, these cells possess poorly defined edges with roughly clustered cells. This morphology was distinct from the colonies observed in A673 cells infected the HVS-iPSC viruses.

Table 4.2 shows representative data of the number of colonies produced for each HVS-iPSC virus combination. The OLN and OL virus combinations produced the highest number of counted colonies with 29 and 23 colonies respectively. ON produced 12 colonies and fewer colonies were observed in O, L, N and LN infected cells.

4.7 Discussion

Due to the current limitations of iPSC-reprogramming vectors, the potential of HVS-based vectors was assessed. HVS is capable of providing prolonged transgene expression without integrating into the host genome. Therefore, this chapter aimed to assess the potential of HVS as a gene delivery system for reprogramming both NSCs to iPSCs and for the generation of ESFT iPCs.

Firstly, HVS was assessed for its ability to infect NSCs, as efficient infection of target cells is essential for gene-delivery vectors. NSCs displayed good infectivity with HVS-GFP, however, extensive cell death was observed upon infection. This high level of cytotoxicity may be caused by the virus predominantly undergoing lytic replication in these cells. Although HVS mostly forms latent infections in human cells, Goodwin et al. demonstrated HVS was capable of latent persistence which could then be reactivated into lytic replication in the lung carcinoma cell line, A549 (Goodwin *et al.*, 2001). Furthermore, HVS has been demonstrated to be capable of specific cytopathic infections in two pancreatic cell lines (Stevenson *et al.*, 2000c). However, another, more likely explanation could be the virus-containing supernatant was responsible for the observed cytotoxicity. Viral supernatants are harvested directly from the permissive OMK cells after the cell sheet has been lysed and will therefore contain various cytokines

caused by the viral infection, which could induce apoptosis. To circumvent this issue, purified virus stock could be produced in the future.

However, HVS may still have potential as a gene-delivery system in brain tissues as NSCs demonstrated very good infection rates with HVS-GFP. Furthermore, the biosafety of HVS-based vectors could be further improved by removing the two genes required for lytic replication, ORF50 and ORF57, and growing HVS particles in a permissive helper cell line stably transfected with constructs expressing these two genes. The resulting viruses would therefore only be capable of lytic replication in these stably transfected permissive helper cell lines.

The second aim of this chapter was to reprogramme ESFT cells into a cancer stem cell-like state. to provide an abundant source of CSCs from various different cancers for further study. A673 cells were selected for reprogramming efforts in this study as they endogenously expressed Sox2, similarly to other ESFT cell lines (Figure 4.3) but displayed the highest rates of transduction (Figure 4.4a). Therefore A673 cells could be transduced efficiently and did not require a functional Sox2 expressing viral construct for successful reprogramming.

Upon transducing cells with HVS-Oct4 alone, colonies were produced after 19 days p.i. possessing morphologies which were distinct from the HVS-GFP infected and uninfected control A673 cells. These colonies possessed densely packed, clearly defined edges; morphologies similar to, but not identical, to those described for iPSCs, and had phase-bright centres. However, they were spheroidal, rather than flat circular colonies, which is the described morphology for iPSCs (data not shown). This could indicate that A673-iPC colonies are not true iPSCs. This is not unexpected as these cells were derived from cancer cell lines rather than primary somatic cells.

Interestingly, these A673-iPC colonies failed to thrive in feeder-free stem cell culture conditions and deformed after 4-6 days in these culture conditions. This deformation and cell death may also indicate incomplete reprogramming, with the A673-iPC cells not sufficiently altered from their original somatic cell state to survive in these culture conditions. Incomplete reprogramming has already been observed upon Somatic Cell Nuclear

Transfer of nuclei from Embryonic Carcinoma cells with enucleated oocytes, where the resulting cells were demonstrated to be an intermediate between cancer cells and ESCs (Chang *et al.*, 2010). Moreover, similar cell death was also reported by Miyoshi et al. during attempts to reprogramme a colorectal cell line, where 98% of cells transfected with reprogramming constructs failed to thrive (Miyoshi *et al.*, 2010).

Another possible explanation for the extensive cell death observed in the resulting A673-iPC colonies could be due to lytic replication of the HVS vectors. HVS was analysed for the presence of lytic replication in A673 cells by infecting with HVS-GFP at MOIs of 0.5 to 3 and although some cell death was observed at higher MOIs, HVS-GFP appeared to predominantly form latent stable infection in A673 cells (unpublished observation) However, it is possible that expression of reprogramming transgenes may be activating expression of lytic genes within the HVS-genome. qRT-PCR analysis to determine the levels of the lytic transcripts, ORF50 and ORF57, in addition to a late transcript such as the glycoprotein, ORF8, would indicate if lytic replication was occurring, and if whole, infectious virus particles were being produced within the A673-iPCs.

Infection with different viral combinations produced colonies displaying morphologies identical to the initial reprogramming effort. Also similarly to the HVS-Oct4 induced A673-iPC colonies, some of these colonies deformed in prolonged feeder-free stem cell culture conditions. However, several colonies survived, indicating that combination of the HVS-iPSC viral vectors sufficiently increased the reprogramming efficiency to reprogramme the A673 cells past a critical point in the reprogramming process. Prolonged culture in these conditions demonstrated the colonies produced from the ON combination were the least stable, indicating this combination of HVS-iPSC viruses is not optimal for reprogramming A673 cells. In contrast, colonies from the OL and OLN infection combinations were viable after prolonged culture in mTeSR, with OL being the most efficient vector combination. From these data, it would be reasonable to assume that Nanog is dispensable for reprogramming in A673 cells. Indeed, cells reprogrammed without exogenous Nanog expression (OL) appeared to be the most efficiently

reprogrammed. Potentially, A673 cells may endogenously express Nanog; perhaps exogenous expression of this reprogramming factor would lead to unfavourable stoichiometries of the reprogramming factors, thereby reducing reprogramming efficiency. Because of this interesting observation, the optimal combination of the HVS-iPSC reprogramming vectors in A673 was determined by infecting cells with various combinations of all three HVS-iPSC vectors.

Reprogramming efficiency trials indicated that infection with all three HVS-iPSC viruses (OLN) produced the most efficient reprogramming of A673 cells, producing 29 colonies, with OL producing the second highest number of colonies. However, colonies produced from the OL vector combination had the highest survival rate during the second reprogramming attempt. This may indicate that, although the OLN combination produces the largest number of colonies, these colonies are less stable in feeder-free stem cell culture conditions, with colonies produced from the OL combination being reprogrammed further. Interestingly, colonies were also observed from cells infected with LN viruses or the L and N viruses individually. This may indicate a very low level of endogenous Oct4 expression in A673 cells, with reprogramming occurring upon the enhancement of Oct4 activity by one or more of the reprogramming factors.

There are, however, certain limitations to this approach to measure reprogramming efficiencies. Firstly, cells were infected with low MOIs of each HVS-iPSC virus because of the cytotoxic effects observed upon transduction with HVS-GFP at an MOI of 1. This means that ultimately, a maximum of 70% of the cells were infected with each individual HVS-iPSC virus. Therefore, the proportion of cells being transduced with all viruses when combinations of the viruses are used is likely to be very low. For example, upon infection with all three viruses, if 70% of the cells are infected with each virus, only around 35% of the cells will receive all three viruses. Hygromycin selection ensures that all the cells remaining carry a copy of the HVS episome, but it is impossible to say if the cells are infected with one or many of the HVS-iPSC viruses, and if so, which ones. Ideally, if the cytotoxicity issues identified could be circumvented, higher MOIs of each

HVS-iPSC viruses would be used to ensure as close to 100% of cells as possible were infected with each of the viruses. This would increase the likelihood of each cell containing all of the viruses. Furthermore, colony identification was further complicated by clusters of cells which did not resemble the colonies observed for the previous reprogramming attempts. Therefore, further investigations into reprogramming efficiencies would require identification of the colonies using alkaline phosphatase staining.

In summary, these data suggest that further work is required to improve the biosafety of HVS-iPSC viruses, and purification of virus particles is essential to reduce the cytotoxicity observed upon infection with HVS-GFP in NSCs and ESFT cells. This would eventually enable further reprogramming attempts on NSCs using the HVS-iPSC viruses. Furthermore, the HVS-iPSC viral vectors were capable of eliciting some degree of reprogramming in the ESFT cell line, A673, with the generation of A673-iPC colonies. However, these colonies are not identical to true iPSC cells. Moreover, the various morphologies and survival rates observed for the A673-iPC colonies indicate that these colonies are at different stages of the reprogramming process, with some being insufficiently reprogrammed to survive.

Chapter 5
Characterisation and Differentiation of A673-iPC Colonies

5 Characterisation and Differentiation of A673-iPC Colonies

5.1 Introduction

Although tumours were originally believed to be monoclonal growths of cells, it has since been demonstrated that they comprise a heterogeneous population of cells, each with their own role in tumour progression and survival (Heppner, 1984). The tumour microenvironment is a complex system of many cell types, including endothelial cells, fibroblasts, smooth-muscle cells, lymphocytes, granulocytes, macrophages and cancer stem cells (CSCs). These cells all play different roles in maintaining the tumour, including releasing cytokines such as VEGF, FGF2 and TGF-β, which can cause fibroses, aberrant angiogenesis and recruitment of various cells, including Mesenchymal Stem Cells (MSCs) to the site of the tumour (Hanahan & Weinberg, 2011).

The discovery of CSCs in tumours helps to explain why many of the current anti-cancer therapies fail, with large numbers of cancers developing drug resistance. Currently, many of the anti-cancer drugs target highly proliferating cells by inhibiting proliferation or by DNA damage. However, these highly proliferating cells tend to be more differentiated, with limited self-renewal and differentiation capacity, thus being unable to seed tumours. Therefore, most chemotherapy treatments remove the bulk of tumours, but are incapable of eradicating the tumour initiating CSCs, which can then cause metastasis and regrowth of the tumour.

CSCs themselves possess many properties which make them more resistant to such forms of anti-cancer therapies. These include their more quiescent nature which prevents DNA damage, high expression levels of multi-drug resistance pumps, and the ability to reform tumours from only a small number of remaining CSCs. Furthermore, there is increasing evidence linking the process of Epithelial to Mesenchymal Transition (EMT) to drug resistance in cancers. Firstly, EMT produces cells more closely resembling MSCs, which are more resistant to drug treatments (Houthuijzen *et al.*, 2012). Secondly, there is increasing evidence to suggest that the EMT

switch may actually produce cells with properties of CSCs (Mani *et al.*, 2008; Morel *et al.*, 2008; Singh & Settleman, 2010). This further complicates the design of anti-cancer therapies as, if only CSCs are targeted, they may be replaced by de-differentiation of other cells within the tumour (Gupta *et al.*, 2009a). In addition, the EMT has roles in metastasis, with the resulting mesenchymal-like cells being able to move to other areas of the body and seed new tumours (Peinado *et al.*, 2007).

Because of their importance in drug-resistance, tumour-initiation, metastasis and patient relapse, there have been several drugs designed to specifically target CSCs. These drugs work by either targeting self-renewal signalling or tumour-initiating cell-surface markers (Zhou *et al.*, 2009a). However, designing drugs against CSCs is further complicated by the necessity to isolate or enrich CSC populations from tumours.

CSCs have a highly unstable nature caused by the genetic and epigenetic perturbations which occur in cancer cells. This can confound typical methods of CSC identification, therefore, isolation of these cells is presently proving difficult. For instance, CSCs are identified by the presence of certain phenotypical markers originally identified in normal stem cells, such as $CD34^{+}/CD38^{low/-}$ for lymphomas, $CD133^{+}$ for brain tumours, and $CD44^{+}/CD24^{low/-}$ for breast cancers. However, evidence suggests that certain cells which do not possess these markers are also capable of initiating cancers (Taussig *et al.*, 2010). Furthermore, *in vivo* xenograft experiments to determine the population of tumour-initiating cells within primary tumour samples provide widely varying results for the same types of cancers (Kelly *et al.*, 2007; Quintana *et al.*, 2008). Transplanting human tumour samples into mice may alter observed CSC numbers as certain factors essential for tumour initiation and growth may not be cross-species reactive (Kelly *et al.*, 2007). In addition, many of the xenograft mice are immunodeficient, thereby facilitating engraftment of human cells. However, these mice will not provide the same cytokine-rich microenvironment which allows tumour growth (Nguyen *et al.*, 2012).

These issues could be circumvented by using iPSC technology to generate pure populations of CSC-like cells from the differentiated cells within the

tumour stroma, thereby allowing the identification of CSC specific anti-cancer therapies. Furthermore, these cells could clarify the link between self-renewal, pluripotency and tumourigenesis, and highlight key factors influencing tumour progression.

5.2 Characterisation of Colonies

5.2.1 A673-iPC Colonies positively stain for Alkaline Phosphatase

Embryonic stem cells (ESCs) express elevated levels of alkaline phosphatase on their cell surface. Therefore, the putative A673-iPC colonies generated in Chapter 4 were screened for ESC-like characteristics initially by staining for alkaline phosphatase activity (Figure 5.1). iPSC cells and uninfected A673 cells served as positive and negative controls respectively. All three iPC colonies stained positive for alkaline phosphatase activity. In contrast, the uninfected A673 cells remained unstained. However, Colony 4 eventually deformed after prolonged culture, and this colony had reduced alkaline phosphatase staining compared with Colonies 2 and 3. These data are a preliminary indication of some degree of reprogramming occurring in the A673-iPC colonies.

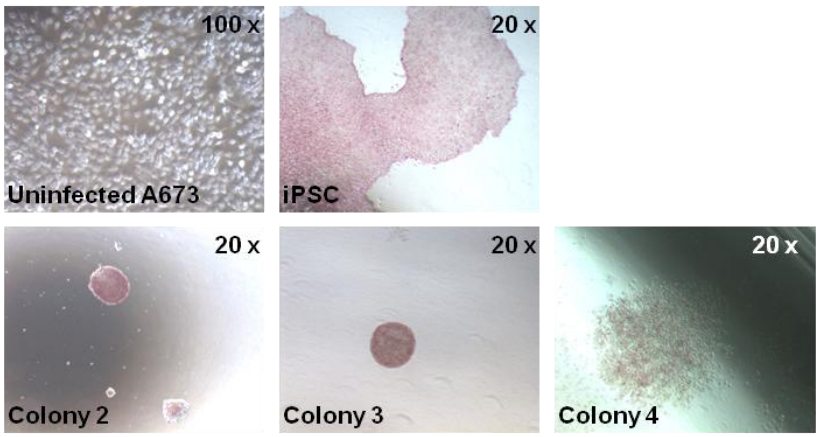


Figure 5.1: Alkaline Phosphatase staining of A673-iPC colonies
Alkaline phosphatase staining was observed for all three A673-iPC colonies, although Colony 4 displayed lower alkaline phosphatase expression, and eventually deformed. Uninfected A673 controls did not possess alkaline phosphatase activity.

5.2.2 A673-iPC colonies demonstrate elevated levels of Embryonic Stem Cell Marker genes

The expression of various genes involved in proliferation and self-renewal are upregulated in ESCs, and their expression is a good indication of successful iPSCs generation. Therefore, to ascertain if reprogramming had been achieved in A673 cells, the expression levels of a number of ESC marker genes were analysed in the generated A673-iPC colonies. To this end, RNA was extracted from Colonies 2 and 3, and control mock and HVS-GFP infected A673 cells. Quantitative RT-PCR was then utilised to assess the expression of endogenous Oct4, Rex1, Klf4 and hTERT within these cells. To ensure only endogenous levels of Oct4 were measured, the 5' primer used for Oct4 transcript detection was designed within the 5' UTR of the Oct4 mRNA. This ensured no amplification of exogenous Oct4 produced from the HVS-based vector would occur.

Figure 5.2 shows the results of the quantitative RT-PCR experiments. Results demonstrate increased levels of Oct4, Rex1 and Klf4 mRNAs in the two A673-iPC colonies compared to uninfected and HVS-GFP infected A673 cells, with Colony 3 displaying the highest increase. Interestingly, hTERT mRNA levels in uninfected and HVS-GFP infected control cells were similar to those observed for the reference standard iPSC cells, with reduced mRNA levels seen in Colonies 2 and 3.

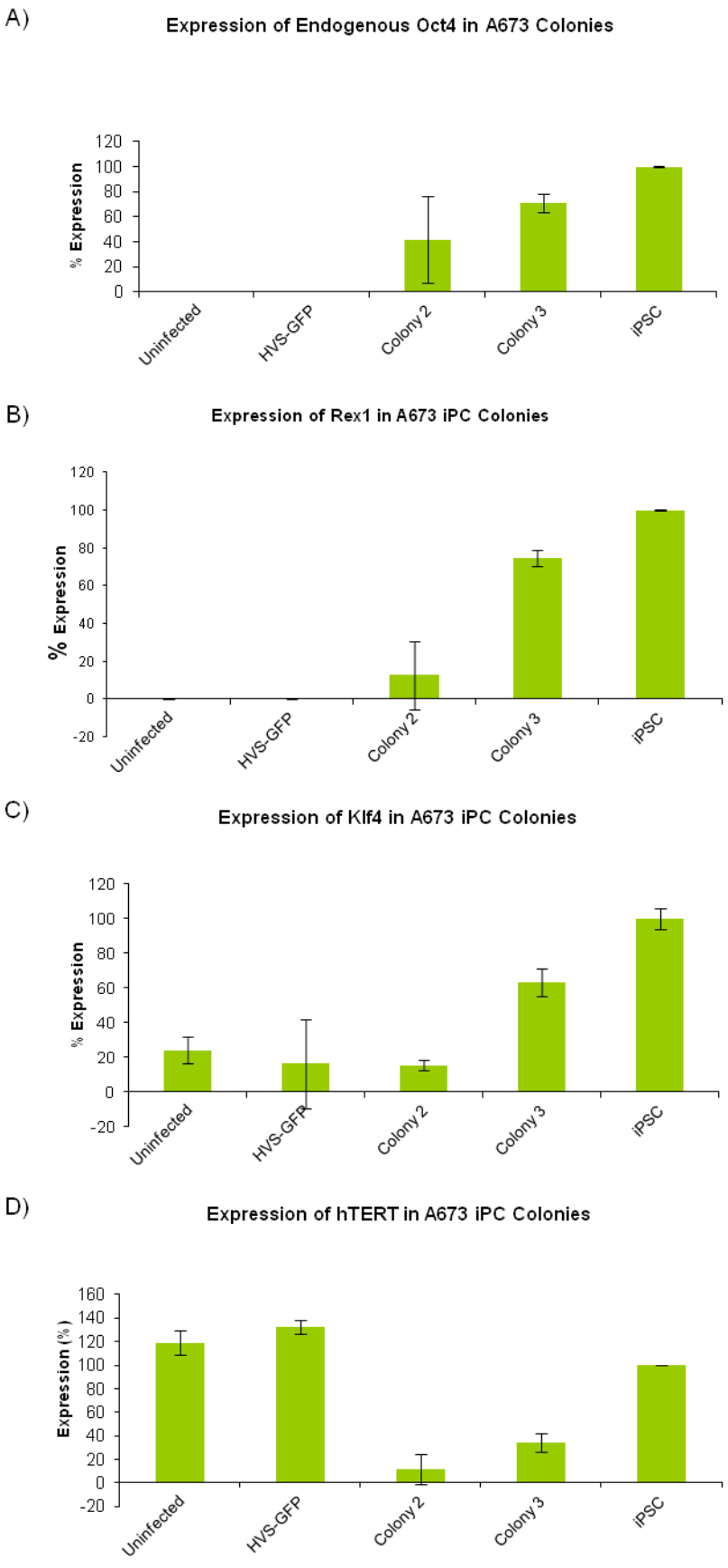
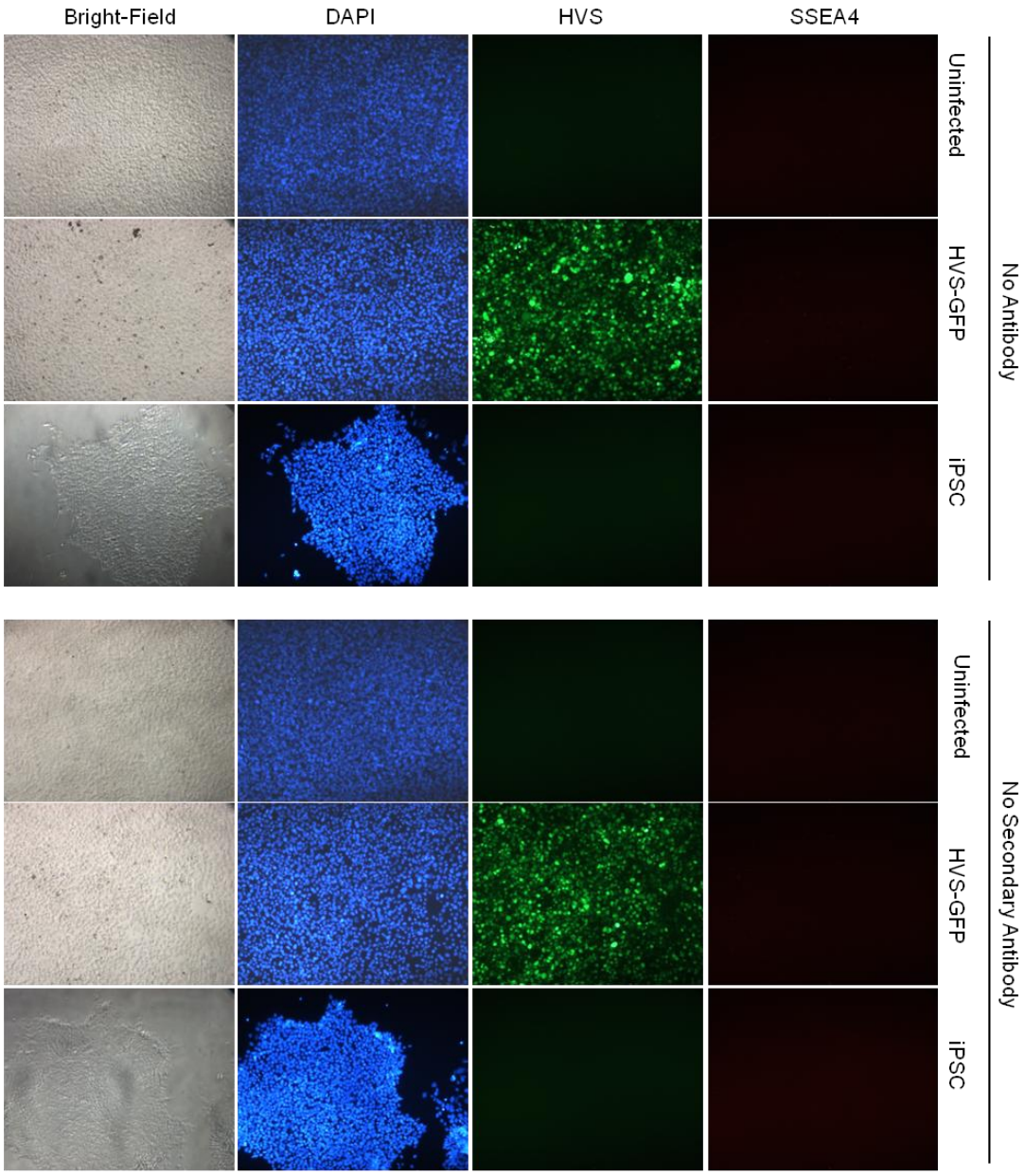


Figure 5.2: qRT-PCR Expression Analysis of ESC marker genes in A673-iPC colonies The expression levels of Oct4 (A), Rex1, (B), Klf4 (C), and hTERT (D) in the A673-iPC colonies were compared with those of iPSCs. iPSCs were used as reference standards. Oct4, Rex1 and Klf4 expression levels are increased in A673-iPC colonies compared to uninfected and HVS-GFP infected A673 cells, with Oct4 and Rex1 levels negligible in these two control cell lines. Colony 3 consistently displayed higher expression levels of the ESC marker gene mRNAs in comparison to Colony 2. Levels of hTERT were higher than iPSCs in uninfected and HVS-GFP infected A673 cells, and were significantly reduced in both A673-iPC colonies.

5.2.3 A673-iPC colonies display differential staining for SSEA4

A further indicator of ESC-like characteristics is the upregulation of the cell surface protein, SSEA4. Therefore, immunofluorescence analysis was performed on A673-iPC Colonies 2 and 3 using an SSEA4 specific antibody provided by Dr. Christian Unger (Centre for Stem Cell Biology, University of Sheffield), and detected using an Invitrogen Alexa Fluor® 546 secondary antibody. Cells were not permeabilised prior to staining to avoid disrupting the cell membranes which would interfere with SSEA4 staining. iPSC cells were used as a positive control, and uninfected and HVS-GFP infected A673 cells served as negative controls.

Figure 5.3 shows SSEA4 expression in A673-iPC Colonies 2 and 3, which is once again higher in Colony 3. This is similar to the positive control iPSC cells. In contrast, SSEA4 staining was not observed in the uninfected and HVS-GFP transduced control cells. Interestingly however, the expression of SSEA4 appeared reduced in some of the colonies split from the same parental A673-iPC colony. This is demonstrated in Figure 5.3, whereby Colony 3a displayed higher SSEA4 expression than Colony 3b. In addition, Colony 3a maintained HVS-iPSC virus GFP reporter gene expression in the absence of hygromycin selection more efficiently than Colony 3b without the presence of hygromycin selection for the HVS-iPSC viral episomes.


Bright-Field
DAPI
HVS
SSEA4
Uninfected
HVS-GFP
iPSC
No Antibody
Uninfected
HVS-GFP
iPSC
No Secondary Antibody

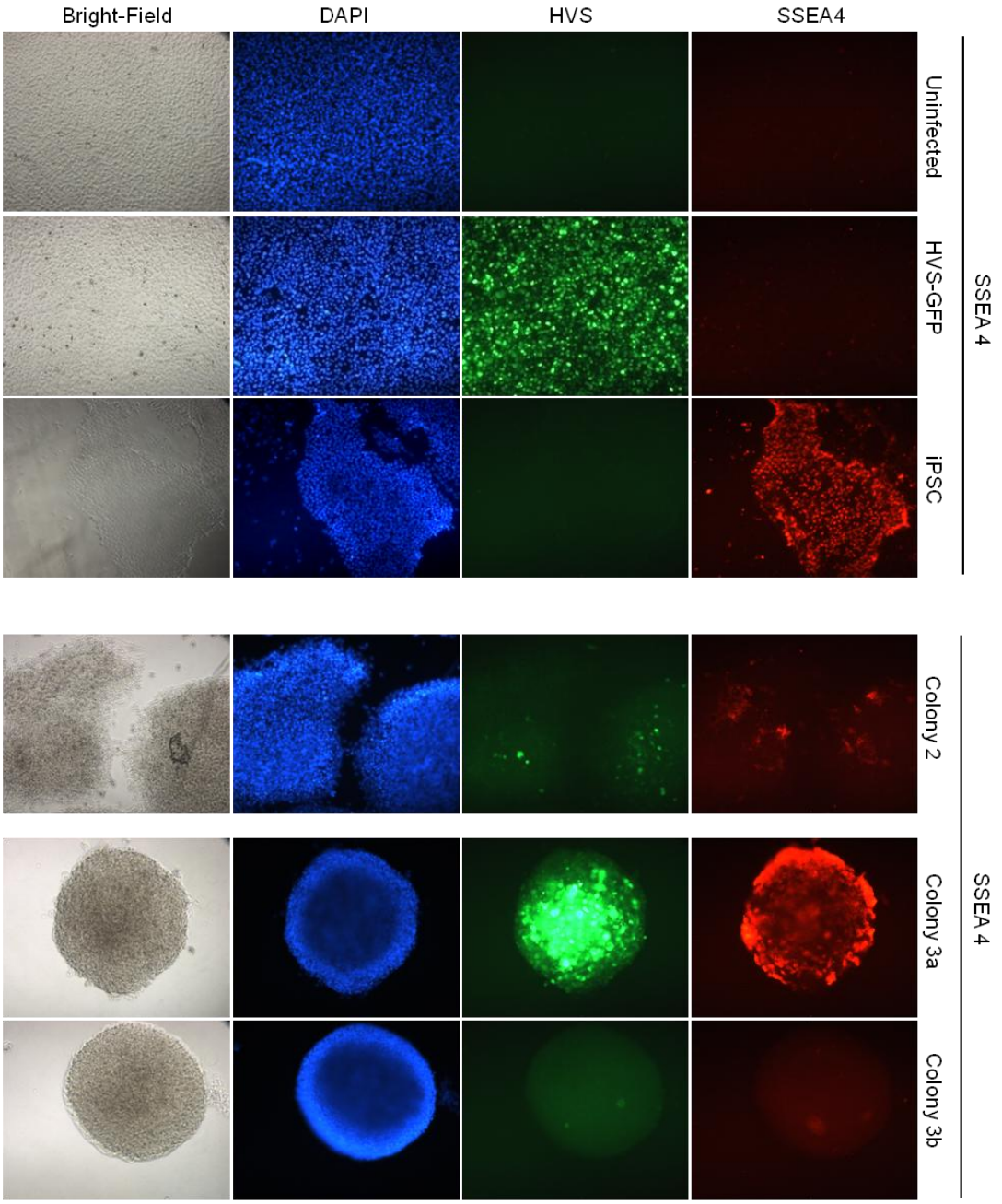


Figure 5.3: SSEA4 staining of A673-iPC colonies

SSEA4 expression is observed on A673-iPC colonies split from Colonies 2 and 3, with Colony 3 displaying higher levels of SSEA4 expression than Colony 2. However, Colony 3a displayed higher levels of SSEA4 expression and maintained GFP expression of the HVS-iPSC episomes better than Colony 3b and Colony 2. No SSEA4 staining was observed in the uninfected and HVS-GFP infected control cells. iPSC cells served as a positive control for SSEA4 staining. All images were taken at 20 x magnification.

