INVESTIGATING THE ROLES OF ZFP322A AND PATZ1 IN ES CELL IDENTITY AND REPROGRAMMING

MA HUI

NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Ma Hui
27 November 2014

ACKNOWLDGEMENTS

It is a great pleasure to express my heartfelt gratitude to all the people who have helped me during my PhD stint.

My deepest thanks go to my supervisor Dr. Wu Qiang, who has been such an incredible mentor to impart his invaluable experience in doing research and help me shape this process into a journey. I am very grateful for his guidance, insights and direction throughout my PhD studies. Special thanks also go to him for his great effort in editing my papers and correcting this thesis.

I am especially grateful to Ng Hui Min and Ow Jin Rong whose work I have referred to in this thesis. I would also like to extend my appreciation to Dr. Henry Yang for his invaluable dedication in the analysis of the ChIP-sequencing data. And also I would like to thank Dr. Jonathan Loh and Dr. Li Hu for their collaborative work on microarray assays and analysis. I would also like to thank Dr. Ng Huck Hui and Dr. Monica Fedele, who generously provided the precious cells.

I am immensely grateful to Dr. Lee Yun Hwa who taught me the critical foundations of laboratory techniques and shared with me her experience of doing research and living a life. I am also indebted to all the members of WQ's lab through the years-in particular Chong Yew Mei, Teh Xiu Wen Vivian, Goh Zi Yi, Yu Su E, Zhao Qian Yi, Loh Wan Ning, Lalchhandami Tochhawng. I would also like to thank my cohorts and friends that I have made during these four years in Singapore. Their companionship and kindness have made my PhD years so much more bearable and enjoyable.

Finally and most importantly, I thank my parents for their constant and unconditional support throughout the years. I am grateful far beyond what words can express for the depth of their love, their understanding and their encouragement.

TABLE OF CONTENTS

SUMMARY

Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) of murine blastocysts are characterised by their ability to self-renew and their potential to differentiate into many different cell types. These unique properties are governed by a complex transcription regulatory network, including master regulators Oct4/Sox2/Nanog and other pluripotency factors. Interestingly, depletion of a single transcription factor of this network will alter ESC pluripotency. Thus it would be vital to decipher the delicate transcription regulation of pluripotent state in mESCs. The importance of this transcription regulatory network was highlighted by the remarkable discovery that overexpression of transcription factors can reprogram the differentiated somatic cells to pluripotent cells, i.e. induced pluripotent stem cells (iPSCs). This process was therefore known as somatic cell reprogramming. Recent studies have found numerous transcription factors important for ESC identity maintainance or pluripotency induction in somatic cell reprogramming, including many zinc finger proteins.

In this thesis, we identified two zinc finger proteins, Zfp322a and Patz1, which are important regulators of ES cell identity Our results revealed Zfp322a as a novel pluripotency factor which is not only required for maintaining ES cell identity but also can enhance somatic cell reprogramming. On the other hand, Patz1, though previously reported as a pluripotency factor in mESCs, modulates the reprogramming process in a dosage-dependent manner, possibly through its regulation of c-Myc, cell senescence, mesenchymal to epithelial transition (MET) and chromatin modification. Together, these results highlight the novel functions of zinc finger proteins in ES cell identity and reprogramming.

LIST OF FIGURES

ABBREVIATIONS

AP	Alkaline phosphatase
bFGF	Basic fibroblast growth factor
BMP	Bone morphogenetic protein
ChIP	Chromatin immunoprecipitaiton
ChIP-seq	ChIP-sequencing
Co-IP	Co-immunoprecipitaion
dpc	Days post coitum
EB	Embryoid body
ECC	Embryonal carcinoma cell
EMT	Epithelial–mesenchymal transition
EpiSC	Epiblast stem cell
ERK	Extracellular signal-regulated kinase
ESC	Embryonic stem cell
GFP	Green fluorescent protein
HAT	Histone acetyltransferase
HDACi	Histone deacetylases inhibitors
ICM	Inner cell mass
iPSC	Induced pluripotent stem cell
JAK	Janus kinases
LIF	Leukemia inhibitory factor
LincRNA	Large intergenic non-coding RNAs
MAPK	Mitogen-activated protein kinase
MEF	Mouse embryonic fibroblast
MET	Mesenchymal–epithelial transition
miRNA	microRNAs
MTL	Multiple transcription factor-binding loci
OE	Over expression
PcG	Polycomb-group
PE	Primitive endoderm
PGC	Primordial germ cell
POK	POZ and krüppel-like zinc finger
qRT-PCR	Quantitative real time PCR
RNAi	RNA interference
RT-PCR	Reverse transcriptase polymerase chain reaction
SCID	Severe combined immune deficient
STAT	Signal transducer and activator of transcription
TE	Trophectoderm
TGFβ	Transforming growth factor β
Wnt	Wingless-type MMTV integration site
ZFP	Zinc finger protein

CHAPTER 1

INTRODUCTION

1.2 Pluripotent stem cells from the early embryo

1.1.1 Early embryo development

During preimplantation development in mammals, the embryo develops from a totipotent zygote into a developing organism that is composed of three germ layers. It has been demonstrated that only one-cell zygote and 2-cell stage blastomeres are totipotent since they are capable to develop all by the cell itself into a fertile organism (Tarkowski, 1959). From the 2-cell stage to the 16-cell stage, each individual blastomere still possesses totipotency, i.e. the ability to contribute to all lineages of an entire organism including the extraembryonic tissues (Posfai et al., 2014; Suwińska et al., 2008). When the blastomeres further divide, the embryo undergoes compaction and polarization to form a blastocyst containing a fluid-filled cavity (blastocoel). Thereafter the first cell fate decision occurs and totipotency no longer persists. At this stage (~E3.5), the embryo is comprised of two distinct populations of cells: the outer layer, which becomes trophectoderm (TE) surrounding a fluid-filled cavity (blastocoel); the inner cells, which forms the inner cell mass (ICM) inside the blastocoel. The inner cell mass can be isolated and outgrown *in vitro*, known as embryonic stem cells (ESCs)(Evans and Kaufman, 1981; Martin, 1981). At the late blastocyst stage (~E4.5), ICM cells further undergo another cell fate decision prior to uterine implantation, whereby cells become restricted to either the primitive endoderm (PE) in contact with the blastocoel or pluripotent epiblast formed in the inner layer. Epiblast cells from post-implantation embryo can also be cultured *in vitro* and are known as epiblast stem cells (EpiSCs) (Tesar et al., 2007). As the embryo implants and enters gastrulation stage, epiblast cells gradually lose their pluripotency,

form primitive streak, and develop to all three distinctive developmental germ layers of the embryo, i.e. mesoderm, endoderm and ectoderm, as well as the primordial germ cells (PGCs). In contrast, PE and TE are exclusively to form extraembryogenic tissues, such as the parietal yolk sacs and the placenta (Arnold and Robertson, 2009).

In general, there are two types of pluripotent cells which can derived from the early embryo: ESCs derived from the preimplantation ICM or early epiblast and EpiSCs isolated after post-implantation epiblast. They are pluripotent since they possess the potential to differentiate into all three germ layers and the germ line *in vitro* and *in vivo*. In addition, they are able to self-renew when cultured with leukemia inhibitory factor (LIF) and bone morphogenetic protein 4 (BMP4) (murine ESCs), or basic fibroblast growth factor (bFGF) and transforming growth factor β (TGFβ) (human ESCs and mouse EpiSCs) (Beattie et al., 2005; Vallier, 2005; Williams et al., 1988; Ying et al., 2003).

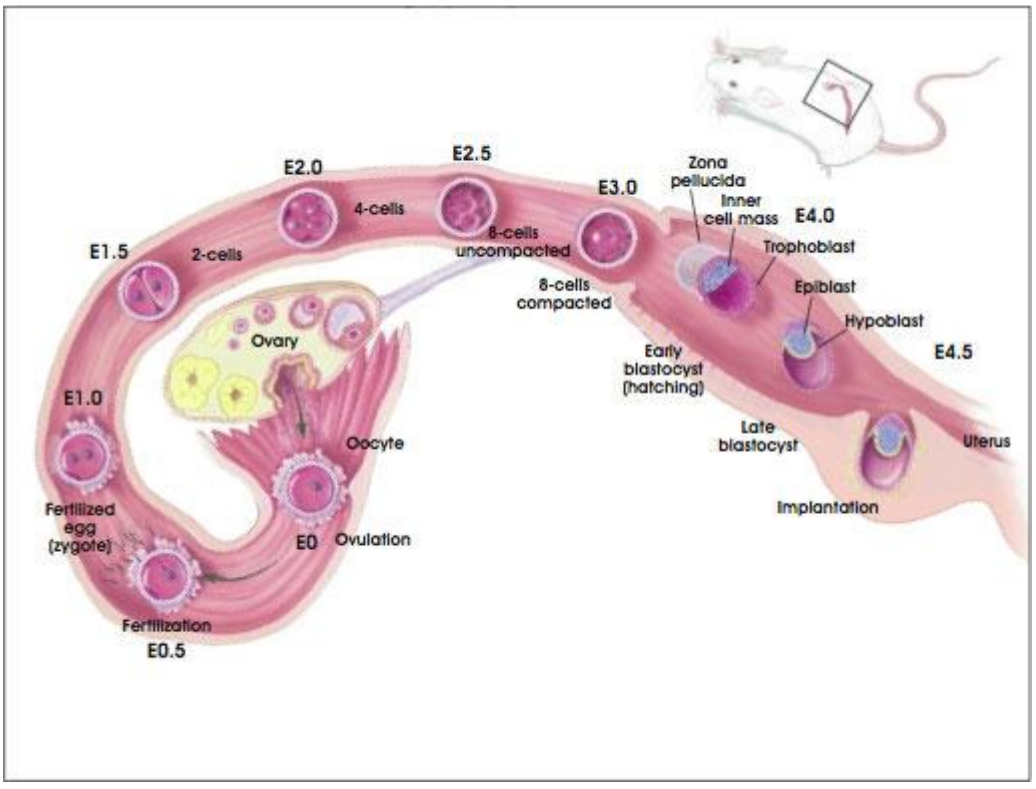

Figure 1.1 Preimplantation embryo development in mouse. (Taken from Appendix A: Early Development. In Stem Cell Information; http://stemcells.nih.gov/info/scireport/pages/appendixA.aspx)

1.1.2 Mouse embryonic stem cells

The isolation of pluripotent embryonic stem cells from mouse blastocysts was first reported by Evans and Kaufman in 1981 (Evans and Kaufman, 1981). They delayed embryo implantation to increase ICM, from which they isolated ICM and expanded them *in vitro*. In the same year, Dr. Gail Martin successfully extracted ICM from late blastocyst using microsurgery (Martin, 1981). These ICM cells were demonstrated to retain infinite self-renewal and pluripotent characteristics by their ability to form teratoma and differentiate *in vitro* and thus are designated as ESCs. Although ESCs are derived from *in vivo* ICM, some differences between ICM and ESCs have been identified, such as that the expression of Pramel5, Pramel6 and Pramel7 are repressed in ESCs but not in ICM (Kaji et al., 2007). It has been reported that mESCs are more similar to 4.5 dpc ICM other than 3.5 dpc ICM, when mESCs are derived (Brook and Gardner, 1997; Kaji et al., 2007). Moreover, ICM cells proliferate rather slowly and no evidence supports that ICM cells can self-renew extensively. How mESCs are derived from ICM (or early epiblast cells) and the molecular differences between these two cell types warrant further investigations.

Mouse ESCs are distinguished for their high nucleus-cytoplasm ratio. They maintain a high proliferative and undifferentiated state *in vitro* by culturing on feeder cells with conventional culture media containing 10-20% fetal bovine serum (FBS) (Evans and Kaufman, 1981). mESCs needs no feeder layer if leukemia inhibitory factor (LIF) is added into the culture media (Williams et al., 1988). When plated at a low density, each undifferentiated ES cell is able to propagate in the same pluripotent state and

forms a rounded, compact and multi-layer clump, which is the characteristic and unique morphology of ESC colonies.

Besides their ability to self-renew, mESCs are also pluripotent, which means they are able to differentiate into all three germ layers in response to developmental cues. This has been demonstrated in numerous studies. For instance, teratoma can be formed if ESCs are injected into mice. Furthermore, when mESCs were cultured on a nonadherent surface in the absence of LIF, they will spontaneously form embryoid body (EB), a spheroid structure comprised of somatic cells from all three germ layers-endoderm, mesoderm and ectoderm. EB generation is an *in vitro* imitation of embryonic development, and, together with teratoma formation, has been considered as defined features of pluripotent cells.

At molecular level, mouse ESCs express a panel of markers that are used to distinguish themselves from other differentiated cell lineages, including cell surface makers SSEA-1 and Thy-1; transcription factors Oct4, Nanog, Sox2, Rex1, Stella *etc*.; signaling pathway-related intracellular markers Smad 1/5/8 ; enzymatic markers alkaline phosphatase, whose expression can be easily visualized by an enzymatic-based reaction and therefore is widely used for pluripotency examination (Zhao et al., 2012).

It has been commonly believed that pluripotent cells cultured *in vitro* fluctuate between ground state pluripotency, a.k.a. naïve pluripotency, and a state primed for lineage commitment (Graf and Stadtfeld, 2008; Takana, 2009). Many studies have established that mESCs are heterogeneous in the expressions of Rex1, Nanog, Klf4 and Stella, hallmarks of naïve pluripotent mESCs (Chambers et al., 2007; Hayashi et al., 2008; Kalmar et al., 2009; Toyooka et al., 2008). Cells which express high levels

of those genes are in a ground pluripotent state while cells with low levels are more prone to exit pluripotency permanently. Interestingly, Ying *et al.* found that inhibitors of Extracelluar-Signal-Regulated Kinases (ERK) and Glycogen Synthase Kinase 3 (GSK3) pathways, which has been termed "2i", eliminate the heterogeneity and robustly support mESCs in a ground state pluripotency, expressing high and homogeneous levels of Oct4, Sox2 and Nanog and conferring ESCs with totipotency (Morgani et al., 2013; Ying et al., 2008). Accordingly, mESCs cultured in 2i display higher colony formation and chimaera contribution rates.

1.1.3 Mouse epiblast stem cells

Mouse EpiSCs are isolated by direct culture of epiblast tissue from early postimplantation embryos (between E4.5 and E8.0) (Brons et al., 2007; Tesar et al., 2007). These cells are capable of infinite propagation in an undifferentiated state when cultured *in vitro* with defined media. The morphology of EpiSCs is different from mouse ESCs. They grow as flattened, monolayered epithelial-like colony with sharp and defined borders. Passaging of EpiSCs using trypsin or other single-cell dissociation methods will induce widespread cell death. Therefore EpiSC colonies are split as small cell clusters by collagenase- or mechanical-mediated passaging (Tesar et al., 2007).

Like mouse ESCs, EpiSCs also express pluripotency markers Oct4, Nanog, Sox2, SSEA-1 and exhibit pluripotency features of forming three primary germ layers *in vivo* teratoma assay and *in vitro* EB formation assay (Brons et al., 2007; Tesar et al., 2007; Zhao et al., 2012). But compared with mESCs, EpiSCs are inefficient to give rise to germ-line contributing chimeras if injected into blastocyst, indicating that they are not naïve pluripotent. Interestingly, female EpiSCs have undergone X

chromosome inactivation, which is not observed in mESCs (Lopes et al., 2008). Furthermore, EpiSC lines cannot be derived in the presence of LIF and BMP4 as mESCs. Instead they are dependent on FGF/ERK and Nodal/Activin signals for long-term maintenance *in vitro* culture (Brons et al., 2007; Greber et al., 2007; Tesar et al., 2007). Given that FGF and Nodal signaling pathways are required for the formation of primitive streak establishment, it is hypothesized that EpiSCs may have acquired some lineage properties of primitive streak cells, thus resulting in a primed state other than naïve pluripotent (F.L. Conlon et al., 1994; Feldman et al., 1995; Verfaillie et al., 2010).

Epiblast, which recapitulates the *in vivo* postimplantation epiblast, is the most immediate pluripotent precursor for all the cell types of the embryo. Therefore EpiSCs provide a unique system to study the transitions from pluripotency to the differentiated lineages. However, the molecular mechanism of regulating EpiSCs has not yet well-defined, which could be attributed to the restriction in accessing this early postimplantation developmental stage and the lack of a robust *in vitro* culture system.

1.1.4 Human embryonic stem cells

In 1998, Thomson *et al.* reported the derivation of human embryonic stem cells (hESC) for the first time (Thomson, 1998). Donated cleavage stage human embryos, produced by i*n vitro* fertilization for clinical purpose, were cultured *in vitro* to the blastocyst stage, from which the ICM were isolated and hESC lines were thus established. Initially, hESCs were cultured on murine or human feeder cells, which served to support hESC expansion and prevent differentiation (Thomson, 1998). Subsequently It was found that hESCs can be maintained in feeder-free condition using optimized culture system, which therefore eliminates the contamination of other

cell types and improves hESC purity (Beattie et al., 2005). Like mESCs, hESCs also grow in colony morphology. However, the morphology of hESCs is similar to mEpiSCs, other than mESCs; and hESCs cannot be derived by clonal expansion of a single cell as mESCs.

The origin of hESCs remains elusive although they are ICM derivatives. In fact hESCs are not fully equivalent to *in vivo* ICM. The transcription profiles of these two cell types are strikingly different (Reijo Pera et al., 2009). There were only 3664 out of 7385 genes expressed in both ICM and hESCs. Indeed, it has been brought up that hESCs actually bear much resemblance to mEpiSCs over mESCs in terms of their morphology, culture condition, growth factor requirements, gene expression profiles and differentiation potentials (Brons et al., 2007; Tesar et al., 2007). Thus it is generally believed that hESCs may not represent the naïve pluripotent state but a rather developmentally primed state. One explanation is that the ICM cells may undergo further differentiation during the derivation of hESCs, and hESCs may be more closely related to a more advanced developmental cell type, such as epiblast cells (Reijo Pera et al., 2009). This notion is supported by the report that hESCs can also be derived from a post-ICM intermediate (PICMI), which is necessary and sufficient for hESC derivation (O'Leary et al., 2012). Interestingly, it has been reported that with ectopic expression of several transcription factors or simply by culturing under physiological oxygen conditions, conventional hESCs can acquire the mESC-like ground pluripotent state (Lengner et al., 2010; Yeo and Ng, 2012).

hESCs express cell surface markers that are characterized for other pluripotent cells, including SSEA-3, SSEA-4, THY-1, TRA1-60, alkaline phosphatase *etc*. They also express high levels of pluripotency genes OCT4, SOX2, NANOG *etc*. (Zhao et al.,

2012). hESCs are able to produce teratoma after injected into severe combined immunodeficient (SCID) mice, but due to legal and ethical reasons, hESCs cannot be injected into a human embryo to test their chimera formation capacity (Thomson, 1998). Notably, hESCs are competent to differentiate into trophectoderm, while mESCs not (Thomson, 1998).

hESCs are significant for their huge potential in translational medicine, human developmental research and drug discovery. Despite the exciting progresses in hESC research in recent years, there is still much basic work required and many obstacles to overcome. Yet due to the limited source of human embryo materials and the ethical restrictions of human embryo manipulations, our current understanding on human preimplantation development and the underlying mechanisms regulating pluripotency and differentiation is rather limited.

Overall, ESCs serve as a powerful model to study cell potential and developmental biology. Genomic modifications by homologous recombination or insertions can be easily manipulated in ESCs. The genetically engineered ESCs then can be used for the generation of knock-in, knock-out or transgenic mice, which are invaluable animal models to study developmental process or genetic diseases. Notably, regarding their robust potential to give rise to all three primary germ cell types, hESCs hold great potential in regenerative medicine.

Property	mESCs	mEpiSCs	hESCs
Origin	ICM ~3.5 dpc	Epiblast 4.5-8.0 dpc	ICM ~5 dpc
Morphology	Domed, multi-layer	Flattened, monolayer	Flattened, monolayer
Clonogenicity	High (single cells)	Low (clumps)	Low (clumps)
Passaging methods	Trypsinization	Mechanical/collagenase-mediated	Mechanical/collagenase-mediated
Response to LIF/STAT3	Self-renewal	None	None
Response to bFGF/Activin	Differentiation	Self-renewal	Self-renewal
Response to BMP	Self-renewal	Differentiation	Differentiation
Teratoma	Yes	Yes	Yes
Chimaera	Yes	No	Not determined
XX status	XaXa	XaXi	XaXa
Typical markers	Oct4, Sox2, Nanog, SSEA-1, Thy-1, Rex1, Stella, SMAD1/5/8, Alkaline phosphatase (AP)	Oct4, Nanog, SSEA-1, Fgf5	OCT4, SOX2, NANOG, SSEA-3, SSEA-4, TRA1-60, Alkaline phosphatase (AP)

Table 1.1 Comparison of mESCs, mEpiSCs and hESCs properties. Table modified from "Pluripotent Stem Cells", book edited by Deepa Bhartiya and Nibedita Lenka, 2013 (Han et al., 2013b). dpc, days postcoitus.

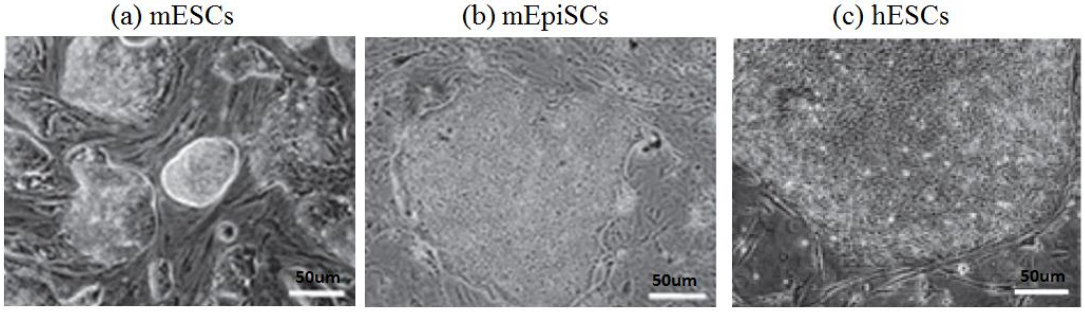

Figure 1.2 Representative figures showing mESCs, mEpiSCs and hESCs morphology (Robinton and Daley, 2012).

1.3 Molecular mechanisms underlying ESC pluripotency

The capacity of self-renewal and pluripotency distinguishes ESCs from differentiated cell types. It is generally believed that the unique properties of ESCs are governed by both extrinsic signaling pathways and intrinsic genetic, epigenetic regulators. The past

decade has witnessed tremendous progress in understanding molecular basis of ESC pluripotency regulation. In this section, I will review these essential signaling and intra-cellular regulators that have been identified to be essential for ESC pluripotency.

1.2.1 Extrinsic signaling pathways required for ESCs

Through binding to cell-membrane receptors, extracellular factors can induce nucleus-directed signaling pathways to modulate gene expression. In ESCs, external growth factors are required to maintain ES pluripotency *in vitro*. The signaling pathways that are required for hESCs are distinctive from mESCs. As shown in Figure 1.3 and Table 1.1, mESCs are responsive to LIF/STAT3 and BMP4 signaling pathways, while hESCs are mainly dependent on two other important signals: FGF/MEK and TGF-β/Activin/Nodal signaling pathways. Interestingly, some of these signaling pathways have completely opposite effect on maintaining ESC phenotype in mouse and human. For instance, FGF/MEK pathway is required to maintain hESC while activation of this pathway drives mESCs towards differentiation; activation of BMP signaling pathway promotes mESC self-renewal but induces hESC differentiation. These differences may reflect the disparities of molecular mechanisms in hESCs and mESCs.

1.2.1.1 LIF/Stat3 signaling pathway

The importance of LIF/Stat3 signaling pathway in mESC has long been recognized. Initially, mouse embryonic fibroblast (MEF) feeder layers were needed to support mESC in an undifferentiated state *in vitro* culture, which indicates that fibroblasts can produce some essential factors for ESC to propagate effectively (Martin, 1981). Later it was revealed that leukemia inhibitor factor (LIF) is the active component that

promotes self-renewal property and maintains the developmental potential of mESCs in the presence of serum (Austin G. Smith et al., 1988; Williams et al., 1988). mESCs differentiate easily without culturing in LIF or on feeder layer. Likewise, MEFs lacking LIF gene are deficient to support mESC culture *in vitro* (Colin L. Stewart et al., 1992).

LIF, a member of Interleukin (IL)-6-type cytokine family, was initially discovered to inhibit mouse leukemia cell proliferation and induce their differentiate into macrophages (Chew et al., 2005). Subsequent studies reported its pleiotropic roles in various biological processes such as embryogenesis, cell proliferation, survival, differentiation and apoptosis depending on cell types (Mathieu et al., 2011). In mESCs, LIF transduces its signal through binding to the transmembrane receptor LIFR, which is a low-affinity ligand specific for LIF, and gp130, a common co-receptor shared by all IL6 family members. The binding of LIF to its receptors leads to the activation of several pathways-the JAK (receptor-associated Janus kinase)/Stat, PI3K (phosphatidylinositol-3-OH kinase) and MAPK (mitogen-activated protein kinases)/ERK pathways. Activated JAKs in turn phosphorylate the tyrosine residues in gp130 and facilitate the recruitment of signal transducer and activator of transcription 3 (Stat3). The following phosphorylation of Stat3 by JAKs leads to Stat3 dimerization and translocation into nucleus, where it binds to specific DNA elements to regulate gene transcriptions, such as activation of Klf4, c-Myc *etc*. (Boeuf et al., 1997; Hitoshi Niwa et al., 1998). Activated PI3K/AKT activates transcription factor Tbx3 (Niwa et al., 2009). Both Klf4 and Tbx3 can activate pluripotency genes *Oct4*, *Nanog* and *Sox2*, to prevent mESC differentiation. Interestingly, MAPK/ERK pathway drives differentiation through negatively regulating Tbx3 and it may be negatively regulated by PI3K, yet the precise mechanism remain elusive (Niwa et al.,

2009; Paling et al., 2004). This suggests that the parallel activations of these pathways are precisely regulated to balance between the pluripotent and primed state.

Among these downstream pathways that LIF signal activates, JAK/Stat has exhibited a predominant role in self-renewal and pluripotency regulation. Mastuda *et al.* reported that constant activation of Stat3 can maintain mESC in self-renewal state independent of LIF, while inactivation of Stat3 drove mESC differentiation (Matsuda et al., 1999). Thus the role of Stat3 as a downstream effector of LIF signal was demonstrated. Interestingly, studies in recent years also suggests that direct downstream targets of Stat3 (such as c-Myc, Klf4, Nanog) can bypass the requirement of LIF when they were ectopically expressed in mESCs (Cartwright, 2005; Niwa et al., 2009; Suzuki et al., 2006). In addition, ChIP-seq analysis have revealed that Stat3 binding sites are co-occupied by pluripotency factors Oct4, Sox2 and Nanog and knockdown of *Stat3* upregulates endoderm and mesoderm genes, implying that Stat3 prevents mESC differentiation by activating pluripotency genes and suppressing lineage specific genes (Bourillot et al., 2009; Chen et al., 2008). Yet the precise mechanism by which activated Stat3 controls its responsive genes and their involvement in pluripotency regulation remains to be further elucidated.

The importance of LIF signaling pathway has led researchers to investigate its roles in preimplantation development. It was found that LIF is only detected in TE while LIFR and gp130 are expressed in ICM cells, which suggests that LIF may contribute to the lineage segregation of differentiated TE and pluripotent ICM (Nichols et al., 1996). Knockout animal models have showed that embryos ablated of LIF/Stat pathways, such as disruption of LIF, LIFR, gp130, Stat3 or Jak1/2, survived beyond blastocyst stage. It has been proposed that there are alternative signaling pathways

that can compensate LIF signaling pathways *in vivo* embryo development (Posfai et al., 2014).

Interestingly, LIF is not sufficient for maintenance of hESCs. If human ES cells are grown without feeder cells but in the presence of LIF, they either differentiate or die (Thomson, 1998). This indicates the distinctive signaling requirements in hESCs and mESCs.

1.2.1.2 BMP/SMAD signaling pathway

In serum-free cultures, LIF alone, however, is insufficient to maintain ESC self-renewal and block neuronal differentiation. It has been found that the requirement of serum can be replaced by BMPs, which hence was identified as another critical component in mESC culture system besides LIF (Ying et al., 2003).

Bone Morphogenetic Protein (BMP) is a subset of the transforming growth factor beta (TGFβ) superfamily, which is comprised of 40 members including BMP, Activin, Nodal *etc*., and regulates various cellular processes in animal development (Itoh et al., 2014). BMP ligand initiates signaling by binding to the type II BMP receptor, followed by recruitment and phosphorylation of type I BMP receptor, which then leads to the phosphorylation of BMP-responsive SMAD 1/5/8. The activated SMAD 1/5/8 form a heteromeric complex with SMAD4 and translocate into the nucleus to regulate gene transcriptions(Itoh et al., 2014; Shi and Massagué, 2003). Ying *et al.* reported that inhibitors of differentiation (Ids), which inhibits basic helix-loop-helix transcription factors, are important mediators of BMP signaling cascade, as forced expression of Id2 can replace BMP signals and inhibit neuronal differentiation in the absence of serum (Ying et al., 2003).

Bmprla is the only type I BMP receptor that has been detected in undifferentiated mESCs and pluripotent ICM (Qi et al., 2004; Ying et al., 2003). Although *Bmprla*-null embryos can survive beyond blastocyst stage, they are deficient to derive mESCs, reaffirming that BMP signaling pathway is essentially needed to sustain ESC pluripotency whereas there may be additional pathways to compensate BMPs for *in vivo* ICM formation. Interestingly, inhibition of MAPK/p38 pathway allows the derivation of ESCs from *Bmprla*$^{-/-}$ blastocysts (Qi et al., 2004). This suggests that BMPs may act through inhibition of MAPK pathway to maintain mESC self-renewal, which needs to be addressed in future studies.

Despite its synergetic roles of LIF pathway in maintaining mESCs, BMPs seem to be implicated in hESC differentiation. Xu *et al.* reported that BMP stimulation induced trophoblast differentiation in the presence of bFGF while inhibition of BMP signaling sustained hESC in undifferentiated state (Pera, 2004; Xu et al., 2005; Xu et al., 2008). Other studies also showed that BMPs are involved in primitive endoderm differentiation and promote mesoderm commitment possibly via activation of SLUG/MSX2 mediated EMT (Pera, 2004; Richter et al., 2014; Teo et al., 2012).

1.2.1.4 TGF-β/Activin/Nodal signaling pathway

TGF- β, Activin and Nodal are other three members of TGF- β family. Nodal, Activin and TGF- β act through binding to their receptors (Nodal and Activin share type I and type II activin-like kinase receptors while TGF- β uses TGF β receptor) and propagate the signals through phosphorylation of downstream effector SMAD2/3. Like SMAD1/5/8 in BMP pathway, SMAD2/3 can also form a complex with SMAD4 and subsequently translocate into nuclear to induce transcription regulations (Schier, 2003).

It was first discovered that TGF- β and Nodal are highly expressed in undifferentiated hESCs and study afterwards showed that Activin A can replace fibroblast feeder layer to sustain hESC *in vitro* culture (Beattie et al., 2005; Sato et al., 2003b; Vallier, 2005). The notion that TGF-β/Activin/Nodal signaling pathway is important in hESC was further demonstrated by the finding that SMAD2/3 activated by this pathway can up-regulate pluripotency factor Nanog, while inhibition of this pathway by follistatin (inhibitor of Activin pathway), Lefty or Cerberus overexpression (known endogenous inhibitors of Nodal pathway), or SB431542 (inhibitor of Activin/Nodal type I receptor) down-regulates pluripotency genes and drives neuroectoderm specification of hESCs(Smith et al., 2008; Vallier, 2005; Vallier et al., 2009; Vallier et al., 2004; Xu et al., 2008). Interestingly, overexpression of Lefty alone could not induce hESC differentiation in the presence of feeder layers and follistatin has no effect on Nodal-overexpressed hESCs, suggesting that Nodal and Activin can act independently to synergistically maintain hESC pluripotency *in vitro* (Vallier, 2005). TGF- β pathway, however, appeared far less efficient in prevention of hESC differentiation in feeder- and serum- free conditions than Activin or Nodal (Vallier, 2005).

The role of Nodal/Activin is much less clear in mouse pluripotent cells. Expression of Activin can be detected in the mouse blastocysts while Nodal starts to be expressed in the epiblast after implantation and has been shown to regulate proper embryo development (Albano et al., 1993; F.L. Conlon et al., 1994; Robertson et al., 2003). Knockout of *Nodal* caused gastrulation lethality in mouse embryo development (Ouwens et al., 2002). Given high similarities between hESC and mEpiSCs, Nodal/Activin pathway has been shown to be important in mEpiSCs. Disruption of Nodal signaling resulted in decreased proliferation of epiblast with very low Oct4 level in mice embryos (Robertson et al., 2003). Nodal/Activin maintained Nanog

expression (Vallier et al., 2009) and inhibition of this pathway drove rapid differentiation of EpiSCs, implying a strict dependence of this pathway for pluripotency maintenance in EpiSCs (Brons et al., 2007).

Nodal/Activin signaling has been proven to promote mESC self-renewal as well. In serum-free cultured mESCs, Nodal/Activin is constitutively activated (Ogawa et al., 2006). Inhibition of Nodal/Activin signals decreased proliferation rate without affecting mESC pluripotency, while stimulation of Nodal/Activin signals, but not TGF-β, increased mESC propagation in a pluripotent state (Ogawa et al., 2006). However, it was reported in another study that overexpression of Nodal in mESCs led to upregulation of mesoderm and endoderm markers but down-regulated neuroectoderm lineages (Pfendler et al., 2005). Galvin *et al.* recently reported that autocrine Nodal/SMAD2 pathway attenuates, but not abolish, BMP signaling pathway, which may contribute to repression of mesoderm and perhaps trophectoderm specification in mESCs (Aloia et al., 2009; Galvin et al., 2010). Thus the multifaceted roles of these pathways in differentiation and self-renewal would require a delicate balance to control the ESC phenotype. Further investigations are needed to examine the distinct roles of Nodal/Activin signaling pathway in human and mouse pluripotent cells and how it cooperates with other pathways or transcription networks to regulate pluripotency.

1.2.1.3 FGF/MEK signal pathway

FGF ligands and heparan sulfate (HS) bind to the FGF receptor (FGFR) forming the FGF/FGFR/HS signalling complex, resulting in receptor dimerization, subsequent autophosphorylation of receptor tyrosine residues and activation of downstream signaling pathways, including classic mitogen-activated protein kinase (MAPK)

(Erk1/2), Jak/Stat, PI3K and phospholipase C (PLC)γ pathways (Mohammadi et al., 2005),

Fibroblast growth factor (FGF) appears to be of central significance to human ESC self-renewal, since FGF2 is essentially required for hESCs in traditional culture medium with fibroblast feeder layers or fibroblast-conditioned medium (Amit et al., 2000; Xu et al., 2005). FGF2 together with a commercially available serum substitute can support hESC clonal propagation on fibroblasts and even in the absence of feeder layers (Amit et al., 2000; Xu et al., 2001). Likewise, withdraw or suppression of FGF signaling will rapidly trigger hESC differentiation (Dvorak et al., 2005).

With gene expression profile analysis, several groups have reported the high expressions of FGF receptors and their downstream effectors in undifferentiated hESCs, including FGFR-1,-2,-3,-4, SOS1, PTPN11, RAF1 (Brandenberger et al., 2004; Dvorak et al., 2005; Sato et al., 2003b; Sperger et al., 2003). In hESCs, it has been found that FGF2 can promote self-renewal and pluripotency in several aspects. Firstly, Vallier *et al.* discovered that FGF cooperates with Nodal/Activin to sustain pluripotency markers for prolonged periods in the absence of serum, Matrigel or feeders. Although FGF signaling is necessary, FGF itself cannot maintain hESC in undifferentiated state if Nodal/Activin signaling pathway is inhibited (Vallier, 2005). More recently, it was found that FGF signaling acts as a direct activator of Nanog promoter in hESCs (Xu et al., 2008; Yu et al., 2011). Thirdly, FGF and TGFβ signaling synergize to suppress BMP pathway to sustain pluripotency genes Oct4, Sox2 and Nanog, preventing hESC from differentiation (Levenstein et al., 2005; Xu et al., 2008; Yu et al., 2011). Moreover, several studies have demonstrated that FGF signals act mainly through FGF/MEK cascade. MEK/ERK pathway is a downstream

cascade of FGF signaling, which functions cooperatively with PI3K/AKT pathway in maintaining hESC pluripotency (Li et al., 2007). In addition, FGF2 indirectly induces TGFβ and IGF2 secretion in feeder cells, which together support hESC pluripotency (Bendall et al., 2007; Greber et al., 2007). Interestingly, FGF2 is also implicated in trophectoderm differentiation and early stage endoderm development in human. This may be attributed to the different roles of FGF2 in different cell context.

In mESCs, FGF signal pathway displays a completely opposing effect. It is not a pluripotency keeper but is well known for its pro-differentiation roles in lineage commitment (Kunath et al., 2007; Lanner and Rossant, 2010). Autocrine FGF-induced ERK1/2 signaling is required for mESCs to exit self-renewal and initiate multi-lineage commitment (Kunath et al., 2007; Stavridis et al., 2007). Fgf4-Fgfr2 pathway drives the epiblast and PE segregation and the PE lineage-primed population is reduced if ERK signaling is blocked (Lanner and Rossant, 2010; Villegas et al., 2010). Indeed, some studies showed that the self-renewal promoting role of LIF and BMP pathways in mESC may act through counterbalancing the autocrine FGF/ERK signaling. Smith and his colleagues discovered that chemicals "2i", which inhibit endogeneous FGF4 and GSK3 signals, can substitute for the growth factor requirement of LIF and BMPs in mESC culture. More recently, another group also reported that inhibitors of FGF/MEK and TGF-β pathways, designated as "R2i", can promote the homogenous, ground state pluripotency in mESCs (Hassani et al., 2013).

1.2.1.5 Other signaling pathways

Although ESCs can be cultured in defined media supplemented with growth factors, it is still recommended to culture ESCs on feeder layers, which will supply complete nutrition, creating a complex microenvironment that favors better pluripotency and

eliminates prevents culture deterioration. In fact, despite these essential growth factors that have been discussed above, fibroblasts secrete multiple other growth factors, which altogether contribute to long-term maintenance of ESCs *in vitro* (Prowse et al., 2007). These multiple signaling pathways crosstalk with each other to form a pluripotent signaling network and maintain a delicate balance among themselves, whereby potently sustaining undifferentiated ESC phenotype.

Eras, for example, is highly expressed in mESCs. *Eras*-null ESCs showed decreased growth rate and tumorigenecity (Takahashi et al., 2003). It activates PI3K/AKT pathway to mediate ESC proliferation, possibly via interacting with mTOR pathway (Murakami et al., 2004; Takahashi et al., 2005). The wingless-related MMTV integration site (Wnt) signaling pathway, has been reported to be active in both mouse and human ESC maintenance (Sato et al., 2003a). Despite its various roles in embryo development, Wnt signaling pathway was shown to promote mESC pluripotency, attenuate neurecotoderm commitment and prevent transition to primed EpiSCs (Berge et al., 2011; Merrill, 2012; Miki et al., 2011; Sato et al., 2003a). Moreover, Wnt pathway has been shown to interconnect with LIF, BMP, GSK, TGF-β/Nodal/Activin pathways to synergize pluripotency regulation (Berge et al., 2011; James, 2005; Lee et al., 2009; Sato et al., 2003a). Therefore Wnt signaling appears to be implicated in keeping ESC pluripotent. However, contradicting results have been published as well, demonstrating that Wnt pathway is inhibited and dispensable in pluripotency maintenance; whereas activation of this pathway stimulates differentiation (Davidson et al., 2012; Kielman et al., 2002; Lyashenko et al., 2011). Hence more efforts are required to define the precise molecular functions of Wnt pathway in ESCs.

