HSV-1 amplicon system for human artificial chromosome formation in human ES/iPS cells and pluripotency induction

A Thesis Submitted for the Degree
of Doctor of Philosophy
Trinity Term 2012

Suhail Khoja
Keble College
Nuffield Department of Clinical Medicine
University of Oxford

Declaration

I, Suhail Khoja, confirm that the work presented in this thesis is conducted by myself. Where information has been derived from other sources, I confirm that this has been mentioned in the thesis. I also confirm that no part of this thesis has been submitted for another degree at this or any other university.

Dedication

This thesis is dedicated to my mother Shireen, my father Late Sikander Ali and my brother Sulleman for their endless love.

Acknowledgements

First of all, I would like to acknowledge Dr. Zoia Larin Monaco for providing me the chance to be a part of her lab, her supervision and continuous support throughout my DPhil studies. I am sincerely grateful to the Clarendon Fund and Keble Sloane Robinson Scholarship at Oxford for their financial support towards my studies and maintenance. I am indeed very appreciative of Dr. Daniela Moralli for all the discussions, assistance in experimental analysis and data interpretation as well as for thesis review and corrections.

Many thanks to my collaborators Dr. Sally Cowley, Prof. William James, Prof. Anthony Monaco and Dr. Emanuela Volpi, to the members of the Larin Monaco's lab Drs. Mohammad Mandegar and David Chan for being a great company and also to the members of Prof. Anthony Monaco's lab.

I am greatly indebted to Dr. Aziz Mithani for his review and helpful comments on the thesis and more importantly for being with me. My sincere acknowledgement to Prof. Chris Tyler-Smith and Prof. Ioannis Ragoussis who provided helpful suggestions and on the basis of these various corrections and improvements were made to the thesis.

Finally, I am very grateful to my family and friends for their constant support, patience and encouragement during the difficult times while reading DPhil at Oxford.

Abstract

HSV-1 amplicon system for human artificial chromosome formation in human ES/iPS cells and pluripotency induction

Suhail Khoja
Keble College

Thesis presented to the University of Oxford for the degree of Doctor of Philosophy
Nuffield Department of Clinical Medicine
Trinity Term 2012

Development of safe and efficient approaches for gene delivery in human embryonic stem cells (hESc) and particularly in human induced pluripotent stem (hiPS) cells, which can be derived in a person-specific manner, is considered to be imperative for harnessing their full potential in both the basic and applied research. The aim of this study was to evaluate the potential of human artificial chromosome (HAC) for gene delivery and expression in hESc and hiPS cells. HAC offers many potential advantages including the provision for carrying large genes with corresponding regulatory elements to obtain long-term regulated gene expression. In addition, they can replicate and segregate independently without integration into the host cell genome. To develop HAC in hiPS cells, the first part of the study was aimed at generating hiPS cells utilising the Herpes Simplex Virus (HSV) -1 amplicon system. With the use of EBNA-1/OriP retention elements incorporated into the HSV-1 amplicon vectors, hiPS cells completely free of vector and transgenes sequences were successfully derived from human embryonic fibroblasts. The hiPS cells exhibited proliferation and differentiation potential similar to that of hESc. In the second part of the study, development of HAC in hESc and hiPS cells was assessed by utilising the HSV-1 amplicon system to deliver the HAC DNA. Analysis of the hESc confirmed the presence of functional HAC which replicated the behaviour of the host chromosomes. Additionally, HAC generation did not lead to impairment in the developmental potential and pluripotency of hESc. The hiPS cells supported HAC at low frequency but DNA also integrated into the host chromosomes. The HAC system, therefore, needs further refinements to improve the frequency of HAC formation and reduce the chromosomal integration of HAC constructs in hiPS cells. Overall, these findings provide a simple and safe way of pluripotency induction and genetic modification of pluripotent stem cells using the HSV-1 amplicon system and represent an important advance towards patient-specific gene and cell therapy.

Table of Contents

List of Tables

List of Figures

Abbreviations

α	Alphoid/alpha DNA
AAV	Adeno Associated Virus
Amp^R	Ampicillin Resistance
BAC	Bacterial Artificial Chromosome
bFGF	basic Fibroblast Growth Factor
BHS	Buffered Hypotonic Solution
BMP	Bone Morphogenetic Protein
bp	base pairs
BSA	bovine serum albumin
ºC	degrees Centigrade
cDNA	complementary DNA
CENP	Centromere Protein
Cm	centimetres
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CHEF1α	Chinese Hamster Elongation Factor-1 α
CHO	Chinese Hamster Ovary
CMV	Cytomegalovirus
CpG	Cytosine-phosphatidyl-Guanosine
C_T	Threshold Cycle
cTc	chlortetracycline
DABCO	1,4-Diazabicyclo(2,2,2)octane
DAPI	4,6-Diamidino-2-Phenylindole
DMD	Duchenne Muscular Dystrophy
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
DNMTs	DNA Methyltransferases
dNTPs	Deoxynucleotide Triphosphates
dUTP	2'-Deoxyuridine 5'-Triphosphate
E. coli	*Escherichia coli*
EB	Embryoid Body
EBNA-1	Epstein Barr Nuclear Antigen-1
EBV	Epstein Barr Virus
EC	Embryonic carcinoma
EGFP	Enhanced Green Fluorescent Protein
FACS	Fluorescence-Activated Cell Sorting
FBS	Foetal Bovine Serum
FCS	Foetal Calf Serum
FISH	Fluorescence *in situ* Hybridisation
FITC	Fluorescein Isothiocyanate
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green Fluorescent Protein
GMP	Good Manufacturing Practice
hAAT	human α1 anti-trypsin
HAC	Human Artificial Chromosomes
HDACs	Histone Deacetylase Enzymes

hEF1α	human Elongation Factor 1 α
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hESc	human embryonic stem cells
HLA	Human leukocyte antigen
HPRT	Hypoxanthine Phosphoribosyltransferase
HSC	Hematopoietic Stem Cells
HSV-1	Herpes Simplex Virus Type 1
ICM	Inner Cell Mass
IFN	Interferons
iMEFs	inactivated-mouse embryonic fibroblasts
INCENP	Inner Centromere Protein
iPSc	induced Pluripotent Stem cells
kb	kilo bases
KCM	Potassium chromosome medium
LB	Luria–Bertani
LIF	Leukaemia inhibitory factor
LTR	Long Terminal Repeat
µg	micrograms
µl	microlitres
µM	micromolar
M	Molar
M-FISH	Multicolour FISH
Mb	Mega base
MEF	Mouse Embryonic Fibroblasts
mESc	mouse Embryonic Stem cells
mg	milligrams
ml	millilitres
mM	millimolar
MMCT	Microcell Mediated Chromosome Transfer
MOI	Multiplicity of Infection
mRNA	messenger RNA
miRNA	microRNA
MSC	Mesenchymal Stem Cells
NDM	Neuronal Differentiation Media
Neo^R	Neomycin Resistance
ng	nanograms
NIH	National Institutes of Health
NT	Nuclear transfer
OHS	Optimal Hypotonic Solution
ORF	Open reading frame
PAC	P1 Artificial Chromosome
PBS	Phosphate Buffered Saline
PcG	Ploycomb group Proteins
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PFGE	Pulsed Field Gel Electrophoresis
pg	picograms
R26	ROSA26
RNA	Ribonucleic Acid
ROCK	Rho-associated Coiled-coil Kinase

rpm	Revolutions per minute
RT	room temperature
SCID	Severe Combined Immunodeficiency
SCNT	Somatic Cell Nuclear Transfer
SSC	Saline sodium citrate
SV40	Simian Virus 40
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TNF	Tumour Necrosis Factor
TRITC	Tetramethyl Rhodamine Isothiocyanate
UbC	Ubiquitin C
UV	Ultraviolet
YAC	Yeast Artificial Chromosome
µl	microlitres
µm	micromitres

Chapter 1. Introduction

During the mammalian development, over 200 specialised cell types are specified from a single totipotent cell zygote. Division and developmental proceeding of the zygote leads to a gradual loss of potency as it undergoes inevitable differentiation events. Throughout these developmental stages, stem cells play a crucial role in the maintenance of tissue homeostasis by generating new cells, differentiating into other cell types and replacing aged or injured cells. For example, multipotent stem cells (which have the ability to self renew and give rise to closely related cells) such as haematopoietic stem cells can differentiate into red and white blood cells (Boiani and Scholer, 2005). Similarly, pluripotent stem cells are the cells that possess the ability for indefinite self renewal and the potential to differentiate into specialised cell types derived from each of the three embryonic germ layers (mesoderm, endoderm and ectoderm) (Donovan and Gearhart, 2001; Thomson et al., 1998). The self renewal ability and differentiation potential has made these cells a valuable research tool for studying early embryonic development as well as a potential therapeutic source to treat various diseases such as Parkinson's disease, juvenile diabetes, heart failure, and spinal cord injuries. Moreover, they hold a great promise for the field of drug screening and determining their associated toxicology (Friedrich Ben-Nun and Benvenisty, 2006; Moon et al., 2006; Thomson et al., 1998).

In mammalian organisms, pluripotent stem cells have been categorised into four types on the basis of their derivation: (1) embryonic stem (ES) cells are derived from the inner cell mass (ICM) of blastocyst-stage (pre-implantation) embryos (Thomson et al., 1998); (2) embryonic germ (EG) cells are derived from primordial germ cells (PGCs) of the post-implantation embryos (Shamblott et al., 1998), (3) reprogrammed cells, which are called

reprogrammed pluripotent stem (RePS) cells in this thesis, are derived from differentiated somatic cells by genetic manipulation or cell fusion procedures (Jaenisch and Young, 2008) and (4) embryonic carcinoma (EC) cells are the stem cells isolated from germ cell tumours (also called teratocarcinomas) (Andrews et al., 2001). These pluripotent stem cell populations have been recognised as a natural reservoir to generate all the cell types in an organism and, have, therefore become a highly valued research tools (Cowan et al., 2004). Human embryonic stem cells (hESc) and the recently generated human induced pluripotent stem (hiPS) cells represent the most widely-known and well-studied types of pluripotent stem cells, as they are considered to render a broad spectrum of clinical prospects and are, therefore, described in detail below.

1.1 Human Embryonic Stem Cells

Human embryonic stem cells (hESc or hES cells) are the pluripotent stem cells that have been derived from the pre-implantation (blastocyst-stage) embryo and are capable of proliferating indefinitely by self-renewal and differentiating into specialised cells types of all the three embryonic germ layers upon induction (Thomson et al., 1998). The cells isolated from the gastrula (the developmental stage after the blastocyst), on the other hand, are multipotent and are only capable of differentiating into the cells representing one of the three germ layers. hESc possess great potential for use in basic and applied clinical research. Their self-renewal ability, high replicative lifespan and differentiation capability renders them very useful for studying early embryogenesis and prospects in drug discovery, tissue regeneration, and cell-replacement therapies (Cowan et al., 2004; Reubinoff et al., 2000; Thomson et al., 1998).

1.1.1 Derivation of Human ES Cells

The fertilisation of a secondary oocyte (resulting after the first meiotic division) by the sperm leads to the formation of a diploid zygote, which contains two pronuclei (Figure 1-1). The resulting zygote (impregnated oocyte) subsequently undergoes cleavage divisions to generate 2-cell embryo, 4-cell embryo and so on to become morula (containing ~8-40 totipotent cells called blastomeres). The blastomeres are encapsulated by a glycoprotein membrane called the zona pellucida, which also surrounds the plasma membrane of the oocyte. The morula undergoes further mitotic cleavage divisions to form a blastocyst. The blastocyst contains ~100-150 cells and consists of a fluid-filled cavity, called the blastocoel, and two distinct layers of cells called the embryoblast and the trophoblast. The embryoblast, also called the inner cell mass (ICM), is formed by the innermost blastomeres in morula and consists of an inner layer of pluripotent cells which contribute to the formation of embryo. The trophoblast is made by the outermost blastomeres in morula cells and forms the surrounding layer of cells (trophectoderm). The trophoblast develops into the extra-embryonic tissues such as placenta, chorion and the umbilical cord (Vazin and Freed, 2010).

hES cells are mainly derived from the ICM of blastocyst stage-embryos (Figure 1-1). Separation of the ICM from the trophoblast is normally performed using the procedure called immuno-surgery, in which the zona pellucida is firstly removed (Cowan et al., 2004; Reubinoff et al., 2000) followed by elimination of the trophoblast cells by exposure to an anti-human primary antibody and secondary complement serum (Thomson et al., 1998). The ICM is then transferred onto inactivated mouse embryonic fibroblasts feeder cells and grown in human ES culture media to establish hES cell lines (Cowan et al., 2004; Reubinoff et al., 2000; Solter and Knowles, 1975). To reduce the use of animal-

derived products in the immunosurgical methods, which could possibly transfer animal pathogens to the hESc, ICM isolation has also been carried out with mechanical and enzymatic digestion of the trophectoderm (Amit and Itskovitz-Eldor, 2002; Genbacev et al., 2005; Strom et al., 2007). However, the derivation of hESc from ICM at blastocyst-stage requires destruction of a viable embryo and has posed serious ethical concerns. To overcome these, recent studies have focused on the derivation of hESc without destruction of embryos, at earlier stages of embryonic development (Figure 1-1) from late morula (30-40 cells) (Strelchenko et al., 2004), arrested embryos (16 to 24 cell embryos) (Zhang et al., 2006) and single blastomeres from 8-cell stage embryos (Chung et al., 2008; Klimanskaya et al., 2006).

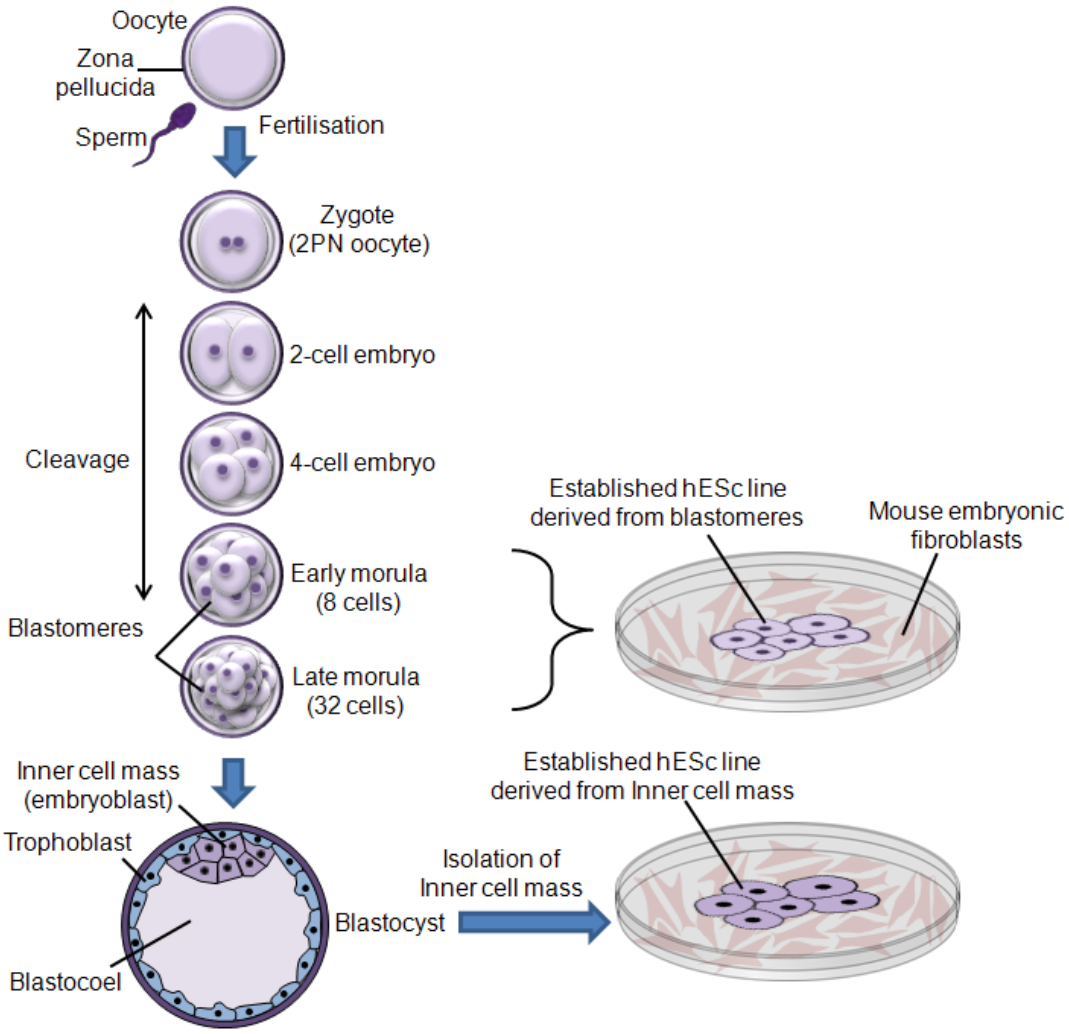

Figure 1-1 Derivation of human embryonic stem cells
Fertilisation of the secondary oocyte by the sperm results in the formation of zygote, which undergoes cleavage divisions to form a structure known as morula. Further divisions result in the formation of the blastocyst, which consists of the trophoblast, the blastocoel and the inner cell mass (ICM; also called embryoblast). Derivation of hESc has been mostly carried out by isolation and culture of ICM in the presence of mouse embryonic fibroblasts feeder cells but can also be achieved from arrested embryos and late morula stage embryos.

1.1.2 Characteristics of hES Cells

hESc grow as tightly-packed distinct colonies with compact borders, have high nucleus to cytoplasmic ratio, and prominent nucleoli. The hESc exhibit high levels of telomerase activity, providing them with an indefinite replicative lifespan (Thomson et al., 1998). They also express high levels of the transcription factors *OCT4* (also called *POU5F1*),

SOX2 and *NANOG*, which play a key role in the maintenance of pluripotency, and suppression of genes that exert differentiation by closely interacting with signalling pathways and epigenetic regulators (Boyer et al., 2005). In addition, the hESc express cell surface glycolipid antigens such as stage specific embryonic antigens (SSEA)-3, SSEA-4, the cell surface keratan sulphate antigens TRA-1-60, TRA-1-81 and the tissue-nonspecific alkaline phosphatase (AP) (Adewumi et al., 2007). The hESc have the ability of continuous self-renewal and have the potential to differentiate into all cell types except extra-embryonic tissues, which are formed from trophoblast cells. This developmental potential can be tested, *in vitro*, by subjecting the embryoid bodies (EBs; three dimensional aggregates of ES cells) to differentiation conditions (Itskovitz-Eldor et al., 2000) and *in vivo*, by transplanting the ES cells into immunocompromised mice to generate teratomas, tumours containing differentiated derivatives of all the three embryonic germ layers (Thomson et al., 1998).

1.1.3 *In vitro* Culture of hES Cells

The hESc require special growth conditions for continuous self renewal and maintenance in an undifferentiated state. Initially, this was accomplished by deriving and propagating hES cells on mitotically inactive mouse embryonic fibroblasts (iMEFs) feeder layers (Reubinoff et al., 2000; Thomson et al., 1998). iMEFs produce the growth factors and signalling factors required for self renewal and undifferentiated growth of hES cells. Mouse embryonic fibroblasts are derived from 13.5 days post-coitum fetuses obtained by sacrificing the pregnant female mice and are prepared freshly at low passage number to avoid senescence. Prior to their use as feeder layers, MEFs are mitotically inactivated by mitomycin C treatment or γ-irradiation to prevent the dilution of hESc with dividing MEFs (Nagy et al., 2003). As the use of animal-derived feeder layers poses the risk of

contamination of hESc with animal proteins and viruses, making these cells unsuitable for clinical purposes (Martin et al., 2005), various efforts have been made to define xeno-free culture conditions such as feeder- and serum-free environments. Extracellular matrices such as Matrigel and fibronectin have been demonstrated to support hESc culture in a feeder-free manner, in MEF conditioned media (Xu et al., 2001). Matrigel is the most commonly used basement membrane matrix, derived from mouse sarcoma cells and mainly consists of laminin, collagen IV, and heparan sulfate proteoglycan (Vukicevic et al., 1992; Xu et al., 2001). Recent studies, based on determining the released factors from feeder layer and serum components, have also made it possible to culture hES cells in completely defined media such as commercially available mTeSR and StemPro formulations to eliminate the need of MEF-conditioned media (Chin et al., 2007; Lim and Bodnar, 2002; Ludwig et al., 2006).

Passaging of hESc is another major challenge because upon cellular dissociation, the cells undergo apoptosis, which leads to massive cell death. Amit and colleagues reported a cloning efficiency of less than 1% when hESc were dissociated in a single cell suspension using trypsin and plated on feeder cells (Amit et al., 2000). Because of these difficulties, initial propagation and expansion of hESc was mostly carried out through mechanical passaging, which involved fragmenting the colonies into small clumps using glass Pasteur pipettes and subsequent transfer of the clumps to MEF feeder cells. The mechanical passaging was not only laborious and time consuming but also posed serious limitations to differentiation, genetic manipulation and subcloning of hES cells (Eiges et al., 2001). Recently, the enzymatic dissociation and passaging of hESc has been facilitated by the application of p160-Rho-associated coiled-coil kinase (ROCK) inhibitor Y-27632 into the hESc medium following trypsinisation. Although the exact mechanism of action of

ROCK inhibitor is not clearly understood, its use has greatly reduced the dissociation-induced apoptosis of hESc, increasing their cloning efficiency to ~27% after trypsinisation (Watanabe et al., 2007).

Presently, cultivation of hESc in suspension culture is also an active area of research as adhesion culture limits the yield of hESc required for clinical and industrial settings. Recent studies have demonstrated the successful development and requirements for suspension culture system of hESc, although it may take considerable time and effort before these methods become a common practice (Amit et al., 2010; Oh et al., 2009).

1.1.4 Karyotypic Integrity of hES Cells

The hESc are expected to maintain long-term chromosomal stability with continuous culture conditions (Hanson and Caisander, 2005). However, many independent studies have reported the development of karyotypic abnormalities during prolonged culture and periodic screening of hES cells (Baker et al., 2007; Draper et al., 2004; Mitalipova et al., 2005; Spits et al., 2008). Some of these chromosomal rearrangements are believed to be recurring such as trisomies of chromosomes 12 (particularly gain of 12p), 17 (particularly gain of 17q) and X, and may provide adaptive and selective advantage during the *in vitro* maintenance of hES cells. As a result, the cells with chromosomal abnormalities proliferate more rapidly than their normal counterparts and gradually prevail and overcome the entire normal population (Baker et al., 2007; Draper et al., 2004; Spits et al., 2008). *BIRC5* (also called *Survivin*; encodes an apoptosis inhibitor), *GRB7* (associated with germ cell tumours), *STAT3* and *GRB2* (both are implicated with regulation of self renewal) are some of the candidate genes believed to confer proliferative advantage, and are all located on chromosome 17q. Similarly *NANOG*, which is one of essential

regulatory genes involved in maintaining pluripotency, lies on chromosome 12 (Draper et al., 2004). In addition, gain of whole or parts of chromosome 1, 3, 7, 8, 9, 14 and 20 have also been observed in long-term hESc culture but most of these and other abnormalities are regarded as secondary as the cells involved also carried trisomy 12, 17 or both. Although the factors responsible for accumulation of these chromosomal rearrangements in hESc have not been clearly understood, some studies have indicated that enzymatic passaging of hESc is responsible for more karyotypic abnormalities as compared to mechanical passaging (Baker et al., 2007; Spits et al., 2008). As enzymatic passaging (involving the use of collagenase or trypsin), which dissociates the hESc colonies into small clumps or single cells, causes massive apoptosis, it could increase the stress or impose selective pressure on individual surviving cells to harbour such chromosomal abnormalities (Baker et al., 2007). In addition to passaging technique, the duration of cells in culture along with the presence and absence of feeder layers may be contributing for acquiring these karyotypic abnormalities (Caisander et al., 2006). Besides this, some hESc lines are inherently more inclined towards karyotypic instabilities compared to others (Catalina et al., 2008). Thus, regular screening of cells for karyotypic abnormalities together with the investigation of factors that contribute and prevent their re-occurrence would be vital if hESc are to be used in any clinical and therapeutic settings.

1.2 Control of Pluripotent State

Molecular mechanisms that control the regulation of pluripotency in hESc and other pluripotent cell types, such as function of transcription factors, epigenetic regulators and signalling pathways are important in developmental studies because their understanding is central to the understanding of development. As defects in development can lead to various diseases, determining and understanding the regulatory mechanisms involved in

governing pluripotency, is crucial to harness the therapeutic potential of pluripotent stem cells.

The molecular machinery of ES cells allows them to maintain pluripotency and self renewal capacity. At the same time, it provides ES cells with the capability for differentiation, which is negatively regulated by pluripotency factors and positively regulated by lineage-specific factors and signals. Core transcription factors, namely OCT4, NANOG and SOX2, appear to be the most important regulatory elements in ES cells. In conjunction with other transcription factors, OCT4, NANOG and SOX2 operate closely to regulate various genes involved in developmental signalling pathways (Young, 2011). In addition, epigenetic modulators, signal transduction pathways and miRNAs play crucial role in the regulation of pluripotency.

1.2.1 Transcription Factors

Oct4 (or POU5F1) and Nanog have been identified as essential regulators of self-renewal and pluripotency in both human and mouse ES cells. Several studies in mouse developmental genetics have indicated their distinguished roles, but at the same time they also function mutually in establishing the pluripotent state. For instance, the loss of *Oct4* and *Nanog* leads to the differentiation of the inner cell mass (ICM) into trophectoderm. Moreover, disruption of *Oct4* and *Nanog* expression is known to cause differentiation of ES cells into extra embryonic endoderm (Chambers et al., 2003; Nichols et al., 1998; Ying et al., 2002). Similarly, disruption of *OCT4* and *NANOG* in human ES cells through RNA interference (RNAi) has been shown to cause marked reduction in the expression of cell surface-specific hES markers, and to lead to hESc differentiation into trophectoderm and early endodermal lineages (Hay et al., 2004; Hyslop et al., 2005; Matin et al., 2004;

Zaehres et al., 2005). Oct4 also forms heterodimer with the high mobility group (HMG) box transcription factor Sox2 in both mouse and human ES cells, thus highlighting the importance of *Sox2* as a key regulator in pluripotency (Masui et al., 2007; Rodda et al., 2005). Moreover, *SOX2* knockdown in hESc by RNAi has been shown to cause differentiation, decreased expression of key stem cells markers including *OCT4* and *NANOG*, and increased expression of trophectoderm markers (Fong et al., 2008).

1.2.2 Core Transcriptional Circuit

To further elucidate the mechanisms by which OCT4, NANOG and SOX2 regulate pluripotency in hESc, various studies have identified the downstream target genes of these three regulators in both mouse and human ES cells (Boyer et al., 2005; Loh et al., 2006). Genome-wide location analysis (chromatin immune-precipitation) coupled with DNA microarray (ChIP-chip) in hESc has revealed that OCT4, NANOG and SOX2 form an interconnected autoregulatory loop by binding to the promoters of their own genes and also to the promoters of the other two genes. Moreover, OCT4, NANOG and SOX2 co-occupy promoters of 353 genes (both transcriptionally active and inactive) and at least two miRNA genes forming a feed-forward loop (Boyer et al., 2005; Loh et al., 2006). These studies indicated that the core regulatory circuit of hESc required the collaborative function of various proteins for the maintenance of self renewal, pluripotency and differentiation capabilities. Of the 353 co-bound genes, hESc-associated transcription factors (e.g., *REST*, *SKIL*, *HESX1*, *STAT3*, *ZIC3*), components of the two central hES cell signalling pathways TGF-β (e.g., *TDGF1*, *LEFTY2*/*EBAF*) and Wnt (e.g., *DKK1*, *FRAT2*), and components of chromatin remodelling (e.g., *SMARCAD1*) and histone-modifying complexes (e.g., *MYST3* and *SET*) are the notable transcriptionally active targets. Transcriptionally inactive targets co-occupied by OCT4, NANOG and SOX2

include important transcription factors involved in lineage specification (extra-embryonic, ectoderm, mesoderm, and endoderm) such as *ESX1L*, *HOXB1*, *MEIS1*, *PAX6*, *LHX5*, *HAND1*, *MYF5*, *ONECUT1*, and *ATBF1*(Boyer et al., 2005; Loh et al., 2006; Young, 2011).

1.2.3 Epigenetic Control

Chromatin structure and the inherited chromatin signatures play a vital role during mammalian development. ES cells generally maintain an open chromatin structure (transcriptionally permissive), through the co-ordinated function of chromatin regulators (proteins involved in histone modification and nucleosome mobilisation), which are mainly recruited to the genes by developmental specific transcription factors (Guenther et al., 2010; Young, 2011).

The Polycomb group (PcG) proteins are known to be involved in chromatin modification and epigenetic silencing by co-occupying several developmental regulators that are also bound by OCT4, NANOG and SOX2 (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006). In ES cells, PcG proteins produce gene silencing by forming two multiprotein complexes: Polycomb repressive complex (PRC) 1 and PRC2. PRC1 facilitates chromatin condensation and inhibition of chromatin remodelling by monoubiquitylation of lysine 119 on histone H2A (H2A-K119). PRC2 catalyses methylation of lysine 27 on histone H3 (H3K27me3), a repressive epigenetic mark that may serve as PRC1 binding site (Lee et al., 2006; Orkin and Hochedlinger, 2011). Recent studies have also revealed that PRC1 and PRC2 bound various genes involved in lineage specification (Boyer et al., 2006; Lee et al., 2006). These genes, though silent, maintained a bivalent chromatin domains in ES cells because of the presence of both histone H3 methylated at lysine 4 (H3K4me3;

associated with active gene expression) and histone H3 methylated at lysine 27 (H3K27me3; associated with repression) marks, and therefore remained poised for expression upon differentiation (Bernstein et al., 2006).

Antagonistic to the role of PcG group proteins in repression (associated with H3K27me3), the Trithorax group (TrxG) complexes help in the maintenance of active gene expression by catalysing H3K4me3 at the promoters of these genes (Schuettengruber et al., 2007). The role of Trx group proteins in the pluripotency and self-renewal mechanisms was recently established in the mouse ES cells through the study of Wdr5, a core member of mammalian TrxG complex, which interacted with Oct4 and collaborated in the transcriptional activation of key regulators involved in self-renewal (Ang et al., 2011).

1.2.4 Signalling Pathways

Extracellular signals render a vital contribution to the maintenance of pluripotency and self-renewal as well as in stimulating differentiation into specific lineages. Soluble factors such as basic fibroblast growth factor (bFGF), leukaemia inhibitory factor (LIF), Wnt and activin/nodal play an important part in maintaining pluripotency of ES cells, but their role could vary depending on the species and culture conditions (Okita and Yamanaka, 2006; Rossant, 2008). In hESc, bFGF signalling appears to be critical in promoting self-renewal (Greber et al., 2010) and has been suggested to promote cell proliferation, inhibit spontaneous differentiation, boost cell survival, and most importantly to contribute in the induction of *NANOG* expression. As a result of bFGF stimulation, mouse embryonic fibroblasts (MEFs) secrete Activin A (a TGF-β family ligand), which assists in the triggering of the transcription factor *SMAD2/3* (James et al., 2005; Vallier et al., 2005).

SMAD2/3 docks at the *NANOG* promoter and thereby drives its expression, indicating the important relationship of signal transduction pathways and transcriptional circuitry in hES cells pluripotency (Greber et al., 2008; Xu et al., 2008).

Wnt-mediated signalling pathway is another important system involved in maintaining the self-renewal and pluripotency in both mouse and human ES cells (Cai et al., 2007; Ogawa et al., 2006; Sato et al., 2004) and has also been shown to enhance proliferation of hESc (Dravid et al., 2005). Similarly, the TGF-β superfamily contains important signalling regulators associated with both maintaining pluripotency and inducing differentiation. Activin, nodal, bone morphogenetic proteins (BMPs), growth differentiation factors (GDFs), LEFTY1 and LEFTY2 are some of the key signalling proteins belonging to the TGF-β family. While the contribution of Activin A in maintaining the ES cell state by the activation of *NANOG* expression has already been described above, nodal signalling is required for both pluripotency and differentiation (Pera and Tam, 2010) where as a BMP member (BMP4) is known for inducing the differentiation of human ES cells (Xu et al., 2002).

1.2.5 miRNA Regulation

Micro RNAs (miRNAs) are short non-coding RNAs that are involved in post-transcriptional repression mainly by binding and destabilising the target coding RNAs. Several studies have proposed vital contribution of miRNAs in the control of early development as well in the maintenance of the stem cell state of both mouse and human ES cells (Houbaviy et al., 2005; Houbaviy et al., 2003; Mineno et al., 2006; Suh et al., 2004). Mouse ES cells deficient for the key miRNA processing enzymes Dicer and Drosha (or its essential cofactor Dgcr8), exhibits severe defects in growth, self-renewal

and differentiation (Pauli et al., 2011). Proliferative defects of *Dgcr8*-knockout ES cells can be rescued by ES cell-specific miR-290-295 cluster and its relative miRNA family members (Wang et al., 2008). These miRNAs and other ES-cell specific miRNAs are also known to be positively regulated by pluripotency transcription factors Oct4, Nanog, Sox2 and Tcf3. The same pluripotency factors co-occupy the genetic loci of differentiation-specific miRNAs such as let-7 and miR-145, along with Polycomb group proteins, thus mediating their transcriptional repression (Marson et al., 2008). A distinct feature of miRNA-based regulation is the establishment of a negative feedback inhibition loop. For example, miR-145 targets the transcripts of pluripotency factors *OCT4*, *SOX2* and *KLF4*, while let-7 targets Lin28 and Sall4, which results in rapid repression of these target mRNAs upon differentiation stimuli (Melton et al., 2010; Xu et al., 2009).

1.3 Clinical Significance of Human ES Cells

The self-renewal ability and differentiation potential has made hESc a valuable research tool for studying early human embryonic development as well as a potential therapeutic tool for various diseases (Friedrich Ben-Nun and Benvenisty, 2006; Moon et al., 2006; Thomson et al., 1998). Disease-specific hESc allow screening for potential drugs, their efficacy and toxicity, before they can be tested in animal models. In therapeutic settings, hESc have been proposed to be used for tissue regeneration and cell-replacement therapies in various pathological conditions such as spinal cord injuries, juvenile-onset diabetes mellitus, heart failure, Parkinson's disease and age-related macular degeneration (Cowan et al., 2004; Li et al., 2008; Reubinoff et al., 2000; Singec et al., 2007; Thomson et al., 1998). To harness the potential of ES cells, various protocols have been developed to successfully differentiate ES cells into a broad spectrum of therapeutically significant cell types, including cardiomyocytes, pancreatic cells, dopamine and motor neurons,

oligodendrocytes, and hematopoietic precursor cells (Murry and Keller, 2008). The therapeutic potential of these ES-induced differentiated cells has successfully been shown in animal models. For example, mouse and human ES cell derived dopamine neurons have been used to alleviate the symptoms in the rat-model of Parkinson disease (Kim et al., 2002; Yang et al., 2008); ES cell derived hematopoietic stem cells have been used to treat severe combined immunodeficiency (SCID)-mice (Rideout et al., 2002); and hESc derived retinal cells have been shown to successfully restore light responses in blind mice (Lamba et al., 2009).

Although the use of ES cells in animal models has given some exciting results, there are significant challenges for the effective translation of these therapeutic applications to human patients. First, there is little understanding of the complex signalling system and factors required for the specific differentiation of hESc into functional cell lineages. Moreover, current methodologies to differentiate ES cells into the desired cell types are not highly efficient and therefore there is a need to improve these efficiencies or design strategies to isolate specific cell populations (Moon et al., 2006). Second, the current methods used to maintain hES cells and differentiate them into specific cell types often use heterologous feeder cells and protein mediators (Schuldiner et al., 2000; Thomson et al., 1998). Third, as described in section 1.1.4, ES cells can develop karyotypic abnormalities during long-term *in vitro* culture. Fourth, the issue of hESc or hESc derived differentiated cells maintaining the required cell status and not converting to teratoma or other unwanted cells after transplantation into patients is yet to be addressed (Moon et al., 2006; Murry and Keller, 2008).

Apart from these technical challenges, other obstructions to hESc therapy persist, including the use of viable human embryos in the derivation of hESc, which has raised

many ethical and political debates over the last decade and has significantly slowed the progress of this field (Evans, 2005; Fischbach and Fischbach, 2004). Finally, tissue rejection remains a significant concern after transplantation of hESc and hESc-derived differentiated cells into patients as non-compatible cells with the host are recognised as 'foreign' and, consequently, rejected (Moon et al., 2006). One way to circumvent these non-technical challenges is to reprogram the nuclei of patient specific somatic cells to an ES-like state and using them for therapeutic applications. Reprogramming not only generates cells that are compatible with the host but also avoids the ethical issues associated with the generation and use of hESc.

1.4 Reprogramming

Loss of potency was formerly considered as a one-way irreversible process, associated with the progressive loss of genetic material from the differentiated cells. This notion is now overturned as many studies on animal-developmental genetics have proved that differentiation occurs as a result of differential gene expression, epigenetic modifications, and signalling environment, thus establishing that all the cells in an individual, with the exception of few cell types, are genetically equivalent. This genetic equivalency can therefore be utilised to induce one cell type into any other cell type (transdifferentiation) or specifically into pluripotent stem cells (reprogramming).

Pluripotent stem cells, particularly hESc are highly significant because of their natural potency to generate all the cell types that can be used as cellular replacements in regenerative diseases or damaged tissues (Cowan et al., 2004). However, derivation of patient-specific hESc is normally not possible as their derivation process and culture prevents the normal proceedings of the embryonic development. Reprogramming

provides an alternative solution as somatic cells can be derived in a patient-specific manner and their epigenetic status can be reversed to produce cells at a state resembling to hESc. Presently, four methods are available to reprogram differentiated somatic cells back to the pluripotent state (Figure 1-2). These include somatic cell nuclear transfer (SCNT), cell fusion with ES cells, cell explantation and induced expression of specific genes (Jaenisch and Young, 2008).

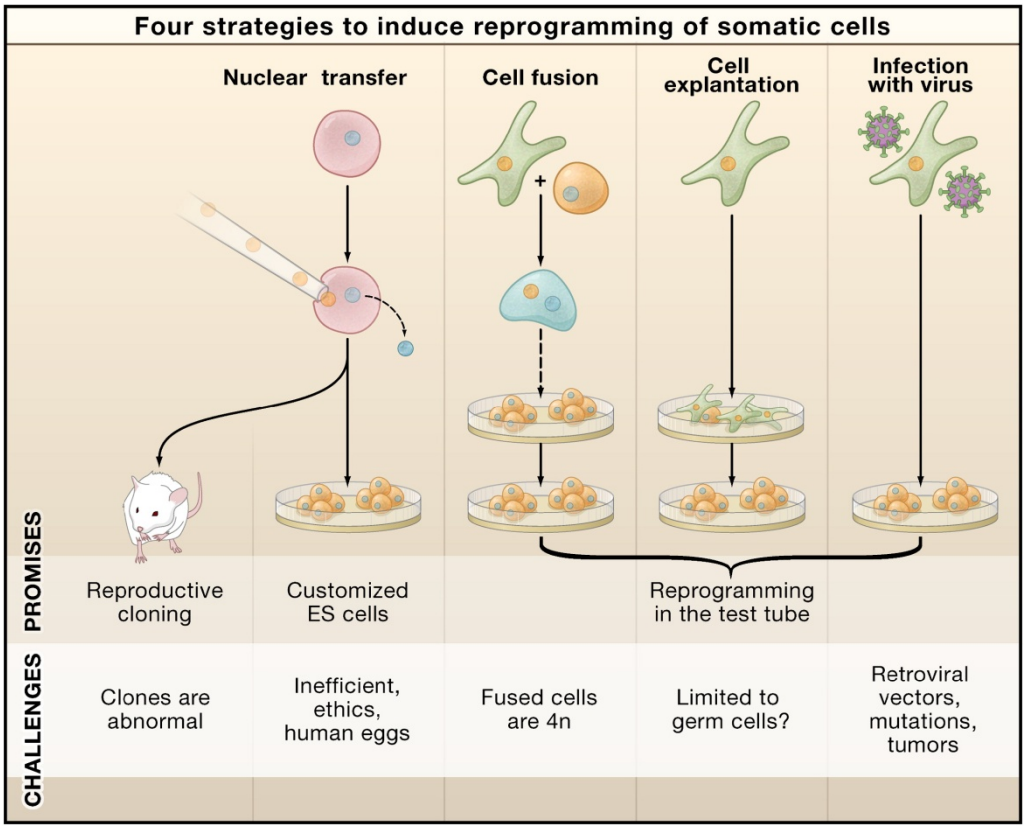

Figure 1-2 Four strategies to induce reprogramming in somatic cells
1) In nuclear transfer, the nucleus from a somatic cell is injected into an enucleated oocyte, which can give rise to reproductive clone upon transfer into a surrogate mother, or, can generate genetically matched ES cells if cultured in ES media. (2) Fusion of somatic cells with ES cells generates cell hybrids that show all the characteristics of pluripotent ES cells. (3) Explantation and subsequent prolonged culture of different types of germ cells results in immortal cell lines that may be pluripotent or multipotent depending on the source of donor cells. (4) Induced expression of defined factors into somatic cells can initiate reprogramming to a pluripotent state.
Figure taken from (Jaenisch and Young, 2008).

1.4.1 Nuclear Transfer

Nuclear transfer (NT) or somatic cell nuclear transfer (SCNT) is a technique in which the nucleus from a donor cell is introduced into enucleated oocytes. The resulting cells can develop into an animal clone if transplanted into the uterus of a surrogate mother or can generate genetically identical ES-like cells upon explantation into culture (Jaenisch and Young, 2008) (Figure 1-2). This technique was devised by Briggs and King, who transferred the blastula cell nuclei into the enucleated eggs of the frog species *Rana pipiens* and generated cloned tadpoles (Briggs and King, 1952). Using a slightly modified e hROSA26 locus followed by the incor*poration of th*e target gene into the first intron of the hROSA26 region. This would allow developing a system t□□□□□□□□□□□□□ture fertile frogs (Gurdon et al., 1958). However, both of these studies indicated that efficiency of obtaining clones was reduced when the donor nuclei were obtained from later stages of development.

Mammalian NT remained technically difficult for over 30 years until Wilmut and colleagues cloned Dolly the sheep by transferring the nuclei from epithelial cells into the enucleated oocytes (Wilmut et al., 1997). Other animal species including mice (Wakayama et al., 1998), cows (Kato et al., 1998), and pigs (Onishi et al., 2000; Polejaeva et al., 2000) were subsequently cloned through NT. However, reproductive cloning of mammals is highly inefficient and most of the embryos die during post-implantation development due to a variety of pathologies. Those that survive also suffer from various abnormalities including premature death, placental defects, obesity and respiratory problems (Cibelli et al., 2002; Hochedlinger and Jaenisch, 2006; Yang et al., 2007). On the other hand, nuclear transfer-generated ES (ntES) cell lines can be derived efficiently and are considered identical to embryo-derived ES cell lines (Brambrink et al.,

2006). Patient-specific ntES cell lines could be induced to differentiate into disease-related cells, which could be utilised for drug screening analysis or could potentially be used for patient-specific cell-based therapy. However, SCNT has never been accomplished to generate ntES cells from enucleated human oocytes. Only last year, was the derivation of nuclear transfer-derived pluripotent stem cells from human oocytes was achieved but the investigators had to leave the oocyte nucleus in place thus resulting in triploid population of cells (Noggle et al., 2011). Although these triploid cells can be utilised for research purposes, they would not be suitable for cell therapy applications. Moreover, ethical issues associated with the use of human oocytes in nuclear transfer have further curtailed the use of nTES cells in therapeutic applications.

1.4.2 Cell Fusion

Somatic cells can be epigenetically reprogrammed upon fusion with pluripotent ES cells (Figure 1-2). The first demonstration of reprogramming in this fashion was by Miller and Ruddle, who fused mouse thymocytes with mouse EC cells (testicular teratoma) to generate pluripotent hybrid cells (Miller and Ruddle, 1976). Similar results were later obtained upon fusion with mouse ES cells (Tada et al., 2001). Importantly, human somatic cells such as fibroblasts and hematopoietic cells have also been reprogrammed upon fusion with hESc (Cowan et al., 2005; Yu et al., 2006). Although these hybrids acquire the epigenetic status resembling ES cells including reactivation of pluripotency genes, it is unclear whether the somatic genome is completely reprogrammed or not (Kimura et al., 2004).

Molecular mechanisms underlying ES cell-fusion mediated reprogramming are not clearly defined and it is still unclear whether the necessary factors responsible for

reprogramming are located in nucleus, cytoplasm or both compartments of ES cells (Do and Scholer, 2004; Strelchenko et al., 2006). As the process involves fusion, the resulting cell-cell hybrids contain two complete sets of chromosomes. Consequently, the use of these tetraploid cells in cell-based therapy prevents their clinical applications. Efforts have been made to selectively remove the ES-chromosome complement, but eliminating the complete ES-genome while retaining the somatic cell set of chromosomes looks unfeasible at this stage (Matsumura et al., 2007).

1.4.3 Cell Explantation

Prolonged *in vitro* culture of cells from the germ cell lineage such as primordial germ cells (EG) or spermatogonial stem cell derived multipotent adult germ-line stem cells (maGSCs) has been shown to result in pluripotent cells (Figure 1-2) (Guan et al., 2006; Kanatsu-Shinohara et al., 2004). At present, only EG cells and spermatogonial stem cells, isolated from postnatal animals, can be used to generate pluripotent cells in this fashion. While these pluripotent cells are capable of *in vitro* differentiation, only a few of the resulting cells form teratomas, generate chimeras and contribute to germline in mice depending on the source of donor cells (Jaenisch and Young, 2008).

1.4.4 Reprogramming with Defined Factors

Differentiated cells can also be reprogrammed by forced expression of a few transcriptional factors. These reprogrammed cells are called induced pluripotent stem (iPS) cells (Figure 1-2). The first study that demonstrated this mechanism of reprogramming used retroviral-mediated transduction of four transcription factors, *Oct4*, *Sox2*, *Klf4* and *c-myc*, into mouse fibroblasts. Upon selection of pluripotency markers, the fibroblast turned into ES-like pluripotent cells (Takahashi and Yamanaka, 2006). Later,

human fibroblasts were reprogrammed to pluripotent state with the use of the human versions of same four transcription factors (Takahashi et al., 2007), and also with slightly different combination of genes including *NANOG* and *LIN28* instead of *KLF4* and *C-MYC* (Yu et al., 2007). The exact role of these six genes in inducing pluripotency is yet to be determined. A recent study of genome-wide associations of these transcription factors and their promoter occupancy during various stages of reprogramming has, however, revealed that regulatory regions of various genes linked to pluripotency are co-occupied by these transcription factors (Sridharan et al., 2009).

Initial studies demonstrated the generation of iPS cells from fibroblasts (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007) but recent reports have documented that iPS cells can be derived from a wide variety of cells types including melanocytes (Utikal et al., 2009a), adipose stem cells (Sun et al., 2009), peripheral blood cells (Loh et al., 2009), pancreatic β-cells (Stadtfeld et al., 2008a), stomach and liver cells (Aoi et al., 2008), keratinocytes (Maherali et al., 2008), astrocytes (Ruiz et al., 2010), neural stem cells (Eminli et al., 2008; Kim et al., 2008b) and mature B lymphocytes (Hanna et al., 2008), thus providing the flexibility with the choice of cell type used for reprogramming. However, these reports also indicate that as in the NT, the reprogramming of terminally differentiated cells is less efficient than cells at earlier developmental stages.

1.5 Human Induced Pluripotent Stem Cells

1.5.1 Generation of iPS Cells

iPS cells were initially generated using retroviral/lentiviral vectors to deliver transcription factors into the fibroblasts (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). These vectors permanently integrate into the host genome and therefore pose the risk of insertional mutagenesis. Moreover, reactivation of transgenes such as *c-myc* (a known oncogene) and LTR-mediated activation of proto-oncogenes could induce tumourigenesis, making the iPS cells unsafe for use in therapeutic applications (Hacein-Bey-Abina et al., 2003b; Okita et al., 2007). Subsequent studies were directed to improve these derivation methods or used alternative methods to avoid genomic integration of viruses, generate iPS cells from various cells types and increase efficiency of iPS derivation. Methods to derive integration-free iPS cells can be classified into three groups: (1) those that use non-integrating vectors; (2) those that use integrating but excisable vectors; and (3) those that use non-DNA delivery approaches (Table 1-1).

Table 1-1 Methods for reprogramming human somatic cells to iPS cells

Type of Vector	Delivery method	Cell types	Advantages	Disadvantages	Ref
Integrating	Retroviral	Fibroblasts, adipose cells, blood cells, hepatocytes, keratinocytes, astrocytes	Reasonably efficient	Integration; incomplete transgene silencing	1
	Lentiviral	Fibroblasts, cord blood cells, adipose stem cells, hepatocytes	Reasonably efficient; also targets non-dividing cells	Integration; incomplete transgene silencing	2
	Inducible lentiviral	Fibroblasts, blood cells, keratinocytes	Reasonably efficient; controlled expression of factors	Integration; requires transactivator expression	3
Excisable	Transposon	Embryonic fibroblasts	Reasonably efficient; integration can be excised	Labour -intensive excision	4
	loxP-flanked lentiviral	Fibroblasts	Reasonably efficient; integration can be excised	Labour- intensive excision; retention of *loxP* sites	5
Non-integrating	Adenoviral	Fibroblasts	No integration	Low efficiency; requires multiple transductions	6
	Plasmids, minicircle vectors	Fibroblasts, blood cells, adipose stem cells	Very rare integration	Low efficiency; possible integration	7
Non-DNA	Sendai Virus	Fibroblasts, blood derived T-cells	No integration	Sequence sensitive viral RNA replicase; difficult to eliminate viruses from the cells	8
	Proteins	Fibroblasts	No integration	Low efficiency; requires multiple treatments;	9
	Synthetic mRNAs	Fibroblasts	High efficiency; no integration; faster kinetics	requires multiple treatments	10
	miRNA (mir-200c, mir-302s, mir-369s)	Fibroblasts, adipose stromal cells	No integration; faster kinetics	Lower efficiency than commonly used methods	11

References: 1 (Aasen et al., 2008; Giorgetti et al., 2009; Kim et al., 2009b; Liu et al., 2010; Loh et al., 2009; Ruiz et al., 2010; Sugii et al., 2010; Takahashi et al., 2007); 2 (Anokye-Danso et al., 2011; Haase et al., 2009; Sun et al., 2009; Yu et al., 2007); 3 (Loh et al., 2010; Maherali et al., 2008; Park et al., 2008a); 4 (Woltjen et al., 2009); 5 (Soldner et al., 2009; Somers et al., 2010); 6 (Zhou and Freed, 2009); 7 (Chou et al., 2011; Hu et al., 2011; Jia et al., 2010; Si-Tayeb et al., 2010; Yu et al., 2009); 8 (Fusaki et al., 2009; Seki et al., 2010); 9 (Kim et al., 2009a); 10 (Warren et al., 2010); 11 (Miyoshi et al., 2011).

The first examples of iPS cells obtained from non-integrating vectors were derived from mouse hepatocytes with adenoviral vectors (Stadtfeld et al., 2008c) and from mouse embryonic fibroblasts by transfection of plasmid vectors (Okita et al., 2008). These experiments demonstrated that integration of transgenes is not essential and transient expression of the reprogramming factors is sufficient to induce reprogramming. Later, human cells were also reprogrammed using adenoviral vectors (Zhou and Freed, 2009), Sendai viruses (Fusaki et al., 2009), minicircle vectors (Jia et al., 2010) and oriP/EBNA1-based self-replicating episomal vectors. However, reprogramming efficiencies with these non-integrating approaches are very low (~0.001%) as compared to methods using integrating systems (0.1%-1%). This is possibly because expression of these transgenes is not maintained for the complete duration of reprogramming. To overcome this problem, methods have been developed that are based on integration and subsequent removal of transgenes, once the cells are completely reprogrammed. In one study, integrated transgenes were successfully excised through Cre/LoxP recombination leaving behind the residual vector sequences (Kaji et al., 2009; Soldner et al., 2009). A second method used introduction and seamless excision of a piggyBac transposon from the host genome using transposase (Woltjen et al., 2009; Yusa et al., 2009). Although iPS cells produced through this later approach are both vector and transgene-free, the removal of multiple transposons is labour-intensive and necessitates rigorous characterisation of integration sites before and after the removal of transposons.

To further avoid integration concerns, recent studies have also reported the derivation of iPS cells with non-DNA based methods. In 2009, Ding and colleagues reprogrammed MEFs with direct transduction of reprogramming factors as recombinant proteins (Zhou et al., 2009). Similarly, human fibroblasts were also reprogrammed by using a human

immunodeficiency virus transactivator of transcription (HIV-TAT) protein segment to deliver the reprogramming factors as recombinant proteins across the cell membrane (Kim et al., 2009a). The main drawbacks of the protein-based reprogramming method are the short half-life of the proteins, necessitating their repeated administration and difficulty in protein purification methods. A safer and more efficient way to generate integration-free iPS cells was recently described by Warren and colleagues, who derived the iPS cells from different human cell types using repeated administration of synthetic modified mRNAs encoding the reprogramming factors (Warren et al., 2010). Last year, iPS derivation was successfully demonstrated by transfecting mature miRNAs into the human fibroblasts and thus completely avoiding the use of reprogramming factors (Miyoshi et al., 2011).

Various groups have performed screening of small molecules and chemical compounds to find substances able to enhance the overall low efficiency of reprogramming. A number of molecules and chemical compounds such as valproic acid (VPA), 5-aza-cytidine and sodium butyrate have been identified that significantly improve the efficiency of iPS generation (Desponts and Ding, 2010; Li and Ding, 2010). Moreover, some of these chemicals have also been shown to replace one or two reprogramming factors, thus raising the possibility of generating iPS cells solely by using small molecules. For example VPA has been demonstrated to replace the function of C-MYC and KLF4 to some extent (Huangfu et al., 2008). However, many of these chemical compounds are known to interfere with DNA and chromatin modifications and therefore could induce genetic and epigenetic aberrations into the resulting iPS cells (Desponts and Ding, 2010; Li and Ding, 2010).

1.5.2 Mechanism of Reprogramming

Various studies on understanding the mechanism of transcription factor-mediated reprogramming of mouse fibroblast cells have demonstrated that it is a gradual and sequential process during which the somatic programme is converted to the ES cell programme (Brambrink et al., 2008; Stadtfeld et al., 2008b). The events occurring during reprogramming are divided into early, middle and late phases. The early reprogramming phase is initiated with an increase in proliferation rate and reduction in size of the cells, and also with down-regulation of somatic cell-specific transcription programme (Koche et al., 2011; Plath and Lowry, 2011). During the middle phase, fibroblasts undergo a process, referred to as mesenchymal to epithelial transition (MET), during which the cells acquire epithelial characteristics such as tight intercellular space and expression of the key epithelial cell surface marker E-cadherin (Li et al., 2010b; Samavarchi-Tehrani et al., 2010). This notion has also been supported by studies showing that inhibition of signalling pathways, particularly that of transforming growth factor-β (TGF-β), suppress MET inhibition and consequently improve the reprogramming efficiency (Li et al., 2010b). Recent studies have indicated that reprogramming factors play a distinct and significant role during MET (Li et al., 2010b; Markoulaki et al., 2009; Nakagawa et al., 2008; Sridharan et al., 2009; Wernig et al., 2008a). For instance, Oct4 and Sox2 have been shown to suppress the *Snai1* gene, which acts as a pro-mesenchymal regulatory factor; Klf4 activates the epithelial gene programme by promoting the up-regulation of multiple epithelial genes including E-cadherin; and c-myc plays a critical role in shutting down the TGF-β signalling pathway by suppressing TGF-β1 and TGF-βR2 (Li et al., 2010b).

The late phase of reprogramming is characterised by the expression of important ES cell markers such as SSEA1 during the generation of mouse iPS cells (Brambrink et al., 2008; Li et al., 2010b; Stadtfeld et al., 2008b). Similar to their contribution in pluripotency of ES cells, Oct4 and Sox2 are found to co-occupy their own promoters and also the promoters of various ES cell specific genes. Klf4 has been recently shown to bind approximately half of the target genes co-bound by Oct4, Sox2 and Nanog, suggesting that it plays a significant role in activating the autoregulatory pluripotency circuit (Chen et al., 2008a; Jiang et al., 2008; Kim et al., 2008a; Loh et al., 2006; Sridharan et al., 2009). Based on the target genes bound by c-myc, it appears that its role is different from that of Oct4, Sox2 and Klf4. Instead of being directly involved in the activation of the pluripotency network, c-myc bound genes are known to be involved in regulation of cell proliferation. Moreover, many of these target genes are already bound by c-myc in partial iPS cells (incompletely reprogrammed cells), suggesting a limited role of *c-myc* during the late reprogramming stage (Kim et al., 2010a; Sridharan et al., 2009). Although these studies have revealed, to some extent, the contribution of reprogramming factors in the process, the detailed mechanism of reprogramming is not clearly understood and many studies are currently underway to clarify these pathways.

1.5.3 Functional Equivalence of iPS and ES Cells

Although initial studies demonstrated that ES and iPS cells shared many features or were even indistinguishable (Guenther et al., 2010; Park et al., 2008b; Takahashi et al., 2007; Yu et al., 2007), later studies identified elusive but significant dissimilarities in gene expression signatures including miRNA expression (Chin et al., 2009), DNA methylation patterns (Deng et al., 2009; Doi et al., 2009), *in vitro* differentiation (Feng et al., 2010; Hu et al., 2010a) and teratoma-forming abilities (Miura et al., 2009). It is still unclear

whether these dissimilarities are inherent or they develop because of differences in the derivation process of iPS and ES cells. Recent studies have, however, indicated that differences in culture conditions or lab-specific variations (Newman and Cooper, 2010), residual transgene expression from viral integrations (Soldner et al., 2009), genetic background (Stadtfeld et al., 2010) and passage number of derived lines (Chin et al., 2009; Polo et al., 2010) significantly contribute to these dissimilarities.

Looking at genetic differences between hESc and hiPS cells, Gore and colleagues identified an increased number of point mutations in hiPS cells, particularly in the genes associated with cancers (Gore et al., 2011). Similarly, another study reported that early-passage hiPS cells carry more copy number variations (CNVs) as compared to intermediate-passage hiPS cells, fibroblasts and hESc (Hussein et al., 2011). Although these and other studies have shown genetic and epigenetic variations in ES cells and iPS cells, extension of the data to all the iPS cells would not be appropriate as these differences could arise due to different culture protocols in different laboratories or could even be inherited from donor cells. Moreover, iPS colonies are believed to derive from a single reprogrammed somatic cell of unknown genetic background whereas ES cells and the donor somatic cells are frequently non-clonal and often carry heterogeneous populations, therefore, comparative analysis of iPS cells with sub-clones of hESc and clonally derived donor cells could provide a meaningful assessment of ES-iPS cells equivalence.

It is also evident that iPS cells, to some extent, retain the somatic memory of their donor cells and therefore their differentiation capabilities are more inclined towards their parental somatic cells as compared to ES cells (Bar-Nur et al., 2011; Hu et al., 2010b; Kim et al., 2010b; Osafune et al., 2008). For example, human iPS cell generated from

pancreatic islet β-cells retained the open chromatin signatures at crucial β-cell genes and exhibited increased propensity of differentiation towards insulin producing cells both *in vivo* and *in vitro*, relative to hESc and isogenic non β-cell derived iPS cells (Bar-Nur et al., 2011). Recently, Bock and colleagues used DNA methylation mapping, gene expression profiling and a quantitative *in vitro* differentiation assay to compare 20 hESc and 12 hiPS cell lines and identified significant variability in epigenetic and transcriptional profiles among iPS cells lines, among ES cell lines and among ES and iPS cell lines (Bock et al., 2011). Importantly, the authors also established a scoring algorithm using which pluripotent stem cells lines can be comprehensively characterised and their differentiation efficiency towards specific lineage can be predicted. These variations among iPS cells, and between hESc and iPS cells, clearly hinder the applications of iPS cells in research and therapeutics but do not decrease their significance. Further evaluation of functional integrity of iPS cells will lead to improvements in iPS cells deriving methods to reduce these discrepancies.

1.5.4 Clinical Significance of iPS Cells

Generating patient-specific pluripotent stem cells without creating ethical controversies associated with destruction of embryos has been a great achievement in the field of regenerative medicine. iPS cells can provide disease-specific and patient-specific donor cells, which can be used in a variety of therapeutic applications such as determining the mechanism of disease development, carrying out drug screening and toxicology and exploring the possibilities of gene correction and cell therapy. In a recent study, Lee and colleagues generated iPS cells from a patient with familial dysautonomia (FD) and demonstrated pathogenesis modelling and drug screening (Lee et al., 2009b). FD is a rare but fatal autosomal recessive genetic disease of the peripheral nervous system,

characterised by the loss of autonomic and sensory neurons, and is caused by a point mutation in the *IKBKAP* gene. The investigators successfully derived the precursors of central and peripheral nervous system from FD-iPS cells and demonstrated *in vitro* pathogenesis of the disease marked by three disease-relevant characteristics: depleted *IKBKAP* transcripts and defects in neuronal differentiation and cell migration behaviour. Moreover, the investigators also carried out *in vitro* screening of various compounds that can alleviate the disease phenotype and found a candidate drug compound, kinetin (a plant hormone), which helped in normalising the above-mentioned characteristics of FD (Lee et al., 2009b). Similarly, other diseases including spinal muscular atrophy (SMA), Parkinson's disease, Down's syndrome, familial hypercholesterolaemia, fragile X syndrome, LEOPARD syndrome, ADA severe combined immunodeficiency (SCID), sickle cell anaemia, β-thalassaemia, Lesch-Nyhan syndrome and type 1 diabetes have been modelled using patient-derived iPS cells [reviewed in (Grskovic et al., 2011; Robinton and Daley, 2012; Wu and Hochedlinger, 2011)].

Although iPS cells have proved to be a remarkable tool for *in vitro* disease modelling, a major goal is to derive patient-specific iPS cells capable of generating disease-relevant cell types and tissues, and utilise them for cell therapies. This seems to be unlikely until the safety issues associated with iPS cells are fully addressed. However, recent studies using mouse models are very encouraging. For example, Hanna and colleagues achieved functional correction of the sickle cell anaemia defect in a humanised mouse-model of the disease. They removed the genetic deficiency of sickle cell anaemia in the mouse-derived iPS cells and introduced a normal copy of the haemoglobin gene by homologous recombination and transplanted the iPS cells-differentiated blood progenitors into the diseased mice. The mice showed successful engraftment of these blood derivatives and

substantial improvement in the pathological features (Hanna et al., 2007). Similarly, Wernig and colleagues transplanted iPS-derived dopamenergic neurons to lessen some symptoms in rat-model of Parkinson's disease, thus demonstrating their potential in transplantation therapy of neurological disorders (Wernig et al., 2008b). These studies have provided proof of principle that cell replacement therapy, exclusive or coupled with gene-correction strategies, could also be developed using human iPS cells and can offer invaluable therapeutic options to treat various diseases.

1.6 Gene Therapy

Gene therapy is broadly defined as the introduction of genetic material into the cells to prevent, block or hinder a pathological process. The genetic material is delivered either *in vivo* or *ex vivo* into the cells through viral vectors such as adenoviruses, retroviruses etc or through non-viral vectors such as episomal DNA. There are three routes through which gene therapy can be performed. The most common clinical way is the addition of a gene that is of therapeutic significance or can supply a missing cellular protein. Gene correction/alteration is another way to fix or produce a mutation using approaches such as DNA recombination along with zinc finger nucleases. The final way is that of gene knockdown to eliminate a gene product responsible for a pathology using RNAi or miRNA (Kay, 2011).

The idea of gene therapy was first presented by Friedmann and Roblin in 1972 as a potential approach for correcting monogenetic diseases such as thalassemia, cystic fibrosis, haemophilia, sickle cell anaemia and Duchenne muscular dystrophy (DMD) (Friedmann and Roblin, 1972). Until 2011, over 1600 studies of gene therapy have received approval for clinical testing, of which the vast majority have targeted the

diseases related to cancer, cardiovascular illnesses and inherited single gene disorders (Edelstein et al., 2007; Sheridan, 2011). In the first gene therapy clinical trial, carried out in 1989-1990, two children with severe combined immunodeficiency (SCID) received the retroviral-mediated gene transfer of adenosine deaminase (*ADA*) gene into T-cells. One of the two children exhibited a temporary response (Blaese et al., 1995). In 1999, a patient suffering from ornithine transcarbamylase (*OTC*) deficiency died in a gene therapy clinical trial because of the adenoviral vector-associated toxicity leading to multiple organ failure (Hollon, 2000; Marshall, 1999). Later, 20 patients suffering from X-linked SCID (SCID-X1) were treated with reasonable success by retroviral transfer of the interleukin-2 receptor, gamma (*IL2RG*) gene into the $CD34^+$ cells. However, 5 of the treated patients developed leukaemia symptoms because of the induction of the oncogene *LMO2* by the retroviral integration in vicinity of the *LMO2* gene (Fischer et al., 2010; Hacein-Bey-Abina et al., 2008; Hacein-Bey-Abina et al., 2003a; Hacein-Bey-Abina et al., 2003b). These studies, and other clinical trials showing gene dysregulation problems associated with genomic integration of vectors caused major setbacks to the promises of gene therapy field and brought attention to the safety of the vectors (Cavazzana-Calvo et al., 2010; Ott et al., 2006; Stein et al., 2010).

1.6.1 Technical Challenges for Gene Therapy

Over the past 10 years, much research has been concentrated on improvements in vector systems, enhancements in delivery efficiencies and vector-host safety issues. However, there are still some obstacles and technical challenges that need to be resolved before gene therapy can be regularly practised in the clinic. One of the main concerns is the risk of insertional mutagenesis as a result of integration of the vectors leading to activation of proto-oncogenes or disruption of tumour-suppressor genes. Another major barrier to safe

gene therapy is the activation of the host immune response and toxicity as a result of viral proteins and other vector particles. As the effects of many gene therapy procedures are temporary, the immunity and toxicity issues associated with viral vectors have severely restricted multiple rounds of treatments and their dosage (Kay, 2011). In contrast to viral vectors, plasmid/naked DNA constructs do not elicit a host immunological response and therefore have gained much attention in the recent studies. Chemical agents such as liposomes and dendrimer complexes, and physical methods such as electroporation, pressure-perfusion and ultrasounds have been used to increase the delivery efficiency of plasmid vectors into the cells and tissues (Wells, 2004). However, the delivery efficiency is lower than with viral vectors. Moreover, as plasmid DNA vectors do not frequently integrate into the genome, they are gradually lost from the dividing cell population thus rendering only short-term benefit (Kay, 2011).

