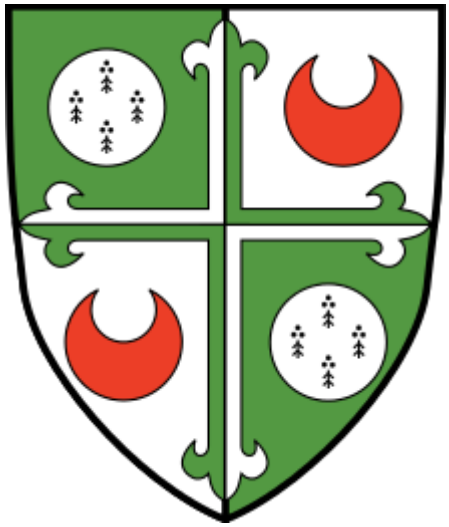
ESTABLISHMENT IN CULTURE OF MOUSE AND HUMAN STEM CELLS WITH EXPANDED FATE POTENTIAL



A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

David John Ryan

Wellcome Trust Sanger Institute

Girton College, University of Cambridge

September 2016

ABSTRACT

ESTABLISHMENT IN CULTURE OF MOUSE AND HUMAN STEM CELLS WITH EXPANDED FATE POTENTIAL

DR DAVID JOHN RYAN

The zygote and blastomeres of cleavage stage mouse embryos have the capacity to differentiate to the embryonic and both extra-embryonic lineages and are considered functionally totipotent. Until now, it has not been possible to establish stable cell lines that resemble these totipotent-like cells. In this work, I hypothesised that by modulating signalling pathways known to be important in early embryonic development it may be possible to capture *in vitro* a self-renewing state that possessed features of pre-implantation blastomeres. I succeeded in formulating a novel hypothesis driven cell culture medium which allowed the establishment of a stem cell state that possessed expanded differentiation potential to the embryonic and both extra-embryonic lineages *in vitro* and *in vivo*. These cells were isolated directly from *in vitro* culture of mouse pre-implantation embryos or single cell blastomeres, reprogrammed from somatic cells or converted from mouse ES cells. With these cells, I generated single cell chimeras which demonstrated extensive contribution to all lineages in the developing organism providing additional evidence that this chemical medium maintained a homogenous stem cell population. I demonstrated that the transcriptome of these cells was enriched with an early pre-implantation blastomere signature, distinct from other rare published totipotent-like cells. Finally, I demonstrated that the same chemical formulation permitted the establishment *in vitro* of a human stem cell state that possessed expanded differentiation potential to the embryonic and extra-embryonic lineage *in vitro*. My work has shown for the first time that through chemical modulation of pathways implicated to be involved in pre-implantation development, a novel homogenous stem cell state that possesses a pre-implantation transcriptional signature and expanded differentiation potential to both the

embryonic and extra-embryonic lineage can be established and maintained *in vitro* in both mouse and human, suggesting a possible interspecies conservation of the signalling networks involved in early embryonic development.

Declaration

I hereby declare that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specially indicated in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution.

In accordance with the Degree Committee for the Faculty of Biology guidelines, this thesis does not exceed 60,000 words.

Signed: .....

Date: .....

David John Ryan

Acknowledgements

I am extremely grateful to my supervisors Dr Pentao Liu and Dr Colin Watts for their scientific guidance, support and direction over the past four years and during compilation of this thesis. I would like to acknowledge my thesis committee members, Professor Allan Bradley and Professor Austin Smith for their astute observations and erudite comments on my project.

I must mention the members, present and past of Team 18 at the Sanger Institute; Dr Wei Wang, Dr Jian Yang, Dr Xuefei Gao, Dr Yong Yu, Dr Cheuk-Ho Tsang, Liliana Antunes, Dr Walid Khalid, Dr Shannon Burke, Dr Su Kit Chew, Dr Juexuan Wang and Rebecca Kiff for their interest in my work, their patience during my learning and their support of my scientific development. Within the Sanger facility, I would like to thank Dr Bee Ling Ng and William Cheng for their assistance with FACS. In the Molecular cytogenetics core facility at the Sanger Institute I would like to pay special thanks to Dr Beiyuan Fu and Dr Fengtang Yang. From the RSF animal facility, Paul Green, and Nick Harman. The advice provided by Dr Frederick J Livesey and Dr Peter Kirwan in the Gurdon Institute and their help in arranging electrophysiological studies is very much appreciated.

My sincerest thanks to my external collaborators Dr Yung-Yao Lin at the Blizzard Institute, Dr Guocheng Lan and Dr Xiangang Zou at Cancer Research UK and the lab members of the Brain Repair Centre, Cambridge. I would like to extend my gratitude to Dr Annabel Smith and Dr Christina Hedberg-Delouka for their administrative assistance during my time at the Sanger Institute. I would to conclude by thanking Professor David Lomas, Professor Krishna Chatterjee and the Wellcome Trust funded PhD programme for Academic Clinicians for giving me this opportunity which has allowed me to take my first steps in an exciting new journey towards a rewarding career as a clinician scientist.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

2C	2-cell stage embryo
2i	MEK inhibitor (PD0325901) + GSK3 inhibitor (CHIR99021)
5hmC	5 hydroxy-methyl-Cytosine
5mC	5-methyl-Cytosine
aCGH	array Comparative Genome Hybridization
APC	Adenomatous Polyposis Coli
B-ME	Beta-mercaptoethanol
BMP	Bone Morphogenetic Protein
BSA	Bovine Serum Albumin
cDNA	complementary Deoxyribonucleic Acid
CB	Cytochalasin B
CM	Conditioned Media
DAPI	4',6-Diamidino-2-Phenylindole, Dihydrochloride
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DOX	Doxycycline
dpc	day post coitum
E3.5	Day 3.5 after fertilization
EGA	Embryonic Gene Activation
Epi	Epiblast
EB	Embryoid Body
Em	Embryo proper
EPSC	Expanded Potential Stem Cell

EPSCM	Expanded Potential Stem Cell Medium
EpiSC	Epiblast Stem cell
ERK	Extracellular signal-Regulated Kinase
ESC	Embryonic Stem Cell
ExEc	Extra-Embryonic Ectoderm
FACS	Fluorescence Assisted Cell Sorting
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
FISH	Fluorescence in Situ Hybridization
FGF	Fibroblast Growth Factor
FGFR	Fibroblast Growth Factor Receptor
GSEA	Gene Set Enrichment Analysis
GSK3	Glycogen Synthase Kinase 3
hCG	human Chorionic Gonadotropin
HDF	Human Dermal Fibroblast
hfNSCs	human fetal Neural Stem Cells
HRP	Horseradish Peroxidase
ICM	Inner Cell Mass
Id	Inhibitor of DNA binding
IGF	Insulin-like Growth Factor
iPSC	induced Pluripotent Stem Cell
KSR	KnockOut Serum Replacement
LEF	Lymphoid-Enhancer binding Factor
LIF	Leukaemia Inhibitory Factor
MAPK	Mitogen Activated Protein Kinase

MEFs	Mouse Embryonic Fibroblasts
MGA	Mid-preimplantation Gene Activation
mRNA	messenger Ribonucleic Acid
NSC-CM	Neural Stem Cell-Culture Media
PB	PiggyBac
PBS	Phosphate Balanced Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
Pl	Placenta
PrE	Primitive Endoderm
PVP	Polyvinylpyrrolidone
RNA	Ribonucleic Acid
RT	Room Temperature
SCNT	Somatic Cell Nuclear Transfer
SFK	Src-Family Kinase
STAT	Signal Transducer and Activator of Transcription
TCF	T-Cell Factor
TE	Trophectoderm
TSC	Trophoblast Stem Cell
TSCLCs	Trophoblast Stem Cell-Like Cells
XCI	X Chromosome Inactivation
XEN	Extra-embryonic Endoderm
Yo	Yolk sac
ZGA	Zygotic Gene Activation

CHAPTER 1

1 INTRODUCTION

1.1 Summary and aims

The overarching aim of the work described in this PhD thesis was to investigate if stable, homogenous, 'totipotent-like' stem cell cultures could be established *in vitro*. To this end, I developed a novel chemically defined media which targeted key signalling pathways known to be important in the development of the pre-implantation mouse embryo. Using this media, I succeeded in capturing a unique stem cell state that I subsequently extensively characterized for the presence of expanded differentiation potential to the embryonic and both extra-embryonic lineages. In the second half of this thesis, I investigated whether there was inter-species conservation of this stem cell state by using the human as a model system. This introductory chapter provides background information concerning these projects and is split into two sections. In the first section, I provide a general overview of totipotency and discuss pre-implantation development in the mouse with specific focus on the regulation of the first and second specialization events. In the second part, I review the current available knowledge on human pre-implantation development and consider the notion that the pathways implicated in pre-implantation development may demonstrate inter-species conservation.

1.2 Development and fate restriction

Development is a process of progressive fate restriction. A conceptual framework to visualise this process is provided in Waddington's epigenetic landscape. In this model, development is conceptualized as a ball descending down the face of a mountain. During its descent, the undulations and variations in the landscape limit the number of paths the ball can ultimately take. Furthermore, the gradient of descent and surface irregularities prevent the ball from either re-ascending (reprogramming) or moving laterally (trans-differentiating) under normal conditions. In this metaphor, the ball sitting at the top of the mountain, prior to descent, has unbridled potential and may be referred to as being totipotent. Totipotency refers to a cell that is capable of developing into a complete organism.

In mammalian systems, embryonic development begins once the sperm and egg fuse to produce the fertilized zygote. Prior to fertilization, the mouse oocyte is arrested in the metaphase of meiosis II. Following fertilization, the cell completes meiosis resulting in the formation of a one-cell embryo containing haploid paternal and maternal pronuclei. The zygote subsequently goes through a series of early cleavage divisions producing multiple equi-potent blastomeres. At the eight-cell stage, the embryo undergoes the process of compaction where individual blastomeres adopt a more flattened morphology secondary to increased intercellular adhesion and development of tight junctions. This process is essential for subsequent development of the three major lineages; the epiblast, trophoblast and primitive endoderm (Morgani and Brickman, 2014; Cockburn and Rossant, 2010; Czechanski et al., 2014) **(Figure 1.1)**. When we discuss the pre-implantation embryo, the choice of terms we use to describe the developmental potential of a particular embryonic stage is important. In the pre-implantation embryo, totipotency is reserved for the zygote and two-cell stage embryo both of which can support the development of a complete organism from a single starting cell (Tarkowski, 1959). In contrast, from the two-cell stage until approximately the 32-cell stage,

each blastomere has the capacity to forms all cell types of the developing organism in the presence of additional carrier blastomeres but, is not capable of forming the entire organism *de novo* (Kelly, 1977; Tarkowski and Wróblewska, 1967).

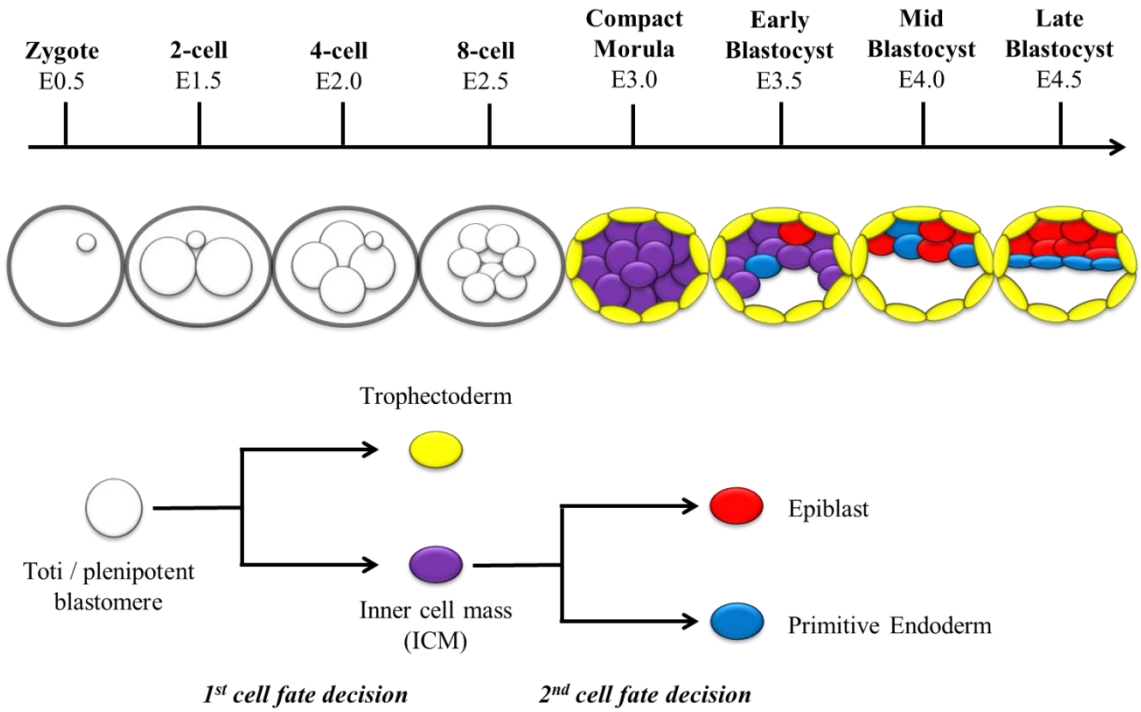


Figure 1.1 Overview of mouse pre-implantation development

In mammalian systems, embryonic development begins once the sperm and egg fuse to produce the fertilized zygote. The zygote subsequently undergoes a series of early cleavage divisions producing multiple equi-potent blastomeres. At the eight-cell stage, the embryo undergoes compaction secondary to increased intercellular adhesion and development of tight junctions. Compaction is essential for the correct development of the three major embryonic lineages; the epiblast, trophoblast and primitive endoderm. The first cell fate decision occurs where the embryo segregates into the trophectoderm which surrounds the developing embryo and will form the extra-embryonic contribution of the placenta and the inner cell mass. As the blastocyst transitions through early, mid and late stages, the inner cell mass undergoes the second cell fate decision forming the epiblast which will form the embryo proper and the hypoblast which produces the primitive endoderm and developing yolk sac.

Adapted from Czechanski et al., 2014.

Blastomeres isolated from four and eight-cell stage mouse embryos that were injected into empty zonae and transferred to the oviduct of pseudopregnant female mice were capable of inducing decidual formation; however, no live embryo was obtained (Rossant, 1976). The conclusion was that experimental failure did not reflect a lack of totipotency per se, but rather could be explained by an insufficient number of cells being present at cavitation preventing normal development form progressing. It is argued; therefore, that we should use the term "plenipotent" to refer to these cells that possess expanded potential but lack all the features to be defined as truly totipotent (Condic, 2014).

The above experiments demonstrated that blastomeres from the four-cell stage onwards had plenipotent potential; however, did the likelihood of identifying this plenipotent potential decrease with later cleavage divisions as would be expected from Waddington's model? Horseradish peroxidase injection (HRP) was traditionally used to track cell lineages in Xenopus embryos (Hirose and Jacobson, 1979). HRP can also be used as a lineage marker in mouse embryos as it can diffuse through the cytoplasm allowing all daughter cells to be marked and it does not exit through gap junctions which increases the specificity of the staining (Lo and Gilua, 1979). HRP was injected into single blastomeres at the two-cell, four-cell, 16-cell and morula-stage and embryos were then allowed to develop until the blastocyst stage. It was observed that single injected blastomeres contributed to both the ICM and the trophectoderm but the probability of observing this contribution pattern decreased at later embryonic stages with > 50% of the outer cells injected at the 16-cell and late morula stages contributing to trophectoderm only in keeping with the notion of temporal developmental fate restriction (Balakier and Pedersen, 1982). Importantly, this loss of plenipotent potential refers to the quantitative reduction in the likelihood of observed plenipotency rather than a reduction in the plenipotent capacity of a random cell as embryo splitting experiments have shown

plasticity in developmental potential as late as the 32-cell stage embryo (Rossant, 1976; Tarkowski et al., 2010).

1.3 Transcriptional and epigenetic dynamics during pre-implantation development

Much of our knowledge concerning pre-implantation development come from studies on the mouse due to the ease associated with its use. Analysis of global gene expression during mouse pre-implantation development has identified three characteristic events which correspond with the morphological and functional changes that are identified in the pre-implantation embryo. Initially, following degradation of maternal RNA, which is 90% complete by the two-cell stage and which comprises the first characteristic event, there are two waves of *de novo* zygotic transcription. The first wave corresponds to zygotic gene activation (ZGA), also referred to as embryonic gene activation (EGA). Correct initiation of this process is critically dependent on maternal RNAs (Hamatani et al., 2004).

In eutherian mammals, including mouse, the onset of ZGA at the two-cell stage coincides with a major epigenetic event necessary for normal development, the onset of X-chromosome inactivation (XCI). In mouse, initial XCI is imprinted with preferential inactivation of the paternal X-chromosome which will subsequently remain silenced throughout pre-implantation development. The paternal X-chromosome will become re-activated in the inner cell mass and will subsequently undergo random XCI during somatic differentiation of the epiblast mediating dosage compensation of gene expression levels between the sexes (Lyon, 1961; Mak et al., 2004). The initial imprinted paternal XCI will persist in descendants of the extra-embryonic lineage (Huynh and Lee, 2003; Takagi and Sasaki, 1975).

Beyond XCI, the pre-implantation embryo undergoes extensive epigenetic remodelling and removal of epigenetic signatures from the maternal and paternal genomes through active and passive DNA demethylation (Mayer et al., 2000, Oswald et al., 2000, Rougier et al., 1998). Active DNA demethylation is mediated by *Tet3* dioxygenase-dependent oxidation of 5-methylcytosine (5mC) (Gu et al., 2011). Passive demethylation occurs secondary to 5mC dilution from the subsequent cleavage divisions due to DNA replication (Rougier et al., 1998). The ultimate goal of this widespread demethylation is the production of a completely hypomethylated genome in the epiblast which is unbiased in its somatic differentiation potential (Dean et al., 2001). Consistent with this stepwise remodelling of the pre-implantation epigenome, the spatiotemporal timeline of pre-implantation development can be reconstructed using the global expression kinetics of chromatin modifiers alone (Burton et al., 2013).

The second wave of *de novo* zygotic transcription which is referred to as mid-preimplantation gene activation (MGA) immediately precedes the first lineage specification event (Hamatani et al., 2004).

1.4 The first lineage specification

After the 4th or 5th cleavage division corresponding to the 16 – 32 cell stage embryo, the morula undergoes a process of compaction and segregation into the developing inner cell mass and trophectoderm. Many theories have been proposed to explain this process; however, the exact mechanism remains elusive.

1.4.1 Hypotheses to explain the first lineage specification

Over the years as our knowledge of mouse pre-implantation development has improved, the working models to explain the first cell fate specification have undergone continual refinement. Originally, the 'early asymmetry model' was suggested which proposed that fate determination i.e.: trophectoderm versus inner cell mass was decided as early as the zygote where spatial positioning relative to the dorsal or ventral cytoplasm would dictate future developmental potential (Gardner, 1996). This mechanism of fate determination exists in non-mammalian species; therefore, it was plausible to consider that such a mechanism may exist in mammals. However, embryo splitting experiments showing plasticity in developmental potential as late as the 32-cell stage embryo challenged this hypothesis (Rossant, 1976; Tarkowski et al., 2010).

An alternative explanation was provided by the 'inside-out' model which based on experimental observation concluded that blastomere position in the morula determined cell fate. The inner cells tended to contribute to the developing inner cell mass whereas the outer cells preferentially populated the trophectoderm **(Figure 1.2A)** (Tarkowski and Wróblewska, 1967; Yamanaka et al., 2006).

The 'polarization-model' supported the inside-out hypothesis and attempted to provide a mechanism to explain this occurrence. Coinciding with the spatial organization of blastomeres into an inside-outside arrangement is the acquisition of cell polarity. It was observed that at the eight-cell stage, cell divisions produced a polar-apolar pair. The polarized cell defined based on the presence of microvilli and surface ligand expression became localised to the outside and committed to the developing trophectoderm. The non-polarized inner cell differentiated into the inner cell mass **(Figure 1.2B)**

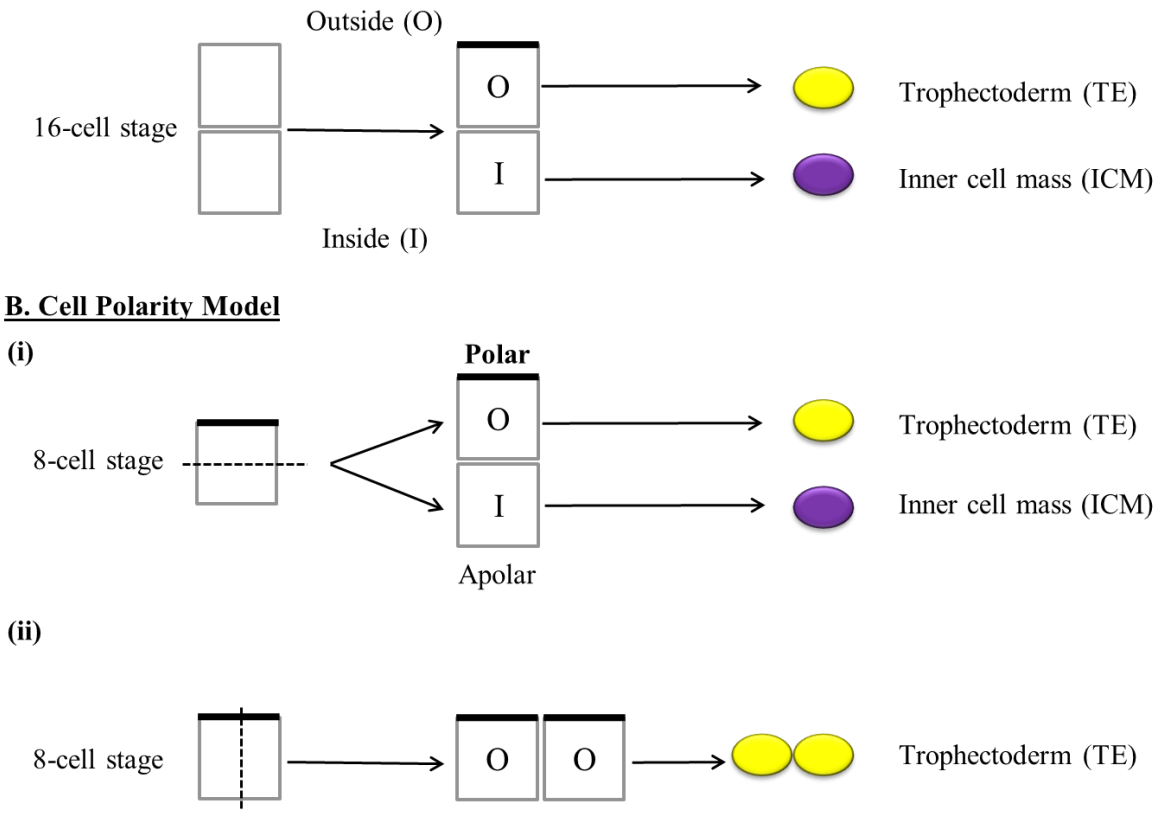


Figure 1.2 1st cell fate decision: Inner cell mass versus trophectoderm.

(A) The 'inside-out' model assumes that in the morula, the inner cells (I) tend to contribute to the developing inner cell mass (ICM) and the outer cells (O) preferentially populate the trophectoderm (TE) (Tarkowski., 1967).

(B) In the 'polarization-model' , an asymmetric cell division in the eight-cell stage embryo produces a polar-apolar pair (B(i)). The polarized outer cell (O) commits to the developing trophectoderm (TE) whereas the non-polarized inner cell (I) differentiates into the inner cell mass (ICM). Alternatively, a symmetric cell division of a polarized cell will produce two polarized outer (O) daughters cells, both of which will contribute to the developing trophectoderm (B (ii)) (Johnson., 1981).

Adapted from Yamanaka et al., 2006.

(Johnson and Ziomek, 1981; Yamanaka et al., 2006). The conclusions of the study from Johnson et al. was that lineage restriction was instituted at the eight-cell stage at the time of acquisition of polarity and the differential inheritance of this polarity was a potential mechanism to explain the first cell fate decision. However, challenging part of this assertion is the finding that outer cells in the morula can still can contribute to the inner cell mass when aggregated with eight-cell stage embryos (Rossant and Vijh, 1980).

The combined inside-outside and polarization models provide a good conceptual framework to explain some aspects of pre-implantation development but it still remains unclear what initiates this spatial organization of cells and subsequent differential inheritance of polarity. Is this a stochastic process, dictated entirely by chance or are there intrinsic and extrinsic determinants present at an earlier stage in development that we are as of now unaware?

Recent work has presented data suggesting that even as early as the four-cell stage embryo, the fate of a single blastomere is deterministic and not stochastic as previously thought (Piotrowska-Nitsche et al., 2005; Tabansky et al., 2013). This difference in developmental potential between individual blastomeres may be explained in part by differential expression of transcription factors and / or epigenetic regulators (Burton et al., 2013; Plachta et al., 2011; Torres-Padilla et al., 2007). PRDM14 is heterogeneously expressed in four-cell stage blastomeres and it was found that forced expression of this transcription factor at the two-cell stage preferentially induced a pluripotent inner cell mass fate (Burton et al., 2013). Epigenetic factors such as the H3-specific arginine methyltransferase CARM1 when overexpressed in blastomeres can direct their fate preferentially to the inner cell mass also (Torres-Padilla et al., 2007). Interestingly, the kinetics of *Oct4* expression in the pre-implantation embryo is thought to be a determinant of blastomere fate (Plachta et al., 2011). Those blastomeres with ‘fast’ and ‘slow’ *Oct4* kinetics are biased to a trophectoderm and

pluripotent inner cell mass fate respectively. *Oct4* and *Sox2* and their target *Sox21* are heterogeneously expressed in the four-cell stage embryo (Goolam et al., 2016). Loss of *Sox21* expression results in an upregulation of *Cdx2* and initiation of trophectoderm specification (Goolam et al., 2016). Conversely, inhibition of *Sox21*, as can be achieved by expression of CARM1 supports an inner cell mass fate. Similar to the binding kinetics of OCT4 in determining blastomere fate, temporal DNA binding of SOX2 in the four-cell embryo also orchestrates the first cell fate decision (White et al., 2016). SOX2 binding was found to be heterogeneous in the four-cell stage embryo and again was regulated by histone 3 arginine 26 methylation (H3R26me). Increasing H3R26 methylation increased long lived DNA binding by SOX2 which was associated with an inner cell mass fate (White et al., 2016).

The preceding paragraphs have provided an overview of the proposed models and mechanisms to explain why a cell may become inherently biased to a trophectoderm or inner cell mass fate but the next logical step is to explore the signalling pathways and networks that are differentially active at the time of the first lineage specification.

1.4.2 Signalling networks involved in the first cell fate decision

Caudal type homeobox 2 (*Cdx2*) is a transcription factor known to be involved in trophectoderm specification. Up until the eight-cell stage embryo, *Cdx2* is ubiquitously expressed in blastomeres (Ralston and Rossant, 2008). Coinciding with the acquisition of polarity, *Cdx2* mRNA becomes localised in the apical segment of the polarized outer cell and subsequent asymmetric cell divisions produce differential inheritance of *Cdx2* which remains localised to the outer cells (Skamagki et al., 2013). *Cdx2* cooperates with additional transcription factors including *Gata3, Eomes and Elf5* to initiate trophectoderm specification and reciprocally inhibit the pluripotency transcriptional network, specifically *Oct4* and *Nanog*

in the trophectoderm (Ng et al., 2008; Niwa et al., 2005; Ralston et al., 2010; Russ et al., 2000; Strumpf et al., 2005). The importance of *Cdx2* is illustrated from combined knockout of maternal and zygotic *Cdx2* which results in failure of appropriate polarization of cells at the eight to 16 cell-stage embryo and subsequent failure of trophectoderm specification which is embryonically lethal (Jedrusik et al., 2010; Jedrusik et al., 2015). *Cdx2* is a downstream effector of trophectoderm specification. So what are the signalling pathways that converge on *Cdx2*?

Firstly, the highly conserved Hippo signalling pathway is differentially active between the inner and outer cells of the pre-implantation embryo. The transcription factor *Tead4* has been shown to be the earliest expressed gene which is essential for specifying the trophectoderm lineage (Yagi et al., 2007). The *Tead4* null phenotype is embryonically lethal due to failed placentation as the morulae do not produce trophectoderm or blastocoel cavities. *Tead4* initiates trophectoderm specification by inducing *Cdx2* (Yagi et al., 2007). This process; however, is modulated by the coactivator protein YAP. In outer cells, YAP localises to the nucleus and leads to changes in *Tead4* and *Cdx2* expression. In the inner cells, YAP is phosphorylated and cytoplasmic which suppresses *Tead4* activity and inhibits trophectoderm differentiation **(Figure 1.3)** (Nishioka et al., 2009; Piccolo et al., 2014).

Phosphorylated YAP is marked for ubiquitination and proteasomal degradation. This degradation occurs in the cytoplasm where phosphorylated *YAP/TAZ* forms part of a complex with AXIN, APC (adenomatous polyposis coli), glycogen synthase kinase 3 (GSK3) and Beta-Catenin. This complex creates an important interaction between the Wnt and Hippo signalling pathways where GSK3 regulates the ubiquitination and degradation of Beta-

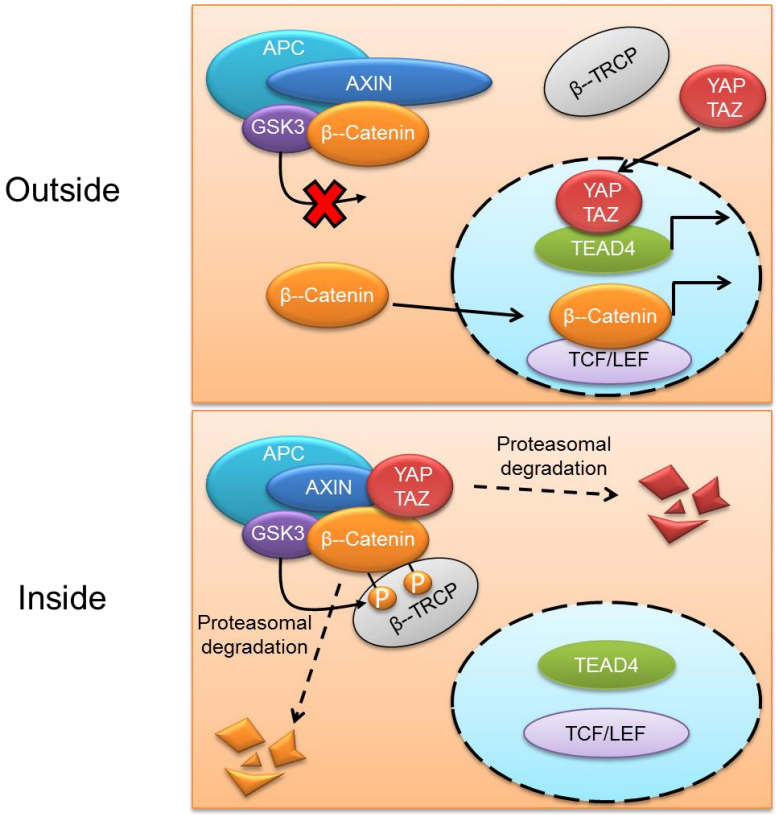


Figure 1.3 WNT and Hippo Signalling in preimplantation development.

Phosphorylated YAP/TAZ and Beta-Catenin are marked for ubiquitation and proteasomal degradation. Phosphorylated YAP/TAZ and Beta-Catenin form a degradation complex with axin, APC (adenomatous polyposis coli) and glycogen synthase kinase 3 (GSK3). This complex creates an important interaction between the Wnt and Hippo signalling pathways as GSK3 regulates the ubiquitation and degradation of Beta-catenin. Beta-transducin repeats-containing proteins (beta-TrCP) is the substrate recognition subunit of E3 ubiquitin ligase.

The Hippo signalling pathway is differentially active between the inner and outer cells of the preimplantation embryo. In the outer cells, which are committed to trophectoderm, the Hippo pathway is inactive with the opposite pattern observed in the inner cells. The transcription factor *Tead4* is essential for specifying the trophectoderm but is modulated by the coactivator protein *YAP*.

In the presence of a Wnt ligand or GSK3 inhibition (outer cells), *β*-catenin remains unphosphorylated and translocates to the nucleus where it binds to *TCF* (T-cell factor) / *LEF* (lymphoid-enhancer binding factor) target genes. Similarly, *YAP* translocates to the nucleus and interacts with *Tead4* initiating trophectoderm specification. In the inner cells, *Yap* is phorphorylated and cytoplasmic which suppresses *Tead4* activity and inhibits trophectoderm differentiation.

Adapted from Piccolo et al., 2014.

catenin (Azzolin et al., 2014). In the presence of a Wnt ligand or GSK3 inhibition, β-catenin remains unphosphorylated and translocates to the nucleus where it binds to *TCF* (T-cell factor) / *LEF* (lymphoid-enhancer binding factor) target genes. Similarly, the GSK3 inhibtion induced dissociation of the degradation complex allows nuclear translocation of YAP and interaction with *Tead4* initiating trophectoderm specification. Despite the classic teaching that Hippo signalling is regulated by the phosphorylation status of YAP as described above, recent studies have discovered an additional level of complexity in the Hippo signalling network whereby stabilisation of AXIN can regulate the Hippo pathway independent of YAP phosphorylation (Wang et al., 2015). Additionally, it has been shown that AMOT family proteins can suppress YAP activity by translocating YAP from the nucleus into the cytoplasm through direct protein-protein interactions independent of YAP phosphorylation status (Chan et al., 2011; Wang et al., 2011a; Zhao et al., 2011). YAP null embryos do not survive beyond E8.5 secondary to defective placentation specifically, failure of chorioallantoic fusion which is the critical event necessary for the establishment of the site of exchange between the maternal and fetal vasculature (Morin-Kensicki et al., 2006).

Notch signalling is active in pre-implantation embryos and acts in parallel with Hippo, both of which converge on the first cell fate specification. In the outer cells, which are committed to trophectoderm, the Notch signalling pathway is active and the Hippo pathway is inactive with the opposite pattern observed in the inner cells (Rayon et al., 2014). This observation was further supported from observations in Drosphilia where inhibition of Hippo resulted in apical accumulation of Notch receptors (Genevet et al., 2009). Beyond the reciprocal relationship that exists between Hippo and Notch in promoting trophectoderm versus ICM specification, it is possible that Notch signalling may influence the initial spatial organization of blastomeres prior to the first cell fate decision. This hypothesis comes from the observation that forced activation of Notch caused inner cells to relocate to an outside position and adopt a

trophectoderm fate. This happened independent of *Cdx2* suggesting that Notch may be acting on additional pathways to influence lineage segregation (Rayon et al., 2014).

The mitogen activated protein kinase (MAPK) signalling network is active in pre-implantation embryos. An overview of the MAPK signalling components active in pre-implantation development is illustrated in **Figure 1.4.** In keeping with the polarization model, Extracellular signal-Regulated Kinases 2 (ERK2), the downstream effector of Ras– MAPK signalling, is asymmetrically expressed in the apical membranes of the eight-cell stage mouse embryo. Inhibition of MAPK signalling reduced the trophectoderm outgrowth and *Cdx2* expression from embryo explants (Lu et al., 2008). Additional components of the MAPK pathway including p38 and JNK have also been implicated in pre-implantation development. Inhibition of these two pathways compromises blastocele formation at the time of the first cell fate decision (Maekawa et al., 2005). The importance of the MAPK kinase pathway in the first cell fate decision was further highlighted in a recent publication that described the interaction of calcineurin-NFAT signalling with Ras-Erk-AP-1 to activate Src which is the terminal downstream effector of the MAPK pathway. In eight-cell stage mouse embryos that were treated with a Src inhibitor, development was arrested at the morula-stage before the first cell fate decision (Li et al., 2011).

The PI3-kinase / Akt pathway is active in pre-implantation mouse embryos and plays an important general role in the promotion of survival and prevention of apoptosis (Gross et al., 2005; Riley et al., 2005). Subunits of the PI3-kinase/Akt pathway are detectable from the one-cell stage up until formation of the blastocyst. Interestingly, phosphorylated PI3K and

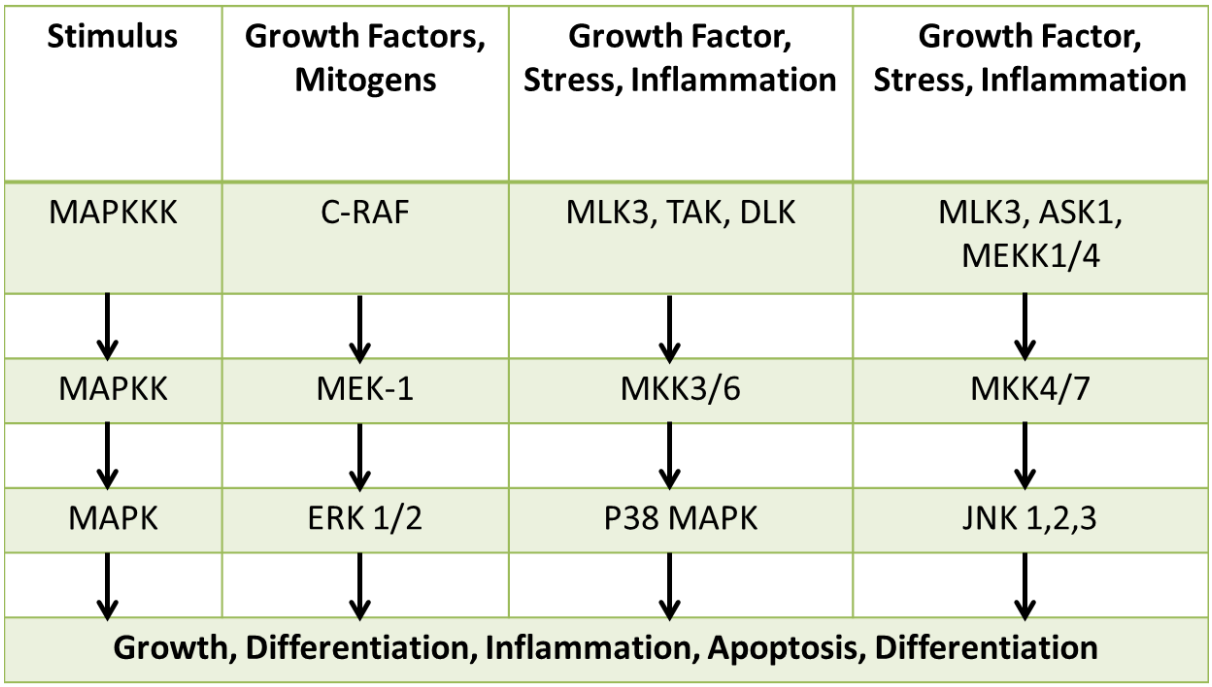


Figure 1.4 Overview of mammalian MAPK signalling

Mammalian Mitogen Activated Protein Kinase (MAPK) signalling begins at the cell membrane where extracellular stimuli (growth factors, mitogens, stress and inflammation) result in the phosphorylation and activation of MAPKKK which subsequently phosphorylates and activates MAPKK and finally MAPK. MAPK orchestrates the cellular response and drives context specific gene expression in response to the initiating extracellular cue. In the preimplantation embryo, at least three distinct MAPK signalling pathways are active, ERK2, JNK and p38 kinase. Ultimately, these pathways converge on Src which is the terminal downstream effector of the MAPK pathway in the preimplantation embryo.

Akt demonstrate an apical distribution pattern in trophectoderm and inhibitors of this signalling pathway not only promote apoptosis but also result in delayed blastocyst hatching (Riley et al., 2005). A functional interaction between Hippo and PI3K/Akt was identified in Drosophila where Akt is negatively regulated by Hippo signalling resulting in growth inhibition (Ye et al., 2012). Although studies are needed to confirm, this would fit with the observed asymmetric activity of PI3K/Akt in the outer cells where the Hippo pathway is known to be inactive.

Bone morphogenetic protein (BMP) signalling is active in pre-implantation mouse embryos from as early as the four-cell stage embryo and is involved in regulating blastomere cleavage divisions (Reyes de Mochel et al., 2015). Moreover, BMPs are required for the appropriate specification of the extra-embryonic lineage including both the trophectoderm and primitive endoderm but interestingly not the embryonic lineage (Graham et al., 2014).

We can conclude that the first cell fate decision is complex with various hypothetical models proposed in an attempt to reconcile the experimental observations to date. The inside-outside and polarization models appear to provide the best explanation to date accounting for asymmetric cell divisions and localisation of transcription factors and proteins known to be essential to trophectoderm specification. However, the shortcoming of this model and all remaining hypothetical frameworks is that the process that dictates blastomere location and acquisition of polarity whether it is deterministic or stochastic remains uncertain.

1.5 The second lineage specification

The second cell fate decision occurs approximately 24 hours after formation of the blastocyst when the inner cell mass segregates into the epiblast and primitive endoderm. The epiblast is

embryonic and will give rise to all somatic cell types in the developing organism. The primitive endoderm is extra-embryonic and gives rise to the developing yolk sac. The primitive endoderm is further divided into the parietal and visceral endoderm with the anterior visceral endoderm playing a major role in anterior patterning of the developing fetus (Beddington and Robertson, 1999).

In the first cell fate decision, *Cdx2* was demonstrated to be a key transcription factor in determining successful trophectoderm specification. The homeobox transcription *Nanog* and the zinc finger transcription factor *Gata6* are the key determinants in the second cell fate decision. *Nanog* null embryos fail to develop an epiblast and only differentiate to parietal endoderm (Chambers et al., 2003; Mitsui et al., 2003). Conversely, *Gata6* null embryos fail to produce visceral or parietal extra-embryonic endoderm (Morrisey et al., 1998).

Similar to the first cell fate decision and reflecting the complexity of the process, there are competing theories to explain the second lineage specification. The first theory pre-supposes that every cell in the inner cell mass has bipotential and can give rise to either the epiblast or primitive endoderm. As an abstraction of the previously described inside-outside model, it was concluded that cells in immediate proximity to the blastocele cavity would become the primitive endoderm and the cells in proximity to the polar trophectoderm would become the nascent epiblast **(Figure 1.5 Positional Epi-PrE segregation)** (Yamanaka et al., 2006).This hypothesis was based on *in vitro* observation of the differentiation behaviour of embryonic stem cells and embryonal carcinoma cells in embryoid body assays (Martin and Evans, 1975).

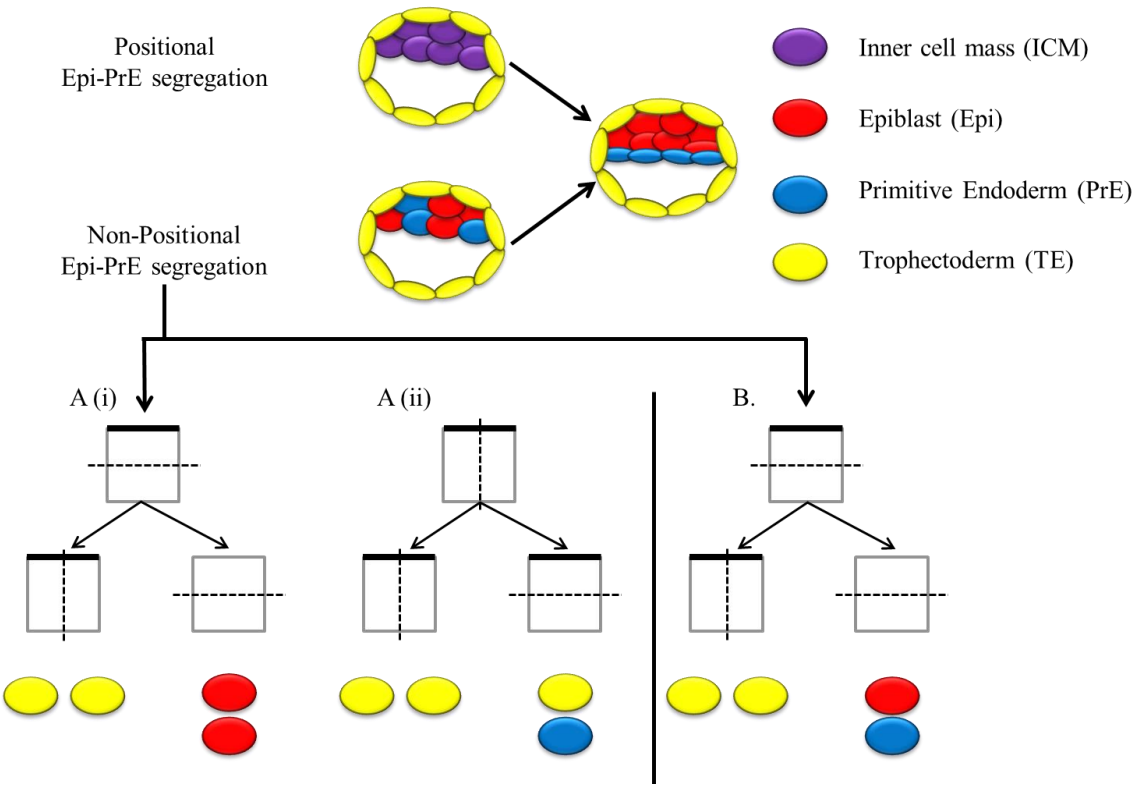


Figure 1.5 Second cell fate decision: Epiblast versus Primitve Endoderm.

Positional Epiblast-Primitive endoderm (Epi-PrE) segregation pre-supposes that every cell in the inner cell mass has the capacity to form either the epiblast or primitive endoderm and that eventual cell fate will be decided based on the cells location in the inner cell mass. Cells in immediate proximity to the blastocele cavity will become the primitive endoderm and conversely, the cells in proximity to the polar trophectoderm will become the nascent epiblast.

Non - positional Epi-PrE segregation suggests that randomly arranged individual cells in the inner cell mass as early as E3.5 already possess a transcriptional profile of the epiblast or primitive endoderm so final fate determination is independent of the cells location in the ICM.

The cleavage hypothesis suggests that there are two waves of inner cell production in the morula at the 8 to 16 cell and 16 to 32 cell stages which introduces heterogeneity into the inner cell mass (A (i) and (ii)). Non-polarized inner cells from the first division are more likely to contribute to the epiblast (A(i)). Inner cells from the second division preferentially develop into primitive endoderm (A (ii)). The alternative model (B) assumes that epiblast - primitive endoderm specification is stochastic with it being entirely random which specification a cell will ultimately adopt. Adapted from Yamanaka et al., 2006.

In recent years, single cell transcriptional profiling and lineage tracing experiments generated evidence which appeared to cast doubt on the preceding hypothesis. It was found that as early as E3.5 the inner cell mass is heterogeneous with a 'salt and pepper' distribution of *Nanog* and *Gata6* in individual cells (Chazaud et al., 2006). Beyond the expression of these two genes, single cell microarray analysis demonstrated that stochastically arranged individual cells in the inner cell mass at this embryonic time point already had a global transcriptional profile of the epiblast or primitive endoderm, but not both **(Figure 1.5 Non-positional Epi-PrE segregation)** (Kurimoto et al., 2006; Yamanaka et al., 2006).

Mechanistically, this may provide an explanation for the actual process of cell specification; but, how does an individual cell in the inner cell mass become selected for an epiblast or primitive endoderm phenotype? This remains unknown, but there are two theories which I will briefly outline. The basis of both hypotheses is that differential activation of Grb2-Ras-MAP kinase signalling will induce primitive endoderm differentiation. In FGFR2/FGF4 null embryos, primitive endoderm does not develop (Arman et al., 1998; Feldman et al., 1995). Grb2 is an adaptor protein essential in this pathway of receptor tyrosine kinase mediated activation of MAPK signalling. Grb2 null embryos phenocopy the FGFR2/FGF4 null embryos with no primitive endoderm formation (Chazaud et al., 2006). The cleavage hypothesis argues that the two waves of inner cell production in the morula at the 8 – 16 cell and 16 – 32 cell stage introduces heterogeneity into the inner cell mass **(Figure 1.5A(i) and (ii)**) (Yamanaka et al., 2006). Inner cells from the first division are more likely to contribute to the epiblast and; therefore, have lower Grb2-Ras-MAPK activity. Whereas, inner cells from the second division have higher activity of this pathway and preferentially develop into primitive endoderm. The basis for this model comes from experimental observation on the spatial distribution of the extra-embryonic trophectodermal cytokeratin, ENDO A, which was found to derive almost entirely from the inner cells from the second division (Chisholm and

Houliston, 1987). The second model, is much more simplistic and assumes that the process of epiblast / primitive endoderm specification is stochastic with it being entirely random which specification a cell will ultimately adopt **(Figure 1.5B)** (Yamanaka et al., 2006).

1.6 Derivation of stem cells from the mouse embryo

By the time of implantation, three distinct cell lineages will have formed in the developing embryo. On the outside, enclosing the embryo lie the trophectoderm which is divided into mural and polar components. The mural trophectoderm surrounds the blastocele cavity and will have a key role in implantation and formation of primary giant cells. The polar trophectoderm located adjacent to the inner cell mass which give rise to the extra-embryonic ectoderm and ectoplacental cone. The inner cell mass consists of the epiblast and hypoblast with the epiblast located adjacent to the polar trophectoderm giving rise to all somatic cell types in the developing organism including the germ cell lineage. The hypoblast located adjacent to the blastocele cavity will form the primitive endoderm from which the yolk sac will develop (Yamanaka et al., 2006). From each of these three anatomical sites we can derive stem cells which retain the expression profile and growth factor requirements characteristic of the lineage from which it was derived (Evans and Kaufman, 1981; Kunath et al., 2005; Martin, 1981; Tanaka et al., 1998). These stem cells have provided a biological resource to study the factors that induce and maintain their respective cellular identities.

1.6.1 Embryonic stem cells

A defining moment in stem cell biology occurred when two independent groups discovered that *in vitro* culture of the epiblast from the pre-implantation mouse blastocyst on a mitotically inactivated feeder layer in the presence of serum allowed derivation *in vitro* of

embryonic stem cells lines (Evans and Kaufman, 1981; Martin, 1981). These embryonic stem cells were pluripotent defined by their ability to differentiate into all somatic lineages.

Since their discovery, experiments on mouse embryonic stem cells have revolutionized our understanding of stem cell biology, the process of self-renewal and the mechanisms controlling differentiation. Genetic manipulation of these cells including the development of knockout technology facilitated the modelling of disease *in vitro*.

The culture condition in which embryonic stem cells were first derived was heterogeneous and the factors responsible for self-renewal were initially unknown. Dissection of the individual components within the culture identified Leukaemia inhibitory factor (LIF) as the key component mediating self-renewal in the feeder layer (Smith et al., 1988; Williams et al., 1988). LIF mediated its effect through downstream activation of signal transducer and activator of transcription 3 (STAT3) (Niwa et al., 1998). Bone morphogenetic protein 4 (BMP4) was identified as the key component in serum which suppressed differentiation through activation of *Id* (inhibitor of DNA binding) genes. The combination of LIF and BMP4 allowed *in vitro* maintenance of undifferentiated ES cells in defined conditions (Ying et al., 2003).

Oct4 and *Sox2* are at the centre of the pluripotency transcriptional network as embryonic stem cells cannot be derived from *Oct4* and *Sox2* null embryos (Avilion et al., 2003; Nichols et al., 1998). Additional transcriptional factors that enhance self-renewal but which are individually dispensable include *Nanog, Klf4*, and *Esrrb* (Chambers et al., 2007; Martello et al., 2013; Martello et al., 2012). In conventional ESC culture conditions many groups observed heterogeneity in transcription factor expression suggesting sub-populations existed within the culture. For example, sub-populations expressing high and low levels of *Nanog* and *Rex1*

were identified (Chambers et al., 2007; Toyooka et al., 2008). *Nanog*-low cells were predisposed to differentiation and *Rex1*-low cells were unable to contribute to the developing embryo following microinjection to blastocyst-stage embryos (Chambers et al., 2007; Toyooka et al., 2008). This observation suggested that our knoweldge of the factors necessary for the self-renewal of a homogenous population of ES cells was sub-optimal with the existing culture conditions permitting the existence of multiple metastable states. The next major advance came with the discovery of a culture condition that maintained a homogenous ES cell culture at the so-called 'ground-state' of self renewal (Ying et al., 2008).

The maintenance of self-renewal is dependent on the inhibition of differentiation inducing signals (Burdon et al., 1999). In ES cells, FGF4 acts in an autocrine fashion to mediate exit from the self-renewal program and initiate lineage committement (Kunath et al., 2007). The donwstream effector of FGF signalling is extracellular-signal-regulated-kinase (ERK) (Kunath et al., 2007). Studies *in vivo* demonstrated that inhibition of FGF signalling in morula-stage embryos blocked formation of the hypoblast and expanded the epiblast compartment (Nichols et al., 2009b). In mouse ES cells treated with MEK (MAPK-ERK) 1/2 inhibitors, differentiation was blocked but clonal expansion was severely compromised. The addition of a glycogen synthase kinase (GSK) inhibtor rescued the clonogenicity and permitted stable derivation and expansion *in vitro* of germline competent mouse ES cells (Ying et al., 2008).

Beyond the anabolic effects, GSK3 plays a critical role in the canonical Wnt pathway. GSK3 regulates the ubiquitination and degradation of Beta-catenin. In the cytoplasm, GSK3 phorphylates β-catenin, marking it for degradation where it forms part of a complex with AXIN and APC (adenomatous polyposis coli). In the nucleus, β-catenin interacts with TCF (T-cell factor) / LEF (lymphoid-enhancer binding factor) target genes. TCF3 is a

transcriptional repressor. In the presence of GSK3 inhibition, β-catenin remains unphosphorylated and translocates to the nucleus where it binds to TCF3 and prevents it binding at key pluripotency loci. By repressing a transcriptional repressor, GSK3 inhibtion serves to promote ES cell self-renewal (Wray et al., 2011).

This combination of a MEK1/2 inhibitor (PD0325901) and a GSK3B inhibitor (CHIR99021) was named 2i media (Nichols et al., 2009a; Nichols and Smith, 2009; Wray et al., 2010; Ying et al., 2008). 2i media permitted the derivation of ES cell lines from STAT3 null embryos so the authors argued that extrinsic stimuli are dispensable for the maintenance of pluripotency and that ES cells in this culture condition have 'an innate programme for self-replication that does not require extrinsic instruction' (Ying et al., 2008). In stark contrast to serum, 2i cultured ES cells had minimal expression of lineage specifiers and reduced amounts of repressive histone modifications at key promoters (Marks et al., 2012). Furthermore, the culture was homogenous with no apparent subpopulations identified. ES cells cultured in 2i based media are referred to as being in the 'ground-state' of self-renewal. Using 2i media, germline competent ES cells were also derived from rat embryos (Buehr et al., 2008; Li et al., 2008). This exciting finding raised the possibility that the ground-state may be a conserved inter-species mechanism of self-renewal. If this were true, it may be possible to derive naïve ES cells from human embryos as human embryonic stem cells are felt to be more representative of a 'primed' pluripotency state.

Primed pluripotency traditionally referred to embryonic stem cells derived from the post-implantation mouse blastocyst referred to as epiblast stem cells (EpiSC) (Brons et al., 2007; Tesar et al., 2007). These cells although pluripotent and with expression of core pluripotency transcription factors (*Oct4, Sox2, Nanog*) were categorically distinct from ES cells derived from pre-implantation embryos. EpiSCs were derived and cultured in the presence of FGF

and Activin, had low clongenicity and although pluripotent *in vitro* and in teratoma assays, these cells contributed poorly in blastocyst chimeras. Epigenetically, female EpiSCs had inactivated one of the X-chromsomes, reflecting dosage compensation during differentiaiton. Naïve mouse ES cells possessed two active X-chromosomes (Nichols and Smith, 2009).

Human ES cells are also derived and cultured in the presence of FGF and possess many similarities with the mouse EpiSC state (Tesar et al., 2007; Thomson et al., 1998). Human ES cells in this 'primed' state are often difficult to culture *in vitro* with a low clonogenicity which hampered genetic manipulation and single cell passaging. X-linked disease modeling was also limited due to the presence of X-chromsome inactivation in human cells (Mekhoubad et al., 2012). The ability to derive human embryonic stem cell with properties more akin to the mouse ground-state would facilitate the use of these cells for disease modelling, gene modification / correction and cell based regenerative medicine.

.

1.6.2 Trophoblast stem cells

The mouse trophoblast stem cell niche is thought to exist in the extra-embryonic ectoderm / ectoplacental cone portion of the polar trophectoderm. The reciprocal expression of FGF2 receptor and its ligand FGF4 led to the hypothesis that embryonic fibroblast growth factor signalling may be promoting trophoblast development (Feldman et al., 1995; Orr-Urtreger et al., 1993). Indeed, self-renewing trophoblast stem cell lines can be derived by *in vitro* culture of E3.5 or E6.5 embryo explants in media containing FGF4, heparin, serum and activators of Activin / Nodal signalling (Tanaka et al., 1998). These trophoblast stem cells are indefinitely self-renewing and demonstrate multipotency *in vitro* following the removal of FGF4 which permits terminal trophoblast differentiation (Tanaka et al., 1998) . Gene expression profiling during *in vitro* differentiation demonstrated an appropriate temporal-spatial pattern with upregulation of *Mash2* known to be expressed in the ectoplacental cone, followed by *Hand1*

in developing giant cells and finally *placental lactogen 1 (Pl-1)* in mature giant cells (Cross et al., 1995; Faria et al., 1991; Guillemot et al., 1994). Moreover, in chimeric embryos, the trophoblast stem cells contributed exclusively to the placenta and secondary to a high percentage of contribution, these cells were able to support normal development to term indicating their functionality in the mature placenta (Tanaka et al., 1998).

Fibroblast growth factor signalling is essential for the maintenance of trophoblast stem cell identity. Targeted deletions of FGF4, FGFR2 or Erk2/MAPK1 result in embryonic lethality due to inability to maintain the trophoblast stem cell niche (Arman et al., 1998; Feldman et al., 1995; Saba-El-Leil et al., 2003). Importantly, however, MAPK is not required for the initial specification of the trophoblast lineage as implantation can still occur in these respective knockouts (Arman et al., 1998; Feldman et al., 1995). As previously discussed, trophoblast specification is thought to reflect differential blastomere polarization, asymmetric activation of signalling pathways including Hippo and Notch with resulting activation of downstream trophoblast lineage specific transcription factors including *Tead4* and *Cdx2*.

Investigating the transcriptional network in mouse trophoblast stem cells allowed an understanding to be attained of the reciprocal inhibition that exists between the inner cell mass and the trophectoderm. Mouse trophoblast stem cells cannot be derived from *Cdx2* null embryos reinforcing the importance of this transcription factor in the maintenance of the niche (Strumpf et al., 2005). Embryonic stem cells cannot be derived from *Oct4* null embryos which again indicated the central role of *Oct4* in maintaining pluripotency (Nichols et al., 1998). In *Oct4* null embryos there is initial specification of the trophectoderm but the trophoblast stem cell compartment does not expand. This phenotype can be rescued by the addition of FGF4 demonstrating the paracrine interaction that exists between the developing inner cell mass and trophectoderm (Nichols et al., 1998).

Many experiments have demonstrated the interaction of *Oct4* and *Cdx2* in ICM versus TE specification. Deletion or repression of *Oct4* expression in ES cells induces differentiation to trophectoderm (Niwa et al., 2000). Similarly, overexpression of *Cdx2* induces trophectoderm differentiation in ES cells (Niwa et al., 2005). This reciprocal inhibitory interaction has been shown to occur at the protein level where inhibition of OCT4 by CDX2 will positively reinforce its own expression and permit trophectoderm specification (Niwa et al., 2005). As OCT4 and CDX2 are initially co-expressed in blastomeres at the morula-stage, this finding may explain how initial lineage segregation occurs (Strumpf et al., 2005).

1.6.3 Extra-embryonic endoderm (XEN) stem cells

Mouse extra-embryonic endoderm cell lines have been derived from blastocyst-stage embryos (Kunath et al., 2005). *In vitro,* these cells lines have a gene expression profile similar to their *in vivo* counterparts and *in vivo* chimeric experiments show exclusive contribution to the extra-embryonic visceral and parietal endoderm. Surprisingly, FGF was not always necessary in the derivation or maintenance of these cells (Kunath et al., 2005). After derivation, XEN cells can be grown feeder free in medium supplemented with EMFI-CM. At present, the exact signalling requirements that are necessary for optimal expansion of these cells remains unknown (Kunath et al., 2005).

1.6.4 Hierarchy of potency – is there an *in vitro* totipotent-like stem cell?

The preceding sections have described the derivation *in vitro* of self-renewing multipotent populations from the three lineages produced during the first and second cell fate decisions. The logical next question to consider is, if a self-renewing state can be isolated from these three lineages is it therefore possible to derive a ‘master’ or ‘totipotent-like’ stem cell? Such a

cell line should be relieved from the lineage restriction of the preceding stem cells states and acquire expanded potential to contribute to all cellular lineages in embryonic development.

It remains unclear whether it is possible to experimentally derive and maintain stable cell lines from early embryos that contain features of cleavage-stage blastomeres. Recently however, a transient subpopulation in ESC culture that is thought to correspond to totipotent two-cell (2C)-stage blastomeres was reported (Macfarlan et al., 2012). These 2C-like ESCs do not express pluripotency factors such as POU5F1 (OCT4), SOX2 or NANOG protein and demonstrate the ability to contribute to both embryonic and extra-embryonic tissues. The authors argue that entry into this 'privileged state' may partially be controlled by histone-modifying enzymes implicating epigenetics as a potential determinant of totipotency.

Similarly, another group reported a rare *Hex* positive subpopulation in ESCs cultured in 2i/LIF, a defined medium that maintains ESCs at an intrinsically stable ground-state (Wray et al., 2010; Ying et al., 2008) which could also apparently contribute to both trophoblast and embryonic lineage development (Morgani et al., 2013). The authors proposed leukaemia inhibitory factor (LIF) as the cytokine responsible for the observed phenotype as the percentage of proliferating HEX positive cells increased in response to increasing concentrations of LIF. Traditionally, LIF is known for its critical role in the self-renewal of serum-cultured mouse ESCs due to its ability to activate the transcription factor STAT3 (Niwa et al., 1998). In serum conditions, STAT3 null ESCs rapidly succumb to differentiation (Ying et al., 2008). Although dispensable for the maintenance of pluripotency in 2i cultured mouse ESCs, LIF is routinely added to improve cell growth and clonal expansion (Ying et al., 2008). As LIF was found to enrich for the HEX positive subpopulation, the authors presented data to implicate LIF as a potential explanation for supporting 'extra-embryonic priming' (Morgani et al., 2013). Additional support for the authors claim came from the observation

that increasing LIF concentrations upregulated the expression of the *Hex* transgene and STAT3 -/- ESCs had decreased extra-embryonic gene expression. Moreover, whole genome expression analysis confirmed that the upregulated gene set in LIF containing media showed the strongest correlation with placental gene expression (Morgani et al., 2013).

A system-wide analysis comparing mouse ESCs cultured in 2i media in the presence or absence of LIF offered further support to the potential role of LIF in mediating extra-embryonic priming by demonstrating relative enrichment of *Zscan4 (zinc finger and SCAN domain containing 4)* expression in the presence of LIF (Cerulo et al., 2014). The *Zscan4* gene cluster has been described as a marker for the 2-cell embryo and mouse ESCs (Falco et al., 2007). One in twenty cells in an *in vitro* ESC culture will express *Zscan4* at any given time. This *Zscan4* positive state is associated with the maintenance of genomic integrity and long term developmental potency (Amano et al., 2013; Zalzman et al., 2010). Repression of *Zscan4* expression in ESCs results in a shortening of telomere length, an increased number of karyotypic abnormalities and a quantitative reduction in cellular proliferation rates (Zalzman et al., 2010). Overexpressing this gene cluster on the other hand, produced ESCs that could contribute to a whole mouse at a much greater frequency than wild type ESCs in tetraploid complementation assays (Amano et al., 2013).

The aforementioned papers led one to the following conclusions. Firstly, cells with expanded potential exist in ESC culture conditions albeit transiently and at extremely low percentages of the total cell population. Secondly, this infrequency may be explained by the fact that the culture condition is not optimized for the capture and expansion of these rare cell types. As stem cell populations can be established from the epiblast, hypoblast and polar trophectoderm through the addition of chemicals that suppress extrinsic developmental cues and promote the inherent self-renewal programme, it may be reasonable to speculate; that if signalling

pathways that are known to be active and critical for pre-implantation development could be selectively targeted using small molecule inhibitors, one may be able to enrich for or capture these cells with expanded potential from standard ESC cultures in the first instance.

1.7 Is the pre-implantation developmental process conserved between mouse and human?

Human pre-implantation development at first glance appears to behave similar to the mouse. Following fertilization, there is the onset of zygotic gene activation (ZGA) followed by blastomere cleavage divisions which ultimately result in the production of the trophectoderm (TE) and inner cell mass (ICM), similar to the mouse. In fact, much of our understanding of human pre-implantation development is from extrapolation from studies on mouse due to the scarcity of available human embryonic material to study and often, the ethical difficulties associated with its use. A potential milestone in human embryo research occurred recently when two independent groups demonstrated that the early stages of post-implantation human development could be successfully modelled *in vitro*. This amazing feat occurred due to the proposed self-organizational capacity of the human embryo. In the absence of maternal proteins, early post-implantation development was observed with expansion of the epiblast, formation of the bilaminar disc, amniotic and yolk sac cavities and trophoblast differentiation. This model, if reproducible, conceivably in the future will permit identification of the key pathways essential for the correct spatio-temporal progression of human development (Deglincerti et al., 2016; Shahbazi et al., 2016).