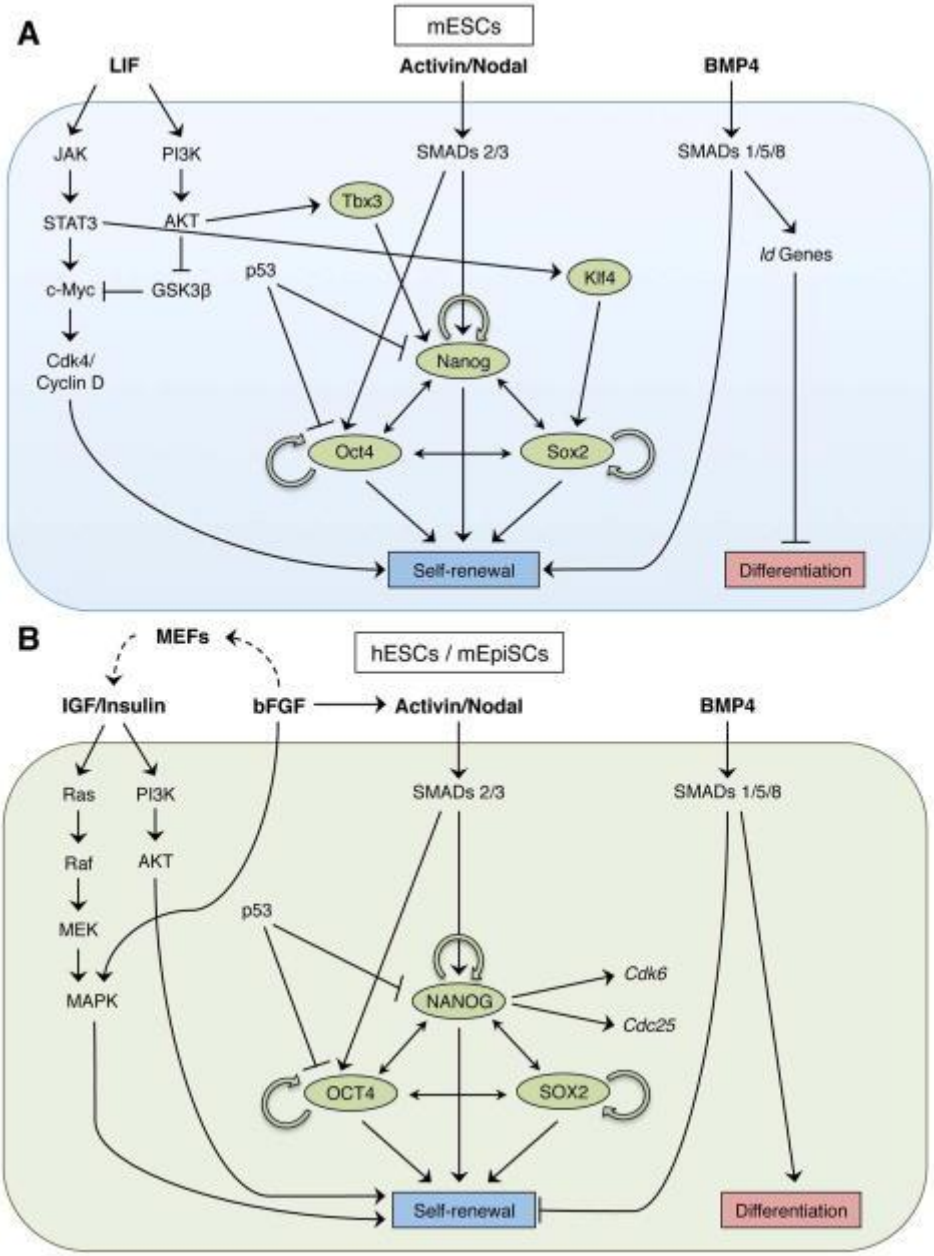

Figure 1.3 Exogenous signaling pathways required for pluripotency maintenance in ESCs. (A) Requirement of LIF and BMP4 signaling in mESCs. (B) Requirement of TGF-β/Nodal/Activin and bFGF signaling in hESCs and mEpiSCs (Saunders et al., 2013).

1.2.2 Intrinsic signaling pathways in ESC maintenance

Besides these signaling pathways which transduce the external signals into the nucleus and lead to transcriptional regulation of their downstream genes, intrinsic regulators, such as transcription factors and epigenetic factors, play crucial roles in

pluripotent cells. Extensive studies have documented that Oct4, Sox2, Nanog are the three master regulators in the core transcriptional network, which mainly contributes to hallmarks of ESC pluripotency and self-renewal.

1.2.2.1 Core transcription factors Oct4/Sox2/Nanog

Oct4 (octamer binding transcription factor-4)

Oct4 (also known as Oct3), encoded by the gene *Pou5f1*, is one of the first transcription factors that have been identified in maintaining ESC pluripotency and proper embryo development. Oct4 belongs to the POU (Pit-Oct-Unc) domain transcription factor family. Like other members of the POU family of transcription factors, Oct4 regulates the expression of their target genes through the binding of its POU-domain to an octameric sequence AGTCAAAT (Pan et al., 2002). Oct4 is known to control downstream genes in synergy with Sox2 via the octamer–sox motif (Oct–Sox enhancer) (Chew et al., 2005) (see more in the following Sox2 section). The other two domains of Oct4, the C-terminal and N-terminal domains, are not responsible for DNA binding but have transactivation activities in transcription regulation (Brehm et al., 1997).

The expression of Oct4 can be initially detected at maternal unfertilized egg, zygote, and all the blastomeres till 8-cell stage. As embryo develops, its expression is restricted to ICM cells, remains in epiblast and then is strictly in PGCs (Palmieri et al., 1994). *In vitro*, Oct4 is expressed at high levels in undifferentiated cells, such as ESCs, embryonal carcinoma cells (ECCs) and embryonic germ cells (EGCs) but limited in lineage cells (Palmieri et al., 1994). Upon differentiation, Oct4 level in ES cells decreases rapidly. Therefore, this expression pattern indicates its essential roles in pluripotency establishment and maintenance. Indeed, knockout of *Oct4* caused

preimplantation lethality at 3.5 dpc. The embryo displayed blastocyst-structure but lacked genuine pluripotent ICM, thus no ESCs could be derived (Nichols et al., 1998). Oct4 is also essential for PGC formation and depletion of Oct4 in PGCs resulted in cell apoptosis and differentiation defect (Kehler et al., 2004; Okamura et al., 2008).

In pluripotent cells, Oct4 functions in a dose-dependent manner. Reduction of Oct4 expression in ESCs triggers trophectoderm differentiation, while a 2-fold overexpression of Oct4 causes ESC differentiate into mesoderm and primitive endoderm (Hough et al., 2006; Niwa et al., 2000b). Furthermore, a recent study revealed that a defined level of Oct4 is required for pluripotency acquisition of induced pluripotent stem cells (iPSCs), as well as their lineage commitment (Radzisheuskaya et al., 2013). Once pluripotency is established in iPSCs, Oct4 is no longer required for self-renewal. Interestingly, Oct4-low iPSCs sustain self-renewal but are deficient to differentiate; the ESC-level Oct4 is required for their *in vitro* and *in vivo* differentiation. Thus an appropriate level of Oct4 is critical to govern the property of pluripotency.

Given its critical role in pluripotency, Oct4 expression is tightly controlled by two distal enhancers (DEs), which are mainly active in ICM, ESCs and PGCs, and its proximal promoter, which drives *Oct4* in epiblast (Yeom et al., 1996). Comparative studies have discovered that there are four conserved regions in *Oct4* promoter, namely, CR1 (immediate upstream of Oct4), CR2, CR3 and CR4 (overlapping with DEs) (Frankenberg et al., 2010). Extensive studies have reported that these conserved regions are occupied by many transcription factors and epigenetic factors to either activate or repress *Oct4* transcription (Chen et al., 2008; Niwa et al., 2005).

Sox2 (sex determining region Y box-2)

Sox2 belongs to the *Sox* transcription factor family which harbors a conserved HMG DNA-binding domain. During embryogenesis, the expression pattern of Sox2 is quite similar to Oct4, which is detected in all the blastomeres till 8-cell stage, subsequently restricted to ICM and epiblast (Avilion, 2003). Unlike to Oct4, Sox2 is also expressed in extraembryonic ectoderm and nervous system (Avilion, 2003). Similarly, disruption of Sox2 expression led to peri-implantation lethality due to the lack of epiblast (Avilion, 2003). It appears that Sox2 plays a role in maintaining epiblast, while Oct4 is essential for ICM development, which is at an earlier developmental stage. Yet both *Sox2*- and *Oct4*-null embryos were unable to give rise to ESCs (Avilion, 2003; Nichols et al., 1998). In ESCs, knockdown of *Sox2* resulted in polyploidy and differentiation into trophectoderm lineage, as *Oct4*-null ESCs do (Chew et al., 2005; Niwa et al., 2000b). These results suggest that Oct4 and Sox2 may function together in embryogenesis and pluripotency maintenance.

Indeed, subsequent studies showed that Sox2 and Oct4 can interact with each other, form an Oct4-Sox2 complex and bind to a specific cis-regulatory element that consists of neighboring Oct (ATGCAAAT) and Sox (CATTGTA) motif (Chew et al., 2005; Loh et al., 2006; Okumura-Nakanishi et al., 2004). Genome-wide studies further revealed that Sox2 and Oct4 occupy same binding sties of their targets, including many silent genes (Chen et al., 2008). Importantly, these oct-sox targets include many pluripotency associated genes, such as Oct4, Sox2, Nanog, Utf1, Fgf4 *etc*.(Okumura-Nakanishi et al., 2004; Remenyi, 2003; Rodda, 2005). However, Sox2 has been found dispensable for the activation of sox-oct enhancers and forced expression of Oct4 was able to restore the loss of pluripotency in *Sox2* null ES cells, suggesting that

Sox2 stabilize the undifferentiated state through synergizing Oct4 in ESCs (Masui et al., 2007).

Nanog

Nanog is another core transcription factor and a key pluripotency marker in ESCs and EpiSCs. The importance of Nanog was first uncovered by the finding that Nanog is highly expressed in ESCs and its overexpression is able to safeguard ESC self-renewal in the absence of LIF (Chambers et al., 2003; Mitsui et al., 2003). In mouse embryos, Nanog is firstly detected in the compacted morula, subsequently localized to ICM and later confined to epiblast (Chambers et al., 2003). *Nanog* deficient embryos failed to generate epiblast and *Nanog*-null ICM cannot proliferate and produce only endoderm-like cells (Mitsui et al., 2003). In ESCs, Nanog faithfully maintains ESC pluripotency and *Nanog* knocked-out in ESCs resulted in differentiation into extra embryonic endoderm (Mitsui et al., 2003). It has been proposed that Nanog can suppress Gata4 and Gata6, which promote lineage commitment of extra embryonic endoderm (Shimosato et al., 2007). Meanwhile, Nanog can activate other pluripotency factors to drive ESC self-renewal, such as Oct4, Rex1 Ids *etc*.. mESCs cultured in 2i/LIF system, which confers ESC ground state pluripotency, increase Nanog expression, indicating that Nanog promotes the transition to naïve pluripotency (Ying et al., 2008).

Although it is widely accepted that Nanog, like Oct4 and Sox2, plays a central role in pluripotency maintenance, Chamber *et al.* subsequently reported that *Nanog*-null ESCs can be propagated in a self-renewal state (Chambers et al., 2007). Nanog is also dispensable for pluripotency reacquisition during reprogramming process (Schwarz et al., 2014). Several other studies reported that Nanog expression is not homogeneous

at single-cell level in ESC population, which may be explained by the variable allelic expression of *Nanog* (MacArthur et al., 2012; Miyanari and Torres-Padilla, 2012; Singh et al., 2007). Biallelic expression of Nanog is important for the establishment of ground state pluripotency of ICM in embryo development. Thus it appears that Nanog is not required for maintaining pluripotency once it has been established. Future studies are of great significance to delineate the exact role of Nanog in ESCs.

Notably, in undifferentiated hESCs and mESCs, a large portion of promoters bound by Oct4/Sox2 are also bound by Nanog, suggesting that Nanog cooperates with Oct4/Sox2 heterodimers to exert its roles (Boyer et al., 2005; Loh et al., 2006). In addition, these three factors form a regulatory circuit to modulate their own transcriptions. It has been proposed that the core transcription factor trio of Oct4/Sox2/Nanog collaboratively establish and govern the pluripotent state by directly and indirectly: 1) repressing lineage specific genes; 2) activating pluripotency associated factors, including each of their own (Young, 2011).

1.2.2.2 Epigenetic regulators

Apart from transcription factors, epigenetic regulators have shown to be important in maintaining ESC pluripotency and cell fate decision for the past decade. Different from somatic cells, ESCs are characterized by a highly plastic and dynamic chromatin configuration with a higher ratio of transcriptionally active marks (e.g., H3K4me3, H3K36me3, H3K27Ac) to repressive marks (e.g., H3K27me3, H3K9me3, 5-mc DNA methylation) (Gaspar-Maia et al., 2011; Zhou et al., 2011). This unique chromatin state is governed by a set of epigenetic regulatory proteins, which are required to support the pluripotency-specific gene expression program.

With large-scale RNAi screening and protein-protein interactome investigations, it was uncovered that the core transcription factors are integrated with multiple epigenetic pathways for pluripotency regulation in ESCs (which will be discussed in a greater detail in the next section). The interplay between central transcription factors and epigenetic regulators has been extensively studied in recent years. Oct4/Sox2/Nanog maintain ESC identity by activating transcription of pluripotency-associated genes via recruitment of co-activators (e.g. p300), chromatin remodelling complexes and the transcriptional machinery, while repressing developmental regulators by engaging Polycomb-group (PcG) protein complexes (e.g., Ring1b, Rybp) as well as other co-repressor complexes, including NuRD, Sin3A and Pml complexes (Chen et al., 2008; Ding et al., 2012; Pardo et al., 2010; van den Berg et al., 2010; Wang et al., 2006). Many transcriptional cofactors, including cohesin, mediators and condensin, have been implicated ESC maintenance as well (Young, 2011). In addition, Oct4, in cooperation with Nanog and Sox2, represses *Xist* (X-inactive specific transcript) and activates Tsix, thus contributing to the X chromosome activation in pluripotent ESCs (Navarro et al., 2011).

Furthermore, Oct4 governs ESC specific chromatin architecture by the direct regulation of other epigenetic modifiers. For example, it can activate H3K9 demethylases Jmjd1a and Jmjd2c to modulate H3K9 methylation status of its target genes (e.g., Nanog, Tcl1, Zfp57) (Loh et al., 2007). On the other hand, Oct4 expression is critically controlled by epigenetic mechanisms as well. G9a, a histone H3 lysine 9 methyltransferase, irreversibly represses *Oct4* expression by H3K9 methylation at its regulatory region (Feldman et al., 2006). Dnmt3a and Dnmt3b are two de novo methyltransferases which establish the DNA methylation patterns during normal embryo development. ESCs lacking of *Dnmt3a* or *Mbd3* were unable to

differentiate properly due to the hypomethylated DNA at *Oct4* promoter (Feldman et al., 2006; Gu et al., 2011). It was proposed that transcription silence by H3K9 methylation and the subsequent incorperation of DNA methylation at *Oct4* promoter regulatory region is required for proper ESC differentiation (Feldman et al., 2006). SNF5, a core subunit of Brahma-associated factor (BAF) complexes, negatively regulates Oct4 and activates Oct4 repressed targets by affecting chromatin landscapes during hESC differentiation (Madhani et al., 2013).

In ESCs, many developmental genes harbor the unique "bivalent" domains, marked by the presence of active H3K4me3 and repressive H3K27me3, which is catalyzed by SET/MLL and PRC2 respectively (Bernstein et al., 2006; Pan et al., 2007; Zhao et al., 2007). This chromatin signature allows a "transcription-ready" state permissive for rapid activation in response to developmental cues. PcG proteins, consisting of PRC1 and PRC2, are important epigenetic players that repress developmental genes by adding repressive chromatin marks at the promoters of these genes (Laugesen and Helin, 2014; Pethe et al., 2014). Deletion of *Eed*, *Suz12* or *Ezh2*, which are components of PRC2, is lethal in embryo development and compromises the ESC self-renewal capacity (O'Carroll et al., 2001; Pasini et al., 2004).

1.2.2.3 The expanded pluripotency network in ESCs

As reviewed above, Oct4/Sox2 and Nanog have been documented to be central to the transcriptional regulatory hierarchy that specifies ESC identity because of their unique expression patterns and vital roles in early development and ESCs. In fact, this core transcription circuit, together with many transcription factors, epigenetic regulators and microRNAs (miRNAs), form a complex and multifaceted regulatory network that fine tunes the delicate pluripotent state in ESCs (Yeo and Ng, 2012).

With large-scale RNAi knockdown studies, many novel important transcription regulators have been screened out. It has been demonstrated that depletion of *Tbx3*, *Esrrb*, *Tcl1*, *Dppa4*, *SetDB1*, *Tip60-400*, *Cnot3*, *Trim28 etc*. in mESCs, and INO80 chromatin remodeling complex, TAF complex, *PRDM14 etc.* in hESCs results in differentiation (Yeo and Ng, 2012). Extensive studies have uncovered that many of these factors act either upstream or downstream of the core transcription trio factors. Esrrb, for example, is activated and recruited by Oct4 (van den Berg et al., 2010) and also a direct downstream target of Nanog and is capable to substitute for Nanog functions in pluripotency regulation (Festuccia et al., 2012). Knockout of *Esrrb* abolished Nanog's ability to guard ESC self-renewal in LIF withdrawal conditions. Taken together, it is implied that these factors are inter-connected to form a network and perturbation of any factors tend to destroy the stabilized pluripotency program to induce differentiation.

Exploring protein interacting partners is another method to expand the transcriptional regulatory network. Proteins that can interact with Oct4/Sox2 dimer or Nanog have been identified (Ding et al., 2012; Pardo et al., 2010; van den Berg et al., 2010; Wang et al., 2006). This protein-protein interactome encompasses other well-documented pluripotency regulators, and particularly many DNA methyltransfereases, chromatin remodeling and modifying factors. Their roles in maintaining ESCs have been supported by multiple independent studies (Yeo and Ng, 2012; Young, 2011).

Moreover genome-wide studies such as ChIP-seq and microarray analysis serve as powerful tools to dissect this transcription regulatory network. Interestingly, it was found that Smad1 and Stat3, the key effector of BMP and LIF pathway respectively, co-bind to many core transcription factors bound sites, which demonstrates the

crosstalk between transcription factors and extrinsic signaling pathways (Chen et al., 2008). Chen *et al.* also observed that these pluripotency transcription factors can be grouped into either Myc- or Oct4-centric module based on their similarities of genomic locations (Chen et al., 2008). c-Myc is dispensable for pluripotency maintenance in ESCs, but it was found to be implicated in chromatin modulations and differentiation suppression (Neri et al., 2011; Varlakhanova et al., 2010). In addition, it was recently reported that c-Myc acts as an amplifier for active transcription rather than activates new genes (Lin et al., 2012; Nie et al., 2012). Thus c-Myc seems to have some roles independently of core transcription trio factors.

Notably, through ChIP-seq analysis, Marson *et al.* uncovered that the core transcription trio factors co-occupy promoters of some stem cell related miRNAs and PcG occupied tissue-specific miRNAs, to activate or repress their transcriptions (Marson et al., 2008). This therefore incorporates miRNAs as components of the ES cell regulatory network. Indeed, recent studies have provided strong evidence that miRNAs have some essential roles in ESC proliferation and pluripotency, as well as in differentiation. Undifferentiated ESCs have a unique set of miRNA expression pattern, including some ESC-preferentially expressed miRNAs (e.g. miR302-367 cluster, miR290) and silent differentiation-associated miRNAs (e.g., let-7, miR 134)(Barroso-delJesus et al., 2008; Marson et al., 2008). ESCs with miRNA processing enzymes disrupted were compromised in their proliferation and differentiation (Fukuda et al., 2007; Kanellopoulou, 2005; Wang et al., 2007a). ESCs lacking Dicer, an enzyme critical for microRNA biogenesis, displayed change in epigenomic landscape and exited self-renewal to differentiate (Asakura et al., 2013). miR290 family regulates G1/S transition in ESCs to promote their rapid proliferation and is able to rescue the DNA methylation defects in *Dnmt*-null ESCs (Sinkkonen et

al., 2008; Wang et al., 2008). miR302, specifically activated by core transcription factors in hESCs, was found to promote BMP signaling and negatively regulate lefty, the inhibitor of Nodal/Activin signaling pathway, thus favoring the pluripotent state of hESCs (Barroso-delJesus et al., 2011; Barroso-delJesus et al., 2008; Lipchina et al., 2011). Moreover, some miRNAs are able to specifically target the core transcription factors and induce transcription silencing of these genes during ESC differentiation(Tay et al., 2008).

In addition, large intergenic non-coding RNAs (lincRNAs) expressed in mESCs are also known targets of pluripotency-associated transcription factors, including Oct4, Sox2, Nanog, cMyc, Klf4 (Guttman et al., 2011). lincRNAs are critical for maintaining Oct4 and Nanog levels in ESCs and suppressing lineage specifications possibly through their associations with multiple epigenetic regulating complexes (e.g. coactivators, mediators) (Lu et al., 2014). It is suggested that lincRNAs serve as a scaffold unit in recruitment of these protein complexes to modulate gene expression (Yeo and Ng, 2012).

The dominant effect of this transcription regulatory network in pluripotency was highlighted by the remarkable discovery that the expression of just four transcription factors, Oct4, Sox2, c-Myc and Klf4 (OSKM) was sufficient to transform somatic MEFs back to pluripotent stem cells, and the expression of OCT4, SOX2, NANOG and LIN28 was sufficient for in human somatic cell reprogramming (Takahashi and Yamanaka, 2006; Yu et al., 2007). This reprogramming process, i.e. the restoration of pluripotency in differentiated cells, serves as a valuable model to decipher the mechanism as how these transcription regulators work together to achieve pluripotency acquisition and maintenance.

1.4 Induced pluripotent stem cells

In the field of developmental biology, cell differentiation is generally thought to be an irreversible process. Once terminally differentiated, specified cell types are non-switchable among each other. However, this concept was challenged by significant breakthroughs that somatic cells can reacquire the pluripotency property by a variety of methods, suggesting that cell fate specification is not a fixed decision, but is a reversible process.

1.3.1 Methods of inducing ESC-like cells

Date back to 1960s, Gurdon and his colleagues first demonstrated that transferring the nucleus of a differentiated frog into an enucleated egg could restore the cell with totipotency (Gurdon, 1962). Subsequent studies have demonstrated the feasibility of somatic nuclear transfer in mammals and the generated pluripotent cells are indistinguishable to ESCs. In 1997, the birth of first cloned mammal Dolly was reported, which was exciting news in stem cell research (I. Wilmut et al., 1997). Cell fusion is another method to regain pluripotency potentials. In 1976, Miller and Ruddler showed that pluripotent hybrids could be derived by fusion of primary thymocyte with pluripotent teractocarcinoma cells (Miller and Ruddle, 1976). Similar finding was observed in fusion of somatic cells with human ESCs (Tada et al., 2001; Terada et al., 2002). Further studies revealed that this method could be simplified by exposure somatic nucleus to ES cell extract (Cowan, 2005). However, these methods have been demonstrated to be very inefficient and the generated pluripotent cells were defective.

In 2006, Takahashi and Yamanaka reported a landmark discovery that reprogramming mouse fibroblasts into pluripotent stem cells can be achieved by viral mediated

transinduction of four transcription factors associated with pluripotency, i.e. Oct4, Klf4, Sox2 and c-Myc (Takahashi and Yamanaka, 2006). These reprogrammed cells resemble ESCs in terms of their morphology, gene expression profile and chromatin configuration. They possess the self-sustained pluripotency to differentiate into many cell types and hence were designated as induced pluripotent stem cells (iPSCs). Yu *et al.* subsequently reported the successful reprogramming in human somatic cells with ectopic expression of Oct4, Sox2, Nanog and Lin28 (Yu et al., 2007). Compared with traditional methods of pluripotency induction, reprogramming with defined transcription factors is easier to manipulate and makes the process less complex to examine. This breakthrough inding has led to blooming progress in stem cell and developmental biology field. Sir John Gurdon and Dr. Shinya Yamanaka were awarded with 2012 Nobel Prize in Physiology or Medicine for these striking achievements.

1.3.2 Mechanism of reprogramming

Studies in past several years have revealed that the cell fate conversion from somatic cells to iPSCs is a dynamic process that involves a cascade of cellular events, such as silencing lineage-specific genes, mesenchymal to epithelial transition (MET), metabolic switch, overcoming cellular senescence and acquisition of cell immortality, reactivation of X-chromosome and reactivation of pluripotency genes, as well as resetting the chromatin signatures (Apostolou and Hochedlinger, 2013; Buganim et al., 2013; David and Polo, 2014). The mature iPSCs generated are transgene-silent, possess self-sustain pluripotency, erase somatic “memories” and acquire all the defined features of ESCs.

MET, a reversed process to EMT, is one of the requisite cellular events during early stage of reprogramming process (Li et al., 2010b; Liu et al., 2013a). Upon OKSM induction, fibroblasts must undergo MET, a process including shutting down the mesenchymal genes, overcoming the EMT epigenetic barrier and epithelial program activation, to successfully initiate the reprogramming. Factors that promote MET, including Klf4, miR302 and E-cad, can enhance pluripotency acquisition; while factors that drive EMT or prevent MET, such as TGF-β and some mesenchymal markers, impede the reprogramming at the initial stage(Li et al., 2010b; Samavarchi-Tehrani et al., 2010; Subramanyam et al., 2011).

Cellular senescence has been reported as another barrier which restricts reprogramming rate at the initial stage (Banito et al., 2009; Li et al., 2009; Utikal et al., 2009b). The *Ink4a*/*Arf* tumor suppressor locus has a critical role in regulating cellular senescence in many types of cells (Collado et al., 2007). For instance, silencing *Ink4a*/*Arf* locus, or ablation of its activators Jmjd3, hasd been shown to reduce cellular senescence and significantly improve reprogramming efficiency (Banito et al., 2009; Zhao et al., 2013). Utikal *et al*. reported that after deletion of *Ink4a*/*Arf* locus, which resulted in immortalisation, almost every fibroblast has the potential to generate iPSCs. Thus cellular senescence is crucial and efficiency-limiting during reprogramming (Utikal et al., 2009b).

Besides, the glycolytic metabotype transition is required for induction of pluripotency. Unlike quiescent somatic cells, pluripotent cells depend on anaerobic glycolysis to meet their high energetic and biosynthetic demands for rapid cell division (Zhang et al., 2012). Somatic cells reprogrammed to pluripotency need to switch from an oxidative to glycolytic state in metabolism (Folmes et al., 2011; Panopoulos et al.,

2011). Studies have revealed that genes involved in glycolysis and oxidative phosphorylation pathways were changed in their epigenetic and gene expression levels during reprogramming process (Panopoulos et al., 2011). Metabolic switching towards glycolysis by chemical compound treatment or HIF1α activation can facilitate reprogramming (Zhou et al., 2012; Zhu et al., 2010). Notably, reprogramming factors Lin28 and c-Myc have been recently shown to stimulate glycolysis and promote metabolism resetting (Singh and Dalton, 2009; Zhu et al., 2011).

Another important issue in reprogramming process is to overcome the "epigenetic memory" of somatic cells. Compare to somatic cells, pluripotent cells possess a highly plastic chromatin structure which is globally more dynamic and decondensed with a higher ratio of active to repressive histone marks. The establishment of unique "bivalent" domains of H3K4me3 and H3K27me3 is prerequisite for successful generation of iPSCs (Polo et al., 2012). Switch of epigenetic landscape is therefore another layer of reprogramming barrier. ChIP-seq results and gene expression profiles uncovered that the epigenetic reprogramming events occur sequentially (Koche et al., 2011). H3K4me2 mark is accumulated at the promoters of many pluripotency genes and lost at repressed somatic genes immediately after induction. These H3K4me2 enriched regions are significantly enriched for Oct4/Sox2 and H3K4me3 targets. H3K4 methylation is also initially deposited at some poised genes, which however is accompanied with a corresponding loss of H3K27me3, thus creating the bivalent domains. ESC-like DNA methylation pattern (i.e. hypomethylated pluripotency gene promoters and hypermethylated somatic gene promoters) and X chromosome reactivation are only fully established at a later stage when the cells have acquired stable pluripotency (Koche et al., 2011; Polo et al., 2012).

The current understanding on reprogramming has brought the model that the induction of pluripotency follows a step-wise process (Buganim et al., 2013; David and Polo, 2014; Yamanaka, 2009). Single-cell cloning experiments have demonstrated that reprogramming starts with a long stochastic phase of gene activation (inducted by reprogramming factors) to adopt one type of the cell fates, such as dedifferentiation, trans-differentiation, senescence, *etc.* This process is very inefficient and is the major rate-limiting step for reprogramming. Cells undergoing dedifferentiation are the ones prepared for pluripotency induction. Next, these reprogrammable cells are rendered with the susceptibility to gain some early features of pluripotency, including rapid proliferation, MET transition, glycolytic metabotype and silencing of MEF-specific genes and even activation of pluripotency genes. Thereby, these cells undergo a second more deterministic phase and eventually become mature iPSCs (Buganim et al., 2012).

The stochastic model at the initial stage of reprogramming indicates that the compacted chromatin of somatic cells has to be destabilized to become more "hyperdynamic", the characteristic of ESC chromatin, whereby rendering the promoters more accessible for the stochastic gene activation. This is evidenced by the observation that epigenetic factors and chemical compounds which promote a globally active chromatin environment can enhance iPSC generation as discussed previously. In addition, c-Myc, a proto-oncogene that enhance cell proliferation, increases the generation of partially reprogrammed iPSCs when combined with OSK, supporting the notion that higher cell proliferation rate accelerates reprogramming at the early stochastic phase (Nakagawa et al., 2007; Wernig et al., 2008). Nevertheless, more efforts are needed to fully elucidate the precise molecular kinetics in pluripotency induction.

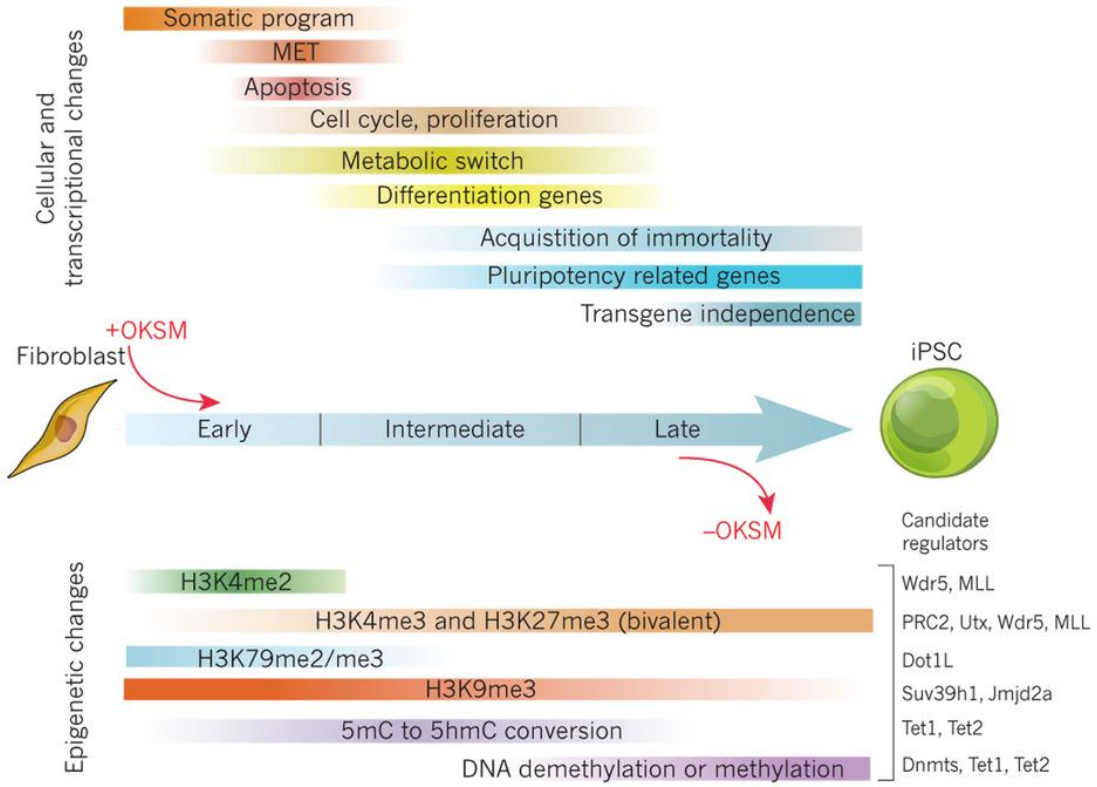

Figure 1.4 Dynamics of key molecular events during reprogramming process (Apostolou and Hochedlinger, 2013).

1.3.3 Reprogramming factors and their replacements

Initially, Takahashi and Yamanaka selected 24 pluripotency genes to assess their capacity to induce pluripotency and found that only the combination of Oct4, Klf4, Sox2 and c-Myc could successfully generate iPSCs, spotlighting their significance in the reprogramming process (Takahashi and Yamanaka, 2006). The OKSM are therefore termed as "Yamanaka factors". Studies in recent year have revealed that OKS function as pioneers for pluripotency induction and maintenance while c-Myc is not absolutely required for reprogramming and seems to act independently of OKS (Soufi et al., 2012; Wernig et al., 2008).

Given its prime requirement in pluripotency (as reviewed in section 1.2.2.1), Oct4 is the pivotal factor in imparting pluripotency. In combination with small molecules,

ectopic induction of Oct4 alone is sufficient for pluripotency induction(Li et al., 2010c). Notably, very few factors can replace Oct4 in iPSC generation, except for DNA hydroxymethylase Tet1 and orphan nuclear receptors Nr5a2, Nr5a1 (Gao et al., 2013; Heng et al., 2010).

As an Oct4-interacting partner, Sox2 generally acts in conjunction with Oct4. Sox2 level is determined in driving reprogrammable cells towards pluripotency acquisition during the later stage of reprogramming (Buganim et al., 2012; Han et al., 2013a). It has been shown to be dispensable for reprogramming using neuronal progenitor cells in which an endogenous Sox2 exists (Eminli et al., 2008). Sox2 could be replaced by other Sox factors or inhibitors of TGFβ-, SFK- pathway and not required for reprogramming of melanocytes and melanoma cells, which collectively suggest a redundant role of Sox2 in reprogramming process (Ichida et al., 2009; Maherali and Hochedlinger, 2009; Nakagawa et al., 2007; Staerk et al., 2011; Utikal et al., 2009a).

Likewise, Klf4 is also not essentially required for pluripotency induction. Klf4 is a Krüppel-like transcription factor and is dispensable for pluripotency maintenance in ESCs (Jiang et al., 2008). It is possible to produce iPSCs without ectopic Klf4 expression in the presence of some small molecules, other Klf members or other factors, but with Klf4 the efficiency could be significantly higher (Chen et al., 2010; Feng et al., 2009; Huangfu et al., 2008b; Lyssiotis et al., 2009; Nakagawa et al., 2007). In reprogramming, Klf4 activates epithelial gene expression program to promote MET at the initial stage (Chen et al., 2010; Li et al., 2010b). However, Klf4 was recently found to counteract Oct4-mediated activation of Mgarp at early reprogramming stage (Tiemann et al., 2014). Hence the role of Klf4 in pluripotency induction needs to be further examined.

c-myc was previously reported to be engaged in pluripotency induction, but it is now accepted that c-myc is only an enhancing, not a necessary factor for producing iPSCs. Its family members, n-Myc and l-Myc are also enhancers for reprogramming (Nakagawa et al., 2007). Although dispensable for reprogramming, the Myc family contribute to multiple cellular processes to facilitate iPSC generation, more likely at the early stage, such as MET initiation, silencing of developmental genes, metabolic switch, cell cycle remodeling, chromatin resetting, etc.(Cho et al., 2010; Nakagawa and Yamanaka, 2012; Neri et al., 2011; Singh and Dalton, 2009; Soufi et al., 2012).

Genomic wide mapping of their binding sites in ESCs and iPSCs reveal that OKS factors are more closely associated, while c-Myc bound promoters are distinct from the OKS group (Chen et al., 2008; Sridharan et al., 2009). In the reprogramming process, it is uncovered that OKS are pioneers in opening and binding chromatin and c-Myc aids in the engagement of OKS to the chromatin (Soufi et al., 2012; Sridharan et al., 2009). Moreover, c-Myc and Klf4 are the major contributors for the initial reprogramming events, particularly the silencing of lineage-specific genes, while Oct4 and Sox2, as well as Klf4, are important for the activation of pluripotency genes and facilitate pluripotency reacquisition at a later stage (Polo et al., 2012; Soufi et al., 2012; Sridharan et al., 2009).

1.3.4 Reprogramming enhancers and inhibitors

The OKSM-mediated reprogramming is a long and low-efficient process and iPSCs that are yielded vary in their differentiation potential. Therefore, many endeavors have been put into improving the quality and efficiency of resulting pluripotent cells. For example, many small molecules have been proven to modulate the reprogramming process. Small molecules, such as HDAC inhibitors (e.g. Valproic

acid), SFK inhibitors (e.g. iPYrazine), GSK3β inhibitors (e.g. Alsterpaullone, CHIR99021), TGFβ inhibitors (e.g. Kenpaullone, E-616452, LY-364947), Vitamin C, PI3K/Akt activator (e.g. PS48) *etc.* can accelerate the reprogramming and are able to replace one or more of the four reprogramming factors (Esteban and Pei, 2012; Huangfu et al., 2008b; Ichida et al., 2009; Li and Rana, 2012; Lyashenko et al., 2011; Staerk et al., 2011; Zhu et al., 2010). Reprogramming with small molecules shows distinctive advantages, such as simpler, transgene-free, controllable, tractable and other features. Furthermore, they make large-scale iPSC production possible. Thus far, they have attracted much interest in steering a faster and efficient reprogramming.

Investigations of the molecular kinetics have demonstrated that reprogramming initiation are largely dependent on pre-existing, accessible chromatin environment so that ectopic inducted reprogramming factors can bind to and exert their specific functions. It is noteworthy that recent studies have highlighted the importance of the interplay between epigenetic factors and reprogramming transcription factors to facilitate the chromatin resetting for pluripotency acquisition. Generally, Inhibition of chromatin condensation, for example, ablation of NuRD/Mbd3 repressor complex, inhibition of HDAC activity, H3K4me3, H3K79me2 (mediated by DOT1L) or H3K9me3 (mediated by SUV39H), strongly increases iPSC induction (Chen et al., 2012; Huangfu et al., 2008a; Luo et al., 2013; Rais et al., 2013). Likewise, factors which have been proposed to promote active chromatin, such as H3K4me3 effector Wdr5 and H3K27me3 "eraser" Utx, facilitate the acquisition of pluripotency (Ang et al., 2011; Mansour et al., 2012).

Besides the canonical "Yamanaka factors" OKSM, many other transcription factors which are involved in pluripotency maintenance, also display emerging roles in

pluripotency acquisition. Nanog, for example, is essential for the establishment of ground state pluripotency during somatic cell reprogramming and its expression is inducted by many reprogramming factors (Silva et al., 2009). Although it is dispensable for somatic cell reprogramming, iPSCs could not be generated from Nanog-null fibroblasts (Schwarz et al., 2014; Silva et al., 2009). On the contrary, factors that are implicated in lineage specification or developmental processes display an inhibitory impact during reprogramming. One example is TGFβ signaling pathway, which is implicated in mESC differentiation and embryo development, counteracts reprogramming (Ichida et al., 2009; Maherali and Hochedlinger, 2009; Puceat, 2007). p53 is also identified as a transcription factor that can inhibit reprogramming process. Studies in recent years have proposed some mechanisms by which p53 inhibits reprogramming, such as inducing cell cycle arrest and apoptosis, inhibiting MET, *etc*.(Banito et al., 2009; Brosh et al., 2012; Wang et al., 2012).

