Establishment and optimization of induced pluripotent stem cell technologies

Amir Morshedi

2013

Amir Morshedi. (2013). Establishment and optimization of induced pluripotent stem cell technologies. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/55295

https://doi.org/10.32657/10356/55295

ESTABLISHMENT AND OPTIMIZATION OF INDUCED PLURIPOTENT STEM CELL TECHNOLOGIES AMIR MORSHEDI 2013

NANYANG TECHNOLOGICAL UNIVERSITY

Establishment and Optimization of Induced Pluripotent Stem Cell Technologies

AMIR MORSHEDI

SCHOOL OF BIOLOGICAL SCIENCES

2013

Establishment and Optimization of Induced Pluripotent Stem Cell Technologies

AMIR MORSHEDI

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in fulfilment of the requirement for the degree of Doctoral of Philosophy

2013

Acknowledgements

Firstly, I wish to express my deepest gratitude to my supervisor, Prof Peter Dröge, for his constant guidance and support during my research which enabled me to develop an understanding of the subject.

My appreciation also goes to my labmates for their help and providing a pleasant environment for during research.

I have furthermore to thank Agency for Science, Technology and Research (A*STAR) for providing the scholarship as well as School of Biological Sciences/Nanyang Technological University for providing the opportunity for my graduate studies as well as financial support.

Last but not least, I would also like to thank my wife, Monireh, for her inspiration and love and my family for their patience and encouragements during my studies.

TABLE OF CONTENTS

Abbreviations

AP	Alkaline Phosphatase
BMP4	Bone Morphogenetic Protein 4
BSA	Bovine Serum Albumin
CPVT	Catecholaminergic polymorphic ventricular tachycardia
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
EB	Embryoid Body
EGFP	Enhanced Green Fluorescent Protein
EGFR	Epidermal Growth Factor Receptor
EMT	epithelial to mesenchymal transition
ESC	Embryonic Stem Cell
FACS	Fluorescence Assisted Cell Sorting
FBS	Fetal Bovine Serum
FSC	Forward Scatter
GFP	Green Fluorescence Protein
GOI	Gene of Interest
HKG	Housekeeping Gene
HMGA2	high mobility group AT-hook 2
hTERT	human telomerase reverse transcriptase gene
ICM	Inner Cell Mass
iMEF	Inactivated MEF
iPSC	Induced Pluripotent Stem Cell
IRES	Internal Ribosome Entry Site
LB	Luria-Bertani

LIF	Leukemia Inhibitory Factor
LTR	Long Terminal Repeat
MEF	Mouse Embryonic Fibroblasts
mESC	Mouse Embryonic Stem Cell
miRNA	MicroRNA
MET	Mesenchymal Epithelial Transition
MMuLV	Moloney Murine Leukemia Virus
mRNA	Messenger RNA
NEA	Non- Essential Amino Acid
NEB	New England Biolabs
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PSi	Packaging Signal
SSC	Side Scatter
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
FSC	Forward Scatter
SSC	Side Scatter
HKG	Housekeeping Genes
GO	Gene Ontology
GOI	Gene Of Interest
SSEA	Stage Specific Embryonic Antigen
TGF-β	Transforming Growth Factor-Beta
tMEF	Transduced MEF
UTF-1	Undifferentiated Transcription Factor-1

Abstract

Reprogramming of adult somatic cells into an embryonic stem cell (ESC) state by various transcription factors has been the start of a new era in the field of biomedical sciences. This finding has had tremendous impact on drug discovery and disease modeling and there is great hope for these cells to replenish body cells for therapy. An important aspect of this is the ability to monitor and acquire cells which are pluripotent, providing suitable cells for therapy. Undifferentiated transcription factor 1 (UTF1) belongs to the core transcriptional network regulating pluripotency and its expression pattern during cell reprogramming and subsequent differentiation appears to be tightly connected to the pluripotent state. Regarding these features we generated a reliable reporter for monitoring induced pluripotent stem cell (iPSC) formation and differentiation. Our construct indicated functionality both transiently or upon integration into the genome. Furthermore fluorescent iPSCs derived from mice carrying the construct indicated that this reporter is a feasible tool for biomedical research. Note that the small size of our cassette enables easy delivery by different means into the cell. In overall these characteristics qualify our reporter as a reliable reporter system to monitor iPSCs.

We looked into the nonhistone chromatin factor called high mobility group AT-hook 2 (HMGA2) normally expressed in ESCs and during early developmental stages. Aberrant expression of this protein has shown to impact body stature, diabetes mellitus and heart development. Furthermore it plays an important role in cancer development and metastasis. Here we studied HMGA2's role in iPSCs to better understand its function regarding pluripotency. Gene profiling of HMGA2 overexpressing iPSCs gave us insight into the biology of HMGA2 in these cells. Gene ontology analysis revealed that anatomical/Developmental processes are strongly affected by HMGA2 with cell adhesion and differentiation process coming next. Furthermore our data indicated that key diabetes susceptibility genes are affected by HMGA2, revealing interesting link to the *Lin28/let-7* pathway regulating mammalian glucose metabolism. Our data support the model that HMGA2 is necessary for maintenance of the pluripotent state and its overexpression predisposes cells into specific lineages during differentiation.

I. INTRODUCTION

I.1. Embryonic stem cells

I.1.1. Overview

Embryonic stem cells (ESCs) are a group of cells that have the potential to develop into all body cell types. They differ from adult stem cells in their ability to differentiate into a wider range of cells and thus are said to be "pluripotent". After birth, adult stem cells remain in parts of the body as a source of cells for replenishment (Liu et al., 2005; Phinney and Prockop, 2007).

ESCs are isolated from the blastocyst of the embryo. The blastocyst, a cluster of cells formed during mammalian embryonic development, is made of trophoblast and inner cell mass (ICM) where trophoblast forms the placenta and ICM is destined to form the embryo. ESCs are initially derived from ICM extracted from the blastocyst during embryonic development, and can grow in culture and differentiate into the three germ layers markedly endoderm, ectoderm and mesoderm (Evans and Kaufman, 1981).

When removed from ICM and cultured in medium containing substances like leukemia inhibitory factor (LIF) and bone morphogenetic protein 4 (BMP4), ESCs sustain pluripotency for prolonged periods until they are destined to differentiate. The remarkable capacity of ESCs to form different germ layers has made them a great candidate for transplant therapies. These cells can act as a replenishment for cells in different diseases like spinal injury, heart failure and stroke, diabetes, Alzheimer's disease and many more (Figure I.1) (Qi et al., 2004; Thomson et al., 1998).

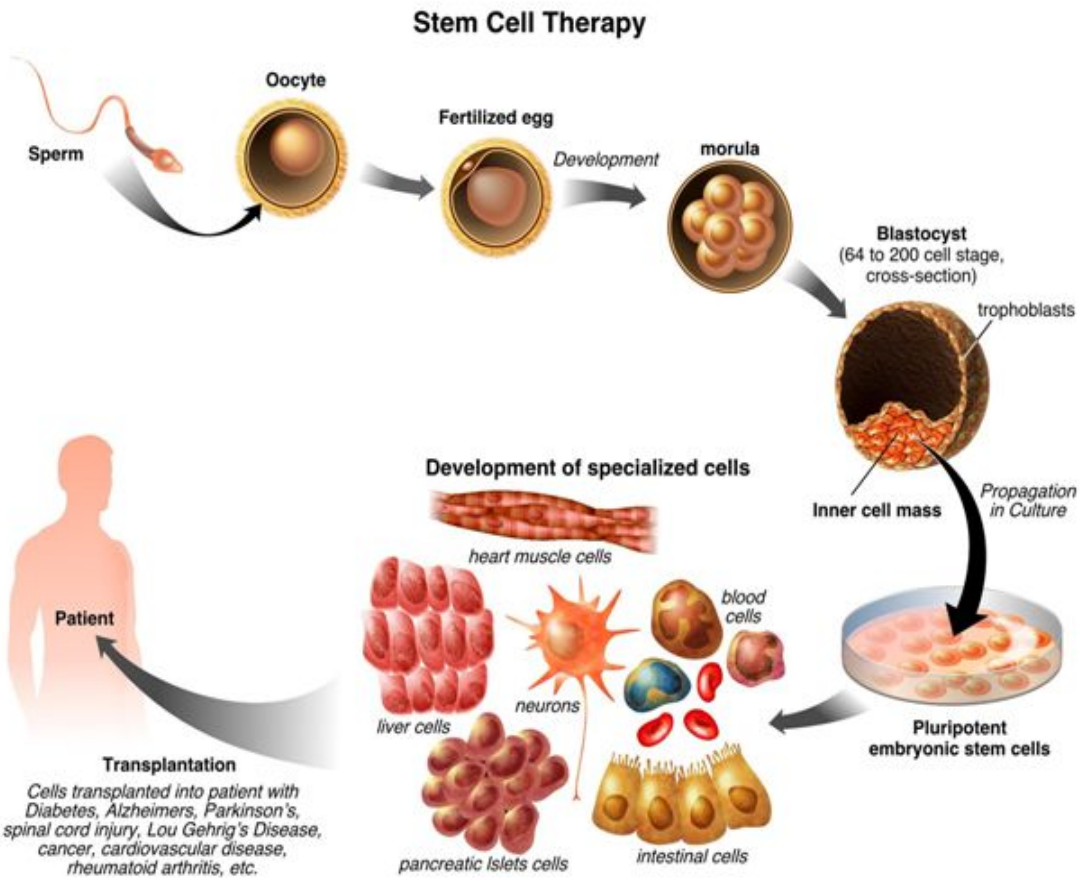


Figure I.1. Embryonic stem cell development and application in therapy**.** The ICM is put into culture to form ESCs. Blastocyst is the embryo in the blastula stage which is one of the stages in mammalian embryonic development after morula. ESC are pluripotent meaning that they have the potential to differentiate into all cell derivatives from three germ layers (mesoderm, ectoderm and endoderm) such as circulatory system cells, nervous system cells and other types of cells which can be used for therapy. (Image is downloaded from "Stem Cells for Hope" website at http://www.stemcellsforhope.com/Stem Cell Therapy)

Other than their therapeutic potentials, ESCs have also shown to be of great importance in research. Once differentiated, these cells can be used for drug screening or the study of particular cell types. These cells have also played a big role in the establishment of knockout mice which has been a powerful tool in biomedical and biological research. This breakthrough owed its founders the Noble Prize in medicine and physiology in 2007 (Cohen-Tannoudji, 2007; Evans and Kaufman, 1981; Thomson et al., 1998).

I.1.2. Characteristics and applications of embryonic stem cells

There are several features that notably mark ESCs as pluripotent. Once cultured under proper conditions, they form distinct colonies with a round shape. Diminished cytoplasm and large nucleolus is generally seen in these cells. Human ESC (hESC) resemble mouse ESCs (mESCs) in morphology with some small difference. For example, mESCs tend to be more round and packed in comparison to their human counterparts (Filip et al., 2004).

ESCs express an array of proteins specific to the pluripotency stage, namely the pluripotency markers. Some of these markers include OCT4, NANOG, SOX2, NR5A2 and UTF1 form the pluripotency core in the ESC transcriptional network which regulate ESC renewal. OCT4 and NANOG shape the basis of different assays to detect pluripotent stem cells (Cavaleri and Scholer, 2003; Constantinescu, 2003; Deb-Rinker et al., 2005). Other ESC markers include ESC specific surface markers like stage specific embryonic antigen 1 (SSEA1) in mESCs and SSEA4 in hESC (Rao et al., 2007).

ESCs display increased amounts of alkaline phosphatase (AP) activity on the surface of their membrane when undifferentiated. Hence, AP staining is commonly used to measure the degree of pluripotency in ESCs (Pease et al., 1990).

Pluripotent ESCs injected into immune deficient mice form tumor structures called teratomas, containing a mixture of cells from the three different germ layers, indicating the ability of ESCs to form different cell types. Teratoma formation assay is used to validate if pluripotent cells can differentiate properly *in vivo* (Li et al., 2008).

More demanding pluripotency assessments include germline transmission and tetraploid complementation which only high quality ESCs are positive for these tests. Germline transmission is the ability of these cells to contribute to offsprings. This capability

enables the generation of transgenic mice from ESCs. In tetraploid complementation assay, ESCs are injected into a tetraploid embryo. The resulting fetus is exclusively derived from the injected ESCs which shows that the ESCs alone are able to give rise to offsprings (Nagy et al., 1990).

As previously mentioned, ESCs have many diverse applications both in medicine and research. They are great candidates for transplant therapy and several test therapies using ESCs have been carried out in humans over the past few years. Furthermore, these cells are worthy models for use in drug discovery and toxicology experiments, (Zou et al., 2009) and as previously described, their usage in generating transgenic mice has made them a powerful tool in biomedical research (Cohen-Tannoudji, 2007; Evans and Kaufman, 1981). Nevertheless, they still retain some limitations which have made their application controversial. Concerns about the rejection of these cells upon transplantation as well as ethical issues regarding the use of embryos, has made their utilization somewhat limiting (Takahashi and Yamanaka, 2006). Although the use of nuclear transferred ESCs overcomes tissue rejection, their ethical issues still remain as a major concern. These restrictions have motivated scientists to look for novel approaches in order to generate pluripotent cells.

I.2. Induced pluripotent stem cells

I.2.1. Overview

Due to designated limitations imposed by the use of ESCs in research and therapy, researchers have been searching for novel approaches to acquire pluripotent cells. Up to date, several methods have been employed to achieve this goal. However, each method

still preserves its own limitations. Somatic cell nuclear transfer (SCNT) is one of the primary methods used for this purpose. In this method the nucleus of the donor cell is fused to a denucleated egg cell and stimulated to form the embryo. Although this method overcomes tissue rejection, it still retains ethical concerns (Wilmut et al., 1997). Another method which is rapid and efficient is cell fusion; however, the cells obtained are tetraploid which limits their usage. In this process somatic cells and pluripotent cells are fused to generate tetraploid pluripotent cells. Yet a different approach is to isolate pluripotent cells from unipotent cells in culture, although the pluripotency of the cells is still controversial (Cowan et al., 2005; Tada et al., 2001). For example, primordial germ cells or spermatogonial stem cells are unipotent *in vivo*; however, pluripotent ES-like cells can be isolated from these cells after prolonged culture *in vitro* under appropriate conditions (Jaenisch and Young, 2008).

Probably the most outstanding method which overcomes most of previous issues is the induction of pluripotent cells from somatic cells by introducing key pluripotency factors into the cells (Figure I.2). *Oct4, Sox2, Klf4* and *cMyc* (4F) are the initial factors used by Yamanaka and colleagues for pluripotent iPSC generation (Takahashi and Yamanaka, 2006).

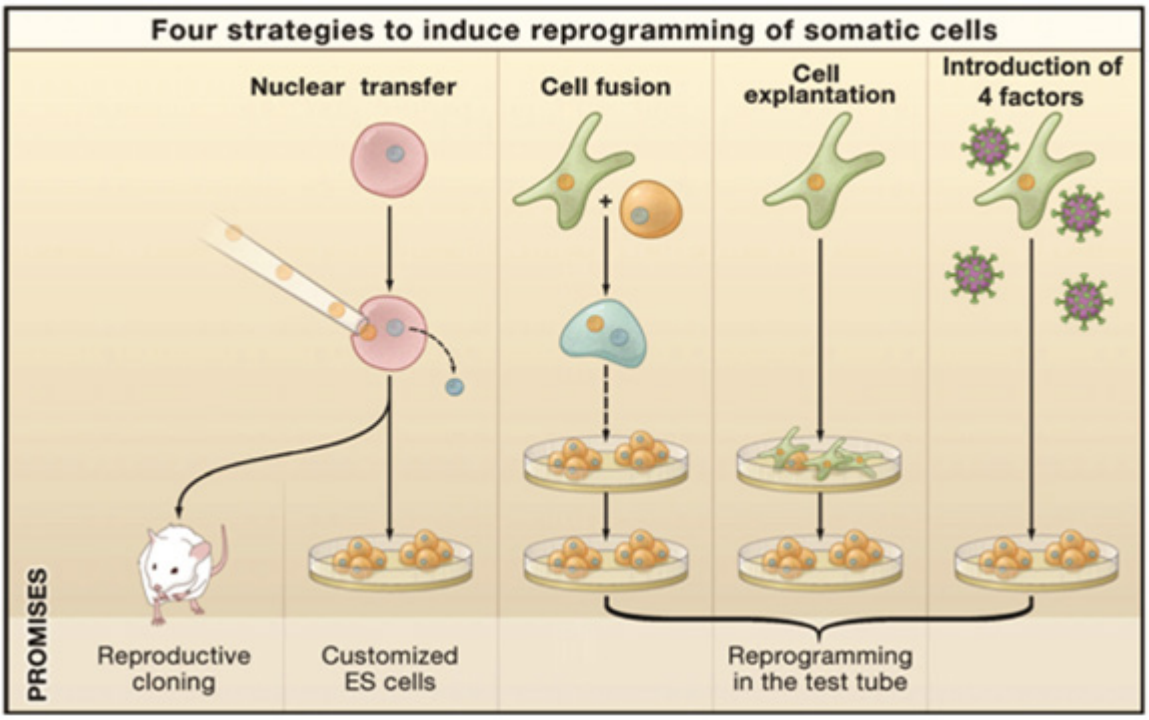


Figure I.2. Various approaches used for pluripotency induction. From left to right in sequence: Somatic cell nuclear transfer: Somatic cell is injected into the embryo to derive ESCs. Fusion of somatic cell with ESC: This method generates fused cells which are 4n and pluripotent. Cell explanation: Specific medium conditions are used to isolate pluripotent-like cells. Pluripotency induction in the presence of four reprogramming factors: Four factors are introduced by different methods to reprogram the cells into a pluripotent state which generate cells so-called iPSCs. (Jaenisch and Young, 2008).

This breakthrough discovery by Yamanaka and colleagues in 2006 has brought tremendous hope into stem cell research, removing many barriers it used to face. Induced pluripotent stem cells (iPSCs) are pluripotent cells resembling ESCs in pluripotency and characteristics. Although excessive research is required to prove these cells fully resemble ESCs, recent studies have shown that at least a group of iPSCs clearly resemble ESCs, which leaves these cells as the best replacement. These cells face no ethical issues and can be acquired much easier than ESCs. Furthermore, because they can be derived from autologus sources they can overcome tissue rejection (Jaenisch and Young, 2008; Takahashi and Yamanaka, 2006). A recent study has showed that iPSCs still show some

tissue rejection due to abnormal expression of some proteins; however, there is still controversy over the quality and the type of iPSCs used. It is not clearly understood whether differences between ESCs and iPSCs are inherited or associated with the generation process of iPSCs itself (Stadtfeld and Hochedlinger, 2010). Indeed, recent studies indicate that factors like genetic background (Stadtfeld et al., 2010), use of viral vectors (Soldner et al., 2009), lab-to-lab variations and passage number (Chin et al., 2009; Polo et al., 2010) can have profound effects on gene expression and function of iPSCs. This could be the result of small differences between some iPSCs's and ESCs's gene expression which is the main reason why scientists are aiming for techniques that optimize iPSC generation or to isolate those iPSCs which fully resemble ESCs (Zhao et al., 2008).

iPSCs have diverse potentials but still face some challenges prior to use in therapy. These challenges include low efficiency, pluripotency characteristics, clinical safety, efficient differentiation and availability in large scale for transplantation purposes (Lu et al., 2012; Saha and Jaenisch, 2009). These challenges will be discussed throughout the text.

I.2.2. Induction of pluripotency

Initially for inducing pluripotency, Yamanaka and colleagues tested a group of 24 factors regarded as eminent ESC related genes. Subsequently, the most essential genes were filtered by withdrawal of unnecessary genes. At last, four genes namely *Oct4, Klf4, Sox2* and *c-Myc* remained which were indispensable to derive iPSC colonies (Takahashi and Yamanaka, 2006). Later studies showed that fewer factors can also be used to derive iPSCs, although iPSC generation efficiency declined considerably. For example reports

have shown that neural stem cells can be reprogrammed with only OCT4, probably, due to the basal expression of other factors in these cells (Kim et al., 2009b). OCT4 has shown to be pivotal for iPSC derivation as it is a core transcription factor for pluripotency.

Retroviral vectors were initially used to carry the factors into the cell enabling stable expression for prolonged periods (Takahashi and Yamanaka, 2006). Retroviruses still carry some features which make them the best choice for efficient iPSC production; they can quickly integrate, which enables their stable expression through cell division. This means that they will retain expression in cells through multiple independent maturation steps (Cepko and Pear, 2001). However they are not suitable for clinical applications and can only be used for research purposes. Use of retro and lentiviruses leads to genome alterations, making these method somewhat problematic for therapeutic purposes (Okita et al., 2008). That is why scientists are moving towards tactics which do not use retro/lentiviruses including RNA, protein and chemical compounds to achieve pluripotency. Some of these methods have been successful; nevertheless, they show a very low efficiency in comparison to retroviral transduction, requiring further optimization or modifications. For example introducing proteins or RNA into the cell eliminates risks of integration and can produce iPSCs which are safe for clinical applications. Due to degradation of proteins and difficulty of introducing proteins into the cell, these methods are further required to be optimized (Kim et al., 2009a; Li et al., 2009; Yakubov et al., 2010). A recent study has shown that iPSCs can be derived solely with the use of chemical compounds, yet the safety of the chemical compounds remains to be evaluated (Hou et al., 2013).

There is an array of different approaches to increase the efficiency of iPSC generation. One is the use of extra factors like LIN28 or UTF1 which increase the number and the pluripotency of the cells (Liao et al., 2008; Zhao et al., 2008). Other factors like TBX3 increases germline transmission of iPSCs making them more useful in research and therapy (Han et al., 2010). Additional approaches include the use of chemical compounds like vitamin C and MEK/ERK inhibitors. Vitamin C seems to ease iPSC generation by inhibiting P53. Inhibitors of MEK/ERK pathway and TGF-β suppress mesenchymal to epithelial transition (MET) which facilitates progression into iPSCs (Esteban et al., 2009; Lin et al., 2009). Amazingly, even environmental conditions like oxygen levels can profoundly affect iPSC generation (Yoshida et al., 2009).

cMyc has long been a concern in iPSCs due to its tumorigenic properties. Nevertheless iPSCs have been derived without *c-Myc* or by replacing *c-Myc* with other factors, although efficiency drops significantly (Nakagawa et al., 2008).

I.2.3. Characteristics of Induced pluripotent stem cells

iPSCs demonstrate great similarity to their counterpart ESCs. These cells, similar to ESCs, are positive in different pluripotency characteristics. Similarly, they form teratomas once injected into mice and can form embryoid bodies (EBs) in appropriate culture conditions *in vitro*. Core pluripotency factors like OCT4 and NANOG are also activated in iPSCs and can be used to assay pluripotency. iPSCs have been isolated which are positive for germline transmission and tetraploid complementation, thus generating fully iPSC derived offspring (Yang et al., 2007; Zhao et al., 2009). Fairly recent studies have demonstrated that a cluster of genes on chromosome 12 (12qF1) in iPSCs are

responsible for germline transmission of these cells (Stadtfeld et al., 2010). iPSCs tend to retain some epigenetic memory of the cells which they were derived from, affecting their application. This epigenetic memory can be used to reproduce iPSCs from differentiated cells with increased efficiencies. On the other hand, the epigenetic state may interfere with normal transplantation of these cells into host by causing immune rejection. Doxycyclin inducible vectors have been key players in iPSC gene expression analysis acting as dynamic tools to control iPSCs generation and differentiation. Studies using these vectors have shown that reprogramming cells go through a MET step before forming iPSCs. This MET is regulated by BMP signaling together with the Yamanaka factors. This process induces microRNA (miRNA) expression signatures and drives the progression through the primary phase. Many pluripotency factors are activated early, including SSEA1 and OCT4 in the initial phase of iPSC generation and NANOG in a later phase. Factors like ESSRB, REX1 and UTF1, which are activated respectively in a later phase are more adherent to the pluripotent state and are only expressed in fully reprogrammed cells (Figure I.3.) (Samavarchi-Tehrani et al., 2010).

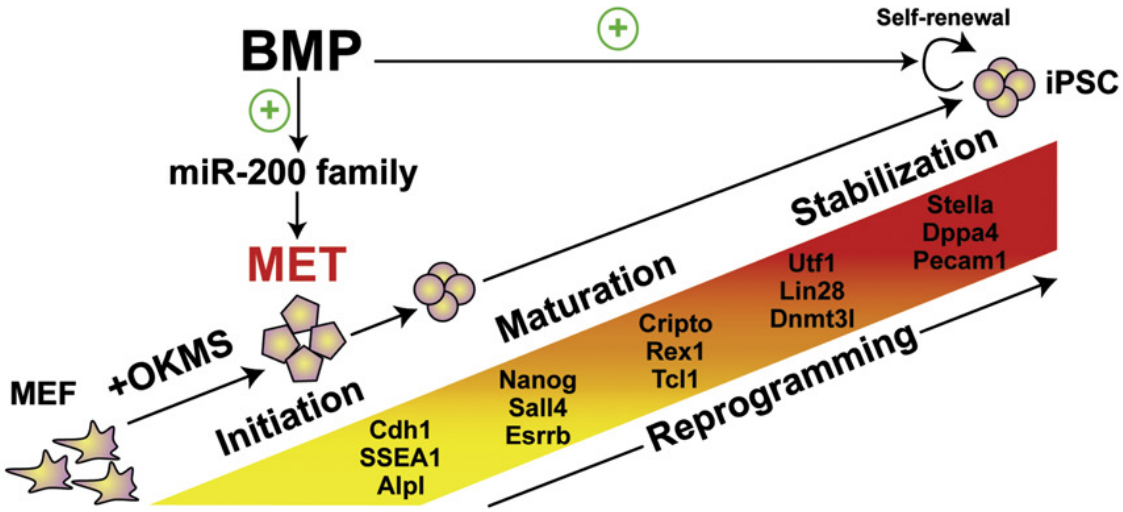


Figure I.3. Course of cellular reprogramming. Model depicts mouse embryonic fibroblast (MEF) cells during reprogramming to iPSCs. The BMP signalling synergizes together with the four Yamanaka factors (OKMS) to drive through the initiation phase (Samavarchi-Tehrani et al., 2010).

Currently, immense effort is being put into pluripotency characteristics and efficiency improvement of these cells by different groups so that they can serve as a "true replacement" to ESCs in the near future.

An important group of elements which control pluripotency are miRNAs. MiRNAs regulate pluripotency in many aspects. For instance, miR-145 expression favors differentiation and is expressed at very low levels in ESCs. Major pluripotency factors such as OCT4, SOX2 and KLF4 are direct targets of this miRNA (Xu et al., 2009). On the other hand, a group of miRNAs are capable of inducing pluripotency. Amazingly, a cluster of miRNAs called miR-302/367 can effectively reprogram both mouse and human cells without the use of other transcription factors. This cluster together with valproic acid is used as an efficient non-integrating technique to produce iPSCs (Anokye-Danso et al., 2011).

I.2.4. Application and advantages of induced pluripotent stem cells

iPSCs are promising candidates in research and therapy. The use of iPSCs in disease modelling and drug discovery has already proven advantageous in many aspects while their application in therapy is still under investigation (Zou et al., 2009). For example iPSCs derived from Catecholaminergic polymorphic ventricular tachycardia (CPVT) patients, have been used for drug discovery. This disease causes rythmic disorder of the heart. The drug Dantrolene binds to an amino terminal region called RYR2 and restores interdomain interaction necessary for closed state, hence improving the arrhthmic behavior in cells derived from these patients. Other drugs like Nifedipine and Pinacidil have been tested on iPSCs-derived from long QT-2 mutant cardiomyocytes which show different characteristics to iPSCs derived from normal individuals (Itzhaki et al., 2011; Jung et al., 2012).

The potential of these cells to be used for transplantation therapy has attracted great attention. Although iPSCs have not yet been used for therapy in humans, they have shown promising results in laboratory animals. Cells isolated from a patient can be modified and corrected for their genetic defects and sent back into the patient through iPSCs (Figure I.4). For example, correction using homologous recombination for the PIG-A gene in iPSCs has been shown previously (Zou et al., 2009). In addition, hematopoietic progenitors acquired from autologus iPSCs *in vitro* were able to rescue mice suffering from sickle cell anemia (Sebastiano et al., 2011). Furthermore gene expression analysis studies have indicated that iPSCs and ESCs are similar and any dissimilarity between these cells is the result of culture conditions or genetic background. Immense number of studies in this field have indicated that iPSCs show all the

pluriptotency characteristics of ESCs and to this day have proven to be a worthy replacement, although methods are required to "pick the best" iPSCs. These finding indicate that it is very likely that these cells can replace human body cells in the near future.

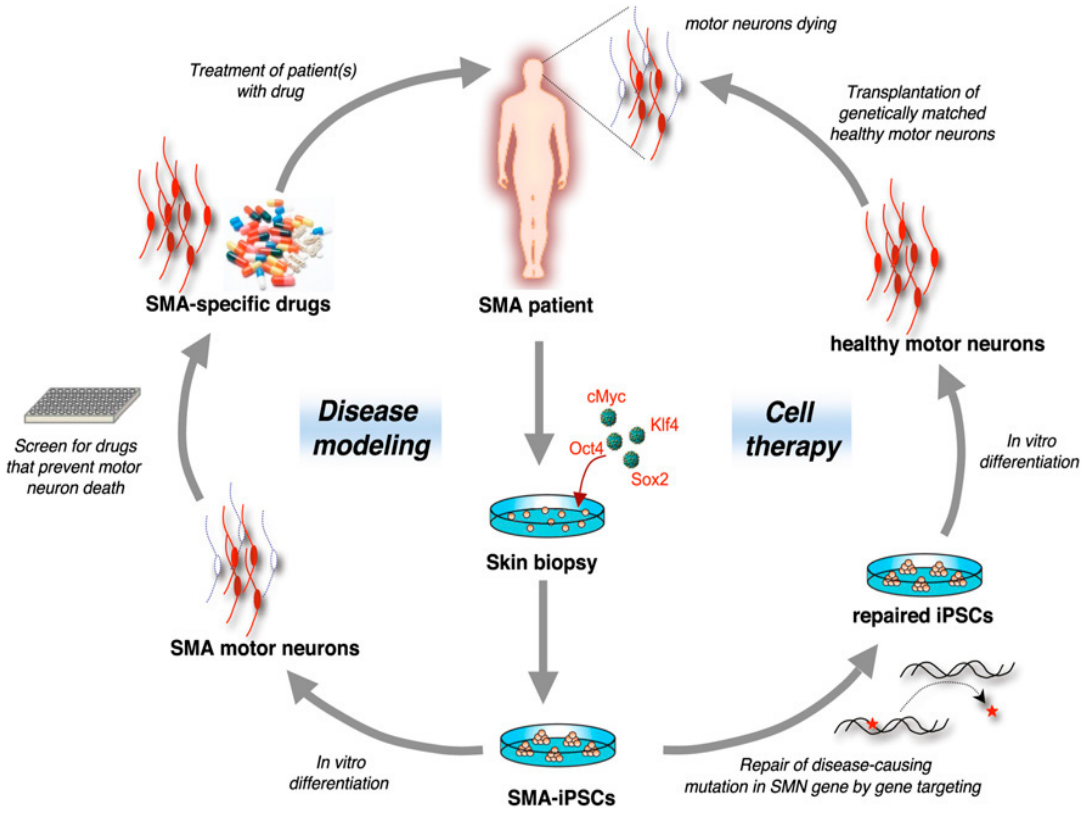


Figure I.4. Potential applications of iPSCs. The figure shows the potential use of iPSCs in therapy and drug research using spinal muscular atrophy (SMA) as an example. In these patients, motor neurons are defective and die resulting in the devastating symptoms of the disease. iPSCs derived from SMA-specific patients can be differentiated into motor neurons *in vitro* in order to create a culture model of the disease that may result in the identification of new drugs to cease abnormal cell death in these patients. Alternatively, if the disease is known, the mutation can be repaired (the SMA gene in this case) before differentiation into healthy motor neurons, followed by transplantation into the patient's brain (Stadtfeld and Hochedlinger, 2010).

I.3. Pluripotency related genes

There are a cluster of genes connected to the pluripotent state which decline upon cell differentiation. One of the core pluripotency transcription factors, undifferentiated transcription factor-1 (UTF1) is tightly restricted to pluripotency. Another protein, high mobility group AT-hook 2 (HMGA2) is a non-histone DNA binding protein which is also expressed in pluripotent cells and early embryonic development. As these proteins were the focus of our work they will be discussed in detail in the following sections.

I.3.1. Undifferentiated transcription factor 1

UTF1 is a transcription factor restricted to pluripotent ESC. This protein is part of the pluripotency transcriptional network, associated with the chromatin (Okuda et al., 1998). Transcriptional activation of many genes is regulated by this factor and is generally carried out by at least the interaction with an upstream factor ATF2 (Fukushima et al., 1998). OCT4 and SOX2 are well known to interact with UTF1 enhancer and NANOG putatively binds to another site near the enhancer in order to sustain pluripotency. These interactions are thought to be carried out through a TATAAT sequence. A recent study has showed that hundreds of genes are co-regulated by UTF1 in ESC and that this gene is essential to prevent chromatin decondensation in ESC (Kooistra et al., 2010). Furthermore, heavy methylation of UTF1 promoter in somatic and partially reprogrammed cells is responsible for regulation of UTF1 expression (Mikkelsen et al., 2008).

An interesting feature of UTF1 is that it shows late expression during reprogramming and rapidly declines upon differentiation, indicating its stringent adherence to the pluripotent

state (Mikkelsen et al., 2008; Samavarchi-Tehrani et al., 2010). UTF1 is expressed later than many pluripotency markers like OCT4, NANOG and SSEA1 during reprogramming and transient states do not seem to show UTF1 expression (Samavarchi-Tehrani et al., 2010). Late expression of this factor is beneficial for distinguishing pluripotent cells from cells which are still in a transient state of reprogramming. This protein presumably has a role in proper differentiation of ESCs. Diminished UTF1 levels in mESCs prevent these cells from ordinary differentiation (van den Boom et al., 2007). Previous studies have shown that hESCs expressing UTF1 show high expression of other pluripotency factors and ESC specific markers like TRA-1-80 and TRA-1-60 (Pfannkuche et al., 2010). Other studies on iPSCs have shown that UTF1 introduced with the four Yamanaka factors for reprogramming of mouse embryonic fibroblasts (MEF), increases AP positive iPSC colonies dramatically.

UTF1 is a part of the chromatin and acts as an intermediate in regulating pluriptoency and differentiation states by linking the main pluripotency core to the myc and PRC2 networks. This factor regulates bivalent genes state by competing with PRC2 a known factor in chromatin silencing. This factor also suppresses gene expression by its association with mRNA decapping factors like Dcp1a (Jia et al., 2012).

The molecular functions of UTF1 are still not clearly understood. However unpublished work by our collaborators showed that homozygous Knockout UTF1 mice are still viable and do not show any apparent phenotypes. The stringency of UTF1 expression in pluripotent cells has made this factor a promising candidate for detecting highly pluripotent stem cells.

I.3.2. High mobility group AT-hook 2

The oncofetal non-histone protein HMGA2 is mainly expressed in ESCs and cancer cells. This factor is expressed during early developmental stages but declines eventually. The protein is known as a chromatin factor which is also linked to the heterochromatin. It is involved in the regulation of different genes through binding the AT rich double stranded DNA of the promoters. The domain mediating interaction with the DNA is called the AT hook DNA binding domain. HMGA2 is regulated through *let-7* miRNA degradation and upon *lin28* expression (Chiappetta et al., 2008; Droge and Davey, 2008; Fedele et al., 2010). It is involved in genome stability maintenance and seems to interact directly with pRB, a tumour suppressor protein (Fedele et al., 2010; Summer et al., 2009). A direct link has been drawn between HMGA2 and the metastatic potential of tumours. Other studies suggest that this gene may well induce the expression of the telomerase gene (Li et al., 2011).

This factor has been primarily associated with body size in different species. Its increased expression correlates with human height in the general population. In this respect, different studies have linked this factor to bone density as well. HMGA2 overexpression in chicken, dog and mouse results in gigantism and increased adipose tissue mass (Battista et al., 1999; Sanna et al., 2008; Song et al., 2011; Winkler et al., 2007). On the other hand, its insufficiency or lack causes dwarfism. Most of these regulations seem to be taking effect during pre- and post-natal developmental stages. We recently provided evidence that SNP regulation of HMGA2 may be a result of miR-196b, possibly during parental development (Tay et al., 2009). Earlier studies indicate that HMGA2 is also

expressed in young neural stem cells as a crucial factor for self-renewal (Tzatsos and Bardeesy, 2008).