So how similar is the process of pre-implantation development in mouse and human? Firstly, even though the spatial pattern of pre-implantation development is similar, there are very distinct temporal differences as regards the timing of blastocyst formation and subsequently

its implantation. In mouse, the blastocyst is formed by E3-3.5 dpc; however, in humans, this doesn't occur until Day 5. Secondly, the mouse blastocyst implants at E4-4.5 but in humans, this process is protracted and doesn't occur until Day 7 to 9 (De Paepe et al., 2014). Furthermore, the process of implantation is different. Human embryos invade into the lining of the uterine wall called the endometrium which is termed interstitial implantation. In contrast, the mouse embryo does not invade the endometrium but merely attaches to it in a process called secondary interstitial implantation (De Paepe et al., 2014; James et al., 2012).

But, do these differences reflect a fundamental biological dissimilarity or rather, can they potentially be explained by the larger size of the human and therefore, the need for more cell divisions to produce an adequate number of cells for successful implantation.

Although there are known species differences in the timing of embryonic gene activation (EGA) and the epigenetic regulation of X-chromosome inactivation (XCI) in the pre-implantation embryo, available studies appear to suggest that the fundamental processes involved in the first cell fate specification may in fact may conserved. Global mRNA expression data from surgically isolated human ICM and TE found expression patterns to suggest that the first cell fate decision is controlled by the same signalling pathways that are active in the mouse including WNT, mitogen-activated protein kinase (MAPK) and NOTCH (Adjaye et al., 2005). Additional evidence to support the claim that developmental processes may be conserved is the shared localisation pattern of transcription factors in the ICM *(Nanog, Oct4, Rex1*) and TE (*Cdx2*) and epigenetically, both species possess global hypomethylation in the TE as compared to the ICM (Adjaye et al., 2005; Santos et al., 2002; Santos et al., 2010; Senner and Hemberger, 2010).

Beyond the specification event, it has been argued that the factors involved in establishment of the trophoblast lineage also appear to be conserved between mouse and man (Senner and Hemberger, 2010). Firstly, there is shared mutual activation of the transcription factors *Cdx2*, *Elf5* and *Eomes* which are thought to represent the trophoblast equivalent of the core pluripotency transcription factor network active in embryonic stem cells (ESCs) (Ng et al., 2008). Secondly, the transcription factors involved in controlling terminal trophoblast differentiation, *GATA2/3* and *TCFAP2A/C* are shared (Senner and Hemberger, 2010). Finally, the epigenetic barriers involved in the control of trophoblast lineage specification are also the same. In mouse and human this epigenetic barrier is *ELF5* hypomethylation which is specific to the trophoblast lineage (Hemberger et al., 2010).

Traditionally, an argument against the conservation of early embryonic states between humans and rodents was the inability to derive ground-state ESCs from human embryos. Established human ESCs differed from mouse ESCs in many biological aspects and were considered to be more akin to mouse EpiSCs (Tesar et al., 2007; Thomson et al., 1998; Yu and Thomson, 2008). However, challenging this hypothesis, several groups have recently reported multiple *in vitro* conditions to maintain human ESCs with transcriptional and metabolic similarities to mouse ESCs in 2i/LIF conditions (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014).

1.8 Thesis Hypothesis and Objectives

The experimental hypothesis was that through chemical modulation of pathways implicated to be involved in mouse pre-implantation development it may be possible to establish *in vitro* a homogenous stem cell state that possesses expanded fate potential to both the embryonic and

extra-embryonic lineage. Secondly, assuming inter-species conservation of early embryonic signalling networks, a similar state may be identifiable in human.

1.8.1 Mouse

I would therefore apply a hypothesis driven chemically defined media to established cultures of mouse ESCs, individual blastomeres and reprogrammed somatic cells in an attempt to establish a stable stem cell line with expanded potential. Expanded potential would be defined based on the cell's ability to differentiate into the embryo proper and both extra-embryonic lineages, the primitive endoderm and trophectoderm. Molecular analyses including single cell RNA sequencing would also be performed to discern whether these cells were enriched in a pre-implantation transcriptional signature.

1.8.2 Human

If successful in mouse, I would then apply the same chemical formulation to established cultures of human ESCs and reprogrammed somatic cells in an attempt to establish a stem cell line with expanded potential. In human, expanded potential would be defined based on the cell's ability to differentiate into all somatic lineages of the embryo proper and trophectoderm *in vitro*. Epigenetic profiling and molecular analyses including RNA sequencing would also be performed to discern whether these cells were enriched in a pre-implantation signature.

CHAPTER 2

2 MATERIALS AND METHODS

2.1 Expanded Potential Stem Cell Medium (EPSCM)

EPSCM media contained DMEM/F12 (Invitrogen), 20% KnockOut serum replacement (Invitrogen), 1X MEM Non-essential amino acid (Invitrogen, 11140), 1X Glutamine-Penicillin-Streptomycin (Invitrogen, 10378), 0.1mM β-ME (Sigma), PD0325901 1μM, CHIR99021 5μM, 1 x 10^6 U/mL human LIF, Recombinant human IGFII 30 ng/mL (R&D, 292-G2), JNK Inhibitor VIII 4μM (Santa-Cruz, sc-202673), SB203580 Hydrochloride 10μM (Tocris, 1402/10), A419259 Trihydrochloride 0.3μM (Santa-Cruz, sc-361094) and XAV939 5μM (Sigma, X3004).

2.2 Mouse Cell Culture

2.2.1 Mouse ESC / iPSC Culture Media

Standard mouse ESCs were cultured in M15 medium which contained Knockout DMEM (Invitrogen), 15% FBS (HyClone), 1X MEM Non-essential amino acid (Invitrogen, 11140), 1X Glutamine-Penicillin-Streptomycin (Invitrogen, 10378), 0.1 mM β -ME (Sigma, M6250), and 1 x 10^6 U/mL recombinant LIF. Alternatively, 2i/LIF media was used which contained NDiff 227 (Stem Cells, SCS-SF-NB-02), CHIR99021 3μM, PD0325901 1μM and 1 x 10^6 U/mL LIF. PD0325901, CHIR99021 and LIF were purchased from Cambridge Stem Cell Institute. Standard ESCs used in this study were AB2.2 (Hooper et al., 1987), E14Tg2a (Ramirez-Solis et al., 1995), *Rex1-GFP* ESCs (Guo et al., 2011) and *Oct4-GFP* reporter ESCs (Silva et al., 2008).

2.2.2 Mouse ESC Expansion and Passaging

EPSCs were cultured feeder-free on 0.1% gelatin coated tissue culture plates. Where indicated, mitomycin C-inactivated SNL feeder cells (STO / SNL) were used. Cells were enzymatically passaged using Accutase (Millipore, SCR005) as per the manufacturer guidelines. For routine expansion, EPSCs were split 1:40.

2.2.3 Mouse ESC Conversion

Conversion of conventional ESCs to EPSCs was performed by culturing ESCs on SNL feeders at low plating density for three passages in EPSCM. The conversion experiments were performed on the publicly available and widely used ES cell lines, AB2.2 and E14Tg2a (Hooper et al., 1987; Ramirez-Solis et al., 1995). To test if EPSCM selected for a rare subpopulation of cells in ESC culture, a single *Rex1-GFP* ESC which was pre-cultured in 2i/LIF or M15 was plated in either EPSCM or in their original media on SNL feeders. The numbers of GFP positive colonies that formed in EPSCM or in 2i/LIF was assessed.

2.2.4 Mouse ESC Cryopreservation

For cryopreservation, EPSCs were resuspended in 500 µl Freezing Buffer (80% EPSCM, 10% FBS (HyClone) and 10% DMSO). Cryovials were placed at -80 degrees Celsius overnight then transferred to liquid nitrogen the following day for long term storage.

2.2.5 Reprogramming and mouse iPSC generation

Reprogramming was performed using the published 6 factor reprogramming platform developed by (Wang et al., 2011b). Primary colonies were established into lines using EPSCM. To reprogram mouse embryonic fibroblasts (MEFs), episomal vectors with *Oct4*, *c-Myc*, *Klf4* and *Sox2* (OCKS 4F, 5µg), and *Rarg*, *Lrh1* (RL 2F, 5.0µg) were first mixed with

1.0 x 10^6 cells in OptiMEM (Invitrogen) and the cells were electroporated with Amaxa Nucleofector (Lonza, Germany). After electroporation, the cells were plated onto 10 cm dishes with STO feeders in M15 for recovery for 24 hours. Medium was changed every two days. At day 7, colonies emerged and at day 14, medium was switched to EPSCM for an additional 6 days. Colonies were picked into 24 wells on STO feeders in EPSCM for expansion and characterization.

2.2.6 *In vitro* Differentiation of mouse ESCs

For three germ layer *in vitro* differentiation, an embryoid body (EB) assay was used. Cells were harvested and transferred to ultra-low attachment 6-well plates (Costar, 3471) at a density of 1 x 10^4 – 1 x 10^5 cells/cm^2. EB differentiation media contained Knockout DMEM (Invitrogen, 10829-018), 15% Fetal Bovine Serum (HyClone), 1X MEM Non-essential amino acid (Invitrogen, 11140), 1X Glutamine-Penicillin-Streptomycin (Invitrogen, 10378) and 0.1 mM β -ME (Sigma, M6250). Media was changed every 48 hours. After 5 days, intact EBs were transferred to gelatinized 6-well plates. Once EBs attached to gelatinized substrate, they were cultured for another 7-10 days before fixation and immunostaining.

In vitro trophoblast differentiation was performed as per the protocol of (Abad et al., 2013). Briefly, cells were plated on mitomycin-C-inactivated fibroblasts (MEFs) at a density of 7×10^4 cells per well of a 6-well plate in EPSCM. After 24 hours, medium was changed to trophoblast stem cell (TSC) medium that contained: 70% MEF conditioned medium (Millipore, AR005), 30% RPMI 1640 (A1049101) with 20% FBS (HyClone), 1mM pyruvate, 2mM L-glutamine, 100μM β-mercaptoethanol, 25 ng/ml FGF4 (R&D Systems, 235-F4-025) and 1 μg/ml of heparin (R&D, 2812). Medium was changed every other day. At day 2, cells were split and differentiation was continued until day 5. Trophoblast stem cell-like cells (TSCLCs) were maintained on MEF feeders in TSC media. For terminal trophoblast

differentiation, TSCLCs and *in vivo* derived trophoblast stem cells were plated on gelatin and cultured in TSC media without heparin and FGF4.

2.2.7 Flow Cytometry

Oct4-GFP and *Rex1-GFP* ESCS were harvested by enzymatic passaging and subsequently centrifuged at 200 × g for 3 min. The medium was aspirated and the ESCs were resuspended in PBS solution and analysed by BD LSRFortessa (BD Bioscience). Wild type unstained E14 ESCs were used as the negative control.

2.2.8 RT-PCR

RNA was isolated using the RNeasy Mini Kit (Qiagen). The samples were subsequently quantified and treated with gDNA WipeOut. First-strand cDNA was prepared by using the QuantiTect Reverse Transcription Kit (Qiagen). For real-time PCR, I used TaqMan Gene Expression Assays. TaqMan probes were purchased from Applied Biosciences. Quantitative PCR was performed in a 7900HT Fast Real-Time PCR System (Applied Biosciences). Mouse gene expression was determined relative to the *Gapdh* gene using the ΔCt method.

2.3 Chemical screen for PD0325901 substitutes in mouse ESCs.

The screen was designed so as to identify potential substitutes for PD0325901 (Mek1 inhibitor). On Day 0, 1 x 10^3 cells were plated in each well of a gelatinized 96-well flat bottomed plate in triplicate. On Day 1, the media was changed to N2B27 plus the GSK3 inhibitor CHIR99021 3μM or N2B27 alone in the presence or absence of the inhibitor of interest. After 4-6 days, 96-well FACS analysis was performed to identify any target compound(s) that could maintain *Rex1-GFP* expression. In each experiment, N2B272i, N2B272i/LIF and N2B27CHIR/LIF were included as positive controls. Cells cultured in

N2B27 alone served as an internal negative control with complete loss of *Rex1-GFP* expression identified. If a potential 'hit' was identified, the compound was first tested in its ability to support undifferentiated ESC self-renewal in the presence of CHIR99021 +/- LIF. Next, the identified pathway was confirmed by testing an additional commercially available selective inhibitor from the same signalling family. The list of chemicals tested is provided in Table S1.

2.4 Pre-implantation Mouse Embryo Culture

2.4.1 Derivation of new lines from pre-implantation mouse embryos

C57B6 females were superovulated and two to four-cell stage embryos were collected and cultured in KSOM (Millipore, MR-121-D) until the 8-cell stage (8C). 8C embryos were transferred to 24-well plates, 1-embryo per well on SNL feeders in M15, 2i/LIF or EPSCM. Partial media change was performed every 48 hours. At day 10, the embryo had attached to the feeder layer and a primary outgrowth was observed. Using enzymatic passaging with Accutase (Millipore, SCR005), each primary outgrowth was split to one well of a six-well SNL feeder plate. Generated lines were expanded for characterization and early cryopreservation was performed. Lines derived from 8C embryos were DR10 (EPSCM), DR25 (EPSCM), DR9 (2i/LIF) and DR47 (2i/LIF).

2.4.2 Derivation of new lines from single blastomeres

For derivation of lines from single blastomeres, 8-cell stage embryos were collected from oviducts of plugged super ovulated females mated by stud males at 1.5 or 2.5 dpc. Before separating the blastomeres, the embryos were cultured in 1.0 ml KSOM containing 3.0 mg/ml BSA and 5.0 μg/ml CB (Sigma, C6762) for at least 30 min. Embryos were suspended in a 100μl Tyrode's Solution (Sigma, T1788) drop for about 15-20 seconds, and then the embryos

were transferred into a new 100μl Tyrode's Solution drop until the zona pellucida had disappeared. The embryos were transferred into a fresh 100 μl M2 drop immediately after zona pellucida removal. The embryos were washed a few times in M2 drops. The embryos were sucked and blown a few times by a blunt needle (inner diameter: 40 μm). Individual blastomeres were collected for seeding to 96-well SNL feeder plates in M15, 2i/LIF or EPSCM.

Genetic backgrounds of ESCs and EPSCs: DR10 and DR25 are from a mixed 129S5 and C57B6 background. Additional EPSC lines from 8C embryos or blastomeres are on a C57B6 background.

2.5 Immunostaining

2.5.1 Embryo immunostaining

Immunostaining of early embryos was performed according to the published protocol by Dr. Jenny Nichols (Cambridge University) with slight modification. Briefly, embryos were washed in PBS / PVP, fixed in 2.5% PFA for 15 minutes at room temperature; permeabilised in PBS / PVP / 0.025% Triton-X-100 for 30 minutes; blocked in PBS / PVP / 0.1% BSA / 0.01% Tween-20 / 2% Donkey serum for 1 hour at room temperature; then incubated with primary antibodies in blocking solution at 4 ℃ overnight; washed with blocking solution 3 times for 15 minutes each; incubated with Alexflor secondary antibodies 1 hour at room temperature; washed with blocking solution 3 times for 15 minutes each; mounted the embryos in a small drop of VectaShield with DAPI; covered with a coverslip and sealed with nail varnish. For the embryo outgrowth, cells in a 96-well were fixed in 4% PFA. Blocking solution is PBS / 0.1%Triton-X-100 / 1% BSA / 3% Donkey serum, the staining procedure is similar to that listed above. After secondary antibody incubation, cell nuclei were stained with

DAPI and observed under fluorescence microscope. The primary antibodies used are Cdx2 (BioGenex); Oct4 (Abcam), Oct4 (Santa Cruz), Cleaved Caspase 3 (Cell Signalling) and Ki-67 (Cell signalling).

2.5.2 Cellular Immunostaining

Cells were grown to the required density on tissue culture plastic for use with fluorescence microscope or chamber slides for use with confocal microscope. Cells were fixed using 4% PFA for 30 mins at room temperature (RT) and subsequently blocked and permeabilised with 3% serum* in PBS (Invitrogen, 14190) + 0.3% Triton-X-100 for 30 minutes. Cells were incubated with the primary antibody that had been diluted to working concentration in 0.3% Triton + 10% serum* in PBS for 2 hours at 37 degrees or overnight at 4 degrees. Secondary antibody was diluted to working concentration in 2% serum* in PBS and cells were incubated for 30 mins at 37 degrees or 1-2 hours at room temperature in the dark. Between each step, cells were washed 3 times with PBS, 5 minutes per wash. Cells were counter stained with DAPI.

Note: * Normal serum from same species that secondary antibody was raised in. See Supplementary Table 4 for list of antibodies that were used in this study.

2.5.3 General Immunohistochemistry

Tissue samples were fixed in 10% buffered formalin phosphate for 24 hours then paraffin-embedded and cut in 3-µm sections. Alternatively, samples were fixed for 24 hours as described previously then placed in 30% sucrose followed by snap freezing in isopentane and embedding in OCT for cryosectioning. Prior to staining, samples were deparaffinised, rehydrated, permeabilized in 0.5% PBS-Triton-X-100 for 5mins and underwent antigen retrieval. This was achieved by incubating slides in boiling 10mM Sodium Citrate Buffer (pH

6.0) for 10mins in a pressure cooker. Samples were blocked for 1 hour in 10% normal goat serum made up in 0.05% PBS-Triton-X-100. Primary antibody in blocking buffer was incubated with sample overnight at 4 degrees Celsius. Conjugated secondary antibody was incubated for 1 hour at room temperature. Counterstaining with VectaShield plus DAPI was performed and the samples were mounted with 30µl 50% PBS-50% glycerol.

2.5.4 Placental and Yolk Sac Immunostaining

Paraffin-embedded E14.5 placenta and yolk sac sections were deparaffinized in Xylene, 100%, 95%, 80%, 70% and 50% ethanol for 3 minutes each and rinsed with running cold water. Antigens were retrieved in sodium citrate buffer (10mM Sodium Citrate, 0.05% Tween-20 pH 6.0) by microwaving for 15 minutes, followed by rinsing with cold water. The slides were then washed in PBS / 0.025% Triton-X-100 (PBST) 2 times, 5 minutes each, blocked in PBST / 1% BSA / 5% Donkey serum for 1 hour at room temperature, briefly rinsed with PBST and incubated with anti-mCherry antibody (Abcam) at 4 degrees Celsius overnight. The slides were washed with PBST 3 times, 5 minutes each, blocked for at least 15 minutes, incubated with secondary antibody for 1 hour at RT and finally washed with PBST 3 times, 5 minutes each. The sections were mounted in a small drop of DAPI (VectaShield), and covered with a coverslip and sealed with nail varnish.

2.6 In Vivo Differentiation

All mouse experiments were performed in accordance with the UK's 1986 Animals and Scientific Procedures Act and local institute ethics committee regulations.

2.6.1 Teratoma Formation

Mouse ESCs were harvested and resuspended in PBS containing 30% matrigel. 5×10^6 cells

in 100µl were injected subcutaneously into the dorsal flank of NSG mice (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ, The Jackson Laboratory) under general anaesthesia. Teratomas developed within 4 - 8 weeks. The mice were culled using Schedule 1 methods once the teratoma reached the legal limit (1.2mm^2) as per Home Office Guidelines. Teratomas were dissected and fixed in 10% buffered formalin phosphate for at least 24 hours before paraffin embedding, sectioning and H & E staining. All slides were evaluated by a histopathologist.

2.6.2 Chimera generation

2.6.2.1 Preparing the embryos and foster recipients

For the production of 8-cell stage embryos, over 8-week-old CD1 females were super ovulated by intraperitoneal injection of Pregnant Mare Serum Gonadotropin (PMSG , 7.5 IU each) (Intervet, PMSG-Intervet) followed 46-48 hours later by injection of human Chorionic Gonadotropin (hCG 7.5 IU each) (Intervet, Chorulon) and set up with proven studs. Vaginal plugs were checked the following morning (0.5 days post coitum, dpc), and 8-cell stage embryos were collected at 2.5 dpc and cultured in 1.0 ml KSOM (Millipore) medium containing 3.0 mg/ml BSA (Sigma) in incubator in humidified atmospheres of 5% CO_2 at 37 °C. CD1 females were mated with vasectomized C57BL6/JCBA F1 males and used as recipients for embryos transfer at 2.5 dpc.

2.6.2.2 Injection of ESCs and EPSCs

EPSCs and conventional ESCs were transfected with a PB-CAG-mCherry-PGK-Puro expression vector using Amaxa nucleofection (Lonza) according to the manufacturer's protocol. Following two rounds of cell sorting, a homogenous mCherry positive cell line was generated. Serial FACS analysis revealed that the cells continued to stably express mCherry. No significant silencing was observed. Injections were carried out with a Nikon Eclipse

TE2000-U inverted microscope unit. ESC/EPSC microinjection was performed at room temperature on the inverted lid of a 100-mm falcon culture dish in an injection drop of 1.0 ml M2 medium (Sigma). For *in vitro* observation, mCherry tagged EPSCs or conventional ESCs were microinjected to morula stage CD1 (ICR) embryos. The injected embryos were allowed to develop to blastocysts *in vitro* for 24-48 hours prior to assessment of contribution to the ICM and trophectoderm.

For *in vivo* chimera generation, the embryo was injected with 6-8 cells per 8-cell stage embryo for bulk injection or 1 cell per embryo for single cell injection. After 45-50 embryos had been injected, the injected embryos were washed three times in KSOM (Millipore) containing 3.0 mg/ml BSA. Embryos were then transferred into a well of a four-well plate containing 1.0 ml KSOM. After overnight culture, 20-25 injected embryos were transferred into both uterine horns of a foster recipient.

2.7 Assessment of Embryonic and Extra-Embryonic Contribution

2.7.1 Microscopy, embryo dissection, placenta dissociation and FACS analysis.

E14.5 embryos were dissected and imaged using Leica M205FA Automated Fluorescence Stereo Microscope (100 ms for phase and 5 seconds for mCherry). Placenta, yolk sac and embryo samples from mCherry positive embryos were taken for genotyping. Dissected placentas were cut into fragments of around 1 mm diameter on ice and digested with 2.35 ml placenta digestion solution (5% FBS, 10mM HEPES, 100 µl Enzyme D, 50 µl Enzyme R and 12.5 µl Enzyme A in HBSS according to the instruction of Lamina Propria Dissociation Kit) for 30 min at 37 °C with shaking. After digestion, the placentas were dissociated into single cells by pipetting vigorously and filtered through a 70 µm cell strainer (BD Bioscience). Single cells from placenta were treated with ACK Lysing Buffer (BioWhittaker) to remove the red blood cells and resuspended in FACS buffer (BD Pharmingen) for FACS analysis or

sorting. E14.5 fetal livers and fetal brains were dissociated in HBSS by pipetting vigorously and filtering through a 70 μm cell strainer. All FACS analysis was performed on BD LSRFortessa Cell Analyzer (BD Biosciences). Placenta cells were sorted on MoFlo™ XDP (Beckman Coulter).

2.7.2 DNA genotyping and placental gene expression analysis

Genomic DNA was extracted from placenta, yolk sac and embryo samples from mCherry positive embryos, quantified and diluted to 10 ng/ml. DR10-EPSCs have a Cre cassette at the Rosa26 locus. A pre-designed TaqMan® Copy Number Assay (Supplementary Table 2) was used to determine the contribution of DR10-EPSCs. AB2.2-EPSCs or ESCs have a non-functional Neo cassette at the Hprt locus. A custom-designed TaqMan® Copy Number Assay (Supplementary Table 2) was used to determine the contribution of these cells. Genomic DNA from parental DR10-EPSCs or AB2.2 ESCs were used as the positive controls, and DNA from a wild type embryo was used as the negative control. Cre/Neo copy number was first normalized to *Tert* using ΔCt method (TaqMan® Copy Number Reference Assay, Life Technologies). The contribution of DR10/AB2.2 in individual placenta, embryo proper, and yolk sac samples was then determined by normalizing them to parental DR10/AB2.2 cells.

In order to check the expression of trophoblast specific genes in sorted placenta cells, SYBR Green ROX qPCR Mastermix (QIagen) was used. Primer pairs are listed in Supplementary Table 6. All qPCR reactions were performed on ABI 7900 HT Sequence Detection System (Life Technologies). Gene expression was determined relative to *Gapdh* using ΔCt method.

2.8 Evaluation of Signalling Pathways

2.8.1 Mouse ESC / EPSC signalling dependence *in vitro*

On Day 0, 100 cells were plated per well of a 24-well gelatinized tissue culture plate in ES

media (M15, 2i/LIF or EPSCM) with four technical repeats per condition. After 24 hours, media was changed to test condition. Media was changed daily until day 10 when colonies were fixed and stained for Leucocyte Alkaline Phosphatase (Sigma). Inhibitors used were JAKi 0.6μM (Santa-Cruz, SC-204021), SB505124 1μM (Tocris, 3263), A83-01 0.5μM (Tocris 2939), and Su5402 2μM (Tocris, 3300).

2.8.2 Western Blot demonstration of Kinase Inhibition

1 x 10^6 ESCs were seeded into gelatinized wells of a 6-well plate in one of the media: M15, 2i/LIF, KSR2iLIF or EPSCM, respectively. After 48 hours, whole cell protein was extracted from harvested cells and quantified using BCA assay (Thermo Scientific). Total protein (40 μg) was fractioned on a 4-12% Bis-Tris Novex gel (Invitrogen) and electroblotted onto PVDF membranes and probed with the following antibodies: pERK, ERK, pSRC, SRC, pp38, p38, pJNK, JNK and AXIN1 (all from Cell Signalling). α-Tubulin (Abcam) was used as loading control. Blots were incubated with horseradish peroxidase-coupled anti-rabbit or anti-mouse IgG and developed with ECL Prime (Amersham).

2.8.3 JAK STAT signalling activity

For LIF stimulation, 1 x 10^6 cells were seeded in EPSCM for 24 hours. Cells were washed with PBS and cultured in DMEM/F12 and 20%KSR for 4 hours, then LIF was added at concentrations of 0.0, 0.5, 1.0, 5.0 and 10.0 ng/ml for 30 minutes. Whole cell protein was extracted for Western blot using pSTAT3 (Cell signalling) and STAT3 (BD Biosciences) antibodies. α-Tubulin (Abcam) was used as loading control.

To analyse the expression of down-stream target genes of Jak-Stat3 and MAPK signalling at transcription level, both E14Tg2a ESCs and EPSCs were deprived of LIF for 4 hours as above, then LIF 1.0 ng/ml was added for 1 hour. Total RNA was extracted and reverse

transcribed to make cDNA. qPCR was performed using Taqman Gene Expression Assays (ABI) (Supplementary Table 5).

2.8.4 Wnt signalling detection

To investigate Wnt signalling in EPSCs, 1 x 10^6 cells were seeded in M15, 2i/LIF, KSR2iLIF and EPSCM for 48 hours. Cytoplasmic and nuclear protein were extracted with a NE-PER nuclear and cytoplasmic extraction kit (Thermo Scientific) for Western blot to detect pβ-catenin (Cell Signalling), β-catenin (Sigma) and active- β-catenin (Millipore). α-Tubulin (Abcam) and H3 were used as loading control for cytoplasmic and nuclear protein respectively.

For TOPflash assay, 4 x 10^6 cells were co-transfected with TCF-LEF1 TOPflash and pRL-TK by Nucleofection (Amaxa). Cells were split 1:9 into a 24 well plate for 24 hours in M15, 2i/LIF and EPSCM for 48 hours, cell lysate was prepared for luciferase assay.

2.9 Bioinformatics

The methodology described herein was prepared in collaboration with Dr Cheuk-Ho J. Tsang (PhD thesis of Dr Cheuk-Ho J. Tsang, Sanger Institute).

2.9.1 Single-cell RNA-seq library preparation

The preparation is previously described by (Brennecke et al., 2013). In brief, five thousands trypsin-dissociated mouse EPSC (DR10) were loaded onto the 10-17-μm C1 Single-Cell Auto Prep IFCs (Fluidigm) and cell capture was performed according to the manufacturer's protocol. Individual capture site was inspected under light microscope to confirm the presence of single cell. The locations of empty capture sites and sites containing multiple cells or burst

cells were noted for downstream quality control during data analysis. The lysis and reverse transcription mixes were then prepared with the SMARTer PCR cDNA Synthesis kit (Clontech) and the Advantage 2 PCR kit (Clontech) according to Fluidigm recommendations. 1.0 µl of the ERCC Spike-In Control Mix (Ambion) in a 1:400 dilution was added to the lysis mix to allow control of technical variation of the library preparation protocol. Lysis and cDNA reverse transcription and PCR were performed and cDNA harvested by the C1 Single-Cell Auto Prep system according to manufacturer's settings. The success of cDNA preparation was confirmed by optimal DNA signal detected by a 2100 Bioanalyzer with high-sensitivity DNA chip (Agilent). Multiplex sequencing libraries were prepared using the Nextera XT DNA Sample Preparation Kit and the Nextera Index Kit (Illumina) according to the recommendation in the C1 Single-Cell Auto Prep manual. The libraries of individual cells were then pooled and sequenced on 4 lanes of HiSeq 2500 (Illumina) to generate 100bp paired-end reads through the Sanger Institute in-house sequencing pipeline.

2.9.2 Reads alignment and quality control of dataset

Pair-end reads from both bulk and single-cell RNA-seq datasets were aligned to the Mus musculus genome (GRCm38.74) by STAR (version STAR_2.3.0e_r291) with specific additional parameter settings: '--outFilterMultimapScoreRange 1 --outSAMstrandField intronMotif' (Dobin., 2013). The quantification of gene expression was performed by the htseq-count module from the HTSeq package (http://www-huber.embl.de/users/anders/HTSeq/) with gene annotation from GTF files (GRCm38.74) with parameter "–s no" in union mode (Anders et al., 2015). The count matrix of individual cells were normalized by size factors with DESeq2. As quality control in the single-cell dataset, cells with less than 500,000 counts in annotated features, expression of less than 4000 genes or high percentage (>10%) of counts mapping to mitochondria-encoded genes were removed from subsequent analysis.

2.9.3 Chromatin immunoprecipitation and analysis

The H3K4me3, H3K27me3 and input ChIP library of EPSCs (DR10) were prepared based on a modified ChIP protocol from (Lee et al., 2006). In short, about 20 million cells were cross-linked 1.1% formaldehyde for 10 mins at room temperature. Cross-linking was then quenched with 0.125 M glycine. Cell pellets were washed with PBS, snap frozen by liquid nitrogen and stored in -80 ℃ until further processing. The chromatins were sheared by Bioruptor Pico (Diagenode) for 5 mins in 30 sec: 30 sec on and off cycles. Immunoprecipitation were performed with 2.5 μg antibodies pre-washed and pre-attached to protein A Dynaebeads (Life Technologies) overnight at 4 ℃. Antibodies: H3K4me3 (Cell Signaling: CST 9751S), H3K27me3 (Millipore: 07-449). The beads were then washed and cross-linking was reversed with the elution buffer at 65 ℃ for 4 hours. Immunoprecipitated DNAs were purified with proteinase K digestion and the Qiagen miniElut kit (Qiagen). The multiplex sequencing libraries were prepared with the microplex library construction kit (Diagenode) following manufacturer's instruction. The DNAs were amplified for 11 cycles and quality of the library was checked on a bioanalyzer (Ailgent) using a high sensitivity DNA kit. The library was sequenced on 2 lanes of HiSeq2500 in multiplex.

The single-end 50 bp reads were then mapped to the mouse reference genome GRCm38 by the Burrows-Wheeler Aligner (Li and Durbin, 2009) (bwa-0.5.10) according to the in-house Sanger DNA pipeline. The H3K4me3 and H3K27me3 ChIP-seq data of TNGA mouse ES cells cultured in 2i/LIF and a TNGA gnomic DNA control were downloaded from the NCBI GEO SuperSeries GSE23943. The data were remapped to the GRCm38 mouse genome by bowtie2 (Langmead and Salzberg, 2012) with default parameters. Read depth between datasets were then normalized by read shuffling and random sampling to 20,664,503 reads (H3K27me3) and 10,311,175 reads (H3K4me3). H3K4me3 and H3K27me3 enriched peaks were detected by MACS (version 2.1.0) with the flag "--broad" turned on for detection of broad histone modification peaks. The signal tracks were generated by MACS (version 1.4.2)

(Zhang et al., 2008) and visualized by the UCSC genome browser. Peaks coordinates were then intersected with the gene promoter windows (+/-1kb from TSS) by Bedtools (Quinlan and Hall, 2010) to identify H3K4me3 and H3K27me3 associated genes. Bivalent genes were defined as genes that were both H3K4me3 and H3K27me3-associated. Comparison of H3K4me3 and H3K27me3 signals at the extended promoter window (TSS+/-3kb) and gene body were performed by ngs.plot (version 2.41.3) (Shen et al., 2014). To compare the H3K4me3 and H3K27me3 signals distribution at genes with different expression levels, expressed genes of EPSCs were first ranked based on their average length corrected mean normalized count level and genes in the 1st, 2nd, 3rd and 4th quartiles were classified into four groups as “Very High”, “High”, “Low” and “Very low”. Signal plot across gene body were then generated by ngs.plot. Length correction was performed by dividing the normalized count by union exon length (in kilobase).

2.9.4 Differential gene expression analysis

For the comparison of the transcriptomes of mouse EPSCs and 2i/LIF ESCs, differential expression analysis was performed with DESeq2. Cooks distance testing was turned off to accommodate the noisy expression nature in single-cell data and counts from the ERCC spikes in the EPSCM datasets were removed from the input matrix. Differential expression with adjusted p-value <0.05 and |log2 fold change| >1 were considered as significantly differentially expressed. Cell lines and passage numbers: DR25 and DR10, in serum, passage 5; in 2i/LIF, passage 5; in EPSCM, passage 15. DR9 and DR4, in 2i/LIF, passage 4 and in EPSCM, passage 5. AB2.2 ESCs, M15 passage 18, 2i/LIF passage 6, EPSCM passage 7. E14 ESCs, M15 passage 20, 2i/LIF passage 5, EPSCM passage 6.

2.9.5 Gene ontology enrichment analysis

Gene ontology term enrichment analysis was performed by DAVID (Huang da et al., 2009) (http://david.abcc.ncifcrf.gov/). Gene symbols were used as input gene list and the mouse genome was chosen as background.

2.9.6 Embryonic stage-enriched gene set enrichment analysis

Gene set enrichment analysis was performed with the javaGSEA application (version 2.1.0) available online (http://www.broadinstitute.org/gsea/downloads.jsp) with "weighted" enrichment statistics setting (Subramanian et al., 2005). The mouse pre-implantation embryonic stage-enriched gene sets were compiled by analysing the published expression count matrix from (Deng et al., 2014). The top 500 genes that showed higher average expression than other stages after size factor normalization by DESeq2 were considered stage-enriched and used as testing gene set. The normalized count matrices of the EPSC and standard 2i/LIF ESC single-cell dataset were used as input expression dataset. Enrichment is considered significant if FDR is below 0.1 and nominal P value is below 0.05.

2.9.7 Estimating the pre-implantation embryonic developmental time frame of EPSCs

The whole transcriptome or the normalized count values of the 36 epigenetic modifiers defined by Burton et al were used to estimate the pre-implantation embryonic developmental time point of EPSCs (Burton et al., 2013). The expression of the selected epigenetic modifiers has been previously shown to provide superior segregation of blastomeres of different pre-implantation embryonic stages to pluripotency-associated transcription factors by single-cell qPCR (Burton et al., 2013). The normalized count matrix of the whole transcriptome or the epigenetic modifiers from the dataset of Deng et al were transformed by log2(nCount+1) to stabilize the variance of the data and principal component analysis were performed by the

prcomp() function in R on the transformed matrix (Deng et al., 2014). The scores were then used to predict the scores of EPSCs or 2i/LIF ES cells by the predict () function in R.

2.9.8 Comparison of EPSCs, Hhex-Venus$^+$ ES cells, MERV-TdTomato$^+$ ES cells and *in vivo* iPS cells

The raw RNA-seq data of *in vivo* reprogrammed iPS cells (iviPS1-6) from Abad et al, Hhex-Venus$^+$ ES cells from Morgani et al (both 2i/LIF and serum/LIF) and MERV-Tdtomato$^+$ reporter ES cells from MacFarlane et al were downloaded from GSE48364, GSE45182 and GSE33923, respectively, and aligned to GRCm38 mouse genome by STAR (Abad et al., 2013; Macfarlan et al., 2012; Morgani et al., 2013). Gene expression was then quantified as described above by HTSeq-count. To normalize the difference between studies and sequencing platforms, the gene expression profile of EPSCs (DR9, DR25, DR4, DR10-EPSCM), MERV-TdTomato+ ES cells and *in vivo* iPS cells were quantified as the relative log2 fold change of expression of the genes compared to 2i/LIF ES cells (DR9, DR25, DR4, DR10-2i/LIF) by DESeq2. For Hhex-Venus$^+$ ES cells, the unsorted 2i/LIF datasets (GSM1098629 and GSM1098630) were used for normalization. Genes with unavailable log2 fold change value in DESeq2 were removed from comparison. Sequences of repeat elements (mouse and shared) were downloaded from (http://www.girinst.org/repbase/index.html). A custom GTF file and genome index was constructed by STAR with individual repeat class treated as individual chromosome. Reads were then realigned to the new index. Expression was quantified by HTSeq-count as described above and normalized by size factors computed by DESeq2. The relative expression of EPSCs and MERV-TdT$^+$ ES cells were then calculated relative to that of MERV-TdT$^-$ ES.

2.9.9 Mass Spectrometry

Genomic DNA was quantified using picogreen assay (Invitrogen), digested using DNA

Degradase plus (Zymo Research) for 90 minutes at 37 degrees and analysed by liquid chromatography-tandem mass spectrometry on a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) fitted with a nanoelectrospray ion-source (Proxeon, Odense, Denmark). Mass spectral data for C, 5mC and 5hmC were acquired as previously described (Ficz., 2013)). 5mC and 5hmC levels are normalized to total cytosine levels. Error bars represent standard deviation of at least 3 technical replicates per sample.

2.10 Human ES cell and Human EPSC Culture

The import and use of human ES cells were approved by The Steering Committee for the UK Stem Cell Bank and for the use of Stem Cell Lines

2.10.1 Human ES and human iPSC conversion to EPSCs

Two widely used human ESC lines H1 and H9 were used for the conversion experiments (Thomson et al., 1998). Individual human ESCs cells were seeded in EPSCM on MEF feeder cells at a low cell density. Approximately 0.3% of H1 cells formed undifferentiated colonies which could be subcloned and expanded in EPSCM to establish stable cell lines. In contrast, none of the colonies of H9 cells that initially survived in EPSCM could be further expanded. The converted H1 cells (H1-EPSCs) were dissociated to single cells for passaging with a 15-20% single cell subcloning efficiency in the presence of a selective ROCK inhibitor (10 μM Y27632) and could be expanded in EPSCM for over 30 passages on SNL feeders. Two standard human iPSC lines from the CGAP facility (BOB-iPSCs and FSPS186-iPSCs) were obtained and the conversion strategy was repeated. BOB survived the conversion process with EPSC colonies emerging from single iPSCs which were plated at low density on SNL feeders. FSPS186 could not be converted.

2.10.2 Picking, expansion and cryopreservation of human EPSCs

Using a P20 pipette (set to 5 µl), a single primary colony was detached from the feeder layer by careful scraping and subsequently aspirated. Each colony was placed in a 96-well round bottomed plate (1 colony / well) that contained 20µl PBS. 30 µl Accutase was added per well and the 50 µl total volume was mixed by pipetting up and down 10 times. The 96-well plate was transferred to the incubator (37 degrees Celsius, 5% CO2) for 5-10 minutes. Using a multichannel pipette set to 50µl, the colony was dissociated to single cells. Each well was checked under microscope to confirm a single cell suspension. 50µl EPSCM media containing 10 µM Y27632. was added to each well and the total 100µl volume (20µl PBS + 30µl Accutase + 50µl EPSCM) was transferred to a pre-prepared 24 well SNL feeder plate with pre-equilibrated EPSCM media plus 10 µM Y27632. One well of a 96-well plate was transferred to one well of a 24-well plate. The plate was placed in the incubator at 37 degrees Celsius. The ROCK inhibitor was removed after 24 hours by changing the media with fresh EPSCM. The media was subsequently changed daily. Colonies re-emerged after 3-5 days however as they were growing from single cells recovery did occasionally take 7-10 days.

For routine bulk culture passaging of human EPSCs, enzymatic dissociation with Accutase was used. Following aspiration of media and washing once with PBS, Accutase was added to cover the base of the well. After five minutes of enzymatic dissociation in an incubator at 37 degrees Celsius, 5% CO2, a pipette was used to assist in cell detachment and create a single cell suspension which was passed through a 100 µm filter into a 15ml falcon tube. The filtered sample was centrifuged at 1200rpm for 3 minutes with subsequent aspiration of the supernatant and resuspension and plating in EPSCM plus 10 µM Y27632. Cell counting was routinely performed using trypan blue exclusion. EPSCs were plated at a density of 1.5 – 5 x 10^4 cells / cm^2.

For cryopreservation, human EPSCs were resuspended in 500 µl Freezing Buffer (80% EPSCM, 10% FBS (HyClone), 10% DMSO and 10 µM Y27632). Cryovials were placed at -80 degrees Celsius overnight then transferred to liquid nitrogen the following day for long term storage.

For cryorecovery, the cryovial was first thawed in a water bath at 37 degrees Celsius. After thawing, the cryovial was centrifuged briefly at 1200 rpm. The contents of the cryovial were then transferred to a 50ml falcon tube which contained 15 ml of pre-equilibrated EPSCM containing 10 µM Y27632. The falcon tube was centrifuged for 3 minutes at 1200rpm. The supernatant was aspirated and the cell pellet was re-suspended in pre-equilibrated ROCK inhibitor containing EPSCM and plated on SNL feeder plates. The media was changed after 12 hours to ensure removal of any residual DMSO.

2.10.3 Differential Adhesion Feeder Removal Strategy

For downstream experiments including differentiation, gene targeting, RNA sequencing etc. feeder removal was often required. Differential adhesion based feeder removal was routinely used as described herein. Firstly, the media was aspirated from the tissue culture plate and washed once with PBS. Following aspiration of the PBS, Accustase was added in a volume sufficient to cover the base of the well. After incubation for 5 minutes at 37 degrees Celsius a pipette was used to assist cell detachment. EPSCM plus 10 µM Y27632 was added to the plate and the total cell suspension was collected and passed through a 100 µm filter into a 15 ml falcon tube which was centrifuged at 1200rpm for 3 minutes. The supernatant was aspirated and the cell pellet was resuspended in EPSCM plus 10 µM Y27632. The cells were transferred to a non - gelatinized 10cm plate containing EPSCM plus 10 µM Y27632 and placed in an incubator for 30 minutes. The supernatant was then collected from the plate and

the cells were counted using trypan blue exclusion. The volume containing the required number of cells was centrifuged at 1200rpm for 3 minutes.

2.11 Culture of human fetal neural stem cells (hfNSCs)

BRC1019, a human fetal neural stem cell line (a gift from Dr. Colin Watts) was routinely expanded in T75 tissue culture flasks in NDiff 227 media (StemCells Inc) supplemented with 10 ng/ml Fibroblast Growth Factor-Basic (Invitrogen) and 10 ng/ml Epidermal Growth Factors (Invitrogen). Prior to cell plating, the tissue culture flasks were coated with ECM gel from Engelbreth-Holm-Swarm murine sarcoma (Sigma). To prepare ECM gel working solution, 1 ml ECM was mixed with 9 ml KnockOut-DMEM (Invitrogen) and the10-ml aliquots were stored at -20°C. Before use, a 10-ml aliquot was thawed at 4°C overnight and used to coat tissue culture flasks. During expansion of BRC1019 cell line, the media was changed every other day. Cells were passaged enzymatically using Accutase.

2.12 Reprogramming and generation of human iPSCs

2.12.1 Reprogramming human fetal neural stem cells to iPSCs

Before electroporation, BRC1019 cells were plated at a density of 0.5–5 x 10^4 cells/cm^2 in a T75 tissue culture flask. Once the cells reached 70% - 80% confluence, they were dissociated using Accutase and harvested for electroporation. Human fetal neural stem cell transfection was performed using an Amaxa machine (Lonza) according to the manufacturer's protocol (murine neural stem cell, program A-033) with a transfection efficiency of 75%. For episomal (integration-free) 6F reprogramming, 1 x 10^6 hfNSCs were usually transfected. After electroporation, hfNSCs were seeded in M15 media supplemented with hLIF (1 x 10^6 U/ml, Merck Millipore) and Vitamin C (50 μg/ml, Sigma) on mitomycin C-inactivated SNL feeder plates. The media was changed every other day. When colonies emerged 10-14 days after

electroporation, the media was switched to EPSCM for a further 5-7 days before picking. Electroporation of 1 x 10^6 hfNSCs cells typically generated 10-20 primary iPSC colonies with a line establishment rate of 50%.

2.12.2 Reprogramming human dermal fibroblasts to iPSCs

Before electroporation, 5 x 10^5 HDFs were plated in a T75 tissue culture flask. When HDFs were 70-80% confluent, they were typsinized and collected for electroporation. Transfection of HDFs was performed using an Amaxa machine according to the manufacturer's protocol (human dermal fibroblast cell, program U-020). After electroporation, HDFs cells were seeded in M15 plus hLIF for 24 hours. From the second day, transfected cells were cultured in M15 media supplemented with hLIF (1 x 10^6 U/ml, Merck Millipore), Vitamin C (50 μg/ml, Sigma) and doxycycline (1.0 μg/ml, Clontech). The media was changed every other day. Human iPSC colonies emerged 16-20 days after electroporation, the media was changed to EPSCM for a further 5-7 days before picking. Electroporation of 1 x 10^6 HDFs typically produced 1,000-2,000 primary iPSC colonies with a 50% stable line establishment rate.

2.13 Gene targeting in human EPSCs

The methodology described herein was prepared in collaboration with Dr Wei Wang PhD, Sanger Institute. Targeting vectors used to generate the ROSA26-SA-H2B-Venus-PGK-Puro, OCT4-2A-H2B-Venus-EF1α-Puro and OCT4-2A-Puro-2A-H2B-Venus and NANOG-2A-Bsd-2A-H2B mCherry reporter EPSC lines were all made using E.coli recombineering. SA-H2B Venus-PGK-Puro, 2A-H2B-Venus-EF1α-Puro and 2A-H2B-mCherry-EF1α-Puro constructs were gifts from Dr. Manousos Koutsourakis and Dr. Bill Skarnes Sanger Institute. To generate 2A-Puro-2A-H2B-Venus and 2A-Bsd-2A-H2B-mCherry constructs, EF1α-Puro selection marker was deleted from 2A-H2B-Venus-EF1α-Puro and 2A-H2B-mCherry-EF1α-

Puro constructs and the 2A-Puro or the 2A-Bsd cassette was inserted in front of 2A-H2B-Venus or 2A-H2B-mCherry, respectively.

For targeting, EPSCs were washed with PBS and dissociated using Accutase. After dissociation, cells were collected, counted, and resuspended in EmbryoMax ES Cell Electroporation Buffer (Merck Millipore). For one electroporation, 5 μg Cas9 expression vector (George Church lab, Addgene), 5 μg gRNA expression vector (George Church lab, Addgene) and 10 μg of targeting vector were mixed with 1 x 10^7 cells, and electroporated with Biorad Gene Pulser using a condition of 320 v, 250 μF. After electroporation, the cells were plated onto SNL feeder plates in EPSCM supplemented 10 μM Y27632 for 24 hours. Drug selection (Puromycin 1.0 μg/ml, Blasticidin 5.0 μg/ml) was performed 48 hours after electroporation. When drug resistant colonies emerged, the medium was switched to EPSCM for an additional 2 days before picking.

2.14 Human EPSC Differentiation

2.14.1 Human embryoid body (EB) differentiation

EPSCs were routinely expanded on 10cm SNL feeder plates. Following enzymatic passaging and feeder removal as previously described, EPSCs were re-suspended in DMEM/F12 / 20% KSR and filtered with a 0.4μM cell strainer to make a single cell suspension. 9 x 10^5 cells were plated into 1 well of an AggreWellTM400 (Stem Cell Technologies) to form embryoid bodies (EB). After 24 hours, the EBs were harvested and re-suspended in DMEM/F12 / 20%KSR. Approximately 1000 EBs were plated into one 6 well of a low attachment plate and grown for 5 days. At day 6, the EBs were collected and plated into a gelatinized 12-well plate in N2B27 for neural differentiation and M10 for mesoderm and endoderm lineage differentiation for 5 more days. At day 11, cells were fixed in 4% PFA and immunostained

with antibodies against neuron-specific class III beta-tubulin (Tuj1), alpha smooth muscle actin (SMA) and alpha fetoprotein (AFP).

2.14.2 Targeted Neural Differentiation of EPSCs

Neural Induction Media contained NDIFF 227 (Stem Cell Sciences, SCS-SF-NB-02), 10ng/ml FGF-BASIC Life Technologies Ltd. (Invitrogen), PHG0021 and 1 μM Dorsomorphin (Sigma, P5499-5MG). Neural Stem Cell Culture media (NSC-CM) contained, NDIFF 227 (Stem Cell Sciences, SCS-SF-NB-02), 10ng/ml FGF-BASIC Life Technologies Ltd. (Invitrogen), PHG0021 and 10ng/ml Epidermal Growth Factors (EGF) Life Technologies Ltd. (Invitrogen), PHG0311.

EPSCs were differentiated on 6 well plate(s) at a cell density of 2.5 – 5 x 10^4 cells / cm^2. Prior to cell plating the well was coated with ECM Gel from Engelbreth-Holm-Swarm murine sarcoma (Sigma, E1270). Briefly this was prepared by thawing the ECM and adding 1ml ECM to 9ml KO-DMEM. This was mixed and stored in 10ml aliquot(s) at -20 degrees. For use, a 10ml tube was thawed at 4 degrees overnight. Once thawed, ECM will form a gel at room temperature. ECM was added to coat the base of the well and then aspirated and discarded. Cells were plated immediately in Neural Induction media plus 10 μM Y27632. The ROCK inhibitor was removed the following day. (**Note:** ROCK inhibitor is only required during the first 24 hours to enhance the single cell survival and plating of EPSCs). During differentiation, the media was changed after the first 24 hours, then, every 48 hours thereafter.

Neural rosettes appeared after 2.5 - 3 weeks of differentiation. At the rosette stage, to establish a neural stem cell line, the culture was passaged enzymatically using Accutase and split 1:1 or 1:2 into NSC-CM onto ECM coated plates. During NSC expansion, the media was changed every 48 hours.

For terminal neuronal differentiation, 35mm tissue culture plates were first coated with laminin at a working concentration of 1-2 μg/cm^2 (Sigma, L6274), 2 hours before cell plating. NSCs were subsequently plated at a density of 5 x 10^4 cells / cm^2 in NSC-CM. The following day, the media was removed, cells were washed once with PBS and NDIFF 227 medium (Stem Cell Sciences, SCS-SF-NB-02) was added. Media change was performed every other day. With prolonged differentiation (~3 months for electrophysiological maturity), the culture was often observed to partially detach from plate. When this was observed a few microliters of laminin was added to the media which resulted in cellular re-attachment to the base matrix. Neuronal electrophysiology was performed in collaboration with Peter Kirwan, Livesey Lab at the Gurdon Institute.

2.14.3 Pancreatic Differentiation

Pancreatic differentiation was performed in collaboration with Mariya Chhatriwala, Vallier Lab at the Sanger Institute using their published monolayer differentiation protocol (Cho et al., 2012).

2.14.4 Differentiation of human EPSCs to trophoblasts

For trophoblast differentiation, H1-ESCs, H1-EPSCs (passage 15) or EC5-EPSCs (passage 20) were plated in 20% KSR media at a density of 5 x 10^5 cells/well in a gelatinized 6-well tissue culture plate. After 24 hours, the media was changed to 20% KSR with 10.0 ng/mL human recombinant BMP4 (R&D Systems) and 10.0 μM SB431542 (Tocris) (Amita et al., 2013; Bernardo et al., 2011). Media was changed daily. Differentiated cells were harvested and used for FACS analysis or qRT-PCR analysis at different time points. All experiments were repeated at least twice on different passages of cells to ensure that the patterns of gene

expression described were reproducible.

2.15 Human EPSCs cell cycle analysis

Click-iT® Plus EdU Alexa Fluor® 647 Flow Cytometry Assay Kit (Life Technologies, C10634) was used for cell-cycle analysis according to manufacturer's instructions. EdU incorporation (Alexa Fluor 647 labelled anti-EdU antibodies) was measured with DNA content (DAPI) in fixed and permeabilized cells.

2.16 Spectral karyotyping

Human and mouse karyotyping including DAPI banding and Multicolour FISH was performed at the molecular cytogenetics core facility at the Sanger Institute (Thanks to Beiyuan Fu and Fengtang Yang for technical assistance). For mouse karyotyping, a 10 cm plate of feeder free EPSCs at 60 – 80% confluence was submitted. For human karyotyping, once the cells reached 60-80% confluence the media was changed with the addition of fresh equilibrated EPSCM containing 10 μM Y27632, 1 hour before sample submission.

2.17 Array CGH

Array comparative genome hybridization of human ESC, human EPSCs and human iPSCs was performed at the Sanger Institute core facility.

2.18 Single Cell RNA FISH

Human EPSCs were passaged as per previously stated method using enzymatic dissociation. After centrifuging at 1200rpm for 3 mins, 1 x 10^5 cells were resuspended in 1ml of PBS plus 10 μM Y27632. For each cytospin, 200μl (2 x 10^4 cells) was used with an optimal

centrifugation speed range of 500-800 rpm for 5 mins. The cell containing slides were fixed and submitted to the Molecular Cytogenetics core facility at the Sanger Institute for further analysis (Thanks to Beiyuan Fu and Fengtang Yang for technical assistance).

2.19 Digital microscopy imaging

Immunofluorescence stained samples were examined with a Leica DM5000B microscope equipped with narrow bandpass filters for Cy3.5, fluorescein and DAPI fluorescence. Images were captured via a monochrome digital camera (ORCA-03G, Hamamatsu) and processed with the SmartCapture software (Digital Scientific UK).

2.20 Human DMR analysis

The methodology described herein was prepared in collaboration with Dr Wei Wang PhD, Sanger Institute. Bisulfite treatment was performed using the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's recommendations. Genomic DNA PCR for promoter regions was performed using primer pairs described in Supplementary Table S3. PCR products were cloned using pGEM-T Easy Vector Systems and sequenced from both ends.

2.21 Human Bioinformatics

The methodology described herein was prepared in collaboration with Dr Cheuk-Ho J. Tsang, Sanger Institute.

2.21.1 Processing of human data

Standard TruSeq mRNA-seq library were generated by in-house Sanger DNA pipeline. At least four million cells were submitted for RNA extraction and multiplex library preparation. The library was sequenced on HiSeq2500 to generate 100bp paired-end reads and were

aligned to the GRCh37 genome using STAR aligner. Gene expression was quantified by Cufflinks2.

2.21.2 Culture condition signature enrichment in different stages of human embryos

The signature genes were defined as the top 1000 mostly expressed differential genes between PSCs (EC2, EC5, H1) cultured in EPSCM and H1 ESC cultured in FGF conditions. Differential gene analysis was performed using Cufflinks 2. The enrichment of the signature genes of the two culture conditions were tested in individual pre-implantation stage against passage 10 human ESC cultured in FGF conditions. The human embryo single-cell RNA-seq data were retrieved from (Yan et al., 2013). Positive end of the x-axis indicates the level of enrichment in passage 10 FGF human ESCs in each pairwise comparison, the negative end indicate enrichment towards the respective embryonic stage. For instance, human 8C embryos showed significant enrichment of EPSCM signature (score: -1.51), compared to passage 10 FGF human ESCs, while the FGF signature is enriched towards p10 FGF ESCs as expected (score 2.44). *$p<0.05$, **$p<0.01$.

2.21.3 Human single-cell RNA sequencing and X-chromosome reactivation analysis

Single-cell RNA-seq data of 96 human EPSCs (EC5-Rosa26-Venus) were generated using Fluidigm C1 AutoPrep system according to manufacturer instruction using medium sized integrated fluidic circuits (10-14um). 100bp paired-ends reads were generated on HiSeq2500 according to in-house Sanger DNA pipeline. Reads were aligned to GRCh37 human genome using STAR aligner with default parameters. In total, 44 cells passed quality control (including removal of lysed cells, doublets, unique aligned reads <500,000, mitochondrial reads >10%, gene detection <8000). Samtools mpileup and Vcftools were used to identify

variant sites at the X-chromosome. Heterozygous SNP sites were then annotated by intersecting their coordinates with that of annotated X-chromosome genes. A SNP is only considered valid if it can be detected in more than 4 cells. The number of probable SNPs in each X-chromosome-linked gene was then summarized and selected for mRNA-FISH validation.

2.22 Statistics analysis

Experimental data are shown as mean and standard deviation. All statistical analyses were either conducted with Prism (GraphPad) or specified in relevant sections.

Supplementary Methods Tables

Table 1. Chemical Screen Compounds

Table 2. QPCR Probes for DNA Copy Number

Table 3. DMR Analysis Primers

Table 4. Antibodies

Table 5. Mouse QPCR Probes for RNA expression

Table 6. Mouse Primers for Trophoblast Gene Expression

Table 7. Human QPCR Probes for RNA Expression

Table 1 Chemical Screen Compounds

Compound	Target	Final_Conc (μM)
GSK1904529A	IGF-1R/IR	0.2
Sunitinib	PDGFRA, KIT	1
VX-680 (MK-0457)	AURK A/B/C, ABL, FLT3	0.2
Imatinib	Abl, KIT, PDGFR	0.2
TAE684	ALK	0.2
PF2341066	MET, ALK	0.2
Dasatanib	Src, Bcr-Abl	0.05
XAV 939	TNKS1/2 (Wnt/b-catenin)	1
Gefitinib	EGFR	0.1
ABT-263	Bcl-2, Bcl-xl	0.5
SAHA (Vorinostat)	HDAC I/II	0.5
AZD2281 (Olaparib)	PARP1/2	1
Lenalidomide	TNF alpha	1
Axitinib	PDGFR, KIT, VEGFR	0.5
AZD 7762	Chk 1/2	0.2
GW 441756	Trk A	0.5
CEP-701 (Lestaurtinib)	JAK2, FLT3	0.1
CHIR-99021	GSK-3	1
17-AAG	Hsp90	0.05
VX-702	p38	1
KU-55933	ATM	0.2
Elesclomol	Hsp70 inducer	0.005
BIBW 2992	EGFR, ERBB2	0.1
GDC-0449	SMO (Hedgehog)	2
BX 795	PDK1, TBK1, IKKe	1
SL 0101-1	Rsk, AURKB, PIM3	2
BIRB 0796	p38, JNK2	2
JNK inhibitor VIII	JNK 1/2/3	2
PD173074	FGFR	0.2
ZM447439	AURK B	0.5
MK-2206	AKT	1
PD-0332991	CDK 4/6	0.5
GW843682X (AN-13)	Plk1	0.05
NVP-BEZ235	PI3K/mTOR	0.01
GDC0941	PI3K (a,b,d)	1
AZD8055	mTORC1/2	0.5
SB590885	BRAF	0.5
AZD6244	MEK 1/2	0.5
WO2009093972	PI3K beta	0.2

AZD6244 + MK-2206	Comb [MEK1/2 + AKT]	0.5/1
GSK 269962A	ROCK1/2	0.2
BMS-708163	Gamma-secretase	0.2
Obatoclax mesylate	Mcl-1	0.05
Nutlin-3	MDM2	2
Embelin	XIAP	2
Rapamycin	mTOR	0.002
BMS-536924	IGF-1R	0.5
SB-505124	TGFβR-I (ALK5)	1
CGP77675	Src	1
GM6001	MMP inhibitor	150
Go6983	PKC inhibitor	5
Cyclosporin A	Calcineurin inhibitor	15
PP2	Src	10
Compound C	AMPK,KDR/VEGFR2, ALK2/BMPR-I	1
PD0325901	ERK	1
PP1	Src	10

Table 2 QPCR Probes for DNA Copy Number

Probe Name	Applied Biosystems Catalog Number	Primer/Probe Name	Primer/Probe Sequence
Tert Copy Number Reference Assay	4458368	N/A	N/A
Cre Copy Number Assay	MR00635245_CN	N/A	N/A
Neo Copy Number Assay	custom-designed	Neo-QPCR-F	CGACCACCAAGCGAAACAT
	custom-designed	Neo-QPCR-R	CGACAAGACCGGCTTCCAT
	custom-designed	Neo-QPCR-Probe	CATCGAGCGAGCACG

Table 3 DMR Analysis Primers

DMR Region	**Primer Name**	**Primer Sequence**	**Use**
Human ELF5 DMR	ELF5-DMR-F1	AAAGTATTGGGTTTAGTTAGTTTGGGAG	Elf5 DMR 1st round PCR
Human ELF5 DMR	ELF5-DMR-R1	ATAACTAAATCCAAAAAAAAAACTTTCAAAAC	Elf5 DMR 1st round PCR
Human ELF5 DMR	ELF5-DMR-F2	TTTAGTTAGTTTGGGAGAGAGGTAG	Elf5 DMR 2nd round nested PCR
Human ELF5 DMR	ELF5-DMR-R2	CCAAAAAAAAAACTTTCAAAACAAAAAAAAAC	Elf5 DMR 2nd round nested PCR

Table 4 Antibodies

Category	Antibody	Company	Catalog Number	Dilution
Primary Antibody	p44/p42 MAPK (ERK)	Cell Signaling	9102	1:1000
	p-p44/p42 MAPK (Thr202/Tyr204)	Cell Signaling	9101	1:1000
	p38	Cell Signaling	9212	1:1000
	p-p38 (Thr180/Tyr182)	Cell Signaling	9215	1:1000
	SAPK/JNK	Cell Signaling	SC-571	1:1000
	p-SAPK/JNK (Thr183/Tyr185)	Cell Signaling	9251	1:1000
	Src	Cell Signaling	2110	1:1000
	p-Src (Tyr416)	Cell Signaling	2101	1:1000
	Axin1	Cell Signaling	2087	1:1000
	Stat3	BD Biosciences	610189	1:1000
	p-Stat3 (Tyr705)	Cell Signaling	9131	1:1000
	B-catenin	Sigma	C7082	1:1000
	p-B-catenin (Ser33/37/Thr41)	Cell Signaling	9561	1:1000
	active B-catenin	Millipore	05-665	1:1000
	Histone H3	Cell Signaling	4499	1:1000
	Alpha-Tubulin	Abcam	Ab7291	1:3000
	Cdx2	Biogenex	MU392A-UC	1:150
	Oct4	Abcam	Ab1985	1:250
	Oct (C10)	Santa-Cruz	SC-5279	1:100
	Cleaved Caspase 3	Cell Signaling	9579	1:100
	Ki-67	Cell Signaling	9129	1:100
	B III Tubulin (Tuj1)	R & D Systems	MAB1159	1:150
	Alpha-Smooth Muscle Actin	R & D Systems	MAB1420	1:150
	Alpha-Fetaprotein	R & D Systems	MAB1368	1:150
	Gata4 (C-20)	Santa-Cruz	SC-1237	1:100
	Gata6	R & D Systems	AF1700	1:6000
	Placenta Lactogen I (P-17)	Santa-Cruz	SC-34713	1:100
	Integrin alpha 5 (H-104)	Santa-Cruz	SC-10729	1:50
	Cytokeratin 7 (5F282)	Santa-Cruz	SC-70936	1:50
	HLA-G	Abcam	Ab52454	1:150
	HCGBeta	Abcam	Ab80786	1:150
	mCherry	Abcam	Ab125096	1:100
	GFP	Abcam	Ab13970	1:200
	NeuN C-A60	Millipore	MAB377	1:100
	Synaptophysin	Abcam	Ab32594	1:100
Secondary Antibody	ECL Mouse IgG, HRP-Linked Whole Ab (for WB)	GE Healthcare	N931V	1:10,000
	ECL Rabbit IgG, HRP-Linked Whole Ab (for WB)	GE Healthcare	N934V	1:10,000
	Alexa Fluor 488 and Alexa Fluor 594 (for IF)	Invitrogen		1:500 -1000

Table 5 Mouse QPCR Probes for RNA expression

Gene Name	Applied Biosystems Catalog Number
Cdx2	Mm01212280_m1
Dppa3	Mm00836373_g1
Egr1	Mm00656714_m1
Elf5	Mm00468732_m1
Eomes	Mm01351985_m1
Esrrb	Mm00442411_m1
Fgf5	Mm00438615_m1
Gapdh	4352339E
Gata6	Mm00802636_m1
Klf4	Mm00516104_m1
Nanog	Mm02384862_g1
Pou4f1	Mm00658129_gH
Rex1	Mm03053975_g
Socs3	Mm00545913_s1
Sox1	Mm00486299_s1
Sox2	Mm03053810_s1
T	Mm01318252_m1

Table 6 Mouse Primers for Trophoblast Gene Expression

Gene Name	Primer Name	Primer Sequence
Ascl2	Ascl2-F	CAGCTGCGAGGGAGAGCTAA
	Ascl2-R	GATGCTCAGTAGCCCCCTAACC
Bmp4	Bmp4-F	GAGTTTCCATCACGAAGAACA
	Bmp4-R	GCTCACATCGAAAGTTTCCC
Cdx2	Cdx2-F	TGGAGCTGGAGAAGGAGTTT
	Cdx2-R	CTGCGGTTCTGAAACCAAAT
Elf5	Elf5-F	AATGAAGACTACTACCCTGCCT
	Elf5-R	CAGCAGAATTGGAGCCATTCC
Eomes	Eomes-F	CAATGTTTTCGTGGAAGTGG
	Eomes-R	GTTAGGAGATTCTGGGTGAA
Fgfr2	Fgfr2-F	GACAAGCCCACCAACTGCACC
	Fgfr2-R	CGTCCCCTGAAGAACAAGAGC
Furin	Furin-F	TGCCAGACCACATGACTAC
	Furin-R	CAAGGACTTGGGGGATGAA
Gcm1	Gcm1-F	CCCCAGCAAGTTCCATCAGA
	Gcm1-R	AAGGCTCACCTCCCGGATT
Hand1	Hand1-F	ATGAACCTCGTGGGCAGGTA
	Hand1-R	TCACTGGTTTAGCTCCAGCG

Otx2	Otx2-F	AAACAGCGAAGGGAGAGGA
	Otx2-R	AGGAGGAGGAAGTGGACAA
Pl1	Pl1-F	TTGGCCGCAGATGTGTATAG
	Pl1-R	TCGTGGACTTCCTCTCGATT
Spc4	Spc4-F	GGCATCAGACCCAACTACA
	Spc4-R	ATCCACCAAGCCAAATCCA

Table 7 Human QPCR Probes for RNA Expression

Gene Name	Applied Biosystems Catalog Number
CDX2	Hs01078080_m1
CGα	Hs00985275_g1
CGβ	Hs00361224_gH
ELF5	Hs01063021_m1
FGF4	Hs00173564_m1
GAPDH	4326317E
GATA6	Hs00232018_m1
GSC	Hs00418279_m1
KLF4	Hs00358836_m1
KRT7	Hs00559840_m1
NANOG	Hs02387400_g1
OCT4	Hs00999634_gH
PAX6	Hs01088114_m1
PGF	Hs00182176_m1
REX1	Hs00399279_m1
SOX1	Hs01057642_s1
SOX17	Hs00751752_s1
SOX2	Hs00602736_s1
STELLA	Hs01931905_g1
T	Hs00610078_m1

CHAPTER 3

3 NOVEL CHEMICALLY DEFINED MEDIA PERMITS ESTABLISHMENT *IN VITRO* OF A UNIQUE MOUSE EMBRYONIC STEM CELL STATE

3.1 Summary

In this chapter, I describe the formulation of a novel chemically defined medium which permitted the maintenance *in vitro* of a unique mouse stem cell state. In contrast to conventional embryonic stem cells, following micro-injection to morula and early blastocyst stage embryos, these cells demonstrated expanded fate potential with donor cell contribution identified in both the embryonic and extra-embryonic compartments including the trophectoderm and primitive endoderm. This chapter begins with a description of the serendipitous sequence of events that resulted in the creation of this medium and the preliminary observations that resulted in the generation of the hypothesis and experimental aims for this thesis. In the Results section, I characterize this *in vitro* state and evaluate the individual components of the medium. Based on published genetic and biochemical data, I speculate on the role of each of the medium constituents in maintaining this culture condition.