Finally, it is noteworthy that microRNAs can be powerful tools to substitute for Yamanaka factors for efficient reprogramming. Ankye-Danso *et al.* and Miyoshi *et al*. reported that expression of *mir302/367* cluster or mir200c/302s/369 cluster can induce iPSCs rapidly and efficiently in both human and mouse somatic reprogramming (Anokye-Danso et al., 2011; Miyoshi et al., 2011). This finding provides a non-viral, transgene-free mediated procedure for pluripotency induction, which permits high throughput generation of iPSCs.

1.3.5 Characterization of iPSCs

As mentioned above, thus far the quality of iPSCs generated by somatic cell reprogramming varies in their pluripotency potential, and hence characterization of

these iPSCs is necessary to ensure that they are completely reprogrammed and equivalent to ESCs. A checklist of assessments includes:

1. ESC-like morphology

Resembling ESCs, iPSCs should be able to form tight, domed colonies with defined edges on feeder cells; or form flat, epithelial-like colonies growing on matrigel-coated surfaces.

2. Positive for alkaline phosphatase

Pluripotent cells, like mESCs, hESCs, but not EpiSCs, express alkaline phosphatase on their cell membrane. This marker is conventionally used as a quick indicator to examine pluripotency. Given that this marker is not exclusive to ESCs, it must be used in conjunction with other tests.

3. Pluripotency marker expression

A panel of molecular markers have been identified for undifferentiated ESCs (As reviewed in section 1.1). The commonly used markers are Oct4, Nanog, Sox2, Rex1, SSEA-1, Thy-1 for mouse iPSCs and OCT4, NANOG, TRA-1-81, SSEA4, REX1, SOX2 and telomerase for human iPSCs. Generally, examination of three or four markers by immunostaining is sufficient to determine pluripotency. In some cases, promoter methylation analysis of key pluripotency genes are needed, which is to confirm that the promoters of key pluripotency genes have been reactivated.

4. Transgene activity

Expression of transgene is silenced in fully reprogrammed iPSCs. Transgene activity should be checked by qRT-PCR, as residual expression of the reprogramming factors may indicate incomplete or impaired reprogramming (Panepucci et al., 2012).

5. Karyotype analysis

Induction of reprogramming factors subjects the cells to high pressure, resulting in genomic instability. Fully reprogrammed healthy iPSCs should have normal karyotype, i.e. correct number of chromosomes and correct arm morphology.

6. Pluripotency evaluation

The ability to differentiate into all three germ layers, i.e. the endoderm, mesoderm and ectoderm is the key feature of pluripotent cells. Demonstration of their differentiation potential can be achieved by *in vitro* EB formation assay or *in vivo* teratoma formation assay and subsequent analysis using immunohistochemistry, qRT-PCR or histological staining method. Chimaera assay is the most stringent and golden method to evaluate the pluripotency of mouse iPSCs. Pluripotent iPSCs should be competent to produce viable chimeras when they are injected into blastocysts. However, considering that it is time-consuming and not applicable to most laboratories, chimera formation is not a commonly used approach.

1.3.6 Application potentials of iPSCs

The discovery of somatic reprogramming allows the generation of patient-specific stem cells, which circumvents the immune response and ethical issues that embryo-derived stem cells are confronted with. iPSCs show the advantage of an unlimited cell source and the potential to give rise to any cell types for specific clinical purposes.

Patient-specific iPSCs provide a unique and promising platform for cell therapy and disease modeling in the field of regenerative medicine, as presented in Figure 5 (Robinton and Daley, 2012). A patient with neurodegenerative disorder is shown as an example. Briefly, with ectopic expressions of transcription factors, the cells isolated from a skin biopsy of the patient can generate patient-specific iPSCs. These

iPSCs can be maintained and propagated with defined medium *in vitro*. If there is any known gene defect, iPSCs can be modified with genetic engineering. Gene-corrected iPSCs would then be transplant into the affected tissue for treatment. Alternatively, iPSCs can be differentiated into the specific neuronal cell type, which can be used as an *in vitro* model of the patient's disease. This model would subsequently be used for pathological study of the disease, *in vitro* drug screening and evaluation of novel therapeutics.

However, there are several concerns to be addressed before iPSCs can be used clinically.

Firstly and most importantly, the safety of iPSC-derived cells should be treated with caution. Up to now, the generation of iPSCs is more or less involved the virus-mediated induction or genomic integration of oncogenes. Indeed, some studies have report the tumorigenicity of the iPSC-derived cells (Liu et al., 2013b). Significant progress has been made to solve this problem. For example, Warren *et al.* introduced a safer, viral-free and nonintegrating strategy for human cell reprogramming. With repeated administrations of synthetic OKSM mRNA cocktails, multiple human cells can be reprogrammed to iPSCs with efficiencies that greatly surpass established viral methods (Warren et al., 2010). Thus the transient RNA-based protein expression may deliver important clinical benefits for application of iPSCs.

Secondly, The low efficiency and slow kinetics of reprogramming and direct differentiation would introduce genetic alterations and delay clinical use. It is necessary to develop a procedure that can generate iPSC-derived cells in a simpler and faster way.

Thirdly, the purity of iPSCs and iPSC-derived cells could be improved. Both the reprogramming and direct differentiation could be incomplete, thus generating immature cell types. Efficient methods should be designed for purification or quality control of the desired cell population.

Lastly, evaluation of the functionality of iPSC-derived cells is necessary. For transplantation purpose, iPSC-derived cells should be able to integrate, survive and engraft in the damaged tissue. Thus it is required to examine whether these cells are functional that can truly contribute to the recovery of affected tissues.

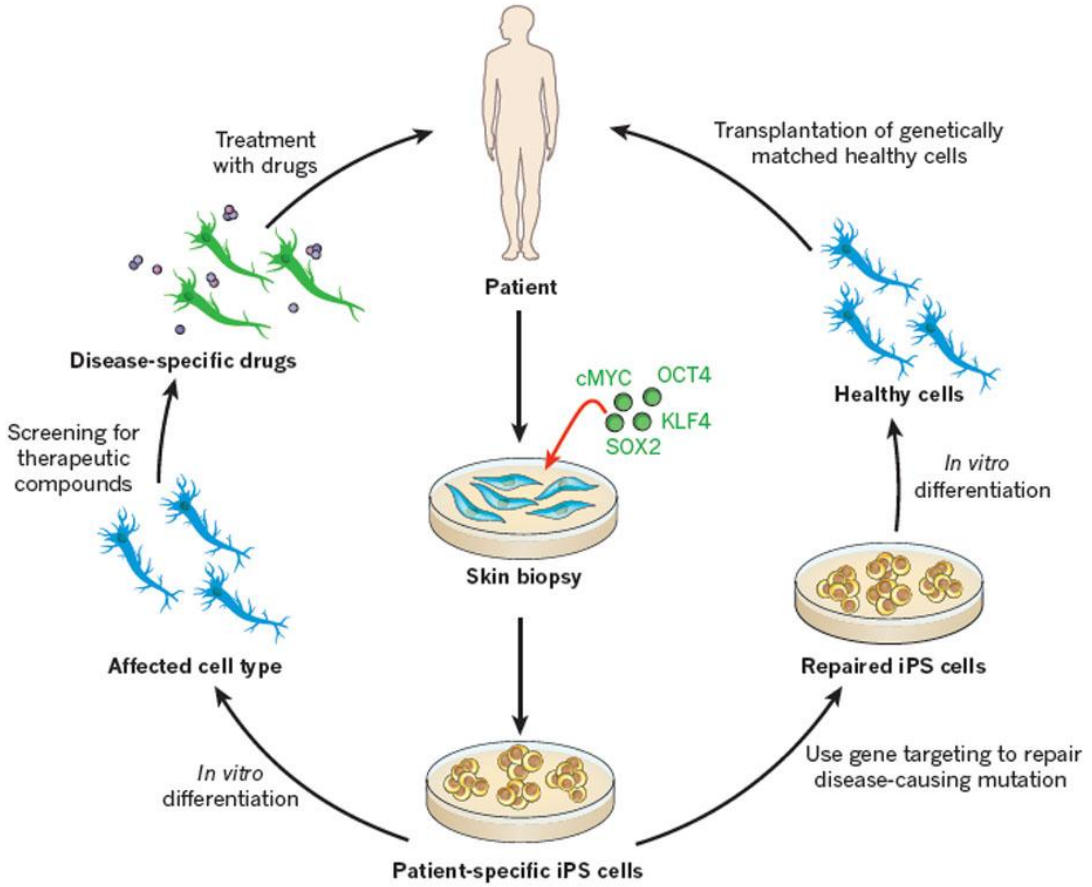

Figure 1.5 Medical application potentials of iPSCs(Robinton and Daley, 2012).

1.3.7 Transdifferentiation

Although iPSCs show remarkable practical implications in regenerative medicine, induced pluripotency is largely restricted by the time length it takes to first reprogram

the cells and then subsequently differentiate them to a specific cell fate, which limits the output of the final cells. In this regard, transdifferentiation arises as a promising alternative approach to solve the issue.

Transdifferentiation, which is also known as direct reprogramming, refers to the conversion from one differentiated cell type to another while bypassing an intermediate pluripotent state or progenitor cell type (Jopling et al., 2011). The first instance of transdifferentiation was reported by Davis *et al.* in 1987. They found that forcing expression of MyoD in MEFs was sufficient to switch these cells to a myoblast cell type (Davis et al., 1987). Yet the progress in this field moved rather slow until the discovery of somatic cell reprogramming by Takahashi and Yamanaka, which has spurred rapid advances in cell transdifferentiation research. Since then there have been multiple reports that achieve such direct lineage conversion in various cell types by transinduction of specific transcription factors, defined medium and even microRNAs (Margariti et al., 2014). Human dermal fibroblasts, for example, can be transdifferentiated to monocytic, erythroid, megakaryocytic and granulocytic lineages via ectopic expression of Oct4 with specific cytokinine treatment (Szabo et al., 2010).

Compared with techniques of somatic reprogramming, transdifferentiation displays several advantages: 1) this process is faster, simpler and higher-yielding, which avoids of the two-step process of reprogramming and lineage commitment; 2) bypassing the pluripotent state, which eliminates the risk of tumorigenicity; 3) potential utilization for cell transdifferentiation *in vivo,* as has been demonstrated in cardiomyocyte induction (Qian et al., 2012). However, so far the underlying mechanism for direct reprogramming remains unclear. Thus detailed analysis of the

properties of directly induced cells and mechanisms of cell transdifferentiation is necessary to advance this technology for future clinical applications.

1.5 Zinc finger proteins

1.4.1 Zinc finger protein family

Zinc finger proteins (Zfps) which usually contain tandem zinc finger domains, are found in DNA binding domains of nearly half of human transcription factors (Messina, 2004). These zinc fingers can be stabilized by a zinc ion, forming a secondary structure for specific cis-regulatory DNA elements binding to regulate gene activity (Dang et al., 2000). Differential use of the Cys and His residues for the Zinc ion interaction gives rise to different subtypes of zinc finger, such as C_2H_2, C_5HC_2, C_2C_2, C_2HC, C_2HC, C_2C_2, *etc.* Of these, C_2H_2 is known as the classical zinc finger and is characterized by the presence of a beta hairpin at the N-terminus and an alpha helix at the C-terminus (Nagashima et al., 2009). A large number of proteins from this family are found to be crucial transcription repressors or activators to regulate gene expressions in diverse cellular processes (Luchi, 2001). However, the functions of majority of these Zfps remain poorly understood.

In the past several years, several studies have reported the novel implications of Zfps in pluripotency in multiple aspects. Zfp206, Zic3 and Sall4, for example, can activate the transcriptions of core regulatory factors *Oct4, Nanog* and have been integrated as key components of the transcription regulatory network (Lim et al., 2007; Wang et al., 2007b; Wu et al., 2006). Zfp296 and Glis1 were reported as enhancers in somatic cell reprogramming via the direct induction of Oct4, Nanog or Myc (Fischedick et al., 2012; Maekawa et al., 2011); while Zfp281 were found to mediate Nanog autorepresssion by recruitment of NuRD complex and thus inhibits the

reprogramming process (Fidalgo et al., 2012). ZFP57, a maternal-zygotic effect protein, is essential for maintaining the DNA methylation imprint in ESCs through interaction with its co-factor KAP1/TRIM28/TIF1β(Zuo et al., 2011). Zfp521, however, has been found negatively regulate pluripotency, as its forced expression in ESCs directs neuronal differentiation (Shen et al., 2011).

Zfp322a and Patz1, our genes of interest in this study, are two proteins from this family, which have shown their emerging roles in ESCs.

1.4.2 Zfp322a

Zfp322a (Zinc finger protein 322a) contains 10 C_2H_2 type zinc finger motifs in total and the amino acid sequences within these domains are highly conserved with other zinc fingers. In human, ZFP322A has a counterparter ZFP322B (Zfp322 psuedogene 1), while mouse not. Using phylogenetic tree analysis, previous studies have shown that mouse Zfp322a is closely related to its human orthologous protein ZFP322A in evolution (Li et al., 2004). Alignment of the amino acid sequences further revealed that human ZFP322A protein is highly identical to mouse Zfp322a (Li et al., 2004). Thus the high conservation of Zfp322a in mammals indicates that Zfp322a could have some essential biological functions. Yet so far little has been known regarding the functions of Zfp322a in mammals.

1.4.3 Patz1

Patz1, also known as Zfp278 or MAZ-related factor (MAZR), contains 7 C_2H_2 type zinc fingers. Besides, it possesses an AT hook DNA binding motif and a BTB/POZ domain, which is essential for protein-protein interactions (Fedele et al., 2000). It

belongs to POK (POZ and Krüppel-like zinc finger) family of transcription repressors. Interestingly, the biological functions of POK proteins are generally associated with other proteins interacting with the POZ domain (Costoya, 2007). Patz1 has been reported as a transcription regulator that can activate *Myc* through its interaction with Bach2, or act as a corepressor that attenuates RNF4-mediated androgen receptor-dependent transcription activation (Kobayashi et al., 2000; Pero, 2001). Through binding with nuclear receptor corepressor (NCoR) complex, Patz1 negatively regulates CD8 expression(Sakaguchi et al., 2010). Patz1 is also found to participate in BCL6-mediated transcription repression by direct interaction with BCL6 (Pero et al., 2012). Therefore Patz1 may function as an architectural transcription factor that can act either as activator or repressor depending on the protein it interacts with.

Indeed Patz1 has displayed diverse functions in various cellular processes. Patz1 may be implicated in carcinogenesis, as its level is up-regulated in colorectal, glioma, testicular and breast tumors (Fedele et al., 2008; Li et al., 2010b; Tian et al., 2008; Tritz et al., 2008). On the other instances, the observations of rearrangement of *PATZ1* allele in small round cell sarcoma, tumor development in Patz1-knockout mice and its involvement in p53 pathway support a potential tumor suppressor role of Patz1 (Mastrangelo et al., 2000; Pero et al., 2012; Valentino et al., 2013a). Besides, Patz1 has an emerging role of inhibiting cellular senescence in endothelial cells and MEFs (Cho et al., 2011; Valentino et al., 2013a). All these results indicate that functions of Patz1 are context dependent.

Interestingly, Patz1 also has a critical role in embryo development. Majority of Patz1-knockout mice underwent prenatal death with severe defects in central nervous system and cardiovascular system; the knockout mice that survived showed a general

growth retardation compared to WT mice (Pero et al., 2012; Valentino et al., 2013b). Patz1 is also a regulator of limb bud development and spermatogenesis (Kobayashi et al., 2000) (Fedele et al., 2008). Given its predominant expression in ICM and ESCs, Patz1 has been identified as an important regulator of pluripotency that is required for maintaining ESC in undifferentiated state (Ow et al., 2014; Yoshikawa et al., 2006). However, the role of Patz1 in reprogramming process is still unknown.

1.5 Purpose and scope

Despite their known functions, here we propose that Patz1 and Zfp322a are potential regulators of pluripotency based on these observations: First, using whole mount in situ hybridization, Yoshikawa *et al.* demonstrated that both Zfp322a and Patz1 are highly expressed in mouse blastocysts and at higher levels in inner cell mass (ICM) than trophoblast (Yoshikawa et al., 2006). Besides, single cell RNA-seq results also suggested higher expression of *Zfp322a* and *Patz1* in $Oct4^{+}$ cells than $Oct4^{-}$ cells (Tang et al., 2010). Secondly, it is noteworthy that ChIP-seq data from Chen *et al.* 's study indicated that many important pluripotency factors, including Oct4, Myc, Klf4 *etc*., bind to the genomic region of Zfp322a and Patz1, which suggests that they are potential targets of other pluripotency factors (Chen et al., 2008). In addition, previous work in our lab has revealed that Patz1 regulates pluripotency genes Oct4 and Nanog to sustain pluripotency in mESCs (Ow et al., 2014).

Together with the known roles of other zinc finger proteins in ES cells that have been reviewed above, we propose that Zfp322a and Patz1 are engaged in pluripotency maintenance or acquisition. However, their functions in pluripotency regulation have been poorly studied.

Therefore, the purpose of this study is to explore novel functions of Zfp322a and Patz1 in pluripotency regulation. In this thesis I aimed to:

1. Investigate the functions of Zfp322a in mESC identity and reprogramming process
    - Examine whether Zfp322a is required for maintaining ESCs in the pluripotent state;
    - Identify the potential targets of Zfp322a through genome-wide ChIP-seq analysis and global gene profile analysis of knocked-down cells;
    - Detect the crosstalk between Zfp322a and other pluripotency factors and dissect its integration within the pluripotency regulatory network;
    - Explore the roles of Zfp322a in the reprogramming process.
2. Intensively study the role of Patz1 in somatic cell reprogramming
    - Determine whether Patz1 overexpression or depletion has an impact on pluripotency induction and check the reprogramming capacity of $Patz1^{+/+}$, $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs;
    - Examine gene expression files of $Patz1^{+/+}$, $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs by microarray analysis and search for novel downstream pathways of Patz1 in MEFs;
    - Investigate the mechanism of Patz1 in reprogramming modulation.

Through investigating the roles of the two novel zinc finger proteins Zfp322a and Patz1, in pluripotency regulation, this study will contribute to a better conception of diverse functions of zinc finger proteins. Genome-wide studies, such as microarray analysis and ChIP-seq, would shed light on their novel functions. Moreover, based on these explorations, we expect to expand the pluripotency regulatory network and gain

more hints on interactions among these pluripotency factors. Overall, the relevant studies on mechanisms of their functions could provide more insights into our current understanding on pluripotency maintenance and reprogramming.

CHAPTER 2

MATERIALS AND METHODS

2.1 Cell culture

Murine ES cells (E14) were cultured in ES cell medium consisting of Glasgow Minimum Essential Medium (GMEM; Invitrogen), 15% ES cell qualified fetal bovine serum (FBS; Invitrogen), 0.055 mM β-mercaptoethanol (Sigma), 100 mM sodium pyruvate (Invitrogen), 0.1 mM MEM nonessential amino acid (NEAA, Invitrogen) and 1,000 units/ml of leukaemia inhibitory factor (LIF) (Millipore).

Platinum-E (Plat-E) cells were maintained in Plat-E medium consisting of Dulbecco's Modified Eagle Medium (DMEM; Invitrogen) containing 10% FBS, 1% penicillin and streptomycin (P/S). For cells transfected with retroviruses, medium were supplemented with 1 μg/ml puromycin (Sigma-Aldrich) and 10 μg/ml blasticidin (Invitrogen).

SNL feeder cells were maintained in GMEM, 10% FBS and 1% Penicillin/Streptomycin. Medium was changed every 2 days and cells were passaged every 2-3 days. Inactivated SNL feeder cells were prepared by incubating the cells in mitomycin C solution (12 μg/ml, Sigma) for 2.5 h at 37 °C in a 5% CO_2 incubator. The inactivated cells were then passaged and seeded at 80% confluence for iPS cell culture.

MEFs were cultured in mESC medium without LIF. For iPSC formation, MEFs that have been infected with retroviruses were maintained in mESC medium without LIF till 5 days post infection and then maintained in Knockout Serum Replacement (KSR) medium. KSR medium contains DMEM, 15% KSR (Invitrogen), 2 mM L- Glutamine

(PAA), 1 mM sodium pyruvate, 1000 units/ml of LIF, 1% P/S, 0.055 mM β-mercaptoethanol and 0.1 mM MEM NEAA.

All the cells were cultured at 37 ℃ in a 5% CO_2 incubator.

2.2 Plasmid construction

For RNAi design and construction of plasmids for shRNA synthesis, Eurofins MWG Operon online software was used to design nucleotide sequence for targeting specific genes. All sequences were analysed by BLAST to ensure specificity and avoid off-target effects. Oligonucleotides were cloned into pSuper.puro vector (Oligoengine).

For overexpression in ES cells, full-length cDNA PCR products were amplified by PCR and inserted into BamH1 and Xho1 site of pPyCAGIP.

For plasmids used in luciferase assays, *Pou5f1* CR4 region and CR1 region was amplified and cloned into the pGL3-Promoter vector (Promega) upstream of the firefly luciferase gene to generate the *Pou5f1* CR4-pSV40-*Luc* and *Pou5f1* CR1-pSV40-*Luc* luciferase reporter plasmids respectively; *Nanog* proximal promoter was amplified and inserted into pGL3-Basic vector to generate the p*Nanog* PP-*Luc* plasmid.

For retrovirus packaging plasmids, full-length cDNA products was amplified by PCR and ligated into MunI and NotI restriction sites of pMX plasmid (Addgene).

The ligated products were transformed into competent DH5α *Escherichia coli* cells by heat shock at 42 ℃ for 90 s and subsequent rescued in 1 ml of Lysogeny broth (LB) medium at 37 ℃ for 45 min before being selected on ampicillin agar plates at 37^{o}C overnight. Single colonies were picked on the next day and plasmids were extracted for sequencing with specific primers to verify the insertion.

Successfully inserted plasmids were isolated using PureYield™ Plasmid Miniprep System (Promega) or PureYield™ Plasmid Midiprep System (Promega) for large scale production according to manuals provided by the kit. In brief, the bacteria were pelleted down, lysed with blue cell lysis buffer, which was stopped by addition of neutralization solution. After centrifuging the cell lysate, the supernatant were transferred to PureYield™ column. After binding and washing, plasmids were then eluted with elution buffer that are provided.

The primers and shRNA sequences being used are listed in the Appendix 1.

2.3 Transfection, RNA extraction, reverse transcription and quantitative real-time PCR

Transfection of ES cells was conducted using Lipofectamine 2000 (Invitrogen) according to the manual provided. For RNAi assays, cells were selected in 6-well culture plate for 3 days using puromycin. ES cells transfected with overexpression vectors were selected using puromycin for 1 week before transferring to 100 mm plates for further selection for another 1 week. Single colonies were picked up and passaged to 6-well dishes. The cells were then harvested for extraction of either protein or RNA.

Total RNAs were isolated using TRIzol Reagent (Invitrogen) followed by purification with RNAeasy Mini Kit (Qiagen) according to the provided protocol. In brief, cells were lysed with TRIzol reagent and protein was removed using chloroform. RNA was precipitated from the water phase by addition of isopropanol. After washed with 75% ethanol, RNA pellet was finally dissolved in diethyl pyrocarbonate (DEPC)-treated water (Ambion). The Superscript III First-Strand Synthesis System with oligo (dT) primer (Invitrogen) was used to convert messenger RNA (mRNA) to complementary

DNA (cDNA). The cDNA was diluted 20 times with nuclease-free water for quantitative real-time PCR, which was performed with CFX96TM Real-Time System (Bio-Rad) and SYBR Green Master Mix (Bio-Rad). Relative expression levels of target genes from sample cDNA were normalized against *β-actin* levels and reflected as a fold change compared to control. For ChIP experiments, relative occupancy values were calculated by determining the apparent IP efficiency (ratios of the amount of ChIP enriched DNA over that of the input sample) and normalized to the level observed at a control region.

All the qPCR primers are listed in Appendix 2.

2.4 Gene expression microarray assay

E14 cells were transfected as described above with plasmids expressing shRNA targeted against either *Zfp322a, Patz1,* or control. Cells were harvested after selection for 4 days. Total RNA was extracted and purified as described above. Then the RNA was diluted to 200 ug/ul and was analysed using Affymetrix Mouse Genome MG430 Plus 2.0 Array according the manufacture's instruction by our collaborator Dr. Li Hu.

Microarray data was processed to extract the representative intensities from each probe set using RMA (Irizarry et al., 2003). Appropriate cut-off values were determined and used to identify differential expression between sample and control groups. Only those differentially expressed genes were subjected for further analysis. Prior to hierarchical clustering, log2 transformation was first performed and the transformed data were subtracted from the mean of the means of the two sample groups. To identify the enriched "Gene Ontology" (GO) terms in the differentially expressed genes, the GO TermFinder was applied (Boyle et al., 2004). For presentation of enriched KEGG pathways in the differentially expressed genes, the

GATHER was used (Chang and Nevins, 2006). The *p* value cut-off of 0.05 was employed for both significant enriched GO terms and KEGG pathways. For overlapping genes between ChIP-seq predicted targets and microarray altered gene targets, two sets of genes were analysed using VLOOKUP functions in Microsoft Excel.

2.5 Chromatin immunoprecipitation (ChIP)

ES cells or MEFs were cross-linked with 1% formaldehyde at 150 rpm for 10 min at room temperature and then quenched by 0.2 M glycine. After washing twice with cold PBS, the cells were harvested by scrapping and centrifuging at 3000 rpm for 15 min at 4^{o}C. The cell pellet was further washed in cold PBS and lysed with SDS cell lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.1 M NaCl, 0.25% Triton X-100) containing protease inhibitor cocktail (Roche). The cell nuclei were pelleted by centrifugation at 2000 rpm for 10 min at 4^{o}C, and then lysed in nuclear lysis buffer (50 mM HEPES-KOH, pH 7.5, 15 0mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1% SDS) with protease inhibitor cocktail. The chromatin was then extracted by spinning down at 20,000 rpm for 30 min at 4^{o}C. The chromatin pellet was subsequently washed twice with ChIP buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) containing protease inhibitor cocktail. Prior to sonication, 1 ml of 0.5 mm glass beads were added to the chromatin extract, which was re-suspended in 8 ml of ChIP buffer. Sonication was carried out at 30% amplitude with pulses of 30 sec on and 30 sec off for about 16 cycles, on the Vibra-Cell VCX750 (Sonics). Size of the sonicated chromatin was determined by de-crosslinking (100 µl chromatin extract, 90 µl TE buffer [10 mM Tris-HCl, pH 8.0, 1 mM EDTA], 30 µl pronase, 200 µl ChIP elution

buffer [50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS]) at 42 ℃ for 2 h followed by 67 ℃ for 6 h. After phenol-chloroform extraction, the size of DNA was determined by agarose gel electrophoresis. If the average size was about 300 to 500 base pairs, the chromatin was continued with immunoprecipitation. Dynabeads Protein G (Invitrogen) beads were washed with ChIP buffer twice before being coated with specific antibody by incubating for 2 h at room temperature. The coated beads were then added to chromatin extract which was pre-cleared by incubation with beads. After overnight incubation, the beads were washed thrice with ChIP buffer, once with ChIP buffer plus 0.35M NaCl, once with ChIP washing buffer (10mM Tris-HCl, pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet P-40 [NP-40], 0.5% sodium deoxycholate), and finally eluted with ChIP elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) while agitating at 1,400 rpm at 68 ℃ for 45 min. The eluent was decrosslinked by pronase. ChIP DNA was then exacted with phenol-chloroform, precipitated and dissolved in 80 ul TE buffer for real-time PCR analysis.

2.6 ChIP-sequencing (ChIP-seq)

For ChIP-seq, ChIP DNA was resupended with 20 μl TE buffer and sent out for sequencing. Briefly, ChIP DNA library was prepared by utilizing the ChIP-seq Sample Prep Kit (Illumina). Sequencing was then performed using the Genome Analyzer IIx (Illumina) and reads were mapped to the M. musculus genome assembly mm9.

ChIP-seq analysis was performed by our collaborative bioinformatics group. In brief, ChIP-seq peak detection was done using Partek software with an average fragment size of 300 bps and 0.05 as the cut-off p-value of Mann-Whitney U test for the separation of forward and reverse reads in a peak. In fact, the Partek software

combined several methods of fragment size estimation, peak identification and peak filtering using the Mann-Whitney U test (Ji et al., 2008; Kharchenko et al., 2008). We further enriched the peaks by using the fold change of Zfp322a peak heights to IgG peak heights (fold change 3 as cut-off), and a minimal Zfp322a peak height at 9 reads as a further cut-off criterion. The final list of the inferred peaks was subjected to motif finding. MEME-ChIP in the MEME suite (http://meme.nbcr.net/meme/cgi-bin/meme-chip.cgi) was applied to the inferred peaks. Clustering of Zfp322a with other transcription factors (TFs) was used to evaluate the similarity of the TF targeting. The co-localization between the TFs was first computed and the correlation coefficients between each pair of co-localization vector were then determined. With the completion of all pair-wise correlation, a correlation matrix was obtained. With the matrix, a heatmap reflecting the hierarchical clustering of the correlation coefficients was generated.

2.7 Co-immunoprecipitation (Co-IP)

Cells were harvested and resuspended in cell lysis buffer (50 mM Tris-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 150 mM sodium chloride (NaCl), 1% NP-40, 10% glycerol) with protease inhibitor cocktail (Roche) for 2 h at 4 ℃. The cold cell lysate obtained was first precleared with Protein G Agarose beads (Invitrogen) and then incubated overnight with beads coated with specific antibody at 4 ℃. The beads were washed for six times using cell lysis buffer before boiled in 80 μl of 2X loading dye for 10 min at 95°C. Western blotting was performed using the supernatant obtained from the boiled beads with specific antibodies. Control IP was performed using anti-IgG antibody (Chemicon).

Antibodies that are used are listed in Appendix 3.

2.8 **Western blotting**

The cells were harvested, washed twice in PBS and re-suspended in Laemmli sample loading buffer (50 mM Tris-HCl, pH 6.8, 2% SDS, 5% glycerol, 1% β-mercaptoethanol, 5 mM EDTA, pH 8.0, 0.002% bromophenol blue) with protease inhibitor cocktail (Roche) by boiling at 95℃ for 5 min. The cell lysate was then centrifuged at 14,000 rpm for 15 min. The supernatant, containing the total protein was harvested. Appropriate amounts of protein were loaded into a 10% sodium dodecyl sulfate (SDS) polyacrylamide gel and ran in SDS-polyacrylamide gel electrophoresis (SDS-PAGE) running buffer (25 mM Tris, pH 8.3, 0.192 M glycine, 0.1% SDS) at 120V. Proteins were then transferred to a methanol-activated polyvinylidene fluoride (PVDF) membrane (Bio-Rad) by running at 320 mA for 3 h in Western Blot transfer buffer (25 mM Tris, pH 8.3, 0.192 M glycine). Subsequently, the membrane was blocked using PBST (0.1% Tween-20 in PBS) with 5% skim milk for 1 h at room temperature, incubated with primary antibody (diluted in PBST with 5% skim milk) overnight at 4^{o}C. The membrane was washed thrice in 0.1% PBST before probed with secondary antibody conjugated to horseradish peroxidase (GE Healthcare) at room temperature for 45 min. After washing thrice with 0.1% PBST, the membrane was then incubated in Immobilon Western Chemiluminescent HRP Substrate (Millipore) for 5 min. The chemiluminescent signal was detected with CL-Xposure Film (Thermo Scientific) in dark.

Antibodies that were used are listed in Appendix 3.

2.9 Immunofluorescence staining

Cells cultured in 24-well dishes were fixed in 4% paraformaldehyde and permeabilized with 0.25% Triton X-100, followed by blocking with 3% BSA in PBS. The cells were then probed with primary antibody in 3% BSA for 1 h at 4 ℃ and secondary antibody conjugates (Life Technology) in 3% BSA for 30 min at room temperature. A drop of Vectashield mounting medium with 4', 6-diamidino-2-phenylindole (DAPI; Vector Laboratories) was placed on the microscope slide and the cover slip was sealed with nail polish in a way that the cells were in contact with the mounting medium. Staining signal was then observed through the Axio Observer A1 inverted light microscope (Zeiss).

Antibodies that were used are listed in Appendix 3.

2.10 Flow cytometry

Cells were collected by centrifugation and resuspended in PBS with 4% formaldehyde for 10 min at 37 ℃. After chilled on ice for 1 min, the cells were resuspended in 90% methanol. The cells were then incubated on ice for 30min for permeabilization. The cells were rinsed with incubation buffer (0.5% BSA in PBS). The cells were incubated with primary antibody overnight at 4 ℃. The cells were rinsed twice and then incubated with secondary antibody for 1 h at room temperature in dark. The cells were rinsed twice before analysed on the flow cytometer (BD FACS Canto). The flow cytometry results were analysed with Flow Software 2.5.0.

2.11 Dual-luciferase assays

Gene-specific shRNA plasmids or overexpression plasmids (600 ng) were cotransfected with *Pou5f1*CR4-*Luc* reporter (600 ng), *Pou5f1* PP-*Luc* or *Nanog* pp-

Luc reporter (600 ng) and an internal control pRL-TK (30 ng, Promega) encoding *Renilla luciferase*. *Firefly* and *Renilla luciferase* activities were measured with the dual-luciferase reporter system (Promega) 72 h post-transfection by Ultra 384 Microplate Reader (Tecan). In brief, after rinsed with cold PBS, the cells were lysed with Passive Lysis Buffer (PLB), and firefly luciferase activity was measured upon addition of the Luciferase Assay Reagent II. Then the Stop & Glo Reagent was added and *Renilla* luciferase activity was measured immediately. The readings generated from cells transfected with RNAi or OE plasmids were calculated as relative to control transfection, after normalization to *Renilla luciferase* readings. Transfections were performed in duplicate and on three independent occasions.

2.12 Alkaline phosphatase (AP) staining

AP staining was performed using Alkaline Phosphatase Detection Kit (Millipore) according to manufacturer's instructions and results were obtained using the Axio Observer A1 inverted light microscope (Zeiss). In brief, cells were fixed in 4% paraformaldehyde in PBS, washed twice in 0.05% PBST (20 mM Tris-HCl, pH 7.4, 0.15 M NaCl, 0.05% Tween-2) and incubated in dark with staining solution (mixture of Fast Red Violet, Naphthol AS-BI phosphate solution and deionized water with a ratio of 2:1:1) for 15 min at room temperature. Cells were then washed with 0.05% PBST and observed with the Axio Observer A1 inverted light microscope (Zeiss).

2.13 Retrovirus packaging, infection and iPSC induction

Plat-E cells were seeded onto a 10-cm tissue culture plate at 50-70% confluency and transfected with specific retrovirus packaging plasmids 4-6 h later. Transfection was performed as normal RNAi assays but in this experiment, 24 μg of plasmid, 60 μl of Lipofectamine 2000, and 1.5 ml of Opti-MEM (Invitrogen) were used instead. Cells

were incubated overnight at 37℃ in a 5% CO2 incubator before changing to fresh medium. Virus-containing medium was collected 48 h post transfection, filtered using a 0.22 μm cellulose acetate filter (TPP) and concentrated 100x using Amicon Centrifugal Filter Units-100 kDa (Millipore) by centrifugation at 3800 rpm for 45 min. The concentrated viruses were stored in -80℃ for infection. *Pou5f1-GFP* MEFs were seeded onto a gelatin-coated 24-well plate at 50-70% confluency 6 h before infection. $Patz1^{+/+}$, $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs were seeded into 24-well pates at a number of $300x10^3$ 6 h before infection. 10 μl of each concentrated retrovirus, supplemented with 8 μg/ml polybrene (Sigma), were added to the MEF cells. MEFs were then passaged onto the inactivated feeder layer 2 days post infection (dpi) and maintained in mESC medium without LIF before replacing with KSR medium at 5 dpi. KSR medium was replaced every day and the number of GFP^+ colonies was counted till the end of the experiment. Alkaline phosphatase staining assay was then performed using Alkaline Phosphatase Detection Kit (Millipore) as described above and results were captured with the Axio Observer A1 inverted light microscope (Zeiss).

2.14 Embryoid body (EB) formation assay

For ES cells, embryoid bodies were formed in suspension culture system. Cells were dissociated as per passaging and re-suspended in LIF withdrawal medium before transferred to Ultra-Low Attachment Surface culture plate (Corning). Embryoid bodies induced from Patz1 over-expressing stable cell line were cultured in LIF withdrawal medium supplemented with puromycin (1μg/ml). Medium was changed every two days and RNAs were extracted from the embryoid bodies at specific days for following quantitative realtime PCR analysis.

For iPS cell differentiation, iPSCs were dissociated as per passaging and re-suspended with LIF withdrawal medium in Ultra-Low Attachment Surface culture plates. After culturing for 4 days in suspension, embryoid bodies were transferred to gelatin-coated coverslips placed in 24-well plates. The adherent EBs were cultured with EB medium (DMEM containing 15% KSR, 2 mM glutamine, 0.1 mM MEM) for another 10 days. Immunostaining was then performed with antibodies against specific lineage markers. Images were captured under a confocal microscope (Olympus FV1000) at 60X magnification. Antibodies that are used are listed in Appendix 3.

2.15 Teratoma assay

iPSCs were suspended at concentration of $1x10^7$ cells/ml and sent to our collaborative lab for teratoma assays. Briefly, 100 µl of suspended cells were injected into the dorsal flanks of SCID mice that were anesthetized with Avertin. Teratomas formed after 2 to 3 months were surgically dissected, fixed, sectioned and analysed with Mallory's Tetrachrome staining. Briefly, the tumors were dissected in PBS, fixed with Bouin's solution for 3 days, and decalcified in Jenkin's solution for three times. The tumors were then embedded in fresh paraffin and sectioned using a microtome. The sections were immersed in toluene, washed with ethanol, stained with Groat solution, Acid Fuschine and Aniline solution, followed by washing with ethanol and toluene. Finally the section was mounted onto the slide for observation under microscope (Zeiss).

CHAPTER 3

ZFP322A REGULATES MESC PLURIPOTENCY AND ENHANCES REPROGRAMMING EFFICIENCY

3.1 Results

3.1.1 Zfp322a is expressed in undifferentiated mESCs.

Chen *et al*. demonstrated that Oct4, Zfx, E2F1, Klf4 and Myc bound to an 800bp-region in the third intron of *Zfp332a* (Chen et al., 2008) (Figure 3.1A). Our ChIP results confirmed the association of Oct4 to this region (Figure 3.1B). This suggested that *Zfp322a* could be a direct target of regulation by these transcription factors. In previous studies Zfp322a was discovered to be expressed at a higher level in ICM compared to trophectoderm (Yoshikawa et al., 2006). We further examined the expression of Zfp322a in mESCs by immunostaining assay (Figure 3.1C), using the antibody specific for Zfp322a protein (Figure 3.1 D). We found that Zfp322a was mainly localized in the nucleus, which indicates that, as a zinc finger protein, Zfp322a may function as a transcription factor in mESCs.

Zfp322a expression during mESC differentiation was next determined. mESCs were induced to differentiate by culturing in LIF withdrawal medium. The level of pluripotency factor Oct4 and Nanog serve as positive controls. As shown in Figure 3.1E, *Zfp322a* mRNA level was reduced during the differentiation process, dropping to 20% at 7 days after LIF removal (Figure 3.1E). Similarly, Zfp322a protein also decreased upon mESC differentiation (Figure 3.1F). The expression of Zfp322a in undifferentiated mESCs and its repression upon ESC differentiation further suggested a possible involvement of Zfp322a in pluripotency maintenance in mESCs.