5.3 Differentiation

One of the defining features of CSCs is their ability to seed new tumours by differentiating into various types of tumour stromal cells. However, little is known about the exogenous signals involved in this differentiation process due to the difficulties in isolating CSCs. Because of this, animal models are commonly used to study CSC differentiation and tumour formation, but this makes it virtually impossible to measure the effects of the complex interplay of signalling pathways involved in CSC mediated tumour progression. However, iPCs could potentially provide an abundant source of cells, thereby allowing the determination of signalling systems involved in the various types of differentiation in both *in vitro* and *in vivo* models. Therefore, to assess the potential of the A673-iPC colonies generated in this study as models for ESFT development and progression, and to further ascertain the extent of their stem-like properties, differentiation trials were performed.

5.3.1 Non-specific differentiation

Firstly, A673-iPC Colonies 2 and 4 were non-specifically differentiated as described in Figure 5.4. Briefly, this process involved the formation of Embryoid Bodies (EBs) by culturing the A673-iPCs in either DMEM containing 10% FCS, or in EB medium (KO). After 8 days in suspension culture, the EBs were transferred onto gelatin coated dishes, and differentiated cells were allowed to grow out from the central colony mass for a further 8 days. Colony 4 was specifically selected for differentiation trials, as it more readily formed EBs.

After 8 days in suspension cultures in EB medium (KO) or DMEM supplemented with 10% FCS, EBs formed from the control iPSC cells (Figure 5.5 C) and the A673-iPC colonies (Figure 5.5 D and E). Upon transferring the EBs to gelatin coated dishes various morphologies were observed in cells which expanded from the attached iPSC EBs (Figure 5.6) and A673-iPC EBs (Figures 5.7 and 5.8). Cells possessing neural (Figure 5.6 B and E), epithelial (Figure 5.6 A) and cobblestone like morphologies (Figure 5.6 C) were observed growing from the iPSC embryoid bodies after 8 days. However, the cells produced from the A673-iPC colonies had predominantly either fibroblastic, (Figures 5.7 C and 5.8 C), or neural

morphologies (Figures 5.7 A and 5.8 B and E), or their morphologies closely resembled that of the parental A673 cells (Figure 5.7 D and E).

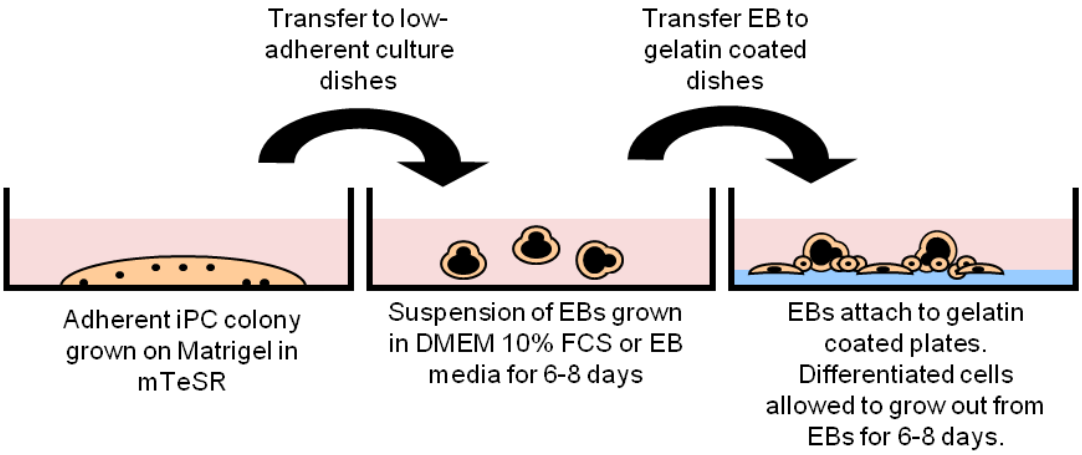


Figure 5.4: Overview of Non-specific Differentiation of iPSCs and A673-iPCs

Adherent colonies were split, the cells transferred to low-adherence tissue culture dishes, and grown in suspension cultures. EBs formed after 6-8 days and were subsequently plated onto gelatin pre-coated tissue culture dishes. Cells were allowed to grow out from the attached EBs and after a further 6-8 days (16 days after initially forming the EBs) various different cell morphologies were observed.

5.3.2 qRT-PCR on non-specifically differentiated Embryoid Bodies

Upon differentiation of ESCs, the genes associated with pluripotency are rapidly downregulated and germline-specific genes are activated. Therefore, to determine the extent of A673-iPC colony differentiation, RNA was extracted from differentiated A673-iPC cells, and qRT-PCR performed to analyse expression of differentiation-associated marker genes were. EBs formed from iPSCs cultured in DMEM with 10% FCS or KO medium, were used as positive controls for the differentiation markers and undifferentiated iPSCs served as reference standards.

Results show a downregulation of endogenous Oct4 expression in A673-iPC EBs formed from Colonies 2 and 4, and iPSC derived EBs. However, small levels of residual Oct4 expression were detected in the A673-iPC EBs (Figure 5.9 A) compared to iPSC derived EBs. In contrast, undifferentiated iPSCs control cells retained high expression of endogenous Oct4.

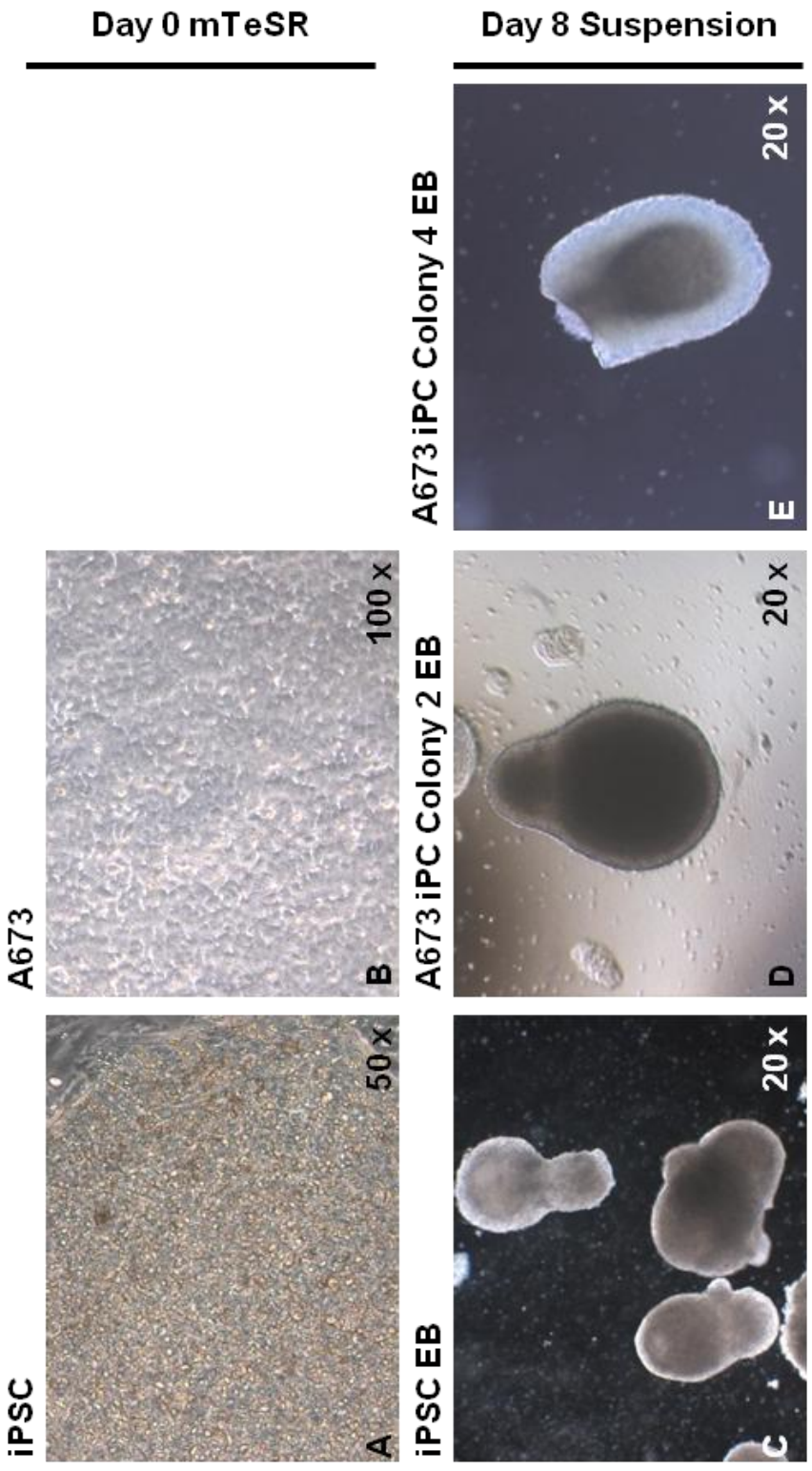


Figure 5.5: Generation of EBs during non-specific differentiation of iPSCs.

A) Undifferentiated iPSC colony, B) Uninfected A673 cells, C) EB formed after 8 days in suspension culture from iPSCs; D and E) EBs formed after 8 days from A673-iPCs.

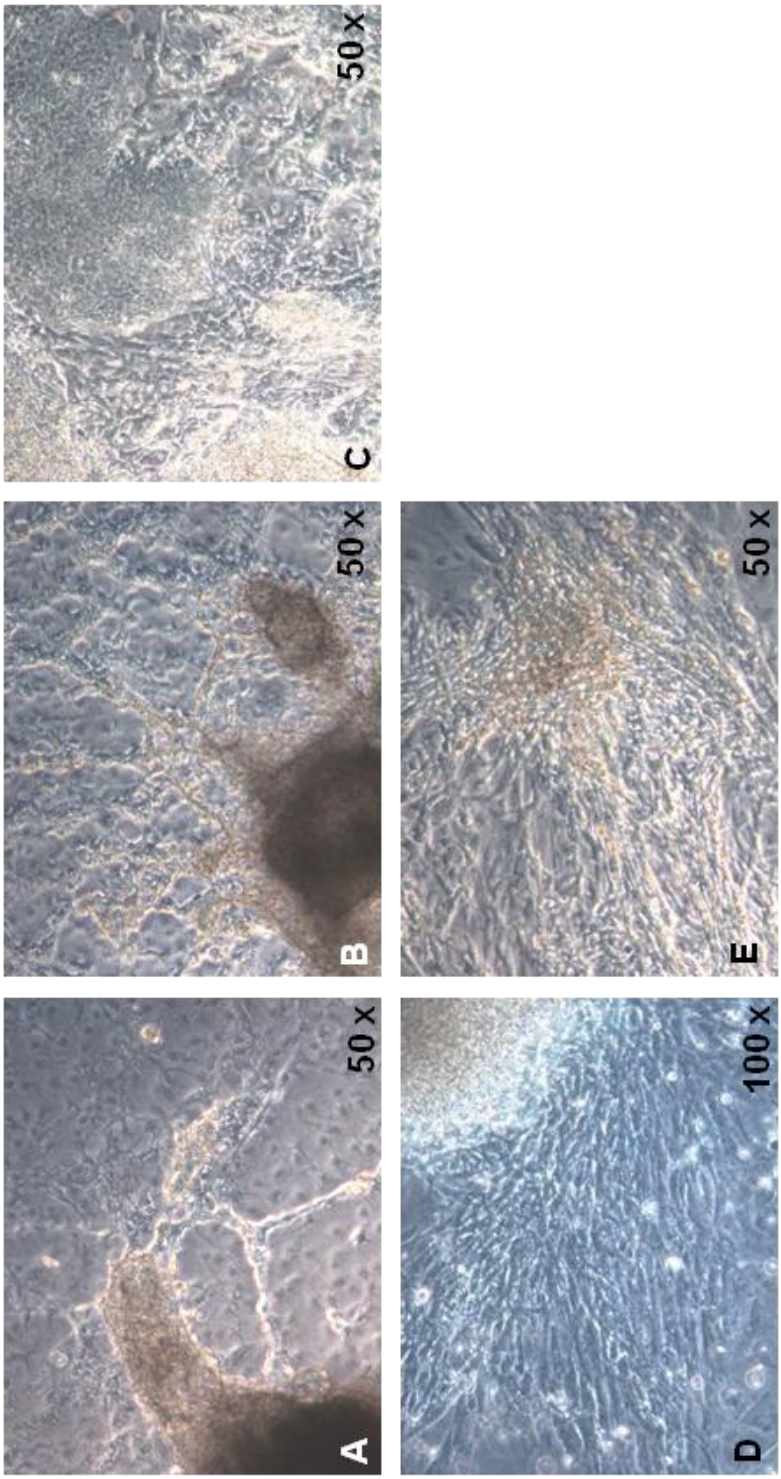


Figure 5.6: Non-specific differentiation of iPSCs at day 16 (8 days on gelatin).

A, B, C, D, E; Differentiated cells growing out from the EB formed from iPSCs 16 days after the EB was formed.

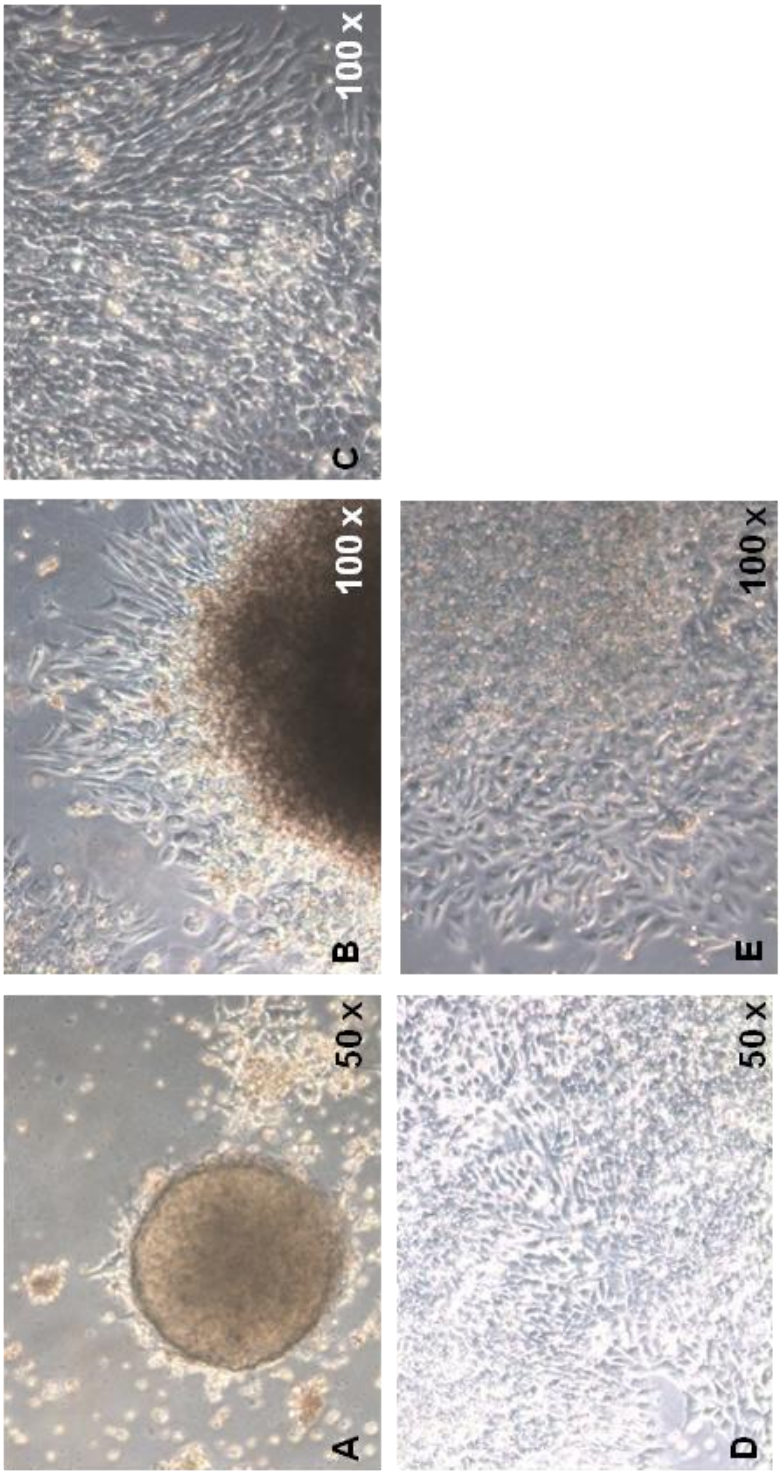


Figure 5.7: Non-specific differentiation of A673-iPC Colony 2 at day 16 (8 days on gelatin).
A, B, C, D, E; Morphology of differentiated cells produced from A673-iPC Colony 2-derived EBs.

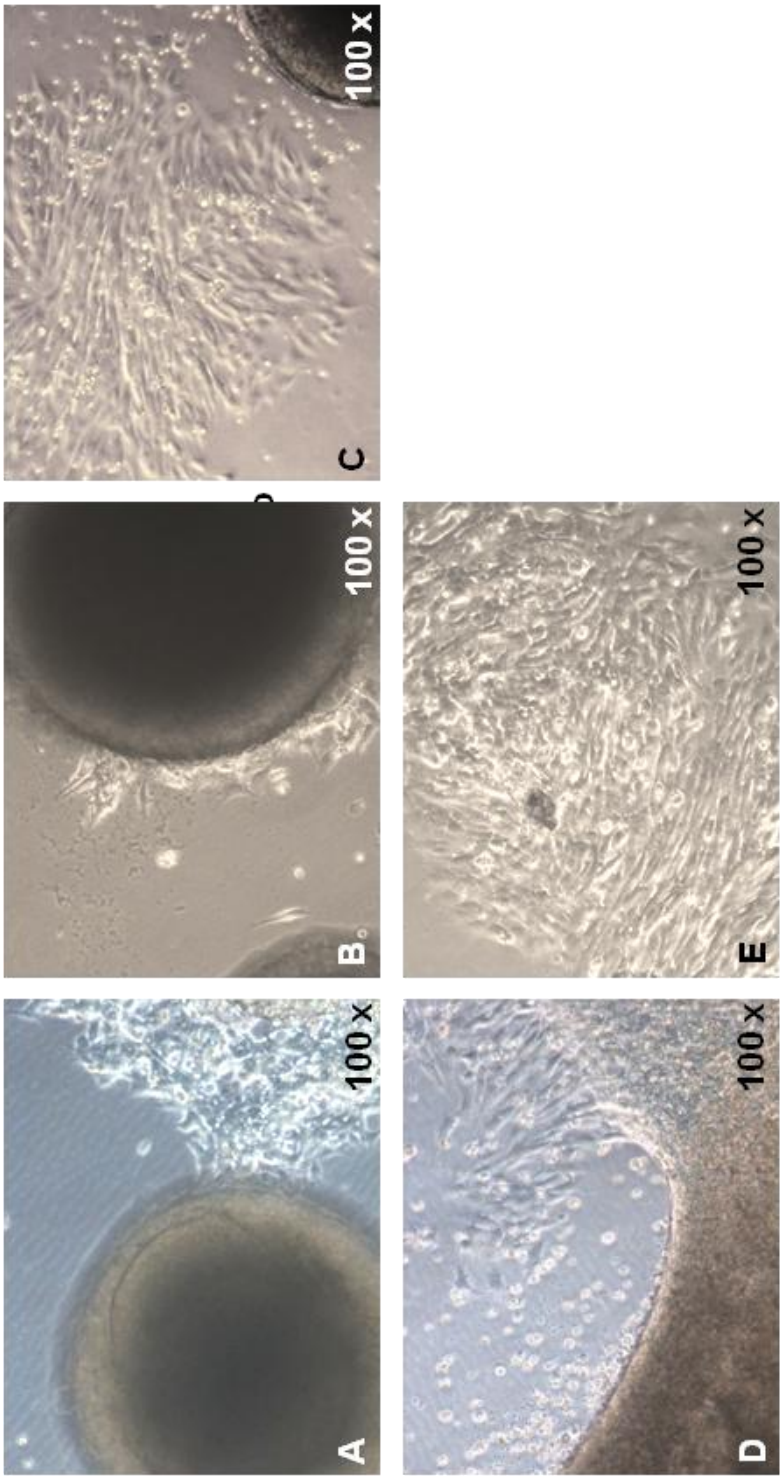


Figure 5.8: Non-specific differentiation of A673-iPCs at day 16 (8 days on gelatin).

A, B, C, D, E; Morphology of differentiated cells produced from A673-iPC Colony 3-derived EBs.

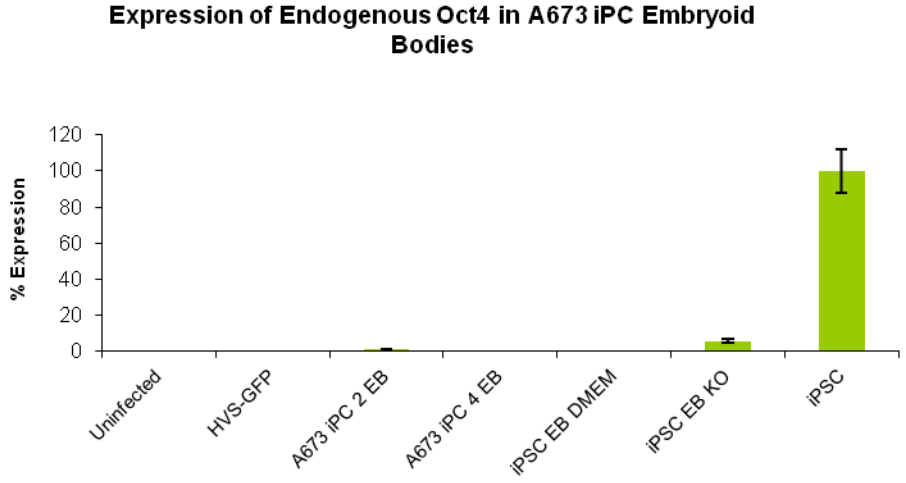
Further to Oct4 depletion, A673-iPC EBs were analysed for expression of lineage specific markers for each of the three germlayers. To this end, the expression levels of MSX1 and MAP2 (ectodermal markers), were analysed, in addition to the endodermal marker, Sox17, and the marker of mesodermal differentiation, Flk1. Figures 5.9 C, D and E show that most of the differentiation markers were not upregulated in the A673-iPC EBs in comparison to the elevated mRNA levels observed for the iPSC EB positive controls. However, A673-iPC EBs demonstrated increased expression of the ectodermal marker, MSX1 (Figure 5.9 B).

5.3.3 Neural Lineage Directed Differentiation

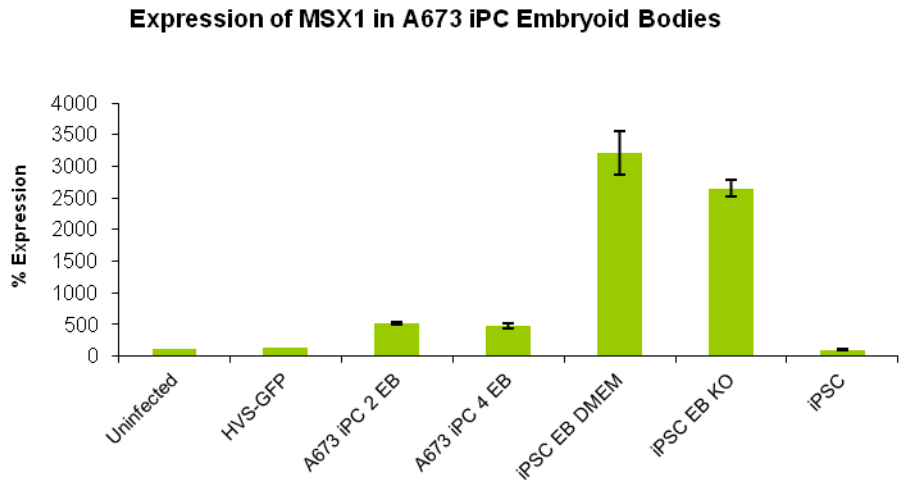
A further test of the stem-like properties of the putative A673-iPC cells is to differentiate the cells towards a specific lineage. Therefore, the A673-iPC colonies were cultured under conditions used to specifically differentiate ESCs down the neural lineage, as summarised in Figure 5.10. A673-iPC colonies were grown in suspension cultures, similar to the non-specific differentiation method, with the exception that the cells were cultured in either Neural Stem Cell Medium from Invitrogen or EB medium containing 1 µM of all-Trans Retinoic Acid (KO RA). Retinoic acid is an inducer of differentiation and promotes differentiation towards a neural lineage.

EBs were formed from iPSC control cells, and A673-iPC Colonies 2 and 4 after 2 days (data not shown). Subsequently, the EBs were transferred to gelatin pre coated dishes and cultured for a further 8 days. Results in Figure 5.11 show that cells possessed various morphologies which had grown from the attached EBs. Control iPSC cells grown in KO medium supplemented with Retinoic Acid (KO RA) or Neural Stem Cell Medium (NSC) developed neuroprogenitor-like morphologies (Figure 5.11 B), with some cells continuing to differentiate and form neuronal networks (Figure 5.11 C), and neurons (Figure 5.11 D). These morphologies were clearly distinct from the undifferentiated iPSC stem cell morphology, (Figure 5.11 A), as the cells were elongated and formed connecting networks. In contrast, the A673-iPC colony grown in KO RA medium (Figure 5.11 H) failed to display any changes in morphology to normal A673 cells grown in DMEM with 10% FCS (Figure 5.11 F).