Silencing of the transgene expression due to the epigenetic modifications of the vector DNA inside the host cells is another major challenge for gene therapy. Viral promoters such as cytomegalovirus immediate early (CMV IE) promoter enhancer, Simian virus 40 (SV40), and Rous sarcoma virus LTR promoters have been frequently used in many gene therapy studies to drive the transgene expression. Although viral promoters are known to facilitate high expression of transgenes, they are prone to transcriptional silencing both *in vitro* and *in vivo* (Papadakis et al., 2004). Host defence mechanisms such as stimulation of tumour necrosis factor (TNFα) and interferons (IFNγ), cytokines, and DNA methylation are commonly identified as responsible for the silencing of viral promoters (Acsadi et al., 1998; Bestor, 2000; Hsu et al., 2010; Papadakis et al., 2004; Sung et al., 2001). Eukaryotic promoters such as those of the human elongation factor 1α (hEF1α), the chicken β-actin coupled with CMV early enhancer (CAG), the phosphoglycerate

kinase (PGK) and Ubiquitin C (UbC) promoters have exhibited prolonged transgene expression as compared to viral promoters and therefore are increasingly being considered for gene therapy studies (Fan et al., 2000; Gill et al., 2001; Papadakis et al., 2004; Zheng and Baum, 2005). In most cases, the choice of promoter can also be governed by the required duration of transgene expression for a specific disease. Genetic diseases generally require a lifelong effect while acquired illnesses such as cancers and infections may require temporary expression of the transgenes.

1.6.2 Gene Therapy with hESc and hiPS Cells

Correction of genetic deficiencies through gene therapy has substantial potential and has already provided many examples where significant clinical success is achieved in various illnesses. However, there are many diseases and pathological conditions where, for example, *ex vivo* genetic correction of somatic cells would not be sufficient as most of the somatic cells lack self-renewal, proliferation and differentiation abilities making it immensely challenging to select, identify and expand the rare corrected clones. Similarly, *in vivo* gene correction of somatic cells for the diseases involving rapid turnover of cells such as those of blood or skin cells would only provide temporary benefits as new cells (unmodified) will replace the corrected cells. Therefore an ideal situation would be to harness the potential of gene therapy in stem cells so that the effects of gene correction can be extended to differentiated lineages. As ES and iPS cells are pluripotent, they become the cell type of choice to be utilised for genetic alteration. Moreover, there are various other applications where genetic alteration in hESc and hiPS cells could provide a valuable set of tools. These include (1) studying the mechanisms involved in early developmental stages governing cell fate decisions; (2) selection of the cells to remove non-genetically altered cells and isolate clonal population; (3) regulating their

differentiation towards particular lineage with the help of transcription factors and specific growth factors (4) reducing or eliminating the immunogenic response during transplant (applicable for ES cells) and (5) modifying the feeder cells to support maintenance of hES and hiPS cells and their differentiation towards specific path (Strulovici et al., 2007).

1.6.3 Methods for Genetic Modification of hES and hiPS Cells

Various methods using viral and non-viral vectors have been reported to perform genetic modification of hES and hiPS cells. Non-viral modes of DNA delivery include the use of commercially available transfection reagents such as Lipofectamine, FuGENE, ExGen 500, GeneJammer, GeneJuice (Braam et al., 2008; Chatterjee et al., 2011; Eiges et al., 2001), physical methods such as electroporation (Braam et al., 2008; Costa et al., 2007) or nucleofection, which uses both a transfection reagent and electroporation through the use of a device referred as Nucleofector (Chatterjee et al., 2011; Moore et al., 2010; Siemen et al., 2005). Viral based DNA delivery systems include lentiviral, retroviral, adenoviral, adeno associated viral, helper dependent adenoviral vectors, foamy viruses (Strulovici et al., 2007) and herpes simplex virus- (HSV)-1 amplicon system [(this study; (Mandegar et al., 2011)). Variable delivery efficiencies have been reported through the use of both viral and non-viral systems. These variations could be attributed, along with different laboratory conditions, to many factors including the size of DNA to be delivered, differences among hES/hiPS cell lines and promoter of the reporter genes. However, viral vectors are generally considered to provide better delivery efficiencies as compared to non-viral vectors (Strulovici et al., 2007).

Although viral vectors can efficiently deliver the target genes, the DNA capacity of most commonly used constructs, with the exception of herpes simplex virus type-1 and helper-dependent adenoviruses, is between 5-10kb (Glorioso et al., 1995; Kay, 2011; Palmer and Ng, 2005). While non-viral methods such as transfection and electroporation can deliver very large DNA fragments, the efficiency of delivery decreases as size of the vector increases. Table 1-2 list various methods of DNA delivery into hESc and iPS cells along with their delivery efficiencies, advantages and disadvantages. Although some of these gene delivery modes have been tested in hESc only and not yet in hiPS cells, the features of these delivery methods can also be applied to hiPS cells as the functional behaviour of hiPS cells is similar to that of hESc.

Table 1-2 Available methods for gene transfer into hESc and hiPS cells

Vector Type	Delivery Method [Ref]	Efficiency	Advantages	Disadvantages
Viral vectors	Retroviral[a] [1]	50-65%	Reasonably efficient; stable expression	Genomic integration; limited capacity (8kb)
	Lentiviral[a] [2]	Up to 100%	Highly efficient; stable expression	Genomic integration; limited capacity (8-10kb)
	Adenoviral[a] [3]	Up to 90%	Reasonably efficient, infects also non-dividing cells; non-integrating	Transient expression; toxicity; limited capacity (8kb)
	Helper-dependent Adenoviral [4]	Up to 100%	Highly efficient; reasonable capacity (~36kb); low immunogenicity	Transient expression; very high MOIs required;
	Adeno Associated Viral [5]	28%	low immunogenicity; infects also non-dividing cells; site specific integration	Very high MOIs required; limited capacity (4.7kb)
	Foamy Virus[a] [6]	15-50%	Reasonably efficient; also infects non-dividing cells; stable expression	Genomic integration; low titres obtained; limited capacity (~9kb)
	HSV-1 Amplicons [7]	25-85%	Non-integrating; large capacity (152kb); low immunogenicity	Gradual silencing of transgene
Non-viral vectors	Lipofectamine[a] [8]; ExGen 500[a] [9]; FuGENE[a] [10]; GeneJammer[a] [11]	10-80% (varies with transfection agent and cell line)	Very rare integration; no size limit; low immunogenicity	Low delivery efficiency especially for larger vectors; loss of vectors during cell division
	Electroporation[a] [12]	18-53%	Very rare integration; no size limit;	High cell death; delivery efficiency decreases for larger vectors
	Nucleofection [13]	30-85%	Very rare integration; no size limit;	Delivery efficiency decreases for larger vectors

[a] The methods have not been reported yet for the human iPS cells. Ref: 1 (Koch et al., 2006; Lebkowski et al., 2001; Strulovici et al., 2007), 2 (Koch et al., 2006; Pfeifer et al., 2002; Strulovici et al., 2007), 3 (Braam et al., 2008; Brokhman et al., 2009), 4 (Aizawa et al., 2012; Suzuki et al., 2008), 5 (Asuri et al., 2012), 6 (Gharwan et al., 2007), 7 [This report; (Mandegar et al., 2011)], 8 (Eiges et al., 2001; Siemen et al., 2005), 9 (Eiges et al., 2001), 10 (Lebkowski et al., 2001; Suzuki et al., 2008; Tan and Droge, 2005), 11 (Braam et al., 2008), 12 (Braam et al., 2008; Costa et al., 2007), 13 (Chatterjee et al., 2011; Hohenstein et al., 2008; Moore et al., 2010; Siemen et al., 2005)

Retroviral vectors are based on retroviruses, a class of RNA viruses that use reverse transcription to copy their RNA into double stranded DNA, which is then integrated into the host cell genome and is stably maintained. They have a capacity to deliver 7-10 kb of DNA at high efficiencies into a variety of cell types including hESc but their host range is restricted to dividing cells (Koch et al., 2006; Lebkowski et al., 2001; Strulovici et al., 2007). Another class of vectors is based on lentiviruses, which belong to the family of retroviruses but are capable of transducing both dividing and non-dividing cells types (Koch et al., 2006; Pfeifer et al., 2002; Strulovici et al., 2007). Although both retroviral and lentiviral vectors exhibit reasonable efficiency of gene delivery, they suffer from limited transgene capacity (8-10kb) and genomic integration.

Adenoviral (Ad) vectors are based on adenoviruses, which are non-integrating double stranded DNA viruses, also capable of infecting both proliferating and quiescent cells. Ad vectors are considered to be the most efficient delivery systems utilised for gene delivery into hESc cells with the delivery efficiency reaching up to 90% (Braam et al., 2008; Brokhman et al., 2009). However, two major limitations of *in vitro* use of adenoviral vectors are their transient transgene expression and high toxicity as a result of viral gene expression (Parks, 2000). Recent improvements in adenoviral vectors have led to the development of helper dependent adenoviral vectors (HdAd), which lack many viral genes and therefore result in reduced toxicity (Palmer and Ng, 2005; Ross et al., 2011). Moreover they have increased transgene capacity of up to 36kb and provide sustained transgene expression.

Another class of non-integrating viral vectors is represented by Herpes Simplex Virus (HSV-1) amplicon system, which is capable of delivering 152kb of DNA into both dividing and non-dividing cells at reasonable efficiencies (Hibbitt and Wade-Martins,

2006; Saeki et al., 2001; Sena-Esteves et al., 2000). Details of the HSV-1 amplicon system will be discussed in Section 1.9.

An alternate vector platform for gene expression is provided by human artificial chromosomes (HAC), which are assembled from chromosomal elements and function as natural chromosomes. Various studies have demonstrated that HAC deliver and express the genes of interest and complement genetic deficiencies in many cell lines (Grimes et al., 2002; Hoshiya et al., 2009; Kazuki et al., 2010; Kazuki et al., 2008; Mejia et al., 2001). The use of HAC offer many potential benefits such as carrying large genomic fragments including their regulatory elements, reducing the risk of integration and insertional mutagenesis, and providing long-term gene expression by overcoming positional effects. In addition, HAC can also be used as a model system to study natural chromosomes and the key roles associated with many chromosomal elements such as the centromere, telomere and the origin of replication. Section 1.8 will provide a detailed overview of HAC.

1.7 Chromosome Structure and Function

Chromosomes are organised structures comprised of coiled DNA and associated proteins that carry the genetic information and regulatory elements required for the function, maintenance and survival of the cells. Chromosomes are also regarded as vehicles of inheritance as they are responsible for conserving the hereditary material throughout an organism's development and augmenting the diversity of that material in the following generations. Eukaryotic chromosomes require three essential structural elements: centromeres, origins of replication and telomeres, for their biological function including

replication and segregation during the cell division (Baird and Farr, 2006; Cleveland et al., 2003; Stinchcomb et al., 1979).

1.7.1 Centromeres

Centromeres are the cis-acting DNA regions responsible for chromosomal segregation during the cell divisions. In higher eukaryotes, a centromere consists of stretches of repetitive DNA units, functions as a site for the assembly of kinetochore and attachment of spindles, and is also responsible for maintaining the attachment of sister chromatids until the beginning of anaphase (the stage of cell division when chromosomes move to opposite poles) (Chan et al., 2005; Cleveland et al., 2003). The kinetochore complex plays an essential role during the segregation of chromosomes as it mediates the attachment of chromosomes to the microtubule spindles, facilitates the positioning of the chromosomes along the spindles and generates a "wait-anaphase" signal until all the chromosomes are properly attached and aligned (Nicklas, 1997; Rieder and Salmon, 1998). Electron microscopy analysis has revealed that the kinetochore has a trilaminar plate structure (Rieder, 1982), which is composed of an inner plate, middle plate and an outer plate as shown in Figure 1-3. The inner plate is adjacent to the centromeric heterochromatin and consists of the outermost centromeric DNA and centromeric proteins such as CENP-A and CENP-C (Craig et al., 1999). The middle plate (or interzone) contains phosphoantigens and is known to regulate spindle assembly at mitotic checkpoint during metaphase to anaphase transition (Campbell and Gorbsky, 1995). The outer kinetochore plate contains many of the microtubule interacting- and mitotic checkpoint-proteins and is known to act as a site for microtubule attachment and regulate checkpoint signalling. When microtubules are not attached to the kinetochore, a region referred as fibrous corona becomes visible that extends from the trilaminar plate, docks

microtubule motor and checkpoint proteins including CENP-E, dynein and BubR1, and contributes in monitoring the integrity of kinetochore attachments along with the outer plate (Figure 1-3) (Cooke et al., 1997; Jablonski et al., 1998).

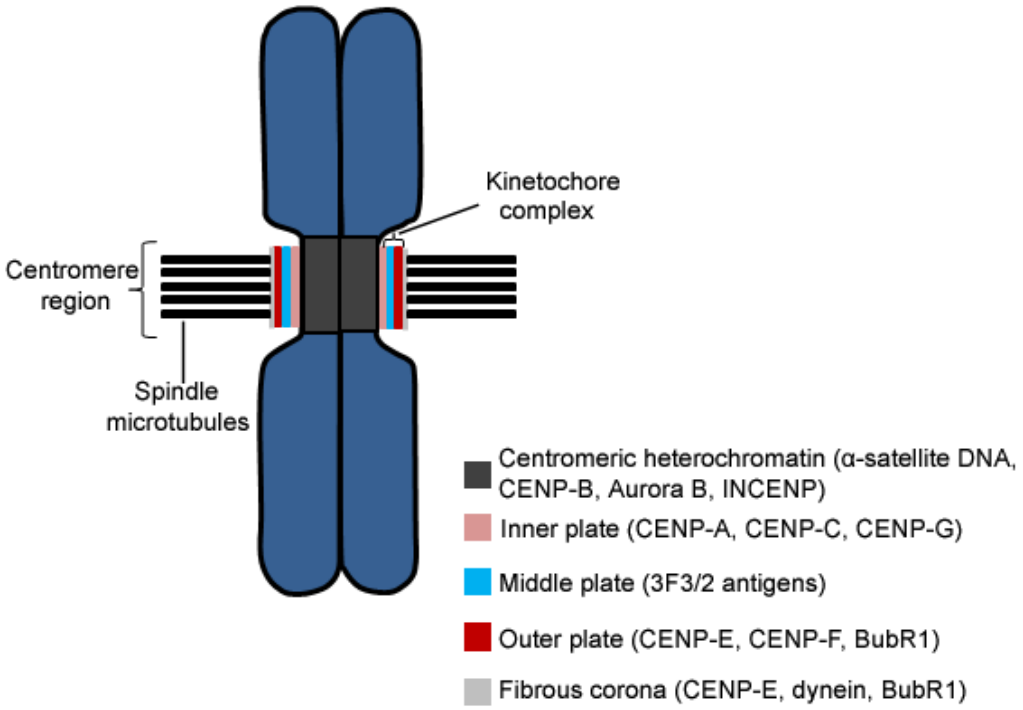

Figure 1-3 Centromere components and associated proteins
The centromere is a region that connects sister chromatids at metaphase and is the site for kinetochore assembly and spindle attachment during cell division. The kinetochore is a trilaminar structure composed of inner plate, middle plate and outer plate. The inner plate is directly adjacent to centromeric heterochromatin. The outer plate and its associated layer fibrous corona are involved in regulating microtubule interaction. The parentheses indicate major protein components of heterochromatin, kinetochore plates and fibrous corona.

While the centromeric proteins (further described in Section 1.7.2) are highly conserved across many species, the DNA sequences at the centromeric regions possess significant diversity (Amor et al., 2004; Henikoff et al., 2001; Oegema et al., 2001; Schueler and Sullivan, 2006; Sullivan and Glass, 1991; Warburton et al., 1996; Willard, 1998). Moreover the centromeric DNA of higher eukaryotes is known to be highly repetitive and

complex, which has made them difficult to be assembled. Therefore they are not clearly understood at the sequence level (Rudd and Willard, 2004). The centromeric DNA of all the human chromosomes consists of alpha-satellite (α-satellite) DNA, which is defined by large arrays of divergent 171bp repeat units. These 171bp units can exist either as monomeric α-satellite DNA (tandemly arranged monomers exhibiting ~70% homology) or as higher order repeats (HOR; monomers arranged in multimeric groups and exhibiting over 97% homology), which are positioned tandemly to generate large arrays of extremely homogenous HOR (Figure 1-4). Variation in monomeric units sequences and homogenous assembly of HOR contributes to chromosome-specific differences in α-satellite DNA (Alexandrov et al., 2001; Rudd and Willard, 2004; Rudd et al., 2006; Warburton and Willard, 1996). HOR are considered to play an important role in the centromere function as some of the main centromeric proteins are known to assemble on the HOR α-satellite DNA but not on the monomeric α-satellite DNA (Rudd and Willard, 2004; Schueler et al., 2001). Moreover, the ability of HOR α-satellite DNA to generate *de novo* human artificial chromosome (HAC, further discussed in Section 1.8.2), when introduced into cultured human cells makes them crucial in the formation of functional centromere (Grimes et al., 2002; Harrington et al., 1997; Henning et al., 1999; Masumoto et al., 1998; Mejia et al., 2002).

Besides the genetic elements, epigenetic factors are also known to contribute to centromere function as it has been shown that in mammalian dicentric chromosomes, despite the presence of α-satellite DNA at two distinct locations of chromosomes, only one of these functions as a centromere (Earnshaw and Migeon, 1985; Earnshaw et al., 1989; Sullivan and Willard, 1998). In addition, a number of studies have also reported the occurrence of neocentromeres in the absence of α-satellite DNA. Presence of these

neocentromeres also implies that α-satellite DNA is not essentially required for a functional centromere and other epigenetic elements such as protein binding could also contribute to the formation of centromere (Amor and Choo, 2002; Barry et al., 1999; Depinet et al., 1997).

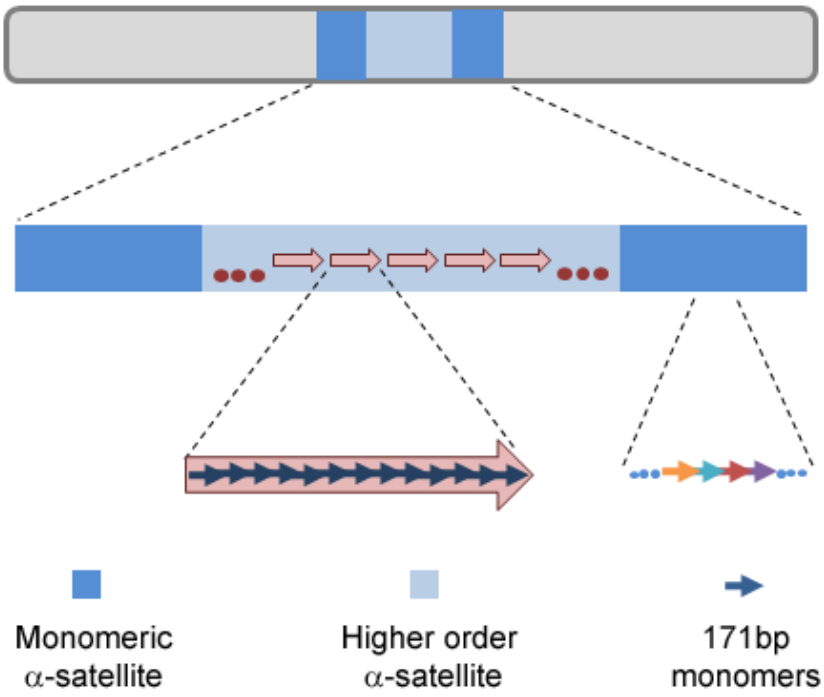

Figure 1-4 Organisation of α-satellite DNA at the centromere of human chromosomes
The centromere of human chromosomes consists of α-satellite DNA, which contains large arrays of 171bp repeating units (small coloured arrows). The α-satellite DNA is organised in (1) higher-order α-satellite arrays in which monomers are aligned in head to tail configuration forming tandem higher order repeat (HOR) units, and (2) monomeric α-satellite, in which monomers are tandemly arranged but do not possess higher order periodicity. HOR units are highly homogenous (over 97%) whereas monomers of monomeric α-satellite exhibit ~70% homology. The kinetochore structure is assembled at the HOR α-satellite DNA.

1.7.2 Centromeric Protein Components

During centromere formation, a histone H3 variant known as centromere protein (CENP) -A replaces a portion of histone H3 in the nucleosomes at core α-satellite sequences. This modification provides a distinct organisation to centromeric chromatin that is different from that of surrounding heterochromatin and helps recruiting other centromeric proteins.

CENP-A is known to associate with both centromeres and neocentromeres, and its presence is a strong epigenetic marker for kinetochore recruitment and active centromere formation (Marshall et al., 2008; Palmer et al., 1991; Sullivan and Karpen, 2004; Warburton et al., 1997). The α-satellite DNA contains a 17bp motif, known as the CENP-B box, which signals CENP-B recruitment. CENP-B is a well-conserved protein and is known to be present at the heterochromatin region underlying the kinetochore complex of all human chromosomes except Y (Cooke et al., 1990; Masumoto et al., 1989). However, its role in centromere function is not clearly understood as it is also known to be present on inactive centromeres and knockout *cenpb* mice do not exhibit any mitotic defects (Hudson et al., 1998; Sullivan and Schwartz, 1995). One of the crucial components of inner kinetochore is the protein CENP-C, which is known to play an important role in the maintaining the stability of kinetochore chromatin, recruiting other kinetochore components and mitotic checkpoint function (Kwon et al., 2007; Ribeiro et al., 2010). CENP-C is only associated with functional centromeres including neocentromeres and is absolutely necessary for chromosome segregation as its disruption cause serious defects in chromosome segregation ultimately leading to mitotic arrest (Kalitsis et al., 1998; Ribeiro et al., 2010; Saffery et al., 2000; Tomkiel et al., 1994). Other protein components such as CENP-E, dynein (both of them referred as microtubule motor proteins) and mitotic centromere-associated kinesin (MCAK) are associated with kinetochore complex and contribute to chromosome motility along the microtubules (Rieder and Salmon, 1998). In addition, Mad2 and Bub1 (also called spindle checkpoint proteins) play a vital role in generating, regulating and releasing the wait-anaphase signal until the kinetochore is attached to the spindle and are therefore critical during chromosome segregation (Shah and Cleveland, 2000). A recent study conducted by Ribeiro and colleagues highlights

additional contributions of these centromeric proteins along with their associations with other proteins (Ribeiro et al., 2010).

1.7.3 Origins of Replication

The origins of replication (Ori) are sites where chromosomal DNA replication is initiated. Ori sites are well characterised in prokaryotes and unicellular eukaryotes, but are less defined in higher eukaryotes. For example, in *S. cerevisiae*, specific Ori have been identified, which contain 125bp of AT-rich DNA sequences and facilitate binding for the origin recognition complex (ORC) to initiate replication (Gilbert, 2001). In multicellular organisms, no such Ori sites have been identified and it is believed that in some systems, any DNA sequence can act as replication origin (Gilbert, 2001). It is also speculated that chromosomes of higher eukaryotes possess many potential initiation sites for replication (also called replicons); however the exact nature and sequence of these regions have not been elucidated (Gilbert, 2004; Norio, 2006). In human cells, it has been demonstrated that a DNA fragment exceeding 15kb size is capable of autonomous replication and therefore replicons must be positioned at least every 10-15kb of DNA (Heinzel et al., 1991). Although the positioning and sequence specificity of replicons are not fully defined, many of the proteins facilitating the replication commencement such as single-stranded DNA binding (SSB) proteins and DNA primase are highly conserved from yeast to higher eukaryotic organisms (Pasero and Schwob, 2000).

1.7.4 Telomeres

Eukaryotic chromosomes contain stretches of cis-acting regions comprising of tandem repeat units at the end of the linear chromosomes called telomeres (Shay, 1999; Shay and Wright, 2007). Telomeres are responsible for providing chromosome stability and

maintaining the chromosome structural integrity by sealing the ends of the chromosome. They also protect the ends of the chromosome from exonuclease degradation, fusion with other broken chromosomes and recombination events. In addition, telomeres are responsible for ensuring complete replication of the entire terminal ends of chromosomes (Neidle and Parkinson, 2003). The tandem repeat units of vertebrates telomeres consists of hexameric motif TTAGGG, which terminates in a brief G-rich overhang of single-stranded DNA region at the utmost 3' ends (Moyzis et al., 1988; Shay, 1999). The length of telomere regions varies from species to species and is also dependent on the age of the cells. For example, human telomeres range from 2-50kb (Osterhage and Friedman, 2009). During DNA replication, the length of telomeres is shortened as a result of the "end-replication" phenomenon, which occurs when the short RNA primer is removed from the lagging strand at 5' end of telomeres resulting in a single-stranded overhang. This overhang is degraded by exonucleases leading to shortened telomere lengths at each cell division (Allsopp et al., 1995; Levy et al., 1992). The enzyme telomerase (an RNA-protein complex comprising of a reverse transcriptase subunit, accessory elements and template RNA) is responsible for maintaining the replicative telomeric sequences by adding $(TTAGGG)_n$ units to the telomeric ends of chromosomes in 5' to 3' direction (Osterhage and Friedman, 2009). Most of the normal somatic cells do not express telomerase and therefore telomere lengths are shortened during each cell division, a process that ultimately leads to replicative senescence or apoptosis of cells (Neidle and Parkinson, 2003). In contrast, stem cells such as ES cells, iPS cells and cancer cells, express high levels of telomerase, an attribute commonly associated with indefinite self-renewal and immortality of cells (Albanell et al., 1999; Takahashi et al., 2007; Thomson et al., 1998).

1.8 Generation of Human Artificial Chromosomes

Human artificial chromosomes (HAC) are autonomous genetic molecules, assembled to function, replicate and segregate as natural chromosomes (Harrington et al., 1997). They can be used to achieve two major goals: first, to understand the structure and behaviour of natural chromosomes; and second, to use them as high-capacity gene carrying vectors for genetic improvement or gene therapy. HAC allow cloning and expression of human transgenes in mammalian systems and can be used as potential delivery vehicles for gene therapy. Several studies have demonstrated that HAC are suitable for carrying large DNA fragments and can complement genetic deficiencies in human cell-culture systems (Grimes et al., 2001; Hoshiya et al., 2009; Mejia et al., 2001). Presently, there are two main approaches available to construct human artificial chromosomes: the top-down approach and the bottom-up approach (Saffery and Choo, 2002).

1.8.1 Top-Down Approach

The top-down method utilises either integration of telomeric sequences to sequentially fragment natural chromosomes through a process referred as telomere associated chromosome fragmentation (TACF; (Barnett et al., 1993; Farr et al., 1991; Itzhaki et al., 1992)) or physical methods such as X-rays irradiation to truncate the natural chromosomes (Guiducci et al., 1999; Voet et al., 2001). The engineered chromosome were initially truncated using a top-down approach, which involved introduction of cloned telomere sequences in a random fashion at several sites along the q-arm of the human X-chromosome, which resulted in chromosome fragmentation (Farr et al., 1992). The size of these fragmented chromosomes was further reduced to remove unwanted elements along the p-arm yet retained the essential components for chromosome function such as the centromere region. The truncated chromosomes were known as

minichromsomes (Farr et al., 1995; Mills et al., 1999). Similarly, other chromosomes such as Y-chromosome and chromosome-21 have been successfully truncated to derivative minichromosomes, and have been transferred to mice and transformed human cells through microcell mediated chromosome transfer (MMCT; (Heller et al., 1996; Katoh et al., 2004)). Chromosome-21 derived minichromosomes have been modified to include transgenes and shown to exhibit prolonged mitotic stability and transgene expression in human HT1080 cells (Katoh et al., 2004).

1.8.2 Bottom-Up Approach

The bottom-up approach, which is the one used in this report, combines chromosomal elements essential for replication and maintenance to generate *de novo* HAC (Harrington et al., 1997). The essential requirement for *de novo* HAC formation is the presence of large alpha satellite arrays, which act as a site for centromere formation (Figure 1-5). Earlier studies also incorporated telomere sequences into HAC vectors (Henning et al., 1999; Ikeno et al., 1998). However, telomeres are dispensable for HAC formation as *de novo* HAC are generally circular. The requirement of origins of replication was also eliminated for generation of *de novo* HAC as most sequences over 15kb length are replication competent (Krysan et al., 1993).

The development of the first-generation *de novo* HAC into the human lung fibrosarcoma cell line HT1080 involved transfection of either synthetic arrays of α-satellite DNA (ranging from 100kb-1Mb) along with genomic DNA and telomere sequences (Harrington et al., 1997) or artificial chromosome vector such as yeast artificial chromosome (YAC; (Ikeno et al., 1998)), bacterial artificial chromosome (BAC; (Mejia et al., 2001)) and P1 artificial chromosome (PAC; (Ebersole et al., 2000)), retrofitted with

α-satellite DNA and telomere sequences. The resulting *de novo* artificial chromosomes (referred as HAC) were estimated to be ~6-10Mb in size, which is approximately one-tenth of the size of smallest human chromosome, chromosome-21 (~50Mb; (Korenberg and Engels, 1978)). Moreover, the HAC bound the antibodies to active centromere-specific proteins (indicating the development of functional centromere) and were mitotically stable in the absence of drug selection. As the size of HAC was 10-20 times larger than the input HAC vectors, further analysis such as extended DNA fibre-FISH revealed that these HAC contained alternated arrangement of α-satellite DNA and vector sequences, thereby indicating that HAC were composed of multimers of input DNA. It was hypothesised that amplification of the input DNA was required to stabilise the HAC, by attaining the minimum size for it to be attached by spindle microtubules (Grimes et al., 2002; Mejia et al., 2001).

Studies on HAC have significantly contributed to the understanding of functional requirements of *de novo* centromere formation. Size and composition of the α-satellite sequences are the most critical features affecting the efficiency of centromere assembly. HAC vectors containing at least 30kb of α-satellite DNA have successfully generated stable HAC and an increase in the size of α-satellite DNA further improved the efficiency (Okamoto et al., 2007). Conversely, HAC vectors with 10kb of α-satellite DNA did not contribute to the formation of HAC (Okamoto et al., 2007). α-satellite sequences from various chromosomes such as 4, 5, 14, 17, 21, 22, X and Y have all been examined for their efficiency to generate stable HAC (Ebersole et al., 2000; Grimes et al., 2002; Harrington et al., 1997; Henning et al., 1999; Ikeno et al., 1998; Laner et al., 2005; Mejia et al., 2002; Ohzeki et al., 2002; Schueler et al., 2001). These studies have provided strong evidence that 17α satellite arrays (D17Z1) exhibit higher efficiency of stable HAC

generation, as compared to other alpha satellite arrays. On the other hand, HAC were not established using Yα satellite arrays (DYZ3), which are devoid of CENP-B boxes (Grimes et al., 2002). The importance of the presence of CENP-B boxes has also been highlighted from other studies where α-satellite arrays with mutated CENP-B boxes have failed to establish HAC (Basu et al., 2005; Ohzeki et al., 2002).

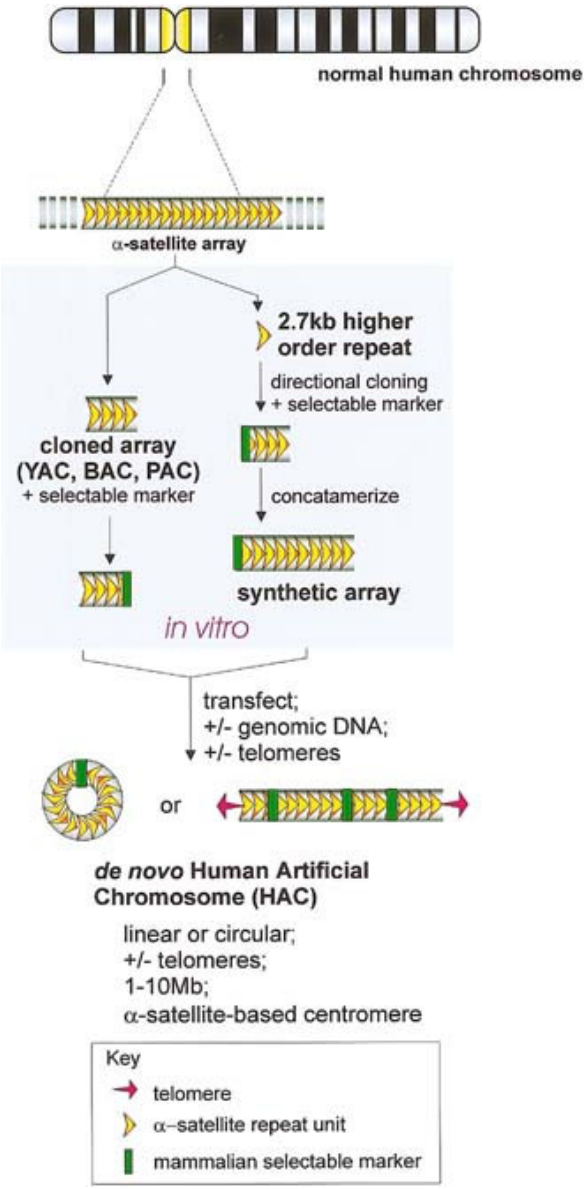

Figure 1-5 Formation of *de novo* HACs
Alphoid arrays (with/without telomeres and genomic DNA) cloned in yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), or P1 artificial chromosomes (PACs), or synthesised as large concatamers are introduced into the cells. The resulting HACs can be linear, or more commonly circular.
Figure taken from (Saffery and Choo, 2002).

1.8.3 Importance of HAC and Remaining Challenges

HAC offers several advantages over conventional gene therapy vectors as they can replicate and segregate independently without integration into the cell's endogenous chromosomes (Grimes et al., 2001; Guiducci et al., 1999; Harrington et al., 1997). Moreover, they have the capacity to hold large genetic elements including multiple genes along with their regulatory regions to assist tissue-specific transgene expression (Hoshiya et al., 2009; Kazuki et al., 2008).

Despite these advantages, there are problems which need to be addressed before HAC can be used as vectors for transgene delivery. For example, *de novo* HAC exhibit more segregation errors, in *in vitro* cell culture, during anaphase than natural human chromosomes (Rudd et al., 2003). However, when transferred to mouse ES cells to form mice, they segregate correctly and are inherited by the progeny (Co et al., 2000). Moreover, HAC vectors containing α-satellite centromeric sequences have been introduced into different cell lines by lipofection, but successful HAC generation has been obtained mainly in the cultured human fibrosarcoma cell line HT1080 (Ebersole et al., 2000; Grimes et al., 2002; Harrington et al., 1997; Ikeno et al., 2002; Mejia et al., 2002). The reason for this is not clear, but it may be because of the ease with which HT1080 cells can be transfected or genetic background of these cells.

1.8.4 Transgene Expression from *de novo* HACs

Previously, our laboratory has demonstrated the efficient generation of *de novo* HAC in human cell lines. A HAC vector, containing chromosome 17 alphoid DNA, a large genomic segment comprising of hypoxanthine phosphoribosyltransferase (*HPRT*) locus and telomere repeats successfully complemented in HT1080 cells the *HPRT* deficiency,

which causes the neurodevelopmental disorder Lesch-Nyhan syndrome (Mejia and Larin, 2000; Mejia et al., 2001). This was the first documented report of the successful complementation of a genetic deficiency using *de novo* HAC. Analysis of the HAC structure showed irregular and alternating stretches of 17-alphoid DNA and *HPRT* genomic DNA, suggesting amplification and rearrangement of the input HAC vector. Recently, a *de novo* HAC vector, containing the human β-globin gene cluster was constructed and introduced into mouse-ES cells (Suzuki et al., 2009). Upon successful formation of HAC, *in vitro* differentiated erythrocytes exhibited natural sequential expression of ε, γ and β-globin genes. Moreover, the chimeric mice generated from ES cells showed tissue-specific and development-specific expression of human globin gene.

These studies highlighted the importance of HAC as high-capacity gene transfer vectors. However, they also highlighted a number of obstacles that are currently limiting their applications in gene therapy. These include (i) unwanted amplification and rearrangement of the introduced vector and, therefore, the target gene; and (ii) inefficient delivery of HAC vectors into target cells, especially human ES cells/iPS cells.

Our group recently adapted the HSV-1 amplicon delivery technique for *de novo* HAC formation. For the first time, high efficiency HAC were obtained (efficiency was up to three orders of magnitude higher than lipofection) in several different cell lines. In addition, an *HPRT* expressing HAC was generated that fully complemented the defect for a prolonged period of time in *HPRT* deficient HT1080 cells, and was retained by the formation of a functional centromere (Moralli et al., 2006).

1.9 Herpes Simplex Virus (HSV)-1 Vectors

Herpes Simplex Virus-1 (HSV-1) belongs to *Herpesviridae*, a family of viruses carrying a linear and double stranded DNA genome, which ranges from 124 to 230 kb. The family also includes important human pathogens such as HSV-2, Epstein-Barr virus, Cytomegalovirus (CMV) and Kaposi' sarcoma-associated Herpesvirus-8 etc (Pellett and Roizman, 2007). Of these, HSV-1 is the most extensively studied and characterised virus. HSV-1 has a tropism for sensory neurons, where it establishes latent infections, and is commonly known for causing orofacial infections and encephalitis (Whitley and Roizman, 2001). The HSV-1 virion consists of four components: (1) an electron-impenetrable core containing the viral genome; (2) an icosahedral capsid; (3) the tegument, an unstructured layer of proteins that surrounds the capsid; (4) and an envelope (Roizman et al., 2007; Whitley and Roizman, 2001) (Figure 1-6). The size of the HSV-1 genome is ~152 kb and it encodes approximately 90 open reading frames (ORFs), of which at least 84 ORFs code for proteins. The DNA consists of two covalently linked regions, U_L (unique long) and U_S (unique short), flanked by large inverted repeats. The inverted repeats flanking U_L are designated *ab* and *a'b'* and those flanking U_S are called *a'c'* and *ac*. Inversion of U_L and U_S regions, relative to each other, results in four isomeric forms of DNA molecules that differs only in the orientation of these DNA sequences (Lehman and Boehmer, 1999; Whitley et al., 1998).

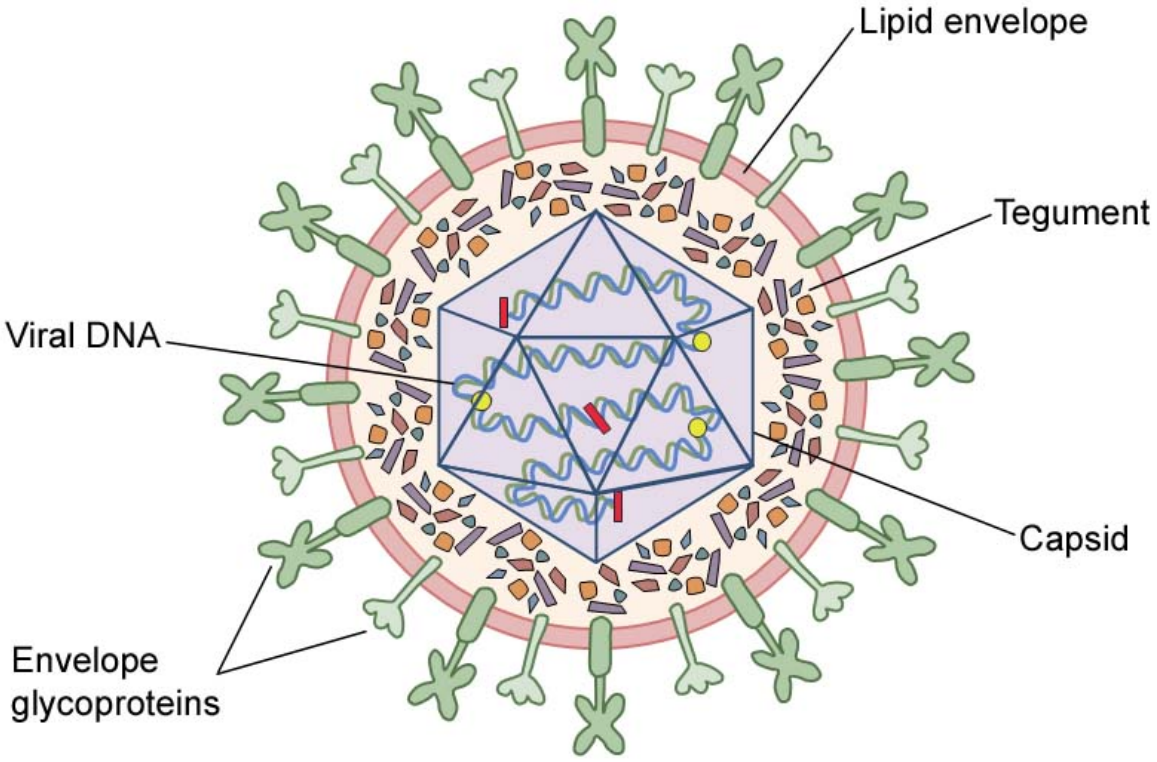

Figure 1-6 Structure of HSV-1 virion
The HSV-1 virion consists of a core containing the DNA, nucleocapsid, tegument proteins and an envelope membrane. The diameter of the capsid is ~100 nm and that of the virion is ~120-300 nm. The envelope membrane is derived from the host cell and it contains viral specific proteins (glycoprotein spikes) on the surface.

1.9.1 HSV-1 Life Cycle

The HSV-1 infection begins with the attachment and subsequent binding of cell-surface receptors to the envelope glycoproteins of the virus. This is followed by the fusion of the viral envelope with the cell plasma membrane (Roizman et al., 2007). The capsid is then transported to the nuclear pores, where the viral DNA is released into the nucleus through the facilitation of tegument proteins (Newcomb and Brown, 1994; Sodeik et al., 1997). Once the viral genome is transported to the nucleus, it circularises and maintains an episomal status (Garber et al., 1993). After circularisation, HSV-1 can either enter the lytic phase, which leads to lytic infection or it can enter the latent phase. However, the mechanisms responsible for dictating the entry of HSV-1 in these two phases are not clearly understood (de Silva and Bowers, 2009). The lytic viral phase is initiated by the

transcription and synthesis of viral gene products in a strictly regulated fashion. The transcription occurs in three sequential phases: (1) the first set of transcribed genes, designated as α or immediate-early (IE) genes, are mainly involved in the regulation of viral replication, (2) β or early genes (E) are associated with replication and packaging of viral DNA, and (3) γ or late (L) genes produce the proteins, most of which become the parts of virion structural components including the capsid, the tegument and envelop glycoproteins (de Silva and Bowers, 2009; Jacobs et al., 1999a). Viral DNA is subsequently packaged into pre-assembled capsids and the virus particles are then transported from the nuclear membrane to the Golgi apparatus, where they are enveloped and released into the extracellular space (Antinone et al., 2006; Campadelli et al., 1993; de Silva and Bowers, 2009).

1.9.2 HSV-1 Vectors

The structure, physiology and function of HSV-1 has been extensively studied and has facilitated the development of methods for diagnosis and efficient treatment of HSV-1 infections as well as the designing of effective vaccines. Moreover, understanding of HSV-1 genetics has enabled the engineering of its genome for the construction of HSV-1-based vectors, to be used as gene transfer vehicles in various experimental settings and clinical gene therapy approaches. There are two types of HSV-1-derived vectors: recombinant vectors and amplicon vectors [(Epstein, 2009a; Jacobs et al., 1999b); (Figure 1-7)]. These are described below.

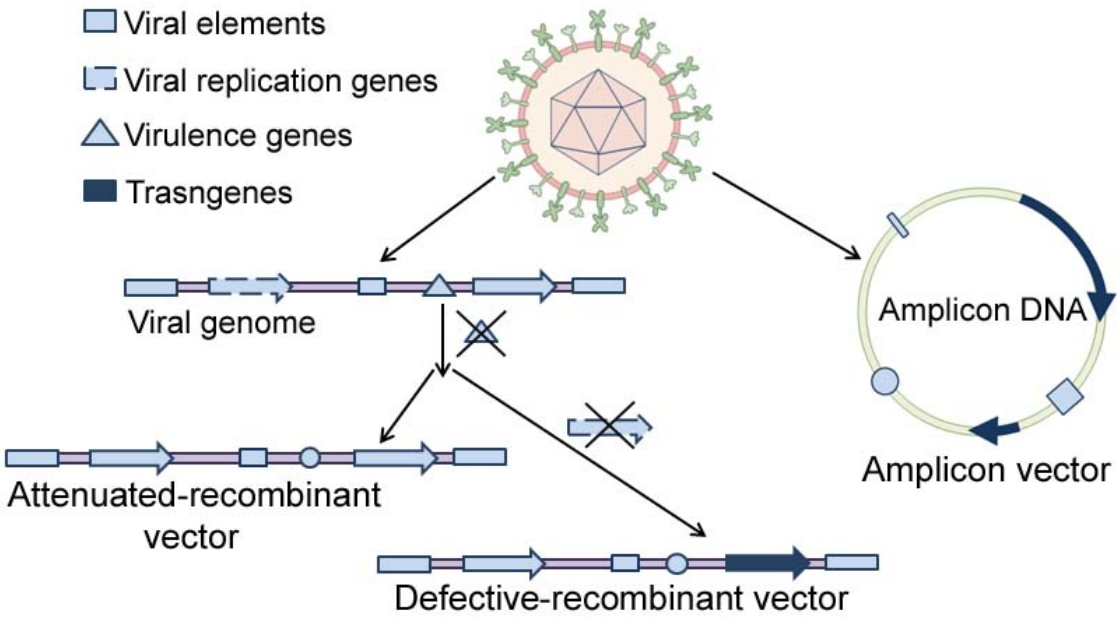

Figure 1-7 Recombinant and amplicon based HSV-1vectors
Recombinant HSV-1 vectors contain the viral genome, which is engineered to carry site specific modifications. They are classified into (a) attenuated-recombinant vectors, which are replication-competent vectors and contain mutation or deletion of virulence genes; and (b) defective-recombinant vectors, which are deficient for genes essential for viral replication and can carry transgene or reporter genes. Amplicon vectors contain HSV-1 origin replication (*oriS*) and packaging site (*pac*) to allow DNA amplification and packaging of the amplicon DNA into the virion particles.

1.9.3 HSV-1 Recombinant Vectors

Recombinant vectors are constructed by homologous recombination between the engineered HSV-1 genome and a shuttle plasmid containing one or several transgenes up to 30 kb in size (Goins et al., 2011; Jacobs et al., 1999b). Recombinant vectors are further classified into two main types (Epstein et al., 2005). The first type is called attenuated-recombinant vectors (Figure 1-7) and corresponds to replication-competent vectors that carry attenuating mutations or deletions to confine the lytic replication of virus to the actively dividing cells (such as tumour cells) but not in the quiescent or normal cells. The viruses containing these vectors are called recombinant oncolytic viruses. The second type of recombinant vector is called defective-recombinant vector (Figure 1-7). These constructs are replication-incompetent and non-pathogenic vectors that lack the essential

genes to prevent the lytic viral replication but retain the ability to express transgenes in the latent phase of infection, for therapeutic gene transfer applications (Epstein, 2009b; Jacobs et al., 1999b).

Recombinant HSV-1 vectors were initially developed to study the role of individual viral genes and for the vaccination purposes. It was only later that their potential as gene transfer vehicles was realised (Jacobs et al., 1999b). Nevertheless, they retain large amount of the HSV-1 genome and therefore viral gene expression can induce considerable immunological and cytotoxic response (Jacobs et al., 1999b).

1.9.4 HSV-1 Amplicon Vectors

HSV-1 amplicon vectors contain concatamers of the parental plasmid DNA, called the amplicon plasmid, instead of the HSV-1 genome (Figure 1-7). The amplicon plasmid carries one origin of replication '*OriS*' to support replication of DNA and one cleavage/packaging signal '*pac*' to facilitate its packaging into the virions. The virion particles are identical to the wild type viruses on the basis of structure, function and immunological response and can, therefore, accommodate and deliver large DNA fragments without compromising the ability to infect a wide range of mammalian cells (Epstein et al., 2005; Jacobs et al., 1999b). As the amplicon plasmid does not carry any viral gene, almost all of the capacity of the viral genome (152 kb) can be utilised to accommodate the foreign DNA. Currently, no other mammalian viral vector can match the ability of the amplicon vector to deliver such a large fragment (up to 150 kb) of DNA (Epstein, 2009a; Epstein et al., 2005). The absence of viral genes and consequently viral proteins also provides a distint advantage to the amplicon system, as they are non-toxic and non-pathogenic to the infected cells. However, the replication, DNA packaging and

production of amplicons are dependent on the helper functions, provided in trans, by either the helper viruses or the recently developed packaging-defective plasmids [reviewed in (Epstein, 2009a; Epstein et al., 2005; Jacobs et al., 1999b)].

a) Helper-free Amplicon Vectors

Conventionally, the amplicons were produced by the transfection of the amplicon plasmid into the packaging cells along with the infection of helper HSV-1 (replication defective) (Jacobs et al., 1999b; Johnson et al., 1992). However, the resulting amplicon preparation was contaminated with helper vectors, making them difficult to be used in gene therapy and vaccination procedures. This was mainly because of the induced cytotoxic effects and immunological response associated with gene expression from helper vectors as well as due to the rare reversion of replication-defective helper HSV-1 to the wild type phenotype. To overcome these challenges, the helper virus-free amplicon system was developed (Saeki et al., 1998), in which the necessary functions for viral replication and packaging were provided by packaging-defective plasmids. One of these approaches utilised the co-transfection into the packaging cell lines of amplicon DNA and bacterial artificial chromosomes (BACs), containing the necessary helper functions. The packaging of helper BAC plasmid was restrained because it lacked the *pac* site, essentially required for the packaging of viral DNA into the virions. An improved and updated version of this system, and also the one that is used in this study, was shown to produce high titer amplicon stock in a helper-free fashion (Saeki et al., 2003). In this system the amplicon vectors were based on a plasmid, which contained the HSV-1 origin of replication '*oriS*' and the packaging signal '*pac*' as shown in Figure 1-8. Amplification and packaging of the amplicon plasmid were assisted by the necessary proteins furnished in trans by a second vector, BAC fHSVΔpacΔ27 0+, which contained the whole HSV-1 genome,

deleted for the *OriS*, the *pac* sequence, and the essential gene *ICP27*. Moreover, the fHSVΔpacΔ27 0+ size was increased to 178 kb, over the HSV-1 capsid packaging capacity, by the addition of inert "stuffer" elements. Amplicons are produced by co-transfection of amplicon DNA, HSV BAC fHSVΔpacΔ27 0+ and pEBHICP27 plasmid into African Green Monkey cells (Vero 2-2). fHSVΔpacΔ27 0+ and pEBHICP27 provide all the essential enzymes and proteins required for replication and packaging of the amplicon DNA (Figure 1-8). This method reduced the risk of generating replication-competent helper viruses (Saeki et al., 2003).

b) Hybrid Amplicon Vectors

Although the helper-free HSV-1 amplicon system has proved to be a promising vector for gene transfer, the system needs to be further improved to prevent silencing of the transgene expression, as the amplicon-mediated expression of the transgene is short-lived and often repressed in most of the cell systems other than the nerve or muscle cells (Epstein, 2009a; Fraefel et al., 1997; Suzuki et al., 2007). To avoid silencing, some modifications have been made into the amplicon system such as placing the transgene under cell type-specific promoters and combining the HSV-1 amplicon vector with DNA from other viruses (Jacobs et al., 1999b). For example: HSV/AAV hybrid vectors have been generated by introducing into the HSV-1 plasmid the *rep* gene and the inverted terminal repeats (ITR) from adeno-associated virus. These features assisted the amplification and integration of the transgene and resulted in prolonged transgene expression (Johnston et al., 1997). Similarly, the *EBNA-1* gene and *OriP* elements of Epstein-Barr virus (EBV) have been inserted into the HSV-1 amplicon vectors to support replication and segregation of the resulting HSV/EBV hybrid vectors into the dividing cells. Wang and colleagues (Wang et al., 1997; Wang and Vos, 1996) have demonstrated

that HSV/EBV hybrid vectors were maintained episomally and sustained efficient transgene expression in wide variety of cell types.

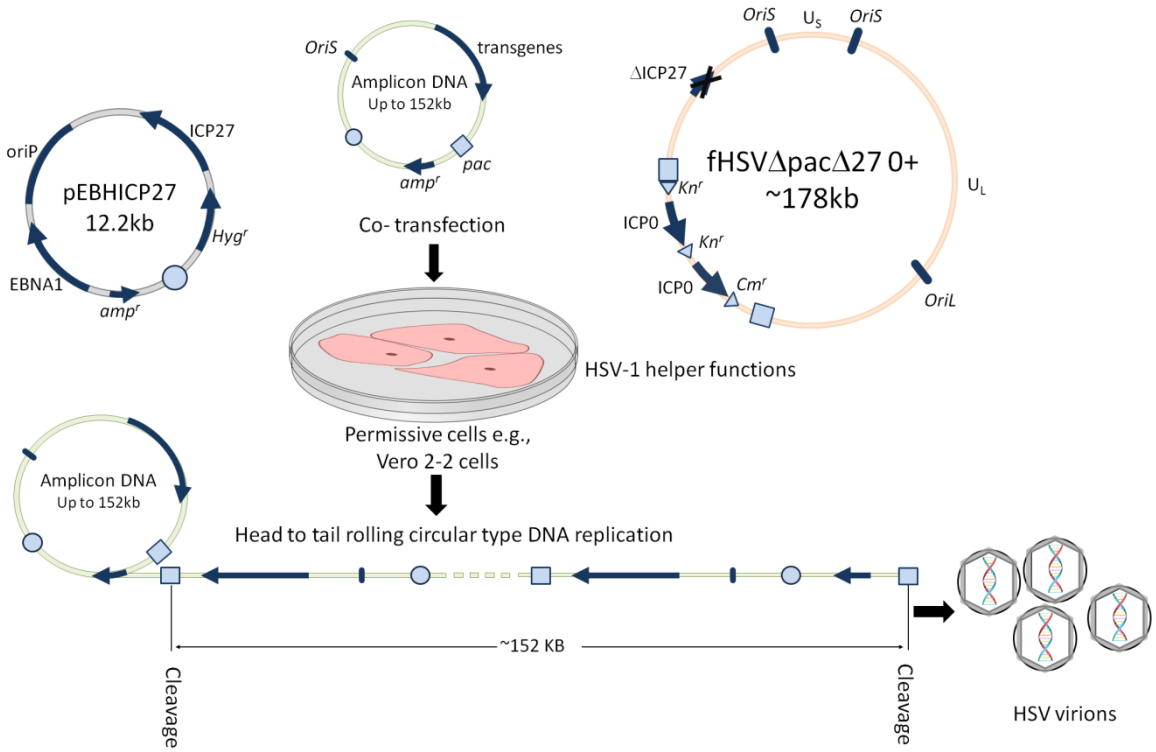

Figure 1-8 Packagining of helper virus free HSV-1 amplicons
Amplicons are produced by co-transfecting, into Vero2-2 cells, the desired amplicon DNA (containing *pac* and *oriS*), a *pac*-delected, *ICP27*-deleted, HSV-BAC (fHSVΔpacΔ27 0+), and an *ICP27* expressing plasmid (pEBHICP27). Helper functions for the amplification and packaging/cleavage of amplicon DNA into HSV-1 virions are supplied in trans by fHSVΔpacΔ27 0+ and pEBHICP27. However, fHSVΔpacΔ27 0+ itself cannot be packaged because it lacks the packaging signal, an essential *ICP27* gene and HSV-1 origin of replication. Moreover, its size is overlarge for the HSV-1 virion.

1.9.5 HSV-1 Amplicon System for Gene Therapy

Herpes simplex virus type 1 (HSV-1) amplicon vectors allow the delivery of very large DNA fragments (~152kb) into most mammalian cells. Several unique features of HSV-1 amplicons have made this technology a promising tool for gene transfer. These include: (1) a very large transgene capacity; (2) they infect a wide variety of mammalian cell types

including dividing and non-dividing cells; (3) immunogenic and cytotoxic effects are absent because of the lack of viral genes (Saeki et al., 2003).

HSV-1 amplicon vectors have been used in a number of gene transfer studies in animal models of ischemic injury (Davis et al., 2007; Hoehn et al., 2003; Lawrence et al., 1995; Yenari et al., 1998), Parkinson's disease (Sun et al., 2005), cancer therapy (Mizuno et al., 2000; Shah et al., 2003; Wang et al., 2002) and hereditary ataxia (Lim et al., 2007). A recent study utilised the oncolytic HSV-1 amplicon vector containing miRNA complementary target sequences for miR-143 or miR-145 to selectively replicate in prostate cancer cells and exhibited over 80% reduction in tumour size in prostate cancer model of mice (Lee et al., 2009a). Similarly, a recent clinical trial of gene therapy for chronic intractable pain has used replication-defective HSV vectors to deliver human preproenkephalin (*PENK)* gene to assess the safety and efficacy of HSV-1 vectors in clinical settings (Fink et al., 2011). Although, HSV-1 amplicon vectors have, so far, not been used in any clinical trials due to scale-up limitations (inability to produce high-titre stocks) and transient and unregulated expression of the transgenes (de Silva and Bowers, 2009), their *in vitro* application to deliver single or multiple transgenes and complete genomic loci along with their endogenous regulatory regions has been established in many studies. For example, Wade-Martins and colleagues used the HSV-1 amplicon BAC vector to deliver and express hypoxanthine phosphoribosyltransferase (*HPRT*) locus in *HPRT*-deficient human fibroblast cells and mouse primary hepatocytes (Wade-Martins et al., 2001). Later, the complete genomic locus of human low-density lipoprotein receptor (*LDLR*) was transferred and shown to complement the *LDLR*-deficiency in both an established cells line ($LDL^{-/-}$) and in human fibroblasts derived from a familial hypercholesterolemia patient using the same system (Wade-Martins et al.,

2003). In both studies, the episomal retention of the HSV-1 amplicon BAC vector was maintained by the incorporation of *OriP* and *EBNA-1* elements of Epstein-Barr virus (EBV). In another study by our laboratory, the episomal retention elements were substituted by higher-order alpha satellite sequences to generate human artificial chromosome (HAC). The HAC carried the *HPRT* gene and exhibited efficient mitotic stability and successful *HPRT*-complementation, in an *HPRT*-deficient HT1080 cell line (Moralli et al., 2006).

1.10 Project Outline

The objective of this project was to investigate the potential of the HSV-1 amplicon system in direct reprogramming and to further extend the system for human artificial chromosome (HAC) development in human embryonic stem (hESc) cells and human induced pluripotent stem (hiPS) cells.

The methods utilised for the generation of hiPS cells suffer from one or many hurdles such as risk of integration, low reprogramming efficiencies, requirement of multiple treatments and removal/screening of transgenes. It is therefore desirable to evaluate the potential of other vector systems to efficiently derive the iPS cells for clinical settings. A potential vector system is represented by the HSV-1 amplicon. It is however unknown whether the HSV-1 amplicon system can accomplish the objectives of iPS generation in a safe and efficient manner.

In addition to the derivation of iPS cells, safe and efficient genetic modification of hESc and hiPS cells pose a significant challenge in progressing towards clinical applications of stem cell-based gene therapy. As previously described, *de novo* HAC have been demonstrated to be a promising candidate for gene transfer into tumour-derived or

transformed cell lines. It is however unknown (i) if HAC can be established in clinically relevant cell systems, particularly hESc and hiPS cells; (ii) whether karyotypically normal stem cells can support the HAC formation and transgene expression; (iii) why the efficiencies of HAC development varies with different cell systems; (iv) can these efficiencies be improved using HSV-1 amplicon system; (v) and whether HAC establishment in pluripotent stem cells affects the pluripotency.

The work presented in this study was conducted to address these queries. Chapter 3 of the thesis builds on the initial work of HAC establishment using the HSV-1 amplicon system by our group, evaluating their development and functional behaviour in hESc. Chapter 4 summarises the initial attempts of reprogramming somatic fibroblasts by using the HSV-1 amplicon system to deliver the reprogramming factors. Chapter 5 focuses on the derivation of iPS cells using the modified HSV-1 amplicon vectors carrying EBNA-1/OriP elements. The efficiency of iPS generation using the modified amplicon system is also assessed in relation to the retroviral system, originally developed by Shinya Yamanaka's group. Chapter 6 addresses the extension of HAC development in the hiPS cells. Both a commercially available iPS cell line and an iPS cell line derived via HSV-1 amplicon system were evaluated for their potential to support HAC establishment. The chapter also deals with the requirements of a modified HAC vector and assesses the efficiency of HAC delivery in iPS cells in comparison to the hESc. Finally, Chapter 7 presents an overview of the important advances that have developed through this study and highlights the prospective directions towards which the presented work can be extended.

Chapter 2. General Materials and Methods

2.1 Suppliers

All general chemicals were purchased from Sigma-Aldrich and all tissue culture reagents were supplied by Gibco (Invitrogen), unless otherwise stated. Restriction enzymes were obtained from New England Biolabs (NEB). Primers and oligonucleotides were synthesised by Eurofins MWG Operon. Phosphate Buffer Saline (PBS) was supplied by PAA.

2.2 Solutions and Buffers

Solutions and Buffers used in the experiments are listed in Table 2-1.

2.3 Growth and Maintenance of *E. coli*

2.3.1 *E. coli* Culture

Escherichia coli (*E. coli*; DH10β, DH5αpir and Stbl4 strains, Invitrogen) were grown at 37ºC overnight in Luria-Bertani (LB) broth with shaking at 300rmp or streaked out on LB agar plates. Where applicable, antibiotic selection was carried out at 100µg/ml ampicillin, 33µg/ml kanamycin and 20µg/ml chloramphenicol. For long-term storage of bacterial stocks, cultures were stored in LB broth containing 25% (v/v) glycerol at -80ºC.

Table 2-1 Media and Solutions

Media	
Luria-Bertani (LB) broth	Prepared using Sigma LB broth powder
Luria-Bertani (LB) agar	Prepared using Sigma LB Agar powder
SOC medium	Sigma Hanahan's broth (SOB), 20mM glucose
Solutions	
KCM buffer	120mM KCl, 20mM NaCl, 10mM Tris-HCl (pH 8.0), 0.5M EDTA (pH 8.0), 0.1% (v/v) Triton X-100.
20x SSC	3 M NaCl, 0.3 M $C_6H_5Na_3O*2H_2O$ (trisodium citrate 2 hydrate).
50x TAE	40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA (pH 8.0).
5x TBE	450 mM Tris base, 450 mM orthoboric Acid, 10 mM EDTA (pH 8.0).
TBS	10 mM Tris-HCl (pH 8.0), 150 mM NaCl.
TE	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0).
TE 50:10	50 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0).
10x TNE	100 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 2 M NaCl.
Antifading solution	0.2M DABCO(1,4-diazabicyclo[2.2.2]octane), 90 % (v/v) glycerol in PBS
Cell lysis buffer	100 mM NaCl, 10mM Tris-HCl (pH 8.0), 25 mM EDTA (pH 8.0), 0.5% (w/v) SDS
FISH denaturation buffer	50mM KCl, 10mM Tris-HCl (pH 8.0) and 5% (v/v) glycerol
Probe hybridisation buffer	2x SSC, 10% (w/v) dextran sulphate, and 50% (v/v) formamide
Immuno-FISH blocking and detection buffer	3% (w/v) BSA, 0.1% (v/v) Triton X-100 in PBS

2.3.2 Plasmid DNA Preparation

E. coli, containing plasmids were cultured by inoculating a single bacterial colony in 10ml of LB broth containing relevant antibiotics at 37ºC overnight with shaking. These cultures were either harvested for preparing DNA as a small scale preparation using the Qiagen Spin Miniprep kit or 1ml was used to inoculate 250ml of LB broth containing antibiotic and the culture was grown in agitation at 300rpm overnight at 37ºC before a

large scale preparation was performed using the Qiagen Plasmid Maxiprep kit. The concentration of plasmid DNA was measured using a NanoDrop ND-1000 spectrophotometer. Plasmid vectors utilised in this study along with their characteristics are listed in Table 2-2. Maps of the HSV-1 amplicon vectors are presented in Figure 2-1.

2.3.3 Preparation of Electrocompetent Bacteria

A single bacterial colony from a fresh streak was picked to seed 1.5ml of culture in LB broth. The culture was grown overnight at 37ºC with agitation. Next day, 100ml of LB broth was inoculated with 1.5ml of starter culture and grown at 37ºC with agitation for approximately 3 hours until the optical density at 600nm (OD_{600}) of the culture reached between 0.4 and 0.5. The culture was chilled on ice for 30 minutes and then pelleted at 4,500rpm for 10 minutes at 4ºC and resuspended in 100ml of ice-cold 10% (v/v) glycerol. This process of centrifugation and resuspension was repeated for two more times. After the final centrifugation the bacterial pellet was resuspended in 600µl of ice-cold 10% (v/v) glycerol, aliquoted and stored at -80ºC.

2.4 Transformation of Electrocompetent Bacteria

Ligation products were desalted using Millipore 0.025µm membrane filters by dialysing on distilled water according to the manufacturer's instructions. Prior to electroporation, electrocompetent cells (DH10β, DH5αpir and Stbl4 strains, Invitrogen) were thawed on ice and 200ng of ligation products or 10pg of plasmid DNA were added to 40µl of electrocompetent cells. The cell/DNA mixture was transferred into a chilled 0.1cm electroporation cuvette (Bio-Rad) and electroporated at 2.0 kilovolts (kV), 200 ohms (Ω) and 25 microfarads (µF) with a Bio-Rad Gene Pulser. 1ml of the SOC medium was added to the cuvette and the combined contents were transferred into a 1.5ml fresh tube before

incubating at 37ºC for 1 hour with agitation. The culture was subsequently spread onto LB agar plate containing the appropriate antibiotics and incubated at 37ºC overnight.