I.4. Reporter systems and stem cell pluripotency

Linking reporters like enhanced green fluorescent protein (EGFP), DsRed, luciferase, etc to different promoters and enhancers to monitor gene expression has been previously carried out for different purposes. These reporter systems have been used to mark different states of the cell or identify cells which express the gene of interest. Following introduction of the reporter construct into a specific cell or animal, the expression of the reporter will give an idea of the level of gene expression induced by the regulatory sequences (Schenborn and Groskreutz, 1999). Different reporter systems have been exploited to detect and monitor pluripotency in cells based on their expression pattern. Primary studies of iPSCs attempted the use of *Fbx15* to isolate iPSC colonies; however, some non-pluripotent cells still expressed this factor (Takahashi and Yamanaka, 2006). *Oct4* and *Nanog* knockin cell lines have been used extensively in this manner. MEF cells derived from *Oct4*-EGFP or *Nanog*-EGFP knockin mice, generated iPSCs which express EGFP (Brambrink et al., 2008; Maherali et al., 2007; Wernig et al., 2007). Transient constructs containing *Rex1* regulatory elements were previously used to isolate pluripotent stem cells (Eiges et al., 2001). Additional studies reported the use of *Tert1*, *Oct4* and *Terf1* promoter driven EGFP or NTR gene cassette to monitor reporter silencing in ESCs (Stewart et al., 2008). UTF1-neo construct has been previously used in our laboratory to enrich highly pluripotent ESCs in culture which showed higher expression of pluripotency markers in compared to non-selected cells (Tan et al., 2007).

An interesting approach is the use of an early transposon promoter with *Oct4* and *Sox2* enhancers by exploiting a lentiviral vector to monitor iPSCs. The integration of lentiviral vector allows monitoring the cells over a long period of time, though it might cause mutations (Hotta et al., 2009a).

I.5. Differentiation of embryonic stem cells and induced pluripotent stem cells

iPSCs have shown to differentiate to different cell types, similar to their counterpart ESCs. ESCs/iPSCs have the capacity to generate hundreds of cell types, creating tremendous opportunities in regenerative and biomedical research. The bases of most diseases lie in cellular deficiency. Heart failure and stroke, diabetes, neurological disorders, blindness, spinal cord injury and many other diseases result from the absence of a particular group of cells which the body is also unable to replace. ESCs/iPSCs provide the basic elements needed to generate specific cell population for tissue repair and regeneration. Although this plasticity enables ESCs/iPSCs to generate various cell types, it also makes the derivation of specific cell types more difficult to control. Generally, there are three methods that form the basis for ESC differentiation:

(1) Formation of round clusters called EBs.

(2) Culture of ESCs on extracellular matrix proteins like gelatine as a monolayer

(3) Culture of ESCs on connective tissue layers as support.

Each method has its own features depending on the application. Random generation of differentiated cells uses fetal calf serum (FCS) (or fetal bovine serum (FBS) which has a variety of stimulating factors. However, recent methods for derivation of specific cell

types use serum free media to minimize variety of factors (Kubo et al., 2004; Ng et al., 2005; Yasunaga et al., 2005). Furthermore, reporters have been generated which signal the start of differentiation making it more convenient to acquire the cell type of interest (Fehling et al., 2003; Gadue et al., 2006).

Generally, formation of different cell types depends upon signalling of different molecules and receptors. Reorganization of embryo into three germ layers takes place through a process called gastrulation which is through a structure called the primitive streak (PS). This structure will define the anterior and posterior and the fate of the embryo through different signalling pathways like Wnt, BMP4, and Activin (Figure I.5) (Fehling et al., 2003; Gadue et al., 2006; Murry and Keller, 2008).

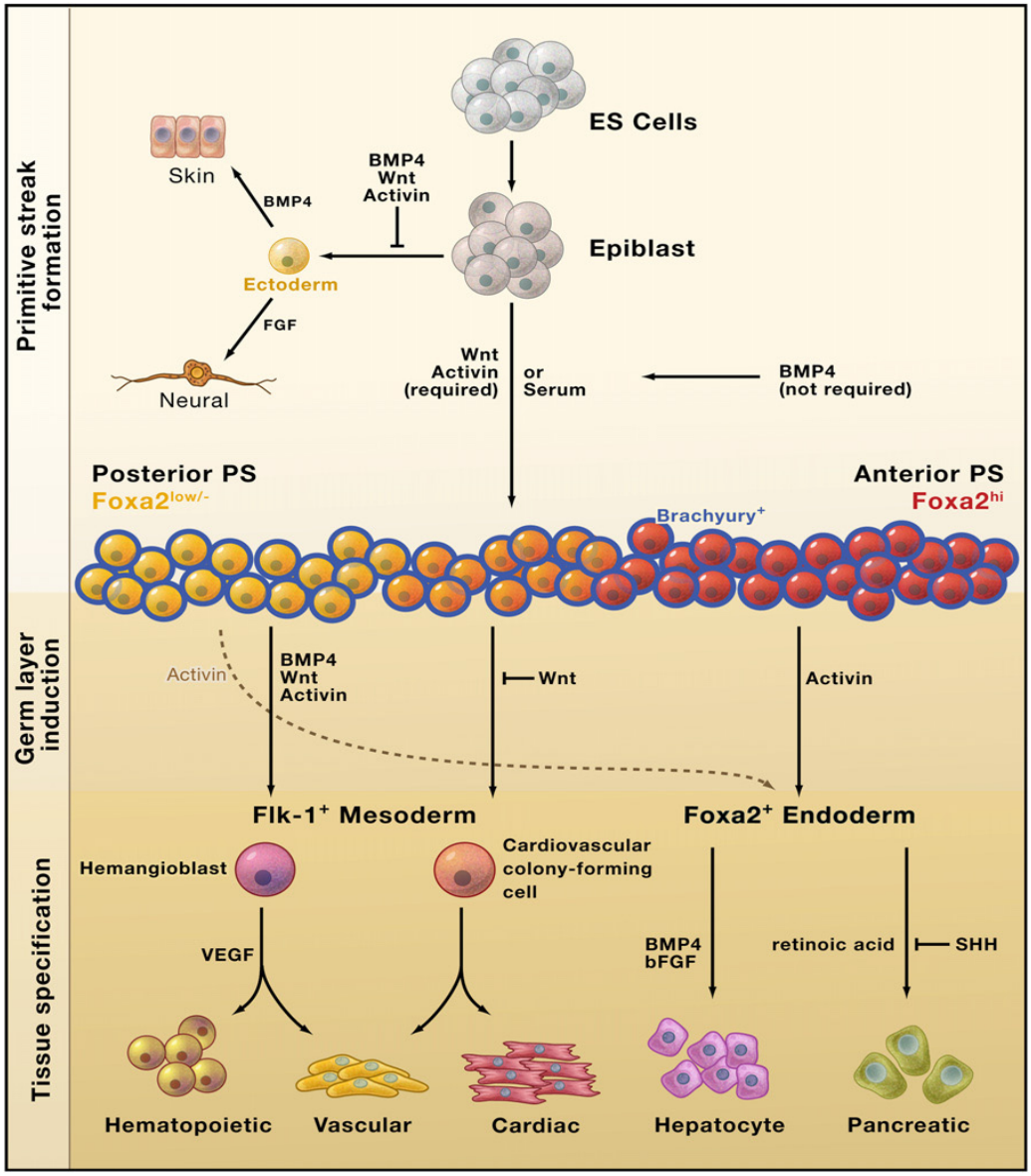


Figure I.5. The model demarcates primary germline induction and tissue specification from differentiated mouse ESCs. The process is initiated by the formation of the primitive streak (PS) like cells. If further activation doesn't take place, cells move into form the ectoderm. Once further activated by BMP4, Wnt and Activin, cells move into forming the PS. The posterior cells (yellow) form the mesoderm while the anterior cells form (orange) endoderm. This is not yet firmly established as Activin can redirect posterior cells into endoderm differentiation. Finally, different cell types are generated from these cells. (Murry and Keller, 2008)

Cardiomyocytes are readily detectable in EBs generated from ESCs stimulated by serum. A low efficiency of less than 3% is usually seen in random differentiation experiments.

(Doetschman et al., 1985). One of the early approaches for directed differentiation of ESCs towards cardiomyocytes involves using medium conditioned with endodermal cell line End-2 which produces activinA and BMP and several other factors which has efficiency of more than 12 percent. Recently this technique was enhanced greatly by using p38 MAP kinase inhibitors, which increases derived cardiomyocytes from hESCs by more than double the amount of previous methods, thus improving mesoderm induction (Graichen et al., 2008). Other studies have shown that addition of vitamin C enhances cardiomyocyte differentiation extensively in iPSCs (Liang et al., 2011).

Primary methods in neuronal differentiation of ESCs either used retinoic acid (Bain et al., 1995), culture in sequential serum and serum free conditions (Okabe et al., 1996), or using stromal cell lines like PA6 to co-culture with the ESCs (Kawasaki et al., 2000). It is well known that ESCs can form neuronal progenitor cells capable of producing neurons, astrocytes and oligodendrocytes (Joannides et al., 2007). Neural progenitors are generally generated from differentiating ESCs under specified conditions including spheroid (neurosphere) formation in medium containing FGF2 or EGF. ESC derived neural progenitors are similar to adult and fetal neural progenitors in their differentiation ability; however, gene profiling and DNA methylation analysis shows that there are some differences between the two. It is necessary to consider this as they can act differently in many settings. It should be noted that although generation of cells with neuronal morphology is relatively easy, many neuronal subtypes exist with various functions. Hence, acquiring a purified subtype for therapy and research goals is still challenging. Generation of different neural populations has been successful by controlling the different signaling pathways like Notch, Sonic hedgehog and Wnt. For example, Sonic

hedgehog analogs were used to generate spinal motor neurons. Other techniques have generated particular neuronal populations, including progenitors for retinal photoreceptors, cerebellar granule neurons, and cerebral type neurons that possess glutamate, GABA and dopamine as their main neurotransmitters (Murry and Keller, 2008; Wichterle et al., 2002).

Differentiation of patient specific iPSCs into relevant disease specific cell types is strengthening our understanding of cell differentiation and prospects for cellular therapies and *in vivo* regeneration. Although most epigenetic signatures and environmental factors are erased during reprogramming, these cells prove handy for disease with strong genetic basis or high penetrance. Nevertheless environmental conditions can be introduced once these cells are in culture to test how these cells react (Cherry and Daley, 2012).

I.6. Statement of goals and prospects

Our main focus is to generate iPSCs and look into pluripotency by studying two pluripotency related genes. As mentioned, iPSCs are pluripotent cells possessing characteristics of ESCs. Upon reprogramming cells into iPSCs a group of different colonies arise. Of all the colonies that appear, only some are fully pluripotent and can be considered as ESC counterparts. The rest are colonies that are either halted in a transitional state or need more time to form pluripotent cells. Hence monitoring these cells through a pluripotent marker will substantially ease their detection for later applications. Our goal is to produce a novel and reliable reporter which can be used to monitor iPSC for pluripotency. Because UTF1 is strictly expressed in the pluripotent state and is rapidly turned off during differentiation, it is a great candidate to monitor

iPSCs. This factor has shown to have a more stringent expression than other factors like OCT4 and NANOG which are expressed during late stages of reprogramming. Hence expression of this factor marks pluripotent cells from non-pluripotent ones. Furthermore, we require a reporter which can be used to monitor cells over a long period of time, due to the fact that reprogramming takes a substantial amount of time. In brief our major goals are:

1- Generation of iPSC using the four "Yamanaka factors"

2- Generation of a reporter construct on the basis of UTF1 regulatory elements

3- Track UTF1 expression during different stages of reprogramming

4- Testing pluripotency in iPSC colonies isolated with this reporter

5- Monitor differentiation of UTF1 reporter containing iPSC

The other pluripotency related gene studied is HMGA2 which is also expressed during early stages of development. Although many previous studies have pointed to the fact that this factor somehow plays a role in ESCs, yet further studies are required to elucidate the function. This factor has been extensively studied in tumour formation and oncogenesis; however, its function in pluripotent cells is still not clearly understood. Previously, our lab has shown that this gene affects the expression of different pluripotency genes. Our goal is to use this gene to monitor differences between HMGA2 overexpressing and normal iPSC, thus more clearly understanding its role in iPSC and pluripotency. This could help better understand the transition state from non-pluripotent to pluripotent cells.

In brief our goals are to:

1- Generate iPSCs overexpressing HMGA2.

2- Compare HMGA2 overexpressing iPSC and normal iPSC in their pluripotency characteristics.

3- Look into differences of gene expression by monitoring the whole genome expression of these cells.

4- Analyze differentiation of HMGA2 overexpressing cells in relation to normal iPSCs.

II. MATERIALS AND METHODS

II.1. Molecular biology

II.1.1. Restriction digests and ligation

Vectors were excised using restriction enzymes from New England Biolabs (NEB). In general, 1 µl (10u/µl) of enzyme was used to digest 1 µg of DNA for a period of 1 hour. DNA products were separated by gel electrophoresis (Biorad) and if needed extracted by gel extraction kit (Qiagen) for further cloning. PCR purification kit (Qiagen) was used to purify DNA from reaction mixtures. For ligation a ratio of 3:1 and 6:1 insert to backbone was used, using 1 µl of T4 DNA ligase (NEB) (10 u/ µl) at room temperature for one hour or at 16^{o}C overnight.

II.1.2. Cloning and purification of plasmids (Vectors)

For replication of vectors, *E.coli* DH5α cells were transformed by using the heat shock transformation method. Competent cells were removed from -80^{o}C and put onto ice to thaw gradually. 100ng of plasmid DNA was added to the cells and placed into 42^{o}C heat bath for a period of exactly 45 seconds. Cells were then immediately put on ice and chilled for 5 minutes for recovery. 900µl of lysogeny broth (LB) medium was added to the cells and placed into the incubator for one hour in 37 ^{o}C, shaking at 200 rpm. Subsequently, 25 µl and 100 µl of the medium were distributed onto two plates and the rest of the medium was centrifuged and mixed with 100 µl of fresh LB and cultured onto the third plate. The plates were incubated in 37 ^{o}C overnight and checked for colonies the following day. Colonies were picked and used for plasmid purification using the miniprep plasmid purification kit (Qiagen) according to the manufactures protocol. LB

medium was made using 1.0 % Bactotryptone (BD biosciences), 0.5% yeast extract (BD biosciences) and 1.0% sodium shloride (Fluka). For agar plates, 1.5% agar was added to the medium. In case of ampicillin plates 200 µg/ml ampicillin (Sigma) was supplemented with the medium.

To acquire higher yields for transfection purposes, endofree maxiprep plasmid purification kit (Qiagen) was used to purify vectors according to protocol.

II.2. Utilized constructs

II.2.1. Retroviral moloney murine leukemia virus (MMLV) based constructs

Vectors containing retroviral backbones were used to transfer the gene of interest into eukaryotic cells. Specific sequences in retroviruses enable them to infect and subsequently package the viral genome. All retroviruses carry the basic *Gag*, *Pol* and *Env* Genes and possess a long terminal repeat (LTR) and a packaging signal (Psi). *Gag*, *Pol* and *Env* are eliminated from retroviral vectors to prevent replication competent viral particles. Usually these genes are located on a separate plasmid for concurrent transfection. Alternatively many different cell lines are available which already contain these genes making them more feasible to use. PLAT-E and Phoenix are two of these cell types, used for Moloney Murine Leukemia Virus (MMLV) based retroviruses. These cell lines are ecotropic (infect a narrow range of cell types) and are basically used to transude murine cells. Upon integration, the 5' LTR region enables expression of introduced genes by acting both as a promoter and enhancer. This region also has the necessary sequences for initiation and termination signalling. The 3'LTR which is very similar to the 5'LTR in sequence acts as the termination and polyadenylation site. The integration of these

plasmids gives them advantage over other plasmids in that they are available and expressed integrated in the cell genome for long periods and without the need of selection (Cepko and Pear, 2001; Coffin et al., 1989; Morita et al., 2000).

pMYC-IRES-*EGFP* was used as the base vector for cloning different genes required for transduction (Figure II.1).

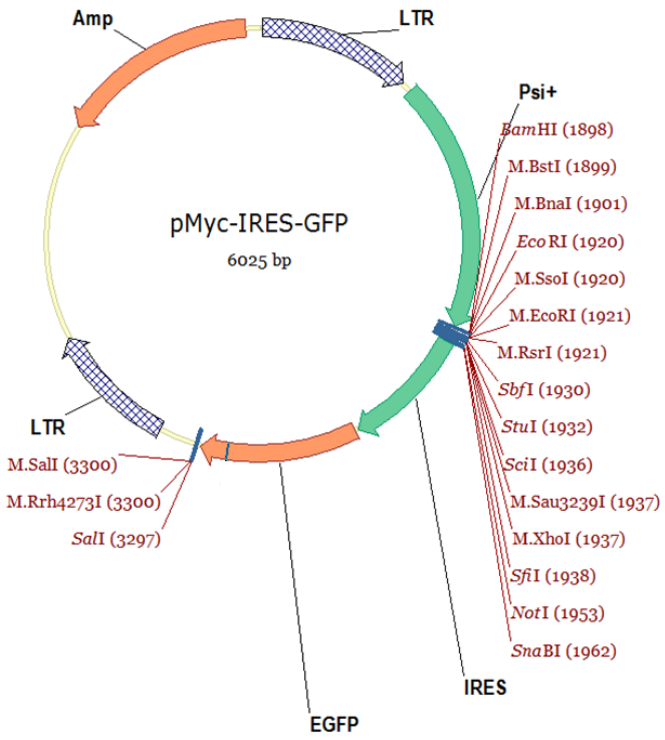


Figure II.1. pMyc-IRES-EGFP vector. The required genes where cloned replacing the IRES-EGFP region. The LTR regions and the packaging signal sequences are marked in the figure.

pMyc plasmids carrying the Yamanaka factors (4F) namely pMyc-IRES-*SOX2*, pMyc-IRES-*OCT4*, pMyc-IRES-*KLF4*, pMyc-IRES-c-*MYC* (Figure II.2) and pMyc-AA-*HMGA2* (Figure II.3) were kindly provided by Klaus Karjalainen lab.

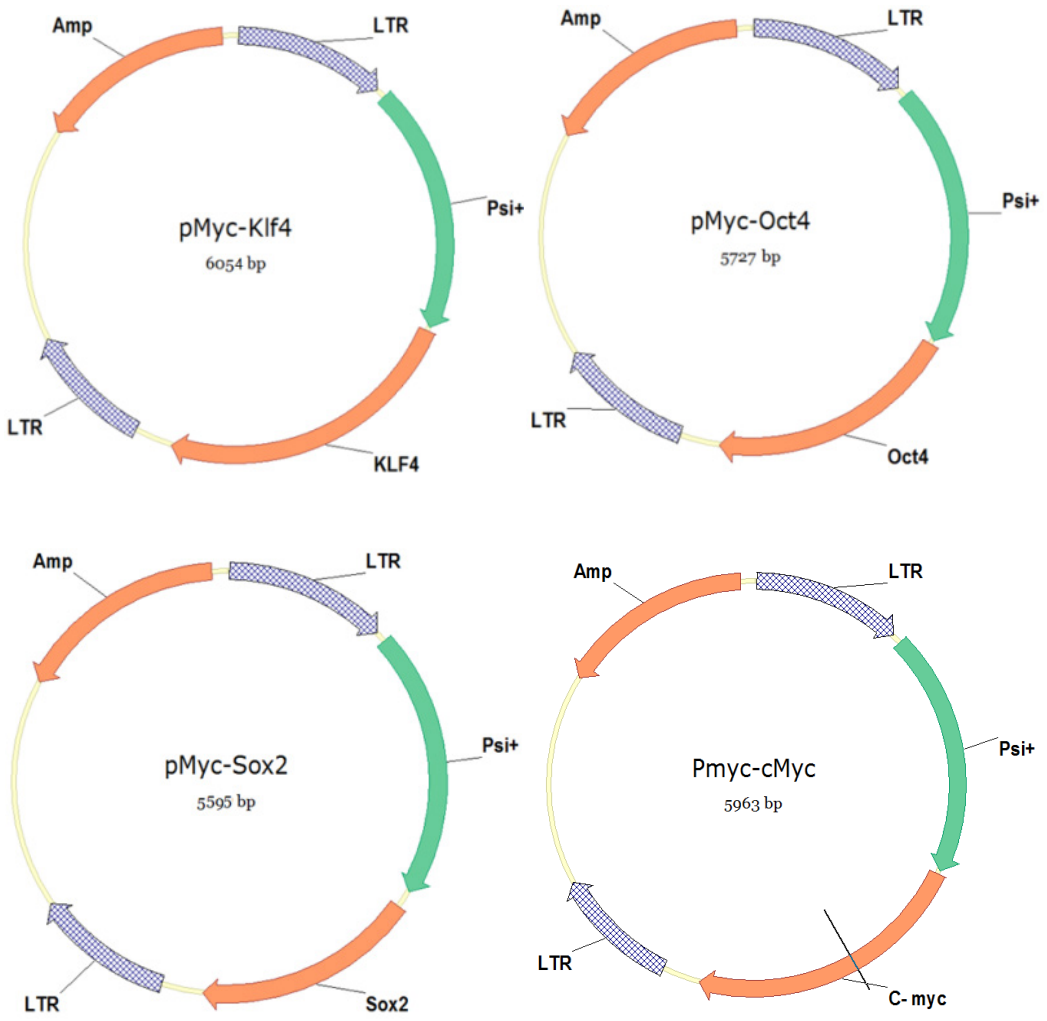


Figure II.2. Genes cloned into the pMYC vector: *OCT4*, *SOX2*, *KLF4* and *c-MYC (4F)* were cloned into pMyc based retroviral vector.

The pMyc-*HMGA2-AA-EGFP* carries a 2A self-cleavage sequence enabling the EGFP to be excised from HMGA2. This allows concurrent expression of both genes from one promoter, hence providing a reporter for exogenous HMGA2 expression (Figure II.3).

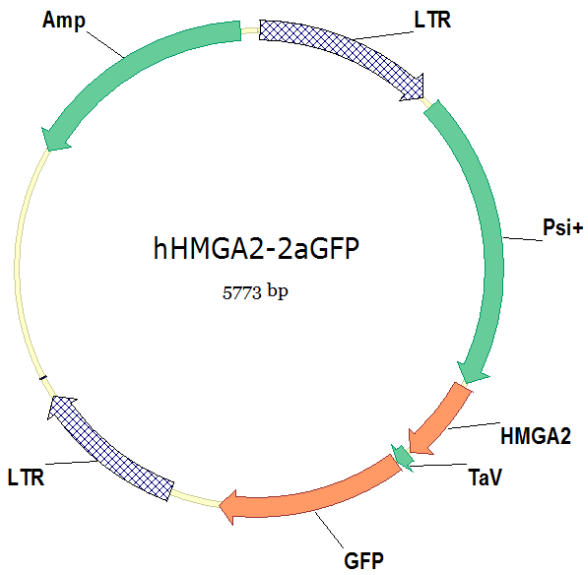


Figure II.3. *HMGA2-AA-EGFP* vector carrying a 2A (TaV) self cleavage sequence enabling concurrent expression of HMGA2 and EGFP.

II.2.2. UTF1 reporter construct

We used a construct containing the UTF1 regulatory elements with EGFP replacing the human *UTF1*. In this construct *UTF1* promoter drives the EGFP which is followed by a poly A signal. OCT4/SOX2 enhancer and a putative NANOG site exist in this construct (Figure II.4).

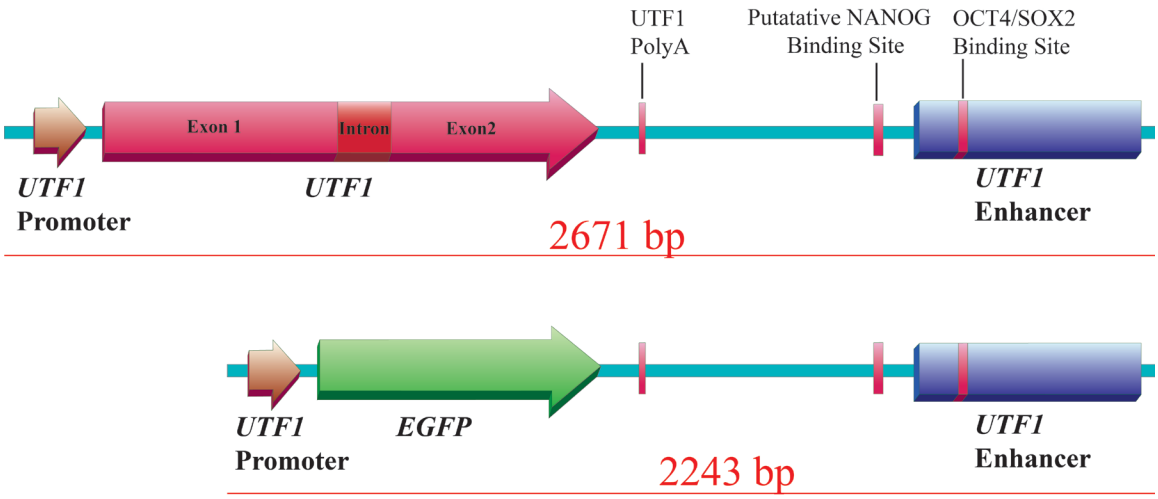


Figure II.4. UE reporter and the generation of UE positive iPSC colonies. Schematic figure showing UTF1 regulatory elements and the UE reporter construct. UTF1 consists of a TATAAT-less promoter, an intron, two exons and a 3' enhancer. A potential NANOG binding site consisting of a conserved octamer sequence is present with an established OCT4/SOX2 binding site following it. The EGFP sequence is replaced by EGFP to show UTF1 activation. The size of the entire UE cassette (2243 bp) is marked below the diagram.

II.2.2.1. pTZ-UTF1-EGFP transient vector

To study ESC and iPSC pluripotency, we utilized a transient expression vector named pTZ-UTF1-EGFP (Figure II.5). This vector carries the UTF1 reporter construct described earlier. The vector was previously constructed by Tan Shen Mynn in our lab.

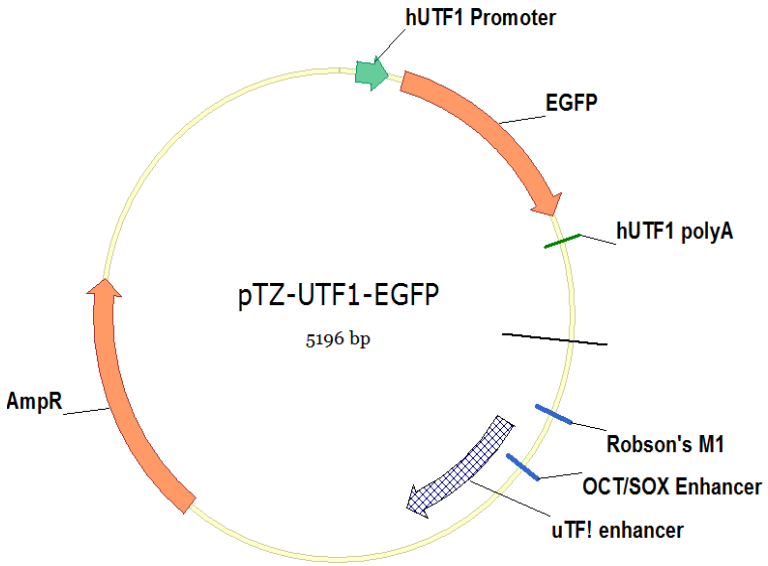


Figure II.5. pTZ-UTF1-EGFP Vector. This vector was used to transiently deliver the UTF1 construct into the cell.

II.2.2.2. pMYC-UTF1-EGFP retroviral vector

To generate the pMYC-UTF1-EGFP vector, the pTZ-UTF1-EGFP vector was digested wit *EcoRI* and *SalI* restriction enzymes (New England Biolabs). Subsequently, the cut fragment containing the *UTF1* regulatory regions and *EGFP* was cloned into the pMYC-EGFP vector replacing the IRES-EGFP fragment cut by the same enzymes. Details are shown in figure II.6.

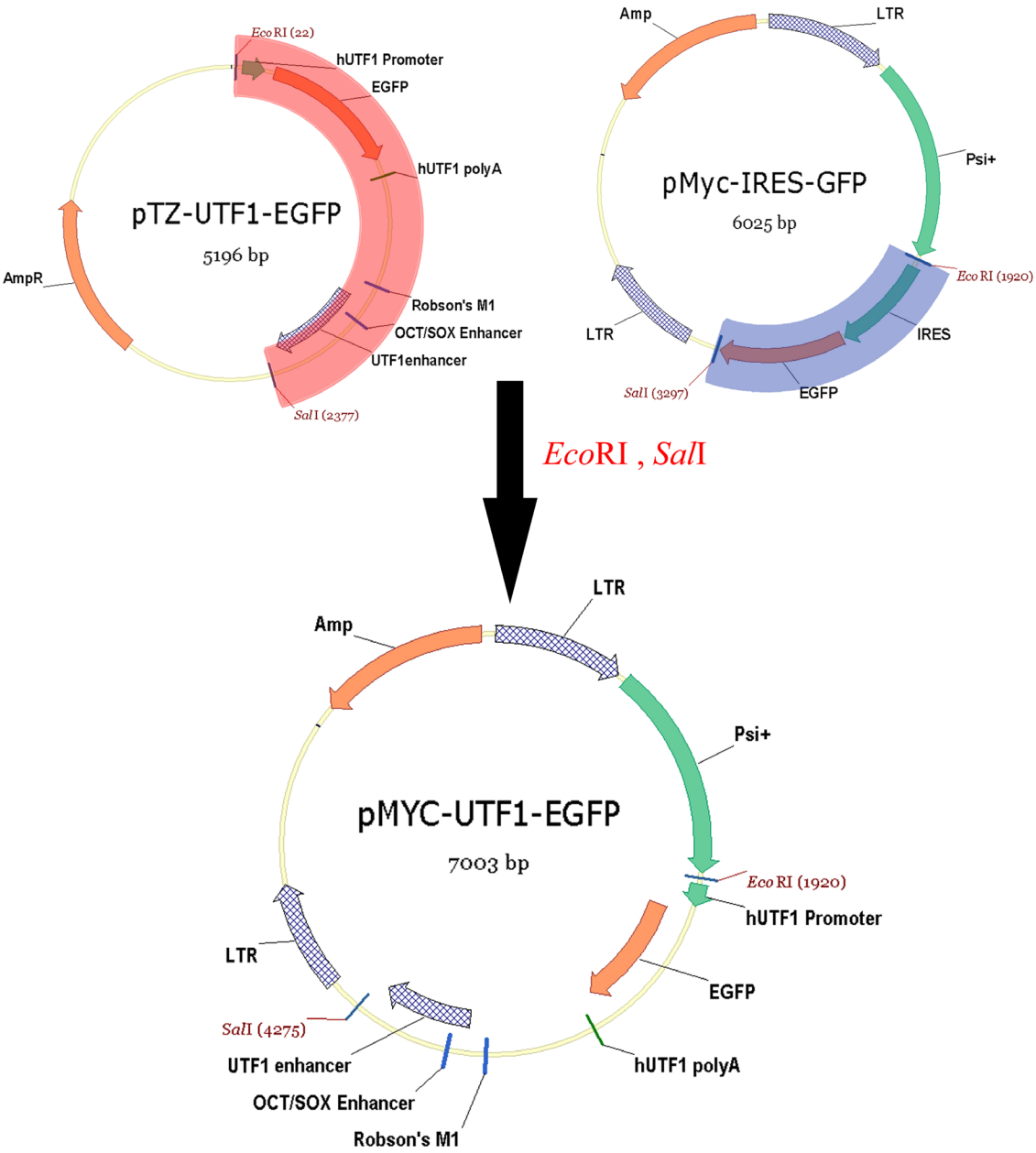


Figure II.6. Cloning of pMYC-UTF1-EGFP. The IRES-EGFP (blue highlight) is replaced by the UTF1-*EGFP*-enhancer (Red Highlight). *Eco*RI and *Sal*I are the restriction enzymes used for digesting both vectors. The final vector is a retroviral vector which contains *EGFP* with the *UTF1* regulatory elements.

II.2.3. UTF1-Tomato knockin construct

In order to produce knockin mouse carrying a reporter for UTF1were used. The *Utf1* knockin reporter mouse was generated by Dr. Yu Wei-Ping from the Biological Resource

Centre (BRC) in Singapore. All the animal procedures were performed in the A*STAR BRC by Dr. Yu and his colleagues according to the Institutional Animal Care and Use Committee guidelines (IACUC).

II.3. Reprogramming mouse embryonic fibroblasts to iPSC

II.3.1. Mouse embryonic fibroblast feeder cell preparation and inactivation

We used Mouse Embryonic Fibroblasts (MEF) as supporting and feeder tissue for culturing ESCs and iPSCs. MEF cells were cultured until passage 6 and subsequently mitototically inactivated to prevent them from affecting cell growth. For inactivation Gamma irradiation using Biobeam 8000(Germany) with 3000 rad (30 Grey) was used. This was advantageous over mytomycin C inactivation which is laborious and time consuming. After irradiation cells were cultured into intended culture plates and used for ESC or iPSC culture. Excess inactivated MEF cells were frozen down for future experiments.

The MEF cells were cultured in routine Dulbecco's modified Eagle’s medium (DMEM) supplemented with fetal bovine serum (FBS), glutamine, 2-mercaptoethenol and penicillin\Streptomycin (all reagents from Gibco). The following table shows all the components and their quantities (Table II.1).

.

Table II.1: Components of Culture Medium for Cell Culture (MEF and Plat-E Cell Lines)

Component	Volume in 500ml
High-glucose DMEM without sodium pyruvate, without L-glutamine, 1X (GIBCO)	439 ml
FBS (GIBCO)	50 ml (10%)
Penicillin/Streptomycin (GIBCO) 10,000 unit/mL penicillin 10,000µg/mL streptomycin	5 ml
L-glutamine, 100X, 200mM (GIBCO)	5 ml
2-Mercaptoethanol (GIBCO)	1ml
Abbreviations: DMEM, Dulbecco's modified Eagle's medium FBS, Fetal Bovine Serum	

II.3.2. Transfection of Plat-E packaging cell line

To generate viruses capable of transduction, we used PLAT-E as the intermediate cell line used to generate retroviral particles. This packaging cell line will generate envelope packaged retroviruses which are able to infect a secondary cell line. PLAT-E is suitable for murine cells and carries the necessary genes needed for retroviral packaging including *Gag*, *Pol* and *Env*. In this cell line only coding sequences of *Gag*, *Pol* and *Env* are used which are positioned on distinct plasmids, making it virtually impossible for recombinant replication-competent retroviruses to arise. These genes are driven by the strong EF1-α promoter enabling higher expression and increased viral yields (Morita et al., 2000). The medium used for culturing PLAT-E cells was similar to MEF medium containing DMEM with 10% FBS and other supplements mentioned in Table II.1.

For transfection of PLAT-E cells, we used the routine lipofectamine 2000 (Invitrogen) transfection method according to the manufactures protocol. When cells reached around 70% confluency they were used for transfection with the appropriate vector. After 12 hours the medium was replaced with fresh medium. The supernatant containing the viral particles was collected at 24 and 48 hours, to be used for transduction of MEF cells. This supernatant was placed in 4^{o}C and used later for transduction.

II.3.3. Transduction of MEF cells using viral supernatant

Viral supernatant from the last step is subsequently filtered through 0.45 μm pore sized filter to remove debris and leftover cells from the medium. To increase transduction efficiencies polybrene (Millipore) is added to the solution with a concentration of 4μg/ml. This compound neutralizes charges between the virion and the cell surface, resulting in decreased repulsion and facilitating adherence and viral entry.

Plates are coated with 10μg/ml fibronectin (r-fibronectin CH-296 from TAKARA) diluted in phosphate buffered saline (PBS) to increase the chance of virus-cell interaction. Virus particles attach to fibronectin which is glycoprotein originating from the extracellular cell matrix. Once fibronectin coated plates are ready they are initially centrifuged with the viral supernatant (no cells) for 60 min at 4000 rpm (4^{o}C).