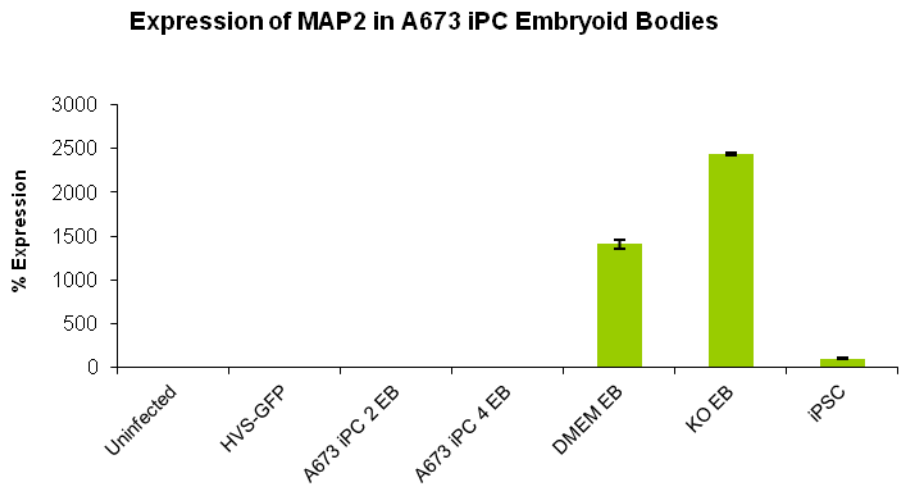
A)


Expression of Endogenous Oct4 in A673 iPC Embryoid Bodies
% Expression
120
100
80
60
40
20
0
Uninfected
HVS-GFP
A673 iPC 2 EB
A673 iPC 4 EB
iPSC EB DMEM
iPSC EB KO
iPSC

B)


Expression of MSX1 in A673 iPC Embryoid Bodies
% Expression
4000
3500
3000
2500
2000
1500
1000
500
0
Uninfected
HVS-GFP
A673 iPC 2 EB
A673 iPC 4 EB
iPSC EB DMEM
iPSC EB KO
iPSC

C)


Expression of MAP2 in A673 iPC Embryoid Bodies
% Expression
3000
2500
2000
1500
1000
500
0
Uninfected
HVS-GFP
A673 iPC 2 EB
A673 iPC 4 EB
DMEM EB
KO EB
iPSC

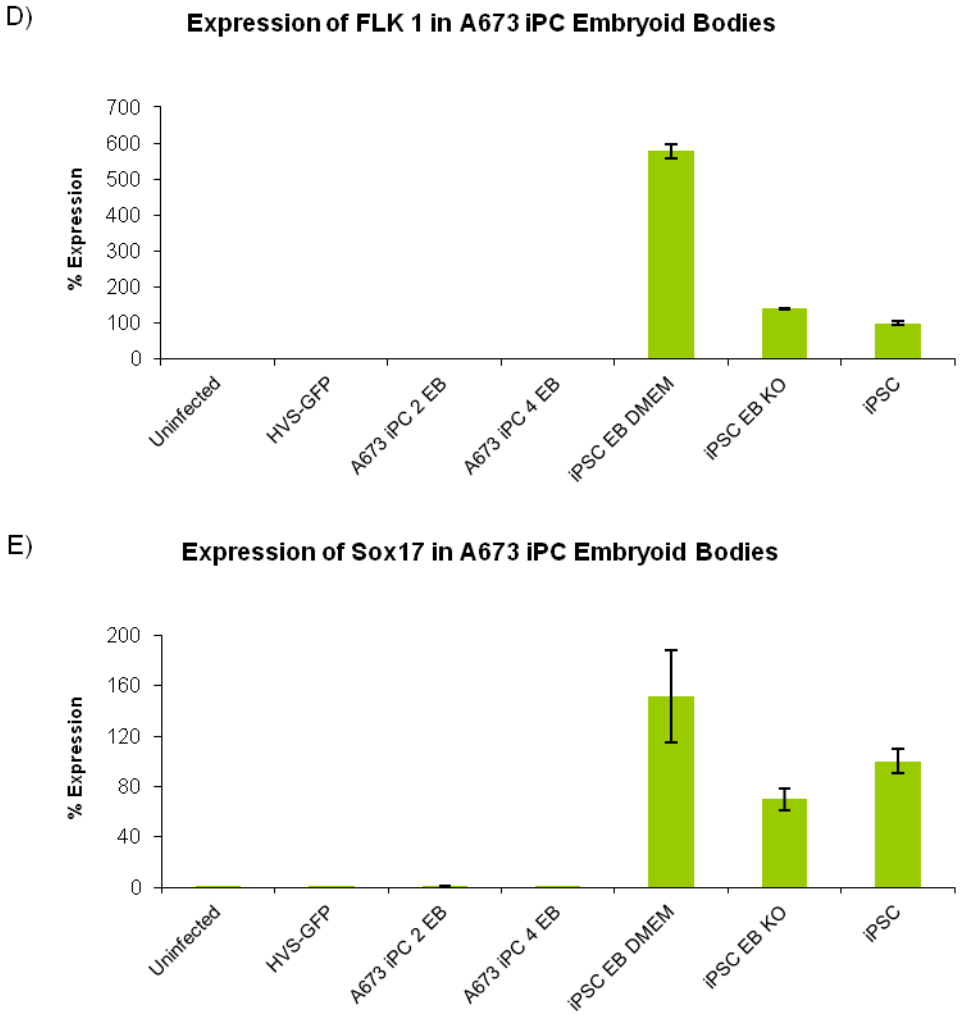


Figure 5.9: Expression of marker genes specific for the three germlayers in A673-iPC Embryoid Bodies
A) Oct4, a marker for pluripotency, is downregulated in A673-iPC EBs and EBs formed from iPSCs; B) The ectodermal marker MSX1 is upregulated in A673-iPC EB cells compared to uninfected A673 and HVS-GFP infected cells; C) Expression of MAP2, another ectodermal marker, is unaltered in A673-iPC colonies compared to negative control cells; D) Flk1, a mesodermal marker gene, and E) Sox17, a marker for endoderm are not upregulated in A673-iPC EBs compared to negative controls.

However, when A673-iPC EBs were grown in NSC medium, there appeared to be neural crests and neuronal networks forming (Figure 5.11 G). Results therefore suggest that the A673-iPC cells are capable of neural differentiation.

5.3.4 qRT-PCR on Directed Neural Lineage Differentiated Embryoid Bodies

During the differentiation of ESCs along the neuronal lineage, there are distinct stages marked by the changing expression levels of specific genes; for example, Nestin and β-III-Tubulin. Nestin is a marker for Neural Progenitor Cells and β-III-Tubulin is a marker for differentiated neurons(Abranches *et al.*, 2006). Therefore, to determine if the morphologies observed in Figure 5.11 truly indicated the differentiation of A673-iPC colonies down a neural lineage, RNA was extracted from differentiated A673-iPC or iPSC cells in addition to mock and control infected A673 cells, and qRT-PCR performed to assess Nestin and β-III-Tubulin expression.

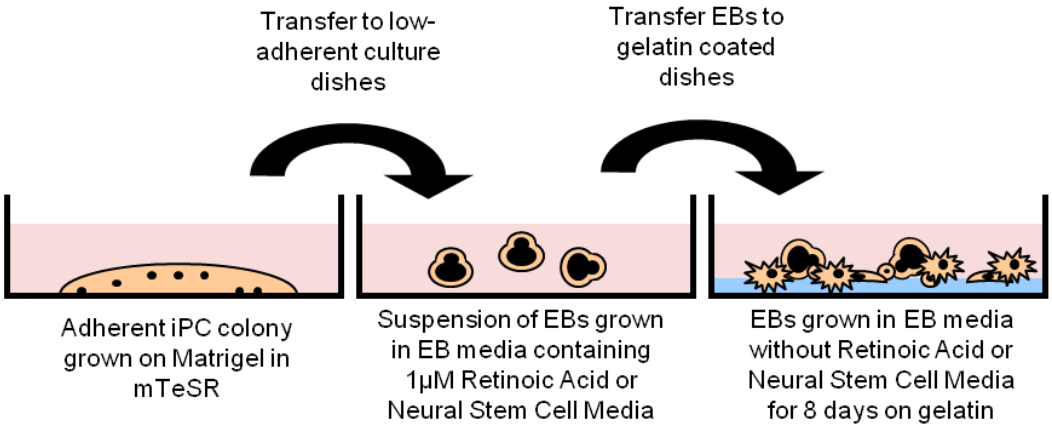


Figure 5.10: Overview of Directed Neural Differentiation of iPSCs and A673-iPCs

Embryoid bodies were formed by splitting the adherent colonies and transferring the cells to low-adherence tissue culture dishes. These suspension cultures of EBs were grown in either EB medium containing 1 μM Retinoic Acid, or Neural Stem Cell (NSC) medium for 2 days, before transferring the EBs to gelatin pre-coated dishes. Cells were cultured for a further 8 days in either EB medium without retinoic acid or NSC medium, and cells possessing neural morphologies were observed.

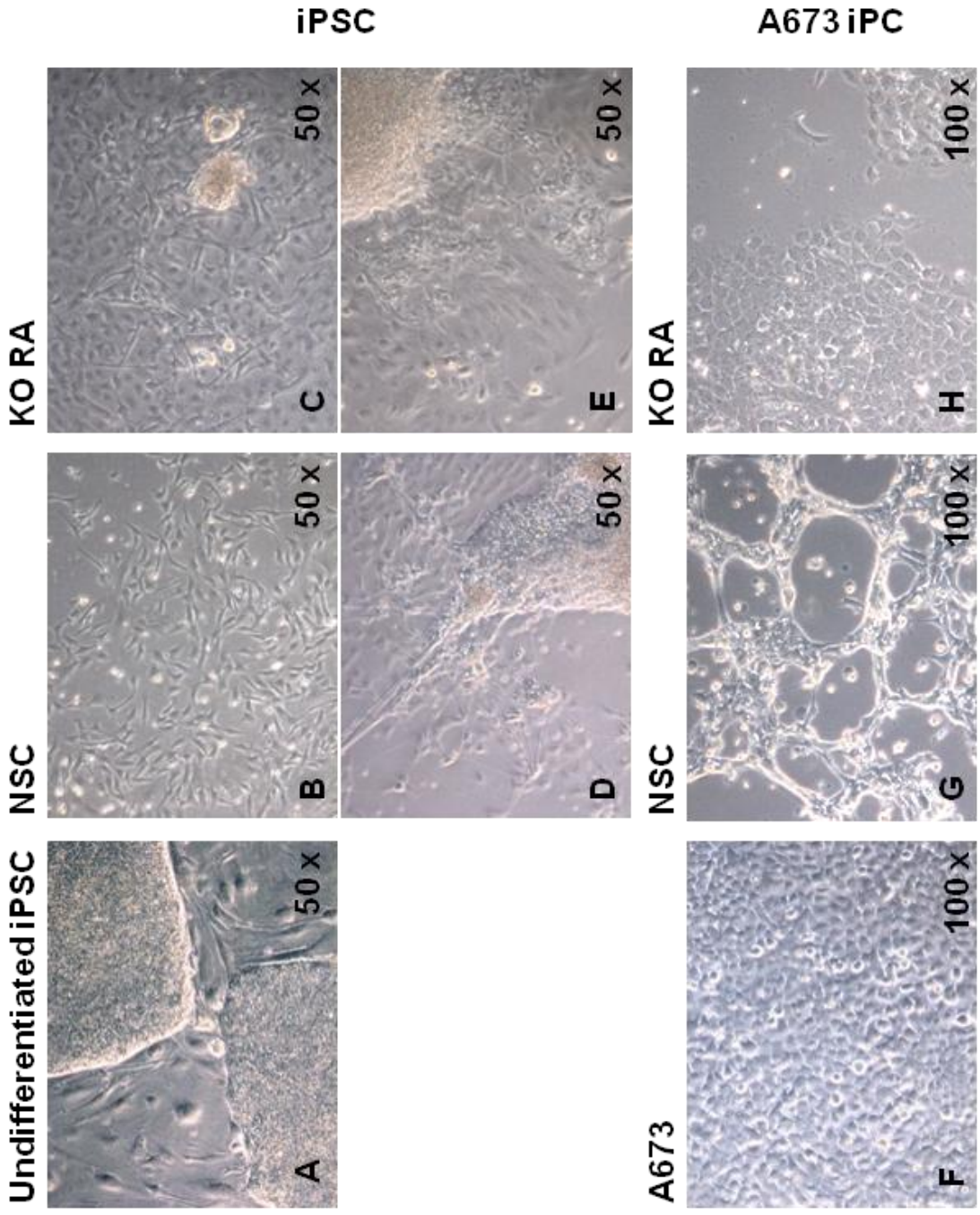
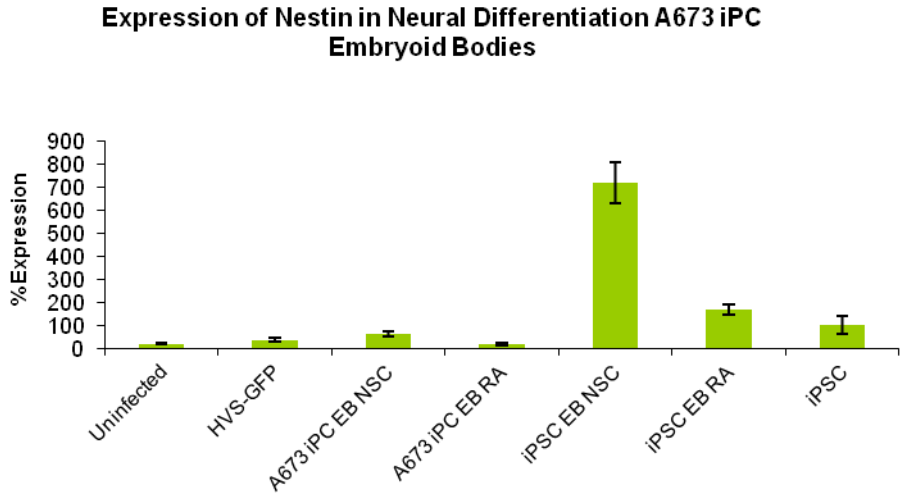


Figure 5.11: Directed Neural Differentiation of iPSCs and A673-iPCs

A) Undifferentiated iPSC cells; B and D) Differentiated iPSCs cultured in NSC medium displaying neuroprogenitor and neural dendrite morphologies respectively; C and E) Differentiated iPSCs cultured in EB medium and retinoic acid (KO RA) display neural networks and neural crests respectively. F) Uninfected A673 control cells cultured in DMEM with 10% serum; G) A673-iPC cells cultured in NSC medium display neural networks and dendritic morphologies; H) A673-iPC cells cultured in KO RA possess morphologies very similar to the uninfected control A673 cells.

Figure 5.12 A shows Nestin mRNA levels are low in undifferentiated iPSCs or uninfected and HVS-GFP infected A673 controls. In contrast, an increase in Nestin expression is observed in iPSC EBs cultured in NSC or KO RA. Slight upregulation of Nestin mRNA levels was also observed in A673-iPC EBs cultured in NSC medium, but not in cells cultured in KO RA medium. β-III-Tubulin expression in the undifferentiated and differentiated iPSC control cells also reflects the results observed for Nestin expression, with the differentiated iPSCs demonstrating elevated β-Tubulin mRNA in comparison to undifferentiated iPSCs (Figure 5.11 B). However, β-III-Tubulin mRNA levels were low in A673-iPC EB colonies cultured in NSC or KO RA medium, with surprisingly high levels observed in the control HVS-GFP infected A673 cells. These results suggest some degree of neural differentiation is occurring in the A673-iPC embryoid bodies, but not to the extent of generating fully differentiated neurons.

A)



B)

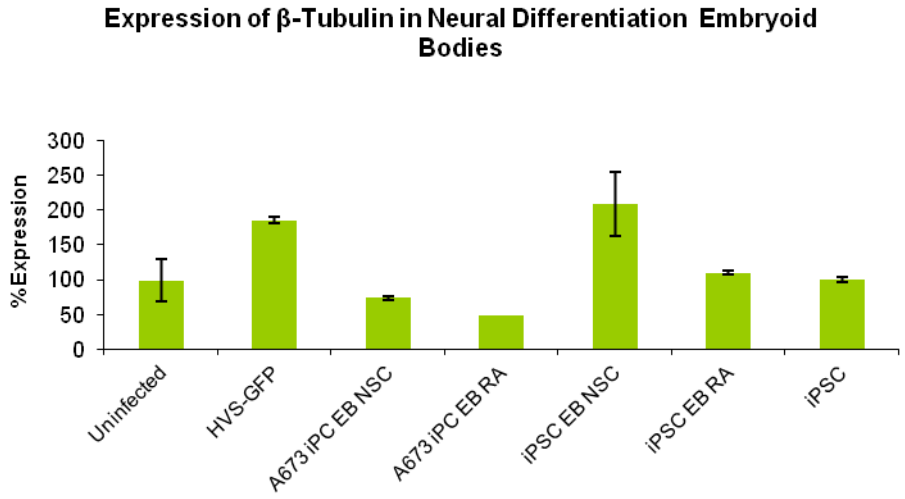


Figure 5.12: Expression of the neural markers Nestin and β-Tubulin in Neural-specific differentiation of iPSCs and A673-iPCs

A) Expression of Nestin is upregulated in A673-iPC EBs and iPSC EBs cultured in NSC medium; B) Slight upregulation of β-Tubulin is observed in iPSC EBs cultured in NSC medium, but the uninfected and HVS-GFP infected A673 control cells display higher levels of β-Tubulin than the A673-iPC EBs.

5.4 Discussion

Chapter 4 detailed the generation of ESFT iPC colonies from the A673 ESFT cell line. However, it was apparent that although there were similarities between these iPCs and ESCs, they were not entirely ESC-like, and their failure to thrive in feeder-free stem cell culture conditions indicated a possible incomplete reprogramming of these cells. Therefore, this chapter aimed to determine the extent of reprogramming which had occurred in the A673-iPC cells, and to characterise their stem cell-like properties.

Initial screening of stem cell-like characteristics was performed here, demonstrating elevated alkaline phosphatase activity, upregulation of ESC-related marker genes Oct4, Rex1 and Klf4 and increased expression of the surface ESC-marker protein, SSEA4. All these results are indicative of reprogramming occurring within these cells. The results obtained for qRT-PCR analysis of the A673-iPC gene expression (Figure 5.2 A) show Oct4 expression was upregulated in the A673-iPC Colonies 2 and 3. Oct4 is a key regulator of pluripotency and binds to a 5′-ATTTGCAT consensus sequence in complexes with various proteins, including Sox2. These complexes allow Oct4 to upregulate certain pluripotency associated genes including Oct4, Sox2, Nanog and FGF4, and repress genes associated with differentiation (Cheong *et al.*, 2011; Pardo *et al.*, 2010; van den Berg *et al.*, 2010). Therefore, upregulation of Oct4 would indicate activation of the pluripotency machinery in the A673-iPC colonies. The results obtained for qRT-PCR analysis of A673-iPC colonies show elevated mRNA levels of the pluripotency regulator Oct4, in addition to two other pluripotency associated marker genes, Rex 1 and Klf4.. These genes are both C2H2 Zinc Finger-containing proteins involved in repressing differentiation, either by Rex 1 mediated inhibition of Notch and STAT3 signalling (Bhandari *et al.*, 2010), or by regulation of Nanog expression by Klf4 (Zhang *et al.*, 2010). The elevated mRNA levels of these two genes further confirm the activation of pluripotency mechanisms within the A673-iPC colonies. Interestingly, the expression of Oct4, Rex1 and Klf4 were consistently higher in A673-iPC Colony 3, despite both colonies being transduced with the same combination of HVS-iPSC viruses (OL). This may indicate that Colony 3 has been

infected with different amounts of the HVS-Oct4 and HVS-Lin28 viral vectors resulting in a more favourable stoichiometry of these two reprogramming factors.

Interestingly, a marked decrease in hTERT mRNA levels was observed in the A673-iPC colonies compared to uninfected and HVS-GFP infected A673 cells and iPSCs. hTERT is an enzyme involved in extending telomeres. Telomeres are gradually shortened upon repeated rounds of cellular division in adult stem cells or differentiated cells until they reach their Hayflick limit, and enter replicative senescence. Some cells may bypass this senescence and will continue dividing until they reach a second growth arrest stage, termed crisis. At this stage, chromosome ends may join together by chromosome bridge-break fusion; a process which usually leads to apoptosis. However, some cells may develop a mutation which upregulates their telomerase activity and enables them to produce stable telomeres and divide indefinitely (Shay & Wright, 2010). CSCs however, are more quiescent and therefore potentially express hTERT at lower levels than their more rapidly dividing progeny (Shay & Wright, 2010). The observed reduction in telomerase activity in the reprogrammed A673-iPC cells may indicate these cells are becoming more similar to CSCs. Furthermore, this observed reduction in telomerase may explain why so many of the A673-iPC colonies that originally formed failed to thrive. Inhibition of telomerase activity with the inhibitor, Imetelstat, causes CSCs to senesce (Djojosubroto *et al.*, 2005; Hochreiter *et al.*, 2006), and if the reprogramming process is also reducing hTERT levels, this may have the same effect in these cells.

Immunofluorescence results for SSEA4 expression showed that both the A673-iPC colonies were SSEA4 positive, indicating some degree of reprogramming occurring in these cells (Figure 5.3). However, some of the cells derived from the same starting colony appeared to have reduced SSEA4 expression. This was demonstrated for A673-iPC Colony 3, whereby Colony 3a expressed more SSEA4 than Colony 3b, despite both colonies being derived from the same original A673-iPC colony. Colony 3a also maintained the HVS-iPSC episomes better than Colony 3b in the absence of hygromycin selection, as demonstrated by GFP expression within these

cells. This highlights the possibility that some of the SSEA4 expression observed in the A673-iPC colonies may be as a result of remaining exogenous expression of the reprogramming factors. Because of the incomplete reprogramming of the A673-iPC cells, removal of exogenous iPSC gene expression by eliminating the HVS-iPSC viral episomes may cause the cells to revert back to their parental cell type, thereby losing SSEA4 expression.

Interestingly, upregulation of ESC-related genes and SSEA4 was greater in A673-iPC Colony 3, despite both colonies being transduced with the same combination of HVS-iPSC viruses (OL). This may indicate that Colony 3 has been infected with different amounts of the HVS-Oct4 and HVS-Lin28 viral vectors resulting in a more favourable stoichiometry of these two reprogramming factors.

Future characterisation of A673-iPCs could include bisulphate sequencing to analyse the methylation status of pluripotency-associated gene promoters, as performed by Carette and colleagues in reprogrammed Chronic Myeloid Leukaemia cells (Carette *et al.*, 2010). Furthermore, it would be interesting to perform micro array analysis comparing A673-iPC colonies to parental A673 cells and ESCs, specifically looking at ESC genes, lineage-specific genes and cancer-associated genes. Lin et al. performed similar microarray analysis on iPCs generated from melanoma cells and interestingly found these cells possessed gene expression profiles that more closely resembled the ESC cell lines H1 ad H9 (89% and 86%, respectively) than their parental melanoma cells (53%) (Lin *et al.*, 2008). Whole genome analysis such as this would also help to identify genes within CSCs which are not present in normal ESCs and would therefore make potential attractive anti-CSC drug targets.

The second aim of this chapter was to address the differentiation potential of the generated A673-iPC colonies, to fully assess their potential as a model of ESFT development. The results of non-specific differentiation of A673-iPC cells (Figures 5.5 to 5.8) demonstrated firstly, that differentiation of the iPCs can occur and secondly, the differentiated cells predominantly displayed morphologies which closely resembled that of their parental somatic A673

cells. This indicates these cells possess a somatic memory causing a bias in their differentiation, which has been reported for iPSCs (Bar-Nur *et al.*, 2011; Ghosh *et al.*, 2010; Kim *et al.*, 2010; Marchetto *et al.*, 2009). Furthermore, if the A673-iPC cells were incompletely reprogrammed, as suggested in Chapter 4, the loss of the HVS-iPSC viral episomes from these cells would allow them to revert to their parental cell type.

Further indication of the A673-iPC somatic cell memory was shown from the other predominant morphologies displayed by these cells, which were fibroblastic and neural. ESFT cells have been shown to express many of the genes related to the neuro-ectodermal (Staege *et al.*, 2004) and mesenchymal (Suvà *et al.*, 2009) lineages. Therefore the differentiation bias of A673-iPC cells to these lineages indicates that expression of these lineage-associated genes remains in the A673-iPC cells, thereby driving their differentiation down these specific lineages. In addition, results from the qRT-PCR analysis of EB marker genes in differentiated A673-iPC EBs provided further evidence for somatic cell memory in these cells (Figure 5.9), whereby only the ectodermal marker MSX1 was upregulated in non-specifically differentiated A673 cells. Future work would involve immunostaining the differentiated progeny of the A673-iPC EBs with markers for various different lineages, particularly Vitronectin and Tuj1 which are specific markers for fibroblasts and neuronal cells respectively. This would verify that the A673-iPC cells were capable of differentiating down mesenchymal and neuroectodermal lineages. Furthermore, as mesenchymal stem cells are implicated as the cells of origin for ESFTs, it would be interesting to stain for the presence of the mesenchymal stem cell marker, CD44, in parental A673 cells and A673-iPCs to determine if reprogramming A673 cells to a more primitive, CSC-like state caused these cells to develop a mesenchymal stem cell phenotype.

Finally, the differentiation of A673-iPC cells down a neural lineage was attempted, as a further assessment of their ESC-like characteristics and differentiation potential. Results of the neural lineage directed differentiation show that A673-iPC cells were capable of differentiating towards the neuro-ectodermal lineage. Specifically, these cells developed neural networks

when cultured in NSC medium (Figure 5.11), and showed upregulated expression of Nestin which is a marker of Neural Progenitor Cells and is expressed in the early stages of neural differentiation (Figure 5.12). However, the upregulation of β-III-Tubulin in these cells was not observed. β-III-Tubulin is expressed in fully differentiated neurons (Abranches *et al.*, 2006). This is in contrast to Nestin which is present in earlier neural progenitors, therefore these results may indicate that the A673-iPC EBs maintain some degree of their primitive nature.

Previously, Suvà and colleagues achieved directed differentiation of ESFT CSCs isolated from primary tumours by sorting for CD133+ cells. Differentiation into three mesenchymal lineages was performed; adipogenic, osteogenic and chondrogenic (Suvà *et al.*, 2009). It would therefore be interesting to apply these differentiation protocols to the A673-iPC cells generated in this study. This would determine if the A673-iPCs would also be capable of differentiating into these lineages, thereby demonstrating the extent to which these cells resemble ESFT CSCs.

In summary, these data indicate the A673-iPC cells have acquired certain stem cell-like characteristics during the process of HVS-iPSC mediated reprogramming. However, this effect is limited; with some ESC marker gene expression levels remaining unaltered in the A673-iPSC colonies, which further indicates incomplete reprogramming.

Differentiation of the A673-iPC cells indicated a strong somatic cell memory, which appeared to limit the differentiation of these cells to ectodermal and mesenchymal lineages. But perhaps the somatic memory of the A673-iPC cells does not present a problem for the further study of tumour formation and progression of the ESFT CSCs, as the A673-iPC cells may still be capable of differentiating into the various cells found in ESFTs. Future work would concentrate on removing the HVS-iPSC episomes from reprogrammed cells, thereby removing iPSC transgene expression which may restrict A673-iPC differentiation. Once the viruses have been removed, the A673-iPC cells would represent a more accurate model of ESFT CSC differentiation.