Table 2-2 Vectors utilised in this study

Vectors (References)	Characteristics	Size	Reporter Gene (Promoter)	Mammalian Resistance Gene (Promoter)	Bacterial Resistance Gene
pSG80A-HG (1)	HSV-1 amplicon vector; contains HSV-1 *oriS* and *pac*; *loxP* site and R6Kγ *ori*	5.9kb	*EGFP* (HSV-1 IE4/5)	-	*Amp*
pHGNeo4 (2)	HSV-1 amplicon vector; based on pSG80A-HG backbone	7.9kb	*EGFP* (HSV-1 IE4/5)	*Neo* (SV40)	*Amp*, *Kan*
pHG-Puro (3)	HSV-1 amplicon vector; based on pSG80A-HG backbone	7.6kb	*EGFP* (HSV-1 IE4/5)	*Pac* (SV40)	*Amp*
pBeloBAC11 (4)	BAC vector; *loxP* site	7.5kb	-	-	*Cm*
pCTP-T (5)	Contains pSC101 *ts ori*; *pir* gene; *Cre* gene (tetracycline inducible)	12.7kb	-	-	*Tet*
fHSVΔpacΔ27 0+ (5)	A PAC vector; contains HSV-1 genome deleted for *oriS*, *pac*, and *ICP27* gene	178kb	-	-	*Cm*, *Kan*
pEBHICP27 (5)	Contains *ICP27* gene	12.2kb	-	-	*Amp*

References: 1 (Wade-Martins et al., 2003); 2 (Mandegar, 2011; Mandegar et al., 2011); 3 (Chan, 2010); 4 (Kim et al., 1996); 5 (Saeki et al., 2001)

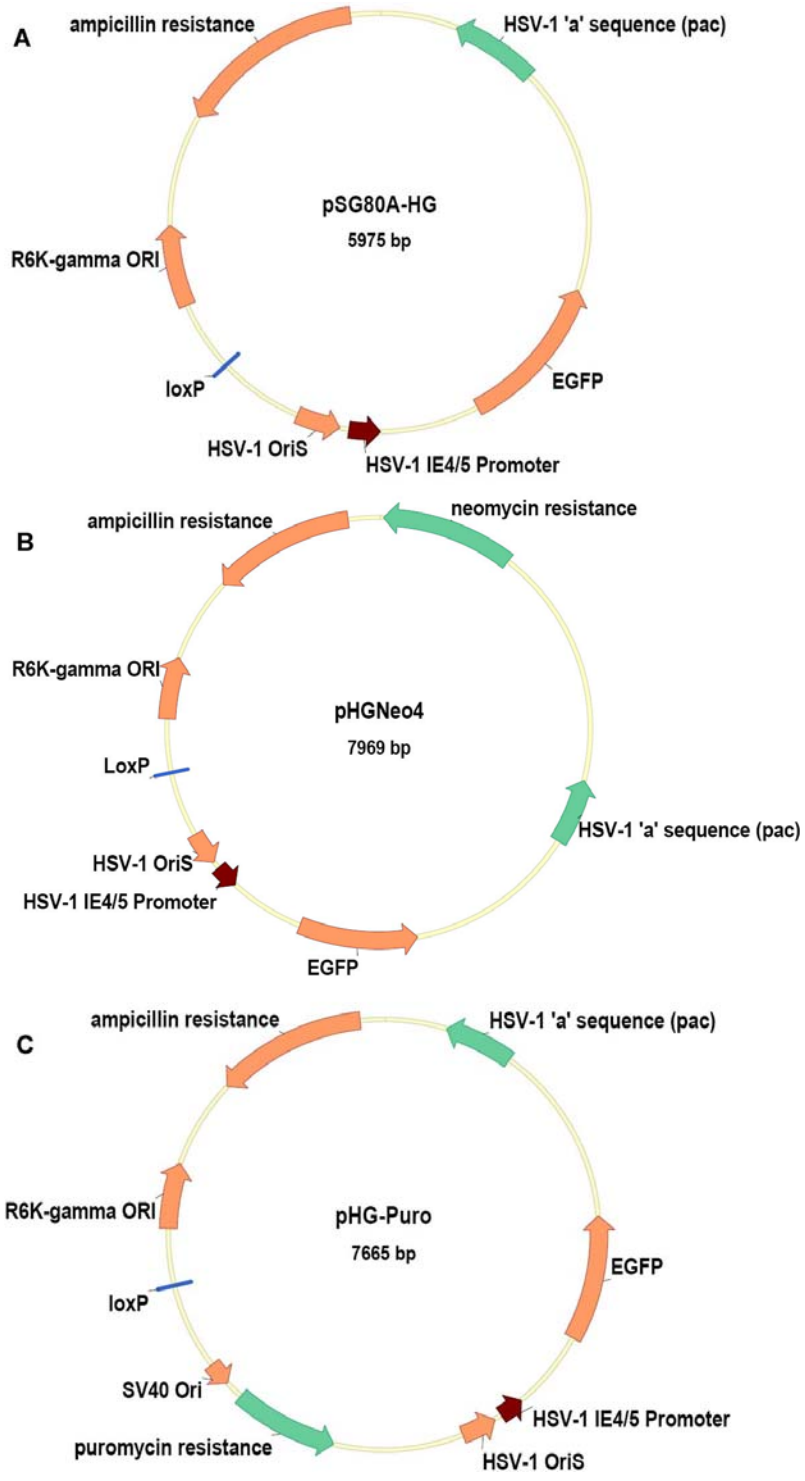

Figure 2-1 Vector maps of pSG80A-HG, pHGNeo4 and pHG-Puro
The HSV-1 amplicon vector pSG80A-HG (A) contains the *R6K-γ* origin of replication, HSV-1 origin of replication (*OriS*), HSV-1 packaging element (*pac*), *EGFP* gene and a *loxP* site (vector a gift from Dr Richard Wade-Martins). The vectors pHGNeo4 (B; constructed by Dr. Mohammad A Mandegar) and pHG-Puro (C; constructed by Dr. David YL Chan) were derived from pSG80A-HG and contain the neomycin resistance and puromycin resistance genes respectively.

2.5 *Cre*/*loxP* Recombination

The method involved introducing 50ng of the tetracycline inducible *Cre* expressing plasmid pCTP-T and 100ng of the retrofitting plasmid carrying the pir dependent *R6Kγ* origin of replication, a *loxP* site and an ampicillin resistance gene, into 40µl of the *pir* negative electrocompetent cells. The cells contained the vector to be retrofitted, carrying a *loxP* site and kanamycin or chloramphenicol resistance gene. Then 500µl of the SOC medium containing 20µg/ml heat-inactivated chlorotetracycline (cTc) was added to the cuvette and the combined contents were transferred into a 1.5ml fresh tube and incubated at 30ºC for 1 hour with shaking. The cell suspension was then diluted 1:10 in 5ml of LB containing the appropriate antibiotics (ampicillin plus kanamycin, or chloramphenicol) and incubated at 30ºC for additional 3 hour with shaking. The culture was spread onto LB agar plate containing antibiotics and incubated at 42ºC overnight. The plasmid pCTP-T expressed the *Cre* recombinase from the tetracycline inducible promoter and, due to the presence of *pSC101* temperature-sensitive origin of replication, replicates at 30ºC but not at 42ºC (Saeki et al., 2001).

2.6 DNA Cloning

2.6.1 Restriction Enzyme Digestion

Plasmid DNA and BAC were digested with one or two restriction enzymes in a final volume of 10-20µl containing 1x digestion buffer and 1x BSA as recommended by the manufacturer. Then 5mM spermidine (Sigma) was added to the reaction mix when the NaCl concentration in the digestion buffer exceeded 50mM; otherwise 2mM spermidine was used. For each µg of DNA, 5-10 units of the restriction enzymes were used at no more than 10% (v/v) of total reaction volume. Restriction digestion was carried out for 2-

3 hours or incubated overnight at the appropriate temperature recommended by the manufacturer.

2.6.2 Agarose Gel Electrophoresis

The DNA fragments were resolved by electrophoresis through 0.7-2% (w/v) agarose gel in 1x TAE buffer containing 0.5µg/ml of ethidium bromide. DNA samples were mixed with 6x loading dye (NEB) prior to loading on gels, a 100bp or 1kb DNA ladder (NEB) was added to each gel for size comparison of DNA fragments and electrophoresis was carried out in 1x TAE buffer containing 0.5µg/ml of ethidium bromide at 80-130V. The DNA fragments were visualised by exposure to ultraviolet (UV) light by using AlphaImager gel documentation system (Alpha Innotech) and images were acquired using AlphaView software version 1.0.3.7.

2.6.3 DNA Isolation from Agarose Gels

DNA fragments on agarose gels were visualised under UV light and gel slices containing the DNA fragments of interest were excised using a clean scalpel blade. DNA was extracted from gel slices using QIAquick gel extraction kit (Qiagen) following the manufacturer's instructions.

2.6.4 Dephosphorylation and Ligation

Digested vectors with compatible restriction sites at the termini were dephosphorylated to avoid self ligation. 0.5-1µg of the desired vector was treated with 2 units of shrimp alkaline phosphatase (Roche) in 1x dephosphorylation buffer at 37ºC for 2 hours, followed by heat inactivation at 65ºC for 25 minutes. For ligation, 50ng of the digested vector DNA and 150ng of the insert DNA were treated with 400 units of T4 DNA ligase

(NEB) in 1x ligation buffer and incubated overnight at 16ºC. Ligation reactions were subsequently used for electroporation into the bacteria.

2.7 Pulsed Field Gel Electrophoresis (PFGE)

Larger DNA fragments were resolved by Pulsed Field Gel Electrophoresis (PFGE) through a 1% (w/v) SeaKem GTG agarose (Cambrex Bioscience) gel in 0.5x TBE buffer using a CHEF Mapper Pulsed Field Gel Electrophoresis system (Bio-Rad). Electrophoresis was carried out at 14ºC for 15 hours at 200V, with an initial pulse time of 0.2 seconds, final pulse time of 22 seconds and a pulse angle of 120º. Lambda ladder PFG Marker (NEB) and HyperLadder VI (Bioline) were used to determine the size of the DNA fragments. After the completion of run, gels were stained with 0.5µg/ml ethidium bromide in water for 1 hour. The DNA fragments were visualised using AlphaImager gel documentation system (Alpha Innotech) and images were acquired using AlphaView software version 1.0.3.7.

2.8 Cell Culture

The cell lines that were utilised in this study along with their characteristic and origin are described in Table 2-3. Unless otherwise mentioned in Table 2-3, all the cells were maintained in standard DMEM (Dulbecco's Modified Eagle's) medium, high glucose, GlutaMAX, pyruvate supplemented with 10% (v/v) foetal bovine serum (FBS; PAA Laboratories) and 1% (v/v) penicillin/streptomycin (100U/ml penicillin, 10mg/ml streptomycin, PAA Laboratories). The cells were grown at 37ºC and 5% CO_2 using standard tissue culture techniques. After reaching 70-90% confluency, the cells were sub-cultured by washing the monolayer with PBS followed by incubation in 1x trypsin/EDTA in PBS for 3-5 minutes at 37ºC and subsequently were plated in new culture dishes/flasks

with 1:2 to 1:10 dilution in their respective growth medium. For long term storage, cells were trypsinised as above, centrifuged at 1000rmp for 5 minutes and resuspended in cryopreservation medium (growth medium containing 10% (v/v) dimethyl sulfoxide, DMSO). 1ml of cell suspension was then aliquoted into a cryovial and stored at -80°C overnight before transferring them to liquid nitrogen.

Table 2-3 Growth conditions and features of cell lines

Cell Line	Growth Medium	Origin of Cell line
HT1080	Standard DMEM	Human cell line derived from fibrosarcoma (Rasheed et al., 1974)
HeLa	Standard DMEM	Human cell line derived from cervix epithelioid carcinoma
G16-9	Standard DMEM (2% FBS) + 200µg/ml hygromycin	Derivative of the human glioma cell line Gli-36 (Kashima et al., 1995) that expresses the HSV-1 *VP16* gene; a kind gift from Dr. Richard Wade-Martins
Vero2-2	Standard DMEM + 500µg/ml G418	A derivative of African Green monkey Vero cells that expresses HSV-1 gene *ICP27* (Smith et al., 1992); a kind gift from Rozanne M. Sandri-Goldin
SNL 76/7	Standard DMEM (7% FBS)	Derivative of mouse STO cell line that expresses neomycin phosphotransferase (*Neo*) and leukaemia inhibitory factor (*LIF*) (McMahon and Bradley, 1990)
PO MEFs	Standard DMEM	Primary mouse embryonic fibroblasts derived from Pathology Oxford outbred strain
CF1 MEFs	Standard DMEM	Primary mouse embryonic fibroblasts derived from CF1 strain
IMR90	Standard DMEM	Primary human foetal lung fibroblasts (Nichols et al., 1977)

2.9 Preparation of MEF Feeder Cells

Primary mouse embryonic fibroblasts (MEFs) used in this study were obtained from either the Pathology Oxford (PO) outbred strain (Sir William Dunn School of Pathology, University of Oxford.) or the CF1 strain (Millipore, catalogue number: PMEF-CFL).

Prior to plating MEFs, the T-175 flask was coated with 0.1% gelatin and incubated at 37ºC for 30 minutes. After the incubation, stock vials of MEFs were thawed at 37ºC, excess gelatin aspirated from the flask and cells were plated in the DMEM medium. Cells were incubated at 37°C and 5% CO_2 until they reached 80-90% confluence, after which they were sub-cultured into three gelatin coated T-175 flasks for expansion. Cells from these flasks were sub-cultured once more into ten gelatin-coated T-175 flasks and grown till they reached 80-90% confluence. Subsequently, cells were inactivated by treating with 10µg/ml mitomycin-C solution in MEF media for 2.5 hours at 37°C and 5% CO_2. After the treatment, the medium containing the mitomycin-C was removed, the inactivated MEFs (iMEFs) were washed twice with PBS and treated with 3ml trypsin for 5 minutes at 37°C. MEF medium (7ml/flask) was added to inactivate trypsin and the cells from all the flasks were pooled and counted using a haemocytometer. Subsequently, cells were centrifuged at 1000rpm for 5 minutes, resuspended in cryopreservation medium and $2x10^6$ cells were aliquoted in cryovials. The vials were stored at -80°C overnight and placed in liquid nitrogen for long-term storage. The SNL 76/7 cells were inactivated in a similar fashion as MEFs but during the inactivation they were treated with 12µg/ml mitomycin-C solution in the DMEM medium (7% FBS).

2.10 Preparation of MEF-Conditioned Medium

To prepare MEF-conditioned medium, iMEFs were plated at a density of $2x10^5$ cells/well on gelatin-coated 6-well plates in the DMEM medium and incubated overnight at 37°C and 5% CO_2. Thereafter, DMEM medium was replaced with 2.5ml hESc medium (without bFGF) and cells were incubated at 37°C at 5% CO_2. After 24 hours, the medium was recovered and fresh hESc medium was added. This cycle was continued for a total of

five days. The recovered medium was filtered using 0.22µm filter and supplemented with 10ng/ml bFGF.

2.11 hESc and hiPS Cell Culture

2.11.1 hESc and hiPS Cell Lines

The human embryonic stem cell (hESc) line HUES10 was obtained from Douglas Melton (Harvard University, USA) under the license from the UK Stem Cell Steering Committee. The human induced pluripotent stem cell (hiPS) line DF19-9-11T.H was obtained from WiCell Research Institute (Wisconsin, USA).

2.11.2 Culture of hES and hiPS Cells

The hESc and hiPS cells were maintained on iMEFs in hESc media consisting of Dulbecco's modified Eagle (DMEM)-F/12 medium supplemented with 20% (v/v) KnockOut Serum Replacement, 0.1mM non-essential amino acids, 1x Glutamax^{TM}-1, 0.1mM ß-mercaptoethanol, 10ng/ml basic fibroblast growth factor (bFGF; PeproTech Inc.) and 1% (v/v) penicillin and streptomycin (P/S). hESc medium was changed every day. Feeder-free culture of hESc and hiPS cells was carried out on Matrigel (BD Biosciences) coated dishes with either mTeSR1 medium (STEMCELL Technologies Inc.) or MEF-conditioned medium supplemented with 10ng/ml bFGF. mTeSR1 medium and MEF-conditioned medium were changed every day.

2.11.3 Passaging of hESc and hiPS Cells

One day before passaging of hESc and hiPS cells, inactivated MEFs were plated onto the dishes as described in section 2.10. For feeder-independent culture, dishes were coated

with Matrigel and incubated for 1 hour at room temperature. Upon reaching confluence, the hESc and hiPS cells were washed with PBS and incubated in TrypLE Express at 37°C for 2-3 minutes. Dissociated cells were transferred to a 15ml tube containing 10ml PBS and centrifuged at 1000rpm for 5 minutes. Cells were then resuspended in either hESc medium for feeder-dependent culture and plated onto dishes pre-coated with inactivated MEFs, or mTeSR1 medium for feeder-free culture and plated onto Matrigel-coated dishes. 10μM ROCK inhibitor Y-27632 (Merck Biosciences) was added into the medium during each passaging step to increase single cell survival.

2.12 Embryoid Body (EB) Formation

The hESc and hiPS cells were dissociated using TrypLE Express and centrifuged at 1000rpm for 5 minutes. The pelleted cells were resuspended in hES media (without bFGF) containing 10μM Y-27632 and allowed to aggregate in 96-well V-bottom dish (Nunc) to form embryoid bodies (EBs). After plating the cells, the dish was centrifuged at 500rpm for 5 minutes to assist settling of the cells. Cells were subsequently incubated for three to four days to form EBs. This method generated large EBs (~100,000 cells/EB), which were used in the preparation of hESc-derived fibroblasts cells, germ layer differentiation and directed neuronal differentiation.

2.13 Germ Layer Differentiation

The hESc and hiPS cells were differentiated into derivatives of the three primary germ layers from EBs. EBs were released from 96-well V-bottom dish and plated onto ultra-low adherent 6-well dishes (Costar, Corning) in fibroblast medium consisting of DMEM-F/12 supplemented with 20% (v/v) FBS, 0.1mM non-essential amino acids, 1X Glutamax^{TM}-1, 0.1mM ß-mercaptoethanol, and 1% penicillin and streptomycin (P/S).

After five days in suspension culture, the EBs were transferred to gelatin-coated 6-well dishes in the fibroblast medium and cultured for an additional nine to ten days. After 10-12 days of cultivation, images of different derivatives of germ layers were acquired with a Nikon Eclipse TE2000U inverted microscope with a Nikon Plan Fluor 4×/0.13 phL DL ∞/1.2 WD 16.4 objective and CFI 10×/22 eyepiece lens using a Hamamatsu C4742-95 Orca low-noise B&W CCD camera and NIS-Elements software version 2.34. EBs and its derivative cells were harvested for RT-PCR analysis. Table 2-4 list the primers used for RT-PCR analysis of germ layer differentiation.

2.14 Neuronal Differentiation

Neuronal cells were derived from hESc and hiPS cells through EB-mediated differentiation. EBs were plated on ultra-low adherent 6-well dishes (Costar, Corning) in neuronal differentiation media-I (NDM-I) consisting of DMEM-F/12 supplemented with 1% (v/v) N2 supplement (GIBCO), 5μg/ml human plasma fibronectin (Sigma), 200ng/ml recombinant human noggin (Fitzgerald), 1x Glutamax^{TM}-1, and 1% (v/v) penicillin and streptomycin (P/S), and left for 5 days in suspension culture. The EBs were then plated onto Matrigel-coated 24-well dishes in NDM-I medium. After 8 days, the medium was switched to neuronal differentiation media-II (NDM-II) consisting of DMEM-F/12 supplemented with 1% (v/v) N2 supplement, 20ng/ml recombinant bovine bFGF (R&D systems), 1x Glutamax^{TM}-1, and 1% (v/v) penicillin and streptomycin (P/S) to expand the neuronal rosettes. The medium was replenished every 2-3 days and the culture continued for a total 8-10 days. Neuronal rosettes were then dissociated using TrypLE Express and passaged onto Matrigel coated dishes in NDM-II medium for further expansion, RT-PCR analysis and immunofluorescence staining.

Table 2-4 Primers for the RT-PCR analysis on markers of all the three germ layers

Genes/Region (Germ layer)	Symbol	Sequence (5'- 3')
α-fetoprotein (Endoderm)	αFP F	CCATGTACATGAGCACTGTTG
	αFP R	CTCCAATAACTCCTGGTATCC
HNF-3α (Endoderm)	HNF-3α F	GAGTTTACAGGCTTGTGGCA
	HNF-3α R	GAGGGCAATTCCTGAGGATT
α1-antitrypsin (Endoderm)	α1AT F	AGACCCTTTGAAGTCAAGGACACCG
	α1AT R	CCATTGCTGAAGACCTTAGTGATGC
GATA-2 (Mesoderm)	GATA-2 F	AGCCGGCACCTGTTGTGCAA
	GATA-2 R	TGACTTCTCCTGCATGCACT
Cytokeratin 14 (Ectoderm)	CK-14 F	ATGATTGGCAGCGTGGAG
	CK-14 R	GTCCAGCTGTGAAGTGCTTG
Cytokeratin 5 (Ectoderm)	CK-5 F	TCAAGGATGCCAGGAACAAG
	CK-5 R	GCTTGCACTGAAGCCAGAG
High sulfur keratin (Ectoderm)	HSK F	AGGAAATCATCTCAGGAGGAAGGGC
	HSK R	AAAGCACAGATCTTCGGGAGCTACC

2.15 Transfection of Cultured Cells

The day before transfection, cells were seeded on a 24-well dish at 70-80% confluency. For each well, 3μg of plasmid or BAC DNA were transfected using ExGen 500 (Fermentas) according to the manufacturer's instructions. The transfection efficiency was monitored 24-48 hours post-transfection by checking the transient *GFP* expression using Nikon Eclipse TE2000U inverted microscope with a Nikon Plan Fluor 4×/0.13 phL DL ∞/1.2 WD 16.4 objective and CFI 10×/22 eyepiece lens. Transfected cells were either subjected to selection for stable clone formation or were harvested for DNA/RNA extraction.

2.16 Cellular and Molecular Analysis

2.16.1 Isolation of Genomic DNA

Genomic DNA was extracted from 1-4x10^6 cells using the "salting out" method. Cells were trypsinised and pelleted at 13,000rpm for 1 minute. Cells were then washed with 1ml of water and centrifuged at 13,000rpm for 1 minute. Subsequently, cells were resuspended in 200µl of cell lysis buffer containing 100-125µg/ml proteinase K (Roche), and incubated at 55°C overnight. After incubation, the cell debris was precipitated by addition of 200ul of 6M NaCl and vigorous shaking. Precipitated contents were removed by centrifugation at 13,000 rpm for 5 minutes and supernatant was transferred to a fresh tube. To the supernatant, an equal volume of isopropanol was added to the precipitate the DNA. The DNA was recovered by centrifugation at 13,000rpm for 5 minutes at 4°C and washed with 1ml of 70% (v/v) ethanol. After a final centrifugation at 13,000rpm for 5 minutes at 4°C, the DNA pellet was allowed to air-dry and resuspended in 100µl of water. DNA was quantified using a NanoDrop ND-1000 spectrophotometer.

2.16.2 RNA Extraction and Reverse Transcription

Total RNA was extracted from 1-4x10^6 cells using an RNeasy mini kit (Qiagen) with on-column DNase digestion according to the manufacturer's instruction. The RNA was eluted in 40-50µl of RNAse-free water and quantified using a NanoDrop ND-1000 spectrophotometer. Extracted RNA was either stored at -80°C for future use or subjected to cDNA synthesis. For the preparation of cDNA, 1-2µg of total RNA was subjected to reverse transcription as described in the qScript cDNA SuperMix kit (Quanta Biosciences). The RNA and 5x reaction buffer were combined in a 0.2ml micro-tube and

the reaction was incubated at 25°C for 5 minutes and 42°C for 1 hour followed by heat inactivation of the reverse transcriptase at 85°C for 5 minutes.

2.16.3 Polymerase Chain Reaction (PCR)

The PCR amplification was performed in a 25µl reaction mixture containing 100ng of template DNA or cDNA, 0.5µM of each forward and reverse primer and 1x BioMix Red (Bioline). For a typical PCR reaction, initial denaturation was carried out at 95°C for 5 minutes followed by 35-40 cycles consisting of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds/kb. Final extension was performed at 72°C for 10 minutes. Reactions were executed on PTC-200 Peltier thermal cycler (MJ Research). Table 2-5 lists the primers sets used for the PCR analysis.

Table 2-5 Primer sets used for PCR amplification

Genes/Region	Symbol	Sequence (5'- 3')
OCT4	OCT4 F	AGTGAGAGGCAACCTGGAGA
	OCT4 R	GCCGGTTACAGAACCACACT
SOX2	SOX2 F	ACTTTTGTCGGAGACGGAGA
	SOX2 R	CATGAGCGTCTTGGTTTTCC
NANOG	NANOG F	TTCCTTCCTCCATGGATCTG
	NANOG R	ATCTGCTGGAGGCTGAGGTA
GAPDH	GAPDH F	TGTTGCCATCAATGACCCCTT
	GAPDH R	CTCCACGACGTACTCAGCG

2.16.4 Quantitative Real-Time RT-PCR

Expression level of the genes and transgenes was quantified by quantitative Real-Time PCR (qPCR) using B-R SYBR® Green SuperMix for iQ™ (Quanta Biosciences) on Bio-Rad iQ^{TM}5 Multicolor Real-Time PCR Detection System. Reactions were performed in triplicate in 25µl reaction mixtures containing 100ng of template cDNA, 0.5µM of each forward and reverse primers and 1x SYBR Green SuperMix. Initial denaturation was carried out at 95°C for 3 minutes followed by 40 cycles consisting of denaturation at 95°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds. The quantification was analysed using the C_T (threshold cycle) values obtained with Bio-Rad iQ5 standard edition software. The expression of the target genes was normalised to that of the endogenous *GAPDH* and the fold enrichment was measured using the $2^{-\Delta\Delta Ct}$ method as follows.

$$\text{Fold Enrichment} = 2^{-\Delta\Delta Ct},$$

$$\text{Where } \Delta\Delta Ct = \Delta Ct_{\text{(Time point 0)}} - \Delta Ct_{\text{(Time point t)}},$$

$$\Delta Ct_{\text{(Time point 0)}} = \Delta Ct_{\text{(Target gene, Time point 0)}} - \Delta Ct_{\text{(}\textit{GAPDH}\text{, Time point 0)}},$$

$$\text{and } \Delta Ct_{\text{(Time point t)}} = \Delta Ct_{\text{(Target gene, Time point t)}} - \Delta Ct_{\text{(}\textit{GAPDH}\text{, Time point t)}}.$$

2.17 HSV-1 Amplicon Preparation and Infection

2.17.1 Packaging of HSV-1 Amplicon Vectors

African Green monkey kidney cells (Vero2-2) were seeded in 6cm dishes at a density of $1x10^6$ and grown overnight to reach >90% confluency. The next day, per dish, 2µg of fHSVΔpacΔ27 0+ DNA, 0.2µg pEBHICP27 DNA and 1.8µg of amplicon DNA (to be packaged) were mixed in 500µl of Opti-MEM medium with 10µl of PLUS reagent (Life Technologies) and 23µl of LipofectAMINE reagent (Life Technologies). The transfection mixture was supplemented with Opti-MEM medium to a final volume of 1.5ml and applied to the Vero2-2 cells. Cells were incubated at 37°C 5% CO_2 for 4 hours, washed thrice with 2.5ml Opti-MEM medium and replenished with 3.5ml DMEM-GlutaMAX medium containing 6% (v/v) FBS, 25mM HEPES and 1% (v/v) P/S. After 60-66 hours of incubation, the cells were harvested by scraping into the medium and frozen at -80°C. Frozen cell suspension was then thawed at 30°C, sonicated for 15-30 seconds at 50% amplitude (Fisher Scientific Sonic Dismembrator Model 500 with a tapered microtip, 2mm tip diameter) and centrifuged at 3,500rpm for 15 minutes at 4°C to remove the cell debris. Amplicons were then concentrated by overlaying the supernatant on a 25% (w/v)

sucrose cushion and spun for 3 hours at 22,000rpm at 4°C using a SW28 rotor (Beckman Coulter). The pellet was resuspended in 250-500µl PBS, aliquoted and stored at -80°C.

2.17.2 Titration of HSV-1 Amplicon Preparation

The titre of HSV-1 amplicon was estimated by infecting G16-9 cells (Kubo et al., 2003) (a derivative of the human glioma cell line Gli36, highly susceptible to HSV-1 infection, which expresses the HSV-1 *VP16* gene). G16-9 cells were seeded at $5x10^5$ cells/well in a 24-well plate. Next day, cells were transduced with 0.5µl, 1µl, 2µl and 5µl of amplicon suspension in a total volume of 250µl of media per well, and incubated at 37°C and 5% CO_2. After 24 hours of transduction, ~100-200 *GFP* expressing cells were scored using a Nikon Eclipse TE2000U inverted microscope with a Nikon Plan Fluor 4×/0.13 phL DL ∞/1.2 WD 16.4 objective and CFI 10×/22 eyepiece lens. Each fluorescent cell was considered as deriving from one transducing unit (TU) and titre of the amplicon preparation was thus calculated by scoring the total number of fluorescent cells per well as a function of the volume of amplicon used in that well. Titres of 10^7-10^8TU/ml were routinely obtained.

2.17.3 Transduction with HSV-1 Amplicons

The day before transduction, cells were seeded in a 24-well dish. On the day of transduction, the cell were infected with HSV-1 amplicons at a multiplicity of infection (MOI) of 1, 2 or 5 in 250µl of media and incubated for 24 hour at 37°C and 5% CO_2. For the transduction of hESc and hiPS cells, the cells were seeded at $5x10^5$ cells/well in a 24-well dish (Matrigel coated) and were infected at desired MOI in 250µl of mTeSR1 media the following day. Transduction efficiency of hESc and hiPS cells was improved by centrifugation of the 24-well dish at 750g for 45 minutes, a method known as

spinoculation. Cells were incubated 37°C and 5% CO_2, and transient expression was monitored at 24 hours post-transduction using fluorescence microscope as described in section 2.17.2.

2.17.4 Drug Selection and Colony Picking

HT1080 and HeLa cells were selected at 350µg/ml and 400µg/ml G418 respectively. Selection was started 48 hours after tranduction and continued for 2-3 weeks until the formation of stable clones. Selection of hESc and hiPS cells was carried out at 100µg/ml G418 on mitomycin C-inactivated SNL 76/7 cells and maintained for 7-10 days. Drug selection was subsequently removed to allow the recovery of hESc and hiPS cells and colony formation. Colonies were mechanically picked using a 20µl pipette under a Motic SMZ 168 Trinocular stereomicroscope using a 2× objective and W10×/23 eyepiece lens and transferred to a 24-well dish pre-coated with inactivated MEFs in hESc media supplemented with 10µM ROCK inhibitor Y-27632.

2.18 Chromosome Harvesting

HT1080 cells were mitotically arrested in metaphase by incubating in 30ng/ml Colcemid at 37°C for four hours. The hESc and hiPS cells were arrested in metaphase by incubating in 0.1µg/ml nocodazole for 16 hours at 37°C and 5% CO_2. After removing the culture medium, cells were washed with PBS and incubated in trypsin at room temperature for 5 minutes followed by centrifugation at 800rpm for 5 minutes. HT1080 cells were resuspended in hypotonic solution (75mM KCl) pre-warmed at 37°C while hESc and hiPS cells were resuspended in pre-warmed buffered hypotonic solution (pH 7.4, Genial Genetics) and incubated at 37°C for 30 minutes. Following incubation, cells were centrifuged at 800rpm for 5 minutes and the pellet resuspended in cold Carnoy's fixative

solution (methanol:acetic acid 3:1) and incubated at 4°C for 1 hour. The cycle of centrifugation and fixation was repeated once more and after the final centrifugation, cells were resuspended in 200-500µl of Carnoy's fixative solution and stored at -20°C.

2.19 Preparation of FISH Probes

FISH probes were labelled by incorporation of biotin-16-dUTP or digoxigenin-11-dUTP (Roche) into 1µg of probe DNA in a total reaction volume of 50µl by nick translation using Nick translation system kit (Invitrogen). For consistency, α-satellite DNA was usually labelled with biotin, and vector DNA was labelled with digoxigenin. The reaction was set up according to the manufacturer's instructions except that the incubation of the nick translation reaction was carried out overnight at 16°C. After incubation, the probe was precipitated by addition of 5x excess of sonicated salmon testes DNA (Sigma), 0.2% (w/v) dextran blue, 3M ammonium acetate and 2 volumes of ethanol. Finally, the probe was pelleted by centrifugation at 13,000rpm for 10 minutes, resuspended in hybridisation buffer (Table 2-1) at a final concentration of 10ng/µl and stored at -20°C.

2.20 Fluorescence *in situ* Hybridisation (FISH)

A 30-50µl quantity of fixed chromosome suspension was dropped onto a clean microscopic slide and fixative solution was allowed to evaporate. Chromosomal DNA was denatured by adding 90µl denaturation buffer (Table 2-1) per slide and incubating it at 95°C for 10 minutes. Following denaturation, the slide was washed in 0.1x SSC for 1 minute and dehydrated for 3 minutes each in 70%, 90% and 100% ethanol. 10-15µl of labelled probes (denatured at 85°C for 10 minutes) were added onto the slide under a cover slip and the slide was incubated overnight in a humid box at 37°C. After the overnight hybridisation, the slide was washed three times in 0.1x SCC for 5 minutes each

at 60°C. Biotin-labelled probes were detected with avidin, Alexa Flour 488 conjugate (1:100, Molecular probes A-21370; [Stock]=1mg/ml) followed by biotinylated anti-avidin D antibody (1:100, Vector Laboratories BA-0300; [Stock]=0.5mg/ml), and a second layer of avidin, Alexa Flour 488 conjugate. Digoxigenin-labelled probes were detected using anti-digoxigenin-rhodamine antibody (1:50, Roche 11207750910; [Stock]=0.2mg/ml), followed by a second layer of rhodamine conjugated anti-sheep antibody (1:100, Chemicon AP147R; [Stock]=0.2mg/ml). Antibodies and avidin were diluted in 4x SSC, 0.1% (v/v) Tween 20 (VWR) and 3% (w/v) BSA. For each slide, 100µl of diluted antibodies or avidin were used and incubated for 30 minutes at 37°C followed by three washes in 4x SSC and 0.1% (v/v) Tween 20, each for 5 minutes at 42°C. Following a final wash in 4x SSC at room temperature, the chromosomes were counterstained with 10ng/ml DAPI (4',6-diamidino-2-phenylindole) and mounted in 0.2M DABCO (1,4-diazabicyclo(2,2,2)octane) antifading solution.

FISH with whole chromosome paints for chromosome 12 and 17 was performed using the FITC-conjugated Star*FISH Paint kit (Cambio) according to the manufacturer's instructions.

FISH slides were analysed with an Olympus BX-51 upright microscope using an Olympus UPlanF1 60×1.25 oil ∞/0.17 objective and Olympus widefield WHN10×/22 eyepiece lens and images were acquired and analysed using JAI CVM4+ progressive-scan 24 fps B&W fluorescence CCD camera and Leica Cytovision Genus software v7.1.

2.21 Immuno-FISH

Metaphase-arrested cells were resuspended in cold 56mM KCl solution and 200-250µl of the cell suspension was cytospun on the slides using a Labofuge 400 cytocentrifuge

(Heraeus Instruments) at 1000g for 10 minutes. Cells were then fixed in 2% (v/v) formaldehyde in PBS and permeabilised in KCM buffer for 10 minutes at room temperature. Non-specific binding was blocked with a 30 minutes incubation in blocking buffer [3% (w/v) BSA, 0.1% (v/v) Triton X-100 in PBS] at room temperature. After blocking, cells were incubated with rabbit anti-human CENP-C antibody (1:100, a gift from Professor William Earnshaw) at 4ºC overnight in a humid box followed by three washes in KCM buffer each for 5 minutes at room temperature. Cells were then treated with FITC-conjugated goat anti-rabbit antibody (1:100, Sigma F0382), and three washes were carried out as indicated above. Primary and secondary antibodies were diluted in KCM buffer and 100µl of them were used per each slide. After washing, cells were post-fixed using 2% (v/v) formaldehyde in PBS for 10 minutes. Subsequently, FISH was performed and slides were analysed as described in section 2.20.

2.22 Fibre-FISH

Cells were harvested by trypsinisation and resuspended in PBS to a final concentration of $1x10^6$ cells/ml. To prepare extended DNA fibres, 10µl of cell suspension was spread near the edge of the slide and dried at 45-50ºC. Subsequently, 150µl of lysis solution [0.05M NaOH, 30% (v/v) ethanol] was added on the slide and allowed to drain to lyse the cells down the slide. The resulting fibres were fixed onto the slide with 200µl of methanol. The slide was then air-dried and dehydrated by successive washes for 2 minutes each in 70%, 90% and 100% ethanol. For FISH, the fibres were denatured in 70% (v/v) formamide and 2x SSC buffer at 72°C for 2 minutes. FISH procedure was performed and slides were analysed as described in section 2.20. Fibre-FISH procedures and analysis were performed by Dr. Daniela Moralli at the Wellcome Trust Centre for Human Genetics (WTCHG) at the University of Oxford.

2.23 Multiplex FISH (mFISH)

Multiplex fluorescence *in situ* hybridisation (mFISH) was performed on hESc and hiPS cells by Dr. Mohammed Yusuf and Dr. Daniela Moralli, and analysed by Dr Emanuela Volpi at the Molecular Cytogenetics and Microscopy core facility, Wellcome Trust Centre for Human Genetics (WTCHG) at the University of Oxford. Fixed metaphase chromosomes were prepared as described in section 2.18 and hybridised using 24XCyte Human Multicolor FISH (mFISH) Probe Kit (MetaSystems) using the manufacturer's recommendations.

2.24 Immunofluorescence Staining

Immunofluorescence staining procedures were performed on cells cultured in 24-well dishes. The cells were washed thrice with PBS and then fixed with 4% (v/v) formaldehyde in PBS for 20 minutes at room temperature, followed by three washings with PBS. For intracellular and nuclear bound proteins, the cells were permeabilised at room temperature with 0.5% (v/v) Triton X-100 in PBS for 15 minutes prior to blocking non-specific binding with 3% (w/v) BSA and 1% (v/v) Triton X-100 in PBS for one hour at RT. Cells were incubated with primary antibodies diluted in blocking buffer at 4°C overnight as mentioned in Table 2-6. After overnight incubation, the cells were washed thrice with PBS followed by incubation with secondary antibodies (Table 2-6) diluted in blocking buffer. This incubation was carried out at room temperature for one hour in the dark. After three washings with PBS, the cells nuclei were counterstained with 1ng/ml DAPI. Stained cells were visualised with Nikon Eclipse TE2000U inverted microscope with a Nikon Plan Fluor 4×/0.13 phL DL ∞/1.2 WD 16.4 objective and CFI 10×/22 eyepiece lens. Images were captured with Hamamatsu C4742-95 Orca low-noise B&W CCD camera using NIS-Elements software version 2.34.

Table 2-6 Primary and secondary antibodies used in immunofluorescence analysis

Antibodies	Company	Catalogue No.	Dilution
Primary Antibodies			
Anti-OCT4 (Goat)	Abcam	ab27985	1:250
Anti-SOX2 (Rabbit)	Abcam	ab69893	1:250
Anti-NANOG (Rabbit)	Abcam	ab21624	1:250
Anti-TRA-1-60 (Mouse)	Chemicon	MAB4360	1:250
Anti-β-III tubulin (Mouse)	R&D Systems	MAB1195	1:500
Secondary Antibodies			
Alexa Fluor 594 rabbit anti-goat	Molecular Probes	A-11080	1:200
Texas Red goat anti-rabbit	Molecular Probes	T-2767	1:200
Rhodamine donkey anti-mouse	Jackson Immunoresearch	715-025-140	1:200
TRITC goat anti-mouse	Sigma	T5393	1:100

2.25 Fluorescence-Activated Cell Sorting (FACS)

Trypsinised hESc were fixed in 4% (v/v) formaldehyde in PBS. Samples were processed through a 4-colour FACSCalibur^{TM} flow cytometer (BDIS, San Jose, CA) using the settings: FL1 449, FL2 499, FL3 500. At least 10,000 events were acquired for each sample preparation. The data were acquired using CellQuest Pro software (BDIS) and analysed using FlowJo software (version 7.6; Tree Star, Inc., Ashland, OR). The percentage of *GFP*-expressing cells was calculated by setting the GFP gate at 0.5% for the negative controls. FACS procedures and data analysis were performed with the assistance of Dr. Sally Cowley, Sir William Dunn School of Pathology at the University of Oxford.

Chapter 3. HAC Development in Human ES Cells

3.1 Introduction

Human embryonic stem cells (hESc) represent a valuable research tool for studying early human embryonic development as well as having potential for various therapeutic purposes (Friedrich Ben-Nun and Benvenisty, 2006; Moon et al., 2006; Thomson et al., 1998). The therapeutic potential of hESc can be further exploited by their safe and efficient *in vitro* genetic manipulation. Most of the genetic manipulation methods that have been employed so far such as chemical transfection, electroporation, Nucleofection and viral delivery systems are either inefficient at transgene delivery or cause genomic □□□□□□□□□□□□□棚靨?□□□升哂Ĉ□脈棚▪?□□□升哂Ĉ□▪味□□絨晰□□□□□□▪▪□□▪穗□□▪鬮▪Ĉ□▪轅□▪轅▪Ĉ□▪轅□□脈棚甬?(Strulovici et al., 2007).

An alternative vector type for genetic manipulation of hESc, which is described and used in this report, is represented by human artificial chromosomes (HAC), which are assembled genetic molecules that replicate the behaviour of natural chromosomes. HAC not only provide an efficient mechanism for transgene delivery but also have the advantage that they remain extra-chromosomal and do not integrate within the cellular genome. Whilst HAC have been successfully developed and used as viable gene expression vectors (Grimes et al., 2001; Ikeno et al., 2002; Mejia et al., 2002), delivery methods such as lipofection and micro-cell mediated chromosome transfer (MMCT) used for introducing HAC constructs into target cells are inefficient (Mejia et al., 2001). Our initial experiments for developing HAC in hESc were also hindered by the low efficiencies of HAC vector delivery by chemical transfection (Mandegar, 2011). In 2006, our laboratory reported the use of the Herpes Simplex Virus (HSV)-1 amplicon system to

successfully transduce HAC vectors into several human cell lines at high efficiencies (Moralli et al., 2006). When applied to the hESc, the HSV-1 mediated transduction was found to be up to 10^3 times more efficient at delivering large DNA than the chemical transfection method (Mandegar, 2011).

The study presented in this chapter aims to determine if HAC delivered via the HSV-1 amplicon system, which allows delivery of large genomic fragments (up to 150kb) into cells, are suitable vectors for long-term gene expression studies in hESc. With the use of the HSV-1 amplicon system, a 55kb HAC vector was delivered into the HUES10 hES cell line and functional HAC were generated, which were mitotically stable for over 90 days. Analysis of the HAC-containing clones did not reveal any integration of the HAC vector into the genome. Most significantly, the cells retained their pluripotency and differentiation capabilities.

3.2 Specific Methods

3.2.1 HSV-1 HAC vector

To generate HAC in hESc, the HSV-1 HAC vector pHSV17α40Neo containing 40kb of 17α DNA was used (Figure 3-1). pHSV17α40Neo was found to be efficient in HAC formation in the HT1080 cells and HUES2 cells in parallel experiments conducted by Mohammad A. Mandegar (Mandegar, 2011; Mandegar et al., 2011). pHSV17α40Neo was assembled by retrofitting a BAC containing 40kb of alpha satellite DNA from chromosome 17 (17α) with the HSV-1 vector pHGNeo4 (Figure 2-1; Table 2-2), which carries the HSV-1 elements (*OriS* and *pac*), the *GFP* reporter gene, and the G418 resistance gene (*Neo*) (Mandegar, 2011; Mandegar et al., 2011). The BAC containing

40kb of alpha satellite DNA from chromosome 17 (17α) was produced through spontaneous recombination and insert deletion of a large portion of hBAC495J24 (Kim et al., 1996), a highly proficient HAC forming vector based on pBeloBAC11-backbone and containing approximately 220kb of genomic 17α satellite DNA (Mejia et al., 2002).

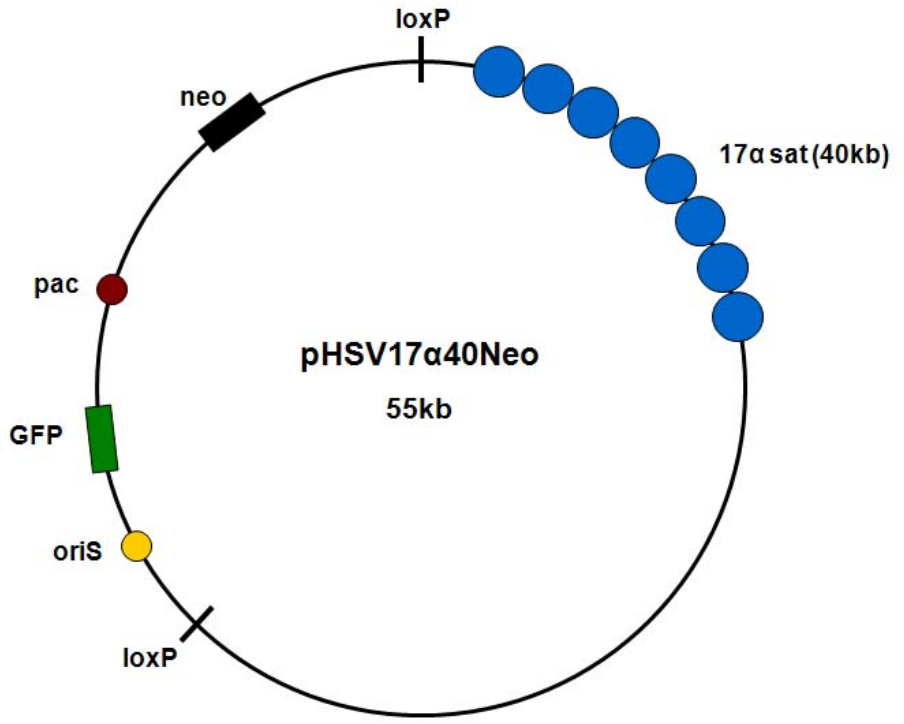

Figure 3-1 Vector map of pHSV17α40Neo
HSV-1 HAC vector pHSV17α40Neo was constructed by Mohammad A. Mandegar and was derived from hBAC495J24 through spontaneous recombination. It contains ~40kb of 17α satellite DNA, HSV-1 origin of replication (*OriS*), HSV-1 packaging site (*pac*), Neomycin resistance gene (*Neo*), *GFP* reporter gene and two *loxP* sites. The vector elements are not drawn to scale.

3.2.2 PCR-screening for HAC vector

The pHSV17α40Neo clones were screened by PCR using primers Neo 8F/8R for neomycin resistance gene and also by nested PCR using the primers pBelo 1F/9R (1st round) and pBelo 1F/6R (2nd round) to amplify regions of pBeloBAC11 vector. Primers are listed in Table 3-1.

Table 3-1 Primer sets used for PCR amplification

Genes/Region	Symbol	Sequence (5'- 3')
pBeloBAC11 vector	pBelo 1F	CGCTGGAGAATAGGTGAAGC
	pBelo 9R	CCGGATCTGACCTTTACCAA
pBeloBAC11 vector	pBelo 1F	CGCTGGAGAATAGGTGAAGC
	pBelo 6R	AATCAGCACATCAGCAGCAC
Neomycin	Neo 8F	CGTTGGCTACCCGTGATATT
	Neo 8R	GCCCAGTCATAGCCGAATAG

3.3 Results

3.3.1 M-FISH and CGH Microarray Analysis of HUES10 Cells

The prolonged culture of human embryonic stem cells can result in acquisition of chromosomal abnormalities. Some of these abnormalities are recurrent, such as gain of extra chromosome 12 and chromosome 17, and appear to provide a selective advantage to the cells where these events occur (Draper et al., 2004; Rosler et al., 2004). With the aim of generating HAC in hES cells exhibiting a normal karyotype, multiplex fluorescence in situ hybridisation (M-FISH) was performed on hES cells HUES10 (passage 29) by the Molecular Cytogenetics and Microscopy core facility at Wellcome Trust Centre for Human Genetics, using the 24XCyte M-FISH probe kit (MetaSystems). Thirty metaphases from these HUES10 cells were analysed and no large structural or numerical aberrations were detected in these cells (Figure 3-2). To further rule out the presence of detectable submicroscopic alterations, the HUES10 cells karyotype was also analysed by whole genome cytogenetic microarrays (Affymetrix Cytogenetics Whole-Genome 2.7M Arrays) by the Molecular Cytogenetics and Microscopy core facility. The karyotype of

the HUES10 cells was found to be completely normal (confidence 90%, 200 markers) and was thus used for HAC-generation experiments.

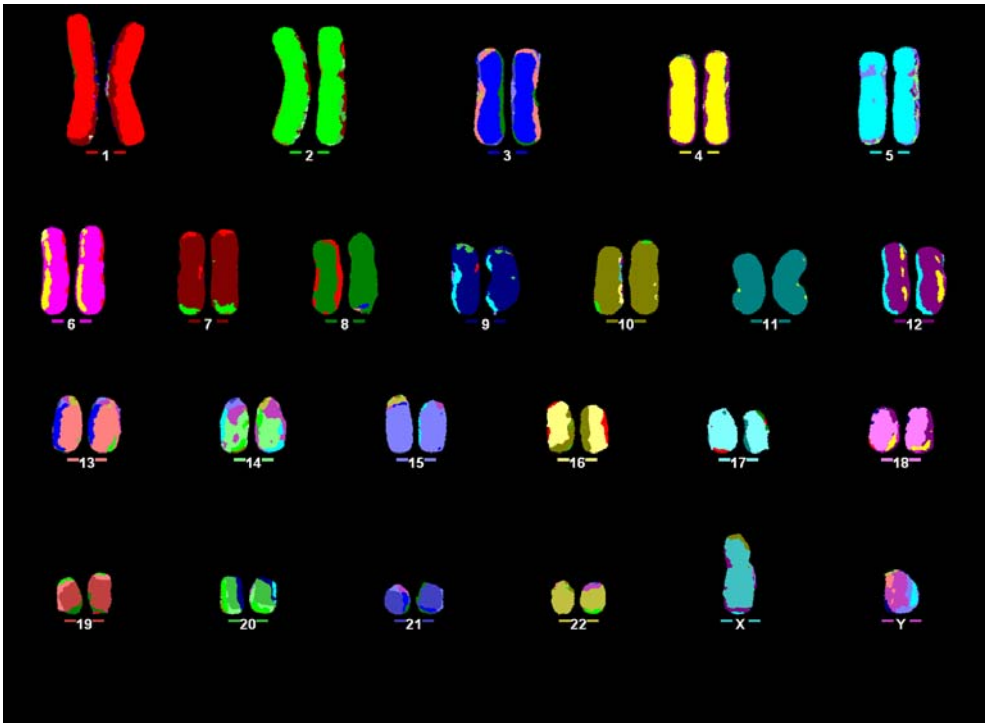

Figure 3-2 M-FISH analysis of HUES10 cells
Cells had a normal karyotype (46, XY) and no structural or numerical aberrations were observed. This image was generated by Molecular Cytogenetics and Microscopy core facility at Wellcome Trust Centre for Human Genetics.

3.3.2 Titration of G418 for Selection of HUES10 Cells

To determine the ideal concentration of G418 selection on HUES10 cells, a range of G418 concentrations was employed to test cell death. The HUES10 cells (passage 20) were seeded on Matrigel-coated wells in 24-well dish in mTeSR1 medium, at two different seeding densities (2.5×10^5 and 5×10^5 per well) and the G418 selection was applied after 24 hours. Table 3-2 lists the range of G418 concentrations that were used and their effects on cell death. The selection of the cells was continued for up to six days and cell death on each day was estimated by visual examination under the microscope.

Table 3-2 Titration of G418 on HUES10 cells

G418 concentration	Seeded cell density	Time to reach 99% Cell Death
0μg/ml	2.5×10^5	No sign of cell death
	5.0×10^5	No sign of cell death
10μg/ml	2.5×10^5	No sign of complete cell death
	5.0×10^5	No sign of complete cell death
25μg/ml	2.5×10^5	6 days
	5.0×10^5	No sign of complete cell death
50μg/ml	2.5×10^5	3 days
	5.0×10^5	4 days
75μg/ml	2.5×10^5	3 days
	5.0×10^5	3 days
100μg/ml	2.5×10^5	2 days
	5.0×10^5	3 days

The observations of cell death were based on the average cell death of two independently treated wells.

These results provided strong evidence that the selection of HUES10 was highly density-dependent as various concentrations of G418, which were sufficient to induce significant cell death at low plating density, had almost no effect on high cell density. For example, 50μg/ml G418 resulted in complete cell death of 2.5×10^5 cells within 3 days. However, using the same concentration of G418, it took 4 days to result in complete cell death of 5.0×10^5 cells. Similarly, 25μg/ml G418 proved to be sufficient to induce complete cell death for 2.5×10^5 cell within 6 days but was found to be inadequate for 5×10^5 cells. Cell-cell interactions and support are possibly the main causes of density-dependent cell resistance in response to selection. Moreover, hESc become tightly packed at higher cell density and, therefore, the surface area of cells exposed to the selecting agent is reduced (Braam et al., 2008).

3.3.3 Transduction of HUES10 with HSV-1 Vectors

In preliminary experiments, the optimal conditions for HSV-1 amplicon transduction in HUES10 cells were determined. Cells were seeded at two densities of $2.5x10^5$ and $5x10^5$ per 24-well dish (Matrigel-coated) in mTeSR-1 medium and were transduced with the pHGNeo4 vector at three different MOIs (1, 2 and 5) followed by spinoculation under low centrifugal force (750 g for 45 minutes) (Scanlan et al., 2005). Through qualitative observations (data not shown), it was established that efficient transduction was achieved with $2.5x10^5$ cells per 24-well dish transduced at MOI 2. Higher cell density ($5x10^5$ cells) led to reduced transduction efficiency possibly owing to cell confluency. Moreover, increasing the traducing units (MOI 5) only contributed to a slight improvement in transduction efficiency as the cells probably reached transduction saturation.

The HSV-1 HAC vector pHSV17α40Neo and the control plasmids pHGNeo4 were transduced into $2.5x10^5$ HUES10 cells at MOI 2 with spinoculation. The efficiency of transduction was determined after 24 hours of transduction by monitoring the percentage of *GFP* expressing cells in comparison to the total number of cells. Figure 3-3 shows *GFP* expression in the HUES10 cells transduced with pHSV17α40Neo and pHGNeo4 vectors. Transduction efficiencies are summarised in Table 3-3. The HSV-1 transduction efficiencies in HUES10 cells (27%) appeared to be relatively higher than that of HT1080 cells (19%), as observed in previous experiments using the same experimental conditions (Mandegar et al., 2011).

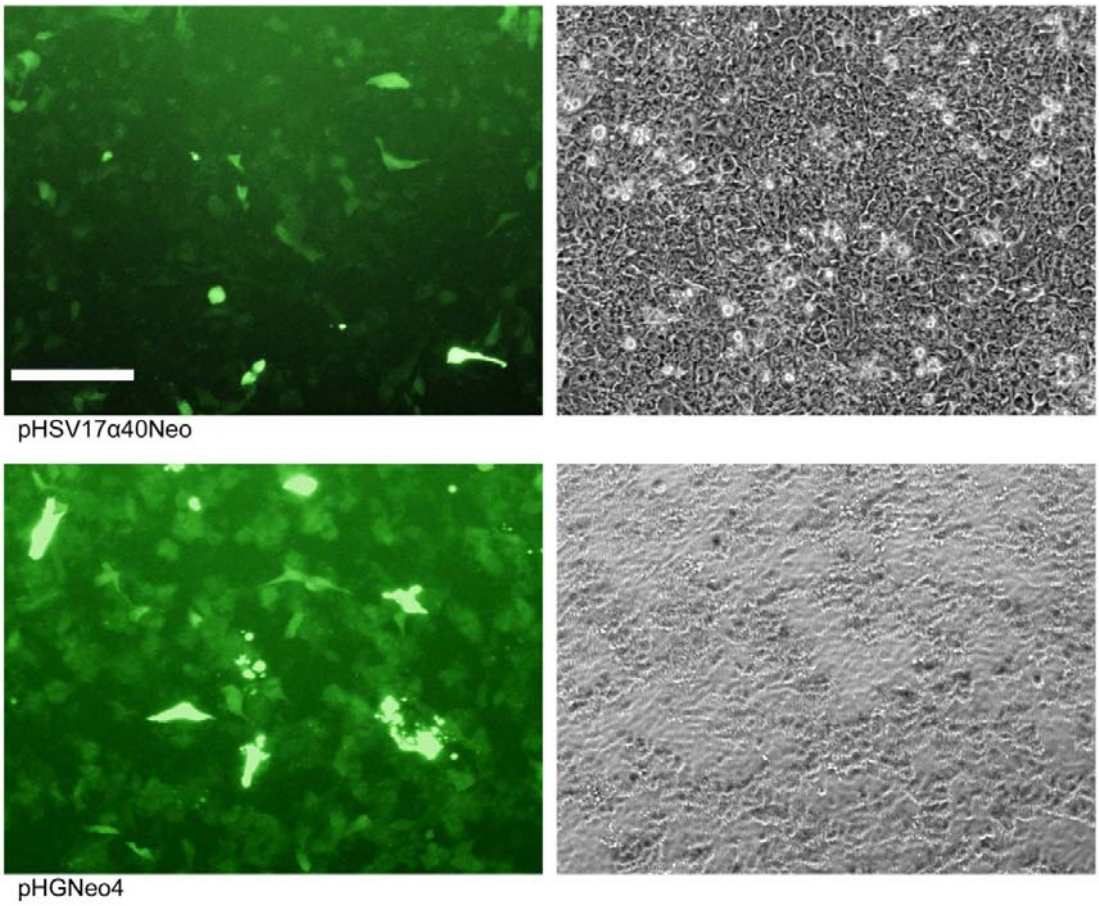

Figure 3-3 *GFP* expression of HUES10 cells after transduction with pHSV17α40Neo and pHGNeo4 amplicons
HUES10 cells were transduced with HSV-1 vectors pHSV17α40Neo and pHGNeo4 at MOI 2. The left panels show *GFP* expression of the transduced cells 24 hours post-transduction. The right panels show the corresponding phase contrast images. Scale bar = 200μm.

Table 3-3: Transduction efficiencies of HSV-1 vector on HUES10 cells

Vector	Transduction Efficiency (% + SE)
pHSV17α40Neo	26 ± 0.70
pHGNeo4	25 ± 0.78

The presented transduction efficiency is based on the average transduction efficiency of two independently transduced wells. SE denotes standard error of the mean.

The transduction of HUES10 cells with HSV-1 amplicons carrying the control and HAC vectors did not led to any immediate visible side effects on cell viability, morphology and growth patterns. Figure 3-4 compares the morphology of untreated HUES10 cells and

HSV-1 transduced HUES10 cells after 24 hours of transduction. This indicated that neither the HSV-1 amplicon, nor other factors present in the amplicon preparation cause gross morphological changes in human embryonic stem cells. Gradual silencing of *GFP* expression was observed 48 hours post-transduction and continued until day six. After six days of transduction, fluorescence microscopy did not reveal any GFP, as expected due to the silencing of HSV-1 IE4/5 promoter controlling *GFP* expression.

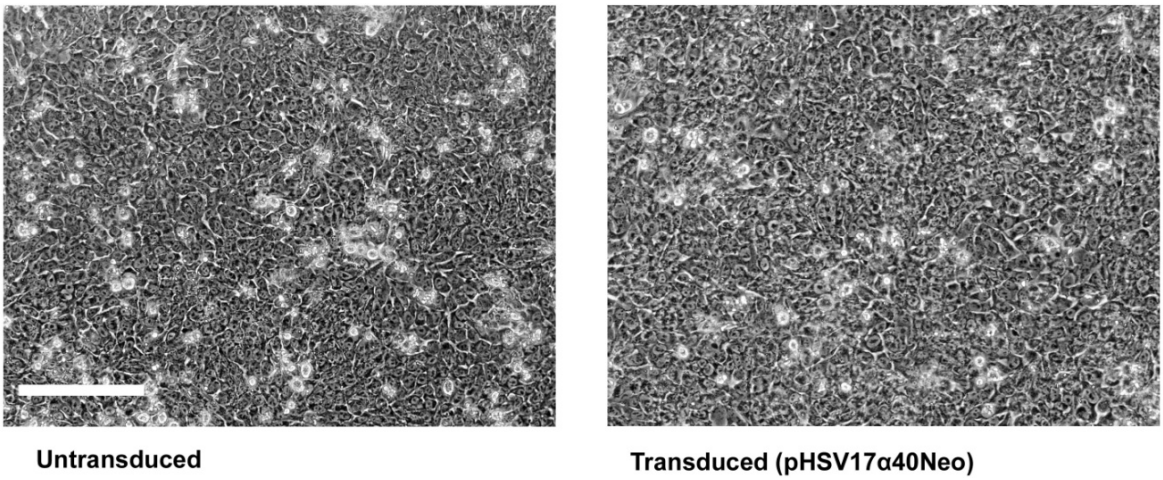

Figure 3-4 Morphology of untransduced and transduced HUES10 cells
No morphological alterations were observed between the HSV-1 transduced and untreated HUES10 cells. Scale bar = 200μm.

3.3.4 Generation of Stable Clones in HUES10 Cells

In the initial attempts, HUES10 cells (grown on Matrigel-coated wells in mTeSR-1 medium) transduced with HSV-1 vectors pHSV17α40Neo and pHGNeo4 were treated with 25μg/ml G418 48 hours post-transduction and the selection was continued for six days. However, all the transduced and non-transduced (negative control) cells died and did not generate any stable clones after additional culture in the absence of selection for 5-6 days. These results were attributed to poor survival of hESc cells as single cells. As hESc require connections with adjacent cells in order to maintain clonogenicity and self-renewal (Li et al., 2010a; Xu et al., 2010), death of the adjacent untransduced cells might

have resulted in loss of those connections and ultimately lead to the death of transduced cells. This phenomenon and growth pattern is generally not observed for hESc plated on the inactivated MEF (iMEF) feeder cells as hESc can establish contacts with iMEF. Moreover, iMEF render extra cell support including growth- and signalling factors for hESc thus reducing their requirement for additional contact with neighbouring hESc. Therefore for subsequent experiments, following transduction the transduced and control cells were plated onto mitotically inactivated SNL 76/7 (G418 resistant MEFs; Table 2-3) feeder cells before subjecting them to drug selection.

Since the experiments summarised in section 3.3.2 indicated that the efficacy of the selection is highly dependent on the cell seeding density, different concentrations of G418 were applied in parallel. Three days after the transduction of $2.5x10^5$ HUES10 cells with pHSV17α40Neo and pHGNeo4, cells were transferred to SNL76/7 feeder cells. Two days later, G418 selection was applied at concentrations of 50µg/ml, 75µg/ml and 100µg/ml and was continued for 7-9 days. Figure 3-5 summarises the effect of the different selective conditions used. Non-transduced (control) cells were completely eliminated by day 7 at 100µg/ml G418. By comparison at the same concentration of G418, individual clones from pHSV17α40Neo and pHGNeo4 transduced cells were visible on day 10 (Figure 3-6). Selection was subsequently removed to allow the clones to expand. Non-transduced (control) cells were not entirely eliminated at the 50µg/ml and 75µg/ml G418, therefore these cells along with transduced cells selected at 50µg/ml, 75µg/ml G418 were discarded, to avoid expanding and analysing possible false positive clones.

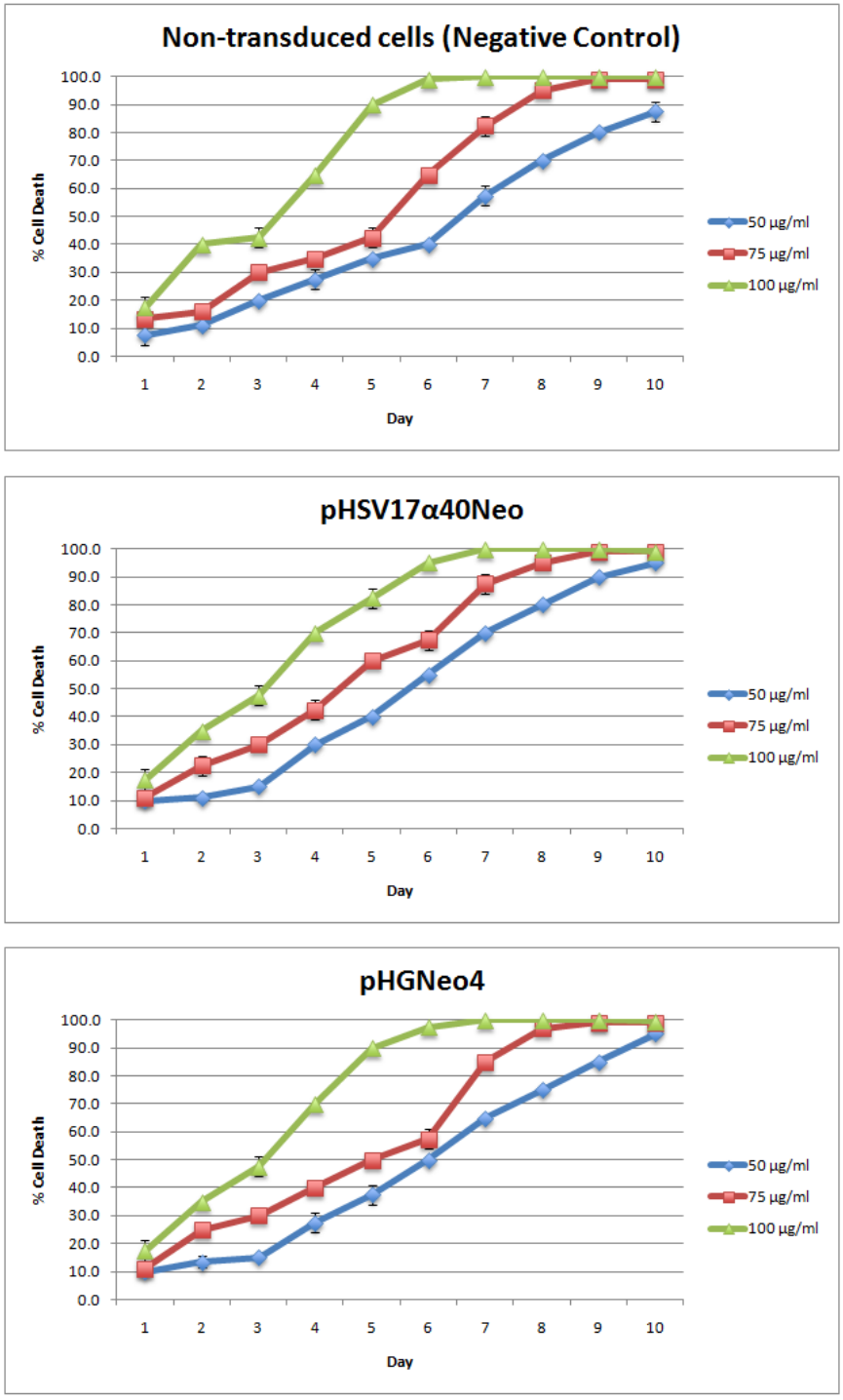

Figure 3-5 Comparison of cell death at different concentrations of G418
Following transduction, the HUES10 cells were selected at different concentrations of G418 and the percent of cell death, as judged subjectively (approximate) by microscopic analysis, was recorded each day. Day 1 corresponds to 24 hours after addition of G418. Non-transduced cells were completely killed by 7th day at 100µg/mL. For the cells transduced with pHSV17α40Neo and pHGNeo4 vectors only a few (about 1%) cells survived and subsequently formed clones. The results are based on duplicates and error bars denote standard error of the mean.

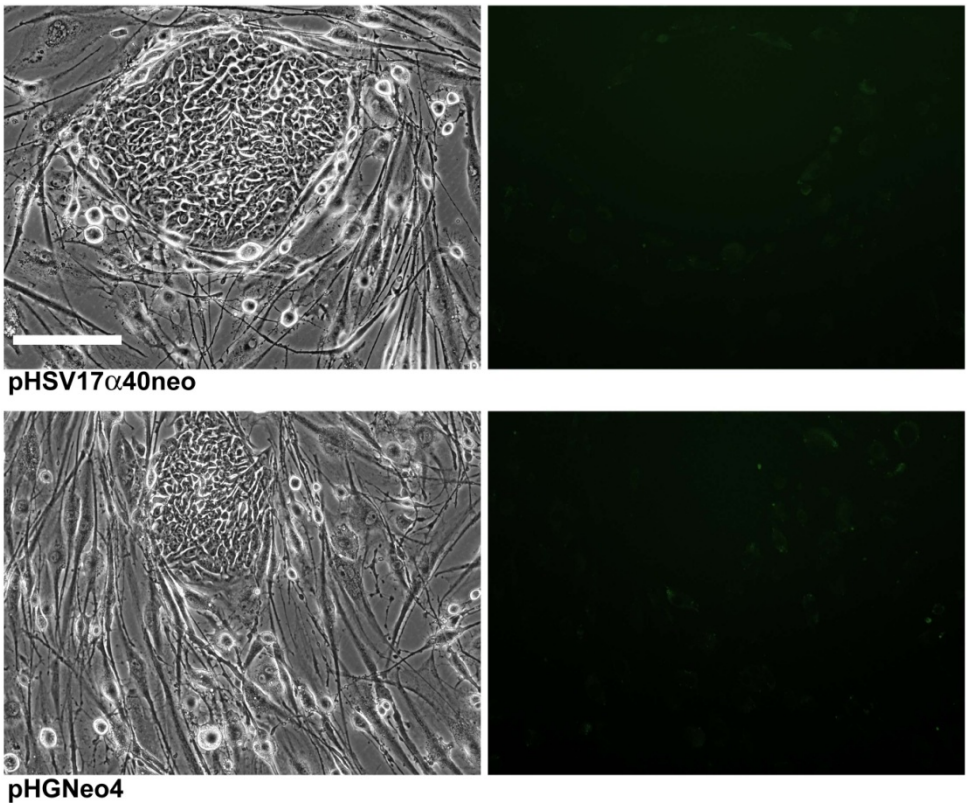

Figure 3-6 Stable HUES10 clones after G418 selection
Phase contrast images of HUES10 stable clones generated 10 days after continuous G418 selection are represented in left panel. The clones didn't exhibit visible *GFP* expression under the microscope. GFP images of the same field as left panel are shown in right panel. Scale bar = 200μm.