3.2 Introduction

3.2.1 Science and serendipity

In keeping with the empirical importance of serendipity in science, initially, this project did not set out to investigate signalling pathways in the early mouse embryo or the presence of self-renewing stem cell states in early mammalian development. Rather, the focus was on further exploring the human naïve state.

In 2006 and 2007, two revolutionary papers were published which challenged our understanding of differentiation induced fate restriction. The authors demonstrated that mouse and human induced pluripotent stem cells (iPSCs) could be generated by 'reprogramming' fully differentiated somatic cells through overexpressing the transcription factors *Oct4, c-Myc, Sox2* and *Klf4* (OCKS) (Takahashi et al., 2007; Takahashi and Yamanaka, 2006). These findings resulted in an entirely new field of scientific research where considerable effort was invested in elucidating the reprogramming process.

With this end in mind, our group performed a genetic screen with the aim of identifying additional or replacement transcription factors which could be used in the reprogramming of mouse embryonic fibroblasts (Wang et al., 2011b). The authors reported that enhancing retinoic acid (RA) signalling by expressing Retinoic acid receptors (RARs) or by adding RA agonists promoted reprogramming and inhibiting RA signalling using a RAR-α dominant-negative form completely blocked the reprogramming process. In addition, co-expressing retinoic acid receptor gamma (RAR-γ) and liver receptor homologue 1 (*Lrh-1* / *Nr5a2*) with the original four Yamanaka factors (OCKS) markedly accelerated reprogramming. Using this six-factor platform, reprogramming of mouse embryonic fibroblast cells to iPSCs required only four days of induction, the overall reprogramming efficiency was increased and >90%

of the primary colonies were ground-state iPSCs (Wang et al., 2011b; Ying et al., 2008).

This six-factor combination also readily reprogrammed primary human neonatal and adult fibroblast cells to iPSCs which resembled ground-state mouse ES cells in growth properties, gene expression and signalling dependency. However, it was consistently observed that when the human cell culture deteriorated and became differentiated, it was always towards a GATA4 positive population despite the presence of 20% Knockout Serum Replacement, a Mek1 inhibitor (PD0325901), a GSK3 inhibitor (CHIR99021), comprising the original 2i inhibitors and Leukemia inhibitory factor (LIF) referred to as KSR2iLIF (Wang et al., 2011b; Ying et al., 2008) **(Figure 3.1A)**. Furthermore, subsequent unpublished experiments demonstrated that new human embryonic stem cells (hESCs) could not be established from pre-implantation human embryos using this medium. Therefore, I sought to stabilize the six-factor reprogrammed human iPSCs by refining the KSR2iLIF medium.

In the mouse embryo, inhibition of MEK (MAPK-ERK) signalling prevents formation of the putative hypoblast and expands the pluripotent epiblast compartment (Nichols et al., 2009). In humans, in contrast, it has been show that segregation of the hypoblast is not dependent on FGF signalling (Roode et al., 2012) suggesting that the Mek1 inhibitor, PD0325901 by itself would be insufficient to prevent differentiation. We speculated that if our 6 factor human iPSCs were similar to the human inner cell mass, perhaps, this differentiation bias was towards a lineage bearing resemblance to the human hypoblast. As the molecular mechanism of hypoblast segregation in the human embryo is unknown, a chemical screen approach was chosen. *Rex1-GFP* mouse embryonic stem cells (Guo et al., 2011) were chosen for the pilot screen as the experimental hypothesis was that a naïve media should allow maintenance of both the mouse and human ground-state.

The screen was designed so as to identify potential substitutes for PD0325901 (Mek1 inhibitor). A chemical library from the CGP at the Sanger Institute (kindly provided by Dr Ultan McDermott) was used for the screen. In total, 56 individual inhibitors were tested in their ability to maintain *Rex1-GFP* expression. Experimental strategy is shown in **Figure 3.1B**. It was found that multiple single and combined inhibitor sets could maintain *Rex1-GFP* expression in mouse ESCs with many known signalling pathways identified including inhibitors of MEK1/2, FGFR and TGFBR-1 (ALK5). The full list of chemicals tested in provided in the material methods section (Supplementary Table 1).

A consistently identified pathway which could maintain *Rex1-GFP* expression in the presence or absence of GSK3 beta inhibition was inhibition of the Src-family kinases (SFKs). A recent publication has presented data supporting a model where calcineurin-NFAT signalling collaborates with Ras-Erk-AP-1 to activate Src which promotes the epithelial-mesenchymal transition (EMT) and subsequently initiates lineage specification in mouse ESCs (Li et al., 2011). Moreover, it was demonstrated that SFK inhibition blocked segregation of the GATA4 positive primitive endoderm. Based on these promising findings, we decided to further characterize mouse ESCs maintained in inhibitors of the SFKs to ensure that the key features of pluripotency were maintained prior to testing SFK inhibitors in six-factor human iPSCs.

mCherry-expressing AB2.2 mouse ES cells maintained in CHIR99021 plus A-419259, a potent pyrrolopyrimidine inhibitor of the SFKs and Leukemia inhibitory factor (LIF) named CHAL media were injected to morula / early blastocysts to assess their ability to contribute to the developing inner cell mass (ICM), a fundamental characteristic of pluripotent cells. mCherry-expressing AB2.2 mouse ESCs were generated by transfecting them with a PB-CAG-mCherry-PGK-Puro expression vector using Amaxa nucleofection according to the manufacturer's protocol (see material methods section Chapter 2). Following two rounds of

cell sorting, a homogenous mCherry positive cell line was generated. Serial FACS analysis revealed that the cells continued to stably express mCherry and no significant silencing was observed. Remarkably, we observed that in contrast to mouse ES cells cultured in serum/LIF or 2i/LIF media, the ES cells cultured in these inhibitors showed significant contribution in the region of the trophectoderm in addition to the expected ICM contribution **(Figure 3.2)**.

In an attempt to explain the apparent *in vitro* contribution to both embryonic and extra-embryonic lineages, the published literature was reviewed specifically focussing on the role of Src inhibition in the first fate specification. Interestingly, it has been shown that the Calcineurin-NFAT pathway is essential for early lineage segregation as its inhibition at the eight-cell stage in mouse embryos arrests development at the morula blocking the first lineage segregation into the trophectoderm and the inner cell mass (Li et al., 2011).

3.3 Hypothesis

If this observation was correct that modulation of Src, a pre-implantation signalling pathway could confer some degree of expanded potential to conventional mouse embryonic stem cells, it was plausible to consider that in order to capture and maintain stem cells that have features of blastomeres, one possible approach is to modulate additional signalling pathways known to be important in pre-implantation development.

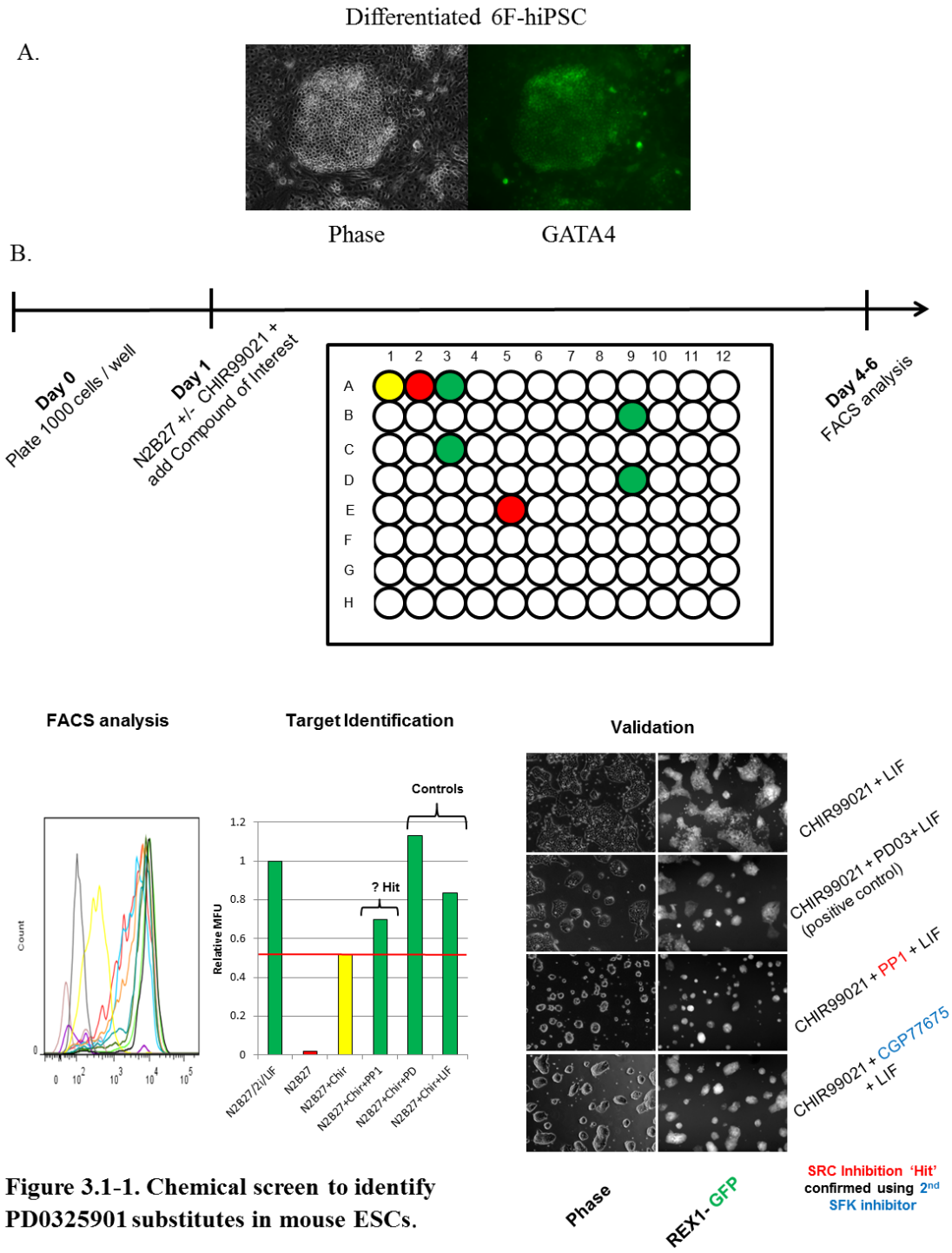


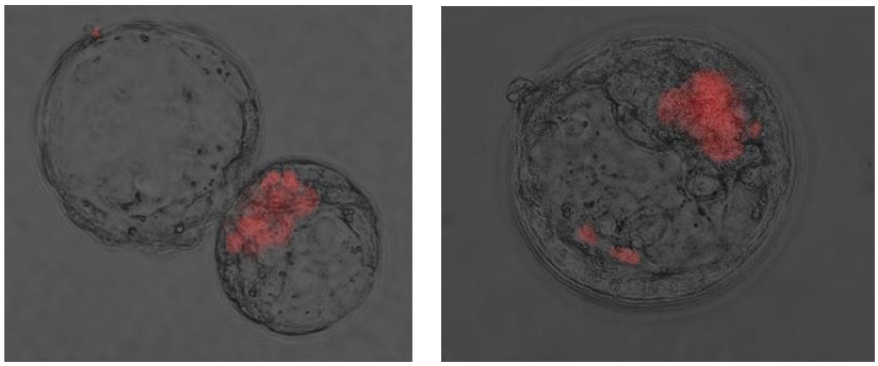
Figure 3.1-1. Chemical screen to identify PD0325901 substitutes in mouse ESCs.

Figure 3.1-1 and 2. Chemical screen to identify PD0325901 substitutes in mouse ESCs.

(A) Primary human neonatal and adult fibroblast cells reprogrammed to iPSCs using the six-factor combination had detectable biased differentiation to GATA4 positive cells despite the presence of 20% Knockout Serum Replacement, a Mek1 inhibitor (PD0325901), a GSK3 inhibitor (CHIR99021) and Leukaemia inhibitory factor (LIF) referred to as KSR2iL (Wang., 2011 and Ying ., 2008).

(B) Schematic representation of the chemical screen strategy used to identify potential substitutes for PD0325901 (Mek1 inhibitor). On Day 0, 1000 cells were plated in each well of a 96-well plate in triplicate. On Day 1, the media was changed to N2B27 plus the GSK3 inhibitor CHIR99021 or N2B27 alone in the presence or absence of the inhibitor of interest. After 4-6 days, 96-well FACS analysis was performed to identify any target compound(s) that could maintain *Rex1-GFP* expression. In each experiment, N2B272i, N2B272iLIF and N2B27CHIR/LIF were included as positive controls. Cells cultured in N2B27 alone served as an internal negative control with complete loss of *Rex1-GFP* expression identified. If a potential 'hit' was identified, the compound was first tested in its ability to support undifferentiated ESC self-renewal in the presence of CHIR99021 +/- LIF. Next, the identified pathway was confirmed by testing an additional commercially available selective inhibitor from the same signalling family. In the target identification panel, PP1, a Src family kinase (SFK) inhibitor was identified as a potential hit so to validate, a second commercially available inhibitor of the SFK family, CGP77675 was used.

A.



KSR2iL CHAL

B.

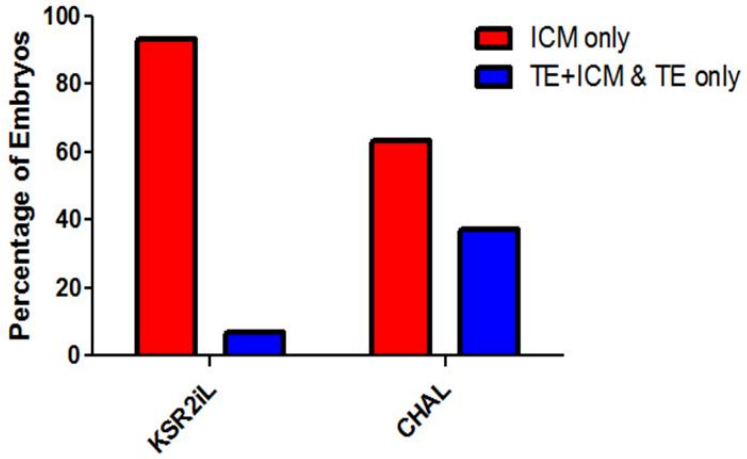


Figure 3.2. CHAL cultured ESCs contribute to the extra-embryonic lineage.

mCherry-expressing AB2.2 cultured in conventional KSR2iL media and CHAL media containing CHIR99021 (CH), A-419259 (A), a potent SFK inhibitor and LIF (L) were injected into morula stage embryos for blastocyst development. Blastocysts were observed for mCherry expression and the localisation of expression using epifluorescence microscopy (A) and (B). For each culture condition, 48-50 embryos were stained and scored. ESCs cultured in CHAL media demonstrated contribution to the trophectoderm and inner cell mass in 37% of embryos compared to ~ 7% in the control KSR2iL cultured ESCs. Thanks to Jian Yang and Lilianes Antunes for assistance.

Genetic and developmental studies have implicated the conserved mitogen-activated protein kinases (MAPKs), Src, and the Hippo pathways in the segregation of the TE and ICM lineages and in the development of pre-implantation embryos (Abad et al., 2013; Li et al., 2011; Lu et al., 2008; Maekawa et al., 2007; Maekawa et al., 2005; Natale et al., 2004; Nishioka et al., 2009; Saba-El-Leil et al., 2003; Strumpf et al., 2005).

As previously described, inhibiting Src completely eliminates CDX2-expressing TE cells in the pre-implantation embryo (Li et al., 2011). In addition, FGF and Wnt signalling pathways are implicated in producing enough pluripotent cells in the embryo for successful implantation (Morris et al., 2012). Furthermore, important functional interactions of the Wnt and MAPK pathways with Yap, the key Hippo pathway downstream effector, have recently been uncovered (Kapoor et al., 2014; Shao et al., 2014). We therefore reasoned that modulating the kinases and pathways mentioned above would interfere with segregation of the TE and ICM or possibly permit establishment of stem cells with expanded potential in culture.

3.4 Results

3.4.1 Media Development

We chose to modulate the pathways that are causally involved in the earliest pre-implantation developmental stages in particular the first lineage segregation, e.g. MAPKs, Src, and Wnt. To target MAPKs, we used Mek1 inhibitor PD0325901, JNK Inhibitor VIII (for Jun N-Terminal Kinase) and SB203580 (for p38). For Src family kinases (SFKs), A-419259, a potent pyrrolopyrimidine inhibitor of the SFKs, was chosen (Wilson et al., 2002). We also selected XAV939, which stabilizes AXIN1, the concentration-limiting component of the β-catenin and Yap destruction complex (Azzolin et al., 2014; Huang et al., 2009). GSK3

inhibitor CHIR99021 was included to promote metabolic and biosynthetic processes (Doble and Woodgett, 2003). The Leukemia inhibitory factor (LIF) was also added, as it enhances ESC self-renewal and may promote the rare totipotent cells in mouse ES cell culture (Morgani et al., 2013). The base media contained DMEM/F12 (Invitrogen) and 20% KnockOut serum replacement (Invitrogen), see methods. Hereafter, this medium that contains all six inhibitors is called the EPSC Medium or EPSCM, which was used in all mouse and human EPSC experiments in this thesis.

In collaboration with Dr Jian Yang PhD, Sanger Institute, mCherry-expressing AB2.2 mouse ESCs were then cultured in various combinations of inhibitors for at least 5 passages and then injected into eight-cell stage embryos for blastocyst development. Blastocysts were stained with antibodies against mCherry and CDX2 for co-staining. Adding A419259 or XAV939 to the KSR/2i/LIF (2i) medium substantially increased the expanded potential with ICM and TE contribution observed in 48% and 46% embryos respectively as compared to <7% in the KSR/2i/LIF control and 37% in Chir/ A419259/LIF (CHAL) media. Combined embryonic and extra-embryonic contribution was identified in 49% of embryos injected with cells cultured in the Chir/ A419259 / XAV939 and LIF without PD0325901. The addition of a JNKi and p38i to 2i media had moderate effects, with ICM and TE contribution in 15% and 24% respectively. Combined, these six inhibitors (EPSCM) produced the largest contribution with TE and ICM contribution observed at a frequency of 53%. **(Figure 3.3)**.

The conversion experiments were performed on two widely used ES cell lines, AB2.2 and E14Tg2a (Hooper et al., 1987; Ramirez-Solis et al., 1995). ESCs previously cultured in either serum-containing medium (15% serum plus LIF or M15) or 2i/LIF survived well in EPSCM

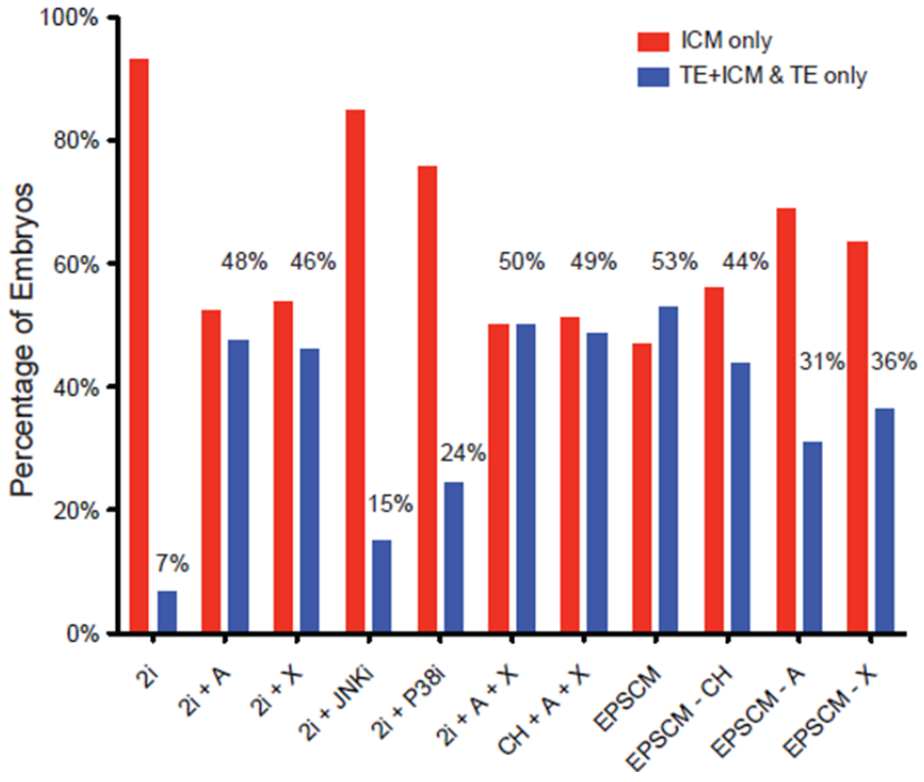


Figure 3.3. Media Development using in vitro embryo contribution assay.

Mouse ESCs (mCherry-expressing AB2.2) were cultured in various combinations of inhibitors for at least 5 passages and then injected into 8C embryos for blastocyst development. Blastocysts were stained with antibodies against mCherry and *Cdx2* for co-staining. A: A419259; X: XAV939; JNKi: JNK Inhibitor VIII; p38i: SB203580; CH: CHIR99021. All culture conditions contained LIF. For each culture condition, 48-50 embryos were stained and scored. Thanks to Jian Yang and Liliana Antunes for assistance.

and proliferated similarly to ESCs in M15 or 2i/LIF. We tested the plating efficiency of seeding a single *Rex1-GFP* ESCs (pre-cultured in 2i/LIF or M15) in either EPSCM or in their original media on SNL feeders. The numbers of GFP positive colonies formed in EPSCM or in 2i/LIF were comparable **(Figure 3.4)**, indicating that EPSCM was permissive for mouse ESCs and did not select for a rare subpopulation of cells. We used ESCs cultured in EPSCM for at least three passages in the subsequent studies. We also investigated whether EPSCM was able to convert EpiSCs to EPSCs. EpiSCs are derived from early post-implantation embryos and are distinct from ESCs in culture properties, gene expression, pluripotency and epigenetic profiles. They are believed to exist at a primed pluripotent state (Brons et al., 2007; Nichols and Smith, 2009; Tesar et al., 2007). Unlike ESCs, EpiSCs rapidly succumbed or differentiated in EPSCM, indicating that conversion or reprogramming by EPSCM is only permissible in ESCs (data not included).

3.4.2 Conventional mouse ESCs cultured in EPSCM contribute to the trophectoderm, and epiblast in blastocyst contribution assays.

In order to state expanded potential, one must demonstrate contribution to the epiblast ($OCT4^+$) and the trophectoderm ($CDX2^+$) therefore, we repeated the contribution experiment in order to co-stain for mCherry and the above lineage markers. mCherry-expressing AB2.2 mouse ESCs were injected into morulas which later developed into blastocysts *in vitro*. In 8 of 17 (47.1%) embryos injected with EPSCM cultured cells, mCherry+ cells were found in both the ICM ($OCT4^+$) and the TE ($CDX2^+$), respectively **(Figure 3.5A - D)** , whereas 2i/LIF ESCs did not contribute to the TE in the injected embryos (n=22) in this experiment.

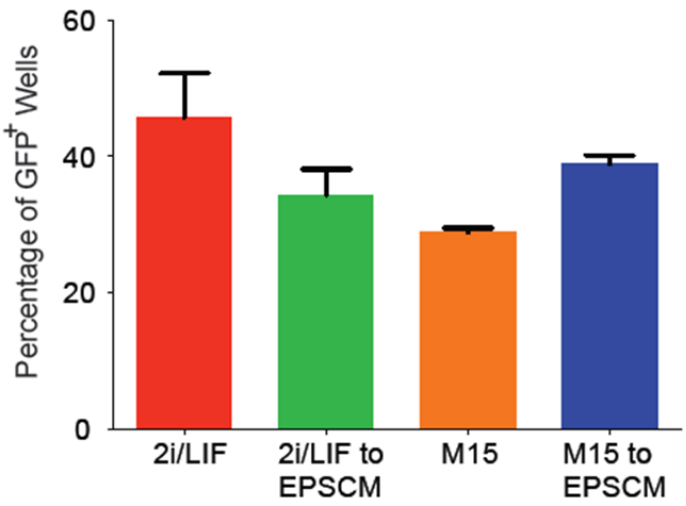
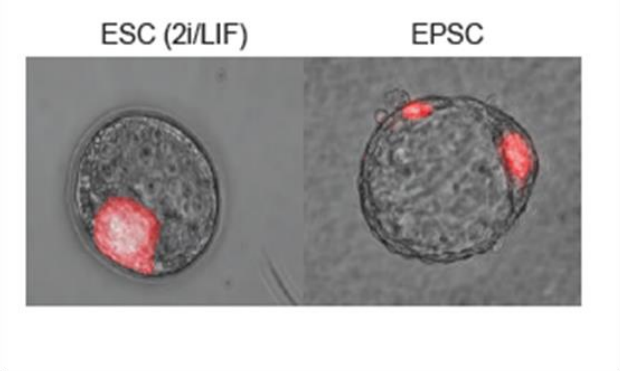


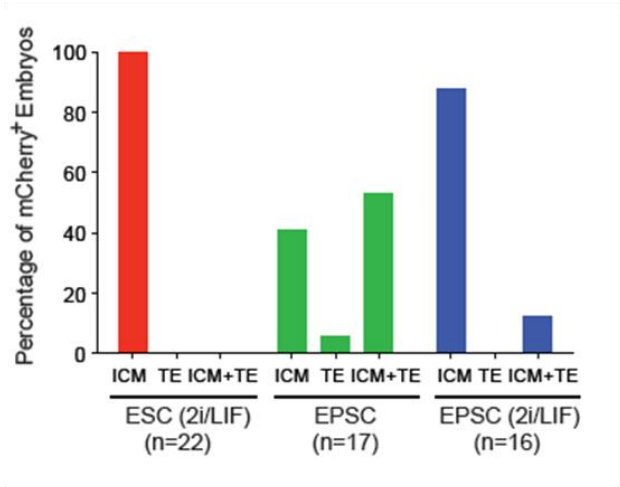
Figure 3.4. Single cell plating of Rex1-GFP ESCs demonstrated that EPSCM was permissive for mouse ESCs and did not select for a rare subpopulation of cells.

Single *Rex1*-GFP ESCs cultured either in 2iLIF or M15 were plated into SNL feeder plates in the original medium or in EPSCM. ESCs cultured in 2iLIF or M15 formed similar numbers of GFP+ colonies in EPSCM. Data are the mean ± s.d. from three experiments.

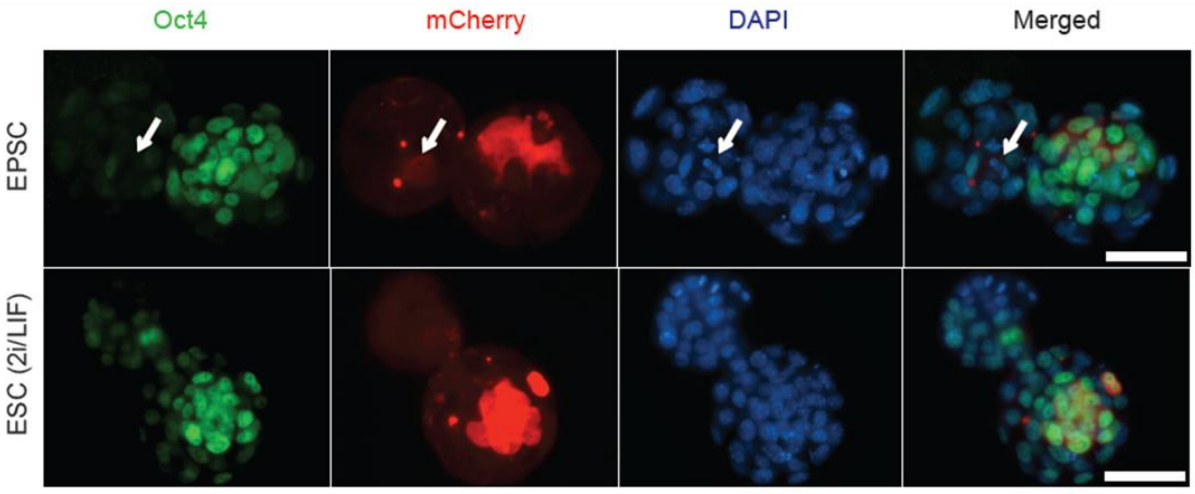
A.


ESC (2i/LIF)
EPSC

B.


Percentage of mCherry+ Embryos
100
80
60
40
20
0
ICM TE ICM+TE ICM TE ICM+TE ICM TE ICM+TE
ESC (2i/LIF) (n=22)
EPSC (n=17)
EPSC (2i/LIF) (n=16)

C.


Oct4
mCherry
DAPI
Merged
EPSC
ESC (2i/LIF)

D.

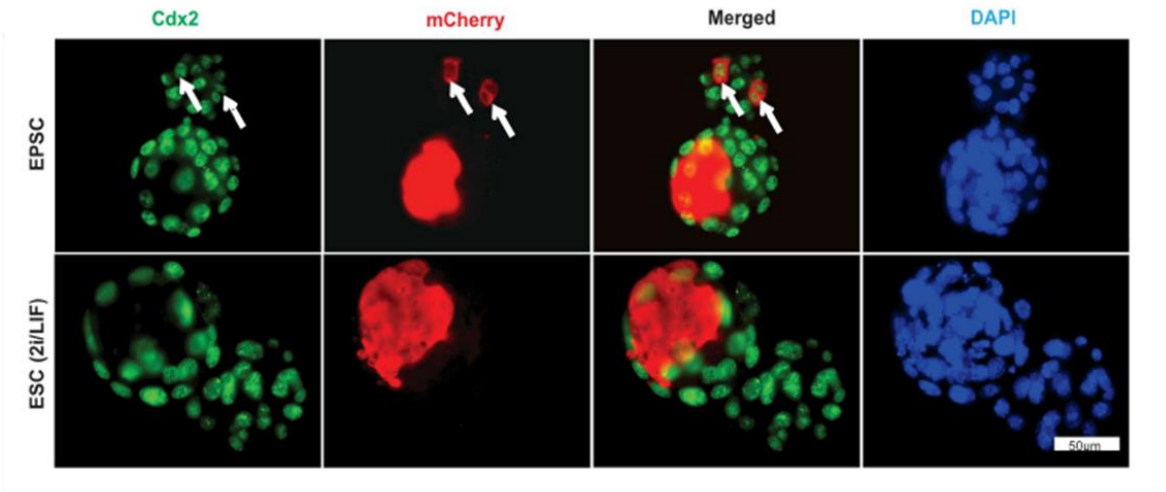


Figure 3.5. EPSCM cultured mouse ESCs contribute to both the ICM and the TE in blastocysts.

(A) Passage 5 mCherry-expressing AB2.2 cultured in conventional KSR2iL media and EPSCM were injected into morula stage embryos for blastocyst development. Contribution of EPSCs (mCherry+) to both the ICM and the TE in the blastocysts is demonstrated. (B) Quantification of contribution and demonstration of the loss of expanded potential of EPSCs once they are cultured in 2i/LIF for three passages. (C) and (D) 8-cell host embryos were injected with mCherry-expressing ESCs or EPSCs. The resultant blastocysts were stained with antibodies to mCherry,, *Oct4* and to *Cdx2*. Embryos were stained with DAPI for the nucleus. The arrows in (C) and (D) indicate cells co-expressing *Oct4* or *Cdx2* and mCherry. respectively Thanks to Jian Yang for embryo culture, fixation and immuno-staining.

Interestingly, once EPSCs were returned to 2i/LIF for three passages, a much lower contribution to the TE was observed **(Figure. 3.5B)**, indicating that EPSCM is necessary to maintain the expanded potential.

The functions of these inhibitors on conferring the capacity for extra-embryonic contribution were further confirmed by the removal of individual inhibitors from EPSCM. Removal of XAV939 or A419259 individually from EPSCM reduced the observed rate of ICM and TE contribution to 36% and 31% respectively **(Figure 3.3)**. The reduction in observed extra-embryonic contribution following the removal of CHIR99021 was less dramatic at 44% compared to 53% in the EPSCM control **(Figure 3.3)**. However, CHIR99021 was essential in maintaining the integrity of the culture. AB2.2 ESCs were cultured in EPSCM or EPSCM minus CHIR99021 for five passages and one thousand cells from each culture condition were re-plated to form single colonies. Removing CHIR99021 substantially decreased AP+ colony numbers **(Figure 3.6)**.

3.4.3 *In vitro* trophoblast differentiation of EPSCM cultured mouse ESCs

Cdx2, *Eomes*, *Gata3* and *Elf5* have been shown to promote or regulate the transcriptional network in TE development (Chen et al., 2010). Rapid up-regulation of these genes indicates the commitment of embryonic cells to trophoblast differentiation. After 2 days in Trophoblast stem cell (TSC) medium, expression of *Cdx2* and *Eomes* was increased 10-20-fold, indicating an increased propensity of these cells toward trophoblast differentiation **(Figure 3.7A)**. By contrast, no obvious *Cdx2* or *Eomes* expression changes were detected in the control ESCs (serum or 2i/LIF) under the same differentiation conditions **(Figure 3.7A)**, even though 2i/LIF ESCs in general have a hypomethylated genome (Leitch et al., 2013), indicating that DNA demethylation alone is not sufficient for the expanded potential. Following a published protocol (Abad et al., 2013), those cells from EPSCs were differentiated into PLACENTAL

LACTOGEN 1 (PL-1) positive trophoblast-like cells following the removal of FGF4 **(Figure 3.7B)**. Trophoblast stem cells were differentiated in parallel as a control.

3.4.4 *In vitro* characterization of converted mouse EPSCs

3.4.4.1 Pluripotency gene expression

Zfp42 (Rex1) and *Pou5f1 (Oct4)* are pluripotency genes (Toyooka et al., 2008). *Rex1-GFP* or *Oct4-GFP* reporter ESCs (Guo et al., 2011; Silva et al., 2008) cultured in EPSCM retained robust GFP expression as seen using microscopy and additionally confirmed by FACS analysis **(Figure 3.8)**. Quantitative PCR was performed which demonstrated that mouse ESCs cultured in EPSCM expressed similar levels of pluripotency genes, and lineage-specific genes, including *T (Brachyury*), *Gata6*, *Sox1*, *Cdx2*, *Eomes* and *Elf5*, as in standard ESCs **(Figure 3.9A and 3.9B)**.

3.4.4.2 Oct4 Enhancer Activity in EPSCs

The distal enhancer of the *Oct4* locus is specifically active in naïve ESCs, whereas the proximal enhancer dominates in epiblast stem cells (EpiSCs) (Bao et al., 2009; Yeom et al., 1996). Functionally, activity of the distal enhancer of the *Oct4* locus is critical in maintaining ES cell identity. Transcription-activator-like effectors (TALEs) activators targeting the distal enhancer of the *Oct4* locus induces epigenetic change, reactivates endogenous expression and can substitute for exogenous *Oct4* in the reprogramming of mouse embryonic fibroblast cells (MEFs) to induced pluripotent stem cells (iPSCs). Furthermore, repression of the *Oct4* distal enhancer induces rapid differentiation of ES cells (Gao et al., 2013). *Oct4* enhancer activity in EPSCM was interrogated using the *Oct4* distal and proximal enhancer luciferase reporter

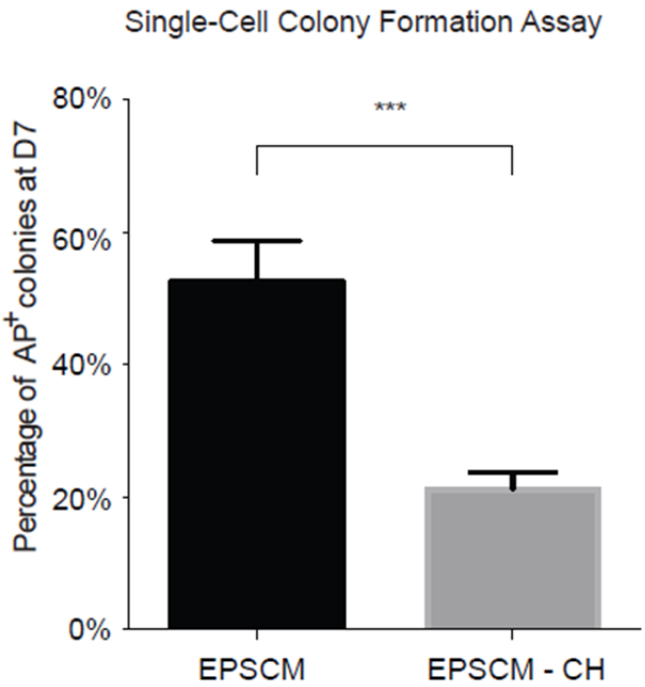
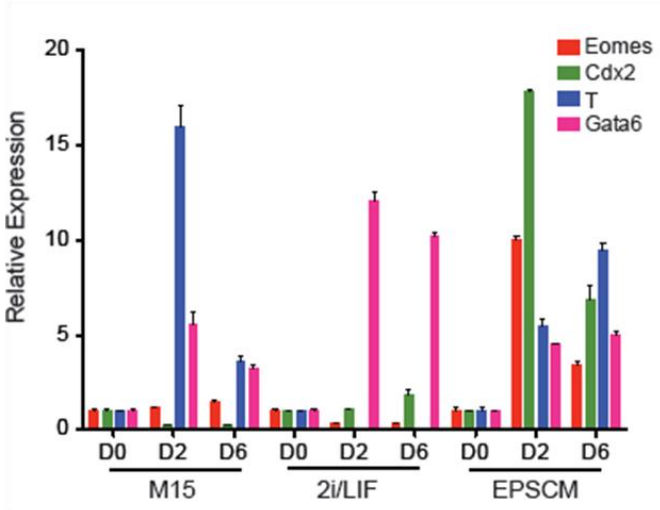


Figure 3.6. CHIR99021 is essential for culture stability in EPSCM.

The effect of CHIR99021 in EPSCM on cells. One thousand EPSCs cultured in either EPSCM or EPSCM minus CHIR99021 were plated for single cell colony formation. Colony numbers were scored on day 7. Data are the mean ± s.d. from three experiments. ***: $p<0.0005$.

A.



B.

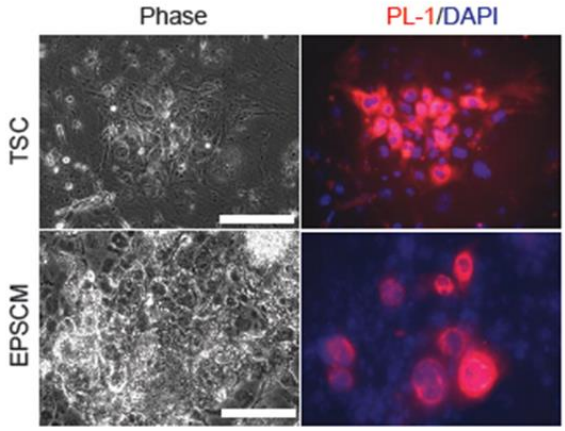
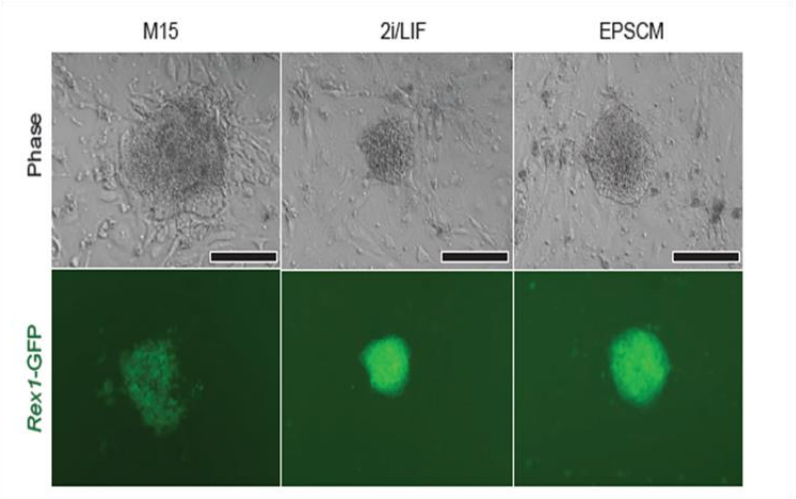


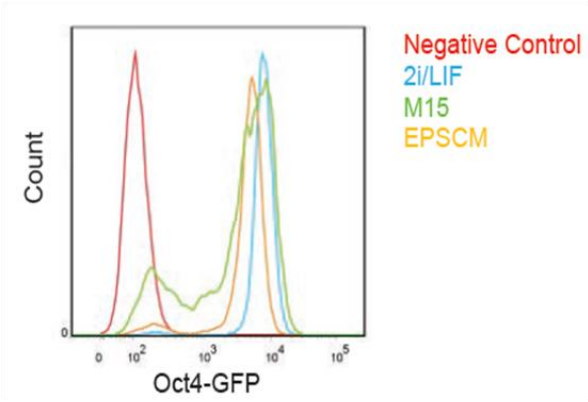
Figure 3.7. Mouse ESCs cultured in EPSCM can be differentiated to trophoblast-like cells in vitro.

(A) Rapid up-regulation of trophoblast marker genes *Cdx2* and *Eomes* when EPSCM converted ESCs are cultured in TSC medium. Only the expression of meso-endoderm markers (*T* and/or *Gata6*) was increased in ESCs (2i/LIF or M15). Expression was normalized to that of un-differentiated ESCs in M15, 2i/LIF or EPSCM, respectively. Data are mean ± s.d. Experiments were repeated three times. (B) Differentiation of EPSCM-cultured ESCs to *Placental Lactogen 1* (*PL-1*)-positive trophoblast-like cells following removal of FGF4 from the TSC medium (scale bar = 25 μm). TSCs were differentiated as the control.

A.



B.



C.

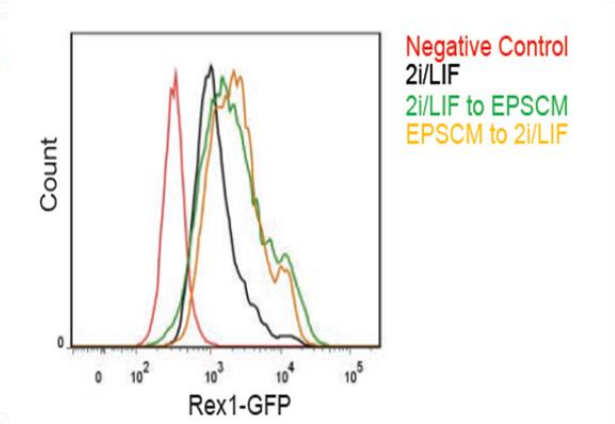
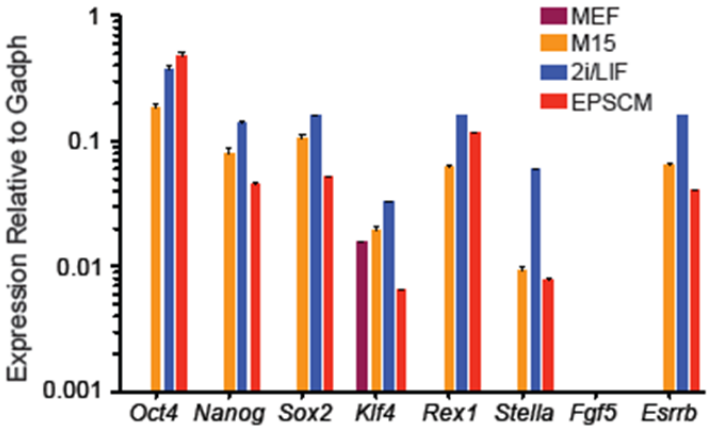


Figure 3.8. Characterization of mouse ESCs cultured in EPSCM.

(A) *Rex1*-GFP reporter ESC colonies grown in M15, 2i/LIF and EPSCM on SNL feeders (scale bar = 50 μm). Top panels, phase images; bottom panels, GFP images. (B) Expression of the pluripotency markers *Oct4* in *Oct4*-GFP ESCs (E14) cultured in M15, 2i/LIF or EPSCM (passage 3). In all conditions, *Oct4* expression detected by GFP is comparable detected by flow cytometry. (C) *Rex1*-GFP ESCs (AB2.2 background) also have comparable levels of *Rex1* expression (GFP) when cultured in EPSCM or when subsequently re-cultured in 2i/LIF. Negative control: wild-type ESCs.

A.



B.

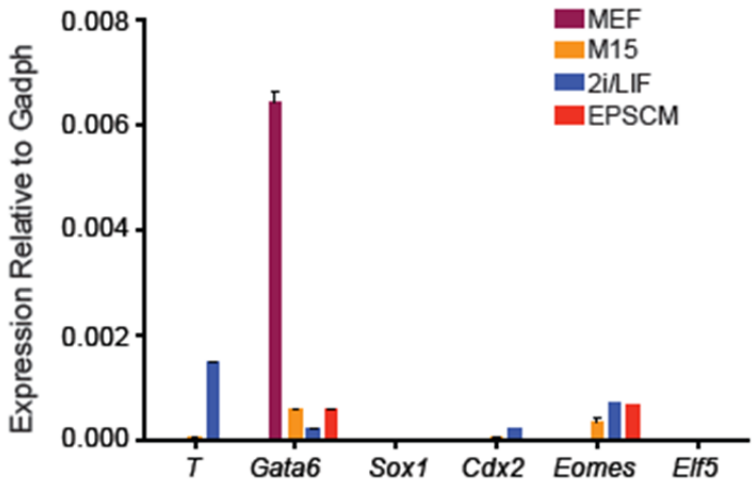


Figure 3.9. Expression of key pluripotency genes and lineage specifiers in EPSCM cultured ESCs.

(A) Expression of pluripotency genes in ESCs cultured in EPSCM. Relative expression of these genes is normalized to *Gadph*. Data are mean ± s.d. Experiments were repeated three times. (B) Expression of lineage-specific genes in ESCs cultured in EPSCM. Relative expression of these genes is normalized to *Gadph*. Data are mean ± s.d. Experiments were repeated three times. Thanks to Wei Wang for assistance with optimization and analysis of the expression data.

constructs schematically represented in **Figure 3.10A**. Similar to standard ESCs, ESCs in EPSCM preferentially used the distal enhancer **(Figure 3.10B)**.

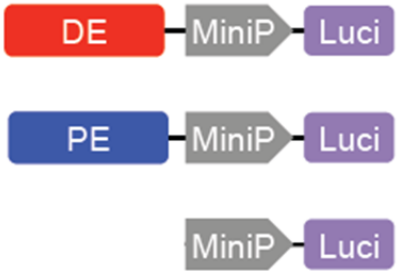
3.4.4.3 Somatic differentiation of EPSCs *in vitro* and *in vivo*

Mouse ESCs were cultured in EPSCM for over 30 passages and were able to differentiate *in vitro* into cell types of all three somatic germ layers **(Figure 3.11A)**. Beta 3-Tubulin expressing neurons, α-smooth muscle actin (SMA) muscle cells and Gata4 endodermal derivatives were the three *in vitro* representative cells types of ectoderm, mesoderm and endoderm differentiation respectively. Following gene targeting of a GFP cassette to the *Rosa26* locus with approximately 60% efficiency (thanks to Dr Wei Wang, Sanger Institute) (Dupuy et al., 2009) these cells were injected into blastocysts. The targeted cells contributed efficiently to live chimeras with no obvious abnormalities at 10 months **(Figure 3.11B)**. Furthermore, donor cells contributed to the germline as confirmed by PCR based detection of the *Rex1-GFP* allele from the injected EPSCM-cultured ESCs **(Figure 3.11C)**. ESCs cultured in EPSCM therefore retained canonical pluripotency features.

3.4.5 Western blot confirms inhibition of targeted kinase activity in EPSCM.

We next confirmed inhibition of the respective kinase activities in EPSCM **(Figure 3.12A)**. When compared to the KSR2iLIF control, the ERK, JNK, p38 and Src pathways were appropriately inhibited with reduction in their overall level of protein phosphorylation. EPSCM contains XAV939, which stabilizes AXIN1 by inhibiting its ubiquitination (Huang et al., 2009). EPSCs had considerably elevated levels of AXIN1 compared with controls **(Figure 3.12A)**. The elevated AXIN1 levels caused accumulation of phosphorylated β-

A.



B.

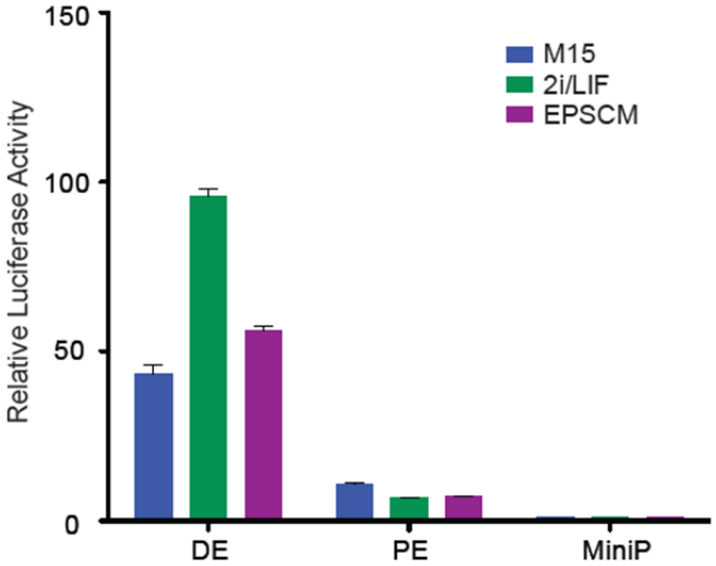
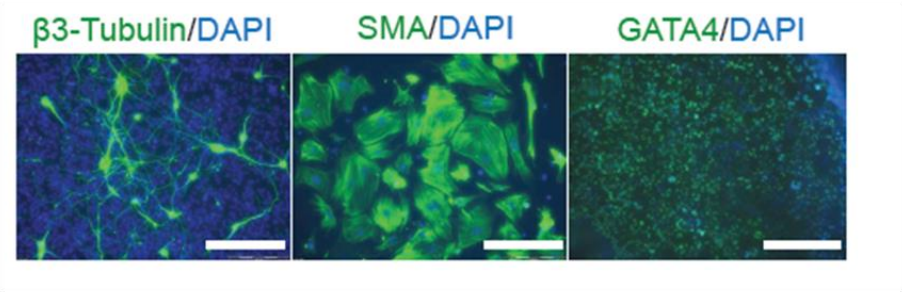


Figure 3.10. *Oct4* Distal Enhancer is preferentially active in EPSCM cultured mouse ESCs.

(A) Schematic representation of *Oct4* distal and proximal enhancer luciferase reporter constructs. DE: the *Oct4* distal enhancer; PE: the *Oct4* proximal enhancer; MiniP: the minimum promoter. (B) The *Oct4* distal enhancer is active in ESCs cultured in EPSCM. Luciferase activities of the *Oct4* DE and PE constructs were normalized to that of the MiniP construct in the same cell type. Data are mean ± s.d. from three experiments. Thanks to Xuefei Gao for assistance with the luciferase assay.

A.



B.

C.

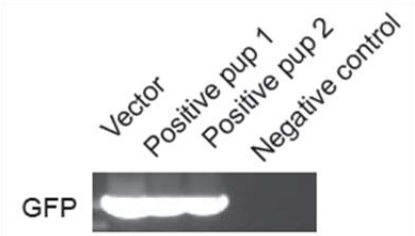


Figure 3.11. EPSCs can differentiate to all somatic lineages in vitro and in vivo

(A) ESCs cultured in EPSCM can differentiate into all three germ layers in vitro. Beta 3-Tubulin, α-smooth muscle actin (SMA) and Gata4 antibodies were used to detect the differentiated cell types representative of the three germ layers. (B) Live chimeras from injecting ESCs cultured in EPSCM into mouse blastocysts. (C) Germline transmission of the *Rex1*-GFP allele from the injected EPSCM-cultured ESCs. Thanks to Wei Wang for assistance with Rex1-GFP allele germline genotyping.

catenin in both the cytoplasm and nuclei **(Figure 3.12B)** and decreased active β-catenin in the nucleus **(Figure 3.12B)**. The TopFlash luciferase assay confirmed reduced β-catenin-LEF/TCF activity **(Figure 3.12C)**.

The signalling dependence of EPSCM-cultured ESCs was assessed by plating one hundred ESCs cultured in EPSCM, 2i/LIF or M15 in the presence of inhibitors targeting the JAK2, FGFR, TGFBR or ALK5 pathways. Alkaline phosphatase (AP) positive colony number was scored after 10 days. Mouse ESCs in EPSCM were dependent on the Jak/Stat pathway, as a Jak inhibitor substantially reduced the number of AP positive colonies from these cells **(Figure 3.13A)**. Adding LIF increased phosphorylation of STAT3 in a concentration dependent fashion **(Figure 3.13B)** and up-regulated the expression of Jak/Stat downstream genes, *Socs3* and *Egr1* **(Figure 3.13C)**. By contrast, ESCs cultured in EPSCM did not appear to be sensitive to FGFR or TGFBR inhibitors, similar to 2i/LIF ESCs **(Figure 3.13A)** (Ying et al., 2008).

3.4.6 Establishing de novo EPSC lines from pre-implantation embryos.

3.4.6.1 Eight-cell stage embryo culture and line derivation

I next explored deriving stable cell lines from pre-implantation embryos using EPSCM. I cultured eight-cell stage embryos in M15, 2i/LIF or EPSCM either on feeder cells or feeder-free **(Figure 3.14A)**. On day two, 19 of 42 (45.2%) embryos in M15 and 17 of 45 (37.7%) in feeder-free EPSCM progressed to blastocysts. In 2i/LIF, only 5 of 54 (9.3%) embryos reached the blastocyst stage, but these embryos died soon after. All blastocysts in M15 hatched by day 5. In striking contrast, the blastocoel in many blastocysts in EPSCM was obliterated and filled with cells, generating a structure reminiscent of an enlarged morula.

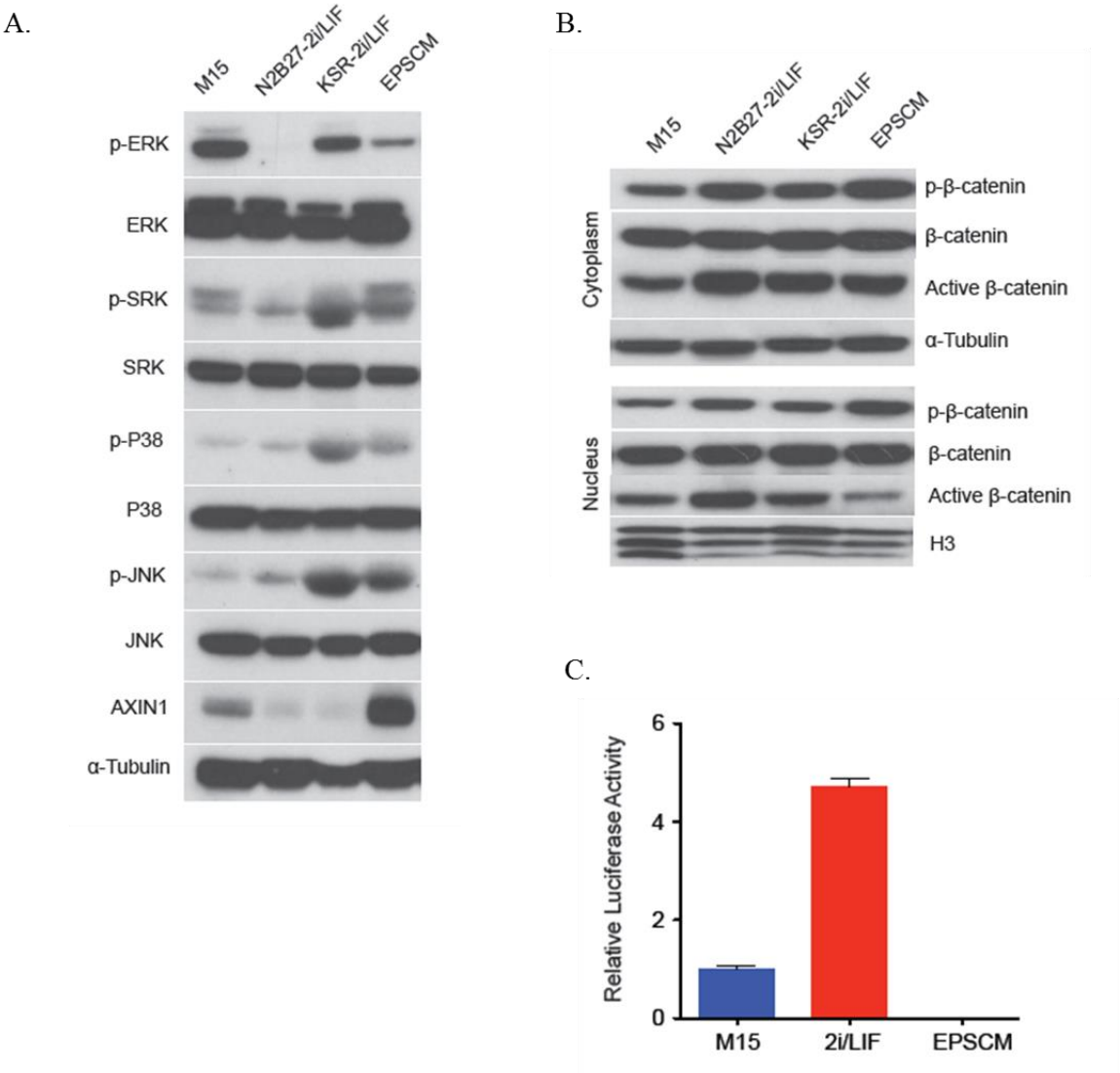
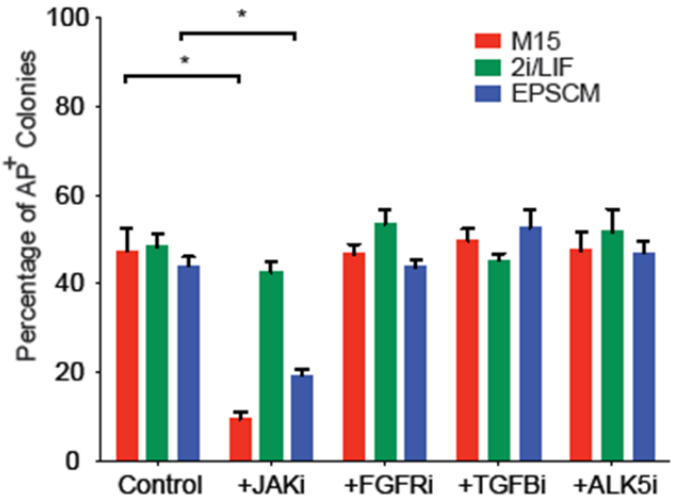


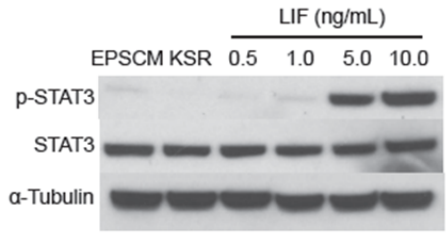
Figure 3.12. Western blot confirmation of inhibition of targeted kinase activity in EPSCM

(A) The effects of small-molecule inhibitors in EPSCM on their respective targets. p-ERK: phosphorylated ERK; p-SRK: phosphorylated SRK; p-P38: phosphorylated P38; p-JNK: phosphorylated JNK. α-tubulin was used as the loading control. (B) Increased levels of phospho-β-catenin and decreased active β-catenin in the nuclei of EPSCM-cultured ESCs. Histone H3 was used as the protein-loading control. (C) TopFlash assay shows lower activity of the canonical Wnt pathway in EPSCM-cultured ESCs. Data are the mean ± s.d. from three independent repeats. Thanks to Jian Yang for assistance with the above assays.

A.



B.



C.

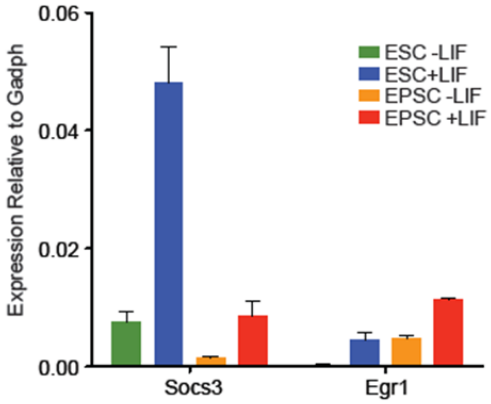


Figure 3.13. Mouse ESCs in EPSCM were dependent on the Jak/Stat pathway.

(A) Signalling pathway dependence in EPSCM-cultured ESCs. One hundred ESCs cultured in EPSCM, 2i/LIF or M15 were plated and cultured with inhibitors of JAK2, FGFR, TGFBR or ALK5. AP+ colonies were scored after 10 days. Data are the mean ± s.d. from four independent repeats. *: $p<0.05$. (B) Increased p-STAT3 in EPSCM-cultured ESCs with increased concentrations of LIF in the medium. (C) Up-regulation of LIF pathway downstream genes in EPSCM-cultured ESCs in response to LIF stimulation. Again, thanks to Jian Yang for help with the STAT3 signalling experiments.

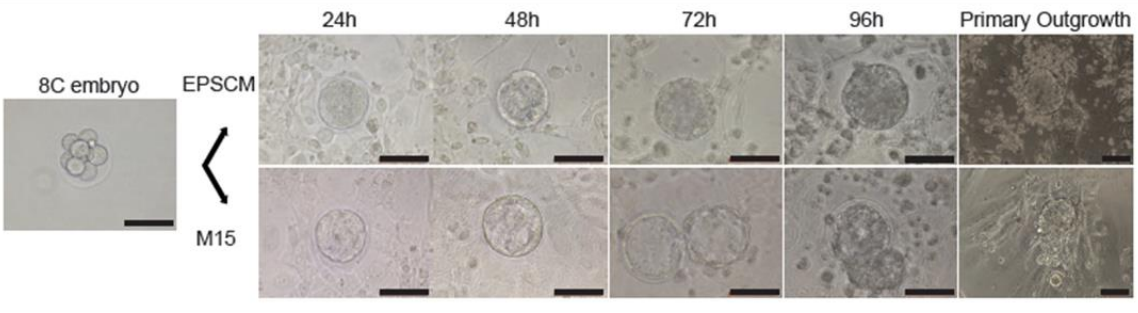
This structure was described as a “filled” blastocyst **(Figure 3.14A and Figure 3.14B)**. All the blastocysts in EPSCM developed into “filled” blastocysts by day 6, and eventually almost all attached and hatched. Primary outgrowths appeared after day 10. Stable cell lines were established from eight-cell stage embryos in EPSCM with an efficiency of approximately 20% under feeder-free condition, and up to 100% on SNL feeder cells **(Figure 3.14A)**.

3.4.6.2 EPSCM induces a complex embryo reprogramming process involving cell death and proliferation accompanied by the disappearance of the TE and ICM cells

To monitor the dynamic changes in pre-implantation embryos in EPSCM, immunostaining with antibodies against OCT4 and CDX2 was performed. From the four-cell stage to the blastocyst, the expression of OCT4 and CDX2 in embryos in EPSCM was similar to that in embryos cultured in M15, where OCT4 protein expression was first detected in 4C and CDX2 protein in 8C embryos **(Figure 3.15A and 3.15B)** (Dietrich and Hiiragi, 2007; Ralston and Rossant, 2008). OCT4 was later restricted to the ICM, and CDX2 to the TE. Interestingly, we observed a gradual loss of both OCT4 positive and CDX2 positive cells in the “filled” blastocysts **(Figure 3.15A).** Eventually, the remaining cells in the embryos had large nuclei, but most expressed neither OCT4 nor CDX2 in the nucleus. Several days later, a population of OCT4 positive cells with small nuclei re-emerged at the periphery of the “filled” blastocysts. Once the embryos hatched, the cells in the early outgrowths expressed OCT4 and/or CDX2 but eventually became uniformly OCT4 positive and CDX2 negative **(Figure 3.15A).**

In an attempt to understand the cause of the loss of OCT4 positive and CDX2 positive cells in the “filled” blastocysts, proliferation and apoptosis were assessed by antibody staining

A.