Chapter 6
Discussion

6 Discussion

The advent of iPSC technology generated immense scientific interest due to its potential to produce large quantities of pluripotent stem cell-like cells for therapeutic use and scientific study, including the ability to produce a variety of disease-in-a-dish models. However, currently retroviral gene delivery vectors are the predominant reprogramming technology used in iPSC generation. These viruses integrate into the host cell, and would therefore disrupt the normal gene profiles of resulting iPSCs a feature which severely limits the value of these cells for therapeutic use or as disease models.

Therefore, this thesis has investigated the potential of Herpesvirus saimiri as a non-integrating, episomally maintained viral vector for iPSC-generation. Three HVS-iPSC gene delivery vectors have successfully been produced, expressing Oct4, Lin28 and Nanog. Their potential as iPSC-gene delivery vectors has been assessed through proof-of-concept experiments to reprogramme neural stem cells (NSCs), or somatic ESFT cancer cells to pluripotent cells.

6.1 HVS as a Gene Delivery Vector for iPSC technology

As described previously, HVS possesses a number of features which makes this virus an attractive gene delivery vector for iPSC generation. Firstly, it has a broad cell tropism, since it has been shown to infect a wide range of primary and cancerous tissues (Griffiths *et al.*, 2006). Secondly, all HVS-gene delivery vectors are based on the A11-S4 virus strain, which contains a natural deletion encompassing the HVS transforming protein, STP. Thirdly, HVS, like other herpesviruses, has a large dsDNA genome which gives it a large capacity for transgenes. Indeed, the generation of the HVS amplicon system allows up to 50kb of exogenous DNA to be inserted into the HVS genome (Macnab *et al.*, 2008). This large capacity potentially enables the cloning of complete genomic loci into HVS-based gene delivery vectors, including regulatory sequences, introns and native promoter elements. Cloning complete genomic loci would ensure prolonged transgene expression without transcriptional silencing, as has been previously demonstrated (Wade-Martins *et al.*, 2000). Moreover, it would also prevent

transgene overexpression, which is a feature often associated with cDNA expression constructs. These features would be particularly beneficial in reprogramming, where overexpression of iPSC transgenes is potentially oncogenic, and also reduces the efficiency of reprogramming. Furthermore, this large capacity would enable the creation of a single gene delivery vector which would encompass all the factors required for efficient iPSC generation. Previous attempts to generate single vectors containing multiple reprogramming factors have utilised polycistronic constructs which contain each reprogramming gene separated by self-cleaving 2A peptides from the foot and mouth disease virus (FMDV) (Carey *et al.*, 2009; Ryan & Drew, 1994). Although these polycistronic constructs improve transduction or transfection efficiencies, the reprogramming efficiencies of these methods remain poor. This may be as a result of incomplete translation and poor cleavage.

Finally, HVS possesses a further attractive feature as an iPSC-gene delivery vector, which is unique to the *gammaherpesvirinae*. It is capable of persisting as an episome in dividing cell populations by means of a virally encoded protein, ORF73 (Calderwood *et al.*, 2005). The potential of episomally maintained gene delivery methods for iPSC generation has already been demonstrated by Yu et al. iPSC cells were successfully generated without disrupting the host cell genome using a plasmid vector containing a polycistronic construct expressing the reprogramming factors, Oct4, Sox2, Nanog, Lin28, Klf4, c-Myc and SV40LT, along with the EBV episomal maintenance elements EBNA1 and OriP (Yu *et al.*, 2009). Although originally this episomal reprogramming method had poor efficiency, further research dramatically increased iPSC generation efficiency by combining this episomal vector with basic Fibroblast Growth Factor (bFGF) and Leukaemia Inhibitor Factor (LIF), in addition to small molecule inhibitors of MEK, GSK3β and TGFβ. These culture conditions enhanced reprogramming efficiency from 0.0001% described for the original reprogramming method, to 0.1% (Yu *et al.*, 2011). Therefore, in future applications, the efficiency of the HVS-based episomal reprogramming method may also be improved by reprogramming cells in these culture conditions.

Removal of reprogramming transgenes upon iPSC generation produces cells which more closely resemble true ESCs in their gene expression profiles. This is a major advantage to future iPSC cells for therapeutic uses and as models of development (Soldner *et al.*, 2009; Sommer *et al.*, 2010). Therefore, reprogramming vectors are often silenced upon iPSC generation by incorporating inducible expression systems, such as the tetracycline inducible systems. These allow tight repression of iPSC transgene expression in the absence of tetracycline or doxycycline. However, although these systems remove transgene expression, the exogenous vector DNA remains in the reprogrammed cells. Other methods to remove the exogenous vector DNA and transgene expression upon reprogramming have included excisable gene delivery vectors such as a Cre-LoxP excisable retrovirus, and the PiggyBac Transposon (Kaji *et al.*, 2009; Soldner *et al.*, 2009; Woltjen *et al.*, 2009; Yusa *et al.*, 2009). However, these systems both required stringent screening of the resulting iPSC cells to ensure all exogenous DNA was removed. Furthermore, in the case of the excisable retrovirus, a viral 'footprint' remained in the cells, which could potentially disrupt normal host gene function.

Once again, episomally maintained vectors have an advantage in this regard, as the episome could easily be removed upon reprogramming. As such, iPSC cells generated using episomal vectors would not require the stringent screening that excisable reprogramming vectors required. Indeed, Yu and colleagues demonstrated their episomal plasmid reprogramming system could be removed through repeated cellular divisions in the absence of selective antibiotic (Yu *et al.*, 2009). Thus, these cells were free of any exogenous DNA and ectopic transgene expression. Therefore, a further favourable characteristic of HVS-based gene delivery vectors for iPSC generation would be the potential to remove the virus from reprogrammed cells. Because HVS episomal persistence is reliant on a single virally encoded protein, it would be feasible to efficiently remove the viral episome by regulation of ORF73 expression. Future work would investigate three possible ways to achieve this; siRNAs targeted against ORF73, generation of a tetracycline-inducible ORF73 HVS vector or creating a virus containing

FLP recombination sites flanking the ORF73 gene, thereby creating an excisable ORF73 cassette.

Using siRNA to knockdown ORF73 expression would require multiple rounds of transfection as siRNA is unstable and readily degraded within cells. In addition, a knockdown of nearly 100% would be required as any remaining episomes would be capable of expressing ORF73, and would therefore persist after siRNA treatment. Since most siRNAs do not achieve 100% knockdown this may not be a feasible approach.

A better way to control ORF73 expression may be to use a tetracycline inducible system (Gossen & Bujard, 1992b). These systems enable tight gene repression in the absence of tetracycline or doxycycline which is alleviated upon addition of either of these drugs. Strong expression of the desired transgene is provided by a CMV promoter located downstream of seven copies of the Tet-O operon which provide tight repression. By replacing the ORF73 endogenous promoter with this Tet-inducible promoter would enable ORF73 expression and episomal persistence upon treatment with tetracycline. Upon withdrawal of the drug, ORF73 expression would be tightly repressed and the viral episome would be lost after repeated rounds of cellular division. However, expression from ORFs 71 – 73 is driven by a common promoter located downstream of ORF73. ORF71 and ORF72 encode vFLIP and a cyclin D homologue, respectively (Chang *et al.*, 1996; Thome *et al.*, 1997), but are not essential for HVS latency (Collins *et al.*, 2002). However, it was observed in ΔORF72 mutants that viral titres were significantly lower that wild-type HVS (unpublished observation). Therefore, although this strategy has significant improvements over siRNA knockdown of ORF73, there are still a few issues with this approach.

Perhaps the most feasible method of producing removable HVS-based vectors is the generation of a HVS vector containing a Cre-excisable ORF73 cassette. A cassette containing ORFs 71-73 and their common promoter could be produced with flanking LoxP sites and recombined using site-specific recombination within the HVS-BAC. This would produce an HVS-based vector capable of producing viral titres similar to those of HVS-GFP, and upon successful reprogramming of target cells, could be removed by

treating the resulting iPSCs with cell-permeable Cre-recombinase. This would remove the ORF71-73 expression cassette from episomes, therefore removing their ability to persist in dividing cell populations.

6.2 Reprogramming Neural Stem Cells

Perhaps the most promising application of iPSC technology is the study and treatment of neurodegenerative disorders. This is mainly due to the difficulties associated with isolating adult neural stem cells from living patients and the limitations of animal models in neurodegenerative disease modelling.

Therefore, the HVS-iPSC vectors were utilised in a proof-of-principle experiment to generate iPSC cells from Neural Stem Cells (NSCs). However, attempts to reprogramme these cells highlighted the need for further work to improve the biosafety of these HVS-iPSC vectors. Infection of NSCs with HVS-GFP caused extreme cytopathic effects in these cells, for which there are two possible explanations. Firstly, the virus-containing supernatant used to transduce cells is harvested directly from lysed cell sheets of the permissive cell line, OMK cells. These supernatants could therefore contain various cytokines and interferons which are released upon viral infection, leading to induction of apoptosis in subsequently infected cells. Purified virus stock could be produced in future applications, which would circumvent this issue. The second explanation for the observed cytotoxicity may be due to lytic expression of viral genes in NSC cells. Although HVS infection predominantly leads to a latent infection in human cells, it is capable of lytic reactivation in certain cell lines, as observed previously in the lung cancer cell line, A549, and in SW480 cells (Goodwin *et al.*, 2001; Smith *et al.*, 2001). Therefore HVS could potentially be capable of lytic replication and infectious virus production in NSCs. To further improve the biosafety of HVS-based reprogramming vectors, future vectors would be generated using the HVS-amplicon-like system (Macnab *et al.*, 2008). The HVS-amplicon has improved biosafety as it lacks the majority of the HVS genome, which removes the ability these recombinant viruses to undergo lytic replication. Significantly, it does not encode the major lytic activation

genes, ORF50 and ORF57. Therefore, future work would involve generating HVS-iPSC gene delivery constructs based on the HVS-amplicon system.

6.3 Reprogramming of ESFT A673 cells

The presence of tumour-initiating Cancer Stem Cells (CSCs) has already been demonstrated for a rapidly increasing number of cancers. Due to their capacity for self-renewal and their ability to seed new tumours, CSCs have important implications in cancer progression and metastasis.

However, the dynamic nature of these cells means they remain difficult to identify and isolate. The use of iPSC technology to generate induced pluripotent cancer stem cells (iPCs) would potentially enable vast quantities of pure CSC populations to be obtained for study. To this end, this thesis examined the potential of HVS-based iPSC gene delivery vectors to reprogramme cancerous cells.

This investigation reports the generation of iPC-like colonies from the A673 ESFT cell line using the HVS-iPSC recombinant viral vectors. Infection with HVS-Oct4 virus alone was capable of generating A673-iPC colonies, but was not sufficient to reprogramme these cells to a state whereby they were able to thrive in feeder-free stem cell culture conditions. The efficiency of reprogramming was improved by transducing the cells with HVS-Oct4 in conjunction with HVS-Nanog, but once again the colonies produced were not viable in stem cell culture conditions.

However, transduction of A673 cells with HVS-Oct4 and HVS-Lin28 viruses, or a combination of all three HVS-iPSC, resulted in greater number of A673-iPC colonies, some of which were stable upon prolonged culture in feeder-free stem cell culture conditions. Analysis of these colonies by screening for alkaline phosphatase activity, qRT-PCR analysis of ESC marker genes and staining for SSEA4 antigen expression demonstrated that these cells exhibited some hallmarks of iPCs, but were probably in a state of incomplete reprogramming. Specifically, these cells expressed alkaline phosphatase, but only displayed increased mRNA levels for a limited number of ESC markers. Additionally, these genes were not expressed to the same levels observed in iPSC control cells. Moreover, significant variation in the resulting

A673-iPC colonies was observed, as demonstrated by the survival of some colonies in feeder-free stem cell culture conditions and differences in ESC-marker gene expression. Perhaps one explanation for this variability could be the inherent aneuploidy and chromosome instability of the parental cancer cells. Aneuploidy is either duplication or removal of chromosomes within cells as a result of incorrect cell division. Cancer cells are frequently aneuploid due to high cell division rates combined with defects in cell-division check points. Because of this, A673 cells may not have a full complement of genes or gene duplications which could affect the reprogramming efficiency, resulting in some cells achieving a more ESC-like state than other cells. This could explain why some of the resulting A673-iPC colonies deformed after prolonged culture.

The partial reprogramming observed within the A673-iPC colonies has been reported previously in the generation of iPCs from cancer cell lines. Hochedlinger and colleagues attempted to generate iPCs from melanoma, leukaemia and lymphoma, by Somatic-cell Nuclear Transfer (SCNT) with enucleated oocytes. However, only cells containing nuclei from the melanoma cancer cells were successfully reprogrammed (Hochedlinger *et al.*, 2004). This evidence hints at some as yet unknown road-block in the reprogramming of cancer cells. This is particularly surprising given the strong links between cancer and pluripotency which have been discussed previously (Bernhardt *et al.*, 2012; Ramos-Mejia *et al.*, 2012). Perhaps the similar gene expression profile of cancers to pluripotent cells makes these cells more difficult to reprogramme. If this is the case, then more primitive cancers which more closely resemble ESCs and iPSCs would be harder to reprogramme than cancers with a less aggressive, less primitive nature.

Although the data presented here suggest that some degree of reprogramming within ESFTs is possible, this report merely details the reprogramming of an ESFT cancer cell line. However, cancer cell lines are homogenous populations of cells, and therefore not truly representative of the heterogeneous tumours they are derived from. Furthermore, cell lines are generally produced from a subpopulation of cancer cells which are better able to survive in prolonged *in vitro* culture. These cells are therefore more

likely to be highly proliferative and capable of unlimited self-renewal, and as such are likely to closely resemble ESCs and CSCs. This may make their reprogramming more difficult due to the endogenous expression of many of the pluripotency genes within these cells. Therefore, future work to investigate the true potential of HVS-iPSC viral vectors in iPC generation would involve reprogramming various cells found within primary tumour samples. Moreover, in addition to highlighting which cells are most amenable to becoming reprogrammed, this would also provide insights into which cells are theoretically capable of de-differentiating into new CSC populations.

6.4 Differentiating A673-iPC cells

One of the main interests of iPCs is as a model for the aberrant signalling pathways observed in cancer, and how these affect proliferation and differentiation of cells within the tumour microenvironment. Therefore, another aspect of this work was to determine if the A673-iPC cells generated in this study were capable of differentiation and thereby providing a suitable model for ESFT cancer progression.

Results from the non-specific differentiation of the A673-iPC cells suggested that these cells were capable of differentiation, as evidenced by the altered cell morphologies observed growing out from the attached EBs. However, it results demonstrate there is an apparent strong bias towards differentiation along the ectodermal lineage.

A possible explanation for the apparent differentiation bias towards the ectoderm could be the presence of the EWS-FLI1 fusion protein. This protein has been demonstrated to upregulate the expression of various genes involved in neural development, including neuronal pentraxin receptor, SMA5, and presenilin (Hu-Lieskovan *et al.*, 2005). However, A673-iPC cells appeared to retain their primitive nature upon differentiation down the neural lineage, as evidenced by their upregulated mRNA levels of Nestin, but low β-Tubullin expression, which is a maker of differentiated neurons (Abranches *et al.*, 2006). Again, the EWS-FLI1 fusion protein may explain this observation, as EWS-FLI1 upregulates expression of EZH2, a histone methyltransferase protein involved in suppressing genes involved in

differentiation. The upregulation of this protein by EWS-FLI1 is capable of blocking endothelial and complete neuro-ectodermal differentiation (Richter *et al.*, 2009). Therefore the EWS-FLI1 protein may prevent complete differentiation of the A673-iPC cells to fully differentiated neurons.

An additional factor which may reduce the A673-iPC cells capacity for differentiation is the residual reprogramming transgene expression from the HVS episomes which persist in these cells, such as Oct4. These genes are key regulators of pluripotency and repress differentiation of stem cells. Therefore residual expression of these genes within the A673-iPC cells would counteract their differentiation. As described previously, future development of the HVS-iPSC vectors would include implementing a system to remove the HVS episomal maintenance element, thereby allowing the removal of any transgenes and exogenous vector DNA from reprogrammed cells. This would potentially produce iPSC and iPC cells which possessed the same capacity for differentiation as true ESCs and CSCs.

6.5 Conclusions

In conclusion, this thesis has demonstrated the potential of HVS as a recombinant viral vector in iPSC generation. HVS-iPSC vectors were created, which were capable of sustained transgene delivery for Oct4, Lin28 and Nanog. Furthermore, these vectors were able to initiate reprogramming in an ESFT cell line.

The A673-iPC cells had been reprogrammed to some degree, as demonstrated by the elevated expression levels of certain ESC marker genes. However, some ESC marker genes remained unaltered or even reduced in the A673-iPC cells in comparison to uninfected and mock-infected control cells, and some colonies failed to thrive in feeder-free stem cell culture conditions. This indicates that, although reprogramming has certainly been initiated in these cells, they remain incompletely reprogrammed. Future work would involve gene arrays to compare ESC maker gene expression levels between uninfected and mock-infected A673 control cells, the A673-iPC cells, and iPSCs or ESCs. This would give a full genome-wide picture of the extent to which these cells are reprogrammed, and which genes remain unaltered. Moreover, generation of HVS-iPSC

vectors with inducible ORF73 expression will enable removal of exogenous vector DNA and transgene expression. This could produce iPCs which more closely resemble true CSCs in their gene expression profiles and differentiation capacity which will be more valuable in the study of cancer progression.

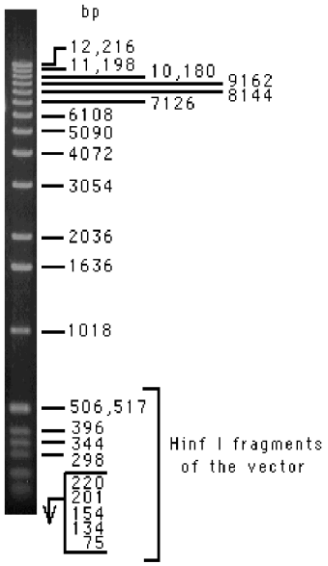
Chapter 7

Appendix

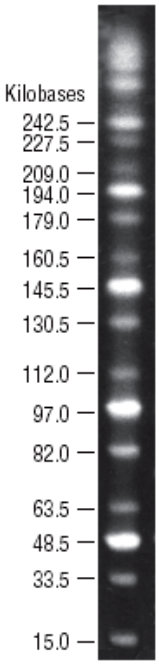
7 Appendix

7.1 DNA ladders

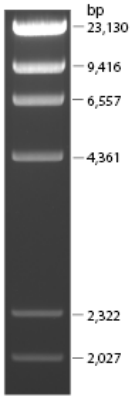
100 bp ladder (Invitrogen)



1 kb ladder (Invitrogen)

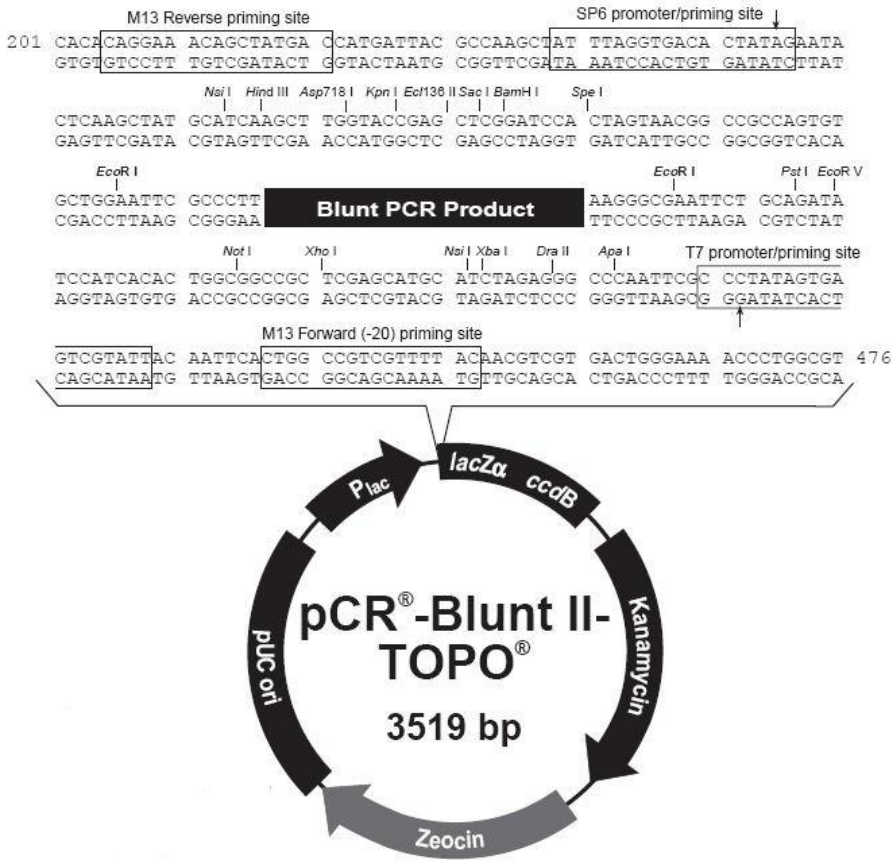


Λ DNA-HindIII digest ladder (NEB)

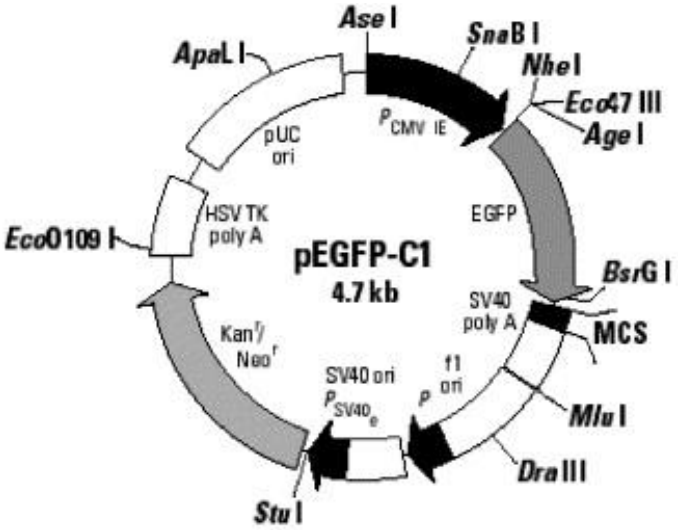


MidRange I Pulse-Field Gel Marker (NEB)

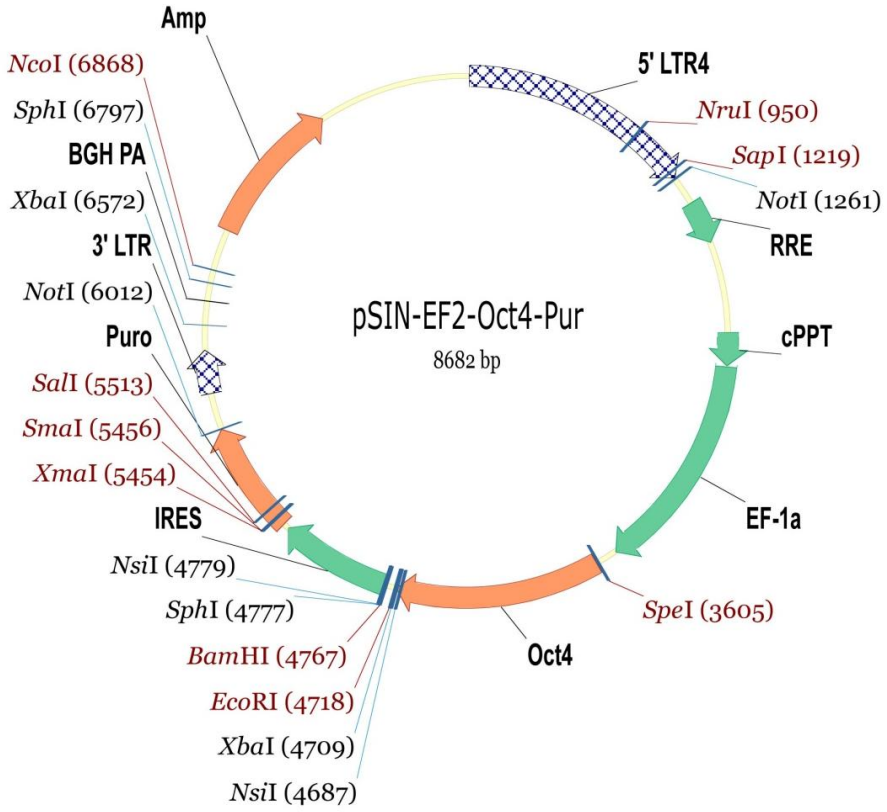
7.2 Constructs



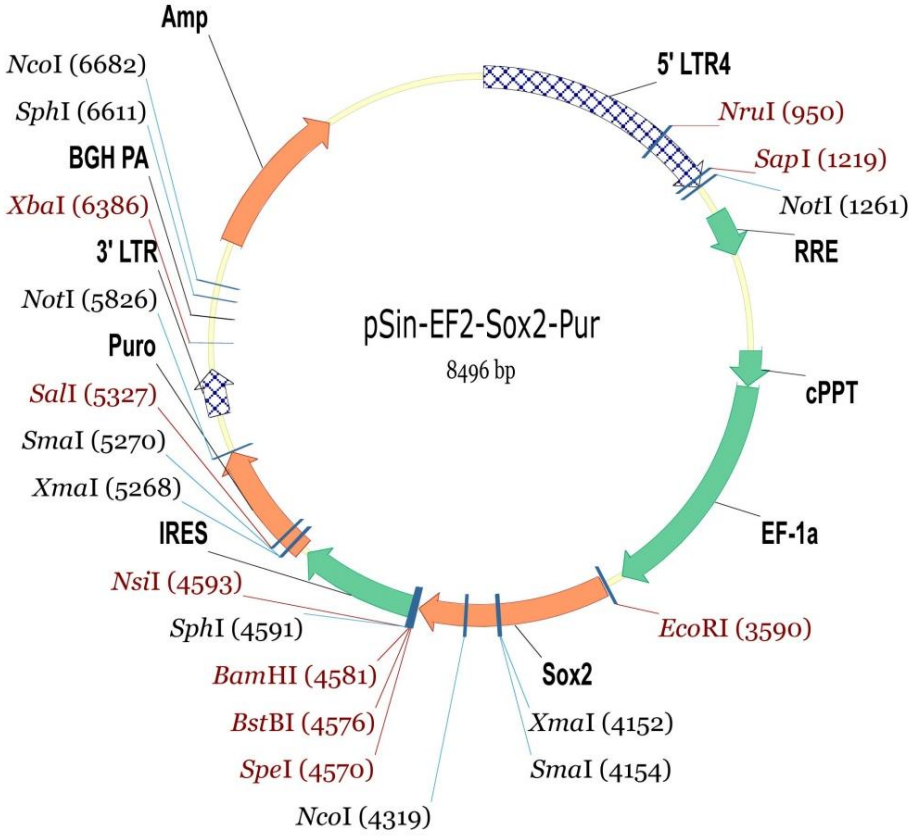
pCR®-Blunt II-TOPO® (Invitrogen)