The total number of clones obtained from HUES10 cells transduced with HSV-1 vectors pHSV17α40Neo and pHGNeo4 were 191 and 48 respectively. Development of these clones in the presence G418 drug selection suggested adequate expression of neomycin resistance gene (SV40 promoter) from the clones. However, none of the clones exhibited visible *GFP* expression under the microscope (Figure 3-6) indicating the silencing of *GFP* gene, which is expressed under the HSV-1 IE4/5 promoter.

Twenty four stable clones were picked from pHSV17α40Neo transduced HUES10 cells and transferred to a 24-well dish in individual wells, on SNL76/7 feeder cells. The clones were maintained for 6-7 days until they grew to larger size. Cells were expanded into

SNL76/7-coated 24-well dishes and subsequently expanded on Matrigel-coated dishes in mTeSR-1 medium for further analysis.

3.3.5 PCR and FISH Analysis of Stable Clones

The pHSV17α40Neo clones were initially screened by PCR using primers for the neomycin gene and pBeloBAC11 sequences (Table 3-1) as described in section 3.2.2. The PCR screening of clones was carried out to remove false-positive clones as the cells were maintained and expanded without G418 selection before they were sufficiently grown for chromosome harvesting and subsequent HAC analysis (about 20-25 days). Out of the twenty four clones analysed, nine were found to be positive with PCR analysis (data not shown).

To identify HAC containing cells, metaphase spreads were then prepared from PCR positive clones and analysed by two colour FISH with a vector probe (HAC-specific) and 17α DNA probe. For each clone, up to 20 metaphases were scored, and the number of HAC containing cells was recorded. The results are presented in Table 3-4. Six of the PCR positive pHSV17α40Neo clones, SK40.10.1, SK40.10.3, SK40.10.17, SK40.10.19, SK40.10.22 and SK40.10.24, were identified to contain HAC in 12% to 50% of cells. Although the three clones SK40.10.4, SK40.10.5 and SK40.10.18 were PCR-positive, FISH analysis did not reveal HAC, possibly due to the presence of very small HAC undetectable through FISH or very low frequency of HAC. No integrations were found in any of these clones as shown in Figure 3-7. Clone SK40.10.19, containing HAC in 50% of the analysed metaphases was selected for further studies.

Table 3-4 PCR and FISH analysis of HUES10 clones

Clone	Vector	PCR Result	FISH result	HAC percentage
SK40.10.1	pHSV17α40Neo	Positive	HAC	40%
SK40.10.2	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.3	pHSV17α40Neo	Positive	HAC	40%
SK40.10.4	pHSV17α40Neo	Positive	Negative	N/A
SK40.10.5	pHSV17α40Neo	Positive	Negative	N/A
SK40.10.6	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.7	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.8	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.9	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.10	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.11	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.12	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.13	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.14	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.15	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.16	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.17	pHSV17α40Neo	Positive	HAC	20%
SK40.10.18	pHSV17α40Neo	Positive	Negative	N/A
SK40.10.19	pHSV17α40Neo	Positive	HAC	50%
SK40.10.20	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.21	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.22	pHSV17α40Neo	Positive	HAC	25%
SK40.10.23	pHSV17α40Neo	Negative	Negative	N/A
SK40.10.24	pHSV17α40Neo	Positive	HAC	12%

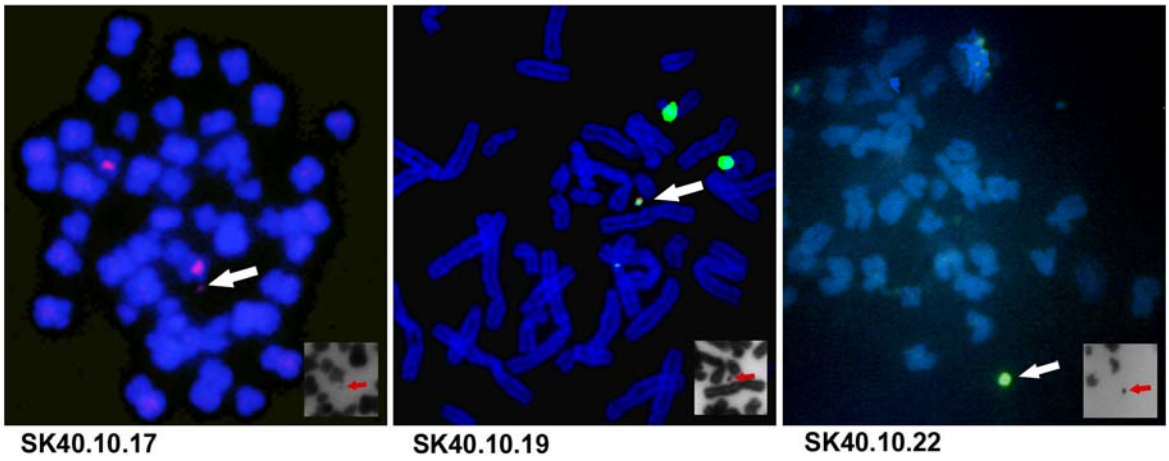

Figure 3-7 Analysis of HAC in HUES10 clones through FISH
FISH analysis of SK40.10.17, SK40.10.19 and SK40.10.22 clones indicated the presence of HAC. For SK40.10.17, 17α DNA probe, which also recognised the endogenous chromosomes 17, is labelled in red. For SK40.10.19 and SK40.10.22 clones, 17α DNA probe is labelled in green. The chromosomes are counterstained by DAPI, blue. The HAC is identified by a white arrow. The insets show DAPI staining only, as a black and white image where the HAC is identified by a red arrow.

3.3.6 CENP-C Immuno-FISH of SK40.10.19 Clone

Immuno-FISH with vector probes (for pBeloBAC11) coupled to centromere protein C (CENP-C) staining was used to determine if the HAC present in the SK40.10.19 clone possesses an active centromere. A positive CENP-C signal was identified on the HAC, at a similar intensity of those observed on endogenous chromosomes (Figure 3-8), and indicated that the HAC in HUES10 cells assembled an active centromere.

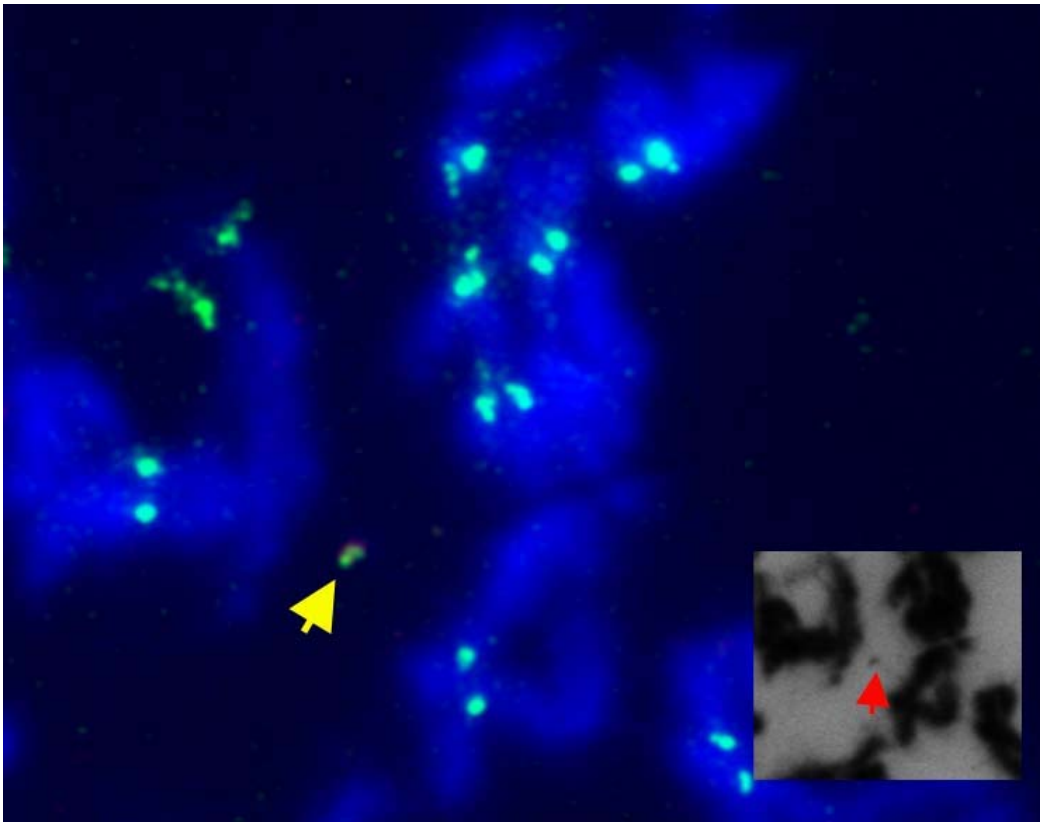

Figure 3-8 CENP-C staining of SK40.10.19 clone
Immunostaining with anti CENP-C antibody (green signal), coupled to FISH with the vector probe (red signal) confirms the presence of an active centromere on the HAC (identified by the yellow arrow). The inset shows DAPI staining only, as a black and white image and the HAC is identified by a red arrow.

3.3.7 Characterisation of HAC Structure

The structure of the HAC in SK40.10.19 clone was analysed by FISH on extended chromatin fibres (fibre-FISH experiment and analysis was carried out by Dr Daniela Moralli). As seen in the previous studies, (Grimes et al., 2002; Mejia et al., 2001), the HAC contained alternating arrangement of 17α-satellite DNA and vector sequences (Figure 3-9) whereas chromatin fibres of the endogenous chromosome 17 were only labelled with 17α-satellite DNA probe. These results indicated that HAC were composed of multimers of input DNA as a result of its amplification during HAC formation.

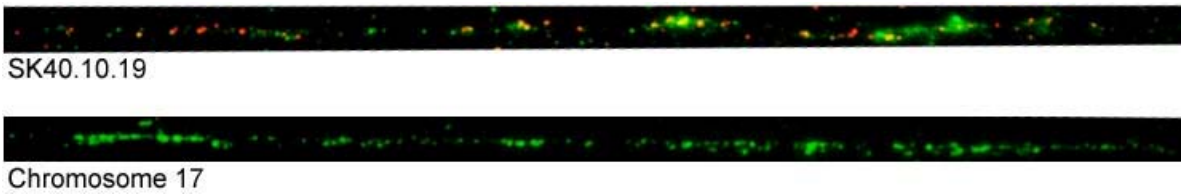

Figure 3-9 Fibre-FISH analysis of SK40.10.19 clone
The structure of the HAC in the SK40.10.19 clone was analysed by FISH on extended chromatin fibres with a 17α-satellite DNA probe (green signal) and backbone vector probe for pBeloBAC11 and pHGNeo4 (red signal). Alternating pattern of red and green signal in SK40.10.19 indicates that the HAC is composed of multimers of 17α-satellite and backbone vector as a result of its amplification. The bottom panel shows the endogenous chromosome 17, in which the fibre is only labelled with 17α-satellite DNA probe. Fibre-FISH analysis was performed by Dr. Daniela Moralli.

3.3.8 HAC Stability

To determine the mitotic stability of the HAC, the SK40.10.19 clone was cultured for 90 days (corresponding to ~30 passages) in the absence of G418 drug selection. At the end of this period, the HAC frequency was assessed by FISH by scoring 20-25 metaphase chromosome spreads. The frequency of HAC did not change significantly during this period with a daily loss rate of 0.24% indicating that the established HAC correctly segregated during the mitotic division along with the endogenous chromosomes. Assessment of daily loss was carried out by Dr. Daniela Moralli using the following equation:

$$R = 1 - \sqrt[n]{N_n / N_0}$$

where n is total number of days (90 in this case) the cells were cultured in the absence of G418 selection, N_0 the percentage of metaphase spreads containing HAC in the cells cultured in the presence of G418 selection, N_n the percentage of metaphase spreads containing HAC after n days of culture in the absence of G418 selection and R denotes the daily rate of loss).

3.3.9 Gene Expression Analysis

Although the SK40.10.19 clone did not exhibit visible *GFP* expression under the microscope, FACS analysis of the cells revealed that 33% of the cells in this clone expressed *GFP* (FACS procedures and data analysis was performed with the help of Dr. Sally Cowley), which was approximately in concordance with the estimated HAC frequency. These results suggested that the overall fluorescence of the cells was too low to be detected with the less-sensitive method of microscope analysis, compared to the FACS analysis. The low *GFP* expression could be attributed to its HSV-1 IE4/5 promoter, which is known to undergo gradual silencing after a short time period (Cuchet et al., 2007) as it requires the HSV-1 transactivator protein VP16 for full activity (Wysocka and Herr, 2003). The gradual silencing of *GFP* expression was also supported by real-time quantitative RT-PCR analysis on RNA, extracted from the cells grown either in the presence and absence of drug selection at initial time point and after forty days. The *GFP* gene expression was compared with that of a ubiquitously expressed reference gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). The analysis revealed that *GFP* expression was decreased by about 50% after 40 days regardless of whether the cells were grown in the presence or absence of G418 (Table 3-5).

Table 3-5 qRT-PCR analysis of *GFP* expression in SK40.10.19 cells

Selection condition	Day 0 (SD)	Day 40 (SD)
On Selection	1.00 (0.02)	0.40 (0.01)
Off selection	1.00 (0.02)	0.52 (0.02)

The relative *GFP* expression after forty days was measured in reference to the initial time point (day 0) and calculated using $2^{-\Delta\Delta Ct}$ method. The average C_T values were normalised to *GAPDH* internal control. The reactions were performed in triplicate. SD denotes standard deviation.

3.3.10 Karyotype Analysis

The karyotype of the SK40.10.19 clone was analysed for the presence of chromosomal aberrations, in particular trisomy 12 and trisomy 17, which are commonly associated with long term culture of hESc. The FISH with 17α-satellite DNA probe confirmed that in all analysed metaphases no extra copy of the entire chromosome 17 was present (Figure 3-7). Moreover, whole chromosome paint experiments indicated that the chromosome 12 number was normal in all the analysed cells as shown in Figure 3-10. Finally, the analysis of the DAPI banded karyotype indicated that the total chromosome number was 46+HAC, and did not reveal the presence of any gross chromosomal rearrangement. These data suggest that the HAC clone is karyotypically normal.

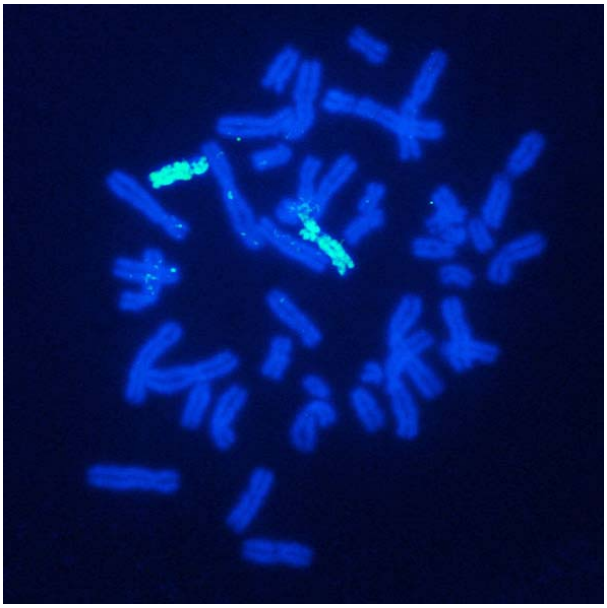

Figure 3-10 Chromosome 12 painting on SK40.10.19 clone
Metaphase spread hybridised with whole chromosome paint probe for chromosome 12 (green) indicates a single pair of chromosome 12. DAPI banding further suggests that total number of endogenous chromosome in the clone is 46.

3.3.11 Pluripotency of the HAC Clone

To verify that the HSV-1 transduction and HAC formation did not affect the pluripotency of HUES10 cells, the SK40.10.19 clone was investigated by immunofluorescence

staining for hESc-specific markers. Cells from this clone were stained with antibodies against the surface stem cell marker TRA-1-60, and the nuclear proteins OCT4 and NANOG along with the parental HUES10 cells (Figure 3-11). The result indicated that the cells from SK40.10.19 clone retained their pluripotency, as they were uniformly stained by the antibodies for hESc markers.

To further evaluate whether the pluripotency of the SK40.10.19 clone was sustained over time, expression of the three hESc markers (*OCT4*, *SOX2* and *NANOG*) was analysed by RT-PCR analysis on the RNA extracted from cells at definite time points of culture in the presence or absence of selection using the primer sets listed in Table 2-5. Figure 3-12 shows that expression of *OCT4*, *SOX2* and *NANOG* was retained at all the time points 0, 30, 60 and 90 days both in the presence and absence of G418 selection confirming that the prolonged culture of the HAC containing hESc didn't impair their pluripotency.

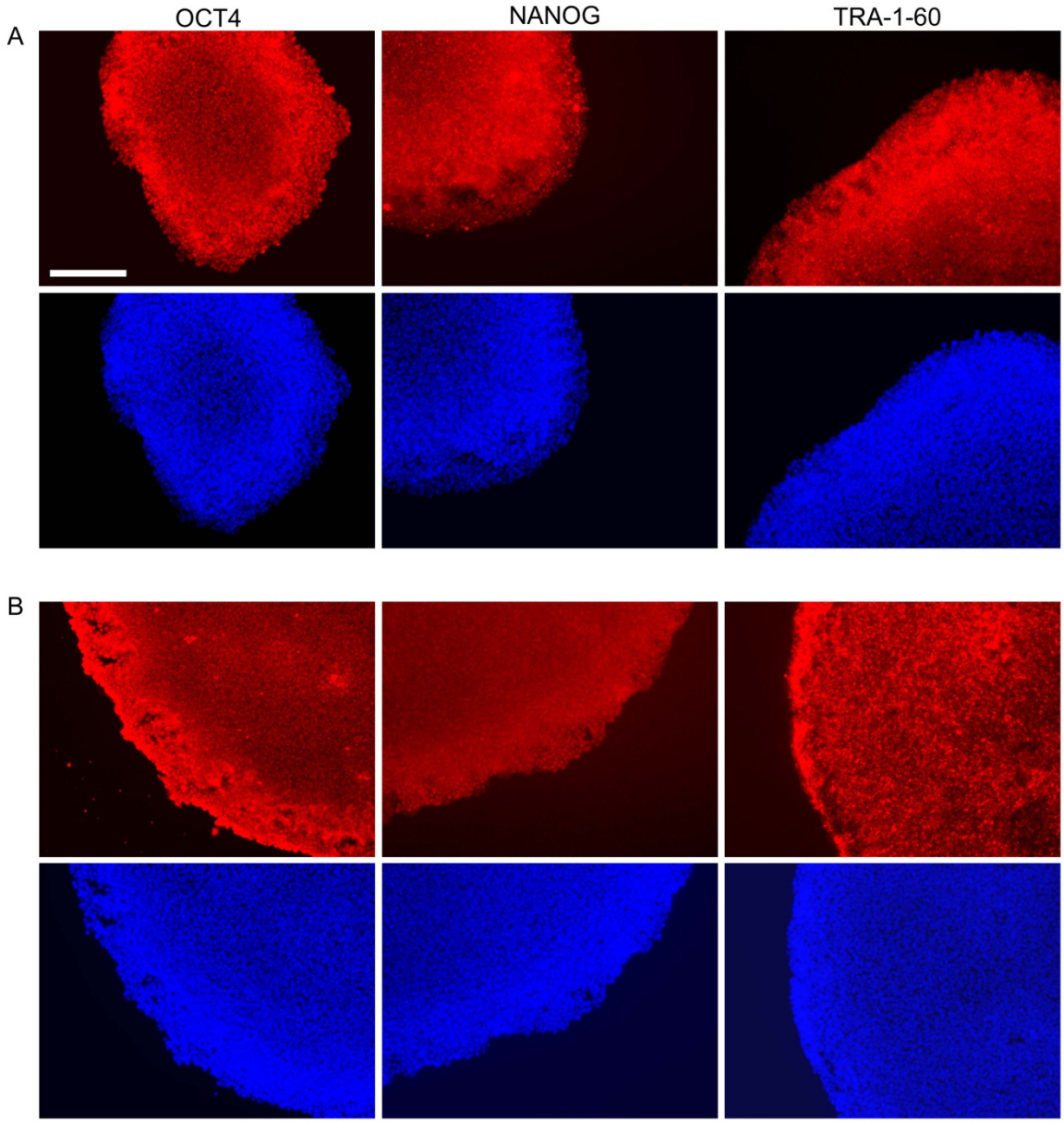

Figure 3-11 Immunofluorescence staining of SK40.10.19 clone and HUES10 cells for pluripotency markers

(A) Clone SK40.10.19 was positive for hESc pluripotency markers *OCT4*, *NANOG* and TRA-1-60 (red signal), suggesting that neither HSV transduction nor HAC establishment led to any loss of pluripotency. (B) HUES10 cells (passage 26) were used as control. Nuclei of the cells were counterstained with DAPI (blue signal). Scale bar = 200μm.

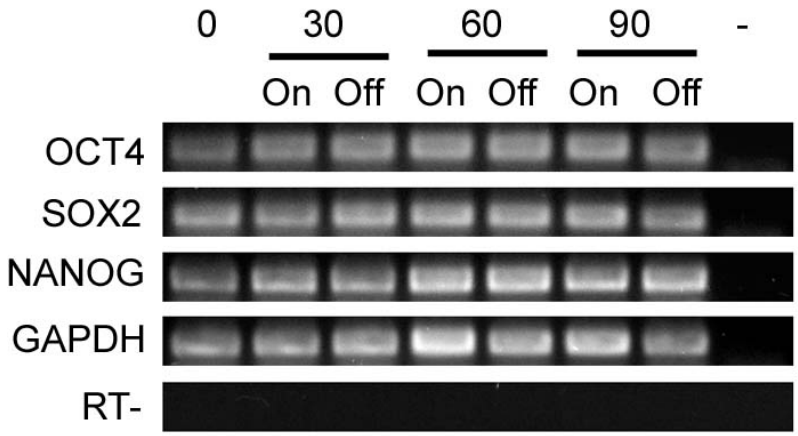

Figure 3-12 RT-PCR analysis of the expression of pluripotency genes in the SK 40.10.19 clone over 90 days in culture
RT-PCR analysis with primer sets specific for the transcripts of pluripotency markers *OCT4*, *SOX2* and *NANAOG* was carried out on the SK40.10.19 cells at time points 0, 30, 60 and 90 days, on and off G418 selection. The time points refer to the number of days after the clones were first transferred and propagated in mTeSR-1 medium on Matrigel-coated dishes. The *GAPDH* gene was used as an internal control. PCR of the RNA without reverse transcriptase (-RT) was done with the *GAPDH* primer set to ensure no amplification of genomic DNA.
(-) = no template control. On = on G418 selection. Off = off G418 selection.

3.3.12 Embryoid Body-Mediated Differentiation

hESc can recapitulate embryogenesis *in vitro* by aggregating into three dimensional (3D)-spherical structures known as embryoid bodies (EBs). To determine the differentiation potential of HAC containing hESc, the clone SK40.10.19 was subjected to forced aggregation in 96-well V-bottom dishes to generate EBs. The EBs were left in suspension culture for four days, after which they were transferred to gelatin-coated dishes to promote their attachment. After cultivation for an additional ten days, the attached EBs underwent differentiation. The deriving cells showed various types of morphologies such as those resembling epithelial cells (hexagonal appearance), hematopoietic precursors (cobblestone-like appearance) and neuronal cells. To confirm that the differentiated cells expressed markers specific for the three germ layers, RNA of the cells was subjected to RT-PCR analysis along with that of undifferentiated SK40.101.19 cells (Figure 3-13)

using the primers listed in the Table 2-4. Transcripts specific for endoderm (*α Feto-protein*, *HNF3α*, *α1 anti-trypsin*), mesoderm (*GATA-2*) and ectoderm (*CK-5*, *CK-14*, *high sulphur keratin*) markers were detected in the differentiated cells. Undifferentiated cells either did not expressed at all or exhibited lower levels of these markers. The overall analysis confirmed that the SK40.10.19 clone retained the ability of EB formation and germ layer differentiation.

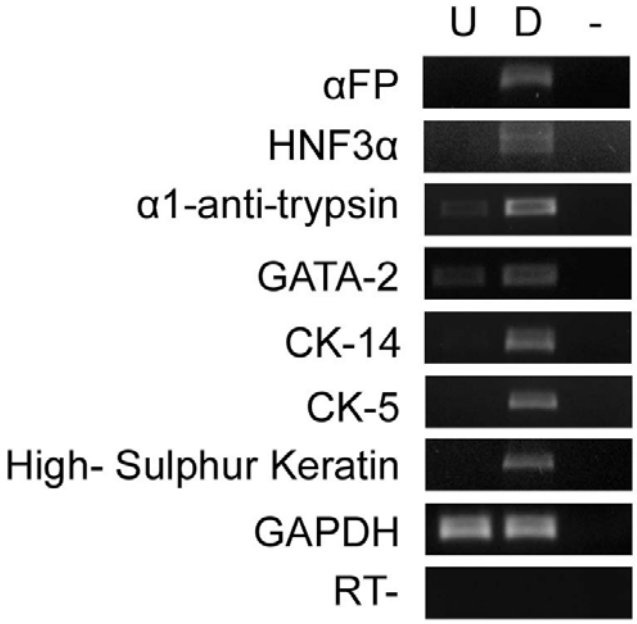

Figure 3-13 RT-PCR analysis of various differentiation markers specific for the three germ layers
SK40.10.19 differentiated and undifferentiated cells were analysed by RT-PCR using primers specific for endoderm (*αFP*, *HFN3α*, *α-1 anti-trypsin*), mesoderm (*GATA-2*) and ectoderm (*CK-14*, *CK-5*, *High-sulphur Keratin*) markers. The *GAPDH* gene was used as an internal control. PCR of the RNA without reverse transcriptase (-RT) was done with the *GAPDH* primer set to ensure no amplification of genomic DNA. (-) = no template control.

3.3.13 Directed Differentiation into Neuronal Cells

In vitro differentiation of hESc into neurons is a well established method and yields high levels of neuronal cells (Iacovitti et al., 2007; Itsykson et al., 2005). To examine if the SK40.10.19 clone could be directed specifically into the neuronal cells differentiation

pathway, a single step, chemically defined strategy, originally established by Iacovitti and colleagues (Iacovitti et al., 2007) was adapted. As control, the HUES10 parental cells were differentiated in the same way. The differentiation involved the treatment of EBs with noggin and fibronectin to repress bone morphogenetic protein (BMP) signalling followed by ectodermal induction using N2-medium (containing bFGF) to promote the expansion of neural progenitor rosettes. After the attachment of EBs to the dishes, gradual outward expansion of differentiated cells was observed followed by the formation of neuronal rosettes about 12 days after EBs attachment. The cells underwent further spreading and some neuronal structures (based on morphology of extensive network of axons and dendrites) were observed at around 20 days of EBs plating. At the end of the neuronal differentiation protocol (about twenty five days starting from EBs formation), the cells were stained with β-III-tubulin antibody to confirm the neuronal differentiation. As shown in Figure 3-14, the neuronal cells derived from the HAC containing clone SK40.10.19 were stained positive for β-III-tubulin with similar intensity as the parental HUES10 cells. The overall analysis indicates that HAC development did not affect the lineage-directed differentiation of hESc.

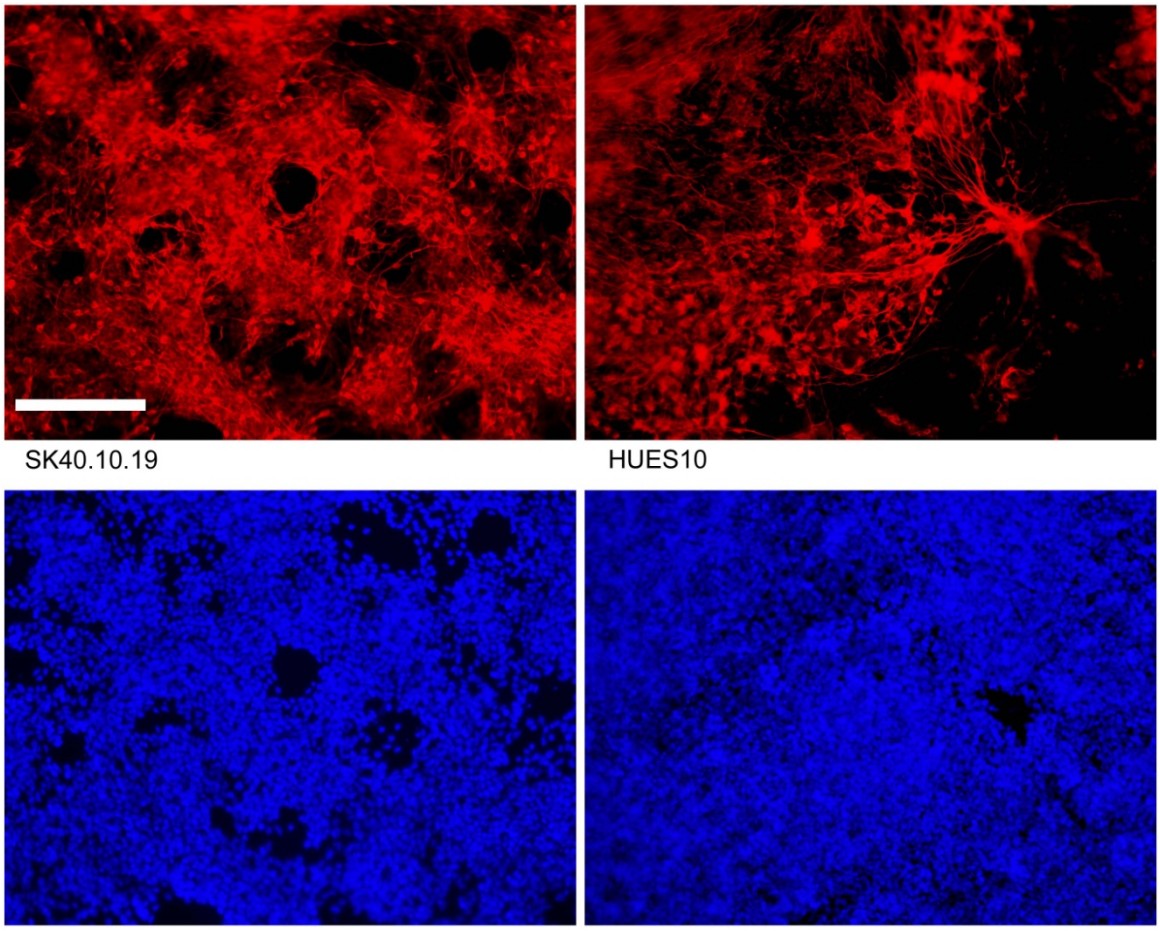

Figure 3-14 β-III tubulin staining of neuronal differentiated cells
EBs derived neuronal cells from the clone SK40.10.19 were stained positive for β-III-tubulin (red signal), a neuronal specific marker, suggesting that HAC containing hES cells retained the ability of lineage-directed differentiation. Nuclei of the cells were counterstained with DAPI (blue signal). HUES10 cells derived neuronal cells were also stained as control. Scale bar = 200μm.

Following the neuronal differentiation of SK40.10.19 clone, the cells were subjected to trypsinisation and cytospun onto poly-L-lysine coated microscopic slides. Subsequently, immuno-FISH with vector-specific probe (pBeloBAC11) and β-III-tubulin staining was conducted to determine whether the neuronal cells retained the HAC during the differentiation events (immuno-FISH experiment was conducted and analysed by Dr. Daniela Moralli). As the differentiated neuronal cells do not divide, the analysis was carried out on interphase cells. As represented in Figure 3-15, HAC was retained in the

neuronal cells stained with β-III-tubulin at the same frequency (50%), as in undifferentiated SK40.10.19 cells.

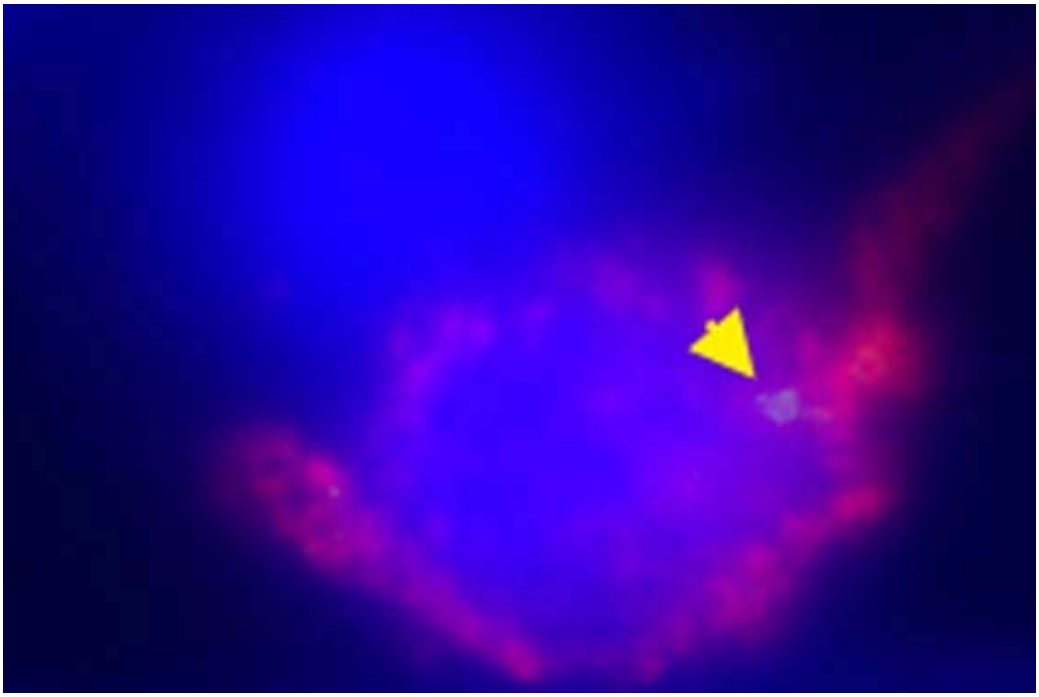

Figure 3-15 Immuno-FISH on neuronal cells
Immuno-FISH with the HAC vector specific probe (green signal) was conducted on neuronal cells derived from SK40.10.19 clone. HAC is identified by yellow arrow in an interphase cell stained with anti-βIII tubulin antibody (red signal). The nuclei were counterstained with DAPI (blue). Immuno-FISH experiment was performed and analysed by Dr. Daniela Moralli.

3.4 Discussion

Our laboratory has previously demonstrated that the HSV-1 amplicon system is highly efficient at delivering large DNA fragments (150kb) into a variety of cells types (~10^3-10^4 fold more than chemical transfection methods) with comparable transduction efficiencies to lentiviruses, without the need of helper virus (Moralli et al., 2006). Thus, the HSV-1 amplicon system is an ideal tool to study the development of human artificial chromosome (HAC). With the aim of generating HAC in hESc, to study their potential application to gene transfer into stem cells, a HSV-1 HAC vector was introduced into

HUES10 cells, which successfully generated HAC as confirmed by FISH and immunostaining for CENP-C. The data from the results on HUES10 has contributed to two manuscripts (Mandegar et al., 2011; Moralli et al., 2010), and the work represents the first example of *de novo* HAC formation in karyotypically normal hESc.

The HSV-1-based HAC constructs were transduced into the HUES10 cells with reasonable efficiencies (~26% of input cells), and led to the generation of several stable clones upon drug selection. The capability of HSV-1 amplicon system to efficiently deliver the HAC vectors was also demonstrated in another hESc line HUES2 in the parallel experiments conducted by Dr. Mohammad A. Mandegar (Mandegar, 2011). The study also established that the HSV-1 amplicon system was about 10^3-10^4 times more efficient than the chemical transfection methods in introducing the large DNA into the HUES2 cells.

The FISH analysis of the stable clones revealed the presence of HAC in most of the analysed clones at a frequency of 12-50% of the metaphase spreads. In one clone (SK40.10.19), the HAC was detected in approximately 50% of the analysed metaphase spreads and therefore this clone was chosen for further analysis. In addition, immunostaining of the SK40.10.19 clone for kinetochore components (CENP-C) also confirmed that the HAC possessed an active and functional centromere.

When the clone SK40.10.19 was grown in the absence of selection for over three months, the HAC frequency did not change significantly, suggesting that the frequency of HAC detection is not due to mitotic instability, as observed in other hESc HAC clones previously generated in the laboratory (Mandegar, 2011; Mandegar et al., 2011). An alternative hypothesis to explain why the HAC is not detected in 100% of the cells is that

different sublines could be present within the clone, due to problems in generation of stable hESc clones. The occurrence of false resistant clones (as indicated by PCR analysis in Table 3-4) suggested incomplete drug selection of cells and further supported the idea that pure populations of clones are difficult to obtain. The hESc have a poor survival rate as individual cells (Watanabe et al., 2007) making it difficult to obtain pure clonal population. This hindered selection for an extended period, since the demise of the majority of untransduced cells can result in the death of any resistant cells, as the total cells number drops below a certain threshold. A short period of drug selection, on the other hand, results in the survival of non-resistant clones as well.

As hESc develop chromosomal abnormalities over long term culture (Baker et al., 2007), and the appearance of trisomy 12 and trisomy 17 has been commonly reported to provide a selective growth advantage to hESc (Draper et al., 2004; Mitalipova et al., 2005), the number of chromosomes was checked in the selected SK40.10.19 clone and found to be normal (46 + HAC). Additionally, FISH analysis with α-17 DNA and chromosome 12 paint probes did not reveal extra copies of either chromosomes 12 and 17. The overall analysis suggested that the HSV-1 amplicon DNA delivery, the G418 selection and the formation of HAC did not trigger any mechanism inducing chromosome instability, an important result in view of gene therapy applications. Moreover, since the HAC are composed of multiple copies of rearranged input DNA, it has been speculated that HAC formation could not be obtained in karyotypically normal cells, as the existence of active pathways preserving the genomic integrity would have triggered apoptosis upon detection of amplification. The data presented here suggests that this hypothesis might not be true and supports the theory that the mechanism underlying the HAC formation is separate

from the breakage-fusion-bridge cycle, a known cause of amplification thus triggering apoptosis (Ma et al., 1993).

When HAC vectors are introduced into other cell types such as HT1080 cells, either by chemical transfection or HSV-1 amplicon transduction, the exogenous DNA frequently integrates into the host chromosomes (Mandegar et al., 2011; Mejia and Larin, 2000; Mejia et al., 2001; Moralli et al., 2006). The HUES10 clones analysed so far did not show any sign of integration, even though the FISH technique is able to efficiently detect integration of sequences as small as 2kb. The lack of integrations constitutes a significant accomplishment in view of the potential application of HAC in gene therapy applications. The difference could be related to the HUES10 normal genetic background, compared to the tumour derived HT1080, but further data, and the analysis of a larger number of clones, will be required to support this hypothesis.

The *GFP* expression of the stable clones, monitored by fluorescence microscopy, revealed that none of the clones exhibited detectable levels of *GFP* expression. However, at least in SK40.10.19, FACS analysis revealed that approximately 33% of the cells were weakly GFP positive. Furthermore, quantitative PCR demonstrated that *GFP* was expressed at low levels. It is possible that the low expression level is due to a dilution effect from non-HAC containing cells. Alternatively, the presence of cells, apparently GFP negative, could be due to the gradual silencing of HSV IE4/5 promoter driving the *GFP* expression. To be fully active, this promoter requires the transactivator protein VP16, which is normally internalised during the transduction process. This phenomenon has also been observed in mouse ES cells where *GFP* expression under HSV IE4/5 promoter gradually reduces over 2-6 days period and is not visible after 6 days (Vicario and Schimmang, 2003).

The FISH on extended chromatin fibres revealed that the HAC resulted from multimerisation of the input DNA and consisted of alternate pattern of α-satellite and vector sequences and therefore it is likely that the generated HAC contained more than a single copy of the transgene. As HAC contains both heterochromatic and euchromatic regions in a similar manner to endogenous chromosomes (Grimes et al., 2004), the phenomenon of multimerisation will lead to the localisation of the transgenes away from the heterochromatic area, and thereby avoid potential silencing of the transgenes. The stochastic distribution of the heterochromatin could possibly explain the variable *GFP* expression observed during the prolonged culture of SK40.10.19 cells. To examine whether physiological expression of the transgene can be achieved it will be important to study the expression analysis of the transgenes placed under their own promoter and appropriate regulatory elements.

Importantly, neither the HSV-1 transduction nor the HAC formation led to a loss of the pluripotency in HUES10 cells, as suggested by the immunofluorescence staining with hESc specific markers TRA-1-60, OCT4, and NANOG. Moreover, EBs derived from the HAC containing hESc demonstrated the capability of differentiation into cells representing all the three germ layers (ectoderm, mesoderm and endoderm), further validating the pluripotent status of the hESc. This differentiation capability was also confirmed by RT-PCR analysis of EB-derived differentiated cells for all the three germ layers markers.

Lineage-directed differentiation of hESc is one of the most active areas of research and can provide a valuable source of donor cells for cell replacement therapies. hESc have been proposed to be used in cells replacement procedures for neurological disorders such as Parkinson's disease and multiple sclerosis. Moreover, hESc mediated *in vitro* derived

neurons could be utilised for drug-based or genetic screenings (Cowan et al., 2004; Kiskinis and Eggan, 2010; Soldner et al., 2009; Thomson et al., 1998). It was therefore significant to examine whether generation of HAC affected the directed differentiation of hESc, particularly into the neuronal cells. As confirmed by neuronal staining, the SK40.10.19 clone exhibited successful differentiation into neuronal cells at high efficiency. Moreover, the differentiation did not cause any reduction in the HAC frequency.

Overall, this study is the first report of *de novo* HAC formation in karyotypically normal hESc and demonstrated that HAC is an important delivery system for large genomic elements into hESc. Moreover, HAC development in hESc has also established a system to efficiently deliver large DNA into the induced pluripotent stem cells (iPS) cells, which represent more relevant pluripotent stem cell type to perform disease- and patient-specific studies.

Chapter 4. Reprogramming of Fibroblasts using the HSV-1 Amplicon System

4.1 Introduction

The ability of human embryonic stem cells to grow indefinitely while maintaining pluripotency and self-renewal ability, has led to their use in understanding disease pathogenesis, screening for drugs and treating various injuries such as spinal cord injuries, or disorders such as juvenile-onset diabetes mellitus and Parkinson's disease (Thomson et al., 1998). However, their derivation is associated with ethical controversies of manipulating human embryos and their use in disease treatment is challenging because of immune-rejection issues. Generating induced pluripotent stem (iPS) cells without the problems associated with the destruction of embryos has been a great achievement in the field of regenerative medicine. The hiPS cells can provide disease-specific and patient-specific donor cells, which can be used in a variety of therapeutic applications such as determining the mechanism of disease development, carrying out drug screening and toxicology and exploring the possibilities of gene correction and cell therapy (Robinton and Daley, 2012; Takahashi et al., 2007).

The generation of hiPS cells was initially achieved by reprogramming somatic cells via the forced expression of the four genes: *OCT4*, *SOX2*, *KLF4* and *C-MYC* (Takahashi et al., 2007) or using *NANOG* and *LIN28* instead of *KLF4* and *C-MYC* (Yu et al., 2007). Since then, many groups have derived iPS cells using variety of gene delivery systems to introduce the reprogramming genes into the somatic cells. These include retroviral/lentiviral vectors (Takahashi et al., 2007; Yu et al., 2007), adenoviral vectors

(Zhou and Freed, 2009), transposons (Kaji et al., 2009; Woltjen et al., 2009), plasmids and episomal vectors (Jia et al., 2010; Yu et al., 2009), protein transductions (Zhou et al., 2009), synthetic mRNAs (Warren et al., 2010) and mature miRNAs (Miyoshi et al., 2011). However, these delivery systems are associated with one or many of the following challenges: transgene integration/reactivation, labour-intensive screening/removal of reprogramming genes, and very low reprogramming efficiencies. Therefore, there is a need to evaluate other systems for simple, efficient and integration-free generation of iPS cells in order to further assess their clinical significance.

The aim of this study was to utilise the Herpes Simplex Virus (HSV)-1 amplicon system to deliver the reprogramming genes into the fibroblasts and generate iPS cells. The HSV-1 amplicon system offers various advantages that distinguish it from other DNA delivery systems. These include integration free delivery of large DNA fragments (up to 152kb), the ability to infect a wide range of mammalian cell types including dividing and non-dividing cells, the high delivery efficiency with low cytotoxicity. In addition, the head-to-tail, rolling-circle DNA replication mechanism of HSV-1 also allows amplification of amplicon plasmid as concatamers leading to increased copy number of reprogramming factors and consequently their expression levels (Saeki et al., 2003; Saeki et al., 2001).

The HSV-1 amplicon vectors were constructed containing the reprogramming genes (*OCT4*, *SOX2*, *KLF4* and *C-MYC*). The transduction of HSV-1 amplicon vectors was optimised on IMR90 fibroblast cells, which were then used in various reprogramming attempts to generate iPS cells. The derivation strategy was improved by incorporating various alterations into the protocol as well as using the small molecules (chemicals that enhance the reprogramming efficiency) to obtain iPS-like colonies. These colonies were analysed in comparison to the hES/iPS colonies reported in previous studies.

4.2 Specific Methods

4.2.1 Generation of HSV-1 Amplicon Based Reprogramming Vectors

Vectors containing cDNAs of the reprogramming genes (Takahashi et al., 2007) *OCT4* (pGEM-Oct4, plasmid 16352), *SOX2* (pGEM-Sox-2, plasmid 16353), *KLF4* (pMXs-hKLF4, plasmid 17219) and *C-MYC* (pBS human c-Myc, plasmid 14971) were obtained from Addgene (Siegel et al., 2003; Takahashi et al., 2007; Yu et al., 2007). To obtain constitutive expression of reprogramming factors, the CAG promoter was excised from pCAG-Oct3/4-IRES-Neo (Addgene, Plasmid 13461) (Tokuzawa et al., 2003) by EcoRI (blunted with Klenow)/NdeI digestion and subsequently cloned into the pIRES vector (Clontech) at SacI (blunted with T4 DNA polymerase)/NdeI restriction sites to produce pIRES-CAG vector (Figure 4-1 A).

Three HSV-1 based reprogramming vectors, pHG-CAG SOKM (Figure 4-1 B), pHG-CAG SOKTM (Figure 4-2 A) and pHG-TOKSM (Figure 4-2 B) were constructed in the following manner. Individual cDNAs of the reprogramming genes *OCT4* (O), *SOX2* (S), *KLF4* (K) and *C-MYC* (M) were excised from their parental constructs and cloned in pIRES-CAG and pTRE-Tight (Clontech) using NotI restriction site for O, S and K and EcoRI restriction site for M. These genes were subsequently excised as an expression cassette (with their respective promoter and polyA site) by BamHI/BglII digestion and ligated one by one into pBeloBAC11 (Kim et al., 1996) at BamHI site to create pBelo CAG-SOKM (constitutive expression) and pBelo T-OKSM (inducible expression). pBelo CAG SOKTM (constitutive expression for SOK and inducible expression for M) was created by excising the *C-MYC* expression cassette (under TRE-Tight promoter) using

BamHI/BglII digestion and cloning it to pBelo CAG-SOK vector at BamHI site. Table 4-1 lists all the vectors and their features utilised in the cloning steps.

Table 4-1 Vectors utilised in cloning reprogramming genes

Vector	Important Characteristics	Size	Mammalian Resistance Gene (Promoter)	Bacterial Resistance Gene
pGEM-Oct4	Human *OCT4* ORF	4.1 kb	-	*Amp*
pGEM-Sox2	Human *SOX2* ORF	3.9 kb	-	*Amp*
pMXs-hKLF4	Human *KLF4* ORF	6.0 kb	-	*Amp*
pBS human c-myc	Human *C-MYC* ORF	4.3 kb	-	*Amp*
pCAG-Oct3/4-IRES-Neo	mouse Oct3/4 (CAG promoter)	7.4 kb	*Neo* (CAG)	*Amp*
pIRES	Backbone vector	6.1 kb	*Neo* (SV40)	*Amp*
pIRES CAG	CAG promoter	7.2 kb	*Neo* (SV40)	*Amp*
pTRE-Tight	P_{tight} Tet-responsive promoter	2.6 kb	-	*Amp*
pTet-On-Advanced	rtTA-Advanced (CMV promoter)	7.1 kb	*Neo* (SV40)	*Amp, Kan*

The rtTA-Advanced transactivator (required for induction of TRE-promoter upon addition of doxycycline) cassette was excised from pTet-On Advanced (Clontech) using EcoRI/BamHI (blunted with Klenow) digestion and cloned into the pIRES-CAG vector at NotI site (blunted with Klenow) for its constitutive expression. The resulting expression cassette (pCAG-rtTA) was then excised using BamHI/BglII digestion and cloned into pBelo TOKSM and pBelo CAG SOKTM at BamHI site to obtain pBelo TOKSM-rtTA and pBelo CAG SOKTM-rtTA respectively. Then the vectors pBelo CAG-SOKM, pBelo TOKSM-rtTA and pBelo CAG SOKTM-rtTA were retrofitted by *LoxP*/*Cre* recombination with pHGNeo4 (Figure 2-1) to include the HSV-1 packaging elements and

the resulting vectors were named pHG-CAG SOKM, pHG-TOKSM and pHG-CAG SOKTM respectively. Maps of these vectors are shown in Figure 4-1and Figure 4-2. Table 4-2 list the properties and important features of these vectors.

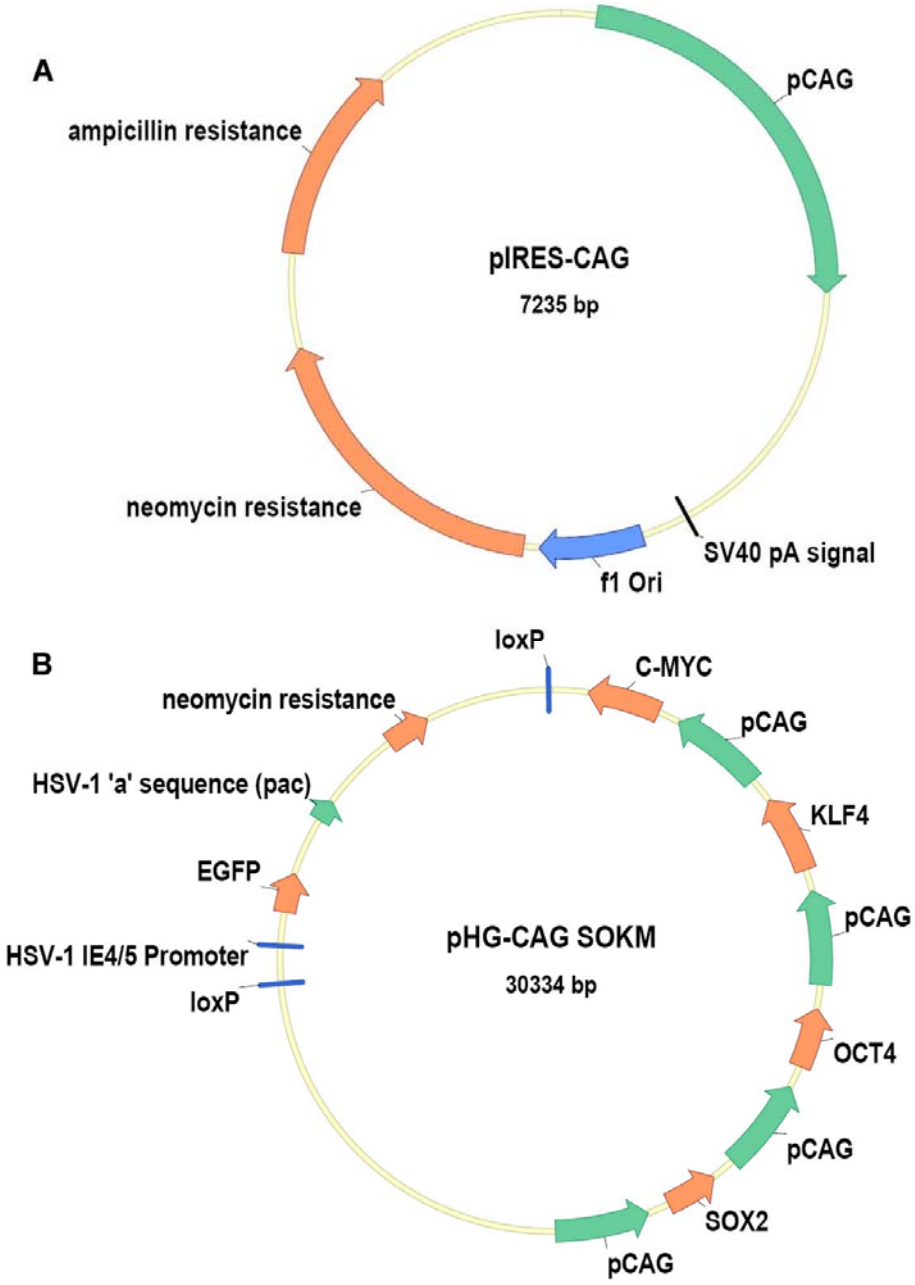

Figure 4-1 Vector maps of pIRES-CAG and pHG-CAG SOKM
pIRES-CAG (A) was derived from the pIRES vector (Clontech) by substituting the CMV IE promoter with CAG promoter. pHG-CAG SOKM (B) was constructed to incorporate the HSV-1 amplicon elements and the CAG promoter-driven reprogramming genes (*SOX2*, *OCT4*, *KLF4* and *C-MYC*).

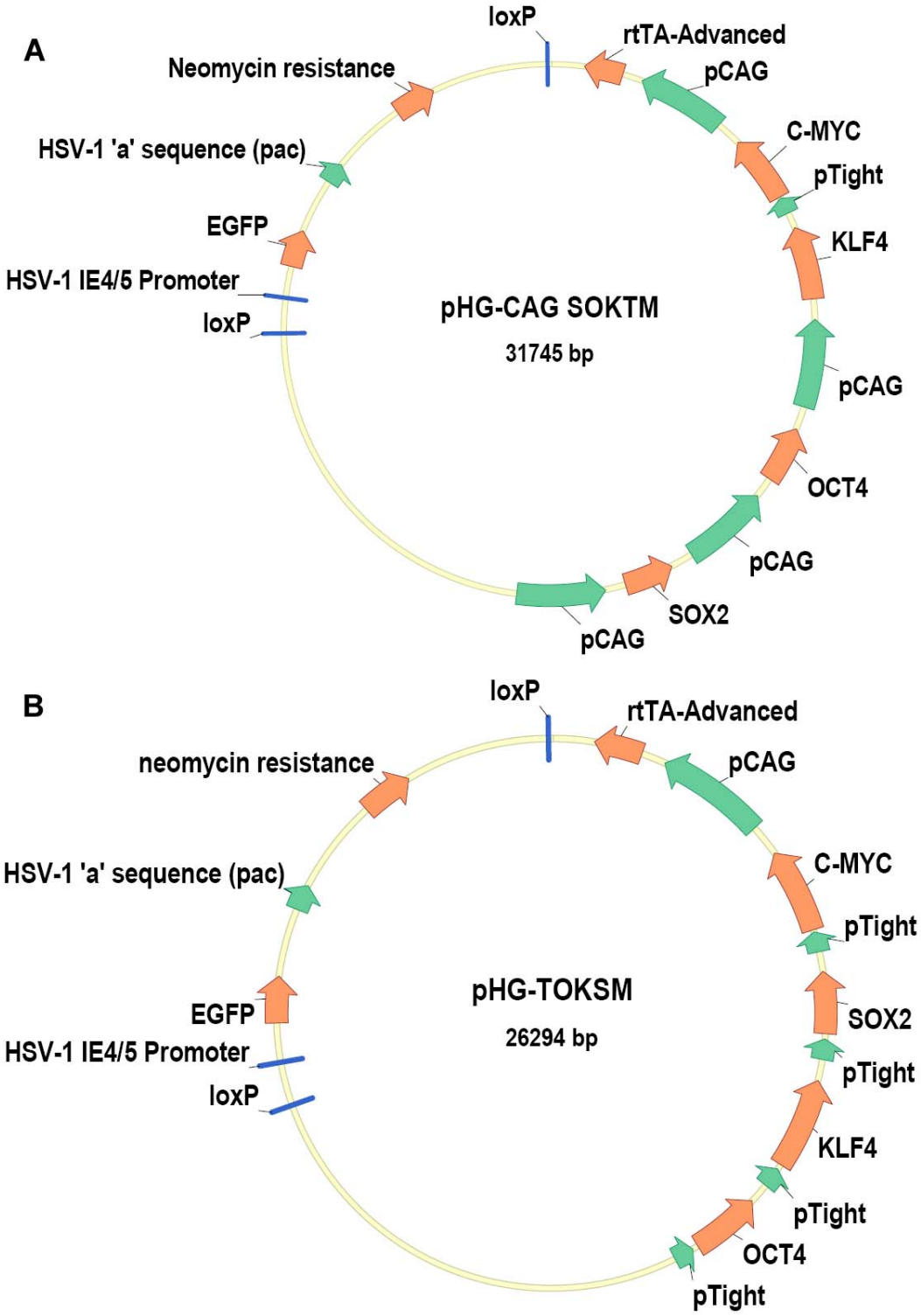

Figure 4-2 Vector maps of pHG-CAG SOKTM and pHG-TOKSM
pHG-CAG SOKTM (A) included the HSV-1 amplicon elements, CAG promoter-driven genes (*SOX2*, *OCT4*, *KLF4* and rtTA-Advanced) and TRE-Tight promoter driven *C-MYC*. pHG-TOKSM (B) contained the HSV-1 amplicon elements, TRE-Tight promoter driven reprogramming factors (*OCT4*, *KLF4*, *SOX2* and *C-MYC*) and rtTA-Advanced gene under the CAG promoter.

Table 4-2 Features of reprogramming vectors

Vector	Size	Reprogramming genes (promoter)	Induction gene (promoter)	Reporter gene (Promoter)	Resistance gene (Promoter)
pHG- CAG SOKM	30.3kb	*OCT4*, *SOX2*, *KLF4*, *C-MYC* (CAG)	N/A	*EGFP* (HSV-1 IE4/5)	*Neo* (SV40)
pHG-TOKSM	26.2kb	*OCT4*, *SOX2*, *KLF4*, *C-MYC* (TRE-Tight)	rtTA-Advanced transactivator (CAG)	*EGFP* (HSV-1 IE4/5)	*Neo* (SV40)
pHG-CAG SOKTM	31.7kb	*OCT4*, *SOX2*, *KLF4* (CAG) and *C-MYC* (TRE-Tight)	rtTA-Advanced transactivator (CAG)	*EGFP* (HSV-1 IE4/5)	*Neo* (SV40)

4.2.2 PCR and RT-PCR Analyses

The PCR analysis of *C-MYC* and *KLF4* gene and the RT-PCR analysis to confirm the expression genes (*OCT4*, *SOX2*, *KLF4* and *C-MYC*) from the reprogramming vectors were carried out using the primer sets listed in Table 4-3.

Table 4-3 Primer sets used for RT-PCR analysis

Genes/Region	Size	Symbol	Sequence (5'- 3')
C-MYC	190bp	c-myc P1 F	CCGAGGAGAATGTCAAGAGG
		c-myc P1 R	AGCTTTTGCTCCTCTGCTTG
C-MYC	159bp	c-myc P2 F	ACTCTGAGGAGGAACAAGAA
		c-myc P2 R	TGGAGACGTGGCACCTCTT
C-MYC	155bp	c-myc P3 F	GGAAACGACGAGAACAGTTGA
		c-myc P3 R	GGTTGTGAGGTTGCATTTGA
C-MYC	123bp	c-myc P4 F	CGGAACTCTTGTGCGTAAGG
		c-myc P4 R	CTCAGCCAAGGTTGTGAGGT
KLF4	136bp	KLF4 P1 F	GTCTCTTCGTGCACCCACTT
		KLF4 P1 R	TGCTCAGCACTTCCTCAAGA
KLF4	127bp	KLF4 P2 F	CCCAATTACCCATCCTTCCT
		KLF4 P2 R	ACGATCGTCTTCCCCTCTTT
KLF4	196bp	KLF4 P3 F	TATGACCCACACTGCCAGAA
		KLF4 P3 R	TCCAGTCACAGACCCCATCT
KLF4	110bp	KLF4 P4 F	TATGACCCACACTGCCAGAA
		KLF4 P4 R	TGGGAACTTGACCATGATTG
GAPDH	202bp	GAPDH F	TGTTGCCATCAATGACCCCTT
		GAPDH R	CTCCACGACGTACTCAGCG
SOX2	121bp	Sox2 F	ACTTTTGTCGGAGACGGAGA
		Sox2 R	CATGAGCGTCTTGGTTTTCC
OCT4	125bp	Oct4 F	AGTGAGAGGCAACCTGGAGA
		Oct4 R	GCCGGTTACAGAACCACACT

4.3 Results

4.3.1 HSV-1 Amplicon Vectors for Reprogramming

In this study, three HSV-1 amplicon vectors were developed for the generation of iPS cells from the human fibroblasts. The first HSV-1 vector pHG-CAG SOKM, incorporated

the four reprogramming genes (*OCT4*, *SOX2*, *KLF4* and *C-MYC*) under the control of a constitutive CAG promoter. This approach was similar to those used in the early reports of direct reprogramming, based on lentiviral/retroviral vectors and constitutive expression of the reprogramming factors (Takahashi et al., 2007; Yu et al., 2007). In subsequent studies, second generation lentiviral vectors carried the reprogramming genes under inducible promoters (Maherali et al., 2008; Park et al., 2008a). Accordingly, the second HSV-1 vector pHG-TOKSM contained the four genes regulated by a doxycycline inducible TRE-Tight promoter. The third vector, pHG-SOKTM, was assembled with three genes (*OCT4*, *KLF4* and *SOX2*) under control of the constitutive CAG promoter while *C-MYC* expression was driven by the inducible TRE promoter with the aim to acquire controlled expression of c-myc. This was a desirable feature since *C-MYC* is an oncogene and its reactivation has been associated with development of tumours in mice (Okita et al., 2007). Secondly, *C-MYC* has been reported to be dispensable for generation of iPS cells although its absence reduces the reprogramming efficiency (Huangfu et al., 2008; Okita et al., 2008).

4.3.2 Choice of Target Cell Line for Reprogramming

Reprogramming experiments have been successfully carried out on several human cell types, such as adult dermal human fibroblasts, adult mesenchymal stem cells, primary foetal fibroblasts (MRC5, Detroit 551, and IMR90), newborn foreskin fibroblasts, ES-cell derived fibroblasts, keratinocytes, and adipose stem cells etc (Kiskinis and Eggan, 2010). For the HSV-1 based reprogramming experiments, IMR90 foetal fibroblasts were used as these cells were readily available through the American Type Culture Collection (ATCC) repository, and their well characterised genetic fingerprint allows confirmation of the origin of iPS clones with ease. These cells proliferate robustly beyond 20 passages but

grow slowly in hESc medium, a characteristic that provides growth advantage to reprogrammed iPS cells and hence aids in their identification on the basis of morphology (e.g., compact border and high nucleus to cytoplasm ratio). Moreover, IMR90 cells have been reported to be successfully reprogrammed by different methods (Banito et al., 2009; Sun et al., 2009; Yu et al., 2007), thus making them an ideal choice of cell type for testing the efficacy of HSV-1 amplicon system.

4.3.3 Optimisation of HSV-1 Transduction on IMR90 Cells

In preliminary experiments, the conditions for HSV-1 amplicon transduction in IMR90 cells were determined. Cells were seeded at a density of $1x10^5$ cells per 24-well dish and were transduced with pHGNeo4 vector at four different MOIs (5, 10, 15 and 20) followed by spinoculation. Based on the number of *GFP* expressing cells 24 hours post-transduction (data not shown), it was established that efficient transduction (30%) was achieved with cells transduced at MOI 10. As shown in Table 4-4, any additional increase in the transducing units (MOI 15 and MOI 20) either only contributed to a slight improvement in transduction efficiency (MOI 15) or even reduced the transduction efficiency (MOI 20) as the cells probably reached transduction saturation. The increase or decrease in transduction efficiency was statistically significant as analysed by Fisher's exact test (2x2 contingency table).

Table 4-4 Transduction efficiencies of IMR90 cells at different MOIs

MOI of HSV-1 amplicon	Transduction Efficiency (%)
5	23
10	30
15	33
20	29

The presented transduction efficiency is based on the average transduction efficiency of two independently transduced wells.

4.3.4 Expression Analysis of Reprogramming Vectors

The three HSV-1 based reprogramming vectors pHG-CAG SOKM (containing *OCT4*, *SOX2*, *KLF4* and *C-MYC*, controlled by a constitutive CAG promoter), pHG-TOKSM (containing *OCT4*, *SOX2*, *KLF4* and *C-MYC*, with a doxycycline inducible TRE promoter), and pHG-CAG SOKTM (where *OCT4*, *SOX2* and *KLF4* are controlled by the CAG promoter, and *C-MYC* by the TRE promoter) were constructed as described in the Section 4.2.1. To verify that the four genes *OCT4*, *SOX2*, *KLF4* and *C-MYC* were efficiently transcribed, their expression was determined through RT-PCR analysis on the cDNA extracted from the SNL76/7 cells transfected with the above mentioned vectors. As two of the genes, *KLF4* and *C-MYC*, are endogenously expressed in human fibroblast cells (Maherali et al., 2008), the analysis was carried out in mouse SNL76/7 cells using primer pairs which specifically amplify the human transcripts of the four genes but not their murine counterparts (Figure 4-3). Four sets of primers were designed (Table 4-3), based on the *KLF4* and *C-MYC* human cDNA sequences (NCBI Accession number NM_004235 and NM_002467, respectively). As a negative control, the primer sets were tested for their capacity to amplify murine cDNA extracted from non transfected SNL76/7 cells. As shown in Figure 4-3, primer sets C-MYCP1, C-MYCP3, C-MYCP4, KLF4P3 and KLF4P4 produced positive bands. Hence, C-MYCP2 and KLF4P1 were

chosen, as they do not recognise the murine targets. Primers for *OCT4* and *SOX2* were not included in this analysis, as these are pluripotent genes and are normally not expressed in fibroblast cells.