After centrifugation the supernatant is removed and fresh viral containing supernatant mixed with 1x10^{5} MEF cells is added to each well of a 6 well plate and centrifuged at 1900 RPM for 50 min in 37 ^{o}C. It has previously been shown that centrifugation of the cells with virus particles increases efficiencies; however the mechanism is still unknown. Subsequently the cells are incubated with the virus in 37 ^{o}C for 6-8 hours. In the next step

cells are incubated overnight with fresh medium and the same process can be repeated several times to increase efficiencies. Different steps of transduction are shown in figure II.7.

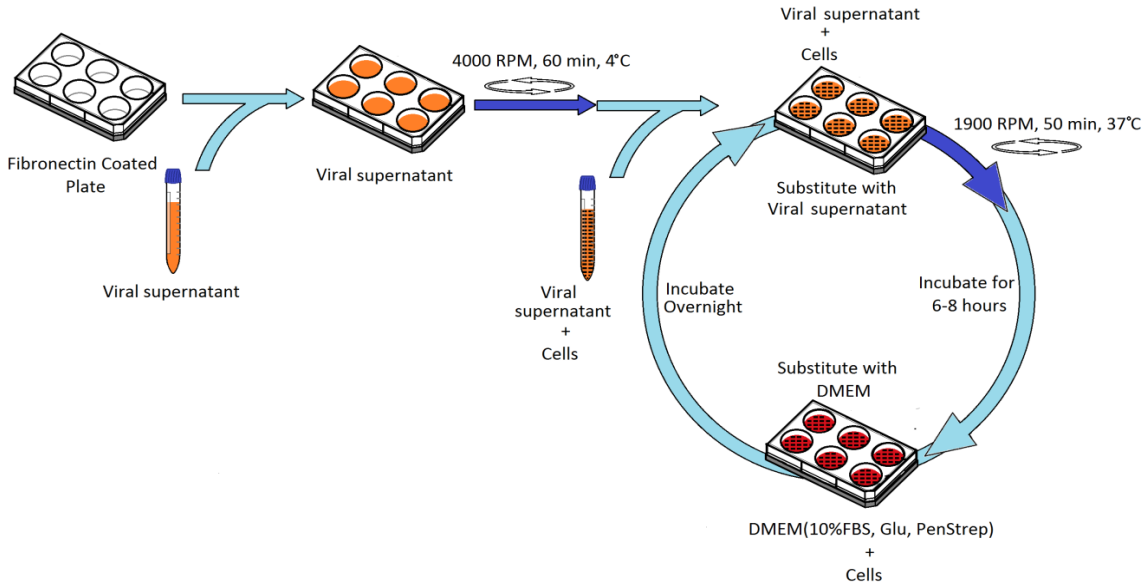


Figure II.7. Stages of Transducing MEF Cells**.** Viral supernatant (without cells) is added to fibronectin coated plate and centrifuged for 60 min at 4000 rpm at 4°C. Viral supernatant is then removed and replaced with fresh viral supernatant with $1x10^5$ cells. Subsequently, the plate is centrifuged at 1900 rpm for 50 min at 37°C and left in incubator for 6-8 hours for transduction. After incubation period, the viral supernatant is replaced with fresh medium (DMEM with supplements) and incubated. The same procedure can be repeated the next day for increased transduction efficiencies.

II.3.3.1. Reprogramming conditions

Specific culture conditions were used to achieve pluripotency and reprogram MEF cells. After transduction on the first and second day cells are left in general MEF culture medium (table II.1) until day 5. At day 5 the cells are trypsinized and moved into a new plate covered with around $2x10^4$ (for 6-well) or $2.4x10^5$ (for 10cm dish) inactivated MEF cells. The same medium is used until day 7 and subsequently the medium is substituted with mouse ESC medium (Table II.2) until iPSC colonies appear.

Table II.2. Components of Mouse ES and iPS Cell Medium

Component	Volume in 500ml
High-glucose DMEM without sodium pyruvate, without L-glutamine, 1X (GIBCO)	404 ml
ES Cell Qualified FBS (Chemicon)	75 ml (15%)
penicillin/streptomycin (GIBCO) 10,000 unit/mL penicillin 10,000µg/mL streptomycin	5 ml
L-glutamine, 200mM, 100X (GIBCO)	5 ml
Sodium Pyruvate, 100mM, 100X (GIBCO)	5 ml
MEM NEA, 100X (GIBCO)	5 ml
LIF, 10^3 U/ml (Chemicon)	1 ml
2-Mercaptoethanol (GIBCO)	1ml
Abbreviation: DMEM, Dulbecco's modified Eagle medium FBS, fetal bovine serum NEA, non-essential amino acids LIF, leukemia inhibitory factor	

Around ten days after stem cell conditions small iPSC colonies start to appear. Colonies can be picked several days after they are visible and passage to create a single cell line. Chart shown in figure II.8 depicts an overview of different steps carried out for reprogramming.

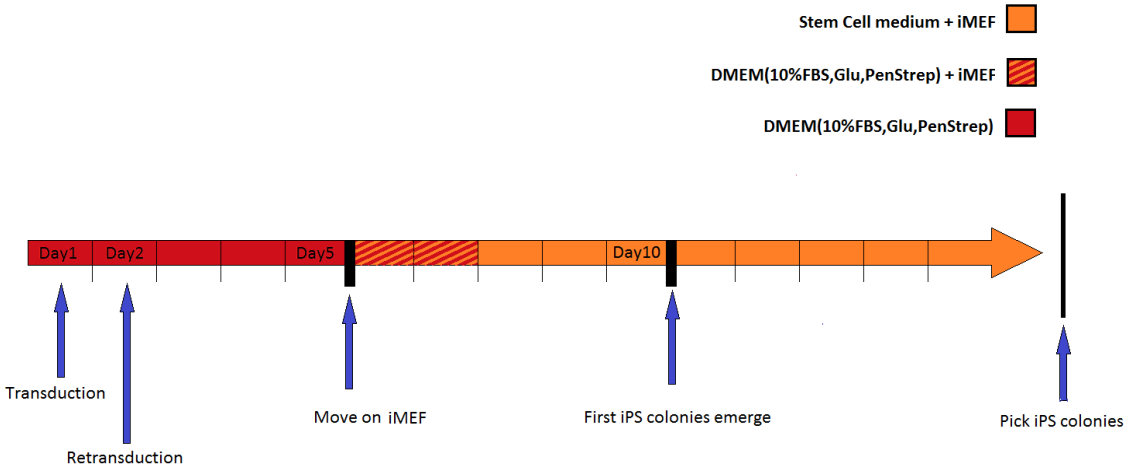


Figure II.8. Chart showing the overall procedure used for reprogramming. MEFs are transduced on day 1 and day2. The cells are incubated for five days in DMEM (with supplements) till day 5 and then moved to MEF (red colour). After two days on MEF the medium is changed to ESC medium (orange colour). Colonies start to appear around 7-10 days later. Subsequently, they can be picked several days after emergence. Important steps are specified with blue arrows.

II.4. Introduction of constructs into iPSCs

II.4.1. Transient transfection of pTZ-UTF1-EGFP

In order to transfect iPSC with the pTZ-UTF1-EGF vector, TransLT1 transfection reagent (Mirus Bio) was used according to provided protocol. For 6 well plates 2.5µg of DNA was diluted in optimum medium. Subsequently 7.5 µl transLT1 transfection reagent was added to form complexes, which was then poured on to the Cells. MEFs initially transduced with four factors were transfected twelve, fourteen and sixteen days after transfection and observed 48 hours later for EGFP expression.

For transfecting previously picked and established iPSCs, cells were cultured to 2–6×10^5 cells per well and subsequently transfected with the same transfection reagent.

II.4.2. Transduction of pMYC-UE simultaneously with 4F

pMyc-UE was tranduced concurrently with the 4F in order to induce iPSCs carrying the UE construct. This enables constant availability of the UE construct during iPSC induction due to its integration.

II.4.3. Transduction of pMYC-AA-HMGA2 simultaneously with 4F

PMyc-AA-HMGA2 was tranduced concurrently with the 4F in order to express HMGA2 stably during iPSC generation. This vector enabled us to monitor expression of the exogenous human HMGA2 after transduction of mouse embryonic fibroblasts (MEFs). A second plate was transduced with a similar EGFP vector lacking the HMGA2 coding region as control.

II.5. iPSC analysis

II.5.1. Microscopy analysis

Cells were observed by light microscopy or fluorescent microscopy using IX71 inverted microscope (Olympus).

II.5.2. Alkaline Phosphates Staining

Detection of alkaline phosphatise activity was assayed using the AP detection kit (Millipore) according to providers protocol. ESCs or iPSCs were primarily cultured for at least five days before AP staining was carried out. Cells were fixed with 4% paraformaldehyde prior to performing the staining. Reagents for AP staining including

Fast Red Violet, naphthol and water were mixed with a ratio of 2:1:1 respectively and poured onto fixed cells and incubated in dark for 15 min at room temperature. In the next step the reagent is washed away and PBS is used to rinse the cells. Finally cells were analyzed under the microscope for AP activity. AP positive colonies stain red and the intensity of the red colour correlates with the AP activity of the colonies. Hence high AP colonies stain dark red, while low AP colonies will manifest a pinkish colour.

II.5.3. Gene expression analysis

II.5.3.1. RNA extraction

Extraction of RNA was performed by the RNAeasy Kit (Qiagen) according to provider's protocol. Around $5x10^6$ cells were used for each RNA extraction. Subsequently RNA was treated with DNase I (DNase I recombinant, RNase-free from Roche) to remove any leftover DNA. For quality of RNA samples the OD 260/280 was measured and checked to be between 1.6-1.8. To measure concentration OD 260 was used for each sample.

II.5.3.2. Real Time PCR (qRT-PCR)

For real time PCR analysis of iPSCs carrying the UTF1-EGFP reporter, we used the SAbiosciences PCR array kit for mouse ESCs. This kit includes 84 pluripotency linked gene primers and various controls (96 in total) for real time PCR analysis. Table II.3 indicates the primer and the location of well in the PCR array kit.

Table II.3. Gene Symbol and Location in 96-well Plate Used for Real Time PCR

	01	02	03	04	05	06	07	08	09	10	11	12
A	*Afp*	*Bxdc2*	*Cd34*	*Cd9*	*Cdh5*	*Cdx2*	*Col1a1*	*Commd3*	*Crabp2*	*Ddx4*	*Des*	*Diap2*
B	*Dnmt3b*	*Ednrb*	*Eomes*	*Fgf4*	*Fgf5*	*Flt1*	*Fn1*	*Foxa2*	*Foxd3*	*Gabrb3*	*Gal*	*Gata4*
C	*Gata6*	*Gbx2*	*Gcg*	*Gcm1*	*Gdf3*	*Grb7*	*Hba*-x	*Hbb*-y	*Hck*	*Iapp*	*Ifitm1*	*Ifitm2*
D	*Igf2bp2*	*Il6st*	*Ins2*	*Kit*	*Krt1*	*Lama1*	*Lamb1*	*Lamc1*	*Lefty1*	*Lefty2*	*Lifr*	*Lin28*
E	*Myf5*	*Myod1*	*Nanog*	*Nes*	*Neurod1*	*Nodal*	*Nog*	*Nr5a2*	*Nr6a1*	*Numb*	*Olig2*	*Pax4*
F	*Pax6*	*Pdx1*	*Pecam1*	*Podxl*	*Pou5f1*	*Pten*	*Ptf1a*	*Rest*	*Runx2*	*Sema3a*	*Serpina1a*	*Sfrp2*
G	*Sox17*	*Sox2*	*Sst*	*Sycp3*	*T*	*Tat*	*Tcfcp2l1*	*Tdgf1*	*Tert*	*Utf1*	*Wt1*	*Zfp42*
H	*Gusb*	*Hprt1*	*Hsp90ab1*	*Gapdh*	*Actb*	MGDC	RTC	RTC	RTC	PPC	PPC	PPC

The selected genes are a set of pluripotency related genes which contain both pluripotency markers and also differentiation markers. The genes are classified in the following categories:

Embryonic Stem Cell-Specific Genes:

- Transcription Factors Maintaining "Stem-ness": Foxd3, Gata6, Gbx2, Nanog, Nr5a2, Nr6a1, Pou5f1, Sox2, Tcfcp2l1, Utf1, Zfp42.
- Signaling Molecules Required for Pluripotency and Self-Renewal: Commd3, Crabp2, Ednrb, Fgf4, Fgf5, Gabrb3, Gal, Grb7, Hck, Ifitm1, Il6st, Kit, Lefty1, Lefty2, Lifr, Nodal, Nog, Numb, Pten, Sfrp2, Tdgf1.
- Cytokines and Growth Factors: Fgf4, Fgf5, Gdf3, Lefty1, Lefty2, Nodal, Tdgf1.
- Other Embryonic Stem Cell-Specific Genes: Brix1, Cd9, Diap2, Dnmt3b, Ifitm2, Igf2bp2, Lin28a, Podxl, Rest, Sema3a, Tert.

Selected Embryonic Stem Cell Differentiation / Lineage Markers:

- Extra-Embryonic Endoderm Markers: Foxa2, Gata4, Ptf1a.
- Trophoblast Markers: Cdx2, Eomes, Gcm1, Krt1.
- Viseral Endoderm Markers: Afp, Serpina1a.
- Parietal Endoderm Markers: Fn1, Lama1, Lamb1, Lamc1, Sox17.
- Mesoderm Markers: T, Wt1.
- Muscle: Des, Myf5, Myod1.
- Blood: Hba-x, Hbb-y.
- Bone: Col1a1, Runx2.
- Neural: Nes, Neurod1, Pax6.
- Endothelial: Cd34, Cdh5, Flt1, Pecam1.
- Germ Cell: Ddx4, Sycp3.
- Pancreas: Gcg, Iapp, Ins2, Pax4, Pdx1, Sst.
- Other Embryonic Stem Cell-Differentiation / Lineage Markers: Olig2, Tat.

Various controls are available which are used for different indications. Wells H1 to H5 contain different housekeeping gene controls (HKGs). Well H6 carries primer sets as genomic DNA controls (GDC) that accurately detect non-transcribed genomic DNA contamination with a high degree of sensitivity and specificity. Wells H7 and H8 contain controls for reverse transcription efficiency. H10 to H12 posses positive PCR controls which tests the PCR efficiency itself by using a pre-dispensed artificial DNA sequence and its specific primer sets. These controls are present for reproducible and trustworthy results. Acquired Ct values from the Real time PCR analysis and the values for the 12 controls are inserted into a readymade excel data analysis worksheet provided by the supplier which automatically calculates and determines all the data and returns the normalized Ct values, delta Ct, DNA contamination and transcription efficiency. This

premade data analysis worksheet can be accessed at (http://www.sabiosciences.com/pcrarraydataanalysis.php).

1µg of RNA was reverse transcribed to cDNA by SABiosciences first strand kit according to manufactures protocol. Equal amounts of RNA from all samples were used for cDNA synthesis. The cDNA was subsequently used for real time PCR analysis.

cDNA was added to a mastermix which contains nucleotides, polymerase and SYBR Green/ROX fluorescent dye. 25 µl of the mastermix is distributed to each well of the 96 well of the PCR plate already containing the specific primers for each gene. In order to remove any bubbles formed due to pipeting, the plate is centrifuged at 1000g for 1 min at room temperature and then loaded into the real time PCR machine.

Applied Biosystems 7500 RealTime PCR system was used with the ABI 7500 sequence detection software (SDS) version 1.4 for analysis of samples. The PCR program used for PCR conditions is specified in Figure II.9.

STAGE1	**Polymerase Activation → 95°C for 10 min**	
STAGE2	Denaturation → 95°C for 15 Sec Annealing → 60°C for 1 min	40 Rounds
STAGE3	Dissociation step → 95°C for 15 Sec , 60°C for 1 min, 95°C for 15 Sec, 60°C for 15 Sec	

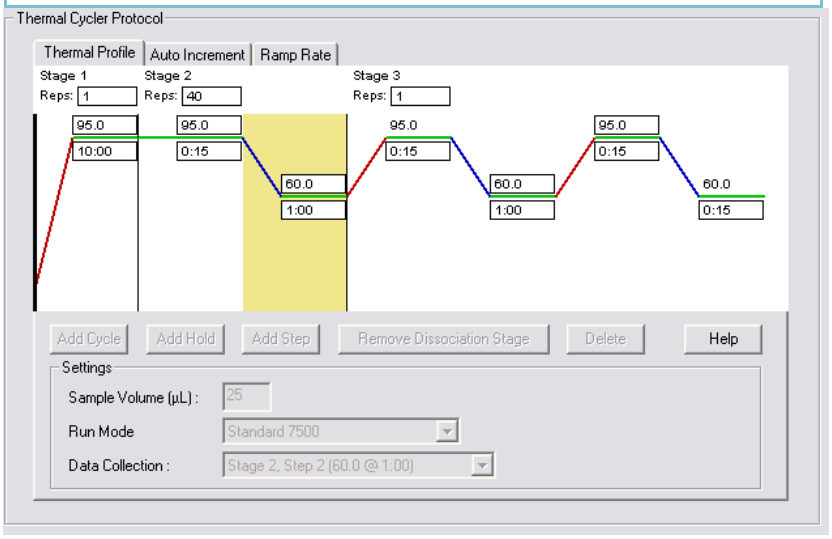


Figure II.9. Conditions Used to Perform Real-Time PCR Analysis in each stage.

Provided primers in the kit are verified primers with uniform annealing temperatures and outstanding amplification efficiencies, avoiding the need to carry out extra analysis for primer verification.

Threshold line is the border where the reaction reaches a fluorescent intensity. The point where the curve crosses the border is indicated as Ct. The threshold line is set in the exponential phase of the amplification curve and the baseline is set between the second cycle and one cycle before the amplification curve rises. The Ct value for each gene is compared to the control for the same gene to determine up or down regulation.

The comparative Ct method was used to analyze the data. These data can be automatically calculated by the Microsoft excel template from SABiosciences which is provided by the manufacturer of the kit. This template calculates the different parameters like ΔCt and $2^{\Delta Ct}$ automatically and constructs the different graphs including a scatter plot which demonstrates up regulation and down regulation of the genes.

Real time analysis was performed three times for each sample for confirming data. The excel datasheet calculates the Ct values according to formula shown in table II.4.

Table II.4. Definition and formula used to calculate the different parameters in the comparative Ct method

Parameter	Definition
Ct	Cycle threshold
ΔCt = Ct(test) – Avg Ct(Control)	Difference between each Ct and average control Ct (for normalizing data)
$2^{-\Delta Ct}$(test)/ $2^{-\Delta Ct}$(control)	Fold difference between test and control
$2^{-(\Delta\Delta Ct)} = 2^{\Delta Ct(test) - \Delta Ct(control)}$	Fold up- or down-regulation(change in gene expression) between test and control

II.5.3.3. Reverse Transcriptase (RT)-PCR analysis

RNA purified with the "RNeasy RNA Purification Kit" (Qiagen) was converted into cDNA using the Superscript First Strand Synthesis kit (Invitrogen). Eaqual quantities of RNA were turned into cDNA for analysing the three germ layer marker expression. The markers include AFP as an endoderm specific marker, brachyury for marking mesoderm and nestin as a marker for ectoderm. PCR was carried out with the cDNA using taq polymerase and specific primers for the mentioned markers with the settings as follows:

Stage	Temperature(^{o}C)	Time	
Initial Denaturation	95	5 min	
Denaturation	95	30 sec	29 Cycles
Annealing	55	30 sec	
Extension	72	45 sec	
Final extension	72	10 min	

Primers used are as follows from 5' to 3':

Afp forward: TGC AGA AAC ACA TCG AGG AGA G, *Afp* reverse: GCT TCA CCA GGT TAA TGA GAA GCT, *T*(Brachyury) forward: ATA ACG CCA GCC CAC CTA CT, *T*(Brachyury) reverse: R: GTC TCA GCA CAT GGA GGA AAG TT, *Nes* (Nestin) forward: AAC TGG CAC ACC TCA AGA TGT, *Nes* (Nestin) reverse: TCA AGG GTA TTA GGC AAG GGG, and *Hprt* forward: GCT GGT GAA AAG GAC CTC T, *Hprt* reverse: GAC AGG ACT AGA ACA CCT GC.

In the final step equal amounts of PCR samples were run on 1.5% agarose gel and visualized with etidium bromide using the Gel Doc from BioRad.

II.5.4. Flow Cytometry (FACS) analysis

In order to analyze fluorescence activity, cells were initially washed with PBS, trypisinized and centrifuged for 5 min 300g. Cells were washed once again with PBS and centrifuged and rediluted in 1 ml of PBS containing 2% FBS. Cell suspension was transferred to appropriate FACS tubes and analyzed with FACS Calibur Flow Cytometer (Becton Dickson) machine and the CELLquest software (Becton Dickson).

For staining cells, after the second wash step cells were diluted in 100 µl PBS with anti SSEA1 antibody (clone MC-480 PE conjugate, Millipore) with a dilution of 1:50 and

incubated in 4° C for 60 min. Subsequently cells were washed twice with PBS and diluted in 1 ml of PBS for analysis.

Dot plot of side scatter (SSC) versus forward scatter (FSC) was used to initially gate live cells as R1. Cells were then analyzed on another dot plot containing FL1 or FL2 vs. SCC (or FSC), looking at R1 gated cells. All experiments were done using a negative sample as control to set the graphs. FL1 was used to monitor EGFP expressing cells, while FL2 was utilised for SSEA1 (Red) expression. A histogram graph was used to monitor intensity of the fluorescence vs FL1 or FL2.

II.5.5. Microarray analysis

Microarray expression assay was performed by taking advantage of the Agilent 8X60 Mouse Whole Genome Expression Microarray System (Agilent Technologies). Agilent 2010 bioanalyzer platform (Agilent Technologies) was used to check quality of RNA samples. The RNA integrity number (RIN) which is a value showing RNA quality was measured for all samples which indicated a RIN of 10 for all samples, showing optimal RNA quality. In order to label and amplify the cRNA with Cy3 and Cy5, Agilent Low input quick amp labelling kit (Agilent technologies) was used according to manufactures protocol. cRNA yields and the amount of dye incorporation rate was measured using the ND-100 spectrophotometer (NanoDrop Technologies). Generally control samples are labelled with Cy3, while experimental samples are hybridized to Cy5. For hybridization 300ng of Cy3 and Cy5 labelled fragmented cRNA were combined and hybridized to Agilent Whole Mouse Genome Oligo Microarrays 8x60K in the appropriate hybridization chamber an oven. This was carried out overnight at 65°C using the Agilent

Gene Expression Hybridization Kit (Agilent Technologies). Detecting fluorescence signals was performed with Agilent DNA microarray scanner (Agilent technologies). To extract data the Agilent feature extraction software (FES) was used to generate and read out the image files. The software calculates and determines data by performing normalization, background subtraction, statistical analysis and removing outliers. Previous steps were carried out by Miltenyi Biotec. Output data from FES were used with the Rosetta Resolver gene expression data analysis system (Rosetta Biosoftware) and Genespring (Agilent Technologies) to further inspect data. Differentially expressed genes were selected with a cut off of 1.8. Common differentially expressed genes in all three were then used for GO analysis with a corrected p-value of 0.01 Benjamini-Hochberg methods.

II.5.6. Immunostaining

To perform OCT4 staining cells were cultured in 12 well plates. Cells were washed with PBS and fixed with 4% paraformaldhyde and then permeablized with 0.2% TX-100 in PBS for 10 min. Nonspecific sites were blocked with 1% BSA in PBS for 1 hour in room temperature. Immunostaining was carried out using primary OCT4 antibody (SC-5279, Santa Cruz) with a dilution of 1:500 at room temperature for one hour. Subsequently cells were washed 3 times for 10 min with PBS and incubated with the secondary antibody goat anti-mouse IgG-TR (SC-2871, Santa Cruz) for one hour. For DAPI staining cells were mounted in DAPI vectashield medium, washed as previously and observed under the microscope.

II.5.7. Western Blot analysis

Cells were trypsinzied, counted by the haematocytometer and centrifuged at 500g for 5 min. Cell pellet was resuspended with lysis buffer and incubated on ice for 10 min. The lysis buffer was made using 150 mM sodium chloride, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS (sodium dodecyl sulphate), 50 mM Tris, pH 8.0. Cells were then centrifuged for 10 min at 16000 g at 4C in order to remove debri. The remaining supernatant contain the protein extract was used for western blot. Equal amounts of $2x10^5$ cells were loaded onto 15 % SDS gels and separated at 200 V. To detect protein the gel was transferred onto a PVDF membrane (Millipore) overnight with 30 V in 4C. For blocking Tris-buffered saline with 0.05% Tween-20 and 5% non-fat dry milk (BioRad) was used. The Primary antibodies used were polyclonal Anti-HMGA2 antibody (ab41878, Abcam) and monoclonal anti β actin (A5449, SIGMA). Goat anti rabbit IgG-HRPsc (SC-2004, Santa Cruz) and Goat anti mouse IgG (ab6823, Abcam) were used as secondary antibody. Primary antibody was incubated overnight and secondary antibody was incubated for 1 hour. 3x10 min was used for washing in-between steps.

II.5.8. SUMOylase assay

To perform SUMOylase assay, lysate extracted from iPSCs was treated with SUMO protease (Invitrogen) according to protocol. Lysate containing 25 µg of protein was used for treatment with the SUMO protease. Lysate extracted from non-treated cells served as control. Samples were incubated in 4^oC for 3 hours for digestion and then loaded onto SDS gel for western blot analysis.

II.5.9. Differentiation of iPSCs

II.5.9.1. DMSO induced random differentiation

In order to differentiate cells randomly, differentiation medium containing 1% DMSO was added every day to cells until analysis was carried out.

II.5.9.2. Embryoid body formation

In order to sepearate MEF and iPSC confluent plates were trypsinized and moved to normal cell culture plates for 30 minutes. MEF cells adhere rapidly to the plate, enabling the removal of iPSC before attachment. The iPSCs which were collected, centrifuged and mixed with differentiation medium (20% FBS) and seeded onto ultra-low attachment plates without MEF and gelatine with a concentration of 2×10^6/dish. Cells were incubated for 6 to 12 days for Embryoid Body (EB) structures to appear while changing the medium every day.

II.5.9.3. Cardiomyocyte directed differentiation

In order to direct cells into cardiomyocytes single cell suspension of iPSCs was prepared at 1000cells/20µl (50.000 cells / ml) in differentiation medium containing 3 mg/ ml L-ascorbic acid 2-phosphate magnesium salt hydrate (Sigma). Subsequently 20 µl drops were put into the lid of a dish and the dish itself was filled with PBS. The lid is carefully put back onto the dish so the drops are in a hanging state. After 2 days small EBs will form. The EBs are collected and put into a gelatine coated plate containing differentiation

medium in order to form cardiomyocytes. It takes around 10 days for beating colonies to appear.

III. RESULTS

III.1. Efficiency of vector uptake

III.1.1. Transfection efficiency of packaging cell lines

Initially we required induced pluripotent stem cells (iPSCs), to use later for testing our UTF1 construct. In order to generate iPSCs, high viral yields are needed to efficiently transduce cells in the next step. Therefore, prior to iPSC generation, transfection efficiency of PLAT-E cells was verified using flowcytometry (FACS). This cell line already carries the required packaging genes, facilitating its use for viral production. PLAT-E cells transfected with pMYC-IRES-EGFP using lipofectamine 2000 were analysed for the amount of EGFP expressing cells. Microscopy observation showed around 80-85% of cells to be EGFP positive, verifying that adequate amounts of viral particles will be produced to be used in the next stage. The efficiency was confirmed by FACS analysis showing 78% of cells expressing EGFP (Figure III.1).

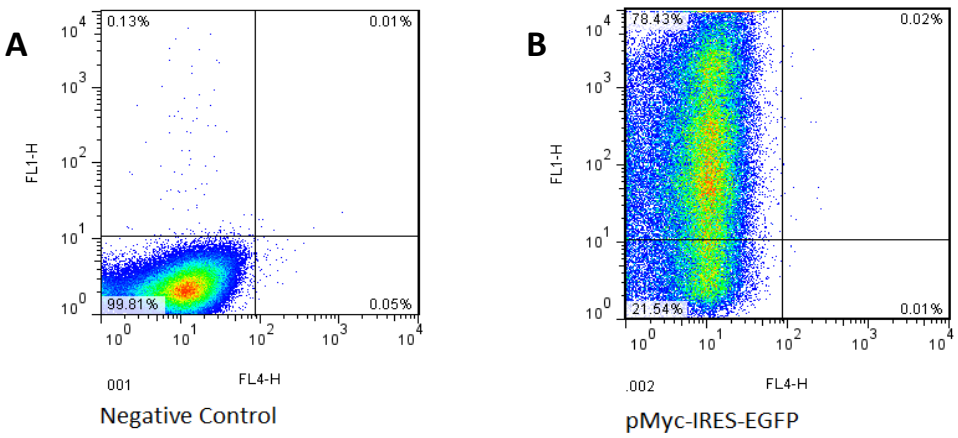


Figure III.1. FACS results showing transfection efficiency of Plat-E cells with pMYC-IRES-EGFP. (A) Cells transfected with pMYC-SOX2 as negative control (B) Transfection of Plat-E cells with pMYC-IRES-EGFP using lipofectamine 2000 showed 78% of the cells are EGFP positive.

III.1.2. Transduction efficiency of MEFs

To verify that viral particles are proficient and can efficiently transduce our MEF cell lines, we performed FACS analysis after transduction. Due to the fact that iPSC generation is an inefficient process (Okita et al., 2007), sufficient amount of cells need to be transduced in order to generate iPSCs. Viral particles obtained from ecotropic PLAT-E cells are capable of trasnducing mouse cell lines; however the efficiency varies for different cell types due to surface receptors. Hence we tested the acquired viral soup on our MEF cells. We also transduced the cells twice in different days for further increase in transduction efficiencies (day 1 and day2). After two days (day 4) more than 80% of the cells were expressing EGFP under microscope visualization. These cells were analyzed by FACS for EGFP expression, indicating 85% of the cells to be EGFP positive (Figure III.2). This convinced us that quite a number of cells will take up the four Yamanaka factors, increasing the possibility of iPSC generation.

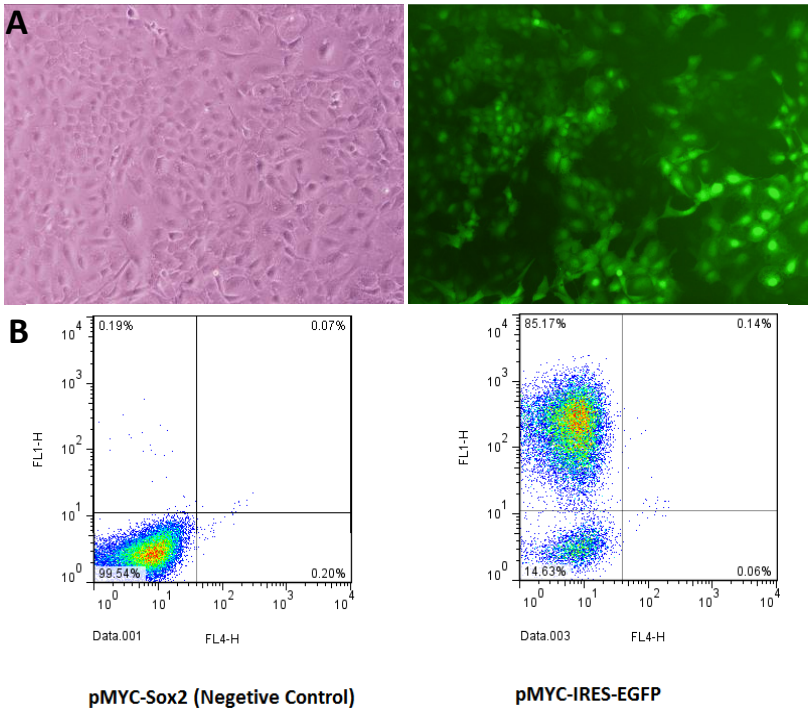


Figure III.2. (A) Light and fluorescence microscopy imaging showing transduced MEF cells 2 days after transduction. (B) FACS results showing transduction efficiency of MEFs with pMYC-IRES-EGFP in compared to pMYC-Sox2 as control. MEF cells were transduced twice and subjected to FACS analysis after second transduction.

III.2. Formation and characterization of iPSCs

III.2.1. Colony formation and morphology

Subsequently, to generate iPSCs, MEF cells transduced with *OCT4, SOX2, CMYC, KLF4* (4F) were put under stem cell conditions according to protocol and monitored for embryonic stem cell (ESC) like colonies. Mouse ESC colonies are generally round with sharp shining edges containing indistinguishable cells with large nucleolus and scant cytoplasm inside (Takahashi and Yamanaka, 2006). Colonies with various morphologies started to appear around two weeks after transduction (Figure III.3).

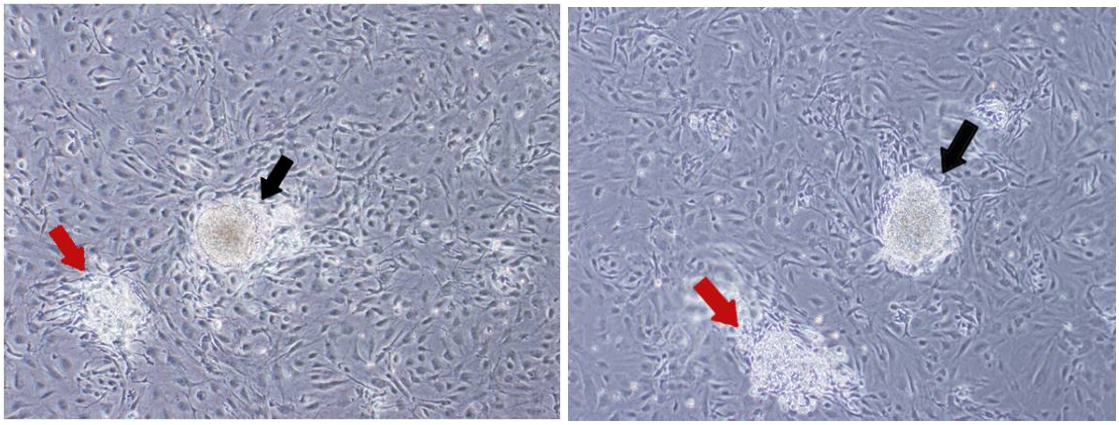


Figure III.3. Morphology of reprogrammed colonies around 15 days after transduction. Black arrows show iPSC colonies resembling ESCs, red arrows indicate colonies not resembling ESC colonies.

For further characterization, colonies closely resembling ESCs were picked and passaged to generate stable iPSC lines (Figure III.4). These established iPSC cell line was used for further experiments.

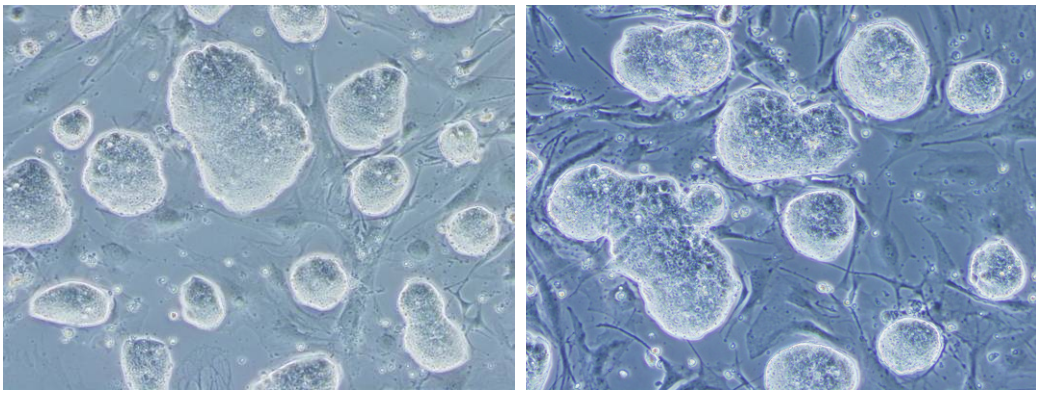


Figure III.4. Morphology of isolated iPSC Colonies after Several Passages. iPSC colonies were picked and passaged for several rounds to produce stable cell lines.