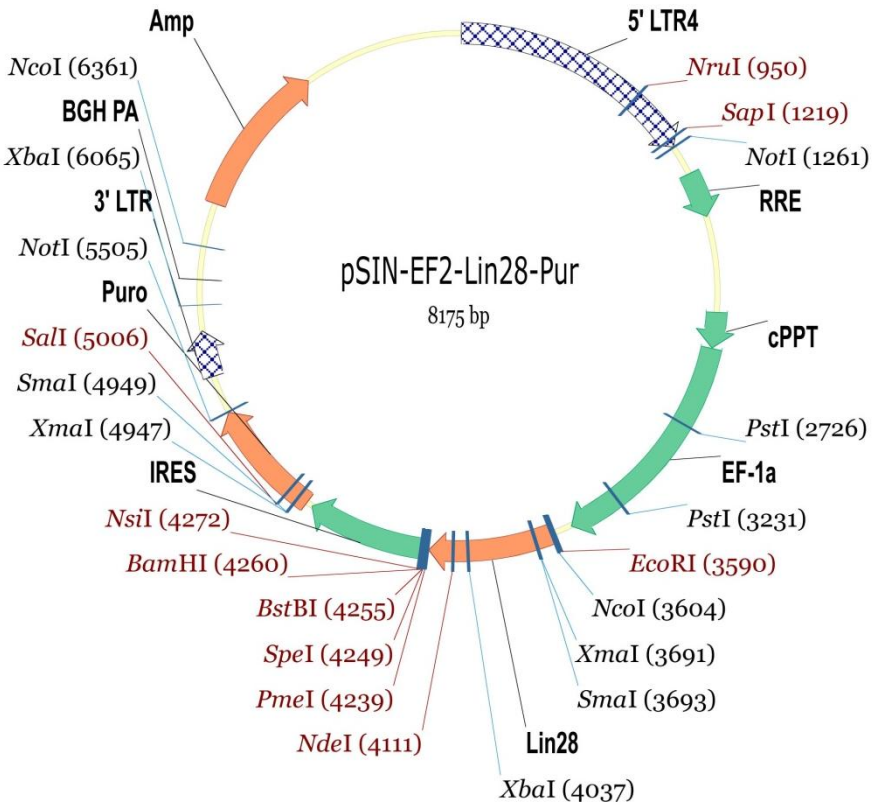
pEGFP-C1 (Clontech)



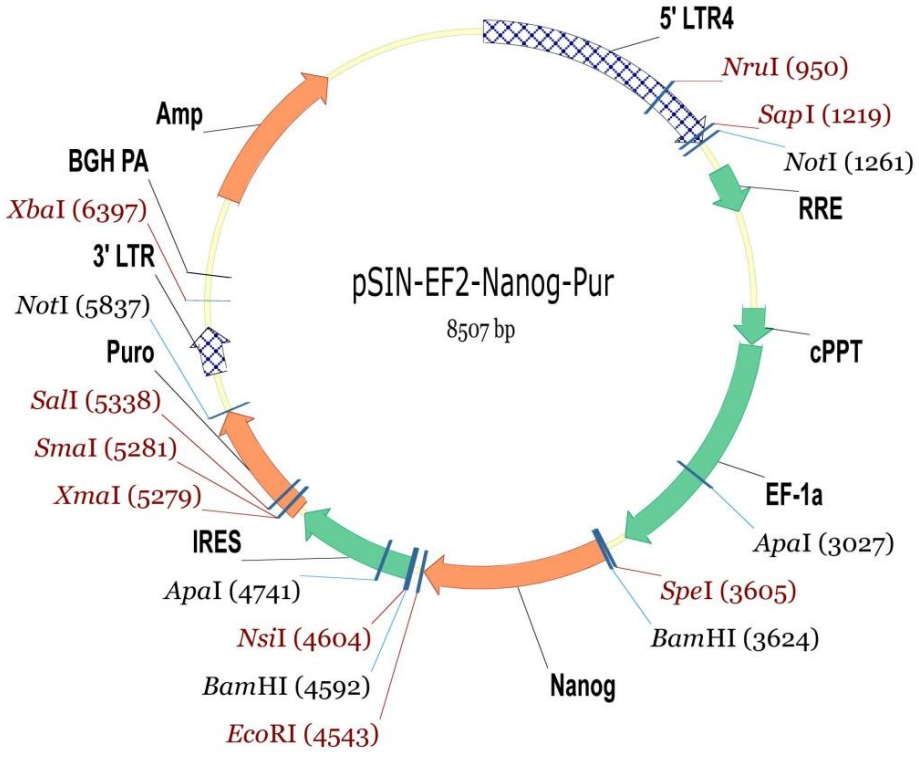
pSIN-EF2-Oct4-Pur (Addgene)



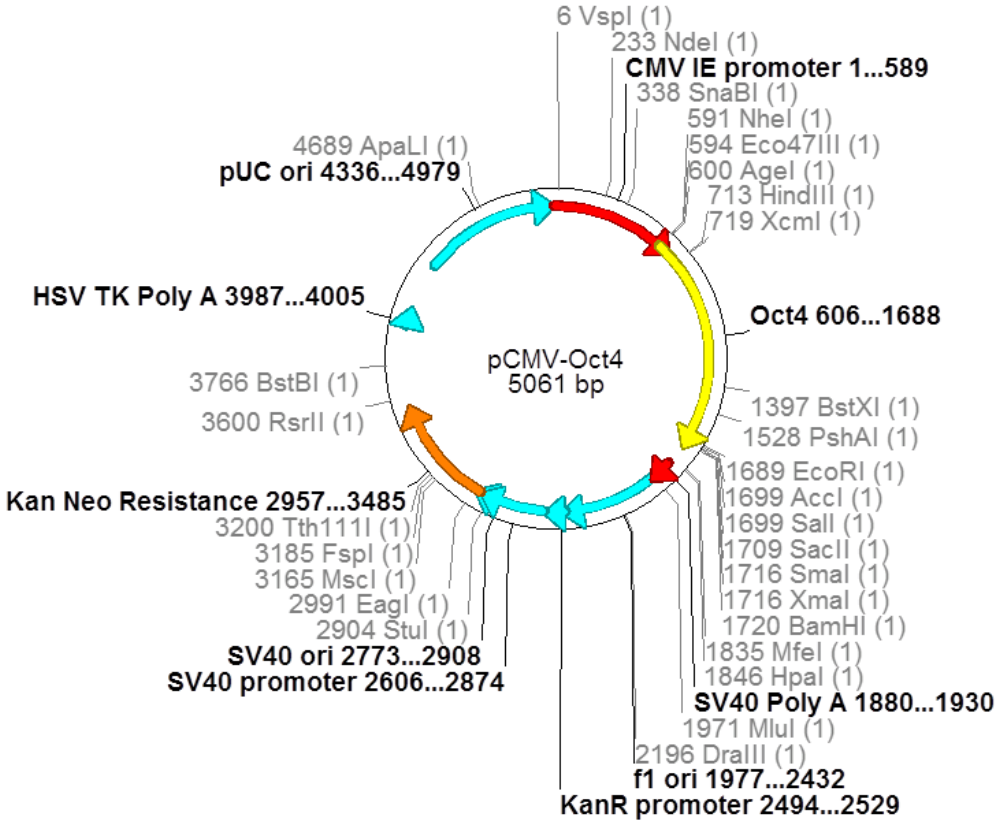
pSIN-EF2-Sox2-Pur (Addgene)



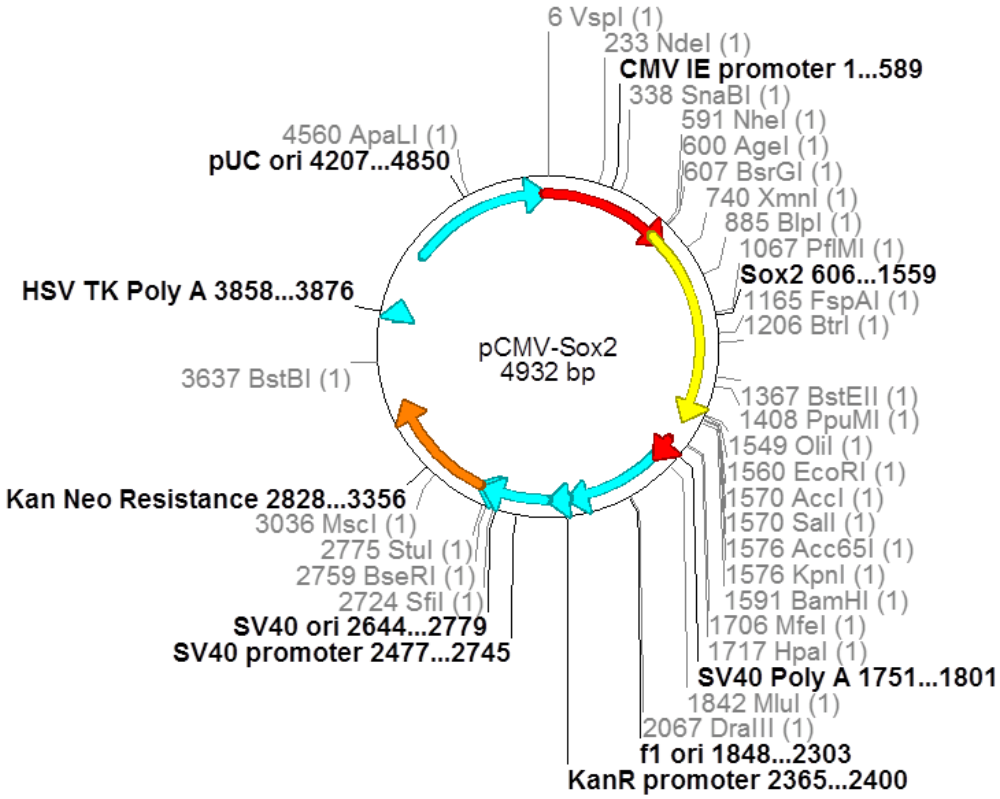
pSIN-EF2-Lin28-Pur (Addgene)



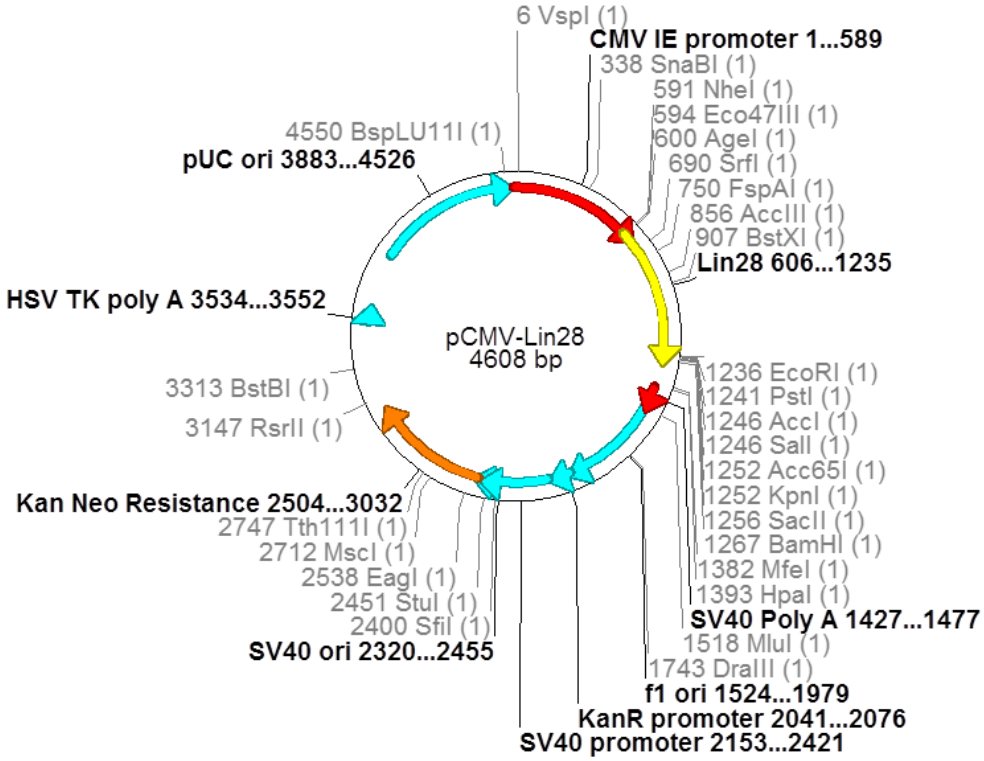
pSIN-EF2-Nanog-Pur (Addgene)



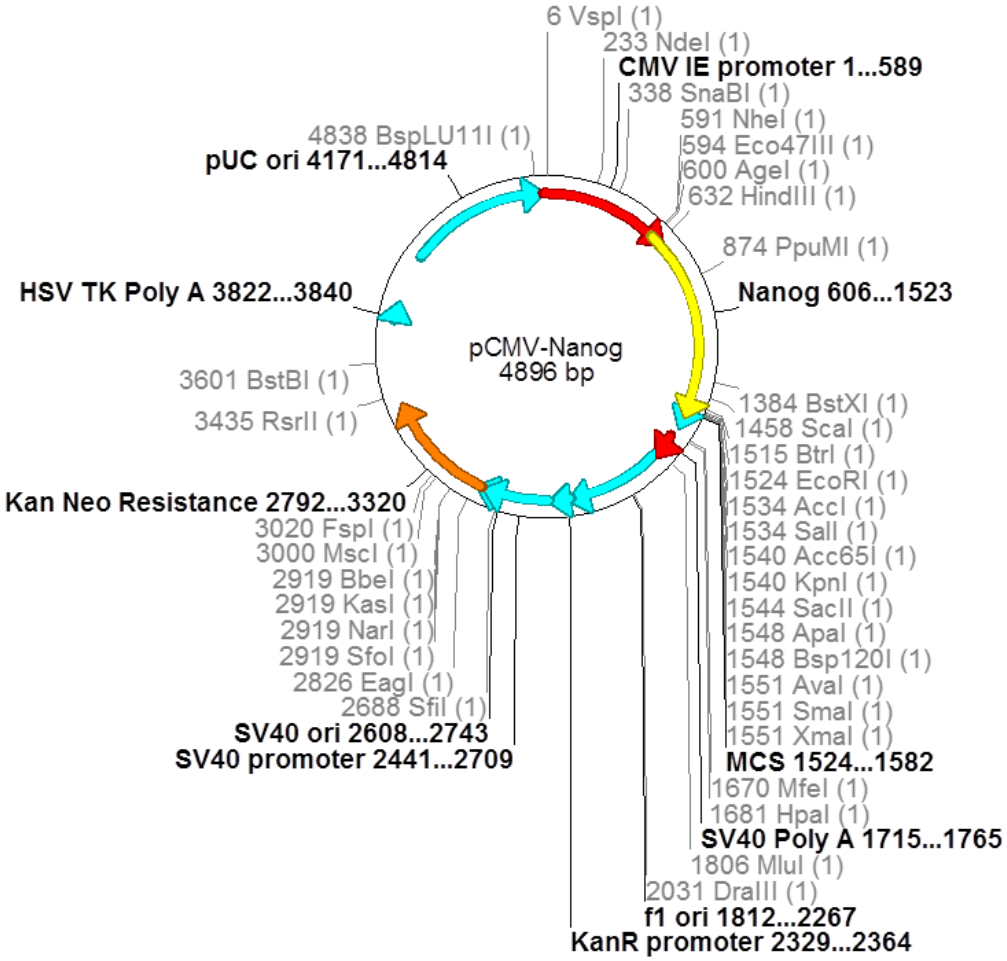
pCMV-Oct4



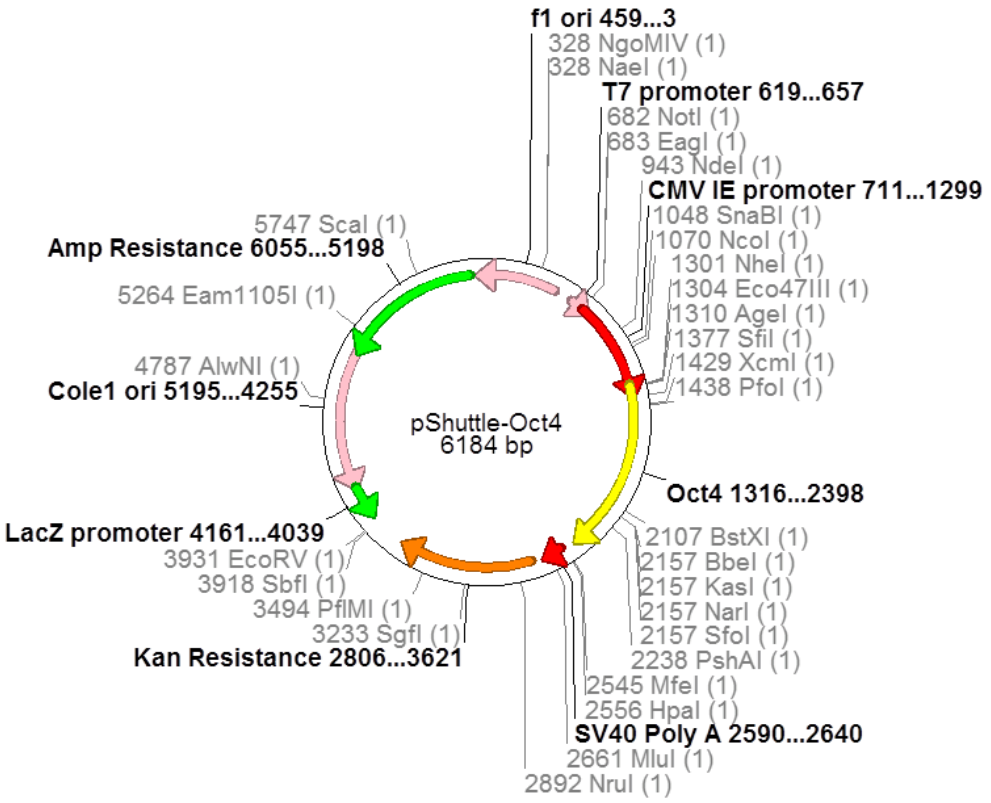
pCMV-Sox2



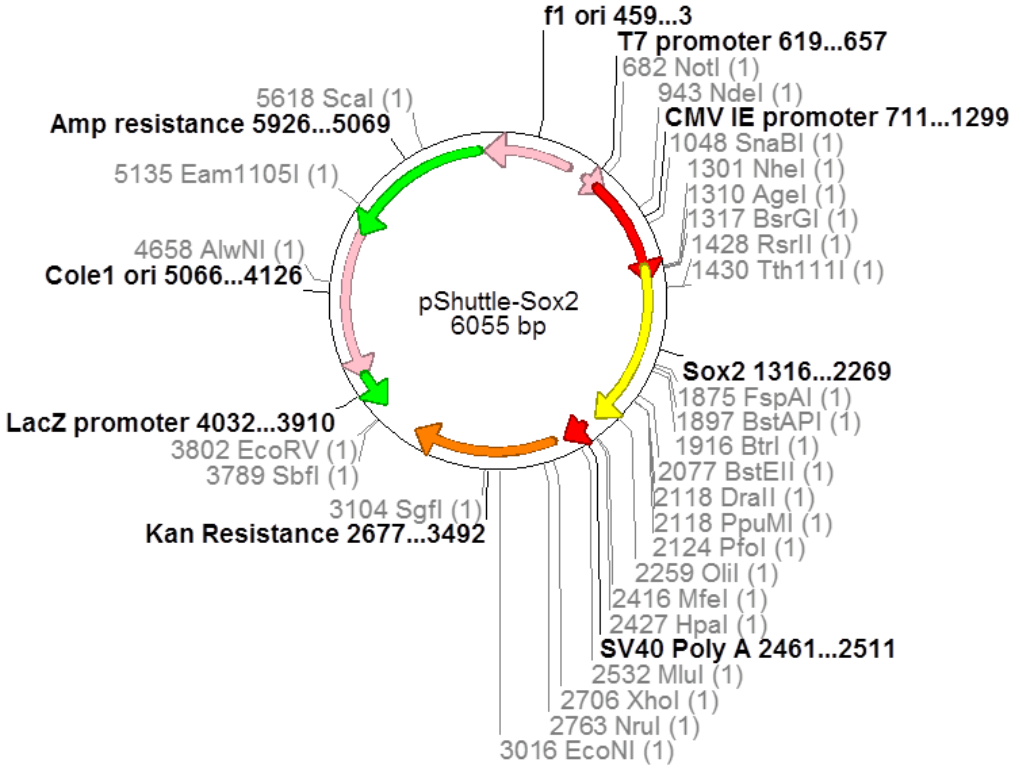
pCMV-Lin28



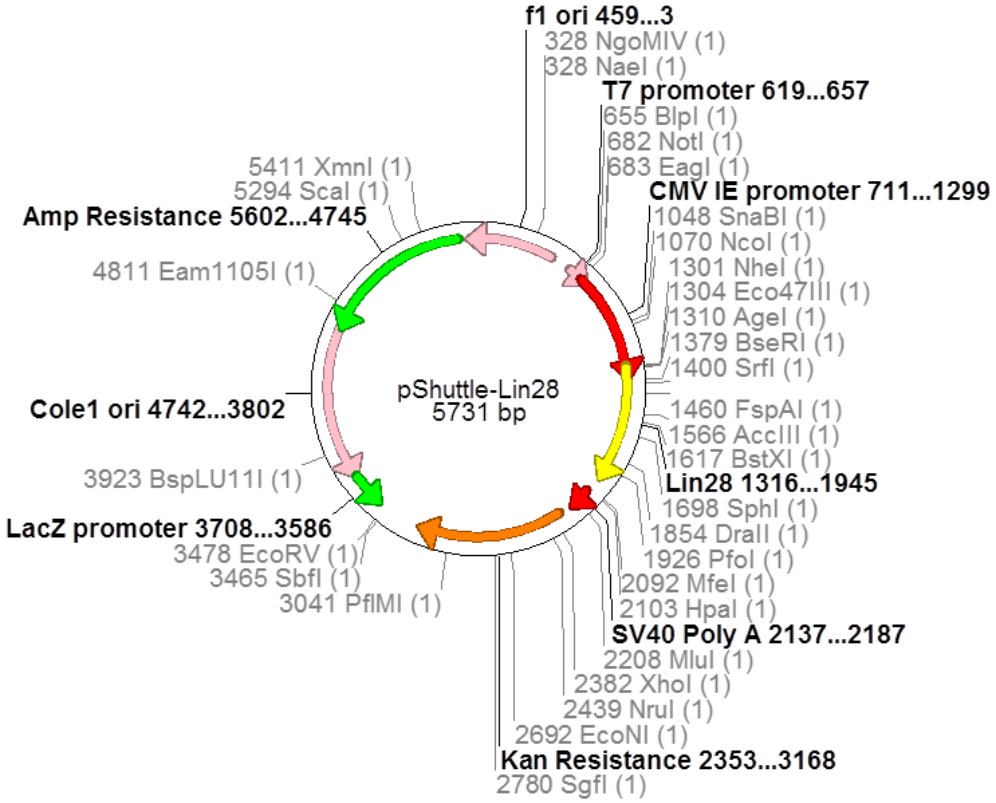
pCMV-Nanog



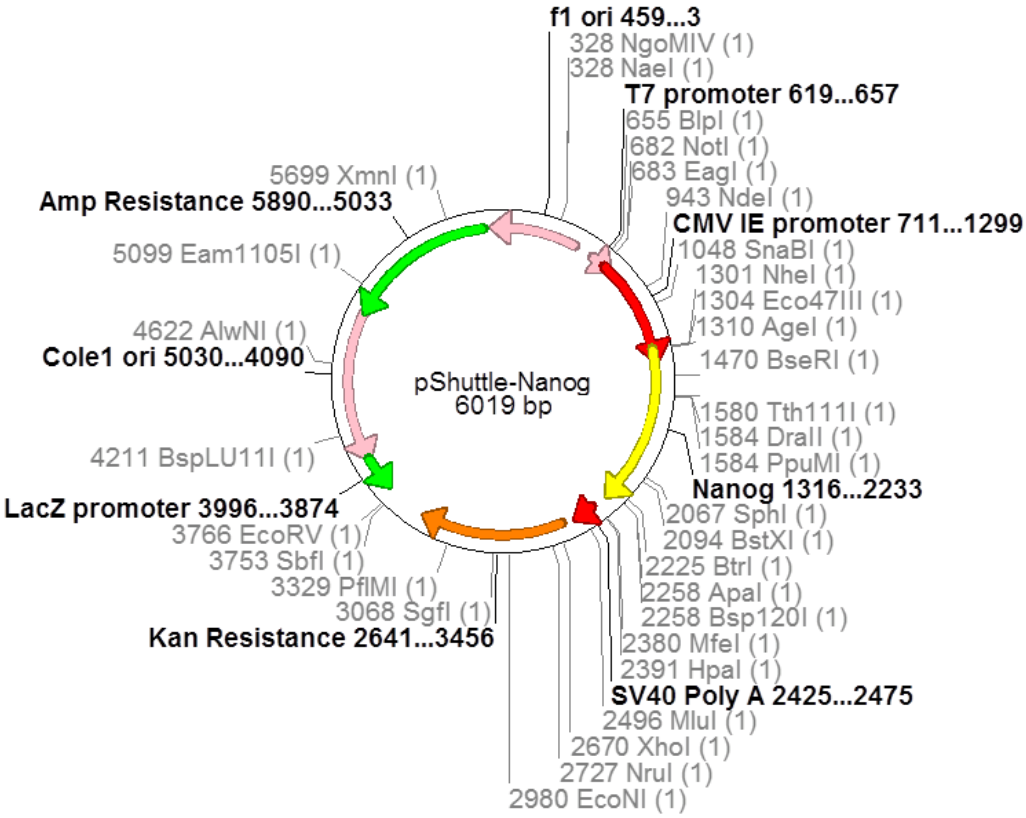
pShuttle-Oct4



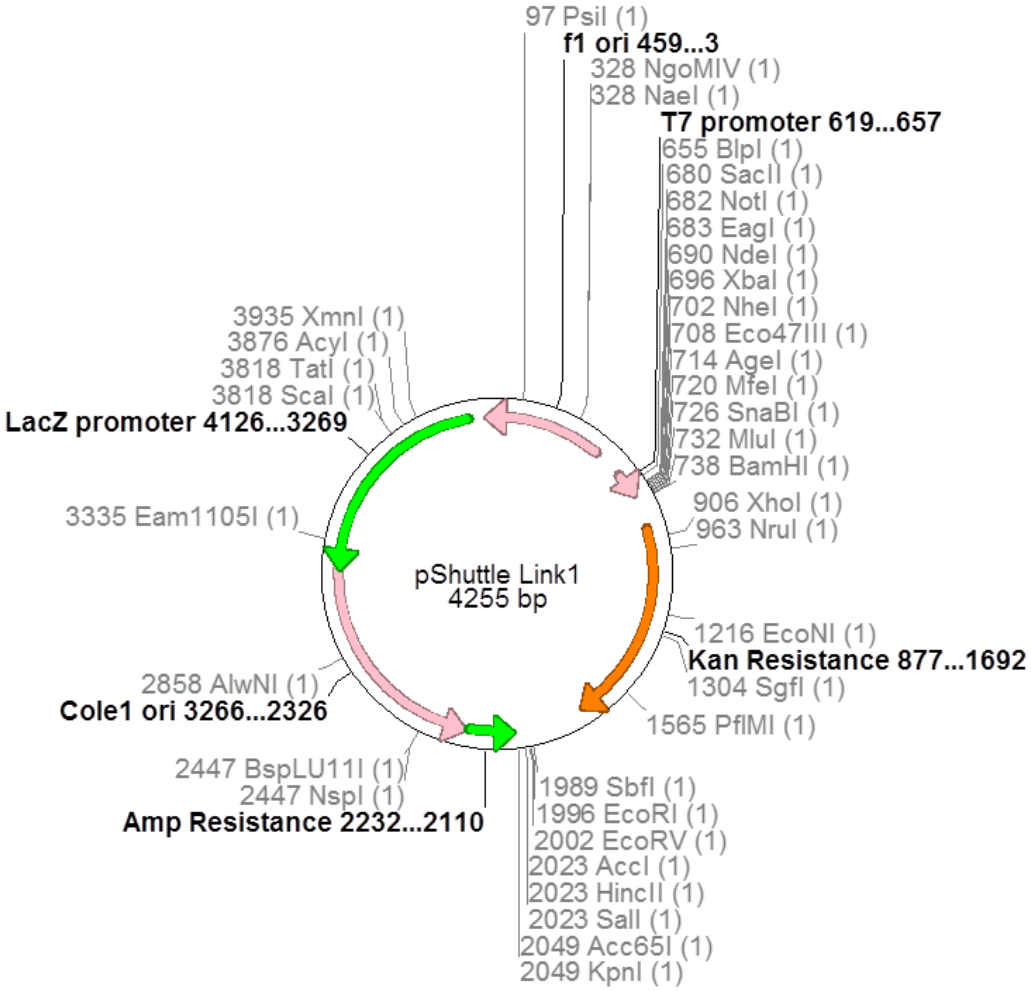
pShuttle-Sox2



pShuttle-Lin28



pShuttle-Nanog



pShuttle Link 1

Chapter 8

References

8 References

Abe, K., Yamashita, T., Takizawa, S., Kuroda, S., Kinouchi, H. & Kawahara, N. (2012). Stem cell therapy for cerebral ischemia: from basic science to clinical applications. *Journal of Cerebral Blood Flow and Metabolism* **32**, 1317-1331.