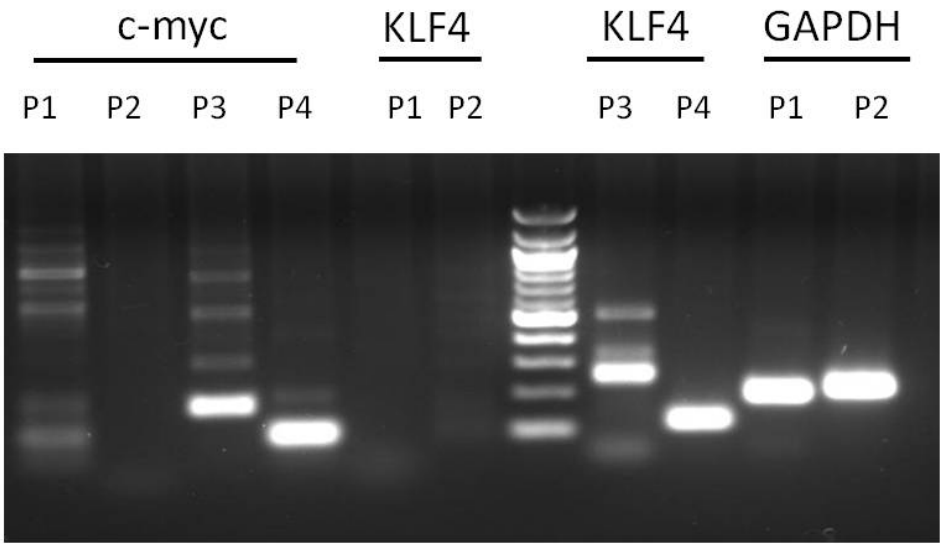

Figure 4-3 PCR analysis of SNL 76/7 (mouse) DNA with different primers for human c-myc and KLF4
cDNA from SNL cells76/7 was subjected to PCR with four different primers pairs (P1-P4) for *C-MYC* and *KLF4*. Primer pair 2 for *C-MYC* and primer pair 1 for *KLF4* do not amplify the mouse DNA. *GAPDH* primers were used as an internal control.

3µg DNA each of the three reprogramming vectors was transfected into $2x10^6$ SNL76/7 cells to evaluate the expression of the four reprogramming genes from the reprogramming vectors. For the inducible vectors pHG-TOKSM and pHG-CAG SOKTM, 1µg/ml doxycycline was added for two days. 48 hours post-transfection, total RNA was extracted, retrotranscribed into cDNA and analysed by PCR for *OCT4*, *SOX2*, *KLF4* and *C-MYC* gene transcripts. These experiments confirmed that the four genes are transcribed in all the three vectors. Figure 4-4 shows the expression analysis of genes from the inducible vector pHG-TOKSM (mostly used for reprogramming experiments as described in Section 4.3.7). However, as shown in Figure 4-4 the doxycycline inducible TRE promoter (in vector pHG-TOKSM) exhibited a leaky (weak) control of transcription, as

positive bands were detected for all genes in cells not treated with doxycyline (D-). As RNA not subjected to cDNA retro-transcription (RT-) failed to generate any amplification product, the presence of bands in the samples not treated with doxycycline is not due to contaminating genomic or vector DNA.

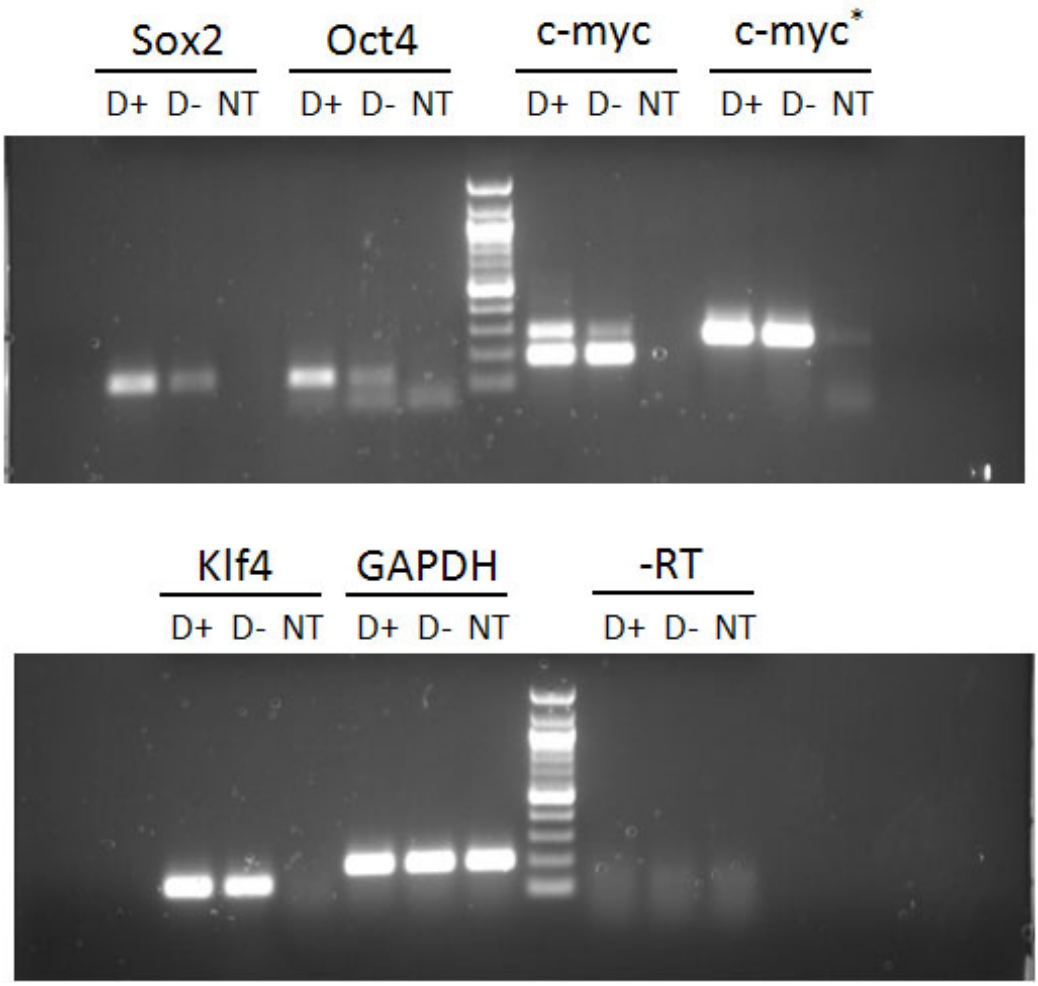

Figure 4-4 RT-PCR analysis of gene expression of *SOX2*, *OCT4*, *C-MYC* and *KLF4* in SNL 76/7 pHG-TOKSM transfected cells

RT-PCR analysis with primers set specific for the human *SOX2*, *OCT4*, *C-MYC* and *KLF4* transcripts in cells transfected with pHG-TOKSM, treated (D+) or untreated (D-) with doxycycline, using parental untransfected cells (NT) as negative controls. The primers used can specifically detect the human version of the genes but not those of the mouse. PCR of the RNA without reverse transcriptase (-RT) was done with primer pair of *SOX2* to ensure no amplification of genomic or vector DNA had occurred. *GAPDH* gene was used as an internal control.

c-myc* = amplification of it with the primer pair 1(C-MYCP1) was carried out as two bands were observed using C-MYCP2 primer set . D+ = added doxycycline (1ug/ml). D- = no doxycycline. NT = Non transfected SNL76/7 cells.

4.3.5 Real-time Quantitative PCR Analysis

To further investigate the leaky (weak) control of expression exhibited by the doxycycline inducible TRE promoter, quantitative real-time PCR experiments were conducted on the pHG-TOKSM transfected cDNA samples with *Sox2*, *Oct4*, *c-myc*P1, *KLF4*P1 and *GAPDH* primer sets as listed in Table 4-3. After verifying on a series of template concentrations that the primers had similar optimal amplification efficiency, the relative amounts of each gene transcripts were measured in the doxycycline induced cells, compared to the un-induced cells, normalising the average C_T values to *GAPDH*. The fold enrichment was measured using the $2^{-\Delta\Delta Ct}$ method. The analysis confirmed that the addition of doxycycline greatly increased the expression of the four genes as shown in Figure 4-5. The C_T values for *GAPDH* were almost the same in the induced and un-induced samples but were different from the four genes.

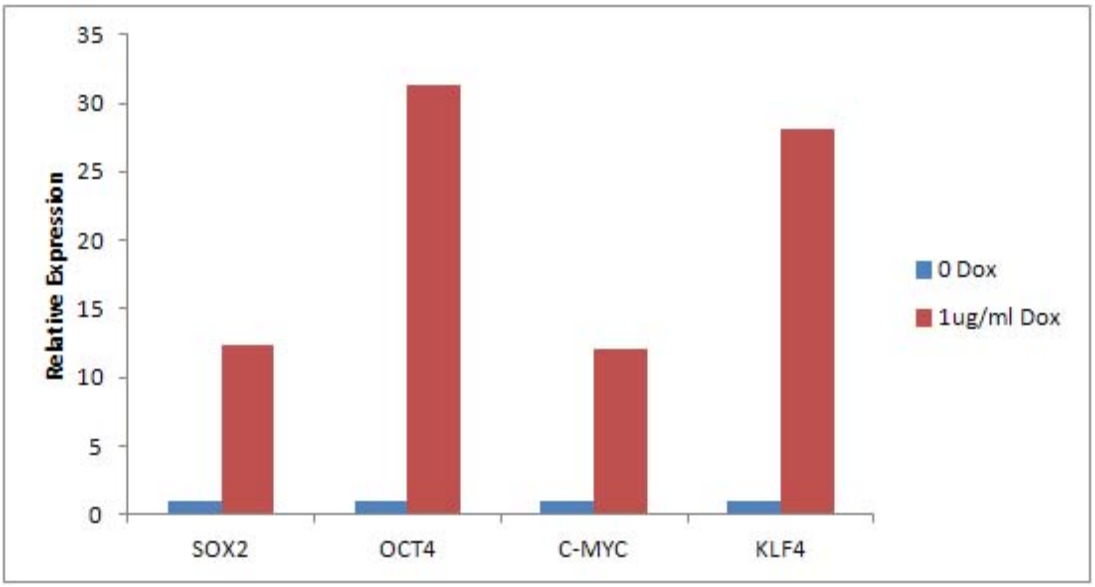

Figure 4-5 Real time quantitative PCR analysis of expression of genes
The graph indicates relative expression of the reprogramming genes from pHG-TOKSM vector with and without the addition of doxycycline as analysed by quantitative RT-PCR. Each sample was run in duplicate. The analysis shows enrichment of 12-30-fold upon the addition of doxycycline.

4.3.6 Reprogramming of IMR90 Cells with HSV-1 Amplicons

In the initial reprogramming experiments (schematically presented in Figure 4-6, Protocol 1), 1 x 10^5 IMR90 cells (passage 15) were independently transduced with pHGNeo4 (control vector), pHG-CAG SOKM, pHG-TOKSM and pHG-CAG SOKTM amplicons, at MOI 10 in 6-well dishes. After 24 hours, the cells were washed with PBS to remove the transduction mix, and then incubated in the standard DMEM medium for five days. On the 6^{th} day, the cells were reseeded on SNL76/7 feeder cells in 10cm dishes and incubated in hESc medium for additional thirty days. The hESc medium was replenished every other day and doxycycline (1μg/ml) was added every day in the pHG-TOKSM and pHG-CAG SOKTM transduced cells. After thirty days of incubation, no hES/iPS like colonies were obtained with any of the three vectors. These unsuccessful experiments could be due to two different reasons. The cells became confluent at the 18-20^{th} day, and as it has been previously reported that confluency of the cells before thirty days can have negative effects on reprogramming (Maherali and Hochedlinger, 2008), it may have prevented iPS cell formation. Alternatively, it is possible that the HSV-1 vectors were lost during the course of experiment as they do not integrate into the cellular genome (Johnston et al., 1997). Hence, the experiment was repeated in the presence of G418 (200μg/ml, added 24 hours post-transduction) in order to give selective advantage to cells receiving the four reprogramming constructs. However, no hES/iPS like colony was observed at the end of the experiment.

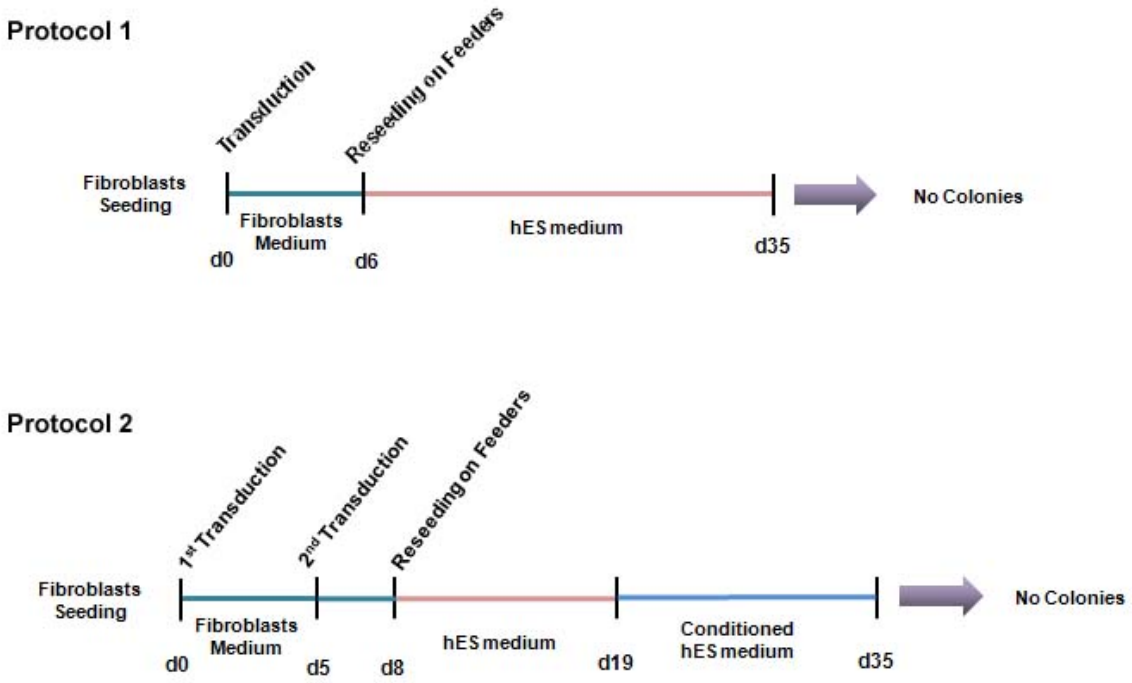

Figure 4-6 Schematic representation of reprogramming Protocols 1 and 2
In protocol 1, a single transduction of IMR90 fibroblasts was conducted and all the transduced cells were transferred to feeder cells on day six, after which the fibroblast medium (standard DMEM medium) was switched with hESc medium. The cells were subsequently cultivated for additional thirty days. In Protocol 2, three things were modified from protocol 1: (1) a second transduction of the cells was carried out on day 5; (2) the transduced cells were plated on feeder cells at 1 x 10^5 cells per 10cm dish; and (3) After 10 days of cultivation on feeder cells, the hESc medium was switched with SNL76/7-conditioned hESc medium.

To ensure that a high number of HSV-1 amplicon constructs were present continuously within the cells, three steps were modified with respect to the previous experiment described above, and summarised in Figure 4-6, Protocol 2. These included (1) on the 5th day, a second transduction was carried out in order to achieve extended expression of reprogramming factors and avoid vector dilution; (2) on day 8, the transduced cells were reseeded in 10cm dishes, at a density of 1x10^5 cells per dish on SNL76/7 feeder layer, so that the cells could not become confluent by the end of the experiment. 3) SNL76/7-conditioned hESc medium was applied to the cells after day 19 to support the deteriorated SNL76/7 feeder cells (Yu et al., 2007). Doxycycline (1μg/ml) was added every day in the

pHG-TOKSM and pHG-CAG SOKTM transduced cells. Despite the described modifications in Protocol 2, no hESc/iPS cell-like colony was observed at the end of the reprogramming experiment.

4.3.7 Reprogramming with Multiple Cycles of Transductions

In addition to the possible loss of the HSV-1 amplicon vectors from the dividing cells over the reprogramming period of 35 days, an alternative explanation for the unsuccessful reprogramming attempts described above is the silencing of the HSV-1 genome in fibroblasts, which would cause short-term provision or transient expression of reprogramming factors (Coleman et al., 2008; Hobbs et al., 2001). Therefore, using an approach similar to the protocol used for iPS generation using reprogramming proteins (Kim et al., 2009a), Protocol 2 was modified further to include several cycles of double transductions as outlined in Figure 4-7-Protocol 3: (1) the IMR90 fibroblasts were seeded on gelatin-coated 6-well dishes and independently transduced at MOI 10 with the three reprogramming vectors (pHG-TOKSM, pHG-CAG SOKM and pHG-CAG SOKTM) and one control vector (pHGNeo4) followed by incubation in hESc medium from the next day; (2) on the day 5, the cells were transduced again with the same amplicons and incubated for four additional days; (3) twenty four hours later, the transduced cells were reseeded at density of $5x10^4$ on gelatin-coated 6-well dish; (4) the cycle of two transductions was repeated twice more, for a total of 6 successive transductions; (5) and finally, all the cells were transferred onto inactivated MEF feeder cells and cultivated for 15 additional days.

Following this protocol, a number of colonies were observed from all the three reprogramming vectors (pHG-TOKSM, pHG-CAG SOKM and pHG-CAG SOKTM),

after each cycle of transduction (Figure 4-8, Table 4-5), however no colonies were detected after the final plating onto the feeder cells. As expected, the control vector pHGNeo4 did not give any colonies.

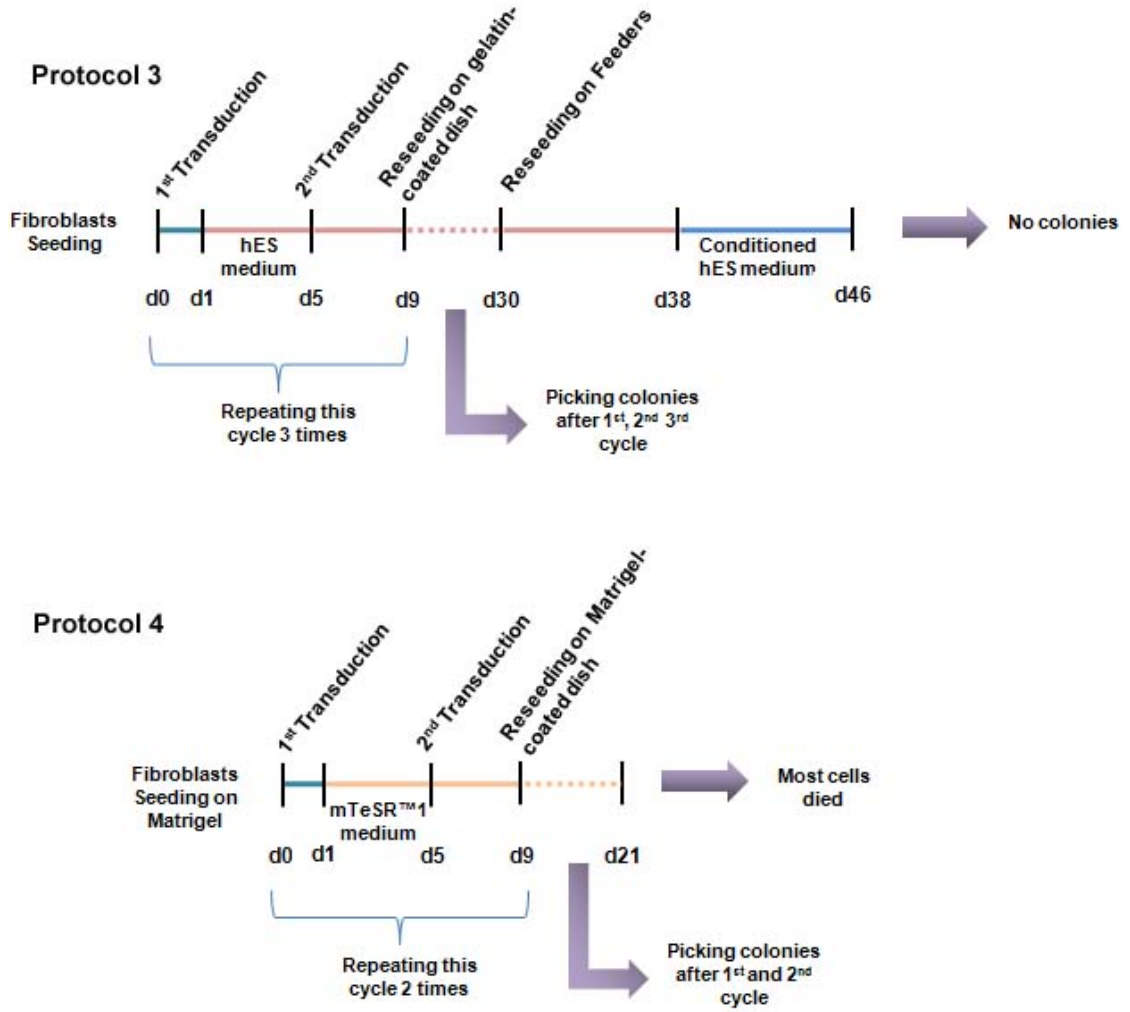

Figure 4-7 Schematic representation of reprogramming Protocols 3 and 4
In protocol 3, IMR90 fibroblasts were transduced twice over the period of six days followed by four days incubation in hESc medium. Then the cells were trypsinised, reseeded on gelatin-coated dishes and the cycle of two transductions and incubation was repeated for two more times. After the final transduction and incubation, the cells were reseeded on an iMEF feeder layer and cultured in hESc medium, which was later switched to conditioned hESc medium. Protocol 4 was the same as Protocol 3 except that mTeSR-1 medium was used instead of hESc medium and the transduced were reseeded on Matrigel-coated dishes instead of gelatin-coated dishes. Both protocols resulted in number of colonies which were picked for expansion.

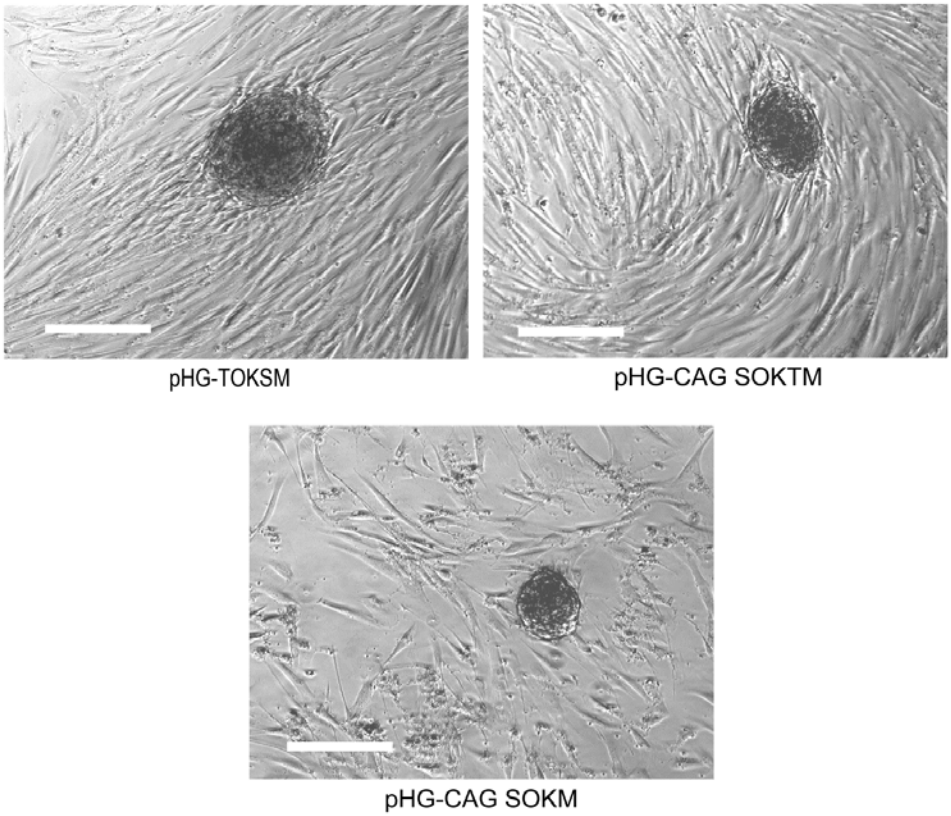

Figure 4-8 Microscopic pictures of initial colonies with Protocol 3
The colonies obtained using the three vectors pHG-TOKSM, pHG-CAG SOKTM and pHG-CAG SOKM appeared to be raised aggregates and exhibited compact appearance, which is unlike the typical appearance of hES/iPS colonies. Scale bar = 200μm.

The colonies were mechanically picked and transferred to 24-well dishes on iMEFs in hESc medium supplemented with 10μM ROCK inhibitor Y-27632. While reattaching readily upon the mechanical transfer, most of the iPS-like clones did not show any further sign of growth. Large colonies were cut in three to four patches under a stereomicroscope and were transferred to fresh feeder cells. The cutting and disaggregating procedures were difficult as the colonies were very firm and compact, with a tendency to form cell clumps upon cutting. These features differed from those of hES/hiPS colonies, suggesting that the iPS-like colonies were not fully reprogrammed. As the colonies did not grow, no further analysis was carried out. Vector pHG-TOKSM (carrying the reprogramming factors under the control of inducible promoter) generated 15 colonies, which were more similar

to hES/hiPS colonies in terms of size, compact appearance, and well defined borders as compared to colonies from the pHG-CAG SOKM and pHG-CAG SOKTM vectors. Therefore, the pHG-TOKSM vector was selected for further round of reprogramming experiments.

Table 4-5 Number of colonies obtained from the Protocol 3

Vectors	1st cycle of transduction	2nd cycle of transduction	3rd cycle of transduction	Total Colonies
pHG-TOKSM	4	4	7	15
pHG-CAG SOKM	3	3	1	7
pHG-CAG SOKTM	4	5	6	15

It is possible that the growth conditions used in Protocol 3, i.e. keeping the cells on gelatin-coated wells, could have hindered full reprogramming, as human ES/iPS cells differentiate in the absence of feeder cells (Takahashi et al., 2007), unless grown on an appropriate protein matrix. Therefore, in the next experiment, (Figure 4-7-Protocol 4) the pHG-TOKSM transduced IMR90 fibroblasts were seeded on Matrigel after the first cycle of transduction (second transduction) and mTeSR™1 medium was used to support iPS generation. The rest of the procedures were the same as described for Protocol 3. Using Protocol 4, forty-nine pHG-TOKSM colonies (Figure 4-9) were obtained after the first and second cycles of transductions (i.e. after second and fourth transduction). As compared to Protocol 3, some of these colonies had an appearance more similar to the hES/iPS colonies while few colonies also exhibited granular appearance (Figure 4-9) as seen in the rudimentary reprogrammed colonies reported in earlier studies of iPS generation (Chan et al., 2009; Takahashi et al., 2007). However, the cells at the end of the second cycle of transduction appeared to be suffering and therefore a third cycle of

transduction was not carried out. The colonies were mechanically picked under the microscope and transferred to fresh feeder cells as described above but failed to grow any further.

It is possible that a gap of five days between the transductions led to a reduced supply of the four reprogramming factors, insufficient for the generation of true iPS cells. To avoid this, further experiments were performed (based on Protocol 3 and 4), in which the cells were transduced with a gap of three and four days. However, following this route, no iPS-like were obtained. In addition, during the course of these experiments, the viability of the cells was significantly impaired (data not shown).

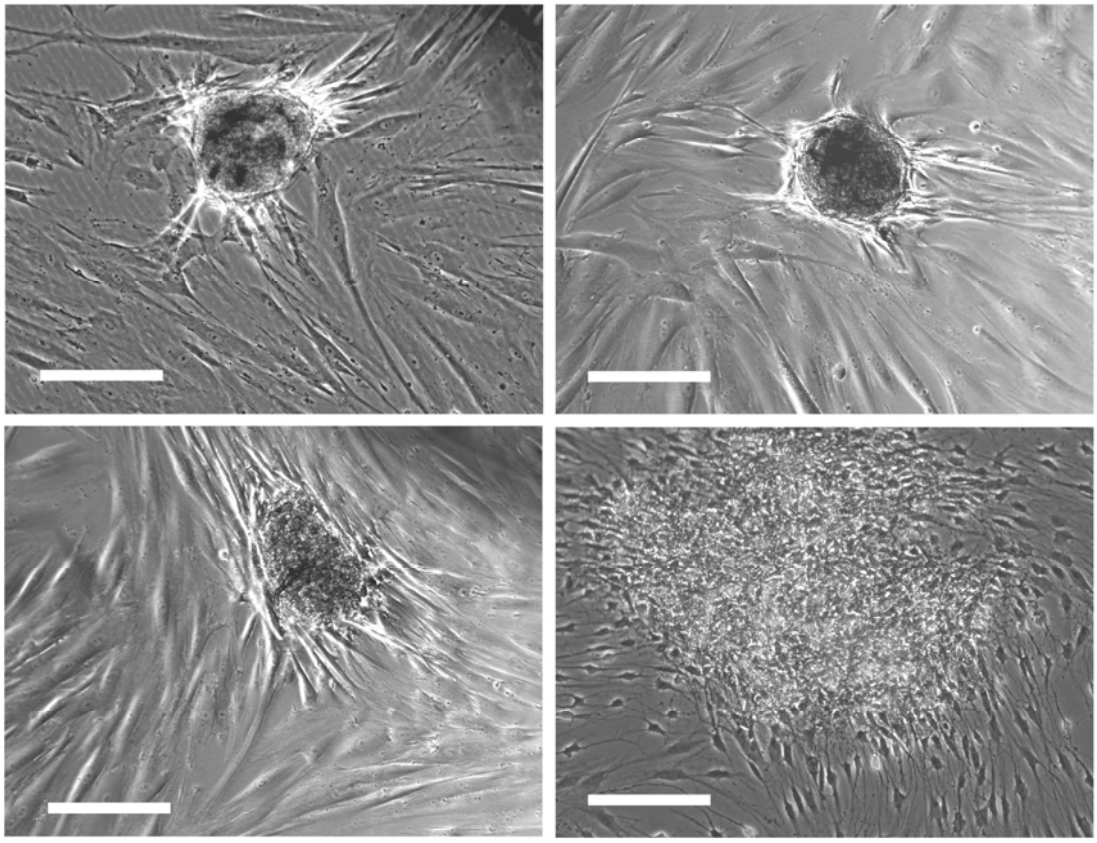

Figure 4-9 Microscopic pictures of colonies obtained using Protocol 4
Most of the colonies were compact and had well-defined borders (top panels) similar to ones observed in Protocol 3, while some colonies showed a morphology similar to that expected for human ES/iPS cells (bottom left image). Some of the colonies had a granular appearance as shown in bottom right image. Scale bar = 200μm.

4.3.8 Reprogramming Using Small Molecules

The slow kinetics and low efficiency of reprogramming have recently been augmented by the identification of various small molecules, which either act as chromatin modifiers or are known to regulate signalling pathways (Huangfu et al., 2008; Lin et al., 2009). One such molecule, valproic acid (VPA; a histone deacetylase inhibitor) has been demonstrated to significantly increase the efficiency of retroviral-mediated reprogramming (~50-fold) and has even lead to the generation of hiPS cells without the use of *KLF4* and *C-MYC* factors, albeit at lower efficiencies as compared to iPS generation in their presence (Huangfu et al., 2008). Similarly, treatment with ascorbic acid (commonly known as vitamin C) resulted in both an increase in the efficiency as well as the speed/pace of iPS generation (Esteban et al., 2010). Moreover, vitamin C was also shown to promote the conversion of pre-iPS cells (incomplete reprogrammed cells) to fully reprogrammed cells, a desirable feature for the incompletely reprogrammed iPS colonies generated with HSV-1 amplicon vectors. The combined use of VPA and ascorbic acid has been reported to increase reprogramming efficiency from 0.02% to ~6.2% with the use of retroviral vectors (Esteban et al., 2010). Therefore, for the next attempt on reprogramming, a method similar to Protocol 3 and Protocol 4 (Figure 4-7) was employed and transduced IMR90 fibroblasts were treated with valproic acid (VPA) and ascorbic acid. On the day after the first transduction with the pHG-TOKSM vector, 0.5mM VPA (Merck Biosciences) and 25µg/ml ascorbic acid (Sigma-Aldrich) were added to the hESc and mTeSR-1 medium. The VPA was removed after ten days from the first transduction, while the ascorbic acid treatment was continued until the end of the experiment. Similar to the colonies observed using Protocol 3 and Protocol 4, the colonies obtained by applying VPA and Vitamin C treatment failed to propagate further upon

transferring to feeder cells and none of them acquired the characteristics of bonafide iPS cells (data not shown).

4.4 Discussion

In this study, the HSV-1 amplicon system was utilised to deliver the reprogramming factors into the fibroblasts and assess the feasibility of this system in the derivation of iPS cells. The system offered a number of potential advantages over the above mentioned delivery methods: larger transgene capacity (~152 kb), higher safety and routinely obtained high titre preparations (Saeki et al., 2001). Thus, it seemed that it might prove useful for generating integration free iPS cells at higher efficiencies.

To investigate the potential of HSV-1 system in reprogramming the cellular identity of cells, three vectors were generated: the first vector (pHG-CAG SOKM), carrying the four reprogramming factors (*OCT4*, *SOX2*, *KLF4* and *C-MYC*) under control of a constitutive promoter; the second vector (pHG-TOKSM) with the same factors, controlled by an inducible promoter and the third vector (pHG-CAG SOKTM) with constitutive expression of three factors (*OCT4*, *SOX2* and *KLF4*) and inducible expression of *C-MYC*. The use of inducible promoters reduces the chance of continued expression of the reprogramming factors and permits the derivation of true iPS cells (completely reprogrammed), since the incompletely reprogrammed cells that rely on the transgene expression for proliferation are eliminated upon withdrawal of doxycycline (Maherali et al., 2008; Stadtfeld and Hochedlinger, 2010). In addition, the inducible systems also helps avoiding the differentiation problems associated with continuous expression of four factors (Maherali and Hochedlinger, 2008).

Initially, the vectors were delivered to fibroblasts using methods similar to those used for lentiviral/retroviral induction of pluripotent stem cells (Park et al., 2008b; Yu et al., 2007), but this did not result in successful generation of iPS cells. This was considered possibly due to the fact that the HSV-1 amplicon vector did not persist inside the cells for the entire duration of the reprogramming (between 20-30 days) (Park et al., 2008b; Takahashi et al., 2007; Yu et al., 2007). As the maintenance of the reprogramming vector in the target cells is an important factor in generating iPS cells, methods based on lentiviral/retroviral transduction, which are highly efficient in generating iPS cells, inherently allow the maintenance of transduced vector since the vector integrates into the host chromosomes. Unlike retroviral/lentiviral vectors, HSV-1 amplicon vectors are not known for integration into the target cells, therefore multiple transductions of the cells might be required to avoid dilution of reprogramming factors from the dividing cells. This is evident from past studies which showed that single transduction of non-integrating adenoviral vectors did not generate iPS cells, but multiple transductions led to successful derivation of iPS cells (Stadtfeld et al., 2008c; Zhou and Freed, 2009). In addition, the generation of iPS cells could possibly have been affected by silencing of the HSV-1 amplicon transgenes leading to inconsistent supply of the reprogramming factors. This idea has also been supported by many studies, which have established that the HSV-1 genome attains a transcriptionally repressed state in fibroblasts and de-repression can be achieved by superinfection with HSV or provision of ICP0 (an HSV transactivator) *in trans* (Coleman et al., 2008; Hobbs et al., 2001). Since the HSV-1 amplicon vectors do not contain any viral gene, except for the packaging signal and origin of replication, it is unclear if the silencing is the product of this pathway, or if it is due to cellular epigenetic phenomena.

To avoid the vector loss and possible silencing of HSV-1 amplicon vectors, various reprogramming attempts were made using multiple transductions and cycles of multiple transductions. These experiments led to the generation of many ball-shaped colonies, however, the colonies lacked the characteristics and growth features of typical hES/iPS colonies. Generation of pseudo-iPS colonies from multiple transductions of HSV-1 vectors could be associated with inconsistent supply of reprogramming factor as a result of the HSV-1 amplicon vectors not being able to transduce the same cells in the subsequent infections. Alternatively, it may be possible that the five day gap between transduction led to a decreased supply of the four reprogramming factors, taking it below the amount sufficient to reprogram the cellular identity. To test this hypothesis, further experiments were carried out in which the cells were transduced with a gap of three and four days. These experiments, however, did not result in any colony. Moreover, the viability of the cells was impaired during the course of these experiments. This might have been caused by the frequent infections, which did not allow sufficient time for the cells to recover. In addition to the challenges associated with vector loss during cell divisions and silencing of the HSV-1 amplicon vectors, other factors such as inconsistent provision of reprogramming factors, and the loss of the cells during their passaging and reseeding in the hESc/mTeSR-1 medium could also be responsible for unsuccessful reprogramming. A method that avoids passaging cells after the DMEM is switched to hESc or mTeSR-1 medium could therefore assist in successful derivation of iPS cells. One possible way to achieve this objective is by modifying the HSV-1 amplicon constructs to incorporate episomal retention elements such as EBNA-1/OriP elements of the Epstein-Barr virus (Yates et al., 1984; Yates et al., 1985). The vectors containing the EBNA-1/OriP elements are maintained as stable episomes in the cells and are known to replicate once per cell division (Yates and Guan, 1991). These features make them a

promising tool to avoid vector loss during the cell division, thus making unnecessary the multiple cycles of transductions. Hence, this approach would also reduce the cell loss associated with passaging during multiple treatments. The use of EBNA-1/OriP elements in the HSV-1 vectors and their feasibility in reprogramming will be further discussed in Chapter 5.

Among other variables that can influence the outcome of reprogramming experiments, the choice of cell type is also vital (Maherali and Hochedlinger, 2008). Fibroblasts have been previously used for successful nuclear transfer-based and cell fusion-based reprogramming (Cowan et al., 2005; Tada et al., 2001; Wakayama et al., 1998). Moreover, fibroblasts are derived in a technically simple manner and can also be obtained from cell banks, making them an ideal choice of cell type for reprogramming experiments. However, recent studies have suggested that although fibroblasts such as IMR90 cells have been successfully used in iPS generation through various methods (Jia et al., 2010; Yu et al., 2007), their reprogramming efficiency is low as compared to other donor cell types such as adipose tissue derived stem cells, keratinocytes and embryonic fibroblasts (Aasen et al., 2008; Jia et al., 2010; Maherali et al., 2008; Park et al., 2008b; Yu et al., 2007; Zhou and Freed, 2009). In order to improve the reprogramming strategy, the subsequent iPS generation experiments were conducted on embryonic fibroblasts, which can be derived from hESc differentiation and therefore can be obtained at low passages. The embryonic fibroblasts were used along with modified HSV-1 amplicon vectors, which incorporated the EBNA-1/OriP elements to successfully generate the iPS cells. The experiment will be described in Chapter 5.

Chapter 5. Generation of Human Induced Pluripotent Stem Cells with Hybrid HSV-1 EBNA-1/OriP Vectors

5.1 Introduction

While the experiments presented in the last chapter were being performed, a study describing the generation of vector and transgene-free iPS cells was published (Yu et al., 2007). In the study, Yu and colleagues utilised the EBNA-1/OriP based-episomal vectors to deliver the reprogramming factors into the fibroblasts and successfully derived integration-free iPS cells. Compared to other non-integrating systems, using EBNA-1/OriP vectors for the generation of iPS cells has several advantages. (1) The vectors are maintained extra-chromosomally through the function of the *EBNA-1* gene and *OriP* elements and are not known to integrate into the host genome (Yates et al., 1984; Yates et al., 1985). (2) The vectors are known to replicate once during each cell cycle and are normally maintained in the cells at low copy number, therefore reducing the risk of DNA rearrangements (Yates and Guan, 1991). (3) The vectors can be removed from the iPS cells because of the silencing of the viral promoter-driven *EBNA-1* gene and therefore vector and transgene-free iPS cells can be obtained (Yu et al., 2009). However, the reprogramming efficiency of these EBNA-1/OriP vectors was found to be low (0.0003% for the foetal foreskin fibroblasts). The low efficiency of the system can be attributed to the gradual loss (5% per cell division) of the EBNA-1/OriP vectors (Nanbo et al., 2007), methylation-mediated silencing of the episomal vectors (Kameda et al., 2006) and the use of the Nucleofection system for delivering the constructs. Nucleofection like

electroporation and transfection, suffers from low transfection efficiency with increasing vector size (Strulovici et al., 2007), especially when two or three vectors are simultaneously nucleofected into the cells as was required for the iPS generation (Yu et al., 2007).

The low delivery efficiency of the Nucleofection system for larger vectors can be addressed by the utilisation of the HSV-1 amplicon system, which possess the ability to efficiently deliver large vectors into a wide variety of cells types (Saeki et al., 2003). However, the HSV-1 amplicon vectors suffer from rapid vector loss during cell division and transient gene expression as a result of silencing (Johnston et al., 1997). Due to these limitations the HSV-1 amplicon vectors (without containing any retention/integration element) could not be utilised for the generation of iPS cells as described in the previous chapter. To overcome the rapid vector loss and possibly extend the transgene expression from HSV-1 amplicon vectors as well as the delivery limitation of nucleofection, the HSV-1 amplicon vectors were modified to incorporate the EBNA-1/OriP retention elements and generate hybrid HSV-1/EBNA-1/OriP vectors. The hybrid system combines the advantages of both the HSV-1 amplicon system for the efficient delivery of large DNA and of the EBNA-1/OriP episomal vectors for the stable maintenance and extended transgene expression. A study conducted by Sia and colleagues also established that the hybrid HSV-1/EBNA-1/OriP amplicon system, in comparison to the HSV-1 amplicon system provided prolonged transgene expression in primary human tumour cells and human bone marrow-derived mesenchymal stem cells (Sia et al., 2010).

In the present study, the hybrid HSV-1/EBNA-1/OriP vectors were constructed containing the reprogramming genes (*OCT4*, *SOX2*, *KLF4*, *C-MYC*, *NANOG* and *LIN28*) and the SV40 large T- antigen (*SV40LT*) gene. The two additional reprogramming genes

NANOG and *LIN28* were included in the cocktail as the combined use of six factors has shown to increase the iPS derivation efficiency as compared to the use of four genes (Yu et al., 2009). The use of *SV40LT* has also been demonstrated to facilitate the generation of iPS cells and hence was also included along with the six reprogramming factors (Yu et al., 2011). The modified HSV-1 amplicon hybrid vectors were subsequently used to generate iPS cells from embryonic fibroblast derived from hESc.

The retroviral iPS generation system (Takahashi et al., 2007), originally developed in Shinya Yamanaka's laboratory, was also employed as a control for the reprogramming experiment. The iPS cells obtained using both the hybrid HSV-1/EBNA-1/OriP system and the retroviral systems were similar to the hESc in morphology, expression of pluripotency markers and differentiation capabilities.

5.2 Specific Methods

5.2.1 Construction of HSV-1 Amplicon Vectors

The EBNA-1/OriP-based reprogramming vectors pEP4EO2SCK2MEN2L (containing *OCT4*, *SOX2*, *KLF4*, *C-MYC*, *NANOG* and *LIN28*) and pEP4EO2SET2K (containing *OCT4*, *SOX2*, *SV40LT* and *KLF4*) (Figure 5-1; (Yu et al., 2009)) were obtained from Addgene. The HSV-1 amplicon vectors pHG puro 6F and pHG puro 3FLT (Figure 5-2) were constructed by retrofitting the entire the pHG-Puro vector (Figure 2-1 C) into the NheI site of the pEP4EO2SCK2MEN2L and pEP4EO2SET2K vectors respectively, taking advantage of the fact that the *R6Kγ* origin of replication is not active in the absence of the pir protein.

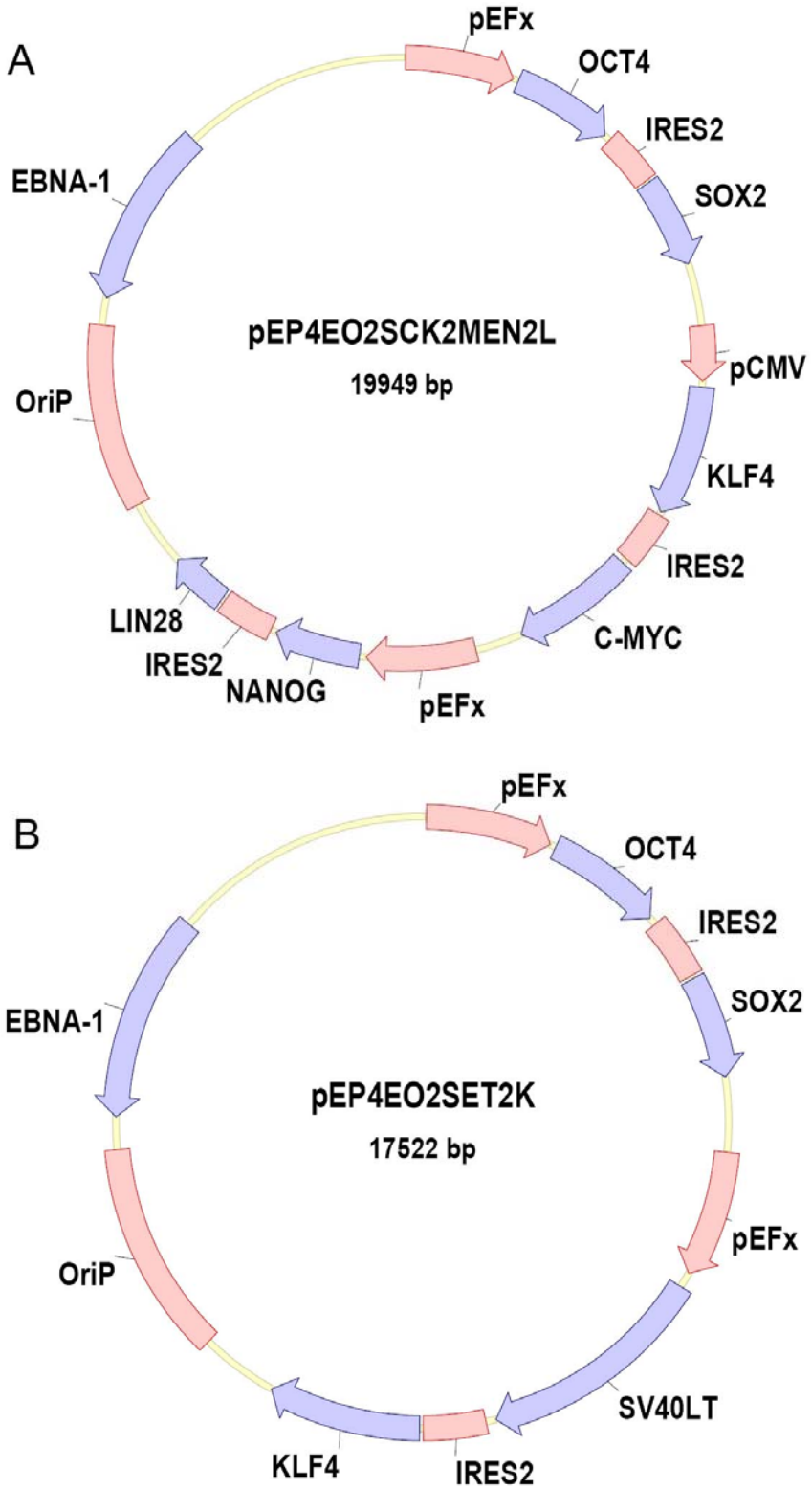

Figure 5-1 Vector maps of pEP4EO2SCK2MEN2L and pEP4EO2SET2K
pEP4EO2SCK2MEN2L (A) is an EBNA-1/OriP episomal vector containing the genes *OCT4*, *SOX2*, *KLF4*, *C-MYC*, *NANOG* and *LIN28*. pEP4EO2SET2K (B) is another EBNA-1/OriP episomal vector containing the genes *OCT4*, *SOX2*, *SV40LT* and *KLF4*. Both vectors were obtained from Addgene.

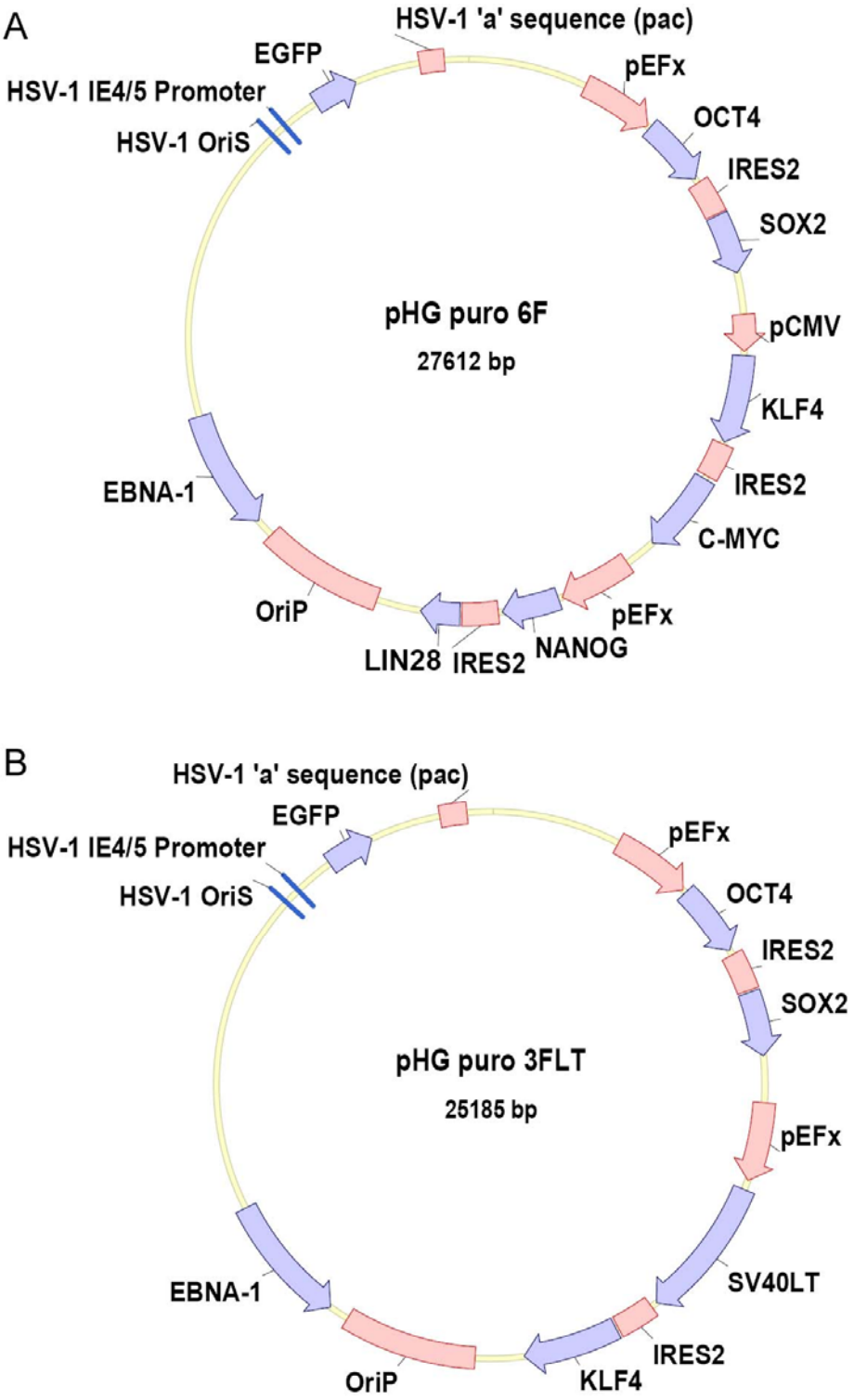

Figure 5-2 Vector maps of pHG puro 6F and pHG puro 3FLT
The vectors pHG puro 6F (A) and pHG puro 3FLT (B) were constructed to retrofit the entire pHG-Puro vector into the pEP4EO2SCK2MEN2L and pEP4EO2SET2K vectors respectively.

5.2.2 Derivation of Embryonic Fibroblasts from HUES10 Cells

Embryonic fibroblasts were derived from HUES10 cells through embryoid body (EB)-mediated differentiation. The EBs were plated onto ultra-low adherent 6-well dishes (Costar, Corning) in fibroblast medium consisting of DMEM-F/12 supplemented with 20% (v/v) FBS, 0.1mM non-essential amino acids, 1X Glutamax^{TM}-1, 0.1mM ß-mercaptoethanol, and 1% penicillin and streptomycin (P/S). After five days in suspension culture, the EBs were transferred to gelatin-coated 6-well dishes in the fibroblast medium and cultured for additional nine to ten days. The outgrowth of the attached EBs were dissociated through trypsinisation and plated onto new gelatin-coated dishes. The cells were passaged four times in total, until the cells homogenously acquired fibroblast-like morphology.

5.2.3 Derivation of iPS Cells Using Retroviral Vectors

For the preparation of retroviruses, the pMXs retroviral expression vectors (Takahashi et al., 2007) containing the human reprogramming factors (*OCT4*, *SOX2*, *NANOG*, *KLF4* and *C-MYC*) were obtained from Addgene. Packaging of the retroviral vectors was carried out as described in the published methodology by Shinya Yamanaka's laboratory (Ohnuki et al., 2009; Takahashi et al., 2007) with two alterations: (1) the transfection of the retroviral vectors into the PLAT-E cells was carried out using polyethylenimine (PEI) instead of Fugene 6; and (2) stocks of viral preparation were subjected to ultracentrifugation for 100x concentration. The day before the retroviral transduction, HUES10-derived embryonic fibroblasts were plated at a density of $2x10^5$ cells/well of a 6-well dish. The cells were infected with the concentrated viral stocks in the DMEM medium containing 5µg/ml polybrene (Sigma-Aldrich) and the cells incubated at 37ºC and 5% CO_2 overnight. After 24 hours, the infection of the cells was repeated once more

and the cells were incubated at 37ºC and 5% CO_2 for three days. Both the production of retroviruses and the infection of cells were performed by Jane Vowles at the Sir William Dunn School of Pathology, University of Oxford. Three days after the second infection, the cells were trypsinised and transferred onto inactivated MEF feeder cells at a density of $8x10^4$ cells/well of a 6-well dish in the DMEM medium. The next day, the DMEM medium was replaced with hESc medium containing 0.5mM valproic acid and 50µg/ml ascorbic acid and changed every alternate day thereafter. After eight days of cultivation on feeder cells, the hESc medium was switched with conditioned hESc medium containing the same amount of chemicals. Valproic acid was withdrawn from the medium after twelve days of cultivation on feeder cells. However, addition of ascorbic acid was continued until the generation of iPS colonies.

5.2.4 Subcloning of the iPS Cells

To isolate subclones from iPS cells derived using the HSV-1/EBNA-1/OriP vectors, the cells were trypsinised using TrypLE Express as described in the Section 2.11.3. The dissociated cells were plated at very low seeding density onto iMEF feeder cells in hESc medium supplemented with 10µM ROCK inhibitor Y-27632. Individual colonies were allowed to develop and were subsequently picked and expanded as described in the Section 2.17.4.

5.2.5 PCR and RT-PCR Analysis of the iPS Clones

Screening for the vector removal from the iPS cells was performed by PCR analysis on genomic DNA extracted from the iPS clones using the primers specific for the transgenes (Tg-*OCT4*, Tg-*SOX2*, Tg-*NANOG*, and Tg-*SV40LT*). To evaluate the expression of the transgenes and endogenous pluripotency genes, total RNA was extracted from the iPS

cells and was subjected to cDNA synthesis. The cDNA was then utilised for RT-PCR analysis using the primers specific for the transgenes (Tg-*OCT4*, Tg-*NANOG* and Tg-*SV40LT*), endogenous *OCT4* (Endo-*OCT4*) and *SOX2* (Endo-*SOX2*), and *NANOG* (Total-*NANOG*). Primers specific for the endogenous genes are located at 3′ untranslated region of the gene and do not amplify the transgenes. The sequence of the primers used for the PCR and RT-PCR analysis is shown in Table 5-1.

Table 5-1 Primer sets used for PCR and RT-PCR analysis

Genes/Region	Size	Primer Name	Sequence (5'- 3')
Tg-*OCT4*	657bp	Oct4-SF1	AGTGAGAGGCAACCTGGAGA
		IRES2-SR	AGGAACTGCTTCCTTCACGA
Tg-*SOX2*	498bp	IRES2-SF	TGGCTCTCCTCAAGCGTATT
		Sox2-SR	GCTTAGCCTCGTCGATGAAC
Tg-*NANOG*	732bp	Nanog-SF2	CAGAAGGCCTCAGCACCTAC
		IRES2-SR	AGGAACTGCTTCCTTCACGA
Tg-*SV40LT*	491bp	SV40T-SF1	TGGGGAGAAGAACATGGAAG
		IRES2-SR	AGGAACTGCTTCCTTCACGA
Endo-*OCT4*	113bp	Oct4-Endo F	AGTTTGTGCCAGGGTTTTTG
		Oct4-Endo R	ACTTCACCTTCCCTCCAACC
Endo-*SOX2*	151bp	Sox2-Endo F	CATCACCCACAGCAAATGAC
		Sox2-Endo R	CCTCCCCAGGTTTTCTCTGT
Total-*NANOG*	214bp	KLF4 P3 F	TTCCTTCCTCCATGGATCTG
		KLF4 P3 R	ATCTGCTGGAGGCTGAGGTA
GAPDH	202bp	GAPDH F	TGTTGCCATCAATGACCCCTT
		GAPDH R	CTCCACGACGTACTCAGCG

Primers for Tg-*OCT4*, Tg-*SOX2*, Tg-*NANOG*, and Tg-*SV40 LT* were obtained from (Yu et al., 2009).

5.3 Results

5.3.1 Choice of HUES10 derived Embryonic Fibroblasts

Based on the unsuccessful attempts at reprogramming IMR90 fibroblasts in Chapter 4, it was speculated that reprogramming of IMR90 fibroblast could be challenging because they represented a developmentally mature somatic status. Therefore other somatic cell types at primitive developmental stages might be more amenable to reprogramming. Various studies have reported that the choice of cell type influence both the efficiency and the pace of reprogramming. For example, reprogramming of human keratinocytes is faster and more efficient than that of fibroblasts (Aasen et al., 2008; Maherali et al., 2008). Similarly, the reprogramming efficiency of minicircle vectors was shown to be ten folds higher in human adipose stem cells compared to that in IMR90 fibroblasts (Jia et al., 2010). Other studies have demonstrated that fibroblast cells derived through *in vitro* differentiation of hESc can be reprogrammed at relatively high efficiencies as compared to developmentally advanced cell types such as foreskin fibroblasts and adult dermal fibroblasts (Chan et al., 2009; Hockemeyer et al., 2008; Maherali et al., 2008; Park et al., 2008b). As hESc differentiated embryonic fibroblast can be derived in a relatively simple manner through EB-mediated differentiation and can be obtained readily at low passages, they were chosen as the target cell type for the subsequent reprogramming experiments.

For the differentiation of hESc, EBs were derived from HUES10 cells (Figure 5-3) and were propagated in fibroblast growth conditions as described in Section 5.2.2 (Xu et al., 2004). As hESc cannot be grown as single cells on normal culture dishes, the outgrowths of the EBs (Figure 5-3 B and C) containing the fibroblast-like cells were harvested and were passaged at least four times using trypsin to eliminate undifferentiated hESc. The

passaging also served in acquiring a homogenous population of fibroblast cells (Figure 5-3 D) as compared to the initial heterogeneous population of cells that were released in the outgrowths of EBs. The embryonic fibroblasts derived from the HUES10 cells exhibited the morphology typical of normal fibroblasts (Figure 5-3 D). In addition, RT-PCR analysis of these fibroblasts demonstrated that the expression of the key pluripotency genes (*OCT4*, *NANOG* and *SOX2*) was completely lost, as shown in Figure 5-4 B.

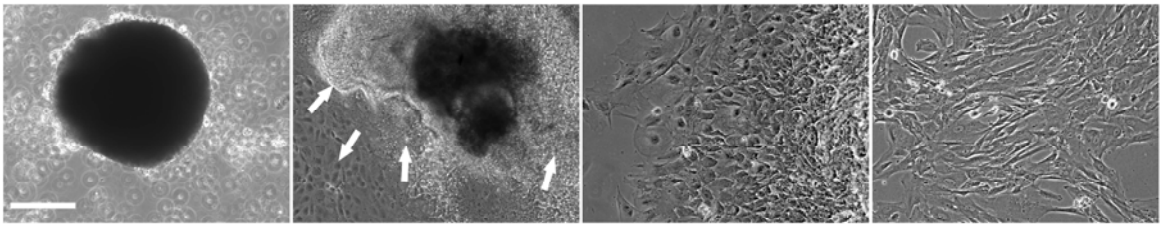

Figure 5-3 Time course of human embryonic fibroblasts derivation from HUES10 cells

(A) HUES10 embryonic stem cells derived embryoid bodies (EBs) cultivated in suspension culture. (B) Outgrowths from attached EBs comprised of highly heterogeneous population of the cells around the central portion of EBs (identified by white arrows) detected after 10 days of plating onto gelatin-coated dishes. (C) A less heterogeneous population towards the edge of the EBs. (D) Homogenous population of fibroblast-like cells, visible after the outgrowths were sub-cultured at least four times during a period of 13-15 days. Scale bar = 200μm.

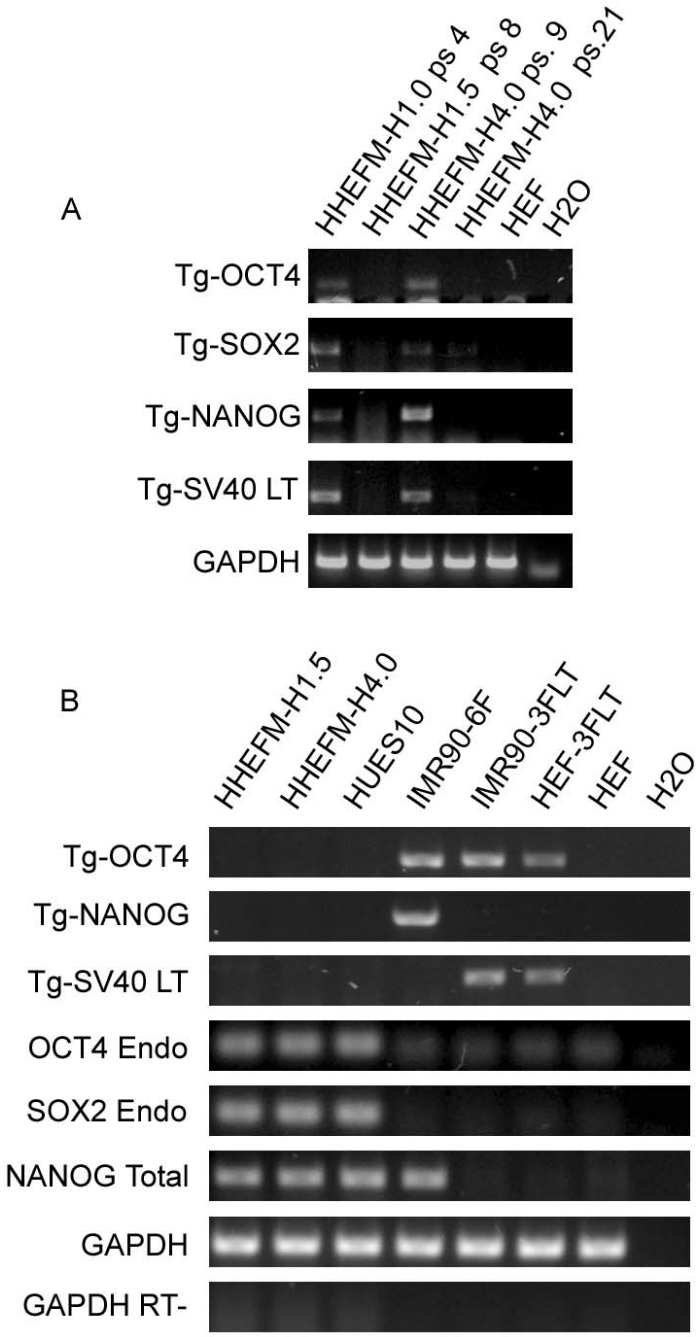

Figure 5-4 Analysis of the vector removal from the iPS clones
(A) PCR analysis on genomic DNA extracted from the iPS cells using the primers sets specific for transgenes. The transgene sequences were not detected in the subclone HHEFM-H1.5 and were detected at a very low amount in the clone HHEFM-H4.0 (passage 21) in contrast to the parental clones HHEFM-H1.0 and low passage HHEFM-H4.0 (passage 9). (B) RT-PCR analysis showing that the clones HHEFM-H1.5 and HHEFM-H4.0 did not express the transgenes. Moreover, endogenous *OCT4*, *SOX2* and *NANOG* genes were expressed indicating the reactivation of pluripotency network.
IMR90 cells transduced with pHG puro 6F (IMR90-6F) and pHG puro 3FLT (IMR90-3FLT) along with HUES10 derived embryonic fibroblasts (HEF) transduced pHG puro 3FLT were used as control. *GAPDH* gene was used as an internal control. PCR of the RNA without reverse transcriptase (-RT) was done with the *GAPDH* primer set to ensure no amplification of genomic DNA. (H_20) = no template control.

5.3.2 Generation of iPS Cells Using the HSV-1 Amplicon System

The method utilised here for the derivation of iPS cells was based on the hybrid HSV-1/EBNA-1/OriP vectors containing the reprogramming factors *OCT4*, *SOX2*, *KLF4*, *C-MYC*, *NANOG* and *LIN28* as well as SV40 large T antigen (*SV40LT*). The EBNA-1/OriP-based episomal vectors pEP4EO2SCK2MEN2L and pEP4EO2SET2K (Figure 5-1) containing the above mentioned factors were initially developed by Yu and colleagues and were successfully used for generating footprint-free iPS cells from foetal foreskin fibroblasts (Yu et al., 2009). To overcome the possible loss of HSV-1 amplicon vectors from the dividing fibroblasts during the reprogramming process, the EBNA-1/OriP-based episomal vector system was extended to the HSV-1 amplicon system. As described in Section 5.2.1, two HSV-1/EBNA-1/OriP vectors pHG puro 6F and pHG puro 3FLT were constructed from the pEP4EO2SCK2MEN2L and pEP4EO2SET2K vectors respectively and tested for their ability to reprogram HUES10-derived embryonic fibroblasts.