III.2.2. Alkaline Phosphatase staining

We moved into characterizing our iPSCs prior to other experiments. One of the primary steps in assaying pluripotency is alkaline phosphatase (AP) staining (Takahashi and Yamanaka, 2006). Alkaline phosphatase expressed on the surface of pluripotent cells is a

reliable method for identifying pluripotent cells. Hence, plates containing the newly generated iPSC colonies were stained for AP activity. Colonies with distinct morphologies were observed, which displayed differences in the intensity of AP expression. Some colonies not resembling ESCs were also AP positive; however the intensity was much lower than cells with ESC morphology (Figure III.5).

AP staining of previously established iPSC cell line was also performed. Cells showed significant AP positive staining with over 90% of the cells being highly AP positive.

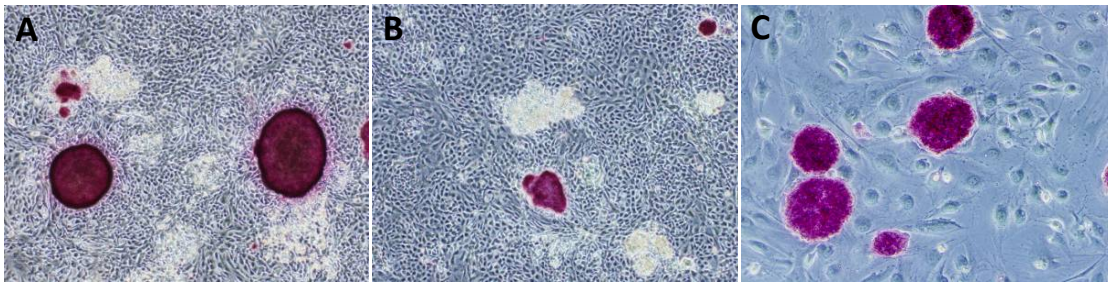


Figure III.5. (A)(B)AP positive iPS colonies in iPSC induction plates. Several days after iPSC colonies appeared they were stained for AP activity. Colonies with different morphologies are observed with AP activity. Non-ESC like colonies can be seen with lower AP staining intensity (C) Picked and passaged cells were stained for AP activity.

III.2.3. Stage-specific embryonic antigen 1 (SSEA1) expression

Stage-specific Embryonic Antigen 1 (SSEA1) is expressed on the surface of mouse ESCs. This receptor is absent in human cells, however human ESCs express a similar marker called SSEA4. To further characterize our cells we assayed our iPSCs for SSEA1 expression. Our established iPSC line was analysed for SSEA1 expression by FACS analysis and was compared to mouse ESCs as control. Our iPSCs showed to be SSEA1 positive with more than 70% of the cells exhibiting SSEA1 expression in compared to mouse ESC control showing around 50% SSEA1 expression (Figure III.6). Note that

mouse ESCs used was passaged for more than 15 rounds, which may explain why they expressed lower amounts of SSEA1.

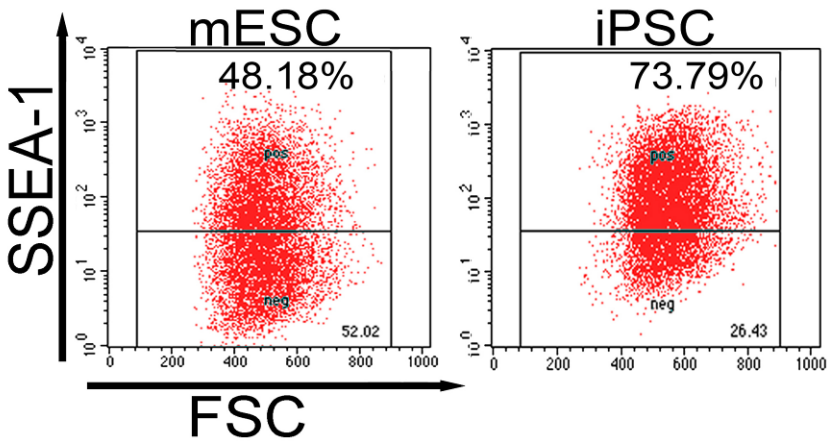


Figure III.6. SSEA1 expression in picked iPSCs in compared to mESC. Our iPSCs showed to be highly positive for SSEA1 expression.

III.2.4. Real time PCR analysis for Pluripotency marker genes

We looked into key stem cell state related genes (containing pluripotency genes and decisive differentiated state genes) to further characterize our cells and better evaluate the degree of their similarity to ESCs. Results from real time PCR of 84 different embryonic stem cell related genes, indicated that expression of well-known stem cell pluripotency markers like *Oct4, Sox2, Nanog, lin28, Nr5a2, Utf1* and *Zfp42* was very similar between mouse ESCs and our iPSCs and was placed in the 2 fold range (Figure III.7). Furthermore, expression of *Utf1* which is strictly expressed in pluripotent cells was similar with a small increase of expression in our iPSC line in compared to mESC cells. These results displayed that the established iPSC line was very similar to ESCs in gene expression.

Several genes were seen which were not in the 2 fold range, however these genes were not key pluripotency genes and their difference may be due to the quality of the ESCs and the conditions in which these cells were grown in.

On the other hand we compared our iPSCs to the primary MEF cells they were derived from. Huge variety in gene expression pattern was witnessed between our iPSCs and MEF cells with an obvious increase of pluripotency genes in our iPSCs (Figure III.7). This indicated that the iPSCs have moved into a different state in compared to MEF cells, displaying a completely distinct gene profile.

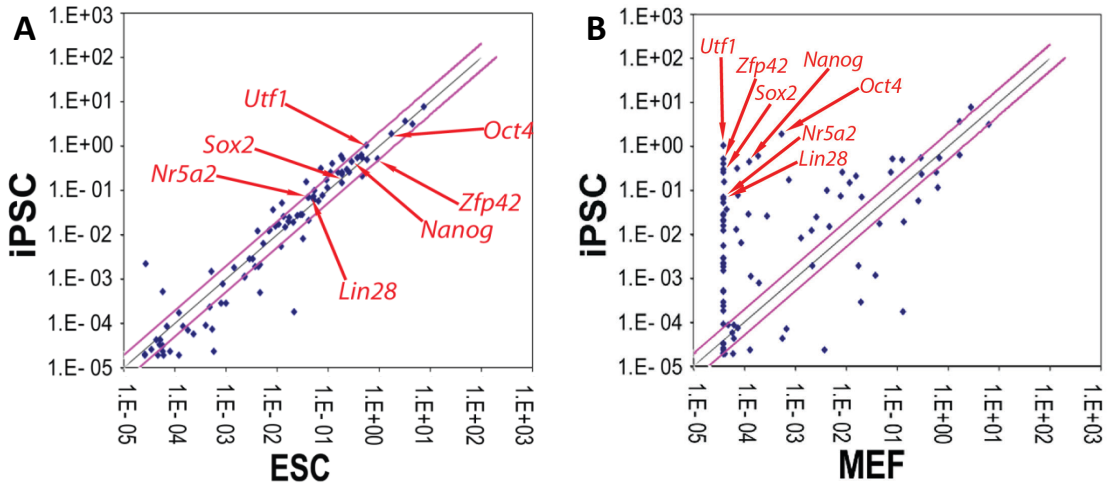


Figure III.7. (A) Scatter plot depicting the $2^{-\Delta CT}$ values of test sample (picked iPSC line) versus control sample (mESC). The graph gives an overall view of how closely the genes in the test sample (picked iPSC line) and control sample (mESC) are expressed. Black line indicates fold changes ($2^{-\Delta CT}$) of 1. The pink line indicates limit of fold change in gene expression threshold. Important pluripotency markers are indicated. (B) Scatter plot showing differences between iPSCs and MEF cells which were used for iPSC generation. Huge difference is seen with most stem cell-linked genes outside the 2 fold threshold.

III.3. Analysis of the UTF1-EGFP construct in iPSCs

III.3.1. EGFP expression in iPSCs transiently carrying the UTF1-EGFP construct

Considering our characterization assay and the reliability of our protocol we moved into analyzing our UTF1-EGFP (UE) construct. We first used a plasmid to transiently deliver UE into iPSCs in order to analyze its functionality. Plates induced to form iPSCs were used for transfection with the transient UE construct on different time points. Transfections were carried out on day 12, 14 and 16 after iPSC induction. Cells were observed 48 hours after each transfection to detect EGFP positive cells. Transfections on day 12 and 14 did not show any green fluorescent cells, however after transfecting the cells on day 16 EGFP expressing colonies could be seen in 48 hours (around 18 days from transduction). While expression was only seen in iPSCs (MEFs did not express any EGFP), only a portion of the cells were EGFP positive (Figure III.8). This is mainly due to transfection limitations, hence only a subset of cells in each colony carry the vector. Vast majority of cells rapidly lost EGFP expression to the point that it was microscopically undetectable four to six days after transfection. This is expected as non-replicating plasmids are degraded and lost over time after mammalian cells divide. Therefore, this makes the transient vector a valuable tool to monitor iPSC pluripotency without genome integrations. This is especially valuable in human cells that require to be assayed without genome alterations for later use in therapeutics.

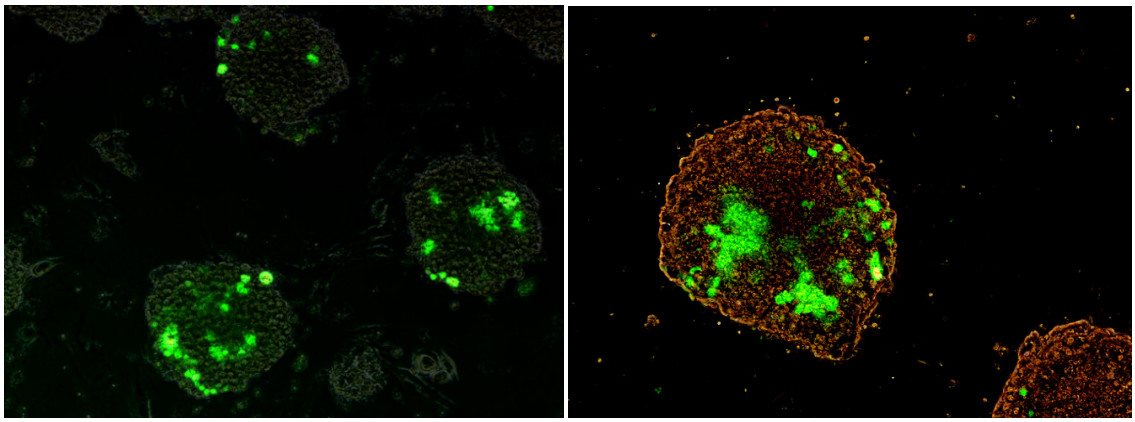


Figure III.8. UTF1 expressing iPSCs transfected with the transient pTZ-UTF1-EGFP vector: UTF-1 expressing colonies show EGFP in a proportion of cells in each colony. Upon transfection of iPSC with pTZ-UTF1-EGFP, UTF1 expressing cells express start to express EGFP. Due to transfection inefficiency not all cells take up the vector and only parts of each colony show expression (Not all cells in one colony are EGFP positive).

III.3.2. EGFP expression in iPSCs carrying the integrated UTF1-EGFP construct

Delivering the UE construct with a transient method into the cell is beneficial as described, however there are some limitations regarding the use of this strategy. This method cannot be used for long term monitoring as the vector signal disappears several days after due to the transient form of the vector. In addition majority of cells remain untransfected due to transfection inefficiencies.

To overcome this issue we proposed to carry the construct into the cells by retroviral transduction. Furthermore transducing the vector into cells allows long term availability of the UTF1-EGFP construct due to its integration. This makes continuous monitoring of cells possible which would allow us to find out the exact timing of UTF1 expression. A further advantage of this method is that all the cells in one colony will possess the construct as it is introduced before iPSC induction (all cells in one colony are derived from one cell carrying the construct). This is not the case in transient transfection which makes consistent monitoring difficult. Hence these characteristics of an integrated

reporter will immensely facilitate UTF1 monitoring. We simultaneously transduced *Oct4, Sox2, Klf4* and *c-Myc (4F)* plus the retroviral pMYC-UE vector and were observed everyday for EGFP expression. Approximately 14 days after transduction appeared colonies showed EGFP expression under microscope (Figure III.9). This indicated early expression of UTF1 as newly risen colonies expressed EGFP. EGFP expression could not be seen in MEFs and the intensity of expression seemed to correlate with the morphology of the colony.

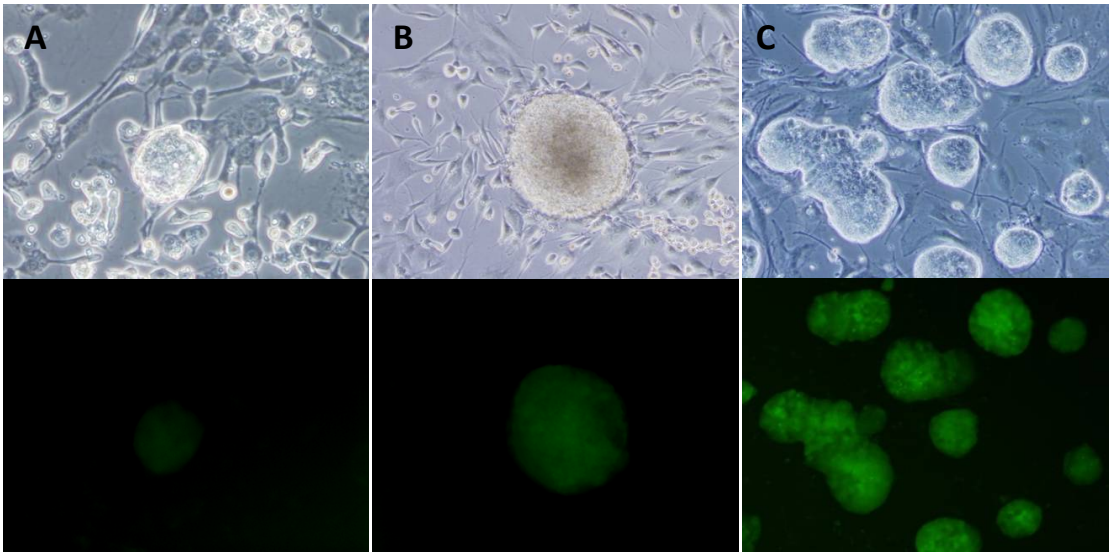


Figure III.9. EGFP Expressing Colonies. (A) EGFP expressing iPSC when first visible at day 14 (B) EGFP expression in a UTF-1 expressing colony carrying the integrated UTF1-EGFP on day 16 (two days after being visible) (C) Picked EGFP positive colony after two rounds of passages.

Generally according to our protocol around 100 colonies are acquired during iPSC formation in one well of a 6-well plate and eighty percent of these colonies were EGFP positive. These EGFP positive colonies morphologically resemble ESCs and no EGFP could be seen in cells which were not reprogrammed, demonstrating that non pluripotent

cells do not express EGFP. Four colonies were chosen randomly for further investigations, which were named GUE-1 to GUE-4.

Transgenes are silenced quite soon after iPSC formation or ESC transfection according to prior studies (Takahashi and Yamanaka, 2006; Yao et al., 2004). This prompted us to monitor our picked UE cell lines in culture to see if the reporter retains its activity for long term. Detectable levels of EGFP were maintained for at least 45 days, proving that the reporter is not subjected to silencing and preserves functionality for prolonged periods (Figure III.10). Furthermore the reporter showed that it's not susceptible to freeze-thawing and iPSCs retained their EGFP expression after freeze thawing.

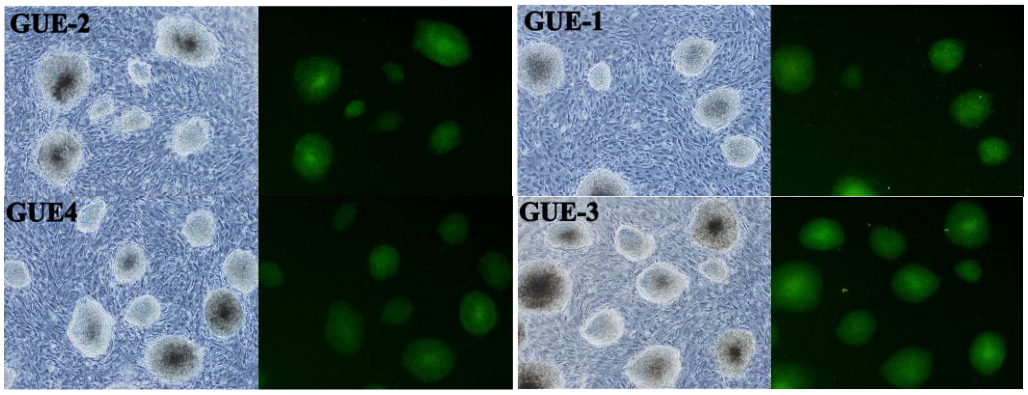


Figure III.10. Long-term functionality and differentiation of UE reporter in iPSCs. Selected UE iPSC lines were established through colony picking, and expression of EGPF was followed until 45 days after co-transduction with 4F +UE. All four cell lines evenly expressed EGFP on day 45.

We sought to assess whether the UE reporter is still functional if there is a long period between introducing the UE reporter into MEFs and induction of iPSCs. This was particularly interesting as to prove that the construct is functional even prior to iPSC induction. For this purpose we transduced the cells first with pMYC-UE and monitored these cells. No UE expression was observed as expected. After 3 weeks in culture, these

cells were used to induce iPSC formation. As seen previously iPSCs expressing the UE reporter started to appear in the plate after two weeks which verifies that the integrated reporter is functional for long periods of time prior to iPSC induction.

III.3.3. SSEA1 expression after 45 days

After maintaining our picked iPSCs in culture for 45 days, we carried out SSEA1 expression to signify that these cells are still pluripotent and that the existing UE expression is due to the pluripotency of these cells. FACS analysis of the four UE positive cell lines (which we called GUE-1 to GUE-4) showed expression of SSEA1 in all these cell lines (Figure III.10), indicating that these cells are still pluripotent.

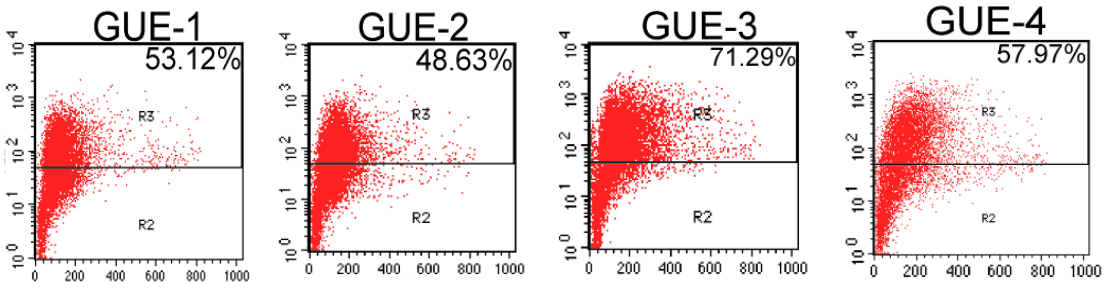


Figure III.10. SSEA1 expression amounts in iPSC lines GUE-1 to GUE-4 was determined after 45 days in culture. All cell lines display SSEA1 activity.

III.3.4. Differentiation of Picked UE iPSCs

Further evidence was required to prove the functionality of our UE reporter upon differentiation; hence we prepared a set of differentiation experiments. Considering UTF1 as a stringent pluripotency factor, its expression should drop immediately upon differentiation. These experiments include differentiating the cells using DMSO, differentiation of cells into EBs and directed differentiation of cells into cardiomyocytes.

These test will also indicate whether our iPSCs are pluripotent due to the fact that an important indicator of pluripotency is the differentiation potential of the cells.

III.3.4.1. DMSO differentiation of UE iPSCs

III.3.4.1.1. Monitoring DMSO differentiation of UE iPSCs by microscopy

We prepared two sets of plates to differentiate in the presence of DMSO and removal of leukaemia inhibitory factor (LIF). One set was monitored and harvested after 6 days of differentiation while the second set was left until day 12. Microscopy observation showed a substantial decrease in EGFP expressing cells (Figure III.11), while some small parts, mainly located in the centre of colonies still retained EGFP expression. A clear change in morphology of the cells was observed after 12 days with most cells differentiated indicating that the reporter goes down after differentiation, following UTF1 downregulation.

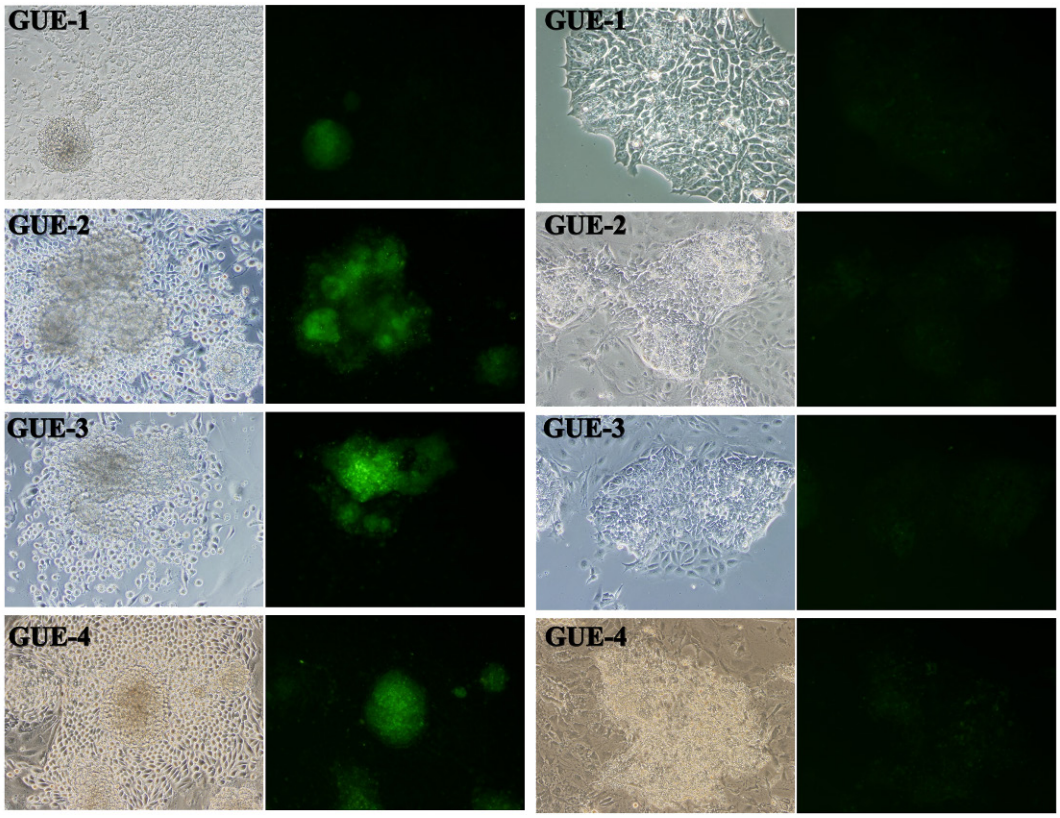


Figure III.11. Microscopy of iPSC lines GUE-1 to GUE-4 at 6 (left column) and 12 days (right column) after induction of differentiation by DMSO. EGFP levels decrease dramatically with some sections still retaining EGFP expression especially in the center of colonies.

III.3.4.1.2. FACS analysis of DMSO differentiated GUE cell lines

In order to confirm that the number of EGFP positive cells has decreased we performed FACS analysis of the differentiated cells at day 6 and day 12 for both sets. We observed a substantial decrease in EGFP expression in all cell lines at both time points. FACS analysis indicated only 10% to 20% of cells remained EGFP positive in comparison to around 90% of cells before treatment. This was accompanied with a decrease in the mean EGFP expression levels which is an indicator of the intensity of EGFP expression (Figure III.12). An exception of the four cell lines was GUE4. Although GUE4 showed a decrease in EGFP expression, this amount was much less in compared to the other cell lines.

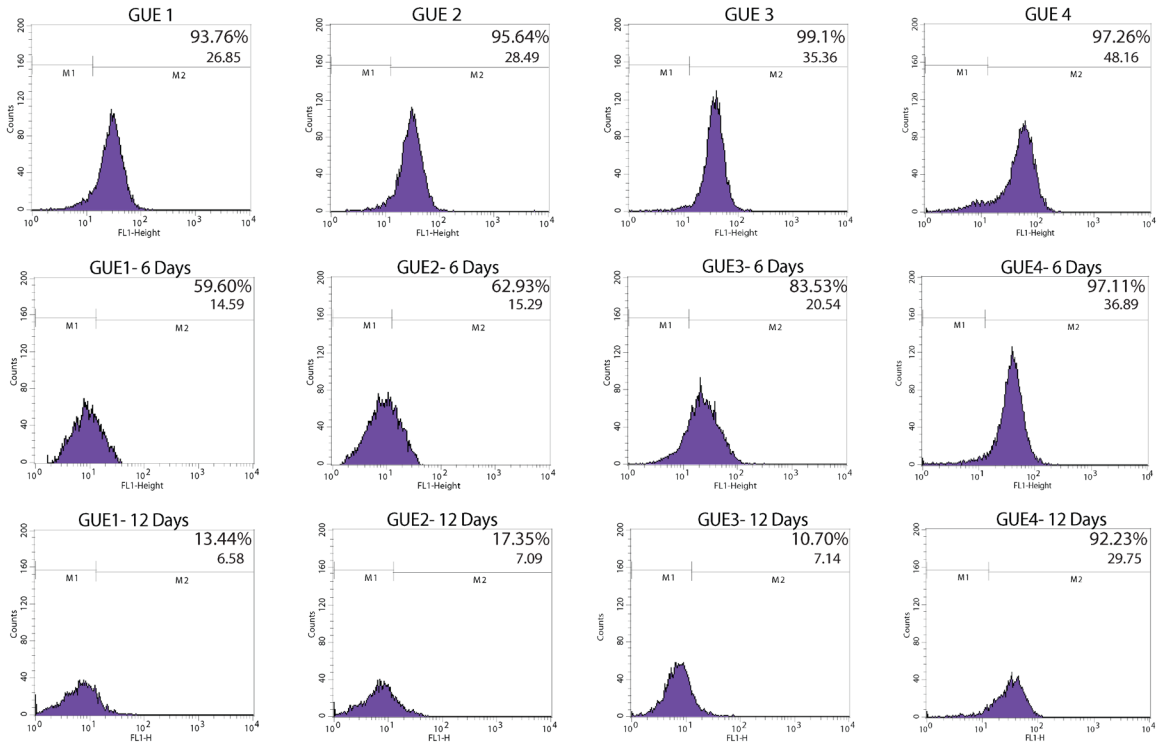


Figure III.12. UE reporter cells were analyzed by FACS before and after treatment with DMSO. Percentage of EGFP positive cells and the Geomean EGFP after 0, 6 and 12 d of DMSO addition are indicated. Geomean in EGFP lacking iPSCs is around 1. Upper number located to the right of each box indicates number of positive cells while the lower number indicates Geomean.

III.3.4.1.3. Real time PCR analysis of differentiated UE cell lines

To confirm differentiation in GUE cell lines and indicate that endogenous UTF1 expression was affected, we carried out real time PCR analysis for different stem cell related genes. This analysis includes quantitative analysis of 84 stem cell related gens as described earlier. On the other hand this test would also indicate whether our iPSCs are able to differentiate properly. The analysis indicated a substantial increase in the expression level of a variety of differentiation markers after 6 and 12 days of exposure to DMSO. Some of these differentiation markers include *T, Sox17, Cdx2, Afp* and *olig2*

which are markers for different cell lineages. In addition, expression of pluripotency markers *Utf1, Oct4*, *Sox2* and *Nanog* concomitantly declined (Figure III.13 and III.14).

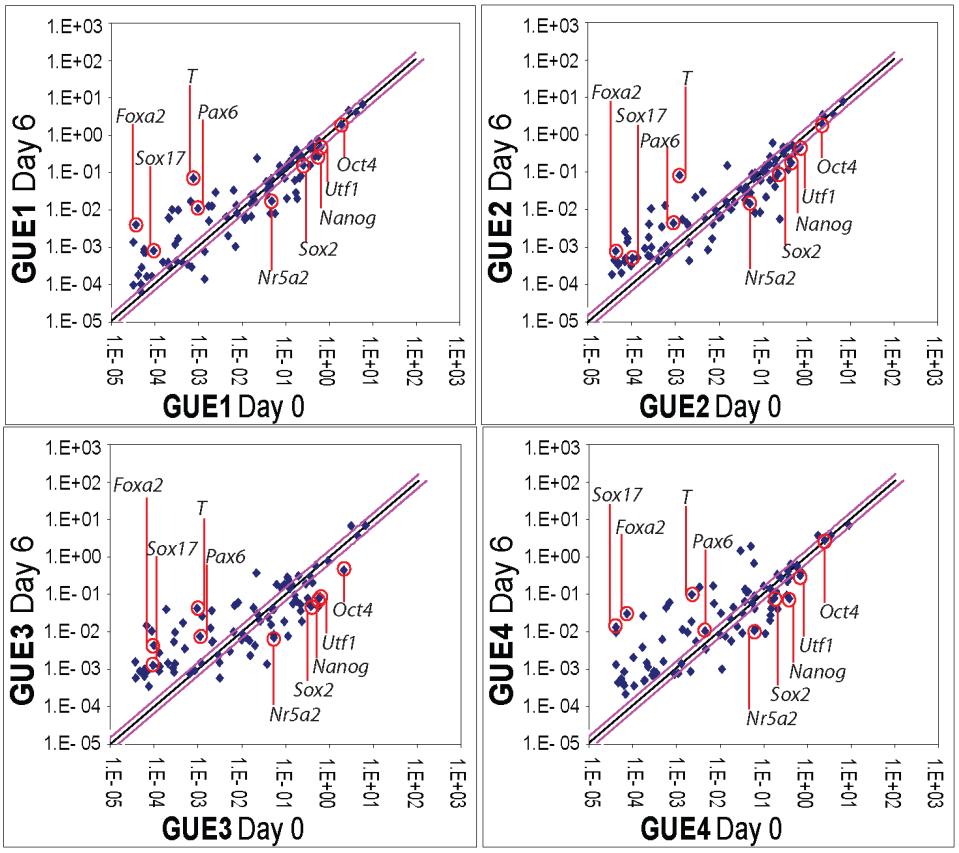


Figure III.13. Quantitative PCR analysis of UE iPSCs after 6 days of differnetiation shown in a scatter plot. The 2−ΔCT values are shown at 6 days of differentiation versus undifferentiated iPSCs as control. The pink lines denote 1.5 fold thresholds for gene expression. Gene expression variation is seen post differentiation, with differentiation markers significantly upregulated and pluripotency markers downregulated. Key pluripotency genes and differentiation markers are marked. The graph displays an average of three repeats for each gene

We observed a lower differentiation rate after 12 days in the parallel plates treated for 6 and 12 days with DMSO (Figure III.14). The overall rate seemed to be slower in cells exposed to DMSO for 12 days in comparison to 6 days, showing less significant changes in gene expression amounts in comparison to 6 days. We think this may be due to a group

of UTF1 positive cell lines which remain undifferentiated in a niche for long periods, still retaining UTF1 expression after 12 days of DMSO addition.

Surprisingly, similar to what we observed in FACS analysis, GUE4 also showed less change in gene expression in compared to the other cell lines.

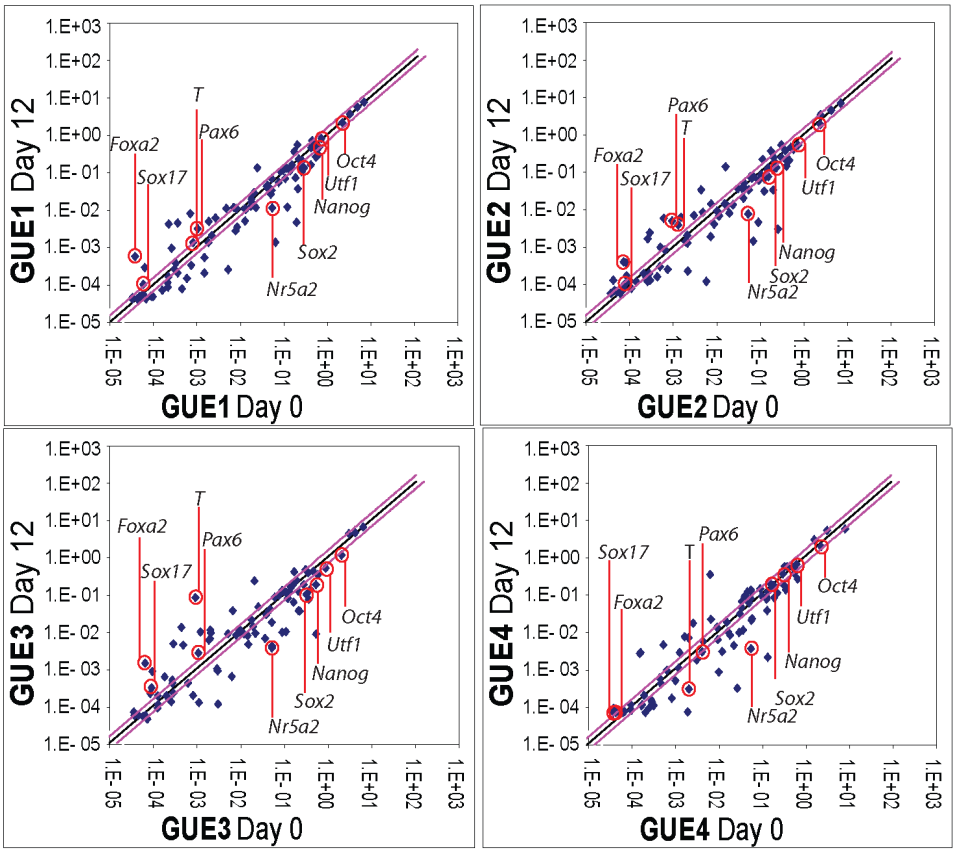


Figure III.14. Quantitative PCR analysis of UE iPSCs after 12 days of differnetiation shown in a scatter plot. The 2−ΔCT values are shown at 12 days of differentiation versus undifferentiated iPSCs as control. The pink lines denote 1.5 fold thresholds for gene expression. Gene expression variation is seen post differentiation, with differentiation markers significantly upregulated and pluripotency markers downregulated. Key pluripotency genes and differentiation markers are marked. The graph displays an average of three repeats for each gene.

III.3.4.2. Differentiating UE iPSCs into embryoid bodies

In order to show that our iPSCs can form EBs and express differentiation markers embryonic body (EB) formation was carried out under appropriate conditions. We observed expression of the three known germ line markers Brachyury, Nestin and AFP between six and twelve days using undifferentiated cells as control (Figure III.15).

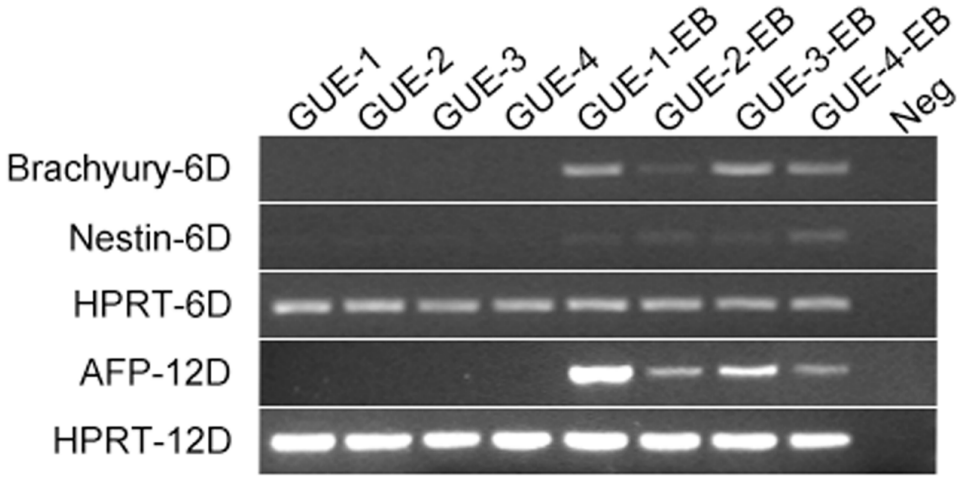


Figure III.15. Formation of EB and expression of the three prominent germ layer markers. EB formation was induced in all GUE cell lines for 6 and 12 days and expression of differntiation markres is shown by RT-PCR. AFP is shown at day 12 as it is known to be expressed a bit later than other markers. HPRT expression is shown as control.