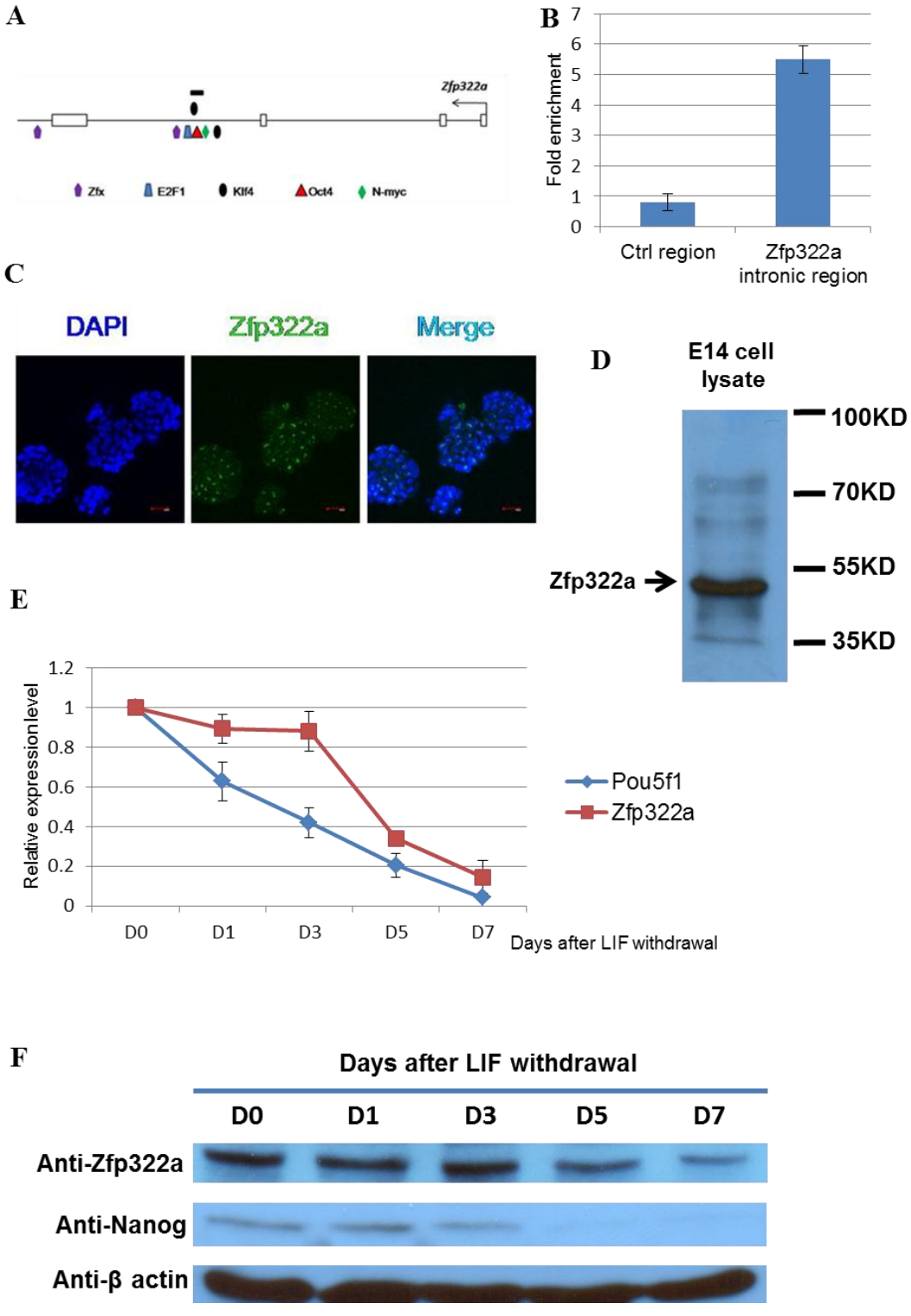

Figure 3.1 Zfp322a is expressed in undifferentiated mESCs. (A) *Zfp322a* intronic region is bound by multiple transcription factors. The black box represents the amplified product from the primer pairs along the intronic region. Open boxes represent exons of *Zfp322*. (B) Oct4 binds to *Zfp322a* intronic region. ChIP DNA with anti-Oct4 antibody was analyzed by qPCR. Fold enrichment was obtained after comparing the values of ChIP DNA to that of the input DNA and normalized against a control region. The control region is an intergenic region on chromosome 6. (C)

Zfp322a is distributed in nucleus and cytoplasm in ESCs. ESCs were stained with anti-Znf322 antibody (green). DAPI (blue) served as nucleus marker. (D) Zfp322a antibody is specific to Zfp322a protein. E14 cell lysate was prepared for western blot assay. (E) qPCR was performed to determine relative levels of *Zfp322a* (red) and *Oct4* (blue) transcripts in ESCs cultured in LIF withdrawal medium for different time periods, compared to ESCs cultured in normal ESC medium and normalised against *β-actin*. (F) Western blot was done to show the change in Zfp322a protein level as mESCs differentiated. β-actin serves as the loading control.

3.1.2 Zfp322a is required for the maintenance of mESC self-renewal and pluripotency

In order to investigate the role of Zfp322a in mESC pluripotency, we examined the effect of *Zfp322a* depletion in mESCs by RNAi. mESC line E14 cells were transfected with two independent *Zfp322a* shRNAs targeting different regions of *Zfp322a* gene. Both shRNAs were effective in depleting the level of *Zfp322a* mRNA to 30% of the control. Upon knock-down of *Zfp322a*, mESCs lost their characteristics, including the round colony-like morphology and AP activity. Instead, RNAi-treated cells exhibited flattened, differentiated cell morphology (Figure 3.2A). These results indicated that *Zfp322a* depletion caused differentiation and impaired self-renewal of mESCs.

We further examined the alteration in gene expression induced by *Zfp322a* depletion. The mRNA levels of pluripotency genes *Pou5f1*, *Nanog*, *Sox2* and *Zfp42* were significantly reduced in RNAi-treated E14 cells (Figure 3.2B). Consistently, protein levels of these pluripotency factors were also reduced upon *Zfp322a* depletion (Figure 3.2C, D). In addition, immunofluorescence (IF) of Oct4 and Nanog were performed to examine their expression at cellular level. The fluorescence intensities of *Zfp322a* knocked-down cells and control cells were measured with flow cytometry. Upon *Zfp322a* RNAi, there was a significant reduction of the fluorescence intensities observed both in anti-Oct4 antibody and anti-Nanog antibody stained cells (Figure

3.2E). Further we examined the population mean value of fluorescence intensities. We found that *Zfp322a* depletion supressed Oct4 and Nanog IF mean intensities by 50%, 40% respectively as compared to control (Figure 3.2F).

Since Oct4 and Nanog are essential for maintenance of pluripotency, these results are consistent with the observation that *Zfp322a* RNAi induces differentiation. The differentiation of mESC was further supported by the dramatic increase in various lineage markers after *Zfp322a* depletion (Figure 3.2G). Upon knock-down of *Zfp322a* in mESCs, we observed an up-regulation of endodermal markers: *Gata6* (4 fold), *Foxa2* (8 fold) and *Sox17* (3 fold), which indicated that Zfp322a could maintain mESC pluripotency by repressing endodermal specification. Trophectoderm maker *Cdx2* displayed a 16 fold increase while mesoderm markers, *Fgf5*, *Hand1* and *Nkx2.5* increased by 7, 8 and 10 fold respectively (Figure 3.2G). Thus consistent with the AP staining results, this suggested that Zfp322a is required to suppress lineage specific gene expressions to maintain mESCs in their undifferentiated state.

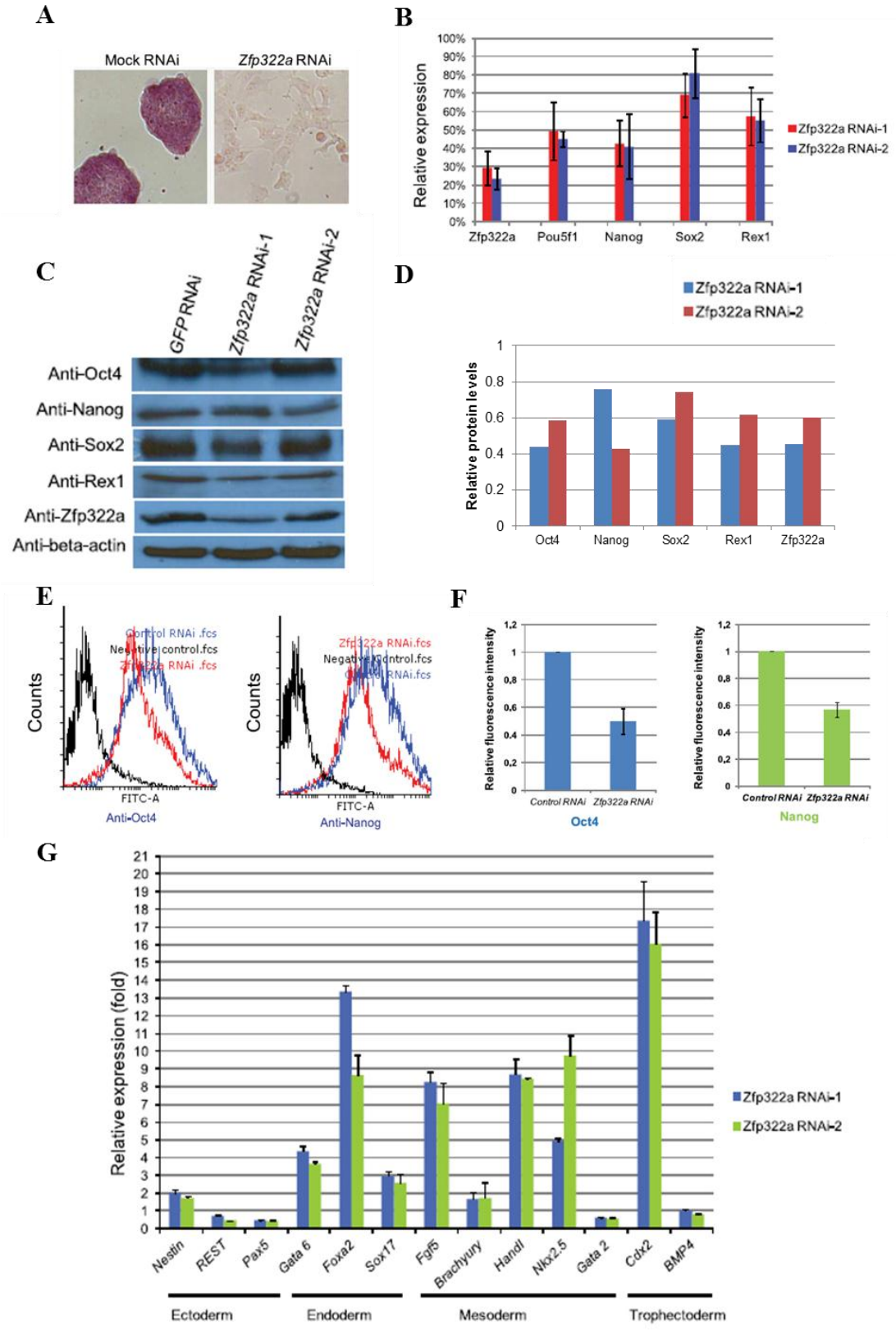

Figure 3.2 Zfp322a is required for maintaining mESC pluripotency. (A) *Zfp322a* RNAi results in AP lose and differentiated morphology. AP staining was conducted on the fourth day of selection after the cells were transfected with *Zfp322a* shRNAs. Zfp322a RNAi cells displayed a lighter colour compared to the dark red colour of mock RNAi cells. (B) Depletion of *Zfp322a* down-regulates mRNA levels of

pluripotency genes Oct4, Sox2, Nanog and Rex1. Two different shRNAs targeting distinct regions of *Zfp322a* (namely *Zfp322a* RNAi-1 and *Zfp322a* RNAi-2) were transfected into ESCs to knock-down *Zfp322a*. ESCs transfected with empty pSUPER.puro vector were used as a control and gene expression levels were normalized against β actin. (C) Knock-down of *Zfp322a* resulted in decreased Oct4, Nanog, Sox2 and Rex1 protein. β-actin served as loading control. (D) Quantification of the protein level changes. Protein levels of *Zfp322a* depleted cells were normalized against β-actin and compared to control RNAi cells using software ImageJ. (E) Representative flow cytometry results showed that fluorescence intensities of Oct4 and Nanog were repressed in *Zfp322a* RNAi cells as compared to control cells. (F) *Zfp322a* depletion led to reductions of Oct4 and Nanog expression at cellular levels. The mean of fluorescence intensity was calculated using software flow software 2.5.0. Relative fluorescent intensities of *Zfp322a* RNAi cells were normalized against control knocked-down cells. Standard deviations were derived from three independent experiments. (G) *Zfp322a* RNAi caused up-regulation of lineage specific markers for endoderm, mesoderm and trophectoderm. Specific primers were used to check the respective gene expression levels by real-time PCR.

3.1.3 Depletion of *Zfp322a* activates developmental genes while repressing pluripotency related genes

To further understand how *Zfp322a* depletion led to mESC differentiation, we used gene expression microarrays to investigate the global gene expression profile changes induced by *Zfp322a* depletion. qPCR experiments were performed to validate the results of the microarray analysis (Figure 3.3A). As an internal control, we examined the level of *Zfp322a* by the microarray, and consistent with real-time PCR results (Figure 3.2B), we found an approximately 3-fold reduction in the mRNA level of *Zfp322a*. Upon *Zfp322a* depletion, 1574 genes were up-regulated (increased by >1.5 fold) and 904 genes were down-regulated (reduced by >1.5 fold) (Figure 3.3B). Importantly, the microarray data analysis revealed that many known pluripotency genes were down-regulated. This indicated that Zfp322a is a high-level regulator in the mESC gene regulatory network, which does not only regulate a subset of genes required for pluripotency, but is an essential component of the core network required for the maintenance of mESC identity.

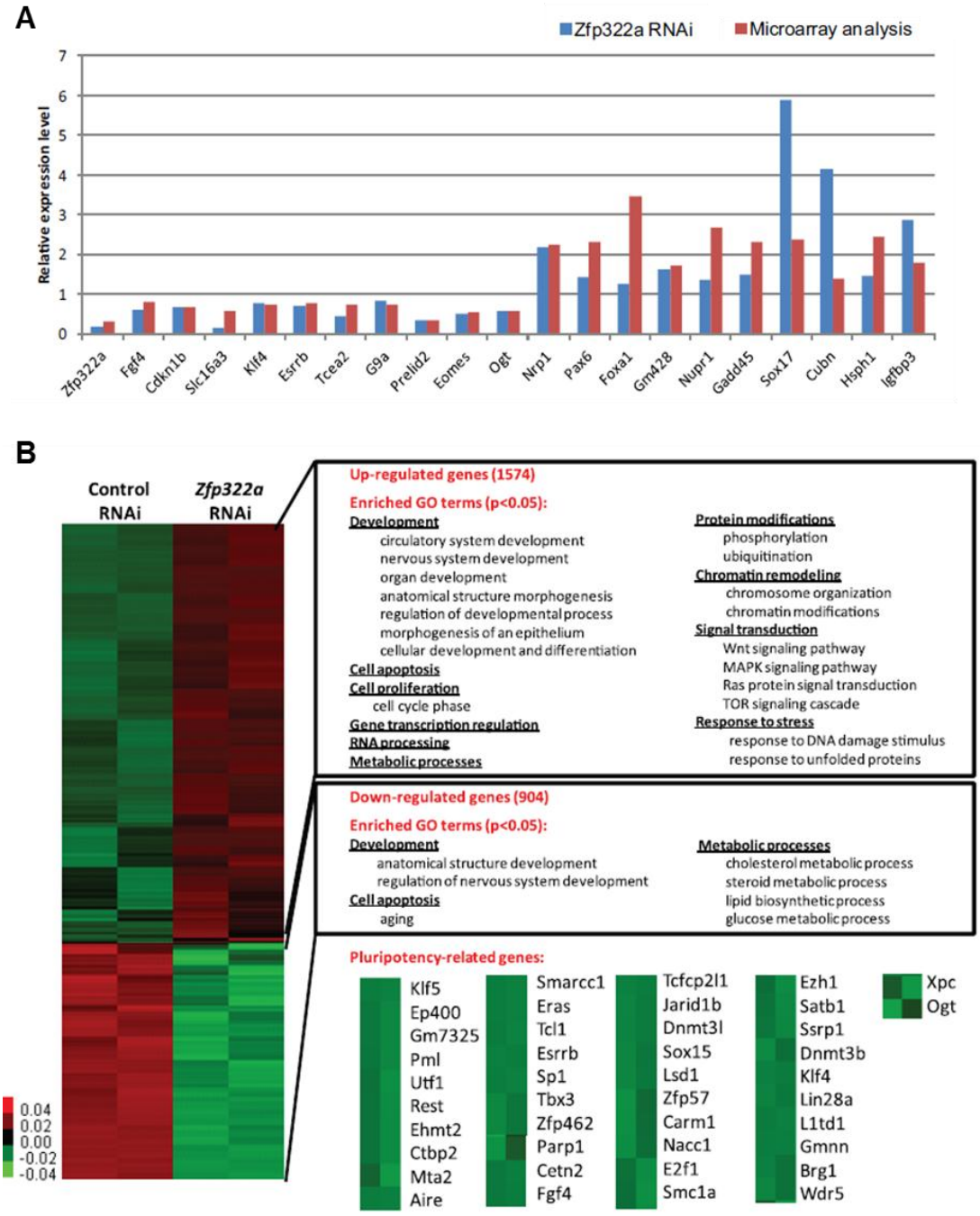

Figure 3.3 Changes of global gene expression upon *Zfp322a* knock-down in mESCs. (A) Validation of gene expression microarray analysis of *Zfp322a* RNAi. 11 down-regulated genes and 10 up-regulated genes were selected from microarray analysis. ESCs transfected with empty pSUPER.puro vector were used as a control and gene expression levels were normalized against *β-actin*. Relative expression level of each gene in qPCR was compared to microarray analysis results. (B) Microarray heat map generated from relative gene expression levels. Zfp322a was knocked down in E14 cells and the cells were selected for 96 hours before whole genome cDNA microarray hybridization was performed. Duplicates were performed to ensure reproducibility of results. Relative highly expressed genes were shown in red and low expressed genes in green. Gene onthology analysis was performed relating to "biological process" for the up- or down-regulated genes respectively. The enriched terms were classified into

several function groups and listed in the figure. Examples of down-regulated pluripotency-related genes upon *Zfp322a* knock-down in ESCs. Genes were selected according to their known functions in pluripotency or ESCs. Each selected gene was taken as individual tiles from the thumbnail-dendogram duplicates.

To determine whether Zfp322a regulated specific types of gene, we conducted gene ontology (GO) analysis of up-regulated and down-regulated genes. The enriched terms were summarized in Figures 3.3B. For the up-regulated genes, many terms related to development were enriched. This is consistent with the role of Zfp322a as a repressor of differentiation. Furthermore, *Zfp322a* depletion activated cell apoptosis related genes, also explained the increased apoptosis in *Zfp322a* RNAi-treated cells. Interestingly, many terms were related to chromosome remodelling, suggesting that Zfp322a may contribute to maintenance of the unique mESC chromatin structure. Notably, several signaling pathways implicated in pluripotency such as MAPK pathway, Wnt signaling pathway, Ras signal cascade were also affected after Zfp322a depletion (Figure 3.3B).

3.1.4 Zfp322a represses MAPK/ERK pathway

It has been shown that inhibition of MAPK/ERK pathway is important for mESC ground state pluripotency (Nichols et al., 2009; Zheng et al., 2008). Activation of Ras-MAPK pathway promotes trophectoderm induction and is specifically suppressed by Oct4 in undifferentiated mESCs(Li et al., 2010a). ERK pathway has been found to promote the onset of mESC differentiation and it is involved in multiple developmental processes (Schübeler et al., 2012). Therefore, Zfp322a may maintain mESC in an undifferentiated state via repression of the MAPK/ERK cascade.

Some important factors involved in MAPK/ERK pathway were found within the gene clusters under the term “MAPK signaling pathway” in KEGG pathway analysis. Fgf1, Kras, Crk and Fgf13, which induces MAPK/ERK pathway, and the MAPK/ERK

pathway downstream targets, Atf2 and Jun, are all up-regulated in Zfp322a depleted cells (Blaukat et al., 1999; Castellano and Downward, 2011; Cohen, 2006; Ouwens et al., 2002; Ramos et al., 2010). In this regard, western blotting was performed to check ERK level in Zfp322a knocked-down cells. Indeed, *Zfp322a* depletion caused an up-regulation of phosphorylated ERK (p-ERK) level while the total levels of ERK (t-ERK) expression was not affected (Figure 3.4B). This substantiated that Zfp322a could be implicated in the repression of MAPK/ERK cascade without changing ERK expression in mESCs.

To further investigate the role of Zfp322a in MAPK pathway, we next examined whether inhibition of the MAPK pathway could rescue the effects caused by *Zfp322a* depletion. *Zfp322a* depleted mESCs were subjected to 50 nM, 250 nM and 1 uM of ERK inhibitors (Ei; PD0325901, Sigma). We found that the addition of ERK inhibitors could not rescue the down-regulated *Pou5f1*, *Sox2* and *Zfp42* upon *Zfp322a* depletion, and the Ei-treated mESCs exhibited differentiated morphology same as DMSO treated control cells (Figure 3.4C). However, with the addition of ERK inhibitors, levels of *Nanog, Sox2 and Rex1* were higher than that of *Zfp322a* depleted cells without ERK inhibitors and the endoderm and ectoderm lineage markers were lower. This is consistent with previous findings in which the blocking of ERK pathway induces elevated expression of Nanog in ESCs (Luo et al., 2012; Wu et al., 2013). Nevertheless, ERK inhibitor did not rescue stem cell differentiation caused by *Zfp322a* knocked-down cells. Interestingly, addition of ERK inhibitors seemed to facilitate *Zfp322a* knocked-down cells to differentiate into mesoderm lineage, other than endoderm or ectoderm lineages cells. This is in consistent with previous finding that *Erk2*-null mESCs tend to differentiate to mesoderm lineage more efficiently than wild type mESCs (Wu et al., 2013).

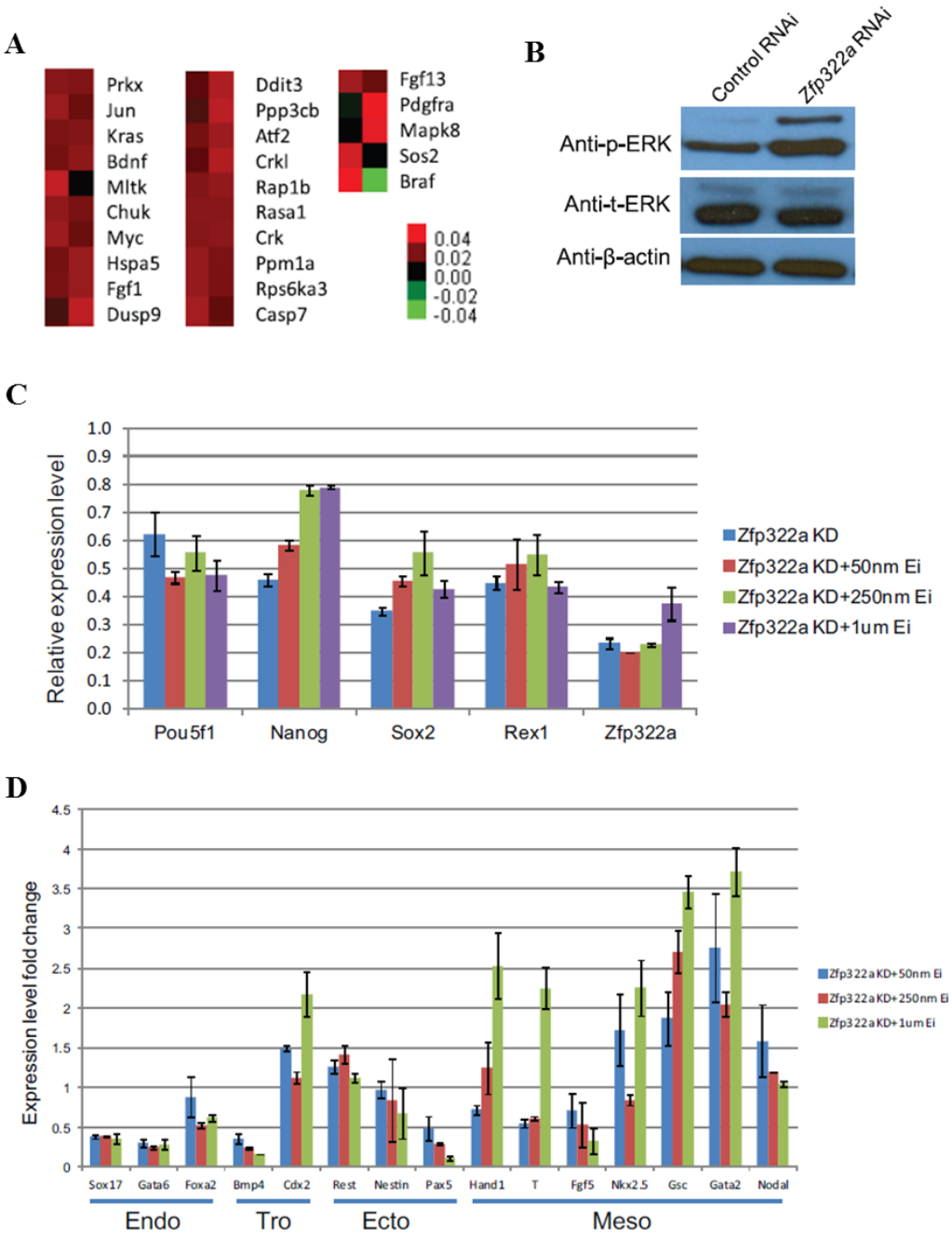

Figure 3.4 Zfp322a represses MAPK pathway in mESCs. (A) List of up-regulated MAPK pathway related genes upon *Zfp322a* RNAi. Genes were selected as they fell into the cluster "MAPK signaling pathway" according to gene ontology analysis for enriched KEGG pathways. Each selected gene was taken as individual tiles from the thumbnail-dendogram duplicates. (B) p-ERK level was elevated in *Zfp322a* depleted cells as compared to control cells, while the total ERK (t-ERK) level was not affected. β-actin served as a control for normalization. (C) ERK inhibitors could not rescue the down-regulated pluripotency genes in *Zfp322a* depleted cells. (**D**) ERK inhibitors could bring down the up-regulated endoderm and ectoderm lineage markers in *Zfp322a* knocked-down cells. Mesoderm markers were elevated with the addition of ERK inhibitors.

Taken together, it appears that *Zfp322a* depletion leads to activation of MAPK/ERK pathway, which could drive mESCs towards differentiation. However, the inhibition of MAPK/ERK pathway could not rescue the differentiation phenotype caused by *Zfp322a* loss, implying that Zfp322a may be implicated in other pathways to sustain mESC pluripotent.

3.1.5 Zfp322a regulates transcriptions of *Pou5f1* and *Nanog*.

Oct4 and Nanog are master regulators of mESC pluripotency (Liu et al., 2007; Niwa et al., 2000a). Many pluripotency factors were found to bind to promoters of *Pou5f1* and *Nanog* to regulate their transcriptions (Chen et al., 2008). Since *Nanog* and *Pou5f1* were down-regulated upon *Zfp322a* depletion, we speculated that Zfp322a may bind to *Pou5f1* and *Nanog* promoters to regulate their transcription.

To test whether Zfp322a binds to cis-regulatory elements of *Pou5f1* and *Nanog*, ChIP experiments were performed using an anti-Zfp322a antibody to pull-down wild type mESC chromatin. Real-time PCR was used to determine whether Zfp322a preferentially bound to known enhancer elements upstream of *Pou5f1* and *Nanog* promoters. We found a clear peak in the *Pou5f1* distal enhancer, which is also known as CR4 (conserved region 4), showing a 28 fold enrichment (Figure 3.5A). CR4 is the main enhancer that drives *Pou5f1* expression in mouse ESCs and early embryos, and it is the site bound by many transcription factors, including Nanog, Sox2 and Oct4 itself (Chew et al., 2005; Loh et al., 2006; Zheng et al., 2008). Similarly, Zfp322a was also shown to bind to the *Nanog* proximal promoter. Strong enrichment for amplicon 4 was found in the multiple transcription factor binding locus of *Nanog* promoter (MTL) (Scotland et al., 2009) (Figure 3.5B). These results showed that Zfp322a could

directly bind to *Pou5f1* distal enhancer and *Nanog* proximal promoter in mESCs and may cooperate with other transcription factors in the regulation of *Pou5f1* and *Nanog* transcription.

To determine whether Zfp322a regulates the transcription of *Pou5f1* and *Nanog*, dual-luciferase assays were performed using two constructs *Pou5f1* CR4-pSV40-*Luc* and p*Nanog* pp-*Luc*. Interestingly, upon knock-down of *Zfp322a*, the luciferase activities were strikingly reduced to 20% and 30% respectively in constructs carrying the CR4 or the *Nanog* proximal promoter (Figure 3.5C, E). To determine whether this reduction was led by *Zfp322a* loss directly, *Pou5f1* CR1 was chosen as negative control. ChIP experiments showed a relatively lower enrichment fold at this region as compared to CR4 (Figure 3.5A). As expected, in the experiment with *Pou5f1* CR1-pSV40-*Luc* construct, the luciferase activity was only reduced by 15% upon *Zfp322a* depletion, much less than the 80% reduction observed in CR4 experiment (Figure 3.5D). These strongly suggested that Zfp322a directly regulated *Pou5f1* and *Nanog* through binding to these cis-regulatory elements. Interestingly, it was also observed that compared to single knock-down of *Pou5f1*, double knock-down of *Pou5f1* and *Zfp322a* further suppressed enhancer activities (Figure 3.5F, G). Given that Oct4 also binds to CR4 and *Nanog* MTL to regulate the transcriptions (Loh et al., 2006), we hypothesized that Zfp322a may cooperate with Oct4 to regulate gene transcriptions.

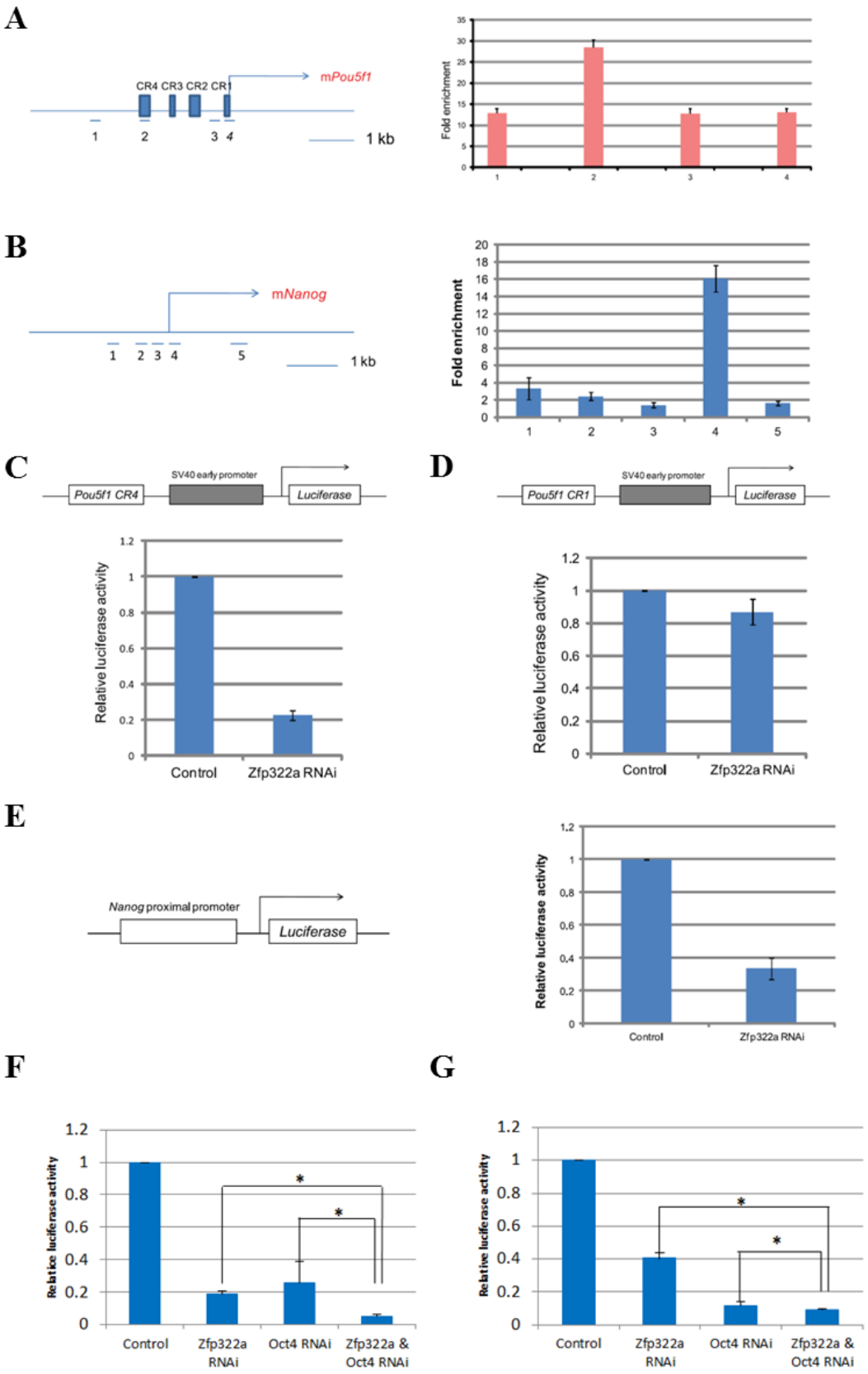

Figure 3.5 Zfp322a positively regulates *Oct4* and *Nanog* transcription. (A) Zfp322a binds to *Oct4* distal enhancer regions. Zfp322a ChIP DNA was analyzed by real-time

PCR. Locations of primers used in qPCR were mapped to the *Pou5f1* genomic region. (B) Zfp322a binds to *Nanog* proximal promoter, with a highest enrichment fold at TSS starting site. Locations of primers were pictured on mouse *Nanog* genomic region. Relative luciferase activities were down-regulated upon *Zfp322a* RNAi using *Pou5f1* CR4-pSV40-*Luc* construct (C) and p*Nanog* PP-*Luc* construct (E), but not *Pou5f1* CR1-pSV40-*Luc* construct (D). Schematic structures of the constructs were presented. Empty pSUPER.puro vector were transfected in ESCs as a control RNAi. Renilla luciferase vector were transfected simultaneously and relative luciferase activities were normalized against Renilla luciferase activity.

3.1.6 Genome-wide mapping of Zfp322a reveals that it is a part of the pluripotency regulatory network

To gain more insights into the downstream pathways through which Zfp322a functions, we identified genome-wide binding sites of Zfp322a in mESCs. Following chromatin immunoprecipitation using anti-Zfp322a antibody to enrich the DNA fragments bound by Zfp322a, we used high-throughput sequencing (ChIP-seq) to analyse the ChIP-enriched DNA. Genomic regions defined by multiple overlapping DNA fragments derived from the ChIP enrichments were considered as putative binding sites. To confirm the validity of these putative binding sites, genomic loci with peaks of various fold changes were arbitrarily selected and tested by qPCR. The final threshold value was determined based on enrichment of 2 fold in qPCR validation (Figure 3.6A), which corresponded to 9-fold or higher enrichment in the ChIP-seq experiment. This gave a total of 4382 putative binding sites of Zfp322a that were associated with 4056 genes.

The location of the binding site within the gene was mapped as well (Figure 3.6B). Notably, after putative Zfp322a binding sites were mapped to nearest genes, 62% fall within 1kb of the nearest gene, showing an obvious preference for TSSs (Figure 3.6C). 19% of the loci were within gene intronic regions, followed by 5’ UTR, distant

promoter (>3kb from TSS) and promoter (<3kb from TSS) which occupied 6% each. Thus, we proposed that Zfp322a is primarily associated with gene promoters.

Among highly enriched binding sites-associated genes, there were many known key components of ESC pluripotency regulatory network, such as *Ino80d, Zfp206, Zfx, Nrf1, Smarce1,* implying that Zfp322a could directly regulate transcription of these pluripotency genes (Appendix 4). To further examine whether Zfp322a targets have preferentially any particular biological functions in mESCs, genes associated with the putative binding sites were subjected to Gene Ontology to search for enriched biological process terms. Large numbers of terms were found to be related to cellular metabolic and biosynthetic processes. Other enriched terms were classified into function groups (Appendix 5). Similar to GO analysis of our microarray data, Zfp322a targets were involved in regulation of gene transcription and translation, especially transcription from RNA polymerase II promoter. Notably, Zfp322a binding sites were found near genes encoding core components of RNA polymerase, such as *Polr2a, Polr2j, Polr3e*. We also found that the targets of Zfp322a were related to developmental processes, implying that Zfp322a may participate in mouse embryo development through regulation of these genes. In addition, many terms surrounding the functions "DNA repair", "protein modifications", "cellular component localization", and "RNA processing", were enriched.

Besides *Pou5f1* and *Nanog*, we sought to refine our prediction of Zfp322a targets by combining ChIP-seq and microarray data in pluripotent mESCs. We analysed the ChIP-seq in concert with microarray dataset. Overlapping genes between these two sets of data indicated that these genes could be potential targets of Zfp322a. We found that 401 of the 1574 up-regulated genes in the *Zfp322a*-RNAi microarray data

analysis were directly repressed by Zfp322a in mESCs ($p<0.05$), while 223 genes were activated directly (Figure 3.6D). Further GO analysis of the directly repressed targets, showed that MAPK pathway related genes was enriched (*p* value=0.006). This reaffirmed our hypothesis that Zfp322a represses MAPK signalling pathway to maintain mESC pluripotency.

Next, we aimed to identify the Zfp322a binding motif. Through bioinformatic computation, we found three different motifs which repeatedly occurred in Zfp322a binding sites; albeit at low frequencies (Figure 3.6E). Motif 1 had the highest frequency, presenting in 9% of all the binding sites. Motif 2 was a 12bp-polyA-sequence with a frequency of 5%. The third motif, which was found in 4% of all binding sites, showed a high similarity to the Oct4/Sox2 binding motif (Chen et al., 2008). This suggested that Zfp322a, Oct4 and Sox2 often bind to the same enhancer element, either acting as a complex or interacting in other ways. Notably, the first and the third element are present in the CR4 element of *Pou5f1* and the proximal promoter of *Nanog*. It is expected that the actual Zfp322a binding site was not identified as a consensus binding motif, given that Zfp322a protein harbours 10 zinc fingers, while only 3-5 zinc fingers were needed for specific DNA binding . Therefore, different zinc fingers of Zfp322a possibly recognise distinct sequences, leading to a wide variety of Zfp322a binding motifs.

In view of the observation that Oct4/Sox2 binding sites tended to be present near Zfp322a binding sites, we compared Zfp322a binding sites with target sequences of other transcription factors mapped in previous studies (Chen et al., 2008). All the transcription factors were clustered according to the similarity of the co-localization with other factors. The results showed that Zfp322a was closer to the Myc cluster

(Figure 3.6F). But with a closer check of the results, Zfp322a actually had a ubiquitous comparable correlation with all the 12 transcription factors, and did not show any significant preference for either Myc or Oct4/Sox2 centred clusters. Indeed Zfp322a showed a slightly higher co-localization frequency with Oct4 cluster target, from which it was inferred that Zfp322a may facilitate or cooperate with Oct4 in mESCs.