The reprogramming strategy used for the generation of iPS cells is presented in Figure 5-5. The method involved the transduction of $1x10^5$ fibroblasts (seeded on 24-well dish) with pHG puro 6F and pHG puro 3FLT amplicons vectors, each at the MOI 10. The next day, the transduction efficiency was assessed by monitoring GFP expression under the fluorescence microscope (Figure 5-6 A and B). Subsequently $5x10^4$ transduced fibroblasts were transferred to inactivated MEF feeder cells in a well of a 6-well dish. The next day, the fibroblasts medium was replaced with hESc medium (containing 20ng/ml bFGF) and the medium was replenished every other day. Ten days after the transduction, the hESc medium was replaced with hESc-conditioned medium (containing 20ng/ml bFGF) to support the deteriorated MEF feeder cells and replenished every alternate day until the end of the experiment. Using this strategy, two colonies were derived after 14 days from

the transduction (Figure 5-6). The reprogramming experiment was continued until day 25 to allow the expansion of the colonies. However, only one out of the two colonies grew and exhibited the morphology of a typical hES/hiPS colony, leading to a reprogramming efficiency of 0.002% of the input cells. The colony (called HHEFM-H1.0) was picked on iMEF feeder cells in a 24-well dish for further expansion in the hESc medium. The entire experiment of iPS generation was repeated for a second time, and it also resulted in the derivation of a single iPS colony (HHEFM-H2.0) thus confirming the reproducibility of the method.

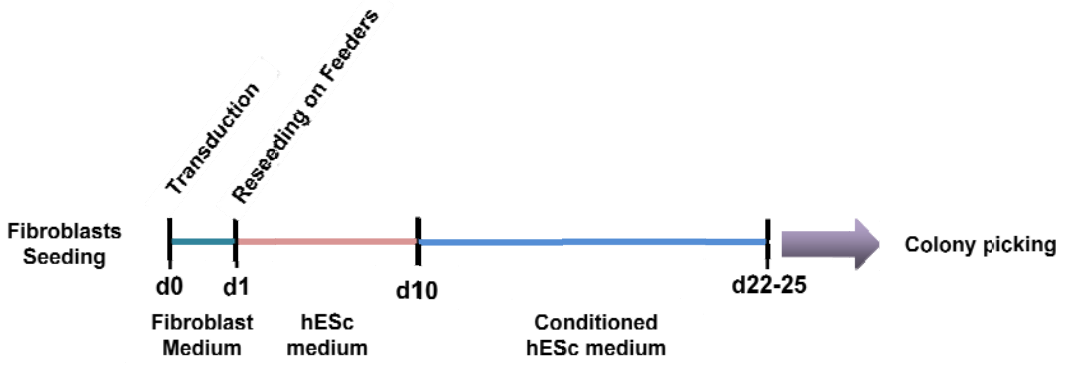

Figure 5-5 Schematic representation of the HSV-1/EBNA-1/OriP reprogramming procedure
In this protocol, fibroblasts were transduced with pHG puro 6F and pHG puro 3FLT amplicons. Next day, the transduced cells were transferred to iMEF feeder cells. A day later the fibroblasts medium was substituted with hESc medium and thereafter, hESc medium was replenished every alternate day. After ten days of transduction, hESc medium was replaced with MEF-conditioned hESc medium and changed every other day until the end of experiment (twenty five to thirty days after transduction).

Recently, many studies have indicated that the addition of small molecules significantly improves the reprogramming efficiency. Although there is a variety of chemicals that have been tested both individually and in combinations with other chemicals, the four chemicals SB431542 (TGF-β1/activin receptor inhibitor), PD0325901 (MEK inhibitor), CHIR99021 (GSK3β inhibitor) and Y-27632 (ROCK inhibitor) have been shown to

significantly enhance the reprogramming efficiency (Hanna et al., 2010; Lin et al., 2009; Shimada et al., 2011; Yu et al., 2011; Zhang et al., 2011). In the next attempt at reprogramming, a similar method was used as described in Figure 5-5 but in order to improve the efficiency of iPS derivation, hESc and conditioned-hESc media were supplemented with 0.5µM SB431542, 0.5µM PD0325901, 3µM CHIR99021 and 10 µM Y-27632. These chemicals were applied on day 2 after the transduction and continued until day 12. Following this strategy, two iPS colonies were obtained from $5x10^4$ fibroblasts leading to an efficiency of 0.004% of the input cells. The two colonies HHEFM-H3.0 and HHEFM-H4.0 were individually picked and expanded as described above. The morphology and the growth features of the two iPS colonies HHEFM-H1.0 and HHEFM-H4.0 were very similar to those of hESc and hiPS cells. Therefore the two clones were transferred to Matrigel/mTeSR-1 culture conditions (Figure 5-6 F-I) and subjected to cellular and molecular analysis to evaluate their pluripotency and differentiation capabilities. The total number of iPS colonies obtained using the HSV-1/EBNA-1/OriP system along with a comparison to retroviral induced iPS colonies are listed in Table 5-2.

In parallel control experiments, no iPS colony was observed, when non-transduced embryonic fibroblasts (negative control) were subjected to the same iPS derivation conditions (both without and with chemicals). Moreover, no iPS colony was derived when only one of the HSV-1 amplicon vectors (either pHG puro 6F or pHG puro 3FLT) was utilised.

Next, the reprogramming potential of hybrid HSV-1/EBNA-1/OriP vectors was assessed in IMR90 fibroblasts. In two successive attempts, pHG puro 6F and pHG puro 3FLT transduced IMR90 fibroblasts were subjected to the derivation method that successfully

generated iPS cells from the embryonic fibroblasts derived from HUES10 cells. However the experiment did not lead to the derivation of any iPS colony from the IMR90 fibroblasts indicating the need of further refinements in the system to reprogram developmentally advanced cells like IMR90 fibroblasts.

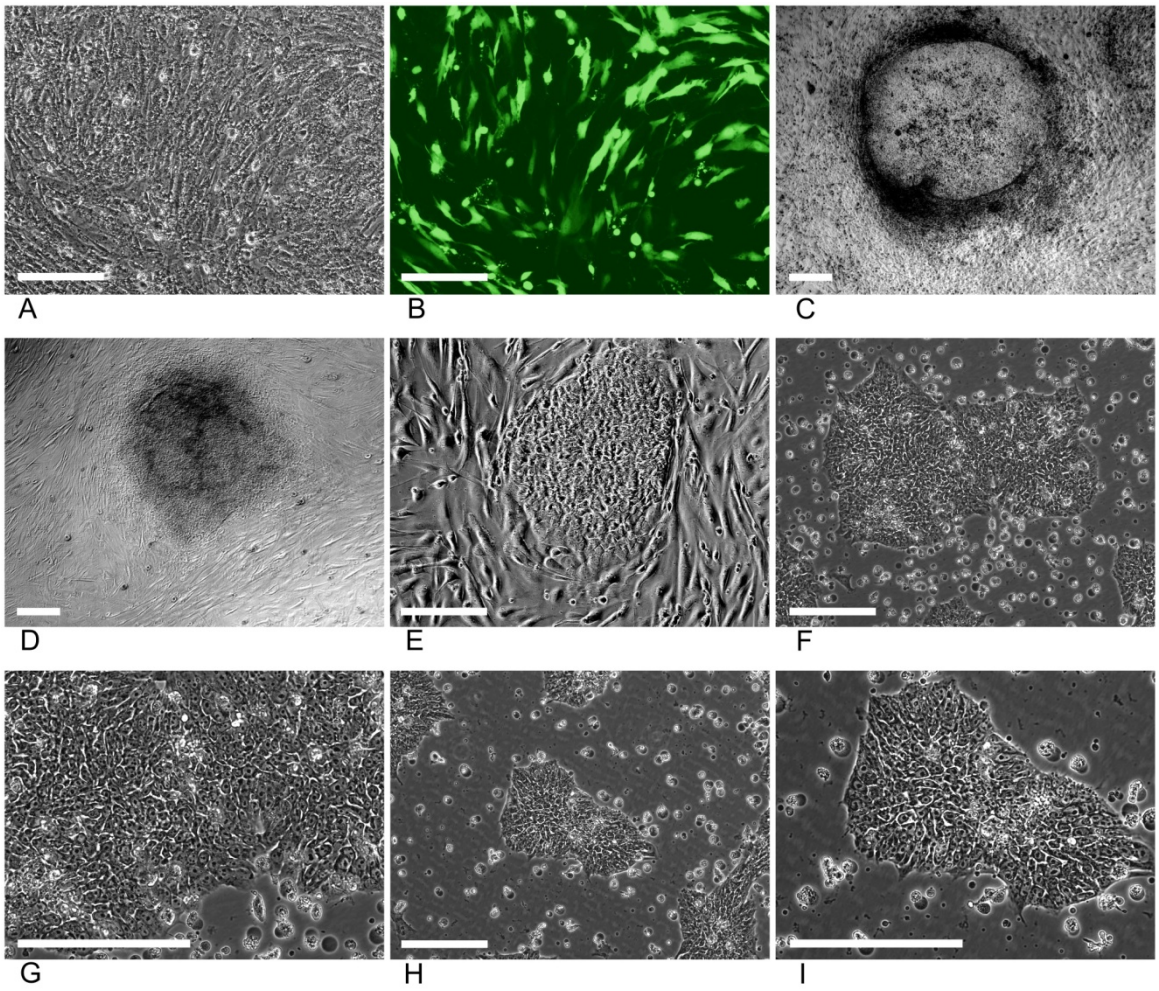

Figure 5-6 iPS generation using HSV-1 amplicon system
HUES10-derived embryonic fibroblasts transduced with pHG puro 6F and pHG puro 3FLT amplicons. (A) Phase contrast image of transduced fibroblasts 24 hours post-transduction. (B) GFP expression in the transduced fibroblasts (corresponding field as shown in A). (C and D) shows typical images of two different iPS colonies derived from fibroblasts on MEF feeder cells. (E) Morphology of a single iPS colony at passage 1 on MEF feeder cells, at higher magnification. (F) Morphology of established HHEFM-H1.5 (subclone of HHEFM-H1.0) iPS cells at passage 25 in Matrigel/mTeSR-1 culture. (G) Higher magnification (20x) of image F showing large nuclei and scant cytoplasm in the iPS cells, which are typical morphological characteristics of hES/hiPS cells. (H) Morphology of established HHEFM-H4.0 iPS cells at passage 15 in Matrigel/mTeSR-1 culture. (I) Higher magnification (20x) of the image shown in H.

Table 5-2 Comparison of reprogramming efficiencies of HSV-1/EBNA-1/OriP and retroviral system

Reprogramming System	Total iPS colonies (Efficiency)	Total iPS colonies with the use of 4C (Efficiency)
HSV-1/EBNA-1/OriP	1 (0.002%)	2 (0.004%)
Retroviral	2 (0.0025%)	3 (0.0037%)

The reprogramming efficiency was determined on $5x10^4$ cells (for HSV-1/EBNA-1/OriP system) and $8x10^4$ cells (for retroviral system). 4C refers to four chemicals (SB431542/PD0325901/CHIR99021/Y-27632).

5.3.3 Analysis of Vector Removal

Since, as described in Section 5.1, the EBNA-1/OriP-based vectors are gradually lost from the dividing cells, the iPS clone HHEFM-H1.0 was subcloned to identify an iPS clone from which the vectors were completely absent. The subclones were analysed for the loss of vector DNA through PCR analysis on genomic DNA for *OCT4*, *SOX2*, *NANOG* and *SV40LT* transgenes. One of the subclones, HHEFM-H1.5, was found to have completely lost the vector at passage 8, as it was negative for all the transgenes in contrast to its parental clone (Figure 5-4 A). In an alternate strategy, the iPS clone HHEFM-H4.0 was passaged several times to ensure the complete removal of the vectors. As shown in Figure 5-4 A, in contrast to the HHEFM-H4.0 at earlier passage (passage 9), the transgene sequences were either not detected or detected at a very low amount at a later stage (passage 21). The overall analysis confirmed that the iPS cells, obtained using the HSV-1 amplicon system, were derived in an integration-free manner. Moreover, it demonstrated that either subcloning or extensive passaging can be used to derive foot-print free iPS cells that are largely free of vector and transgene sequences.

Subsequently, RT-PCR analysis of the iPS clones HHEFM-H1.5 and HHEFM-H4.0 for the *OCT4*, *NANOG* and *SV40LT* transgenes was performed to investigate whether there

was any residual transgene expression in the clones (Figure 5-4 B) using transgene specific primers. The analysis demonstrated that the *OCT4*, *NANOG* and *SV40LT* transgenes were not expressed from the iPS cells. Moreover, using the primers specific for the endogenous *OCT4* and *SOX2* genes, and the total *NANOG* gene, the RT-PCR analysis revealed that in comparison to the parental fibroblast, the endogenous *OCT4*, *SOX2* and *NANOG* were reactivated in the iPS clones, further validating the activation of their pluripotency genes. Although the primers specific for the total *NANOG* (endogenous and transgene) were used, the lack of expression in Tg-*NANOG* indicated that expression of the total *NANOG* was based on the endogenous expression.

5.3.4 Immunofluorescence Staining for Pluripotency Markers

To demonstrate that the iPS cells generated through the HSV-1/EBNA-1/OriP amplicon system were pluripotent, immunofluorescence staining for hESc-specific pluripotency markers was performed on the clones HHEFM-H1.5 and HHEFM-H4.0. Cells from these clones were stained with antibodies against the nuclear proteins OCT4 and NANOG and the surface stem cell marker TRA-1-60 along with HUES10 cells as control. Both HSV-1 amplicon-derived iPS clones were uniformly stained by the antibodies for OCT4, NANOG and TRA-1-60 with a similar intensity as HUES10 cells, as shown in Figure 5-7.

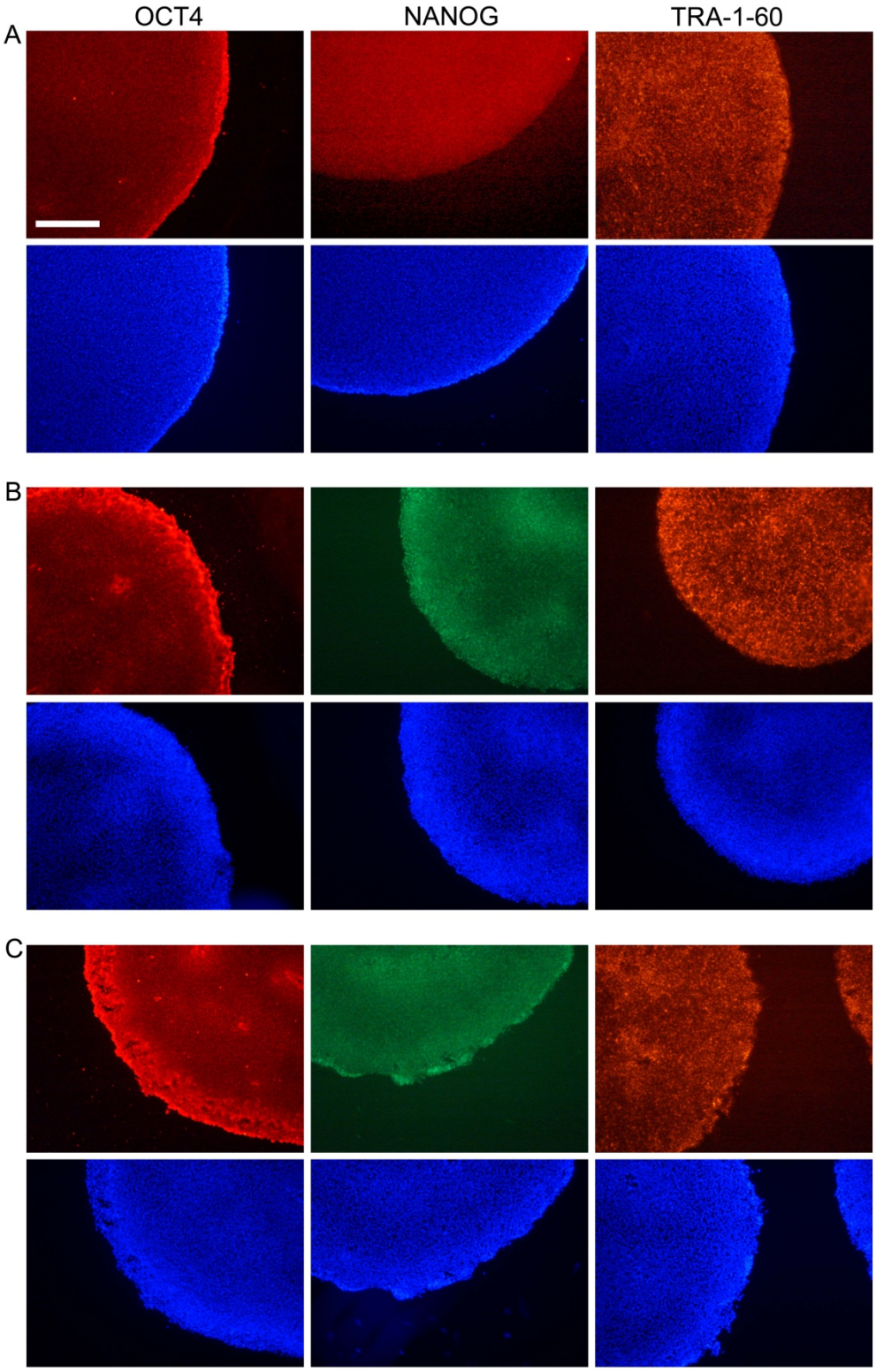

Figure 5-7 Immunofluorescence staining of HSV-1 amplicon-derived iPS colonies for pluripotency markers

The iPS clones HHEFM-H1.5 (A) and HHEFM-H4.0 (B) derived using HSV-1 amplicon system were positive for hESc pluripotency markers OCT4 (red signal), NANOG (red signal for HHEFM-H1.5 clone and green signal for HHEFM-H4.0 clone) and TRA-1-60 (red signal). HUES10 cells (passage 26) were used as control (C). Nuclei of the cells were counterstained with DAPI (blue signal). Scale bar = 200μm.

5.3.5 Generation of iPS Cells with Retroviral Vectors

HUES10-derived embryonic fibroblasts were also reprogrammed using the retroviral system, originally developed by Shinya Yamanaka's group (Takahashi et al., 2007) as a control for the HSV-1 amplicon-mediated iPS generation. The experiment was performed with the help of Dr. Sally Cowley and Jane Vowles and was based on the methodology modified from Shinya Yamanaka's protocol and established in Prof. William James laboratory, University of Oxford. Details of the procedure are described in Section 5.2.3. The reprogramming strategy is presented in Figure 5-8. Briefly $2x10^6$ fibroblasts were transduced twice with the pMXs retroviral vectors on successive days. As the pMXs retroviral vectors lack the *GFP* gene, the transduction efficiency was not determined and hence cannot be correlated with HSV-1 amplicon transduction of the fibroblasts. Four days after the first transduction, $8x10^4$ of the total transduced cells were transferred to MEF feeder cells in fibroblast medium. A day later, fibroblast medium was switched with hESc medium (containing 20ng/ml bFGF) and the medium was replenished every other day. Eight days after the cells were transferred to MEF feeder cells, hESc medium was substituted with MEF-conditioned medium and changed every other day thereafter. The media were also supplemented with valproic acid (VPA) and ascorbic acid to evaluate the effect of these chemicals on the efficiency of iPS generation as described in Section 5.2.3. About 10 days after the first transduction of fibroblast, some small colonies were observed which later increased in size. However, these colonies lacked the typical morphology of a hES/hiPS colony and exhibited irregular borders and/or granular appearance as shown in Figure 5-9 C. Between 12-15 days, some additional colonies were obtained which closely resembled hES/hiPS colonies in terms of well-defined borders, compact appearance and high nucleus to cytoplasm ratio as shown in Figure 5-9 A and B. Overall, ~2 hES/hiPS like colonies and ~7 irregular-shaped colonies (non

hES/hiPS colonies) were obtained from three independent wells of a 6-well-dish (Table 5-2). In a parallel experiment, the hESc and conditioned-hESc media were supplemented with the four chemicals SB431542, PD0325901, CHIR99021 and Y-27632 instead of the VPA and ascorbic acid. This treatment resulted in ~3 hES/hiPS like colonies and ~10 non hES/iPS-like colonies from two independent wells of a 6-well-dish (Table 5-2). The efficiency of iPS generation was determined to be 0.0025% with VPA/ascorbic acid treatment and 0.0037% with SB431542/PD0325901/CHIR99021/Y-27632 treatment.

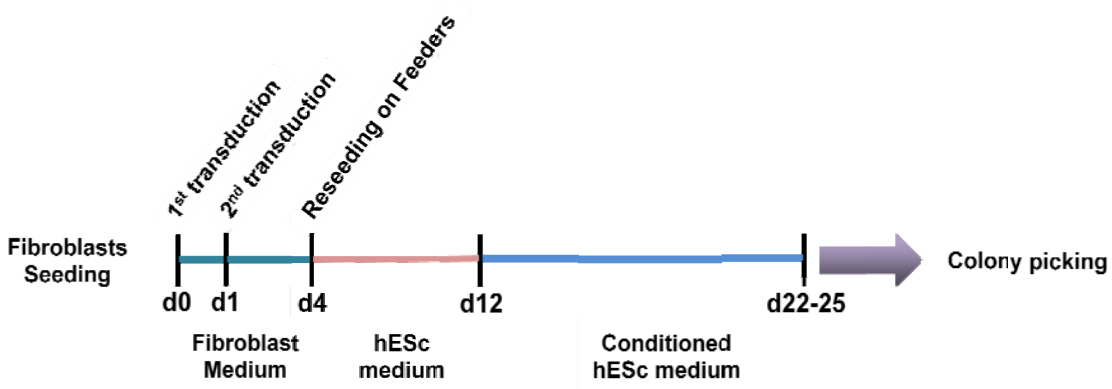

Figure 5-8 Schematic representation of the retroviral reprogramming procedure
In this procedure, fibroblasts were transduced with the pMXs retroviruses on two successive days. Four days after the 1st transduction, cells were transferred to iMEF feeder cells. A day later the fibroblasts medium was substituted with hESc medium. Twelve days after the 1st transduction, hESc medium was replaced with MEF-conditioned hESc medium.

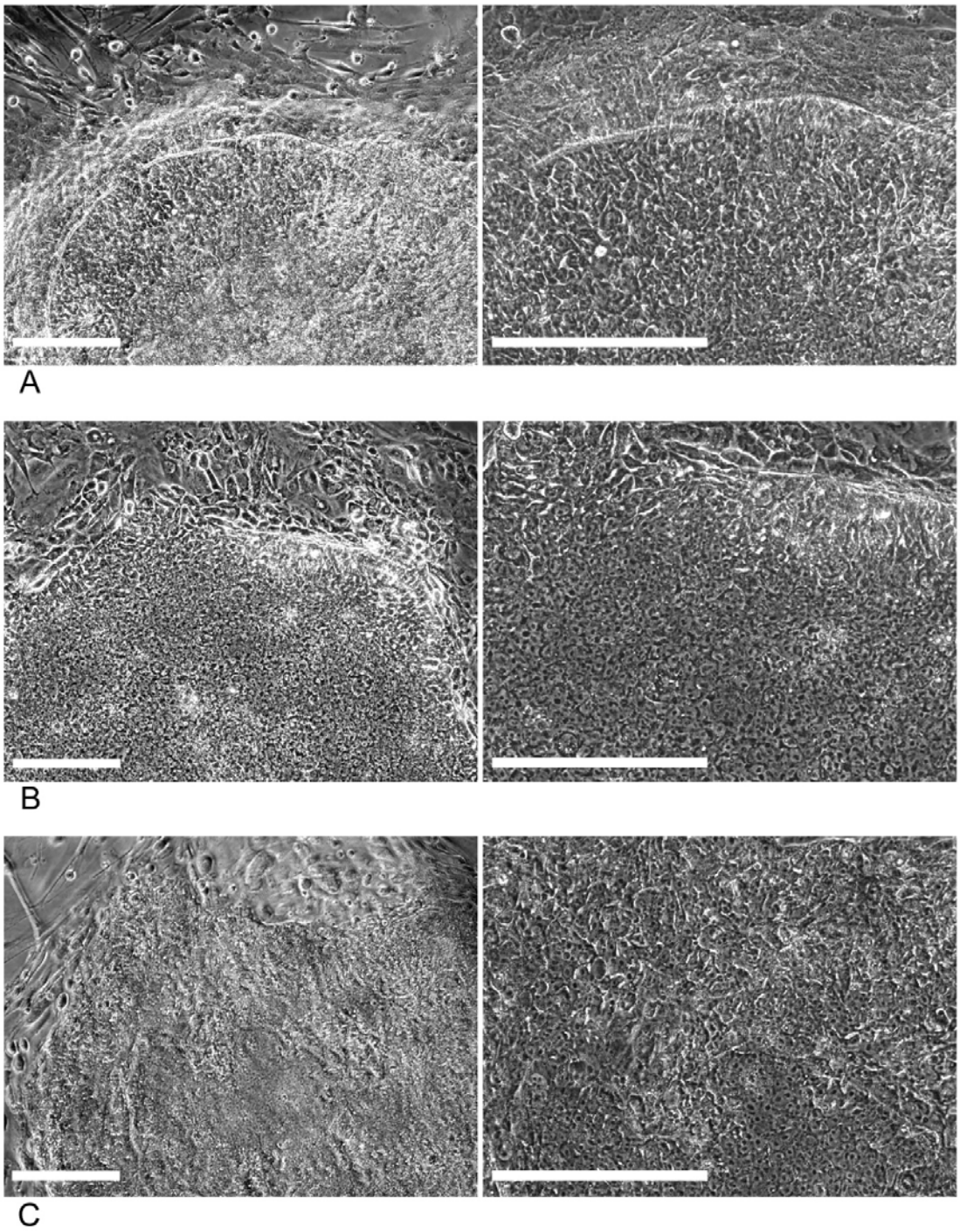

Figure 5-9 iPS colonies derived using the retroviral system
(A and B) Left Panel, phase contrast images of two retroviral system-induced iPS colonies, showing compact appearance and defined borders. Right panel, higher magnification (20x) of the same colonies exhibiting prominent nucleoli in the cells. (C) An irregular-shaped (non-iPS) colony obtained due to partial reprogramming is shown on the left panel. The cellular appearance of non-iPS cells is similar to that of iPS cells as shown in higher magnification (20x) on the right panel.

The non-iPS like colonies, which are reported to arise due to incomplete reprogramming (Maherali et al., 2008; Takahashi et al., 2007) were not analysed. The hES/hiPS like colonies were individually picked and expanded as described above. Two colonies, RHFC-HMVV and RHFC-HV1 were subsequently established in Matrigel/mTeSR-1 culture and were analysed for pluripotency markers and differentiation capabilities. To perform immunofluorescence analysis for the hESc-specific pluripotency markers, the cells from these clones were stained with antibodies against the nuclear proteins OCT4 and NANOG, and the surface stem cell marker TRA-1-60 along with the HUES10 cells as control. Both retroviral method-derived iPS clones RHFC-HMVV and RHFC-HV1 were uniformly stained by the antibodies for OCT4, NANOG and TRA-1-60 with a similar intensity as HUES10 cells as shown in Figure 5-10.

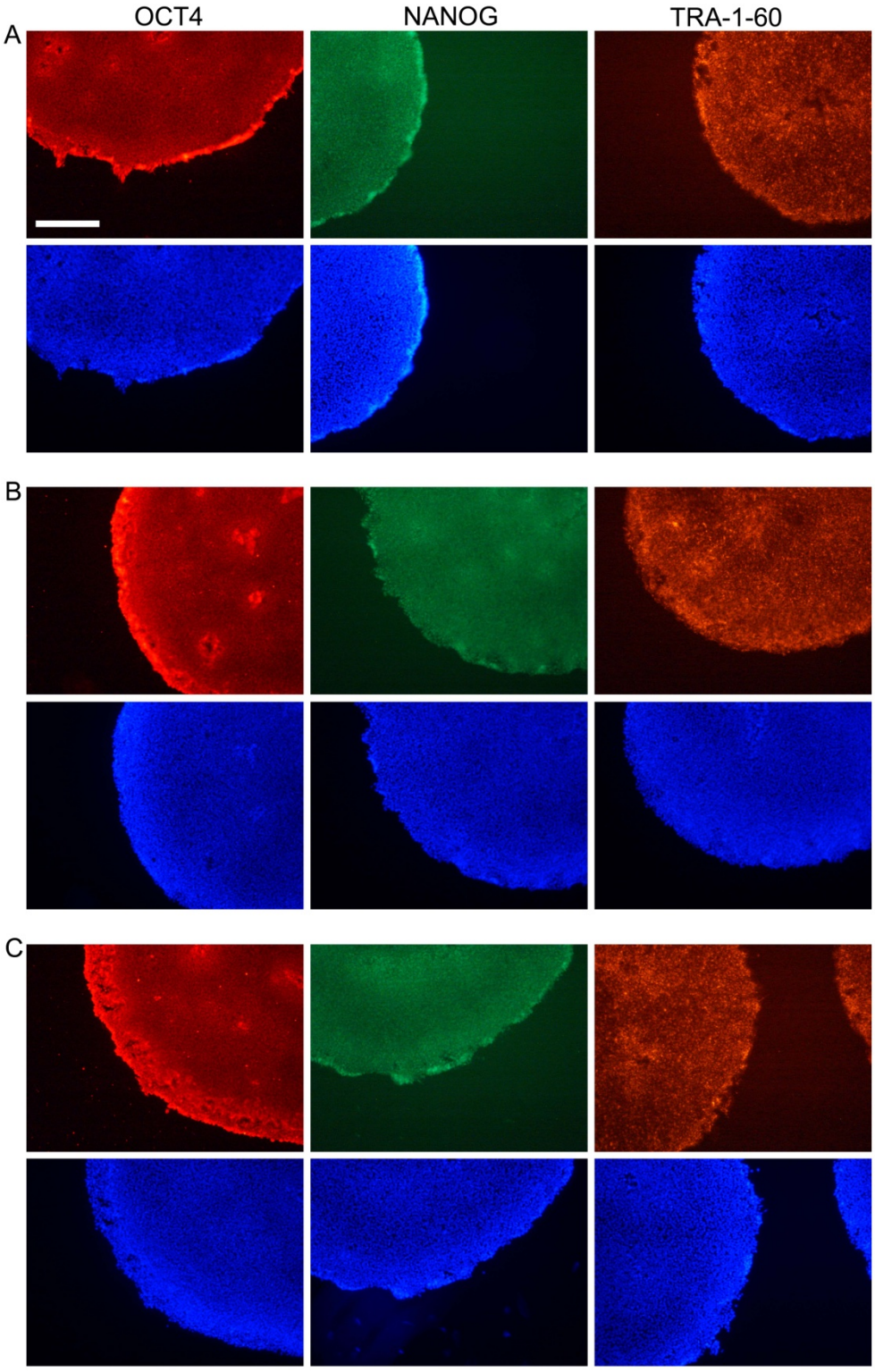

Figure 5-10 Immunofluorescence staining of retroviral vector-derived iPS colonies for pluripotency markers

The iPS clones RHFC HV1 (passage 10; A) and RHFC HMVV (passage 10; B) derived using the retroviral system were positive for hESc pluripotency markers OCT4 (red signal), NANOG (green signal) and TRA-1-60 (red signal). HUES10 cells (passage 26) were used as control (C). Nuclei of the cells were counterstained with DAPI (blue signal). Scale bar = 200µm.

5.3.6 Karyotype Analysis

Similarly to hESc, long-term *in vitro* culture of human iPS cells is associated with recurrent chromosomal abnormalities such as trisomies 17 and 8 (Taapken et al., 2011). To determine whether the HSV-1 amplicon-derived iPS clones along with the retroviral-generated iPS clones carried any karyotypic abnormality, M-FISH was performed on HHEFM-H1.5 (passage 12), HHEFM-H4.0 (passage 8), RHFC HMVV (passage 8) and RHFC HV1 (passage 8) using the 24XCyte M-FISH probe kit (MetaSystems). Detailed M-FISH analysis of thirty metaphases from these iPS cells indicated that majority of these cells did not contain large structural or numerical aberrations such as translocations, deletions or duplications (Figure 5-11). The clones HHEFM-H4.0 and RHFC HMVV were completely normal (both 46, XY). However, 3% of the analysed metaphases from the clones HHEFM-H1.5 and RHFC HV1 carried the trisomy 12 and 17 respectively (data not shown). The M-FISH on the iPS clones was performed and analysed by Dr. Daniela Moralli, Molecular Cytogenetics and Microscopy core facility, WTCHG.

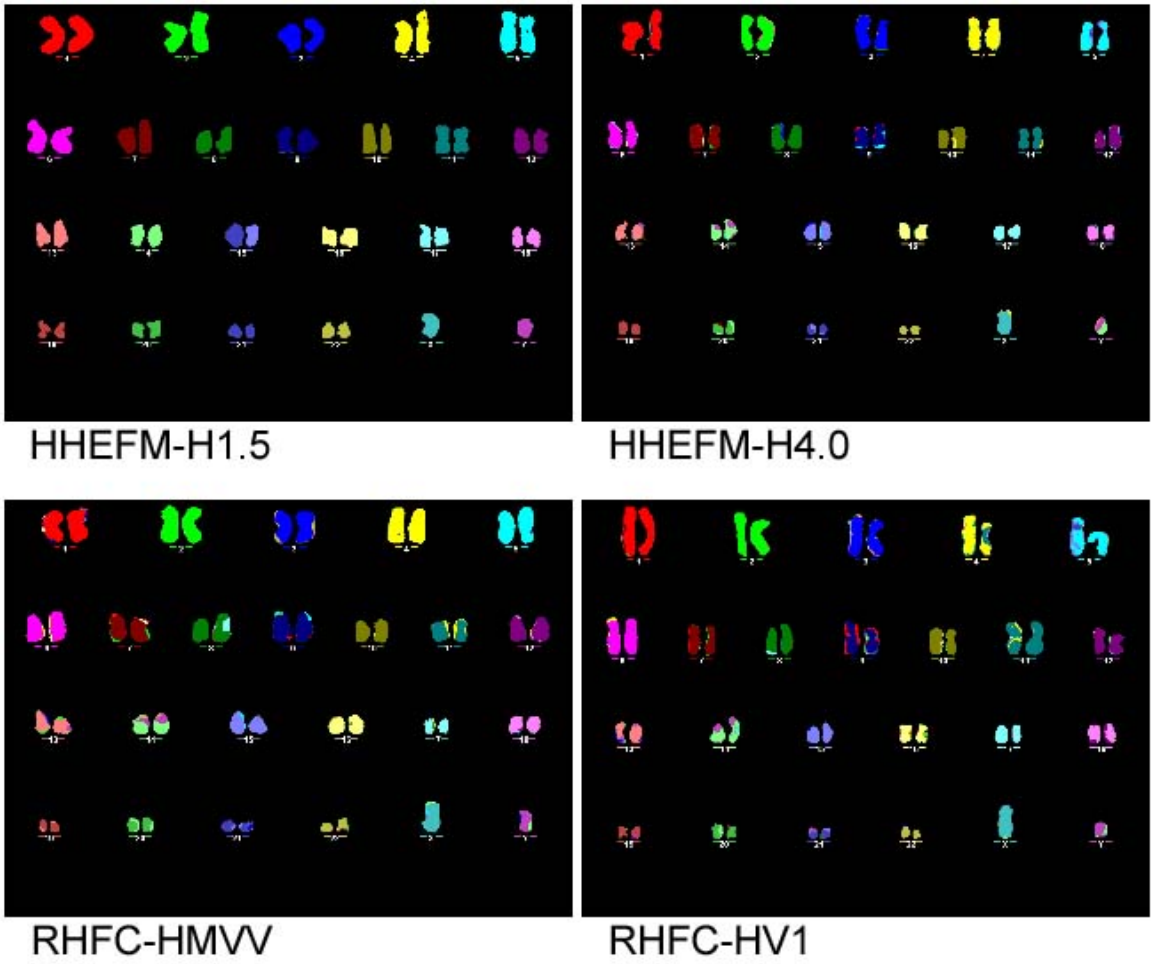

Figure 5-11 M-FISH analysis of the iPS clones
iPS cells generated through the HSV-1 amplicon system (HHEFM-H1.5 and HHEFM-H4.0) and the retroviral system (RHFC HMVV and RHFC HV1) showing normal karyotype (46, XY) and no structural or numerical aberrations. M-FISH images were generated by Dr. Daniela Moralli.

5.3.7 *In vitro* Differentiation and Germ Layer Analysis

To analyse the differentiation potential of the HSV-1 amplicon-iPS cells as well as retroviral-iPS cells, embryoid bodies were derived from the clones HHEFM-H1.5, HEFM-H4.0, RHFC HMVV and RHFC HV1 and subjected to differentiation as described in Section 2.13. The EB-derived differentiated cells showed various types of morphologies such as those resembling epithelial cells (hexagonal appearance), hematopoietic precursors (cobblestone-like appearance) and neuronal cells. To investigate if the differentiated cells expressed markers specific for the three germ layers, RT-PCR

analysis of the germ layer markers was performed on total RNA extracted from the differentiated derivatives of the iPS cells along with the hES cells HUES10 as control. As shown in Figure 5-12, transcripts specific for endoderm (*α Feto-protein*, *HNF3α*, *α1 anti-trypsin*), mesoderm (*GATA-2*) and ectoderm (*CK-5* and *CK-14*) were detected in the differentiated cells from all the iPS clones. Undifferentiated cells either did not express or exhibit lower levels of these markers. The overall analysis confirmed that these iPS clones retained the capacity for EB formation and germ layer differentiation.

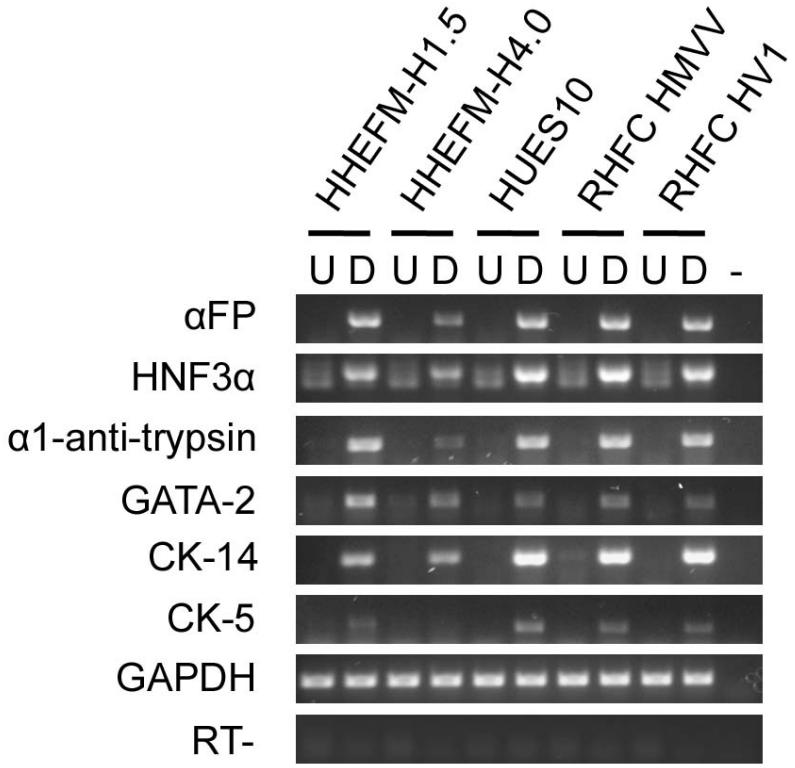

Figure 5-12 RT-PCR analysis of differentiation markers for the three germ layers in the iPS clones

The capability of germ layer differentiation of the HSV-1 amplicon-iPS clones (HHEFM-H1.5 and HHEFM-H4.0) and the Retroviral-iPS clones (RHFC HMVV and RHFC HV1) analysed by RT-PCR using primers specific for endodermal genes (*αFP*, *HFN3α*, *α-1 anti-trypsin*), mesodermal gene (*GATA-2*) and ectodermal genes (*CK-14* and *CK-5*). HUES10 cells were used as a control. *GAPDH* gene was used as an internal control. PCR of the RNA without reverse transcriptase (-RT) was done with the *GAPDH* primer set to ensure no amplification of genomic DNA. (-) = no template control.

Note: Due to a handling error, some RNA from the differentiated HHEFM-H4.0 clone was lost during the RNA preparation and as a result its expression for all the germ layer markers is low as compared to other clones.

5.3.8 Directed Differentiation of iPS Cells into Neuronal Cells

To examine if the iPS cells derived through HSV-1/EBNA-1/OriP amplicon system and retroviral vectors could be directed specifically into the neuronal cells, the strategy discussed in Section 3.3.13 was adopted. The differentiation involved the treatment of EBs with Noggin and fibronectin to repress bone morphogenetic protein (BMP) signalling followed by ectodermal induction using N2-media (containing bFGF) to promote the expansion of the neural progenitor rosettes. After the attachment of the EBs to the dishes, gradual outward expansion of differentiated cells was observed, followed by the formation of neuronal rosettes about 12 days after EBs attachment, as shown in Figure 5-13. The cells underwent further spreading and some neuronal structures (based on morphology of extensive network of axons and dendrites) were observed at around 20 days of EBs plating. At the end of the neuronal differentiation protocol (about twenty five days starting from EBs formation), the cells were stained with β-III-tubulin antibody to confirm the neuronal differentiation. As shown in Figure 5-14 the neuronal cells derived from all the iPS clones were stained positive for β-III-tubulin with similar intensity as the HUES10 control cells. The overall analysis indicated that HSV-1 amplicon derived iPS cells along with the retroviral-mediated iPS cells were capable of lineage-directed differentiation into neuronal cells.

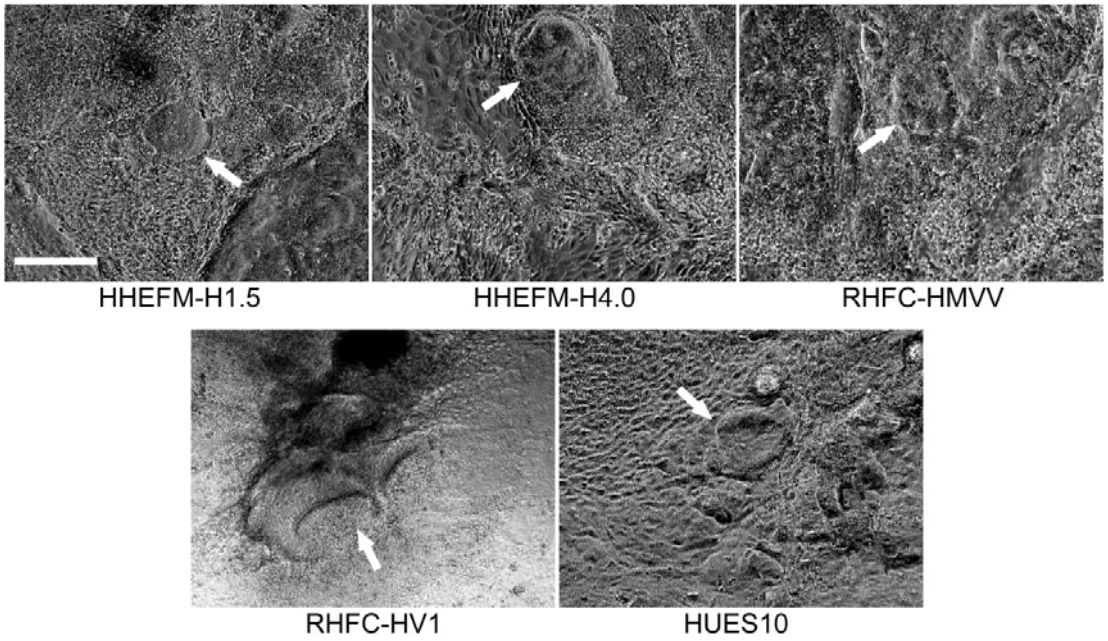

Figure 5-13 Neuronal differentiated cells from iPS clones
Phase contrast images of neuronal-directed differentiation of HSV-1 amplicon-iPS clones HHEFM-H1.5 and HHEFM-H4.0, retroviral-iPS clones RHFC-HMVV and RHFC-HV, and HUES10 cells (control). Neuronal rosettes are identified by a white arrow. Scale bar = 200μm.

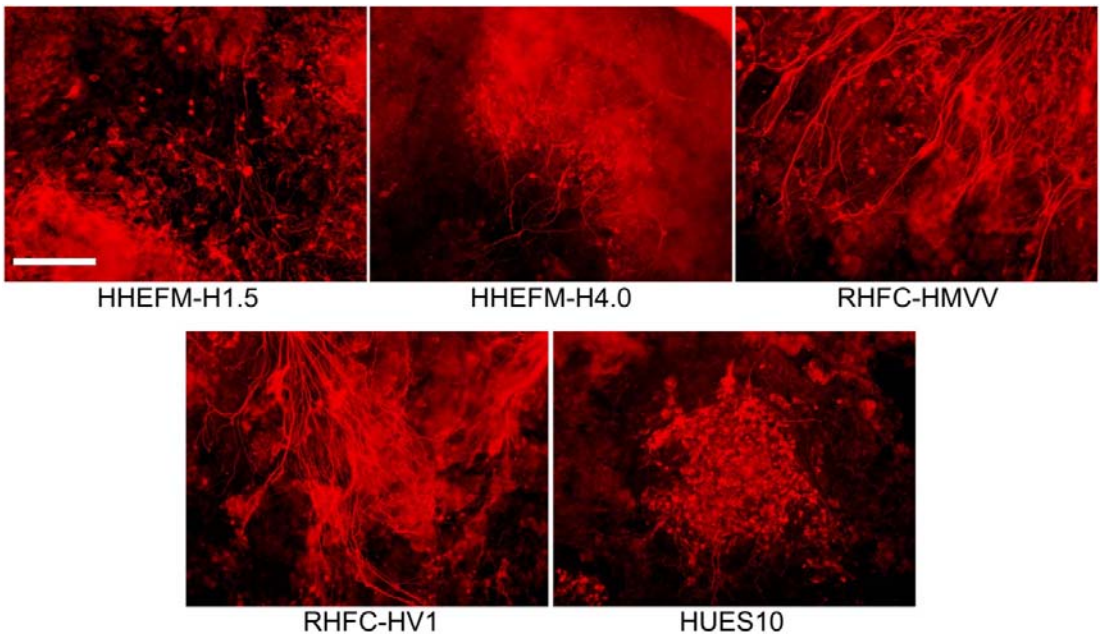

Figure 5-14 β-III tubulin staining of neuronal differentiated cells from iPS clones
Embryoid bodies-derived neuronal cells from the HSV-1 amplicon-iPS clones HHEFM-H1.5 and HHEFM-H4.0, and retroviral-iPS clones RHFC-HMVV and RHFC-HV1 were stained for β-III-tubulin (red), a neuronal specific marker suggesting that these iPS cells were capable of lineage-directed differentiation. HUES10 cells-derived neuronal cells were also stained as control. Scale bar = 200μm.

5.4 Discussion

This study established an integration-free strategy for generating iPS cells utilising the hybrid HSV-1/EBNA-1/OriP amplicon system. The system involved the transduction of hES cell-derived embryonic fibroblasts with two HSV-1/EBNA-1/OriP amplicon vectors, pHG puro 6F and pHG puro 3FLT containing the seven genes *OCT4*, *SOX2*, *KLF4*, *C-MYC*, *NANOG*, *LIN28* and *SV40LT*, and subsequent cultivation in hESc culture conditions, which resulted in the derivation of iPS cells with a reprogramming efficiency of 0.002-0.004% of the input cells. PCR analysis of the genomic DNA, extracted from these iPS clones, for the presence of vector elements indicated that the vector DNA did not integrate into the host genome. Moreover, the iPS cells were subjected to subcloning and extended passaging to remove the pHG puro 6F and pHG puro 3FLT vectors from the cells. The resulting iPS cells were completely free of vector and transgene sequences, a significant accomplishment for their potential use in clinical settings. An important aspect of the vector removal from the iPS cells is also established from another study (Maherali et al., 2008). This work demonstrated that the isolation and expansion of the true iPS cells (completely reprogrammed) can be facilitated by the removal of transgene expression of reprogramming factors, since incompletely reprogrammed cells rely on the continued expression of transgene factors, and thus will be eliminated upon their withdrawal.

The iPS cells generated using the HSV-1 amplicon system were pluripotent as indicated by the activation of endogenous pluripotency genes, immunoflourescence staining for pluripotency marker, germ layer differentiation and directed differentiation into neuronal cells. Nevertheless, the possibility that these iPS cells did not acquire the true iPS cell state cannot be ruled out as these cells were not subjected to teratoma analysis, which is

regarded as the most stringent assay for the assessment of pluripotency in hESc/hiPS cells (Takahashi et al., 2007; Yu et al., 2007). However, in a recent study, Chan and colleagues identified that the ability to form teratoma alone could not be used to evaluate whether true reprogrammed state is achieved or not (Chan et al., 2009). Instead the authors suggested that evaluating proviral silencing (applicable to their system), and expression of other markers such as TRA-1-60, DNMT3B and REX1 in conjunction with teratoma formation capability was a more precise way to distinguish fully reprogrammed cells. These analyses would provide a better assessment of the iPS colonies obtained through HSV-1/EBNA-1/OriP amplicon system, but they could not be carried out due to time constraints.

Though the reprogramming efficiency of the HSV-1 amplicon system was found to be low as compared to other non-integrating methods such as adenoviral vectors, protein transduction and mRNA transfection (Kim et al., 2009a; Warren et al., 2010; Zhou and Freed, 2009), a major advantage of the HSV-1/EBNA-1/OriP system was that it accomplished reprogramming with a single infection instead of multiple rounds of treatments required for the above mentioned non-integrating approaches.

The present episomal/plasmid vectors incorporate various bacterial elements such as antibiotic resistance gene and bacterial origin of replication for their maintenance , which carry many CpG islands associated with transgene silencing (Chen et al., 2004). These elements, also present in the HSV-1 amplicon vectors, might have contributed to the overall low reprogramming efficiency of the system. In order to reduce the transgene silencing, some improvements can be made in the presented HSV-1 amplicon vectors. For example, bacterial elements from the HSV-1 amplicon vectors can be removed through site-specific recombination system to develop minicircle HSV-1 amplicon vectors, which

may prove useful for increasing the reprogramming efficiency. Compared to plasmid/episomal vectors, minicircle vectors are known to provide relatively extended transgene expression as they are less susceptible to silencing mechanisms (Chen et al., 2003) and have also been successfully used in iPS generation (Jia et al., 2010).

As described in Section 5.3.1 the reprogramming efficiency of a particular system (both viral and non-viral methods) has been found to be associated with the developmental stage of the donor cells and therefore varies with the choice of target cells (Aasen et al., 2008; Jia et al., 2010; Maherali et al., 2008). This was also evident in the HSV-1/EBNA-1/OriP -based reprogramming attempt of IMR90 cells, which did not generate iPS cells. One way to circumvent the low HSV-1-mediated reprogramming efficiency of developmentally mature cells such as IMR90 and MRC5 fibroblasts, that have also demonstrated low amenability to reprogramming by other vector systems (Jia et al., 2010; Park et al., 2008b), could be achieved by scaling-up the reprogramming method and increasing the number of target fibroblasts. The low reprogramming efficiency of the EBNA-1/OriP episomal vectors was also demonstrated in the method initially employed by Yu and colleagues, which only resulted in three iPS colonies from $1x10^6$ foetal foreskin fibroblasts (0.0003%) (Yu et al., 2009). Therefore the HSV-1 amplicon method described here can be scaled-up to include $1x10^6$-$2x10^6$ cells instead of $5x10^4$ cells, which will formulate a 20-40 fold increase in the starting cell numbers and could possibly assists in the generation of iPS cells from developmentally advanced cells types such as IMR90 fibroblasts.

Initial studies on derivation of mouse iPS cells have documented that the chimeric mice developed from mouse iPS cells have increased occurrence of tumour formation, attributed to the reactivation of the *c-myc* gene, which is a well-known proto-oncogene

(Nakagawa et al., 2008; Okita et al., 2007). To avoid the use of *c-myc*, some studies have reported the successful derivation of both mouse and human iPS cells without *c-myc* albeit with lower efficiency (Huangfu et al., 2008; Nakagawa et al., 2008; Wernig et al., 2008a). Recently, efforts on avoiding the use of *c-myc* but at the same time maintaining/increasing the reprogramming efficiency have lead to the identification of *L-myc*, another member of *myc* family (Nakagawa et al., 2010). *L-myc* exhibits significantly lower transformation activity compared to other *myc* members but most importantly has been shown to provide a stronger and more specific effect on generating iPS in contrast to *c-myc* (Nakagawa et al., 2010; Yu et al., 2011). Thus, the use of *L-MYC* instead of the *C-MYC* in the presented HSV-1/EBNA-1/OriP amplicon system may possibly enhance the reprogramming efficiency as well as reduce the transformation risk posed by *C-MYC*, therefore deriving clinically-relevant hiPS cells.

The inability of the HSV-1 amplicon vector pHG-puro 6F on its own (i.e. without employing the pHG puro 3FLT transduction) to generate iPS cells was also supported by the studies conducted by Yu and colleagues (Yu et al., 2011; Yu et al., 2009), in which the combinations of EBNA-1/OriP vectors not containing the *SV40LT* gene were not sufficient to generate the iPS cells. Both the use and the role of *SV40LT* gene in the generation of iPS cells is controversial as in addition to its positive effect on reprogramming, it is known to cause cellular transformation and inhibit *p53*, a tumour suppressor gene (Cheng et al., 2009). However, the first study incorporating the *SV40LT* gene in the reprogramming factors cocktail suggested that its expression facilitated but did not contribute to the generation of iPS cells (Park et al., 2008b). Similarly, another study reported that *SV40LT* enhanced both the pace and the efficiency of iPS generation (Mali et al., 2008). Yu and colleagues suggested that the *SV40LT* gene used in EBNA-

1/OriP vectors counteracted the toxic effect induced by *C-MYC* during the reprogramming process. Although it is not clearly understood how *SV40LT* facilitates the iPS generation, its established role in inhibiting the *p53* gene and road blocking effect of *p53* in reprogramming (Marion et al., 2009; Utikal et al., 2009b) could possibly be interconnected.

The retroviral system of iPS generation (Takahashi et al., 2007), originally developed in Shinya Yamanaka's lab, was also employed along with the iPS generation through the HSV-1 amplicon system to evaluate the relative efficiency of both systems. The retroviral system is known for efficient generation of iPS cells (0.02% of input cells). However, in this study the efficiency of the retroviral system (0.0025%) was found to be almost similar to that of the HSV-1 amplicon system (0.002%). The discrepancies in the iPS generation efficiency of retroviral system used in this study and the study led by Shinya Yamanka could be due to the different modifications in both the system as described in the Section 5.2.3. Most importantly, the transduction efficiency of the retroviral system for the embryonic fibroblasts used in this study was expected to be low because these fibroblasts do not express Slc7a1 (mouse receptor for retroviruses), which increases the transduction efficiencies of retroviruses in human fibroblasts (Takahashi et al., 2007). Nevertheless, the HSV-1 amplicon iPS cells were derived in an integration free manner as compared to those derived through retroviral vectors. In addition, non-iPS cells (incompletely reprogrammed cells), which are represented by colonies with loose borders and/or granular appearance, were not observed in the HSV-1 amplicon-mediated reprogramming in contrast to the retroviral system. These non-iPS colonies are known to develop from cells which only acquired *OCT4* and *C-MYC* retroviruses (Maherali et al., 2008). In contrast, the two HSV-1 amplicon vectors, developed in this study, harboured

all the reprogramming genes and thereby possibly reduced the likely imbalance of the reprogramming genes and as a result the occurrence of non-iPS cells.

As hiPS cells are cultured in conditions which are possibly responsible for selecting karyotypic abnormalities in hESc (Draper et al., 2004), karyotype analysis of the iPS cells was conducted. The karyotypes of the iPS cells generated through the HSV-1 amplicon system and the retroviral system were found to be mostly normal as analysed by M-FISH. However, there was low incidence in both the HSV-1 iPS clone HHEFM-H1.5 and the retroviral iPS clone RHFC HV1, of trisomy 12 and trisomy 17 respectively. It is, however, not determined whether the frequency of cells exhibiting chromosome gains will remain the same or will increase during additional sub-culturing of the cells. In a recent study, Taapken and colleagues analysed over 1700 hESc and hiPS cell lines for their karyotype (Taapken et al., 2011). Although the trend of chromosomal abnormalities was found to be similar for both hESc and hiPS cells, trisomy 8 was more frequently observed in hiPS cells as compared to hESc. In addition trisomy 17 (commonly found in hESc) was not observed in any of the 219 hiPS cell lines and 552 hiPS cultures that were karyotyped. It is, therefore, interesting that one retroviral-iPS clone RHFC HV1 had one occurrence of trisomy 17. However, as mentioned above karyotype analysis of the cells at higher passages could possibly furnish a more meaningful assessment. Nevertheless, the iPS clone HHEFM-H4.0 retained a completely normal karyotype, indicating that the iPS cells generated through HSV-1 amplicon system can potentially be utilised in clinical and therapeutic settings.

Overall, the development of the HSV-1 amplicon system for the generation of iPS cells is a significant advantage towards utilising the iPS cells for disease modelling and drug screening purposes. Moreover, the vector and integration-free iPS cells could also be

utilised for the better assessment of functional equivalence of ES and iPS cells. Carrying out this assessment is otherwise difficult using the iPS cells generated through retroviral system or other integration-based approaches because the residual expression of the transgene sequences in the iPS cells can affect their molecular signatures and differentiation potential (Sommer et al., 2010; Stadtfeld et al., 2008c).

Chapter 6. HAC Development in Human Induced Pluripotent Stem Cells

6.1 Introduction

Human induced pluripotent stem (hiPS) cells, like hESc, have the capability of indefinite proliferation and differentiation into all the cell types within an individual, and are regarded as an invaluable tool for studying *in vitro* disease modelling, screening drugs and generating patient-specific/immune-compatible cells for autologous transplantation therapies for a wide variety of diseases (Takahashi et al., 2007; Yu et al., 2007). Approaches for safe and efficient genetic modification of hiPS cells would be crucially required for studying specific genes in the context of personalised medicine, correcting the mutations in hereditary diseases and to some extent for directing lineage-specific differentiation into disease relevant cell types. Similar to hESc, the methods available for genetic modification of hiPS cells such as retroviral/lentiviral vectors, adenoviral vectors, transfection, electroporation and Nucleofection are either associated with the risk of insertional mutagenesis or exhibit very low delivery efficiencies.

Chapter 3 described the experiments in which hES cells HUES10 were efficiently modified in an integration-free manner utilising the HSV-1 amplicon vector (pHSV17α40Neo) to develop HAC. In the study described here, the same approach was utilised and extended to generate HAC in hiPS cells. Using the HSV-1 amplicon system, hiPS cells were transduced with the pHSV17α40Neo HAC vector to evaluate whether hiPS cells can support the generation of mitotically stable HAC and provide sustainable gene expression in the cells. Initial attempts to generate HAC in hiPS cells, however,

were hindered by drug selection problems after the delivery of HAC constructs into the cells. Moreover, gradual silencing of the *GFP* expression further curtailed the tracking of the transduced cells, 5-6 days post-transduction. To overcome the limitation imposed by the drug selection and silencing of both the neomycin resistance gene (*neo*) and the *GFP* gene, the pHSV17α40Neo vector was improved to incorporate the *neo* gene and the *GFP* gene under the control of the CAG promoter, which is known to sustain stable gene expression in hESc and hiPS (Chatterjee et al., 2011; Liew et al., 2007). The modified HAC vector not only improved the selection of the hiPS cells but also facilitated the development of the stable clones that expressed GFP. The overall study represents the first report in which the hiPS cells have been targeted via the HSV-1 amplicon vectors with the aim of developing mitotically stable and non-integrating gene expressing HAC.

6.2 Specific Methods

6.2.1 Human iPS Cell Lines

To investigate the potential of establishing HAC in hiPS cells, two hiPS cell lines, a commercially available DF19-9-11T.H (Yu et al., 2009) and the other derived in this study HHEFM-H1.5 (Chapter 5) were chosen. The DF19-9-11T.H iPS cells were derived from human foreskin fibroblasts using OriP/EBNA1-based episomal vectors in an integration-free manner and are devoid of any transgene or vector elements. These iPS cells exhibited typical hESc characteristics including the ability to form teratomas when injected into immunocompromised mice and their gene expression pattern was highly similar to those observed in common hES cell lines (Yu et al., 2009). Moreover, the retention of normal karyotype after their generation and subsequent culture made these cells an ideal choice to determine if HAC can be generated in karyotypically normal hiPS

cells. The second hiPS cell line, HHEFM-H1.5, was generated from HUES10-derived embryonic fibroblasts as described in Chapter 5.

6.2.2 Assembly of HAC Vector

Multiple cloning steps were performed to construct the HSV-1 17α40kb HAC vector containing *neo* and *GFP* cassettes under control of the CAG promoter: (1) The *AcGFP* cassette was cloned from the pIRES2-AcGFP1 vector (Clontech) into the pIRES-CAG vector (Figure 4-1) using the NheI and NotI restriction sites to construct the pIRES CAG-GFP vector. (2) pSG80A-HG (or pHG; Figure 2-1) vector was digested with PmeI/NruI and subsequently blunt-end ligated (referred as pHG GFP^-) to remove the *GFP* gene; (3) The *neo* cassette was excised from the pIRESneo2 vector (Clontech) through XmaI/XbaI digestion and cloned back into the pIRESneo2 at AgeI/XbaI restriction sites; (4) Subsequently, the *neo* gene from the resulting vector was excised through NheI/XbaI digestion and inserted at NheI site of pIRES CAG-GFP vector (generated in step 1), to produce pIRES CAG-N2G; (5) The CAG Neo-IRES2-GFP fragment from pIRES CAG-N2G was then excised through SpeI/PvuI digestion and cloned into the SpeI/PacI restriction sites of the pHG GFP^- vector (developed in step 2) to form pHG CAG-N2G (Figure 6-1A); (6) finally, the pHG 17α40 CAG N2G vector (HSV-1 HAC vector; Figure 6-1B) was created by performing Cre *loxP* recombination between pHG CAG-N2G and BACα40 HAC vector containing 40kb of 17α DNA (Mandegar, 2011; Mandegar et al., 2011).

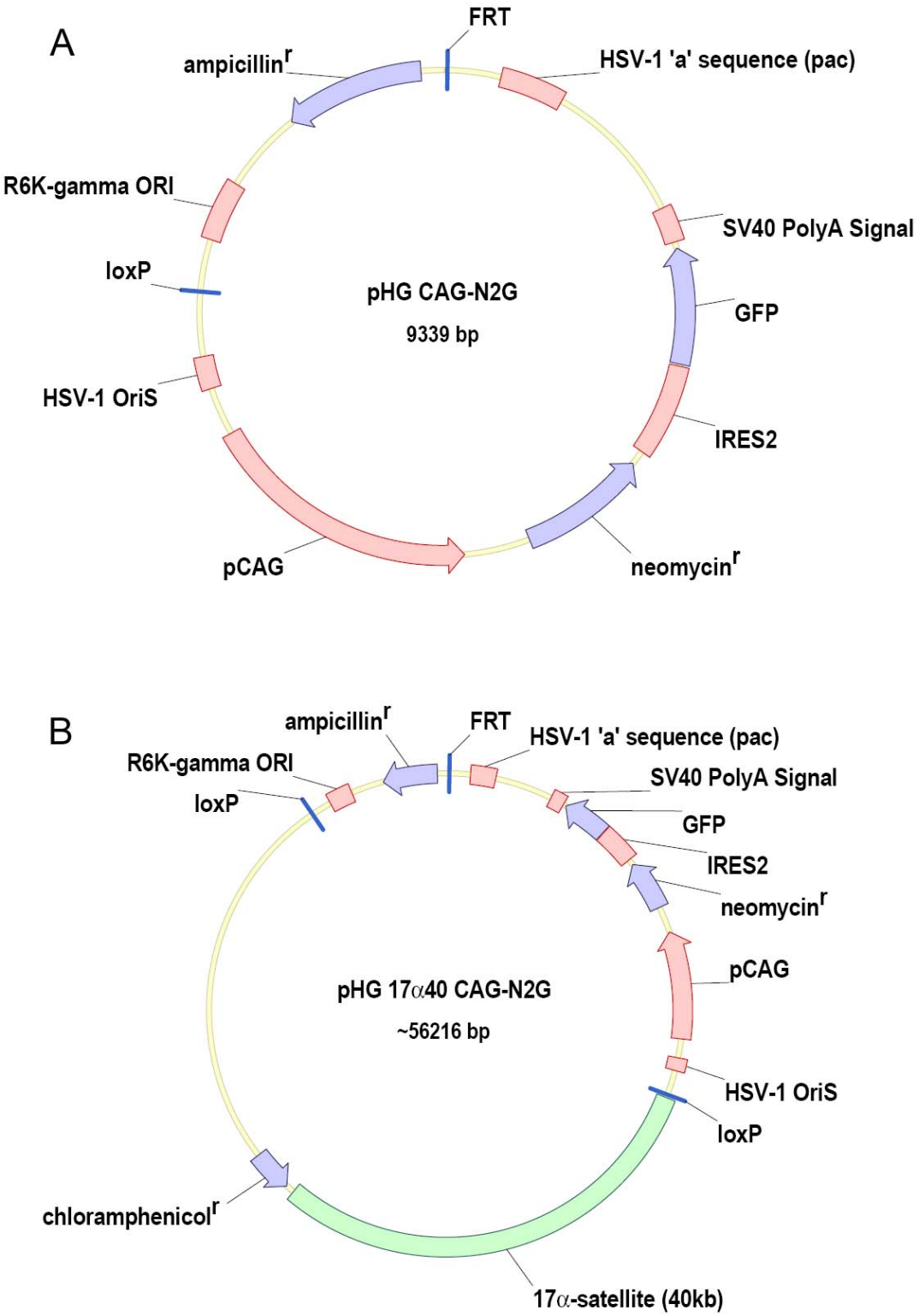

Figure 6-1 Vector maps of pHG CAG-N2G and pHG 17α40 CAG-N2G
The HSV-1 amplicon vector pHG CAG-N2G (A) contains *neo* and *GFP* cassettes under a single CAG promoter co-linked by IRES2. The HSV-1 HAC vector pHG 17α40 CAG-N2G (B; not drawn to scale) contains the 40kb17α DNA.

6.2.3 PCR Analysis of the Stable Clones

The hiPS and the hESc stable clones generated using the pHG 17α40 CAG N2G HSV-1 HAC vector were screened by the PCR on the genomic DNA extracted from the clones. The PCR was performed on the *AcGFP* gene using the primers AcGFP 1F 5'-CACATGAAGCAGCACGACTT-3' and AcGFP 1R 5'-ATGTTGTGGCGGATCTTGA-3'.

6.3 Results

6.3.1 M-FISH analysis of the iPS Cells

With the aim of generating HAC in hiPS cells exhibiting a normal karyotype, M-FISH was performed on the DF19-9-11T.H iPS cells (passage 29) using 24XCyte M-FISH probe kit (MetaSystems). Detailed M-FISH analysis of thirty metaphases from the iPS cells indicated that none of these thirty cells contained large structural or numerical aberrations such as translocations, deletions or duplications, and gain or loss of chromosomes (Figure 6-2). The M-FISH experiment was performed by the Molecular Cytogenetics and Microscopy core facility, Wellcome Trust Centre for Human Genetics (WTCHG), University of Oxford. The karyotype of the iPS cell line HHEFM-H1.5 derived using HSV-1 amplicon system was shown in Chapter 5 (Figure 5-11).