B.

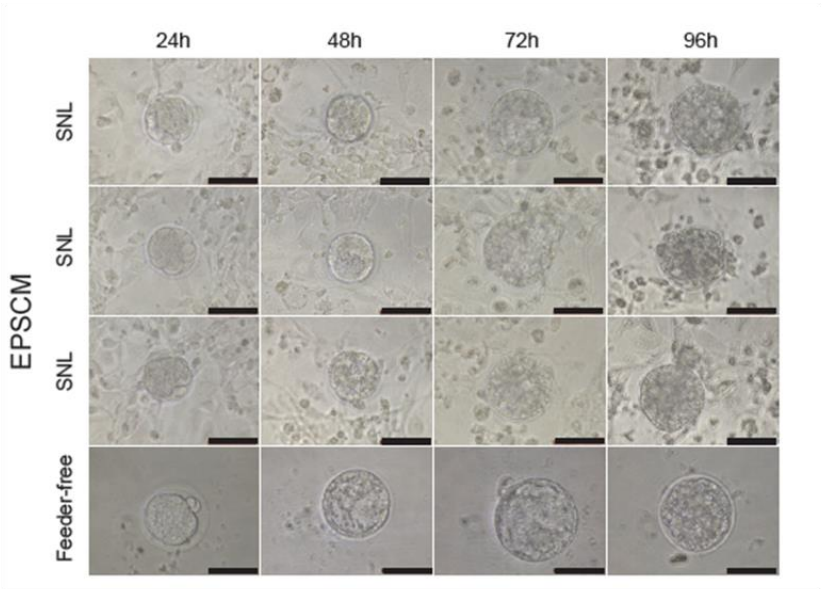


Figure 3.14. Derivation of EPSC lines from preimplantation embryos

(A) Development of 8C mouse embryos in EPSCM or in M15 (scale bar = 50 μm). (B) 8C embryo development in EPSCM on SNL feeders or feeder-free. Scale bar=50 μm.

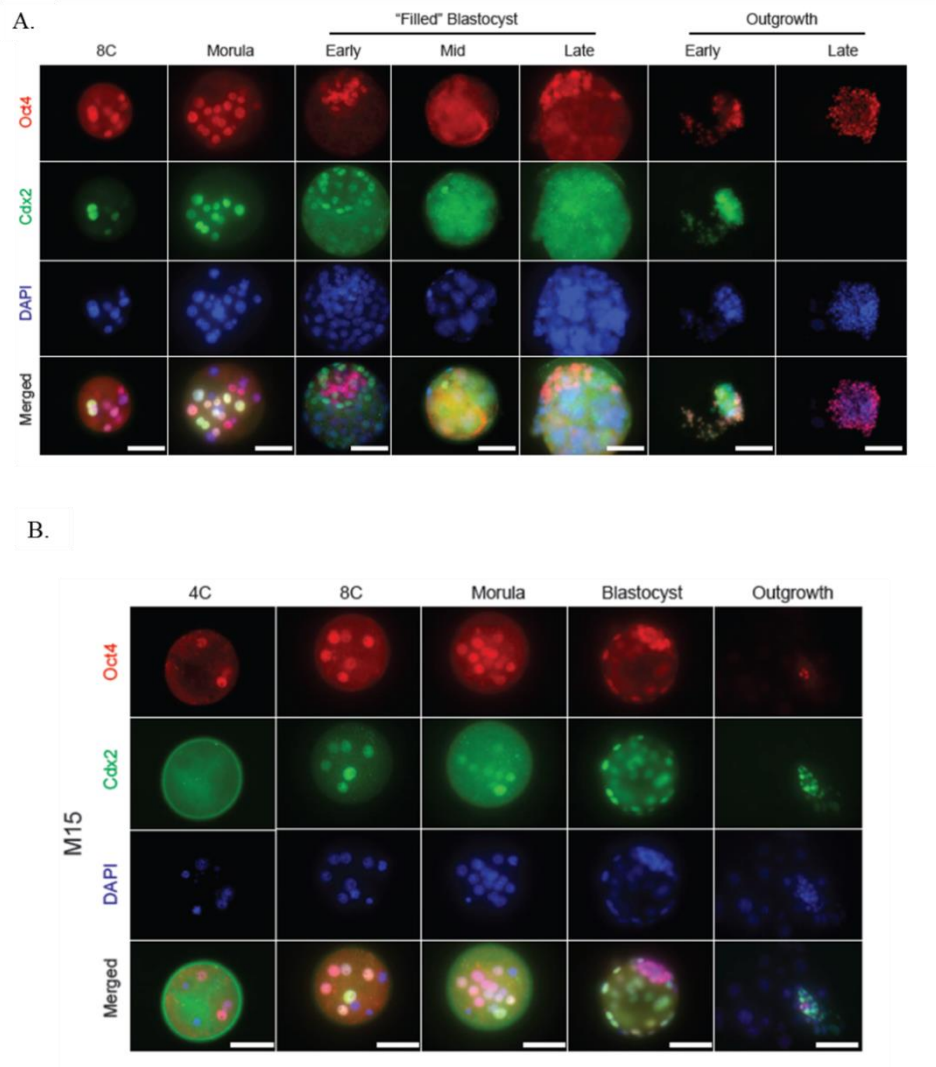


Figure 3.15. Immunofluorescent detection of OCT4 and CDX2 in embryos cultured in EPSCM.

(A) Most if not all cells in the mid-filled-blastocysts do not express either OCT4 or CDX2 (staining for both antibodies is background). OCT4+ cells emerged from the edge of the late "filled" blastocysts. In the outgrowths, many cells initially expressed CDX2, but subsequently all cells became uniformly OCT4+ and CDX2-. Early "filled" blastocyst: approximately 72-96 hours in EPSCM; mid, 5-6 days; late, approximately 7 days in EPSCM (Scale bar = 25 µm, and 100 µm for the late outgrowths). n=70 embryos in EPSCM were immunostained. (B) Immunofluorescent staining for OCT4 and CDX2 in pre-implantation embryos cultured in M15 as the control. 4C to 8C mouse embryos were cultured in M15 media, and embryos representative of each developmental stage were stained with OCT4 and CDX2 antibodies and imaged with fluorescence microscopy (scale bar=25 µm). n=50 embryos in M15 were immunostained.

against Ki67 and cleaved Caspase 3 respectively. Many cells in embryos cultured in EPSCM were Ki67 positive **(Figure 3.16A)**, indicating that they were still proliferating. However, apoptotic cells as detected by Caspase 3 antibody were clearly observed **(Figure 3.16B)**, particularly between the time when the OCT4 positive / CDX2 positive cells disappeared and when the OCT4 positive cells re-emerged. Therefore, the small OCT4 positive cells did not appear to directly arise from the blockade of the initial TE and ICM segregation. Nevertheless, EPSCM induced a complex reprogramming process involving cell death and proliferation accompanied by the disappearance of the TE and ICM cells, which warrants future investigation.

3.4.6.3 Derivation of new lines from single blastomeres using EPSCM

Single blastomeres from 8C embryos were plated into a well of a 96-well plate with SNL feeders **(Figure 3.17A)**. In the following days, in 9 of 32 (28.1%) wells with EPSCM, blastomeres proliferated and by day 12 formed large outgrowths consisting predominantly of OCT4 positive and CDX2 negative cells **(Figure 3.17A and Figure 3.17B)**, whereas the other blastomeres did not form obvious outgrowths. By contrast, no outgrowths were obtained in either M15 (n=32) or 2i/LIF (n=32) on feeders **(Figure 3.17C)**, or when blastomeres were cultured in feeder-free conditions in any medium (n=96).

3.4.6.4 Characterization of newly derived EPSC lines.

Two lines derived *de novo* from pre-implantation embryos in EPSCM: DR10 and DR25-EPSCs were chosen for further characterization. Two pre-implantation embryo-derived 2i/LIF ESC lines (DR4 and DR9) were included in the analysis. DR10-EPSCs and DR25-EPSCs expressed pluripotency genes at levels comparable to standard ESCs **(Figure 3.18A**

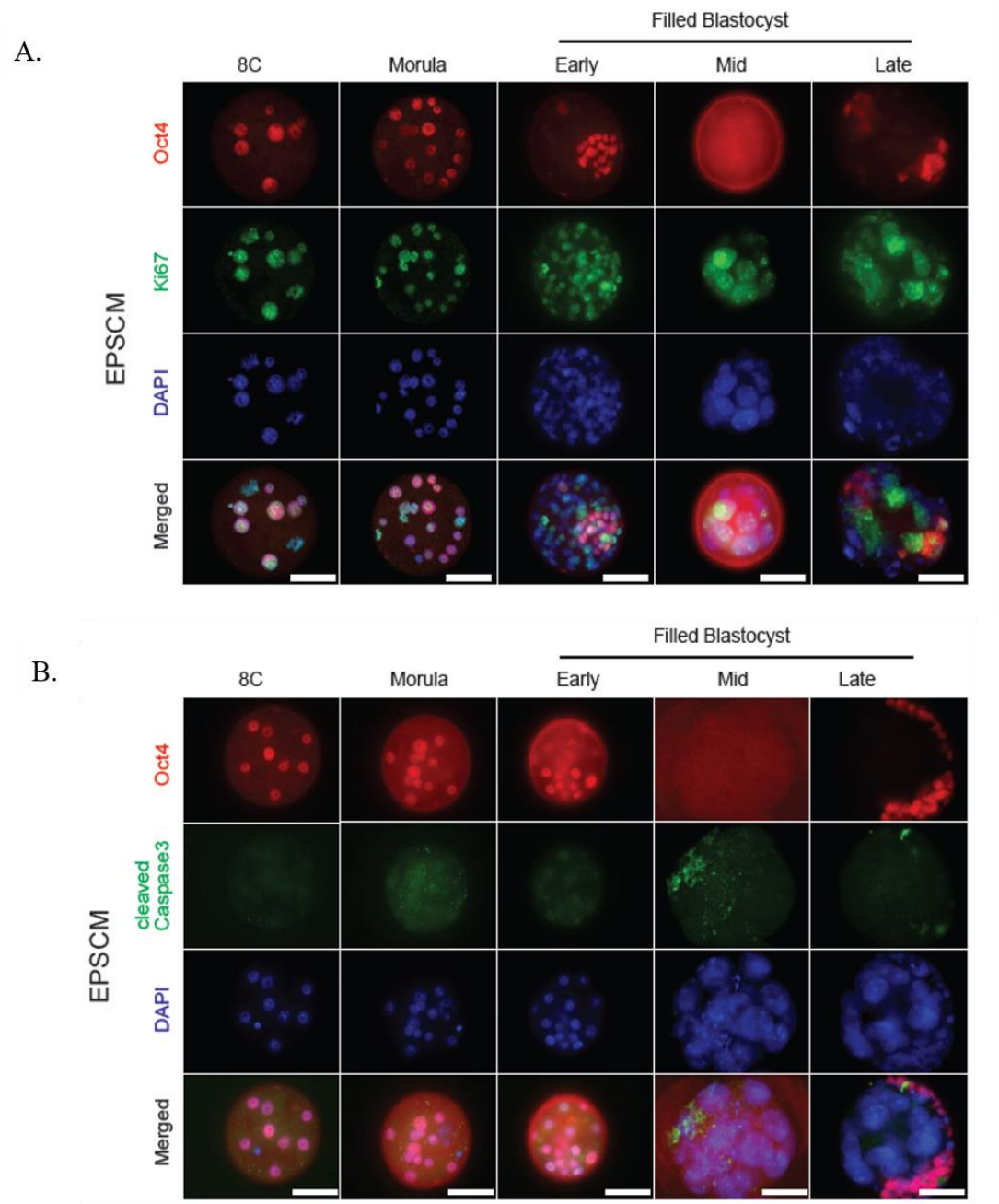
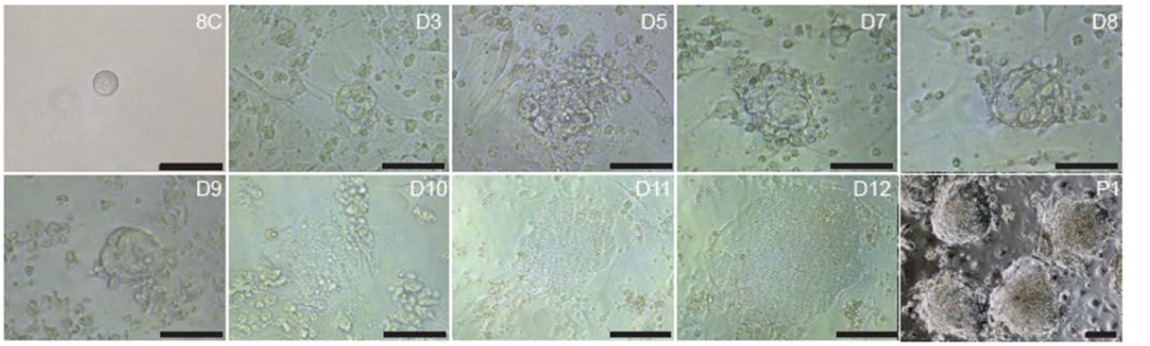


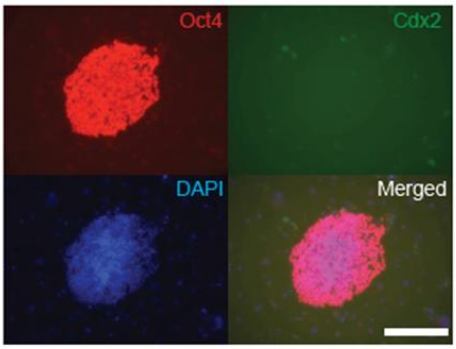
Figure 3.16. Immunofluorescent detection of Ki67 and Caspase 3 in embryos cultured in EPSCM.

(A) Immunofluorescent detection of OCT4 and Ki67 in 8C embryos cultured in EPSCM (scale bar = 25 μm) for proliferative cells. n=56 embryos. (B) Immunofluorescent detection of OCT4 and cleaved Caspase 3 in 8C embryos cultured in EPSCM (scale bar = 25 μm) for apoptotic cells. n=50 embryos analysed. As in Fig 3.15, thanks to Jian Yang for assistance with embryo culture, fixation and immuno-staining.

A.



B.



C.

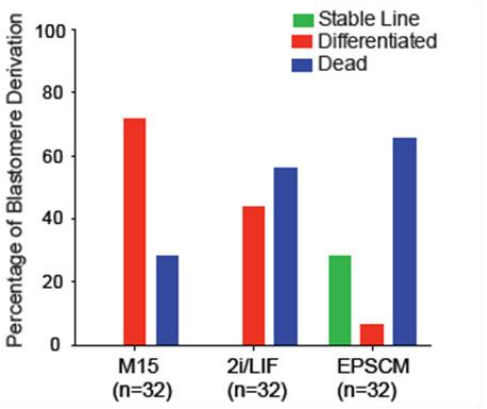


Figure 3.17. Derivation of new EPSC lines from single blastomeres from 8 cell stage embryos

(A) Single 8C blastomeres were seeded in individual wells of a 96-well plate on feeders in EPSCM. Temporal progression of the blastomere was imaged (scale bar = 50 µm). P1: passage 1 of a primary outgrowth. (B) Expression of *Oct4* or *Cdx2* in the primary outgrowth from a single blastomere (scale bar = 100 µm). Most cells were *Oct4*+ and *Cdx2*-. (C) Efficiency of derivation of stable EPSC lines from single blastomeres.

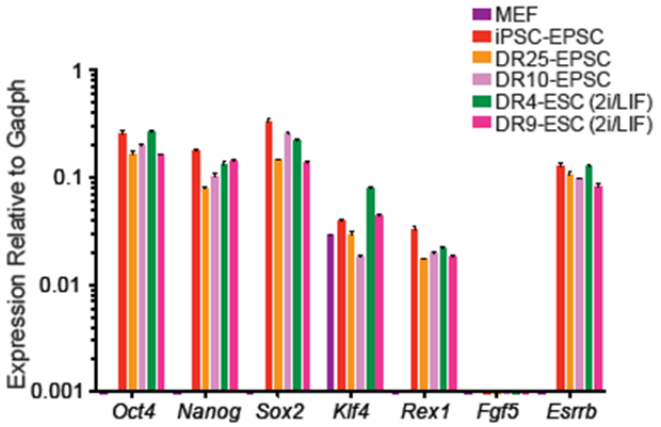
and 3.18B), had a normal karyotype **(Figure 3.19)**, were able to form mature teratomas **(Figure 3.20A)** and contributed to both somatic **(Figure 3.20B)** and germline lineages **(Figure 3.20C)** in chimeras. In keeping with best practice and satisfying the principles of the 3Rs strategy (reduce, refine, replace), germline contribution was determined by assessing contribution of mCherry positive EPSCs to the 14.5dpc gonad relative to the control 2i/LIF ESCs **(Figure 3.20C).**

Importantly, once injected into morulas, DR10-EPSCs contributed to both the inner cell mass and the mural trophectoderm in the formed blastocysts **(Figure 3.21A and 3.21B)**. We also asked whether we could obtain EPSCs directly from somatic cells (MEFs) by reprogramming them to iPS cells. We took advantage of the efficient six-factor reprogramming technology (Wang et al., 2011b; Yang et al., 2015) and cultured primary iPS cell colonies directly in EPSCM. Established iPS cells cultured in EPSCM for over 5 passages expressed levels of core pluripotency factors comparable to those of ESCs **(Figure 3.18A and 3.18B)**, and also contributed to both the ICM and TE in the blastocysts developed from injected morulas **(Figure 3.21A and 3.21B)**.

3.5 Discussion

In this chapter I have described the fortuitous sequence of events where a chemical screen designed to identify PD0325901 substitutes for the maintenance of naïve like human iPSCs identified the SFK pathway, inhibition of which conferred expanded potential to conventional mouse embryonic stem cells. With a knowledge of the role of Src in pre-implantation development based on published data, we speculated that inhibition of additional pathways known to be active in the pre-implantation embryo may permit establishment of a stable stem

A.



B.

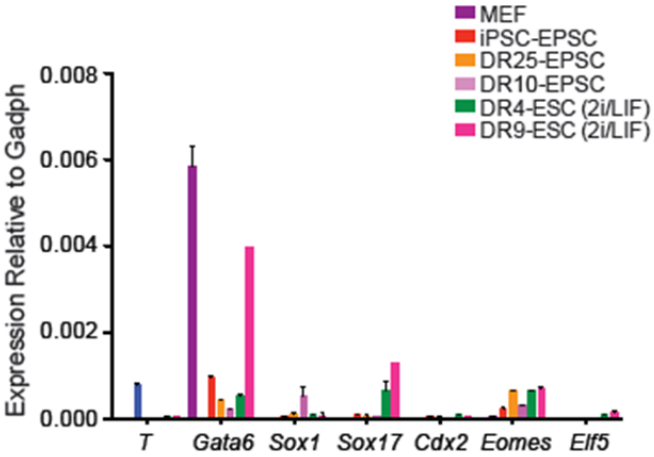


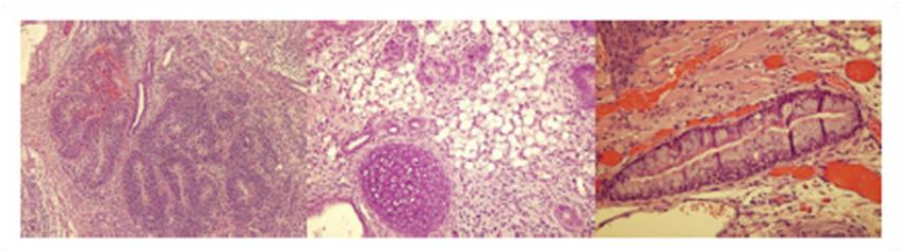
Figure 3.18. Pluripotency marker expression in EPSCs derived from preimplantation embryos

(A) Analysis of pluripotency markers in two preimplantation embryo-derived EPSC lines (DR25 and DR10), one somatically reprogrammed iPSC-EPSC line and two preimplantation embryo-derived 2i/LIF ESC lines (DR4 and DR9). Relative expression of pluripotency genes was normalized to *Gadph* in the qRT-PCR assay. Data are the mean ± s.d. from three experiments. (B) Low levels of lineage-specific genes in ESC or EPSC lines described in (A). Relative expression of the lineage-specific genes is normalized to *Gadph* in qRT-PCR assays. Data are the mean ± s.d. from three experiments. Thanks to Wei Wang for assistance with the quantitative PCR.

Figure 3.19. EPSCs derived from pre-implantation embryos are genetically stable

Spectral karyotyping of DR10-EPSCs (n=200 cells analysed).

A.



B.

C.

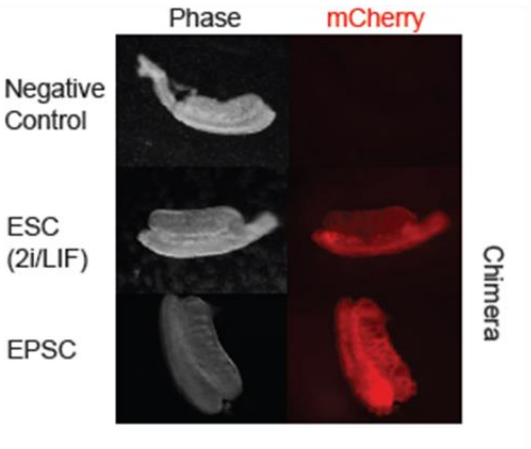
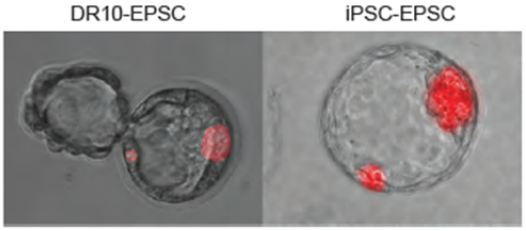


Figure 3.20. EPSCs derived from pre-implantation embryos contribute to somatic lineages in vivo.

(A) Mature teratomas from DR10-EPSCs. Left panel: neural tube-like structures from the ectoderm (neuroectoderm); middle panel: cartilage from the mesoderm; right panel: a gland of possible gastrointestinal type arising from the endodermal layer and muscle fibers derived from the mesoderm. The sections were stained with hematoxylin-eosin, and the photographs were taken at 200X magnification. (B) Live chimeras from blastocysts injected with DR10-EPSCs. (C) The mCherry+ EPSCs also show extensive contribution to the 14.5dpc gonad, similar to 2i/LIF ESCs.

A.



B.

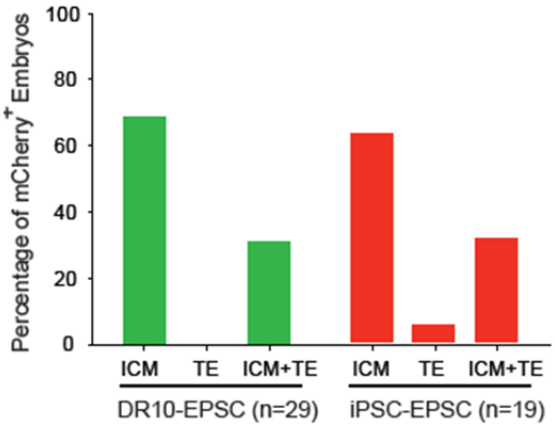


Figure 3.21. Newly derived EPSC lines can contribute to both the ICM and the TE in blastocysts.

(A) Contribution of EPSCs (mCherry+) derived from preimplantation embryos or iPSCs to both the ICM and TE in the blastocysts. (B) Percentage of mCherry+ embryos that have donor contribution to the ICM, the TE or both.

cell line that possessed expanded potential in its ability to contribute to both the embryonic and extra-embryonic lineages in the developing mouse embryo. This resulted in the formulation of a novel cell culture medium, EPSCM, which enabled the conversion of conventional mouse ESCs to expanded potential stem cells (EPSCs), the derivation of new lines with expanded potential from pre-implantation embryos and from single blastomeres of eight-cell stage embryos. Mouse iPSC lines could also be maintained in EPSCM and were similarly conferred with expanded differentiation potential.

EPSCM contained six inhibitors targeting the MAP kinase pathway, Src, and Wnt. The MAPK and SFK pathways have clearly established roles in pre-implantation development with genetic studies implicating these pathways in the segregation of the trophectoderm and inner cells mass (Abell et al., 2009; Li et al., 2011; Lu et al., 2008; Maekawa et al., 2007; Maekawa et al., 2005; Natale et al., 2004; Nishioka et al., 2009; Saba-El-Leil et al., 2003; Strumpf et al., 2005). The microinjection of mCherry-expressing AB2.2 mouse ESCs cultured in the presence or absence of MAPK (ERK, p38 and JNK) and Src inhibition into eight-cell stage embryos for blastocyst development confirmed the individual and additive role of inhibiting these pathways and conferring expanded potential. Correct inhibition of the downstream effectors of the MAPK and Src pathway was confirmed by Western blot where the phosphorylated proteins of p38, JNK, ERK and Src were all reduced relative to the KSR2iL control.

Modulation of Wnt signalling appears to be more complicated. EPSCM contains XAV939, which stabilizes AXIN1 by inhibiting its ubiquitination (Huang et al., 2009). EPSCs had considerably elevated levels of AXIN1 compared with the control which caused accumulation of phosphorylated β-catenin in both the cytoplasm and nuclei and decreased active β-catenin

in the nucleus. The TopFlash luciferase assay confirmed the reduced β-catenin-LEF/TCF activity. Functionally, Wnt signalling appears to be a balancing act in EPSCs. The addition of a WNT inhibitor, XAV939 confers expanded potential, its removal causes a significant loss of expanded potential. One may then consider that CHIR99021, a GSK3B inhibitor and stimulant of canonical WNT signalling therefore would not be required. However, the removal of CHIR99021 causes a modest reduction in expanded potential but more significantly its removal causes a deterioration in the robustness of the culture with substantially decreased AP positive colony numbers.

In pre-implantation development, Wnt and FGF signalling pathways are implicated in producing enough pluripotent cells in the embryo for successful implantation (Morris et al., 2012). This may explain in part the reduction in colony number following the removal of CHIR99021. In ground-state embryonic stem cells, part of the self-renewal network relies on the presence of canonical WNT signalling. In EPSCs, despite the presence of GSK3B inhibition, the stabilisation of AXIN1 results in the accumulation of phosphorylated β-catenin in both the cytoplasm and nuclei and decreased active β-catenin in the nucleus. Therefore, the cells cannot be using the canonical WNT signalling network. Although speculative to consider, perhaps the pre-implantation embryo is reliant on non-canonical WNT signalling. One study in bovine embryos has suggested that this may be the case however, further studies are required (Denicol et al., 2013).

Furthermore, important functional interactions of the Wnt and MAPK pathways with YAP, the key Hippo pathway downstream effector, have recently been uncovered (Kapoor et al., 2014; Shao et al., 2014). Nuclear translocation of YAP and interaction with TEAD4 is instrumental in initiating trophectoderm commitment and the first fate decision. Although provisional data on EPSCs did not find a difference in total or phosphorylated YAP levels in

EPSCs, a recent publication has shown that XAV939 targets YAP independent of YAP phosphorylation (Wang et al., 2015). As the authors highlighted, these findings are consistent with previous studies showing that AMOT family proteins suppressed YAP activity by translocating YAP from nucleus into cytoplasm through direct protein-protein interaction but independent of YAP phosphorylation status (Chan et al., 2011; Hirate et al., 2013; Hirate and Sasaki, 2014; Wang et al., 2011a; Zhao et al., 2011). Further experiments are required to precisely elucidate the role of WNT in EPSCs.

A recent challenge to our understanding of ground-state pluripotency was the controversial description of the presence of 2C-like cells and HEX positive subpopulations in standard ESC culture conditions (Macfarlan et al., 2012; Morgani et al., 2013). These subpopulations were shown to possess expanded potential being able to contribute to both embryonic and extra-embryonic lineages. However, these subpopulations are rare and only existed transiently. In light of these published findings, we considered, what is the origin of EPSCs? Are EPSCs a subpopulation of conventional ESC cultures? Single cell plating of mouse ESCs in EPSCM suggested that EPSCM was permissive for ESCs and did not select for any rare subpopulation of cells. However, EPSCs could not be derived from mouse epistem cells. Unlike ESCs, EpiSCs rapidly succumbed or differentiated in EPSCM, indicating that conversion or reprogramming by EPSCM is only permissible in ESCs. Intriguingly, derivation of EPSCs from embryos reveals the transient loss of OCT4 and CDX2 expression in the "filled" blastocysts. This transient "blank" state bears certain similarities with the 2C-like ESCs (Macfarlan et al., 2012), where OCT4 protein expression is absent. This phenomenon could be due to global translational repression by *Eif1a* as in 2C embryos (Hung et al., 2013). The fact that this "blank" state is only transient and OCT4 expression is reactivated in the cellular outgrowths indicates that eventual activation of the canonical pluripotency network is

necessary for stable *in vitro* adaptation and propagation of EPSCs. Future single-cell experiments are warranted to provide deeper characterization of these molecular events.

Despite the presence of expanded potential, EPSCs also show activation of the ground-state embryonic stem cell pluripotency network expressing similar levels of pluripotency genes and lineage-specific genes as standard ESCs. EPSCs preferentially use the *Oct4* distal enhancer, which is specifically active in naïve ESCs. Moreover, EPSCs were genetically stable and retained the ability to robustly differentiate into the three somatic germ layers *in vitro* and *in vivo* and could contribute to the germline. The signalling dependence of conventional ESCs and EPSCs was similar in terms of FGFR and TGFBR inhibition, to which both were insensitive (Ying et al., 2008). A difference between the two cell states was identified with regards to dependence on the Jak/Stat pathway. Mouse ESCs in EPSCM were dependent on the Jak/Stat pathway, as a Jak inhibitor substantially reduced the number of alkaline phosphatase positive colonies from these cells and adding LIF increased phosphorylation of STAT3 and up-regulated the expression of Jak/Stat downstream genes. By contrast, conventional ESCs cultured in 2i/LIF were insensitive to Jak inhibition. This finding may add further evidence for the maintenance of a distinct stem cell state as a recent publication has suggested that the Leukemia inhibitory factor (LIF) and STAT3 activation was necessary to promote the self-renewal and expansion of the rare totipotent HEX positive subpopulation in mouse ES cell culture (Morgani et al., 2013).

EPSCs demonstrated the capacity for trophoblast specification *in vitro* and consistent contribution to the embryonic and extra-embryonic lineages in blastocyst contribution assays. However, multiple *in vitro* differentiation attempts failed to produce a self-renewing EPSC derived trophoblast stem cell. Speculatively, this may be explained by a lack of knowledge of the cues that are necessary for trophoblast induction. Published data has demonstrated the role

of FGF4/FGFR2 interaction in mediating trophoblast stem cell renewal but FGF is not required for initial trophectoderm specification (Arman et al., 1998; Feldman et al., 1995). Further experiments are necessary to precisely define the signalling requirements necessary for efficient trophoblast induction.

In conclusion, I have described herein the development of a novel chemically defined medium which permitted the establishment of stable embryonic stem cell lines that possess expanded fate potential *in vitro*. In Chapter 4, I proceed to evaluate the capacity of these cells to contribute to all lineages of the developing organism *in vivo.*

CHAPTER 4

4 EPSCS DEMONSTRATE EXPANDED FATE POTENTIAL *IN VIVO*

4.1 Summary

In this chapter, I demonstrate that EPSCs can contribute to the embryo proper, placental trophoblasts and yolk sac endoderm in chimeras. To formally exclude the presence of unipotent subpopulations within EPSCM, I proceeded to generate chimeras from injection of a single EPSC. Generated single cell chimeras demonstrated extensive contribution to all lineages in the developing organism providing additional evidence that EPSCM permits maintenance of a homogenous stem cell population that possesses expanded fate potential *in vitro* and *in vivo.*

4.2 Introduction

Recent reports have described the existence of a rare subpopulation of cells within ground state culture conditions that possess totipotent-like properties (Macfarlan et al., 2012; Morgani et al., 2013). This claim was supported by the demonstration of inner cell mass (ICM) and trophectoderm (TE) contribution in blastocyst stage embryos. Additionally, these cells had expanded fate potential in late gestation chimeric embryos with contribution identified in the embryo proper, yolk sac and placenta.

Currently, there is inconsistency in the choice of assays that are used to claim extra-embryonic contribution. Published studies may define expanded potential based on apparent contribution on whole mount fluorescence microscopy, DNA genotyping of extra-embryonic membranes or immunostaining. Exceptional claims require exceptional data in support and multiple independent assays are often needed to confirm or refute such findings.

The presence of donor cells in the vicinity of the placenta or yolk sac does not equal expanded potential as both of these organs have cellular descendants from the embryonic and extra-embryonic lineages. An understanding of this cellular composition is a necessary prelude to designing any experiments assessing *in vivo* expanded fate potential.

4.2.1 Placenta

The placenta is located at the maternal-fetal interface and has many critical functions during pregnancy including gas, nutrient and waste exchange between the fetus and mother, production of hormones to support pregnancy and is involved in immunoprotection of the developing fetus. Descendants of the trophoblast lineage occupy the majority of the placenta

but there is a fundamental contribution from the embryonic mesoderm which gives rise to the placental vasculature and umbilical cord.

Placental development begins at the first cell fate specification which gives rise to the inner cell mass and trophectoderm (TE). Schematic representation of murine placental development is shown in **Figure 4.1A**. First, the trophectoderm surrounding the blastocyst cavity which is the mural TE differentiates almost immediately to form the primary giant cells. Giant cells are formed by endoreplication which involves multiple rounds of DNA replication but no intervening mitosis which results in cellular polyploidy. At the embryonic pole of the blastocyst, the polar trophectoderm is identified which will give rise to the extra-embryonic ectoderm (ExEc) and ectoplacental cone. Trophoblast stem cells can be derived from the polar trophectoderm and ExEc in FGF4 containing media as described previously (Tanaka et al., 1998).

As development progresses, the extra-embryonic ectoderm will form the chorionic epithelium which is an extra-embryonic membrane lined by mesothelial cells. Simultaneously, at the posterior end of the embryo, a focus of embryonic mesoderm will expand to form the allantois (Rossant and Cross, 2001). At E8.5 the chorion and allantois come in close proximity which is termed chorioallantoic fusion. This event is critical for a pregnancy to progress uneventfully to term. Through the allantois, the fetal vasculature migrates into the placenta and the blood vessels along with the surrounding trophoblast undergo extensive branching to form the labyrinth which will be the future site of exchange between the maternal and fetal vasculature. The labyrinth is supported by spongiotrophoblast, derived from the ectoplacental cone (Rossant and Cross, 2001). The labyrinthine trophoblast differentiates to produce two layers of synctiotrophoblast which lie in external contact with

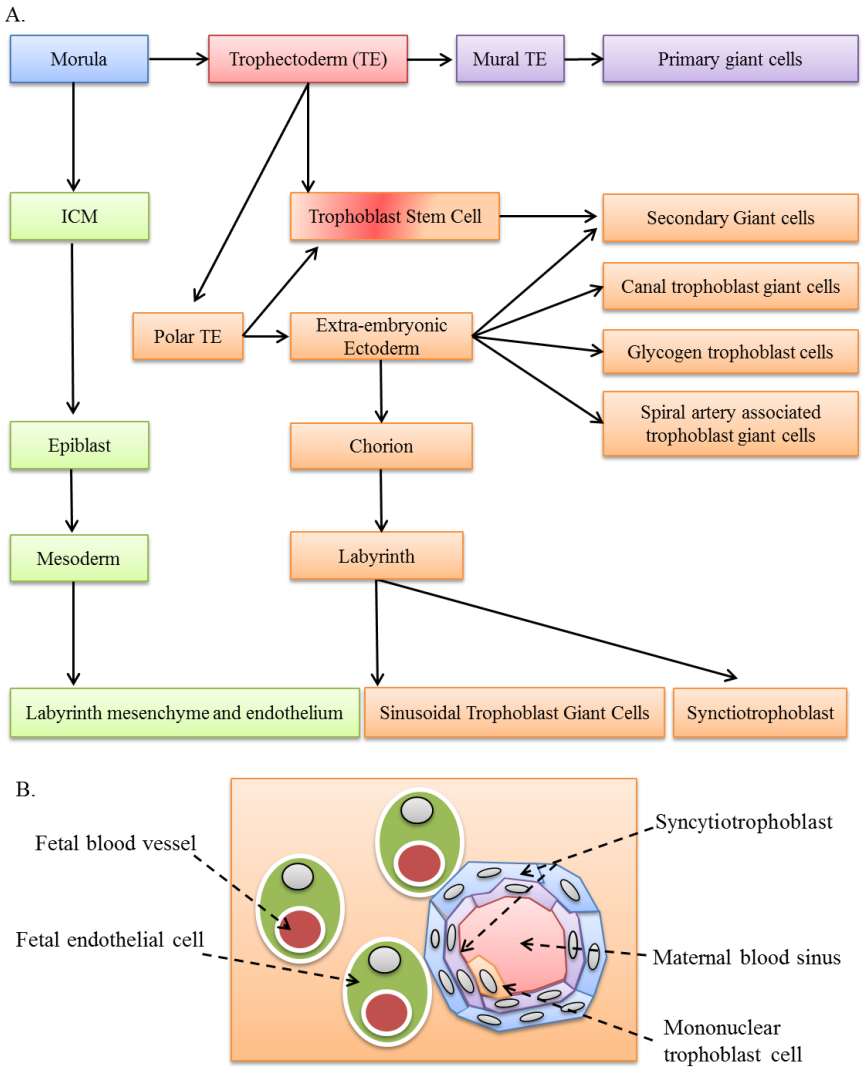


Figure 4.1. Mouse Placental Development

(A)Schematic showing the steps involved in the first cell fate specification which gives rise to the inner cell mass (ICM) and trophectoderm (TE). The cellular composition of the placenta is then illustrated showing contribution from both the extra-embryonic and embryonic lineages. (B) Graphic representation of a high resolution microscopic view of the placenta. Two layers of synctiotrophoblast derived from the labyrinthine trophoblast and a single inner layer of mononuclear trophoblast surround the maternal vascular sinusoids. Surrounding the synctiotrophoblast lined maternal sinusoids are the fetal blood vessels. This arrangement is the maternal-fetal interface which is the site of gas, nutrient and waste exchange between the fetus and mother during pregnancy.

the surrounding fetal blood vessels and an inner layer of mononuclear trophoblast which will surround the maternal vascular sinusoids **(Figure 4.1B)**.

4.2.2 Yolk Sac

Yolk sac formation begins at the second fate specification where the ICM segregates into the epiblast and hypoblast, the latter of which will give rise to the primitive endoderm and extra-embryonic portions of the developing yolk sac. By E5.0 the primitive endoderm will have segregated into the visceral and parietal endodermal portions of the yolk sac (Enders et al., 1978). Parietal endoderm grows on the inner surface of the primary trophoblast giant cell layer and secretes large amounts of basement membrane proteins which in combination with the giant cell layer form Reichert's membrane. The primitive endoderm in contact with the epiblast becomes the visceral endoderm which plays an important role in anterior patterning of the developing fetus (Beddington and Robertson, 1999). By E8.0 the epiblast derived extra-embryonic mesoderm in combination with the visceral endoderm will form the definitive yolk sac. The visceral endoderm is also thought to secrete proteins including vascular endothelial growth factor and Indian Hedgehog which directs haematopoiesis in the overlying extra-embryonic mesoderm (Damert et al., 2002; Dyer et al., 2001).

4.3 Aims of this Chapter

The primary experimental aim of this chapter was to evaluate for *in vivo* embryonic and extra-embryonic contribution from mouse EPSCs. If successful, the secondary aim was to demonstrate expanded fate potential *in vivo* from injection of a single EPSC which would refute the presence of unipotent populations within EPSCM and provide additional evidence

to support a claim that EPSCM permits establishment and maintenance of a homogenous population of cells that possess expanded fate potential.

4.4 Results

4.4.1 EPSCs contribute to the embryo proper and placental trophoblasts in chimeras.

I proceeded to demonstrate the *in vivo* differentiation potency of EPSCs by injecting mCherry positive EPSCs (DR10, DR25, AB2.2) and the control cells (AB2.2 in 2i/LIF or in M15) into pre-implantation embryos. EPSCs (DR10 and DR25) were microinjected into morula stage embryos to generate chimeras which would then be assessed for their contribution to the embryonic and extra-embryonic lineage, particularly the placental cells and yolk sac. In the 14.5dpc chimera embryos, whole-mount fluorescence examination revealed extensive mCherry positive donor cell contribution to the embryo proper and the extra-embryonic tissues (20/32 (63%) embryos for DR10-EPSCs, and 14/18 (78%) for DR25-EPSCs) **(Figure 4.2)**. Flow cytometry analysis of dissociated cells from the embryos and the placentas confirmed mCherry positive cells in both lineages **(Figure 4.3).**

To independently confirm this result, I repeated the experiment at another transgenic facility (Cancer Research UK-Cambridge Institute), where 6-8 cells were injected into eight-cell host embryos for chimera production. mCherry positive chimeras (14.5dpc) were found in 13/39 (33%) embryos for DR10-EPSCs, 9/23 (39%) for DR25-EPSCs, 21/38 (55%) for AB2.2-EPSCs and 16/28 (57%) for AB2.2-2i/LIF ESCs **(Figure 4.4)**. DNA genotyping again confirmed the contribution of EPSCs in the embryo proper, in the yolk sac and in a number of placentas **(Figure 4.5A)**. Flow cytometry analysis of dissociated single placenta, brain and liver cells further confirmed the presence of descendants of mCherry positive EPSCs **(Figure**

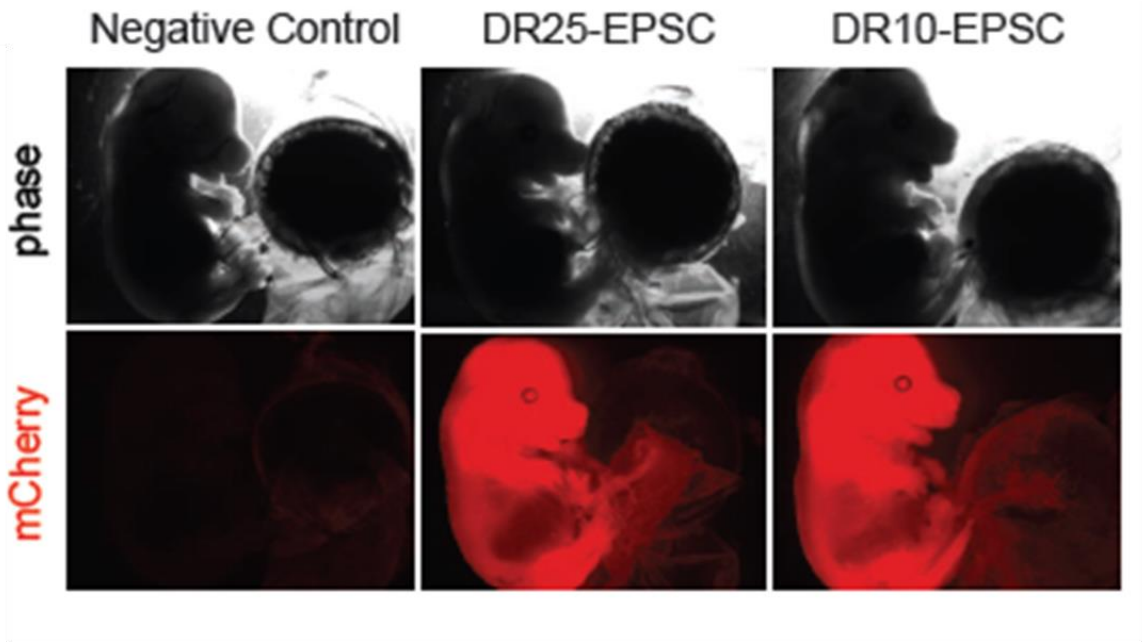


Figure 4.2. Qualitative assessment of EPSC contribution in 14.5dpc chimeras from morula injection.

Whole-mount fluorescent imaging of 14.5dpc chimeras from morula embryos injected with EPSCs. Extensive contribution appeared to be present in the embryonic and extra-embryonic tissues in DR25-EPSC and DR10-EPSC chimeric embryos. Negative control: wild-type embryo. Imaging was performed using a Leica M205FA Automated Fluorescence Stereo Microscope (100 ms for phase and 5 seconds for mCherry).

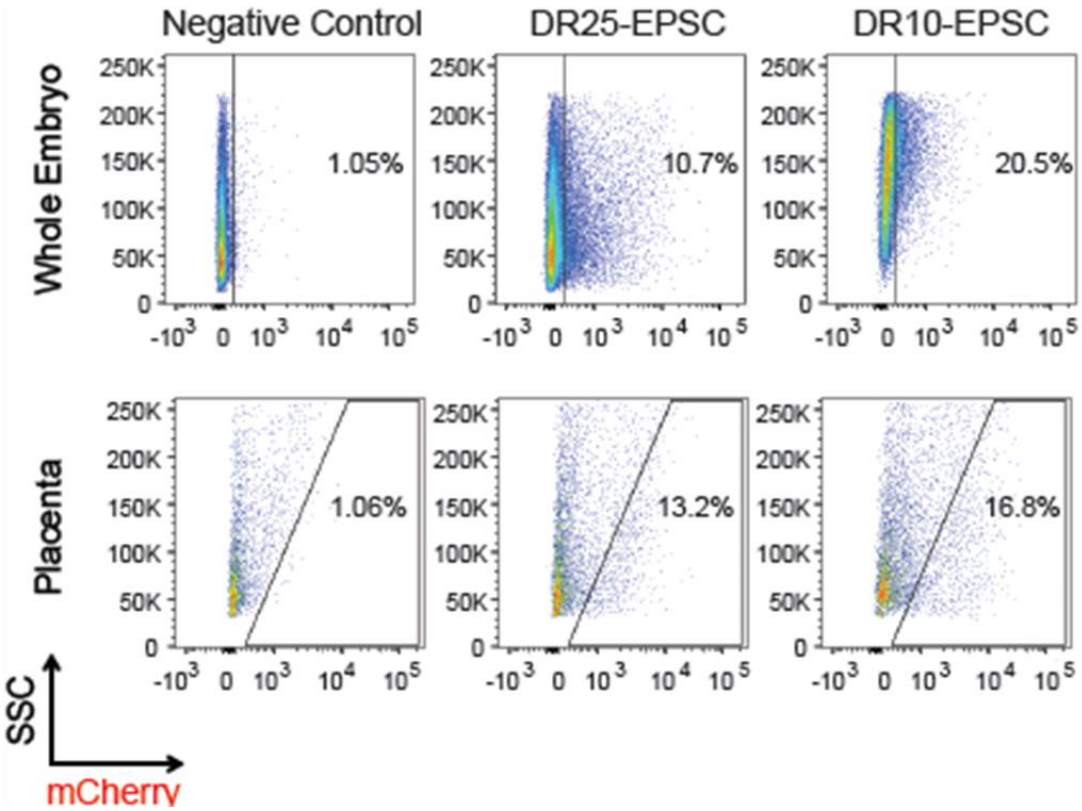


Figure 4.3. Quantitative assessment of EPSC contribution in 14.5dpc chimeras from morula injection.

Flow cytometry analysis of dissociated embryonic and placental cells from two mCherry positive chimeras derived from DR10 and DR25 EPSCs, respectively. Negative control: wild-type embryo. Thanks to Wei Wang and Yong Yu for help with FACS.

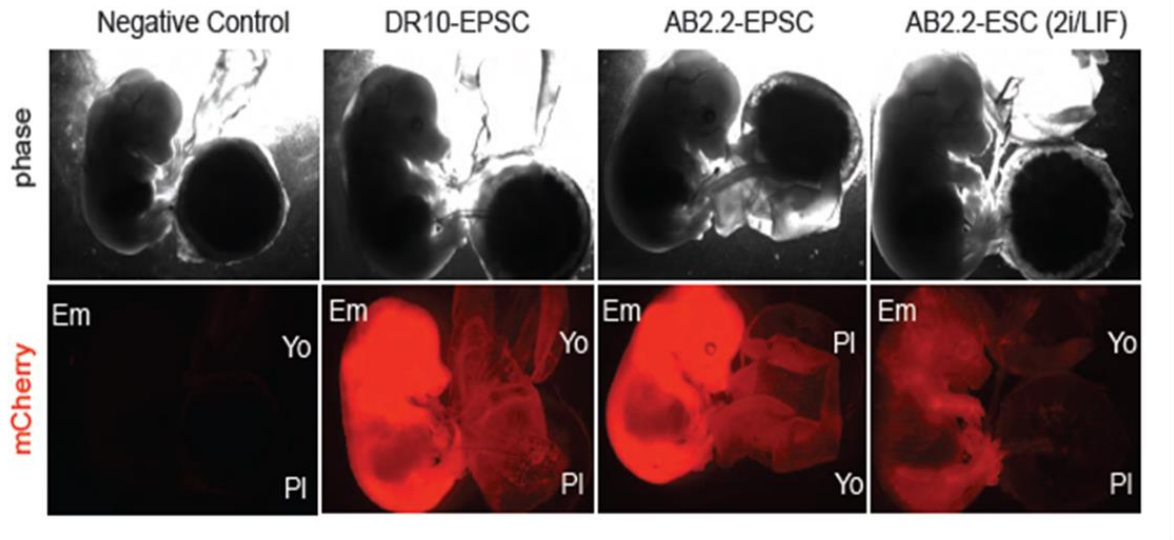
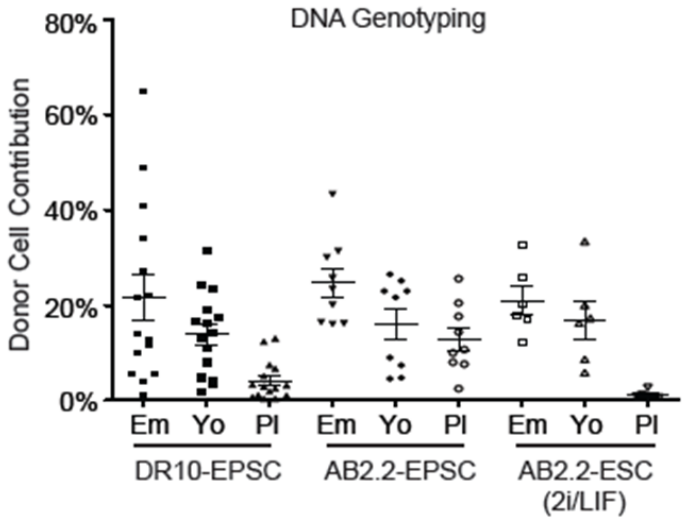


Figure 4.4. Contribution of EPSCs and ESCs in chimera embryos – Qualitative.

Whole-mount fluorescent imaging of representative 14.5dpc chimera embryos from 8 cell stage embryos injected with EPSCs or 2i/LIF ESCs. Negative control: wild-type embryo. DR10-EPSCs at passage 15 or AB2.2-EPSCs at passage five in EPSCM were used for injection.. Qualitatively, extensive contribution was identified in the embryo proper (Em), yolk sac (Yo) and placenta (Pl) of DR10-EPSC and AB2.2-EPSC generated chimeras. In contrast, although contribution was extensive in the embryo proper in AB2.2-ESC from 2iLIF, no definite contribution could be identified in the yolk sac or placenta. Imaging was performed using a Leica M205FA Automated Fluorescence Stereo Microscope with a fixed exposure parameter applied to all conditions, 100 ms for phase and 5 seconds for mCherry.

A.



B.

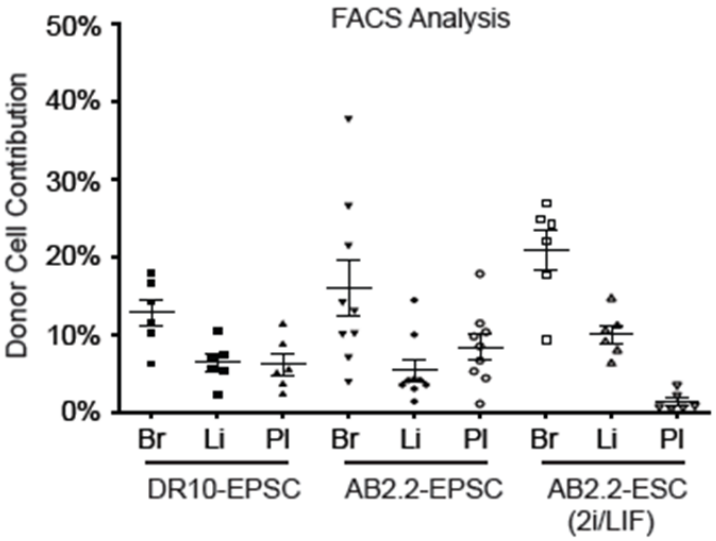


Figure 4.5 Contribution of EPSCs and ESCs in chimera embryos -Quantitative

(A) Genotyping of donor cells in the embryo proper (Em), the yolk sac (Yo) and the placenta (Pl). The DR10-EPSCs have a Cre cassette at the Rosa26 locus. A Cre-specific TaqMan probe was therefore used to determine the contribution of these cells. AB2.2-EPSCs or ESCs have a non-functional Neo cassette at the Hprt locus. A Neo-specific TaqMan probe was used to determine the contribution of these cells. Genomic DNA from parental DR10-EPSCs or AB2.2 ESCs was used as the positive control in qPCR, and DNA from a wild-type embryo was used as the negative control. The internal control is a Tert TaqMan probe. (B) Quantification of mCherry positive donor cell contribution in chimeras detected by flow cytometry. Br: fetal brain; Li: fetal liver; Pl: placenta. Thanks to Wei Wang and Yong Yu for help with genotyping and FACS analysis.

4.5B). By contrast, both DNA genotyping and FACS failed to detect obvious descendants of AB2.2-2i/LIF ESCs in the placenta **(Fig. 4.5A and 4.5B).**

The placenta contains derivatives of both the extra-embryonic and embryonic lineages therefore; it is essential to discriminate to which cell types the donor cells are contributing. Two approaches were taken to answer this question. First, the mCherry positive placental cells were sorted for gene expression analysis **(Figure 4.6A and 4.6B)**. The genes chosen for analysis represented multiple cellular subsets within the mouse placenta including the trophoblast stem cell niche (*Cdx2, Eomes, Elf5, Fgfr*), extraembryonic ectoderm (*Furin* and *Spc4*), ectoplacental cone / spongiotrophoblast (*Mash2 / Ascl2*), synctiotrophoblast (*Gcm1*) and trophoblast giant cells (*Hand1 and Pl1*). Excitingly, the EPSC donor cells expressed comparable levels of trophoblast genes, such as *Ascl2* (*Mash2*), *Gcm1, Pl-1*, and *Hand1*, as the mCherry negative placental cells (Hemberger et al., 2010) **(Figure 4.6B)**. By contrast, very few mCherry positive placental cells were found in the chimeras of AB2.2-2i/LIF ESCs, and we were unable to recover enough mCherry positive placental cells for qRT-PCR **(Figure 4.6A)**.

Secondly, to further establish the identity of the mCherry positive placenta cells from EPSC chimeras, these cells were stained with DAPI for DNA content analysis. By flow cytometry, trophoblasts could be detected as 2N, 4N, 8N and so on due to endoreplication and cell fusion during placenta development (El-Hashash et al., 2010). mCherry positive 8N cells were unambiguously detected in the EPSC chimera placenta cells, similar to the flow pattern in mCherry negative placental cells **(Figure 4.7)**. At the cellular level, immunostaining of placental sections identified mCherry positive multinucleated trophoblasts in EPSC chimeras, but not in the chimeras of AB2.2-2i/LIF ESCs **(Figure 4.8A and Figure 4.8B)**.

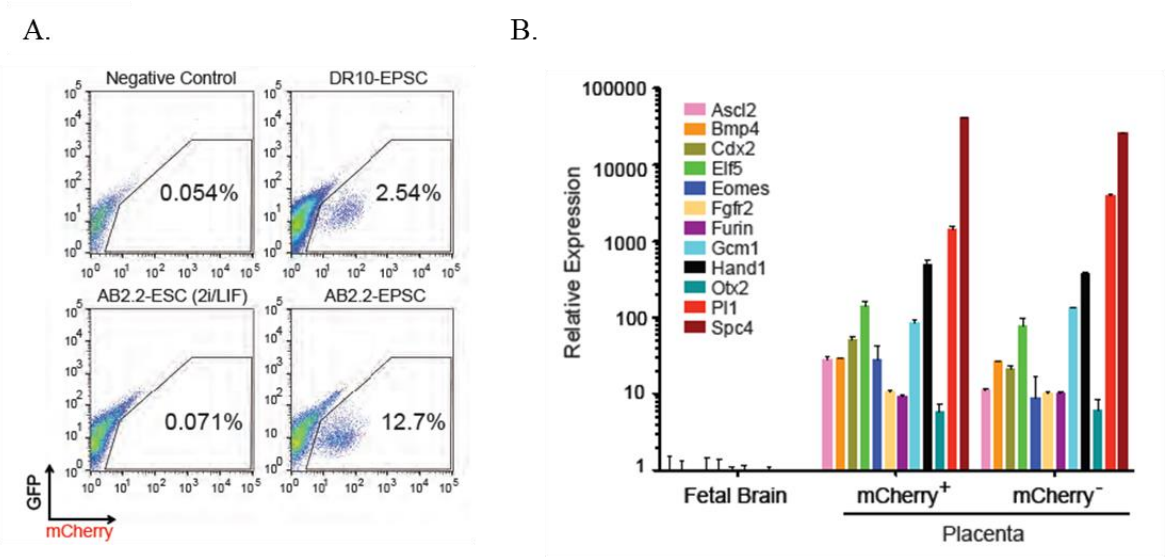


Figure 4.6. Sorted mCherry positive EPSC derived placental cells express key trophoblast genes.

(A) Sorting of mCherry positive placenta cells from 14.5dpc chimeras. The contribution to the placenta from 2iLIF ESCs was too low for sorting. To minimize residual background cell sorting, a GFP channel was also used to exclude autofluorescence. (B) Expression of trophoblast genes in sorted mCherry positive and mCherry negative placenta cells from an EPSC chimera. Expression was normalized to fetal brain expression. Data are the mean ± s.d.. Thanks to Wei Wang and Yong Yu for assistance with cell sorting and expression analysis.

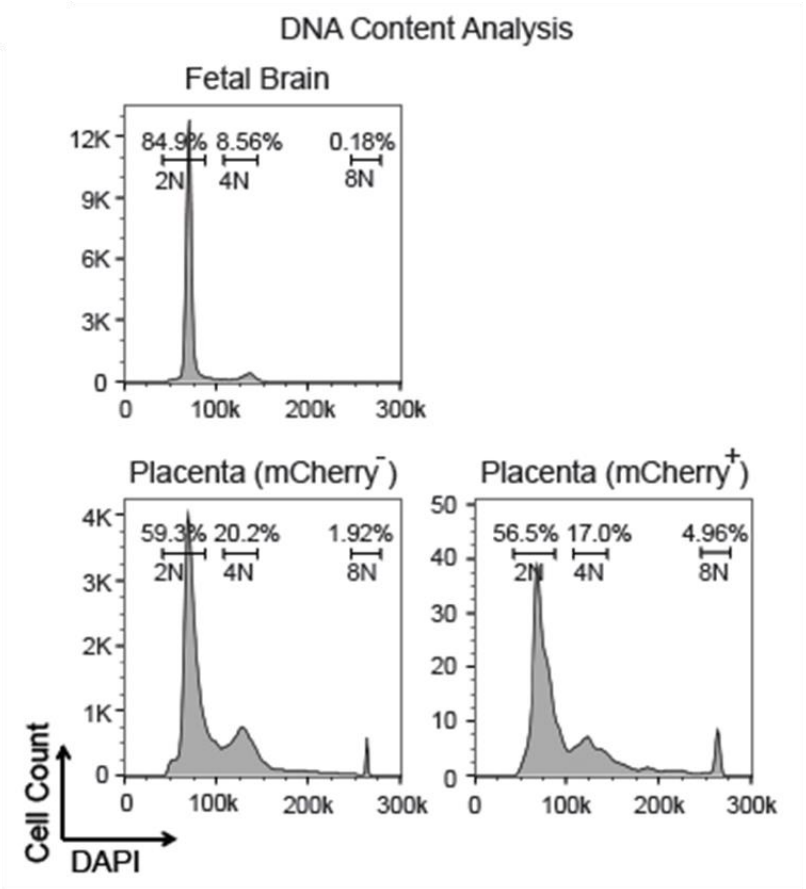


Figure 4.7. Detection of mCherry positive EPSC derived polyploid placental cells.

A distinct population of 8N cells was present in both mCherry positive and mCherry negative placental cells from an AB2.2-EPSC chimera. Fetal brain cells were used as the control. Similar to Figure 4.6A, no mCherry positive 8N cells could be detected in the placenta from 2iLIF ESCs due to the low number of mCherry positive cells. Thanks to Wei Wang and Yong Yu for technical assistance.

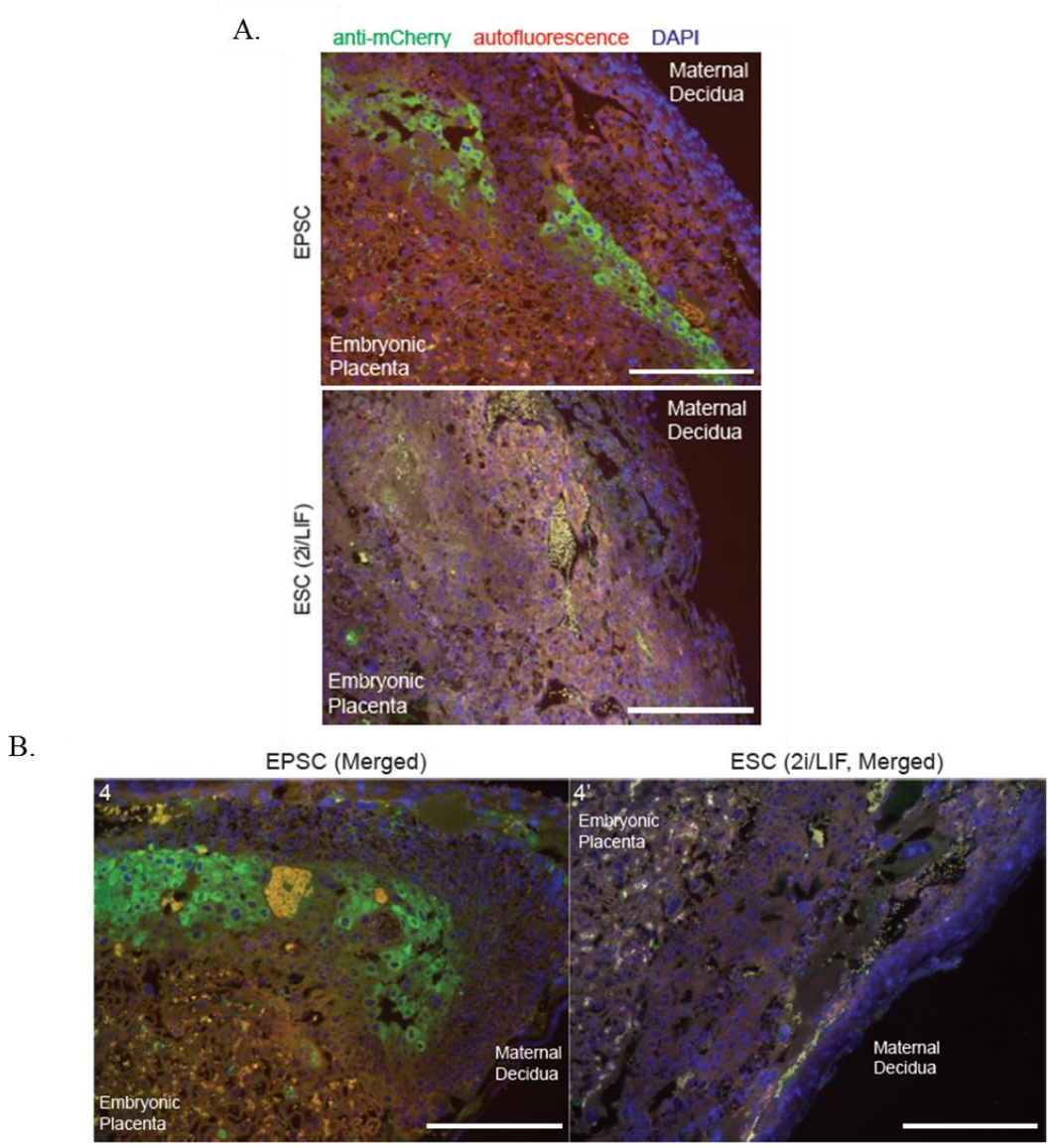


Figure 4.8. Immunofluorescent staining of EPSC derivatives in the placenta.

(A & B) Immunofluorescent staining of EPSC derivatives in the placenta using a FITC-conjugated antibody against mCherry (scale bar = 100 μm). Shown are images of sections of the mouse placenta from a 14.5dpc EPSC or ESC chimera. The same sections were also imaged in the red channel to detect background autofluorescence. DAPI staining was used to detect the nucleus. The original mCherry fluorescence signal was quenched due to fixation. Cells that appeared positive in both the red and the green channels (yellow in the merged panels) are background. The true mCherry positive trophoblast cells with low levels of background are only found in EPSC chimeras (green in the merged panel). Thanks to Jian Yang and Lia Campos for help with histology.