Given the correlation between Zfp322a targets and Oct4 targets, similarities between Oct4/Sox2 and Zfp322a binding motifs, together with our observation that both Zfp322a and Oct4 bind to the CR4 region of *Pou5f1* distal promoter and *Nanog* proximal promoter at the same regions, we sought to determine whether Zfp322a could physically interact with Oct4. Co-IP experiments were performed with either anti-Zfp322a antibody or anti-Oct4 antibody. Western blots were then carried out with anti-Oct4 antibody or anti-Zfp322a antibody. We observed an Oct4 band in Zfp322 IP lane and Zfp322a band in Oct4 IP lane, indicating that Zfp322a physically interacts with Oct4 in mESCs (Figure 3.6G). This confirmed our hypothesis that Zfp322a functions as a partner of Oct4 in the regulation of gene transcription, though previous studies did not list Zfp322a as an Oct4 partner (Pardo et al., 2010; van den Berg et al., 2010).

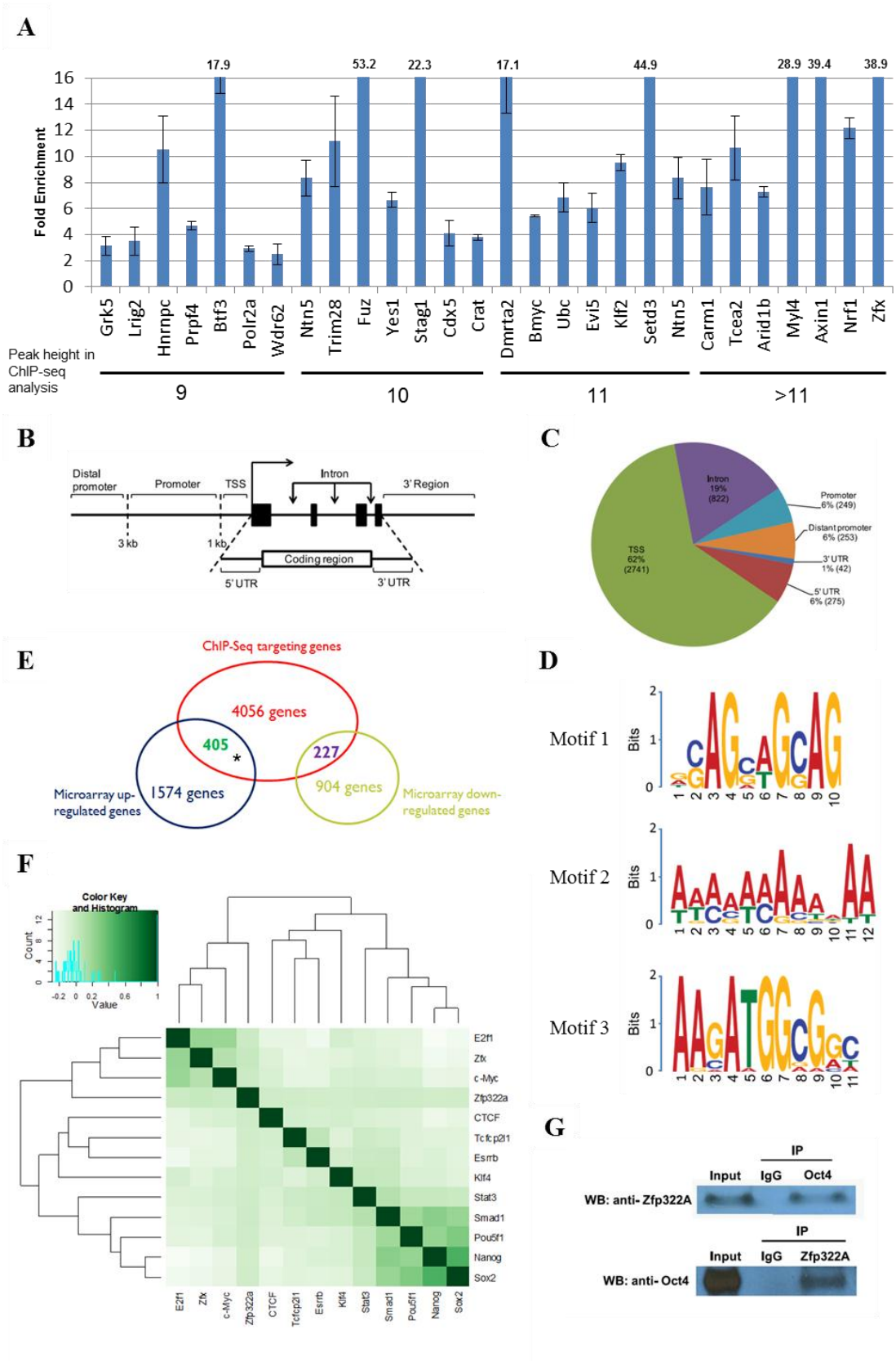

Figure 3.6 Genomic-wide analyses of Zfp322a binding sites. (A) Validation of ChIP-seq data to determine fold change threshold. Genomic loci harbouring peaks with various fold changes were randomly selected from the ChIP-seq data and categorized

into three groups: peak height with 9, 11 or more (>11). These selected loci were validated using qPCR. The resultant enrichment fold were shown in the vertical axis of the graph. (B) Schematic definitions of locations of the putative Zfp322a binding sites relative to the nearest transcriptional unit. TSS referred to -1000 to +100 bp from 5'-end of annotated RNA. (C) Genomic distributions of Zfp322a binding loci. (D) Identification of genes that were predicted to be directly regulated by Zfp322a. The datasets from microarray analysis and ChIP-Seq targets were calculated for overlapping genes. The results revealed 1574 tentative genes that likely were activated ($p<0.05$) and 223 tentative genes repressed directly by Zfp322a. *p* value highlights statistical significance as compared to random chance (CGI scripts based website: Statistical significance of the overlap between two groups of genes) (E) Predicted binding motifs for Zfp322a. Motifs were computationally determined based on the ChIP-Seq data. Three different motifs were identified, namely motif 1, motif 2, motif 3, with frequencies of 9%, 5%, 4% respectively. (F) Zfp322a can be integrated within ESC transcription regulatory network. Shown was co-occurrence of transcription factors at the multiple binding loci. Colours in the heat map reflected the co-localization frequency of each pair of transcription factors (the darker the color was, the more frequently colocalized). All the transcription factors were clustered according to the colocalization frequency with other factors, which was calculated based on their co-occurrence at the same binding loci. (G) Zfp322a can interact with Oct4. Cell lysate of wild type ESCs were immunoprecipitated using either anti-ZNF322A antibody or anti-Oct4 antibody. Western blot was subsequently carried out with anti-Oct4 antibody or anti-ZNF322a antibody. Control IP was performed using anti-IgG antibody.

3.1.7 Zfp322a can enhance OKSM-induced reprogramming of MEFs into iPSCs

Since Zfp322a is involved in mESC self-renewal and pluripotency regulation, it would be interesting to investigate whether overexpression of Zfp322a can enhance OSKM-induced reprogramming or act as a novel reprogramming factor to replace any of the OSKM factors in generating iPSCs. MEFs transfected with a *Pou5f1*-GFP reporter were used to identify putative iPSC colonies (Kim et al., 2008). It was observed that MEFs infected with OSKM plus Zfp322a showed a more efficient and faster reprogramming process than OSKM alone (Figure 3.7A). Addition of Zfp322a could enhance the kinetics of OKSM-induced reprogramming as GFP expressing colonies were detected earlier than OKSM control. The number of iPSCs, when counted as GFP^+ colonies formed by OKSM plus Zfp322a was higher than OKSM

throughout the whole reprogramming process. By day 14 of reprogramming process, the number of GFP^+ colonies generated from OKSM plus Zfp322a was 1.4 fold higher than control. Further examination of these iPSC colonies by AP staining also showed more AP positive colonies formed by OKSM plus Zfp322a as compared to OKSM alone (Figure 3.7B).

Next, the pluripotency of the iPSCs generated by OKSM plus Zfp322a were evaluated. The expression of GFP suggested that *Pou5f1* promoter was strongly reactivated in iPSCs generated from OKSM plus Zfp322a. IF staining results confirmed that these iPSCs expressed pluripotency markers Oct4, Nanog, Sox2 and Rex1. mESC marker SSEA-1 was also expressed in the iPSCs generated from OKSM plus Zfp322a (Figure 3.7C). To further characterize the pluripotency of these iPSCs, embryoid body (EB) formation assays were performed to examine whether these reprogrammed cells were able to differentiate into three germ layers. iPSCs were cultured in suspension to form EBs and then transferred to coated plate with EB differentiation media for 14 days before they were stained with lineage markers. We found that EBs derived from those iPSCs were able to express endoderm marker Gata4, mesoderm marker alpha smooth muscle actin (SMA) and ectoderm marker Nestin (Figure 3.7D). Therefore, it was demonstrated that iPSCs generated from OKSM plus Zfp322a were pluripotent and closely resembled mESCs. Taken together, we concluded that Zfp322a could enhance OKSM-induced reprogramming of MEFs into iPSCs more efficiently and faster.

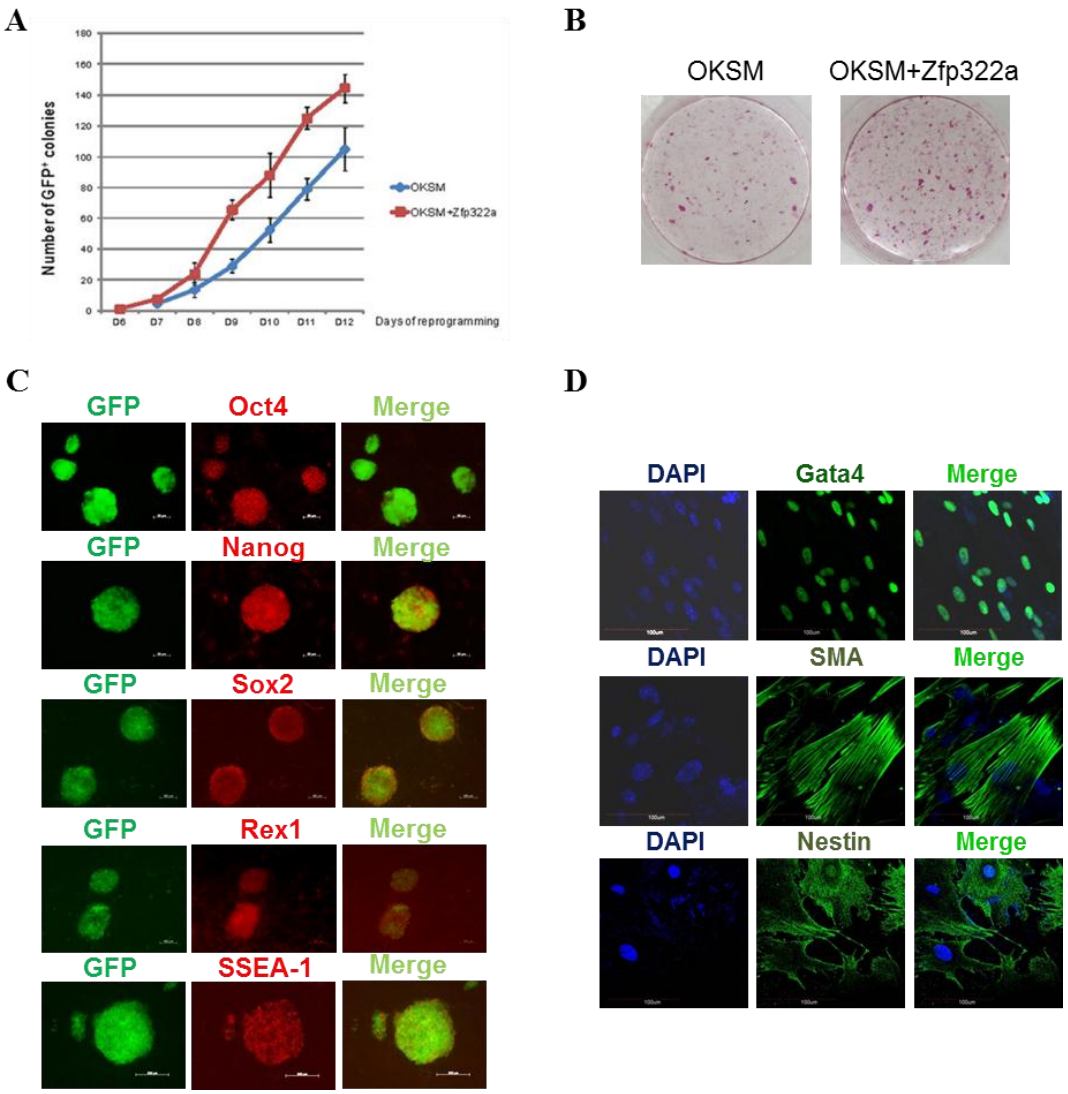

Figure 3.7 Zfp322a can enhance reprogramming inducted by OKSM. (A) Zfp322a enhanced reprogramming efficiency and accelerated the onset of reprogramming process. OSKM serves as control experiment. (B) The iPSCs generated from OSKM plus Zfp322a presented alkaline phosphatase activity. There were more AP stained colonies generated from OKSM+Zfp322a compare to OKSM. (C) The iPSCs expressed endogenous Oct4, Nanog, Sox2, Rex1 and SSEA-1, indicating that they were ES-cell like. Immunostaining using anti-Oct4, anti-Nanog anti-Sox2, anti-Rex1 and anti-SSEA-1 antibodies were performed with GFP^+ iPSCs generated from OKSM+Zfp322a. (D) GFP^+ iPSCs generated by OKSM+Zfp322a were able to express ectoderm, mesoderm and endoderm lineage markers in the EB formation assay. iPSCs were stained with anti-Nestin, anti-Gata4 and anti-alpha smooth muscle actin (SMA) antibodies and pictures were taken at 60X magnification. DAPI (blue) served as nucleus marker.

3.1.8 Zfp322a can replace Sox2 in the OKSM-induced reprogramming

We then investigated whether Zfp322a can replace the core reprogramming factors in addition to enhancement of reprogramming efficiency. Given that c-Myc is dispensable for reprogramming (Nakagawa et al., 2007; Wernig et al., 2008), we only investigated whether Zfp322a can replace any of the OKS factors to generate iPSC colonies from MEFs. Zfp322a was unable to replace Oct4 or Klf4, however, there were GFP^+ colonies observed in the MEFs infected with Zfp322a plus OKM (Figure 3.8A). This indicated that exogenous Zfp322a could replace Sox2 in OKSM-induced reprogramming, albeit at a lower efficiency than OKSM. The first GFP^+ colony generated from Zfp322a plus OKM was observed later as compared to OKSM control. In addition, the expression of GFP was weaker and the number of GFP^+ colonies was fewer (Figure 3.8A, B). Similar to iPSCs formed from OKSM plus Zfp322a, these iPSC colonies were positive for AP staining (Figure 3.8B, D). Further examination of the pluripotency profile of these iPSCs showed that these iPSC colonies could be stained with anti-SSEA-1, anti-Oct4, anti-Nanog, anti-Sox2 and anti-Rex1 antibodies, and were able to express all three lineage markers when they were induced to differentiate in the EB formation assays (Figure 3.8E, F). These suggested that Zfp322a could replace Sox2, but the combination of OKM plus Zfp322a might have relatively slower kinetics in generating iPSCs than OKSM.

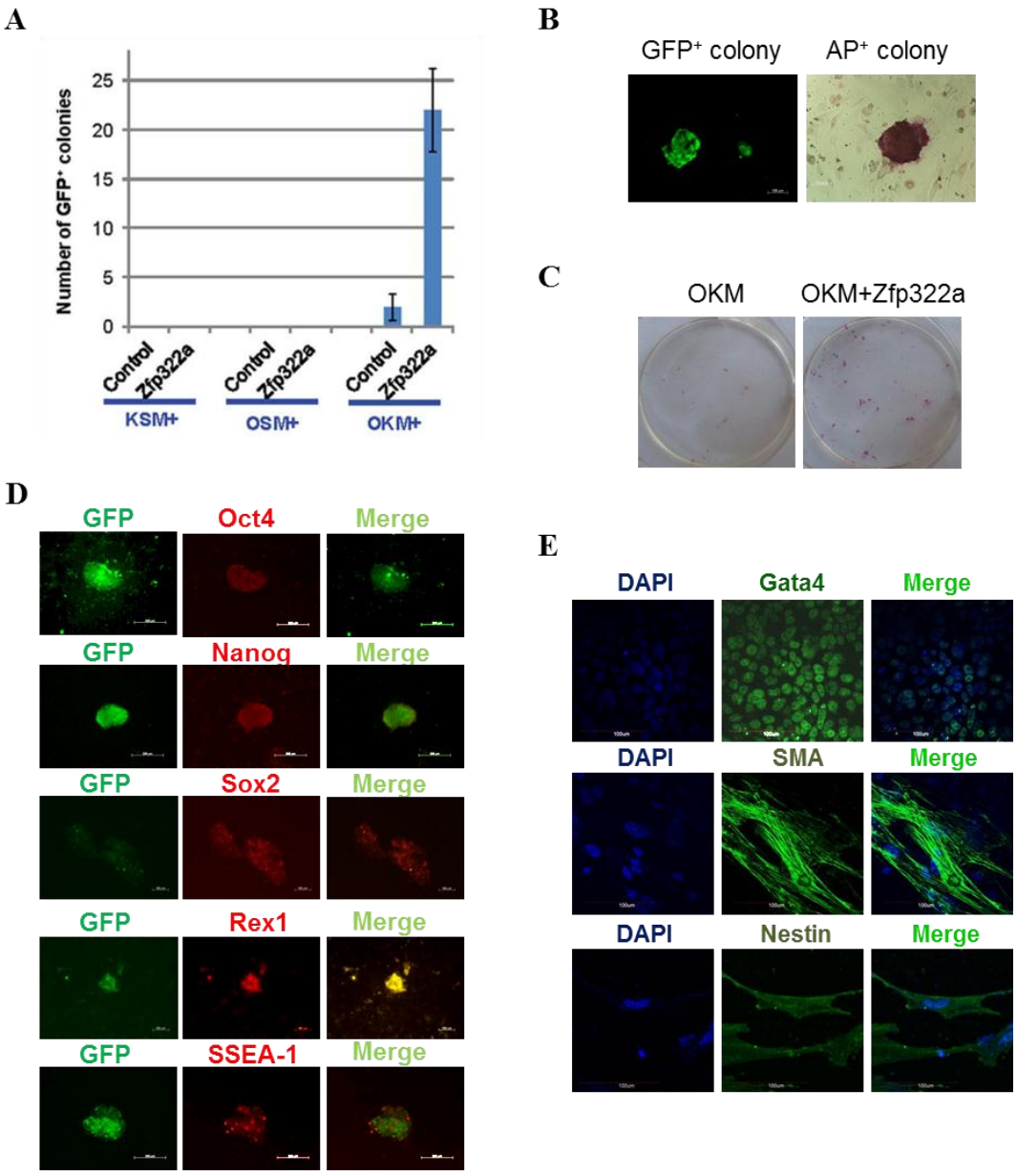

Figure 3.8 Zfp322a can replace Sox2 in OKSM-mediated somatic reprogramming. (A) Zfp322a was able to replace Sox2, but not Oct4 or Klf4 in OSKM reprogramming process. Results from three independent experiments were presented. (B) iPSCs generated from OKM plus Zfp322a expressed weak GFP and were positive for AP activity. (C) iPSCs generated from OKM plus Zfp322a were positive with AP staining and more AP positive colonies were observed in OKM+Zfp322a as compared to OKSM. **(D)** iPSCs generated by OKM plus Zfp322a expressed pluripotency markers Oct4, Nanog, Sox2, Rex1 and SSEA-1. **(E)** iPSCs derived from OKM+Zfp322a could differentiate into ectoderm, mesoderm and endoderm lineages, which were showed by anti-Nestin, anti-Gata4, anti-SMA staining respectively.

3.2 Discussion

3.2.1 Zfp322a is a novel pluripotency factor

The unique properties of ESCs are governed by the master regulators Oct4, Nanog and Sox2, along with a variety of transcription factors (Chen et al., 2008). These transcription factors form a complex network to regulate ESC identity. So far, a lot of transcription factors have been identified to be important for mESC pluripotency and self-renewal, such as Esrrb, Zfp281, Zic3, Sall4, Nr5a2, *etc.* (Festuccia et al., 2012; Fidalgo et al., 2012; Lim et al., 2007; Wu et al., 2006; Heng et al., 2010). Interestingly, depletion of a single transcription factor will alter mESC pluripotency (Chambers and Tomlinson, 2009; Yeo and Ng, 2012). Thus it would be vital to identify novel transcription factors to decipher the delicate transcription regulation of pluripotent state in mESCs. Our research revealed for the first time that Zfp322a is a transcription factor which is important in maintaining mESCs in an undifferentiated state.

Our results have demonstrated that depletion of *Zfp322a* through RNAi induced differentiation of mESCs. The differentiation of mESCs could be due to the suppression of *Pou5f1* and *Nanog* expression after *Zfp322a* was depleted. Zfp322a was shown to actively regulate *Pou5f1* transcription through binding to the CR4 region, and activate *Nanog* transcription via *Nanog* MTL, which are the co-binding sites of other pluripotency factors. Our ChIP-seq and microarray analysis further revealed the binding of Zfp322a to the locus of many other key pluripotency genes that were down-regulated upon *Zfp322a* depletion. It is interesting that Zfp322a can regulate *Pou5f1* and *Nanog* transcription, while Zfp322a itself may also be a target of Oct4 and other pluripotency genes. It appears that these pluripotency factors can form

a regulatory loop within the transcriptional network controlling the pluripotency of mESCs. Thus it can be inferred that Zfp322a is a regulator of mESCs by targeting or possibly cooperating with other pluripotency factors.

Notably over-expression of Zfp322a in mESCs did not significantly change cell morphology. Interestingly, *Nanog* level was activated via Nanog proximal promoter in Zfp322a over-expressing mESCs, while *Pou5f1* level showed no significant change over control (Figure 3.9). Nanog is a well-known core regulator of ESCs, which can sustain ESC pluripotency in the absence of LIF (Chen et al., 2008; Liu et al., 2007). The activation of *Nanog* was also observed in a lot of other pluripotency factors, such as Zfp296, Nr5a2, Zic3 etc. (Heng et al., 2010; Lim et al., 2010; Muotri et al., 2012). All of these factors are required for maintaining ESCs in their undifferentiated state and can induce *Nanog* expression when over-expressed in ESCs. Moreover, our microarray data and ChIP-seq results showed that Zfp322a can repress MAPK/ERK pathway. It is highly possible that Zfp322a, when over-expressed, may serve as an ERK pathway repressor which results in elevated Nanog expression, mimicking high and homogeneous *Nanog* expression in "2i+LIF" ESC culturing media containing ERK inhibitors (Luo et al., 2012; Wu et al., 2013; Zheng et al., 2008). Therefore we hypothesize that Zfp322a can activate Nanog expression either directly or indirectly via MARK/ERK pathway inhibition to maintain the ground pluripotency in mESCs.

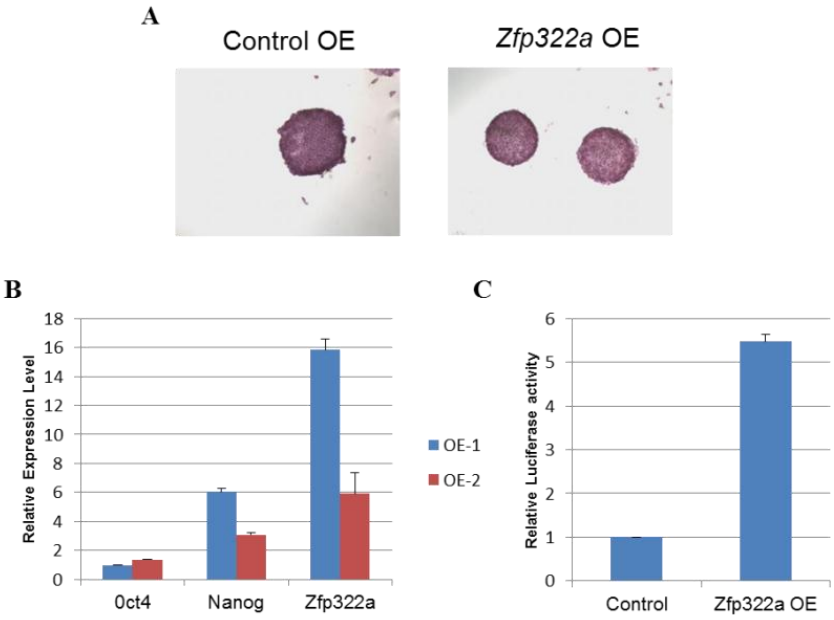

Figure 3.9 Zfp322a overexpression activates *Nanog* transcription. (A) Zfp322a overexpressing cells maintained ESC undifferentiated, displaying ESC typical morphology and AP activity. (B) Zfp322a overexpressing cells showed elevated Nanog expression. ESCs transfected with empty pPCAGIP vector were used as a control and gene expression levels were normalized against β-actin. (C) Zfp322a activates Nanog expression via Nanog proximal promoter. Dual luciferase assay were performed using *Nanog* pp-*Luc* construct in control (empty pPyCAGIP vectors transfected) and Zfp322a overexpressing cells. Renilla luciferase vector was transfected simultaneously and relative luciferase activities were normalized against Renilla luciferase activity.

3.2.2 Zfp322a is integrated within the pluripotency regulatory network

In our results, we also discovered the interaction between Oct4 and Zfp322a and the presence of a similar binding motif for Zfp322a and Oct4/Sox2. This is in concert with the observed higher co-occurrence frequency between Oct4 and Zfp322a binding loci based on the ChIP-seq analysis. Indeed, when we compared the gene expression profile changes after *Zfp322a* RNAi with *Pou5f1* RNAi, a large number of high overlapping targets were identified ($p<0.05$; Figure 3.10) (Loh et al., 2006). Gene ontology analysis of these co-targeting genes displayed a large number of terms related to cellular, organic, embryonic development, cell proliferation and apoptosis,

chromatin remodelling, DNA transcription *etc.* (Appendix 6). In addition, many Oct4-interacting proteins were also affected in the microarray data analysis (Appendix 7). These implied such a close correlation between Zfp322a and Oct4 that Zfp322a may synergize Oct4 functions in maintaining ESCs pluripotency.

Furthermore, the observation of replacement of Zfp322a to Sox2 in OKSM-induced reprogramming verified Zfp322a facilitation of Oct4 functions. Sox2 was discovered as a transcription factor that binds next to Oct4 motif, thus acting synergistically to active gene transcriptions (Boyer et al., 2005; Loh et al., 2006). But subsequent studies have indicated that Sox2 functions redundantly in the activation of Oct-Sox element (Masui et al., 2007). It is then suspected that Zfp322a could have roles partially redundant with Sox2 to interact with Oct4 and participate in Oct4/Sox2 gene regulations. Besides, it is noteworthy that Zfp322a knock-down also depleted expression of *Sox2*, which suggested that Zfp322a is required for *Sox2* activation (Figure 1D, 1E). Therefore, Zfp322a may play a similar function of Sox2 or activate Sox2 expression to mediate its functions.

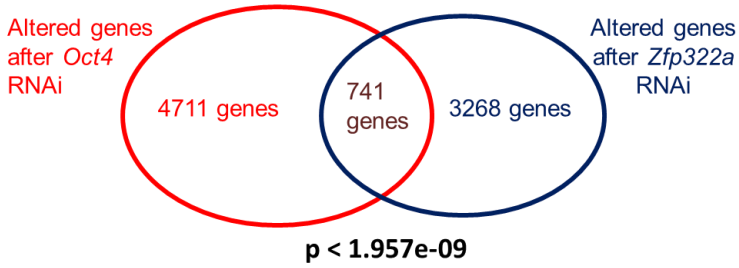

Figure 3.10 Zfp322a share many targets with Oct4. Genes that displayed altered expression levels in gene expression microarray analysis upon *Zfp322a* RNAi were compared to genes altered upon *Oct4* RNAi in previous study. *p* value highlights statistical significance as compared to random chance (CGI scripts based website: Statistical significance of the overlap between two groups of genes).

However, as a Krüppel-like zinc finger transcription factor, Zfp322a may have more complex roles in mESCs since these zinc finger transcription factors have evolved to fill roles in many different biological processes (Huntley, 2006). Indeed, our further analysis of genome-wide Zfp322a target sequence revealed that Zfp322a displayed a non-preferential consensus sequence binding and ubiquitously co-localized with other key pluripotency regulators. Therefore, the ubiquitous association with other factors and diverse implicated functions from microarray and ChIP-seq analysis render Zfp322a functions more complicated. Given that Zfp322a protein contains 10 zinc finger motifs, it might associate with a wide variety of co-factors through different fingers, and therefore its function is depending on the co-factors it interacts with. This pattern may be similar to other C2H2 Zfps harbouring multiple zinc fingers (Brayer and Segal, 2008; Nowick et al., 2013). Taken together, we hypothesize that Zfp322a function as a coordinator that fine tunes the association and recruitments of various factors, including Oct4. It would be interesting to further examine the association of Zfp322a with other components of Oct4/Sox2 regulatory cluster and also Myc cluster proteins to examine their interactions.

3.2.3 Zfp322a may enhance reprogramming efficiency in multiple ways

Considering the relatively low efficiency and considerable time for the OKSM-induced reprogramming process, many studies focused on finding new factors or developing new methods that can accelerate the kinetics of reprogramming process or defining the reprogramming mechanism. Our results from iPSC formation assays identified Zfp322a as a novel reprogramming enhancer that can replace Sox2, thus expanding the current reprogramming code. Sridharan *et al* reported that Zfp322a expression levels were comparable between miPSC and mESC, which was

significantly higher than MEFs and partially reprogrammed iPSCs (Sridharan et al., 2009). Addition of Zfp322a make OKSM-induced reprogramming more efficiently and faster than OKSM alone, suggesting that Zfp322a, as a partner of Oct4 and a key regulator of mESC pluripotency, can accelerate and enhance the efficiency of this process. To our understanding, Zfp322a seems to have a more remarkable role in increasing the portion of GFP^+ colonies than increasing the number of AP^+ colonies. Since the expression of GFP indicates the reactivation of ESC marker *Pou5f1*, this implies that Zfp322a could facilitate the transition of partially reprogrammed AP^+ colonies towards fully reprogrammed GFP^+ colonies.

There are several possible ways in which Zfp322a can enhance reprogramming. Firstly, overexpression of *Zfp322a* has been shown to activate *Nanog* expression. Although Nanog is not one of the canonical quartets of transcription factors used for reprogramming, it is essential for the transition from dedifferentiated intermediates to ground state pluripotency (Silva et al., 2009). Thus the enhancement and acceleration of reprogramming brought by Zfp322a could be partially facilitated by Nanog induction. This may also be the same mechanism shared by Nr5a2 and Zfp296 in their enhancement of reprogramming efficiency (Heng et al., 2010; Muotri et al., 2012). Secondly, reprogramming process consists of down-regulation of lineage specific markers, activation of ESC genes and widespread chromatin remodelling to re-establish the unique chromosomal confirmation of ESCs. As mentioned previously, Zfp322a was shown to repress the lineage specific markers and act as an activator for mESC pluripotency genes. ChIP-seq analysis revealed that Zfp322a has many targets involved in chromosome assembly and modifications. Recruitment of epigenetic modifiers, such as histone acetyltransferases, and inhibition of DNA methytransferases and histone deacetylases, can promote reprogramming by

loosening the condensed chromatin and thus enabling the exogenous reprogramming factors to access and transcribe pluripotency genes and jumpstart the pluripotency transcriptional network. Therefore, Zfp322a could possibly aid in reprogramming by activating directly or establishing a permissive chromatin state to allow the transcriptions of mESC-specific genes. Thirdly, Zfp322a may promote the reprogramming via facilitating Oct4/Sox2 functions. Given our observation that Zfp322a is an interacting partner of Oct4 that can replace Sox2 in the OKSM-induced reprogramming, it can be inferred that Zfp322a has similar functions of Sox2. Fourthly, the suppression of MAPK/ERK pathway is implicated in the predicted Zfp322a direct repressed targets. Although the total ERK level was not affected, the elevated p-ERK level upon *Zfp322a* RNAi indicates that Zfp322a could repress ERK pathway but not ERK expression. ERK pathway has been shown to trigger mESC differentiation (Lanner and Rossant, 2010). Inhibition of this pathway is important for maintaining the ground pluripotent state of mESCs and can improve somatic cell reprogramming efficiency as well(Nichols et al., 2009). Thus Zfp322a could also possibly enhance the reprogramming efficiency via the suppression of MAPK/ERK cascade. Therefore, Zfp322a can maintain mESC properties and promote reprogramming process in many aspects, yet the underlying mechanisms warrant further investigations.

3.3 Future work

Zfp322a is a novel protein which function remains to be further investigated. Therefore more efforts are needed to unravel the deeper mechanism of Zfp322a in pluripotency regulation, as well as its potential roles in other biological processes as

indicated by our genome wide studies, such as transcription initiation, chromatin regulation, MAPK/ERK pathway *etc.*

Our gene ontology analysis of ChIP-seq and microarray data strongly suggested that Zfp322a targets are involved in embryonic development process and related pathways. Zfp322a is implicated as a pluripotency factor because of its relatively high expression in ICM and undifferentiated mESCs. Since the general reduction trend of Zfp322a expression correlates with a commitment to differentiation and a transient up-regulation of *Zfp322a* at third day of differentiation possibly indicates a specific role of Zfp322a in regulating early lineage commitments. We speculate that Zfp322a not only simply maintains mESC in their undifferentiated state but also has certain roles in lineage specifications in embryo development. Therefore, it is important to knockout *Zfp322a* in mESCs, which serves as a better model to study intensively whether the absence of Zfp322a could alter mESC self-renewal and differentiation. Moreover, it would be interesting to generate Zfp322a knockout mice and explore the mutant phenotype in embryos. These investigations would deepen our understanding regarding the functions of Zfp322a.

Although human and mouse ESCs are differed in the signaling networks and epigenetic landscapes, it has been revealed that they share the same core regulators Oct4/Sox2/Nanog and similar transcriptional regulatory network, and the well-known Yamanaka factors OKSM are able to drive reprogramming of both human and mouse somatic cells (Jaenisch and Young, 2008; Schnerch et al., 2010; Takahashi and Yamanaka, 2006; Whitworth et al., 2014). Given that Zfp322a is a conserved zinc finger protein in human and mouse (Li et al., 2004), we propose that Zfp322a is extremely possible to have similar functions regarding the maintenance and

acquaintance of pluripotency in human cells, which is very worthy to be elucidated in future studies.

CHAPTER 4

THE DOSAGE OF PATZ1 MODULATES REPROGRAMMING PROCESS

4.1 Results

4.1.1 Patz1 inhibits reprogramming process

Patz1 has been found predominantly expressed in ICM and ESCs (Tang et al., 2010; Yoshikawa et al., 2006). Previously we have identified Patz1 as an important regulator of pluripotency that is required for maintaining ESC in undifferentiated state (Ow et al., 2014). In view of the versatility of Patz1 functions and its critical role in embryo development (as discussed in chapter 1.4.3), we question its role in reprogramming process.

To explore the role of Patz1 in the reprogramming process, we first added Patz1 together with the canonical reprogramming factors OKSM to infect MEFs. The same *Pou5f1*-GFP MEFs which will express GFP if Pou5f1 promoter is reactivated were used. The number of GFP^+ colonies served as an evaluation of reprogramming efficiency (Kim et al., 2008).

As shown in Figure 4.1A, the number of GFP^+ colonies that were generated with OKSMP infection is 70% less than OKSM control. Alkaline phosphatase (AP) staining also showed a significant reduction of iPS colony formation with the addition of Patz1 (Figure 4.1B). This indicates that Patz1 represses reprogramming process. To confirm the inhibitory role of Patz1 in reprogramming, *Patz1*-knockdown retrovirus was generated and infected MEFs together with OKSM. We observed that MEFs deficient in *Patz1* were reprogrammed with higher efficiency than in WT MEFs, as indicated by both the yield of GFP^+ colonies and AP staining results (Figure 4.1C, D).

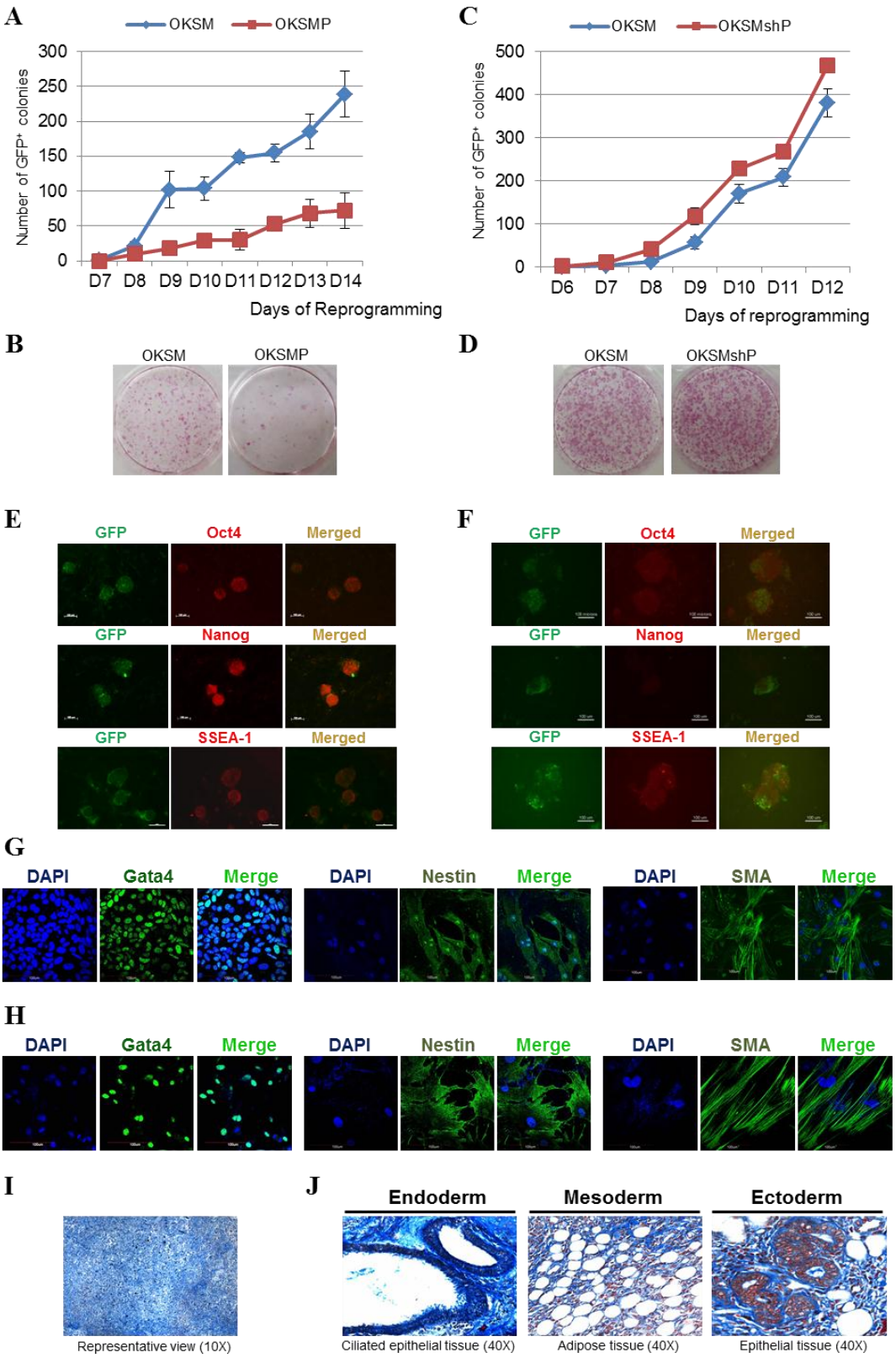

Figure 4.1 Patz1 inhibits OKSM reprogramming process. (A) Addition of Patz1 significantly reduces *Pou5f1* promoter activity induction, measured as GFP$^+$ colonies number, in OKSM mediated reprogramming. The numbers of GFP$^+$ colonies were counted daily from D7 till D14. Data represent means ± SD of two independent experiments. (B) AP staining results indicate that there are less iPSC generated with OKSMP. (C) OKSMshP showed a higher reprogramming efficiency than OKSM control. The numbers of GFP$^+$ colonies were counted

daily from D6 till D12. Data represent means ± SD of two independent experiments. (D) More AP colonies were obtained from OKSMshP than OKSM. Please note that the numbers of MEFs used for each assay were different, therefore giving rise to different number of colonies. Thus OKSM serves as a control for each assay. (E) iPSCs derived from OKSMP express pluripotency marker genes Oct4, Nanog and SSEA1. (F) iPSCs derived from OKSMshP express pluripotency marker genes Oct4, Nanog and SSEA1. (G)iPSCs derived from OKSMP were able to express lineage markers when induced to differentiation in EB formation assays. Green fluorescence showing the expression of Nestin (ectoderm), SMA (mesoderm) and Gata4 (endoderm) were merged with DAPI (blue). The scale bar represents 100 um. (H) iPSCs derived from OKSMshP were able to express lineage markers Nestin (ectoderm), SMA (mesoderm) and Gata4 (endoderm) when induced to differentiation in EB formation assays. The scale bar represents 100 um. (I) Representative view of teratoma sections generated from OKSMP-derived iPSCs. iPSCs derived from OKSMP were able to form tumors in teratoma assays, which however did not have representative tissue structures. (J) iPSCs derived from OKSMshP were able to form teratomas with distinctive structures of three lineages, represented by ciliated epithelial tissue (endoderm), Adipose tissue (mesoderm), epithelial tissue (ectoderm).

iPSCs generated from OKSMP and OKSMshP expressed pluripotency marker genes Oct4, Nanog and SSEA-1 (Figure 4.1E, F). The pluripotency of these iPSCs were next examined by EB formation assays, showing that all these iPSCs were able to express all three germ layer markers upon differentiation (Figure 4.1G, H). However, further teratoma assay found that although OKSMshP-derived iPSCs were able to produce teratomas with representative structure of all three developmental layers, iPSCs generated from OKSMP were defective to differentiate, as they generated tumors containing no distinctive structures of germ layers (Figure 4.1I, J). This is consistent with the previous finding that Patz1 is required to prevent ESC differentiation and its overexpression represses lineage markers during EB formation from mESCs (Ow et al., 2014). Given that teratoma assay, which shows the presence of well-differentiated cells other than the mere detection of lineage-specific markers in EBs, is a more stringent assessment of pluripotency, we suspect that Patz1 overexpression may lead iPSCs more resistant to differentiate, thereby limiting germ layer derivation in teratoma formation (Daley et al., 2009).