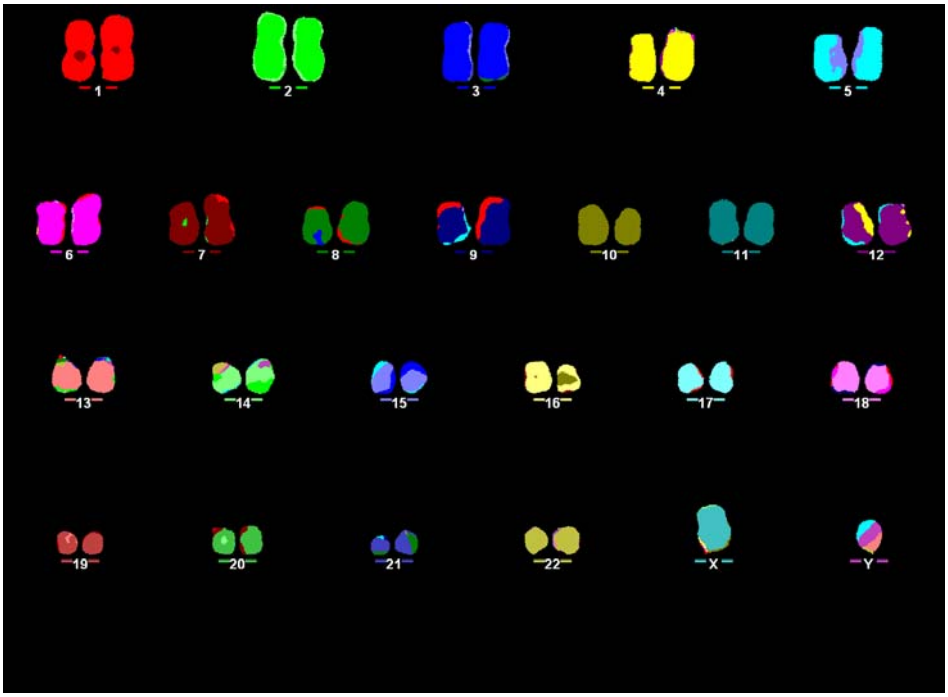

Figure 6-2 M-FISH analysis of DF19-9-11T.H iPS cells
Karyotype of the iPS cell line DF19-9-11T.H (passage 29) by M-FISH shows that the cells did not contain any translocations and other structural or numerical aberrations, and exhibited a normal karyotype (46, XY). This image was generated by Molecular Cytogenetics and Microscopy core facility at Wellcome Trust Centre for Human Genetics.

6.3.2 Transduction of HSV-1 HAC Amplicon in iPS Cells

The HSV-1 HAC vector pHSV17α40Neo (Figure 3-1) was transduced into $2.5x10^5$ iPS cells from lines DF19-9-11T.H (passage 29) and HHEFM-H1.5 (passage 30) at MOI 5, with spinoculation. The efficiency of transduction was determined 24 hours post-transduction by monitoring the percentage of GFP expressing cells in comparison to the total number of cells (Figure 6-3). The transduction efficiencies in the iPS cells DF19-9-11T.H and HHEFM-H1.5 were determined to be 25% and 32% respectively, which were roughly the same as those observed for HUES10 cells, using the same experimental conditions.

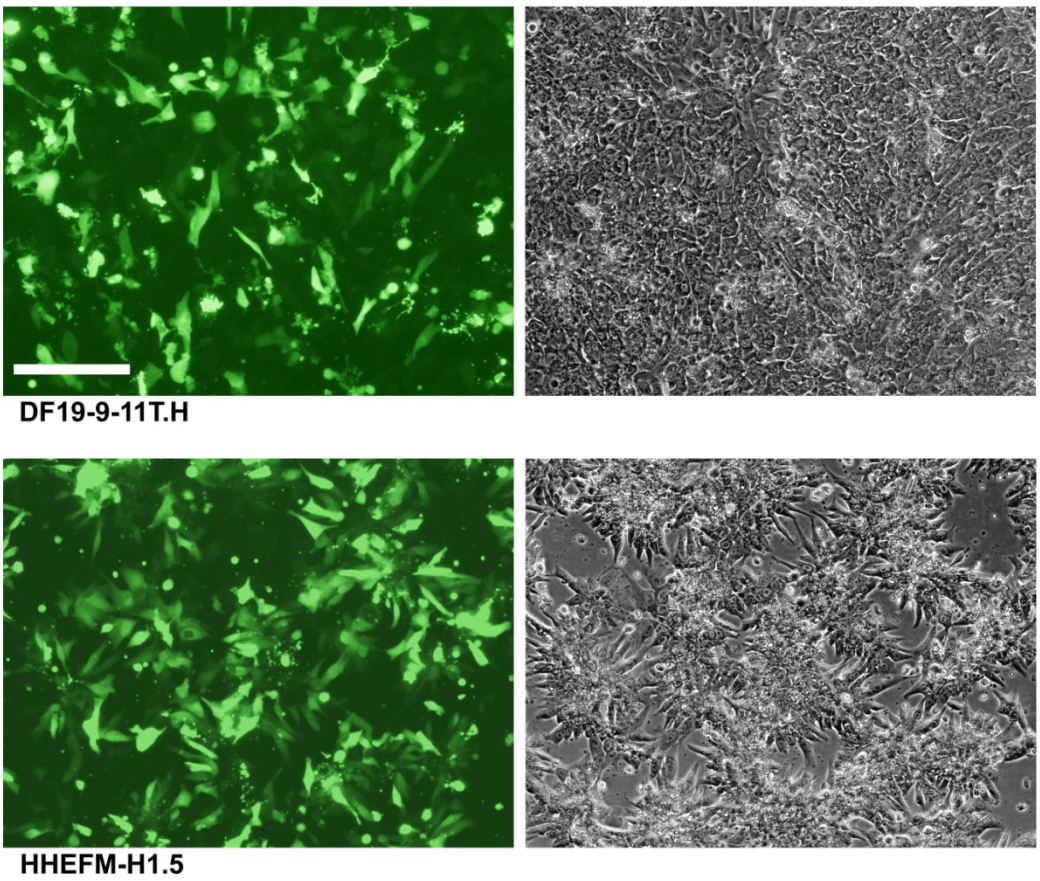

Figure 6-3 hiPS cells DF19-9-11T.H and HHEFM-H1.5 at 24 hours post pHSV17α40Neo transduction
hiPS cells DF19-9-11T.H (passage 29) and HHEFM-H1.5 (passage 30) were transduced with pHSV17α40Neo HSV amplicon vector at MOI 5. *GFP* expression (left panel) in these iPS cells was monitored 24 hours post-transduction. The right panel contains the corresponding phase contrast images. Scale bar = 200μm.

The transduction of iPS cell lines with the HSV-1 HAC vector did not lead to any immediate visible side effects on cell viability, morphology and growth patterns as shown in the Figure 6-3 (bright field images). Similar to the results obtained in the transduction of HUES10 cells, gradual silencing of *GFP* expression was observed 48 hours post-transduction and continued until day six. After six days of transduction, fluorescence microscopy did not reveal any GFP, as expected due to the silencing of HSV-1 IE4/5 promoter controlling *GFP* expression.

6.3.3 Drug Selection and Stable Clone Formation

Three days after the transduction of DF19-9-11T.H and HHEFM-H1.5 cells with the pHSV17α40Neo vector, the cells were transferred to SNL76/7 feeder cells. Two days later, G418 selection was applied at the concentration of 100μg/ml (as previously used for HUES10 cells) and was continued for 7 days. After the removal of the G418 selection, the cells were cultivated for additional 5-7 days to allow the development of individual clones. However, no clones were obtained from either of the iPS cell lines, indicating that the cells didn't withstand the G418 selection. These results suggested that the iPS cells, in comparison to HUES10 cells, acted differently in response to the same drug selection conditions.

Subsequent experiments were performed to improve the drug selection conditions for iPS cells by either reducing the G418 concentration or the number of days the cells were kept under selection. Various attempts were performed in which the G418 concentration was reduced to 50μg/ml and 75μg/ml and/or selection was applied for 5-6 days. However, the reduction in the stringency of selection condition also lead to the development of clones in the non-transduced cells (negative control) suggesting that the selection conditions were not sufficient to eliminate all the non-transduced cells. Nevertheless, sixteen clones obtained from pHSV17α40Neo transduced DF19-9-11T.H cells and selected at 100μg/ml G418 for 5 days as shown in Figure 6-4, were chosen for further analysis. None of the clones exhibited visible *GFP* expression under the microscope suggesting that either the clones developed as a result of incomplete drug selection or the *GFP* gene was silenced. Individual clones were picked and transferred to separate wells of a 24-well dish on inactivated SNL76/7 feeder cells and maintained for 6-7 days until the clones grew to

larger size. Cells were later expanded on Matrigel-coated dishes in mTeSR-1 medium for further analysis.

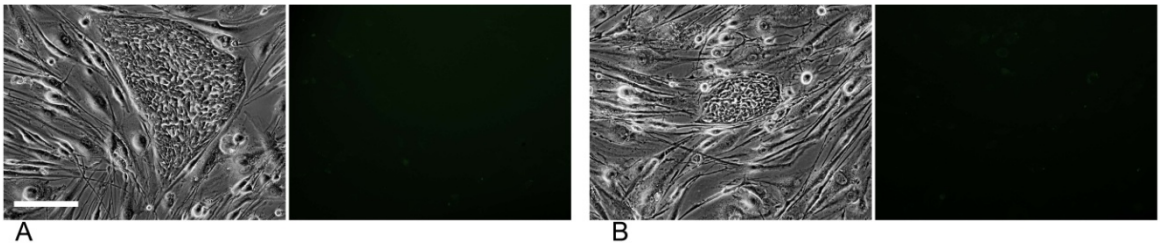

Figure 6-4 Stable clones of DF19-9-11T.H cells generated after G418 selection
Phase contrast images of two DF19-9-11T.H stable clones (A and B) generated after G418 selection for five days. The clones didn't exhibit visible *GFP* expression under the microscope (right panels). Scale bar = 200μm.

PCR analysis on genomic DNA extracted from the DF19-9-11T.H clones using the primers set specific for the *neo* gene (Section 3.2.2) gave a positive band in only one out of the sixteen clones (data not shown). Metaphase spreads were prepared from the cells of the PCR-positive clone and analysed by two colour FISH with a vector probe (HAC-specific) and 17α DNA probe to determine whether the clone contained HAC. However the FISH results did not reveal either HAC or integrations of the HAC vector in the host chromosomes. As the clone was found to be positive for the PCR of *neo* gene, the inability to identify HAC could be associated with the presence of a very small HAC undetectable through FISH. Alternatively, DNA from residual SNL76/7 feeder cells (which also express *neo* gene) contaminating the DNA from the clone could possibly be responsible for the detection of the *neo* gene through the PCR.

6.3.4 Modification of the HSV-1 Amplicon HAC Vector

In the pHSV17α40Neo transduction experiments described above, the generation of stable clones was hindered by problems with selection, possibly due to the gradual

silencing of the neomycin resistance (*neo*) gene, which is controlled by the SV40 promoter (Table 2-2). Moreover, the *GFP* reporter gene (controlled by HSV IE4/5 promoter) also exhibited silencing within 5-6 days post-transduction and therefore restricted the tracking of the transduced cells. As viral promoters such as SV40 and CMV are known to be prone to silencing in the hESc (Chan et al., 2008; Liew et al., 2007) and possibly in the hiPS cells, the HSV-1 amplicon vector pSG80A-HG (pHG) was modified to incorporate the *neo* and the *GFP* gene under the CAG promoter, which is known to confer efficient transgene expression in hESc as well as hiPS cells (Chatterjee et al., 2011; Liew et al., 2007). The *neo* gene was placed directly under the CAG promoter and the *AcGFP* gene was connected to the same CAG promoter utilising the IRES2 linker to obtain bicistronic expression of the *neo* and the *AcGFP* genes, resulting in the vector pHG CAG-N2G (Figure 6-1 A). The functional performance of the pHG CAG-N2G vector was initially assessed in the HeLa cells by transfecting the pHG CAG-N2G DNA and the control vector (pHGNeo4) into the cells using ExGen500. The transfected cells with both vectors expressed the *GFP* gene and supported the formation of stable clones upon G418 selection (data not shown). Subsequently, the HSV-1 HAC vector pHG 17α40 CAG-N2G (Figure 6-1 B) was assembled by retrofitting the BAC containing 40kb 17α DNA (Mandegar, 2011) with the HSV-1 amplicon vector pHG CAG-N2G using the Cre/LoxP recombination as described in Section 6.2.2.

6.3.5 HSV-1 HAC Transduction of pHG 17α40 CAG-N2G in HT1080 Cells

The modified HSV-1 HAC amplicon vector pHG 17α40 CAG-N2G was initially transduced into the control cell line HT1080 cells to assess the *GFP* expression, stable clone formation capability and HAC forming capacity of the vector. The pHG 17α40 CAG-N2G was transduced at MOI 8 into $1x10^5$ cells in one well of a 24-well dish

followed by spinoculation. The transduction efficiency was determined after 48 hours of transduction by monitoring the percentage of GFP expressing cells in comparison to the total number of cells and was found to be 33.68%. Figure 6-5 A shows *GFP* expression in the HT1080 cells transduced with pHG 17α40 CAG-N2G vector. On the same day the transduced cells, along with the non-transduced cells (control) were trypsinised and transferred to 10cm dishes. A day later, the transduced and the non-transduced cells were subjected to 350µg/ml G418 to assess the formation of stable clones. By day 10 after the application of G418 selection, all the non-transduced cells were eliminated while stable clones were generated from the transduced cells (Figure 6-5 B). The majority of the clones were GFP positive, however, the level of *GFP* expression varied between the clones. Overall 86 stable clones were obtained leading to stable clone formation efficiency of $8.6x10^{-4}$ of the overall cells and $2.5x10^{-3}$ of the total transduced cells. The stable clones were pooled and subsequently expanded for FISH analysis to determine whether they contained HAC.

FISH analysis of the metaphase spreads of the HT1080 stable clones (pooled) with the 17α DNA probe and the pHG CAG-N2G vector probe indicated the presence of HAC as shown in Figure 6-6 A. The analysis established that the HSV-1 HAC vector pHG 17α40 CAG-N2G was capable of generating HAC in HT1080 cells. However, in some metaphases, the HAC vector DNA also integrated into the host chromosomes (Figure 6-6 B) in a similar fashion as previously observed in the HT1080 cells transduced with the pHSV17α40Neo vector (Mandegar, 2011).

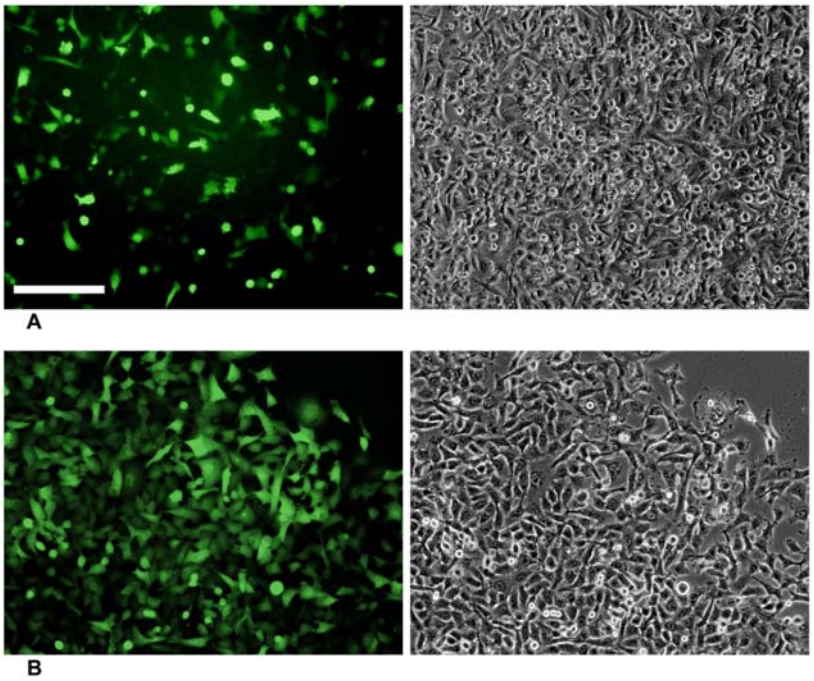

Figure 6-5 pHG 17α40 CAG-N2G transduction of HT1080 cells
HT1080 cells were transduced with the HSV-1 HAC vector pHG 17α40 CAG-N2G. (A) *GFP* expression (left panel) in the transduced cells was monitored 48 hours post-transduction. (B) A GFP expressing stable clone obtained after application of G418 selection. Right panels contain phase contrast images of the corresponding fields. Scale bar = 200μm.

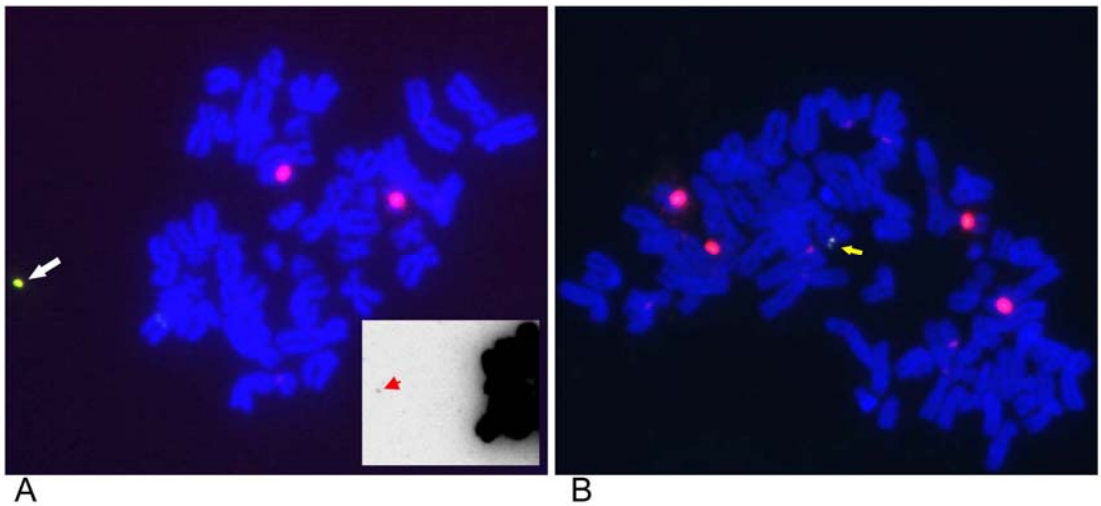

Figure 6-6 FISH analysis of stable clones of HT1080 cells
FISH analysis of stable clones (pooled) of HT1080 cells generated with pHG 17α40 CAG-N2G vector for the presence of HAC. The 17α DNA probe, which also recognised the endogenous chromosome 17, is labelled in red and the pHG CAG-N2G vector probe is labelled in green. The chromosomes are counterstained by DAPI, blue. (A) A metaphase spread showing the presence of HAC (identified by a white arrow). The inset shows DAPI staining only, and the HAC is identified by a red arrow. (B) A metaphase spread showing integration (identified by a yellow arrow) of the HAC DNA into the host chromosome.

6.3.6 Transduction of pHG 17α40 CAG-N2G vector in iPS Cells

After establishing the capability of the vector pHG 17α40 CAG-N2G for stable clone formation, HAC generation and *GFP* expression in HT1080 cells, pHG 17α40 CAG-N2G HSV-1 amplicons were transduced into the iPS cells DF19-9-11T.H and HHEFM-H1.5 as well as HUES10 cells (control) to investigate HAC formation. $2.5x10^5$ iPS cells DF19-9-11T.H (passage 31) and HHEFM-H1.5 (passage 17), and hES cells HUES10 (passage 22), all seeded on Matrigel-coated 24-well dishes in mTeSR-1 medium, were transduced with the pHG 17α40 CAG-N2G HSV-1 amplicons at MOI 2 followed by spinoculation. The *GFP* expression of the transduced iPS cells and hESc (48 hours post-transduction) is shown in Figure 6-7. Transduction efficiency of the cells was measured by recording the percentage of GFP expressing cells in comparison to the total number of cells and is listed in Table 6-1. DF19-9-11T.H cells were transduced at a relatively low efficiency as compared to the HHEFM-H1.5 and HUES10 cells. This was possibly due to the overall low number of cells as can be seen in the phase contrast images in Figure 6-7.

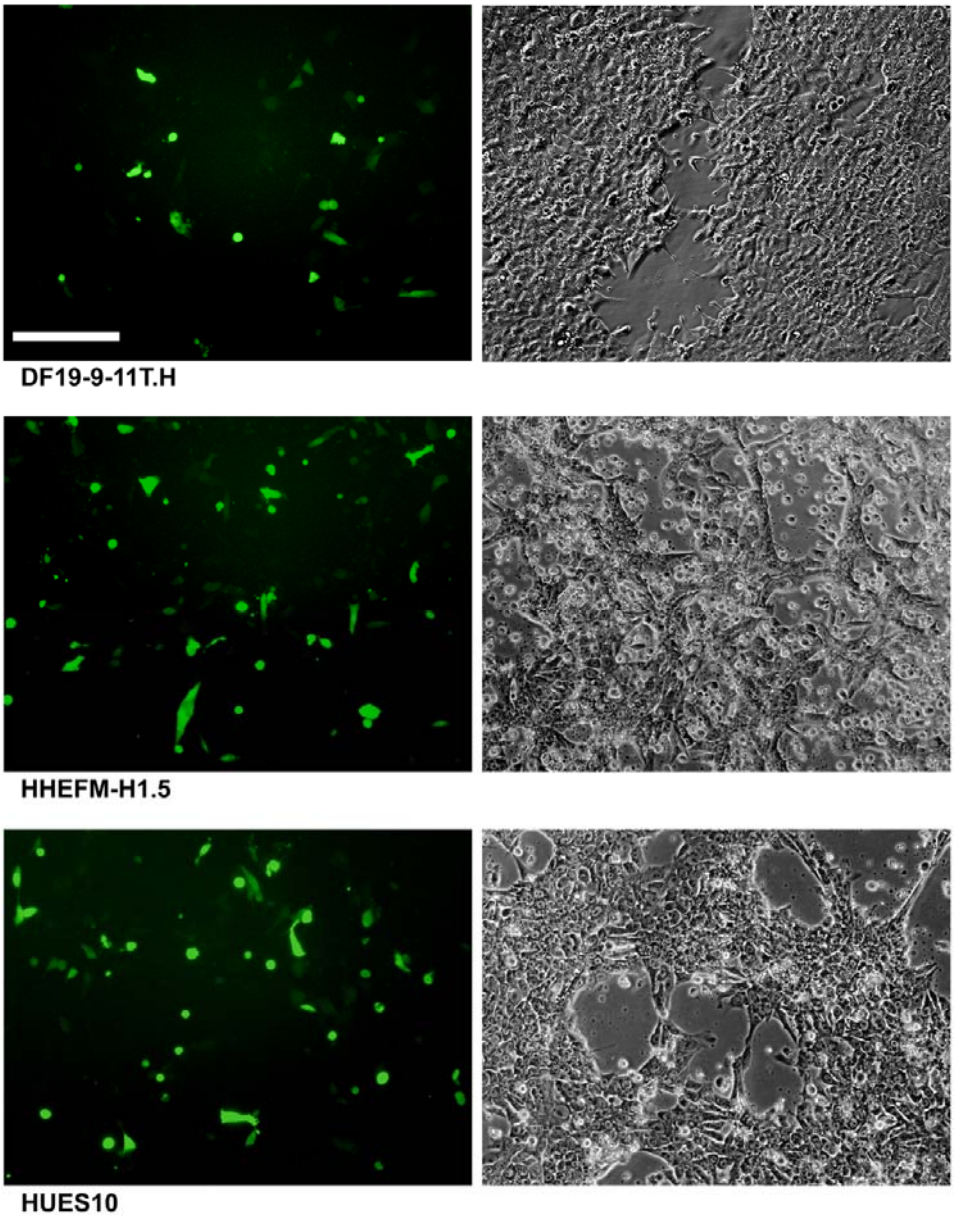

Figure 6-7 Transduction of DF19-9-11T.H, HHEFM-H1.5 and HUES10 with pHG 17α40 CAG-N2G HSV-1 amplicons

The iPS cells DF19-9-11T.H and HHEFM-H1.5, and the hES cells HUES10 were transduced with the pHG 17α40 CAG-N2G HAC vector at MOI 2 using the HSV-1 amplicon system. The left panel shows *GFP* expression of the transduced cells 48 hours after transduction. Corresponding phase contrast images are shown in the right panel. Scale bar = 200μm.

Table 6-1 Transduction efficiencies on DF19-9-11T.H, HHEFM-H1.5 and HUES10 cells

Cells	Transduction Efficiency (% + SE)
DF19-9-11T.H	15 ± 1.63
HHEFM-H1.5	25 ± 2.60
HUES10	18 ± 1.95

The presented transduction efficiency is based on the average transduction efficiency of two independently transduced wells. SE denotes standard error of the mean.

6.3.7 Generation of Stable Clones

Three days post-transduction, the transduced and the non-transduced cells (control) were transferred to inactivated SNL76/7 feeder cells in 6-well dishes in the hESc medium. Two days later, G418 selection was applied at concentrations of 50μg/ml, 100μg/ml and 200μg/ml to obtain stable clones. 10-12 days after the application of selection, the non-transduced cells were entirely eliminated from the wells containing 100μg/ml and 200μg/ml G418 but not from the wells containing 50μg/ml G418. The transduced iPS cells and hESc generated several stable clones in all the three G418 concentrations (Figure 6-8). Unlike the previous experiments of G418 selection of hiPS and HUES10 cells transduced with pHSV17α40Neo HSV-1 amplicons, in which the selection had to be removed in order to obtain the stable clones, the pHG 17α40 CAG-N2G HSV-1 amplicons transduced iPS cells and HUES10 cells resulted in the formation of stable clones while the G418 selection was maintained. Moreover, the stable clones also exhibited varying levels of GFP expression under the microscope as shown in the Figure 6-8. Table 6-2 lists the total number of clones that were obtained from the DF19-9-11T.H, HHEFM-H1.5 and HUES10 cells along with the levels of *GFP* expression that these clones exhibited.

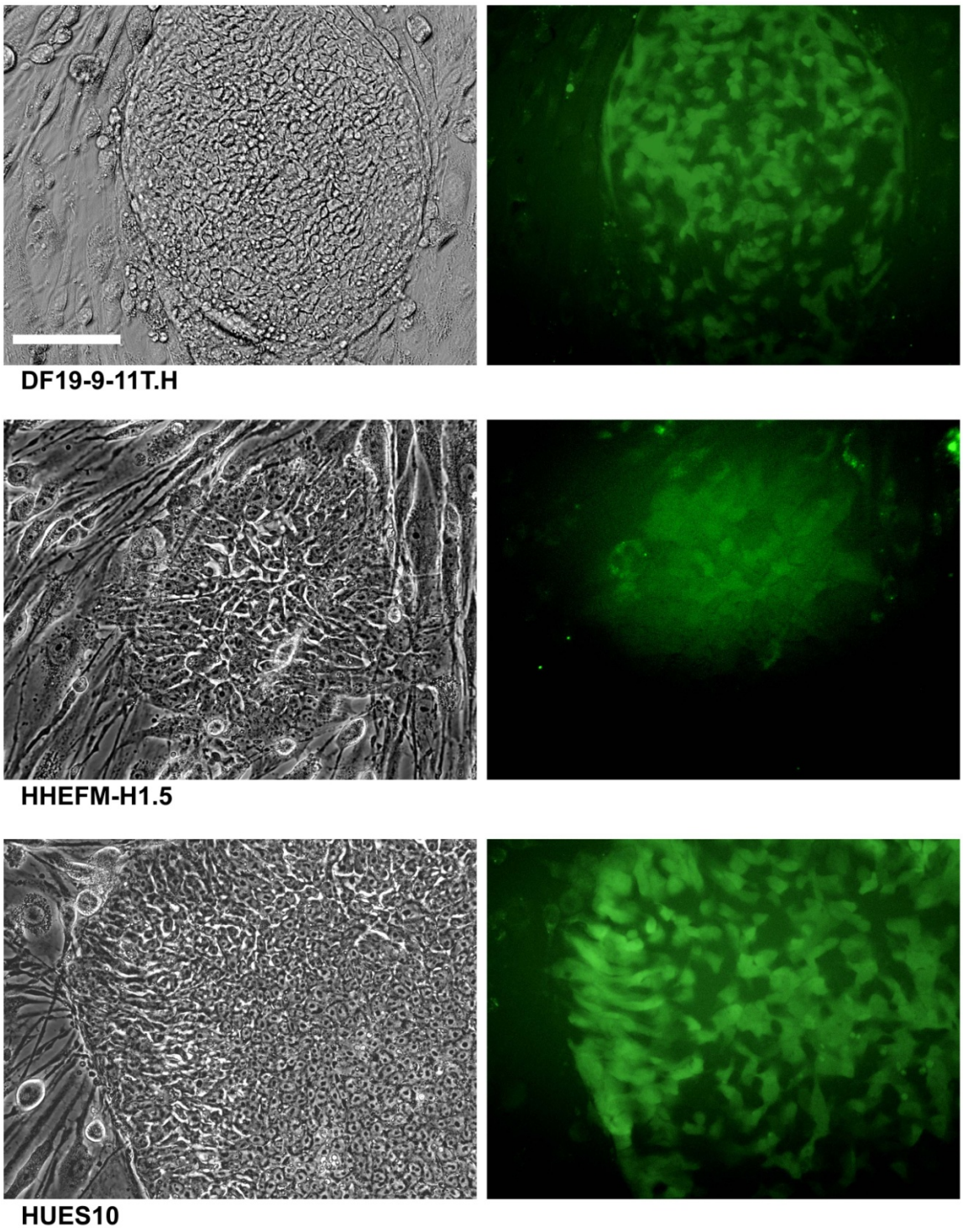

Figure 6-8 Stable hiPS and hESc clones from pHG 17α40 CAG-N2G transduction
Phase contrast images (left panel) of the clones obtained from the iPS cells DF19-9-11T.H and HHEFM-H1.5, and the hES cells HUES10. The clones exhibited variable levels of *GFP* expression under the microscope as shown in the right panel images. Scale bar = 100μm.

The clones were subsequently picked and transferred to a 24-well dish in individual wells, on inactivated SNL76/7 feeder cells in hESc medium. After their maintenance for 6-7 days, during which the colonies grew in size, cells were passaged onto SNL76/7-coated 24-well dishes for their expansion and subsequently grown on Matrigel-coated dishes in mTeSR-1 medium for further analysis. During the expansion, the cells were continuously maintained under the selection at 50μg/ml and 25μg/ml G418 on the feeder/hESc medium and Matrigel/mTeSR-1 medium conditions respectively.

To screen the clones containing the pHG 17α40 CAG-N2G HAC vector and to eliminate the false positive clones, genomic DNA isolated from the cells of each clone was subjected to PCR analysis using the primers sets specific for the *AcGFP* gene as described in the Section 6.2.3. As shown in the Figure 6-9, all the clones were found to be positive by PCR indicating that the stringency of drug selection was significantly improved with the modified HSV-1 HAC vector pHG 17α40 CAG-N2G. To identify the HAC containing cells, metaphase spreads from three stable clones derived from each of the three parental cell lines (DF19-9-11T.H, HHEFM-H1.5 and HUES10) were prepared and analysed by FISH. As shown in Figure 6-10, the DF19-9-11T.H clone DF.40.100G1 and the HHEFM-H1.5 clone 1.5.40.50G contained HAC. The HAC was detected in 20% and 13% of the metaphase spreads of DF.40.100G1 and 1.5.40.50G clones respectively. However, these two iPS clones and most of the other iPS clones also exhibited integration of HAC components into the host chromosomes as shown in Figure 6-10 and summarised in Table 6-3. Moreover, the three HUES10 clones contained integration and no HAC were identified in these clones. The overall analysis indicated that *de novo* HAC can be generated in the human iPS cells. However, the low frequency of HAC and the presence of integration events suggest that HAC formation was inefficient and therefore lead to the

requirement of further refinements in the HAC vector to improve HAC centromere formation as well as eliminate the possibility of integration.

Table 6-2 Analysis of the hiPS and hES clones obtained using the pHG 17α40 CAG-N2G vector

Cells	Clone	GFP expression (under microscope)	PCR Analysis
DF19-9-11T.H	DF.40.50G1[a]	-	N/A
	DF.40.50G2[a]	+/-	N/A
	DF.40.50G	+	Positive
	DF.40.100G1	++	Positive
	DF.40.100G2	+	Positive
	DF.40.200G	++	Positive
HHEFM-H1.5	1.5.40.50G	-	Positive
	1.5.40.100G	+/-	Positive
	1.5.40.200G[a]	+	N/A
	1.5.40.50G MG	++	Positive
HUES10	H10.40.100G1	+	Positive
	H10.40.100G2	++	Positive
	H10.40.100G3	+++	Positive
	H10.40.100G4[b]	-	N/A
	H10.40.200G1	++	Positive
	H10.40.200G2	+	Positive
	H10.40.200G3	+/-	Positive
	H10.40.200G4	+++	Positive

[a]Clone failed to grow after picking or passaging; [b]Clone was not picked; N/A = Not applicable; '-' = negative; '+/-' = faintly positive; '+' = positive.

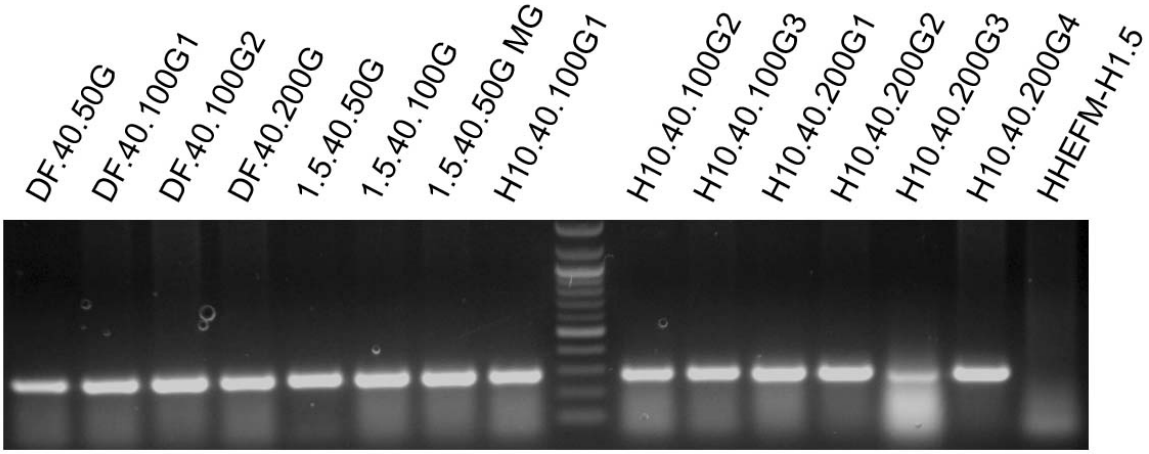

Figure 6-9 Genomic DNA PCR analysis of the stable clones
Genomic DNA of the stable clones of the iPS clones (DF19-9-11T.H and HHEFM-H1.5) and hES clones (HUES10) was subjected to the PCR for the *AcGFP* gene. The analysis indicated that all the analysed clones were positive by PCR. Genomic DNA of the iPS cells HHEFM-H1.5 was used as a negative control.

Table 6-3 Summary of FISH analysis of iPS and HUES10 stable clones

Cells	Clone	HAC (% of spreads)	Integration (% of spreads)
DF19-9-11T.H	DF.40.100G1	HAC (20%)	4
	DF.40.100G2	-ve	100
	DF.40.200G	-ve	0
HHEFM-H1.5	1.5.40.50G	HAC (13%)	80
	1.5.40.50G MG	-ve	100
	1.5.40.100G	-ve	100
HUES10	H10.40.100G3	-ve	100
	H10.40.200G2	-ve	100
	H10.40.200G3	-ve	100

'-ve' = negative for HAC.

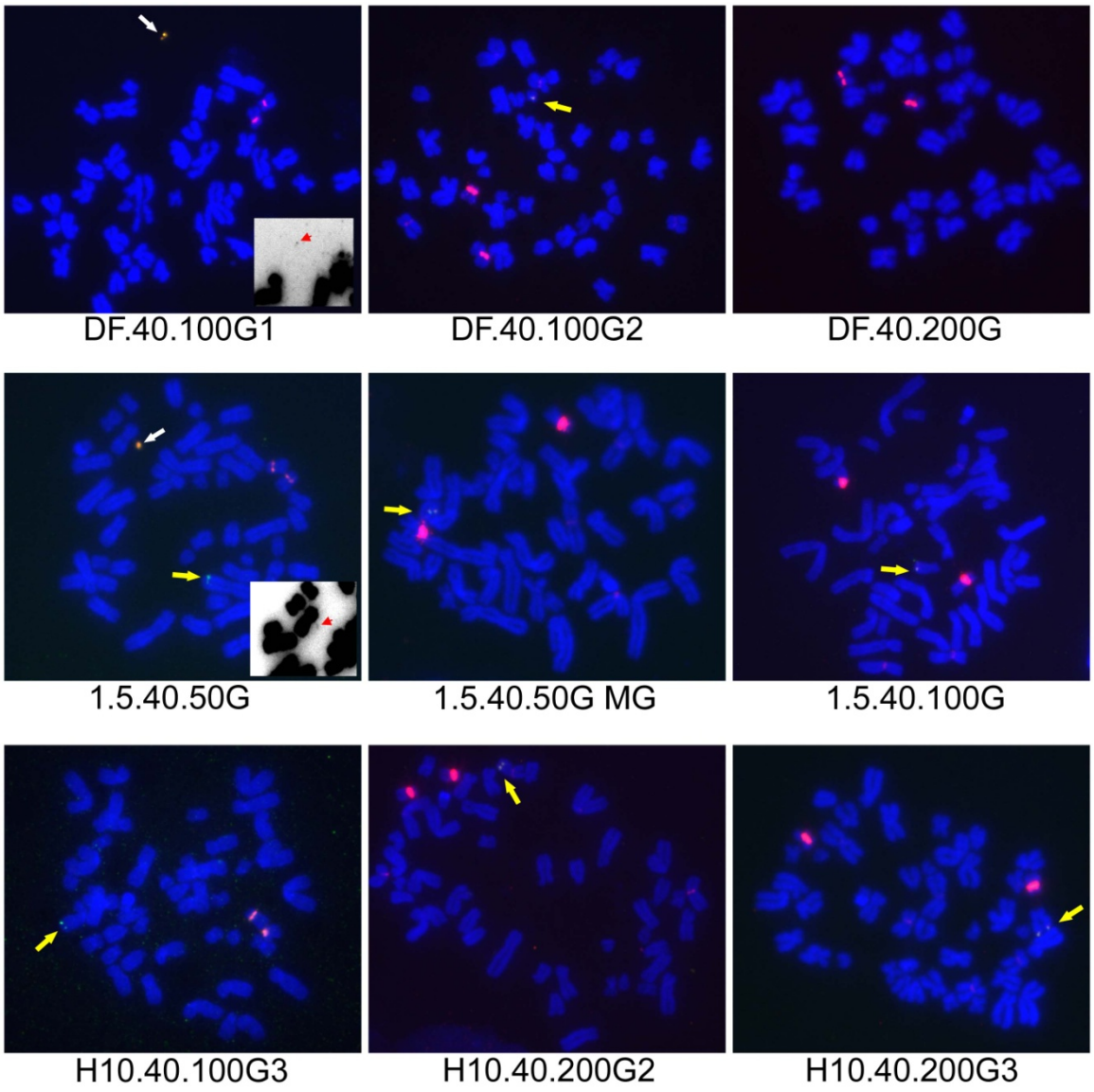

Figure 6-10 Analysis of the stable clones of iPS cells and hESc through FISH
FISH analysis of iPS and HUES10 clones generated with pHG 17α40 CAG-N2G vector. The 17α DNA probe is labelled in red and the pHG CAG-N2G vector probe is labelled in green. The chromosomes are counterstained by DAPI, blue. In the DF19-9-11T.H clone DF.40.100G1 and the HHEFM-H1.5 clone 1.5.40.50G, HAC (identified by white arrow) were detected. The inset shows DAPI staining only, where the HAC is identified by a red arrow. All the other clones, except the DF19-9-11T.H clone DF.40.200G, were shown to contain integrations (identified by yellow arrow) of HAC DNA into the host chromosome.

6.4 Discussion

In this study, the feasibility of human artificial chromosome (HAC) development in human induced pluripotent (hiPS) stem cells was assessed. The 17α HAC vector

(pHSV17α40Neo), which was shown to form mitotically stable and non-integrating HAC in the hESc HUES10 (Chapter 3) and HUES2 (Mandegar, 2011), was delivered into hiPS cells using the HSV-1 amplicon system. Transduction of the two hiPS cell lines, DF19-9-11T.H and HHEFM-H1.5, led to reasonable delivery efficiencies of up to 26-32% of the input cells.

Although the HAC vector was efficiently delivered into the hiPS cells, the determination of HAC development in the hiPS cells was affected by the selection of the HAC containing cells. During various attempts to recover stable clones after application of G418 selection, either all the cells died (at stringent selection conditions) or non-transduced cells survived (at less stringent selection conditions) and therefore resulted in many false-positive clones. The inability to derive stable clones from the pHSV17α40Neo transduction of hiPS cells suggested that the functional performance of hiPS cells was different from the hESc. Recent studies (Doi et al., 2009; Feng et al., 2010; Miura et al., 2009) have also supported this notion and have identified subtle but significant differences in gene expression signatures, DNA methylation patterns, *in vitro* differentiation and teratoma-forming abilities of the hiPS cells and hESc (discussed in Section 1.5.3). The inclusion of the hiPS cells HHEFM-H1.5 in the HAC generation experiments should have minimised the differences in the genetic background of the cells in comparison to HUES10 cells as HHEFM-H1.5 cells are derived from the embryonic fibroblasts differentiated from the HUES10 cells. Since there is no difference in the genetic background between the HUES10 and the HHEFM-H1.5 cells, the dissimilar behaviour observed in drug selection response could be due to (i) epigenetic effects (e.g. methylation, chromatin modification etc), (ii) differences in the methods, using which the hiPS cells and hESc cells are derived (Grskovic et al., 2011), (iii) and subtle differences

in the transduction methodology/efficiency etc. Any of these factors could have resulted in the instability of the HSV-1 HAC amplicons vectors, or hampered the amplification step required for the HAC formation, or more likely, increased the effect of silencing of the viral promoters such that the *neo* gene was switched off too early to allow selection of the transduced cells.

Silencing of the viral promoter that control the expression of the *neo* gene could be one of the major factors contributing to the inability to generate stable clones as described in the Section 6.3.4. To overcome the promoter mediated silencing, the HSV-1 HAC vector was modified to accommodate both the *neo* and the *GFP* reporter genes under a CMV early enhancer/chicken β actin (CAG) promoter that is known to drive stable expression of transgenes in hiPS and hESc (Chatterjee et al., 2011; Liew et al., 2007). In the HT1080 cells, the stable clone formation efficiency of the pHG 17α40 CAG-N2G vector was found to be significantly improved ($2.5x10^{-3}$) in comparison to that of pHSV17α40Neo vector ($6.3x10^{-4}$) (Mandegar, 2011).

With the use of the modified HAC vector (pHG 17α40 CAG-N2G), G418 selection of the hiPS cells as well as hESc was significantly improved as the generation and the propagation of the stable clones was possible. Thus, the pHG 17α40 CAG-N2G transduced cells were maintained in G418 selection for the overall duration of the experiment, while in the previous experiments with pHSV17α40Neo in the two hESc lines HUES10 (Chapter 3) and HUES2 (Mandegar, 2011), and the two hiPS cell lines DF19-9-11T.H and HHEFM-H1.5 (Section 6.3.3), the selection had to be removed (10-12 days post-transduction) to allow the expansion of the clones. Most importantly, the hiPS and hESc clones obtained using the pHG 17α40 CAG-N2G HAC vector also exhibited continuous *GFP* expression. Generation and particularly the propagation of these clones

in the presence of G418 selection and the visible *GFP* expression under the microscope indicated that the *neo* and the *GFP* gene, both of which are regulated by a single CAG promoter and the IRES2 linker, did not become silent for the entire duration of the experiment (from transduction to picking of the clones). The study comprises the first report in which the hiPS cells have been successfully modified via the HSV-1 amplicon system to deliver a HAC vector and express the gene of interest supporting the use of this system in deriving effective transgene expression in these clinically relevant pluripotent stem cells. The pHG 17α40 CAG-N2G HAC can be further improved by incorporation of the transgenes under the pCAG promoter (in which the CAG promoter is linked to the polyoma virus mutant enhancer PyF101), as the pCAG promoter, in comparison to the CAG promoter, is known for better circumventing the silencing event in hESc (Liew et al., 2007).

The stable clones obtained from the hiPS and hESc using the pHG 17α40 CAG-N2G HSV-1 HAC vector were also confirmed to contain vector DNA by PCR analysis of the *AcGFP* gene. These results indicated that none of the hiPS and hES clones were derived due to incomplete drug selection unlike several false-positive clones that were obtained using the previous HSV-1 HAC vector pHSV17α40Neo. The two iPS cell clones DF.40.100G1 and 1.5.40.50G were identified to contain HAC in a proportion ranging from 13-20%. However, integration events were observed in both HAC- and non- HAC containing clones of the iPS cells as well the control cells HUES10. The low frequency of HAC formation could be attributed to the instability of the generated HAC in the iPS cells. This could be confirmed with further analysis of the mitotic stability of the HAC in the iPS clones. In addition, the HAC containing iPS clones could be subjected to subcloning in order to isolate clones with possibly all HAC containing cells. However,

these experiments were not performed due to time constraints and therefore could not be included in this dissertation. Alternatively, it is possible that the presence of a strong promoter on the HAC construct may have interfered with the centromere seeding and maintenance, as indicated by a study conducted by Nakashima and colleagues (Nakashima et al., 2005). According to this study, the transcriptionally active regions flanking the α-satellite DNA were found to be deleterious for HAC formation. This phenomenon, if true, could explain why the pHSV17α40Neo vector did not cause any integration event in the HUES10 cells (Chapter 3) as opposed to the pHG 17α40 CAG-N2G vector. The reason for the integration of the modified HAC vector (pHG 17α40 CAG-N2G) in the iPS clones as well as HUES10 clones is not clearly understood but could be due to the same reason. Prolonged drug selection possibly resulted in the survival of the cells in which the HAC vector integrated into the chromosome. To improve the drug selection procedure, further refinements can be made to reduce the number of days the cells are exposed to the drug selection. Moreover, the formation and stability of the HAC is also known to be affected by the size of the α-satellite DNA (Okamoto et al., 2007) and therefore could be improved by increasing the size of the α-satellite DNA.

A recent study has reported the successful introduction of HAC carrying the distrophin gene in hiPS cells to correct Duchenne muscular dystrophy (DMD) genetic deficiency in these iPS cells (Kazuki et al., 2010). However, the HAC were not directly generated in the hiPS cells, but in CHO cells and subsequently transferred by microcell-mediated chromosome transfer (MMCT). The MMCT procedure, widely used to transfer chromosomes between cell lines, can sometime induce rearrangements which could potentially modify HAC structure and stability in the HAC (Alazami et al., 2004).

Therefore the work presented in this study to generate *de novo* HAC in hiPS cells could be more beneficial in terms of their utilisation in clinical settings for patient-specific gene therapy.

Chapter 7. Conclusions and Future Directions

7.1 Conclusion

The reprogramming of differentiated cells identity to a pluripotent state presents a wide spectrum of biomedical applications including disease modelling and drug screening on disease-specific iPS cells, and offers hopes for cell and tissue transplantation therapies. The potential promises offered by the pluripotent stem cells including both hESc and hiPS cells can be magnified by developing safe and efficient methods with which the cell can be genetically modified. The study presented here has rendered important contributions in establishing an integration free strategy for direct reprogramming as well as developing HAC mediated gene transfer into the hESc and hiPS cells utilising the HSV-1 amplicon system.

In chapter 3, the HSV-1 amplicon system was employed to establish gene expressing HAC in hESc, which were mitotically stable for up to 90 days without drug selection. The established HAC were shown to contain an active centromere as judged by CENP-C staining. Analysis of the HAC containing clones did not reveal any integration event indicating an important advancement towards gene therapy procedures, alleviating the risk of insertional mutagenesis. In addition the hESc cells containing the HAC retained their pluripotent capabilities and differentiation potential into all the three germ layers as well as into neuronal cells. The HSV-1 HAC vector could be modified with ease to accommodate ~100kb of additional DNA content, thus making the system very useful for incorporating large genes along with their regulatory elements and utilising it for more relevant cell types for clinical purposes such as iPS cells.

In Chapter 4, the feasibility of the HSV-1 amplicon system was evaluated to induce pluripotency in fibroblast cells. To achieve the objective, HSV-1 amplicon vectors were constructed to incorporate all the four reprogramming genes *OCT4*, *SOX2*, *KLF4* and *C-MYC* and were delivered to the IMR90 fibroblasts. With various attempts and slight modifications in the methodology, several iPS-like colonies were derived, which appeared to be blocked in a stage between somatic status and completely reprogrammed cells. These incompletely reprogrammed colonies were speculated to be derived following the loss of HSV-1 amplicon vectors from the proliferating cells and also due to the silencing of HSV-1 amplicon vectors in the fibroblasts. Both of these limitations were addressed by improvement of the HSV-1 amplicon vectors in the next chapter.

In Chapter 5, derivation of hiPS cells was successfully accomplished from embryonic fibroblasts using a hybrid HSV-1/EBNA-1/OriP amplicon system. The incorporation of EBNA-1/OriP elements facilitated the episomal retention of the HSV-1 amplicon vectors in the dividing cells during the reprogramming period as well as their removal from the cells after the reprogramming was completed to derive vector and transgene-free iPS cells. The overall study represents the first report of HSV-1 amplicon mediated generation of footprint free iPS cells, an important step towards their clinical use. As a control for the reprogramming experiments, the retroviral system was also employed to generate the iPS cells with nearly similar efficiency as compared to that of HSV-1 amplicon system.

In Chapter 6, the HSV-1 HAC system discussed in the Chapter 3 was extended to the hiPS cells. For this purpose the vector/transgene-free iPS cells derived in Chapter 5 and commercially available hiPS cells were subjected to HSV-1 HAC mediated gene delivery. To improve drug selection of the transduced hiPS cells, the HSV-1 HAC vector was

modified to incorporate CAG promoter driven neomycin resistance (*neo*) and green fluorescent protein (*GFP*) genes. The modified HAC vector not only significantly improved the selection procedure but also proved to be effective in obtaining *GFP* expressing iPS clones. Although, it is yet to be confirmed whether mitotically stable HAC were generated in the iPS cells, the successful employment of the HSV-1 HAC mediated gene delivery system in the hiPS cells has established its potential application in correcting the genetic defects in the disease-relevant iPS cells.

7.2 Future Implications: Basic Research

With the successful generation of pluripotent stem cells from a patient's somatic cells, the research area of pluripotent stem cells has flourished to a great extent. The ability to transform cell fate by either reprogramming or transdifferentiation has identified new dimensions of stem cell research into disease modelling and directed differentiation. However, a major concern is the equivalency of hiPS cell and hESc and therefore much research is still focused on evaluating molecular and functional differences between the two cell types, and improving the iPS derivation methods to minimise these variations (Robinton and Daley, 2012).

As the currently available strategies for generating integration-free iPS cells, including the one described in this study, suffer from low efficiencies, a possible extension of the HSV-1 amplicon system could be developed in generating iPS cells. The HSV-1 HAC vector presented in this study is capable of incorporating over 100kb of the foreign DNA and thus can be easily employed to incorporate all the reprogramming genes and its reprogramming potential can be investigated. Both the HSV-1/EBNA/OriP system and HSV-1HAC system utilised in this study can also be employed to deliver

transdifferentiation factors into the fibroblasts. Such a system would allow derivation of a desired cell type in an integration-free manner without involving the iPS generation step. For example, forced expression of the three transcription factors *BRN2*, *ASCL1* and *MYT1L* into the fibroblasts could be implemented to transdifferentiate them into induced neuronal cells (Pang et al., 2011). A similar strategy could be followed to transdifferentiate fibroblasts into induced cardiomyocytes (Ieda et al., 2010; Qian et al., 2012) and blood derivatives (Szabo et al., 2010).

There is considerable interest in the development of patient-derived disease models of human iPS cells, which can be subjected to differentiation into disease-specific lineages such as hepatic, cardiac and neural lineages. Upon differentiation into the desired lineage, one can study the relevant disease phenotype, investigate its pathophysiology and provide mechanistic insights into the disease development and its progression (Cherry and Daley, 2012). For example, iPS cells can be derived from the patients with Lesch-Nyhan syndrome (NLS), which is caused by *HPRT* deficiency and is associated with various neurological implications such as mental retardation, self-injury and motor dysfunction (Jinnah, 2009). Although it is known that mutations in the *HPRT* gene impair the purine salvage pathway (Baumeister and Frye, 1985; McDonald and Kelley, 1971), it is unknown why the nervous system is mainly affected by this imparity. Unfortunately, *Hprt*-deficient mouse models do not exhibit similar neurological symptoms to LNS-affected individuals (Engle et al., 1996; Finger et al., 1988). Therefore, iPS cells derived from the LNS-patients could allow the establishment of a robust disease model, which could be utilised for differentiation into neural lineages and to study the genotype-phenotype connections of the disease (Mekhoubad et al., 2012).

In addition to providing disease-specific iPS cell models, genetic modification of both the hESc and hiPS cells has also received wide interest within the scientific community. An ideal gene delivery method should be based on the efficient and integration-free system that not only allows stable gene expression in the targeted ES and iPS cells but also in various differentiated lineages derived from those ES and iPS cells (Irion et al., 2007). The ability of HSV-1-HAC to efficiently deliver single or multiple gene of interests in hESc and hiPS cells with reduced risk of integration could serve to accomplish these objectives in both the targeted cells and their differentiated lineages. As the HSV-1-HAC system exhibited gradual silencing of the reporter genes in the hESc and hiPS cells, various strategies can be adapted to overcome this loss of gene expression. For example, the gene of interest can be placed under ubiquitously expressed promoter or a tissue-specific promoter to derive its expression in a particular lineage to be studied (Lee et al., 2009c; Matrai et al., 2010). Perhaps, a better way to avoid transgene silencing could be achieved by targeting a gene of interest in a specific region that can escape the silencing. One such region is the *hROSA26* locus, which is an established targeting region to resist the transgene silencing (Irion et al., 2007). To achieve this objective, the HSV-1 HAC construct can be modified to incorporate the *hROSA26* locus followed by the incorporation of the target gene into the first intron of the *hROSA26* region. This would allow developing a system that can efficiently deliver the gene of interest in an integration-free fashion as well as avoiding the transgene silencing. The system can further be modified to include specific promoters if tissue-specific expression is required and hence can be utilised in both basic and therapeutic settings (Irion et al., 2007). An alternative way to prevent transgene silencing could be achieved through the conservation of an open chromatin state by flanking the transgene and its regulatory sequences with

genetic insulators such as chicken beta-globin HS4 insulator (cHS4) or gamma-satellite arrays (Chen et al., 2008b; Gaszner and Felsenfeld, 2006; Kim et al., 2009c).

Presently, one of the active areas of stem cell biology is the identification of the transcription factors and small molecules which can facilitate the differentiation of ES and iPS cells into the desired lineages (Cherry and Daley, 2012; Dontu et al., 2004; Livesey and Cepko, 2001; Stern and Fraser, 2001). To accomplish this objective, HSV-1-HAC system could be employed to target the ES and iPS cells by incorporating reporter genes under the tissue-specific promoters. These reporter cell lines can assist in lineage tracing and could be utilised to determine the factors that can stimulate specific lineage commitment, through high throughput screening methods (Bushway and Mercola, 2006; Stegmaier et al., 2004).

In recent studies, our group has established dual transduction of HSV-1 amplicon system into HT1080 cells and demonstrated that by employing this system the capacity of the HSV-1 amplicon system can theoretically be increased two fold (unpublished data; (Chan, 2010)). Preliminarily experiments have indicated that two independently transduced HSV-1 amplicon vectors, for example, a HAC vector containing α-satellite DNA and a second vector bearing a gene of interest can co-localise to form a single HAC in double drug selected HT1080 clones and HUES2. Although the current efficiency of this system is low, the findings are potentially important and it would be even more significant if the system can be further developed and extended to the clinically relevant cell types such as hiPS cells.

7.3 Future Implications: Clinical Applications

The ability to induce pluripotency in somatic cells using the forced-expression of reprogramming factors has revolutionised the medical approaches in terms of creating opportunities for disease modelling and developing prospects for personalised cell therapies. Although significant progress has been made to derive iPS cells without interfering with the somatic genome, there are crucial challenges that need to be resolved before the use of iPS cells can be harnessed in therapeutic applications (Robinton and Daley, 2012). These include (1) developing more-uniform methodologies and more stringent controls to acquire experimental consistency among different laboratories; (2) following practices that could help maintaining their karyotypic integrity; (3) defining better strategies to evaluate their functional equivalence to hESc; (4) establishing new and streamlined procedures for their directed differentiation. Another important accomplishment in employing the iPS cells in cell transplantation therapies would be to improve the current methods for introducing genes or repairing genetic defects responsible for a particular disease. The HSV-1 HAC system presented in this study can be applied to deliver single or multiple genes of interest along with their regulatory elements into the disease relevant iPS cells to provide the missing functions.

To date, the feasibility of *de novo* HAC establishment using the HSV-1 amplicon system has only been examined in two pluripotent cell types hESc (HUES2 and HUES10) (Mandegar et al., 2011) and hiPS cells (DF19-9-11T.H and HHEFM-H1.5). However, research on the development of HAC, their characterisation and their functional behaviour is still in the study and development phase, and therefore the employment of HAC in the clinical settings is still far off. In order to advance this system toward therapeutic purposes, various steps need to be taken to comply with the standards and

guidelines of good manufacturing practice (GMP) (Hewitt et al., 2007). Primarily, the safety and the efficacy of the HAC system would need to be rigorously examined in other cell systems such as additional hES cell lines, disease-specific iPS cells and other multipotent stem cell populations. Safety assessment would include extensive analysis of the HAC containing cells to measure the risk of integration events. As some instances in this study have suggested that the generated HAC could be too small to be detected with the current FISH techniques, improvements in the FISH procedures such as incorporation of quantum dots fluorophores could enhance the sensitivity of detecting small HAC signals (Ioannou et al., 2009). In addition, high throughput screening through splinkerette-PCR method (Uren et al., 2009) could be utilised to detect any integration events in the HAC containing cells so that any such cells can be eliminated. To evaluate the efficacy of the HSV-1 HAC system, the mitotic stability of the HAC and the physiological expression of the transgenes would have to be investigated for an extended period of time. Moreover, derivation of the stable clones, their screening and functional validity of the HAC containing cells would need to be essentially optimised to satisfy GMP measures for the employment of HAC system in gene-based cells therapies.

Once the safety and the efficacy challenges associated with both the iPS cells and the HSV-1 HAC system are addressed, HAC can be potentially employed in patient-specific *ex vivo* gene therapy coupled to cell therapy. For example gene therapy for Hemophilia A is hampered by the size of the factor VIII (*FVIII*) cDNA (>8kb) and therefore the cDNA is generally split into heavy and light chain fragments to accommodate it into viral vectors (Murphy and High, 2008). Using the capacity that the HSV-1 HAC system provides, the *FVIII* can be easily incorporated into the HAC vector along with a liver-specific promoter such as human α1 antitrypsin (hAAT) promoter (Matsui et al., 2011) to

acquire physiological expression of *FVIII*. The vector can then be administered *ex vivo* either into the patient derived iPS cells followed by differentiation into liver cells or into the mature hepatocytes and subsequently the modified cells expressing physiological levels of *FVIII* can be transplanted into the Hemophilia A patients for therapy.

Our group is also aiming to utilise the HSV-1 HAC system to express the Xeroderma Pigmentosum, complementation group A (*XPA*) gene in XPA patients-derived iPS cells. XPA is one of the crucial proteins involved in DNA excision repair and its deficiency is associated with accumulation of bulky adducts on the DNA molecule upon exposure to the UV light (Cleaver, 2005; Kraemer et al., 1994; Magnaldo, 2004). Due to impaired nucleotide excision repair system, the XPA patients are extremely sensitive to the UV light. The corrected *XPA* iPS cells can be used as disease model to understand the disease pathology and progression, tested for differentiation into skin progenitor cells for confirmation of regulated *XPA* expression and utilised for potential application in gene therapy.

Chapter 8. References

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G.*, et al.* (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol *26*, 1276-1284.

Acsadi, G., O'Hagan, D., Lochmuller, H., Prescott, S., Larochelle, N., Nalbantoglu, J., Jani, A., and Karpati, G. (1998). Interferons impair early transgene expression by adenovirus-mediated gene transfer in muscle cells. J Mol Med (Berl) *76*, 442-450.

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S.*, et al.* (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat Biotechnol *25*, 803-816.

Aizawa, E., Hirabayashi, Y., Iwanaga, Y., Suzuki, K., Sakurai, K., Shimoji, M., Aiba, K., Wada, T., Tooi, N., Kawase, E.*, et al.* (2012). Efficient and accurate homologous recombination in hESCs and hiPSCs using helper-dependent adenoviral vectors. Mol Ther *20*, 424-431.

Alazami, A.M., Mejia, J.E., and Monaco, Z.L. (2004). Human artificial chromosomes containing chromosome 17 alphoid DNA maintain an active centromere in murine cells but are not stable. Genomics *83*, 844-851.

Albanell, J., Bosl, G.J., Reuter, V.E., Engelhardt, M., Franco, S., Moore, M.A., and Dmitrovsky, E. (1999). Telomerase activity in germ cell cancers and mature teratomas. J Natl Cancer Inst *91*, 1321-1326.

Alexandrov, I., Kazakov, A., Tumeneva, I., Shepelev, V., and Yurov, Y. (2001). Alpha-satellite DNA of primates: old and new families. Chromosoma *110*, 253-266.

Allsopp, R.C., Chang, E., Kashefi-Aazam, M., Rogaev, E.I., Piatyszek, M.A., Shay, J.W., and Harley, C.B. (1995). Telomere shortening is associated with cell division in vitro and in vivo. Exp Cell Res *220*, 194-200.

Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev Biol *227*, 271-278.

Amit, M., Chebath, J., Margulets, V., Laevsky, I., Miropolsky, Y., Shariki, K., Peri, M., Blais, I., Slutsky, G., Revel, M.*, et al.* (2010). Suspension culture of undifferentiated human embryonic and induced pluripotent stem cells. Stem Cell Rev *6*, 248-259.

Amit, M., and Itskovitz-Eldor, J. (2002). Derivation and spontaneous differentiation of human embryonic stem cells. J Anat *200*, 225-232.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H. (2004). Human centromere repositioning "in progress". Proc Natl Acad Sci U S A *101*, 6542-6547.

Amor, D.J., and Choo, K.H. (2002). Neocentromeres: role in human disease, evolution, and centromere study. Am J Hum Genet *71*, 695-714.

Andrews, P.W., Przyborski, S.A., and Thomson, J.A. (2001). Embryonal Carcinoma Cells as Embryonic Stem Cells. In Stem cell biology, D.R. Marshak, R.L. Gardner, and D. Gottlieb, eds. (New York, Cold Spring Harbor Laboratory Press), pp. 231–266.

Ang, Y.S., Tsai, S.Y., Lee, D.F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B.*, et al.* (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell *145*, 183-197.

Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A.*, et al.* (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell *8*, 376-388.

Antinone, S.E., Shubeita, G.T., Coller, K.E., Lee, J.I., Haverlock-Moyns, S., Gross, S.P., and Smith, G.A. (2006). The Herpesvirus capsid surface protein, VP26, and the majority of the tegument proteins are dispensable for capsid transport toward the nucleus. J Virol *80*, 5494-5498.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science *321*, 699-702.

Asuri, P., Bartel, M.A., Vazin, T., Jang, J.H., Wong, T.B., and Schaffer, D.V. (2012). Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. Mol Ther *20*, 329-338.

Baird, D.M., and Farr, C.J. (2006). The organization and function of chromosomes. EMBO Rep *7*, 372-376.

Baker, D.E., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., and Andrews, P.W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. Nat Biotechnol *25*, 207-215.

Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M.*, et al.* (2009). Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev *23*, 2134-2139.

Bar-Nur, O., Russ, H.A., Efrat, S., and 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.

Barnett, M.A., Buckle, V.J., Evans, E.P., Porter, A.C., Rout, D., Smith, A.G., and Brown, W.R. (1993). Telomere directed fragmentation of mammalian chromosomes. Nucleic Acids Res *21*, 27-36.

Barry, A.E., Howman, E.V., Cancilla, M.R., Saffery, R., and Choo, K.H. (1999). Sequence analysis of an 80 kb human neocentromere. Hum Mol Genet *8*, 217-227.

Basu, J., Stromberg, G., Compitello, G., Willard, H.F., and Van Bokkelen, G. (2005). Rapid creation of BAC-based human artificial chromosome vectors by transposition with synthetic alpha-satellite arrays. Nucleic Acids Res *33*, 587-596.

Baumeister, A.A., and Frye, G.D. (1985). The biochemical basis of the behavioral disorder in the Lesch-Nyhan syndrome. Neurosci Biobehav Rev *9*, 169-178.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bestor, T.H. (2000). Gene silencing as a threat to the success of gene therapy. J Clin Invest *105*, 409-411.

Blaese, R.M., Culver, K.W., Miller, A.D., Carter, C.S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P.*, et al.* (1995). T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. Science *270*, 475-480.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H.*, et al.* (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144*, 439-452.

Boiani, M., and Scholer, H.R. (2005). Regulatory networks in embryo-derived pluripotent stem cells. Nat Rev Mol Cell Biol *6*, 872-884.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Braam, S.R., Denning, C., van den Brink, S., Kats, P., Hochstenbach, R., Passier, R., and Mummery, C.L. (2008). Improved genetic manipulation of human embryonic stem cells. Nat Methods *5*, 389-392.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

Brambrink, T., Hochedlinger, K., Bell, G., and Jaenisch, R. (2006). ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. Proc Natl Acad Sci U S A *103*, 933-938.

Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proc Natl Acad Sci U S A *38*, 455-463.

Brokhman, I., Pomp, O., Fishman, L., Tennenbaum, T., Amit, M., Itzkovitz-Eldor, J., and Goldstein, R.S. (2009). Genetic modification of human embryonic stem cells with adenoviral vectors: differences of infectability between lines and correlation of infectability with expression of the coxsackie and adenovirus receptor. Stem Cells Dev *18*, 447-456.

Bushway, P.J., and Mercola, M. (2006). High-throughput screening for modulators of stem cell differentiation. Methods Enzymol *414*, 300-316.

Cai, L., Ye, Z., Zhou, B.Y., Mali, P., Zhou, C., and Cheng, L. (2007). Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res *17*, 62-72.

Caisander, G., Park, H., Frej, K., Lindqvist, J., Bergh, C., Lundin, K., and Hanson, C. (2006). Chromosomal integrity maintained in five human embryonic stem cell lines after prolonged in vitro culture. Chromosome Res *14*, 131-137.