We monitored these cells for EGFP expression to observe UTF1 expression pattern in EBs. EGFP expression decreased during EB formation however some EBs were still EGFP positive after 12 days mainly in centre parts similar to what we observed in DMSO differentiation of these iPSCs. As stated previously, this may be due to the presence of some cells within the colonies which always remain in a stem cell condition and may take prolonged periods to differentiate.

III.3.4.3. Cardiomyocytes directed differentiation of UE iPSCs

To rule out any specific effect that DMSO might have on the UE reporter regulation, we monitored UE expression during cardiomyocyte differentiation. Beating colonies started to come up around day 10 and most cells had lost their EGFP expression (Figure III.16).

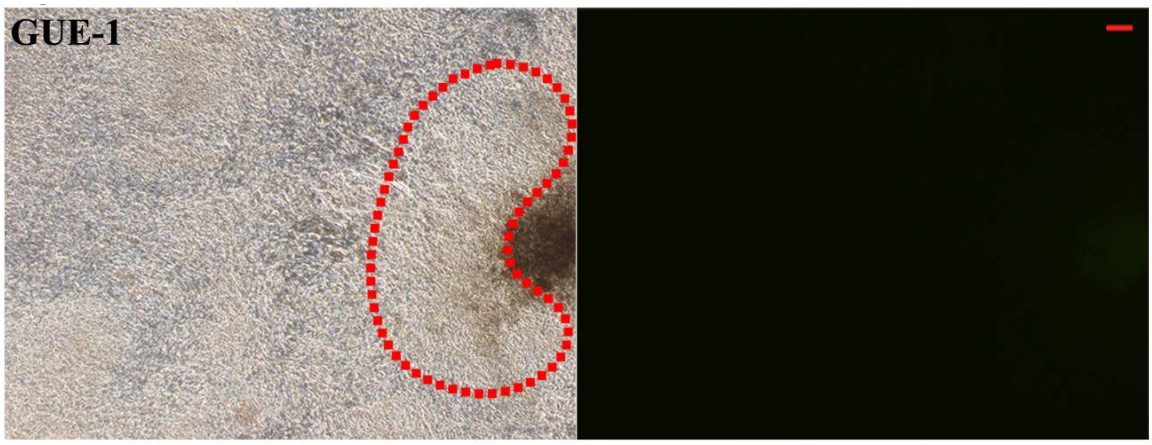


Figure III.16. Cardiomyocyte-directed differentiation. An example of cardiomyocyte induced differentiation is shown in GUE-1 around 14 days after differentiation. Highlighted cluster shows a beating section of cells (left). UE expression is lost in differentiating cells (Right). Scale bar indicates 100 μm.

Interestingly, GUE4 did not show any beating colonies, although the EGFP amounts decreased. One possibility is that this cell line may need more time to differentiate and may explain why GUE4 retained more EGFP expression and less gene expression changes after 12 days in compared to the other cell lines after DMSO treatment. In brief our data indicate that this reporter is a reliable sensor of the pluripotent state.

III.4. Analysis of mouse tail fibroblasts carrying a *Utf1*-tomato reporter

The knockin mice were generated by introducing a tandem repeat of the *tomato* gene into the *Utf1* locus. This strategy enables higher tomato protein expression, hence easing

detection. The location where the gene is inserted is 519 bp from the 5' end of *Utf1* disrupting exon1. An optimized Kozak sequence gccaccATGG was inserted at the start of the UTF1 gene in the target vector. A selection marker flanks the dtmoato region which can be excised using the Cle or Flp recombinase.

The *Utf1* knockin reporter mouse was generated by Dr. Yu Wei-Ping from the Biological Resource Centre (BRC) in Singapore. All the animal procedures were performed in the A*STAR BRC by Dr. Yu and his colleagues according to the Institutional Animal Care and Use Committee guidelines (IACUC).

The detailed strategy used to generate the knocking mouse is as follows (Figure III.16):

1) The 519bp at the 5'-end of UTF1 coding sequence is replaced by tdTomato (Clonetech). tdTomato encodes two tandem copies of Tomato protein, which allows intramolecular dimerization. tdTomato gives rise to strong fluorescence signal, which facilitates detection of weaker signals and in vivo animal imaging.
2) The Kozak sequence in the targeting vector is optimized by replacing the 6 bp (tcaggg) prior to ATG in the original UTF1 gene by a 4 bp sequence of cacc, which together generates the optimal Kozak sequence, gccaccATGG.
3) The selection marker can be excised by expression of Cre or Flpe either in ES cells or animals (by crossing with respective recombinase-carrying deleter strain).

4) Genotyping by Southern with XhoI:

Mouse type / Probe	Wild type	Mut before excision of selection marker	Mut after excision of selection marker
5' external probe	7904 bp	4963 bp	3229 bp
3' external proble	7904 bp	5662 bp	5662 bp

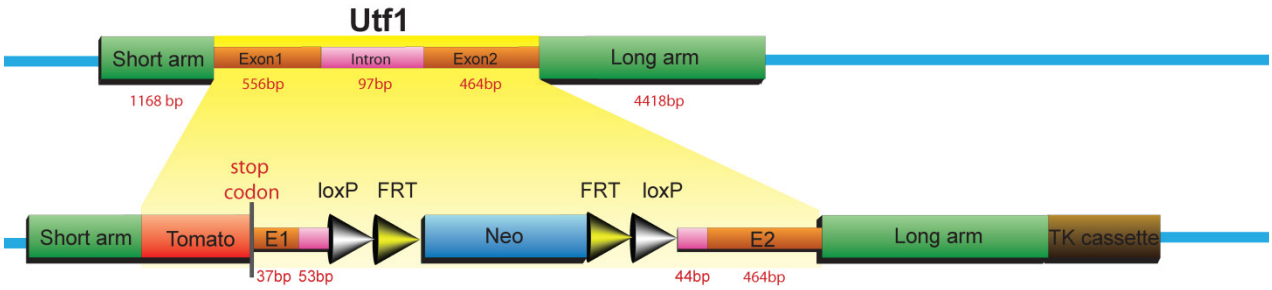


Figure III.16. Outline for UTF1 Targeting Vector: pTGV-UTF1

As stated, the tomato gene interrupts *Utf1*; hence derived homozygous mice should lack UTF1 expression. A preliminary inspection revealed that homozygous mice are still viable, and *Utf1* knockout does not affect viability. Tail tips were collected from these mice to test the functionality of this reporter in iPSCs

III.4.1. Monitoring mouse tail fibroblasts and formation of iPSCs

Fibroblasts derived from tail tips of mice carrying the dtomato gene knocked into the *Utf1* locus were cultured *in vitro*, to show that the reporter is functional. Cells started to grow around the mouse tail tips which were put into gelatin coated plates. Tail tips were removed several days after. Cells were grown and were monitored for expression of tomato for several days prior to iPSC induction. No tomato expression could be seen in

these fibroblasts even after several passages. Subsequently, cells were transduced with 4F and monitored every day for fluorescence activity. Red fluorescent colonies started to appear 14 days after transduction, indicating that the reporter is functional. Subsequently, iPSC colonies were monitored for several days to see their expression. Remarkably, Colonies with different intensities of tomato expression could be seen with the fluorescence intensity correlating with the morphology of these colonies, i.e. how closely they resemble ESCs. Colonies not resembling ESCs indicated low or no expression of tomato (Figure III.17). This indicates that the "reporter intensity" provides information about the relative amount of UTF1 expression in these cells.

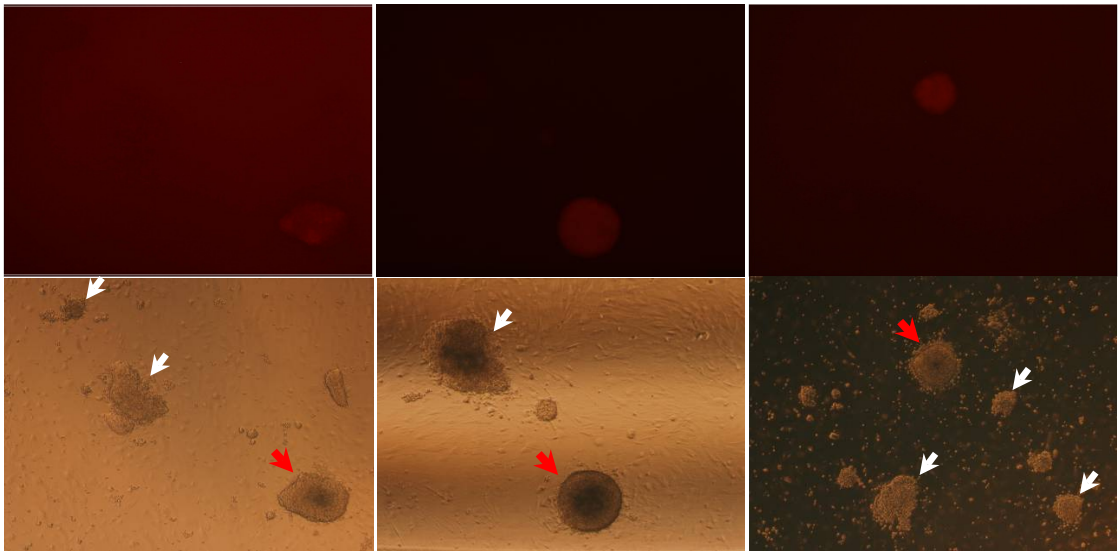


Figure III.17. Images of newly appearing iPSC colonies. The red arrows indicated colonies which resemble ESCs and show tomato expression, while white arrows point to colonies which do not resemble ESCs and have no tomato expression. Intensity of red fluorescence is also differnet in positive iPSC colonies.

After several days of iPSC growth, strong red fluorescence iPSCs could clearly be seen (Figure III.18). One colony was isolated and propagated under ESC conditions for further investigation.

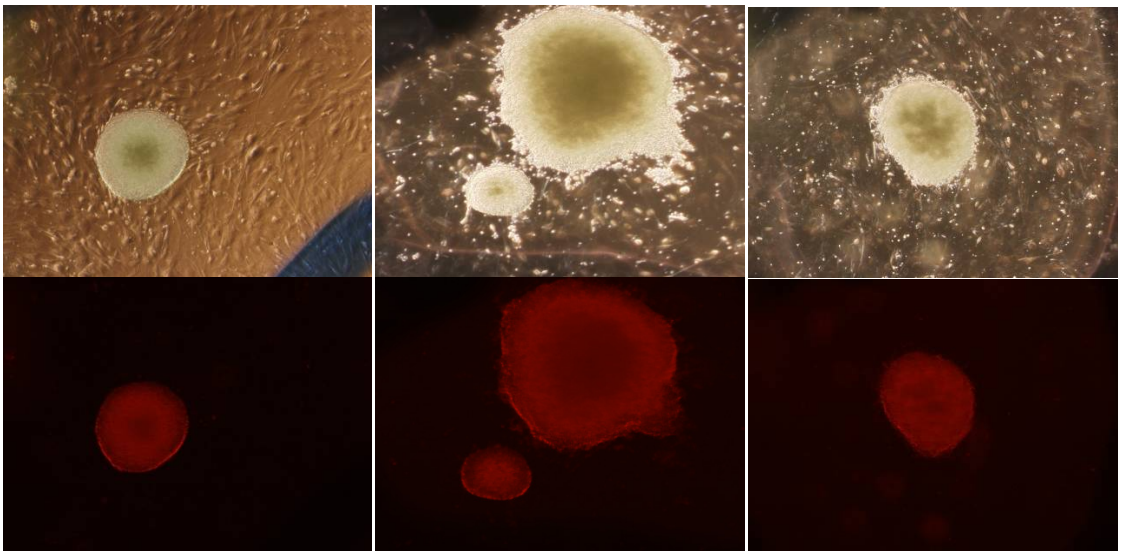


Figure III.18. Different iPSC colonies after several days of growth with red fluorescence expression.

III.4.2. Random differentiation of mouse tail iPSCs

The derived iPSCs were left to differentiate in differentiation medium (20% FBS, without LIF) for 12 days (Figure III.19). This would indicate if the reporter is still functional after differentiation. Notably, fluorescent cells in the centre of colonies remained fluorescent until day 12 and the size of the red fluorescent area did not change; however the differentiated cells surrounding the fluorescent central part expanded and had no tomato expression.

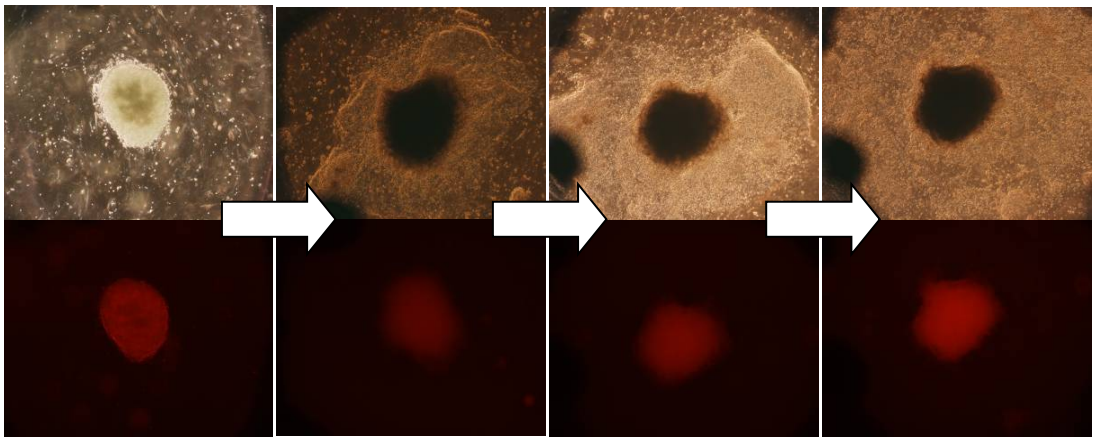


Figure III.19. iPSC colonies differntiating during 12 days. Panel shows an iPSC colony after 12 days which still retaines red fluorescence in the center of the colony. Note that the size of the fluoeresent section has not changed.

III.4.3. Differentiation of *UTF1*-tomato iPSCs into EB

Next, we differentiated the *Utf1-Tomato* carrying iPSCs into EBs under specified conditions using non adherent plates. These cells were continuously monitored for red fluorescence expression. Most differentiated cells lost tomato expression after 12 days; however some of the EBs retained red fluorescent signals particularly in their central regions until day 12 (figure III.20). This is in agreement with a previous study using an *Oct4-Egfp* construct reporter, indicating that groups of EBs remain OCT4 positive after 14 days in culture. In fact the number of OCT4 positive cells appeared to rise by 30% after 25 days in culture confirmed by RT-PCR analysis. The cluster of OCT4 positive cells indicated chromosomal abnormalities by either containing an additional copy of chromosome 8 or 9 (Ensenat-Waser et al., 2006).

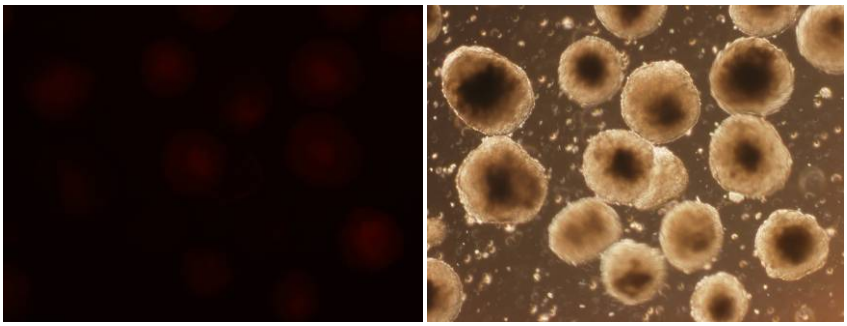


Figure III.20. EBs at day 12 still maintaining some tomato expression with a stronger signal in the central sections of the colony.

III.4.4. Cardiomyocyte differentiation of *UTF1-tomato* iPSCs

Considering that some cells remained fluorescent during random differentiation of iPSCs, we induced-cardiomyocyte formation using vitamin C. These iPSCs were monitored every day for fluorescence activity. Expression declined until beating cardiomyocytes appeared, however, small parts, particularly in the center of the colonies retained tomato expression. (Figure III.21).

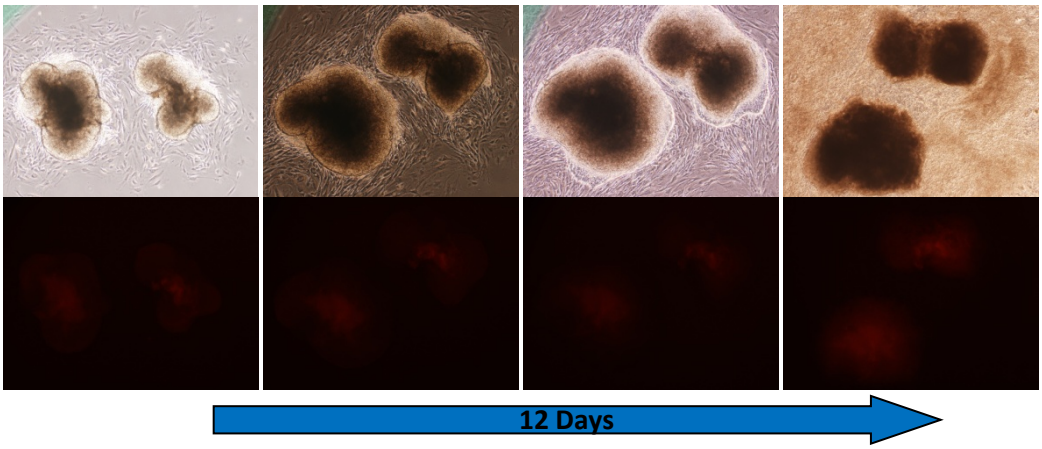


Figure III.21. Cardiomyocyte directed differentiation of iPSCs using vitamin C. Note that centre of the colonies remains fluorescent after 12 days. Surrounding cells started to beat around 10 days after differentiation in the presence of vitamin C.

We hypothesize that this is due to the fact that cells located in the center do not replicate and remain in a resting state which preserves a pluripotent(-like) state and prevents them from differentiation.

To verify our assumption, we added DMSO to the cells to force differentiation. After 5 days colonies retained the red fluorescence in the center. We then trypsinized the cells on day 17 and passaged them at lower density in the presence of DMSO to trigger cell proliferation. Remarkably, observation of these cells after passaging did not show fluorescent cells. After several days in culture, clusters of cells re-appeared which partially resembled ESCs morphology; however, they did not show red fluorescence (Figure III.22). This supports our hypothesis that non-dividing cells retain their pluripotency and cells need to replicate in order to fully enter differentiation pathways.

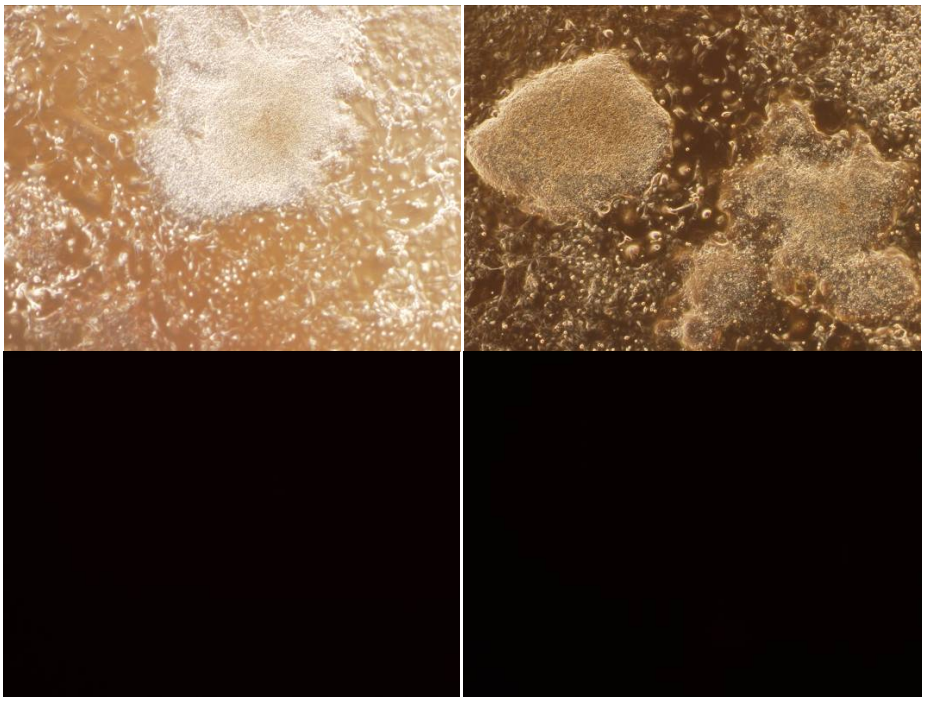


Figure III.22. No fluorescence expression can be seen in tomato iPSCs after passaging these cells. Colonies of cells are still seen, however no tomato expression is present.

III.5. Effect of HMGA2 on iPSC formation

III.5.1. Colony formation in HMGA2 iPSCs

The HMGA2-AA-EGFP retroviral based expression vector enabled us to observe translation of exogenous human HMGA2 protein after transduction of MEFs. Mouse and human HMGA2 proteins are highly conserved and share 95% identity at the aa level. A comparable vector missing the HMGA2 coding region and carrying only the EGFP expression sequence was utilised as control. iPSCs were generated with the 4F plus either HMGA2 or EGFP according to previously described protocol. Following iPSC colony count in HMGA2 transduced cells after 18 days AP staining was performed for both plates (Figure III.23).

This assay was carried out to see if HMGA2 affects the amount of AP positive iPSC colonies which is an indication of the number of pluripotent iPSCs. A slight decrease of about 30 percent was seen in the number of colonies due to the presence of HMGA2. One reason may be that HMGA2 which is not usually expressed in fibroblasts promotes cell cycle arrest in these cells (Narita et al., 2006).

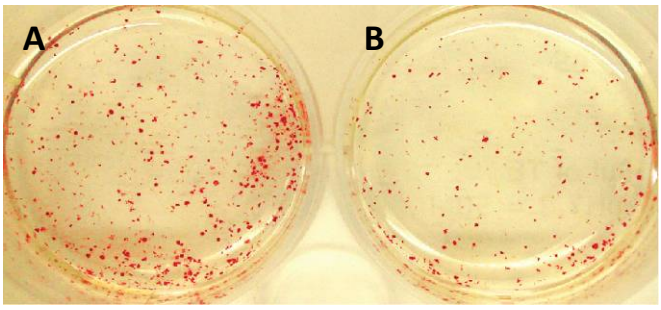


Figure III.23. Effect of HMGA2 on iPSC colony numbers. A slight but reproducible difference in iPSC colony number is seen after alkaline phosphatase staining. Figure shows Comparison of (A) AP positive colonies in HMGA2 and (B) 4F plus HMGA2 plates.

III.5.2. Expression of HMGA2 in transduced iPSC

After iPSC carrying HMGA2-2A-EGFP and EGFP were generated, four colonies were randomly picked based on ESC morphology and established as iPSC lines. Picked colonies from both HMGA2-2A-EGFP and EGFP plates were paired randomly for analysis of HMGA2 expression and subjected to western blot analysis after five rounds of passage. HMGA2 carries a modified form which is SUMOylated. This form was also considered in quantifying the amount of HMGA2 expression. This posttranscriptional modification of HMGA2 has been previously shown in COS cells, however its functional associations have not yet been clearly understood (Cao et al., 2008).

We detected 1.6 to 4 fold overexpression in the first three cell lines (H1-H3) compared to their respective controls (G1-3). The amount of HMGA2 in cell line G4 was higher than H4 due to higher expression of endogenous HMGA2 expression. Our data also revealed that our MEFs do not express HMGA2 and expression must have been induced during reprogramming (Figure III.24). This is consistent with the fact that ESCs express high amounts of HMGA2 (Li et al., 2006).

To verify that the upper band is SUMO-HMGA2 we performed digestion of lysates collected form iPSCs with SUMO protease. Primary digestion result showed a decrease in SUMO-HMGA2 with an increase in HMGA2 itself (Figure III.24).

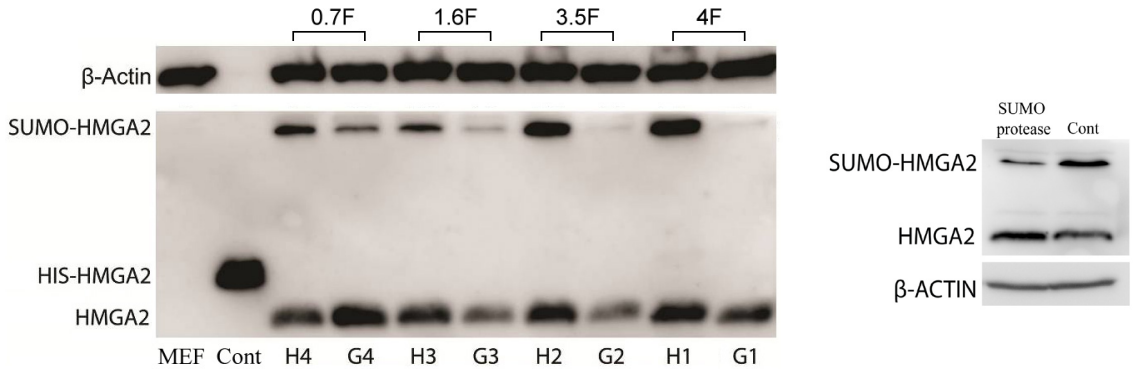


Figure III.24. Western blot analysis showing 4F+HMGA2 iPSC lines (H1-4) paired randomly with 4F+EGFP (G1-4). Controls include cell lysate of MEF cells (MEF) and purified Histagged HMGA2 from *E.coli* (Cont). The relative fold change for HMGA2 is shown above the figure (Normalized to β-actin). Both SUMOylated and unmodified HMGA2 were considered for quantification. The right panel shows SUMO protease treated HMGA2.

We repeated the SUMO protease experiment this time with different amounts of enzyme. Data showed a decrease in SUMO-HMGA2 upon increase in the amount of enzyme (Figure III.25).

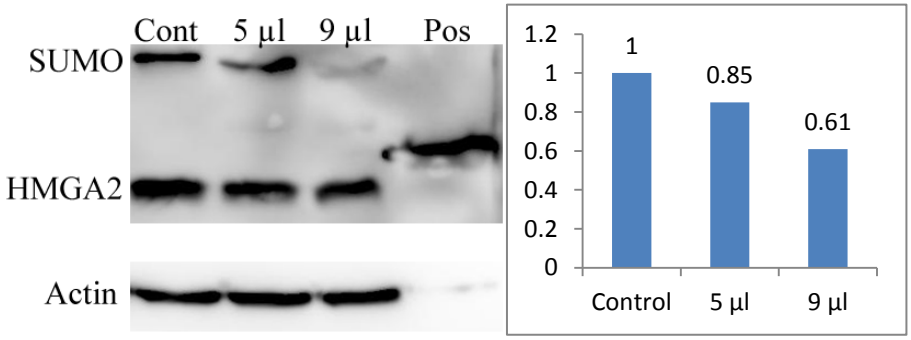


Figure III.25. Treatment of iPSC lysate with 5 and 9 µl of SUMO protease. The untreated sample served as control (cont). Positive band (pos) shows purified HMGA2 from *E.coli*. SUMO bands were quantified and normalized. The graph on the right shows quantification of SUMO bands after normalization and equalizing of the HMGA2 bands. Decrese in SUMO amounts can be observed after treatment with differnet amounts of enzyme.

III.5.3. Pluripotency characteristics of HMGA2 iPSCs

III.5.3.1. SSEA1 expression in HMGA2 expressing cell lines

We sought to analyze our HMGA2 positive cell lines for some pluripotency tests and see if these cells differ to normal iPSCs in pluripotency. The four HMGA2 overxpressing iPSC lines together with EGFP controls were assayed for SSEA1 expression. Data indicated that HMGA2 overexpressing cell lines express SSEA1 similarly to controls indicating that the HMGA2 iPSCs are also pluripotent. SSEA1 expression amounts did not appear to correlate with HMGA2 levels. Both groups of iPSCs express this factor in similar amounts (Figure III.26).

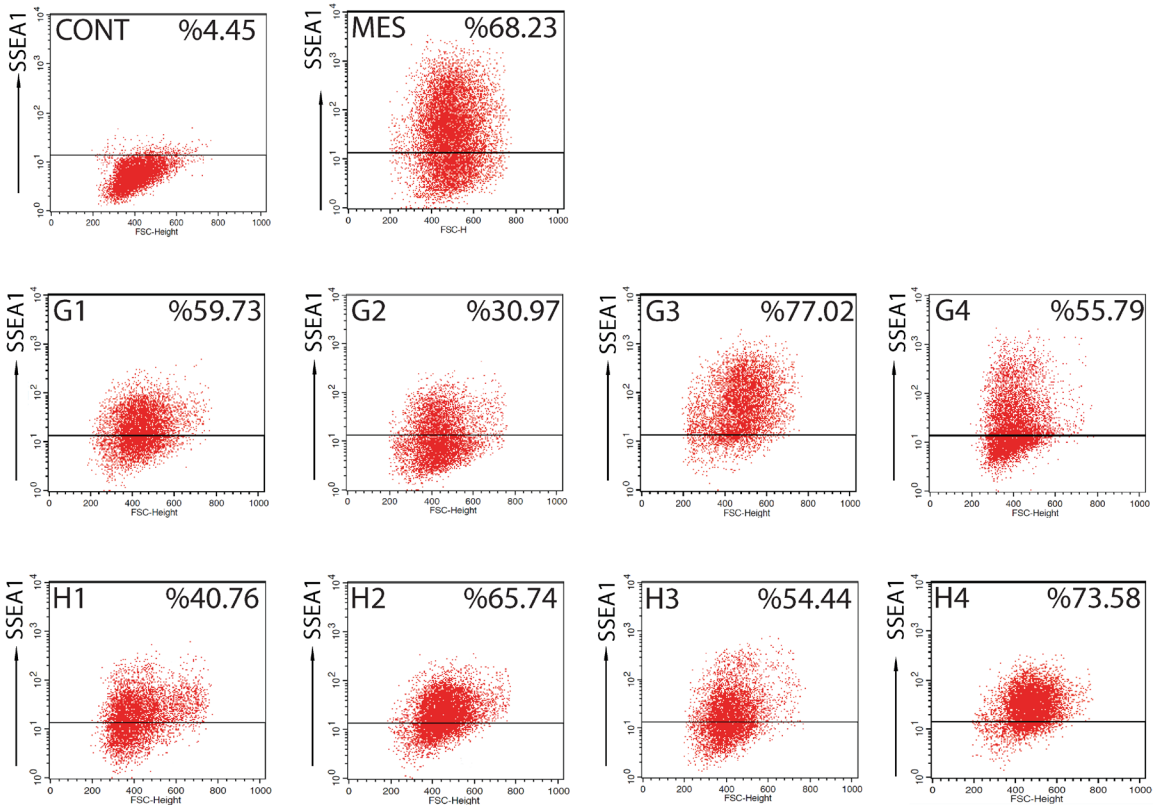


Figure III.26. FACS analysis was used to assay all cell lines for SSEA1 expression. mESCs (MES in figure) and nonstained mESC served as controls.

III.5.3.2. OCT4 expression in HMGA2 expressing cell lines

In the next step we looked into OCT4 expression in these cells. All cell lines from both group showed to express OCT4, indicating that they are pluripotent (Figure III.27).

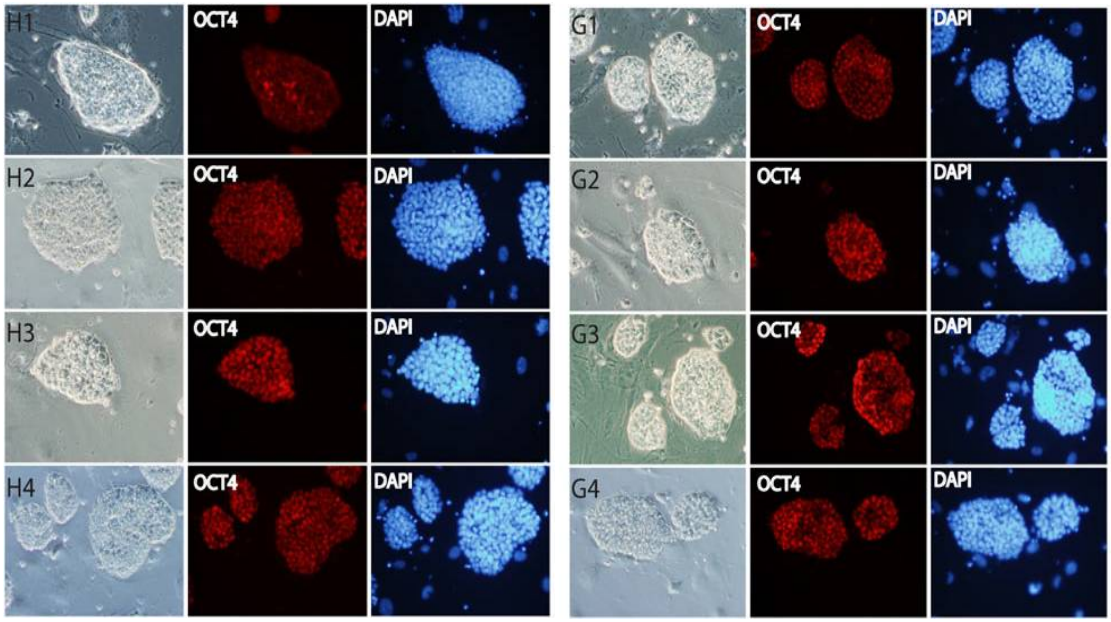


Figure III.27. Strong nuclear immune-staining is seen for pluripotency marker OCT4 in all eight cell lines. Figure shows presence of OCT4 in random colonies selected from each plate and does not indicate quantitative amounts.

III.5.3.3. Alkaline phosphatase expression

Both groups of cells were also positive for SSEA1 expression, further indicating that all cell lines are pluripotent (Figure III.28). No difference was seen in the intensity of AP staining.

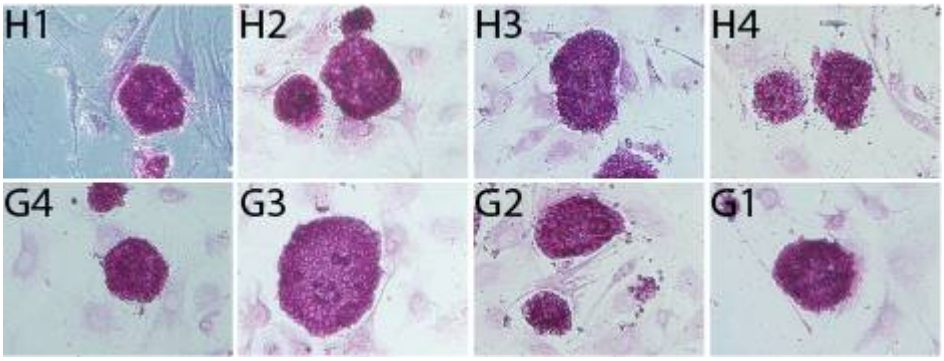


Figure III.28. AP staining of 4F plus EGFP and 4F plus HMGA2. Both group of cells showed AP activity with similar intensity of AP staining in their cells.

III.5.4. Microarray analysis

III.5.4.1. Biological processes affected by HMGA2

The *HMGA2-AA-EGFP* and *EGFP* iPSC lines were used for microarray analysis. This was done to see how HMGA2 affects the overall gene analysis in iPSCs. Cell line H4/G4 was not examined any further due to the absence of differential HMGA2 expression. At cell passage five after picking the iPSCs, RNA was extracted from G1/H1, G2/H2 and G3/H3 to perform microarray analysis. Note that this assay was carried out prior to silencing of *HMGA2-AA-EGFP* and *EGFP* transgenes.