4.1.2 Addition of Patz1 blocks reactivation of *Pou5f1* promoter in OKSM-induced reprogramming

To further investigate Patz1's inhibitory role during reprogramming process, we infected *Pou5f1-GFP* MEFs with Patz1 overexpressing retrovirus 0 h, 24 h, 48 h and 72 h after OKSM infection (Figure 4.2A). GFP^+ colonies were counted daily until D12 when AP staining was performed. As shown in Figure 4.2B, compared to OKSM+P/0h, there were more AP^+ colonies formed if Patz1 was added at a later stage of reprogramming process. OKSM+P/72h gave rise to the highest number of AP^+ colonies in Patz1 overexpressed MEFs, but it was still obviously less than OKSM control (Figure 4.2B). This implies that Patz1 is a strong inhibitor for the initiation of reprogramming.

We observed more than 70% reduction in *Pou5f1* promoter activity with addition of Patz1 into OKSM, which was consistent with previous results. Moreover, Pou5f1 promoter activity is still reduced and tends to be further decreased when Patz1 was added 24 h, 48 h or 72 h later, suggesting that Patz1 can significantly repress *Pou5f1* promoter regardless of its addition point (Figure 4.2C). Previous studies have revealed that AP marker can be gained during the initial stage while expression of endogenous-Oct4 is an indicator for a maturation phase of reprogramming (David and Polo, 2014). Therefore, it could be inferred that Patz1 acts not only as an inhibitor for the initiation of reprogramming, but also as a potential roadblock for the progression from the early-intermediates to the maturate iPSCs during reprogramming.

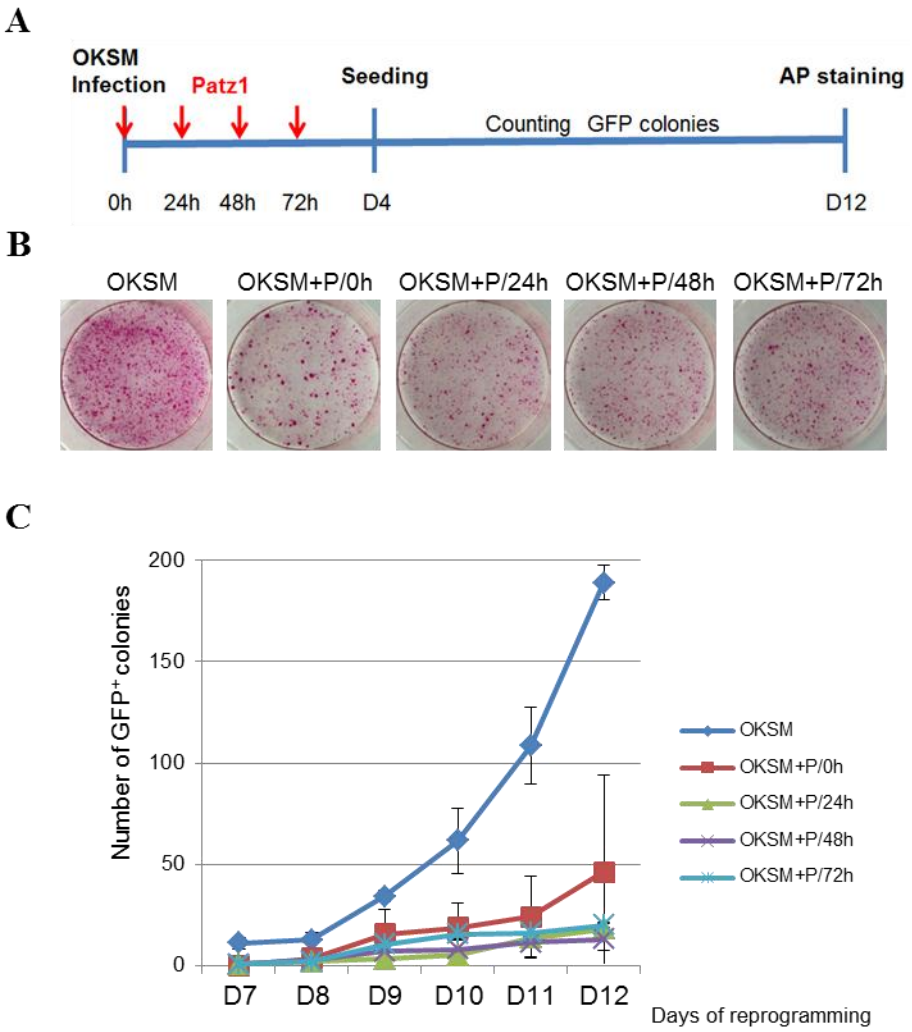

Figure 4.2 Patz1 represses *Pou5f1* reactivation during reprogramming process. (A) Schematic figure to show that *Pou5f1-GFP* MEFs were infected with Patz1 overexpressing retrovirus 0h, 24h, 48h and 72h after OKSM infection. The numbers of GFP^+ colonies were counted everyday till D12, at which time point AP staining assays were performed. (B) Addition of Patz1 at different reprogramming points inhibits AP colony formation. (C) Overexpression of Patz1 represses activation of *Pou5f1* promoter, measured as GFP^+ colonies number, in OKSM reprogramming. Data represent means ± SD of two independent experiments.

4.1.3 Heterozygous knockout of *Patz1* promotes reprogramming

Given that overexpression of Patz1 inhibits reprogramming while knockdown of *Patz1* enhances reprogramming, we surmised that lower Patz1 may facilitate iPSC generation and deletion of Patz1 may further significantly enhance reprogramming.

To test our hypothesis, we next investigated the impact of heterozygous knockout and homozygous knockout of *Patz1* on cellular reprogramming.

Patz1$^{+/-}$ MEFs and Patz1$^{-/-}$ MEFs were generated as previously described (Valentino et al., 2013b). The expression levels of Patz1 were confirmed at both protein and mRNA levels (Figure 4.3A, B). We infected these three types of MEFs with OKSM. As expected, Patz1$^{+/-}$ MEFs produced the greatest number of Oct4^{+} colonies and AP^{+} colonies (Figure 4.3C, D). Surprisingly, iPSC formation was obviously repressed in Patz1$^{-/-}$ MEFs. Similar to previous observations, when Patz1 is overexpressed in WT MEFs, the generation of iPS colonies was inhibited (Figure 4.3E, F). This again reaffirmed the repressive impact of Patz1 in cellular reprogramming.

Patz1$^{-/-}$ MEFs derived iPSCs displayed characteristics of pluripotent cells. They expressed pluripotency markers and were able to differentiate to express all three lineage markers in EB formation assay (Figure 4.3G, H). However, the iPSCs derived from Patz1$^{+/-}$ MEFs were difficult to passage and died gradually. This fragility of iPSCs suggests that the rapid cell division of Patz1$^{+/-}$ MEFs may dilute the effect of virus infection and introduce more variations in the generated iPSCs, which, however, are necessary for a successful reprogramming. Yet how Patz1 dosage impacts on the derivation of healthy iPSCs is a critical issue that warrants further investigations.

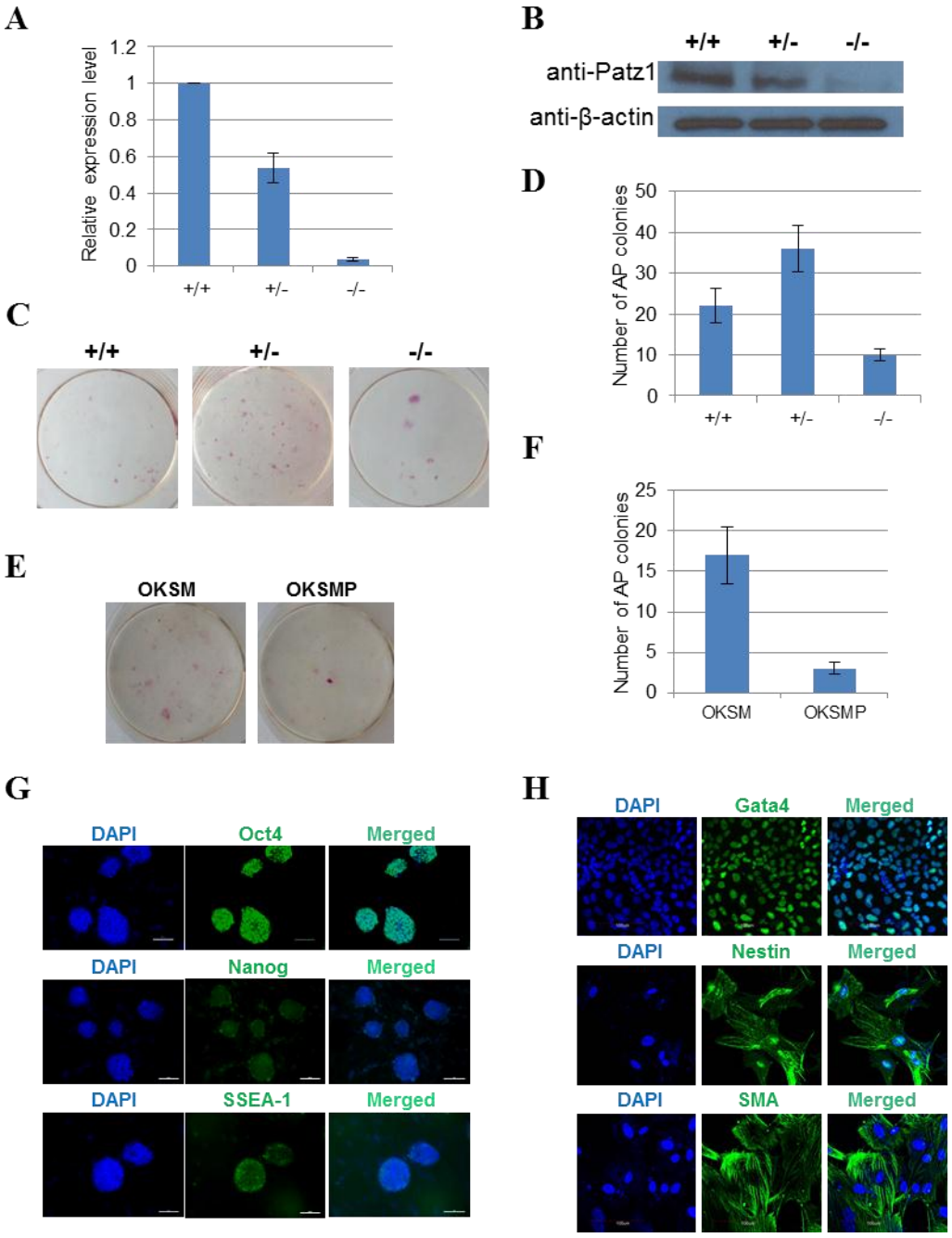

Figure 4.3 Heterozygous knockout of *Patz1* promotes iPSC generation while Patz1-/- MEFs are difficult to be reprogrammed. (A) Relative mRNA level of Patz1 in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs as demonstrated by real time PCR. The relative expression level was normalized against β-actin and Patz1$^{+/+}$ MEFs were used as control. (B) Patz1 protein expression in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs as assessed by western blot. β-actin protein levels were analyzed as a loading control. (C) AP staining results of iPSCs inducted from Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs. Patz1$^{+/-}$ MEFs showed the highest reprogramming efficiency, while lowest number of iPSCs was generated from Patz1$^{-/-}$ MEFs. (D) Quantification of AP colonies generated from Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs. The numbers of AP colonies were counted under microscope after AP staining. Data represent means ± SD of two independent experiments. (E) Representative results of AP staining of iPSC colonies generated by OKSM reprogramming in Patz1$^{+/+}$ MEFs with or without overexpression of Patz1 (F)

Quantification of the experiment shown in e. Data represent means ± SD of two independent experiments.) (G) iPSCs derived from Patz1$^{-/-}$ MEFs expressed pluripotency marker genes Oct4, Nanog and SSEA-1. The green fluorescence indicates expressions of specific proteins. The lower panel shows the figures merged with DAPI (blue). The scale bar represents 100 um. (H) iPSCs derived from Patz1$^{-/-}$ MEFs were differentiated to express lineage markers Gata4 (endoderm), SMA (mesoderm) and Nestin (ectoderm). The scale bar represents 100 um.

4.1.4 Patz1 acts through repression of c-Myc to inhibit iPSC induction

Previous studies have shown that Patz1 acts either as an activator or a repressor for *c-Myc* transcription based on the cellular context. c-Myc, which is involved in many cellular process during reprogramming, can robustly enhance the reprogramming efficiency (David and Polo, 2014; Wernig et al., 2008). We therefore asked whether the inhibitory effect of Patz1 is mediated through regulating c-Myc. We first infected Pou5f1-GFP MEFs with OKSP or OKSshP to examine the impact of Patz1 on reprogramming in the absence of c-Myc retrovirus. As indicated in Figure 4.4, more iPSCs were generated from OKSshP and less iPSCs were derived from OKSP. In addition, the repressive effect of Patz1 seems to be more striking in the absence of c-Myc. This suggests that c-Myc may counteract the inhibitory effect of Patz1 overexpression during reprogramming.

Our previous ChIP-seq results in mESCs revealed that c-Myc is one of the binding targets of Patz1 (OW et al., 2014). ChIP using anti-Patz1 antibody showed that indeed Patz1 was bound to an intronic region of *c-Myc* in MEFs (Figure 4.4E). We next examined the c-Myc level in Patz1 OE, Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs, showing that c-Myc was significantly repressed in Patz1 OE MEFs, whereas it was obviously up-regulated in Patz1 $^{+/-}$ and Patz1$^{-/-}$ MEFs (Figure 4.4 F). Interestingly, the level of p16, a major marker for cellular senescence, and p53 were up-regulated in Patz1 OE MEFs (Figure 4.4G). p53 and c-Myc are known to have opposite roles in diverse

cellular processes (Ceballos et al., 2005; Ho et al., 2005; Zheng et al., 2008; Sachdeva et al., 2009). Based on these results, it appears that the inhibitory effect of Patz1 in the reprogramming process could be partially rescued by overexpression of c-Myc, implying that Patz1 could act through repressing c-Myc to inhibit iPS cell induction.

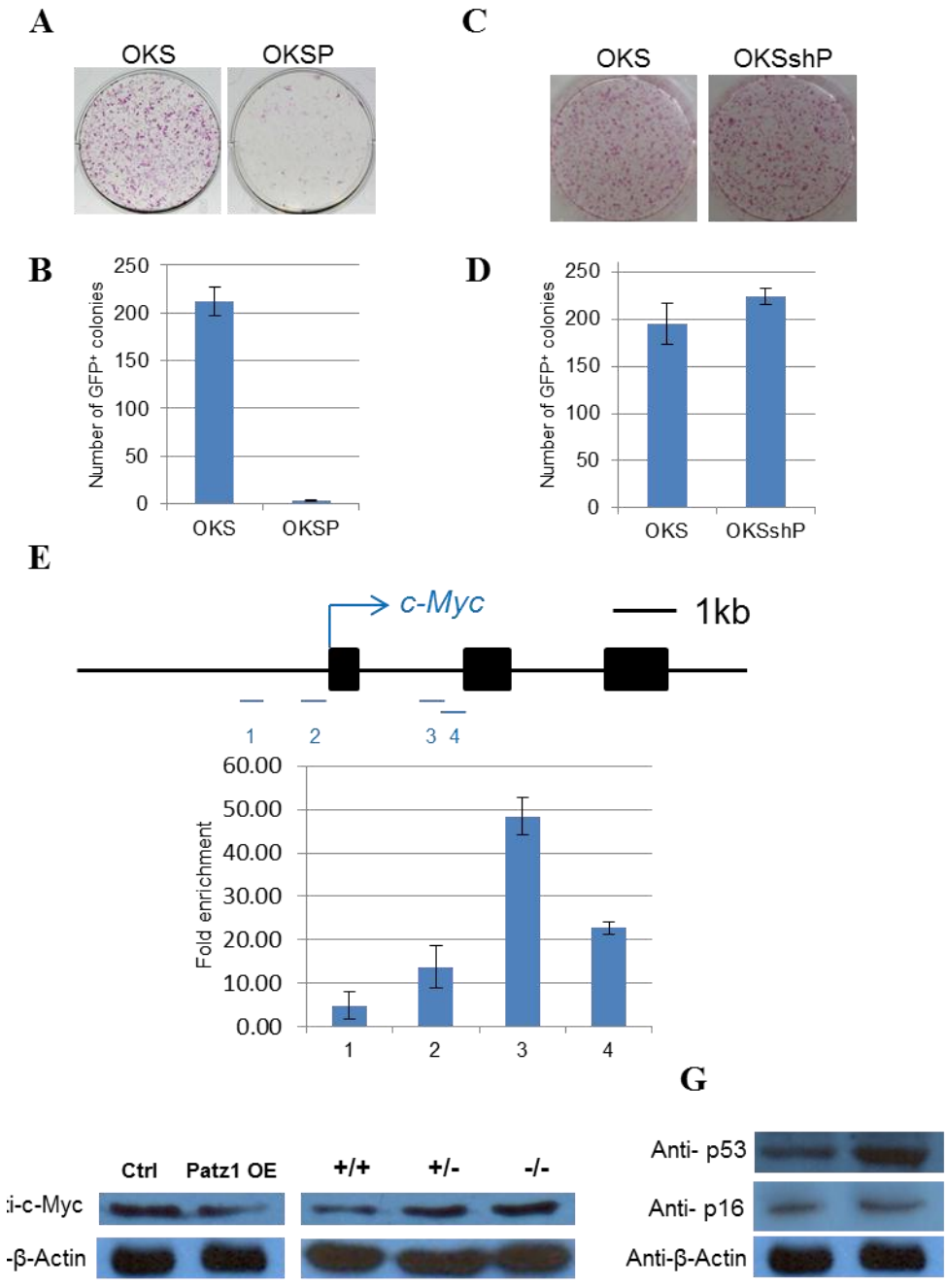

Figure 4.4 Patz1 acts through c-Myc to repress reprogramming. (A) Representative results of AP staining after reprogramming of *Pou5f1-GFP* MEFs with OKS and OKSP. There were less iPSCs generated when Patz1 was added to the OKS cocktail.

(B) *Pou5f1* promoter activity, as assessed by GFP^+ colonies number, was highly decreased in OKSP *versus* OKS inducted iPSCs. (C) *Patz1* knockdown increased AP colony formation in OKS reprogramming. (D) More GFP^+ colonies were generated with OKSshP compare with OKS. (E) Patz1 binds to *c-Myc* in WT MEFs. (F) Overexpression of Patz1 repressed c-Myc protein level in MEFs; while c-Myc expression was up-regulated in $Patz1^{+/-}$ or $Patz1^{-/-}$ MEFs. (G) Overexpression of Patz1 induced up-regulation of p16 and p53. β-actin served as loading control.

4.1.5 Global gene expression analysis of $Patz1^{+/+}$, $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs

To deeper understand how Patz1 is involved in the reprogramming process, we performed gene expression microarray analysis to examine the gene expression profiles in $Patz1^{+/+}$ (WT), $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs. As shown in Figure 4.5A, compared to WT MEFs, 1286 genes were down-regulated by more than 40% and 740 genes were up-regulated by more than 2 fold in $Patz1^{+/-}$ MEFs. The same cutoff values gave 1353 down-regulated genes and 677 up-regulated genes in $Patz1^{-/-}$ MEFs (Appendix). These up-regulated or down-regulated genes were further grouped into 2 clusters according to their different levels in $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs respectively. Gene ontology (GO) analysis was conducted for each cluster of genes (Figure 4.5A). Transcription regulation was one of the enriched terms in both up-regulated and down-regulated genes, which supports the role of Patz1 as a transcription regulator. Many genes related to development, cell specification or differentiation was found to be altered, reaffirming the requirement of Patz1 for a proper embryo development. Many neurophysiological terms are enriched, which is consistent with previous finding that Patz1 is essentially involved in nervous system development and functions (Valentino et al., 2013b). Interestingly, a number of MET inducing genes were up-regulated and genes facilitating EMT were down-regulated in $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs, indicating that Patz1 could be a potential regulator in MET induction during reprogramming process (Figure 4.5B).

Moreover, some genes related to chromatin organization and epigenetic regulation of gene expressions were found within the up-regulated cluster. Interestingly, expressions of many histone deacetylases and acetylases were altered (Figure 4.5C). Histone deacetylases, such as Hdac2, Hdac4, Hdac11, were generally down-regulated upon Patz1 loss, while histone acetylase Hat1, Kat2a were up-regulated. This implies that Patz1 may crosstalk with many epigenetic factors and influence chromatin modification.

It is noteworthy that there were a group of genes that were reversely changed between $Patz1^{+/-}$ and $Patz1^{-/-}$ MEFs, as indicated in cluster iii in Figure 4.5A. GO analysis indicated that these genes, which were down-regulated in $Patz1^{+/-}$ MEFs but up-regulated in $Patz1^{-/-}$ MEFs, were enriched in programmed cell death and apoptosis (Figure 4.5D). Indeed previous studies have shown that $Patz1^{-/-}$ MEFs usually underwent premature cellular senescence and grew significantly slower compared to their WT counterparts, while $Patz1^{+/-}$ MEFs showed a much faster growth rate than WT MEFs (Figure 4.6A). Overexpression of Patz1 did not show any significant change in cell proliferation rate (data not shown). Therefore, we concluded that $Patz1^{-/-}$ MEFs are distinguished from $Patz1^{+/-}$ MEFs by severe cell senescence.

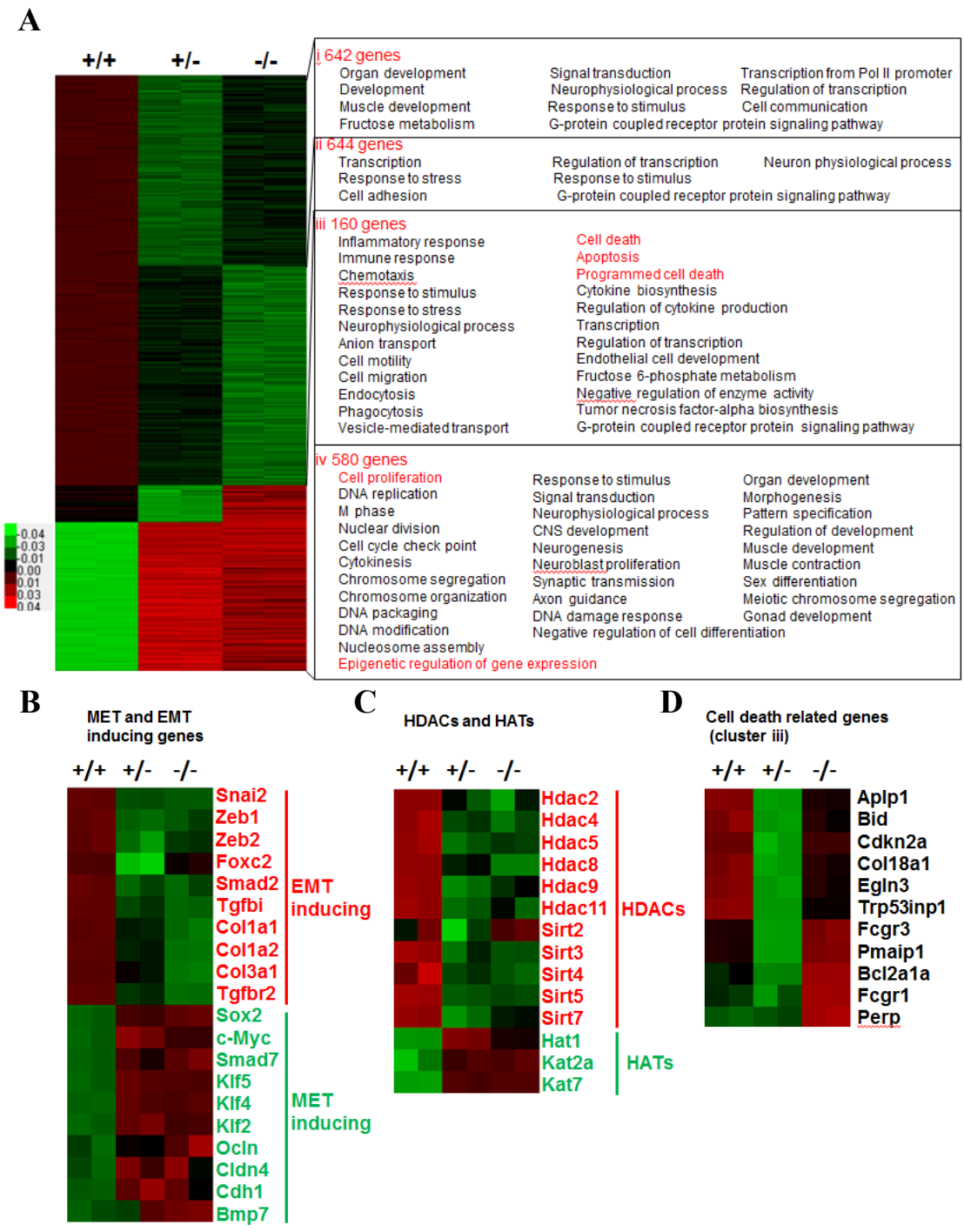

Figure 4.5 Gene expression profiles in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs. (A) Heatmaps of genes selected according to the fold change (>2.0 or <0.6) in Patz1$^{+/-}$ compared to Patz1$^{+/+}$ MEFs. Relative highly expressed genes are shown in red and low expressed genes in green. Compared to WT MEFs, 1286 genes (clusters i and ii) were down-regulated by more than 40% and 740 genes (clusters iii and iv) were up-regulated by more than 2 fold in Patz1$^{+/-}$ MEFs. The two different clusters for each up-regulated or down-regulated group of genes were created with Cluster 3.0 according to their different levels in Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs. Heatmaps were visualized using Java Treeview. GO analysis (GATHER) was performed with each cluster of genes and enriched GO terms ($p<0.05$) were selected and classified into groups accordingly. (B) List of EMT or MET inducing genes, which expressions were

changed in Patz1$^{+/-}$ and Patz1$^{-/-}$ compared to WT MEFs. The trend indicates that EMT inducing genes were generally repressed whereas MET inducing genes were up-regulated upon Patz1 loss. (C) List of histone deacetylase and acetylase genes, which expression is changed in Patz1$^{+/-}$ and Patz1$^{-/-}$ compared to WT MEFs. Histone deacetylases were generally down-regulated, while histone acetylases were up-regulated, in Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs compared to WT control. (D) List of cell death related genes differentially expressed in Patz1$^{+/-}$ and Patz1$^{-/-}$ compared to WT MEFs. All these genes belong to the cluster iii, i.e. up-regulated in Patz1$^{-/-}$ MEFs whereas down-regulated in Patz1$^{+/-}$ compared to WT MEFs. Each selected gene was taken as individual tiles from the thumbnail-dendogram duplicates.

4.1.6 Patz1$^{+/-}$ MEFs but not Patz1$^{-/-}$ MEFs surpass *Ink4a/Arf* locus barrier in reprogramming

Among those reversely changed genes, we found that Cdkn2a was significantly up-regulated in Patz1$^{-/-}$ MEFs, but not in Patz1$^{+/-}$ MEFs. Cdkn2a (also known as p16), encoded by *Ink4a*, is a key effector of cellular senescence. Recent studies have revealed that H3K27me3 level at *Ink4a/Arf* is decreased in response to expressions of reprogramming factors and silencing of this locus allows for an efficient reprogramming (Banito et al., 2009; Li et al., 2009; Utikal et al., 2009b). Therefore we hypothesized that Patz1 may regulate this *Ink4a/Arf* locus to induce cell senescence, thus inhibiting reprogramming process.

We examined the protein level of Cdkn2a in Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs by WB. Indeed the level of p16 was reduced in Patz1$^{+/-}$ MEFs, but up-regulated in Patz1$^{-/-}$ MEFs (Figure 4.6B). We then performed ChIP experiments using anti-H3K27me3 and anti-H3K4me3 antibodies to check their levels at this locus. As presented in Figure 4.6C, in Patz1$^{+/-}$ MEFs the active histone mark H3K4me3 was significantly repressed, whereas the level of repressive mark H3K27me3 was robustly increased than in wt MEFs. This reaffirmed that the roadblock of *Ink4a/Arf* locus was bypassed in Patz1$^{+/-}$ MEFs, which therefore can be easily reprogrammed.

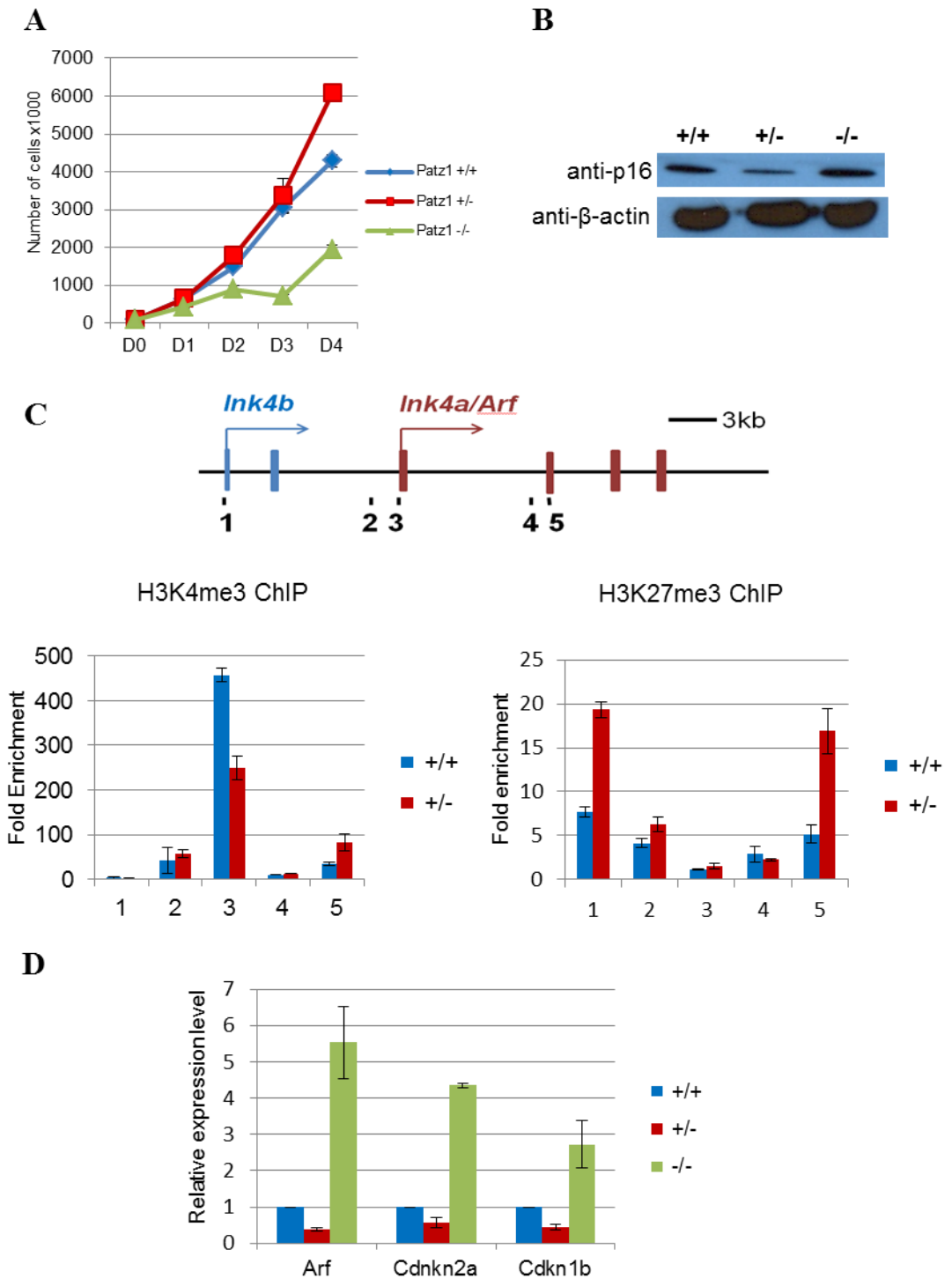

Figure 4.6 Ink4a/Arf locus is activated in Patz1$^{-/-}$ MEFs. (A) Patz1$^{+/-}$ MEFs showed the highest cell proliferation rate whereas Patz1$^{-/-}$ MEFs was the lowest. The number of cells increases fastest in Patz1$^{+/-}$ MEFs. Same number of cells were seeded into 96-well plates and cell numbers were counted and calculated every day. Data represent means ± SD of three independent experiments. (B) WB results showed increased expression of CDkn2a in Patz1$^{-/-}$ MEFs, but lower level in Patz1$^{+/-}$ MEFs. B-actin was used as a loading control. (C) ChIP results showed a lower level of active histone mark H3K4me3 and elevated level of repressive histone mark H3K27me3 in Patz1$^{+/-}$ MEFs. Realtime PCR primers were designed along the Ink4a/Arf locus as the

schematic figure shows. The fold enrichment was quantified as relative to input, and the fold changes were normalized against an intragenic control region on chromosome 17. (D) Relative mRNA levels of senescence markers in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs 5 days after OKSM infection. Note that at day 5 of reprogramming, Cdkn2a and Cdkn1b mRNA levels were significantly increased in Patz1$^{-/-}$ MEFs compared to the WT counterparts, whereas their levels were clearly reduced in Patz1$^{+/-}$ MEFs.

We next harvested RNAs from Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs 5 days after OKSM infection to examine the levels of Cdkn2a in response to reprogramming factors. We found that at day 5 of reprogramming Cdkn2a mRNA levels were significantly increased in Patz1$^{-/-}$ MEFs compared to the WT counterparts, whereas their levels were clearly reduced in Patz1$^{+/-}$ MEFs (Figure 4.6D). Similar trend was observed for another senescence related protein Cdkn1b (also known as p27). This reinforced our hypothesis that the *Ink4a/Arf* locus is silenced in Patz1$^{+/-}$ MEFs, thus contributing to the faster cell division and higher reprogramming rate of these cells compared to wild-type controls. Conversely, Patz1$^{-/-}$ MEFs cannot overcome the Ink4a/Arf barrier and undergo cellular senescence during reprogramming process, resulting in the limited formation of iPS colonies.

4.1.7 Patz1 negatively regulates reprogramming possibly via modulating global histone modifications in MEFs

As a number of terms related to chromatin assembly, organization and epigenetic regulations of gene expressions were enriched among the up-regulated genes in Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs (Figure 4.5A, C), we next explored the roles of Patz1 in the epigenetic regulation.

Patz1 belongs to POK family of transcription repressors and the POZ domain is known to interact with corepressor complexes to negatively regulate gene

transcriptions. Since Patz1 has been shown to interact with Ncor1 and Sirt1, we first examined the histone acetylation levels (Cho et al., 2011; Sakaguchi et al., 2010). H3Ac level was higher in Patz1$^{+/-}$ MEFs and was significantly repressed in Patz1 OE MEFs (Figure 4.7A). This suggested that the level of Patz1 is correlated with histone acetylation level. In addition, our microarray results also revealed that a number of histone acetylases and deacetylases were altered in Patz1$^{+/-}$ MEFs (Figure 4.5C). Using ChIP analysis, we also found that Patz1 could bind to many histone deacetylation related genes, indicating that Patz1 may potentially regulate these genes to modulate histone acetylation (Figure 4.7B). Indeed when Patz1 was overexpressed, the level of Hdac2, Hdac3 and Sin3a were significantly increased, suggesting that Patz1 overexpression increases histone deacetylation via activation of these epigenetic factors (Figure 4.7C).

We also examined the level of marks associated with euchromatin- H3K36me3 and H3K4me3, and the repressive histone marks H3K9me3 in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1 $^{-/-}$ MEFs. As shown in Figure 4.7D, the levels of active histone marks were generally increased, while the repressive histone mark were decreased, in Patz1$^{+/-}$ compared to wt. HP1α, well-known for its role in heterochromatin formation, was also decreased in Patz1$^{+/-}$ MEFs. Notably, the level of H3K4me2 was significantly upregulated in Patz1$^{+/-}$ MEFs and reduced in Patz1 overexpressed MEFs (data not shown). Acquisition of H3K4me2 at the pluripotency genes occurs at the initial stage of reprogramming that primes the gene for activation later on, which is a prerequisite for pluripotency induction (Koche et al., 2011). Thus, heterozygous loss of *Patz1* may promote a globally more open chromatin state accessible for activation of pluripotency transcriptional network, thus enhancing reprogramming.