Campadelli, G., Brandimarti, R., Di Lazzaro, C., Ward, P.L., Roizman, B., and Torrisi, M.R. (1993). Fragmentation and dispersal of Golgi proteins and redistribution of glycoproteins and glycolipids processed through the Golgi apparatus after infection with herpes simplex virus 1. Proc Natl Acad Sci U S A *90*, 2798-2802.

Campbell, M.S., and Gorbsky, G.J. (1995). Microinjection of mitotic cells with the 3F3/2 anti-phosphoepitope antibody delays the onset of anaphase. J Cell Biol *129*, 1195-1204.

Catalina, P., Montes, R., Ligero, G., Sanchez, L., de la Cueva, T., Bueno, C., Leone, P.E., and Menendez, P. (2008). Human ESCs predisposition to karyotypic instability: Is a matter of culture adaptation or differential vulnerability among hESC lines due to inherent properties? Mol Cancer *7*, 76.

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K.*, et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature *467*, 318-322.

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

Chan, D.Y.L. (2010). Analysis of Artificial Chromosomes and Factors Affecting Stability in Murine and Human Cultured and Embryonic Stem Cells. In DPhil Thesis (University of Oxford).

Chan, E.M., Ratanasirintrawoot, S., Park, I.H., Manos, P.D., Loh, Y.H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T.A.*, et al.* (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol *27*, 1033-1037.

Chan, G.K., Liu, S.T., and Yen, T.J. (2005). Kinetochore structure and function. Trends Cell Biol *15*, 589-598.

Chan, K.K., Wu, S.M., Nissom, P.M., Oh, S.K., and Choo, A.B. (2008). Generation of high-level stable transgene expressing human embryonic stem cell lines using Chinese hamster elongation factor-1 alpha promoter system. Stem Cells Dev *17*, 825-836.

Chatterjee, P., Cheung, Y., and Liew, C. (2011). Transfecting and nucleofecting human induced pluripotent stem cells. J Vis Exp.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J.*, et al.* (2008a). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Chen, Z.Y., He, C.Y., Ehrhardt, A., and Kay, M.A. (2003). Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. Mol Ther *8*, 495-500.

Chen, Z.Y., He, C.Y., Meuse, L., and Kay, M.A. (2004). Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. Gene Ther *11*, 856-864.

Chen, Z.Y., Riu, E., He, C.Y., Xu, H., and Kay, M.A. (2008b). Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. Mol Ther *16*, 548-556.

Cheng, J., DeCaprio, J.A., Fluck, M.M., and Schaffhausen, B.S. (2009). Cellular transformation by Simian Virus 40 and Murine Polyoma Virus T antigens. Semin Cancer Biol *19*, 218-228.

Cherry, A.B., and Daley, G.Q. (2012). Reprogramming cellular identity for regenerative medicine. Cell *148*, 1110-1122.

Chin, A.C., Fong, W.J., Goh, L.T., Philp, R., Oh, S.K., and Choo, A.B. (2007). Identification of proteins from feeder conditioned medium that support human embryonic stem cells. J Biotechnol *130*, 320-328.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J.*, et al.* (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell *5*, 111-123.

Chou, B.K., Mali, P., Huang, X., Ye, Z., Dowey, S.N., Resar, L.M., Zou, C., Zhang, Y.A., Tong, J., and Cheng, L. (2011). Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. Cell Res *21*, 518-529.

Chung, Y., Klimanskaya, I., Becker, S., Li, T., Maserati, M., Lu, S.J., Zdravkovic, T., Ilic, D., Genbacev, O., Fisher, S.*, et al.* (2008). Human embryonic stem cell lines generated without embryo destruction. Cell Stem Cell *2*, 113-117.

Cibelli, J.B., Campbell, K.H., Seidel, G.E., West, M.D., and Lanza, R.P. (2002). The health profile of cloned animals. Nat Biotechnol *20*, 13-14.

Cleaver, J.E. (2005). Cancer in xeroderma pigmentosum and related disorders of DNA repair. Nat Rev Cancer *5*, 564-573.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Co, D.O., Borowski, A.H., Leung, J.D., van der Kaa, J., Hengst, S., Platenburg, G.J., Pieper, F.R., Perez, C.F., Jirik, F.R., and Drayer, J.I. (2000). Generation of transgenic mice and germline transmission of a mammalian artificial chromosome introduced into embryos by pronuclear microinjection. Chromosome Res *8*, 183-191.

Coleman, H.M., Connor, V., Cheng, Z.S., Grey, F., Preston, C.M., and Efstathiou, S. (2008). Histone modifications associated with herpes simplex virus type 1 genomes during quiescence and following ICP0-mediated de-repression. J Gen Virol *89*, 68-77.

Cooke, C.A., Bernat, R.L., and Earnshaw, W.C. (1990). CENP-B: a major human centromere protein located beneath the kinetochore. J Cell Biol *110*, 1475-1488.

Cooke, C.A., Schaar, B., Yen, T.J., and Earnshaw, W.C. (1997). Localization of CENP-E in the fibrous corona and outer plate of mammalian kinetochores from prometaphase through anaphase. Chromosoma *106*, 446-455.

Costa, M., Dottori, M., Sourris, K., Jamshidi, P., Hatzistavrou, T., Davis, R., Azzola, L., Jackson, S., Lim, S.M., Pera, M.*, et al.* (2007). A method for genetic modification of human embryonic stem cells using electroporation. Nat Protoc *2*, 792-796.

Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369-1373.

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D.*, et al.* (2004). Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med *350*, 1353-1356.

Craig, J.M., Earnshaw, W.C., and Vagnarelli, P. (1999). Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Exp Cell Res *246*, 249-262.

Cuchet, D., Potel, C., Thomas, J., and Epstein, A.L. (2007). HSV-1 amplicon vectors: a promising and versatile tool for gene delivery. Expert Opin Biol Ther *7*, 975-995.

Davis, A.S., Zhao, H., Sun, G.H., Sapolsky, R.M., and Steinberg, G.K. (2007). Gene therapy using SOD1 protects striatal neurons from experimental stroke. Neurosci Lett *411*, 32-36.

de Silva, S., and Bowers, W.J. (2009). Herpes Virus Amplicon Vectors. Viruses *1*, 594-629.

Deng, J., Shoemaker, R., Xie, B., Gore, A., LeProust, E.M., Antosiewicz-Bourget, J., Egli, D., Maherali, N., Park, I.H., Yu, J.*, et al.* (2009). Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. Nat Biotechnol *27*, 353-360.

Depinet, T.W., Zackowski, J.L., Earnshaw, W.C., Kaffe, S., Sekhon, G.S., Stallard, R., Sullivan, B.A., Vance, G.H., Van Dyke, D.L., Willard, H.F., *et al.* (1997). Characterization of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. Hum Mol Genet *6*, 1195-1204.

Desponts, C., and Ding, S. (2010). Using small molecules to improve generation of induced pluripotent stem cells from somatic cells. Methods Mol Biol *636*, 207-218.

Do, J.T., and Scholer, H.R. (2004). Nuclei of embryonic stem cells reprogram somatic cells. Stem Cells *22*, 941-949.

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., *et al.* (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet *41*, 1350-1353.

Donovan, P.J., and Gearhart, J. (2001). The end of the beginning for pluripotent stem cells. Nature *414*, 92-97.

Dontu, G., Jackson, K.W., McNicholas, E., Kawamura, M.J., Abdallah, W.M., and Wicha, M.S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. Breast Cancer Res *6*, R605-615.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat Biotechnol *22*, 53-54.

Dravid, G., Ye, Z., Hammond, H., Chen, G., Pyle, A., Donovan, P., Yu, X., and Cheng, L. (2005). Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. Stem Cells *23*, 1489-1501.

Earnshaw, W.C., and Migeon, B.R. (1985). Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma *92*, 290-296.

Earnshaw, W.C., Ratrie, H., 3rd, and Stetten, G. (1989). Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. Chromosoma *98*, 1-12.

Ebersole, T.A., Ross, A., Clark, E., McGill, N., Schindelhauer, D., Cooke, H., and Grimes, B. (2000). Mammalian artificial chromosome formation from circular alphoid input DNA does not require telomere repeats. Hum Mol Genet *9*, 1623-1631.

Edelstein, M.L., Abedi, M.R., and Wixon, J. (2007). Gene therapy clinical trials worldwide to 2007--an update. J Gene Med *9*, 833-842.

Eiges, R., Schuldiner, M., Drukker, M., Yanuka, O., Itskovitz-Eldor, J., and Benvenisty, N. (2001). Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. Curr Biol *11*, 514-518.

Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. Stem Cells *26*, 2467-2474.

Engle, S.J., Womer, D.E., Davies, P.M., Boivin, G., Sahota, A., Simmonds, H.A., Stambrook, P.J., and Tischfield, J.A. (1996). HPRT-APRT-deficient mice are not a model for lesch-nyhan syndrome. Hum Mol Genet *5*, 1607-1610.

Epstein, A.L. (2009a). HSV-1-derived amplicon vectors: recent technological improvements and remaining difficulties--a review. Mem Inst Oswaldo Cruz *104*, 399-410.

Epstein, A.L. (2009b). Progress and prospects: biological properties and technological advances of herpes simplex virus type 1-based amplicon vectors. Gene Ther *16*, 709-715.

Epstein, A.L., Marconi, P., Argnani, R., and Manservigi, R. (2005). HSV-1-derived recombinant and amplicon vectors for gene transfer and gene therapy. Curr Gene Ther *5*, 445-458.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S.*, et al.* (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71-79.

Evans, M. (2005). Ethical sourcing of human embryonic stem cells--rational solutions? Nat Rev Mol Cell Biol *6*, 663-667.

Fan, X., Brun, A., and Karlsson, S. (2000). Adenoviral vector design for high-level transgene expression in primitive human hematopoietic progenitors. Gene Ther *7*, 2132-2138.

Farr, C., Fantes, J., Goodfellow, P., and Cooke, H. (1991). Functional reintroduction of human telomeres into mammalian cells. Proc Natl Acad Sci U S A *88*, 7006-7010.

Farr, C.J., Bayne, R.A., Kipling, D., Mills, W., Critcher, R., and Cooke, H.J. (1995). Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. EMBO J *14*, 5444-5454.

Farr, C.J., Stevanovic, M., Thomson, E.J., Goodfellow, P.N., and Cooke, H.J. (1992). Telomere-associated chromosome fragmentation: applications in genome manipulation and analysis. Nat Genet *2*, 275-282.

Feng, Q., Lu, S.J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., Honig, G.R., Kim, K.S., and Lanza, R. (2010). Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. Stem Cells *28*, 704-712.

Finger, S., Heavens, R.P., Sirinathsinghji, D.J., Kuehn, M.R., and Dunnett, S.B. (1988). Behavioral and neurochemical evaluation of a transgenic mouse model of Lesch-Nyhan syndrome. J Neurol Sci *86*, 203-213.

Fink, D.J., Wechuck, J., Mata, M., Glorioso, J.C., Goss, J., Krisky, D., and Wolfe, D. (2011). Gene therapy for pain: results of a phase I clinical trial. Ann Neurol *70*, 207-212.

Fischbach, G.D., and Fischbach, R.L. (2004). Stem cells: science, policy, and ethics. J Clin Invest *114*, 1364-1370.

Fischer, A., Hacein-Bey-Abina, S., and Cavazzana-Calvo, M. (2010). 20 years of gene therapy for SCID. Nat Immunol *11*, 457-460.

Fong, H., Hohenstein, K.A., and Donovan, P.J. (2008). Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. Stem Cells *26*, 1931-1938.

Fraefel, C., Jacoby, D.R., Lage, C., Hilderbrand, H., Chou, J.Y., Alt, F.W., Breakefield, X.O., and Majzoub, J.A. (1997). Gene transfer into hepatocytes mediated by helper virus-free HSV/AAV hybrid vectors. Mol Med *3*, 813-825.

Friedmann, T., and Roblin, R. (1972). Gene therapy for human genetic disease? Science *175*, 949-955.

Friedrich Ben-Nun, I., and Benvenisty, N. (2006). Human embryonic stem cells as a cellular model for human disorders. Mol Cell Endocrinol *252*, 154-159.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Jpn Acad Ser B Phys Biol Sci *85*, 348-362.

Garber, D.A., Beverley, S.M., and Coen, D.M. (1993). Demonstration of circularization of herpes simplex virus DNA following infection using pulsed field gel electrophoresis. Virology *197*, 459-462.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet *7*, 703-713.

Genbacev, O., Krtolica, A., Zdravkovic, T., Brunette, E., Powell, S., Nath, A., Caceres, E., McMaster, M., McDonagh, S., Li, Y.*, et al.* (2005). Serum-free derivation of human embryonic stem cell lines on human placental fibroblast feeders. Fertil Steril *83*, 1517-1529.

Gharwan, H., Hirata, R.K., Wang, P., Richard, R.E., Wang, L., Olson, E., Allen, J., Ware, C.B., and Russell, D.W. (2007). Transduction of human embryonic stem cells by foamy virus vectors. Mol Ther *15*, 1827-1833.

Gilbert, D.M. (2001). Making sense of eukaryotic DNA replication origins. Science *294*, 96-100.

Gilbert, D.M. (2004). In search of the holy replicator. Nat Rev Mol Cell Biol *5*, 848-855.

Gill, D.R., Smyth, S.E., Goddard, C.A., Pringle, I.A., Higgins, C.F., Colledge, W.H., and Hyde, S.C. (2001). Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1alpha promoter. Gene Ther *8*, 1539-1546.

Giorgetti, A., Montserrat, N., Aasen, T., Gonzalez, F., Rodriguez-Piza, I., Vassena, R., Raya, A., Boue, S., Barrero, M.J., Corbella, B.A.*, et al.* (2009). Generation of induced

pluripotent stem cells from human cord blood using OCT4 and SOX2. Cell Stem Cell *5*, 353-357.

Glorioso, J.C., DeLuca, N.A., and Fink, D.J. (1995). Development and application of herpes simplex virus vectors for human gene therapy. Annu Rev Microbiol *49*, 675-710.

Goins, W.F., Krisky, D.M., Wechuck, J.B., Wolfe, D., Huang, S., and Glorioso, J.C. (2011). Generation of replication-competent and -defective HSV vectors. Cold Spring Harb Protoc *2011*, pdb prot5615.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E.*, et al.* (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63-67.

Greber, B., Lehrach, H., and Adjaye, J. (2008). Control of early fate decisions in human ES cells by distinct states of TGFbeta pathway activity. Stem Cells Dev *17*, 1065-1077.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P.*, et al.* (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. Cell Stem Cell *6*, 215-226.

Grimes, B.R., Babcock, J., Rudd, M.K., Chadwick, B., and Willard, H.F. (2004). Assembly and characterization of heterochromatin and euchromatin on human artificial chromosomes. Genome Biol *5*, R89.

Grimes, B.R., Rhoades, A.A., and Willard, H.F. (2002). Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation. Mol Ther *5*, 798-805.

Grimes, B.R., Schindelhauer, D., McGill, N.I., Ross, A., Ebersole, T.A., and Cooke, H.J. (2001). Stable gene expression from a mammalian artificial chromosome. EMBO Rep *2*, 910-914.

Grskovic, M., Javaherian, A., Strulovici, B., and Daley, G.Q. (2011). Induced pluripotent stem cells--opportunities for disease modelling and drug discovery. Nat Rev Drug Discov *10*, 915-929.

Guan, K., Nayernia, K., Maier, L.S., Wagner, S., Dressel, R., Lee, J.H., Nolte, J., Wolf, F., Li, M., Engel, W.*, et al.* (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. Nature *440*, 1199-1203.

Guenther, M.G., Frampton, G.M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., and Young, R.A. (2010). Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. Cell Stem Cell *7*, 249-257.

Guiducci, C., Ascenzioni, F., Auriche, C., Piccolella, E., Guerrini, A.M., and Donini, P. (1999). Use of a human minichromosome as a cloning and expression vector for mammalian cells. Hum Mol Genet *8*, 1417-1424.

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182*, 64-65.

Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., Zweigerdt, R., Gruh, I., Meyer, J., Wagner, S.*, et al.* (2009). Generation of induced pluripotent stem cells from human cord blood. Cell Stem Cell *5*, 434-441.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K.*, et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest *118*, 3132-3142.

Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M.*, et al.* (2003a). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med *348*, 255-256.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E.*, et al.* (2003b). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415-419.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A *107*, 9222-9227.

Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creyghton, M.P., Steine, E.J., Cassady, J.P., Foreman, R.*, et al.* (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell *133*, 250-264.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M.*, et al.* (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science *318*, 1920-1923.

Hanson, C., and Caisander, G. (2005). Human embryonic stem cells and chromosome stability. APMIS *113*, 751-755.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet *15*, 345-355.

Hay, D.C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. Stem Cells *22*, 225-235.

Heinzel, S.S., Krysan, P.J., Tran, C.T., and Calos, M.P. (1991). Autonomous DNA replication in human cells is affected by the size and the source of the DNA. Mol Cell Biol *11*, 2263-2272.

Heller, R., Brown, K.E., Burgtorf, C., and Brown, W.R. (1996). Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. Proc Natl Acad Sci U S A *93*, 7125-7130.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098-1102.

Henning, K.A., Novotny, E.A., Compton, S.T., Guan, X.Y., Liu, P.P., and Ashlock, M.A. (1999). Human artificial chromosomes generated by modification of a yeast artificial chromosome containing both human alpha satellite and single-copy DNA sequences. Proc Natl Acad Sci U S A *96*, 592-597.

Hewitt, Z.A., Amps, K.J., and Moore, H.D. (2007). Derivation of GMP raw materials for use in regenerative medicine: hESC-based therapies, progress toward clinical application. Clin Pharmacol Ther *82*, 448-452.

Hibbitt, O.C., and Wade-Martins, R. (2006). Delivery of large genomic DNA inserts >100 kb using HSV-1 amplicons. Curr Gene Ther *6*, 325-336.

Hobbs, W.E., Brough, D.E., Kovesdi, I., and DeLuca, N.A. (2001). Efficient activation of viral genomes by levels of herpes simplex virus ICP0 insufficient to affect cellular gene expression or cell survival. J Virol *75*, 3391-3403.

Hochedlinger, K., and Jaenisch, R. (2006). Nuclear reprogramming and pluripotency. Nature *441*, 1061-1067.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell *3*, 346-353.

Hoehn, B., Yenari, M.A., Sapolsky, R.M., and Steinberg, G.K. (2003). Glutathione peroxidase overexpression inhibits cytochrome C release and proapoptotic mediators to protect neurons from experimental stroke. Stroke *34*, 2489-2494.

Hohenstein, K.A., Pyle, A.D., Chern, J.Y., Lock, L.F., and Donovan, P.J. (2008). Nucleofection mediates high-efficiency stable gene knockdown and transgene expression in human embryonic stem cells. Stem Cells *26*, 1436-1443.

Hollon, T. (2000). Researchers and regulators reflect on first gene therapy death. Nat Med *6*, 6.

Hoshiya, H., Kazuki, Y., Abe, S., Takiguchi, M., Kajitani, N., Watanabe, Y., Yoshino, T., Shirayoshi, Y., Higaki, K., Messina, G.*, et al.* (2009). A highly stable and nonintegrated human artificial chromosome (HAC) containing the 2.4 Mb entire human dystrophin gene. Mol Ther *17*, 309-317.

Houbaviy, H.B., Dennis, L., Jaenisch, R., and Sharp, P.A. (2005). Characterization of a highly variable eutherian microRNA gene. RNA *11*, 1245-1257.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. Dev Cell *5*, 351-358.

Hsu, C.C., Li, H.P., Hung, Y.H., Leu, Y.W., Wu, W.H., Wang, F.S., Lee, K.D., Chang, P.J., Wu, C.S., Lu, Y.J.*, et al.* (2010). Targeted methylation of CMV and E1A viral promoters. Biochem Biophys Res Commun *402*, 228-234.

Hu, B.Y., Weick, J.P., Yu, J., Ma, L.X., Zhang, X.Q., Thomson, J.A., and Zhang, S.C. (2010a). Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. Proc Natl Acad Sci U S A *107*, 4335-4340.

Hu, K., Yu, J., Suknuntha, K., Tian, S., Montgomery, K., Choi, K.D., Stewart, R., Thomson, J.A., and Slukvin, II (2011). Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. Blood *117*, e109-119.

Hu, Q., Friedrich, A.M., Johnson, L.V., and Clegg, D.O. (2010b). Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. Stem Cells *28*, 1981-1991.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol *26*, 1269-1275.

Hudson, D.F., Fowler, K.J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N.G., de Kretser, D.M., Cancilla, M.R.*, et al.* (1998). Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. J Cell Biol *141*, 309-319.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Narva, E., Ng, S., Sourour, M., Hamalainen, R., Olsson, C.*, et al.* (2011). Copy number variation and selection during reprogramming to pluripotency. Nature *471*, 58-62.

Hyslop, L., Stojkovic, M., Armstrong, L., Walter, T., Stojkovic, P., Przyborski, S., Herbert, M., Murdoch, A., Strachan, T., and Lako, M. (2005). Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages. Stem Cells *23*, 1035-1043.

Iacovitti, L., Donaldson, A.E., Marshall, C.E., Suon, S., and Yang, M. (2007). A protocol for the differentiation of human embryonic stem cells into dopaminergic neurons using only chemically defined human additives: Studies in vitro and in vivo. Brain Res *1127*, 19-25.

Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375-386.

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N.I., Cooke, H., and Masumoto, H. (1998). Construction of YAC-based mammalian artificial chromosomes. Nat Biotechnol *16*, 431-439.

Ikeno, M., Inagaki, H., Nagata, K., Morita, M., Ichinose, H., and Okazaki, T. (2002). Generation of human artificial chromosomes expressing naturally controlled guanosine triphosphate cyclohydrolase I gene. Genes Cells *7*, 1021-1032.

Ioannou, D., Tempest, H.G., Skinner, B.M., Thornhill, A.R., Ellis, M., and Griffin, D.K. (2009). Quantum dots as new-generation fluorochromes for FISH: an appraisal. Chromosome Res *17*, 519-530.

Irion, S., Luche, H., Gadue, P., Fehling, H.J., Kennedy, M., and Keller, G. (2007). Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat Biotechnol *25*, 1477-1482.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., and Benvenisty, N. (2000). Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol Med *6*, 88-95.

Itsykson, P., Ilouz, N., Turetsky, T., Goldstein, R.S., Pera, M.F., Fishbein, I., Segal, M., and Reubinoff, B.E. (2005). Derivation of neural precursors from human embryonic stem cells in the presence of noggin. Mol Cell Neurosci *30*, 24-36.

Itzhaki, J.E., Barnett, M.A., MacCarthy, A.B., Buckle, V.J., Brown, W.R., and Porter, A.C. (1992). Targeted breakage of a human chromosome mediated by cloned human telomeric DNA. Nat Genet *2*, 283-287.

Jablonski, S.A., Chan, G.K., Cooke, C.A., Earnshaw, W.C., and Yen, T.J. (1998). The hBUB1 and hBUBR1 kinases sequentially assemble onto kinetochores during prophase with hBUBR1 concentrating at the kinetochore plates in mitosis. Chromosoma *107*, 386-396.

Jacobs, A., Breakefield, X.O., and Fraefel, C. (1999a). HSV-1-based vectors for gene therapy of neurological diseases and brain tumors: part I. HSV-1 structure, replication and pathogenesis. Neoplasia *1*, 387-401.

Jacobs, A., Breakefield, X.O., and Fraefel, C. (1999b). HSV-1-based vectors for gene therapy of neurological diseases and brain tumors: part II. Vector systems and applications. Neoplasia *1*, 402-416.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567-582.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273-1282.

Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., Panetta, N.J., Chen, Z.Y., Robbins, R.C., Kay, M.A.*, et al.* (2010). A nonviral minicircle vector for deriving human iPS cells. Nat Methods *7*, 197-199.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Jinnah, H.A. (2009). Lesch-Nyhan disease: from mechanism to model and back again. Dis Model Mech *2*, 116-121.

Johnson, P.A., Miyanohara, A., Levine, F., Cahill, T., and Friedmann, T. (1992). Cytotoxicity of a replication-defective mutant of herpes simplex virus type 1. J Virol *66*, 2952-2965.

Johnston, K.M., Jacoby, D., Pechan, P.A., Fraefel, C., Borghesani, P., Schuback, D., Dunn, R.J., Smith, F.I., and Breakefield, X.O. (1997). HSV/AAV hybrid amplicon vectors extend transgene expression in human glioma cells. Hum Gene Ther *8*, 359-370.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature *458*, 771-775.

Kalitsis, P., Fowler, K.J., Earle, E., Hill, J., and Choo, K.H. (1998). Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. Proc Natl Acad Sci U S A *95*, 1136-1141.

Kameda, T., Smuga-Otto, K., and Thomson, J.A. (2006). A severe de novo methylation of episomal vectors by human ES cells. Biochem Biophys Res Commun *349*, 1269-1277.

Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S.*, et al.* (2004). Generation of pluripotent stem cells from neonatal mouse testis. Cell *119*, 1001-1012.

Kashima, T., Vinters, H.V., and Campagnoni, A.T. (1995). Unexpected expression of intermediate filament protein genes in human oligodendroglioma cell lines. J Neuropathol Exp Neurol *54*, 23-31.

Kato, Y., Tani, T., Sotomaru, Y., Kurokawa, K., Kato, J., Doguchi, H., Yasue, H., and Tsunoda, Y. (1998). Eight calves cloned from somatic cells of a single adult. Science *282*, 2095-2098.

Katoh, M., Ayabe, F., Norikane, S., Okada, T., Masumoto, H., Horike, S., Shirayoshi, Y., and Oshimura, M. (2004). Construction of a novel human artificial chromosome vector for gene delivery. Biochem Biophys Res Commun *321*, 280-290.

Kay, M.A. (2011). State-of-the-art gene-based therapies: the road ahead. Nat Rev Genet *12*, 316-328.

Kazuki, Y., Hiratsuka, M., Takiguchi, M., Osaki, M., Kajitani, N., Hoshiya, H., Hiramatsu, K., Yoshino, T., Kazuki, K., Ishihara, C.*, et al.* (2010). Complete genetic correction of ips cells from Duchenne muscular dystrophy. Mol Ther *18*, 386-393.

Kazuki, Y., Hoshiya, H., Kai, Y., Abe, S., Takiguchi, M., Osaki, M., Kawazoe, S., Katoh, M., Kanatsu-Shinohara, M., Inoue, K.*, et al.* (2008). Correction of a genetic defect in multipotent germline stem cells using a human artificial chromosome. Gene Ther *15*, 617-624.

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.*, et al.* (2009a). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell *4*, 472-476.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008a). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. (2010a). A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell *143*, 313-324.

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

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M.*, et al.* (2008b). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature *454*, 646-650.

Kim, J.H., Auerbach, J.M., Rodriguez-Gomez, J.A., Velasco, I., Gavin, D., Lumelsky, N., Lee, S.H., Nguyen, J., Sanchez-Pernaute, R., Bankiewicz, K.*, et al.* (2002). Dopamine neurons derived from embryonic stem cells function in an animal model of Parkinson's disease. Nature *418*, 50-56.

Kim, J.H., Ebersole, T., Kouprina, N., Noskov, V.N., Ohzeki, J., Masumoto, H., Mravinac, B., Sullivan, B.A., Pavlicek, A., Dovat, S.*, et al.* (2009c). Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. Genome Res *19*, 533-544.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.*, et al.* (2010b). Epigenetic memory in induced pluripotent stem cells. Nature *467*, 285-290.

Kim, U.J., Birren, B.W., Slepak, T., Mancino, V., Boysen, C., Kang, H.L., Simon, M.I., and Shizuya, H. (1996). Construction and characterization of a human bacterial artificial chromosome library. Genomics *34*, 213-218.

Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2004). Histone code modifications on pluripotential nuclei of reprogrammed somatic cells. Mol Cell Biol *24*, 5710-5720.

Kiskinis, E., and Eggan, K. (2010). Progress toward the clinical application of patient-specific pluripotent stem cells. J Clin Invest *120*, 51-59.

Klimanskaya, I., Chung, Y., Becker, S., Lu, S.J., and Lanza, R. (2006). Human embryonic stem cell lines derived from single blastomeres. Nature *444*, 481-485.

Koch, P., Siemen, H., Biegler, A., Itskovitz-Eldor, J., and Brustle, O. (2006). Transduction of human embryonic stem cells by ecotropic retroviral vectors. Nucleic Acids Res *34*, e120.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96-105.

Korenberg, J.R., and Engels, W.R. (1978). Base ratio, DNA content, and quinacrine-brightness of human chromosomes. Proc Natl Acad Sci U S A *75*, 3382-3386.

Kraemer, K.H., Levy, D.D., Parris, C.N., Gozukara, E.M., Moriwaki, S., Adelberg, S., and Seidman, M.M. (1994). Xeroderma pigmentosum and related disorders: examining the linkage between defective DNA repair and cancer. J Invest Dermatol *103*, 96S-101S.

Krysan, P.J., Smith, J.G., and Calos, M.P. (1993). Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences. Mol Cell Biol *13*, 2688-2696.

Kubo, S., Saeki, Y., Chiocca, E.A., and Mitani, K. (2003). An HSV amplicon-based helper system for helper-dependent adenoviral vectors. Biochem Biophys Res Commun *307*, 826-830.

Kwon, M.S., Hori, T., Okada, M., and Fukagawa, T. (2007). CENP-C is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. Mol Biol Cell *18*, 2155-2168.

Lamba, D.A., Gust, J., and Reh, T.A. (2009). Transplantation of human embryonic stem cell-derived photoreceptors restores some visual function in Crx-deficient mice. Cell Stem Cell *4*, 73-79.

Laner, A., Goussard, S., Ramalho, A.S., Schwarz, T., Amaral, M.D., Courvalin, P., Schindelhauer, D., and Grillot-Courvalin, C. (2005). Bacterial transfer of large functional genomic DNA into human cells. Gene Ther *12*, 1559-1572.

Lawrence, M.S., Ho, D.Y., Dash, R., and Sapolsky, R.M. (1995). Herpes simplex virus vectors overexpressing the glucose transporter gene protect against seizure-induced neuron loss. Proc Natl Acad Sci U S A *92*, 7247-7251.

Lebkowski, J.S., Gold, J., Xu, C., Funk, W., Chiu, C.P., and Carpenter, M.K. (2001). Human embryonic stem cells: culture, differentiation, and genetic modification for regenerative medicine applications. Cancer J *7 Suppl 2*, S83-93.

Lee, C.Y., Rennie, P.S., and Jia, W.W. (2009a). MicroRNA regulation of oncolytic herpes simplex virus-1 for selective killing of prostate cancer cells. Clin Cancer Res *15*, 5126-5135.

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.*, et al.* (2009b). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature *461*, 402-406.

Lee, J.H., Park, I.H., Gao, Y., Li, J.B., Li, Z., Daley, G.Q., Zhang, K., and Church, G.M. (2009c). A robust approach to identifying tissue-specific gene expression regulatory variants using personalized human induced pluripotent stem cells. PLoS Genet *5*, e1000718.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of

developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lehman, I.R., and Boehmer, P.E. (1999). Replication of herpes simplex virus DNA. J Biol Chem *274*, 28059-28062.

Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere end-replication problem and cell aging. J Mol Biol *225*, 951-960.

Li, J.Y., Christophersen, N.S., Hall, V., Soulet, D., and Brundin, P. (2008). Critical issues of clinical human embryonic stem cell therapy for brain repair. Trends Neurosci *31*, 146-153.

Li, L., Wang, B.H., Wang, S., Moalim-Nour, L., Mohib, K., Lohnes, D., and Wang, L. (2010a). Individual cell movement, asymmetric colony expansion, rho-associated kinase, and E-cadherin impact the clonogenicity of human embryonic stem cells. Biophys J *98*, 2442-2451.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q.*, et al.* (2010b). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell *7*, 51-63.

Li, W., and Ding, S. (2010). Small molecules that modulate embryonic stem cell fate and somatic cell reprogramming. Trends Pharmacol Sci *31*, 36-45.

Liew, C.G., Draper, J.S., Walsh, J., Moore, H., and Andrews, P.W. (2007). Transient and stable transgene expression in human embryonic stem cells. Stem Cells *25*, 1521-1528.

Lim, F., Palomo, G.M., Mauritz, C., Gimenez-Cassina, A., Illana, B., Wandosell, F., and Diaz-Nido, J. (2007). Functional recovery in a Friedreich's ataxia mouse model by frataxin gene transfer using an HSV-1 amplicon vector. Mol Ther *15*, 1072-1078.

Lim, J.W., and Bodnar, A. (2002). Proteome analysis of conditioned medium from mouse embryonic fibroblast feeder layers which support the growth of human embryonic stem cells. Proteomics *2*, 1187-1203.

Lin, T., Ambasudhan, R., Yuan, X., Li, W., Hilcove, S., Abujarour, R., Lin, X., Hahm, H.S., Hao, E., Hayek, A.*, et al.* (2009). A chemical platform for improved induction of human iPSCs. Nat Methods *6*, 805-808.

Liu, H., Ye, Z., Kim, Y., Sharkis, S., and Jang, Y.Y. (2010). Generation of endoderm-derived human induced pluripotent stem cells from primary hepatocytes. Hepatology *51*, 1810-1819.

Livesey, F.J., and Cepko, C.L. (2001). Vertebrate neural cell-fate determination: lessons from the retina. Nat Rev Neurosci *2*, 109-118.

Loh, Y.H., Agarwal, S., Park, I.H., Urbach, A., Huo, H., Heffner, G.C., Kim, K., Miller, J.D., Ng, K., and Daley, G.Q. (2009). Generation of induced pluripotent stem cells from human blood. Blood *113*, 5476-5479.

Loh, Y.H., Hartung, O., Li, H., Guo, C., Sahalie, J.M., Manos, P.D., Urbach, A., Heffner, G.C., Grskovic, M., Vigneault, F.*, et al.* (2010). Reprogramming of T cells from human peripheral blood. Cell Stem Cell *7*, 15-19.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S.*, et al.* (2006). Derivation of human embryonic stem cells in defined conditions. Nat Biotechnol *24*, 185-187.

Ma, C., Martin, S., Trask, B., and Hamlin, J.L. (1993). Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. Genes Dev *7*, 605-620.

Magnaldo, T. (2004). Xeroderma pigmentosum: from genetics to hopes and realities of cutaneous gene therapy. Expert Opin Biol Ther *4*, 169-179.

Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C., and Hochedlinger, K. (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. Cell Stem Cell *3*, 340-345.

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. Cell Stem Cell *3*, 595-605.

Mali, P., Ye, Z., Hommond, H.H., Yu, X., Lin, J., Chen, G., Zou, J., and Cheng, L. (2008). Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts. Stem Cells *26*, 1998-2005.

Mandegar, M.A. (2011). Analysis of Artificial Chromosomes in Human Embryonic Stem Cells. In DPhil Thesis (University of Oxford).

Mandegar, M.A., Moralli, D., Khoja, S., Cowley, S., Chan, D.Y., Yusuf, M., Mukherjee, S., Blundell, M.P., Volpi, E.V., Thrasher, A.J.*, et al.* (2011). Functional human artificial chromosomes are generated and stably maintained in human embryonic stem cells. Hum Mol Genet *20*, 2905-2913.

Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149-1153.

Markoulaki, S., Hanna, J., Beard, C., Carey, B.W., Cheng, A.W., Lengner, C.J., Dausman, J.A., Fu, D., Gao, Q., Wu, S.*, et al.* (2009). Transgenic mice with defined combinations of drug-inducible reprogramming factors. Nat Biotechnol *27*, 169-171.

Marshall, E. (1999). Gene therapy death prompts review of adenovirus vector. Science *286*, 2244-2245.

Marshall, O.J., Marshall, A.T., and Choo, K.H. (2008). Three-dimensional localization of CENP-A suggests a complex higher order structure of centromeric chromatin. J Cell Biol *183*, 1193-1202.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Martin, M.J., Muotri, A., Gage, F., and Varki, A. (2005). Human embryonic stem cells express an immunogenic nonhuman sialic acid. Nat Med *11*, 228-232.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A.*, et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Masumoto, H., Ikeno, M., Nakano, M., Okazaki, T., Grimes, B., Cooke, H., and Suzuki, N. (1998). Assay of centromere function using a human artificial chromosome. Chromosoma *107*, 406-416.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J Cell Biol *109*, 1963-1973.

Matin, M.M., Walsh, J.R., Gokhale, P.J., Draper, J.S., Bahrami, A.R., Morton, I., Moore, H.D., and Andrews, P.W. (2004). Specific knockdown of Oct4 and beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells. Stem Cells *22*, 659-668.

Matrai, J., Chuah, M.K., and VandenDriessche, T. (2010). Recent advances in lentiviral vector development and applications. Mol Ther *18*, 477-490.

Matsui, H., Hegadorn, C., Ozelo, M., Burnett, E., Tuttle, A., Labelle, A., McCray, P.B., Jr., Naldini, L., Brown, B., Hough, C.*, et al.* (2011). A microRNA-regulated and GP64-pseudotyped lentiviral vector mediates stable expression of FVIII in a murine model of Hemophilia A. Mol Ther *19*, 723-730.

Matsumura, H., Tada, M., Otsuji, T., Yasuchika, K., Nakatsuji, N., Surani, A., and Tada, T. (2007). Targeted chromosome elimination from ES-somatic hybrid cells. Nat Methods *4*, 23-25.

McDonald, J.A., and Kelley, W.N. (1971). Lesch-Nyhan syndrome: altered kinetic properties of mutant enzyme. Science *171*, 689-691.

McMahon, A.P., and Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell *62*, 1073-1085.

Mejia, J.E., Alazami, A., Willmott, A., Marschall, P., Levy, E., Earnshaw, W.C., and Larin, Z. (2002). Efficiency of de novo centromere formation in human artificial chromosomes. Genomics *79*, 297-304.

Mejia, J.E., and Larin, Z. (2000). The assembly of large BACs by in vivo recombination. Genomics *70*, 165-170.

Mejia, J.E., Willmott, A., Levy, E., Earnshaw, W.C., and Larin, Z. (2001). Functional complementation of a genetic deficiency with human artificial chromosomes. Am J Hum Genet *69*, 315-326.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of Dosage Compensation Impacts Human iPSC Disease Modeling. Cell Stem Cell *10*, 595-609.

Melton, C., Judson, R.L., and Blelloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature *463*, 621-626.

Miller, R.A., and Ruddle, F.H. (1976). Pluripotent teratocarcinoma-thymus somatic cell hybrids. Cell *9*, 45-55.

Mills, W., Critcher, R., Lee, C., and Farr, C.J. (1999). Generation of an approximately 2.4 Mb human X centromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. Hum Mol Genet *8*, 751-761.

Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M.*, et al.* (2006). The expression profile of microRNAs in mouse embryos. Nucleic Acids Res *34*, 1765-1771.

Mitalipova, M.M., Rao, R.R., Hoyer, D.M., Johnson, J.A., Meisner, L.F., Jones, K.L., Dalton, S., and Stice, S.L. (2005). Preserving the genetic integrity of human embryonic stem cells. Nat Biotechnol *23*, 19-20.

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M.*, et al.* (2009). Variation in the safety of induced pluripotent stem cell lines. Nat Biotechnol *27*, 743-745.

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., Nishikawa, S., Tanemura, M., Mimori, K., Tanaka, F.*, et al.* (2011). Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell Stem Cell *8*, 633-638.

Mizuno, H., Yanoma, S., Nishimura, G., Hattori, S., Ito, T., Okudera, K., and Tsukuda, M. (2000). Therapeutic efficiency of IL-2 gene transduced tumor vaccine for head and neck carcinoma. Cancer Lett *152*, 175-185.

Moon, S.Y., Park, Y.B., Kim, D.S., Oh, S.K., and Kim, D.W. (2006). Generation, culture, and differentiation of human embryonic stem cells for therapeutic applications. Mol Ther *13*, 5-14.

Moore, J.C., Atze, K., Yeung, P.L., Toro-Ramos, A.J., Camarillo, C., Thompson, K., Ricupero, C.L., Brenneman, M.A., Cohen, R.I., and Hart, R.P. (2010). Efficient, high-throughput transfection of human embryonic stem cells. Stem Cell Res Ther *1*, 23.

Moralli, D., Simpson, K.M., Wade-Martins, R., and Monaco, Z.L. (2006). A novel human artificial chromosome gene expression system using herpes simplex virus type 1 vectors. EMBO Rep *7*, 911-918.

Moralli, D., Yusuf, M., Mandegar, M.A., Khoja, S., Monaco, Z.L., and Volpi, E.V. (2010). An improved technique for chromosomal analysis of human ES and iPS cells. Stem Cell Rev *7*, 471-477.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc Natl Acad Sci U S A *85*, 6622-6626.

Murphy, S.L., and High, K.A. (2008). Gene therapy for haemophilia. Br J Haematol *140*, 479-487.

Murry, C.E., and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell *132*, 661-680.

Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, B. (2003). Manipulating the mouse embryo: a laboratory manual, 3rd edn (New York, Cold Spring Harbor Laboratory Press).

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.

Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T., and Yamanaka, S. (2010). Promotion of direct reprogramming by transformation-deficient Myc. Proc Natl Acad Sci U S A *107*, 14152-14157.

Nakashima, H., Nakano, M., Ohnishi, R., Hiraoka, Y., Kaneda, Y., Sugino, A., and Masumoto, H. (2005). Assembly of additional heterochromatin distinct from centromere-kinetochore chromatin is required for de novo formation of human artificial chromosome. J Cell Sci *118*, 5885-5898.

Nanbo, A., Sugden, A., and Sugden, B. (2007). The coupling of synthesis and partitioning of EBV's plasmid replicon is revealed in live cells. EMBO J *26*, 4252-4262.

Neidle, S., and Parkinson, G.N. (2003). The structure of telomeric DNA. Curr Opin Struct Biol *13*, 275-283.

Newcomb, W.W., and Brown, J.C. (1994). Induced extrusion of DNA from the capsid of herpes simplex virus type 1. J Virol *68*, 433-440.

Newman, A.M., and Cooper, J.B. (2010). Lab-specific gene expression signatures in pluripotent stem cells. Cell Stem Cell *7*, 258-262.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nichols, W.W., Murphy, D.G., Cristofalo, V.J., Toji, L.H., Greene, A.E., and Dwight, S.A. (1977). Characterization of a new human diploid cell strain, IMR-90. Science *196*, 60-63.

Nicklas, R.B. (1997). How cells get the right chromosomes. Science *275*, 632-637.

Noggle, S., Fung, H.L., Gore, A., Martinez, H., Satriani, K.C., Prosser, R., Oum, K., Paull, D., Druckenmiller, S., Freeby, M.*, et al.* (2011). Human oocytes reprogram somatic cells to a pluripotent state. Nature *478*, 70-75.

Norio, P. (2006). DNA replication: the unbearable lightness of origins. EMBO Rep *7*, 779-781.

Oegema, K., Desai, A., Rybina, S., Kirkham, M., and Hyman, A.A. (2001). Functional analysis of kinetochore assembly in Caenorhabditis elegans. J Cell Biol *153*, 1209-1226.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun *343*, 159-166.

Oh, S.K., Chen, A.K., Mok, Y., Chen, X., Lim, U.M., Chin, A., Choo, A.B., and Reuveny, S. (2009). Long-term microcarrier suspension cultures of human embryonic stem cells. Stem Cell Res *2*, 219-230.

Ohnuki, M., Takahashi, K., and Yamanaka, S. (2009). Generation and characterization of human induced pluripotent stem cells. Curr Protoc Stem Cell Biol *Chapter 4*, Unit 4A 2.

Ohzeki, J., Nakano, M., Okada, T., and Masumoto, H. (2002). CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J Cell Biol *159*, 765-775.

Okamoto, Y., Nakano, M., Ohzeki, J., Larionov, V., and Masumoto, H. (2007). A minimal CENP-A core is required for nucleation and maintenance of a functional human centromere. EMBO J *26*, 1279-1291.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science *322*, 949-953.

Okita, K., and Yamanaka, S. (2006). Intracellular signaling pathways regulating pluripotency of embryonic stem cells. Curr Stem Cell Res Ther *1*, 103-111.

Onishi, A., Iwamoto, M., Akita, T., Mikawa, S., Takeda, K., Awata, T., Hanada, H., and Perry, A.C. (2000). Pig cloning by microinjection of fetal fibroblast nuclei. Science *289*, 1188-1190.

Orkin, S.H., and Hochedlinger, K. (2011). Chromatin connections to pluripotency and cellular reprogramming. Cell *145*, 835-850.

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among human embryonic stem cell lines. Nat Biotechnol *26*, 313-315.

Osterhage, J.L., and Friedman, K.L. (2009). Chromosome end maintenance by telomerase. J Biol Chem *284*, 16061-16065.

Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kuhlcke, K., Schilz, A., Kunkel, H.*, et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med *12*, 401-409.

Palmer, D.J., and Ng, P. (2005). Helper-dependent adenoviral vectors for gene therapy. Hum Gene Ther *16*, 1-16.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc Natl Acad Sci U S A *88*, 3734-3738.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C.*, et al.* (2011). Induction of human neuronal cells by defined transcription factors. Nature *476*, 220-223.

Papadakis, E.D., Nicklin, S.A., Baker, A.H., and White, S.J. (2004). Promoters and control elements: designing expression cassettes for gene therapy. Curr Gene Ther *4*, 89-113.

Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008a). Disease-specific induced pluripotent stem cells. Cell *134*, 877-886.

Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008b). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141-146.

Parks, R.J. (2000). Improvements in adenoviral vector technology: overcoming barriers for gene therapy. Clin Genet *58*, 1-11.

Pasero, P., and Schwob, E. (2000). Think global, act local--how to regulate S phase from individual replication origins. Curr Opin Genet Dev *10*, 178-186.

Pauli, A., Rinn, J.L., and Schier, A.F. (2011). Non-coding RNAs as regulators of embryogenesis. Nat Rev Genet *12*, 136-149.

Pellett, P.E., and Roizman, B. (2007). The Family Herpesviridae: A Brief Introduction. In Fields Virology, D.M. Knipe, and P.M. Howley, eds. (Philidelphia, PA, Lippincott Williams & Wilkins).

Pera, M.F., and Tam, P.P. (2010). Extrinsic regulation of pluripotent stem cells. Nature *465*, 713-720.

Pfeifer, A., Ikawa, M., Dayn, Y., and Verma, I.M. (2002). Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. Proc Natl Acad Sci U S A *99*, 2140-2145.

Plath, K., and Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. Nat Rev Genet *12*, 253-265.

Polejaeva, I.A., Chen, S.H., Vaught, T.D., Page, R.L., Mullins, J., Ball, S., Dai, Y., Boone, J., Walker, S., Ayares, D.L.*, et al.* (2000). Cloned pigs produced by nuclear transfer from adult somatic cells. Nature *407*, 86-90.

Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T.*, et al.* (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol *28*, 848-855.

Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.D., and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature.

Rasheed, S., Nelson-Rees, W.A., Toth, E.M., Arnstein, P., and Gardner, M.B. (1974). Characterization of a newly derived human sarcoma cell line (HT-1080). Cancer *33*, 1027-1033.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol *18*, 399-404.

Ribeiro, S.A., Vagnarelli, P., Dong, Y., Hori, T., McEwen, B.F., Fukagawa, T., Flors, C., and Earnshaw, W.C. (2010). A super-resolution map of the vertebrate kinetochore. Proc Natl Acad Sci U S A *107*, 10484-10489.

Rideout, W.M., 3rd, Hochedlinger, K., Kyba, M., Daley, G.Q., and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. Cell *109*, 17-27.

Rieder, C.L. (1982). The formation, structure, and composition of the mammalian kinetochore and kinetochore fiber. Int Rev Cytol *79*, 1-58.

Rieder, C.L., and Salmon, E.D. (1998). The vertebrate cell kinetochore and its roles during mitosis. Trends Cell Biol *8*, 310-318.

Robinton, D.A., and 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., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem *280*, 24731-24737.

Roizman, B., Knipe, D.M., and Whitley, R.J. (2007). Herpes Simplex Viruses. In Fields Virology, D.M. Knipe, and P.M. Howley, eds. (Philidelphia, PA, Lippincott Williams & Wilkins).

Rosler, E.S., Fisk, G.J., Ares, X., Irving, J., Miura, T., Rao, M.S., and Carpenter, M.K. (2004). Long-term culture of human embryonic stem cells in feeder-free conditions. Dev Dyn *229*, 259-274.

Ross, P.J., Kennedy, M.A., Christou, C., Risco Quiroz, M., Poulin, K.L., and Parks, R.J. (2011). Assembly of helper-dependent adenovirus DNA into chromatin promotes efficient gene expression. J Virol *85*, 3950-3958.

Rossant, J. (2008). Stem cells and early lineage development. Cell *132*, 527-531.

Rudd, M.K., Mays, R.W., Schwartz, S., and Willard, H.F. (2003). Human artificial chromosomes with alpha satellite-based de novo centromeres show increased frequency of nondisjunction and anaphase lag. Mol Cell Biol *23*, 7689-7697.

Rudd, M.K., and Willard, H.F. (2004). Analysis of the centromeric regions of the human genome assembly. Trends Genet *20*, 529-533.

Rudd, M.K., Wray, G.A., and Willard, H.F. (2006). The evolutionary dynamics of alpha-satellite. Genome Res *16*, 88-96.

Ruiz, S., Brennand, K., Panopoulos, A.D., Herrerias, A., Gage, F.H., and Izpisua-Belmonte, J.C. (2010). High-efficient generation of induced pluripotent stem cells from human astrocytes. PLoS One *5*, e15526.

Saeki, Y., Breakefield, X.O., and Chiocca, E.A. (2003). Improved HSV-1 amplicon packaging system using ICP27-deleted, oversized HSV-1 BAC DNA. Methods Mol Med *76*, 51-60.

Saeki, Y., Fraefel, C., Ichikawa, T., Breakefield, X.O., and Chiocca, E.A. (2001). Improved helper virus-free packaging system for HSV amplicon vectors using an ICP27-deleted, oversized HSV-1 DNA in a bacterial artificial chromosome. Mol Ther *3*, 591-601.

Saeki, Y., Ichikawa, T., Saeki, A., Chiocca, E.A., Tobler, K., Ackermann, M., Breakefield, X.O., and Fraefel, C. (1998). Herpes simplex virus type 1 DNA amplified as bacterial artificial chromosome in Escherichia coli: rescue of replication-competent virus progeny and packaging of amplicon vectors. Hum Gene Ther *9*, 2787-2794.

Saffery, R., and Choo, K.H. (2002). Strategies for engineering human chromosomes with therapeutic potential. J Gene Med *4*, 5-13.

Saffery, R., Irvine, D.V., Griffiths, B., Kalitsis, P., Wordeman, L., and Choo, K.H. (2000). Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. Hum Mol Genet *9*, 175-185.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64-77.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Scanlan, P.M., Tiwari, V., Bommireddy, S., and Shukla, D. (2005). Spinoculation of heparan sulfate deficient cells enhances HSV-1 entry, but does not abolish the need for essential glycoproteins in viral fusion. J Virol Methods *128*, 104-112.

Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. (2001). Genomic and genetic definition of a functional human centromere. Science *294*, 109-115.

Schueler, M.G., and Sullivan, B.A. (2006). Structural and functional dynamics of human centromeric chromatin. Annu Rev Genomics Hum Genet *7*, 301-313.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745.

Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D.A., and Benvenisty, N. (2000). Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A *97*, 11307-11312.

Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M.*, et al.* (2010). Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell Stem Cell *7*, 11-14.

Sena-Esteves, M., Saeki, Y., Fraefel, C., and Breakefield, X.O. (2000). HSV-1 amplicon vectors--simplicity and versatility. Mol Ther *2*, 9-15.

Shah, J.V., and Cleveland, D.W. (2000). Waiting for anaphase: Mad2 and the spindle assembly checkpoint. Cell *103*, 997-1000.

Shah, K., Tang, Y., Breakefield, X., and Weissleder, R. (2003). Real-time imaging of TRAIL-induced apoptosis of glioma tumors in vivo. Oncogene *22*, 6865-6872.

Shamblott, M.J., Axelman, J., Wang, S., Bugg, E.M., Littlefield, J.W., Donovan, P.J., Blumenthal, P.D., Huggins, G.R., and Gearhart, J.D. (1998). Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc Natl Acad Sci U S A *95*, 13726-13731.

Shay, J.W. (1999). At the end of the millennium, a view of the end. Nat Genet *23*, 382-383.

Shay, J.W., and Wright, W.E. (2007). Hallmarks of telomeres in ageing research. J Pathol *211*, 114-123.

Sheridan, C. (2011). Gene therapy finds its niche. Nat Biotechnol *29*, 121-128.

Shimada, H., Hashimoto, Y., Nakada, A., Shigeno, K., and Nakamura, T. (2011). Accelerated generation of human induced pluripotent stem cells with retroviral transduction and chemical inhibitors under physiological hypoxia. Biochem Biophys Res Commun *417*, 659-664.

Si-Tayeb, K., Noto, F.K., Sepac, A., Sedlic, F., Bosnjak, Z.J., Lough, J.W., and Duncan, S.A. (2010). Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. BMC Dev Biol *10*, 81.

Sia, K.C., Chong, W.K., Ho, I.A., Yulyana, Y., Endaya, B., Huynh, H., and Lam, P.Y. (2010). Hybrid herpes simplex virus/Epstein-Barr virus amplicon viral vectors confer enhanced transgene expression in primary human tumors and human bone marrow-derived mesenchymal stem cells. J Gene Med *12*, 848-858.

Siegel, P.M., Shu, W., and Massague, J. (2003). Mad upregulation and Id2 repression accompany transforming growth factor (TGF)-beta-mediated epithelial cell growth suppression. J Biol Chem *278*, 35444-35450.

Siemen, H., Nix, M., Endl, E., Koch, P., Itskovitz-Eldor, J., and Brustle, O. (2005). Nucleofection of human embryonic stem cells. Stem Cells Dev *14*, 378-383.

Singec, I., Jandial, R., Crain, A., Nikkhah, G., and Snyder, E.Y. (2007). The leading edge of stem cell therapeutics. Annu Rev Med *58*, 313-328.

Smith, I.L., Hardwicke, M.A., and Sandri-Goldin, R.M. (1992). Evidence that the herpes simplex virus immediate early protein ICP27 acts post-transcriptionally during infection to regulate gene expression. Virology *186*, 74-86.

Sodeik, B., Ebersold, M.W., and Helenius, A. (1997). Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. J Cell Biol *136*, 1007-1021.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M.*, et al.* (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell *136*, 964-977.

Solter, D., and Knowles, B.B. (1975). Immunosurgery of mouse blastocyst. Proc Natl Acad Sci U S A *72*, 5099-5102.

Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean, J.M., Smith, B.W.*, et al.* (2010). Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells *28*, 1728-1740.

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., and 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.

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., and Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. Nat Biotechnol *26*, 1361-1363.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364-377.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature *465*, 175-181.

Stadtfeld, M., Brennand, K., and Hochedlinger, K. (2008a). Reprogramming of pancreatic beta cells into induced pluripotent stem cells. Curr Biol *18*, 890-894.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev *24*, 2239-2263.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008b). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230-240.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008c). Induced pluripotent stem cells generated without viral integration. Science *322*, 945-949.

Stegmaier, K., Ross, K.N., Colavito, S.A., O'Malley, S., Stockwell, B.R., and Golub, T.R. (2004). Gene expression-based high-throughput screening(GE-HTS) and application to leukemia differentiation. Nat Genet *36*, 257-263.

Stein, S., Ott, M.G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Kramer, A., Schwable, J., Glimm, H.*, et al.* (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat Med *16*, 198-204.

Stern, C.D., and Fraser, S.E. (2001). Tracing the lineage of tracing cell lineages. Nat Cell Biol *3*, E216-218.

Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterisation of a yeast chromosomal replicator. Nature *282*, 39-43.

Strelchenko, N., Kukharenko, V., Shkumatov, A., Verlinsky, O., Kuliev, A., and Verlinsky, Y. (2006). Reprogramming of human somatic cells by embryonic stem cell cytoplast. Reprod Biomed Online *12*, 107-111.

Strelchenko, N., Verlinsky, O., Kukharenko, V., and Verlinsky, Y. (2004). Morula-derived human embryonic stem cells. Reprod Biomed Online *9*, 623-629.

Strom, S., Inzunza, J., Grinnemo, K.H., Holmberg, K., Matilainen, E., Stromberg, A.M., Blennow, E., and Hovatta, O. (2007). Mechanical isolation of the inner cell mass is effective in derivation of new human embryonic stem cell lines. Hum Reprod *22*, 3051-3058.

Strulovici, Y., Leopold, P.L., O'Connor, T.P., Pergolizzi, R.G., and Crystal, R.G. (2007). Human embryonic stem cells and gene therapy. Mol Ther *15*, 850-866.

Sugii, S., Kida, Y., Kawamura, T., Suzuki, J., Vassena, R., Yin, Y.Q., Lutz, M.K., Berggren, W.T., Izpisua Belmonte, J.C., and Evans, R.M. (2010). Human and mouse adipose-derived cells support feeder-independent induction of pluripotent stem cells. Proc Natl Acad Sci U S A *107*, 3558-3563.

Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., Moon, S.Y.*, et al.* (2004). Human embryonic stem cells express a unique set of microRNAs. Dev Biol *270*, 488-498.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol *11*, 1076-1083.

Sullivan, B.A., and Schwartz, S. (1995). Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. Hum Mol Genet *4*, 2189-2197.

Sullivan, B.A., and Willard, H.F. (1998). Stable dicentric X chromosomes with two functional centromeres. Nat Genet *20*, 227-228.

Sullivan, K.F., and Glass, C.A. (1991). CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. Chromosoma *100*, 360-370.

Sun, M., Kong, L., Wang, X., Lu, X.G., Gao, Q., and Geller, A.I. (2005). Comparison of the capability of GDNF, BDNF, or both, to protect nigrostriatal neurons in a rat model of Parkinson's disease. Brain Res *1052*, 119-129.

Sun, N., Panetta, N.J., Gupta, D.M., Wilson, K.D., Lee, A., Jia, F., Hu, S., Cherry, A.M., Robbins, R.C., Longaker, M.T.*, et al.* (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. Proc Natl Acad Sci U S A *106*, 15720-15725.

Sung, R.S., Qin, L., and Bromberg, J.S. (2001). TNFalpha and IFNgamma induced by innate anti-adenoviral immune responses inhibit adenovirus-mediated transgene expression. Mol Ther *3*, 757-767.

Suzuki, K., Mitsui, K., Aizawa, E., Hasegawa, K., Kawase, E., Yamagishi, T., Shimizu, Y., Suemori, H., Nakatsuji, N., and Mitani, K. (2008). Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent adenoviral vectors. Proc Natl Acad Sci U S A *105*, 13781-13786.

Suzuki, M., Chiocca, E.A., and Saeki, Y. (2007). Early STAT1 activation after systemic delivery of HSV amplicon vectors suppresses transcription of the vector-encoded transgene. Mol Ther *15*, 2017-2026.

Suzuki, N., Itou, T., Hasegawa, Y., Okazaki, T., and Ikeno, M. (2009). Cell to cell transfer of the chromatin-packaged human beta-globin gene cluster. Nucleic Acids Res *38*, e33.

Szabo, E., Rampalli, S., Risueno, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. Nature *468*, 521-526.

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., and Montgomery, K.D. (2011). Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. Nat Biotechnol *29*, 313-314.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tan, S.M., and Droge, P. (2005). Comparative analysis of sequence-specific DNA recombination systems in human embryonic stem cells. Stem Cells *23*, 868-873.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. Mol Cell Biol *23*, 2699-2708.

Tomkiel, J., Cooke, C.A., Saitoh, H., Bernat, R.L., and Earnshaw, W.C. (1994). CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. J Cell Biol *125*, 531-545.

Uren, A.G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A.H., Wilson, C.H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A*., et al.* (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat Protoc *4*, 789-798.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009a). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. J Cell Sci *122*, 3502-3510.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009b). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145-1148.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci *118*, 4495-4509.

Vazin, T., and Freed, W.J. (2010). Human embryonic stem cells: derivation, culture, and differentiation: a review. Restor Neurol Neurosci *28*, 589-603.

Vicario, I., and Schimmang, T. (2003). Transfer of FGF-2 via HSV-1-based amplicon vectors promotes efficient formation of neurons from embryonic stem cells. J Neurosci Methods *123*, 55-60.

Voet, T., Vermeesch, J., Carens, A., Durr, J., Labaere, C., Duhamel, H., David, G., and Marynen, P. (2001). Efficient male and female germline transmission of a human chromosomal vector in mice. Genome Res *11*, 124-136.

Vukicevic, S., Kleinman, H.K., Luyten, F.P., Roberts, A.B., Roche, N.S., and Reddi, A.H. (1992). Identification of multiple active growth factors in basement membrane Matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components. Exp Cell Res *202*, 1-8.

Wade-Martins, R., Saeki, Y., and Chiocca, E.A. (2003). Infectious delivery of a 135-kb LDLR genomic locus leads to regulated complementation of low-density lipoprotein receptor deficiency in human cells. Mol Ther *7*, 604-612.

Wade-Martins, R., Smith, E.R., Tyminski, E., Chiocca, E.A., and Saeki, Y. (2001). An infectious transfer and expression system for genomic DNA loci in human and mouse cells. Nat Biotechnol *19*, 1067-1070.

Wakayama, T., Perry, A.C., Zuccotti, M., Johnson, K.R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature *394*, 369-374.

Wang, S., Di, S., Young, W.B., Jacobson, C., and Link, C.J., Jr. (1997). A novel herpesvirus amplicon system for in vivo gene delivery. Gene Ther *4*, 1132-1141.

Wang, S., Qi, J., Smith, M., and Link, C.J. (2002). Antitumor effects on human melanoma xenografts of an amplicon vector transducing the herpes thymidine kinase gene followed by ganciclovir. Cancer Gene Ther *9*, 1-8.

Wang, S., and Vos, J.M. (1996). A hybrid herpesvirus infectious vector based on Epstein-Barr virus and herpes simplex virus type 1 for gene transfer into human cells in vitro and in vivo. J Virol *70*, 8422-8430.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet *40*, 1478-1483.

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F., *et al.* (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol *7*, 901-904.

Warburton, P.E., Haaf, T., Gosden, J., Lawson, D., and Willard, H.F. (1996). Characterization of a chromosome-specific chimpanzee alpha satellite subset: evolutionary relationship to subsets on human chromosomes. Genomics *33*, 220-228.

Warburton, P.E., and Willard, H.F. (1996). Evolution of centromeric alpha satellite DNA: molecular organization within and between human and primate chromosomes. In Human Genome Evolution, M.S. Jackson, T. Strachan, and G. Dover, eds. (Oxford, BIOS Scientific Publishers), pp. 121-145.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., *et al.* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell *7*, 618-630.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Muguruma, K., and Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol *25*, 681-686.

Wells, D.J. (2004). Gene therapy progress and prospects: electroporation and other physical methods. Gene Ther *11*, 1363-1369.

Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008a). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell *2*, 10-12.

Wernig, M., Zhao, J.P., Pruszak, J., Hedlund, E., Fu, D., Soldner, F., Broccoli, V., Constantine-Paton, M., Isacson, O., and Jaenisch, R. (2008b). Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. Proc Natl Acad Sci U S A *105*, 5856-5861.

Whitley, R.J., Kimberlin, D.W., and Roizman, B. (1998). Herpes simplex viruses. Clin Infect Dis *26*, 541-553; quiz 554-545.

Whitley, R.J., and Roizman, B. (2001). Herpes simplex virus infections. Lancet *357*, 1513-1518.

Willard, H.F. (1998). Centromeres: the missing link in the development of human artificial chromosomes. Curr Opin Genet Dev *8*, 219-225.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., *et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature *458*, 766-770.

Wu, S.M., and Hochedlinger, K. (2011). Harnessing the potential of induced pluripotent stem cells for regenerative medicine. Nat Cell Biol *13*, 497-505.

Wysocka, J., and Herr, W. (2003). The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. Trends Biochem Sci *28*, 294-304.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. Nat Biotechnol *19*, 971-974.

Xu, C., Jiang, J., Sottile, V., McWhir, J., Lebkowski, J., and Carpenter, M.K. (2004). Immortalized fibroblast-like cells derived from human embryonic stem cells support undifferentiated cell growth. Stem Cells *22*, 972-980.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell *137*, 647-658.

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol *20*, 1261-1264.

Xu, R.H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R.*, et al.* (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell *3*, 196-206.

Xu, Y., Zhu, X., Hahm, H.S., Wei, W., Hao, E., Hayek, A., and Ding, S. (2010). Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. Proc Natl Acad Sci U S A *107*, 8129-8134.

Yang, D., Zhang, Z.J., Oldenburg, M., Ayala, M., and Zhang, S.C. (2008). Human embryonic stem cell-derived dopaminergic neurons reverse functional deficit in parkinsonian rats. Stem Cells *26*, 55-63.

Yang, X., Smith, S.L., Tian, X.C., Lewin, H.A., Renard, J.P., and Wakayama, T. (2007). Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning. Nat Genet *39*, 295-302.

Yates, J., Warren, N., Reisman, D., and Sugden, B. (1984). A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc Natl Acad Sci U S A *81*, 3806-3810.

Yates, J.L., and Guan, N. (1991). Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. J Virol *65*, 483-488.

Yates, J.L., Warren, N., and Sugden, B. (1985). Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature *313*, 812-815.

Yenari, M.A., Fink, S.L., Sun, G.H., Chang, L.K., Patel, M.K., Kunis, D.M., Onley, D., Ho, D.Y., Sapolsky, R.M., and Steinberg, G.K. (1998). Gene therapy with HSP72 is neuroprotective in rat models of stroke and epilepsy. Ann Neurol *44*, 584-591.

Ying, Q.L., Nichols, J., Evans, E.P., and Smith, A.G. (2002). Changing potency by spontaneous fusion. Nature *416*, 545-548.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., and 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, and Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. Science *324*, 797-801.

Yu, J., Vodyanik, M.A., He, P., Slukvin, II, and Thomson, J.A. (2006). Human embryonic stem cells reprogram myeloid precursors following cell-cell fusion. Stem Cells *24*, 168-176.

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.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Yusa, K., Rad, R., Takeda, J., and Bradley, A. (2009). Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat Methods *6*, 363-369.

Zaehres, H., Lensch, M.W., Daheron, L., Stewart, S.A., Itskovitz-Eldor, J., and Daley, G.Q. (2005). High-efficiency RNA interference in human embryonic stem cells. Stem Cells *23*, 299-305.

Zhang, X., Stojkovic, P., Przyborski, S., Cooke, M., Armstrong, L., Lako, M., and Stojkovic, M. (2006). Derivation of human embryonic stem cells from developing and arrested embryos. Stem Cells *24*, 2669-2676.

Zhang, Z., Gao, Y., Gordon, A., Wang, Z.Z., Qian, Z., and Wu, W.S. (2011). Efficient generation of fully reprogrammed human iPS cells via polycistronic retroviral vector and a new cocktail of chemical compounds. PLoS One *6*, e26592.

Zheng, C., and Baum, B.J. (2005). Evaluation of viral and mammalian promoters for use in gene delivery to salivary glands. Mol Ther *12*, 528-536.

Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y.*, et al.* (2009). Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell *4*, 381-384.

Zhou, W., and Freed, C.R. (2009). Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. Stem Cells *27*, 2667-2674.