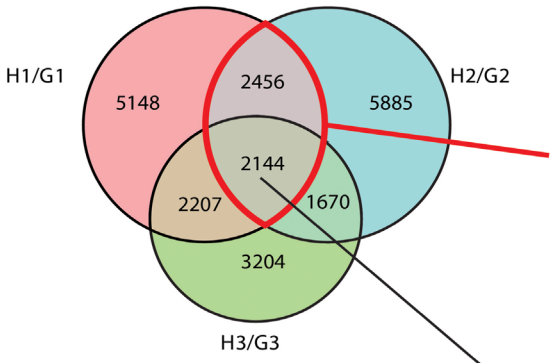
Within each of the three pairs, control iPSCs (G) were compared to HMGA2-iPSC (H). Analysis of data showed that around 2000 entities are expressed differentially in all three cell lines which contain the lincRNAs as well. Differentially expressed genes within these entities were then analysed and clustered with Genespring and grouped into Gene Ontology (GO) terms according to a corrected p-value of (<0.01). Strikingly, developmental/anatomical processes were shown to have the lowest P-value and affected mostly (Figure III.28), indicating that HMGA2 mostly influences anatomical related genes which comes into agreement with earlier findings on HMGA2's role in

development and body size (Hirning-Folz et al., 1998; Weedon et al., 2007; Zhou et al., 1995).

For a more in comprehensive analysis we selected the first two pairs of cell lines, which showed the highest differential HMGA2 expression in compared to H3/G3. Data showed that similarly anatomical/developmental genes were still primarily affected in two pair analysis (Figure III.28), signifying HMGA2 as a critical factor in regulating body size (Hirning-Folz et al., 1998; Weedon et al., 2007; Zhou et al., 1995).

Note that skeletal muscle developmet genes like Col6a6, Col6a2 and Col6a1 showed differential expression (Fitzgerald et al., 2008). Col6a6 was upregulated in all three cell lines, while Col6a2 and Col6a1 were only upregulated in the first two cell lines (Supplementary 3). In addition Bmp2, a desicive bone and cartliage development gene was elevated in H1 and H2 cell lines (Supplementary 3) (Chen et al., 2004).

Aside from developmental/anatomical processes, we observed cell and biological adhesion and nervous system development to be affected next (Figure III.29).



GO ACCESSION	GO Term(Biological Process)	P-Value(0.01)
GO:0007275	multicellular organismal development	2.89E-15
GO:0032502	developmental process	4.62E-15
GO:0048856	anatomical structure development	8.69E-14
GO:0048731	system development	3.94E-13
GO:0048513	organ development	2.84E-12
GO:0030154	cell differentiation	4.30E-09
GO:0048869	cellular developmental process	4.30E-09
GO:0009653	anatomical structure morphogenesis	6.11E-08
GO:0032501|GO:0050874	multicellular organismal process	1.03E-04
GO:0007155	cell adhesion	6.45E-04
GO:0022610	biological adhesion	6.45E-04
GO:0048511	rhythmic process	6.82E-04
GO:0007399	nervous system development	6.82E-04
GO:0048468	cell development	8.85E-04
GO:0050896|GO:0051869	response to stimulus	1.07E-03
GO:0042127	regulation of cell proliferation	4.56E-03
GO:0009887	organ morphogenesis	8.34E-03
GO:0050801	ion homeostasis	8.49E-03

GO ACCESSION	GO Term(Biological Process)	P-Value(0.01)
GO:0032502	developmental process	7.39E-07
GO:0007275	multicellular organismal development	3.37E-06
GO:0048856	anatomical structure development	1.08E-05
GO:0048513	organ development	2.58E-05
GO:0048731	system development	2.91E-05
GO:0009605	response to external stimulus	4.22E-04
GO:0009888	tissue development	3.40E-03
GO:0009611|GO:0002245	response to wounding	3.81E-03
GO:0050896|GO:0051869	response to stimulus	7.87E-03

Figure III.29. Selected genes from cell line pairs H1/G1, H2/G2, and H3/G3 were analyzed for Gene Ontology (GO). The image shows area proportional venn diagram of common entities differentially expressed in the three pairs. The table indicates GO analysis of common entities differentially expressed between all three cell line pairs, and GO analysis of common entities differentially expressed only in pairs H1/G1 and H2/G2, applying a p-value cut-off of 0.01.

The effect of HMGA2 on cell adhesion may be due to the tumour invasive properties of this factor. According to prior studies HMGA2 is an inevitable factor in tumour growth and epithelial to mesenchymal transition (EMT). EMT is triggered by HMGA2 by the RAS signalling pathway. Its depletion represses cell proliferation and drives the transition back to the epithelial state by increasing E-cadherin expression and promoting cell to cell adhesion (Watanabe et al., 2009). Remarkably, Our in depth scrutiny of the differentially expressed genes also showed that Epidermal Growth Factor Receptor (EGFR) a prominent modulator of EMT (Thiery et al., 2009) is also influenced by HMGA2 expression (Supplementary 3). EGFR is one of the factors long known to induce a variety of cancers like lung cancer and brain tumours. This factor plays a major part in cancer

outset and serves as a key target for drug development and therapeutic purposes (Zhang et al., 2007). Our data indicated that mRNA levels of *EGFR* increased in all three cell lines while the greatest change was witnessed in H1 and H2 which displayed higher HMGA2 expression in compared to H3 (Supplementary 3). Genes affected in the Cell\biological adhesion category showed other key EMT genes like *RhoB* to be upregualted and *Cdh4* being downregulated in H1 and H2 (Supplementary 3) (Thiery et al., 2009).

Our data showed that mouse *Tert* expression is not affected significantly (Supplementary 3), although a previous report by Ann and co-workers found that HMGA2 expression modulates human telomerase reverse transcriptase gene (*hTERT*) and in turn possibly inducing tumorigenesis (Li et al., 2011) .

III.5.4.2. Major pluripotency factors are down-regulated in HMGA2 over-expressing cell lines

As mentioned earlier HMGA2 overexpression does not affect iPSC formation as HMGA2 over-expressing colonies resembled normal iPSC colonies in several pluripotency tests, however HMGA2 transudced plates displayed a minor decrease in iPSC colony amounts. This comes into agreement with our previous data that HMGA2 may be an influential factor in pluripotency. Remarkably we observed that core pluripotency related genes from our microarray data like *Nr5a2, Utf1, Oct4, Sox2* and *Nanog* were all reduced in H1 and H2 cell lines. On the other hand in H3 cells, which only had a 1.6-fold increase in HMGA2 over G3 cells, expression of these pluripotency genes was much less affected (Figure III.30).

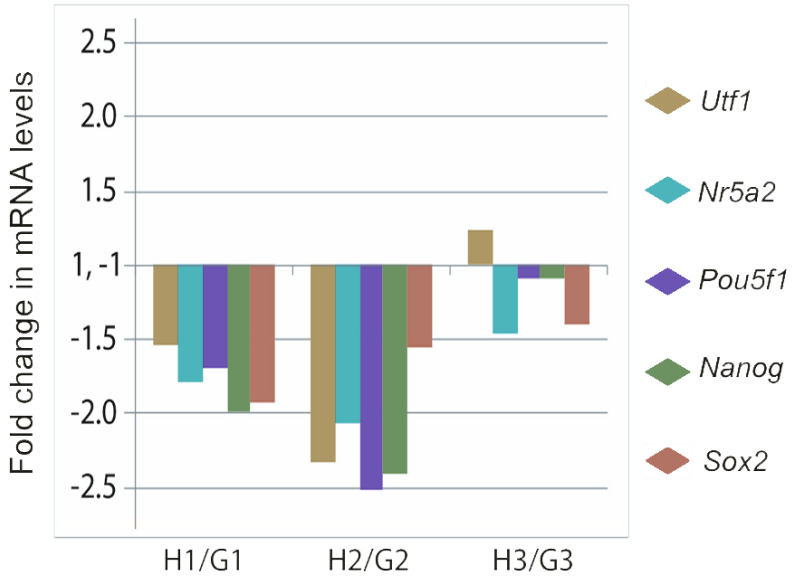


Figure III.30. Core pluripotency markers differentially expressed in three iPSC pairs. Graphical display depicts indepth analysis of microarray data revealing that five key pluripotency marker genes are down-regulated in cell lines with high HMGA2 amounts. Moreover, the HMGA2-expressing H1 and H2 cells show a higher decrease in mRNA levels for these markers in comparison to H3 cells.

Together these data imply that HMGA2 may be a notable factor controlling cell differentiation. This idea was supported by our microarray data GO analysis showing that HMGA2 plays an important role in regulating cell differentiation processes. For example previous studies have reported HMGA2 to be crucial in heart development of *Xenopus laevis* (Monzen et al., 2008).

III.5.4.3. Other key regulatory genes affected by HMGA2

We further looked into other key regulatory genes that may be influenced by HMGA2 overexpression and concentrating on the fact that HMGA2 is one of the factors conferring susceptibility to type II diabetes (Ohshige et al., 2011). In addition, the recent study on the effect of *Lin28-let7* on glucose metabolism connects HMGA2 to diabetes (Brants et al., 2004; Ohshige et al., 2011; Voight et al., 2010; Zhu et al., 2011). Known

diabetes linked genes *Igf2bp2, Irs1 and PiK3ip1* were all seen to be upregulated in our data set for all three cell lines. Our analysis revealed that previously identified diabetes related genes like *Igf2bp2*, *Irs1* and *PiK3ip1* increased in cells with higher HMGA2 (Supplementary 3). Strikingly, similar to HMGA2 these genes are targets of let-7 miRNA (Zhu et al., 2011). Our study points to the existence of a possible positive feedback loop where HMGA2 further enhances the effect of *lin28-let7* mediated *Igf2bp2*, *Irs1* and *PiK3ip1* expression (Zhu et al., 2011). This may also designate that HMGA2 induces the "Warburg effect" whereby a metabolic shift is seen in cancer cells with increase in glucose uptake.

Other genes linked to diabetes like *Igfbp1* and *Kcnq1* also displayed a large increase in their expression amounts (Supplementary 3) (Firth and Baxter, 2002; Ohshige et al., 2011). Suprisingly, *Lepr* which is the gene for leptin receptor, one of the major regulators of adipocytes, also increased substantially in H1-3 lines (Supplementary 3). Previous reports have also linked HMGA2 to metabolism of adipocytes, which might also be linked to the onset of diabetes (Arlotta et al., 2000).

III.5.5. EB formation in HMGA2 expressing cell lines

All eight HMGA2 cell lines were cultured in embryoid body EB conditions to assay for different differentiation markers. We discovered that all cell lines were capable of generating EBs and both HMGA2 and control cell lines expressed chosen markers after 6 days of EB formation (Fig. III.31). This suggests that likewise normal iPSCs, overexpressed HMGA2 iPSC lines differentiate in a similar manner. Interestingly, increased levels of *Nes* (nestin) mRNA were observed in iPSCs expressing higher levels

of HMGA2 (Figure III.31). This corresponds with previous studies in our lab showing a decrease in *Nes* amounts in HMGA2 knockdown human ESC (Li et al., 2007).

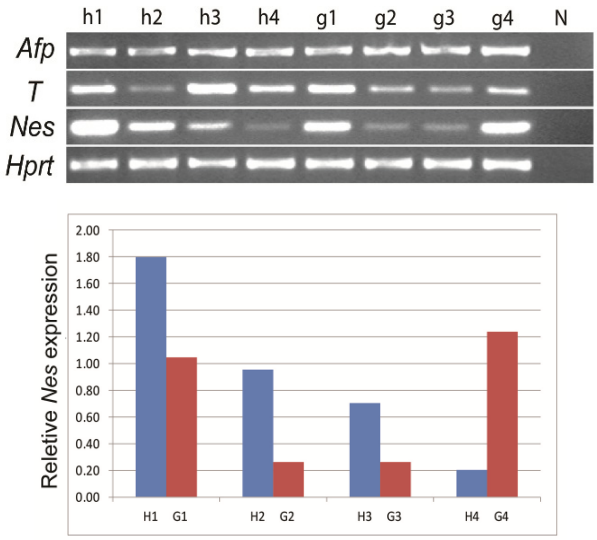


Figure III.31. Formation of embryoid body (EB) and expression of germ layer markers. (a) EBs were induced over a period of 6 days. Reverse transcriptase (RT-) PCR was then performed to look into the expression of the three major germ layer markers, as shown. Expression of Hprt served as control. N: Negative control without added template. (b) Relative expression levels of *Nes* (nestin) mRNA in HMGA2-expressing cell lines after EB formation. Quantification of RT-PCR reveals that cell lines with elevated HMGA2 levels show higher expression levels of *Nes* after EB formation. The corresponding signals acquired for *Nes* mRNA expression after ethidium bromide staining were normalized to *Hprt* mRNA

In accordance, previous publications indicate HMGA2 as a major factor required for selfrenewal of juvenile neural stem cells (Nishino et al., 2008). This indication in conjugation with GO analysis data showing that nervous system development is affected by HMGA2, illustrates that HMGA2 is involved in differentiation control towards ectodermal cell fates.

IIII. DISCUSSION

IIII.1. Induced pluripotent stem cells

Reverting terminally differentiated cells back to pluripotent stem cells by introducing several transcription factors was achieved by Yamanaka and colleagues in 2006 (Takahashi and Yamanaka, 2006). These reprogrammed cells called induced pluripotent stem cells (iPSC) have brought tremendous hope for drug discovery and therapy over the past few years. However, iPSCs still face many limitations which need to be overcome before they can replace ESCs. One of the major limitations is the pluripotency characteristics of iPSCs generated meaning that different iPSC colonies show different degree of pluripotency. One strategy to distinguish fully pluripotent cells is to select for those cells expressing essential pluripotency genes, enabling easy access to fully reprogrammed iPSCs. Identifying pluripotent cells form a bulk of cells containing non pluripotent cells can greatly facilitate the use of these cells in research and therapy (Zhao et al., 2009). In this study, we generated mouse iPSCs using the famous Yamanaka factors and used these reprogrammed cells to study our UTF1 reporter construct.

We utilized a UTF1 promoter driven EGFP construct to distinguish between pluripotent and non-pluripotent iPSCs. This construct enabled UTF1 monitoring during iPSC formation and its differentiation. UTF1 is a stringent pluripotency gene making it a great candidate for selecting highly pluripotent cells. Previous studies have shown that UTF1 increases pluripotent iPSC colonies dramatically when introduced into the cell concurrently with other four factors (Zhao et al., 2008). UTF1 which has also been a target of previous research in our laboratory has proven to be a very important factor in pluripotency (Tan et al., 2007; Zhao et al., 2008). This factor has shown to be expressed

in iPSCs derived from different species including human, mice, rat and pigs. Studies carried out on the timing of UTF1 expression during reprogramming indicated that this factor is expressed quite late during the transitional stage into iPSC. Previously known pluripotency markers like OCT4 and NANOG are expressed earlier than UTF1 which makes them less stringent. These features make UTF1 a great tool for identifying highly pluripotent stem cells. Studies by Jaenisch and colleagues indicated that cells expressing UTF1 at early stages of reprogramming are destined to form iPSCs. Hence, UTF1 acts as a definitive marker for iPSC formation, further indicating the importance of this factor in pluripotent cells (Buganim et al., 2012). By forming a link between PRC2 and myc pathways, UTF1 acts as an important regulator of pluripotency and selfrenewal (Jia et al., 2012).

In addition we looked into the oncofetal, non-histone chromatin factor HMGA2, expressed generally in pluripotent cells and during early developmental stages (Chiappetta et al., 2008; Droge and Davey, 2008; Fedele et al., 2010). This factor has previously shown to be a prominent factor in ESCs (Pfannkuche et al., 2009) , yet no studies have shown its role in iPSCs. Using iPSCs to study HMGA2 is a great model to study this factor both as a pluripotency factor and an oncogene. This approach can greatly increase our understanding of HMGA2.

IIII.2. Transient UTF1-EGFP construct

We observed that our construct can function when transiently transfected into iPSCs. Primary studies in our lab showed that this vector can be used to identify highly

pluripotent human ESCs making it feasible to be used in iPSCs likewise (Tan et al., 2007).

This is advantageous for use in human cell lines as there is low risk of integration and mutation in the genome. Similarly transient constructs containing Rex1 regulatory elements have previously been used to isolate pluripotent stem cells (Eiges et al., 2001). Furthermore iPSC lines selected with UTF1 promoter attached to neomycin had higher pluripotency and life span in compared to control iPSCs (Pfannkuche et al., 2010).

IIII.3. Integrated UTF1-EGFP construct

Cloning this reporter into pMYC (retroviral vector), showed that this reporter also works when integrated into the genome. Primary MEFs transduced with pMYC-*UTF1-EGFP* did not show any EGFP expression; however iPSC colonies with ESC morphology indicated EGFP expression upon appearance, displaying that this vector starts to express in pluripotent cells. Besides, iPSCs which did not resemble ESCs had no UTF1 expression denoting that this factor distinguishes highly pluripotent and low (or non-) pluripotent iPSCs.

We demonstrated that our reporter is functional as an integrating cassette and maintains functionality after several passages, after freeze thawing and even after long term integration before reprogramming. This indicated that our cassette is less prone to silencing in compared to for example, viral-based promoters (Takahashi and Yamanaka, 2006). These characteristics of the reporter system qualify it as versatile tool for use in iPSC research. In addition due to the small size of the vector, it may be possible to carry this cassette using adenoviral vectors or other non-integrating methods. Moreover its

small size makes it easily customizable. A notable example is that the EGFP can be virtually replaced by any sequence like a cell surface marker, which then can be used to enrich pluripotent iPSCs from differentiated cells or vice versa. Other studies have shown the use of integrating retroviral reporters for selecting highly pluripotent stem cells. For example Ellis and colleagues developed an EOS (Early Transposon promoter with Oct4 and Sox2 enhancers) lentiviral based vector distinctly expressed in both mouse and human embryonic stem cells and iPSCs but not in fibroblasts (Hotta et al., 2009b). Our integrating UE reporter can be used in high throughput screening of human cell lines for different substances which induce pluripotency. In overall we demonstrated that our reporter could act as a great tool for identifying highly pluripotent iPSC colonies, either by transfection or transduction.

IIII.4. UTF1-tomato knockin mice

We demonstrated that this construct can be used when knocked into mice and tail tips of these mice generate iPSCs with tomato expression. This is a valuable tool for research as these mice can be used to study different functions of UTF1.

For example expression of UTF1 during embryonic development has not yet been thoroughly studied. Monitoring UTF1 with this construct during embryonic development is useful for better understanding the mechanisms underlying this process, precisely indicating when UTF1 comes up and the time point it is downregulated. Furthermore various tissues in these mice can be screened for UTF1 expression easily with this method. For example previous studies have indicated that UTF1 is expressed in rat testis (van Bragt et al., 2008). Interestingly, UTF1 and another known pluripotency maker

REX-1 have been found recently to be expressed in epidermis, making them a potential marker for this cell type (Reinisch et al., 2011). This protein is also expressed in human spermatogonial cells derived from iPSCs and ESCs (Easley et al., 2012). Cells derived from these mice are also valuable sources for ESC and iPSC studies. For example mouse cervical cancer cells carry hypermethylated UTF1 promoters and express increased amounts of utf1 protein and mRNA (Guenin et al., 2012). Other cancer types like OCT4 knockin mice have previously been used extensively, however OCT4 comes up very early during the course of iPSC formation which makes its use somewhat questionable (Samavarchi-Tehrani et al., 2010). Considering that UTF1 is a sensitive pluripotency marker, it may prove to be more useful than OCT4 in detecting pluripotent cells.

Upon differentiation of our tomato-iPSCs, some colonies still preserved fluorescent activity especially in central parts. This was somewhat surprising as this expression still remained even after forced differentiation by addition of DMSO. However once we passaged these cells no fluorescent activity could be seen even where clusters of cells were seen. This indicated that cell division is necessary for these cells to differentiate. According to previous studies it seems that cellular proliferation is required for reprogramming (Hanna et al., 2009).

The knockin tomato sequence also inactivates UTF1 expression, enabling us to generate knockout UTF1 mouse. We observed that knockout homozygous UTF1 mice seem to be viable. Yet further studies are required to understand how absence of UTF1 may affect these mice. A recent study has shown that UTF1 knockout mice are smaller than normal when born in compared to normal mice, due to placental insufficiency caused by loss of

UTF1. Nevertheless iPSCs derived from these mice were pluripotent and formed teratomas (Nishimoto et al., 2013).

Downstream applications of the UTF1-EGFP construct would be to isolate pluripotent cells from non-pluripotent cells for therapeutic and research purposes. Furthermore the EGFP can be replaced with a Neomycin selection cassette which can be used to produce pure populations of highly pluripotent cells. Replacing the EGFP with a suicidal gene on the other hand can be a possibility when only differentiated cells are required without remaining pluripotent cells. In contrast to the previous approach the suicidal gene will cause the pluripotent cells to die, hence retaining a pure population of differentiated cells. As mentioned earlier the sensitivity of this reporter in compared to other reporters like Oct4 and Nanog makes this reporter a much more reliable reporter for detecting pluripotent cells for future applications.

The transient reporter can be used to detect pluripotent human iPSCs without genome integration, leaving these cells safe for clinical purposes.

One possible approach is to utilize this reporter for high throughput screening for discovering chemical compounds capable of reprogramming.

IIII.5. Role of HMGA2 in iPSCs

We looked into the developmental and ESC regulating gene HMGA2, which has not been studied in iPSCs. Previously HMGA2 has been shown to be an important factor in ESC development. Many studies have been carried out to see the role of HMGA2 in ESCs; however no study has shown the role of this important ESC factor in iPSCs.

HMGA2 is expressed during early developmental stages and declines eventually upon formation of different cell types. This implies that this factor is playing a role in pluripotent cells and early developmental stages (as it is not expressed in fully differentiated cells). Previous reports from our lab have indicated that association of HMGA2 with the chromatin imposes a global effect on the chromatin structure which defines ESC identity during pluripotency and early differentiation. Moreover HMGA2 overexpression is associated with mesenchymal and adipose differentiation from ESCs (Li et al., 2007; Li et al., 2006).

We thought to study this factor in iPSCs as it is a more effective approach of looking into its functions due to the fact that iPSC formation is a sequential process from non-pluripotent to pluripotent cells. Furthermore this factor plays an important part in epithelial to mesenchymal transition (EMT) and is modulated by *let7-lin28*, crucial ESC regulating factors (Samavarchi-Tehrani et al., 2010; Tzatsos and Bardeesy, 2008; Watanabe et al., 2009).

In this study we examined whether it's possible to generate iPSCs in conjugation with exogenous HMGA2 for better understanding the role of this chromatin factor, recognized to be associated with different molecular phenotypes and disease. Firstly our data revealed that increased HMGA2 expression does not suppress iPSC formation itself, although it slightly reduces the colony numbers. The decline in colonies may be due to previously identified role of HMGA2 on cell cycle arrest in fibroblasts yet this needs further studies to validate (Narita et al., 2006).

It is known that HMGA2 is expressed during early developmental points (Hirning-Folz et al., 1998; Li et al., 2007) which suggests that the presence of HMGA2 may be necessary

for pluripotency and differentiation. This claim is clearly supported by our data in combination with prior reports on HMGA2 which indicated that association of HMGA2 with the chromatin defines cell identity in ESCs and during early differentiation. (Li et al., 2007; Li et al., 2006).

For example, we demonstrated that MEFs which normally do not express HMGA2 start to continuously express endogenous HMGA2 to a particular degree in some stage during reprogramming. When this amount exceeds due to exogenous HMGA2 expression from the lentiviral promoter, pluripotency marker genes start to decrease, as indicated by their mRNA levels.

Interestingly, previous studies in our lab indicated that HMGA2 knockdown in human ESCs also leads to down regulation of pluripotency regulating factors like UTF1 and OCT4. This is supported by previous research showing iPSC efficiency increases in cells treated with anti-cancer drugs which downregulate HMGA2 expression (Yang et al., 2011). Together these findings show that a particular HMGA2 threshold in ESCs may be critical for modulating pluripotency. This also comes with the finding that those iPSCs which displayed higher HMGA2 levels had elevated *Nes* mRNA levels upon differentiation, yet expression of other differentiation markers like *T* and *Afp* seemed to be unaffected. Microarray data also points towards ectoderm differentiation as GO analysis shows that nervous system development is affected by HMGA2 overexpression.

Therefore it seems ectodermal differentiation is favoured by increased HMGA2 levels, though it does not exclude the fact that HMGA2 might play a role in mesodermal or endodermal differentiation.

Using HMGA2 iPSCs for *invivo* differentiation tests may gradually ease the way for more interesting prospects of this hypothesis.

Our GO analysis clearly indicates developmental and anatomical genes to be affected mostly by HMGA2. Earlier reports have unveiled HMGA2 as a significant developmental gene which regulates body height and size in different species (Battista et al., 1999; Sanna et al., 2008; Song et al., 2011; Weedon et al., 2007; Winkler et al., 2007). Our findings may be worthy for elucidating the functions of certain genes affected by HMGA2 and appear to be crucial for the respective developmental pathways. One of the other biological processes affected is cell adhesion which is very important in cancer development and invasion. This supports the concept that HMGA2 is a factor modulating invasiveness in cancer cells (Watanabe et al., 2009).

Our study indicated that crucial diabetes related gene IGF2BP2 is differentially expressed in HMGA2 overexpressed cells. Seemingly this factor which was previously known to be crucial in diabetes mellitus has recently been shown to regulate myoblast and skeletal muscle development. Recent study has shown that IGF2BP2 is regulated by HMGA2 levels. While Knockout HMGA2 mice indicated sever muscle defects due to IGF2BP2 deficiency, overexpressing IGF2BP2 partially rescued these mice (Li et al., 2012).

Our data provides the basis for further inspection into HMGA2s role in diabetes by finding key susceptibility genes regulated by HMGA2.

V. REFERENCES

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.

Arlotta, P., Tai, A.K.-F., Manfioletti, G., Clifford, C., Jay, G., and Ono, S.J. (2000). Transgenic Mice Expressing a Truncated Form of the High Mobility Group I-C Protein Develop Adiposity and an Abnormally High Prevalence of Lipomas. Journal of Biological Chemistry *275*, 14394-14400.

Bain, G., Kitchens, D., Yao, M., Huettner, J.E., and Gottlieb, D.I. (1995). Embryonic stem cells express neuronal properties in vitro. Dev Biol *168*, 342-357.

Battista, S., Fidanza, V., Fedele, M., Klein-Szanto, A.J.P., Outwater, E., Brunner, H., Santoro, M., Croce, C.M., and Fusco, A. (1999). The Expression of a Truncated HMGI-C Gene Induces Gigantism Associated with Lipomatosis. Cancer Research *59*, 4793-4797.

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.

Brants, J.R., Ayoubi, T.A., Chada, K., Marchal, K., Van de Ven, W.J., and Petit, M.M. (2004). Differential regulation of the insulin-like growth factor II mRNA-binding protein genes by architectural transcription factor HMGA2. FEBS Lett *569*, 277-283.

Buganim, Y., Faddah, Dina A., Cheng, Albert W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, Sandy L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. Cell *150*, 1209-1222.

Cao, X., Clavijo, C., Li, X., Lin, H.H., Chen, Y., Shih, H.M., and Ann, D.K. (2008). SUMOylation of HMGA2: selective destabilization of promyelocytic leukemia protein via proteasome. Molecular cancer therapeutics *7*, 923-934.

Cavaleri, F., and Scholer, H.R. (2003). Nanog: a new recruit to the embryonic stem cell orchestra. Cell *113*, 551-552.

Cepko, C., and Pear, W. (2001). Overview of the Retrovirus Transduction System. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

Chen, D., Zhao, M., and Mundy, G.R. (2004). Bone morphogenetic proteins. Growth Factors *22*, 233-241.

Cherry, Anne B.C., and Daley, George Q. (2012). Reprogramming Cellular Identity for Regenerative Medicine. Cell *148*, 1110-1122.

Chiappetta, G., Ferraro, A., Vuttariello, E., Monaco, M., Galdiero, F., De Simone, V., Califano, D., Pallante, P., Botti, G., Pezzullo, L.*, et al.* (2008). HMGA2 mRNA expression correlates with the malignant phenotype in human thyroid neoplasias. European journal of cancer *44*, 1015-1021.

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 *5*, 111-123.

Coffin, J.M., Stoye, J.P., and Frankel, W.N. (1989). Genetics of endogenous murine leukemia viruses. Ann N Y Acad Sci *567*, 39-49.

Cohen-Tannoudji, M. (2007). [Nobel Prize 2007 for Medicine to Mario Capecchi, Martin Evans and Oliver Smithies: the mutant mice to order]. Med Sci (Paris) *23*, 1159-1161.

Constantinescu, S. (2003). Stemness, fusion and renewal of hematopoietic and embryonic stem cells. J Cell Mol Med *7*, 103-112.

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.

Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M., and Walker, P.R. (2005). Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. J Biol Chem *280*, 6257-6260.

Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J Embryol Exp Morphol *87*, 27-45.

Droge, P., and Davey, C.A. (2008). Do cells let-7 determine stemness? Cell Stem Cell *2*, 8-9.

Easley, Charles A., Phillips, Bart T., McGuire, Megan M., Barringer, Jennifer M., Valli, H., Hermann, Brian P., Simerly, Calvin R., Rajkovic, A., Miki, T., Orwig, Kyle E.*, et al.* (2012). Direct Differentiation of Human Pluripotent Stem Cells into Haploid Spermatogenic Cells. Cell Reports *2*, 440-446.

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. Current biology : CB *11*, 514-518.

Ensenat-Waser, R., Santana, A., Vicente-Salar, N., Cigudosa, J.C., Roche, E., Soria, B., and Reig, J.A. (2006). Isolation and characterization of residual undifferentiated mouse embryonic stem cells from embryoid body cultures by fluorescence tracking. In Vitro Cell Dev Biol Anim *42*, 115-123.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S.*, et al.* (2009). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71-79.

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

Fedele, M., Palmieri, D., and Fusco, A. (2010). HMGA2: A pituitary tumour subtype-specific oncogene? Molecular and Cellular Endocrinology *326*, 19-24.

Fehling, H.J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., and Kouskoff, V. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. Development *130*, 4217-4227.

Filip, S., English, D., and Mokry, J. (2004). Issues in stem cell plasticity. J Cell Mol Med *8*, 572-577.

Firth, S.M., and Baxter, R.C. (2002). Cellular actions of the insulin-like growth factor binding proteins. Endocr Rev *23*, 824-854.

Fitzgerald, J., Rich, C., Zhou, F.H., and Hansen, U. (2008). Three novel collagen VI chains, alpha4(VI), alpha5(VI), and alpha6(VI). The Journal of biological chemistry *283*, 20170-20180.

Fukushima, A., Okuda, A., Nishimoto, M., Seki, N., Hori, T.A., and Muramatsu, M. (1998). Characterization of functional domains of an embryonic stem cell coactivator UTF1 which are conserved and essential for potentiation of ATF-2 activity. J Biol Chem *273*, 25840-25849.

Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc Natl Acad Sci U S A *103*, 16806-16811.

Graichen, R., Xu, X., Braam, S.R., Balakrishnan, T., Norfiza, S., Sieh, S., Soo, S.Y., Tham, S.C., Mummery, C., Colman, A*., et al.* (2008). Enhanced cardiomyogenesis of human embryonic stem cells by a small molecular inhibitor of p38 MAPK. Differentiation *76*, 357-370.

Guenin, S., Mouallif, M., Deplus, R., Lampe, X., Krusy, N., Calonne, E., Delbecque, K., Kridelka, F., Fuks, F., Ennaji, M.M*., et al.* (2012). Aberrant Promoter Methylation and Expression of UTF1 during Cervical Carcinogenesis. PLoS ONE *7*, e42704.

Han, J., Yuan, P., Yang, H., Zhang, J., Soh, B.S., Li, P., Lim, S.L., Cao, S., Tay, J., Orlov, Y.L*., et al.* (2010). Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature *463*, 1096-1100.

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

Hirning-Folz, U., Wilda, M., Rippe, V., Bullerdiek, J., and Hameister, H. (1998). The expression pattern of the Hmgic gene during development. Genes, chromosomes & cancer *23*, 350-357.

Hotta, A., Cheung, A.Y.L., Farra, N., Garcha, K., Chang, W.Y., Pasceri, P., Stanford, W.L., and Ellis, J. (2009a). EOS lentiviral vector selection system for human induced pluripotent stem cells. Nat Protocols *4*, 1828-1844.

Hotta, A., Cheung, A.Y.L., Farra, N., Vijayaragavan, K., Seguin, C.A., Draper, J.S., Pasceri, P., Maksakova, I.A., Mager, D.L., Rossant, J.*, et al.* (2009b). Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat Meth *6*, 370-376.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K.*, et al.* (2013). Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. Science *341*, 651-654.

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

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

Jia, J., Zheng, X., Hu, G., Cui, K., Zhang, J., Zhang, A., Jiang, H., Lu, B., Yates, J., Liu, C.*, et al.* (2012). Regulation of Pluripotency and Self- Renewal of ESCs through Epigenetic- Threshold Modulation and mRNA Pruning. Cell *151*, 576-589.

Joannides, A.J., Fiore-Heriche, C., Battersby, A.A., Athauda-Arachchi, P., Bouhon, I.A., Williams, L., Westmore, K., Kemp, P.J., Compston, A., Allen, N.D.*, et al.* (2007). A scaleable and defined system for generating neural stem cells from human embryonic stem cells. Stem Cells *25*, 731-737.

Jung, C.B., Moretti, A., Mederos y Schnitzler, M., Iop, L., Storch, U., Bellin, M., Dorn, T., Ruppenthal, S., Pfeiffer, S., Goedel, A.*, et al.* (2012). Dantrolene rescues arrhythmogenic RYR2 defect in a patient-specific stem cell model of catecholaminergic polymorphic ventricular tachycardia. EMBO Molecular Medicine *4*, 180-191.

Kawasaki, H., Mizuseki, K., Nishikawa, S., Kaneko, S., Kuwana, Y., Nakanishi, S., Nishikawa, S.I., and Sasai, Y. (2000). Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. Neuron *28*, 31-40.

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.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.

Kooistra, S.M., van den Boom, V., Thummer, R.P., Johannes, F., Wardenaar, R., Tesson, B.M., Veenhoff, L.M., Fusetti, F., O'Neill, L.P., Turner, B.M.*, et al.* (2010). Undifferentiated embryonic cell transcription factor 1 regulates ESC chromatin organization and gene expression. Stem Cells *28*, 1703-1714.

Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., and Keller, G. (2004). Development of definitive endoderm from embryonic stem cells in culture. Development *131*, 1651-1662.

Li, A.Y., Lin, H.H., Kuo, C.Y., Shih, H.M., Wang, C.C., Yen, Y., and Ann, D.K. (2011). High-mobility group A2 protein modulates hTERT transcription to promote tumorigenesis. Molecular and cellular biology *31*, 2605-2617.

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, O., Li, J., and Droge, P. (2007). DNA architectural factor and proto-oncogene HMGA2 regulates key developmental genes in pluripotent human embryonic stem cells. FEBS Lett *581*, 3533-3537.

Li, O., Vasudevan, D., Davey, C.A., and Droge, P. (2006). High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. Genesis *44*, 523-529.

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009). Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. Cell Stem Cell *4*, 16-19.

Li, Z., Gilbert, J.A., Zhang, Y., Zhang, M., Qiu, Q., Ramanujan, K., Shavlakadze, T., Eash, J.K., Scaramozza, A., Goddeeris, M.M.*, et al.* (2012). An HMGA2-IGF2BP2 Axis Regulates Myoblast Proliferation and Myogenesis. Dev Cell.

Liang, Y., Xiao-Gang, Z., and Xin-Wei, W. (2011). Vitamin C promotes differentiation of mouse induced pluripotent stem cells into cardiomyocyte in vitro. Heart *97*, A8-A9.

Liao, J., Wu, Z., Wang, Y., Cheng, L., Cui, C., Gao, Y., Chen, T., Rao, L., Chen, S., Jia, N.*, et al.* (2008). Enhanced efficiency of generating induced pluripotent stem (iPS) cells from human somatic cells by a combination of six transcription factors. Cell Res *18*, 600-603.

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 Meth *6*, 805-808.

Liu, S., Dontu, G., and Wicha, M.S. (2005). Mammary stem cells, self-renewal pathways, and carcinogenesis. Breast Cancer Res *7*, 86-95.

Lu, H.F., Narayanan, K., Lim, S.-X., Gao, S., Leong, M.F., and Wan, A.C.A. (2012). A 3D microfibrous scaffold for long-term human pluripotent stem cell self-renewal under chemically defined conditions. Biomaterials *33*, 2419-2430.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. Cell Stem Cell *1*, 55-70.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49-55.

Monzen, K., Ito, Y., Naito, A.T., Kasai, H., Hiroi, Y., Hayashi, D., Shiojima, I., Yamazaki, T., Miyazono, K., Asashima, M.*, et al.* (2008). A crucial role of a high mobility group protein HMGA2 in cardiogenesis. Nat Cell Biol *10*, 567-574.