Abranches, E., O'Neill, A., Robertson, M. J., Schaffer, D. V. & Cabral, J. M. (2006). Development of quantitative PCR methods to analyse neural progenitor cell culture state. *Biotechnology and Applied Biochemistry* **44**, 1-8.

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences* **100**, 3983-3988.

Alberts, B. J., Alexander; Lewis, Julian; Raff, Martin, Roberts, Keith; Walter, Peter, (ed) (2002). *Molecular Biology of the Cell*. New York: Garland Science.

Albrecht, J. C., Nicholas, J., Biller, D., Cameron, K. R., Biesinger, B., Newman, C., Wittmann, S., Craxton, M. A., Coleman, H., Fleckenstein, B. & Honess, R. W. (1992). PRIMARY STRUCTURE OF THE HERPESVIRUS SAIMIRI GENOME. *Journal of virology* **66**, 5047-5058.

Alex, D. & Lee, K. A. (2005). RGG-boxes of the EWS oncoprotein repress a range of transcriptional activation domains. *Nucleic Acids Research* **33**, 1323-1331.

Ambrosetti, D. C., Basilico, C. & Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Molecular and cellular biology* **17**, 6321-6329.

Angervall, L. & Enzinger, F. M. (1975). Extraskeletal neoplasm resembling Ewing's sarcoma. *Cancer* **36**, 240-251.

Ansieau, S. (2012). EMT in breast cancer stem cell generation. *Cancer Letters*.

Askin, F. B., Rosai, J., Sibley, R. K., Dehner, L. P. & McAlister, W. H. (1979). Malignant small cell tumor of the thoracopulmonary region in childhood: a distinctive clinicopathologic entity of uncertain histogenesis. *Cancer* **43**, 2438-2451.

Aurich, H., Sgodda, M., Kaltwasser, P., Vetter, M., Weise, A., Liehr, T., Brulport, M., Hengstler, J. G., Dollinger, M. M., Fleig, W. E. & Christ, B. (2009). Hepatocyte differentiation of mesenchymal stem cells from human adipose tissue in vitro promotes hepatic integration in vivo. *Gut* **58**, 570-581.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. & Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & development* **17**, 126-140.

Babcock, G. J., Decker, L. L., Volk, M. & Thorley-Lawson, D. A. (1998). EBV persistence in memory B cells in vivo. *Immunity* **9**, 395-404.

Bach, F., Albertini, R., Joo, P., Anderson, J. & Bortin, M. (1968). BONE-MARROW TRANSPLANTATION IN A PATIENT WITH THE WISKOTT-ALDRICH SYNDROME. *The Lancet* **292**, 1364-1366.

Bao, S., Wu, Q., McLendon, R. E., Hao, Y., Shi, Q., Hjelmeland, A. B., Dewhirst, M. W., Bigner, D. D. & Rich, J. N. (2006). Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760.

Bar-Nur, O., Russ, Holger A., Efrat, S. & Benvenisty, N. (2011). Epigenetic Memory and Preferential Lineage-Specific Differentiation in Induced Pluripotent Stem Cells Derived from Human Pancreatic Islet Beta Cells. *Cell stem cell* **9**, 17-23.

Barker, N., Ridgway, R. A., van Es, J. H., van de Wetering, M., Begthel, H., van den Born, M., Danenberg, E., Clarke, A. R., Sansom, O. J. & Clevers, H. (2009). Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608-611.

Baron, U., Gossen, M. & Bujard, H. (1997). Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Research* **25**, 2723-2729.

Barroso-del Jesus, A., Lucena-Aguilar, G. & Menendez, P. (2009). The miR-302-367 cluster as a potential stemness regulator in ESCs. *Cell Cycle* **8**, 394-398.

Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jaker, C., Hock, J., Meister, G. & Grasser, F. A. (2008). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Res* **36**, 666-675.

Ben-David, Y., Giddens, E. B., Letwin, K. & Bernstein, A. (1991). Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes & development* **5**, 908-918.

Bernhardt, M., Galach, M., Novak, D. & Utikal, J. (2012). Mediators of induced pluripotency and their role in cancer cells – current scientific knowledge and future perspectives. *Biotechnology Journal* **7**, 810-821.

Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V. & Hannon, G. J. (2003). Dicer is essential for mouse development. *Nature genetics* **35**, 215-217.

Bertolotti, A., Melot, T., Acker, J., Vigneron, M., Delattre, O. & Tora, L. (1998). EWS, but not EWS-FLI-1, is associated with both TFIID and RNA polymerase II: interactions between two members of the TET family, EWS and hTAFII68, and subunits of TFIID and RNA polymerase II complexes. *Molecular and cellular biology* **18**, 1489-1497.

Bhandari, D. R., Seo, K. W., Roh, K. H., Jung, J. W., Kang, S. K. & Kang, K. S. (2010). REX-1 expression and p38 MAPK activation status can determine proliferation/differentiation fates in human mesenchymal stem cells. *PLoS ONE* **5**, e10493.

Biesinger, B., Trimble, J. J., Desrosiers, R. C. & Fleckenstein, B. (1990). THE DIVERGENCE BETWEEN 2 ONCOGENIC HERPESVIRUS-SAIMIRI STRAINS IN A GENOMIC REGION RELATED TO THE TRANSFORMING PHENOTYPE. *Virology* **176**, 505-514.

Biesinger, B., Tsygankov, A. Y., Fickenscher, H., Emmrich, F., Fleckenstein, B., Bolen, J. B. & Broker, B. M. (1995). THE PRODUCT OF THE HERPESVIRUS-SAIMIRI OPEN READING FRAME-1 (TIP) INTERACTS WITH T-CELL-SPECIFIC KINASE P56(LCK) IN TRANSFORMED-CELLS. *Journal of Biological Chemistry* **270**, 4729-4734.

Birkmann, A., Mahr, K., Ensser, A., Yaguboglu, S., Titgemeyer, F., Fleckenstein, B. & Neipel, F. (2001). Cell surface heparan sulfate is a receptor for human herpesvirus 8 and interacts with envelope glycoprotein K8.1. *Journal of virology* **75**, 11583-11593.

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**, 475-480.

Bonnet, D. & Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine* **3**, 730-737.

Brinkmann, M. M. & Schulz, T. F. (2006). Regulation of intracellular signalling by the terminal membrane proteins of members of the Gammaherpesvirinae. *Journal of General Virology* **87**, 1047-1074.

Brumbaugh, J., Hou, Z., Russell, J. D., Howden, S. E., Yu, P., Ledvina, A. R., Coon, J. J. & Thomson, J. A. (2012). Phosphorylation regulates human OCT4. *Proceedings of the National Academy of Sciences* **109**, 7162-7168.

Bussolati, B., Bruno, S., Grange, C., Ferrando, U. & Camussi, G. (2008). Identification of a tumor-initiating stem cell population in human renal carcinomas. *The FASEB Journal* **22**, 3696-3705.

Buzzard, J. J., Gough, N. M., Crook, J. M. & Colman, A. (2004). Karyotype of human ES cells during extended culture. *Nature Biotechnology* **22**, 381-382.

Calderwood, M., White, R. E., Griffiths, R. A. & Whitehouse, A. (2005). Open reading frame 73 is required for herpesvirus saimiri A11-S4 episomal persistence. *Journal of General Virology* **86**, 2703-2708.

Calderwood, M. A., Hall, K. T., Matthews, D. A. & Whitehouse, A. (2004). The herpesvirus saimiri ORF73 gene product interacts with host-cell mitotic chromosomes and self-associates via its C terminus. *Journal of General Virology* **85**, 147-153.

Card, D. A., Hebbar, P. B., Li, L., Trotter, K. W., Komatsu, Y., Mishina, Y. & Archer, T. K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Molecular and cellular biology* **28**, 6426-6438.

Carette, J. E., Pruszak, J., Varadarajan, M., Blomen, V. A., Gokhale, S., Camargo, F. D., Wernig, M., Jaenisch, R. & Brummelkamp, T. R. (2010). Generation of iPSCs from cultured human malignant cells. *Blood* **115**, 4039-4042.

Carey, B. W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M. & Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 157-162.

Carpenter, M. K., Frey-Vasconcells, J. & Rao, M. S. (2009). Developing safe therapies from human pluripotent stem cells. *Nature Biotechnology* **27**, 606-613.

Cazalla, D., Yario, T. & Steitz, J. A. (2010). Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* **328**, 1563-1566.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. & Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655.

Chan, K. S., Espinosa, I., Chao, M., Wong, D., Ailles, L., Diehn, M., Gill, H., Presti, J., Chang, H. Y., van de Rijn, M., Shortliffe, L. & Weissman, I. L. (2009). Identification, molecular characterization, clinical prognosis, and therapeutic targeting of human bladder tumor-initiating cells. *Proceedings of the National Academy of Sciences* **106**, 14016-14021.

Chang, G., Miao, Y.-L., Zhang, Y., Liu, S., Kou, Z., Ding, J., Chen, D.-Y., Sun, Q.-Y. & Gao, S. (2010). Linking Incomplete Reprogramming to the Improved Pluripotency of Murine Embryonal Carcinoma Cell-Derived Pluripotent Stem Cells. *PLoS ONE* **5**, e10320.

Chang, Y., Moore, P. S., Talbot, S. J., Boshoff, C. H., Zarkowska, T., GoddenKent, D., Paterson, H., Weiss, R. A. & Mittnacht, S. (1996). Cyclin encoded by KS herpesvirus. *Nature* **382**, 410-410.

Chen, J., Li, Y., Yu, T.-S., McKay, R. M., Burns, D. K., Kernie, S. G. & Parada, L. F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **advance online publication**.

Chen, S. J., Chang, C. M., Tsai, S. K., Chang, Y. L., Chou, S. J., Huang, S. S., Tai, L. K., Chen, Y. C., Ku, H. H., Li, H. Y. & Chiou, S. H. (2010). Functional improvement of focal cerebral ischemia injury by subdural transplantation of induced pluripotent stem cells with fibrin glue. *Stem cells and development* **19**, 1757-1767.

Cheng, L. T., Sun, L. T. & Tada, T. (2012). Genome editing in induced pluripotent stem cells. *Genes Cells* **17**, 431-438.

Cheong, C. Y., Lon Ng, P. M., Ponnampalam, R., Tsai, H. H., Bourque, G. & Lufkin, T. (2011). In silico tandem affinity purification refines an Oct4 interaction list. *Stem cell research & therapy* **2**, 26.

Chiang, A. K., Tao, Q., Srivastava, G. & Ho, F. C. (1996). Nasal NK- and T-cell lymphomas share the same type of Epstein-Barr virus latency as nasopharyngeal carcinoma and Hodgkin's disease. *International Journal of Cancer* **68**, 285-290.

Chin, M. H., Mason, M. J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J., Khvorostov, I., Ott, V., Grunstein, M., Lavon, N., Benvenisty, N., Croce, C. M., Clark, A. T., Baxter, T., Pyle, A. D., Teitell, M. A., Pelegrini, M., Plath, K. & Lowry, W. E. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell stem cell* **5**, 111-123.

Choi, J.-K., Ishido, S. & Jung, J. U. (2000). The Collagen Repeat Sequence Is a Determinant of the Degree of Herpesvirus saimiri STP Transforming Activity. *Journal of virology* **74**, 8102-8110.

Choi, S. M., Liu, H., Chaudhari, P., Kim, Y., Cheng, L., Feng, J., Sharkis, S., Ye, Z. & Jang, Y. Y. (2011). Reprogramming of EBV-immortalized B-lymphocyte cell lines into induced pluripotent stem cells. *Blood* **118**, 1801-1805.

Chua, H. L., Bhat-Nakshatri, P., Clare, S. E., Morimiya, A., Badve, S. & Nakshatri, H. (2007). NF-kappaB represses E-cadherin expression and enhances epithelial to mesenchymal transition of mammary epithelial cells: potential involvement of ZEB-1 and ZEB-2. *Oncogene* **26**, 711-724.

Chung, Y. H., Cho, N. H., Garcia, M. I., Lee, S. H., Feng, P. H. & Jung, J. U. (2004). Activation of Stat3 transcription factor by Herpesvirus saimiri STP-A oncoprotein. *Journal of virology* **78**, 6489-6497.

Colgan, K. J., Boyne, J. R. & Whitehouse, A. (2009a). Identification of a response element In a herpesvirus saimiri mRNA recognized by the ORF57 protein. *Journal of General Virology* **90**, 596-601.

Colgan, K. J., Boyne, J. R. & Whitehouse, A. (2009b). Uncoupling of hTREX demonstrates that UAP56 and hTHO-complex recruitment

onto herpesvirus saimiri intronless transcripts is required for replication. *Journal of General Virology* **90**, 1455-1460.

Collins, A. T., Berry, P. A., Hyde, C., Stower, M. J. & Maitland, N. J. (2005). Prospective Identification of Tumorigenic Prostate Cancer Stem Cells. *Cancer Research* **65**, 10946-10951.

Collins, C. M., Medveczky, M. M., Lund, T. & Medveczky, P. G. (2002). The terminal repeats and latency-associated nuclear antigen of Herpesvirus saimiri are essential for episomal persistence of the viral genome. *Journal of General Virology* **83**, 2269-2278.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D. & van Roy, F. (2001). The Two-Handed E Box Binding Zinc Finger Protein SIP1 Downregulates E-Cadherin and Induces Invasion. *Molecular Cell* **7**, 1267-1278.

Cooper, M., Goodwin, D., Hall, K., Stevenson, A., Meredith, D., Markham, A. & Whitehouse, A. (1999). The gene product encoded by ORF 57 of Herpesvirus saimiri regulates the redistribution of the splicing factor SC-35. *Journal of General Virology* **80**, 1311-1316.

Countryman, J., Jenson, H., Seibl, R., Wolf, H. & Miller, G. (1987). Polymorphic proteins encoded within BZLF1 of defective and standard Epstein-Barr viruses disrupt latency. *Journal of virology* **61**, 3672-3679.

Countryman, J. & Miller, G. (1985). Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 4085-4089.

Damania, B. (2004). Oncogenic gamma-herpesviruses: comparison of viral proteins involved in tumorigenesis. *Nature Reviews Microbiology* **2**, 656-668.

Davison, A. J., Dolan, A., Akter, P., Addison, C., Dargan, D. J., Alcendor, D. J., McGeoch, D. J. & Hayward, G. S. (2003). The human cytomegalovirus genome revisited: comparison with the

chimpanzee cytomegalovirus genome. *Journal of General Virology* **84**, 17-28.

Deacon, E. M., Pallesen, G., Niedobitek, G., Crocker, J., Brooks, L., Rickinson, A. B. & Young, L. S. (1993). Epstein-Barr virus and Hodgkin's disease: transcriptional analysis of virus latency in the malignant cells. *Journal of Experimental Medicine* **177**, 339-349.

Dean, M., Fojo, T. & Bates, S. (2005). Tumour stem cells and drug resistance. *Nature Reviews Cancer* **5**, 275-284.

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A. & Thomas, G. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162-165.

Desrosiers, R., Bakker, A., Kamine, J., Falk, L., Hunt, R. & King, N. (1985a). A region of the Herpesvirus saimiri genome required for oncogenicity. *Science* **228**, 184-187.

Desrosiers, R. C., Kamine, J., Bakker, A., Silva, D., Woychik, R. P., Sakai, D. D. & Rottman, F. M. (1985b). Synthesis of bovine growth hormone in primates by using a herpesvirus vector. *Molecular and Cellular Biolology* **5**, 2796-2803.

Dhara, S. K. & Benvenisty, N. (2004). Gene trap as a tool for genome annotation and analysis of X chromosome inactivation in human embryonic stem cells. *Nucleic Acids Research* **32**, 3995-4002.

Diehn, M., Cho, R. W., Lobo, N. A., Kalisky, T., Dorie, M. J., Kulp, A. N., Qian, D., Lam, J. S., Ailles, L. E., Wong, M., Joshua, B., Kaplan, M. J., Wapnir, I., Dirbas, F. M., Somlo, G., Garberoglio, C., Paz, B., Shen, J., Lau, S. K., Quake, S. R., Brown, J. M., Weissman, I. L. & Clarke, M. F. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* **458**, 780-783.

Dimmock, N. J., Easton, A. & Leppard, K. (2001). *An Introduction to Modern Virology*: Blackwell Science Ltd.

Dimmock, N. J., Easton, A. & Leppard, K. (2006). *Introduction to Modern Virology*.

Djojosubroto, M. W., Chin, A. C., Go, N., Schaetzlein, S., Manns, M. P., Gryaznov, S., Harley, C. B. & Rudolph, K. L. (2005). Telomerase antagonists GRN163 and GRN163L inhibit tumor growth and increase chemosensitivity of human hepatoma. *Hepatology* **42**, 1127-1136.

Doody, G. M., Leek, J. P., Bali, A. K., Ensser, A., Markham, A. F. & de Wynter, E. A. (2005). Marker gene transfer into human haemopoietic cells using a herpesvirus saimiri-based vector. *Gene Therapy* **12**, 373-379.

Driessens, G., Beck, B., Caauwe, A., Simons, B. D. & Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. *Nature* **advance online publication**.

Ebert, A. D., Yu, J., Rose, F. F., Mattis, V. B., Lorson, C. L., Thomson, J. A. & Svendsen, C. N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**, 277-280.

Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., Lund, E. & Dahlberg, J. E. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3627-3632.

Ensser, A., Pfinder, A., Muller-Fleckenstein, I. & Fleckenstein, B. (1999). The URNA Genes of Herpesvirus Saimiri (Strain C488) Are Dispensable for Transformation of Human T Cells In Vitro. *J Virol* **73**, 10551-10555.

Ensser, A., Thurau, M., Wittmann, S. & Fickenscher, H. (2003). The genome of Herpesvirus saimiri C488 which is capable of transforming human T cells. *Virology* **314**, 471-487.

Episkopou, V. (2005). SOX2 functions in adult neural stem cells. *Trends in Neurosciences* **28**, 219-221.

Erkizan, H. V., Uversky, V. N. & Toretsky, J. A. (2010). Oncogenic Partnerships: EWS-FLI1 Protein Interactions Initiate Key Pathways of Ewing's Sarcoma. *Clinical Cancer Research* **16**, 4077-4083.

Evans, M. J. & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.

Ewing, J. (1921). Diffuse endothelioma of bone. *Proceedings of the New York Pathological Society* **21**, 17.

Falk, L. A., Wolfe, L. G. & Deinhardt, F. (1972). Isolation of Herpesvirus saimiri from blood of squirrel monkeys (Saimiri sciureus). *Journal of the National Cancer Institute* **48**, 1499-1505.

Fang, D., Nguyen, T. K., Leishear, K., Finko, R., Kulp, A. N., Hotz, S., Van Belle, P. A., Xu, X., Elder, D. E. & Herlyn, M. (2005). A Tumorigenic Subpopulation with Stem Cell Properties in Melanomas. *Cancer Research* **65**, 9328-9337.

Feinberg, A. P., Ohlsson, R. & Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. *Nat Rev Genet* **7**, 21-33.

Feng, D., Peng, C., Li, C., Zhou, Y., Li, M., Ling, B., Wei, H. & Tian, Z. (2009). Identification and characterization of cancer stem-like cells from primary carcinoma of the cervix uteri. *Oncol Rep* **22**, 1129-1134.

Fickenscher, H. & Fleckenstein, B. (2001). Herpesvirus saimiri. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **356**, 545-567.

Fingeroth, J. D., Weis, J. J., Tedder, T. F., Strominger, J. L., Biro, P. A. & Fearon, D. T. (1984). Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 4510-4514.

Fleckenstein, B. & Desrosiers, R. C., (eds) (1982). *Herpesvirus saimiri and herpesvirus ateles*. New York: Plenum Press.

Forster, K., Helbl, V., Lederer, T., Urlinger, S., Wittenburg, N. & Hillen, W. (1999). Tetracycline-inducible expression systems with reduced

basal activity in mammalian cells. *Nucleic Acids Research* **27**, 708-710.

Frolova-Jones, E. A., Ensser, A., Stevenson, A. J., Kinsey, S. E. & Meredith, D. M. (2000). Stable marker gene transfer into human bone marrow stromal cells and their progenitors using novel herpesvirus saimiri-based vectors. *Journal of Hematotherapy and Stem Cell Research* **9**, 573-581.

Fu, X. D. & Maniatis, T. (1990). FACTOR REQUIRED FOR MAMMALIAN SPLICEOSOME ASSEMBLY IS LOCALIZED TO DISCRETE REGIONS IN THE NUCLEUS. *Nature* **343**, 437-441.

Gal, A., Sjoblom, T., Fedorova, L., Imreh, S., Beug, H. & Moustakas, A. (2008). Sustained TGF beta exposure suppresses Smad and non-Smad signalling in mammary epithelial cells, leading to EMT and inhibition of growth arrest and apoptosis. *Oncogene* **27**, 1218-1230.

Garcia, M. I., Kaserman, J., Chung, Y.-H., Jung, J. U. & Lee, S.-H. (2007). Herpesvirus saimiri STP-A Oncoprotein Utilizes Src Family Protein Tyrosine Kinase and Tumor Necrosis Factor Receptor-Associated Factors To Elicit Cellular Signal Transduction. *Journal of virology* **81**, 2663-2674.

Gatti, R., Meuwissen, H., Allen, H., Hong, R. & Good, R. (1968). IMMUNOLOGICAL RECONSTITUTION OF SEX-LINKED LYMPHOPENIC IMMUNOLOGICAL DEFICIENCY. *The Lancet* **292**, 1366-1369.

Ghosh, Z., Mallick, B. & Chakrabarti, J. (2009). Cellular versus viral microRNAs in host–virus interaction. *Nucleic Acids Research* **37**, 1035-1048.

Ghosh, Z., Wilson, K. D., Wu, Y., Hu, S., Quertermous, T. & Wu, J. C. (2010). Persistent Donor Cell Gene Expression among Human Induced Pluripotent Stem Cells Contributes to Differences with Human Embryonic Stem Cells. *PLoS ONE* **5**, e8975.

Gidekel, S., Pizov, G., Bergman, Y. & Pikarsky, E. (2003). Oct-3/4 is a dose-dependent oncogenic fate determinant. *Cancer cell* **4**, 361-370.

Goodwin, D. J., Hall, K. T., Stevenson, A. J., Markham, A. F. & Whitehouse, A. (1999). The open reading frame 57 gene product of herpesvirus saimiri shuttles between the nucleus and cytoplasm and is involved in viral RNA nuclear export. *Journal of virology* **73**, 10519-10524.

Goodwin, D. J., Walters, M. S., Smith, P. G., Thurau, M., Fickenscher, H. & Whitehouse, A. (2001). Herpesvirus saimiri Open Reading Frame 50 (Rta) Protein Reactivates the Lytic Replication Cycle in a Persistently Infected A549 Cell Line. *Journal of virology* **75**, 4008-4013.

Gossen, M. & Bujard, H. (1992a). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences* **89**, 5547-5551.

Gossen, M. & Bujard, H. (1992b). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 5547-5551.

Grassmann, R. & Fleckenstein, B. (1989). SELECTABLE RECOMBINANT HERPESVIRUS SAIMIRI IS CAPABLE OF PERSISTING IN A HUMAN T-CELL LINE. *Journal of virology* **63**, 1818-1821.

Gray, W. L., Starnes, B., White, M. W. & Mahalingam, R. (2001). The DNA sequence of the simian varicella virus genome. *Virology* **284**, 123-130.

Griffiths, R., Harrison, S. M., Macnab, S. & Whitehouse, A. (2008). Mapping the minimal regions within the ORF73 protein required for herpesvirus saimiri episomal persistence. *Journal of General Virology* **89**, 2843-2850.

Griffiths, R. & Whitehouse, A. (2007). Herpesvirus saimiri episomal persistence is maintained via interaction between open reading frame 73 and the cellular chromosome-associated protein MeCP2. *Journal of virology* **81**, 4021-4032.

Griffiths, R. A., Boyne, J. A. & Whitehouse, A. (2006). Herpesvirus saimiri-based gene delivery vectors. *Current Gene Therapy* **6**, 1-15.

Gupta, P. B., Chaffer, C. L. & Weinberg, R. A. (2009a). Cancer stem cells: mirage or reality? *Nature Medicine* **15**, 1010-1012.

Gupta, P. B., Onder, T. T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R. A. & Lander, E. S. (2009b). Identification of Selective Inhibitors of Cancer Stem Cells by High-Throughput Screening. *Cell* **138**, 645-659.

Hall, K. T., Stevenson, A. J., Goodwin, D. J., Gibson, P. C., Markham, A. F. & Whitehouse, A. (1999). The activation domain of Herpesvirus saimiri R protein interacts with the TATA-binding protein. *Journal of virology* **73**, 9756-9763.

Hanahan, D. & Weinberg, Robert A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646-674.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C. J., Creyghton, M. P., van Oudenaarden, A. & Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595-601.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.-W., Meissner, A., Cassady, J. P., Beard, C., Brambrink, T., Wu, L.-C., Townes, T. M. & Jaenisch, R. (2007). Treatment of Sickle Cell Anemia Mouse Model with iPS Cells Generated from Autologous Skin. *Science* **318**, 1920-1923.

Hardwick, J. M., Lieberman, P. M. & Hayward, S. D. (1988). A new Epstein-Barr virus transactivator, R, induces expression of a cytoplasmic early antigen. *Journal of virology* **62**, 2274-2284.

Hargus, G., Cooper, O., Deleidi, M., Levy, A., Lee, K., Marlow, E., Yow, A., Soldner, F., Hockemeyer, D., Hallett, P. J., Osborn, T., Jaenisch, R. & Isacson, O. (2010). Differentiated Parkinson patient-derived induced pluripotent stem cells grow in the adult rodent brain and reduce motor asymmetry in Parkinsonian rats. *Proceedings of the National Academy of Sciences* **107**, 15921-15926.

Hartley, D. A. & Cooper, G. M. (2000). Direct binding and activation of STAT transcription factors by the herpesvirus saimiri protein tip. *Journal of Biological Chemistry* **275**, 16925-16932.

Heck, E., Friedrich, U., Gack, M. U., Lengenfelder, D., Schmidt, M., MÃ¼ller-Fleckenstein, I., Fleckenstein, B., Ensser, A. & Biesinger, B. (2006). Growth Transformation of Human T Cells by Herpesvirus Saimiri Requires Multiple Tip-Lck Interaction Motifs. *Journal of virology* **80**, 9934-9942.

Heng, J. C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J. H., Orlov, Y. L., Huss, M., Yang, L., Lufkin, T., Lim, B. & Ng, H. H. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell stem cell* **6**, 167-174.