4.4.2 EPSCs contribute to the yolk sac in chimeras

I next sought to examine EPSC contribution to the second extra-embryonic lineage, the primitive endoderm in the fetal yolk sac. Similar to the placenta, the yolk sac contains cell types which are derived from the embryonic (extra-embryonic mesoderm) and extra-embryonic lineage (extra-embryonic endoderm). Demonstration of contribution to both lineages would be required in order to support a claim of expanded fate potential in EPSCs.

In the 14.5dpc yolk sac of chimeras, cells from EPSCs could clearly be found in both the extra-embryonic endoderm layer and the extra-embryonic mesoderm cells (endothelial cells and mesothelial cells) **(Figure 4.9)**, whereas ESCs, as expected, only contributed to extra-embryonic mesoderm cells (Nagy., 1990). Therefore, EPSCs have developmental potential to the embryo proper and to all extra-embryonic lineages.

4.4.3 A single EPSC demonstrates expanded fate potential in chimeras

It was still possible that bulk injection of EPSCs produced contribution in the embryonic and extra-embryonic lineages due to the presence of unipotent subpopulations within the culture. For example, the culture may have possessed TS, XEN and ES-like cells which contribute to the placenta, yolk sac and embryo proper respectively in chimeras. This was felt to be unlikely, but in order to formally exclude this possibility and to conclusively demonstrate the potency of EPSCs, single-cell injection in eight-cell stage embryos was performed. To this end, each eight-cell stage host embryo was injected with a single EPSC (DR10, DR25 and AB2.2) or with a control ESC (AB2.2-2i/LIF) for chimera production. Embryos were harvested at 14.5dpc. Under the fluorescence microscope, mCherry positive cell contribution was found in 7 out of 85 (8%) embryos for DR10 cells, 8 of 77 (10.4%) for DR25 cells, 7 of 51 (13.7%) for AB2.2-EPSCs, and 10 of 82 (12.2%) for AB2.2-2i/LIF ESCs **(Figure. 4.10)**.

DNA genotyping and flow cytometry analysis of mCherry positive chimeras showed EPSC descendants in the embryo proper, the yolk sac and the placenta **(Figure 4.11A and 4.11B)**. FACS purified mCherry positive placental cells from EPSC chimeras **(Figure 4.12)** were analysed for gene expression. Again, trophoblast lineage markers were clearly detected in these cells by qRT-PCR, comparable to the host mCherry negative placenta cells **(Figure 4.12)**. Furthermore, flow cytometry analysis of the mCherry positive placental cells detected 8N trophoblasts, similar to the host mCherry negative placental cells **(Figure 4.13)**. In a couple of embryos, mCherry positive EPSCs contributed up to 40% of the placental cells by DNA genotyping. The embryos were morphologically normal, suggesting that placental cells from EPSCs were likely functional for normal embryo development, which warrants future investigation. In contrast, AB2.2-ESCs cultured in 2i/LIF contributed poorly to the embryo proper and little to the placenta in chimeras, as revealed by genotyping. FACS analysis failed to detect enough mCherry positive placental cells for sorting or for 8N trophoblast analysis in these chimeras. Single AB2.2 cells cultured in M15 medium were also injected into 8C host embryos but no chimeras were recovered (0/18). This result reflects the fact that in practice, efficient chimera production usually requires injecting 6-8 standard ESCs into an eight-cell stage host embryo.

4.5 Discussion

In this chapter I have demonstrated that consistent with the *in vitro* observations in Chapter 3, EPSCM permits maintenance *in vitro* of a stem cell state that can contribute to the embryonic and extra-embryonic lineage *in vivo*.

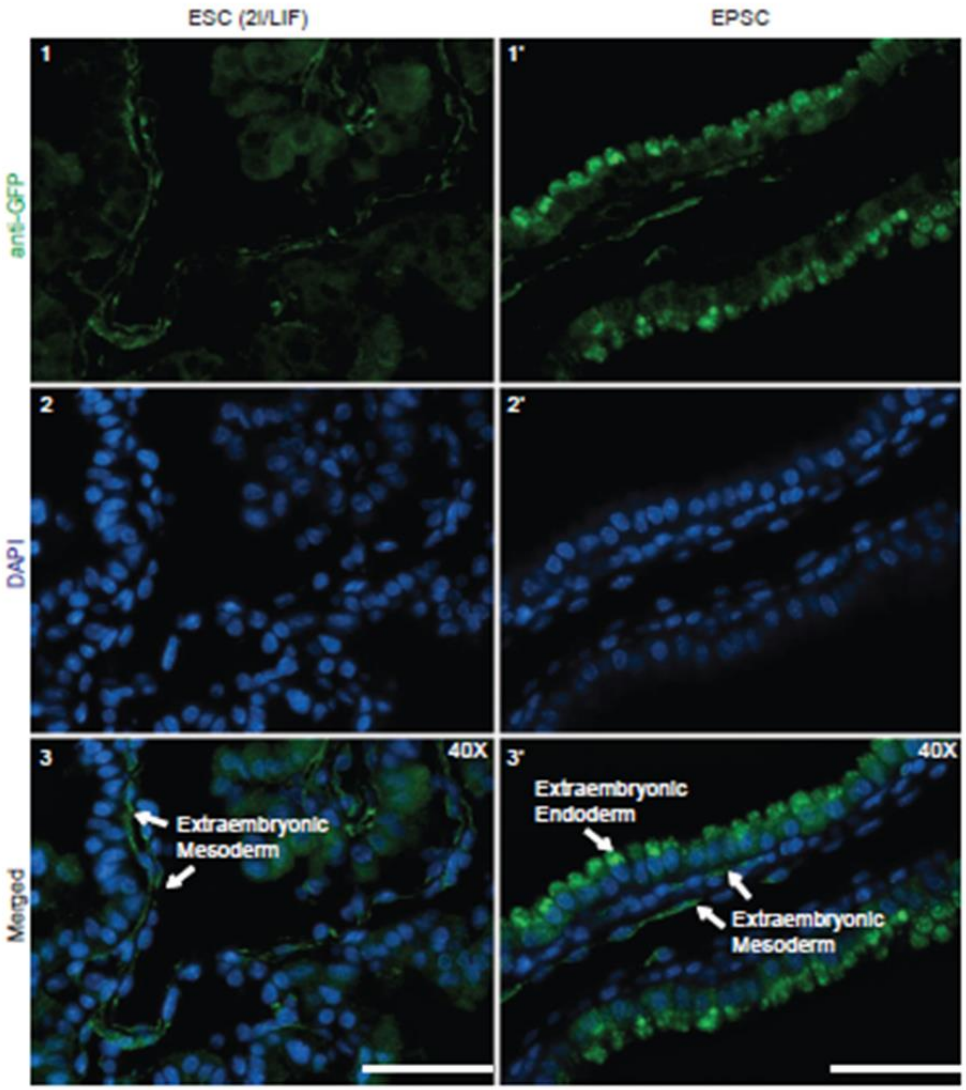


Figure 4.9. Yolk Sac Immunostaining

Yolk sac sections of 14.5dpc chimeras of either ESCs (Panels 1-3) or EPSCs (Panels 1'-3'). Due to high autofluorescence of FITC-conjugated antibody for mCherry in the yolk sac, I analysed yolk sac sections of chimeras from injecting Rosa26-GFP reporter EPSCs or ESCs. The original GFP fluorescence signal was quenched due to fixation. A CF660C conjugated antibody against GFP was used to detect donor cells. DAPI staining for the nucleus. GFP+ cells were found in extraembryonic mesoderm cells (endothelial and mesothelial) in the yolk sac of chimeras of both EPSCs and ESCs. In contrast, GFP+ cells were only detected in the extraembryonic endoderm layer (extraembryonic-visceral-endoderm-derived) in the yolk sac of EPSC chimeras. Scale bar=50 μm..

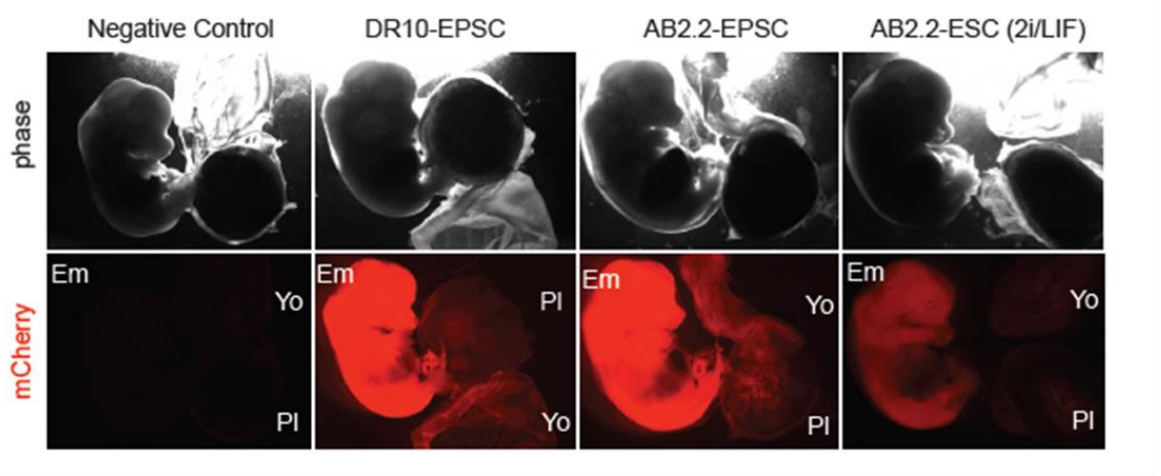


Figure 4.10. Contribution of a single EPSC to the 14.5dpc chimera.

Whole-mount fluorescence imaging of representative 14.5dpc chimeras from 8 cell stage embryos injected with a single EPSC or ESC. Negative control: wild-type embryo. Similar to the bulk injection experiments, extensive contribution was identified in the embryo proper (Em), yolk sac (Yo) and placenta (Pl) of DR10-EPSC and AB2.2-EPSC generated chimeras. Contribution in AB2.2 2iLIF cultured ESCs appeared to be entirely embryonic. Imaging was performed using a Leica M205FA Automated Fluorescence Stereo Microscope with a fixed exposure parameter applied to all conditions, 100 ms for phase and 5 seconds for mCherry.

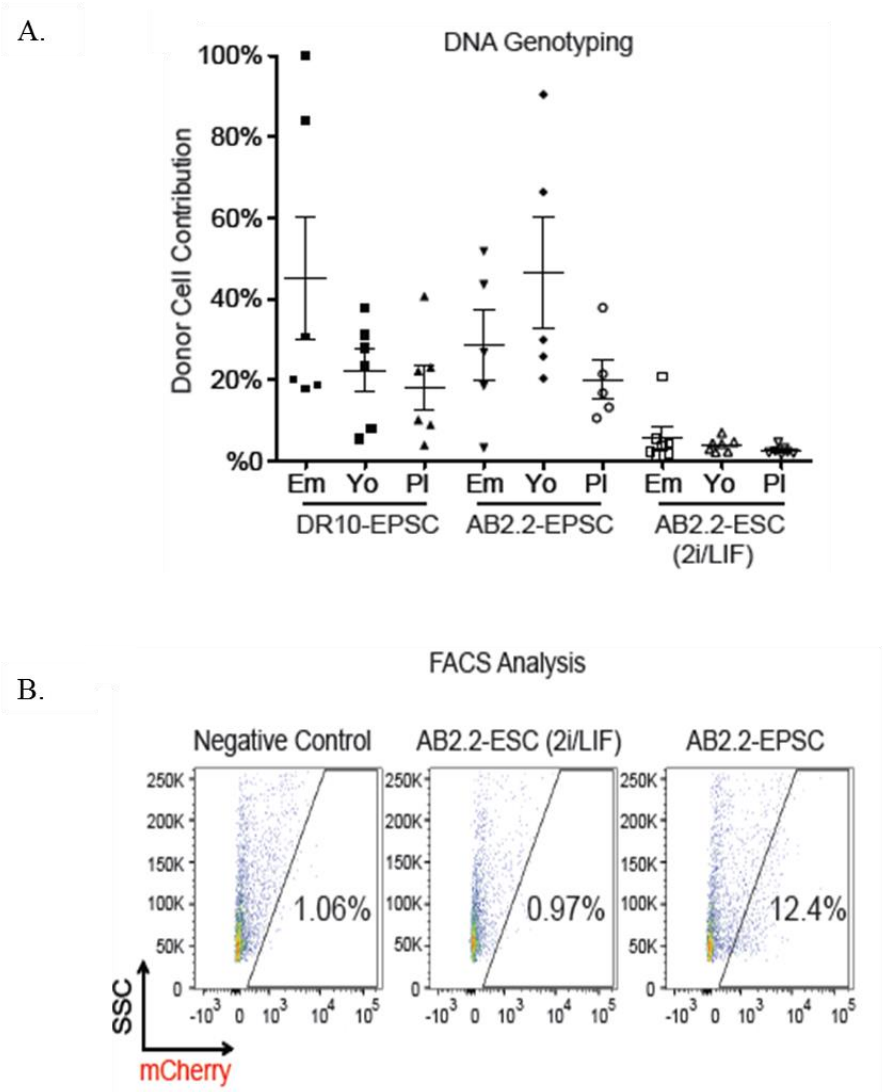


Figure 4.11. DNA Genotyping and FACS of single cell contribution to the 14.5dpc chimera.

(A)Genotyping of EPSC or ESC contribution in chimeras. The same genotyping strategy was used as in Figure 4.5A. (B) FACS analysis of mCherry positive placental cells. Thanks to Wei Wang and Yong Yu for cell sorting and expression analysis.

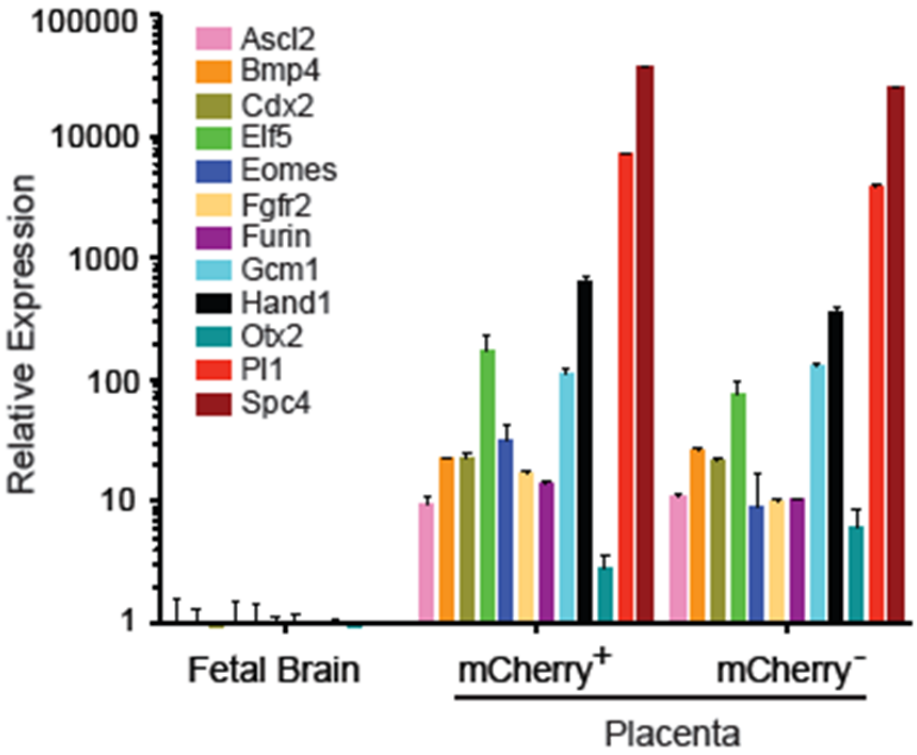


Figure 4.12. EPSC cells in the 14.5dpc placenta express trophoblast specific genes.

Expression of trophoblast genes in sorted mCherry positive and mCherry negative placenta cells from an EPSC chimera generated from a single EPSC. Expression was normalized to fetal brain expression. Data are the mean ± s.d. Cell sorting was performed as in Figure 4.6A.

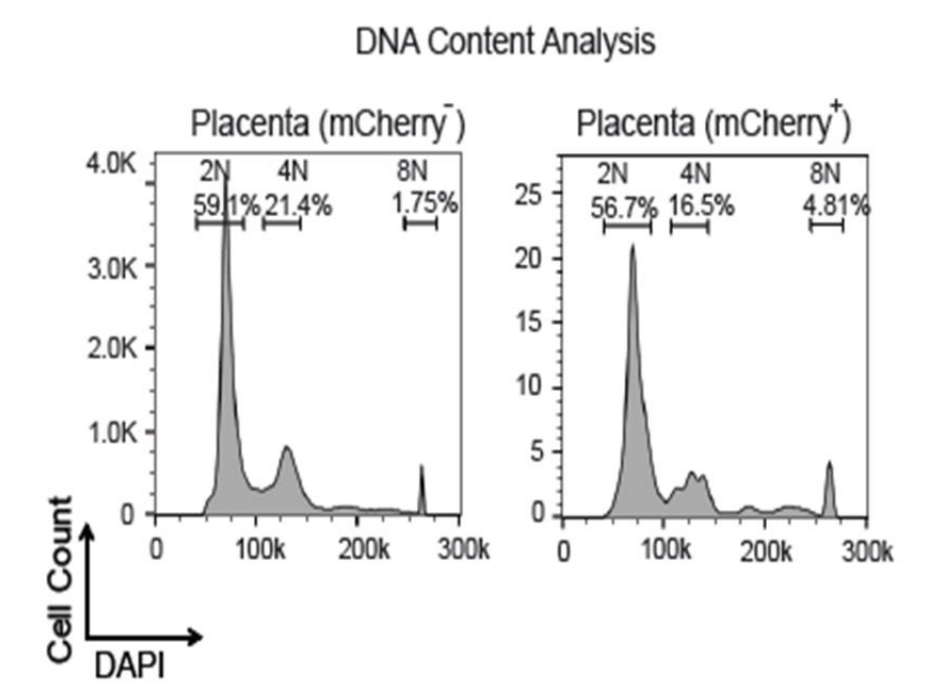


Figure 4.13. A single EPSC contributes to trophoblast giant cells in the mature placenta. Detection of polyploid placenta cells. A distinct population of 8N cells was found in both mCherry positive and mCherry negative placenta cells in a 14.5dpc EPSC chimera.

Multiple independent assays were performed in an attempt to prove bona fide contribution in the placenta and yolk sac. In line with previous publications I began by presenting whole mount fluorescent microscopy data showing EPSC contribution in the embryo proper and extra-embryonic tissues with additional experimental support from tissue genotyping and FACS sorting of the respective tissues. However, as previously discussed, these methods can by themselves lead to erroneous conclusions due to normal embryonic and extra-embryonic contribution in the placenta and yolk sac.

In the placenta, many terminally differentiated trophoblast cells undergo endoreplication and produce giant cells. This cellular descendant has no contribution from the epiblast being entirely extra-embryonic in origin. FACS sorting of the wild type placenta and mCherry negative placental cells revealed a subset of cells that were polypoid (>8N) consistent with the presence of trophoblast giant cells or syncytiotrophoblast. The mCherry positive fraction in chimeric placenta generated from EPSCs was found to contain polypoid cells indicating the presence of giant cells / syncytiotrophoblast and therefore definitive extra-embryonic contribution.

In order to gain a surrogate pan-placenta indication of cell type contribution in the placenta, the mCherry positive and negative fractions were sorted and gene expression analysis was performed. The genes chosen for analysis represented multiple cellular subsets within the mouse placenta. Both cellular fractions expressed comparable levels of these genes indirectly suggesting unbiased trophoblast EPSC differentiation. Further support for the above claim was provided by assessing the anatomical localisation of contribution by immunohistochemistry. The labyrinthine trophoblast was clearly identified and donor cells were seen to contribute to the layers of syncytiotrophoblast lining the maternal vascular sinusoids. As expected, 2i/LIF ESC did not contribute to trophoblast.

Similarly, in the yolk sac, EPSC contribution was identified at the cellular level in both the extra-embryonic endoderm which is derived from the extra-embryonic primitive endoderm and the extra-embryonic mesoderm, an epiblast derivative. In stark contrast, 2i/LIF ESCs, as expected demonstrated contribution to the extra-embryonic mesoderm only.

Although the above experiments provide multiple independent sources of evidence to support extra-embryonic contribution in the placenta and yolk sac, additional experiments including single cell transcriptional profiling and co-staining of donor cells in the placenta and yolk sac with antibodies against tissue specific antigens are needed to further support this claim.

An additional unresolved question from the placental contribution experiments is whether EPSC can produce functional trophoblast. The experiments described above show integration and expression of key placental related genes but could these donor cells support unaided, a pregnancy to term? This is a difficult question to answer as in chimeric assays; the donor cell contribution is never 100%. In an attempt to solve this problem, future experiments are being designed to evaluate the capacity of EPSCs to rescue placentation in parthenogenetic mouse embryos. Parthenogenetic embryos die by mid gestation due to placental failure. This is thought to be secondary to loss of the normal imprinting pattern of the paternal genome which is required for the normal development of the extra-embryonic membranes and trophoblast (Surani et al., 1984). A second perhaps more straight forward future experiment to test whether EPSCs can produce functional trophoblast in SOCS3 deficient mice. Suppressor of cytokine signalling 3 (SOCS3) is a negative regulator of LIF mediated trophoblast differentiation in the embryo. The SOCS3 mutation is embryonically lethal at E12.5 due to disrupted placentation (Roberts et al., 2001; Takahashi et al., 2006). This phenotype can be rescued by tetraploid complementation assays or more intriguingly by the injection of mouse

trophoblast stem cells (Takahashi et al., 2003; Takahashi et al., 2006). Therefore, based on the experimental data presented thus far, we hypothesise that EPSCs should be able to rescue placentation in SOCS3 deficient embryos.

Claims of totipotency require demonstration that a single cell can form an entire organism. Plenipotency on the other hand requires that an individual cell can contribute to all cell types in the developing organism (Condic, 2014). Following injection of a single EPSC to the eight-cell stage embryo, extensive donor cell contribution was identified in both the embryonic and extra-embryonic lineages using the assays described above. Furthermore, when EPSC contribution was observed, it was always in both lineages providing strong evidence that EPSCM maintains a homogenous population of plenipotent stem cells. Following injection of a single EPSC or group of EPSCs to an empty zona pellucida or two-cell stage embryo, embryonic development did not progress (data not included) indicating that EPSCs are not by the strictest definition totipotent. EPSCs contribute extensively to all lineages in the developing organism when injected at the eight-cell stage. This may reflect the fact that the cells require a critical number of carrier blastomeres in order to allow normal development to proceed. An alternative explanation is that EPSCs may have biological similarity to the eight-cell stage embryo. As has been shown in recent studies, synchronization of the donor cell and host embryo temporal developmental timelines is critical for embryonic contribution (Huang et al., 2012; Mascetti and Pedersen, 2016).

In conclusion, EPSCs possess totipotent-like characteristics *in vitro* and *in vivo* but from the data presented thus far, there was no definite mechanism to explain their expanded fate potential. In Chapter 5, I perform transcriptomic and epigenetic profiling of EPSCs to evaluate for the presence of a pre-implantation signature which may provide an explanation for the observed pheno

CHAPTER 5

5 TRANSCRIPTOMIC AND EPIGENETIC PROFILING OF EPSCS

5.1 Summary

In this chapter, I demonstrate that the transcriptome of mouse EPSCs is enriched with an early pre-implantation blastomere signature, distinct from other rare published totipotent-like cells and epigenetically, EPSCs possess bivalent domains at multiple loci encoding key developmental regulators. Moreover, I show that the transcriptome of EPSCs shares a pluripotency foundation and transcriptomic homogeneity comparable to ground-state ESCs. I then discuss how these findings may provide a potential mechanistic insight into the expanded fate potential characteristic of this cell state.

5.2 Introduction

Ground-state embryonic stem cells are pluripotent, contributing to all lineages in the embryo proper and importantly, have an *in vivo* counterpart in the nascent epiblast (Nichols and Smith, 2009). Contribution experiments in the preceding chapters have provided evidence to support the claim that EPSCs represent a distinct self-renewing state with expanded fate potential to all lineages in the developing organism. Therefore, I questioned whether EPSCs may have biological similarity with a specific stage of pre-implantation development. However, initial *in vitro* characterization in Chapter 3 indicated that EPSCs possessed a core pluripotency signature, aspects of which were comparable to ground-state ESCs. Therefore, transcriptomic profiling, DNA methylation content and histone methylation analysis were deemed essential to discern what self-renewal network was active in EPSCs, whether this network had features in common with naive pluripotency and importantly, if EPSCs possessed a transcriptional and / or epigenetic signature suggesting a pre-implantation origin.

5.2.1 The transcriptional and epigenetics characteristics of pluripotency

Mouse ESCs are pluripotent being ability to contribute to all somatic lineages (ectoderm, mesoderm and endoderm) in the developing organism. Mouse ESCs maintained in serum albeit pluripotent are considered to represent a 'metastable' state with heterogenous expression of pluripotency related factors and differentiation related genes, the latter of which is referred to as lineage priming (Marks et al., 2012). Epigenetically, serum cultured ESCs possess bivalent domains at the promoters of key developmental loci. This co-localisation of H3K4me3 and H3K27me3 is thought to mark genes that are poised for activation or repression in response to developmental cues (Azuara et al., 2006). The combination of a MEK1 inhibitor and GSK3B inhibition permits establishment and maintenance *in vitro* of a substantially homogenous mouse ESC population which is thought to represent the ground

state of self-renewal (Ying et al., 2008). In contrast to serum cultured ESCs, 2i cultured ESCs possess homogenous expression of key pluripotency regulators and reduced lineage priming with minimal expression of lineage specifiers. Epigenetically, there is reduced prevalence of the repressive H3K27me3 modification at promoters and fewer bivalent domains are detected in naive ES cells (Marks et al., 2012). Naive ESC possess global DNA hypomethylation which is thought to be mediated by high expression of *Prdm14* which downregulates the expression of *Dnmt3a* and *Dnmt3b* (Leitch et al., 2013).

5.2.2 The transcriptional and epigenetics characteristics of totipotency

Epigenetic remodelling in the pre-implantation embryo involves the erasure of epigenetic memories from the maternal and paternal genomes and the establishment of new epigenetic signatures including imprinting related gene expression and X-chromosome inactivation. Debates on the role of post-zygotic demethylation and subsequent remethylation have also considered that this de novo methylation may in fact be linked to the first-lineage decisions during mammalian development (Reik et al., 2001).

Mechanistically, this process involves active and passive demethylation of both genomes (Guo et al., 2014). Active demethylation is mediated by *Tet3* dioxygenase-dependent oxidation of 5-methylcytosine (5mC) (Gu et al., 2011). Passive demethylation occurs secondary to 5mC dilution from the following cleavage divisions due to DNA replication (Rougier et al., 1998).

Pre-implantation development proceeds in a stepwise fashion from the fertilized zygote until the first and second cell fate decisions. Analysis of global gene expression during this process identified three characteristic events which correspond with the morphological and functional

changes that are identified in the pre-implantation embryo. Following degradation of maternal RNA, which comprises the first characteristic event, there are two waves of *de novo* zygotic transcription. The first wave corresponds to zygotic gene activation, also referred to as embryonic gene activation (EGA). The second wave which is referred to as mid-pre-implantation gene activation (MGA) immediately precedes the first lineage specification event (Hamatani et al., 2004). Moreover, recent studies have succeeded in reconstructing the individual steps of pre-implantation development by performing global analysis of allelic expression across individual cells of the mouse pre-implantation embryo (Deng et al., 2014).

The existence of subpopulations of totipotent-like cells in ESC culture has recently been described. The endogenous retrovirus *MuERV-L* is restricted in temporo-spatial expression to the zygote and two-cell stage embryo. A group recently reported that a small fraction (<2%) of mouse ESCs expressed this element. These *MuERV-L* expressing ES cells expressed additional genes known to be associated with the two-cell stage of development and did not express pluripotency factors at the protein level. These cells demonstrated expanded potential to both the embryonic and extra-embryonic lineage in chimeras (Macfarlan et al., 2012).

A second study described the existence of a totipotent *Hex1* expressing population in ground state culture conditions. These cells co-expressed pluripotency genes and trophoblast determinants which the authors argued was similar to blastomeres during pre-implantation development (Morgani et al., 2013).

Another report described the production of a totipotent like cell following *in vivo* reprogramming. Transcriptionally, these *in vivo* iPSCs (iviPSCs) were very similar to conventional ES cells but some of the genes that were differentially highly expressed in the iviPSCs were morula specific such as *Nlrp4mf* (Abad et al., 2013).

5.3 Aims of this Chapter

In this chapter, firstly, I aim to perform bulk and single cell RNA sequencing to characterise the transcriptomic properties of EPSCs. Second, using published sequencing data on the pre-implantation developmental timeline in mouse embryos, I aim to assess for the presence of a pre-implantation signature in EPSCs. Finally, I aim to interrogate the epigenetic properties of EPSCs by analysing the genome-wide distribution of H3K4me3 and H3K27me3 modifications.

5.4 Results

5.4.1 The transcriptome of EPSCs shares a pluripotency foundation with ground-state ESCs

These experiments were conducted in collaboration with Dr Cheuk-Ho J. Tsang, Sanger Institute. Where indicated, reference is made to the thesis of Dr Cheuk-Ho J. Tsang who performed extensive transcriptional and epigenetic characterization of EPSCs lines which I generated. In an attempt to understand the molecular characteristics of EPSCs, the transcriptomes of EPSCs and ESCs in different culture conditions were profiled by population RNA-sequencing. In hierarchical clustering, it is evident that the transcriptomes of the cell lines segregated by their maintenance conditions, irrespective of their original derivation methods or culture history **(Figure 5.1)**. This suggests that the transcriptomes of EPSCs and ESCs are convertible, consistent with the experimental results.

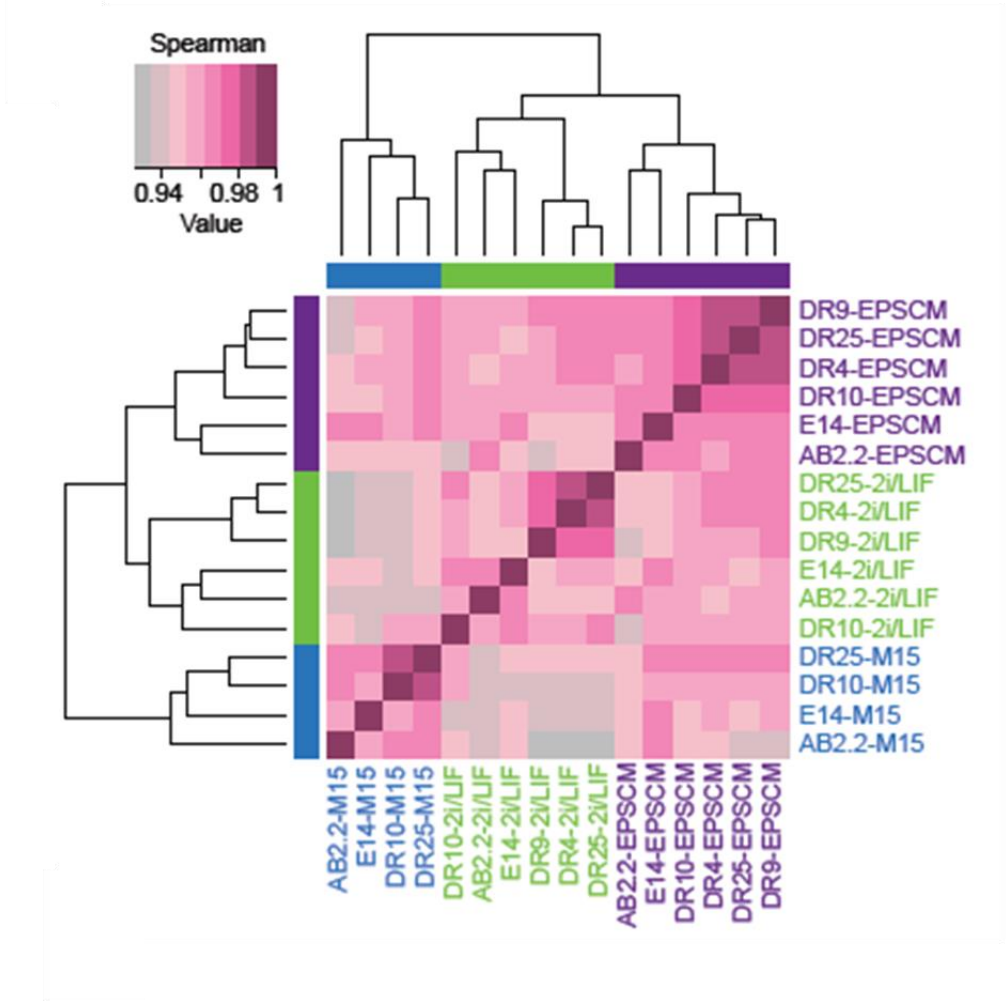


Figure 5.1. Cell line transcriptomes segregate by their maintenance culture conditions. Unsupervised hierarchical clustering of the population transcriptomes. Standard mouse ESC lines (AB2.2 and E14), ESC lines derived in 2i/LIF media (DR4 and DR9) or EPSC lines derived in EPSCM (DR25 and DR10) from preimplantation embryos were cultured in M15, 2i/LIF and EPSCM, respectively. The heatmap represents the inter-sample correlation coefficient calculated by the Spearman rank correlation method.

In addition, 84 individual EPSCs (DR10) were profiled to study the molecular heterogeneities of the EPSC culture. The data was compared with the single-cell RNA sequencing dataset of 2i/LIF and M15 ESCs generated on the same platform with an ES cell line (G4) of similar background (Kolodziejczyk et al., 2015). In principal component analysis (PCA), individual cells were segregated by their culture conditions **(Figure 5.2)**, demonstrating the global differences between EPSCs, 2i/LIF and M15 ESCs at the single-cell level. 2843 and 1465 genes were detected that were up-regulated in EPSCs and 2i/LIF ESCs, respectively. Gene ontology term enrichment analysis revealed that 2i/LIF ESCs were enriched in genes of metabolic processes such as oxidative reduction and the electron transport chain **(Figure 5.3)** concordant with a previous report (Marks et al., 2012). By contrast, biological terms related to transcriptional regulation and embryonic development, particularly placental development, were preferentially featured in EPSCs **(Figure 5.3)**.

Genes instrumental to placental development, such as *Eomes* (Russ et al., 2000), *Peg10* (Ono et al., 2006), *Ascl2* (*Mash2*) (Oh-McGinnis et al., 2011), *Esx1* (Li and Behringer, 1998) and *Gata3* (Home et al., 2009), were differentially up-regulated in EPSCs **(Figure 5.4, see PhD thesis of Dr Cheuk-Ho J. Tsang)**. In addition, multiple lineage-specific genes such as *Gata4*, *Gata6*, *Pax6*, *Sox1* and *Sox7* were also detected in EPSCs at low levels **(Figure 5.4)**. Despite the differences, pluripotency related genes such as *Oct4*, *Sox2*, *Nanog*, *Esrrb*, *Klf2* and *Klf4* were not differentially expressed between EPSCs and 2i/LIF ESCs **(Figure 5.4 and Figure 5.5)** consistent with the *in vitro Rex1-GFP*, *Oct4-GFP* reporter experiments. These results again confirmed that EPSCs and ESCs share the core pluripotency module despite their distinct transcriptome profiles.

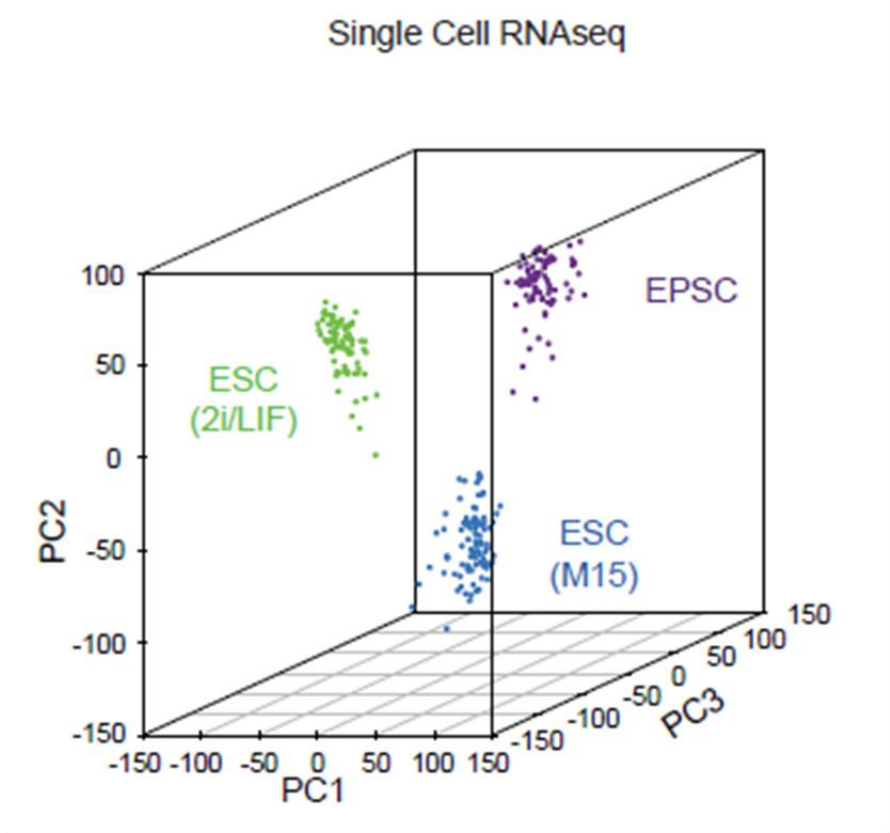


Figure 5.2. EPSCs have a distinct transcriptome from standard ESCs.

PCA of the single-cell transcriptomes of EPSCs and ESCs (M15 and 2i/LIF) were plotted in 3D spaces defined by the first three principal components.

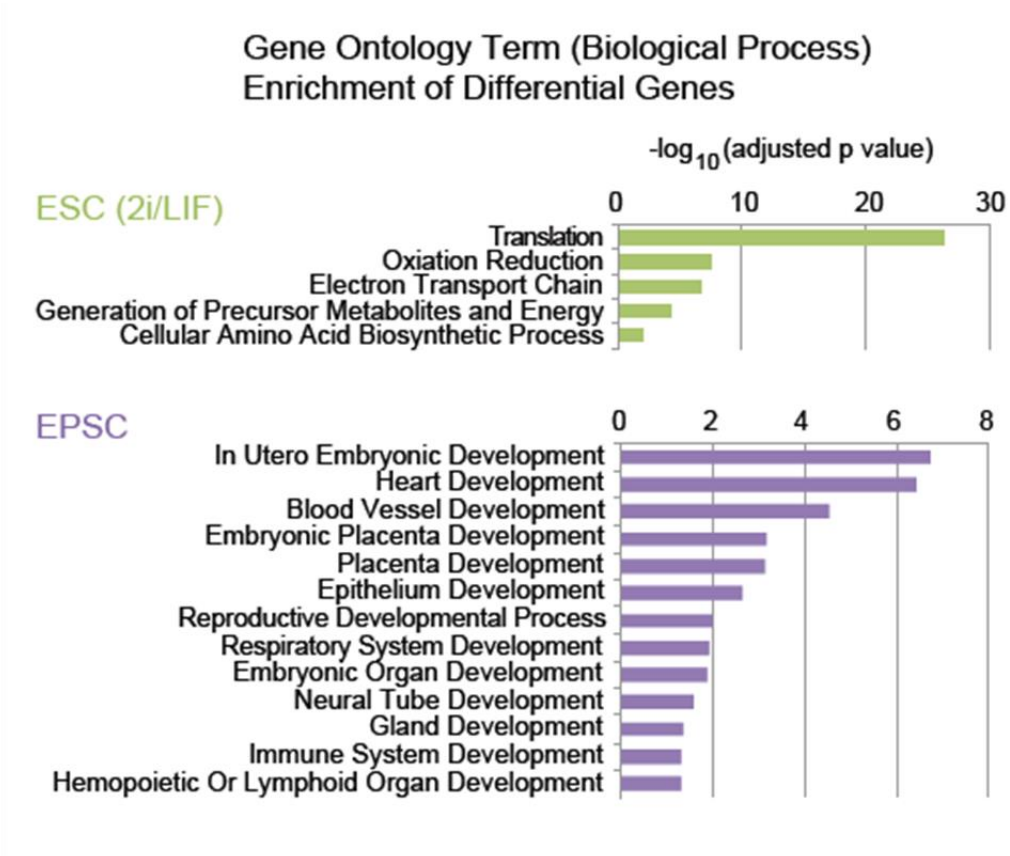


Figure 5.3. EPSCs are enriched in terms related to placental development.

Gene ontology term enrichment analysis of differential genes in EPSCs and 2i/LIF ESCs. Terms related to developmental and metabolic processes are shown (adjusted p-value <0.05) in purple (EPSC) or green (2i/LIF ESC).

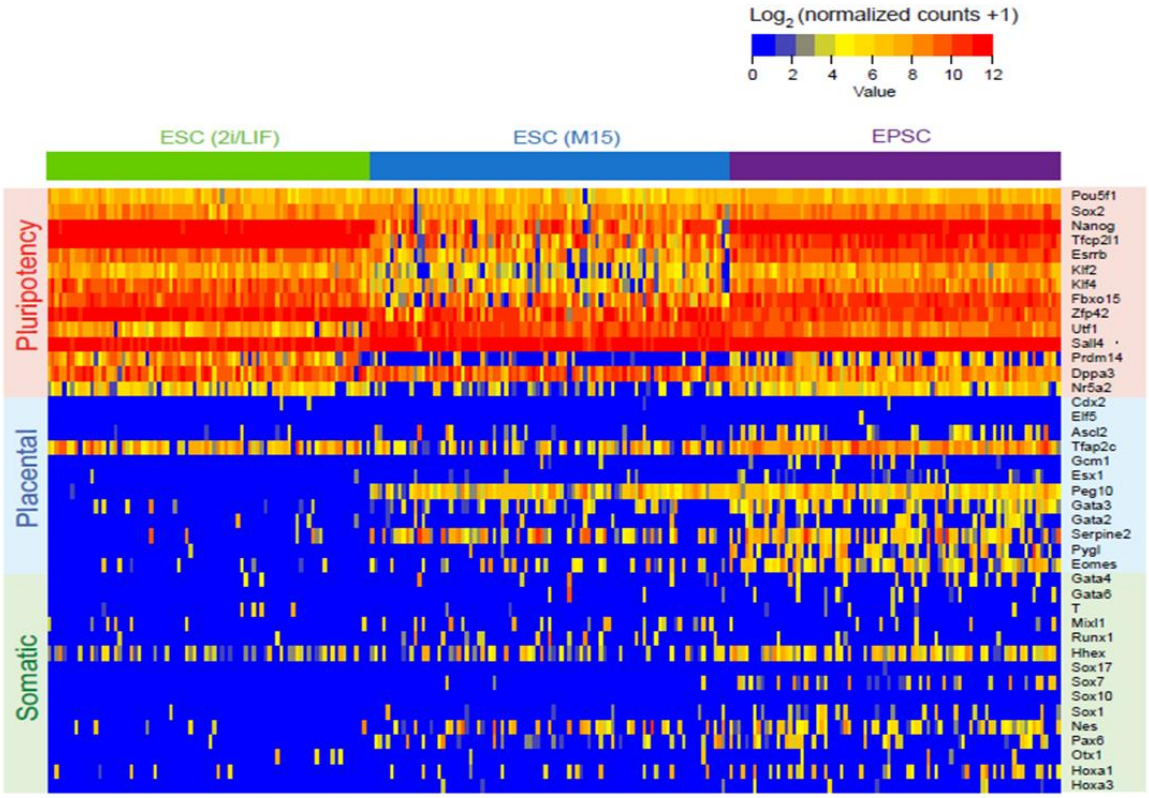


Figure 5.4. Differential up-regulation of genes critical for placental development in EPSCs.

Expression level heatmap showing the expression of pluripotency genes, lineage-specific genes in EPSCs, M15 ESCs and 2i/LIF ESCs. Individual column represents data of a single cell. See PhD thesis of Dr Cheuk-Ho J. Tsang.

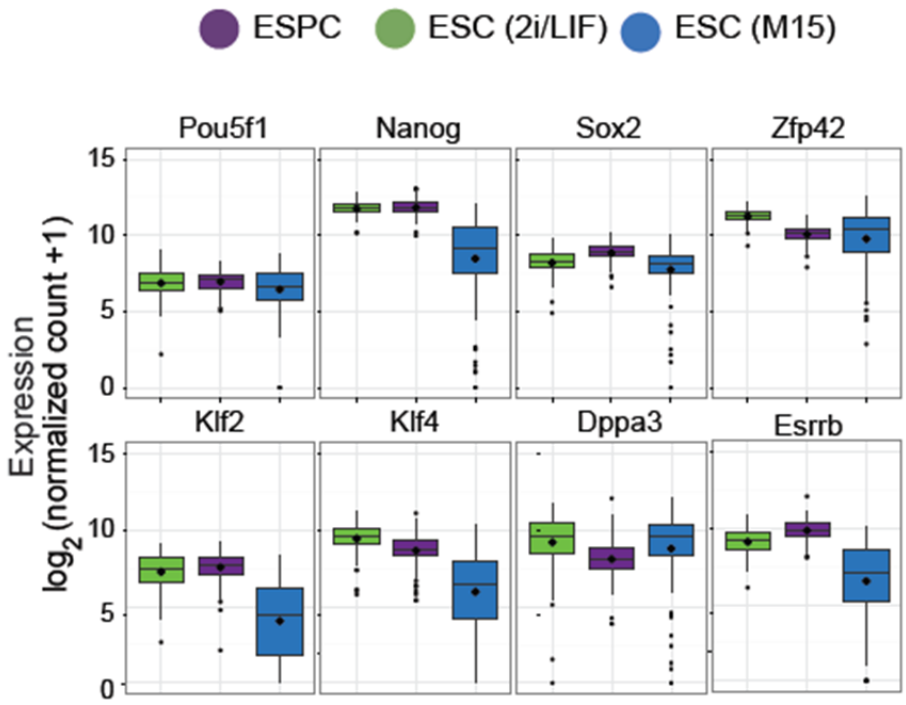


Figure 5.5. Comparable expression of pluripotency related genes in EPSCs and 2i/LIF ESCs.

Boxplots comparing the expression of pluripotency factors in EPSCs and in ESCs. A higher expression fluctuation than that in 2i/LIF ESCs or in EPSCs can be seen in M15 ESCs. The bar and diamond in the box mark the median and mean of each group.

.

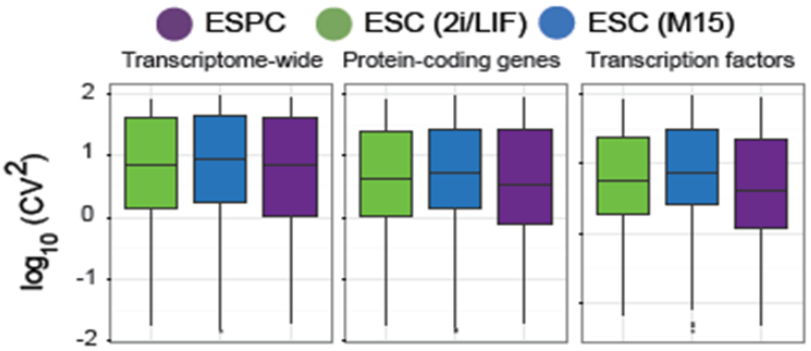
5.4.2 EPSCs and ground-state ESCs share transcriptomic homogeneity

It was previously shown that mouse ESCs cultured in serum-containing media exist in a "metastable" state and display fluctuation of the pluripotency related genes (Marks et al., 2012). For instance, *Nanog* and *Rex1* are heterogeneously expressed in ESCs cultured in serum-containing medium, whereas their expression in 2i/LIF ESCs is more homogeneous. In collaboration with Cheuk-Ho J. Tsang, the single-cell expression variability of key pluripotency factors was compared in EPSCs to 2i/LIF and M15 ESCs. It was found that the expression variability (quantified by the coefficient of variation) of several pluripotency genes remained low (highly homogeneous) and was comparable in EPSCs and 2i/LIF ESCs, unlike the high variability observed in M15 ESCs **(Figure 5.6A and Figure 5.6B)**.

Notably, rare sporadic expressions of lineage-specific genes in 2i/LIF ESCs were observed but this phenomenon was more frequent and pervasive in M15 and EPSCM **(Figure 5.4)**. The expression of the lineage-specific genes in M15 ESCs was associated with variable expression of pluripotency genes, suggesting that the pluripotent states in M15 ESCs are unstable and that they are poised for differentiation. By contrast, EPSCs maintained uniform expression of pluripotency factors. Therefore, the low levels of lineage-specific gene expression observed in EPSCs are not due to *in vitro* differentiation or pluripotent state fluctuation **(Figure 5.4)**. Rather, this transcriptomic feature is consistent with the EPSC's expanded and enhanced developmental potential as described hereby and may represent a salient molecular feature of EPSCs.

The transcriptional variability of global gene expression, as well as those protein-coding genes and putative transcription factors was also assessed and in all cases, EPSCs were similar to 2i/LIF ESCs and had significantly reduced transcriptional variability compared to

A.



B.

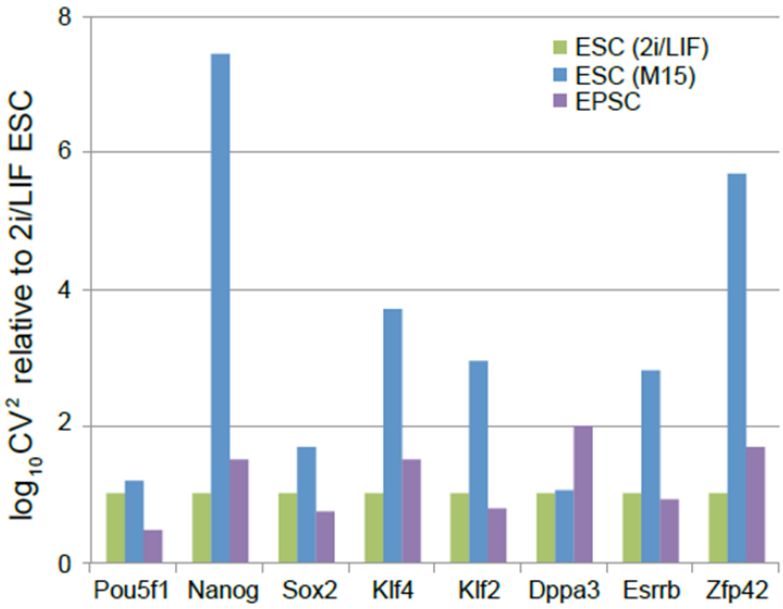


Figure 5.6. EPSCs and ground-state ESCs share a transcriptomic homogeneity.

(A) Boxplots comparing the expression variability of the global transcriptome, protein-coding genes and putative transcription factors in EPSCs and ESCs. The variability is quantified by the log10 (squared coefficient of variation) of individual genes in these cells. The bar in the boxes marks the median of each group. Wilcoxon rank sum test with continuity correction showed significant reduction of squared coefficient of variation in EPSCs compared to M15 ESCs ($p < 0.001$). (B) Comparison of the expression variability of pluripotency genes in EPSCs and ESCs. The variability is compared by the ratio of the squared coefficient of variation (CV2) of these genes compared with that of 2i/LIF ESCs.

M15 ESCs **(Figure 5.6A)**. Collectively, this data indicated that EPSCs have similar transcriptomic homogeneity to 2i/LIF ESCs.

5.4.3 The transcriptome of EPSCs is enriched with early pre-implantation blastomere signatures and is distinct from other rare totipotent-like cells.

I next assessed the biological similarities of EPSCs to *in vivo* pre-implantation totipotent blastomeres. In collaboration with Cheuk-Ho J. Tsang, mouse embryonic development time course single-cell data from Deng et al was retrieved for comparison (Deng et al., 2014). EPSCs were separated from the native developmental trajectory of totipotent blastomeres with the scores of EPSCs at the range of 4 cell stage blastomeres in PC1 **(Figure 5.7, see PhD thesis of Cheuk-Ho J. Tsang)**, the component representing the major embryonic development axis. To test whether EPSCs have transcriptomic features of pre-implantation embryos compared with standard ESCs, the top 500 stage-specific genes of each blastomeric stage from Deng et al were compiled and compared with the expression of these signature gene sets in EPSCs and 2i/LIF ESCs by GSEA. The result showed significantly higher enrichment of early pre-implantation (zygote, 2C and 4C) signatures in EPSCs **(Figure 5.8A)**. This data suggests that the transcriptome of EPSCs contains molecular features of early blastomeres.

In addition, comparison with the published population RNA-seq data showed that EPSCs were distinct in PCA from the recently reported rare totipotent-like ESC subpopulations (MERV-TdTomato+ and Hex-Venus+) (Macfarlan et al., 2012; Morgani et al., 2013) or *in vivo* reprogrammed iPSCs (iviPSCs) (Abad et al., 2013) **(Figure 5.9A)**. For instance, unlike in MERV-TdTomato+ 2C-like cells, only limited up-regulation of endogenous retroviral transcripts was observed in EPSCs **(Figure 5.9B)**.

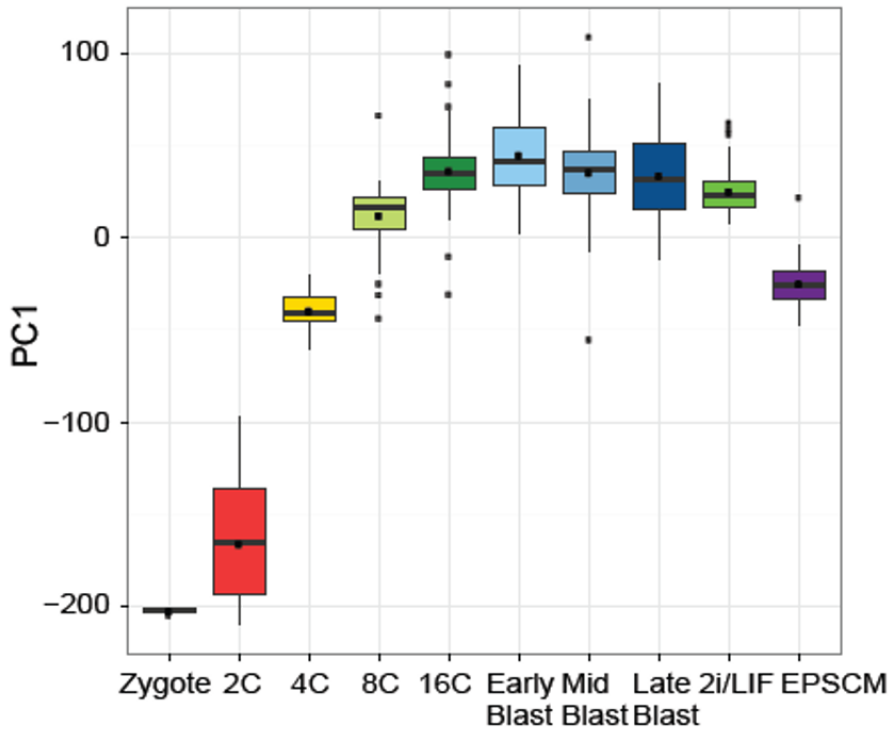


Figure 5.7. EPSCs have transcriptomic similarity to the 4 to 8 cell stage mouse embryo. The zygote to 16-cell (16C) stage embryo raw count matrix from Deng et al was merged with the EPSC and 2i/LIF raw count data and normalized by size factors using DESeq2. Genes that were found to be differentially expressed between EPSCs and 2i/LIF ESCs were used as the testing matrix. The PCA boxplot shown above compared the distribution of scores for each stage of pre-implantation development in the first principal component (PC1), which is the component representing the major embryonic development axis. This distribution was then used to predict the scores (position in the pre-implantation timeframe) of EPSCs and 2i/LIF ESCs. EPSC scores occupied the range of four (4C) to eight (8C) cell stage embryos and 2i/LIF ESCs were situated at the range of 16-cell to blastocyst stages. The bar and diamond in the boxes mark the median and mean scores of each group. See PhD thesis of Dr Cheuk-Ho J. Tsang

A.

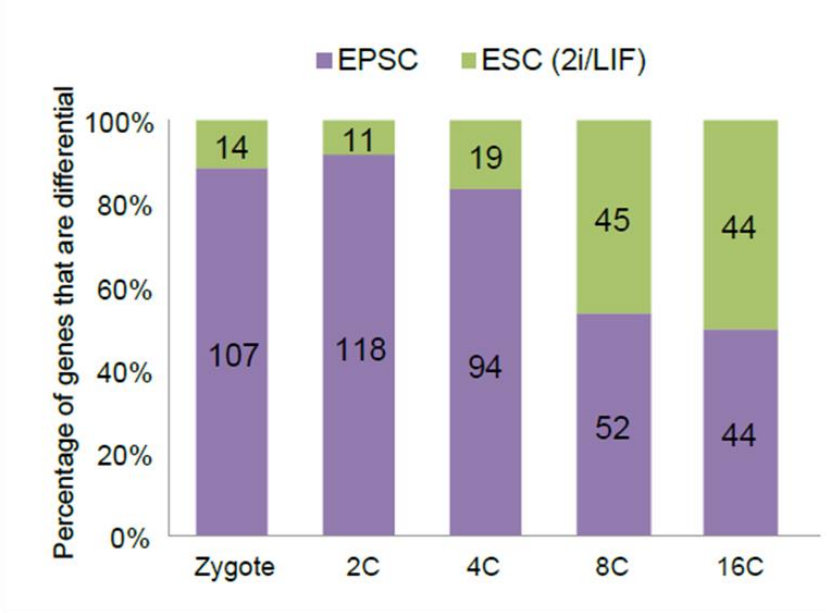
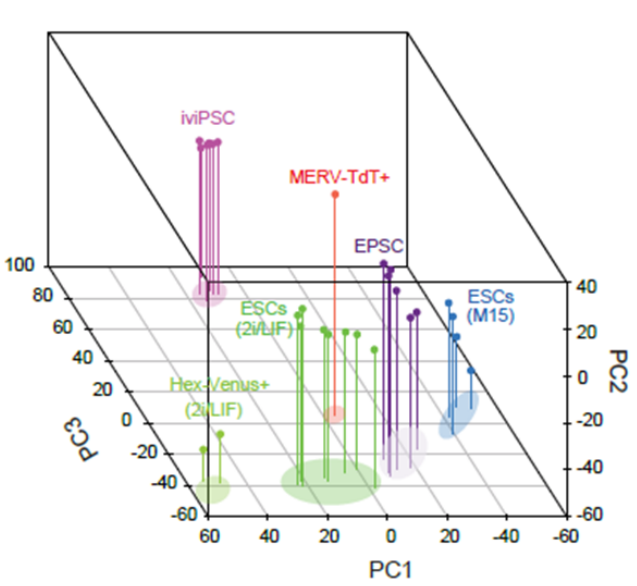


Figure 5.8. Gene set enrichment analysis suggests that the transcriptome of EPSCs contains molecular features of early blastomeres.

(A) The fraction of differential genes in the top 500 pre-implantation embryo stage-enriched genes between EPSCs and 2i/LIF ESCs. The green and purple fractions indicate the percentages of genes that were up-regulated in 2i/LIF ESCs and EPSCs, respectively. The number in each box indicates the number of genes in that fraction. If EPSCs are more similar to blastomeres of a particular stage than 2i/LIF ESCs, majority of the genes specific to that stage would show significantly higher expression in EPSCs compared to 2i/LIF ESCs. On the other hand, if the fraction in a particular stage is equal between EPSCs and ESCs, it suggested neutral similarity of that stage between the two cell types.

A.



B.

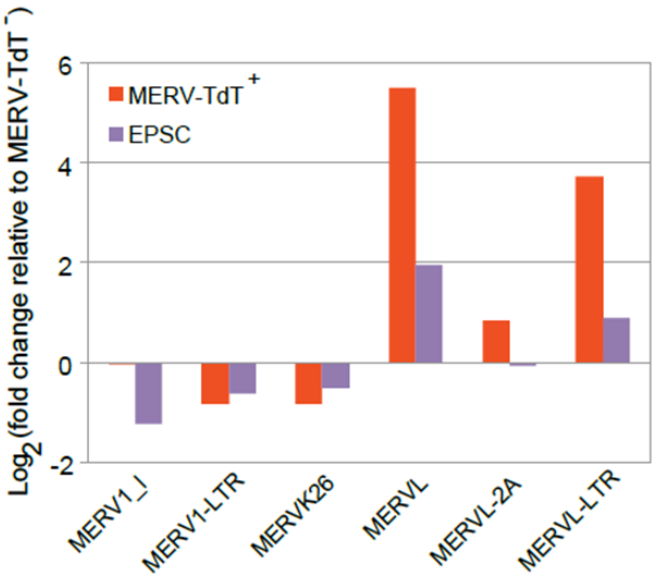


Figure 5.9. EPSCs are distinct from recently reported rare totipotent-like ESC subpopulations.

(A) Three-dimensional PCA visualization of EPSCs, ESCs and other totipotent-like cells. (B) Relative expression of selected endogenous retroviral transcripts in EPSCs and MERV-TdT+ ESCs.

5.4.4 EPSCs have a dynamic DNA methylome

It has been suggested that ERK signalling inhibition in ESCs induces global DNA hypomethylation, which is associated with *Prdm14*-mediated down-regulation of *Dnmt3a* and *Dnmt3b* (Leitch et al., 2013). In collaboration with Cheuk-Ho J. Tsang, intriguingly, we observed significant up-regulation of methyltransferases (*Dnmt1, Dnmt3a* & *Dnmt3b*) and reduction of *Prdm14* despite the presence of ERK inhibition in EPSCM **(Figure 5.10, see PhD thesis of Cheuk-Ho J. Tsang)**.

Genetic components involved in active DNA demethylation such as *Tet1*, *Tet2* and *Tdg* were also up-regulated in EPSCs. The gene expression pattern is reflected on the global cytosine methylation and hydroxymethylation levels quantified by mass spectrometry **(Figure 5.11A and Figure 5.11B, see PhD thesis of Cheuk-Ho J. Tsang)**. Consistent with previous reports (Ficz et al., 2013; Leitch et al., 2013), substantially lower levels of DNA methylation were found in 2i/LIF ESCs, compared to that in M15 **(Figure 5.11A)**, whereas EPSCs showed an intermediate level **(Figure 5.11A)**. Strikingly, much higher levels of hydroxymethylcytosine were found in EPSCs **(Figure 5.11B)**. Albeit at one point in time, this data suggests that the DNA methylome of EPSCs is highly dynamic with active DNA methylation and demethylation.

5.4.5 Prevalent bivalent domains in EPSCs at loci encoding key developmental regulators

In collaboration with Cheuk-Ho J. Tsang, to further understand the epigenetic status of EPSCs, the genome-wide distribution of H3K4me3 and H3K27me3 modifications were profiled in an EPSC line (DR10) and compared with that in 2i/LIF ESCs (Marks et al., 2012). As expected, the H3K4me3 signal in EPSCs predominantly marked promoter regions and

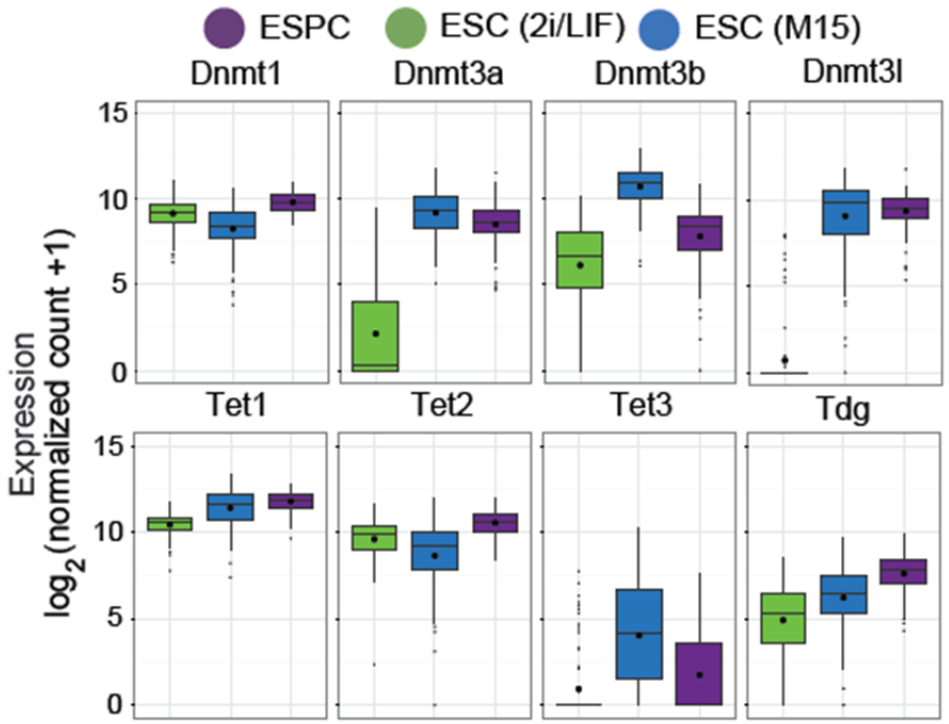


Figure 5.10. Genetic components involved in active DNA methylation and demethylation are up-regulated in EPSCs.

Boxplots comparing the expression of genetic components involved in methylation and demethylation of cytosine in EPSCs and in ESCs. The bar and diamond in the box mark the median and mean of each group. See PhD thesis of Dr Cheuk-Ho J. Tsang.