Morita, S., Kojima, T., and Kitamura, T. (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther *7*, 1063-1066.

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., Gocza, E., Diaz, E.M., Prideaux, V.R., Ivanyi, E., Markkula, M., and Rossant, J. (1990). Embryonic stem cells alone are able to support fetal development in the mouse. Development *110*, 815-821.

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.

Narita, M., Krizhanovsky, V., Nunez, S., Chicas, A., Hearn, S.A., Myers, M.P., and Lowe, S.W. (2006). A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. Cell *126*, 503-514.

Ng, E.S., Azzola, L., Sourris, K., Robb, L., Stanley, E.G., and Elefanty, A.G. (2005). The primitive streak gene Mixl1 is required for efficient haematopoiesis and BMP4-induced ventral mesoderm patterning in differentiating ES cells. Development *132*, 873-884.

Nishimoto, M., Katano, M., Yamagishi, T., Hishida, T., Kamon, M., Suzuki, A., Hirasaki, M., Nabeshima, Y., Nabeshima, Y.-i., Katsura, Y.*, et al.* (2013). <italic>In Vivo</italic> Function and Evolution of the Eutherian-Specific Pluripotency Marker UTF1. PLoS ONE *8*, e68119.

Nishino, J., Kim, I., Chada, K., and Morrison, S.J. (2008). Hmga2 Promotes Neural Stem Cell Self-Renewal in Young but Not Old Mice by Reducing p16Ink4a and p19Arf Expression. Cell *135*, 227-239.

Ohshige, T., Iwata, M., Omori, S., Tanaka, Y., Hirose, H., Kaku, K., Maegawa, H., Watada, H., Kashiwagi, A., Kawamori, R.*, et al.* (2011). Association of new loci identified in European genome-wide association studies with susceptibility to type 2 diabetes in the Japanese. PLoS One *6*, e26911.

Okabe, S., Forsberg-Nilsson, K., Spiro, A.C., Segal, M., and McKay, R.D. (1996). Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. Mech Dev *59*, 89-102.

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.

Okuda, A., Fukushima, A., Nishimoto, M., Orimo, A., Yamagishi, T., Nabeshima, Y., Kuro-o, M., Boon, K., Keaveney, M., Stunnenberg, H.G*., et al.* (1998). UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra-embryonic cells. EMBO J *17*, 2019-2032.

Pease, S., Braghetta, P., Gearing, D., Grail, D., and Williams, R.L. (1990). Isolation of embryonic stem (ES) cells in media supplemented with recombinant leukemia inhibitory factor (LIF). Developmental Biology *141*, 344-352.

Pfannkuche, K., Fatima, A., Gupta, M.K., Dieterich, R., and Hescheler, J.r. (2010). Initial Colony Morphology-Based Selection for iPS Cells Derived from Adult Fibroblasts Is Substantially Improved by Temporary UTF1-Based Selection. PLoS ONE *5*, e9580.

Pfannkuche, K., Summer, H., Li, O., Hescheler, J., and Droge, P. (2009). The high mobility group protein HMGA2: a co-regulator of chromatin structure and pluripotency in stem cells? Stem Cell Rev *5*, 224-230.

Phinney, D.G., and Prockop, D.J. (2007). Concise review: mesenchymal stem/multipotent stromal cells: the state of transdifferentiation and modes of tissue repair--current views. Stem Cells *25*, 2896-2902.

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. Nature biotechnology *28*, 848-855.

Qi, X., Li, T.G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., and Zhao, G.Q. (2004). BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. Proc Natl Acad Sci U S A *101*, 6027-6032.

Rao, R.R., Johnson, A.V., and Stice, S.L. (2007). Cell Surface Markers in Human Embryonic Stem Cells. In Stem Cell Assays, pp. 51-61.

Reinisch, C.M., Mildner, M., Petzelbauer, P., and Pammer, J. (2011). Embryonic stem cell factors undifferentiated transcription factor-1 (UFT-1) and reduced expression protein-1 (REX-1) are widely expressed in human skin and may be involved in cutaneous differentiation but not in stem cell fate determination. Int J Exp Pathol *92*, 326-332.

Saha, K., and Jaenisch, R. (2009). Technical Challenges in Using Human Induced Pluripotent Stem Cells to Model Disease. Cell stem cell *5*, 584-595.

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.

Sanna, S., Jackson, A.U., Nagaraja, R., Willer, C.J., Chen, W.-M., Bonnycastle, L.L., Shen, H., Timpson, N., Lettre, G., Usala, G*., et al.* (2008). Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet *40*, 198-203.

Schenborn, E., and Groskreutz, D. (1999). Reporter gene vectors and assays. Molecular Biotechnology *13*, 29-44.

Sebastiano, V., Maeder, M.L., Angstman, J.F., Haddad, B., Khayter, C., Yeo, D.T., Goodwin, M.J., Hawkins, J.S., Ramirez, C.L., Batista, L.F.Z*., et al.* (2011). In Situ Genetic Correction of the Sickle Cell Anemia Mutation in Human Induced Pluripotent Stem Cells Using Engineered Zinc Finger Nucleases. Stem Cells *29*, 1717-1726.

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.

Song, C., Gu, X., Feng, C., Wang, Y., Gao, Y., Hu, X., and Li, N. (2011). Evaluation of SNPs in the chicken HMGA2 gene as markers for body weight gain. Animal Genetics *42*, 333-336.

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.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev *24*, 2239-2263.

Stewart, R., Yang, C., Anyfantis, G., Przyborski, S., Hole, N., Strachan, T., Stojkovic, M., Keith, W.N., Armstrong, L., and Lako, M. (2008). Silencing of the expression of pluripotent driven-reporter genes stably transfected into human pluripotent cells. Regenerative Medicine *3*, 505-522.

Summer, H., Li, O., Bao, Q., Zhan, L., Peter, S., Sathiyanathan, P., Henderson, D., Klonisch, T., Goodman, S.D., and Droge, P. (2009). HMGA2 exhibits dRP/AP site cleavage activity and protects cancer cells from DNA-damage-induced cytotoxicity during chemotherapy. Nucleic Acids Res *37*, 4371-4384.

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., 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., Wang, S.T., Hentze, H., and Droge, P. (2007). A UTF1-based selection system for stable homogeneously pluripotent human embryonic stem cell cultures. Nucleic Acids Res *35*, e118.

Tay, Y., Peter, S., Rigoutsos, I., Barahona, P., Ahmed, S., and Droge, P. (2009). Insights into the regulation of a common variant of HMGA2 associated with human height during embryonic development. Stem Cell Rev *5*, 328-333.

Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. Cell *139*, 871-890.

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.

Tzatsos, A., and Bardeesy, N. (2008). Ink4a/Arf regulation by let-7b and Hmga2: a genetic pathway governing stem cell aging. Cell Stem Cell *3*, 469-470.

van Bragt, M.P., Roepers-Gajadien, H.L., Korver, C.M., Bogerd, J., Okuda, A., Eggen, B.J., de Rooij, D.G., and van Pelt, A.M. (2008). Expression of the pluripotency marker UTF1 is restricted to a subpopulation of early A spermatogonia in rat testis. Reproduction *136*, 33-40.

van den Boom, V., Kooistra, S.M., Boesjes, M., Geverts, B., Houtsmuller, A.B., Monzen, K., Komuro, I., Essers, J., Drenth-Diephuis, L.J., and Eggen, B.J. (2007). UTF1 is a chromatin-associated protein involved in ES cell differentiation. J Cell Biol *178*, 913-924.

Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G*., et al.* (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet *42*, 579-589.

Watanabe, S., Ueda, Y., Akaboshi, S., Hino, Y., Sekita, Y., and Nakao, M. (2009). HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells. The American journal of pathology *174*, 854-868.

Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R.B., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B*., et al.* (2007). A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet *39*, 1245-1250.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. Cell *110*, 385-397.

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.

Winkler, S., Escobar, H.M., Meyer, B., Simon, D., Eberle, N., Baumgartner, W., Loeschke, S., Nolte, I., and Bullerdiek, J. (2007). HMGA2 expression in a canine model of prostate cancer. Cancer genetics and cytogenetics *177*, 98-102.

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.

Yakubov, E., Rechavi, G., Rozenblatt, S., and Givol, D. (2010). Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. Biochemical and Biophysical Research Communications *394*, 189-193.

Yang, C.S., Lopez, C.G., and Rana, T.M. (2011). Discovery of nonsteroidal anti-inflammatory drug and anticancer drug enhancing reprogramming and induced pluripotent stem cell generation. Stem Cells *29*, 1528-1536.

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.

Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R.R., Pasceri, P., and Ellis, J. (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. Mol Ther *10*, 27-36.

Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L.M., Nishikawa, S., Chiba, T., and Era, T. (2005). Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. Nature biotechnology *23*, 1542-1550.

Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T., and Yamanaka, S. (2009). Hypoxia enhances the generation of induced pluripotent stem cells. Cell Stem Cell *5*, 237-241.

Zhang, H., Berezov, A., Wang, Q., Zhang, G., Drebin, J., Murali, R., and Greene, M.I. (2007). ErbB receptors: from oncogenes to targeted cancer therapies. J Clin Invest *117*, 2051-2058.

Zhao, X.Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.L., Ma, Q.W., Wang, L*., et al.* (2009). iPS cells produce viable mice through tetraploid complementation. Nature *461*, 86-90.

Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z*., et al.* (2008). Two Supporting Factors Greatly Improve the Efficiency of Human iPSC Generation. Cell Stem Cell *3*, 475-479.

Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. (1995). Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature *376*, 771-774.

Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G*., et al.* (2011). The Lin28/let-7 axis regulates glucose metabolism. Cell *147*, 81-94.

Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q*., et al.* (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell *5*, 97-110.

Appendix

iPSCs vs ESCs complete list of gene expression analayzed by real time PCR

iPSC vs ESC					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	2.00E-05	2.47E-05	0.81	-1.23
Bxdc2	A02	5.00E-01	5.61E-01	0.89	-1.12
Cd34	A03	1.92E-03	4.16E-03	0.46	-2.17
Cd9	A04	5.56E-01	3.69E-01	1.51	1.51
Cdh5	A05	9.05E-05	3.85E-04	0.23	-4.26
Cdx2	A06	7.67E-05	5.33E-04	0.14	-6.94
Col1a1	A07	2.10E-01	4.50E-01	0.47	-2.14
Commd3	A08	5.86E-02	6.42E-02	0.91	-1.10
Crabp2	A09	6.48E-03	5.27E-03	1.23	1.23
Ddx4	A10	1.62E-02	9.00E-03	1.79	1.79
Des	A11	1.33E-02	7.14E-03	1.86	1.86
Diap2	A12	8.00E-04	8.57E-04	0.93	-1.07
Dnmt3b	B01	1.72E-01	8.99E-02	1.91	1.91
Ednrb	B02	8.78E-05	1.42E-04	0.62	-1.62
Eomes	B03	2.84E-03	2.79E-03	1.02	1.02
Fgf4	B04	2.67E-01	1.12E-01	2.39	2.39
Fgf5	B05	5.09E-04	4.51E-03	0.11	-8.86
Flt1	B06	1.11E-03	2.23E-03	0.50	-2.00
Fn1	B07	6.57E-01	4.40E-01	1.49	1.49
Foxa2	B08	2.40E-05	5.57E-04	0.04	-23.22
Foxd3	B09	2.06E-02	1.65E-02	1.25	1.25
Gabrb3	B10	1.21E-02	6.88E-03	1.76	1.76
Gal	B11	2.85E-04	9.79E-04	0.29	-3.44
Gata4	B12	1.48E-03	4.99E-04	2.98	2.98
Gata6	C01	1.16E-03	2.29E-03	0.51	-1.97
Gbx2	C02	2.88E-02	3.06E-02	0.94	-1.06
Gcg	C03	2.00E-05	4.48E-05	0.45	-2.24
Gcm1	C04	2.41E-04	4.70E-04	0.51	-1.95
Gdf3	C05	2.82E-02	2.48E-02	1.14	1.14
Grb7	C06	3.67E-02	8.05E-03	4.56	4.56
Hba-x	C07	4.28E-05	5.05E-05	0.85	-1.18

Hbb-y	C08	3.39E-05	4.83E-05	0.70	-1.43
Hck	C09	5.57E-03	1.16E-02	0.48	-2.08
Iapp	C10	8.70E-05	6.86E-05	1.27	1.27
Ifitm1	C11	2.58E-01	1.84E-01	1.40	1.40
Ifitm2	C12	2.59E-01	2.54E-01	1.02	1.02
Igf2bp2	D01	2.59E-01	1.55E-01	1.68	1.68
Il6st	D02	1.18E-01	9.33E-02	1.26	1.26
Ins2	D03	2.38E-05	5.54E-05	0.43	-2.33
Kit	D04	7.77E-02	7.53E-02	1.03	1.03
Krt1	D05	5.29E-04	5.59E-05	9.46	9.46
Lama1	D06	5.15E-02	1.21E-02	4.24	4.24
Lamb1-1	D07	7.44E-02	5.15E-02	1.44	1.44
Lamc1	D08	7.05E-02	4.04E-02	1.74	1.74
Lefty1	D09	6.05E-01	1.79E-01	3.38	3.38
Lefty2	D10	3.15E-01	7.19E-02	4.38	4.38
Lifr	D11	2.45E-02	1.68E-02	1.46	1.46
Lin28	D12	6.58E-02	5.11E-02	1.29	1.29
Myf5	E01	2.00E-05	5.77E-05	0.35	-2.88
Myod1	E02	2.06E-05	2.47E-05	0.83	-1.20
Nanog	E03	4.55E-01	2.88E-01	1.58	1.58
Nes	E04	8.45E-03	3.17E-02	0.27	-3.75
Neurod1	E05	1.89E-03	1.42E-03	1.33	1.33
Nodal	E06	2.97E-02	2.87E-02	1.03	1.03
Nog	E07	7.03E-05	1.76E-04	0.40	-2.50
Nr5a2	E08	7.05E-02	3.87E-02	1.82	1.82
Nr6a1	E09	9.89E-02	5.26E-02	1.88	1.88
Numb	E10	1.78E-02	1.07E-02	1.67	1.67
Olig2	E11	2.27E-03	2.62E-05	86.84	86.84
Pax4	E12	2.67E-05	3.37E-05	0.79	-1.26
Pax6	F01	2.92E-03	3.27E-03	0.89	-1.12
Pdx1	F02	2.38E-05	7.94E-05	0.30	-3.34
Pecam1	F03	1.57E-01	3.70E-02	4.25	4.25
Podxl	F04	2.66E-02	1.31E-02	2.04	2.04
Pou5f1	F05	1.96E+00	1.72E+00	1.14	1.14
Pten	F06	2.32E-01	1.86E-01	1.25	1.25
Ptf1a	F07	3.25E-05	5.61E-05	0.58	-1.73
Rest	F08	5.36E-01	1.88E-01	2.84	2.84
Runx2	F09	2.91E-04	7.82E-04	0.37	-2.69
Sema3a	F10	1.97E-03	3.70E-03	0.53	-1.88
Serpina1a	F11	4.36E-05	4.21E-05	1.03	1.03
Sfrp2	F12	1.53E-02	1.39E-02	1.10	1.10
Sox17	G01	2.00E-05	1.18E-04	0.17	-5.92
Sox2	G02	3.02E-01	2.31E-01	1.31	1.31
Sst	G03	5.85E-05	2.27E-04	0.26	-3.89

Sycp3	G04	2.15E-02	4.01E-02	0.53	-1.87
T	G05	1.83E-04	2.11E-02	0.01	-115.48
Tat	G06	1.76E-04	1.16E-04	1.51	1.51
Tcfcp2l1	G07	1.49E-01	1.81E-01	0.82	-1.22
Tdgf1	G08	4.12E-01	1.19E-01	3.46	3.46
Tert	G09	1.24E-02	4.06E-03	3.06	3.06
Utf1	G10	1.08E+00	5.48E-01	1.96	1.96
Wt1	G11	2.13E-03	4.47E-03	0.48	-2.10
Zfp42	G12	5.18E-01	9.18E-01	0.56	-1.77
Gusb	H01	1.92E-02	2.07E-02	0.93	-1.08
Hprt1	H02	5.44E-01	4.51E-01	1.21	1.21
Hsp90ab	H03	7.97E+00	7.46E+00	1.07	1.07
Gapdh	H04	3.74E+00	3.20E+00	1.17	1.17
Actb	H05	3.22E+00	4.50E+00	0.72	-1.40

Complete list of gene expression for GUE cell lines after 0, 6 and 12 days of differentiation

GUE-1 (0 days vs 6 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	1.29E-03	3.25E-05	39.55	39.55
Bxdc2	A02	2.36E-01	2.34E-01	1.01	1.01
Cd34	A03	7.22E-04	1.40E-03	0.51	-1.94
Cd9	A04	3.94E-01	4.60E-01	0.85	-1.17
Cdh5	A05	6.75E-04	6.19E-05	10.90	10.90
Cdx2	A06	9.56E-03	2.52E-04	37.89	37.89
Col1a1	A07	2.71E-02	1.73E-01	0.16	-6.38
Commd3	A08	6.35E-02	4.73E-02	1.34	1.34
Crabp2	A09	7.63E-03	7.76E-03	0.98	-1.02
Ddx4	A10	6.27E-03	1.64E-02	0.38	-2.62
Des	A11	5.42E-03	1.21E-02	0.45	-2.24
Diap2	A12	3.64E-04	4.35E-04	0.84	-1.19
Dnmt3b	B01	1.82E-01	1.24E-01	1.47	1.47
Ednrb	B02	1.60E-03	3.47E-04	4.61	4.61
Eomes	B03	2.75E-02	2.50E-03	11.01	11.01
Fgf4	B04	6.66E-02	2.09E-01	0.32	-3.15
Fgf5	B05	9.27E-03	3.48E-04	26.67	26.67
Flt1	B06	1.59E-02	6.70E-04	23.67	23.67
Fn1	B07	7.17E-01	5.40E-01	1.33	1.33
Foxa2	B08	3.76E-03	3.73E-05	100.60	100.60
Foxd3	B09	1.38E-02	1.60E-02	0.86	-1.17
Gabrb3	B10	5.17E-03	1.18E-02	0.44	-2.27
Gal	B11	1.11E-03	2.94E-04	3.77	3.77
Gata4	B12	1.18E-02	1.73E-03	6.83	6.83
Gata6	C01	1.19E-02	2.06E-03	5.76	5.76
Gbx2	C02	2.47E-02	1.88E-02	1.32	1.32
Gcg	C03	9.73E-05	4.84E-05	2.01	2.01
Gcm1	C04	4.05E-04	3.15E-04	1.29	1.29
Gdf3	C05	3.85E-02	4.73E-02	0.81	-1.23
Grb7	C06	5.36E-02	3.66E-02	1.46	1.46
Hba-x	C07	5.74E-03	2.09E-04	27.46	27.46
Hbb-y	C08	6.94E-04	5.70E-05	12.18	12.18
Hck	C09	9.80E-04	7.15E-03	0.14	-7.29
Iapp	C10	2.69E-04	5.04E-05	5.32	5.32
Ifitm1	C11	2.67E-01	1.53E-01	1.75	1.75
Ifitm2	C12	7.56E-02	2.26E-01	0.33	-2.99

Igf2bp2	D01	1.76E-01	1.35E-01	1.30	1.30
Il6st	D02	1.33E-01	1.01E-01	1.31	1.31
Ins2	D03	3.77E-04	1.55E-04	2.43	2.43
Kit	D04	1.90E-02	1.07E-01	0.18	-5.64
Krt1	D05	2.77E-04	7.04E-04	0.39	-2.55
Lama1	D06	7.49E-03	5.76E-02	0.13	-7.70
Lamb1-1	D07	1.40E-01	7.45E-02	1.88	1.88
Lamc1	D08	6.54E-02	9.58E-02	0.68	-1.46
Lefty1	D09	3.74E-01	1.80E-01	2.08	2.08
Lefty2	D10	2.74E-01	2.26E-01	1.22	1.22
Lifr	D11	1.53E-02	2.07E-02	0.74	-1.35
Lin28	D12	9.20E-02	6.80E-02	1.35	1.35
Myf5	E01	6.31E-05	4.76E-05	1.33	1.33
Myod1	E02	1.54E-04	6.62E-05	2.33	2.33
Nanog	E03	2.58E-01	5.67E-01	0.46	-2.19
Nes	E04	3.13E-02	6.47E-03	4.83	4.83
Neurod1	E05	1.32E-04	1.40E-03	0.09	-10.60
Nodal	E06	3.32E-02	3.79E-02	0.88	-1.14
Nog	E07	5.50E-03	2.17E-04	25.39	25.39
Nr5a2	E08	1.61E-02	4.85E-02	0.33	-3.01
Nr6a1	E09	1.15E-01	8.56E-02	1.35	1.35
Numb	E10	2.34E-02	1.64E-02	1.43	1.43
Olig2	E11	6.13E-03	4.40E-03	1.39	1.39
Pax4	E12	9.13E-05	3.25E-05	2.81	2.81
Pax6	F01	1.03E-02	9.89E-04	10.37	10.37
Pdx1	F02	1.70E-04	4.22E-05	4.04	4.04
Pecam1	F03	1.87E-02	1.86E-01	0.10	-9.92
Podxl	F04	2.27E-01	2.23E-02	10.16	10.16
Pou5f1	F05	1.82E+00	1.90E+00	0.96	-1.04
Pten	F06	2.50E-01	1.94E-01	1.29	1.29
Ptf1a	F07	1.56E-04	8.66E-05	1.80	1.80
Rest	F08	2.58E-01	4.76E-01	0.54	-1.85
Runx2	F09	3.27E-04	5.61E-04	0.58	-1.71
Sema3a	F10	1.90E-03	4.86E-03	0.39	-2.56
Serpina1a	F11	7.58E-04	9.49E-05	7.99	7.99
Sfrp2	F12	1.23E-02	7.49E-03	1.65	1.65
Sox17	G01	8.24E-04	5.76E-05	14.31	14.31
Sox2	G02	1.48E-01	2.56E-01	0.58	-1.74
Sst	G03	3.68E-04	2.05E-04	1.80	1.80
Sycp3	G04	7.52E-03	4.52E-02	0.17	-6.01
T	G05	6.62E-02	7.74E-04	85.45	85.45
Tat	G06	3.03E-04	2.14E-04	1.42	1.42
Tcfcp2l1	G07	5.75E-02	1.64E-01	0.35	-2.86
Tdgf1	G08	5.35E-01	3.91E-01	1.37	1.37
Tert	G09	6.45E-03	1.15E-02	0.56	-1.78
Utf1	G10	4.50E-01	6.40E-01	0.70	-1.42
Wt1	G11	4.71E-03	1.76E-03	2.67	2.67

Zfp42	G12	1.45E-01	3.66E-01	0.40	-2.52
Gusb	H01	2.40E-02	3.87E-02	0.62	-1.61
Hprt1	H02	4.13E-01	3.72E-01	1.11	1.11
Hsp90ab1	H03	6.21E+00	5.89E+00	1.05	1.05
Gapdh	H04	3.81E+00	4.19E+00	0.91	-1.10
Actb	H05	4.27E+00	2.82E+00	1.52	1.52

GUE-2 (0 days vs 6 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	8.15E-04	4.53E-05	17.98	17.98
Bxdc2	A02	3.17E-01	2.96E-01	1.07	1.07
Cd34	A03	1.34E-03	2.11E-03	0.64	-1.57
Cd9	A04	4.98E-01	4.88E-01	1.02	1.02
Cdh5	A05	1.40E-03	8.28E-05	16.95	16.95
Cdx2	A06	3.55E-03	5.06E-04	7.01	7.01
Col1a1	A07	5.70E-03	2.47E-01	0.02	-43.27
Commd3	A08	8.86E-02	5.74E-02	1.54	1.54
Crabp2	A09	9.66E-03	8.55E-03	1.13	1.13
Ddx4	A10	1.19E-02	1.68E-02	0.71	-1.41
Des	A11	7.01E-03	1.09E-02	0.64	-1.56
Diap2	A12	4.54E-04	4.29E-04	1.06	1.06
Dnmt3b	B01	1.43E-01	1.10E-01	1.29	1.29
Ednrb	B02	8.28E-04	4.96E-04	1.67	1.67
Eomes	B03	2.16E-02	2.96E-03	7.28	7.28
Fgf4	B04	6.97E-02	1.94E-01	0.36	-2.79
Fgf5	B05	1.08E-02	2.90E-04	37.14	37.14
Flt1	B06	1.31E-02	6.04E-04	21.62	21.62
Fn1	B07	3.46E-01	5.56E-01	0.62	-1.61
Foxa2	B08	2.63E-03	7.34E-05	35.81	35.81
Foxd3	B09	1.58E-02	2.31E-02	0.69	-1.46
Gabrb3	B10	6.56E-03	1.12E-02	0.59	-1.70
Gal	B11	1.88E-03	2.56E-04	7.35	7.35
Gata4	B12	3.72E-03	2.08E-03	1.79	1.79
Gata6	C01	4.45E-03	1.90E-03	2.35	2.35
Gbx2	C02	3.72E-02	2.99E-02	1.24	1.24
Gcg	C03	3.98E-04	5.23E-05	7.60	7.60
Gcm1	C04	9.60E-04	2.09E-04	4.59	4.59
Gdf3	C05	2.83E-02	5.32E-02	0.53	-1.88
Grb7	C06	3.99E-02	4.16E-02	0.96	-1.04
Hba-x	C07	1.02E-03	2.95E-04	3.47	3.47
Hbb-y	C08	3.33E-04	1.01E-04	3.30	3.30

Hck	C09	1.72E-03	7.12E-03	0.24	-4.14
Iapp	C10	4.27E-04	7.34E-05	5.82	5.82
Ifitm1	C11	2.53E-01	1.48E-01	1.71	1.71
Ifitm2	C12	6.12E-02	1.48E-01	0.41	-2.42
Igf2bp2	D01	1.59E-01	1.42E-01	1.11	1.11
Il6st	D02	1.03E-01	1.19E-01	0.86	-1.16
Ins2	D03	5.43E-04	1.44E-04	3.78	3.78
Kit	D04	8.62E-03	8.76E-02	0.10	-10.16
Krt1	D05	1.13E-03	6.44E-04	1.76	1.76
Lama1	D06	8.31E-03	6.71E-02	0.12	-8.08
Lamb1-1	D07	9.44E-02	6.78E-02	1.39	1.39
Lamc1	D08	6.04E-02	8.99E-02	0.67	-1.49
Lefty1	D09	5.89E-01	1.70E-01	3.47	3.47
Lefty2	D10	1.43E-01	2.06E-01	0.69	-1.44
Lifr	D11	1.03E-02	1.91E-02	0.54	-1.85
Lin28	D12	6.74E-02	6.29E-02	1.07	1.07
Myf5	E01	1.95E-04	3.73E-05	5.23	5.23
Myod1	E02	2.15E-04	6.11E-05	3.53	3.53
Nanog	E03	1.78E-01	4.47E-01	0.40	-2.51
Nes	E04	2.95E-02	4.31E-03	6.84	6.84
Neurod1	E05	4.53E-04	2.14E-03	0.21	-4.72
Nodal	E06	3.32E-02	3.70E-02	0.90	-1.12
Nog	E07	4.18E-03	2.60E-04	16.07	16.07
Nr5a2	E08	1.46E-02	5.14E-02	0.29	-3.51
Nr6a1	E09	1.03E-01	1.02E-01	1.02	1.02
Numb	E10	1.66E-02	1.49E-02	1.12	1.12
Olig2	E11	6.18E-03	4.73E-03	1.31	1.31
Pax4	E12	3.52E-04	5.15E-05	6.83	6.83
Pax6	F01	4.44E-03	9.33E-04	4.75	4.75
Pdx1	F02	4.64E-04	4.23E-05	10.97	10.97
Pecam1	F03	2.18E-02	1.20E-01	0.18	-5.48
Podxl	F04	1.48E-01	1.56E-02	9.48	9.48
Pou5f1	F05	1.84E+00	2.30E+00	0.80	-1.25
Pten	F06	2.12E-01	1.97E-01	1.07	1.07
Ptf1a	F07	5.35E-04	1.10E-04	4.88	4.88
Rest	F08	2.43E-01	3.88E-01	0.63	-1.60
Runx2	F09	7.11E-04	7.40E-04	0.96	-1.04
Sema3a	F10	8.75E-04	5.68E-03	0.15	-6.49
Serpina1a	F11	1.76E-03	8.53E-05	20.65	20.65
Sfrp2	F12	4.21E-03	7.29E-03	0.58	-1.73
Sox17	G01	5.25E-04	7.63E-05	6.88	6.88
Sox2	G02	9.13E-02	2.27E-01	0.40	-2.48
Sst	G03	6.18E-04	2.48E-04	2.50	2.50
Sycp3	G04	1.83E-02	4.00E-02	0.46	-2.18
T	G05	8.22E-02	1.29E-03	63.68	63.68
Tat	G06	5.77E-04	2.80E-04	2.06	2.06
Tcfcp2l1	G07	4.15E-02	1.91E-01	0.22	-4.62

Tdgf1	G08	4.52E-01	2.94E-01	1.54	1.54
Tert	G09	6.30E-03	9.98E-03	0.63	-1.58
Utf1	G10	4.51E-01	7.10E-01	0.63	-1.58
Wt1	G11	5.62E-03	1.50E-03	3.75	3.75
Zfp42	G12	1.17E-01	3.85E-01	0.30	-3.28
Gusb	H01	2.70E-02	4.40E-02	0.61	-1.63
Hprt1	H02	4.51E-01	3.61E-01	1.25	1.25
Hsp90ab1	H03	7.45E+00	6.69E+00	1.11	1.11
Gapdh	H04	3.36E+00	4.07E+00	0.83	-1.21
Actb	H05	3.29E+00	2.31E+00	1.42	1.42

GUE-3 (0 days vs 6 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	1.55E-03	3.70E-05	41.80	41.80
Bxdc2	A02	1.84E-01	2.75E-01	0.67	-1.49
Cd34	A03	2.27E-03	2.91E-03	0.78	-1.28
Cd9	A04	5.64E-01	4.53E-01	1.25	1.25
Cdh5	A05	4.27E-03	9.76E-05	43.73	43.73
Cdx2	A06	1.69E-02	4.48E-04	37.70	37.70
Col1a1	A07	4.45E-02	5.25E-01	0.08	-11.82
Commd3	A08	1.77E-01	5.99E-02	2.95	2.95
Crabp2	A09	5.87E-02	7.97E-03	7.37	7.37
Ddx4	A10	2.28E-03	1.51E-02	0.15	-6.61
Des	A11	5.68E-03	1.45E-02	0.39	-2.55
Diap2	A12	9.07E-04	8.28E-04	1.10	1.10
Dnmt3b	B01	5.44E-02	1.40E-01	0.39	-2.57
Ednrb	B02	5.28E-03	2.79E-04	18.90	18.90
Eomes	B03	5.20E-03	2.05E-03	2.54	2.54
Fgf4	B04	1.84E-02	2.23E-01	0.08	-12.08
Fgf5	B05	3.47E-03	4.92E-04	7.04	7.04
Flt1	B06	2.54E-02	1.22E-03	20.85	20.85
Fn1	B07	7.50E-01	8.70E-01	0.86	-1.16
Foxa2	B08	1.43E-02	6.68E-05	214.50	214.50
Foxd3	B09	4.19E-03	1.92E-02	0.22	-4.59
Gabrb3	B10	5.95E-03	9.64E-03	0.62	-1.62
Gal	B11	1.38E-03	2.95E-04	4.70	4.70
Gata4	B12	1.22E-02	1.89E-03	6.47	6.47
Gata6	C01	3.59E-02	2.02E-03	17.78	17.78
Gbx2	C02	1.86E-02	2.54E-02	0.74	-1.36
Gcg	C03	5.74E-04	3.79E-05	15.14	15.14
Gcm1	C04	8.22E-04	2.81E-04	2.92	2.92
Gdf3	C05	8.28E-03	5.14E-02	0.16	-6.21

Grb7	C06	3.83E-02	4.75E-02	0.81	-1.24
Hba-x	C07	3.86E-02	1.86E-04	206.99	206.99
Hbb-y	C08	1.02E-02	8.75E-05	116.32	116.32
Hck	C09	1.48E-03	3.01E-03	0.49	-2.03
Iapp	C10	9.92E-04	6.04E-05	16.42	16.42
Ifitm1	C11	9.15E-02	1.52E-01	0.60	-1.66
Ifitm2	C12	6.78E-02	1.62E-01	0.42	-2.39
Igf2bp2	D01	2.48E-01	1.10E-01	2.26	2.26
Il6st	D02	1.81E-01	1.28E-01	1.41	1.41
Ins2	D03	1.17E-03	1.33E-04	8.80	8.80
Kit	D04	1.83E-02	8.90E-02	0.21	-4.85
Krt1	D05	7.53E-04	6.35E-04	1.19	1.19
Lama1	D06	3.90E-02	5.19E-02	0.75	-1.33
Lamb1-1	D07	2.78E-01	1.04E-01	2.66	2.66
Lamc1	D08	1.38E-01	1.20E-01	1.16	1.16
Lefty1	D09	6.33E-02	1.67E-01	0.38	-2.64
Lefty2	D10	5.17E-02	2.08E-01	0.25	-4.03
Lifr	D11	2.94E-02	2.18E-02	1.35	1.35
Lin28	D12	1.44E-01	7.47E-02	1.93	1.93
Myf5	E01	6.40E-04	5.96E-05	10.73	10.73
Myod1	E02	3.45E-04	7.52E-05	4.59	4.59
Nanog	E03	7.13E-02	5.21E-01	0.14	-7.30
Nes	E04	4.88E-02	6.64E-03	7.35	7.35
Neurod1	E05	5.72E-04	3.07E-03	0.19	-5.36
Nodal	E06	1.50E-02	4.45E-02	0.34	-2.96
Nog	E07	9.44E-03	2.94E-04	32.08	32.08
Nr5a2	E08	6.77E-03	5.18E-02	0.13	-7.64
Nr6a1	E09	1.77E-01	8.89E-02	1.99	1.99
Numb	E10	4.24E-02	1.78E-02	2.38	2.38
Olig2	E11	2.96E-03	5.17E-03	0.57	-1.75
Pax4	E12	8.74E-04	4.70E-05	18.62	18.62
Pax6	F01	7.36E-03	1.09E-03	6.75	6.75
Pdx1	F02	8.79E-04	1.28E-04	6.84	6.84
Pecam1	F03	7.46E-03	1.08E-01	0.07	-14.49
Podxl	F04	3.58E-01	2.07E-02	17.30	17.30
Pou5f1	F05	4.26E-01	2.00E+00	0.21	-4.69
Pten	F06	3.01E-01	1.87E-01	1.61	1.61
Ptf1a	F07	8.57E-04	6.81E-05	12.58	12.58
Rest	F08	1.98E-01	3.89E-01	0.51	-1.97
Runx2	F09	1.48E-03	1.13E-03	1.31	1.31
Sema3a	F10	9.77E-03	6.31E-03	1.55	1.55
Serpina1a	F11	2.69E-03	1.29E-04	20.80	20.80
Sfrp2	F12	8.58E-03	1.01E-02	0.85	-1.18
Sox17	G01	1.25E-03	9.37E-05	13.35	13.35
Sox2	G02	5.54E-02	3.16E-01	0.18	-5.70
Sst	G03	1.43E-03	2.04E-04	7.02	7.02
Sycp3	G04	1.34E-03	3.62E-02	0.04	-26.91

T	G05	4.09E-02	9.37E-04	43.67	43.67
Tat	G06	1.06E-03	2.63E-04	4.03	4.03
Tcfcp2l1	G07	3.52E-02	2.09E-01	0.17	-5.94
Tdgf1	G08	7.63E-02	3.01E-01	0.25	-3.95
Tert	G09	7.00E-03	1.18E-02	0.59	-1.69
Utf1	G10	8.35E-02	5.80E-01	0.14	-6.95
Wt1	G11	5.46E-03	2.10E-03	2.60	2.60
Zfp42	G12	4.62E-02	3.58E-01	0.13	-7.75
Gusb	H01	2.38E-02	3.85E-02	0.62	-1.62
Hprt1	H02	2.87E-01	3.48E-01	0.83	-1.21
Hsp90ab1	H03	6.54E+00	6.22E+00	1.05	1.05
Gapdh	H04	3.47E+00	4.10E+00	0.85	-1.18
Actb	H05	6.44E+00	2.93E+00	2.20	2.20