Henle, G., Henle, W. & Diehl, V. (1968). Relation of Burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences of the United States of America* **59**, 94-101.

Heppner, G. H. (1984). Tumor heterogeneity. *Cancer Research* **44**, 2259-2265.

Herold, B. C., Visalli, R. J., Susmarski, N., Brandt, C. R. & Spear, P. G. (1994). Glycoprotein C-independent binding of herpes simplex virus to cells requires cell surface heparan sulphate and glycoprotein B. *Journal of General Virology* **75 (Pt 6)**, 1211-1222.

Herr, W. & Cleary, M. A. (1995). The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes & development* **9**, 1679-1693.

Hirschmann-Jax, C., Foster, A. E., Wulf, G. G., Nuchtern, J. G., Jax, T. W., Gobel, U., Goodell, M. A. & Brenner, M. K. (2004). A distinct “side population” of cells with high drug efflux capacity in human tumor cells. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14228-14233.

Hochedlinger, K., Blelloch, R., Brennan, C., Yamada, Y., Kim, M., Chin, L. & Jaenisch, R. (2004). Reprogramming of a melanoma genome by nuclear transplantation. *Genes & development* **18**, 1875-1885.

Hochreiter, A. E., Xiao, H., Goldblatt, E. M., Gryaznov, S. M., Miller, K. D., Badve, S., Sledge, G. W. & Herbert, B.-S. (2006). Telomerase Template Antagonist GRN163L Disrupts Telomere Maintenance, Tumor Growth, and Metastasis of Breast Cancer. *Clinical Cancer Research* **12**, 3184-3192.

Hoffman, L. M., Hall, L., Batten, J. L., Young, H., Pardasani, D., Baetge, E. E., Lawrence, J. & Carpenter, M. K. (2005). X-inactivation status varies in human embryonic stem cell lines. *Stem Cells* **23**, 1468-1478.

Hong, Y., Macnab, S., Lambert, L. A., Turner, A. J., Whitehouse, A. & Usmani, B. A. (2011). Herpesvirus saimiri-based endothelin-converting enzyme-1 shRNA expression decreases prostate cancer cell invasion and migration. *International Journal of Cancer* **129**, 586-598.

Hotta, A. & Ellis, J. (2008). Retroviral vector silencing during iPS cell induction: An epigenetic beacon that signals distinct pluripotent states. *Journal of Cellular Biochemistry* **105**, 940-948.

Houthuijzen, J. M., Daenen, L. G. M., Roodhart, J. M. L. & Voest, E. E. (2012). The role of mesenchymal stem cells in anti-cancer drug resistance and tumour progression. *British Journal of Cancer*.

Hu-Lieskovan, S., Zhang, J., Wu, L., Shimada, H., Schofield, D. E. & Triche, T. J. (2005). EWS-FLI1 Fusion Protein Up-regulates Critical Genes in Neural Crest Development and Is Responsible for the Observed Phenotype of Ewing's Family of Tumors. *Cancer Research* **65**, 4633-4644.

Hu, K., Yu, J., Suknuntha, K., Tian, S., Montgomery, K., Choi, K.-D., Stewart, R., Thomson, J. A. & Slukvin, I. I. (2011). Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. *Blood* **117**, e109-e119.

Hu, Y. & Fu, L. (2012). Targeting cancer stem cells: a new therapy to cure cancer patients. *American journal of cancer research* **2**, 340-356.

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. *Nature Biotechnology* **26**, 1269-1275.

Isakov, N. & Biesinger, B. (2000). Lck protein tyrosine kinase is a key regulator of T-cell activation and a target for signal intervention by Herpesvirus saimiri and other viral gene products. *European Journal of Biochemistry* **267**, 3413-3421.

Israel, M. A., Yuan, S. H., Bardy, C., Reyna, S. M., Mu, Y., Herrera, C., Hefferan, M. P., Van Gorp, S., Nazor, K. L., Boscolo, F. S., Carson, C. T., Laurent, L. C., Marsala, M., Gage, F. H., Remes, A. M., Koo, E. H. & Goldstein, L. S. (2012). Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216-220.

Itzhaki, I., Maizels, L., Huber, I., Zwi-Dantsis, L., Caspi, O., Winterstern, A., Feldman, O., Gepstein, A., Arbel, G., Hammerman, H., Boulos, M. & Gepstein, L. (2011). Modelling the long QT syndrome with induced pluripotent stem cells. *Nature* **471**, 225-229.

Jackson, B. R., Noerenberg, M. & Whitehouse, A. (2012). The Kaposi's Sarcoma-Associated Herpesvirus ORF57 Protein and Its Multiple Roles in mRNA Biogenesis. *Frontiers in Microbiology* **3**, 59.

Jaffe, R., Santamaria, M., Yunis, E. J., Tannery, N. H., Agostini, R. M., Jr., Medina, J. & Goodman, M. (1984). The neuroectodermal tumor of bone. *The American journal of surgical pathology* **8**, 885-898.

Jang, H., Kim, Tae W., Yoon, S., Choi, S.-Y., Kang, T.-W., Kim, S.-Y., Kwon, Y.-W., Cho, E.-J. & Youn, H.-D. (2012). O-GlcNAc Regulates Pluripotency and Reprogramming by Directly Acting on Core Components of the Pluripotency Network. *Cell stem cell* **11**, 62-74.

Jones, J. F., Shurin, S., Abramowsky, C., Tubbs, R. R., Sciotto, C. G., Wahl, R., Sands, J., Gottman, D., Katz, B. Z. & Sklar, J. (1988). T-

cell lymphomas containing Epstein-Barr viral DNA in patients with chronic Epstein-Barr virus infections. *N Engl J Med* **318**, 733-741.

Jung, J. U. & Desrosiers, R. C. (1995). ASSOCIATION OF THE VIRAL ONCOPROTEIN STP-C488 WITH CELLULAR RAS. *Molecular and cellular biology* **15**, 6506-6512.

Jung, J. U., Lang, S. M., Friedrich, U., Jun, T., Roberts, T. M., Desrosiers, R. C. & Biesinger, B. (1995a). IDENTIFICATION OF LCK-BINDING ELEMENTS IN TIP OF HERPESVIRUS SAIMIRI. *Journal of Biological Chemistry* **270**, 20660-20667.

Jung, J. U., Lang, S. M., Jun, T., Roberts, T. M., Veillette, A. & Desrosiers, R. C. (1995b). Downregulation of Lck-mediated signal transduction by tip of herpesvirus saimiri. *Journal of virology* **69**, 7814-7822.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P. & Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* **458**, 771-U112.

Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M. & Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & development* **19**, 489-501.

Kawaguchi, A., Kanai, K., Satoh, Y., Touge, C., Nagata, K., Sairenji, T. & Inoue, Y. (2009). The evolution of Epstein-Barr virus inferred from the conservation and mutation of the virus glycoprotein gp350/220 gene. *Virus Genes* **38**, 215-223.

Kelly, P. N., Dakic, A., Adams, J. M., Nutt, S. L. & Strasser, A. (2007). Tumor Growth Need Not Be Driven by Rare Cancer Stem Cells. *Science* **317**, 337.

Kim, C. F. B., Jackson, E. L., Woolfenden, A. E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R. T. & Jacks, T. (2005). Identification of Bronchioalveolar Stem Cells in Normal Lung and Lung Cancer. *Cell* **121**, 823-835.

Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., Ko, S., Yang, E., Cha, K. Y., Lanza, R. & Kim, K.-S. (2009a). Generation of Human Induced Pluripotent Stem Cells by Direct Delivery of Reprogramming Proteins. *Cell stem cell* **4**, 472-476.

Kim, J. B., Greber, B., Arauzo-Bravo, M. J., Meyer, J., Park, K. I., Zaehres, H. & Scholer, H. R. (2009b). Direct reprogramming of human neural stem cells by OCT4. *Nature* **461**, 649-643.

Kim, J. B., Zaehres, H., Arauzo-Bravo, M. J. & Scholer, H. R. (2009c). Generation of induced pluripotent stem cells from neural stem cells. *Nature Protocols* **4**, 1464-1470.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M. J., Ji, H., Ehrlich, L. I. R., Yabuuchi, A., Takeuchi, A., Cunniff, K. C., Hongguang, H., McKinney-Freeman, S., Naveiras, O., Yoon, T. J., Irizarry, R. A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S. H., Weissman, I. L., Feinberg, A. P. & Daley, G. Q. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285-290.

Kim, K., Lu, Z. & Hay, E. D. (2002). DIRECT EVIDENCE FOR A ROLE OF β-CATENIN/LEF-1 SIGNALING PATHWAY IN INDUCTION OF EMT. *Cell Biology International* **26**, 463-476.

Kim, V. N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol* **6**, 376-385.

Kingston, D., Chang, H., Ensser, A., Lee, H.-R., Lee, J., Lee, S.-H., Jung, J. U. & Cho, N.-H. (2011). Inhibition of Retromer Activity by Herpesvirus Saimiri Tip Leads to CD4 Downregulation and Efficient T Cell Transformation. *Journal of virology* **85**, 10627-10638.

Klein, E., Kis, L. L. & Klein, G. (2007). Epstein-Barr virus infection in humans: from harmless to life endangering virus-lymphocyte interactions. *Oncogene* **26**, 1297-1305.

Kleinsmith, L. J. & Pierce, G. B., Jr. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Research* **24**, 1544-1551.

Knoop, L. L. & Baker, S. J. (2000). The splicing factor U1C represses EWS/FLI-mediated transactivation. *Journal of Biological Chemistry* **275**, 24865-24871.

Knoop, L. L. & Baker, S. J. (2001). EWS/FLI Alters 5′-Splice Site Selection. *Journal of Biological Chemistry* **276**, 22317-22322.

Kumano, K., Arai, S., Hosoi, M., Taoka, K., Takayama, N., Otsu, M., Nagae, G., Ueda, K., Nakazaki, K., Kamikubo, Y., Eto, K., Aburatani, H., Nakauchi, H. & Kurokawa, M. (2012). Generation of induced pluripotent stem cells from primary chronic myelogenous leukemia patient samples. *Blood* **119**, 6234-6242.

Lai, Y. S., Chang, C. W., Pawlik, K. M., Zhou, D., Renfrow, M. B. & Townes, T. M. (2012). SRY (sex determining region Y)-box2 (Sox2)/poly ADP-ribose polymerase 1 (Parp1) complexes regulate pluripotency. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3772-3777.

Laker, C., Meyer, J., Schopen, A., Friel, J., Heberlein, C., Ostertag, W. & Stocking, C. (1998). Host cis-Mediated Extinction of a Retrovirus Permissive for Expression in Embryonal Stem Cells during Differentiation. *Journal of virology* **72**, 339-348.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M. A. & Dick, J. E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-648.

Laquerre, S., Argnani, R., Anderson, D. B., Zucchini, S., Manservigi, R. & Glorioso, J. C. (1998). Heparan sulfate proteoglycan binding by herpes simplex virus type 1 glycoproteins B and C, which differ in their contributions to virus attachment, penetration, and cell-to-cell spread. *Journal of virology* **72**, 6119-6130.

Laurent, L. C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J. V., Lee, S., Barrero, M. J., Ku, S., Martynova, M., Semechkin, R., Galat, V., Gottesfeld, J., Belmonte, J. C. I., Murry, C., Keirstead, H. S., Park, H.-S., Schmidt, U.,

Laslett, A. L., Muller, F.-J., Nievergelt, C. M., Shamir, R. & Loring, J. F. (2011). Dynamic Changes in the Copy Number of Pluripotency and Cell Proliferation Genes in Human ESCs and iPSCs during Reprogramming and Time in Culture. *Cell stem cell* **8**, 106-118.

Law, W. J., Cann, K. L. & Hicks, G. G. (2006). TLS, EWS and TAF15: a model for transcriptional integration of gene expression. *Briefings in functional genomics & proteomics* **5**, 8-14.

Lee, G., Papapetrou, E. P., Kim, H., Chambers, S. M., Tomishima, M. J., Fasano, C. A., Ganat, Y. M., Menon, J., Shimizu, F., Viale, A., Tabar, V., Sadelain, M. & Studer, L. (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461**, 402-406.

Lee, H., Trimble, J., Yoon, D., Regier, D., Desrosiers, R. & Jung, J. (1997). Genetic variation of herpesvirus saimiri subgroup A transforming protein and its association with cellular src. *Journal of virology* **71**, 3817-3825.

Lee, J., Rhee, B. K., Bae, G. Y., Han, Y. M. & Kim, J. (2005). Stimulation of Oct-4 activity by Ewing's sarcoma protein. *Stem Cells* **23**, 738-751.

Lessnick, S. L., Dacwag, C. S. & Golub, T. R. (2002). The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer cell* **1**, 393-401.

Levine, P. H., Ablashi, D. V., Berard, C. W., Carbone, P. P., Waggoner, D. E. & Malan, L. (1971). Elevated antibody titers to Epstein-Barr virus in Hodgkin's disease. *Cancer* **27**, 416-421.

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F. & Simeone, D. M. (2007). Identification of Pancreatic Cancer Stem Cells. *Cancer Research* **67**, 1030-1037.

Li, W., Zhou, H., Abujarour, R., Zhu, S., Young Joo, J., Lin, T., Hao, E., Scholer, H. R., Hayek, A. & Ding, S. (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. *Stem Cells* **27**, 2992-3000.

Li, Y., Li, X., Fan, G., Fukushi, J.-i., Matsumoto, Y., Iwamoto, Y. & Zhu, Y. (2012). Impairment of p53 acetylation by EWS-Fli1 chimeric protein in Ewing Family Tumors. *Cancer Letters* **320**, 14-22.

Li, Y., Tanaka, K., Fan, X., Nakatani, F., Li, X., Nakamura, T., Takasaki, M., Yamamoto, S. & Iwamoto, Y. (2010). Inhibition of the transcriptional function of p53 by EWS-Fli1 chimeric protein in Ewing Family Tumors. *Cancer Letters* **294**, 57-65.

Liang, H., Olejniczak, E. T., Mao, X., Nettesheim, D. G., Yu, L., Thompson, C. B. & Fesik, S. W. (1994). The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 11655-11659.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S. Y., Qin, J., Wong, J., Cooney, A. J., Liu, D. & Songyang, Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature Cell Biology* **10**, 731-739.

Lin, S. L., Chang, D. C., Chang-Lin, S., Lin, C. H., Wu, D. T., Chen, D. T. & Ying, S. Y. (2008). Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* **14**, 2115-2124.

Lister, R., Pelizzola, M., Kida, Y. S., Hawkins, R. D., Nery, J. R., Hon, G., Antosiewicz-Bourget, J., O/'Malley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J. A., Evans, R. M. & Ecker, J. R. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73.

Liu, P. & Speck, S. H. (2003). Synergistic autoactivation of the Epstein-Barr virus immediate-early BRLF1 promoter by Rta and Zta. *Virology* **310**, 199-206.

Lo, A. K., To, K. F., Lo, K. W., Lung, R. W., Hui, J. W., Liao, G. & Hayward, S. D. (2007). Modulation of LMP1 protein expression by

EBV-encoded microRNAs. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 16164-16169.

Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C. L., Ruan, Y., Lim, B. & Ng, H. H. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics* **38**, 431-440.

Longnecker, R. & Neipel, F. (2007). Introduction to the human gamma-herpesviruses. In *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*. Edited by A. Arvin, G. Campadelli-Fiume, E. Mocarski, P. S. Moore, B. Roizman, R. Whitley & K. Yamanishi. Cambridge.

Lund, T. C., Garcia, R., Medveczky, M. M., Jove, R. & Medveczky, P. G. (1997). Activation of STAT transcription factors by herpesvirus Saimiri Tip-484 requires p56lck. *Journal of virology* **71**, 6677-6682.

Lunn, J. S., Sakowski, S. A., Federici, T., Glass, J. D., Boulis, N. M. & Feldman, E. L. (2011). Stem cell technology for the study and treatment of motor neuron diseases. *Regenerative Medicine* **6**, 201-213.

Macfarlan, T. S., Gifford, W. D., Driscoll, S., Lettieri, K., Rowe, H. M., Bonanomi, D., Firth, A., Singer, O., Trono, D. & Pfaff, S. L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **advance online publication**.

Macnab, S., White, R., Hiscox, J. & Whitehouse, A. (2008). Production of an infectious Herpesvirus saimiri-based episomally maintained amplicon system. *Journal of Biotechnology* **134**, 287-296.

Macnab, S. A., Turrell, S. J., Carr, I. M., Markham, A. F., Coletta, P. L. & Whitehouse, A. (2011). Herpesvirus saimiri-mediated delivery of the adenomatous polyposis coli tumour suppressor gene reduces

proliferation of colorectal cancer cells. *International journal of oncology* **39**, 1173-1181.

Maitra, A., Arking, D. E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D. J., Liu, Y., Brimble, S. N., Noaksson, K., Hyllner, J., Schulz, T. C., Zeng, X., Freed, W. J., Crook, J., Abraham, S., Colman, A., Sartipy, P., Matsui, S., Carpenter, M., Gazdar, A. F., Rao, M. & Chakravarti, A. (2005). Genomic alterations in cultured human embryonic stem cells. *Nature genetics* **37**, 1099-1103.

Mani, S. A., Guo, W., Liao, M.-J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M., Campbell, L. L., Polyak, K., Brisken, C., Yang, J. & Weinberg, R. A. (2008). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell* **133**, 704-715.

Mao, X., Miesfeldt, S., Yang, H., Leiden, J. M. & Thompson, C. B. (1994). The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. *The Journal of biological chemistry* **269**, 18216-18222.

Marchetto, M. C. N., Yeo, G. W., Kainohana, O., Marsala, M., Gage, F. H. & Muotri, A. R. (2009). Transcriptional Signature and Memory Retention of Human-Induced Pluripotent Stem Cells. *PLoS ONE* **4**, e7076.

Marcinowski, L., Tanguy, M., Krmpotic, A., Radle, B., Lisnic, V. J., Tuddenham, L., Chane-Woon-Ming, B., Ruzsics, Z., Erhard, F., Benkartek, C., Babic, M., Zimmer, R., Trgovcich, J., Koszinowski, U. H., Jonjic, S., Pfeffer, S. & Dolken, L. (2012). Degradation of cellular mir-27 by a novel, highly abundant viral transcript is important for efficient virus replication in vivo. *PLoS pathogens* **8**, e1002510.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7634-7638.

Martin, G. R. & Evans, M. J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. *Cell* **2**, 163-172.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S. H. & Niwa, H. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nature Cell Biology* **9**, 625-U626.

Matsui, W., Wang, Q., Barber, J. P., Brennan, S., Smith, B. D., Borrello, I., McNiece, I., Lin, L., Ambinder, R. F., Peacock, C., Watkins, D. N., Huff, C. A. & Jones, R. J. (2008). Clonogenic Multiple Myeloma Progenitors, Stem Cell Properties, and Drug Resistance. *Cancer Research* **68**, 190-197.

McGeoch, D. J., Rixon, F. J. & Davison, A. J. (2006). Topics in herpesvirus genomics and evolution. *Virus Research* **117**, 90-104.

McTaggart, S. & Al-Rubeai, M. (2002). Retroviral vectors for human gene delivery. *Biotechnology Advances* **20**, 1-31.

Means, R. E. (2004). Characterization of the Herpesvirus saimiri Orf51 protein. *Virology* **326**, 67-78.

Medveczky, P., Szomolanyi, E., Desrosiers, R. C. & Mulder, C. (1984). CLASSIFICATION OF HERPESVIRUS SAIMIRI INTO 3 GROUPS BASED ON EXTREME VARIATION IN A DNA REGION REQUIRED FOR ONCOGENICITY. *Journal of virology* **52**, 938-944.

Melendez, L. V., Daniel, M. D., Hunt, R. D. & Garcia, F. G. (1968). An apparently new herpesvirus from primary kidney cultures of the squirrel monkey (Saimiri sciureus). *Laboratory Animal Care* **18**, 374-381.

Min, C., Eddy, S. F., Sherr, D. H. & Sonenshein, G. E. (2008). NF-kappaB and epithelial to mesenchymal transition of cancer. *Journal of Cellular Biochemistry* **104**, 733-744.

Mitchell, J. L., Trible, R. P., Emert-Sedlak, L. A., Weis, D. D., Lerner, E. C., Applen, J. J., Sefton, B. M., Smithgall, T. E. & Engen, J. R. (2007). Functional characterization and conformational analysis of the Herpesvirus saimiri Tip-C484 protein. *Journal of Molecular Biology* **366**, 1282-1293.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. & Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642.

Miyoshi, N., Ishii, H., Nagai, K.-i., Hoshino, H., Mimori, K., Tanaka, F., Nagano, H., Sekimoto, M., Doki, Y. & Mori, M. (2010). Defined factors induce reprogramming of gastrointestinal cancer cells. *Proceedings of the National Academy of Sciences* **107**, 40-45.

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D. L., Kano, Y., Nishikawa, S., Tanemura, M., Mimori, K., Tanaka, F., Saito, T., Nishimura, J., Takemasa, I., Mizushima, T., Ikeda, M., Yamamoto, H., Sekimoto, M., Doki, Y. & Mori, M. (2011). Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell stem cell* **8**, 633-638.

Morel, A. P., Lievre, M., Thomas, C., Hinkal, G., Ansieau, S. & Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE* **3**, e2888.

Mullor, J. L., Sanchez, P. & Ruiz i Altaba, A. (2002). Pathways and consequences: Hedgehog signaling in human disease. *Trends in Cell Biology* **12**, 562-569.

Murthy, S., Kamine, J. & Desrosiers, R. C. (1986). VIRAL-ENCODED SMALL RNAS IN HERPESVIRUS SAIMIRI INDUCED-TUMORS. *EMBO Journal* **5**, 1625-1632.

Muylaert, I. & Elias, P. (2007). Knockdown of DNA ligase IV/XRCC4 by RNA interference inhibits herpes simplex virus type I DNA replication. *Journal of Biological Chemistry* **282**, 10865-10872.

Naranatt, P. P., Krishnan, H. H., Smith, M. S. & Chandran, B. (2005). Kaposi's sarcoma-associated herpesvirus modulates microtubule dynamics via RhoA-GTP-diaphanous 2 signaling and utilizes the dynein motors to deliver its DNA to the nucleus. *Journal of virology* **79**, 1191-1206.

Neipel, F., Albrecht, J. C. & Fleckenstein, B. (1997). Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity? *Journal of virology* **71**, 4187-4192.

Nelson, T. J., Martinez-Fernandez, A., Yamada, S., Perez-Terzic, C., Ikeda, Y. & Terzic, A. (2009). Repair of Acute Myocardial Infarction by Human Stemness Factors Induced Pluripotent Stem Cells. *Circulation* **120**, 408-416.

Nguyen, H. N., Byers, B., Cord, B., Shcheglovitov, A., Byrne, J., Gujar, P., Kee, K., Schule, B., Dolmetsch, R. E., Langston, W., Palmer, T. D. & Pera, R. R. (2011). LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell stem cell* **8**, 267-280.

Nguyen, L. V., Vanner, R., Dirks, P. & Eaves, C. J. (2012). Cancer stem cells: an evolving concept. *Nature Reviews Cancer* **12**, 133-143.

Nicholas, J., Cameron, K. R. & Honess, R. W. (1992). HERPESVIRUS SAIMIRI ENCODES HOMOLOGS OF G-PROTEIN-COUPLED RECEPTORS AND CYCLINS. *Nature* **355**, 362-365.

Nicholas, J., Coles, L. S., Newman, C. & Honess, R. W. (1991). REGULATION OF THE HERPESVIRUS SAIMIRI (HVS) DELAYED-EARLY 110-KILODALTON PROMOTER BY HVS IMMEDIATE-EARLY GENE-PRODUCTS AND A HOMOLOG OF THE EPSTEIN-BARR VIRUS-R TRANS ACTIVATOR. *Journal of virology* **65**, 2457-2466.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. & Smith, A. (1998). Formation of

pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391.

Nishimoto, M., Fukushima, A., Okuda, A. & Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Molecular and cellular biology* **19**, 5453-5465.

Nordhoff, V., Hubner, K., Bauer, A., Orlova, I., Malapetsa, A. & Scholer, H. R. (2001). Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mammalian genome : official journal of the International Mammalian Genome Society* **12**, 309-317.

Nori, S., Okada, Y., Yasuda, A., Tsuji, O., Takahashi, Y., Kobayashi, Y., Fujiyoshi, K., Koike, M., Uchiyama, Y., Ikeda, E., Toyama, Y., Yamanaka, S., Nakamura, M. & Okano, H. (2011). Grafted human-induced pluripotent stem-cell-derived neurospheres promote motor functional recovery after spinal cord injury in mice. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16825-16830.

Noseda, M., McLean, G., Niessen, K., Chang, L., Pollet, I., Montpetit, R., Shahidi, R., Dorovini-Zis, K., Li, L., Beckstead, B., Durand, R. E., Hoodless, P. A. & Karsan, A. (2004). Notch Activation Results in Phenotypic and Functional Changes Consistent With Endothelial-to-Mesenchymal Transformation. *Circulation Research* **94**, 910-917.

O'Brien, C. A., Pollett, A., Gallinger, S. & Dick, J. E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106-110.

Ohno, T., Ouchida, M., Lee, L., Gatalica, Z., Rao, V. N. & Reddy, E. S. (1994). The EWS gene, involved in Ewing family of tumors, malignant melanoma of soft parts and desmoplastic small round cell tumors, codes for an RNA binding protein with novel regulatory domains. *Oncogene* **9**, 3087-3097.

Okita, K., Nakagawa, M., Hong, H. J., Ichisaka, T. & Yamanaka, S. (2008). Generation of Mouse Induced Pluripotent Stem Cells Without Viral Vectors. *Science* **322**, 949-953.

Okumura-Nakanishi, S., Saito, M., Niwa, H. & Ishikawa, F. (2005). Oct-3/4 and Sox2 Regulate Oct-3/4 Gene in Embryonic Stem Cells. *Journal of Biological Chemistry* **280**, 5307-5317.

Ovitt, C. E. & Schöler, H. R. (1998). The molecular biology of Oct-4 in the early mouse embryo. *Molecular Human Reproduction* **4**, 1021-1031.

Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M. M. & Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell stem cell* **6**, 382-395.

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**, 877-886.

Park, J., Cho, N. H., Choi, J. K., Feng, P., Choe, J. & Jung, J. U. (2003). Distinct roles of cellular Lck and p80 proteins in herpesvirus saimiri Tip function on lipid rafts. *Journal of virology* **77**, 9041-9051.

Pasi, C. E., Dereli-Oz, A., Negrini, S., Friedli, M., Fragola, G., Lombardo, A., Van Houwe, G., Naldini, L., Casola, S., Testa, G., Trono, D., Pelicci, P. G. & Halazonetis, T. D. (2011). Genomic instability in induced stem cells. *Cell Death and Differentiation* **18**, 745-753.

Pathmanathan, R., Prasad, U., Sadler, R., Flynn, K. & Raab-Traub, N. (1995). Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. *New England Journal of Medicine* **333**, 693-698.

Pedro Simas, J. & Efstathiou, S. (1998). Murine gammaherpesvirus 68: a model for the study of gammaherpesvirus pathogenesis. *Trends in Microbiology* **6**, 276-282.

Peinado, H., Olmeda, D. & Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nature Reviews Cancer* **7**, 415-428.

Pellet, P. & Roizman, B., (eds) (2007). *The Family Herpesviridae: A Brief Introduction*. Philadelphia: Lippincott, Williams and Wilkins.

Pick, M., Stelzer, Y., Bar-Nur, O., Mayshar, Y., Eden, A. & Benvenisty, N. (2009). Clone- and Gene-Specific Aberrations of Parental Imprinting in Human Induced Pluripotent Stem Cells. *Stem Cells* **27**, 2686-2690.