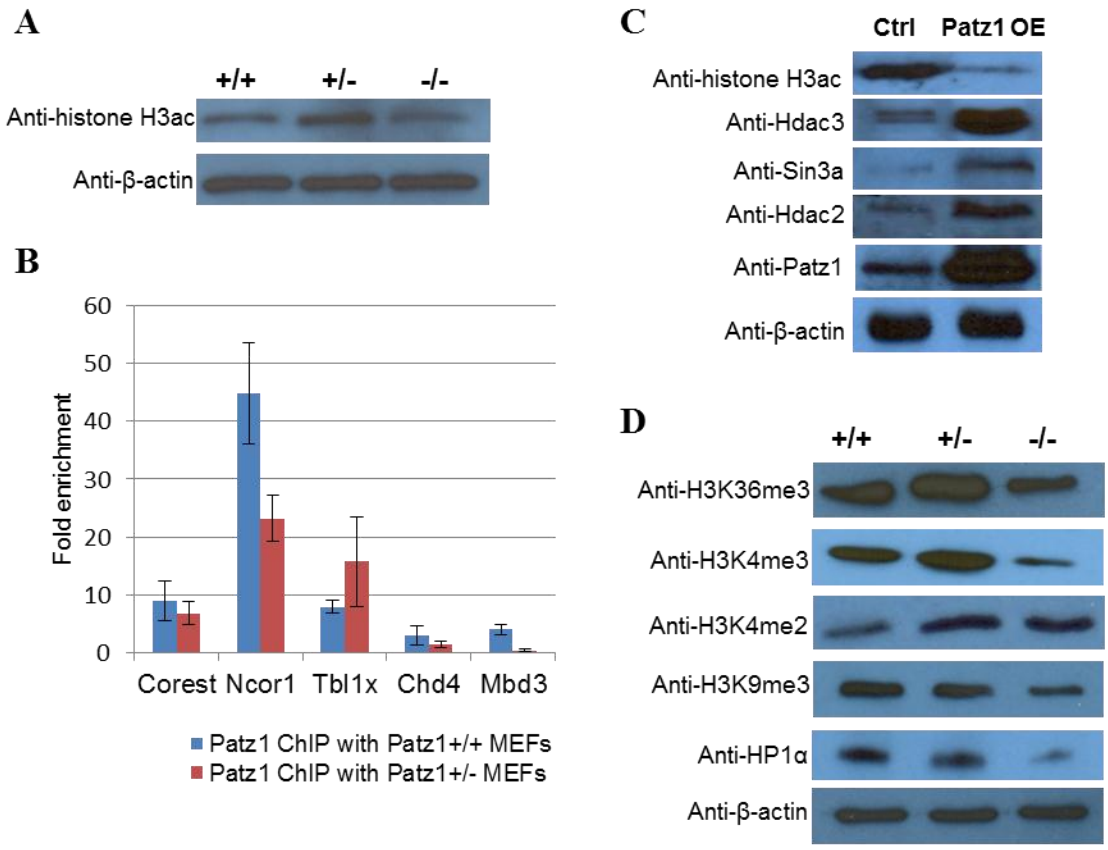

Figure 4.7 Level of Patz1 modulates chromatin modifications in MEFs. (A) H3 acetylation level was increased in Patz1$^{+/-}$ MEFs. (B) ChIP using Patz1 antibody showed that Patz1 can bind to histone deacetylation related genes, such as Corest, Ncor1 and Chd4. The fold enrichments of Patz1 on these potential sites were higher in WT MEFs than Patz1-heterozygous knockout MEFs. (C) H3 acetylation was repressed by Patz1 overexpression. H3Ac was significantly repressed, while the levels of histone deacetylation related factors Hdac2, Hdac3, Sin3a were increased in Patz1 OE MEFs. (D) Up-regulated H3K4me3, H3K4me2, H3K36me3 levels and repressed H3K9me3 level in heterozygous knockout MEFs. B-actin was used as a loading control.

4.2 Discussion

We previously demonstrated that Patz1 is an essential pluripotency factor. It is therefore intricate to investigate whether it has a role in pluripotency reacquisition. In this study, we revealed the first evidence for the involvement of Patz1 in reprogramming process: (i) overexpression of Patz1 inhibits the acquisition of pluripotency, while interference or heterozygous loss of Patz1 enhances iPSC

generation; (ii) knockout of *Patz1* hinders the reprogramming process by inducing cellular senescence. This suggests that a critical control of Patz1 dosage is essential for the generation of iPSCs (Figure 4.8).

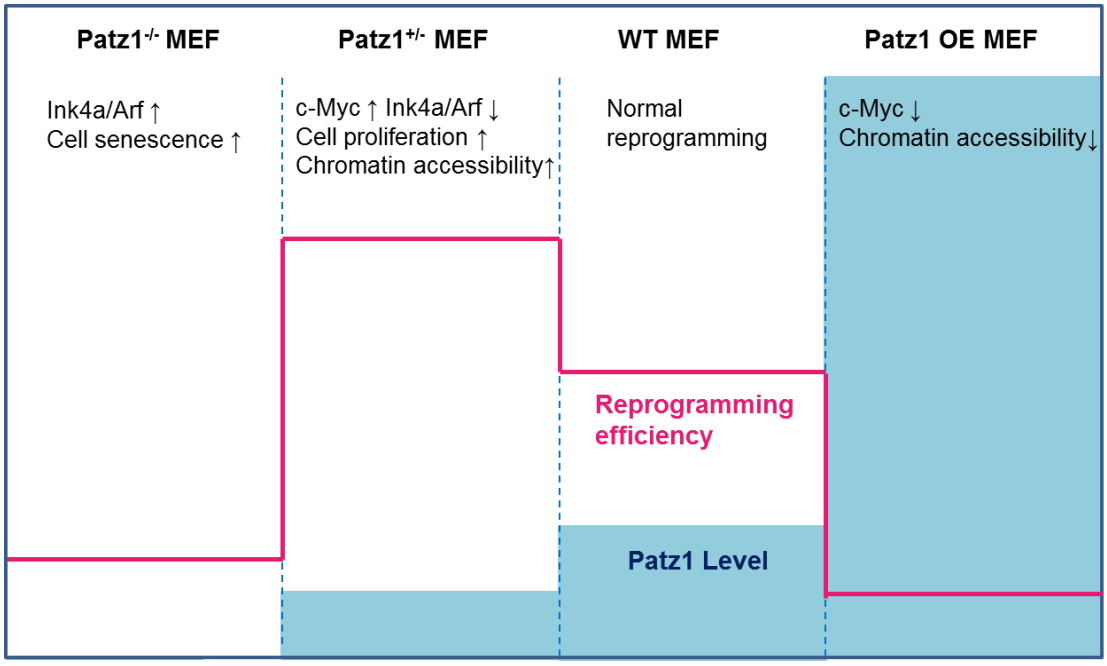

Figure 4.8 Proposed role of Patz1 during somatic cell reprogramming. Briefly, overexpression of Patz1 creates a condensed chromatin which represses the reprogramming process; Patz1 overexpression also represses c-Myc and induces cell senescence to inhibit reprogramming. Heterozygous knockout of *Patz1* can promote MET, activate c-Myc, overcome *Ink4a/Arf* barrier to surpass senescence and also create an open, hyperdynamic chromatin structure accessible for pluripotency gene reactivation, thus enhancing cellular reprogramming. Patz1$^{-/-}$ MEFs undergo cell senescence and are hard to be reprogrammed.

4.2.1 The role of Patz1 is dosage- and possibly p53- dependent

Patz1 has emerged as a fascinating transcription factor that have been implicated in various cellular or developmental processes, such as transcription regulation, carcinogenesis, spermatogenesis, thymocyte development, neurological process, pluripotency maintenance, cellular senescence and apoptosis. The involvement of PATZ1 in tumor suppression or carcinogenesis has been a debating issue for decades. Patz1 is previously known as MAZR (Myc-associated zinc finger related protein),

which can activates *c-Myc* promoter (Kobayashi et al., 2000). But subsequent studies reported the implications of Patz1 in tumor suppression. Here we showed that Patz1 possibly acts as a repressor of c-Myc in MEFs. c-Myc protein level was the highest in $Patz1^{-/-}$ MEFs. Overexpression of Patz1 significantly repressed c-Myc, but induces p53 and p16. Previous studies have found that c-Myc and p53 negatively regulates each other in tumorigenesis, cell differentiation and cell apoptosis etc. (Ceballos et al., 2005; Ho et al., 2005; Zheng et al., 2008; Sachdeva et al., 2009). p53 is known to negatively regulate *c-Myc* transcription through a mechanism that involves histone deacetylation (Ho et al., 2005). Given that Patz1 is an interacting partner of Ncor1 and Sirt1 and modulates histone acetylation level (Sakaguchi et al., 2010; Cho et al., 2011), it is therefore suggested that ectopic Patz1 could act via repression of c-Myc, possibly via cooperation with p53, to inhibit reprogramming efficiency.

A more recent finding revealed that the tumor suppressor role of PATZ1 is p53-dependent (Valentino et al., 2013a). In the presence of p53, PATZ1 interacts with p53 to enhance its functions in tumor suppression; in the absence of p53, PATZ1 is more likely to enhance cell survival and proliferation. Previous studies have proposed some mechanisms by which p53 inhibits reprogramming, such as inducing cell cycle arrest and apoptosis, inhibiting mesenchymal to epithelial transition (Banito et al., 2009; Brosh et al., 2012; Wang et al., 2012). Here we show that the level of Patz1 is also critical for reprogramming process. In the absence of Patz1, p53/p16 axis is activated, and the cells undergo cellular senescence. When only one of *Patz1* alleles is disrupted, p53 level is not that affected but *Ink4a/Arf* locus is repressed, whereby preventing the cells from senescence induction. Overexpression of Patz1, however, robustly activates p53 and p16, whereby eliciting cell cycle arrest in MEFs. Given that p53 is a known inhibitor of reprogramming, we therefore speculate that excessive expression of Patz1

may act through p53/p16 to induce cell senescence, and thus inhibiting reprogramming process. The observation of higher proliferation rate in Patz1$^{+/-}$ MEFs indicates that there are other pathways involved. It could not be ruled out that when Patz1 level is reduced in Patz1$^{+/-}$ MEFs, the anti-proliferative effects of p53 could no longer be stabilized. This may somewhat indicates that the role of Patz1 relies on its expression levels and possibly is p53-dependent as well.

4.2.2 Patz1 is implicated in cell senescence

We also propose that the restriction of iPSC generation in Patz1$^{-/-}$ MEFs may be associated with cellular senescence. We have shown that knockout of *Patz1* up-regulates senescence effectors p16 and p21, driving cellular senescence in MEFs. Patz1 may be involved in cellular senescence by regulating epigenetic status of *Ink4a/Arf* locus. The absence of *Patz1* induces higher H3K4me3 in this locus; but with a haploinsufficient level of Patz1, H3K27me3 is significantly increased, resulting in repressed expression of *Ink4a/Arf*. *Ink4a/Arf* locus needs to be repressed for the transition from somatic cell to pluripotent state (Banito et al., 2009; Li et al., 2009; Utikal et al., 2009b). Jmjd3 for example, which possesses tumor suppressor character, inhibits the reprogramming process by demethylation of H3K27me3 at *Ink4a/Arf* (Agger et al., 2009; Zhao et al., 2013). Similarly, haploinsufficient loss of Patz1 may accelerate the reprogramming kinetics by pre-repressing of *Ink4a/Arf* locus. This is consistent with the concept that somatic cells need to overcome cellular senescence to acquire pluripotency. But it is not clear whether Patz1 is directly involved in the recruitment of epigenetic factors to this locus, which would be a future research direction.

4.2.3 Patz1 may be a novel negative regulator of MET

Since many MET inducing genes were up-regulated while EMT inducing related genes were down-regulated in Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs, we propose that Patz1 could be implicated in MET inhibition. MET, a reversed process to EMT, is one of the key cellular events during early stage of reprogramming process (Li et al., 2010b). Factors that promote MET, including Klf4 and E-cad can enhance pluripotency acquisition; while factors that drive EMT or prevent MET, such as TGF-β and some mesenchymal markers, impede the reprogramming at the initial stage (Li et al., 2010b; Samavarchi-Tehrani et al., 2010). Interestingly our Patz1 ChIP results revealed that many MET and EMT associated genes are bound by Patz1 in ES cells (Figure 4.9A). Thus Patz1 may help to activate the epithelial program and suppress mesenchymal genes to overcome the EMT epigenetic barrier of fibroblasts. Since Klf4 is one of the major effector to drive MET during reprogramming process, we checked Klf4 expression level in Patz1$^{+/+}$, Patz1$^{+/-}$ and Patz1$^{-/-}$ MEFs. Klf4 levels seemed to be similar in those MEFs (Figure 4.9B). Hence we reckon that Patz1 may act independently from Klf4, perhaps directly regulating MET related genes or their regulators.

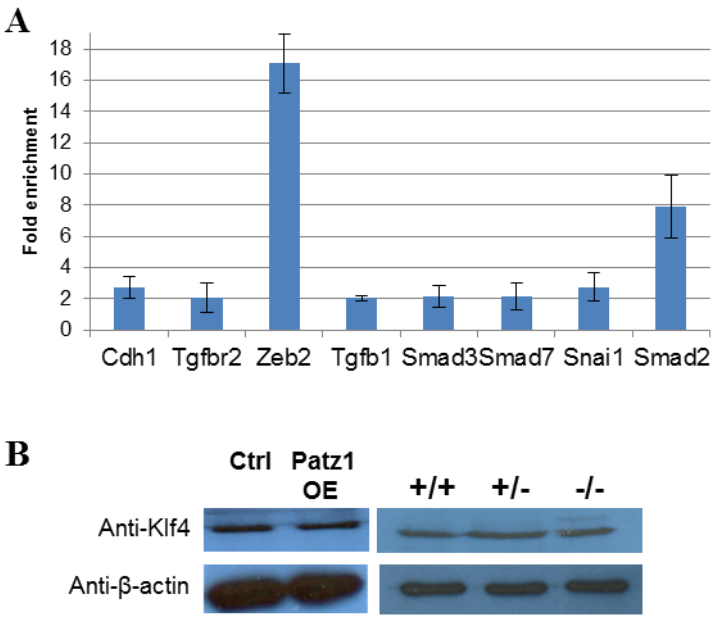

Figure 4.9 Patz1 may be implicated in MET. (A) Many EMT and MET related genes were bound by Patz1 in MEFs. ChIP using anti-Patz1 antibody was performed with WT Patz1 MEFs. Primers were designed based on Patz1 ChIP-seq results in mESCs. The fold enrichment was quantified as relative to input, and the fold changes were normalized against an intragenic control region on chromosome 17. (B) Klf4 protein levels were not affected by Patz1 in MEFs. B-actin was used as a loading control.

4.2.4 Patz1 modulate histone acetylation to enhance reprogramming

Somatic cells utilize a large number of epigenetic regulations to prevent the expression of unwanted genes. Pluripotent cells however possess a rather open chromatin configuration which allows transcriptional programs to switch rapidly upon induction of differentiation. How to overcome the epigenetic barriers of somatic cells in reprogramming remains a key question. In this report we provide some evidence that heterozygous loss of *Patz1* enhances reprogramming possibly through promoting a more active chromatin state ready for reprogramming onset.

One possible way for Patz1 modulating chromatin status is via histone acetylation. Previous studies have reported the interaction of Patz1 and NCoR complex and Sirt1. Our results show that Patz1 is possibly a supreme factor regulating histone acetylation,

not only by participating directly to recruit HDAC complexes, but also as a regulator of HDAC related genes. The acetylation of histones serves as a marker for active or open chromatin, and its level is higher in plastic pluripotent chromatin. In view of previous finding that inhibitors of HDAC activity significantly improve the reprogramming efficiency (Huangfu et al., 2008a), we propose that the inhibitory effect of Patz1 OE could be attributed to the repressed histone acetylation level, which hinders the chromatin resetting for efficient reprogramming. This is consistent with the general notion that Patz1 is a transcriptional repressor. Moreover, the repression of Patz1 OE on histone acetylation could be one of the possible mechanism by which Patz1 OE inhibits expression of lineage genes during differentiation, as histone acetylation are increased during differentiation, leading to rises in gene expression from all germ layers.

4.2.5 The stochastic model of reprogramming and Patz1 dosage

A model has been proposed that reprogramming is stochastic at the early stage (Buganim et al., 2012; Yamanaka, 2009). At reprogramming initiation, the reprogramming factors induce global gene activation and in response to that somatic cells acquire one type of cell fate, which could be senescence, transdifferentiation, dedifferentiation *etc.* Only those cells which tend to dedifferentiate, *i.e.* reprogrammable cells, will subsequently undergo a second more deterministic phase, during which a cascade of molecular events occurs and finally iPSCs are generated. The stochastic cell fate transition, which has a long latency and a low efficiency, has been considered as the major rate-limiting step for reprogramming. Reprogramming factors are known to induce senescence, thus the cells are more likely to undergo cell senescence other than dedifferentiation (Banito et al., 2009). Hence the

reprogramming efficiency is very low. Here we suspect the dosage of Patz1 may affect the stochastic phase of reprogramming.

As shown in Figure 4.6, Patz1$^{+/-}$ MEFs expressed lower level of Cdkn2a at day 5 of reprogramming, while Cdkn2a was robustly activated in Patz1$^{-/-}$ MEFs. It seems that Patz1$^{-/-}$ MEFs have a higher tendency to acquire cell senescence for the stochastic transition, thus very limited iPSCs were generated. But once these cells have surpassed the stochastic phase and become reprogrammable cells that are deterministic to be reprogrammed, they can generate healthy iPSCs. This indicates that although Patz1 is not essentially required for pluripotency induction, its absence will drive cell senescence and impairs the stochasticity of cell fate transition at the early reprogramming stage.

On the other hand, Patz1$^{+/-}$ MEFs display a distinct story. Based on our observation of Patz1's role in chromatin modulation, heterozygous knockout of *Patz1* creates a hyperdynamic chromatin state, rendering the global promoters more permissive for the stochastic gene activation. Thus the cells can be rapidly activated for cell fate conversion in response to the reprogramming factors, thus the reprogramming process is accelerated. Moreover, the lower level of Cdkn2a observed at day 5 of reprogramming implies that there was a smaller percentage of cells undergoing cell senescence compared with wt and knockout MEFs. Thus the stochasiticity of cell fate transition is biased against cell senescence in Patz1$^{+/-}$ MEFs, leading to a higher reprogramming efficiency. Yet, the defective iPSCs generated from Patz1$^{+/-}$ MEFs indicates that a proper level of Pazt1 might be required for the deterministic reprogramming phase. This is consistent with our previous report that depletion of *Patz1* impairs mESC pluripotency and self-renewal (Ow et al., 2014). Therefore,

Patz1 may affect some key molecular events for pluripotency induction, which yet warrant further studies.

Overexpression of Patz1 seems to result in a globally closed chromatin configuration, which impedes the binding of reprogramming factors to the promoters for gene activation. The stochastic transition at single-cell level is therefore inefficient. This is support by our finding that if Patz1 was added at a later time point after OKSM induction, there were more cells that had passed through the stochastic phase, and gained some early features of pluripotency, such as AP activity. But the reactivation of pluripotency factors such as Oct4, which occurs at a later stage, remained blocked by addition of Patz1. Moreover, it seems that overexpression of Patz1 hinders the deterministic phase of reprogramming as well. When Patz1 was added at a later stage, a further suppression of GFP^+ colony formation was observed than Patz1 co-inducted with OKSM, indicating that these cells were subjected to the inhibitory effect of ectopic Patz1 all along the road towards pluripotency.

4.3 Future work

It has been proposed that the functions of Patz1 are cellular context- and dosage-dependent, largely because it is an architecture protein which may interact with different proteins to exert specific roles in different types of cells. This has been reaffirmed by our studies of Patz1 in mESC identity and somatic reprogramming, which uncover that Patz1 has different roles in pluripotency regulation. Besides, it is quite obscure regarding why ectopic expression and homozygous knockout of *Patz1* result in a similar phenotype that is distinct to that of Patz1-heterozygous knockout. Hence, it would be necessary to identify the interactome of Patz1 and their targets by

genomic wide studies in different Patz1 level background, which may give a better understanding of its functions in various cellular processes.

Through the exploration of Patz1's roles in somatic reprogramming, we provide new evidence that the level of Patz1 is closely related to cell cycle regulation, chromatin conformation and MET, thus modulating reprogramming process. These results shed some lights on novel functions of Patz1, especially in chromatin regulation. Our results revealed a strong evidence of the interplay between Patz1 dosage and histone acetylation level, which points out the direction of future studies. As a transcription factor, Patz1 may potentially recruit epigenetic factors to a specific locus to regulate gene transcriptions. Notably, Patz1 seems to be able to modulate the chromatin configuration and epigenetic modifications, leading to global gene transcription repressions. This will shed some light on future studies of Patz1 in cancer cells. Since altered epigenetic regulations have been found in many cancer types, especially at the early stage of cancer, dissection of the relationship between Patz1 and chromatin would be one of the new directions for future research. In regard of these, it would be of great significance and interest to illustrate the deeper mechanisms of how Patz1 is implicated in these processes.

CHAPTER 5

CONCLUDING REMARKS

ESC is a unique type of cells, featured by by their ability to self-renew and by their potential to differentiate into many different cell types. This specific state is delicately governed by a complex network of various factors. With the remarkable breakthrough of somatic cell reprogramming, pluripotent cells have shown emerging potentials in regenerative medicine. However, the mechanism underlying the transformation from somatic cells to pluripotent cells remains elusive.

Zinc finger proteins are a big family of proteins that can regulate gene activities by interacting directly with specific cis-regulatory DNA elements through their zinc fingers; they have evolved to fill roles in many different biological processes (Dang et al., 2000; Huntley et al., 2006). Recent studies have revealed that members of this family, such as ZSCAN4, Zfp296, Zfp206, Zfp57, and Zfp42 *etc*., are key components of the ESC transcriptional network and are crucial for maintaining pluripotent ESCs (Fischedick et al., 2012; Hirata et al., 2012; Scotland et al., 2012; Yu et al., 2009; Zuo et al., 2011). In this study, we discovered that Zfp322a and Patz1, two zinc finger proteins highly expressed in ICM and ESCs, are critically engaged in pluripotency regulation.

Zfp322a is a highly conserved protein, the functions of which yet remained unclear. Here we demonstrated that Zfp322a plays an important role in the maintenance and acquisition of pluripotency. Depletion of *Zfp322a* impairs mESC self-renewal and induces them to differentiate. It is suggested that Zfp322a prevents mESC from differentiation possibly by activating *Pou5f1* and *Nanog* expression while repressing MAPK/ERK pathway. Zfp322a is also an interacting partner of Oct4 and is integrated

as a component within the ESC transcriptional network. Further genome wide studies identified the targets of Zfp322a which are involved in a variety of biological processes, including DNA transcription and translation, chromosome organization, development, DNA repair, cell cycle and apoptosis. Additionally, Zfp322a is found to be a novel reprogramming factor that can replace Sox2 in the classical Yamanaka's factors. It can be used in combination with Yamanaka's factors and that addition leads to a higher reprogramming efficiency and to acceleration of the onset of the reprogramming process. Together, these results established Zfp322a as a novel pluripotency factor that can enhance reprogramming process, filling the functions of this newly identified protein.

Patz1, on the other hand, appears to regulate pluripotency in a cellular context-dependent and dosage-dependent manner. Similar to Zfp322a, Patz1 is required for mESC pluripotency maintenance by activating Oct4 and Nanog. However, Patz1 displayed a distinct role in the induction of pluripotency. It is observed that overexpression of Patz1 inhibits OKSM-mediated process, while depletion of *Patz1* enhances iPSC generation, indicating an inhibitory role of Patz1 in the reprogramming process. Complete loss of *Patz1*, however, resulted in cellular senescence that significantly blocks iPSC formation. Through the exploration of Patz1's roles in somatic reprogramming, we provide new evidence that the level of Patz1 is closely related to cell cycle regulation, chromatin conformation and MET, thereby modulating the induction of pluripotency. We believed our results provide some insightful hints on study how genetic factors crosstalk with epigenetic factors in reprogramming process. We proposed that the functions of Patz1 are cellular context- and dosage- dependent, largely because it may interact with different proteins to exert specific roles in different types of cells. It would be of great significance and interest

to illustrate more underlying mechanisms in future studies. Nevertheless, this study is among a few studies in which different dosages of pluripotency factors play diverse roles in reprogramming process.

Taken together, these data revealed that zinc finger proteins could regulate pluripotency in distinct manners, expanding the current knowledge of pluripotency maintenance and acquisition, as well as the diverse functions of zinc finger protein family. Our studies also shed light on how somatic zinc finger proteins affect reprogramming efficiency, opening a new route to reboot pluripotency via modulating the level of zinc finger proteins. Moreover, we believe that our findings would point new directions for future research on Zfp322a and Patz1, and provide some hints to the studies of other zinc finger proteins as well.

BIBLIOGRAPHY

Agger, K., Cloos, P.A.C., Rudkjaer, L., Williams, K., Andersen, G., Christensen, J., and Helin, K. (2009). The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. Genes & Development *23*, 1171-1176.

Albano, R., Groome, N., and Smith, J. (1993). Activins are expressed in preimplantation mouse embryos and in ES and EC cells and are regulated on their differentiation. Development *117*, 711-723.

Aloia, L., Parisi, S., Fusco, L., Pastore, L., and Russo, T. (2009). Differentiation of Embryonic Stem Cells 1 (Dies1) Is a component of bone morphogenetic protein 4 (BMP4) signaling pathway required for proper differentiation of mouse embryonic stem cells. Journal of Biological Chemistry *285*, 7776-7783.

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

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

Anokye-Danso, F., Trivedi, Chinmay M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, Peter J., Epstein, Jonathan A., *et al.* (2011). Highly Efficient miRNA-Mediated Reprogramming of Mouse and Human Somatic Cells to Pluripotency. Cell Stem Cell 4, 376-388.

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature *502*, 462-471.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nature Reviews Molecular Cell Biology *10*, 91-103.

Asakura, A., Tennakoon, J.B., Wang, H., Coarfa, C., Cooney, A.J., and Gunaratne, P.H. (2013). Chromatin changes in dicer-deficient mouse embryonic stem cells in response to retinoic acid induced differentiation. PLoS ONE, e74556.

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

Avilion, A.A. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes & Development *17*, 126-140.

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

Barroso-delJesus, A., Lucena-Aguilar, G., Sanchez, L., Ligero, G., Gutierrez-Aranda, I., and Menendez, P. (2011). The Nodal inhibitor Lefty is negatively modulated by the microRNA miR-302 in human embryonic stem cells. The FASEB Journal *25*, 1497-1508.

Barroso-delJesus, A., Romero-Lopez, C., Lucena-Aguilar, G., Melen, G.J., Sanchez, L., Ligero, G., Berzal-Herranz, A., and Menendez, P. (2008). Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. Molecular and Cellular Biology *28*, 6609-6619.

Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. Stem Cells *23*, 489-495.

Bendall, S.C., Stewart, M.H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bossé, M., *et al.* (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. Nature *448*, 1015-1021.

Berge, D.t., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nature Cell Biology *13*, 1070-1075.

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

Blaukat, A., Ivankovic-Dikic, I., Gronroos, E., Dolfi, F., Tokiwa, G., Vuori, K., and Dikic, I. (1999). Adaptor proteins Grb2 and Crk couple Pyk2 with activation of specific mitogen-activated protein kinase cascades. Journal of Biological Chemistry *274*, 14893-14901.

Boeuf, H., Hauss, C., Graeve, F.D., Baran, N., and Kedi, C. (1997). Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. Journal of Cel Biology *138*, 1207-1217.

Bourillot, P.-Y., Aksoy, I., Schreiber, V., Wianny, F., Schulz, H., Hummel, O., Hubner, N., and Savatier, P. (2009). Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. Stem Cells *27*, 1760-1771.

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

Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics *20*, 3710-3715.

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., *et al.* (2004). Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation. Nature Biotechnology *22*, 707-716.

Brayer, K.J., and Segal, D.J. (2008). Keep your fingers off my DNA: Protein–protein interactions mediated by C2H2 zinc finger domains. Cell Biochemistry and Biophysics *50*, 111-131.

Brehm, A., Ohbo, K., and Schöler, H. (1997). The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain. Molecular and Cellular Biology *17*, 154-162.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., 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.

Brook, F.A., and Gardner, R.L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. Proceedings of the National Academy of Sciences *94*, 5709-5712.

Brosh, R., Assia-Alroy, Y., Molchadsky, A., Bornstein, C., Dekel, E., Madar, S., Shetzer, Y., Rivlin, N., Goldfinger, N., Sarig, R., *et al.* (2012). p53 counteracts reprogramming by inhibiting mesenchymal-to-epithelial transition. Cell Death and Differentiation *20*, 312-320.

Buganim, Y., Faddah, Dina A., Cheng, Albert W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, Sandy L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell *150*, 1209-1222.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. Nature Reviews Genetics *14*, 427-439.

Cartwright, P. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. Development *132*, 885-896.

Castellano, E., and Downward, J. (2011). RAS interaction with PI3K: more than just another effector pathway. Genes & Cancer *2*, 261-274.

Chambers, I., Colby, D., Morag Robertson, Jennifer Nichols, Sonia Lee, Susan Tweedie, 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.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311-2322.

Chang, J.T., and Nevins, J.R. (2006). GATHER: a systems approach to interpreting genomic signatures. Bioinformatics *22*, 2926-2933.

Chen, J., Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., Liang, H., Chen, Y., Chen, J., Wu, Y., *et al.* (2012). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nature Genetics *45*, 34-42.

Chen, J., Liu, J., Yang, J., Chen, Y., Chen, J., Ni, S., Song, H., Zeng, L., Ding, K., and Pei, D. (2010). BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone. Cell Research *21*, 205-212.

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

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Molecular and Cellular Biology *25*, 6031-6046.

Cho, J.H., Kim, M.J., Kim, K.J., and Kim, J.R. (2011). POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) inhibits endothelial cell senescence through a p53 dependent pathway. Cell Death and Differentiation *19*, 703-712.

Cho, K.B., Cho, M.K., Lee, W.Y., and Kang, K.W. (2010). Overexpression of c-myc induces epithelial mesenchymal transition in mammary epithelial cells. Cancer Letters *293*, 230-239.

Cohen, A. (2006). An essential role of ERK signalling in TPA-induced reactivation of Kaposi's sarcoma-associated herpesvirus. Journal of General Virology *87*, 795-802.

Colin L. Stewart, Petr Kaspar, Lisa J. Brunet, Harshida Bhatt, Inder Gadi, Frank Köntgen, and Abbondanzo, S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature, *4*, 76-79.

Collado, M., Blasco, M.A., and Serrano, M. (2007). Cellular senescence in cancer and aging. Cell *130*, 223-233.

Costoya, J.A. (2007). Functional analysis of the role of POK transcriptional repressors. Briefings in Functional Genomics and Proteomics *6*, 8-18.

Cowan, C.A. (2005). Nuclear reprogramming of somatic cells After fusion with human embryonic stem cells. Science *309*, 1369-1373.

Daley, G.Q., Lensch, M.W., Jaenisch, R., Meissner, A., Plath, K., and Yamanaka, S. (2009). Broader implications of defining standards for the pluripotency of iPSCs. Cell Stem Cell *4*, 200-201.

Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Krüppel-like family of transcription factors. Int J Biochem Cell Biol *32*, 1103-1121.

David, L., and Polo, J.M. (2014). Phases of reprogramming. Stem Cell Research *12*, 754-761.

Davidson, K.C., Adams, A.M., Goodson, J.M., McDonald, C.E., Potter, J.C., Berndt, J.D., Biechele, T.L., Taylor, R.J., and Moon, R.T. (2012). Wnt/ -catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. Proceedings of the National Academy of Sciences *109*, 4485-4490.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected DNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.

Ding, J., Xu, H., Faiola, F., Ma'ayan, A., and Wang, J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Research *22*, 155-167.

Dvorak, P., Dvorakova, D., Koskova, S., Vodinska, M., Najvirtova, M., Krekac, D., and Hampl, A. (2005). Expression and potential role of fibroblast growth factor 2 and its receptors in human embryonic stem cells. Stem Cells *23*, 1200-1211.

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

Esteban, M.A., and Pei, D. (2012). Vitamin C improves the quality of somatic cell reprogramming. Nature Genetics *44*, 366-367.

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

F.L. Conlon, K.M. Lyons, N. Takaesu, K.S. Barth, A. Kispert, and, B.H., and Robertson, E.J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. Development *120*, 10.

Fedele, M., Benvenuto, G., Pero, R., Majello, B., Battista, S., Lembo, F., Vollono, E., Day, P.M., Santoro, M., Lania, L., *et al.* (2000). A novel member of the BTBPOZ family, PATZ, associates with the RNF4 RING finger protein and acts as a transcriptional repressor. The Journal of Biological Chemistry *275*, 7894-7901.

Fedele, M., Franco, R., Salvatore, G., Paronetto, M.P., Barbagallo, F., Pero, R., Chiariotti, L., Sette, C., Tramontano, D., Chieffi, G., *et al.* (2008). PATZ1 gene has a critical role in the spermatogenesis and testicular tumours. The Journal of Pathology *215*, 39-47.

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

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nature Cell Biology *8*, 188-194.

Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., Yaw, L.-P., Zhang, W., Loh, Y.-H., Han, J., *et al.* (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nature Cell Biology *11*, 197-203.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, Simon R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell *11*, 477-490.

Fidalgo, M., Faiola, F., Pereira, C.-F., Ding, J., Saunders, A., Gingold, J., Schaniel, C., Lemischka, I.R., Silva, J.C.R., and Wang, J. (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. Proceedings of the National Academy of Sciences *109*, 16202-16207.

Fischedick, G., Klein, D.C., Wu, G., Esch, D., Höing, S., Han, D.W., Reinhardt, P., Hergarten, K., Tapia, N., Schöler, H.R., *et al.* (2012). Zfp296 is a novel, pluripotent-specific reprogramming factor. PLoS ONE *7*, e34645.

Folmes, Clifford D.L., Nelson, Timothy J., Martinez-Fernandez, A., Arrell, D.K., Lindor, Jelena Z., Dzeja, Petras P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metabolism *14*, 264-271.

Frankenberg, S., Pask, A., and Renfree, M.B. (2010). The evolution of class V POU domain transcription factors in vertebrates and their characterisation in a marsupial. Developmental Biology *337*, 162-170.

Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., *et al.* (2007). DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. Nature Cell Biology *9*, 604-611.

Galvin, K.E., Travis, E.D., Yee, D., Magnuson, T., and Vivian, J.L. (2010). Nodal signaling regulates the bone morphogenic protein pluripotency pathway in mouse embryonicstem cells. Journal of Biological Chemistry *285*, 19747-19756.

Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., *et al.* (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell *12*, 453-469.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nature Reviews Molecular Cell Biology *12*, 36-47.

Graf, T., and Stadtfeld, M. (2008). Heterogeneity of embryonic and adult stem cells. Cell Stem Cell *3*, 480-483.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast Growth Factor 2 Modulates transforming growth factor β signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. Stem Cells *25*, 455-464.

Gu, P., Xu, X., Le Menuet, D., Chung, A.C.K., and Cooney, A.J. (2011). Differential recruitment of methyl CpG-binding domain factors and DNA methyltransferases by the orphan receptor germ cell nuclear factor initiates the repression and silencing of Oct4. Stem Cells *29*, 1041-1051.

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

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., *et al.* (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295-300.

Han, D.W., Tapia, N., Araúzo-Bravo, M.J., Lim, K.T., Kim, K.P., Ko, K., Lee, H.T., and Schöler, H.R. (2013a). Sox2 level is a determinant of cellular reprogramming Potential PLoS ONE, 10, e67594.

Han, G., Wang, H., and Hao, J., eds. (2013b). Molecular mechanisms of embryonic stem cell pluripotency (InTech), pluripotent stem cells, Dr. Deepa Bhartiya (Ed.), ISBN: 978-953-51-1192-4, InTech, DOI: 10.5772/54365. Available from:

http://www.intechopen.com/books/pluripotent-stem-cells/molecular-mechanisms-of-embryonic-stem-cell-pluripotency

Hassani, S.-N., Totonchi, M., Sharifi-Zarchi, A., Mollamohammadi, S., Pakzad, M., Moradi, S., Samadian, A., Masoudi, N., Mirshahvaladi, S., Farrokhi, A., *et al.* (2013). Inhibition of TGFβ signaling promotes ground state pluripotency. Stem Cell Reviews and Reports *10*, 16-30.

Hayashi, K., Lopes, S.M.C.d.S., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391-401.

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

Hitoshi Niwa, Tom Burdon, Ian Chambers, and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes and Development *12*, 2048-2060.

Hough, S.R., Clements, I., Welch, P.J., and M.S, K.A.W. (2006). Differentiation of Mouse Embryonic stem cells after RNA interference-mediated silencing of OCT4 and Nanog. STEM CELLS *24*, 1467-1475.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nature Biotechnology *26*, 795-797.

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

Huntley, S. (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: Insights into the evolutionary history of a large family of transcriptional repressors. Genome Research *16*, 669-677.

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

Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K., *et al.* (2009). A small-molecule inhibitor of Tgf-β signaling replaces Sox2 in reprogramming by inducing Nanog. Cell Stem Cell *5*, 491-503.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). Summaries of affymetrix GeneChip probe level data. Nucleic Acid Research *31*, e15.

Itoh, F., Watabe, T., and Miyazono, K. (2014). Roles of TGF-β family signals in the fate determination of pluripotent stem cells. Seminars in Cell & Developmental Biology 32, 98-106.

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

James, D. (2005). TGF /activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273-1282.

Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nature Biotechnology *26*, 1293-1300.

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

Jopling, C., Boue, S., and Belmonte, J.C.I. (2011). Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. Nature Reviews Molecular Cell Biology *12*, 79-89.

Kaji, K., Nichols, J., and Hendrich, B. (2007). Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. Development *134*, 1123-1132.

Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Arias, A.M. (2009). Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biology *7*, e1000149.

Kanellopoulou, C. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes & Development *19*, 489-501.

Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomelí, H., Nagy, A., McLaughlin, K.J., Schöler, H.R., *et al.* (2004). Oct4 is required for primordial germ cell survival. EMBO reports *5*, 1078-1083.

Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nature Biotechnology *26*, 1351-1359.

Kielman, M.F., Rindapää, M., Gaspar, C., van Poppel, N., Breukel, C., van Leeuwen, S., Taketo, M.M., Roberts, S., Smits, R., and Fodde, R. (2002). Apc modulates embryonic stem-cell differentiation by controlling the dosage of β-catenin signaling. Nature Genetics *32*, 594-605.

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

Kobayashi, A., Yamagiwa, H., Hoshino, H., Muto, A., Sato, K., Morita, M., Hayashi, N., Yamamoto, M., and Igarashi, K. (2000). A combinatorial code for gene expression generated by transcription factor Bach2 and MAZR (MAZ-related factor) through the BTB/POZ domain. Molecular and Cellular Biology *20*, 1733-1746.

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

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.

Lanner, F., and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. Development *137*, 3351-3360.

Laugesen, A., and Helin, K. (2014). Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell *14*, 735-751.

Lee, M.Y., Lim, H.W., Lee, S.H., and Han, H.J. (2009). Smad, PI3K/Akt, and Wnt-dependent signaling pathways are involved in BMP-4-induced eSC self-renewal. Stem Cells *27*, 1858-1868.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., and Muffat, J. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141*, 872-883.

Levenstein, M.E., Ludwig, T.E., Xu, R.-H., Llanas, R.A., VanDenHeuvel-Kramer, K., Manning, D., and Thomson, J.A. (2005). Basic fibroblast growth factor support of human embryonic stem cell self-renewal. Stem cells *24*, 568-574.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136-1139.

Li, J., Wang, G., Wang, C., Zhao, Y., Zhang, H., Tan, Z., Song, Z., Ding, M., and Deng, H. (2007). MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. Differentiation *75*, 299-307.

Li, L., Sun, L., Gao, F., Jiang, J., Yang, Y., Li, C., Gu, J., Wei, Z., Yang, A., Lu, R., *et al.* (2010a). Stk40 links the pluripotency factor Oct4 to the Erk/MAPK pathway and controls extraembryonic endoderm differentiation. Proceedings of the National Academy of Sciences *107*, 1402-1407.

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

Li, Y., Wang, Y., Zhang, C., Yuan, W., Wang, J., Zhu, C., Chen, L., Huang, W., Zeng, W., Wu, X., *et al.* (2004). ZNF322, a novel human C2H2 Krüppel-like zinc-finger protein, regulates transcriptional activation in MAPK signaling pathways. Biochemical and Biophysical Research Communications *325*, 1383-1392.

Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., *et al.* (2010c). Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. Cell Research *21*, 196-204.

Li, Z., and Rana, T.M. (2012). A kinase inhibitor screen identifies small-molecule enhancers of reprogramming and iPS cell generation. Nature Communications *3*, 1085.