GUE-4 (0 days vs 6 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	1.25E-03	4.18E-05	29.89	29.89
Bxdc2	A02	2.65E-01	3.81E-01	0.70	-1.44
Cd34	A03	1.32E-03	1.30E-03	1.02	1.02
Cd9	A04	5.98E-01	4.09E-01	1.46	1.46
Cdh5	A05	9.04E-03	4.18E-05	215.92	215.92
Cdx2	A06	5.87E-03	4.77E-04	12.30	12.30
Col1a1	A07	3.92E-02	1.46E-01	0.27	-3.72
Commd3	A08	1.29E-01	3.71E-02	3.47	3.47
Crabp2	A09	7.20E-03	1.31E-02	0.55	-1.82
Ddx4	A10	8.68E-03	3.21E-02	0.27	-3.70
Des	A11	5.34E-03	1.79E-02	0.30	-3.35
Diap2	A12	4.63E-04	1.84E-04	2.52	2.52
Dnmt3b	B01	1.55E-01	9.80E-02	1.58	1.58
Ednrb	B02	1.51E-03	5.06E-04	2.98	2.98
Eomes	B03	1.76E-02	2.22E-03	7.96	7.96
Fgf4	B04	6.12E-02	1.45E-01	0.42	-2.37
Fgf5	B05	3.63E-03	1.68E-04	21.57	21.57
Flt1	B06	4.12E-02	5.83E-04	70.62	70.62
Fn1	B07	5.50E-01	5.63E-01	0.98	-1.02
Foxa2	B08	2.85E-02	7.46E-05	381.60	381.60
Foxd3	B09	2.93E-02	1.04E-02	2.81	2.81
Gabrb3	B10	3.21E-03	4.72E-03	0.68	-1.47
Gal	B11	6.51E-04	2.81E-04	2.32	2.32
Gata4	B12	1.46E-01	3.40E-03	42.96	42.96
Gata6	C01	5.68E-02	7.00E-04	81.10	81.10
Gbx2	C02	1.31E-02	2.43E-02	0.54	-1.85

Gcg	C03	4.17E-04	4.67E-05	8.94	8.94
Gcm1	C04	8.31E-04	3.08E-04	2.70	2.70
Gdf3	C05	3.89E-02	6.12E-02	0.64	-1.57
Grb7	C06	5.55E-02	5.44E-02	1.02	1.02
Hba-x	C07	7.55E-04	3.51E-04	2.15	2.15
Hbb-y	C08	3.42E-04	1.08E-04	3.16	3.16
Hck	C09	9.52E-04	1.10E-02	0.09	-11.53
Iapp	C10	4.08E-04	6.08E-05	6.72	6.72
Ifitm1	C11	1.09E-01	1.89E-01	0.58	-1.73
Ifitm2	C12	3.79E-02	2.06E-01	0.18	-5.42
Igf2bp2	D01	1.10E-01	4.33E-02	2.55	2.55
Il6st	D02	1.32E-01	1.38E-01	0.96	-1.04
Ins2	D03	4.78E-04	1.90E-04	2.51	2.51
Kit	D04	1.09E-01	3.97E-02	2.74	2.74
Krt1	D05	7.27E-04	1.11E-03	0.66	-1.52
Lama1	D06	1.36E+00	2.91E-02	46.69	46.69
Lamb1-1	D07	1.78E+00	5.01E-02	35.48	35.48
Lamc1	D08	6.25E-01	5.91E-02	10.59	10.59
Lefty1	D09	3.69E-01	7.11E-01	0.52	-1.92
Lefty2	D10	2.57E-01	2.42E-01	1.06	1.06
Lifr	D11	4.23E-02	1.99E-02	2.12	2.12
Lin28	D12	8.39E-02	6.01E-02	1.40	1.40
Myf5	E01	3.50E-04	6.42E-05	5.46	5.46
Myod1	E02	2.05E-04	6.92E-05	2.97	2.97
Nanog	E03	8.76E-02	3.27E-01	0.27	-3.73
Nes	E04	5.32E-02	7.23E-03	7.36	7.36
Neurod1	E05	7.03E-04	9.12E-04	0.77	-1.30
Nodal	E06	6.02E-02	3.63E-02	1.66	1.66
Nog	E07	2.49E-02	1.75E-04	142.77	142.77
Nr5a2	E08	1.00E-02	6.02E-02	0.17	-6.00
Nr6a1	E09	8.39E-03	1.07E-01	0.08	-12.81
Numb	E10	2.17E-02	1.68E-02	1.29	1.29
Olig2	E11	9.08E-03	1.80E-03	5.05	5.05
Pax4	E12	4.14E-04	5.72E-05	7.23	7.23
Pax6	F01	9.89E-03	4.53E-03	2.18	2.18
Pdx1	F02	5.27E-04	5.52E-05	9.54	9.54
Pecam1	F03	2.57E-02	1.41E-01	0.18	-5.49
Podxl	F04	1.57E-01	6.93E-03	22.69	22.69
Pou5f1	F05	2.59E+00	2.44E+00	1.06	1.06
Pten	F06	3.15E-01	2.01E-01	1.56	1.56
Ptf1a	F07	8.09E-04	6.44E-05	12.58	12.58
Rest	F08	2.04E-01	3.43E-01	0.60	-1.68
Runx2	F09	1.09E-03	2.26E-04	4.84	4.84
Sema3a	F10	8.20E-04	2.11E-03	0.39	-2.58
Serpina1a	F11	1.93E-03	1.04E-04	18.51	18.51
Sfrp2	F12	5.15E-03	2.82E-03	1.82	1.82
Sox17	G01	1.25E-02	4.18E-05	299.57	299.57

Sox2	G02	7.35E-02	1.72E-01	0.43	-2.34
Sst	G03	6.34E-04	3.20E-04	1.98	1.98
Sycp3	G04	1.16E-02	6.36E-02	0.18	-5.47
T	G05	9.24E-02	2.25E-03	41.01	41.01
Tat	G06	1.40E-03	2.11E-04	6.62	6.62
Tcfcp2l1	G07	1.88E-02	1.37E-01	0.14	-7.31
Tdgf1	G08	3.85E-01	3.54E-01	1.09	1.09
Tert	G09	7.89E-03	9.94E-03	0.79	-1.26
Utf1	G10	2.96E-01	6.67E-01	0.44	-2.26
Wt1	G11	1.08E-02	1.38E-03	7.82	7.82
Zfp42	G12	7.19E-02	3.78E-01	0.19	-5.26
Gusb	H01	2.83E-02	3.36E-02	0.84	-1.19
Hprt1	H02	4.56E-01	5.96E-01	0.77	-1.31
Hsp90ab1	H03	6.52E+00	8.71E+00	0.75	-1.34
Gapdh	H04	3.45E+00	3.36E+00	1.03	1.03
Actb	H05	3.44E+00	1.71E+00	2.02	2.02

GUE-1 (0 days vs 12 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	4.45E-05	3.25E-05	1.37	1.37
Bxdc2	A02	1.76E-01	2.34E-01	0.75	-1.33
Cd34	A03	5.02E-04	1.40E-03	0.36	-2.79
Cd9	A04	4.50E-01	4.60E-01	0.98	-1.02
Cdh5	A05	2.75E-04	6.19E-05	4.44	4.44
Cdx2	A06	8.81E-04	2.52E-04	3.49	3.49
Col1a1	A07	4.22E-03	1.73E-01	0.02	-41.06
Commd3	A08	5.95E-02	4.73E-02	1.26	1.26
Crabp2	A09	3.17E-03	7.76E-03	0.41	-2.45
Ddx4	A10	1.44E-02	1.64E-02	0.88	-1.14
Des	A11	9.21E-03	1.21E-02	0.76	-1.32
Diap2	A12	7.61E-05	4.35E-04	0.17	-5.71
Dnmt3b	B01	9.60E-02	1.24E-01	0.78	-1.29
Ednrb	B02	1.98E-04	3.47E-04	0.57	-1.75
Eomes	B03	5.52E-03	2.50E-03	2.21	2.21
Fgf4	B04	1.14E-01	2.09E-01	0.54	-1.84
Fgf5	B05	4.06E-03	3.48E-04	11.69	11.69
Flt1	B06	7.15E-03	6.70E-04	10.68	10.68
Fn1	B07	2.53E-01	5.40E-01	0.47	-2.13
Foxa2	B08	5.36E-04	3.73E-05	14.35	14.35
Foxd3	B09	1.12E-02	1.60E-02	0.70	-1.43
Gabrb3	B10	4.58E-03	1.18E-02	0.39	-2.57
Gal	B11	2.51E-04	2.94E-04	0.85	-1.17

Gata4	B12	1.25E-03	1.73E-03	0.72	-1.39
Gata6	C01	1.89E-03	2.06E-03	0.92	-1.09
Gbx2	C02	2.53E-02	1.88E-02	1.35	1.35
Gcg	C03	4.14E-05	4.84E-05	0.86	-1.17
Gcm1	C04	3.55E-04	3.15E-04	1.13	1.13
Gdf3	C05	2.93E-02	4.73E-02	0.62	-1.61
Grb7	C06	3.70E-02	3.66E-02	1.01	1.01
Hba-x	C07	8.43E-05	2.09E-04	0.40	-2.48
Hbb-y	C08	5.51E-05	5.70E-05	0.97	-1.03
Hck	C09	2.53E-03	7.15E-03	0.35	-2.82
Iapp	C10	4.69E-05	5.04E-05	0.93	-1.08
Ifitm1	C11	3.03E-01	1.53E-01	1.98	1.98
Ifitm2	C12	9.04E-02	2.26E-01	0.40	-2.50
Igf2bp2	D01	1.35E-01	1.35E-01	1.00	-1.00
Il6st	D02	1.12E-01	1.01E-01	1.11	1.11
Ins2	D03	7.00E-05	1.55E-04	0.45	-2.22
Kit	D04	1.06E-02	1.07E-01	0.10	-10.13
Krt1	D05	3.15E-04	7.04E-04	0.45	-2.24
Lama1	D06	1.26E-03	5.76E-02	0.02	-45.67
Lamb1-1	D07	9.15E-02	7.45E-02	1.23	1.23
Lamc1	D08	5.97E-02	9.58E-02	0.62	-1.61
Lefty1	D09	4.56E-01	1.80E-01	2.53	2.53
Lefty2	D10	1.03E-01	2.26E-01	0.46	-2.19
Lifr	D11	1.03E-02	2.07E-02	0.50	-2.00
Lin28	D12	4.72E-02	6.80E-02	0.69	-1.44
Myf5	E01	4.73E-05	4.76E-05	0.99	-1.01
Myod1	E02	4.14E-05	6.62E-05	0.63	-1.60
Nanog	E03	3.96E-01	5.67E-01	0.70	-1.43
Nes	E04	1.04E-02	6.47E-03	1.61	1.61
Neurod1	E05	1.97E-04	1.40E-03	0.14	-7.11
Nodal	E06	3.96E-02	3.79E-02	1.04	1.04
Nog	E07	3.81E-03	2.17E-04	17.59	17.59
Nr5a2	E08	1.01E-02	4.85E-02	0.21	-4.80
Nr6a1	E09	9.87E-02	8.56E-02	1.15	1.15
Numb	E10	1.72E-02	1.64E-02	1.05	1.05
Olig2	E11	9.79E-03	4.40E-03	2.22	2.22
Pax4	E12	4.14E-05	3.25E-05	1.27	1.27
Pax6	F01	2.87E-03	9.89E-04	2.90	2.90
Pdx1	F02	4.14E-05	4.22E-05	0.98	-1.02
Pecam1	F03	3.20E-02	1.86E-01	0.17	-5.80
Podxl	F04	1.14E-01	2.23E-02	5.12	5.12
Pou5f1	F05	1.75E+00	1.90E+00	0.92	-1.09
Pten	F06	1.93E-01	1.94E-01	0.99	-1.01
Ptf1a	F07	4.86E-05	8.66E-05	0.56	-1.78
Rest	F08	2.19E-01	4.76E-01	0.46	-2.17
Runx2	F09	1.42E-04	5.61E-04	0.25	-3.94
Sema3a	F10	2.43E-04	4.86E-03	0.05	-19.98

Serpina1a	F11	4.72E-05	9.49E-05	0.50	-2.01
Sfrp2	F12	9.65E-03	7.49E-03	1.29	1.29
Sox17	G01	9.93E-05	5.76E-05	1.72	1.72
Sox2	G02	1.13E-01	2.56E-01	0.44	-2.27
Sst	G03	1.10E-04	2.05E-04	0.54	-1.86
Sycp3	G04	2.31E-02	4.52E-02	0.51	-1.95
T	G05	1.24E-03	7.74E-04	1.60	1.60
Tat	G06	1.58E-04	2.14E-04	0.74	-1.36
Tcfcp2l1	G07	6.24E-02	1.64E-01	0.38	-2.63
Tdgf1	G08	5.88E-01	3.91E-01	1.50	1.50
Tert	G09	9.52E-03	1.15E-02	0.83	-1.21
Utf1	G10	6.70E-01	6.40E-01	1.05	1.05
Wt1	G11	4.45E-03	1.76E-03	2.52	2.52
Zfp42	G12	1.37E-01	3.66E-01	0.37	-2.68
Gusb	H01	3.27E-02	3.87E-02	0.85	-1.18
Hprt1	H02	3.84E-01	3.72E-01	1.03	1.03
Hsp90ab1	H03	6.02E+00	5.89E+00	1.02	1.02
Gapdh	H04	4.50E+00	4.19E+00	1.08	1.08
Actb	H05	2.94E+00	2.82E+00	1.04	1.04

GUE-2 (0 days vs 12 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	6.76E-05	4.53E-05	1.49	1.49
Bxdc2	A02	2.29E-01	2.96E-01	0.77	-1.30
Cd34	A03	1.24E-03	2.11E-03	0.59	-1.70
Cd9	A04	4.30E-01	4.88E-01	0.88	-1.14
Cdh5	A05	3.68E-04	8.28E-05	4.45	4.45
Cdx2	A06	9.53E-04	5.06E-04	1.88	1.88
Col1a1	A07	2.89E-03	2.47E-01	0.01	-85.35
Commd3	A08	7.09E-02	5.74E-02	1.23	1.23
Crabp2	A09	3.92E-03	8.55E-03	0.46	-2.18
Ddx4	A10	1.23E-02	1.68E-02	0.73	-1.37
Des	A11	6.42E-03	1.09E-02	0.59	-1.70
Diap2	A12	1.58E-04	4.29E-04	0.37	-2.72
Dnmt3b	B01	2.04E-01	1.10E-01	1.85	1.85
Ednrb	B02	2.93E-04	4.96E-04	0.59	-1.69
Eomes	B03	6.67E-03	2.96E-03	2.25	2.25
Fgf4	B04	1.09E-01	1.94E-01	0.56	-1.79
Fgf5	B05	2.32E-03	2.90E-04	7.99	7.99
Flt1	B06	5.59E-03	6.04E-04	9.25	9.25
Fn1	B07	3.80E-01	5.56E-01	0.68	-1.46
Foxa2	B08	3.91E-04	7.34E-05	5.33	5.33

Foxd3	B09	1.28E-02	2.31E-02	0.55	-1.81
Gabrb3	B10	6.49E-03	1.12E-02	0.58	-1.72
Gal	B11	1.19E-04	2.56E-04	0.47	-2.15
Gata4	B12	1.98E-03	2.08E-03	0.95	-1.05
Gata6	C01	1.34E-03	1.90E-03	0.71	-1.42
Gbx2	C02	3.42E-02	2.99E-02	1.14	1.14
Gcg	C03	6.33E-05	5.23E-05	1.21	1.21
Gcm1	C04	1.90E-04	2.09E-04	0.91	-1.10
Gdf3	C05	3.19E-02	5.32E-02	0.60	-1.67
Grb7	C06	4.79E-02	4.16E-02	1.15	1.15
Hba-x	C07	1.27E-04	2.95E-04	0.43	-2.31
Hbb-y	C08	1.25E-04	1.01E-04	1.24	1.24
Hck	C09	1.83E-03	7.12E-03	0.26	-3.90
Iapp	C10	7.35E-05	7.34E-05	1.00	1.00
Ifitm1	C11	1.21E-01	1.48E-01	0.82	-1.22
Ifitm2	C12	7.15E-02	1.48E-01	0.48	-2.07
Igf2bp2	D01	1.09E-01	1.42E-01	0.76	-1.31
Il6st	D02	6.39E-02	1.19E-01	0.53	-1.87
Ins2	D03	1.14E-04	1.44E-04	0.80	-1.26
Kit	D04	4.38E-03	8.76E-02	0.05	-19.98
Krt1	D05	8.14E-04	6.44E-04	1.26	1.26
Lama1	D06	1.39E-03	6.71E-02	0.02	-48.32
Lamb1-1	D07	7.10E-02	6.78E-02	1.05	1.05
Lamc1	D08	5.72E-02	8.99E-02	0.64	-1.57
Lefty1	D09	3.42E-01	1.70E-01	2.01	2.01
Lefty2	D10	1.78E-01	2.06E-01	0.87	-1.15
Lifr	D11	7.69E-03	1.91E-02	0.40	-2.49
Lin28	D12	6.47E-02	6.29E-02	1.03	1.03
Myf5	E01	5.70E-05	3.73E-05	1.53	1.53
Myod1	E02	5.76E-05	6.11E-05	0.94	-1.06
Nanog	E03	1.94E-01	4.47E-01	0.43	-2.30
Nes	E04	3.20E-02	4.31E-03	7.41	7.41
Neurod1	E05	2.23E-04	2.14E-03	0.10	-9.61
Nodal	E06	3.53E-02	3.70E-02	0.95	-1.05
Nog	E07	1.60E-03	2.60E-04	6.15	6.15
Nr5a2	E08	7.30E-03	5.14E-02	0.14	-7.04
Nr6a1	E09	7.92E-02	1.02E-01	0.78	-1.28
Numb	E10	1.35E-02	1.49E-02	0.91	-1.10
Olig2	E11	8.08E-03	4.73E-03	1.71	1.71
Pax4	E12	6.31E-05	5.15E-05	1.23	1.23
Pax6	F01	4.74E-03	9.33E-04	5.08	5.08
Pdx1	F02	1.27E-04	4.23E-05	3.00	3.00
Pecam1	F03	2.93E-02	1.20E-01	0.24	-4.08
Podxl	F04	1.31E-01	1.56E-02	8.35	8.35
Pou5f1	F05	1.74E+00	2.30E+00	0.76	-1.32
Pten	F06	2.31E-01	1.97E-01	1.17	1.17
Ptf1a	F07	7.74E-05	1.10E-04	0.71	-1.42

Rest	F08	1.91E-01	3.88E-01	0.49	-2.03
Runx2	F09	2.47E-04	7.40E-04	0.33	-3.00
Sema3a	F10	1.18E-04	5.68E-03	0.02	-48.16
Serpina1a	F11	1.66E-04	8.53E-05	1.95	1.95
Sfrp2	F12	1.34E-02	7.29E-03	1.84	1.84
Sox17	G01	9.84E-05	7.63E-05	1.29	1.29
Sox2	G02	1.24E-01	2.27E-01	0.55	-1.83
Sst	G03	1.41E-04	2.48E-04	0.57	-1.76
Sycp3	G04	1.92E-02	4.00E-02	0.48	-2.09
T	G05	3.74E-03	1.29E-03	2.90	2.90
Tat	G06	1.92E-04	2.80E-04	0.69	-1.46
Tcfcp2l1	G07	3.15E-02	1.91E-01	0.16	-6.08
Tdgf1	G08	4.96E-01	2.94E-01	1.69	1.69
Tert	G09	8.75E-03	9.98E-03	0.88	-1.14
Utf1	G10	4.90E-01	7.10E-01	0.69	-1.45
Wt1	G11	5.97E-03	1.50E-03	3.98	3.98
Zfp42	G12	1.36E-01	3.85E-01	0.35	-2.83
Gusb	H01	3.50E-02	4.40E-02	0.79	-1.26
Hprt1	H02	3.50E-01	3.61E-01	0.97	-1.03
Hsp90ab1	H03	6.34E+00	6.69E+00	0.95	-1.06
Gapdh	H04	4.11E+00	4.07E+00	1.01	1.01
Actb	H05	3.14E+00	2.31E+00	1.36	1.36

GUE-3 (0 days vs 12 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	6.87E-05	3.70E-05	1.86	1.86
Bxdc2	A02	1.75E-01	2.75E-01	0.64	-1.57
Cd34	A03	9.81E-04	2.91E-03	0.34	-2.96
Cd9	A04	3.95E-01	4.53E-01	0.87	-1.14
Cdh5	A05	8.32E-04	9.76E-05	8.52	8.52
Cdx2	A06	1.25E-02	4.48E-04	27.98	27.98
Col1a1	A07	8.27E-03	5.25E-01	0.02	-63.53
Commd3	A08	7.99E-02	5.99E-02	1.34	1.34
Crabp2	A09	7.84E-03	7.97E-03	0.98	-1.02
Ddx4	A10	4.40E-03	1.51E-02	0.29	-3.42
Des	A11	7.57E-03	1.45E-02	0.52	-1.91
Diap2	A12	1.91E-04	8.28E-04	0.23	-4.34
Dnmt3b	B01	1.52E-01	1.40E-01	1.08	1.08
Ednrb	B02	4.91E-04	2.79E-04	1.76	1.76
Eomes	B03	1.06E-02	2.05E-03	5.19	5.19
Fgf4	B04	3.41E-02	2.23E-01	0.15	-6.54
Fgf5	B05	4.21E-03	4.92E-04	8.55	8.55

Flt1	B06	9.47E-03	1.22E-03	7.78	7.78
Fn1	B07	4.79E-01	8.70E-01	0.55	-1.82
Foxa2	B08	1.38E-03	6.68E-05	20.70	20.70
Foxd3	B09	1.32E-02	1.92E-02	0.69	-1.45
Gabrb3	B10	1.08E-02	9.64E-03	1.12	1.12
Gal	B11	2.21E-04	2.95E-04	0.75	-1.33
Gata4	B12	5.72E-03	1.89E-03	3.02	3.02
Gata6	C01	6.22E-03	2.02E-03	3.09	3.09
Gbx2	C02	4.41E-02	2.54E-02	1.74	1.74
Gcg	C03	5.59E-05	3.79E-05	1.47	1.47
Gcm1	C04	4.06E-04	2.81E-04	1.44	1.44
Gdf3	C05	1.70E-02	5.14E-02	0.33	-3.02
Grb7	C06	3.78E-02	4.75E-02	0.80	-1.25
Hba-x	C07	5.99E-04	1.86E-04	3.21	3.21
Hbb-y	C08	2.15E-04	8.75E-05	2.45	2.45
Hck	C09	8.58E-04	3.01E-03	0.28	-3.51
Iapp	C10	5.31E-05	6.04E-05	0.88	-1.14
Ifitm1	C11	1.63E-01	1.52E-01	1.07	1.07
Ifitm2	C12	5.87E-02	1.62E-01	0.36	-2.76
Igf2bp2	D01	1.06E-01	1.10E-01	0.96	-1.04
Il6st	D02	1.11E-01	1.28E-01	0.86	-1.16
Ins2	D03	1.10E-04	1.33E-04	0.83	-1.21
Kit	D04	9.27E-03	8.90E-02	0.10	-9.59
Krt1	D05	4.07E-04	6.35E-04	0.64	-1.56
Lama1	D06	3.96E-03	5.19E-02	0.08	-13.11
Lamb1-1	D07	1.09E-01	1.04E-01	1.04	1.04
Lamc1	D08	7.44E-02	1.20E-01	0.62	-1.61
Lefty1	D09	3.83E-01	1.67E-01	2.29	2.29
Lefty2	D10	1.15E-01	2.08E-01	0.55	-1.81
Lifr	D11	8.65E-03	2.18E-02	0.40	-2.52
Lin28	D12	9.08E-02	7.47E-02	1.22	1.22
Myf5	E01	5.27E-05	5.96E-05	0.88	-1.13
Myod1	E02	4.43E-05	7.52E-05	0.59	-1.70
Nanog	E03	1.73E-01	5.21E-01	0.33	-3.00
Nes	E04	3.48E-02	6.64E-03	5.25	5.25
Neurod1	E05	1.12E-04	3.07E-03	0.04	-27.44
Nodal	E06	3.51E-02	4.45E-02	0.79	-1.27
Nog	E07	4.54E-03	2.94E-04	15.41	15.41
Nr5a2	E08	3.53E-03	5.18E-02	0.07	-14.68
Nr6a1	E09	1.22E-01	8.89E-02	1.38	1.38
Numb	E10	1.99E-02	1.78E-02	1.11	1.11
Olig2	E11	8.46E-03	5.17E-03	1.63	1.63
Pax4	E12	5.55E-05	4.70E-05	1.18	1.18
Pax6	F01	2.59E-03	1.09E-03	2.38	2.38
Pdx1	F02	8.81E-05	1.28E-04	0.69	-1.46
Pecam1	F03	9.65E-03	1.08E-01	0.09	-11.20
Podxl	F04	2.18E-01	2.07E-02	10.51	10.51

Pou5f1	F05	1.06E+00	2.00E+00	0.53	-1.88
Pten	F06	2.27E-01	1.87E-01	1.22	1.22
Ptf1a	F07	5.54E-05	6.81E-05	0.81	-1.23
Rest	F08	1.30E-01	3.89E-01	0.33	-3.00
Runx2	F09	1.23E-04	1.13E-03	0.11	-9.18
Sema3a	F10	6.08E-04	6.31E-03	0.10	-10.37
Serpina1a	F11	1.85E-04	1.29E-04	1.43	1.43
Sfrp2	F12	9.16E-03	1.01E-02	0.91	-1.10
Sox17	G01	3.04E-04	9.37E-05	3.25	3.25
Sox2	G02	9.51E-02	3.16E-01	0.30	-3.32
Sst	G03	1.86E-04	2.04E-04	0.91	-1.10
Sycp3	G04	6.35E-03	3.62E-02	0.18	-5.69
T	G05	7.82E-02	9.37E-04	83.54	83.54
Tat	G06	1.30E-04	2.63E-04	0.49	-2.03
Tcfcp2l1	G07	3.48E-02	2.09E-01	0.17	-6.02
Tdgf1	G08	4.37E-01	3.01E-01	1.45	1.45
Tert	G09	8.37E-03	1.18E-02	0.71	-1.41
Utf1	G10	3.21E-01	5.80E-01	0.55	-1.81
Wt1	G11	8.82E-03	2.10E-03	4.20	4.20
Zfp42	G12	7.01E-02	3.58E-01	0.20	-5.11
Gusb	H01	2.73E-02	3.85E-02	0.71	-1.41
Hprt1	H02	3.61E-01	3.48E-01	1.04	1.04
Hsp90ab1	H03	6.02E+00	6.22E+00	0.97	-1.03
Gapdh	H04	4.20E+00	4.10E+00	1.03	1.03
Actb	H05	4.01E+00	2.93E+00	1.37	1.37

GUE-4 (0 days vs 12 days)					
Symbol	Well	2^-ΔCt	2^-ΔCt	Fold Difference	Fold Up- or Down-Regulation
		Test Sample	Control Sample	Test Sample /Control Sample	Test Sample /Control Sample
Afp	A01	7.39E-05	4.18E-05	1.77	1.77
Bxdc2	A02	1.55E-01	3.81E-01	0.41	-2.45
Cd34	A03	1.12E-03	1.30E-03	0.86	-1.16
Cd9	A04	5.61E-01	4.09E-01	1.37	1.37
Cdh5	A05	7.39E-05	4.18E-05	1.77	1.77
Cdx2	A06	6.10E-04	4.77E-04	1.28	1.28
Col1a1	A07	2.18E-03	1.46E-01	0.01	-66.96
Commd3	A08	6.71E-02	3.71E-02	1.81	1.81
Crabp2	A09	3.09E-03	1.31E-02	0.24	-4.24
Ddx4	A10	9.88E-03	3.21E-02	0.31	-3.25
Des	A11	1.04E-02	1.79E-02	0.58	-1.72
Diap2	A12	1.12E-04	1.84E-04	0.61	-1.63
Dnmt3b	B01	1.45E-01	9.80E-02	1.48	1.48
Ednrb	B02	4.41E-04	5.06E-04	0.87	-1.15

Eomes	B03	7.16E-03	2.22E-03	3.23	3.23
Fgf4	B04	1.19E-01	1.45E-01	0.82	-1.22
Fgf5	B05	2.73E-03	1.68E-04	16.23	16.23
Flt1	B06	4.65E-03	5.83E-04	7.98	7.98
Fn1	B07	5.22E-01	5.63E-01	0.93	-1.08
Foxa2	B08	7.85E-05	7.46E-05	1.05	1.05
Foxd3	B09	6.42E-03	1.04E-02	0.62	-1.62
Gabrb3	B10	5.81E-03	4.72E-03	1.23	1.23
Gal	B11	1.00E-04	2.81E-04	0.36	-2.80
Gata4	B12	1.84E-03	3.40E-03	0.54	-1.84
Gata6	C01	9.77E-04	7.00E-04	1.39	1.39
Gbx2	C02	4.48E-03	2.43E-02	0.18	-5.42
Gcg	C03	7.39E-05	4.67E-05	1.58	1.58
Gcm1	C04	1.59E-04	3.08E-04	0.52	-1.94
Gdf3	C05	2.11E-02	6.12E-02	0.34	-2.90
Grb7	C06	6.11E-02	5.44E-02	1.12	1.12
Hba-x	C07	1.13E-04	3.51E-04	0.32	-3.09
Hbb-y	C08	8.46E-05	1.08E-04	0.78	-1.28
Hck	C09	8.79E-04	1.10E-02	0.08	-12.48
Iapp	C10	7.39E-05	6.08E-05	1.22	1.22
Ifitm1	C11	1.56E-01	1.89E-01	0.82	-1.22
Ifitm2	C12	8.33E-02	2.06E-01	0.41	-2.47
Igf2bp2	D01	9.24E-02	4.33E-02	2.14	2.14
Il6st	D02	8.58E-02	1.38E-01	0.62	-1.61
Ins2	D03	1.19E-04	1.90E-04	0.62	-1.60
Kit	D04	1.15E-02	3.97E-02	0.29	-3.45
Krt1	D05	5.19E-04	1.11E-03	0.47	-2.13
Lama1	D06	3.26E-04	2.91E-02	0.01	-89.29
Lamb1-1	D07	1.21E-01	5.01E-02	2.41	2.41
Lamc1	D08	8.04E-02	5.91E-02	1.36	1.36
Lefty1	D09	2.62E-01	7.11E-01	0.37	-2.71
Lefty2	D10	1.12E-01	2.42E-01	0.46	-2.15
Lifr	D11	8.43E-03	1.99E-02	0.42	-2.36
Lin28	D12	1.34E-01	6.01E-02	2.22	2.22
Myf5	E01	7.39E-05	6.42E-05	1.15	1.15
Myod1	E02	7.39E-05	6.92E-05	1.07	1.07
Nanog	E03	3.62E-01	3.27E-01	1.11	1.11
Nes	E04	2.23E-02	7.23E-03	3.09	3.09
Neurod1	E05	1.77E-04	9.12E-04	0.19	-5.17
Nodal	E06	4.65E-02	3.63E-02	1.28	1.28
Nog	E07	2.92E-03	1.75E-04	16.70	16.70
Nr5a2	E08	3.68E-03	6.02E-02	0.06	-16.35
Nr6a1	E09	9.34E-02	1.07E-01	0.87	-1.15
Numb	E10	1.88E-02	1.68E-02	1.12	1.12
Olig2	E11	7.66E-03	1.80E-03	4.27	4.27
Pax4	E12	7.39E-05	5.72E-05	1.29	1.29
Pax6	F01	3.02E-03	4.53E-03	0.67	-1.50

Pdx1	F02	7.39E-05	5.52E-05	1.34	1.34
Pecam1	F03	3.81E-02	1.41E-01	0.27	-3.70
Podxl	F04	3.47E-01	6.93E-03	50.06	50.06
Pou5f1	F05	1.90E+00	2.44E+00	0.78	-1.28
Pten	F06	1.79E-01	2.01E-01	0.89	-1.13
Ptf1a	F07	7.93E-05	6.44E-05	1.23	1.23
Rest	F08	1.82E-01	3.43E-01	0.53	-1.88
Runx2	F09	7.39E-05	2.26E-04	0.33	-3.06
Sema3a	F10	7.64E-05	2.11E-03	0.04	-27.67
Serpina1a	F11	1.19E-04	1.04E-04	1.14	1.14
Sfrp2	F12	1.76E-02	2.82E-03	6.23	6.23
Sox17	G01	9.40E-05	4.18E-05	2.25	2.25
Sox2	G02	1.92E-01	1.72E-01	1.11	1.11
Sst	G03	1.41E-04	3.20E-04	0.44	-2.26
Sycp3	G04	2.32E-02	6.36E-02	0.36	-2.74
T	G05	3.11E-04	2.25E-03	0.14	-7.25
Tat	G06	2.10E-04	2.11E-04	1.00	-1.00
Tcfcp2l1	G07	7.65E-02	1.37E-01	0.56	-1.80
Tdgf1	G08	5.64E-01	3.54E-01	1.59	1.59
Tert	G09	1.37E-02	9.94E-03	1.37	1.37
Utf1	G10	6.52E-01	6.67E-01	0.98	-1.02
Wt1	G11	2.91E-03	1.38E-03	2.11	2.11
Zfp42	G12	1.36E-01	3.78E-01	0.36	-2.79
Gusb	H01	3.10E-02	3.36E-02	0.93	-1.08
Hprt1	H02	3.76E-01	5.96E-01	0.63	-1.59
Hsp90ab1	H03	5.66E+00	8.71E+00	0.65	-1.54
Gapdh	H04	5.16E+00	3.36E+00	1.54	1.54
Actb	H05	2.93E+00	1.71E+00	1.72	1.72

Selected genes from microarray analysis and their fold change gene expression in all three lines

GeneSymbol (Probe ID)	H1/G1	H2/G2	H3/G3
Col6a6 (A_51_P105847)	6.39	2.77	6.16
Col6a2 (A_51_P279639)	1.82	2.14	-4.33
Col6a1 (A_55_P2128153)	1.94	3.41	-3.99
Bmp2 (A_55_P2041961)	29.61	2.80	-12.77
Egfr (A_52_P106259)	3.15	3.41	1.65
RhoB (A_52_P89567)	6.62	3.59	-1.61
Cdh4 (A_52_P402897)	-3.71	-3.05	1.54
Nr5a2 (A_51_P514449)	-1.79	-2.06	-1.46
Utf1 (A_55_P2104769)	-1.54	-2.33	1.24
Pou5f1 (A_51_P202340)	-1.69	-2.52	1.09
Sox2 (A_52_P307739)	-1.92	-1.55	-1.40
Nanog (A_55_P2007708)	-1.98	-2.41	-1.09
Igf2bp2 (A_51_P315682)	1.67	1.93	1.52
Irs1 (A_52_P175242)	5.06	2.94	-2.64
Pik3ip1 (A_51_P463428)	3.12	3.67	-1.99
Igfbp1 (A_51_P447545)	12.48	5.01	1.04
Kcnq1 (A_55_P2086203)	13.01	9.56	-1.39
Lepr (A_55_P2177911)	5.52	9.10	2.37
Prdm9 (A_55_P2027057)	1.99	2.39	4.22
Fcgbp (A_51_P380078)	7.76	1.94	-2.01
Fam107b (A_55_P2023772)	2.69	4.16	-1.12
Dcdc2a (A_52_P640922)	1.83	2.01	-2.56
Tert (A_51_P433778)	1.03	1.30	-1.14

Publications (PhD)

1- Morshedi, A., Ren, Z., Li, J., and Dröge, P. (2012). Probing into the Biological Processes Influenced by ESC Factor and Oncoprotein HMGA2 Using iPSCs. Stem Cell Reviews and Reports, 1-9.

2- Morshedi, A., Soroush Noghabi, M., and Dröge, P.(2011). Use of UTF1 Genetic Control Elements as iPSC Reporter Stem cell Reviews and Reports, 1-8.

3- Pfannkuche, K., Hannes, T., Khalil, M., Noghabi, M.S., Morshedi, A., Hescheler, J., and Droge, P. (2010). Induced pluripotent stem cells: a new approach for physiological research. Cell Physiol Biochem *26*, 105-124.