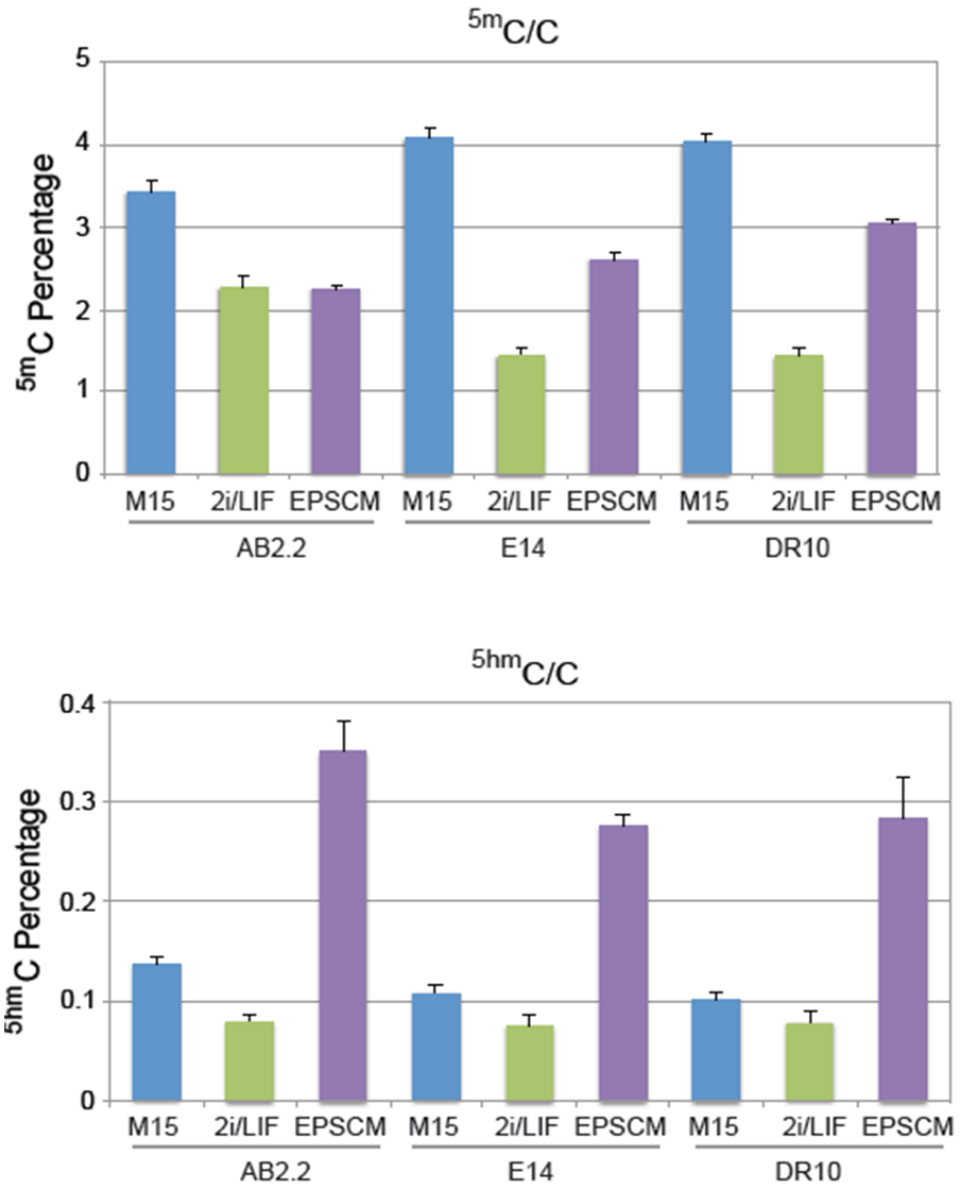


Figure 5.11. EPSCs have a highly dynamic methylome.

Column charts comparing the fractional levels of methylated (**A**) and hydroxymethylated (**B**) cytosine in EPSCs and in ESCs as quantified by mass spectrometry. See PhD thesis of Dr Cheuk-Ho J. Tsang, Sanger Institute.

correlated with the level of active gene expression **(Figure 5.12A)**, whereas H3K27me3 peaks were distributed throughout the promoters and gene bodies and inversely correlated with gene expression **(Figure 5.12B)**.

Compared with 2i/LIF ESCs, EPSCs contain more genes (6224 vs. 3968) associated with both H3K4me3 and H3K27me3 (bivalent) histone modifications **(Figure 5.13)**. Bivalency could be detected at the genetic loci encoding key lineage regulators such as *Sox1*, *Pax6*, *Gata6*, *T*, *Cdx2* and *Eomes* **(Figure 5.14)**. By contrast, the H3K4me3 and H3K27me3 signals of EPSCs at key pluripotency loci such as *Pou5f1*, *Sox2* and *Nanog* were highly similar to 2i/LIF ESCs **(Figure 5.14)**, consistent with the fact that these pluripotency genes were expressed at similar levels in both cell types **(Figure 5.4 and Figure 5.5)**. Gene ontology term enrichment analysis showed that EPSC specific bivalent genes were enriched in biological processes of somatic lineage and placenta development **(Figure 5.15)**, including genetic loci such as *Cdx2*, *Eomes*, *Esx1*, *Hand1*, *Ascl2*, *Sox1*, *Sox17*, and *Pax6* **(Figure 5.14)**, which provide an epigenetic basis for the low expression of lineage-specific genes detected in single cell RNA-seq **(Figure 5.4)**. **See PhD thesis of Cheuk-Ho J. Tsang for extensive characterization of the epigenetic state in EPSCs**.

5.5 Discussion

In this chapter I have demonstrated that established EPSC lines are molecularly similar, independent of their derivation origins, suggesting that EPSCM is capable of maintaining a convergent cell state. Despite the commonality of self-renewal network in EPSCs with ESCs, molecular characterization reveals that EPSCs have a distinct transcriptome, DNA methylation content and histone methylation pattern compared with ESCs.

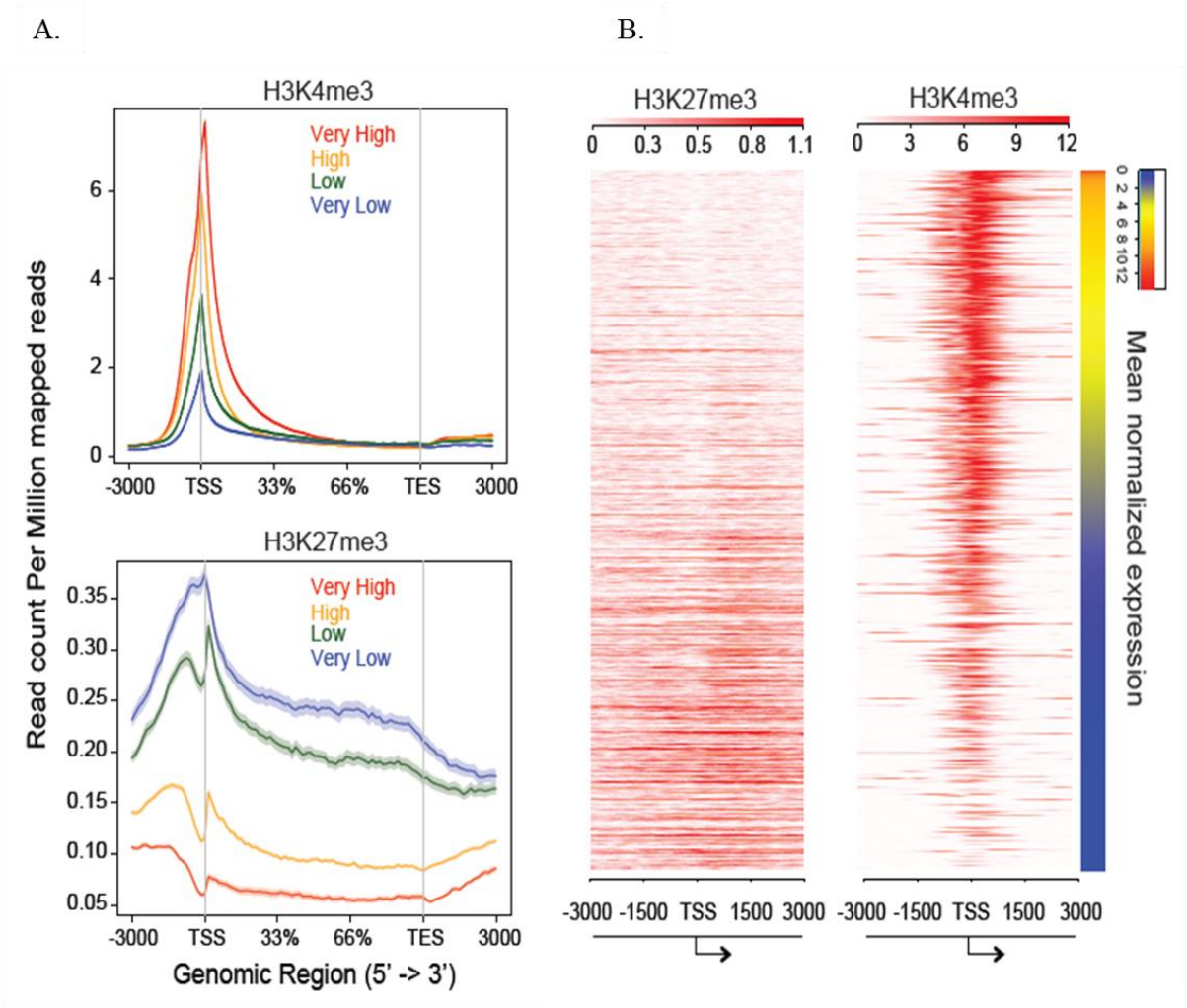


Figure 5.12. Epigenetic profiling of EPSCs.

(A) Distribution of H3K4me3 (upper panel) or H3K27me3 (lower panel) signals across gene bodies in the EPSC epigenome. Genes are classified into "Very High", "High", "Low" and "Very Low" depending on their length-corrected mean-normalized count quartiles in the single-cell RNA-seq dataset. (B) Distribution of H3K27me3 (left) and H3K4me3 (right) signals at gene promoters (+/-3kb from TSS). The promoters were ranked based on their length-corrected mean-normalized count levels in the single-cell RNA-seq dataset. The signals were quantified as read count per million mapped reads.

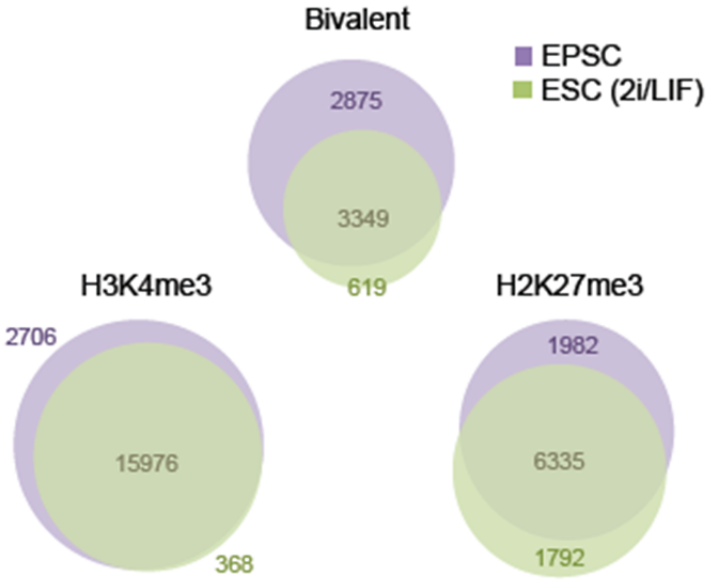


Figure 5.13. Prevalent bivalent histone domains in EPSCs.

Pie charts comparing the number of H3K4me3, H3K27me3 or bivalent histone domains associated genes in EPSCs and 2*i*/LIF ESCs.

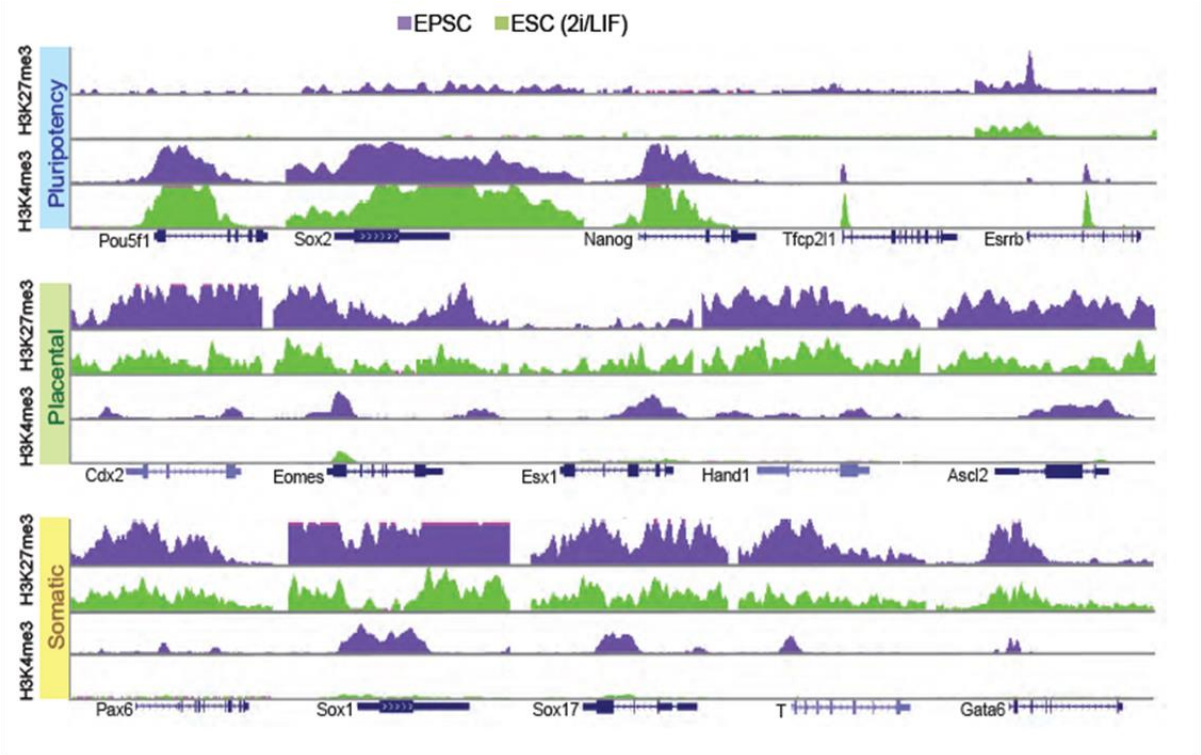


Figure 5.14. Prevalent bivalent histone domains in EPSCs are associated with developmental loci.

Profiles of H3K4me3 and H3K27me3 signals of 2i/LIF ESCs and EPSCs (DR10) in selected pluripotency and developmental genes. The ranges of signals in H3K4me3 and H3K27me3 are (0-100) and (0-20), respectively. See PhD thesis of Dr Cheuk-Ho J. Tsang, Sanger Institute.

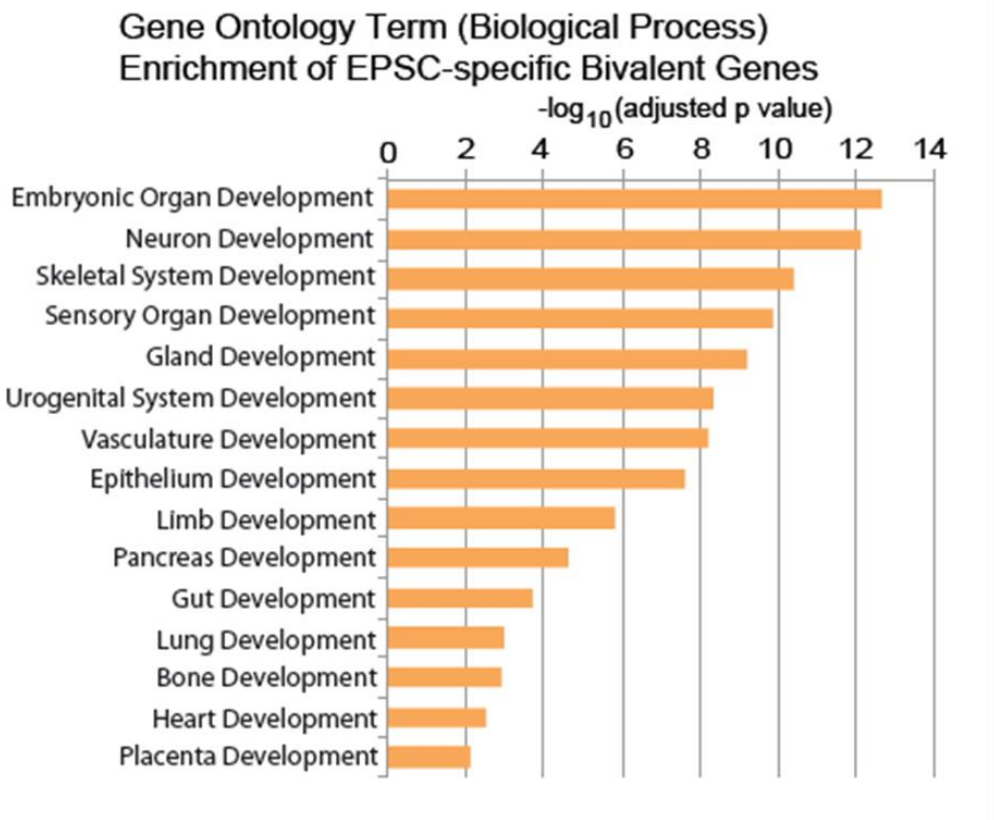


Figure 5.15. EPSC specific bivalent genes are enriched in biological processes of somatic lineage and placenta development.

Bar chart showing selected biological gene ontology terms enriched in EPSC-specific bivalent genes. Only terms relevant to development are shown here.

By comparing with the single-cell RNA-sequencing data of pre-implantation embryos in principal component analysis, it is clear that EPSCs deviate from the native blastocyst cells and 2i/LIF ESCs in the embryonic development trajectory towards the direction of earlier blastomeres. Moreover, EPSCs contain increased bivalent domains at the genetic loci of embryonic and extra-embryonic lineage specifiers.

It has been suggested that "promiscuous" transcription at bivalent loci is a hallmark of "metastable" ESCs cultured in serum-containing media, and these genes are 'poised' for activation upon differentiation induction (Marks et al., 2012; Takashima et al., 2014). In spite of the increased bivalent domains in EPSCs, single-cell RNA-sequencing shows that the transcriptome of EPSCs has reduced expression variability compared with ESCs maintained in serum-containing media. The presence of 'poised' promoters at key developmental genes in EPSCs may account in part for the expanded potential of these cells by permitting rapid gene expression in response to appropriate *in vivo* or *in vitro* developmental cues. As a proof of concept experiment and in support of this hypothesis it has been shown that EPSCs have markedly enhanced hematopoiesis potential which may be explained in part by the observation of bivalent domains at key transcription factors involved in blood lineage specification (Performed by Adam Wilkinson, Gottgen's Lab, Cambridge University).

During mouse pre-implantation development there is progressive global DNA demethylation which culminates in the formation of the hypomethylated inner cell mass (Leitch et al., 2013). The ICM possesses more 5 hydroxymethylcytosine (5hmC) than the trophectoderm with as expected a similar finding in ESCs and TS cells, respectively (Ruzov et al., 2011). The global expression of *Tet1* and *Tet2* mRNAs and 5hmC abundance is less in the trophectoderm (Ito et al., 2010; Senner et al., 2012). Moreover, loss of *Tet1* expression and resulting loss of 5hmC in one of the blastomeres in the two-cell stage embryo biases that blastomere to a TE fate (Ito

et al., 2010). In contrast to ESCs, EPSCs express genes related to both methylation and de-methylation which is functionally supported by the presence of elevated 5mC and 5hmC on mass spectrometry. As EPSCs are a homogenous population, this co-expression of methylation-demethylation components in each cell may explain in part the observed plenipotent or unbiased embryonic and extra-embryonic differentiation potential of an individual cell as demonstrated in the single cell chimeric contribution assays in Chapter 4.

The rare 2C-like and Hhex^{+} ESCs were reported to transiently exist in standard ESC culture (Macfarlan et al., 2012; Morgani et al., 2013). Meanwhile, *in vivo* reprogrammed iPSCs were shown to be highly similar to ESCs transcriptionally but express several morula-stage genes (Abad et al., 2013). It remains unclear whether the expanded potential observed in the *in vivo* iPSCs is contributed by similar rare ESC subpopulations in culture since no single cell genomics or experimental data are available. Nevertheless, EPSCs are categorically different from these reported totipotent-like cells. EPSCs do not have up-regulation of endogenous retroviral transcripts and express high levels of OCT4 protein.

In chapter 4, I demonstrated that EPSCs possess functional homogeneity of expanded potential in single-cell chimera formation experiments. In this chapter, single-cell transcriptomic experiments have confirmed this molecular homogeneity further refuting the existence of rare privileged subpopulations within EPSCM. Nevertheless, it remains possible that stable homogenous long-term 2C-like cell cultures can be established by modulating additional pathways or epigenetic modifiers in the future (Macfarlan et al., 2012).

CHAPTER 6

6 DERIVATION AND MAINTENANCE OF HUMAN EPSCS

6.1 Summary

In this chapter, I demonstrate that EPSCM permits the establishment *in vitro* of a human stem cell state that possesses expanded differentiation potential to both the embryonic and extra-embryonic lineage *in vitro*. Similar to mouse EPSCs, human EPSCs possess a core pluripotency transcriptional self-renewal network and are easily amenable to gene targeting. Epigenetic profiling of X-chromosome activation status in EPSCs provides evidence to suggest that these cells may correspond to a pre-implantation stage of development and provisional gene set enrichment analysis demonstrates enrichment in genes related to the eight-cell stage embryo.

6.2 Introduction

6.2.1 Conservation of signalling pathways in the pre-implantation embryo

Despite difference in the timing of onset of zygotic gene activation (ZGA) and the temporal rate of embryo development, available studies appear to suggest that the fundamental processes involved in the first cell fate specification may be conserved between mouse and human (Adjaye et al., 2005; De Paepe et al., 2014; Santos et al., 2002; Santos et al., 2010; Senner and Hemberger, 2010; Telford et al., 1990). Beyond the specification event, it appears that the factors involved in establishment of the trophoblast lineage also appear to be shared between mouse and man (Senner and Hemberger, 2010).

This knowledge and recent reports from several groups describing *in vitro* conditions to maintain human ESCs with transcriptional and metabolic similarities to mouse ESCs in 2i/LIF conditions led me to consider that if a cell state(s) similar to the mouse ground-state of self-renewal exists in humans and if many of the signalling processes involved in pre-implantation development are conserved, perhaps an EPSC state can also be established in human (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014).

Beyond establishing such a line *in vitro*, an obvious challenge in humans is evaluating for the presence of expanded differentiation potential as *in vivo* chimeric contribution assays are not ethically permissible. Therefore, surrogate functional *in vitro* assays are necessary for evaluating the presence of a pre-implantation signature and secondly, the capacity for expanded differentiation potential during differentiation. Epigenetic regulation of X chromosome inactivation (XCI) and *in vitro* trophoblast differentiation are potentially useful assays for answering these respective questions.

6.2.2 Epigenetic regulation of X-chromosome inactivation (XCI)

In terminally differentiated somatic cells of female mammals, one of the X-chromosomes is inactive and can be identified as a dense condensation of heterochromatin called the Barr body (Lyon, 1961). This process of XCI can be random or imprinted and has evolved to mediate dosage compensation of gene expression between the sexes.

In eutherian mammals, the onset of XCI in the pre-implantation embryo coincides with embryonic gene activation. In the mouse, this corresponds to the two-cell stage embryo and XCI is imprinted with preferential inactivation of the paternal X-chromosome which remains silent throughout pre-implantation development and in descendants of the extra-embryonic lineage (Huynh and Lee, 2003; Takagi and Sasaki, 1975). The paternal X-chromosome becomes re-activated in the inner cell mass and subsequently the cell will undergo random XCI during somatic differentiation of the epiblast (Mak et al., 2004). The importance of early imprinted XCI in mouse has been demonstrated in parthenogenetic embryos where the lack of paternal imprinting due to two maternal genomes is embryonically lethal (Nesterova et al., 2001; Shao and Takagi, 1990).

Mammals employ different strategies for the initiation and maintenance of XCI. Zygotic gene activation occurs at later stages in non-rodent species, specifically the four to eight-cell stage in humans and the eight to 16-cell stage in rabbits (Telford et al., 1990). Therefore, and not surprisingly, it has been shown that XCI occurs at later stages in these species (Okamoto et al., 2011; Van den Berg et al., 2009). An additional difference between human and mouse is that initial XCI in humans is not imprinted and XCI is not necessary for correct specification of trophectoderm in blastocyst stage embryos (Okamoto et al., 2011).

In mouse, initiation of XCI is controlled by the long non-coding RNA, X inactive specific

transcript (*XIST*) which coats the inactive X-chromosome (Penny et al., 1996). The antisense transcription factor *Tsix* prevents inactivation of the remaining active X-chromosome (Lee, 2005). Interestingly, *Tsix* is poorly conserved and in humans, although *XIST* is present and does coat the inactive X-chromosome, it is not thought to be responsible for the initiation of XCI (Chang and Brown, 2010; Okamoto et al., 2011). Humans also possess another long non-coding RNA, *XACT* which coats the active X-chromosome and which is not present in the mouse. *XACT* is thought to be pluripotency specific and in the absence of *XIST* will coat both active X-chromosomes (Vallot et al., 2015). In the mouse inner cell mass, both X chromosomes are active with no repressive *XIST* transcript identified. In the human inner cell mass, *XIST* is detected but the cells still possess two active X-chromosomes (Okamoto et al., 2011).

With this understanding of human XCI, it should therefore be possible to temporally place a cell of interest at a specific locus in the pre-implantation development timeline based on their X activation - inactivation status alone.

6.2.3 Human Trophoblast Specification

The human placenta just like the mouse is comprised of cell types derived from both the embryonic and extra-embryonic lineages. The extra-embryonic portion is contributed by the trophoblast cells which are sub-divided into the villous and extra-villous cytotrophoblast and syncytiotrophoblast all of which have different biological functions in placental formation. The embryonic portion as in the mouse comes from the extra-embryonic mesoderm, a derivative of the epiblast, which forms the placental vasculature. At the time of initial derivation of human ESCs from embryos, it was noted that spontaneous differentiation to cell types that secreted β-HCG occurred suggesting that human ESCs had a capacity to form

trophoblast-like cells (Thomson et al., 1998). Further studies elucidated that this phenotype could be reproduced by the addition of BMP4 which permitted the establishment *in vitro* of cell lines that had similarity with post-mitotic syncytiotrophoblast and cytotrophoblast (Harun et al., 2006; Xu et al., 2002). As human ESCs are thought to represent a 'primed' or post-implantation stage of development, this finding was at odds with our current understanding of normal development (Tesar et al., 2007). Numerous debates have centred on whether this human ESC differentiation represents bona fide trophoblast induction or whether this phenotype may be explained by somatic contribution from the extra-embryonic mesoderm (Roberts et al., 2014). Whilst attempting to reconcile this dichotomy, *ELF5* was identified as a key transcription factor whose epigenetic status would allow identification of genuine trophoblast (Hemberger et al., 2010).

ELF5 is a transcription factor expressed in first trimester villous trophoblast but not in extra-villous or syncytiotrophoblast. In the first trimester, there is a transient highly proliferative ELF5, CDX2 double positive sub-population within the placenta which is thought to represent the putative human trophoblast stem cell niche (Hemberger et al., 2010). *ELF5* along with *CDX2* and *EOMES* forms a 'trophoblast-reinforcing transcriptional network' (Senner and Hemberger, 2010). Critically, the methylation status of *ELF5* creates an epigenetic signature for the trophoblast compartment. *ELF5* is hypomethylated in the trophoblast lineage but is significantly hypermethylated in human ESCs. Moreover, differentiated 'trophoblast' derived from human ESCs remains hypermethylated at *ELF5* indicating that this *in vitro* differentiation process may not produce bona fide trophoblast (Hemberger et al., 2010).

Therefore, one can conclude that in order to prove that a cell type has expanded differentiation potential to the trophoblast lineage *in vitro*, demonstration of hypomethylation at the *ELF5* promoter is paramount.

6.3 Aims of this Chapter

The aim of this chapter was to determine if EPSCM permitted the maintenance *in vitro* of a human stem cell population that possessed expanded differentiation potential *in vitro*. This would be attempted by first converting human ESCs to human EPSCs and by reprogramming human somatic cells to iPSCs. After initial characterization of these cells, expanded potential would be assessed by evaluating the capacity of these cells to produce bona fide trophoblast *in vitro,* epigenetic profiling of their XCI status and transcriptionally by the presence of a unique pre-implantation gene expression signature.

6.4 Results

6.4.1 Derivation of human EPSCs by conversion of standard human ESCs and iPSCs

We chose two human ESC lines H1 and H9, which are widely used in the research community (Thomson et al., 1998). Individual H1 cells were seeded in EPSCM on MEF feeder cells at a low cell density. Approximately 0.3% of cells formed undifferentiated colonies **(Figure 6.1B)**. These colonies could be subcloned again and expanded in EPSCM to establish stable cell lines. In contrast, none of the colonies of H9 cells that initially survived in EPSCM could be further expanded, consistent with the known heterogeneity among human ESC lines. The converted H1 cells (H1-EPSCs) were dissociated to single cells for passaging with a 15-20% single cell subcloning efficiency in the presence of a selective ROCK inhibitor and could be expanded in EPSCM for over 30 passages on SNL feeders. Two standard human iPSC lines from the CGAP facility were obtained (BOB and FSPS186) and the conversion strategy was repeated **(Figure 6.1A)**. BOB iPSCs survived the conversion process with EPSC colonies emerging from single iPSCs which were plated at low density on SNL feeders **(Figure 6.1C)**. FSPS186 iPSCs could not be converted. As with hESCs there is known

heterogeneity among human iPSC lines which may in part explain the variation in the capacity for conversion.

H1-EPSCs proliferated faster than H1 ESCs and had a shorter cell cycle time with 69.1% of H1-EPSCs in the S phase of the cell cycle as compared to 41.9% in H1 ESCs. Less than 13.1% of H1-EPSCs were found in the G0-G1 checkpoint in contrast to 44.2% of H1 ESCs **(Figure 6.2)**. Converted lines were genetically stable as confirmed by spectral karyotyping **(Figure 6.3)**. H1-EPSCs expressed pluripotency genes (*OCT4, SOX2, NANOG, STELLA and PRDM14*) at similar levels with the parental H1 ESCs **(Figure 6.4A)**, but lower levels of lineage markers (*SOX1, PAX6, T, GSC, GATA6 and SOX17*) **(Figure 6.4B)**. H1-EPSCs were efficiently differentiated to cells of all three germ layers **(Figure 6.5)**.

6.4.2 Derivation and maintenance of human EPSCs by reprogramming somatic cells

Using the six factor reprogramming technology (Wang et al., 2011), human neural stem cells were reprogrammed to iPSCs by expressing reprogramming factors from episomal vectors, or dermal fibroblast cells by Doxycycline (DOX) - inducible factors from *piggyBac* (PB) transposon vectors **(Figure 6.6A)**. Stable iPSC lines could be established directly in EPSCM from about half of the primary iPSC colonies **(Figure 6.6B)**. Episome clone 5 iPSC line (EC5-EPSC) expressed key pluripotency markers **(Figure 6.4A)**, had minimal expression of lineage specific genes **(Figure 6.4B)** and was further confirmed to have lost the exogenous reprogramming vectors after five passages **(Figure 6.6C)**. Spectral karyotyping **(Figure 6.7A)** and aCGH **(Figure 6.7B)** demonstrated genetic stability in EC5-EPSCs.

EC5-EPSCs could form mature teratomas *in vivo* containing derivatives of the three germ layers and foci of melanin expression were identified indicating the presence of mature

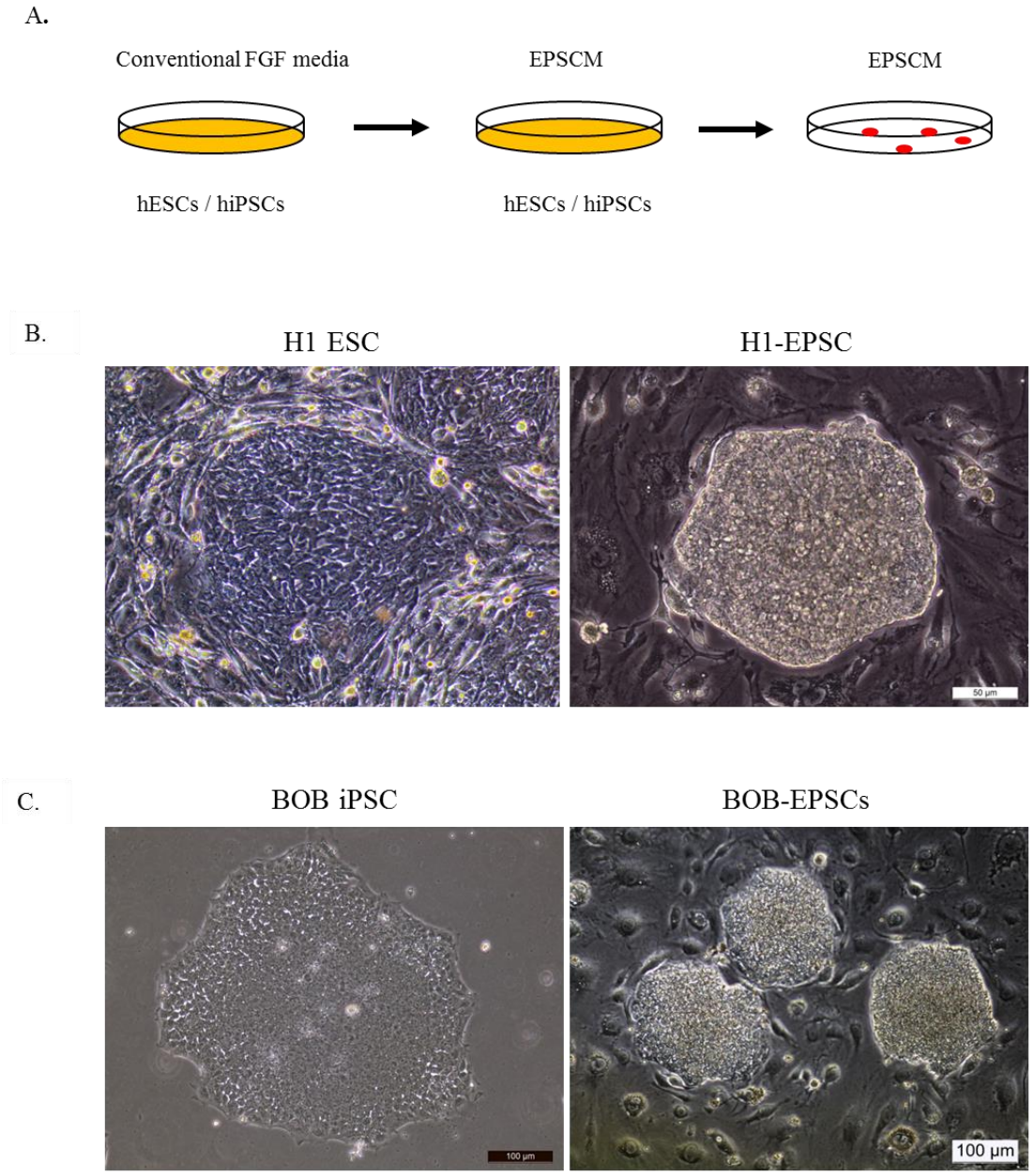


Figure 6.1. Human ESC (hESC) and human iPSC (hiPSC) conversion to EPSCs.

(A) Conversion strategy. (B) Images of colonies of H1-ESCs in FGF containing medium or after they were cultured in EPSCM for 6 passages (H1-EPSC) (scale bar = 50 µm). (C) Images of colonies of BOB-iPSCs in FGF containing medium or after they were cultured in EPSCM for 10 passages (BOB-EPSCs) (scale bar = 100 µm).

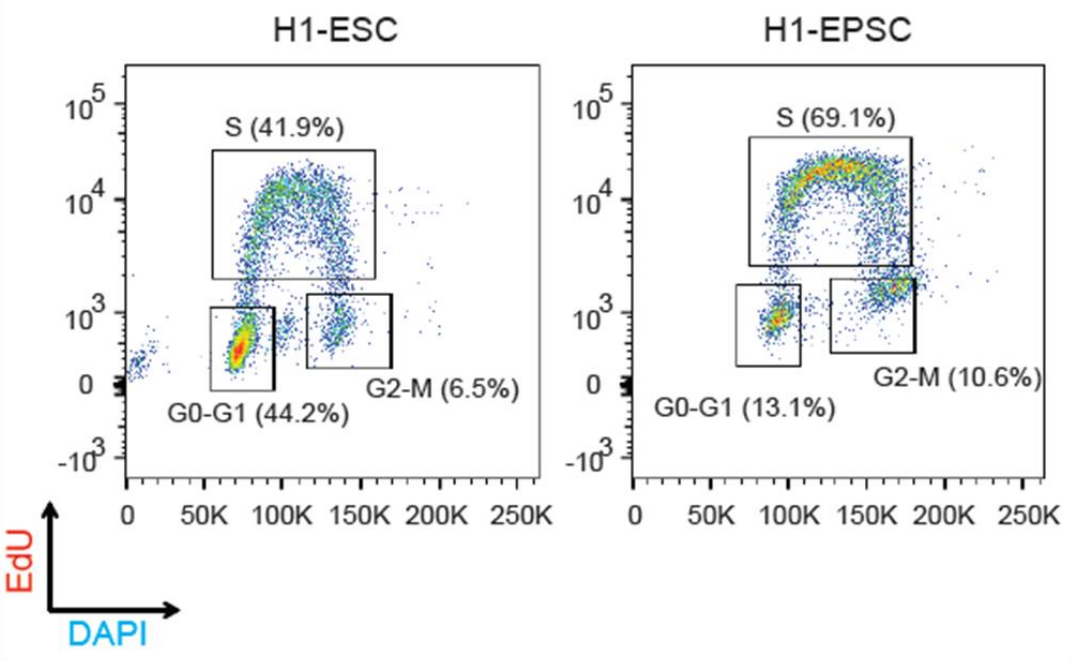


Figure 6.2. Cell cycle analysis in converted human ESCs.

Cell cycle analysis of H1 ESCs and H1-EPSCs (passage 15 in EPSCM). Substantially more EPSCs are in G2/S phases. Thanks to Wei Wang for the cell cycle analysis.

H1-EPSC, 46 XY, 10/10 correct

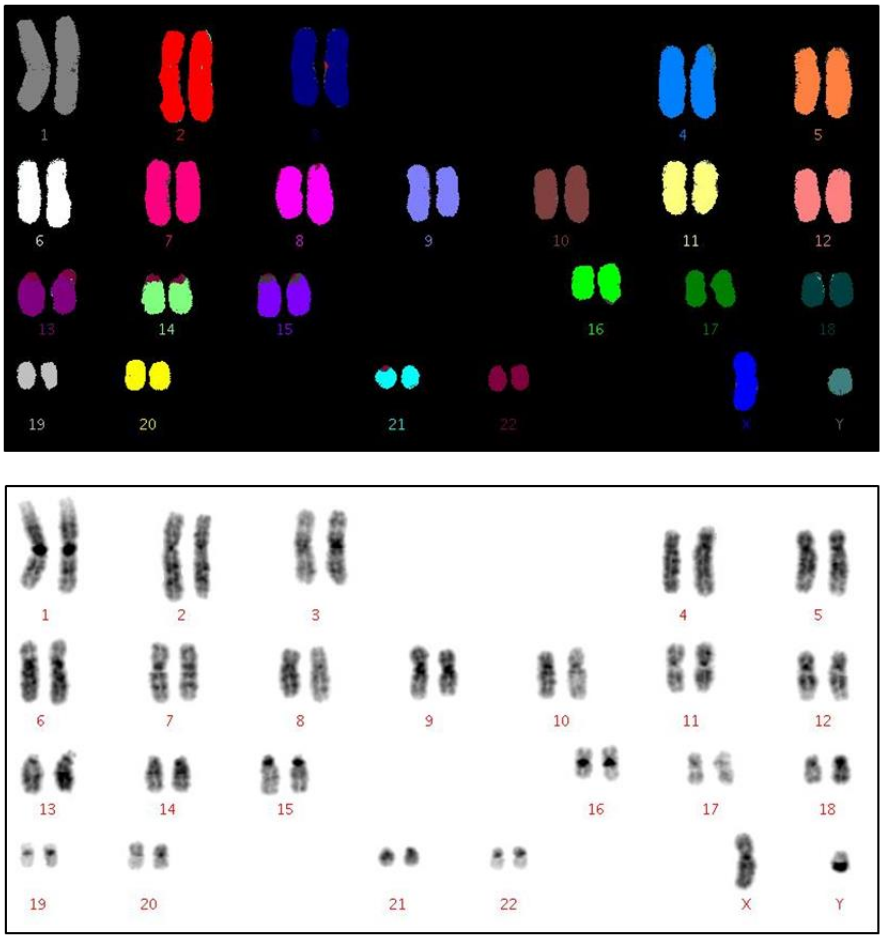
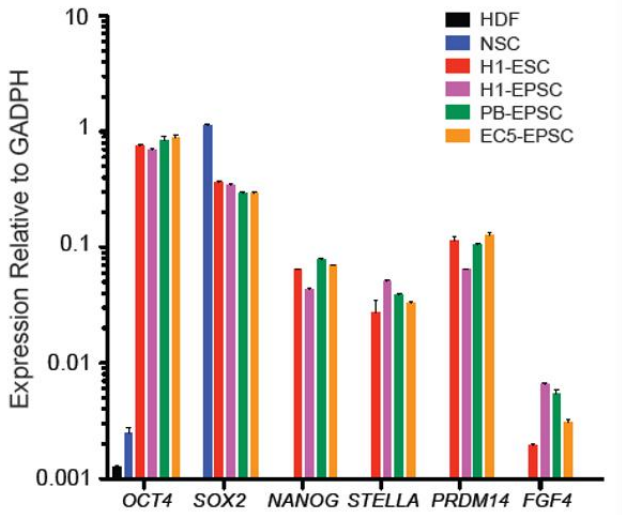


Figure 6.3. H1-EPSCs are genetically stable.

Spectral karyotyping of H1-EPSCs. **10** randomly selected metaphases were karyotyped by M-FISH and DAPI- banding.

A.



B.

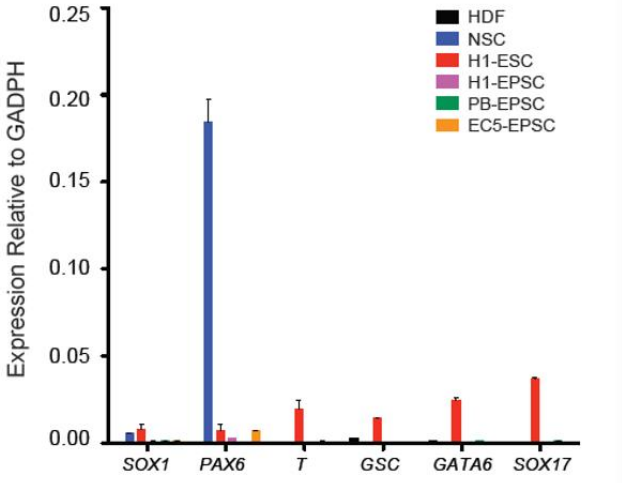


Figure 6.4 H1-EPSCs express comparable levels of pluripotency related genes but lower levels of lineage specifiers relative to conventional H1-ESCs.

(A) and (B) Quantitative PCR expression profile of pluripotency and lineage specific genes in H1-ESCs, H1-EPSCs (passage 15 in EPSCM) and EC5-EPSCs (passage 20 in EPSCM). Included in the analysis are also PB-EPSCs reprogrammed from human dermal fibroblasts (HDF) using the piggyBac (PB) transposition. Parental HDFs and neural stem cells (NSCs) were used as control. Relative expression of these genes was normalized to *GAPDH*. Data are mean ± s.d. Experiments were repeated three times. Performed in collaboration with Wei Wang.

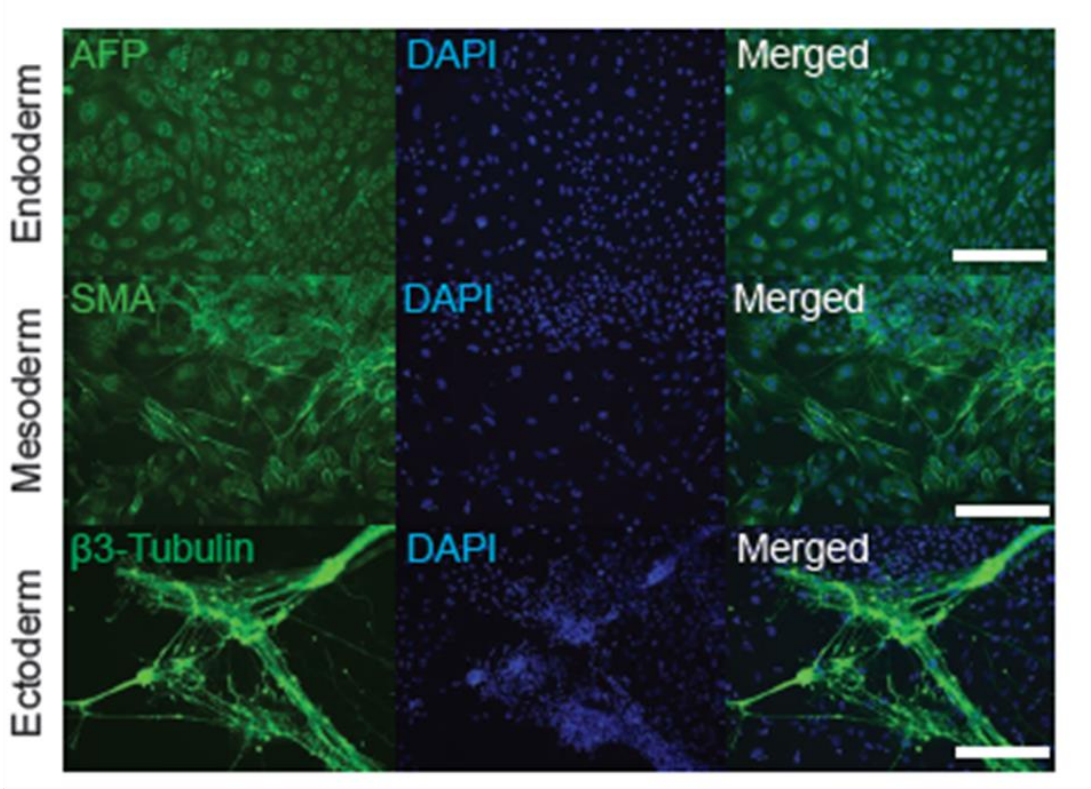


Figure 6.5. H1-EPSCs were efficiently differentiated to cells of all three germ layers.
H1-EPSCs were differentiated according to published embryoid body induction assays and subsequently plated on gelatin coated plates for immunostaining. Antibodies against alpha-feta protein, smooth muscle actin and Beta-3 tubulin were chosen representing the endoderm, mesoderm and ectoderm respectively. Derivatives of all three germ layers were detected in H1-EPSCs embryoid bodies further confirming their in vitro pluripotency. Performed in collaboration with Jian Yang.

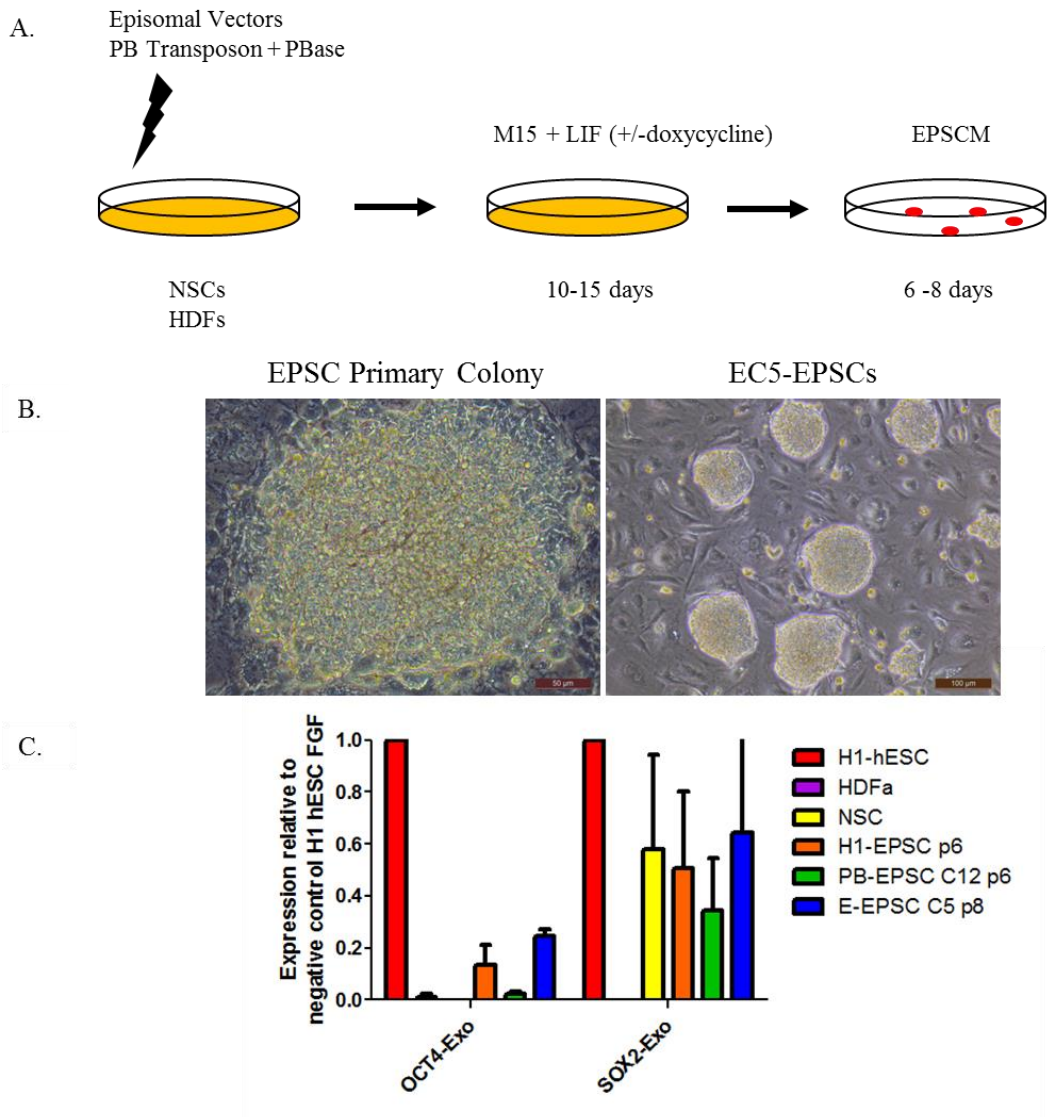
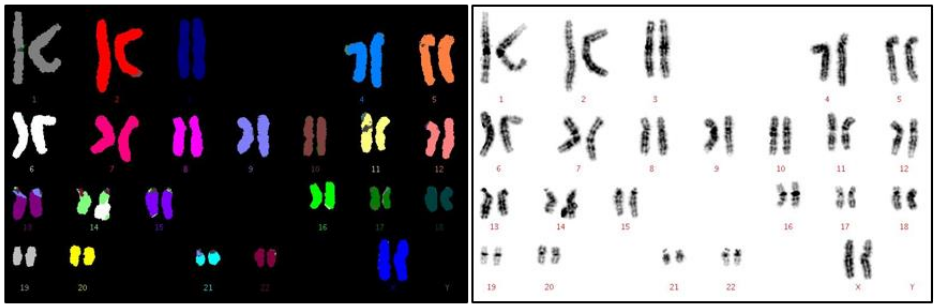


Figure 6.6. EPSCs can be generated by somatic cell reprogramming and are transgene independent.

(A) Episomal vectors and PB vectors were used to reprogram NSCs and HDFs respectively with a reprogramming induction period of 10-15 days. At this point, the media was changed to EPSCM for another 6-8 days before colony picking and clonal expansion (B) Top left panel shows a primary EPSC colony generated by episomal reprogramming of a human fetal neural stem cell line (BRC1019). Top right panel shows morphology of an established EPSC line from episomal reprogramming (C) qPCR expression profile showing no exogenous factor expression relative to wild type negative control (HDFa, NSC and H1-hESCs) in C5-EPSCs, iPSCs generated by PiggyBac transposition (PB-EPSCs) and as expected H1-EPSCs. Data are mean ± SEM.

A.



B.

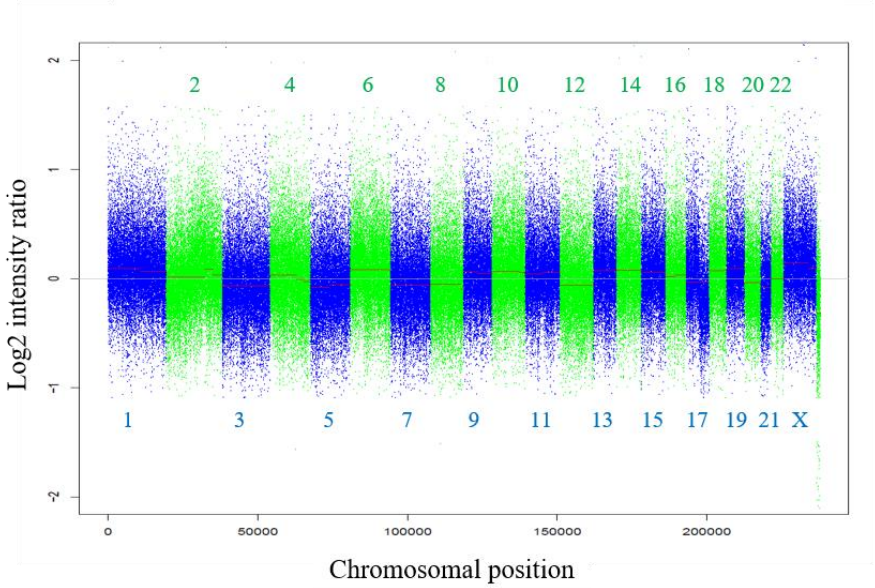
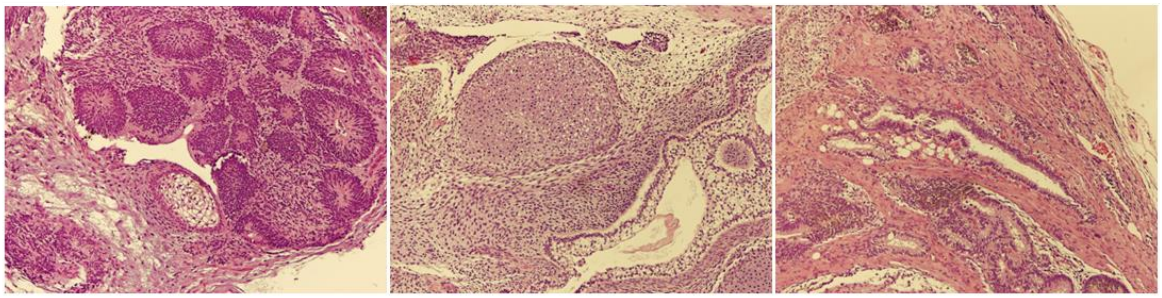


Figure 6.7. EPSCs generated by reprogramming somatic cells are genetically stable.

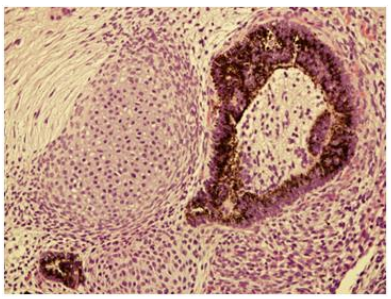
(A) Spectral karyotyping of EC5-EPSCs. 10 randomly selected metaphases were karyotyped by M-FISH and DAPI- banding. (B) Whole genome array CGH in EC5-EPSCs. Log2 intensity ratio on the y-axis is plotted against chromosomal position on the x-axis. No gain or loss of any chromosomal region is demonstrated.

A.



Ectoderm Mesoderm Endoderm

B.

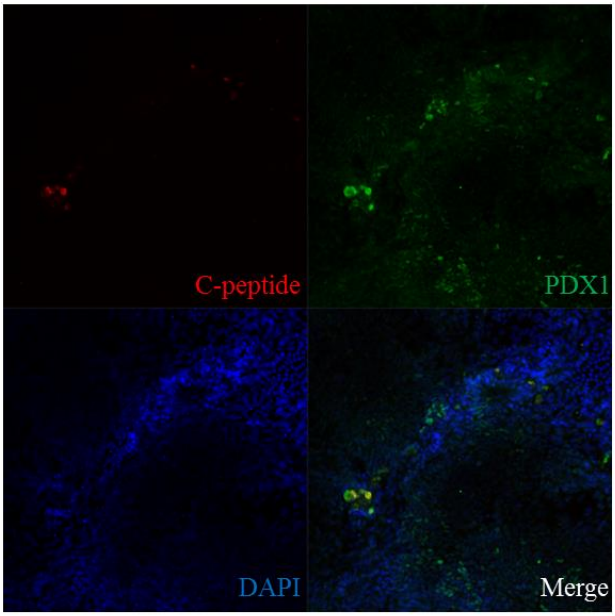


Melanin

Figure 6.8. In vivo differentiation of EC5-EPSC.

(A) EC5-EPSCs were injected into the dorsal flanks of NSG mice. Within 2 months teratomas became evident which contained derivatives of the three germ layers. Left panel:neural tube-like structures from the ectoderm (neuroectoderm); middle panel: cartilage from the mesoderm; right panel: a gland of possible gastrointestinal type arising from the endodermal layer. (B) melanin expression was detected indicating that the teratomas were mature.

A.



B.

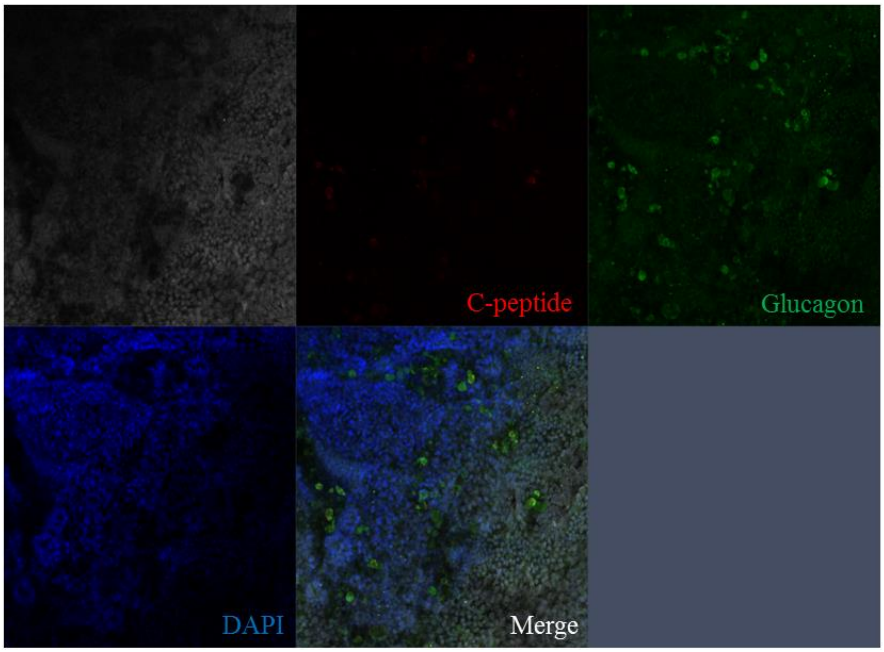


Figure 6.9. EC5-EPSC can be differentiated to pancreatic cell types in vitro.

Using a monolayer differentiation strategy, EC5-EPSCs were differentiated to pancreatic cell types (*PDX1*) including insulin secreting cells using C-peptide as a surrogate marker (A) and glucagon secreting alpha cells (B). (Collaboration with Mariya Chhatriwala, Vallier Lab).

teratomal elements **(Figure 6.8A and Figure 6.8B)**. In a collaborative effort with Ludovic Vallier's lab, EC5-EPSCs were differentiated to pancreatic cell types (PDX1) including insulin secreting C-peptide beta cells and glucagon secreting alpha cells using their published monolayer differentiation protocol **(Figure 6.9A and Figure 6.9B)** (Cho et al., 2012).

6.4.3 Human EPSCs can be genetically modified efficiently

In contrast to conventional human ESCs, human EPSCs were easy to culture and proliferated robustly **(Figure 6.2)**; therefore, we attempted to genetically modify the genome. These experiments were conducted in collaboration with Dr Wei Wang, Sanger Institute.

The ubiquitously expressed *ROSA26* was chosen as the first locus to target. A splice acceptor-H2B-Venus cassette was targeted to *ROSA26* using standard homologous recombination. The correct targeting efficiency was 47% which is comparable to that of targeting a similar cassette to the *Rosa26* locus in mouse ES cells **(Figure 6.10A)**. The CRISPR/Cas9 system has made genome editing much simpler (Doudna and Charpentier, 2014). *OCT4* and *NANOG* H2B-Venus and H2B-mCherry reporter human EPSC lines were generated using CRISPR/Cas9-facilitated targeting. The targeting efficiencies at these loci ranged from about 50% to 80% **(Figure 6.10B and 6.10C)**.

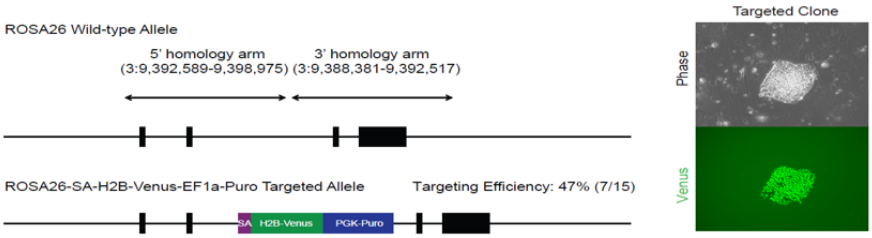
The availability of reporter cell lines in mouse ESCs has greatly facilitated studies on pluripotency and differentiation. As a proof of concept, a double targeted *OCT4* H2B-Venus, *SOX2* H2B-mCherry reporter line (EC5-EPSCs) was used for neural differentiation. EC5-EPSCs were rendered feeder-free and subsequently differentiated using a monolayer neural induction protocol. The reporter line allowed visualisation of the differentiation kinetics using epifluorescent microscopy and FACS **(Figure 6.11A and Figure 6.11B)**. During the 28 day induction, there was a progressive loss of *OCT4* expression and maintenance of *SOX2*

expression which was highly expressed at the neural rosette stage. The *SOX2* positive, *OCT4* negative population in the neural rosette culture was sorted and expanded in neural stem cell conditions containing EGF and FGF2. Bulk RNA sequencing confirmed the neural stem cell identity of the newly derived line **(Figure 6.11C)**. To functionally confirm the neural stem cell identity, C5-derived neural stem cells (C5-dNSCs) were terminally differentiated on laminin by removing the mitogens, FGF2 and EGF from the neural stem cell culture. After 3 months in culture, mature neuronal morphologies were evident. Immunocytochemistry confirmed the presence of mature neuronal markers, *NeuN* and *SYNAPTOPHYSIN* **(Figure 6.12A)**. In collaboration with Peter Kirwan at the Livesey Lab, Gurdon Institute, Cambridge, these neurons were shown to be mature with synaptic activity and action potential transmission identified on electrophysiology **(Figure 6.12B)**.

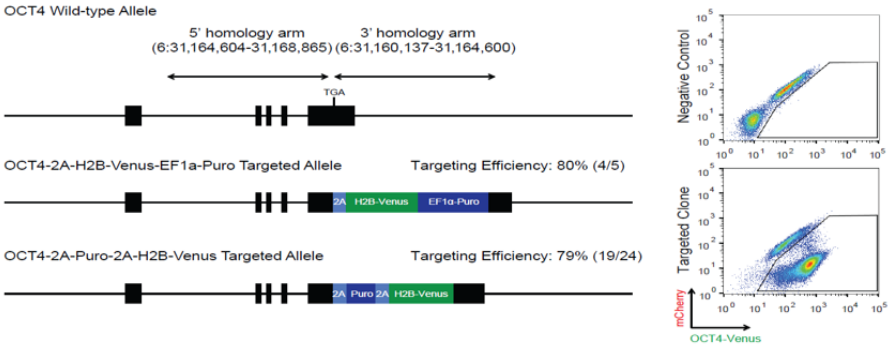
6.4.4 All components in EPSCM are essential for maintenance of human EPSCs

As in mouse EPSCs, the presence of a core pluripotency network is necessary for the maintenance of human EPSCs. If mouse EPSCs lose pluripotency gene expression and become differentiated then, as expected, the expanded potential phenotype is also lost. Therefore, I used the human *OCT4*-H2B-Venus reporter EPSCs as a convenient readout to evaluate the inhibitor requirement in maintaining human EPSCs. I found that all inhibitors in EPSCM are essential to prevent deterioration of human EPSC colony morphology and the *OCT4*-H2B-Venus positive cellular fraction in culture **(Figure 6.13).**

A.



B.



C.

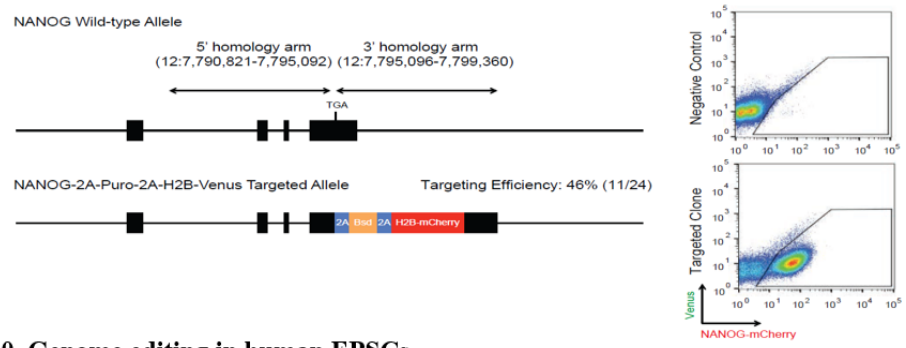


Figure 6.10. Genome editing in human EPSCs.