Postigo, A. A., Depp, J. L., Taylor, J. J. & Kroll, K. L. (2003). Regulation of Smad signaling through a differential recruitment of coactivators and corepressors by ZEB proteins. *EMBO Journal* **22**, 2453-2462.

Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., Weissman, I. L., Clarke, M. F. & Ailles, L. E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceedings of the National Academy of Sciences* **104**, 973-978.

Quintana, E., Shackleton, M., Sabel, M. S., Fullen, D. R., Johnson, T. M. & Morrison, S. J. (2008). Efficient tumour formation by single human melanoma cells. *Nature* **456**, 593-598.

Ragoczy, T. & Miller, G. (1999). Role of the epstein-barr virus RTA protein in activation of distinct classes of viral lytic cycle genes. *Journal of virology* **73**, 9858-9866.

Ramos-Mejia, V., Fraga, M. F. & Menendez, P. (2012). iPSCs from cancer cells: challenges and opportunities. *Trends in Molecular Medicine* **18**, 245-247.

Ramunas, J., Montgomery, H. J., Kelly, L., Sukonnik, T., Ellis, J. & Jervis, E. J. (2007). Real-time Fluorescence Tracking of Dynamic Transgene Variegation in Stem Cells. *Molecular Therapy* **15**, 810-817.

Rao, V. N., Ohno, T., Prasad, D. D., Bhattacharya, G. & Reddy, E. S. (1993). Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene* **8**, 2167-2173.

Redmond, D. E., Jr., Bjugstad, K. B., Teng, Y. D., Ourednik, V., Ourednik, J., Wakeman, D. R., Parsons, X. H., Gonzalez, R., Blanchard, B. C., Kim, S. U., Gu, Z., Lipton, S. A., Markakis, E. A., Roth, R. H., Elsworth, J. D., Sladek, J. R., Jr., Sidman, R. L. & Snyder, E. Y. (2007). Behavioral improvement in a primate Parkinson's model is associated with multiple homeostatic effects of human neural stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 12175-12180.

Reske, A., Pollara, G., Krummenacher, C., Chain, B. M. & Katz, D. R. (2007). Understanding HSV-1 entry glycoproteins. *Reviews in Medical Virology* **17**, 205-215.

Reya, T. & Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature* **434**, 843-850.

Ricci-Vitiani, L., Lombardi, D. G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C. & De Maria, R. (2007). Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111-115.

Richart, S. M., Simpson, S. A., Krummenacher, C., Whitbeck, J. C., Pizer, L. I., Cohen, G. H., Eisenberg, R. J. & Wilcox, C. L. (2003). Entry of herpes simplex virus type 1 into primary sensory neurons in vitro is mediated by Nectin-1/HveC. *Journal of virology* **77**, 3307-3311.

Richter, G. H. S., Plehm, S., Fasan, A., Rössler, S., Unland, R., Bennani-Baiti, I. M., Hotfilder, M., Löwel, D., von Luettichau, I., Mossbrugger, I., Quintanilla-Martinez, L., Kovar, H., Staege, M. S., Müller-Tidow, C. & Burdach, S. (2009). EZH2 is a mediator of EWS/FLI1 driven tumor growth and metastasis blocking endothelial and neuro-ectodermal differentiation. *Proceedings of the National Academy of Sciences* **106**, 5324-5329.

Riggi, N., Cironi, L., Provero, P., Suva, M. L., Kaloulis, K., Garcia-Echeverria, C., Hoffmann, F., Trumpp, A. & Stamenkovic, I. (2005). Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Research* **65**, 11459-11468.

Riggi, N. & Stamenkovic, I. (2007). The Biology of Ewing sarcoma. *Cancer Letters* **254**, 1-10.

Riggi, N., Suvà , M.-L., Suvà , D., Cironi, L., Provero, P., Tercier, S. p., Joseph, J.-M., Stehle, J.-C., Baumer, K., Kindler, V. & Stamenkovic, I. (2008). EWS-FLI-1 Expression Triggers a Ewing's Sarcoma Initiation Program in Primary Human Mesenchymal Stem Cells. *Cancer Research* **68**, 2176-2185.

Riggi, N., Suva, M. L., De Vito, C., Provero, P., Stehle, J. C., Baumer, K., Cironi, L., Janiszewska, M., Petricevic, T., Suva, D., Tercier, S., Joseph, J. M., Guillou, L. & Stamenkovic, I. (2010). EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. *Genes and Development* **24**, 916-932.

Robinton, D. A. & Daley, G. Q. (2012). The promise of induced pluripotent stem cells in research and therapy. *Nature* **481**, 295-305.

Rodda, D. J., Chew, J. L., Lim, L. H., Loh, Y. H., Wang, B., Ng, H. H. & Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. *The Journal of biological chemistry* **280**, 24731-24737.

Ryan, M. D. & Drew, J. (1994). FOOT-AND-MOUTH-DISEASE VIRUS 2A OLIGOPEPTIDE MEDIATED CLEAVAGE OF AN ARTIFICIAL POLYPROTEIN. *EMBO Journal* **13**, 928-933.

Sahlgren, C., Gustafsson, M. V., Jin, S., Poellinger, L. & Lendahl, U. (2008). Notch signaling mediates hypoxia-induced tumor cell migration and invasion. *Proceedings of the National Academy of Sciences* **105**, 6392-6397.

Sambrook, J., Maniatis, T. & Fritsch, E. (1989). *Molecular Cloning - A Laboratory Manual*. New York: Cold Spring Harbour Laboratory Press.

Sanchez, G., Delattre, O., Auboeuf, D. & Dutertre, M. (2008). Coupled alteration of transcription and splicing by a single oncogene: Boosting the effect on cyclin D1 activity. *Cell Cycle* **7**, 2299-2305.

Sayan, A. E., Griffiths, T. R., Pal, R., Browne, G. J., Ruddick, A., Yagci, T., Edwards, R., Mayer, N. J., Qazi, H., Goyal, S., Fernandez, S., Straatman, K., Jones, G. D., Bowman, K. J., Colquhoun, A., Mellon, J. K., Kriajevska, M. & Tulchinsky, E. (2009). SIP1 protein protects cells from DNA damage-induced apoptosis and has independent prognostic value in bladder cancer. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14884-14889.

Schafer, A., Lengenfelder, D., Grillhosl, C., Wieser, C., Fleckenstein, B. & Ensser, A. (2003). The latency-associated nuclear antigen homolog of herpesvirus saimiri inhibits lytic virus replication. *Journal of virology* **77**, 5911-5925.

Scharenberg, C. W., Harkey, M. A. & Torok-Storb, B. (2002). The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood* **99**, 507-512.

Schepers, A. G., Snippert, H. J., Stange, D. E., van den Born, M., van Es, J. H., van de Wetering, M. & Clevers, H. (2012). Lineage Tracing Reveals Lgr5+ Stem Cell Activity in Mouse Intestinal Adenomas. *Science*.

Scherdin, U., Rhodes, K. & Breindl, M. (1990). Transcriptionally active genome regions are preferred targets for retrovirus integration. *Journal of virology* **64**, 907-912.

Schirm, S., Müller, I., Desrosiers, R. C. & Fleckenstein, B. (1984). Herpesvirus saimiri DNA in a lymphoid cell line established by in vitro transformation. *Journal of virology* **49**, 938-946.

Shay, J. W. & Wright, W. E. (2010). Telomeres and telomerase in normal and cancer stem cells. *FEBS Letters* **584**, 3819-3825.

Shieh, M. T., Wudunn, D., Montgomery, R. I., Esko, J. D. & Spear, P. G. (1992). CELL-SURFACE RECEPTORS FOR HERPES-SIMPLEX VIRUS ARE HEPARAN-SULFATE PROTEOGLYCANS. *Journal of Cell Biology* **116**, 1273-1281.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992). CLONING AND STABLE MAINTENANCE OF 300-KILOBASE-PAIR FRAGMENTS OF HUMAN DNA IN ESCHERICHIA-COLI USING AN F-FACTOR-BASED VECTOR. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8794-8797.

Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X. M., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D. & Spear, P. G. (1999). A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **99**, 13-22.

Simmer, B., Alt, M., Buckreus, I., Berthold, S., Fleckenstein, B., Platzer, E. & Grassmann, R. (1991). PERSISTENCE OF SELECTABLE HERPESVIRUS SAIMIRI IN VARIOUS HUMAN HEMATOPOIETIC AND EPITHELIAL-CELL LINES. *Journal of General Virology* **72**, 1953-1958.

Singh, A. & Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* **29**, 4741-4751.

Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., Henkelman, R. M., Cusimano, M. D. & Dirks, P. B. (2004). Identification of human brain tumour initiating cells. *Nature* **432**, 396-401.

Skalsky, R. L., Samols, M. A., Plaisance, K. B., Boss, I. W., Riva, A., Lopez, M. C., Baker, H. V. & Renne, R. (2007). Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155. *Journal of virology* **81**, 12836-12845.

Smale, S. T. & Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annual Review of Biochemistry* **72**, 449-479.

Smith, P. G., Burchill, S. A., Brooke, D., Coletta, P. L. & Whitehouse, A. (2004). Efficient infection and persistence of a herpesvirus saimiri-based gene delivery vector into human tumor xenografts and multicellular spheroid cultures. *Cancer Gene Therapy* **12**, 248-256.

Smith, P. G., Coletta, P. L., Markham, A. F. & Whitehouse, A. (2001). In vivo episomal maintenance of a herpesvirus samiri-based gene delivery vector. *Gene Therapy* **8**, 1762-1769.

Smith, P. G., Oakley, F., Fernandez, M., Mann, D. A., Lemoine, N. R. & Whitehouse, A. (2005). Herpesvirus saimiri-based vector biodistribution using noninvasive optical imaging. *Gene Therapy* **12**, 1465-1476.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G. W., Cook, E. G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O. & Jaenisch, R. (2009). Parkinson's Disease Patient-Derived Induced Pluripotent Stem Cells Free of Viral Reprogramming Factors. *Cell* **136**, 964-977.

Sommer, C. A., Sommer, A. G., Longmire, T. A., Christodoulou, C., Thomas, D. D., Gostissa, M., Alt, F. W., Murphy, G. J., Kotton, D. N. & Mostoslavsky, G. (2010). Excision of Reprogramming Transgenes Improves the Differentiation Potential of iPS Cells Generated with a Single Excisable Vector. *Stem Cells* **28**, 64-74.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. (2008). Induced Pluripotent Stem Cells Generated Without Viral Integration. *Science* **322**, 945-949.

Staege, M. S., Hutter, C., Neumann, I., Foja, S., Hattenhorst, U. E., Hansen, G., Afar, D. & Burdach, S. E. G. (2004). DNA Microarrays Reveal Relationship of Ewing Family Tumors to Both Endothelial and Fetal Neural Crest-Derived Cells and Define Novel Targets. *Cancer Research* **64**, 8213-8221.

Stevenson, A. J., Clarke, D., Meredith, D. M., Kinsey, S. E., Whitehouse, A. & Bonifer, C. (2000a). Herpesvirus saimiri-based gene delivery vectors maintain heterologous expression throughout mouse embryonic stem cell differentiation in vitro. *Gene Therapy* **7**, 464-471.

Stevenson, A. J., Cooper, M., Griffiths, J. C., Gibson, P. C., Whitehouse, A., Jones, E. F., Markham, A. F., Kinsey, S. E. & Meredith, D. M. (1999). Assessment of Herpesvirus saimiri as a potential human gene therapy vector. *Journal of Medical Virology* **57**, 269-277.

Stevenson, A. J., Frolova-Jones, E., Hall, K. T., Kinsey, S. E., Markham, A. F., Whitehouse, A. & Meredith, D. M. (2000b). A herpesvirus saimiri-based gene therapy vector with potential for use in cancer immunotherapy. *Cancer Gene Therapy* **7**, 1077-1085.

Stevenson, A. J., Giles, M. S., Hall, K. T., Goodwin, D. J., Calderwood, M. A., Markham, A. F. & Whitehouse, A. (2000c). Specific oncolytic activity of herpesvirus saimiri in pancreatic cancer cells. *British Journal of Cancer* **83**, 329-332.

Studer, L., Tabar, V. & McKay, R. D. (1998). Transplantation of expanded mesencephalic precursors leads to recovery in parkinsonian rats. *Nature neuroscience* **1**, 290-295.

Sun, C. & Liu, Y. K. (2011). Induced pluripotent cancer cells: progress and application. *Journal of Cancer Research and Clinical Oncology* **137**, 1-8.

Suvà , M.-L., Riggi, N., Stehle, J.-C., Baumer, K., Tercier, S. p., Joseph, J.-M., Suvà , D., ClÃ©ment, V., Provero, P., Cironi, L., Osterheld, M.-C., Guillou, L. & Stamenkovic, I. (2009). Identification of Cancer Stem Cells in Ewing's Sarcoma. *Cancer Research* **69**, 1776-1781.

Swijnenburg, R. J., Schrepfer, S., Cao, F., Pearl, J. I., Xie, X., Connolly, A. J., Robbins, R. C. & Wu, J. C. (2008a). In vivo imaging of embryonic stem cells reveals patterns of survival and immune rejection following transplantation. *Stem cells and development* **17**, 1023-1029.

Swijnenburg, R. J., Schrepfer, S., Govaert, J. A., Cao, F., Ransohoff, K., Sheikh, A. Y., Haddad, M., Connolly, A. J., Davis, M. M., Robbins, R. C. & Wu, J. C. (2008b). Immunosuppressive therapy mitigates immunological rejection of human embryonic stem cell xenografts. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 12991-12996.

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**, 861-872.

Takahashi, K. & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676.

Taussig, D. C., Vargaftig, J., Miraki-Moud, F., Griessinger, E., Sharrock, K., Luke, T., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S. G., Lister, T. A., Gribben, J. G. & Bonnet, D. (2010). Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34− fraction. *Blood* **115**, 1976-1984.

Tchieu, J., Kuoy, E., Chin, M. H., Trinh, H., Patterson, M., Sherman, S. P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J. A., Clark, A. T., Pyle, A. D., Lowry, W. E. & Plath, K. (2010). Female Human iPSCs Retain an Inactive X Chromosome. *Cell stem cell* **7**, 329-342.

Terry, J. & Nielsen, T. (2010). Expression of CD133 in synovial sarcoma. *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* **18**, 159-165.

Thiery, J. P. (2003). Epithelial–mesenchymal transitions in development and pathologies. *Current Opinion in Cell Biology* **15**, 740-746.

Thomas, G. R., Faulkes, D. J., Gascoyne, D. & Latchman, D. S. (2004). EWS differentially activates transcription of the Brn-3a long and short isoform mRNAs from distinct promoters. *Biochemical and Biophysical Research Communications* **318**, 1045-1051.

Thome, M., Schneider, P., Hofmann, K., Fickenscher, H., Meinl, E., Neipel, F., Mattmann, C., Burns, K., Bodmer, J. L., Schroter, M., Scaffidi, C., Krammer, P. H., Peter, M. E. & Tschopp, J. (1997). Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. *Nature* **386**, 517-521.

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**, 1145-1147.

Tirino, V., Desiderio, V., Paino, F., De Rosa, A., Papaccio, F., Fazioli, F., Pirozzi, G. & Papaccio, G. (2011). Human primary bone sarcomas contain CD133+ cancer stem cells displaying high tumorigenicity in vivo. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **25**, 2022-2030.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M. & Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4–Sox-2 complex. *Nucleic Acids Research* **30**, 3202-3213.

Tsuji, O., Miura, K., Okada, Y., Fujiyoshi, K., Mukaino, M., Nagoshi, N., Kitamura, K., Kumagai, G., Nishino, M., Tomisato, S., Higashi, H., Nagai, T., Katoh, H., Kohda, K., Matsuzaki, Y., Yuzaki, M., Ikeda, E., Toyama, Y., Nakamura, M., Yamanaka, S. & Okano, H. (2010). Therapeutic potential of appropriately evaluated safe-induced pluripotent stem cells for spinal cord injury. *Proceedings of the National Academy of Sciences* **107**, 12704-12709.

Tunnicliffe, R. B., Hautbergue, G. M., Kalra, P., Jackson, B. R., Whitehouse, A., Wilson, S. A. & Golovanov, A. P. (2011). Structural Basis for the Recognition of Cellular mRNA Export Factor REF by Herpes Viral Proteins HSV-1 ICP27 and HVS ORF57. *PLoS pathogens* **7**, e1001244.

Turrell, S. J. (2010). Development of Herpesvirus saimiri as a cancer gene therapy vector. In *Faculty of Biological Sciences*: University of Leeds.

Turrell, S. J., Macnab, S. A., Rose, A., Melcher, A. A. & Whitehouse, A. (2012). A Herpesvirus saimiri-based vector expressing TRAIL induces cell death in human carcinoma cell lines and multicellular spheroid cultures. *International journal of oncology* **40**, 2081-2089.

Turrell, S. J. & Whitehouse, A. (2011). Mutation of herpesvirus Saimiri ORF51 glycoprotein specifically targets infectivity to hepatocellular carcinoma cell lines. *Journal of Biomedicine and Biotechnology* **2011**, 785158.

Utikal, J., Maherali, N., Kulalert, W. & Hochedlinger, K. (2009). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. *Journal of Cell Science* **122**, 3502-3510.

van den Berg, D. L., Snoek, T., Mullin, N. P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I. & Poot, R. A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. *Cell stem cell* **6**, 369-381.

van Es, J. H., van Gijn, M. E., Riccio, O., van den Born, M., Vooijs, M., Begthel, H., Cozijnsen, M., Robine, S., Winton, D. J., Radtke, F. & Clevers, H. (2005). Notch/[gamma]-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* **435**, 959-963.

Varnum-Finney, B., Xu, L., Brashem-Stein, C., Nourigat, C., Flowers, D., Bakkour, S., Pear, W. S. & Bernstein, I. D. (2000). Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling. *Nature Medicine* **6**, 1278-1281.

Varnum, S. M., Streblow, D. N., Monroe, M. E., Smith, P., Auberry, K. J., Pasa-Tolic, L., Wang, D., Camp, D. G., 2nd, Rodland, K., Wiley, S., Britt, W., Shenk, T., Smith, R. D. & Nelson, J. A. (2004). Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *Journal of virology* **78**, 10960-10966.

Verma, S. C. & Robertson, E. S. (2003). ORF73 of herpesvirus saimiri strain C488 tethers the viral genome to metaphase chromosomes and

binds to cis-acting DNA sequences in the terminal repeats. *Journal of virology* **77**, 12494-12506.

Verschueren, K., Remacle, J. E., Collart, C., Kraft, H., Baker, B. S., Tylzanowski, P., Nelles, L., Wuytens, G., Su, M.-T., Bodmer, R., Smith, J. C. & Huylebroeck, D. (1999). SIP1, a Novel Zinc Finger/Homeodomain Repressor, Interacts with Smad Proteins and Binds to 5′-CACCT Sequences in Candidate Target Genes. *Journal of Biological Chemistry* **274**, 20489-20498.

Wade-Martins, R., White, R. E., Kimura, H., Cook, P. R. & James, M. R. (2000). Stable correction of a genetic deficiency in human cells by an episome carrying a 115 kb genomic transgene. *Nature Biotechnology* **18**, 1311-1314.

Wakamatsu, Y., Sakamoto, N., Oo, H. Z., Naito, Y., Uraoka, N., Anami, K., Sentani, K., Oue, N. & Yasui, W. (2012). Expression of cancer stem cell markers ALDH1, CD44 and CD133 in primary tumor and lymph node metastasis of gastric cancer. *Pathology International* **62**, 112-119.

Walter, D., Satheesha, S., Albrecht, P., Bornhauser, B. C., D'Alessandro, V., Oesch, S. M., Rehrauer, H., Leuschner, I., Koscielniak, E., Gengler, C., Moch, H., Bernasconi, M., Niggli, F. K. & Schafer, B. W. (2011). CD133 positive embryonal rhabdomyosarcoma stem-like cell population is enriched in rhabdospheres. *PLoS ONE* **6**, e19506.

Walters, M. S., Hall, K. T. & Whitehouse, A. (2004). The Herpesvirus Saimiri Open Reading Frame (ORF) 50 (Rta) Protein Contains an AT Hook Required for Binding to the ORF 50 Response Element in Delayed-Early Promoters. *Journal of virology* **78**, 4936-4942.

Wang, F. Z., Akula, S. M., Pramod, N. P., Zeng, L. & Chandran, B. (2001). Human herpesvirus 8 envelope glycoprotein K8.1A interaction with the target cells involves heparan sulfate. *Journal of virology* **75**, 7517-7527.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W. & Orkin, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368.

Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. (2012). Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells. *Cell stem cell* **10**, 440-454.

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**, 618-630.

Wasungu, L. & Hoekstra, D. (2006). Cationic lipids, lipoplexes and intracellular delivery of genes. *Journal of Controlled Release* **116**, 255-264.

Wernig, M., Zhao, J.-P., Pruszak, J., Hedlund, E., Fu, D., Soldner, F., Broccoli, V., Constantine-Paton, M., Isacson, O. & Jaenisch, R. (2008). Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. *Proceedings of the National Academy of Sciences* **105**, 5856-5861.

Whitbeck, J. C., Peng, C., Lou, H., Xu, R., Willis, S. H., Ponce de Leon, M., Peng, T., Nicola, A. V., Montgomery, R. I., Warner, M. S., Soulika, A. M., Spruce, L. A., Moore, W. T., Lambris, J. D., Spear, P. G., Cohen, G. H. & Eisenberg, R. J. (1997). Glycoprotein D of herpes simplex virus (HSV) binds directly to HVEM, a member of the tumor necrosis factor receptor superfamily and a mediator of HSV entry. *Journal of virology* **71**, 6083-6093.

White, R. E., Calderwood, M. A. & Whitehouse, A. (2003). Generation and precise modification of a herpesvirus saimiri bacterial artificial chromosome demonstrates that the terminal repeats are required for

both virus production and episomal persistence. *Journal of General Virology* **84**, 3393-3403.

Whitehouse, A., Carr, I. M., Griffiths, J. C. & Meredith, D. M. (1997). The herpesvirus Saimiri ORF50 gene, encoding a transcriptional activator homologous to the Epstein-Barr virus R protein, is transcribed from two distinct promoters of different temporal phases. *Journal of virology* **71**, 2550-2554.

Whitehouse, A., Cooper, M. & Meredith, D. M. (1998). The Immediate-Early Gene Product Encoded by Open Reading Frame 57 of Herpesvirus Saimiri Modulates Gene Expression at a Posttranscriptional Level. *Journal of virology* **72**, 857-861.

Wieser, C., Stumpf, D., Grillhosl, C., Lengenfelder, D., Gay, S., Fleckenstein, B. & Ensser, A. (2005). Regulated and constitutive expression of anti-inflammatory cytokines by nontransforming herpesvirus saimiri vectors. *Gene Therapy* **12**, 395-406.

Williams, B. J. L., Boyne, J. R., Goodwin, D. J., Roaden, L., Hautbergue, G. M., Wilson, S. A. & Whitehouse, A. (2005). The prototype gamma-2 herpesvirus nucleocytoplasmic shuttling protein, ORF 57, transports viral RNA through the cellular mRNA export pathway. *Biochemical Journal* **387**, 295-308.

Wilson, K. D., Venkatasubrahmanyam, S., Jia, F., Sun, N., Butte, A. J. & Wu, J. C. (2009). MicroRNA profiling of human-induced pluripotent stem cells. *Stem cells and development* **18**, 749-758.

Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **11**, 228-234.

Woltjen, K., Michael, I. P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P. T., Gertsenstein, M., Kaji, K., Sung, H. K. & Nagy, A. (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**, 766-U106.

Xu, D., Alipio, Z., Fink, L. M., Adcock, D. M., Yang, J., Ward, D. C. & Ma, Y. (2009). Phenotypic correction of murine hemophilia A using an iPS cell-based therapy. *Proceedings of the National Academy of Sciences* **106**, 808-813.

Xue, Z., Yan, H., Li, J., Liang, S., Cai, X., Chen, X., Wu, Q., Gao, L., Wu, K., Nie, Y. & Fan, D. (2012). Identification of cancer stem cells in vincristine preconditioned SGC7901 gastric cancer cell line. *Journal of Cellular Biochemistry* **113**, 302-312.

Yamashita, T., Ji, J., Budhu, A., Forgues, M., Yang, W., Wang, H. Y., Jia, H., Ye, Q., Qin, L. X., Wauthier, E., Reid, L. M., Minato, H., Honda, M., Kaneko, S., Tang, Z. Y. & Wang, X. W. (2009). EpCAM-Positive Hepatocellular Carcinoma Cells Are Tumor-Initiating Cells With Stem/Progenitor Cell Features. *Gastroenterology* **136**, 1012-1024.e1014.

Young, L. S., Dawson, C. W., Clark, D., Rupani, H., Busson, P., Tursz, T., Johnson, A. & Rickinson, A. B. (1988). Epstein--Barr Virus Gene Expression in Nasopharyngeal Carcinoma. *Journal of General Virology* **69**, 1051-1065.

Yu, J., Chau, K. F., Vodyanik, M. A., Jiang, J. & Jiang, Y. (2011). Efficient Feeder-Free Episomal Reprogramming with Small Molecules. *PLoS ONE* **6**, e17557.

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, II & Thomson, J. A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**, 797-801.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I. & Thomson, J. A. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* **318**, 1917-1920.

Yuan, H., Corbi, N., Basilico, C. & Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes and Development* **9**, 2635-2645.

Yusa, K., Rad, R., Takeda, J. & Bradley, A. (2009). Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nature Methods* **6**, 363-369.

Zhang, L., Rayner, S., Katoku-Kikyo, N., Romanova, L. & Kikyo, N. (2007). Successful co-immunoprecipitation of Oct4 and Nanog using cross-linking. *Biochemical and Biophysical Research Communications* **361**, 611-614.

Zhang, P., Andrianakos, R., Yang, Y., Liu, C. & Lu, W. (2010). Kruppel-like Factor 4 (Klf4) Prevents Embryonic Stem (ES) Cell Differentiation by Regulating Nanog Gene Expression. *Journal of Biological Chemistry* **285**, 9180-9189.

Zhang, S., Balch, C., Chan, M. W., Lai, H.-C., Matei, D., Schilder, J. M., Yan, P. S., Huang, T. H.-M. & Nephew, K. P. (2008). Identification and Characterization of Ovarian Cancer-Initiating Cells from Primary Human Tumors. *Cancer Research* **68**, 4311-4320.

Zhao, T., Zhang, Z.-N., Rong, Z. & Xu, Y. (2011). Immunogenicity of induced pluripotent stem cells. *Nature* **474**, 212-215.

Zhou, B. B., Zhang, H., Damelin, M., Geles, K. G., Grindley, J. C. & Dirks, P. B. (2009a). Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nature reviews Drug discovery* **8**, 806-823.

Zhou, H., Wu, S., Joo, J. Y., Zhu, S., Han, D. W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Schöler, H. R., Duan, L. & Ding, S. (2009b). Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins. *Cell stem cell* **4**, 381-384.

Zhu, F. X., Chong, J. M., Wu, L. & Yuan, Y. (2005). Virion proteins of Kaposi's sarcoma-associated herpesvirus. *Journal of virology* **79**, 800-811.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K. & Ding, S. (2010). Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. *Cell stem cell* **7**, 651-655.

Zielke, K., Full, F., Teufert, N., Schmidt, M., Muller-Fleckenstein, I., Alberter, B. & Ensser, A. (2012). The insulator protein CTCF binding sites in the orf73/LANA promoter region of herpesvirus saimiri are involved in conferring episomal stability in latently infected human T cells. *Journal of virology* **86**, 1862-1873.