(A) The *ROSA26*-SA-H2B-Venus-EF1α-Puro knock-in allele. Left panel: A SA-H2B-Venus-EF1α-Puro cassette was targeted to the second intron of the *ROSA26* locus; right panel: a targeted clone show nuclear expression of Venus protein. (B) The *OCT4*-2A-H2B-Venus-EF1α-Puro (with promoter targeting) and *OCT4*-2A-Puro-2A-H2B-Venus (promoter-less) knock-in alleles. Left panel: A 2A-H2B-Venus-EF1α-Puro or a 2A-Puro-2A-H2B-Venus cassette was targeted to the last exon of the *OCT4* locus, stop codon was deleted to allow Venus to be expressed as a fusion protein with *OCT4* before the cleavage of 2A peptide separate them; right panel: a targeted clone show expression of Venus protein by FACS. (C) The *NANOG*-2A-Bsd-2A-H2B-mCherry knock-in allele. Left panel: A 2A-Bsd-2A-H2B-mCherry cassette (promoter-less) was targeted to the last exon of the *NANOG* locus, stop codon was deleted to allow mCherry to be expressed as a fusion protein with *NANOG* before the cleavage of 2A peptide separate them; right panel: a targeted clone show expression of mCherry protein by FACS. Collaboration with Dr Wei Wang.

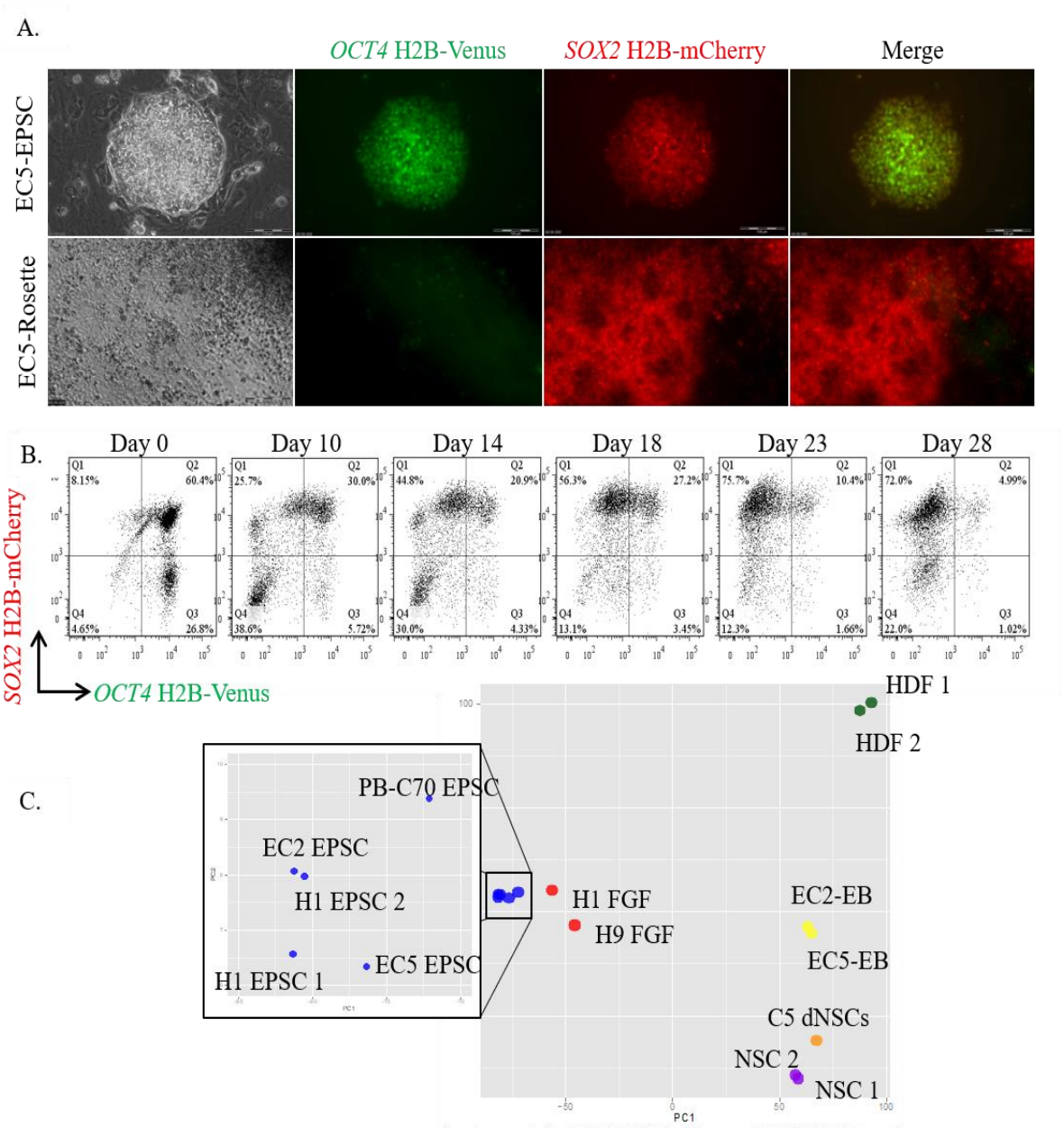
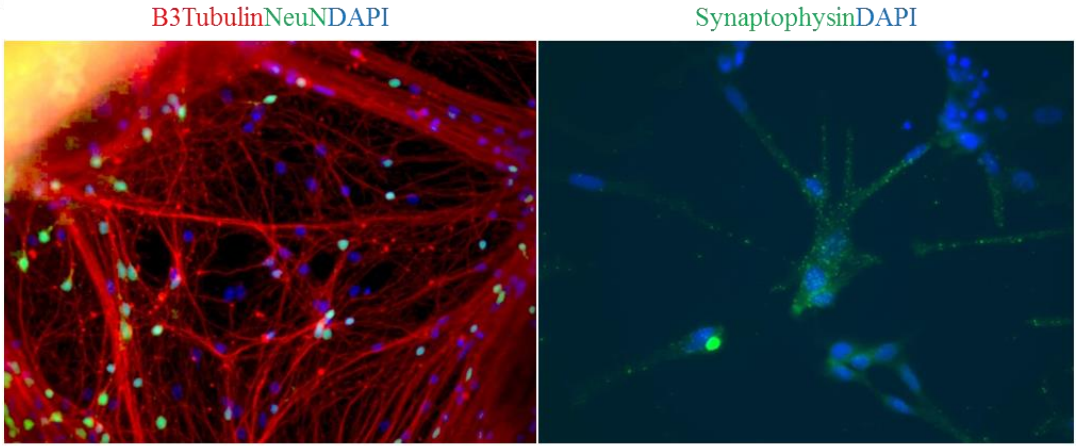


Figure 6.11. Neural differentiation of double-targeted *OCT4* H2B-Venus *SOX2* H2B-mCherry EPSCs.

(A) As expected, EC5-EPSCs express *OCT4* and *SOX2*. During the 28 day neural induction, there was a progressive loss of *OCT4* expression and maintenance of *SOX2* expression which was highly expressed at the neural rosette stage at day 26. (B) FACS showing neural differentiation kinetics. (C) PCA analysis demonstrating the homogenous and distinct global transcriptional profile of multiple EPSC subclones (EC2, PB-C70, EC5 and H1 EPSCs) relative to human ESCs (H1 and H9 FGF) and that C5 derived neural stem cells (C5 dNSCs) cluster with the parental NSCs (NSC1 and NSC2). Embryoid bodies differentiated from EPSCs (EC2 and EC5-EBs) and human dermal fibroblasts (HDF1 and 2) were also included in the analysis.

A.



B.

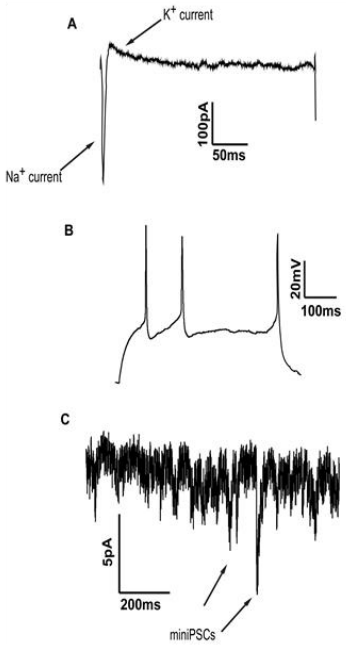


Figure 6.12. Terminal differentiation of EPSC derived neural stem cells.

(A) C5-derived neural stem cells (C5-dNSCs) were terminally differentiated on laminin in N2B27 following the removal of the mitogens EGF and FGF. After 3 months in culture, mature neuronal morphologies were evident which stained positively for the mature neuronal markers, *NeuN* (Top left panel, 200X)and *synaptophysin* (Top right panel, 400X). (B) Electrophysiological competence was seen with spontaneous sodium and potassium currents, action potential generation and miniPSCs indicating synaptic activity. Thanks to Peter Kirwan at the Livesey Lab for electrophysiological assistance.

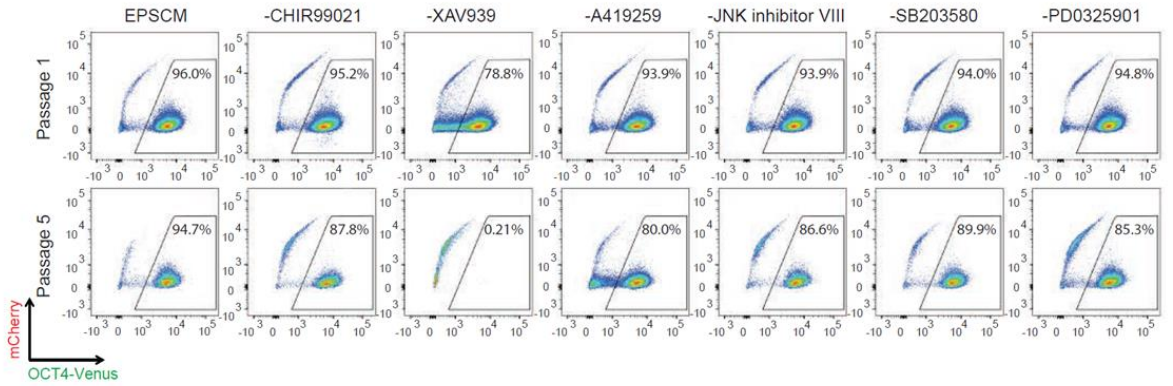


Figure 6.13. All six inhibitors are required for maintenance of human EPSCs.

(A) Signalling dependence of EPSCs. The *Oct4*-2A-Puro-2A-H2B-Venus reporter EPSCs were passaged in EPSCM or EPSCM minus each of the 6 inhibitors for 5 passages. FACS was used to determine the stability of EPSCs in different medium. As in mouse EPSCs, removal of individual inhibitors of WNT and Src signalling produced the most dramatic phenotype with marked reduction in the Oct4 positive population but, all inhibitors are required, albeit perhaps to a lesser extent, to maintain the integrity of the culture over extended passaging.

6.4.5 Transcriptomic analysis of newly derived human EPSCs

These experiments were conducted in collaboration with Dr Cheuk-Ho J. Tsang, Sanger Institute. Population RNA-sequencing was performed to compare the transcriptome of human EPSCs and human ESCs. Hierarchical clustering showed that the transcriptome of H1-EPSCs segregated from H1 and H9 human ESCs **(Figure 6.14)**. Consistent with findings on quantitative PCR, human EPSCs showed similar expression of core pluripotency factors such as *POU5F1 (OCT4)*, *SOX2* and *NANOG* as in human ESCs **(Figure 6.15A)**. Interestingly, multiple genes associated with mouse ESC pluripotency such as *KLF4*, *UTF1* and *TFCP2L1* were significantly up-regulated in H1-EPSCs **(Figure. 6.15A)**. In addition, genes encoding key lineage regulators such as *EOMES*, *GATA4*, *GATA6, T*, *SOX17*, *SOX1* and *RUNX1* were all significantly down-regulated **(Figure. 6.15B)**.

Recently, there has been significant momentum to derive mouse ESC-like human pluripotent stem cells. Several groups have reported multiple *in vitro* conditions to maintain human ESCs with transcriptional and metabolic similarities to mouse ESCs in 2i/LIF conditions (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014). As mouse EPSCs possess a core pluripotency transcriptional network that has many similarities with the mouse ground-state of self-renewal, I sought to investigate if human EPSCs clustered with the proposed human naïve state.

Population RNA-sequencing data were retrieved from respective studies and re-analysed. On three-dimensional principal component analysis and hierarchical clustering, the transcriptome in EPSCM was found to be distinct from standard FGF human ESCs and all proposed human naïve states **(Figure 6.16 and Figure 6.17)**. Therefore, similar to mouse EPSCs, human EPSCs possessed a unique transcriptome that had an active pluripotency transcriptional

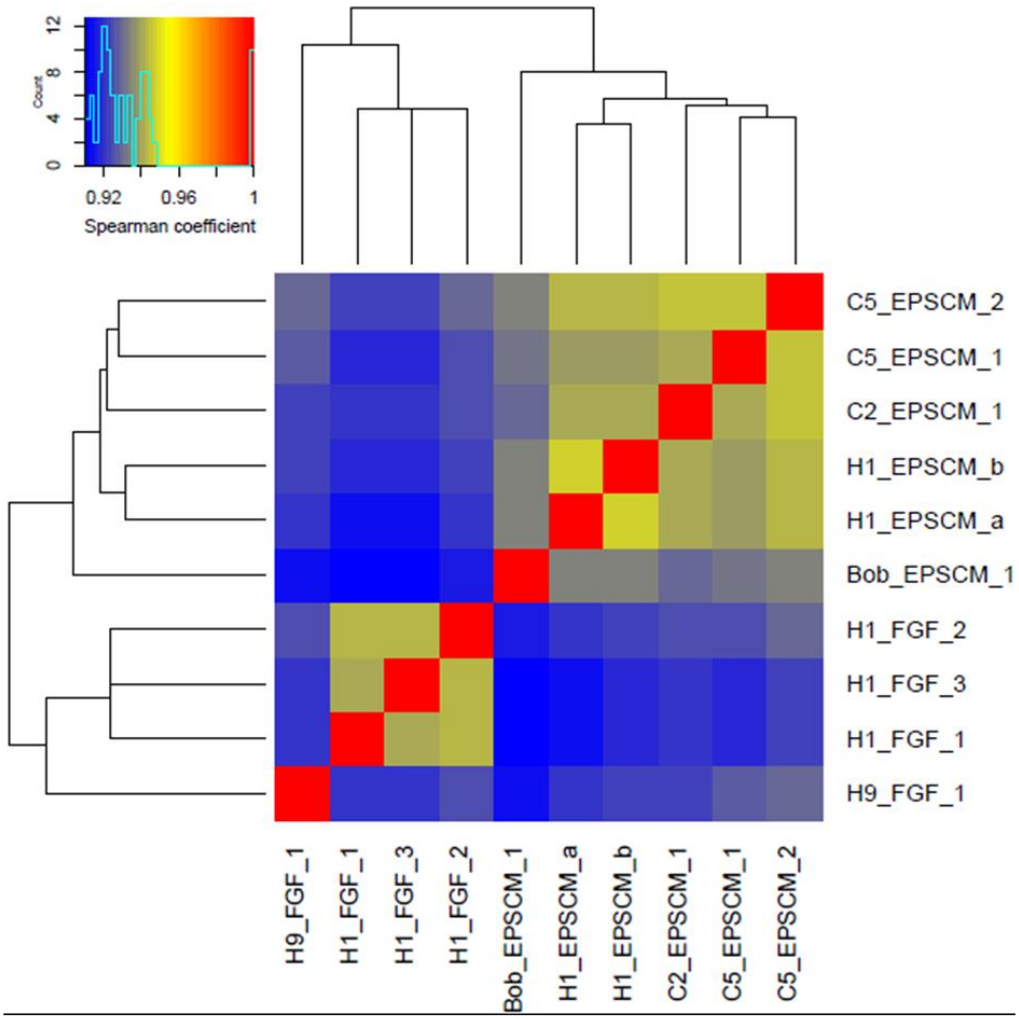
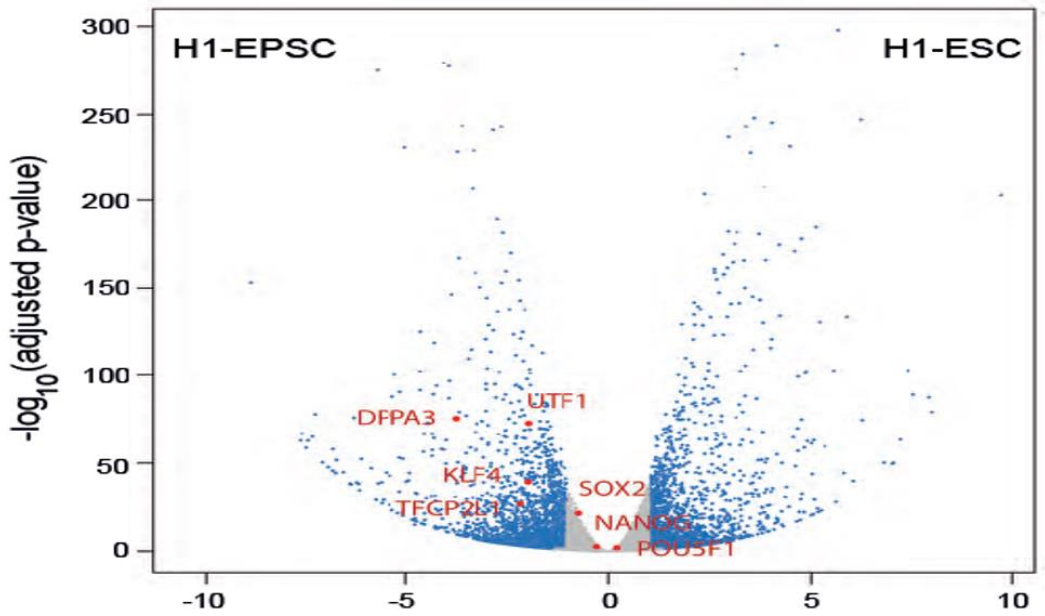


Figure 6.14. Transcriptome of human EPSCs is distinct from conventional human ESCs.

Heatmap of correlations and hierarchical clustering of the transcriptomes of ESCs (H1 FGF and H9 FGF) and EPSCs (C5 EPSCM, C2 EPSCM, H1 EPSCM and BOB EPSCM). Correlations were calculated by Spearman correlation coefficient method. Genes with expression of normalized count less than 100 were removed. Clustering was done with complete linkage. Bioinformatic analysis was in collaboration with Cheuk-Ho J. Tsang.

A.



B.

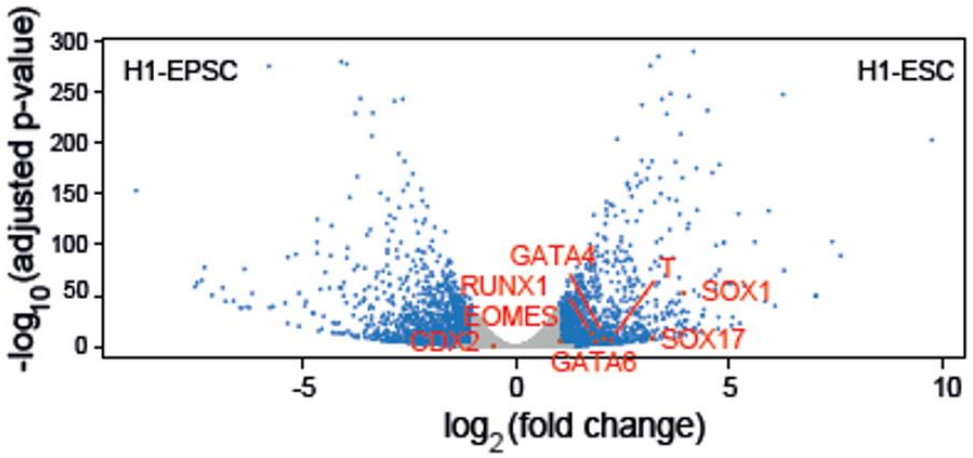


Figure 6.15. Transcriptomic analysis of newly derived human EPSCs.

(A) and (B) Volcano plots showing the relationship of the relative expression fold change and statistical p-values of individual genes between H1 human ESCs in EPSCM and bFGF conditions. The p-values and fold changes were computed by DESeq2. Genes with adjusted p-values <0.05 and fold changes greater than 1 were highlighted in blue. Individual pluripotency genes were labelled. Genes with adjusted p-values <0.05 and fold change greater than 1.0 were highlighted in blue. Pluripotency-related genes and lineage-specific genes were labelled in red. Bioinformatic analysis was performed in collaboration with Cheuk-Ho J. Tsang.

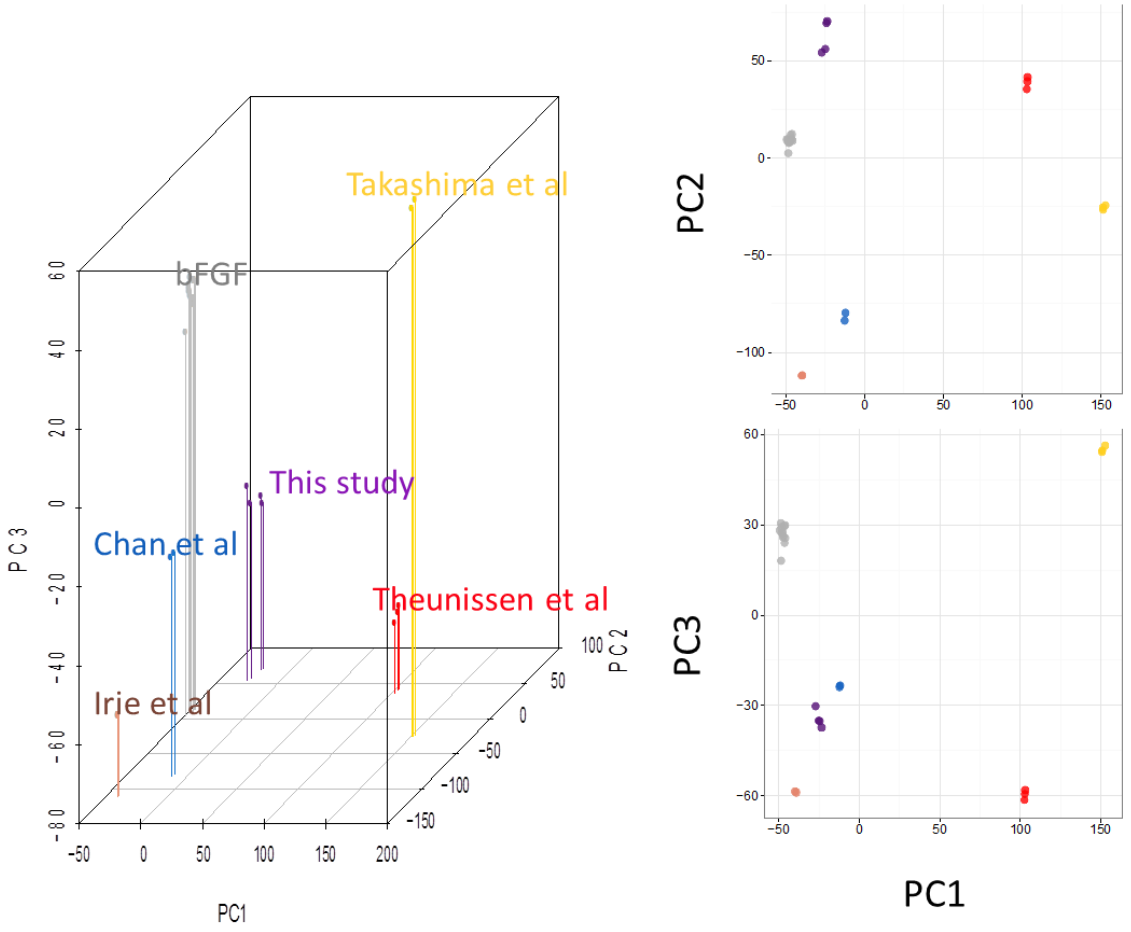


Figure 6.16. Human EPSCs have a unique transcriptome.

Three-dimensional visualization of PCA on the transcriptomes of human pluripotent stem cells (PSCs) maintained in different conditions. Population RNA-seq data were retrieved from respective studies and re-analyzed. The expression matrix is quantified as the relative expression of normalized count to the mean expression of that of hPSCs cultured in bFGF condition from the same study. The expression matrix was then transformed by log2(normalized count +1) and genes with relative expression change less than 1 were removed. Bioinformatic analysis was performed in collaboration with Cheuk-Ho J. Tsang.

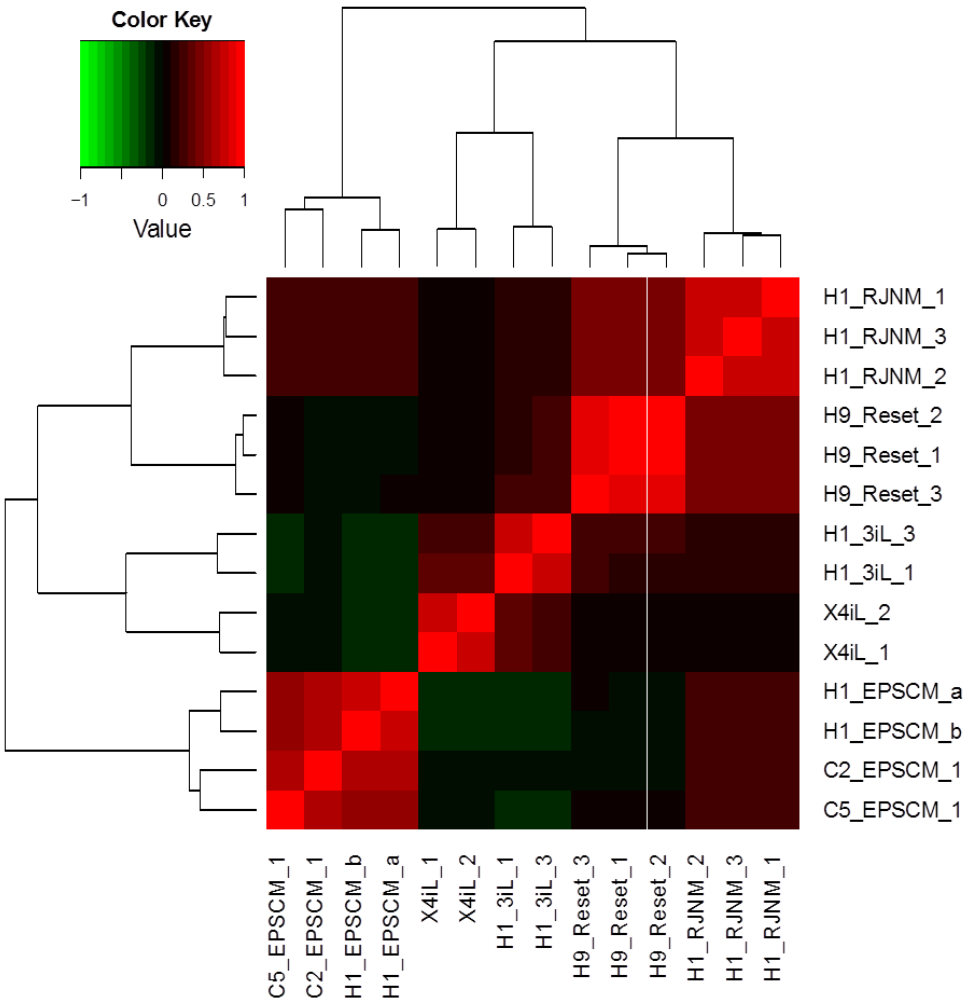


Figure 6.17. The transcriptome of human EPSCs is distinct from the proposed human naïve state

Heatmap of correlations and hierarchical clustering of the transcriptomes of EPSCs and published 'naïve' human cell lines. Correlations were calculated by Spearman correlation coefficient method. Genes with expression of normalized count less than 100 were removed. Clustering was done with complete linkage.. Bioinformatic analysis was in collaboration with Cheuk-Ho J. Tsang.

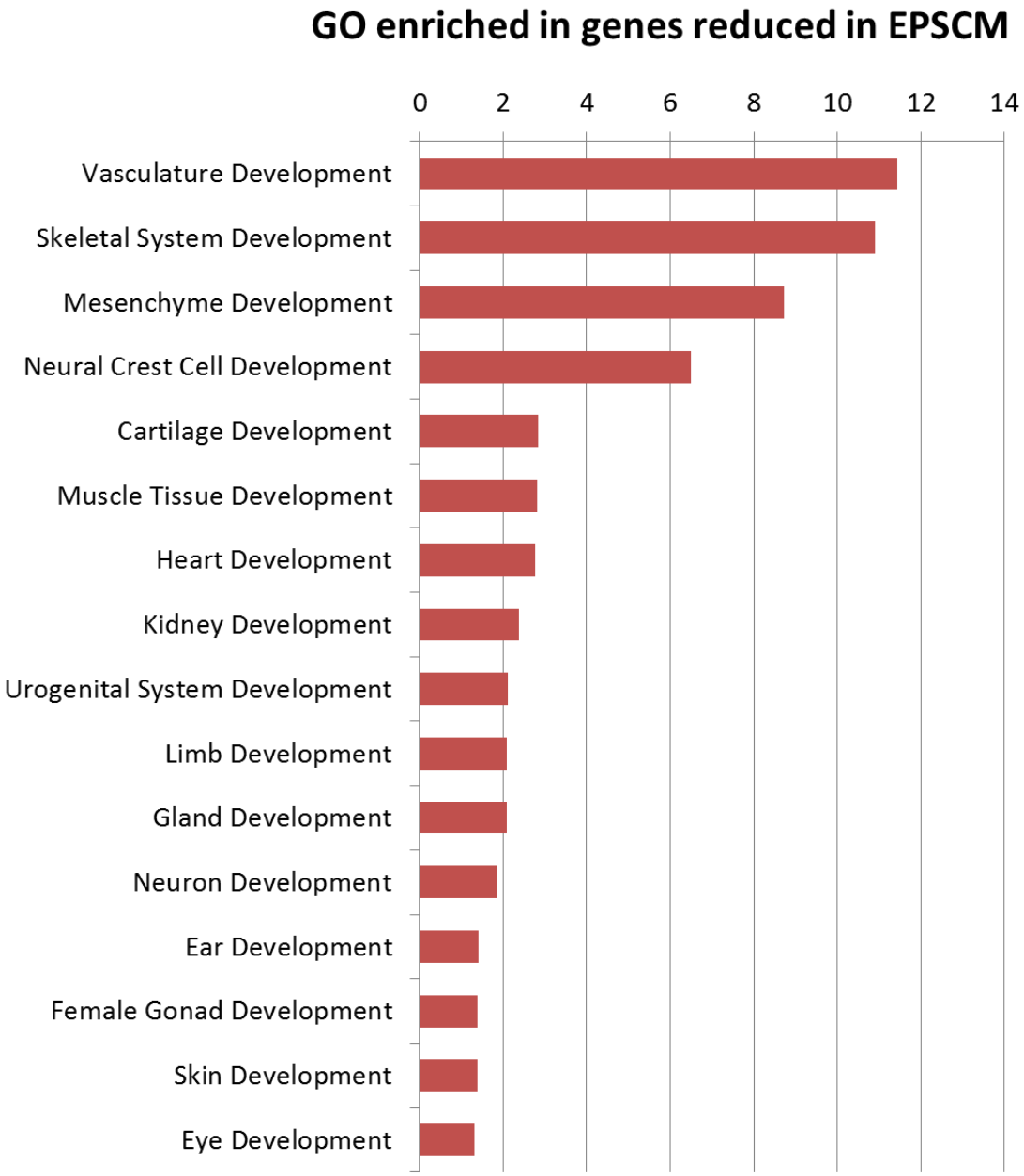


Figure 6.18. H1-EPSCs gene ontology analysis.

Gene ontology terms enrichment analysis confirmed that the down-regulated genes in H1-EPSCs are enriched with somatic lineage development. Bioinformatic analysis was in collaboration with Cheuk-Ho J. Tsang.

network as a dominant finding. Gene ontology analysis confirmed that the down-regulated genes in H1-EPSCs were enriched with terms related to somatic lineage development **(Figure 6.18)**.

6.4.6 Epigenetic status of X-chromosome suggests that human EPSCs may have similarity to the pre-implantation embryo.

Using single cell RNA FISH, female EPSCs (<10 passages) were found to have lost expression of the repressive X-inactive specific transcript (*XIST*) and had gained biallelic expression of the long non-coding RNA, *XACT* which is known to coat the entire active X chromosome **(Figure 6.19-1A,B and C)** (Vallot et al., 2015). These findings suggested that EPSCs had two active X-chromosomes which I further confirmed by assessing the expression of individual X-linked genes as, if both X-chromosomes were active then expression should be bi-allelic. Indeed, single cell RNA FISH confirmed bi-allelic expression of X-linked genes in EC5-EPSCs **(Figure 6.20-1A and B).** In collaboration with Dr Cheuk-Ho J. Tsang, Sanger Institute, bulk RNA-sequencing was performed which further supported this finding by revealing global increased X-linked gene expression with associated down-regulation of *XIST* in female EPSCs as compared to the parental NSC **(Figure 6.19-2D)**. As XCI begins at the eight-cell stage in humans at which time *XIST* starts to become detectable, the absence of *XIST* and presence of two active X-chromosomes tentatively may indicate that human EPSCs have epigenetic similarity with the eight-cell stage of human embryo development.

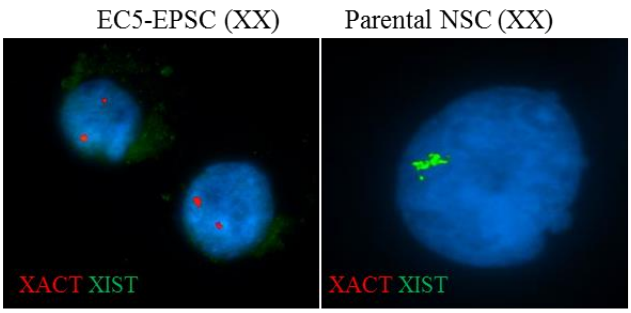
If the loss of *XIST* expression and the presence of two active X-chromosomes indicated that EPSCs were corresponding to a stage of pre-implantation development, then, X-chromosome inactivation and dosage compensation should be observed during differentiation. To assess whether EPSCs underwent this normal developmental process, differentiation was performed using *in vitro* embryoid body differentiation and targeted neural differentiation as previously

described. Embryoid bodies were made from EC5-EPSCs and were allowed to differentiate for 14 days before testing for X-chromosome inactivation. Using single-cell RNA FISH, differentiated EC5 was found to have inactivated one X-chromosome with re-expression of X-inactive specific transcript (*XIST*) and had lost bi-allelic expression of the X-linked gene, *HPRT1* **(Figure 6.21-1A,B and C)**. RNA FISH also confirmed that EC5-EPSCs lost expression of *XACT* during differentiation. Interestingly, 3% of differentiated EC5 cells (C5 dEB Day 14) still had detectable *XACT* expression which is most likely explained by some residual undifferentiated cells in the embryoid body culture as *XACT* expression is pluripotency specific **(Figure 6.21-2D)** (Vallot et al., 2015). Further supporting that EPSCs underwent bona fide X-chromosome inactivation, I detected evidence for dosage compensation with down-regulation of X-linked genes at the pan-X chromosome level during the differentiation of EC5-EPSCs to embryoid bodies (EC5 dEB) and neural stem cells (EC5 dNSC) **(Figure 6.21-2E)**.

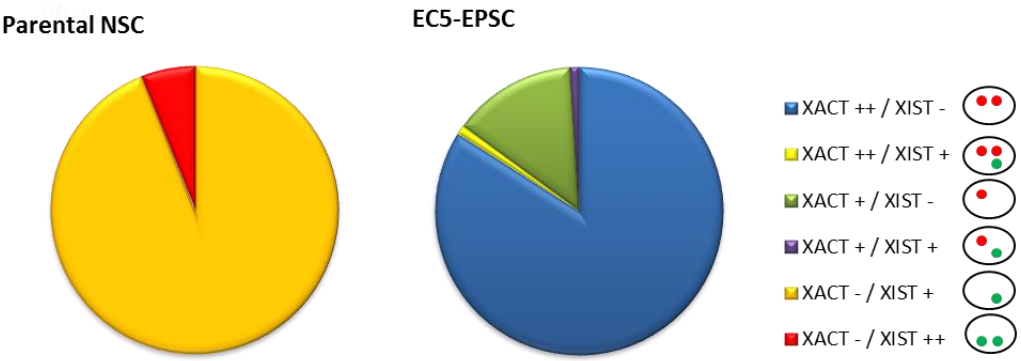
6.4.7 Human EPSCs have a pre-implantation transcriptional signature

In Chapter 4, the transcriptome of mouse EPSCs was shown to be enriched with an early pre-implantation blastomere signature. Therefore, the analysis was repeated on the converted human EPSCs to determine if these cells also possessed a unique signature. In collaboration with Cheuk-Ho J. Tsang, the top stage-specific genes of each pre-implantation stage were compiled and a gene set enrichment analysis (GSEA) was performed. Although a crude and provisional representation, this analysis demonstrated that in addition to possessing a pluripotent signature, close to that of the native epiblast, converted H1-EPSCs also showed

A.



B.



C.

Frequency Observed (%)	Parental NSC (XX) (n=100 cells)	EC5-EPSC (XX) (n=100 cells)
XACT ++ / XIST -	0	84
XACT ++ / XIST +	0	1
XACT + / XIST -	0	14
XACT + / XIST +	0	1
XACT - / XIST +	94	0
XACT - / XIST ++	6	0

Fig 6.19-1. Pan-X chromosome re-activation in female human EPSCs

D.

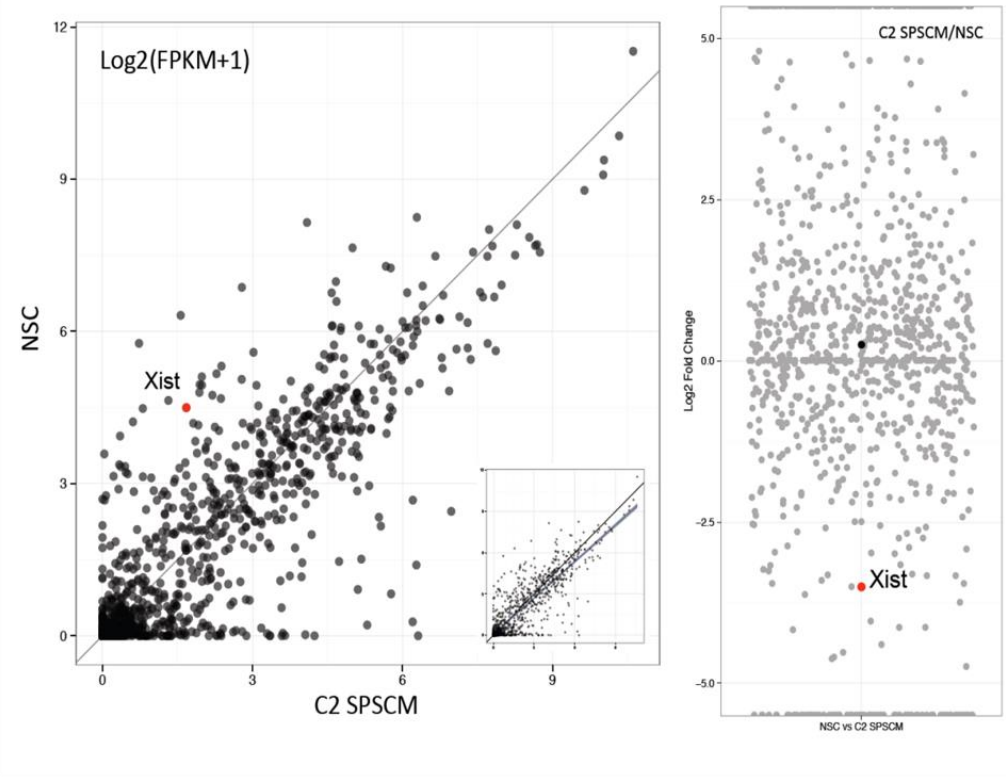


Figure 6.19-1 and 2. Pan-X chromosome re-activation in female human EPSCs

(A) Using single cell RNA FISH, female EPSCs were found to have X chromosome reactivation with loss of the repressive X-inactive specific transcript (XIST) and expression of the long noncoding RNA, XACT which coats the entire active X chromosome in pluripotent cells. In the top left panel, there are two XACT signals reflecting, two active X chromosomes. (B) and (C) show a visual representation and quantification of the FISH signals detected in female EC5-EPSCs and the parental female neural stem cell (NSC). N=100 cells were counted in each condition. (D) Further supporting the RNA-FISH data in (A) – (C), bulk RNA-seq showed increased X-linked gene expression with associated down-regulation of XIST in female EPSCs (C2 SPSCM) as compared to the parental NSC. Bioinformatic analysis was performed in collaboration with Cheuk-Ho J. Tsang.

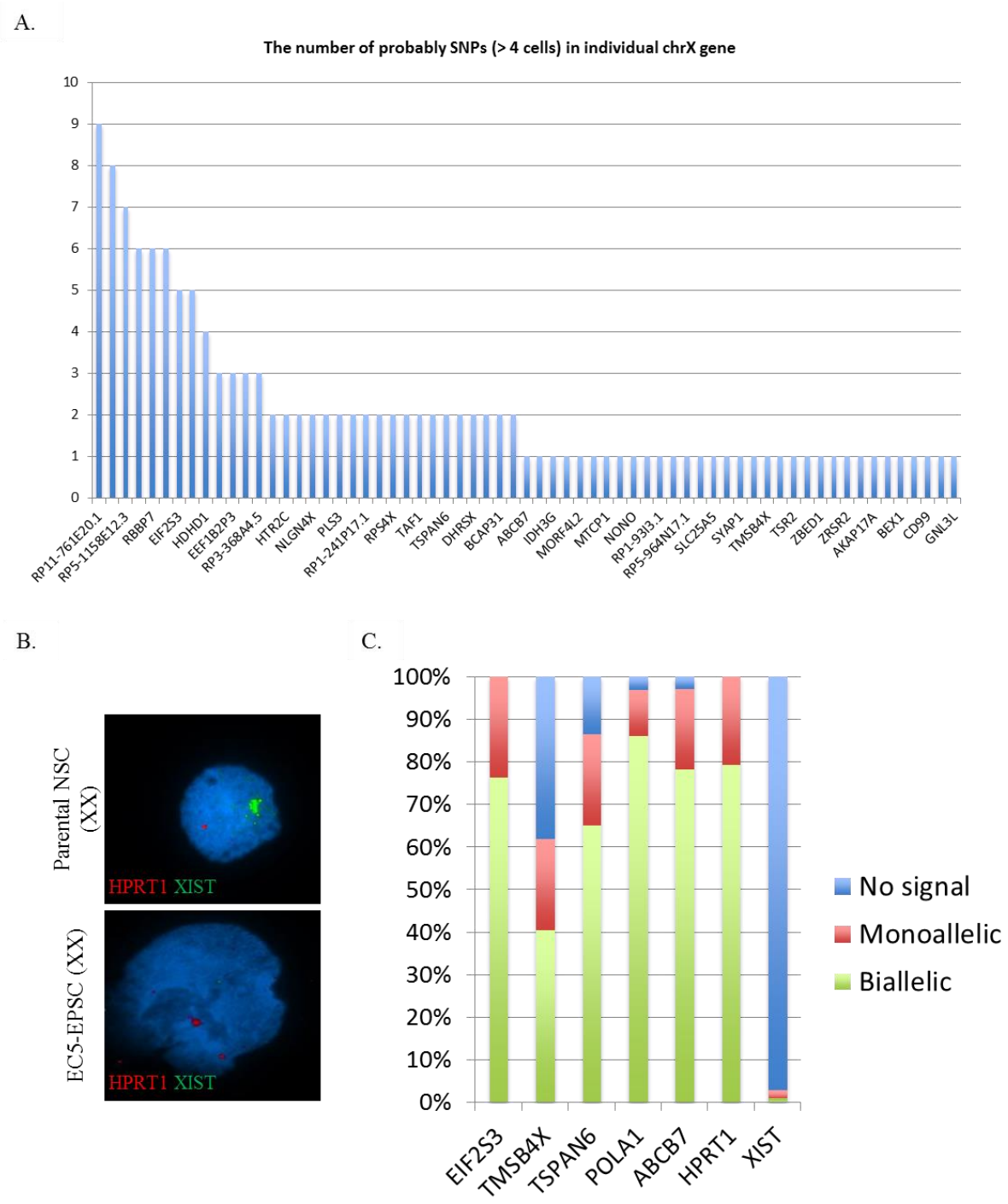
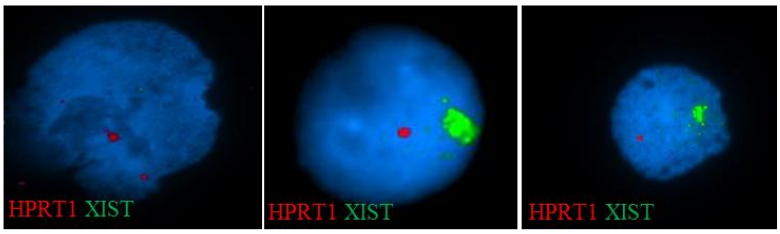


Figure 6.20-1. X-linked genes show bi-allelic expression in female human EPSCs

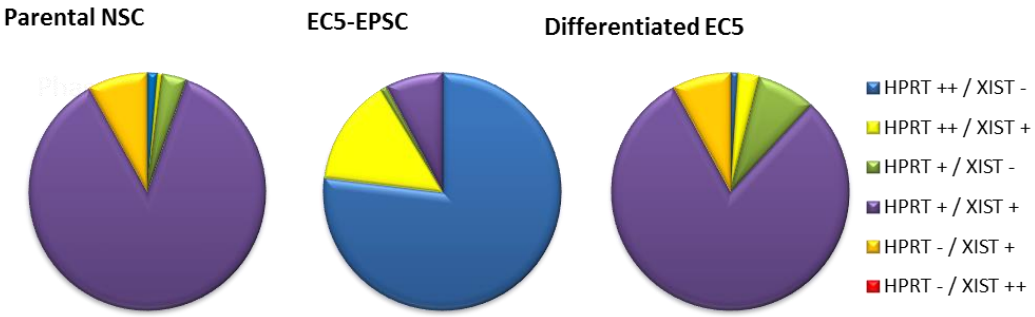
Figure 6.20-1 and 2. X-linked genes show bi-allelic expression in female human EPSCs

The preceding single cell RNA FISH data showing XACT expression from both X chromsomes and loss of the repressive XIST cloud suggested that both X-chromosomes were active. To further confirm this we looked at the expression of individual X-linked genes. If both X chromosomes were active then the expression should become bi-allelic. Single-cell RNA-seq data of 96 human EPSCs (EC5-ROSA26-Venus) were generated using Fluidigm C1 AutoPrep system according to manufacturer instruction using medium sized integrated fluidic circuits (10-14um). 100bp paired-ends reads were generated on HiSeq2500 according to in-house Sanger DNA pipeline. Reads were aligned to GRCh37 human genome using STAR aligner with default parameters. In total, 44 cells passed quality control (including removal of lysed cells, doublets, unique aligned reads <500,000, mitochondrial reads >10%, gene detection <8000). Samtools mpileup and Vcftools were used to identify variant sites at the X chromosome. Heterozygous SNP sites were then annotated by intersecting their coordinates with that of annotated X chromosome genes. A SNP is only considered valid if it can be detected in more than 4 cells. The number of probable SNPs in each X chromosome-linked gene were then summarized and selected for mRNA-FISH validation. (A) Number of probable SNPs identified in X chromosome genes. Probable SNPs are defined as SNPs that can be detected in the single-cell RNA-seq data in more than 4 cells. B) Detection of allelic expression dosage of HPRT1 and XIST in parental neural stem cells (NSCs) and reprogrammed human EPSCs by mRNA-FISH. C) Validation of additional X-linked genes identified in (A) using single cell RNA FISH which confirmed bi-allelic expression in EC5-EPSCs providing further evidence that there is re-activation of the inactive X chromosome in human EPSCs. N=100 cells were counted per probe. Bioinformatic analysis was performed in collaboration with Cheuk-Ho J. Tsang.

A.



B.



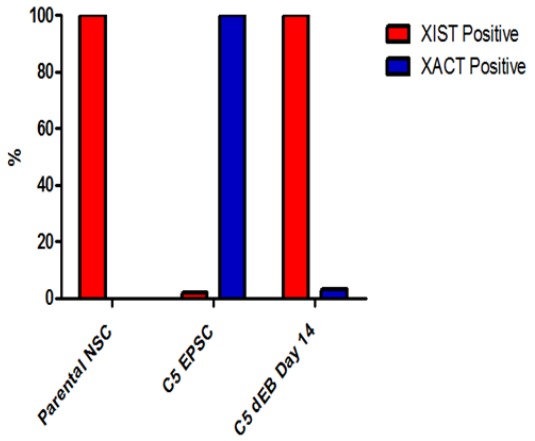
C.

Frequency Observed (%)

	Parental NSC (n=148)	EC5-EPSC (n=104)	Differentiated EC5 (n=101)
HPRT ++ / XIST -	1.4	76.9	1.0
HPRT ++ / XIST +	0.7	14.4	3.0
HPRT + / XIST -	3.4	1.0	7.9
HPRT + / XIST +	86.5	7.7	80.2
HPRT - / XIST +	8.1	0.0	7.9
HPRT - / XIST ++	0.0	0.0	0.0

Figure 6.21-1 Reversible X chromosome inactivation and evidence for dosage compensation during differentiation of female human EPSCs.

D.



E.

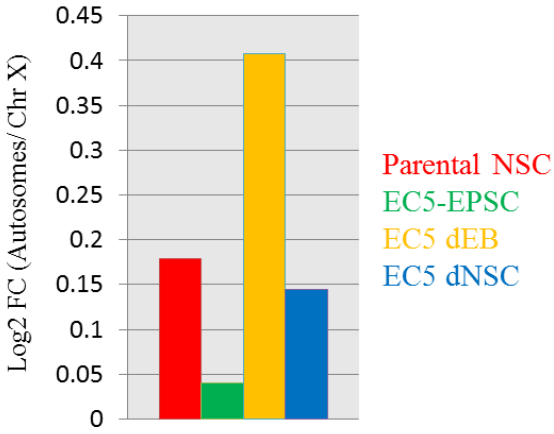


Figure 6.21-1 and 2. Reversible X chromosome inactivation and evidence for dosage compensation during differentiation of female human EPSCs.

(A)Embryoid bodies made from EC5-EPSCs were allowed to differentiate for 14 days before testing for X chromosome inactivation. Using single cell RNA FISH, differentiated EC5 was found to have inactivated one X chromosome with re-expression of X-inactive specific transcript (*XIST*) and loss of bi-allelic expression of the X linked gene, *Hprt1*. (B) and (C) show a visual representation and quantification of the FISH signals detected in female EC5-EPSCs, the parental female neural stem cell (NSCs) and differentiated EC5. (D) RNA FISH confirmed that C5-EPSCs lost expression of *XACT* during differentiation. 3% of cells in differentiated EC5 cells (C5 dEB Day 14) still had detectable XACT expression which most is most likely explained by some residual undifferentiated cells in the embryoid body culture as *XACT* expression is pluripotency specific. (E) Further supporting our RNA FISH data suggesting X chromosome inactivation, we detected dosage compensation with down-regulation of X linked genes at the pan-X chromosome level during differentiation of EC5-EPSCs to embryoid bodies (EC5 dEB) and neural stem cells (EC5 dNSC). This was produced by calculating the Log2 fold change of the median autosomal gene expression divided by the median X linked gene expression.

enrichment for genes that are expressed in the morula and eight-cell stage of development **(Figure 6.22)**.

These findings suggested that human EPSCs were distinct from standard human ESCs and had some transcriptional and epigenetic characteristics that are known to be associated with a pre-implantation stage of human development. If this were true, then these cells in addition to being able to differentiate to somatic lineages as seen previously, should also be able to produce bona fide trophoblast. Since mouse EPSCs have expanded potential to placental trophoblasts, it was logical to next examine the trophoblast differentiation potential of human EPSCs.

6.4.8 Trophoblast induction potential of human EPSCs

These experiments were conducted in collaboration with Dr Wei Wang, Sanger Institute. In the first trimester of human placental trophoblast development, *ELF5* is expressed in villous cytotrophoblast and the *ELF5* locus is demethylated. Furthermore, ELF5, in conjunction with CDX2, is thought to represent a trophoblast stem cell population within the placental villous cytotrophoblast layer (Hemberger et al., 2010). It is known that human ESCs have an intriguing property to differentiate to cells that express many trophoblast genes in the presence of BMP4 (Amita et al., 2013; Xu et al., 2002). However, in these trophoblast-like cells from human ESCs, *ELF5* was not found to be highly expressed and the locus was still substantially methylated, which is different from the placental trophoblasts of early developmental stages (Bernardo et al., 2011; Hemberger et al., 2010).

To directly compare H1 ESCs and EPSCs (H1 and EC5) for their trophoblast potential, they were differentiated using published protocols (BMP4 plus a TGFB or NODAL inhibitor)

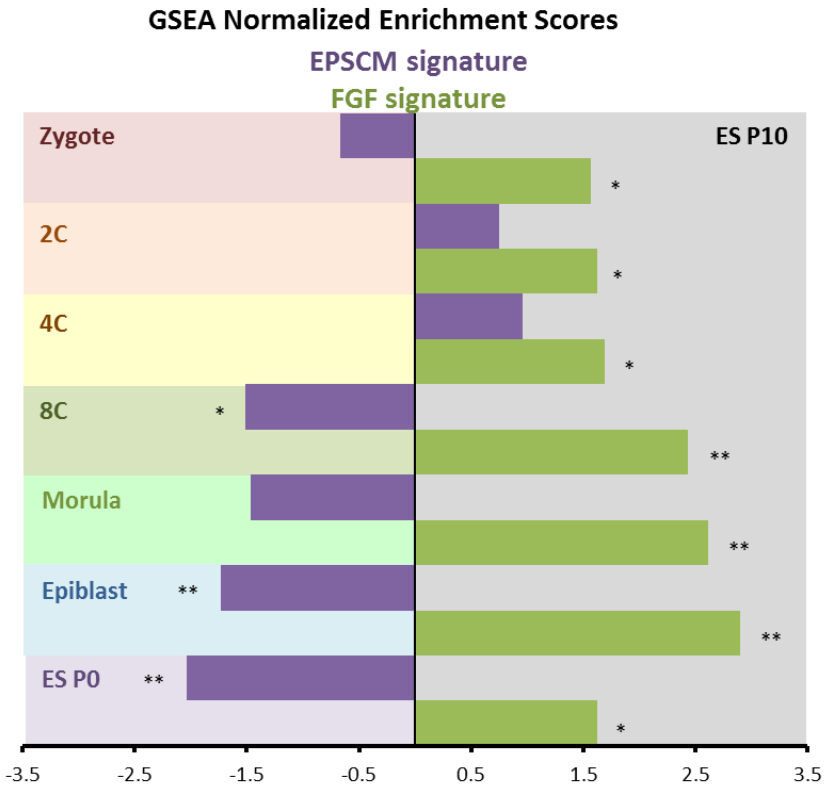


Figure 6.22. Pre-implantation transcriptomic signature in H1-EPSCs.

Gene set enrichment analysis showing EPSCM/FGF signature gene expression enrichment in different stages of development of preimplantation human embryos comparing to standard human ESC in FGF condition (passage 10).

The signature genes were defined as the top 1000 mostly expressed differential genes between PSCs (EC2, EC5, H1) cultured in EPSCM and H1 ESC cultured in FGF conditions. Differential gene analysis was performed using Cufflinks 2. The enrichment of the signature genes of the two culture conditions were tested in individual preimplantation stage against passage 10 human ESC cultured in FGF conditions. The human embryo single-cell RNA-seq data were retrieved from Yan et al., 2013. Positive end of the x-axis indicates the level of enrichment in passage 10 FGF human ESCs in each pairwise comparison, the negative end indicate enrichment towards the respective embryonic stage. For instance, human 8C embryos showed significant enrichment of EPSCM signature (score: -1.51), compared to passage 10 FGF human ESCs, while the FGF signature is enriched towards p10 FGF ESCs as expected (score 2.44). *$p<0.05$, **$p<0.01$

(Amita et al., 2013; Bernardo et al., 2011). The temporal expression of *CDX2* and *ELF5* was monitored. Expression of *CDX2* was quickly up-regulated and peaked at day 4 **(Figure 6.23)**, which was confirmed by immuno-staining of CDX2 protein **(Figure 6.24)**. The peaking of *CDX2* expression was followed by a robust up-regulation of *ELF5* in differentiated EPSCs **(Figure 6.23)**. In both H1-EPSCs and EC5-EPSCs, the levels of *ELF5* were at least 10-fold higher than those achieved in differentiated H1-ESCs.

As was described above, the *ELF5* locus is demethylated in human early placental trophoblast, but is almost fully methylated in human ESCs, human iPSCs, and in human ESC-derived trophoblast-like cells (Hemberger et al., 2010). Concordantly, the *ELF5* locus was confirmed to be heavily methylated in H1-ESCs in our study, however; human EPSCs already had lower DNA methylation **(Figure 6.25A)**. Moreover, once EPSCs differentiated toward trophoblast lineages, DNA methylation at the *ELF5* locus was more drastically decreased as compared to cells differentiated from H1-ESCs **(Figure 6.25A)**. Similarly, the *ELF5* locus was also demethylated in cells differentiated from EC5-EPSCs **(Figure 6.25B)** under the same trophoblast differentiation conditions.

Besides these two key trophoblast transcription factors *CDX2* and *ELF5*, the expression of other trophoblast markers was also monitored using immuno-staining **(Figure. 6.26)** and qRT-PCR **(Figure 6.23)**. Following sequential activation of *CDX2* and *ELF5*, differentiating EPSCs later expressed high levels of *CYTOKERATIN 7* (*CK7*) and *HLA-G* **(Figure 6.23, 6.24, 6.26 and 6.27)**. FACS analysis demonstrated high expression of FGFR4, INTEGRIN a5 (ITGA5) and permitted quantitative analysis of CYTOKERATIN 7 expression **(Figure 6.27)**. The expression of chorionic gonadotropin (*CGα* and *CGβ*) and placenta growth factor (*PGF*) was also up-regulated to levels that were substantially higher than that of the differentiated H1-ESCs **(Figure 6.23)**. Collectively, these results demonstrate that human EPSCs have distinct *in vitro* differentiation potential towards trophoblasts.

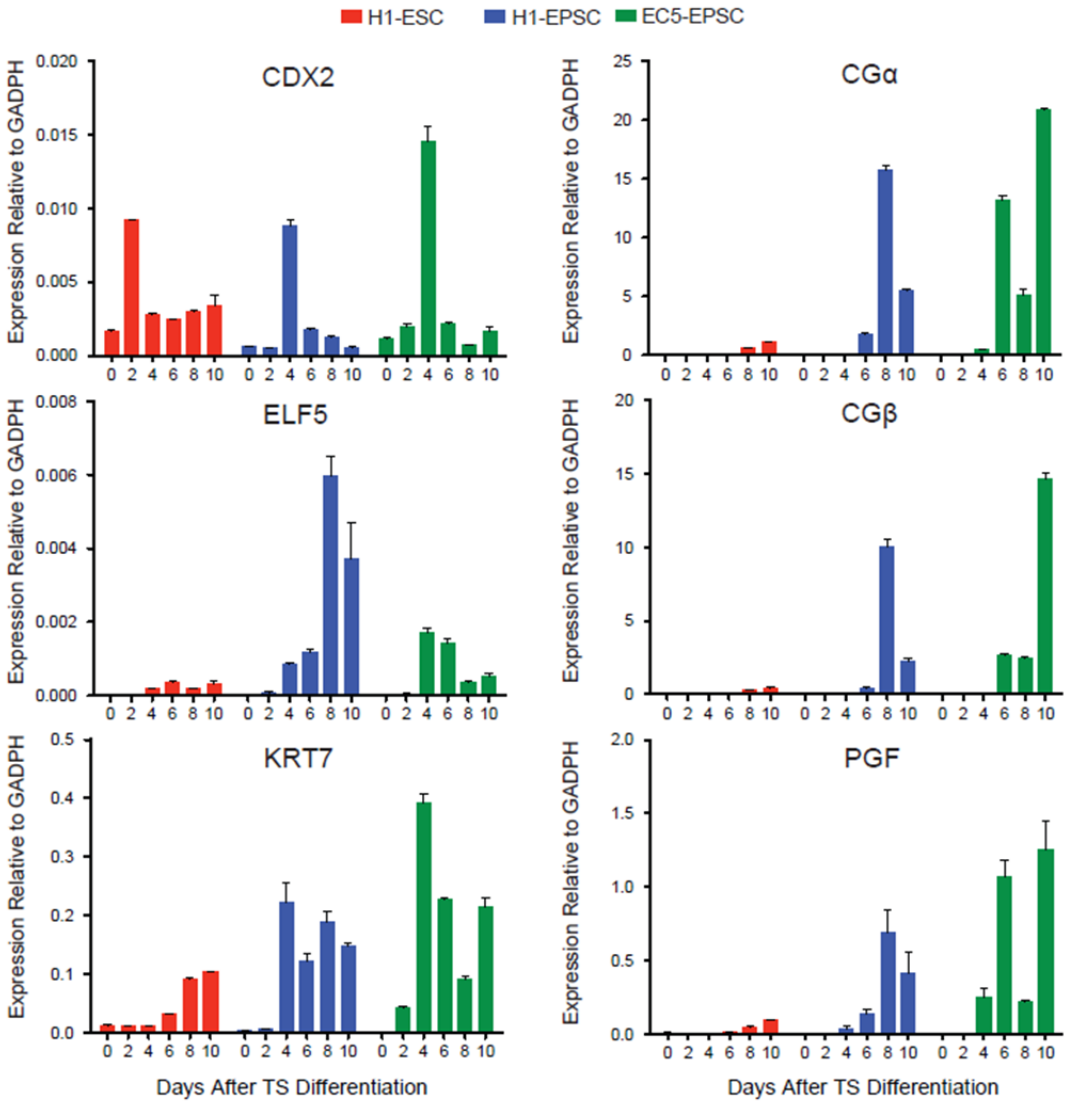


Figure 6.23. Temporal expression of trophoblast lineage specific markers during in vitro differentiation of human EPSCs.

Expression kinetics of trophoblast lineage specific markers in H1-ESCs, H1-EPSCs and EC5-EPSCs during in vitro trophoblast induction.

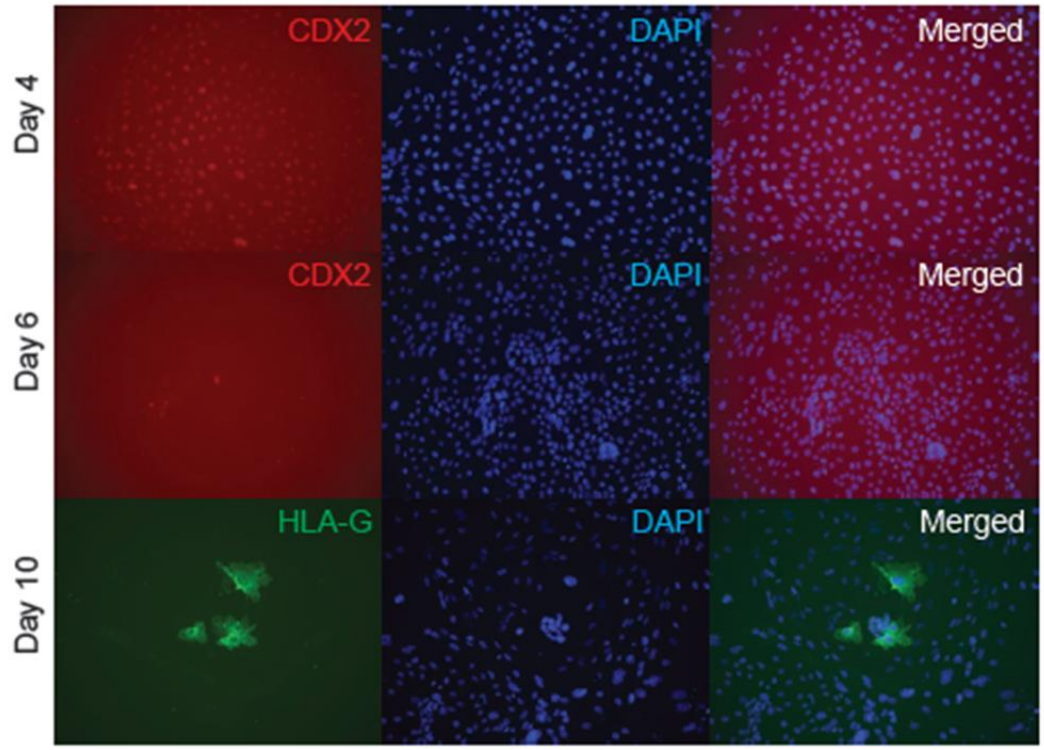
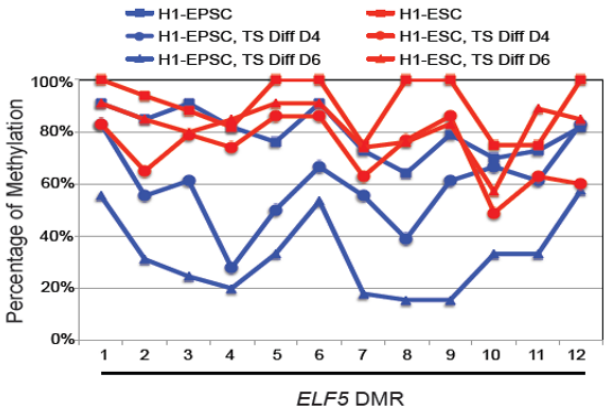


Figure 6.24. Expression of trophoblast markers at protein level in differentiated human EPSCs.

Immuno-staining of cells differentiated from EC5-EPSCs for trophoblast specific markers, *Cdx2* (at day 4 and 6) and *HLA-G*. (Trophoblast induction was performed in collaboration with Dr Wei Wang PhD, Sanger Institute).

A.



B.

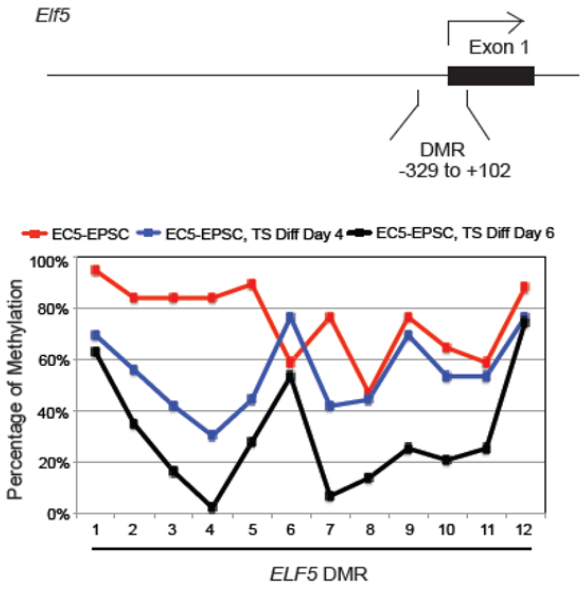


Figure 6.25. Dynamic changes in *ELF5* methylation during trophoblast induction strongly suggest that human EPSCs are producing bona-fide trophectoderm.

(A) Bisulfite analysis of the promoter region of the *ELF5* in H1-EPSCs (passage 10 in EPSCM) during trophoblast differentiation. The methylation percentage of *ELF5* promoter of undifferentiated H1-EPSCs was significantly down-regulated compared to H1-ESCs under the same condition. (B) Bisulfite analysis of the promoter region of the *ELF5* locus in EC5-EPSCs during differentiation towards trophoblasts. The top panel is the diagram of the DMR at the *ELF5* promoter region.. After differentiation for 6 days, the methylation levels of the *ELF5* promoter in EC5-EPSCs was further decreased. Methylation experiments were done in collaboration with Dr Wei Wang, Sanger Institute.

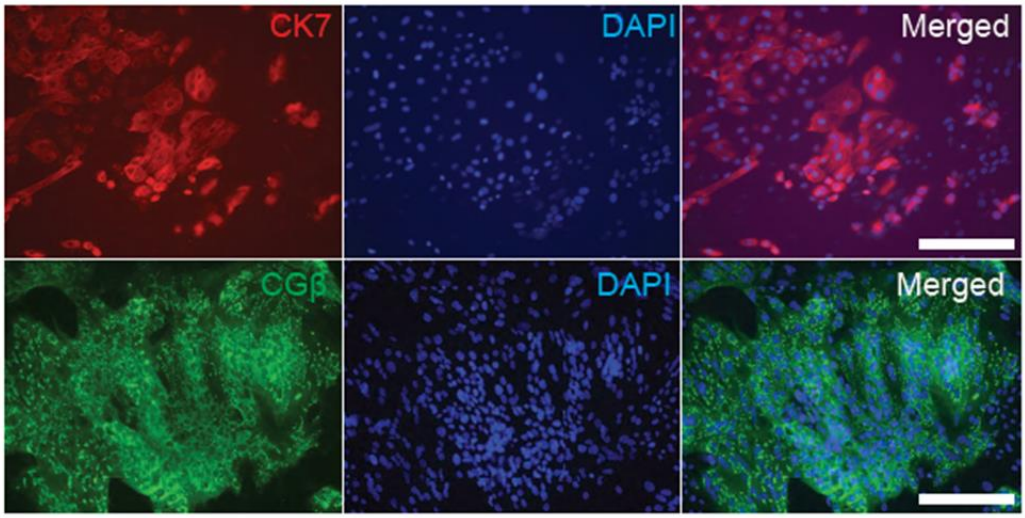


Figure 6.26. Human EPSCs can produce terminally differentiated trophoblast.

Immuno-staining of day 10 differentiated cells from EC5-EPSCs for trophoblast specific markers (scale bar = 50 μm). CK7, Cytokeratin 7; CGβ, Chorionic Gonadotropin, β peptide. Thanks to Jian Yang and Wei Wang for immunostaining.

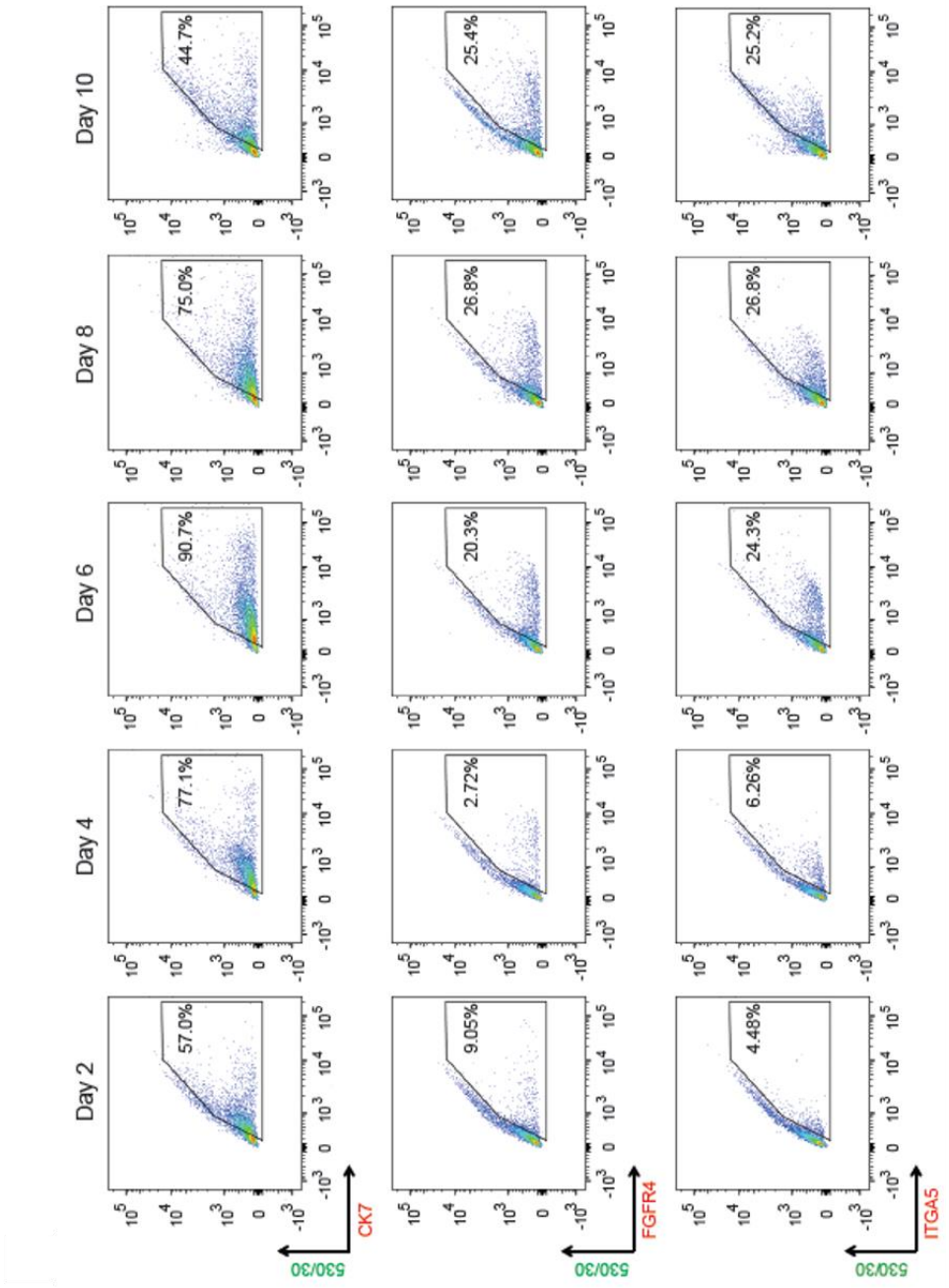


Figure 6.27. Timeline of CK7, FGFR4 and ITGA5 expression in differentiating human EPSCs.

FACS analysis demonstrated high expression of FGFR4, Integrin α5 (ITGA5) and Cytokeratin 7 during trophoblast differentiation of human EPSCs.

6.5 Discussion

In this chapter, I reported that the same chemical formulation EPSCM permitted the conversion of human ESCs and iPSCs to a unique human stem cell state that possessed distinct differentiation properties *in vitro* and combined transcriptional and epigenetic features associated with the pre-implantation stage human embryo, similar to what was described in the mouse system in Chapter(s) 3, 4 and 5.

6.5.1 EPSCM permits successful conversion of human ESCs and human iPSCs to human EPSCs.

In mouse, successful conversion in EPSCM required embryonic stem cells as the starting material as mouse EpiSCs were found not to be 'convertible'. As human ESCs are thought to be more akin to the mouse EpiSC or primed state of self-renewal, how can successful conversion of human ESCs be explained assuming the interspecies comparison is valid? (Yu and Thomson, 2008). It is known that human ESCs are heterogeneous in terms of marker gene expression and differentiation potency with some lines having features close to mouse ESCs in particular H1 human ESCs, clones of which we have found to possess preferential activation of the distal enhancer of the *Oct4* locus (internal communication), a characteristic associated with naïve mouse ESCs (Adewumi et al., 2007; Bao et al., 2009; Koyanagi-Aoi et al., 2013; Thomson et al., 1998; Yeom et al., 1996). The fact that H1, but not H9 human ESCs, which are thought to be more akin to the 'primed' or mouse EpiSC state could be converted in EPSCM is not unexpected in view of this known line-to-line heterogeneity in human ESCs (Cahan and Daley, 2013; Kojima et al., 2014; Koyanagi-Aoi et al., 2013).

In Chapters 3 and 5, following successful conversion, mouse EPSCs activated a core pluripotency transcriptional network that had features in common with the transcriptomic

constitution of ground-state mouse ESCs (Marks et al., 2012). Similarly, human EPSCs expressed high levels of pluripotency related genes and had low expression of lineage specifiers. Interestingly, human EPSCs upregulated multiple genes known to be associated with naïve pluripotency in the mouse including *KLF4*, *UTF1* and *TFCP2L1*. Despite some of the shared transcriptional features with ground-state pluripotency, the transcriptome of mouse EPSCs was unique and clustered independently of 2i/LIF mouse ESCs. Similarly, the human EPSC transcriptome was also found to be unique and clustered independently of conventional human ESCs and all published cell types claimed to represent the naïve state of human pluripotency (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014).

Following conversion, human EPSCs adopted cell culture characteristics more akin to the mouse ES / EPSC state include shortened cell cycle time, capacity for single cell subcloning, ease of gene targeting and robust non-targeted differentiation to all three lineages in embryoid body and teratoma assays. Efficient targeted differentiation to functional neuronal and pancreatic cell types *in vitro* was also reported. The ease of cell culture maintenance, differentiation and efficiency of gene targeting with generation of lineage specific reporters as hereby demonstrated may provide a useful biological resource for future studies on *in vitro* differentiation.

6.5.2 The X-chromosome reactivation status in human EPSCs is unique and may share some similarity with a pre-implantation stage of development

In humans, zygotic gene activation occurs at the four to eight-cell stage which coincides with the initiation of XCI, specifically, the first appearance of *XIST*. Pre the eight-cell stage, *XIST* is not expressed and after this time point, *XIST* will always be detectable on at least one X

chromosome. Single cell RNA FISH revealed that >75% of human EPSCs had no detectable *XIST* expression but had two active X-chromosomes as evidenced by biallelic X-lined gene expression on single molecule RNA FISH and pan-X chromosome re-activation on RNA sequencing. This epigenetic data may provide indirect evidence to support the notion that human EPSCs may possess some features in common with at least the eight-cell stage of development. Gene set enrichment analysis again although indirect evidence suggested that a pre-implantation signature corresponding to the eight-cell stage of development was enriched in human EPSCs.

An important caveat that must be considered when analysing human XCI data *in vitro* is the knowledge that X-chromosome reactivation can be an abnormal finding in cultured cells due to erosure of repressive X-chromosomal marks. These abnormal cells can be identified as during differentiation there is a failure of dosage compensation and re-inactivation of the spuriously reactivated X-chromosome (Mekhoubad et al., 2012; Vallot et al., 2015). Therefore, to exclude this possibility in human EPSCs, they were subjected to differentiation using non-selective embryoid body induction and targeted neural differentiation protocols, respectively. Excitingly, during differentiation, as should be observed in normal development, there was reappearance of *XIST* and silencing of one of the X-chromosomes.

6.5.3 Hypomethyation at the *ELF5* locus during trophectoderm induction suggests that human EPSCs produce bona fide trophoblast *in vitro*

The data reported thus far on human EPSCs suggested a core similarity between mouse and human EPSCs. A critical final experiment would be the demonstration of functional expanded potential by differentiating human EPSCs to bona fide trophoblast *in vitro*.

Following initiation of trophoblast induction, human EPSCs in contrast to standard human ESCs showed rapid and robust upregulation of the TE related genes, *CDX2* and *ELF5* with a greater than 10 fold difference in the expression level of the TE gatekeeper, *ELF5* between differentiating EPSCs and H1 ESCs (Amita et al., 2013; Bernardo et al., 2011; Hemberger et al., 2010). Incredibly, during differentiation of EPSCs, the *ELF5* promoter rapidly became demethylated and by day 6, two sites within the *ELF5* promoter were completely demethylated. In stark contrast, *ELF5* promoter methylation did not appreciably change in differentiating H1 human ESCs. As the human *ELF5* locus is demethylated in early human placental trophoblasts, this data strongly suggested that EPSCs were differentiating to bona fide trophoblast (Hemberger et al., 2010). The lack of demethylation in differentiating human ESCs is consistent with published data indicating that BMP mediated human ESC differentiation may not produce genuine extra-embryonic trophoblast (Hemberger et al., 2010; Roberts et al., 2014).

Beyond the changes in promoter methylation, differentiating EPSCs later expressed high levels of markers known to be expressed in syncytiotrophoblast and extra-villous cytotrophoblast suggesting that representative cell types from the full diversity of extra-embryonic placental subtypes were present in differentiating EPSCs, although further studies are necessary to confirm this claim.

In conclusion, the chemical formulation EPSCM, similar to what was demonstrated in the mouse, permits the establishment *in vitro* of a unique stem cell state that possesses expanded differentiation potential to both the embryonic and extra-embryonic lineage. Epigenetic and transcriptional profiling indirectly suggests that EPSCs possess some characteristics that are thought to be associated with a pre-implantation stage of human development which may provide a mechanism for the observed phenotype. However, multiple future experiments are necessary to confirm or refute this claim.

CHAPTER 7

7 GENERAL DISCUSSION

7.1 Summary

This chapter begins by summarizing the thesis on the establishment in culture of mouse and human stem cells with expanded fate potential. I will then evaluate the potential significance of my findings in the context of pre-implantation development, stem cell biology and totipotency. I will also briefly discuss proposed future experiments and finally, I will make the conclusions of my study in this thesis.

7.1.2 Hypothesis driven media formulation, EPSCM, permits the establishment and maintenance *in vitro* of a unique stem cell state

In Chapter 3, I reported the development of a novel chemically defined medium, EPSCM which permitted the establishment and maintenance *in vitro* of a unique stem cell state that demonstrated expanded fate potential to both the embryonic and extra-embryonic lineage in chimeric blastocysts.

Preceding studies had described the existence of rare subpopulations of cells within conventional ESC culture that possessed totipotent-like features (Abad et al., 2013; Macfarlan et al., 2012; Morgani et al., 2013). However, these subpopulations were a transient phenomenon and could not be established into stable lines *in vitro*. Following my chance observation that inhibition of Src, a pre-implantation signalling pathway conferred some consistent expanded potential to conventional mouse embryonic stem cells, I hypothesised

that further modulation of signalling pathways known to be important in pre-implantation development may allow 'capture' of an expanded potential stem cell state (Li et al., 2011). As totipotent-like cells were present in ES cell culture, albeit at low frequency, I chose ESCs as the platform to test the novel chemical formulation (Macfarlan et al., 2012; Morgani et al., 2013).

Excitingly, I reported that stable stem cell lines with expanded potential could be derived from mouse ESCs, by reprogramming somatic cells and from individual blastomeres of mouse pre-implantation embryos. These cells could be differentiated to mature trophoblast *in vitro* and chimeric blastocyst assays confirmed a high frequency of observed contribution in the epiblast and trophectoderm in contrast to ESCs which contributed to the epiblast only. Low plating density experiments provided evidence to suggest that the establishment of EPSCs in EPSCM cultured mouse ESCs may involve a conversion or reprogramming process rather than capture of a pre-existing cell type within the culture. Additional evidence in support of this claim came from experiments designed to derive *de novo* EPSC lines from embryos. Here, EPSCM substantially altered the development of pre-implantation embryos. The derivation of EPSCs appeared to be a complex reprogramming process involving embryo cell proliferation and apoptosis. The transient loss of OCT4 and CDX2 expression in the "filled" blastocysts and the re-emergence of OCT4 positive or CDX2 positive cells may resemble the normal dynamic expression observed in the four to eight-cell cleavage-stage embryo. The mechanisms underlying the embryonic cell reprogramming process in EPSCM requires further investigation.

Despite their expanded potential, EPSCs retained the core features associated with pluripotency including expression of pluripotency related transcription factors, low expression of lineage specifiers and interestingly, preferential activity of the *Oct4* distal enhancer

suggesting that these cells had features in common with the ground-state of self-renewal (Ying et al., 2008). However, the signalling requirements of EPSCs were different from 2i/LIF ESCs with EPSCs demonstrating sensitivity to JAK inhibition. Consistent with reported data on STAT3 null ESCs, JAK-STAT signalling was not required for the maintenance of 2i/LIF cultured ESCs (Ying et al., 2008). In contrast, LIF has been shown to be involved in extra-embryonic lineage priming and a key component in the promotion of the reported totipotent-like HEX1 population in ESC culture (Morgani et al., 2013). Therefore, LIF may be responsible for some of the observed expanded potential in EPSCs.

The chemical inhibitors in EPSCM were chosen based on their known role in pre-implantation development mediating the first lineage specification which came from genetic and biochemical studies. Our biochemical data confirmed inhibition of the respective pathways in cells cultured in EPSCM. Single chemical addition and removal experiments confirmed the necessity of each individual component although; the magnitude of effect was greatest for inhibitors of WNT and SFK signalling. Furthermore, EPSCs that were re-cultured in conventional ESC media rapidly lost their expanded potential differentiation capacity indicating that EPSCM was essential for maintaining the observed phenotype.

If a cell has totipotent-like potential, then, by definition, it has pluripotent potential. Indeed, mouse EPSCs readily contributed to the three somatic germ layers in embryoid body assays *in vitro* and formed mature teratomas *in vivo*. Spectral karyotyping confirmed genetic stability and EPSCs were able to contribute to the germline fulfilling the requirements for assigning pluripotent potential.

7.1.3 *In vivo* chimeric contribution confirms the expanded fate potential of EPSCs

mCherry expressing EPSCs were injected to morula, eight-cell and blastocyst stage embryos and the E14.5 embryo was subsequently assessed for donor cell contribution. As expected, 2i/LIF ESCs contributed extensively to the embryo proper but donor cell contribution was not detected in the extra-embryonic compartment. Incredibly, EPSC contribution was detected in the embryo proper, yolk sac and placenta. As the placenta and yolk sac have contribution from both the embryonic (extra-embryonic mesoderm) and extra-embryonic lineage, multiple assays were performed to confirm the nature of the EPSC contribution in both compartments (Rossant and Cross, 2001). Definitive trophoblast or extra-embryonic contribution was confirmed by detecting mCherry positive donor cells in the trophoblast giant cell compartment and histologically with the identification of donor cells lining the maternal vascular sinusoids. Furthermore, gene expression analysis of a random population of donor cells from the placenta confirmed a comparable trophoblast lineage gene expression profile to a random sample of wild type placental cells (Rossant and Cross, 2001). Yolk sac sectioning and high resolution microscopy similarly confirmed the presence of EPSC donor cells in the extra-embryonic endoderm, a true extra-embryonic cell type and the extra-embryonic mesoderm, the epiblast derivative which will form the allantois. This data provided strong evidence to support the claim that EPSCs possessed bona fide totipotent-like potential *in vivo*.

7.1.4 Single cell chimeras confirmed that EPSCM maintains a homogenous population of cells that possess expanded fate potential

It was possible that the observed contribution in EPSCM may have been secondary to the presence of subpopulations within the culture that possessed unipotent potential such as a trophoblast stem cell-like cell, an ES cell and XEN cell which could in theory give rise to all three lineages in contribution assays. This was considered unlikely but in order to refute such claims, a single EPSC was injected into eight-cell embryos for chimera production.

Remarkably, a single EPSC could contribute extensively to both the embryonic and extra-embryonic compartments as previously described. Furthermore, when contribution was observed, it was always present in both compartments confirming that EPSCM maintains a homogenous population of totipotent-like cells in culture, the first ever demonstration of this possibility.

7.1.5 EPSCs have a unique transcriptome and possess a gene expression signature with some features of eight-cell stage pre-implantation blastomeres

Independent of their derivation origins, EPSCs were demonstrated to be functionally similar. Not surprisingly, EPSCs were also found to be molecularly similar with their transcriptomes clustering together regardless of their derivation origin. Consistent with the experimental results, EPSCs shared a pluripotency foundation with mouse ESCs and possessed many transcriptional features known to be associated with ground-state pluripotency including homogenous expression of key pluripotency related factors and low expression of lineage specifiers (Marks et al., 2012). Interestingly, in contrast to ground-state ESCs which were enriched in gene ontology terms related to metabolism, EPSCs were enriched in terms related to placental development (Marks et al., 2012). Single cell RNA sequencing demonstrated molecular homogeneity supporting the functional homogeneity observed in the single cell chimera experiments. Intriguingly, EPSCs were enriched with the molecular signatures of cleavage-stage embryos, particularly the four to eight-cell embryo which may provide a molecular basis for their expanded potential. Importantly, the EPSC transcriptome was found to cluster independently of the recently reported 2C-like cells, *Hex1* positive ESCs and *in vivo* reprogrammed iPSCs all of which were shown to have some degree of expanded potential *in vivo*, further supporting the claim that EPSCs represent a unique cell state (Abad et al., 2013; Macfarlan et al., 2012; Morgani et al., 2013).

7.1.6 A dynamic methylome and bivalent domains in EPSCs may be further support for a pre-implantation signature

Epiblast and ground-state mouse ESCs which are the derived from the former represent a 'tabula-rasa' or 'blank-slate', a site of global DNA hypomethylation prior to the initiation of somatic lineage commitment (Leitch et al., 2013). This effect is mediated through elevated levels of *Prdm14* which in turn down-regulates the *de novo* methyltransferases *Dnmt3a/3b* (Leitch et al., 2013). This ground-state possesses minimal repressive histone marks such as H3K27me3 and very few bivalent domains (Marks et al., 2012). Bivalent domains refer to the co-localisation of repressive (H3K27me3) and activator (H3K4me3) histone marks at a promoter. This occurrence, which is prevalent in serum cultured ESCs, is thought to represent a site with poised transcription prior to the onset of somatic differentiation (Marks et al., 2012). In support of this reported data, I presented the same findings from serum and 2i cultured mouse ESCs in this thesis.

In contrast, EPSCs possess a unique epigenetic signature. Firstly, multiple bivalent domains were detected at the promoters of key developmental loci including at those of the extra-embryonic lineage. As EPSCs have minimal expression of lineage specifiers, similar to 2i cultured ESCs and a homogenous expression of pluripotency genes, this bivalency is unlikely to represent so called 'promiscuous transcription' (Marks et al., 2012). Serum cultured ESCs are poised to differentiate to the somatic lineage in response to appropriate environmental cues which is thought to be reflective of the bivalent domains it possesses which allows rapid activation or repression of a specific lineage determinant. It is reasonable; therefore, to speculate that if mouse EPSCs have some features of an eight-cell stage of development at the cusp of the first fate decision, the presence of bivalent domains at key trophoblast and ICM determinant loci would allow the cell to react quickly to developmental cues, a necessity in pre-implantation development.

A second unique feature of mouse EPSCs is that enzymes for DNA methylation and active DNA demethylation are both active. This is corroborated by elevated levels of 5-methylcytosine (mC) and 5 hydroxy-methylcytosine (5hmC) in EPSCs. Our knowledge of pre-implantation development describes a process of rapid and progressive DNA demethylation leading to the blank state that constitutes the epiblast (Guo et al., 2014; Leitch et al., 2013; Reik et al., 2001). The global expression of *Tet1* and *Tet2* mRNAs and 5hmC abundance is less in the trophectoderm than in the ICM (Ito et al., 2010; Senner et al., 2012). Furthermore 5hmC levels may have a functional role in dictating the first cell fate specification as loss of *Tet1* expression and 5 hmC in one of the blastomeres in the two-cell stage embryo biases that blastomere to a TE fate (Ito et al., 2010). The co-existence of DNA methylation and demethylation activity in EPSCs may explain the plenipotent potential of EPSCs. Speculatively, based on the loss of function experiments on *Tet1* in the two-cell embryo, an imbalance in 5mC or 5hmC in EPSCs would produce TE or ICM, respectively. Although further analyses are needed to confirm, such as interrogating the spatiotemporal methylation / demethylation pattern of individual blastomeres, the dynamic methylome in EPSCs may be additional evidence to support the claim that, these cells possess a pre-implantation signature corresponding to the eight-cell stage blastomere.

7.1.7 EPSCM permits the establishment and maintenance of a novel stem cell state in human which has biological, transcriptional and functional similarity with mouse EPSCs

Experiments on human pre-implantation development are limited due to a scarcity of available material to study; however, transcriptional data suggests that many of the signalling pathways involved in the first lineage specification are shared between mouse and human

(Adjaye et al., 2005; Senner and Hemberger, 2010). Therefore, if these pathways are conserved, I speculated that it may be possible to establish a human stem cell state that has similarities with mouse EPSCs. Similar to mouse, we chose human ESCs as the platform for initial testing. Excitingly, human ESCs were capable of being converted and maintained in EPSCM although; there was variable success with marked line to line variation. Tentatively, lines that possessed properties more akin to mouse ESCs such as activity of the distal enhancer of *OCT4*, as was identified in H1 ESCs were successfully converted in comparison to more primed lines such as H9 ESCs which could not be converted. This is entirely speculative based on the mouse conversion experiments and further validation is required in human. I also successfully converted human iPSC lines and generated *de novo* iPSC lines in EPSCM.

Following conversion, similar to mouse, human EPSCs possessed a core pluripotency network with high expression of pluripotency related genes and minimal expression of lineage specifiers. Preferential expression of genes such as *KLF4, UTF1* and *TFCP2L1* raised the possibility that the cells were adopting some transcriptional features known to be associated with the mouse ground-state. Beyond signalling, the cells behaved more closely to mouse ESCs than human ESCs with a shortened cell cycle, single cell subcloning and demonstrated efficient gene targeting with generation of single and dual reporter lines including OCT4, NANOG and dual OCT4 / SOX2 reporter lines which may be a useful biological resource for future studies on pluripotency and / or differentiation.

Human EPSCs retained the core functional characteristics of a pluripotent cell being easily differentiated to three germ layers *in vitro* and produced mature teratomas *in vivo*. Furthermore, the cells could be selectively differentiated to functional pancreatic and electrophysiologically competent terminally differentiated neurons.

Transcriptionally, as in mouse, EPSCM cultured human cells clustered together regardless of their derivation origin and possessed a unique transcriptome independent of conventional FGF cultured human ESCs and all published human naïve cells (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014). Assessment of totipotent-like properties is difficult in human due to ethical constraints. Therefore, I sought to evaluate for the presence of a pre-implantation signature using surrogate assays, specifically, gene set enrichment analysis and epigenetic profiling of the status of X chromosome inactivation in human EPSCs. Intriguingly, although a crude assay, gene set enrichment analysis suggested that EPSCs possessed a signature that had similarity with the eight-cell stage of development.

In pre-implantation development, zygotic gene activation coincides with the initiation of XCI. In humans, this corresponds to the eight-cell stage of development when *XIST* first becomes detectable (Okamoto et al., 2011; Telford et al., 1990; Van den Berg et al., 2009). Therefore, if human EPSCs were in fact representative of the eight-cell stage or pre-XCI phase of development, two active X-chromosomes and no *XIST* should be detectable (Van den Berg et al., 2009). Incredibly, > 75% of human EPSCs possessed this pattern which was further confirmed by the detection of biallelic gene expression and pan X-chromosome reactivation on RNA sequencing. However, it is known that spurious X-chromosome reactivation can occur in cultured cells, which is thought to be partly explained by the erosure of *XIST* which is not reversible during differentiation (Mekhoubad et al., 2012; Vallot et al., 2015). Amazingly, following differentiation of early passage human EPSCs, *XIST* reappeared, X-linked gene expression became mono-allelic and pan X-chromosome dosage compensation was detected confirming that epigenetic status of the X-chromosome in EPSCs was not an artefact of the culture condition. This is the first demonstration that a stably cultured line can

maintain X-chromosome activation and retain the capacity to undergo normal dosage compensation during differentiation. Such a cell type may be a useful resource for investigating the process of random and imprinted XCI in humans and for modelling X-linked disease which has been hampered up until now by the lack of a cell line with the epigenetic characteristics described herein (Mekhoubad et al., 2012).

Finally, as a proof of concept of expanded potential, trophoblast differentiation was performed. Human ESCs are often claimed to have a capacity to form trophoblast *in vitro* following treatment with BMP4 (Amita et al., 2013; Xu et al., 2002). Multiple debates suggest that this process may not reflect bona fide trophoblast induction but, is likely more indicative of extra-embryonic mesoderm differentiation, the somatic contribution to the placenta (Roberts et al., 2014). Evidence for this statement comes from experiments which demonstrate failure to reactivate the endogenous *ELF5* locus during BMP4-mediated differentiation. *ELF5* is a key transcription factor in the trophoblast self-renewal network and is highly expressed in first trimester villous trophoblast (Hemberger et al., 2010). Remarkably following trophoblast induction, as compared to human ESCs, human EPSCs showed rapid and marked up-regulation of key trophoblast lineage determinants including *ELF5* and critically, differentiated EPSCs rapidly became de-methylated at the *ELF5* locus suggesting that EPSCs were differentiating to bona fide extra-embryonic trophoblast, the first ever demonstration of this possibility.

7.2 Significance

7.2.1 Mouse

This study demonstrates for the first time that substantially homogenous, genetically stable stem cell lines with expanded fate potential to both the embryonic and extra-embryonic

lineage can be derived from individual blastomeres of mouse pre-implantation embryos and from reprogramming mouse ESCs or somatic cells in a novel chemically defined media called EPSCM. EPSCM and EPSCs may provide a useful resource to study the first lineage specification *in vitro*.

The establishment of EPSCs lends support to my original hypothesis that stem cell states with blastomere-like potential may be isolated *in vitro* through the selective modulation of pathways known to be involved in pre-implantation development. Therefore, modulation of additional pre-implantation signalling pathways may permit establishment of cell states that possess some features of the two-cell or zygote stage embryo and ultimately, the acquisition of a totipotent line, which in theory would allow creation of an entire mouse from a single starting cell.

7.2.2 Human

I have also demonstrated for the first time, that a stem cell state with features associated with pre-implantation blastomeres can be maintained in human using the same chemical formulation, EPSCM. Human EPSCs which are similar to mouse EPSCs may provide a useful biological resource due to their distinct biological properties including robust culture, genetic stability and ease of genetic modification which should accelerate disease modelling and ultimately cell based therapy. Moreover, the X-reactivation in early passage human EPSCs may facilitate studies on the mechanisms involved in the initiation and maintenance of random and imprinted human XCI. Finally and perhaps the most significant finding, is that human EPSCs can be differentiated to apparent bona fide trophoblast, the first ever demonstration of this possibility.

7.2.3 Clinical relevance

The placenta is a critical organ at the maternal-fetal interface that supports the fetus throughout pregnancy being involved in gas and nutrient exchange, excretion, hormone secretion and immunotolerance. In humans, miscarriage, intra-uterine growth restriction and pre-eclampsia to name but a few, are disorders of placentation. The mechanisms involved in pre-eclampsia for instance have remained unknown due to the inability to model human trophoblast development *in vitro*. Although additional studies are required, the present data suggests that human EPSCs can be differentiated to bona fide trophoblast *in vitro* as evidenced by *ELF5* promoter demethylation, a key epigenetic regulator of lineage restriction and a tropohoblast specific determinant. Therefore, human EPSCs may permit insights to be obtained into disorders of human placentation which may hasten the path to its future prevention.

7.3 Future Experiments

With any scientific study, the obtained results are not absolute and as was described throughout my thesis, many areas require additional investigation to confirm or refute the made observation. Beyond these experiments, the major areas that will be interrogated in immediate future experiments focus on firstly, trying to prove that EPSCs represent a conserved state across multiple species by attempting to derive EPSCs from species such as pig, dog and monkey. Gain and loss of function studies will be performed in an attempt to provide mechanistic insight into which component(s) and / or signalling pathway(s) in EPSCM are causal for the expanded fate potential phenotype. It is possible that an optimised medium which contains fewer inhibitors or compounds that modulate alternative pathways known to be active in pre-implantation development may similarly allow maintenance of an expanded potential stem cell state. In addition, continuing with the original hypothesis of this

thesis, additional chemicals will be tested in an attempt to capture an even earlier stage of development *in vitro*.

At present, a major effort is being placed on optimizing the process of mouse and human trophoblast induction from EPSCs, characterization of differentiated trophoblast cell types and attempting to isolate a self-renewing trophoblast stem cell population,

An interesting platform which I am planning to use to further evaluate the developmental capacity of mouse EPSCs is somatic cell nuclear transfer (SCNT). The nucleus of a differentiated cell type once transferred into an enucleated oocyte will be successfully reprogrammed to a totipotent state which has the capacity to form an entire developing organism *de novo* (Gurdon, 1962). 'Dolly' the sheep was the first example of successful mammalian cloning (Wilmut et al., 1997).

However, although at first glance, an apparent panacea, SCNT is an inefficient process with very few of the cloned embryos being able to develop into the adult organism (Blelloch et al., 2006). This inefficiency is influenced by the differentiation and methylation status of the donor nucleus (Blelloch et al., 2006). Embryonic stem cells are the most efficient choice of donor nucleus for successful SCNT with a progressive reduction in efficiency as somatic differentiation progresses (Blelloch et al., 2006; Wakayama et al., 1999). Therefore, if EPSCs have features in common with a pre-implantation stage of development and represent an earlier temporal location in the developmental timeline, they should be an even more efficient donor nucleus for SCNT than conventional embryonic stem cells.

In mouse, in addition to conversion of conventional ESCs, *de novo* EPSCs were derived by culturing pre-implantation embryos and single blastomeres in EPSCM. The human EPSCs presented in this thesis were generated through conversion of conventional human ESCs and

human iPSCs only. Therefore, the next set of experiments will focus on attempting to derive a human EPSC line directly from the pre-implantation human embryo.

Mouse EPSCs possess a unique epigenetic signature with multiple bivalent domains detected at the promoters of key developmental loci including at those of the extra-embryonic lineage. Secondly, the enzymes for DNA methylation and active DNA demethylation are both active. Single cell epigenomics including DNA methylation and Chip seq analysis at each stage of mouse pre-implantation development are planned which will allow interrogation of the spatiotemporal epigenetic signature of individual blastomeres. This data will allow EPSCs to be chronologically positioned in the pre-implantation timeline. The gene set enrichment analysis and functional data presented in this thesis suggests that EPSCs have features in common with the eight-cell stage blastomere.

The mechanism involved in the derivation of EPSCs from the pre-implantation embryo is at present unknown. As discussed in Chapter 3, the derivation process appears to be a complex reprogramming process that involves embryo cell proliferation and apoptosis. Speculatively, I argued that the transient loss of OCT4 and CDX2 expression in the "filled" blastocysts and the re-emergence of OCT4 positive or CDX2 positive cells may resemble the normal dynamic expression observed in the four to eight-cell embryo. In an attempt to reconcile my observations, *in vivo* single cell tracking and live embryo imaging experiments are being designed so as to observe the derivation process in EPSCM in real time.

X-chromosome reactivation status, *in vitro* trophoblast induction and gene set enrichment analysis were the surrogate assays used in this thesis to assess for expanded differentiation potential in human EPSCs. In human, *in vitro* assays are traditionally used as *in vivo* assessment of such potential is not ethically permissible. However, with the recent

demonstration that the early stages of post-implantation human development can be successfully modelled *in vitro* it is conceivable, pending ethical approval, that the differentiation capacity of human EPSCs may be testable in a chimeric assay using this model system (Deglincerti et al., 2016; Shahbazi et al., 2016).

7.4 Conclusion

This thesis has shown for the first time that through chemical modulation of pathways implicated to be involved in pre-implantation development, a novel homogenous stem cell state that possesses a pre-implantation transcriptional signature and functional expanded fate potential to both the embryonic and extra-embryonic lineage can be established and maintained *in vitro* in both the mouse and human, suggesting a possible interspecies conservation of the signalling networks involved in early embryonic development.

CHAPTER 8

8 REFERENCES

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., Graña, O., Megías, D., Domínguez, O., Martínez, D., et al. (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature 502, 340-345.

Abell, A.N., Granger, D.A., Johnson, N.L., Vincent-Jordan, N., Dibble, C.F., and Johnson, G.L. (2009). Trophoblast stem cell maintenance by fibroblast growth factor 4 requires MEKK4 activation of Jun N-terminal kinase. Mol Cell Biol 29, 2748-2761.

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

Adjaye, J., Huntriss, J., Herwig, R., BenKahla, A., Brink, T., Wierling, C., Hultschig, C., Groth, D., Yaspo, M.L., Picton, H.M., et al. (2005). Primary differentiation in the human blastocyst: comparative molecular portraits of inner cell mass and trophectoderm cells. Stem Cells 23, 1514-1525.

Amano, T., Hirata, T., Falco, G., Monti, M., Sharova, L.V., Amano, M., Sheer, S., Hoang, H.G., Piao, Y., Stagg, C.A., et al. (2013). Zscan4 restores the developmental potency of embryonic stem cells. Nat Commun 4, 1966.

Amita, M., Adachi, K., Alexenko, A.P., Sinha, S., Schust, D.J., Schulz, L.C., Roberts, R.M., and Ezashi, T. (2013). Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. Proc Natl Acad Sci U S A 110, E1212-1221.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J.K., and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signalling in pregastrulation mammalian development. Proc Natl Acad Sci U S A 95, 5082-5087.

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

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., et al. (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8, 532-538.

Azzolin, L., Panciera, T., Soligo, S., Enzo, E., Bicciato, S., Dupont, S., Bresolin, S., Frasson, C., Basso, G., Guzzardo, V., et al. (2014). YAP/TAZ incorporation in the β-catenin destruction complex orchestrates the Wnt response. Cell 158, 157-170.

Balakier, H., and Pedersen, R.A. (1982). Allocation of cells to inner cell mass and trophectoderm lineages in preimplantation mouse embryos. Dev Biol 90, 352-362.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature 461, 1292-1295.

Beddington, R.S., and Robertson, E.J. (1999). Axis development and early asymmetry in mammals. Cell 96, 195-209.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell 9, 144-155.

Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells 24, 2007-2013.

Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. Nature Methods 10, 1093-1095.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191-195.

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.L., and Smith, A. (2008). Capture of authentic embryonic stem cells from rat blastocysts. Cell 135, 1287-1298.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev Biol 210, 30-43.

Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., and Torres-Padilla, M.E. (2013). Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. Cell Rep 5, 687-701.

Cahan, P., and Daley, G.Q. (2013). Origins and implications of pluripotent stem cell variability and heterogeneity. Nat Rev Mol Cell Biol 14, 357-368.

Cerulo, L., Tagliaferri, D., Marotta, P., Zoppoli, P., Russo, F., Mazio, C., DeFelice, M., Ceccarelli, M., and Falco, G. (2014). Identification of a novel gene signature of ES cells self-renewal fluctuation through system-wide analysis. PLoS One 9, 83235.

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

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature 450, 1230-1234.

Chan, S.W., Lim, C.J., Chong, Y.F., Pobbati, A.V., Huang, C., and Hong, W. (2011). Hippo pathway-independent restriction of TAZ and YAP by angiomotin. J Biol Chem 286, 7018-7026.

Chan, Y.S., Göke, J., Ng, J.H., Lu, X., Gonzales, K.A., Tan, C.P., Tng, W.Q., Hong, Z.Z., Lim, Y.S., and Ng, H.H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell 13, 663-675.

Chang, S.C., and Brown, C.J. (2010). Identification of regulatory elements flanking human XIST reveals species differences. BMC Mol Biol 11, 20.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell 10, 615-624.

Chen, L., Wang, D., Wu, Z., Ma, L., and Daley, G.Q. (2010). Molecular basis of the first cell fate determination in mouse embryogenesis. Cell Res 20, 982-993.

Chisholm, J.C., and Houliston, E. (1987). Cytokeratin filament assembly in the preimplantation mouse embryo. Development 101, 565-582.

Cho, C.H., Hannan, N.R., Docherty, F.M., Docherty, H.M., João Lima, M., Trotter, M.W., Docherty, K., and Vallier, L. (2012). Inhibition of activin/nodal signalling is necessary for pancreatic differentiation of human pluripotent stem cells. Diabetologia 55, 3284-3295.

Cockburn, K., Rossant, J. (2010). Making the blastocyst: lessons from the mouse. J Clin Invest 120, 995-1003.

Condic, M.L. (2014). Totipotency: what it is and what it is not. Stem Cells Dev 23, 796-812.

Cross, J.C., Flannery, M.L., Blanar, M.A., Steingrimsson, E., Jenkins, N.A., Copeland, N.G., Rutter, W.J., and Werb, Z. (1995). Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development 121, 2513-2523.

Czechanski, A., Byers, C., Greenstein, I., Schrode, N., Donahue, L.R., Hadjantonakis, A.K., and Reinholdt, L.G. (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nat Protoc 9, 559-574.

Damert, A., Miquerol, L., Gertsenstein, M., Risau, W., and Nagy, A. (2002). Insufficient VEGFA activity in yolk sac endoderm compromises haematopoietic and endothelial differentiation. Development 129, 1881-1892.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci U S A 98, 13734-8.

De Paepe, C., Krivega, M., Cauffman, G., Geens, M., and Van de Velde, H. (2014). Totipotency and lineage segregation in the human embryo. Mol Hum Reprod 20, 599-618.

Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. Nature 533, 251-254.

Deng, Q., Ramsköld, D., Reinius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 343, 193-196.

Denicol, A.C., Dobbs, K.B., McLean, K.M., Carambula, S.F., Loureiro, B., and Hansen, P.J. (2013). Canonical WNT signaling regulates development of bovine embryos to the blastocyst stage. Sci Rep 3, 1266.

Dietrich, J.E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development 134, 4219-4231.

Doble, B.W., and Woodgett, J.R. (2003). GSK-3: tricks of the trade for a multi-tasking kinase. J Cell Sci 116, 1175-1186.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.

Dupuy, A.J., Rogers, L.M., Kim, J., Nannapaneni, K., Starr, T.K., Liu, P., Largaespada, D.A., Scheetz, T.E., Jenkins, N.A., and Copeland, N.G. (2009). A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. Cancer research 69, 8150-8156.

Dyer, M., Farrington, S.M., Mohn, D., Munday, J.R., and Baron, M.H. (2001). Indian hedgehog activates hematopoiesis and vasculogenesis and can respecify prospective neurectodermal cell fate in the mouse embryo. Development 128, 1717-1730.

El-Hashash, A.H., Warburton, D., and Kimber, S.J. (2010). Genes and signals regulating murine trophoblast cell development. Mech Dev 127, 1-20.

Enders, A.C., Given, R.L., and Schlafke, S. (1978). Differentiation and migration of endoderm in the rat and mouse at implantation. Anat Rec 190, 65-77.

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

Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T., and Ko, M.S. (2007). Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. Dev Biol 307, 539-550.

Faria, T.N., Ogren, L., Talamantes, F., Linzer, D.I., and Soares, M.J. (1991). Localization of placental lactogen-I in trophoblast giant cells of the mouse placenta. Biol Reprod 44, 327-331.

Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., and Goldfarb, M. (1995). Requirement of FGF-4 for postimplantation mouse development. Science 267, 246-249.

Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y., Walter, J., et al. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell 13, 351-359.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature 504, 282-286.

Gao, X., Yang, J., Tsang, J.C., Ooi, J., Wu, D., and Liu, P. (2013). Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. Stem Cell Reports 1, 183-197.

Gardner, R.L. (1996). Can developmentally significant spatial patterning of the egg be discounted in mammals? Hum Reprod Update 2, 3-27.

Genevet, A., Polesello, C., Blight, K., Robertson, F., Collinson, L.M., Pichaud, F., and Tapon, N. (2009). The Hippo pathway regulates apical-domain size independently of its growth-control function. J Cell Sci 122, 2360-2370.

Goolam, M., Scialdone, A., Graham, S.J., Macaulay, I.C., Jedrusik, A., Hupalowska, A., Voet, T., Marioni, J.C., and Zernicka-Goetz, M. (2016). Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. Cell 165, 61-74.

Graham, S.J., Wicher, K.B., Jedrusik, A., Guo, G., Herath, W., Robson, P., and Zernicka-Goetz, M. (2014). BMP signalling regulates the pre-implantation development of extra-embryonic cell lineages in the mouse embryo. Nat Commun 5, 5667.

Gross, V.S., Hess, M., and Cooper, G.M. (2005). Mouse embryonic stem cells and preimplantation embryos require signaling through the phosphatidylinositol 3-kinase pathway to suppress apoptosis. Mol Reprod Dev 70, 324-332.

Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477, 606-610.

Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A.L. (1994). Essential role of Mash-2 in extraembryonic development. Nature 371, 333-336.

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.P., Hu, B., et al. (2014). Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell 15, 447-458.

Guo, G., Huang, Y., Humphreys, P., Wang, X., and Smith, A. (2011). A PiggyBac-based recessive screening method to identify pluripotency regulators. PLoS One 6, e18189.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J Embryol Exp Morphol 10, 622-640.

Hamatani, T., Carter, M.G., Sharov, A.A., and Ko, M.S. (2004). Dnamics of global gene expression changes during mouse preimplantation development. Dev Cell 6, 117-131.

Harun, R., Ruban, L., Matin, M.D., J., Jenkins, N.M., Liew, G.C., Andrews, P.W., Li, T.C., Laird, S.M., and Moore, H.D. (2006). Cytotrophoblast stem cell lines derived from human embryonic stem cells and their capacity to mimic invasive implantation events. Hum Reprod 21, 1349-1358.

Hemberger, M., Udayashankar, R., Tesar, P., Moore, H., and Burton, G.J. (2010). ELF5-enforced transcriptional networks define an epigenetically regulated trophoblast stem cell compartment in the human placenta. Hum Mol Genet 19, 2456-2467.

Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K., et al. (2013). Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. Curr Biol 23, 1181-1194.

Hirate, Y., and Sasaki, H. (2014). The role of angiomotin phosphorylation in the Hippo pathway during preimplantation mouse development. Tissue Barriers 2, e28127.

Hirose, G., and Jacobson, M. (1979). Clonal organization of the central nervous system of the frog. I. Clones stemming from individual blastomeres of the 16-cell and earlier stages. Dev Biol 71, 19.

Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M., and Paul, S. (2009). GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. J Biol Chem 284, 28729-28737.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature 326, 292-295.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57.

Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature 461, 614-620.

Huang, Y., Osorno, R., Tsakiridis, A., and Wilson, V. (2012). In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. Cell Rep 2, 1571-1578.

Hung, S.S., Wong, R.C., Sharov, A.A., Nakatake, Y., Yu, H., and Ko, M.S. (2013). Repression of global protein synthesis by Eif1a-like genes that are expressed specifically in the two-cell embryos and the transient Zscan4-positive state of embryonic stem cells. DNA Res 20, 391-402.

Huynh, K.D., and Lee, J.T. (2003). Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature 426, 857-862.

Irie, N., Weinberger, L., Tang, WW., Kobayashi, T., Viukov, S., Manor, YS., Dietmann, S., Hanna, JH., Surani, MA. (2015). SOX17 is a critical specifier of human primordial germ cell fate. Cell 160, 253-68.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129-1133.

James, J.L., Carter, A.M., and Chamley, L.W. (2012). Human placentation from nidation to 5 weeks of gestation. Part I: What do we know about formative placental development following implantation? Placenta 33, 327-334.

Jedrusik, A., Bruce, A.W., Tan, M.H., Leong, D.E., Skamagki, M., Yao, M., and Zernicka-Goetz, M. (2010). Maternally and zygotically provided Cdx2 have novel and critical roles for early development of the mouse embryo. Dev Biol 344, 66-78.

Jedrusik, A., Cox, A., Wicher, K., Glover, D.M., and Zernicka-Goetz, M. (2015). Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition. Dev Biol 398, 147-152.

Johnson, M.H., and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. Cell 24, 71-80.

Kapoor, A., Yao, W., Ying, H., Hua, S., Liewen, A., Wang, Q., Zhong, Y., Wu, C.J., Sadanandam, A., Hu, B., et al. (2014). Yap1 activation enables bypass of oncogenic Kras addiction in pancreatic cancer. Cell 158, 185-197.

Kelly, S.J. (1977). Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. J Exp Zool 200, 365-376.

Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A., Jones, V., Hor, A., de Alencastro, G., Logan, G.J., et al. (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell Stem Cell 14, 107-120.

Kolodziejczyk, A.A., Kim, J.K., Tsang, J.C., Ilicic, T., Henriksson, J., Natarajan, K.N., Tuck, A.C., Gao, X., Bühler, M., Liu, P., et al. (2015). Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. Cell Stem Cell 17, 471-485.

Koyanagi-Aoi, M., Ohnuki, M., Takahashi, K., Okita, K., Noma, H., Sawamura, Y., Teramoto, I., Narita, M., Sato, Y., Ichisaka, T., et al. (2013). Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. Proc Natl Acad Sci U S A 110, 20569-20574.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development 132, 1649-1661.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development 134, 2895-2902.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K.D., Yamada, R.G., Ueda, H.R., and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res 34, 42.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359.

Lee, J.T. (2005). Regulation of X-chromosome counting by Tsix and Xite sequences. Science 309, 768-771.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nature protocols 1, 729-748.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol 20, 311-316.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.L., et al. (2008). Germline competent embryonic stem cells derived from rat blastocysts. Cell 135, 1299-1310.

Li, X., Zhu, L., Yang, A., Lin, J., Tang, F., Jin, S., Wei, Z., Li, J., and Jin, Y. (2011). Calcineurin-NFAT signaling critically regulates early lineage specification in mouse embryonic stem cells and embryos. Cell Stem Cell 8, 46-58.

Li, Y., and Behringer, R.R. (1998). Esx1 is an X-chromosome-imprinted regulator of placental development and fetal growth. Nat Genet 20, 309-311.

Lo, C.W., and Gilua, N.B. (1979). Gap junctional communication in the preimplantation mouse embryo. Cell 18, 399-409.

Lu, C.W., Yabuuchi, A., Chen, L., Viswanathan, S., Kim, K., and Daley, G.Q. (2008). Ras-MAPK signaling promotes trophectoderm formation from embryonic stem cells and mouse embryos. Nat Genet 40, 921-926.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L). Nature 190, 372-373.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature 487, 57-63.

Maekawa, M., Yamamoto, T., Kohno, M., Takeichi, M., and Nishida, E. (2007). Requirement for ERK MAP kinase in mouse preimplantation development. Development 134, 2751-2759.

Maekawa, M., Yamamoto, T., Tanoue, T., Yuasa, Y., Chisaka, O., and Nishida, E. (2005). Requirement of the MAP kinase signaling pathways for mouse preimplantation development. Development 132, 1773-1783.

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science 303, 666-669.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A.S., A.F., Smith, A., and Stunnenberg, H.G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149, 590-604.

Martello, G., Bertone, P., and Smith, A. (2013). Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. EMBO J 32, 2561-2574.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell 11, 491-504.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 78, 7634-7638.

Martin, G.R., and Evans, M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. Proc Natl Acad Sci U S A 72, 1441-1445.

Mascetti, V.L., and Pedersen, R.A. (2016). Human-Mouse Chimerism Validates Human Stem Cell Pluripotency. Cell Stem Cell 18, 67-72.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., Haaf T. (2000). Demethylation of the zygotic paternal genome. Nature 403, 501-2.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A., and Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modelling. Cell Stem Cell 10, 595-609.

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

Morgani, S.M., Canham, M.A., Nichols, J., Sharov, A.A., Migueles, R.P., Ko, M.S., and Brickman, J.M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. Cell Rep 3, 1945-1957.

Morgani, SM., Brickman, JM. (2014). The molecular underpinnings of totipotency. Philos Trans R Soc Lond B Biol Sci 5, 369(1657).

Morin-Kensicki, E.M., Boone, B.N., Howell, M., Stonebraker, J.R., Teed, J., Alb, J.G., Magnuson, T.R., O'Neal, W., and Milgram, S.L. (2006). Defects in yolk sac vasculogenesis, chorioallantoic fusion, and embryonic axis elongation in mice with targeted disruption of Yap65. Mol Cell Biol 26, 77-87.

Morris, S.A., Guo, Y., and Zernicka-Goetz, M. (2012). Developmental plasticity is bound by pluripotency and the Fgf and Wnt signaling pathways. Cell Rep 2, 756-765.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev 12, 3579-3590.

Nagy, A., Gócza, E., Diaz, EM., Prideaux, VR., Iványi, E., Markkula,M., Rossant, J. (1990). Embryonic stem cells alone are able to support fetal development in the mouse. Development 110(3):815-21.

Natale, D.R., Paliga, A.J., Beier, F., D'Souza, S.J., and Watson, A.J. (2004). p38 MAPK signaling during murine preimplantation development. Dev Biol 268, 76-88.

Nesterova, T.B., Barton, S.C., Surani, M.A., and Brockdorff, N. (2001). Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos. Dev Biol 235, 343-350.

Ng, R.K.D., W., Dawson, C., Lucifero, D.M., Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol 10, 1280-1290.

Nichols, J., Jones, K., Phillips, J.M., Newland, S.A., Roode, M., Mansfield, W., Smith, A., and Cooke, A. (2009). Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nat Med 15, 814-818.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development 136, 3215-3222.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell 4, 487-492.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor. Cell 95, 379-391.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., et al. (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell 16, 398-410.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev 12, 2048-2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 24, 372-376.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell 123, 917-929.

Oh-McGinnis, R., Bogutz, A.B., and Lefebvre, L. (2011). Partial loss of Ascl2 function affects all three layers of the mature placenta and causes intrauterine growth restriction. Dev Biol 351, 277-286.

Okamoto, I., Patrat, C., Thépot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.P., Renard, J.P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 472, 370-374.

Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., Hino, T., Suzuki-Migishima, R., Ogonuki, N., Miki, H., et al. (2006). Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. Nat Genet 38, 101-106.

Orr-Urtreger, A., Bedford, M.T., Burakova, T., Arman, E., Zimmer, Y., Yayon, A., Givol, D., and Lonai, P. (1993). Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR2). Dev Biol 158, 475-486.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., Walter J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr Biol 10, 475-8.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature 379, 131-137.

Piccolo, S., Dupont, S., and Cordenonsi, M. (2014). The biology of YAP/TAZ: hippo signaling and beyond Physiol Rev 94, 1287-1312.

Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S., and Zernicka-Goetz, M. (2005). Four-cell stage mouse blastomeres have different developmental properties. Development 132, 479-490.

Plachta, N., Bollenbach, T., Pease, S., Fraser, S.E., and Pantazis, P. (2011). Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. Nat Cell Biol 13, 117-123.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.

Ralston, A., Cox, B.J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P.D., J.S., and Rossant, J. (2010). Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. Development 137, 395-403.

Ralston, A., and Rossant, J. (2008). Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Dev Biol 313, 614-629.

Ramirez-Solis, R., Liu, P., and Bradley, A. (1995). Chromosome engineering in mice. Nature 378, 720-724.

Rayon, T., Menchero, S., Nieto, A., Xenopoulos, P., Crespo, M., Cockburn, K., Cañon, S., Sasaki, H., Hadjantonakis, A.K., de la Pompa, J.L., et al. (2014). Notch and hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. Dev Cell 30, 410-422.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science 293, 1089-1093.

Reyes de Mochel, N.S., Luong, M., Chiang, M., Javier, A.L., Luu, E., Toshihiko, F., MacGregor, G.R., Cinquin, O., and Cho, K. (2015). BMP signaling is required for cell cleavage in preimplantation-mouse embryos. Dev Biol 397, 45-55.

Riley, J.K., Carayannopoulos, M.O., Wyman, A.H., Chi, M., Ratajczak, C.K., and Moley, K.H. (2005). The PI3K/Akt pathway is present and functional in the preimplantation mouse embryo. Dev Biol 284, 377-386.

Roberts, A.W., Robb, L., Rakar, S., Hartley, L., Cluse, L., Nicola, N.A., Metcalf, D., Hilton, D.J., and Alexander, W.S. (2001). Placental defects and embryonic lethality in mice lacking suppressor of cytokine signaling 3. Proc Natl Acad Sci U S A 98, 9324-9329.

Roberts, R.M., Loh, K.M., Amita, M., Bernardo, A.S., Adachi, K., Alexenko, A.P., Schust, D.J., Schulz, L.C., Telugu, B.P., Ezashi, T., et al. (2014). Differentiation of trophoblast cells from human embryonic stem cells: to be or not to be? . Reproduction 147, D1-12.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev Biol 361, 358-363.

Rossant, J. (1976). Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. . J Embryol Exp Morphol 36, 283-290.

Rossant, J., and Cross, J.C. (2001). Placental development: lessons from mouse mutants. Nat Rev Genet 2, 538-548.

Rossant, J., and Vijh, K.M. (1980). Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. Dev Biol 76, 475-482.

Rougier, N., Bourc'his, D., Gomes, D.M., Niveleau, A., Plachot, M., Pàldi, A., and Viegas-Péquignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. Genes Dev 12, 2108-2113.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C., et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature 404, 95-99.

Ruzov, A., Tsenkina, Y., Serio, A., Dudnakova, T., Fletcher, J., Bai, Y., Chebotareva, T., Pells, S., Hannoun, Z., Sullivan, G., et al. (2011). Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res 21, 1332-1342.

Saba-El-Leil, M.K., Vella, F.D., Vernay, B., Voisin, L., Chen, L., Labrecque, N., Ang, S.L., and Meloche, S. (2003). An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. EMBO Rep 4, 964-968.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol 241, 172-182.

Santos, F., Hyslop, L., Stojkovic, P., Leary, C., Murdoch, A., Reik, W., Stojkovic, M., Herbert, M., and Dean, W. (2010). Evaluation of epigenetic marks in human embryos derived from IVF and ICSI. Hum Reprod 25, 2387-2395.

Senner, C.E., and Hemberger, M. (2010). Regulation of early trophoblast differentiation - lessons from the mouse. Placenta 31, 944-950.

Senner, C.E., Krueger, F., Oxley, D., Andrews, S., and Hemberger, M. (2012). DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary. Stem Cells 30, 2732-2745.

Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.M., Campbell, A., Devito, L.G., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat Cell Biol 18, 700-708.

Shao, C., and Takagi, N. (1990). An extra maternally derived X chromosome is deleterious to early mouse development. Development 110, 969-975.

Shao, D.D., Xue, W., Krall, E.B., Bhutkar, A., Piccioni, F., Wang, X., Schinzel, A.C., Sood, S., Rosenbluh, J., Kim, J.W., et al. (2014). KRAS and YAP1 converge to regulate EMT and tumor survival. Cell 158, 171-184.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 15, 284.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol 6, e253.

Skamagki, M., Wicher, K.B., Jedrusik, A., Ganguly, S., and Zernicka-Goetz, M. (2013). Asymmetric localization of Cdx2 mRNA during the first cell-fate decision in early mouse development. Cell Rep 3, 442-457.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature 336, 688-690.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development 132, 2093-2102.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545-15550.

Surani, M.A., Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308, 548-550.

Tabansky, I., Lenarcic, A., Draft, R.W., Loulier, K., Keskin, D.B., Rosains, J., Rivera-Feliciano, J., Lichtman, J.W., Livet, J., Stern, J.N., et al. (2013). Developmental bias in cleavage-stage mouse blastomeres. Curr Biol 23, 21-31.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature 256, 640-642.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676.

Takahashi, Y., Carpino, N., Cross, J.C., Torres, M., Parganas, E., and Ihle, J.N. (2003). SOCS3: an essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. EMBO J 22, 372-384.

Takahashi, Y., Dominici, M., Swift, J., Nagy, C., and Ihle, J.N. (2006). Trophoblast stem cells rescue placental defect in SOCS3-deficient mice. J Biol Chem 281, 11444-11445.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. Cell 158, 1254-1269.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science 282, 2072-2075.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. Nature 184, 1286-1287.

Tarkowski, A.K., and Wróblewska, J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. J Embryol Exp Morphol 18, 155-180.

Tarkowski, A.K., Suwińska, A., Czołowska, R., and Ożdżeński, W. (2010). Individual blastomeres of 16- and 32-cell mouse embryos are able to develop into foetuses and mice. Dev Biol 348, 190-198.

Telford, N.A., Watson, A.J., and Schultz, G.A. (1990). Transition from maternal to embryonic control in early mammalian development: a comparison of several species. Mol Reprod Dev 26, 90-100.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196-199.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell 15, 471-487.

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.

Torres-Padilla, M.E., Parfitt, D.E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature 445, 214-218.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development 135, 909-918.

Vallot, C., Ouimette, J.F., Makhlouf, M., Féraud, O., Pontis, J., Côme, J., Martinat, C., Bennaceur-Griscelli, A., Lalande, M., and Rougeulle, C. (2015). Erosion of X Chromosome Inactivation in Human Pluripotent Cells Initiates with XACT Coating and Depends on a Specific Heterochromatin Landscape. Cell Stem Cell 16, 533-546.

Van den Berg, I.M., Laven, J.S., Stevens, M., I., J., Galjaard, R.J., Gribnau, J., and van Doorninck, J.H. (2009). X chromosome inactivation is initiated in human preimplantation embryos. Am J Hum Genet 84, 771-779.

Wakayama, T., Rodriguez, I., Perry, A.C., Yanagimachi, R., and Mombaerts, P. (1999). Mice cloned from embryonic stem cells. Proc Natl Acad Sci U S A 96, 14984-14989.

Wang, W., Huang, J., and Chen, J. (2011). Angiomotin-like proteins associate with and negatively regulate YAP1. J Biol Chem 286, 4364-4370.

Wang, W., Li, N., Li, X., Tran, M.K., Han, X., and Chen, J. (2015). Tankyrase Inhibitors Target YAP by Stabilizing Angiomotin Family Proteins. Cell Rep 13, 524-532.

Wang, W., Yang, J., Liu, H., Lu, D., Chen, X., Zenonos, Z., Campos, L.S., Rad, R., Guo, G., Zhang, S., et al. (2011). Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. Proc Natl Acad Sci U S A 108, 18283-18288.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naive human embryonic stem cells. Proc Natl Acad Sci U S A 111, 4484-4489.

White, M.D., Angiolini, J.F., Alvarez, Y.D., Kaur, G., Zhao, Z.W., Mocskos, E., Bruno, L., Bissiere, S., Levi, V., and Plachta, N. (2016). Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. Cell 165, 75-87.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336, 684-687.

Wilmut, I., Schnieke, A., McWhir, J., Kind, A., and Campbell, K. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature 385, 810-813.

Wilson, M.B., Schreiner, S.J., Choi, H.J., Kamens, J., and Smithgall, T.E. (2002). Selective pyrrolo-pyrimidine inhibitors reveal a necessary role for Src family kinases in Bcr-Abl signal transduction and oncogenesis. Oncogene 21, 8075.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol 13, 838-845.

Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. Biochem Soc Trans 38, 1027-1032.

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat biotechnology 20, 1261-1264.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development 134, 3827-3836.

Yamanaka, Y., Ralston, A., Stephenson, R.O., and Rossant, J. (2006). Cell and molecular regulation of the mouse blastocyst. Dev Dyn 235, 2301-2314.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 20, 1131-1139.

Yang, J., Wang, W., Ooi, J., Campos, L.S., Lu, L., and Liu, P. (2015). Signalling Through Retinoic Acid Receptors is Required for Reprogramming of Both Mouse Embryonic Fibroblast Cells and Epiblast Stem Cells to Induced Pluripotent Stem Cells. Stem Cells 33, 1390-1404.

Ye, X., Deng, Y., and Lai, Z.C. (2012). Akt is negatively regulated by Hippo signaling for growth inhibition in Drosophila. Dev Biol 369, 115-123.

Yeom, Y., Fuhrmann, G., Ovitt, C., Brehm, A., Ohbo, K., Gross, M., Hübner, K., and Schöler, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development 122, 881-894.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115, 281-292.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature 453, 519-523.

Yu, J., and Thomson, J.A. (2008). Pluripotent stem cell lines. Genes Dev 22, 1987-1997.

Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.L., Stagg, C.A., Hoang, H.G., Yang, H.T., Indig, F.E., et al. (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature 464, 858-863.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., BE., B., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137.

Zhao, B., Li, L., Lu, Q., Wang, L.H., Liu, C.Y., Lei, Q., and Guan, K.L. (2011). Angiomotin is a novel Hippo pathway component that inhibits YAP oncoprotein. Genes Dev 25, 51-63.