Lim, L.S., Hong, F.H., Kunarso, G., and Stanton, L.W. (2010). The Pluripotency Regulator Zic3 is a direct activator of the nanog promoter in ESCs. Stem Cells *28*, 1961-1969.

Lim, L.S., Loh, Y.H., Zhang, W., Li, Y., Chen, X., Wang, Y., Bakre, M., Ng, H.H., and Stanton, L.W. (2007). Zic3 is required for maintenance of pluripotency in embryonic stem cells. Molecular Biology of the Cell *18*, 1348-1358.

Lin, Charles Y., Lovén, J., Rahl, Peter B., Paranal, Ronald M., Burge, Christopher B., Bradner, James E., Lee, Tong I., and Young, Richard A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56-67.

Lipchina, I., Elkabetz, Y., Hafner, M., Sheridan, R., Mihailovic, A., Tuschl, T., Sander, C., Studer, L., and Betel, D. (2011). Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. Genes & Development *25*, 2173-2186.

Liu, N., Lu, M., Tian, X., and Han, Z. (2007). Molecular mechanisms involved in self-renewal and pluripotency of embryonic stem cells. Journal of Cell Physiology *211*, 279-286.

Liu, X., Sun, H., Qi, J., Wang, L., He, S., Liu, J., Feng, C., Chen, C., Li, W., Guo, Y., *et al.* (2013a). Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT–MET mechanism for optimal reprogramming. Nature Cell Biology *15*, 829-838.

Liu, Z., Tang, Y., Lü, S., Zhou, J., Du, Z., Duan, C., Li, Z., and Wang, C. (2013b). The tumourigenicity of iPS cells and their differentiated derivates. Journal of Cellular and Molecular Medicine *17*, 782-791.

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

Loh, Y.H., Zhang, W., Chen, X., George, J., and Ng, H.H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes & Development *21*, 2545-2557.

Lopes, S.M.C.d.S., Hayashi, K., Shovlin, T.C., Mifsud, W., Surani, M.A., and McLaren, A. (2008). X chromosome activity in mouse XX primordial germ cells. PLoS Genetics *4*, e30.

Lu, X., Sachs, F., Ramsay, L., Jacques, P.-É., Göke, J., Bourque, G., and Ng, H.-H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. Nature Structural & Molecular Biology *21*, 423-425.

Luchi, S. (2001). Three classes of zinc finger proteins. Cell Mol Life Sci *58*, 625-635.

Luo, M., Ling, T., Xie, W., Sun, H., Zhou, Y., Zhu, Q., Shen, M., Zong, L., Lyu, G., Zhao, Y., *et al.* (2013). NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. Stem Cells *31*, 1278-1286.

Luo, Y., Lim, C.L., Nichols, J., Martinez-Arias, A., and Wernisch, L. (2012). Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. Journal of The Royal Society Interface *10*, 20120525-20120525.

Lyashenko, N., Winter, M., Migliorini, D., Biechele, T., Moon, R.T., and Hartmann, C. (2011). Differential requirement for the dual functions of β-catenin in embryonic stem cell self-renewal and germ layer formation. Nature Cell Biology *13*, 753-761.

Lyssiotis, C.A., Foreman, R.K., Staerk, J., Garcia, M., Mathur, D., Markoulaki, S., Hanna, J., Lairson, L.L., Charette, B.D., Bouchez, L.C., *et al.* (2009). Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. Proceedings of the National Academy of Sciences *106*, 8912-8917.

MacArthur, B.D., Sevilla, A., Lenz, M., Müller, F.-J., Schuldt, B.M., Schuppert, A.A., Ridden, S.J., Stumpf, P.S., Fidalgo, M., Ma'ayan, A., *et al.* (2012). Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. Nature Cell Biology *14*, 1139-1147.

Madhani, H.D., You, J.S., De Carvalho, D.D., Dai, C., Liu, M., Pandiyan, K., Zhou, X.J., Liang, G., and Jones, P.A. (2013). SNF5 is an essential executor of epigenetic regulation during differentiation. PLoS Genetics *9*, e1003459.

Maekawa, M., Yamaguchi, K., Nakamura, T., Shibukawa, R., Kodanaka, I., Ichisaka, T., Kawamura, Y., Mochizuki, H., Goshima, N., and Yamanaka, S. (2011). Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. Nature *474*, 225-229.

Maherali, N., and Hochedlinger, K. (2009). Tgfβ Signal Inhibition Cooperates in the Induction of iPSCs and replaces Sox2 and cMyc. Current Biology *19*, 1718-1723.

Mansour, A.A., Gafni, O., Weinberger, L., Zviran, A., Ayyash, M., Rais, Y., Krupalnik, V., Zerbib, M., Amann-Zalcenstein, D., Maza, I., *et al.* (2012). The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature *488*, 409-413.

Margariti, A., Kelaini, S., and Cochrane, A. (2014). Direct reprogramming of adult cells: avoiding the pluripotent state. Stem Cells and Cloning: Advances and Applications *7*, 19-29.

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

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

Mastrangelo, T., Modena, P., Tornielli, S., Bullrich, F., Testi, M.A., Mezzelani, A., Radice, P., Azzarelli, A., Pilotti1, S., Croce, C.M., *et al.* (2000). A novel zinc finger gene is fused to EWS in small round cell tumor. Oncogene *19*, 3799-3804.

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

Mathieu, M.-E., Saucourt, C., Mournetas, V., Gauthereau, X., Thézé, N., Praloran, V., Thiébaud, P., and Bœuf, H. (2011). LIF-dependent signaling: New pieces in the lego. Stem Cell Reviews and Reports *8*, 1-15.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. The EMBO Journal *18*, 4261-4269.

Merrill, B.J. (2012). Wnt pathway regulation of embryonic stem cell self-renewal. Cold Spring Harbor Perspectives in Biology *4*, a007971.

Messina, D.N. (2004). An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. Genome Research *14*, 2041-2047.

Miki, T., Yasuda, S.-y., and Kahn, M. (2011). Wnt/β-catenin signaling in embryonic stem cell self-renewal and somatic cell reprogramming. Stem Cell Reviews and Reports *7*, 836-846.

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

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.

Miyanari, Y., and Torres-Padilla, M.-E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470-473.

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

Mohammadi, M., Olsen, S.K., and Ibrahimi, O.A. (2005). Structural basis for fibroblast growth factor receptor activation. Cytokine & growth factor reviews *16*, 107-137.

Morgani, Sophie M., Canham, Maurice A., Nichols, J., Sharov, Alexei A., Migueles, Rosa P., Ko, Minoru S.H., and Brickman, Joshua M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. Cell Reports *3*, 1945-1957.

Muotri, A.R., Fischedick, G., Klein, D.C., Wu, G., Esch, D., Höing, S., Han, D.W., Reinhardt, P., Hergarten, K., Tapia, N., *et al.* (2012). Zfp296 is a novel, pluripotent-specific reprogramming factor. PLoS ONE *7*, e34645.

Murakami, M., Ichisaka, T., Maeda, M., Oshiro, N., Hara, K., Edenhofer, F., Kiyama, H., Yonezawa, K., and Yamanaka, S. (2004). mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. Molecular and Cellular Biology *24*, 6710-6718.

Nagashima, T., Hayashi, F., Umehara, T., and Yokoyama, S. (2009). Molecular structures of kruppel-like factors, The biology of krupple-like factors. New York, Springer-Japan, p21-p31.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2007). Generation of induced pluripotent

stem cells without Myc from mouse and human fibroblasts. Nature Biotechnology *26*, 101-106.

Nakagawa, M., and Yamanaka, S. (2012). Function of Myc for generation of induced pluripotent stem cells. 79-85.

Navarro, P., Moffat, M., Mullin, N.P., and Chambers, I. (2011). The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. Human Genetics *130*, 255-264.

Neri, F., Zippo, A., Krepelova, A., Cherubini, A., Rocchigiani, M., and Oliviero, S. (2011). Myc Regulates the transcription of the PRC2 gene to control the expression of developmental genes in embryonic stem cells. Molecular and Cellular Biology *32*, 840-851.

Nichols, J., Davidson, D., Taga, T., Yoshida, K., Chambers, I., and Smith, A. (1996). Complementary tissue-specific expression of LIF and LIF-receptor mRNAs in early mouse embryogenesis. Mechanisms of Development *57*, 123-131.

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., 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 Oct4. Cell *95*, 379-391.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, Douglas R., Tessarollo, L., Casellas, R., *et al.* (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell *151*, 68-79.

Niwa, H., Miyazaki, J.-i., and Smith, A.G. (2000a). Quantitative expression of Oct4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature Genetics *24*, 372-376.

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118-122.

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.

Nowick, K., Carneiro, M., and Faria, R. (2013). A prominent role of KRAB-ZNF transcription factors in mammalian speciation? Trends in Genetics *29*, 130-139.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Molecular and Cellular Biology *21*, 4330-4336.

O'Leary, T., Heindryckx, B., Lierman, S., van Bruggen, D., Goeman, J.J., Vandewoestyne, M., Deforce, D., de Sousa Lopes, S.M.C., and De Sutter, P. (2012). Tracking the progression of the human inner cell mass during embryonic stem cell derivation. Nature biotechnology *30*, 278-282.

Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., and Miyazono, K. (2006). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. Journal of Cell Science *120*, 55-65.

Okamura, D., Tokitake, Y., Niwa, H., and Matsui, Y. (2008). Requirement of Oct3/4 function for germ cell specification. Developmental Biology *317*, 576-584.

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2004). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. Journal of Biological Chemistry *280*, 5307-5317.

Ouwens, D.M., Ruiter, N.D.d., Zon, G.C.M.v.d., Carter, A.P., Schouten, J., Burgt, C.v.d., Kooistra, K., Bos, J.L., Maassen, J.A., and Dam, H.v. (2002). Growth factors can activate

ATF2 via a two-step mechanism phosphorylation of Thr71 through the Ras-MEK-ERK pathway and of Thr69 through RalGDS-Src-p38. The EMBO Journal *21*, 3782-3793.

Ow, J.R., Ma, H., Jean, A., Goh, Z., Lee, Y.H., Chong, Y.M., Soong, R., Fu, X.-Y., Yang, H., and Wu, Q. (2014). Patz1 regulates embryonic stem cell identity. Stem Cells and Development *23*, 1062-1073.

Paling, N.R.D., Wheadon, H., Bone, H.K., and Welham, M.J. (2004). Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling. Journal of Biological Chemistry *279*, 48063-48070.

Palmieri, S.L., Peter, W., Hess, H., and Schöler, H.R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. Developmental Biology *166*, 259-267.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 jysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell *1*, 299-312.

Pan, G.J., CHANG, Z.Y., SCHÖLER, H.R., and PEI, D. (2002). Stem cell pluripotency and transcription factor Oct4. Cell Research *12*, 321-329.

Panepucci, R.A., Ramos-Mejía, V., Montes, R., Bueno, C., Ayllón, V., Real, P.J., Rodríguez, R., and Menendez, P. (2012). Residual expression of the reprogramming factors prevents differentiation of iPSC generated from human fibroblasts and cord blood CD34+ progenitors. PLoS ONE *7*, e35824.

Panopoulos, A.D., Yanes, O., Ruiz, S., Kida, Y.S., Diep, D., Tautenhahn, R., Herrerías, A., Batchelder, E.M., Plongthongkum, N., Lutz, M., *et al.* (2011). The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Cell Research *22*, 168-177.

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

Pasini, D., Bracken, A.P., Jensen, M.R., Denchi, E.L., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. The EMBO Journal *23*, 4061-4071.

Pera, M.F. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. Journal of Cell Science *117*, 1269-1280.

Pero, R. (2001). PATZ attenuates the RNF4-mediated enhancement of androgen receptor-dependent transcription. Journal of Biological Chemistry *277*, 3280-3285.

Pero, R., Palmieri, D., Angrisano, T., Valentino, T., Federico, A., Franco, R., Lembo, F., Klein-Szanto, A.J., Del Vecchio, L., Montanaro, D., *et al.* (2012). POZ-, AT-hook-, and Zinc finger-containing protein (PATZ) interacts with human oncogene B cell lymphoma 6 (BCL6) and is required for its negative autoregulation. Journal of Biological Chemistry *287*, 18308-18317.

Pethe, P., Nagvenkar, P., and Bhartiya, D. (2014). Polycomb group protein expression during differentiation of human embryonic stem cells into pancreatic lineage in vitro. BMC Cell Biology *15*, 18.

Pfendler, K.C., Catuar, C.S., Meneses, J.J., and Pedersen., R.A. (2005). Overexpression of Nodal promotes differentiation of mouse embryonic stem cells into mesoderm and endoderm at the expense of neuroectoderm formation. Stem Cells and Development *14*, 162-172.

Polo, Jose M., Anderssen, E., Walsh, Ryan M., Schwarz, Benjamin A., Nefzger, Christian M., Lim, Sue M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., *et al.* (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617-1632.

Posfai, E., Tam, O.H., and Rossant, J. (2014). Mechanisms of pluripotency in vivo and in vitro. Current Topics in Developmental Biology. *107*, 1-37.

Prowse, A.B.J., McQuade, L.R., Bryant, K.J., Marcal, H., and Gray, P.P. (2007). Identification of potential pluripotency determinants for human embryonic stem cells following proteomic analysis of human and mouse fibroblast conditioned media. Journal of Proteome research *6*, 3796-3807.

Puceat, M. (2007). TGFβ in the differentiation of embryonic stem cells. Cardiovascular Research *74*, 256-261.

Qi, X., Li, T.G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., and Zhao, G.Q. (2004). BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. Proceedings of the National Academy of Sciences *101*, 6027-6032.

Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.-d., and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature *485*, 593-598.

Radzisheuskaya, A., Le Bin Chia, G., dos Santos, R.L., Theunissen, T.W., Castro, L.F.C., Nichols, J., and Silva, J.C.R. (2013). A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. Nature Cell Biology *15*, 579-590.

Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I., Krupalnik, V., Zerbib, M., *et al.* (2013). Deterministic direct reprogramming of somatic cells to pluripotency. Nature *502*, 65-70.

Ramos, C., Becerril, C., Montaño, M., Garcí-De-Alba, C., Ramírez, R., Checa, M., Pardo, A., and Selman, M. (2010). FGF-1 reverts epithelial-mesenchymal transition induced by TGF-β1 through MAPK/ERK kinase pathway. American Journal of Physiology - Lung Cellular and Molecular Physiology *229*, 222-231.

Reijo Pera, R.A., DeJonge, C., Bossert, N., Yao, M., Hwa Yang, J.Y., Asadi, N.B., Wong, W., Wong, C., and Firpo, M.T. (2009). Gene expression profiles of human inner cell mass cells and embryonic stem cells. Differentiation *78*, 18-23.

Remenyi, A. (2003). Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes & Development *17*, 2048-2059.

Richter, A., Valdimarsdottir, L., Hrafnkelsdottir, H.E., Runarsson, J.F., Omarsdottir, A.R., Oostwaard, D.W.-v., Mummery, C., and Valdimarsdottir, G. (2014). BMP4 promotes EMT and mesodermal commitment in human embryonic stem cells via SLUG and MSX2. Stem Cells *32*, 636-648.

Robertson, E.J., Norris, D.P., Brennan, J., and Bikoff, E.K. (2003). Control of early anterior-posterior patterning in the mouse embryo by TGF- signalling. Philosophical Transactions of the Royal Society B: Biological Sciences *358*, 1351-1358.

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

Rodda, D.J. (2005). Transcriptional regulation of Nanog by OCT4 and SOX2. Journal of Biological Chemistry *280*, 24731-24737.

Sakaguchi, S., Hombauer, M., Bilic, I., Naoe, Y., Schebesta, A., Taniuchi, I., and Ellmeier, W. (2010). The zinc-finger protein MAZR is part of the transcription factor network that controls the CD4 versus CD8 lineage fate of double-positive thymocytes. Nature Immunology *11*, 442-448.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.-k., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64-77.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2003a). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nature Medicine *10*, 55-63.

Sato, N., Sanjuan, I.M., Heke, M., Uchida, M., Naef, F., and Brivanlou, A.H. (2003b). Molecular signature of human embryonic stem cells and its comparison with the mouse. Developmental Biology *260*, 404-413.

Saunders, A., Faiola, F., and Wang, J. (2013). Concise Review: Pursuing self-renewal and pluripotency with the stem cell factor Nanog. Stem Cells *31*, 1227-1236.

Schübeler, D., Yang, S.-H., Kalkan, T., Morrisroe, C., Smith, A., and Sharrocks, A.D. (2012). A genome-wide RNAi screen reveals MAP kinase phosphatases as key ERK pathway regulators during embryonic stem cell differentiation. PLoS Genetics, e1003112.

Schier, A.F. (2003). Nodalsignaling Invertebratedevelopment. Annual Review of Cell and Developmental Biology *19*, 589-621.

Schnerch, A., Cerdan, C., and Bhatia, M. (2010). Distinguishing between mouse and human pluripotent stem cell regulation: The best laid plans of mice and men. Stem Cells *28*, 419-430.

Schwarz, Benjamin A., Bar-Nur, O., Silva, José C.R., and Hochedlinger, K. (2014). Nanog Is Dispensable for the Generation of Induced Pluripotent Stem Cells. Current Biology *24*, 347-350.

Scotland, K.B., Chen, S., Sylvester, R., and Gudas, L.J. (2009). Analysis of Rex1 (zfp42) function in embryonic stem cell differentiation. Developmental Dynamics *238*, 1863-1877.

Shen, S., Pu, J., Lang, B., and McCaig, C.D. (2011). A zinc finger protein Zfp521 directs neural differentiation and beyond. Stem Cell Research and Therapy *2*, 20.

Shi, Y., and Massagué, J. (2003). Mechanisms of TGF-β signaling from cell membrane to the nucleus. Cell *113*, 685-700.

Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. BMC Developmental Biology *7*, 80.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722-737.

Singh, A.M., and Dalton, S. (2009). The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. Cell Stem Cell *5*, 141-149.

Singh, A.M., Hamazaki, T., Hankowski, K.E., and Terada, N. (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells *25*, 2534-2542.

Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C.G., Zavolan, M., Svoboda, P., and Filipowicz, W. (2008). MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nature Structural & Molecular Biology *15*, 259-267.

Smith, J.R., Vallier, L., Lupo, G., Alexander, M., Harris, W.A., and Pedersen, R.A. (2008). Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Developmental Biology *313*, 107-117.

Soufi, A., Donahue, G., and Zaret, Kenneth S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O., and Thomson, J.A. (2003). Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proceedings of the National Academy of Sciences *100*, 13350-13355.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364-377.

Staerk, J., Lyssiotis, C.A., Medeiro, L.A., Bollong, M., Foreman, R.K., Zhu, S., Garcia, M., Gao, Q., Bouchez, L.C., Lairson, L.L., *et al.* (2011). Pan-Src family kinase inhibitors replace Sox2 during the direct reprogramming of somatic cells. Angewandte Chemie International Edition *50*, 5734-5736.

Stavridis, M.P., Lunn, J.S., Collins, B.J., and Storey, K.G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development *134*, 2889-2894.

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. Nature Biotechnology *29*, 443-448.

Suwińska, A., Czołowska, R., Ożdżeński, W., and Tarkowski, A.K. (2008). Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: Expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos. Developmental Biology *322*, 133-144.

Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F.H., Rodriguez-Esteban, C., and Izpisua Belmonte, J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proceedings of the National Academy of Sciences *103*, 10294-10299.

Szabo, E., Rampalli, S., Risueño, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. Nature *468*, 521-526.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Current Biology *11*, 1553-1558.

Takahashi, K., Mitsui, K., and Yamanaka, S. (2003). Role of ERas in promoting tumour-like properties in mouse embryonic stem cells. Nature *423*, 541-545.

Takahashi, K., Murakami, M., and Yamanaka, S. (2005). Role of the phosphoinositide 3-kinase pathway in mouse embryonic stem (ES) cells. Biochemical Society Transactions *33*, 1522-1525.

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.

Takana, T.S. (2009). Transcriptional heterogeneity in mouse embryonic stem cells. Reproduction, Fertility and Development *21*, 67-75.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M.A. (2010). Tracing the Derivation of Embryonic Stem Cells from the inner cell mass by single-cell RNA-seq analysis. Cell Stem Cell *6*, 468-478.

Tarkowski, A.K. (1959). Totipotency and lineage segregation in the human embryo. Development *184*, 1286-1287.

Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature *455*, 1124-1128.

Teo, A.K.K., Ali, Y., Wong, K.Y., Chipperfield, H., Sadasivam, A., Poobalan, Y., Tan, E.K., Wang, S.T., Abraham, S., Tsuneyoshi, N., *et al.* (2012). Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. Stem Cells *30*, 631-642.

Terada, N., Hamazaki, T., Oka, M., Hoki, M., Mastalerz, D.M., Nakano, Y., Meyer, E.M., Morel, L., Petersen, B.E., and Scott, E.W. (2002). Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. Nature *416*, 542-545.

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.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.

Thomson, J.A. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tian, X., Sun, D., Zhang, Y., Zhao, S., Xiong, H., and Fang, J. (2008). Zinc finger protein 278, a potential oncogene in human colorectal cancer. Acta Biochimica et Biophysica Sinica *40*, 289-296.

Tiemann, U., Marthaler, Adele G., Adachi, K., Wu, G., Fischedick, Gerrit Ulf L., Araúzo-Bravo, Marcos J., Schöler, Hans R., and Tapia, N. (2014). Counteracting activities of OCT4 and KLF4 during reprogramming to pluripotency. Stem Cell Reports *2*, 351-365.

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.

Tritz, R., Mueller, B.M., Hickey, M.J., Lin, A.H., Gomez, G.G., Hadwiger, P., Sah, D.W.V., Muldoon, L., Neuwelt, E.A., and Kruse, C.A. (2008). siRNA down-regulation of the PATZ1 gene in human glioma cells increases their sensitivity to apoptotic stimuli. Cancer Therapy *6*, 865-876.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009a). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. Journal of Cell Science *122*, 3502-3510.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009b). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145-1148.

Valentino, T., Palmieri, D., Vitiello, M., Pierantoni, G.M., Fusco, A., and Fedele, M. (2013a). PATZ1 interacts with p53 and regulates expression of p53-target genes enhancing apoptosis or cell survival based on the cellular context. Cell Death and Disease *4*, e963.

Valentino, T., Palmieri, D., Vitiello, M., Simeone, A., Palma, G., Arra, C., Chieffi, P., Chiariotti, L., Fusco, A., and Fedele, M. (2013b). Embryonic defects and growth alteration in mice with homozygous disruption of thePatz1gene. Journal of Cellular Physiology *228*, 646-653.

Vallier, L. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. Journal of Cell Science *118*, 4495-4509.

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.H., Martinez, A., Rugg-Gunn, P., *et al.* (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development *136*, 1339-1349.

Vallier, L., Reynolds, D., and Pedersen, R.A. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. Developmental Biology *275*, 403-421.

Van den Berg, D.L.C., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369-381.

Varlakhanova, N.V., Cotterman, R.F., deVries, W.N., Morgan, J., Donahue, L.R., Murray, S., Knowles, B.B., and Knoepfler, P.S. (2010). Myc maintains embryonic stem cell pluripotency and self-renewal. Differentiation *80*, 9-19.

Verfaillie, C.M., Zheng, Z., de Iongh, R.U., Rathjen, P.D., and Rathjen, J. (2010). A Requirement for FGF Signalling in the formation of primitive streak-like intermediates from primitive ectoderm in culture. PLoS ONE *5*, e12555.

Villegas, S.N., Canham, M., and Brickman, J.M. (2010). FGF signalling as a mediator of lineage transitions-Evidence from embryonic stem cell differentiation. Journal of Cellular Biochemistry *110*, 10-20.

Wang, J., He, Q., Han, C., Gu, H., Jin, L., Li, Q., Mei, Y., and Wu, M. (2012). p53-facilitated miR-199a-3p regulates somatic cell reprogramming. Stem Cells *30*, 1405-1413.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. (2008). Embryonic stem cell–specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nature Genetics *40*, 1478-1483.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007a). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nature Genetics *39*, 380-385.

Wang, Z.-X., Kueh, J.L.L., Teh, C.H.-L., Rossbach, M., Lim, L., Li, P., Wong, K.-Y., Lufkin, T., Robson, P., and Stanton, L.W. (2007b). Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. stem cells *25*, 2173-2182.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., *et al.* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell *7*, 618-630.

Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell *2*, 10-12.

Whitworth, D.J., Ovchinnikov, D.A., Sun, J., Fortuna, P.R.J., and Wolvetang, E.J. (2014). Generation and characterization of leukemia inhibitory factor-dependent equine induced pluripotent stem cells from adult dermal fibroblasts. Stem Cells and Development *23*, 1515-1523.

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.

Wu, Q., Chen, X., Zhang, J., Loh, Y.H., Low, T.Y., Zhang, W., Sze, S.K., Lim, B., and Ng, H.H. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. Journal of Biological Chemistry *281*, 24090-24094.

Wu, Q., Hamilton, W.B., Kaji, K., and Kunath, T. (2013). ERK2 suppresses self-renewal capacity of embryonic stem cells, but is not required for multi-lineage commitment. PLoS ONE, e60907.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. Nature Biotechnology *19*, 971-974.

Xu, R.-H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nature Methods *2*, 185-190.

Xu, R.-H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., *et al.* (2008). NANOG is a direct target of TGFβ/Activin-mediated SMAD signaling in human ESCs. Cell Stem Cell *3*, 196-206.

Yamanaka, S. (2009). Elite and stochastic models for induced pluripotent stem cell generation. Nature *460*, 49-52.

Yeo, J.-C., and Ng, H.-H. (2012). The transcriptional regulation of pluripotency. Cell Research *23*, 20-32.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Michael Gross, Hübner, K., and Schöler, H.R. (1996). Germline regulatory element of Oct4-specific for the totipotency cycle of embryonal cells. Development *122*, 881-894.

Ying, Q.-L., Jennifer Nichols, Ian Chambers, 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.

Yoshikawa, T., Piao, Y., Zhong, J., Matoba, R., Carter, M.G., Wang, Y., Goldberg, I., and Ko, M.S.H. (2006). High-throughput screen for genes predominantly expressed in the ICM of mouse blastocysts by whole mount in situ hybridization. Gene Expression Patterns *6*, 213-224.

Young, Richard A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Yu, P., Pan, G., Yu, J., and Thomson, James A. (2011). FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stemcell differentiation. Cell Stem Cell *8*, 326-334.

Zhang, J., Nuebel, E., Daley, George Q., Koehler, Carla M., and Teitell, Michael A. (2012). Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. Cell Stem Cell *11*, 589-595.

Zhao, W., Ji, X., Zhang, F., Li, L., and Ma, L. (2012). Embryonic stem cell markers. Molecules *17*, 6196-6236.

Zhao, W., Li, Q., Ayers, S., Gu, Y., Shi, Z., Zhu, Q., Chen, Y., Wang, Helen Y., and Wang, R.-F. (2013). Jmjd3 inhibits reprogramming by upregulating expression of INK4a/Arf and targeting PHF20 for ubiquitination. Cell *152*, 1037-1050.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.-K., Shahab, A., Kuznetsov, V.A., *et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell *1*, 286-298.

Zheng, H., Ying, H., Yan, H., Kimmelman, A.C., Hiller, D.J., Chen, A.-J., Perry, S.R., Tonon, G., Chu, G.C., Ding, Z., *et al.* (2008). p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. Nature *455*, 1129-1133.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., *et al.* (2012). HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. The EMBO Journal *31*, 2103-2116.

Zhou, Y., Kim, J., Yuan, X., and Braun, T. (2011). Epigenetic modifications of stem cells: A paradigm for the control of cardiac progenitor cells. Circulation Research *109*, 1067-1081.

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

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell Stem Cell *7*, 651-655.

Zuo, X., Sheng, J., Lau, H.T., McDonald, C.M., Andrade, M., Cullen, D.E., Bell, F.T., Iacovino, M., Kyba, M., Xu, G., *et al.* (2011). Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. Journal of Biological Chemistry *287*, 2107-2118.

APPENDICES

Appendix 1. Sequences of primers for plasmid constructions

Zfp322a shRNA-1 into pSuper.puro	F	gatccccacagaggagctccattatattcaagagatataatggagctcctctgtttttta
	R	agcttaaaaaacagaggagctccattatatctcttgaatataatggagctcctctgtggg
Zfp322a shRNA-2 into pSuper.puro	F	gatccccacttagcacgaatgtgtgattcaagagatcacacattcgtgctaagtttttta
	R	agcttaaaaacttagcacgaatgtgtgatctcttgaatcacacattcgtgctaagtggg
Zfp322a cDNA into pPyCAGIP	F	atatactcgagatgtatccttatgacgtgcctgactatgccgaaaaccacttaggagagaacc
	R	taatagcggccgctcaagacacacgtgaggttt
Oct4 CR4 into pGL3-promoter vector	F	atataggatcccacaatccataagacaaggttgg
	R	tatatgtcgacagcttcctcaatagcagattaag
Oct4 CR1 into pGL3-promoter vector	F	atatactcgagagctggggaagtcttgtgtg
	R	tatatagatctggtgggaggtgggtagagag
Nanog promoter into pGL3-promoter vector	F	cgcgtcgactaaagtgaaatgaggtaaagcc
	R	cgcggatccggaaagatcatagaaagaagag
Zfp322a cDNA into pMXs	F	atatacaattgatggaaaaccacttaggagagaac
	R	taatagcggccgctcaagacacacgtgaggtttcc
Patz1 shRNA into pSuper.retro.puro	F	gatccccctggagatgcacaccatcattcaagagatgatggtgtgcatctccagttttta
	R	agcttaaaaactggagatgcacaccatcatctcttgaatgatggtgtgcatctccagggg
Patz1 cDNA into pMXs	F	atatacaattgatggagcgggtcaacgacgcttc
	R	taatagcggccgctcacttcccttcaggccccatg

Appendix 2. Sequences of primers for qPCR

qPCR primers for *Zfp322a* RNAi assays		
β-actin	accaactgggacgacatggagaa	tacgaccagaggcatacagggac
Pou5f1	ttgggctagagaaggatgtggtt	ggaaaagggactgagtagagtgtgg
Sox2	ccaggagaaccccaagatgcacaact	aagcctccgggaagcgtgtacttatcctt
Rex1	agatggcttccctgacggatacctagagt	ctcttgcttcgtcccctttgtcatgtac
Nanog	ggttgaagactagcaatggtctga	tgcaatggatgctgggatactc
Zfp322a	tcaaaagggtgaaaaagatttatattcg	tcacatctataaggtttctcctcggtata
Bmp4	gttcctggacacctcatcacacgactact	gtaacgatcggctgattctgacatgct
Cdx2	cgcagaactttgtcagtcctccgcagtacc	gtattcggcggggctgctgtagcccatagc
Pax5	gggcgcagacggcatgtatgataaa	agtcgcatctgagcttcatccgagtcttc
Nestin	agaggaagagcagcaaggccatgac	tccctgactctgctccttcttcttcat
Rest	cccttccgttgtaagccatgccagtatga	tggtgcttcaggtgtgccgtgtagtgat
Fgf5	gcagatctacccggatggcaaagtcaatg	gttctgtggatcgcggacgcataggtatt
Hand1	cctgcccaaacgaaaaggctcaggacccaa	cgaccgccatccgtctttttgagttcagcc
Brachyury	gagctgtggctgcgcttcaaggagctaac	ccccaggtacccactccccgttcacata
Gata2	ggcctcttcttctgcaggggtagtgtag	gcacataggagggataggtgggtatcgg
Nkx2-5	gaaggcagtggagctggacaaagccgaga	ggaaccagatcttgacctgcgtggacgtg
Gata6	tgtgcaatgcatgcggtctctacagca	ttcatagcaagtggtcgaggcaccc
Sox17	tgaaaggcgaggtggtggcgagtag	caacgccttccaagacttgcctagcatct
Foxa2	cctacgccaacatgaactcgatga	gtagaaagggaagaggtccatgatccact
qPCR primers for *Zfp322a* RNAi gene expression microarray results validation		
Fgf4	gggcatcggattccacctg	gctgctcatagccacgaagaa
Esrrb	aaccgaatgtcgtccgaagac	gtggctgagggcatcaatg
Klf4	ggcgagtctgacatggctg	gctggacgcagtgtcttctc
Slc16a3	tcacgggtttctcctacgc	gccaaagcggttcacacac
Tcea2	aggctggacaaaatggtgacc	tgcccctcccttgattcct
G9a	ccgccgagagagttcatagc	ggttcgtccccgatgagtg
Prelid2	cagtacccctttcgagcagg	gactgagccatgactcttcttc
Ogt	tgagttggcacatcgagaatatc	gaacaccagtattgtcaggctc
Eomes	ggcccctatggctcaaattcc	cctgccctgtttggtgatg
Cdkn1b	tcaaacgtgagagtgtctaacg	ccgggccgaagagatttctg
Foxa1	acattcaagcgcagctaccc	tgctggttctggcggtaatag
Nrp1	gacaaatgtggcgggaccata	tggattagccattcacacttctc
Pax6	taccagtgtctaccagccaat	tgcacgagtatgaggaggtct
Gadd45	aatatgactttggaggaattc	attcggatgccatcaccgttc
Gm428	acccaactgactgcctta	tctgctttctccttcctg
Nupr1	cccttcccagcaacctctaaa	tcttggtccgacctttccga
Sox17	tgaaaggcgaggtggtggcgagtag	caacgccttccaagacttgcctagcatct
Hsph1	caggtacaaactgatggtcaaca	tgaggtaagttcaggtgaaggg
Igfbp3	ccaggaaacatcagtgagtcc	ggatggaacttggaatcggtca
Cubn	tcctcaggaaattaccgcagt	gattgccaaatgctgggtgta
qPCR primers for Zfp322a ChIP		
Amplicon 1 of *Pou5f1* regulatory region	gtggtggagagtgctgtctaggccttag	agcagattaaggaagggctaggacgagag
Amplicon 2 of *Pou5f1* regulatory region	tgctctgggctttttgaggctgtgtgatt	tggcggaaagacactaaggagacgggatt

Amplicon 3 of *Pou5f1* regulatory region	ggggaggggtgggtgacgaggatga	tactcaacccttgaatgggccaggatggct
Amplicon 4 of *Pou5f1* regulatory region	gggggtggttagtgtctaatctaccaacct	acccagtatttcagcccatgtccaa
Amplicon 1 of *Nanog* promoter region	atttcttcttccattgcttagacggctgag	ctaccaccatgcccaatttaaggagtgttt
Amplicon 2 of *Nanog* promoter region	ccaggtttcccaatgtgaagagcaagcaa	tggcgatctctagtgggaagtttcaggtca
Amplicon 3 of *Nanog* promoter region	gaggatgccccctaagctttccctccc	cctcctaccctacccaccccctattctccc
Amplicon 4 of *Nanog* promoter region	ctctttctgtgggaaggctgcggctcactt	catgtcagtgtgatggcgagggaaggga
Amplicon 5 of *Nanog* promoter region	gcgggtgtccttatcactcttctggaaa	tccaagctaggatgttaggtctccctgcta
qPCR primers for Zfp322a ChIP-seq results validation		
Nrf1	gcccagccaaaccgccacc	gccgggagaacccgaggc
Zfx	cgcagcgaggccactgggct	agctgacaaaaagcggccca
Arid1b	gccatcgagccacggtcg	gcgggagggtttcggaga
Myl4	agggctcgccgaaaggga	atgctagcccctgctccgag
Polr2a	tctcggcgcttctgaggag	agggagccggagccctag
Tcea2	tccccctggcattcggtg	gaccacaactcccagacgactg
Axin1	gctggtgagtagtgggagc	cttggatctcaggtctccctac
Btf3	cctgagactcccgtcctcg	cgaggcagggaggggacg
Grk5	gccgaaccccttgctgcc	cggaaagcacaagccaagg
Lrig2	agatgcccggaggctaaacc	cgattggctggctgctga
Hnrnpc	ggccagatcaccgcatttc	ctagagataactctcctcccttc
Prpf4	gccagtgacgcacttccact	ccggcctctctagggcag
Wdr62	ccagcacccactcgaatagc	acccagccagactcagtaagc
Nrtn	atccgcatctacgacctggg	agctcttgcagcgtgtggtag
Bmyc	gcttctcggcctccgccaa	tcttacgcccaggatttgag
Ubc	cgtcggagactgtggggtg	cctacaaccgagggaaagcta
Evi5	gatctccccgcctccgtt	acgcctcctgctcaggaacc
Klf2	gctgagcccggagctcgt	aaagtggcaaaggacggcaa
Setd3	ccctcttccaccgggacg	caccggaggaaggaaacca
Ntn5	gagaacctgggcagagaaga	tctccactgagtcttccaggc
Trim28	tgtctcagaagcgaggga	caacggccaggcagccg
Fuz	ggacccacaggtgaggcgg	tgccaccagcgttccatta
Cdk5	ttctcgtatttctgcattgc	acaagtcccagatgccagtg
Crat	caagcaaacagaacccgacg	ccggtgcgaggcactca
Yes1	taggcagctgacccggacc	caggagacgagacactcacg
Stag1	cagagggcggtcgggacc	ccaaaagtctccggtgtgt
Carm1	gacaaaaagatgcagagcacc	cctaggctggcgtgctgc
Cage1	gccctgctggtggcatactt	ctgcttcccacctagaggatc
qPCR primers for H3K4me3 and H3K27me3 ChIP		
Amplicon 1of *Arf/Ink4a* locus	ctcttgcttcgtccccttttgtcatgtac	ggcggagaaatggcttatttc
Amplicon 2 of *Arf/Ink4a* locus	tgtgacaagcgaggtgagaagc	atgggcgtggagcaaagatg
Amplicon 3 of *Arf/Ink4a* locus	ttcccaggagctgaaattccag	aaaaattcccaacacccacttgc
qPCR primers for Patz1 ChIP		
Corest	cccagtagtgtgtggaaccg	gaccggaatggggcgcac
Ncor1	gcgatggctgccacggc	agaacggagtgcggctctt
Tbl1x	tgggcgcgctccctgca	tgggacgaactgcagagcgt
Chd4	caggagcagctgggccaat	agatgttacctggatggggtg

Mbd3	aaggcgggacgtagccaact	gctcgcaggaccagtctgac
Cdh1	tttggagcttggcggctggtt	aaagtaagcaaactggtgacttgag
Tgfbr2	cctctcaggatgcgggccagatgt	tcggcgcccgggtaaagttgatgag
Zeb2	ttaggactccccccccaagc	ggagcggctgtgaaagttaagg
Tgfb1	gctaatggtggaccgcaacaacg	cagcctctttgggacacccac
Smad3	ctaggggggctgggccagtgc	caggaggagaagtggtgcgaga
Smad7	taaaacaaaacgaatgaatgaagc	ggaggtgggaggccgagacg
Snail1	ctgcggggagcctttaccttc	gtgggcagtgcctggcaagg
Smad2	ccagccctaataccaaccgcac	ctctgaagcagcctgggtcctg
qPCR primers for gene expression determination		
Patz1	gagcttcttccgttctaagtcctacttga	actaaagatgatgcaaacgctgactg
Cdkn2a	cgcaggttcttggtcactgt	tgttcacgaaagccagagcg
Arf	gccgcaccggaatcct	ttgagcagaagagctgctacgt
Cdkn1b	tcaaacgtgagagtgtctaacg	ccgggccgaagagatttctg

Appendix 3. List of antibodies

Primary antibodies		
Protein specificity	Company	Catalog#
β-actin	Santa Cruz	sc-81178
Zfp322a	Santa Cruz	Sc-102205
Oct4	Santa Cruz	sc-8628
Nanog	Santa Cruz	sc-33760
Sox2	Santa Cruz	sc-99000
Rex1	Santa Cruz	sc-377095
t-Erk	Cell Signaling	sc-7383
p-Erk	Santa Cruz	137F5
SSEA-1	Millipore	mab34301
SMA	Abcam	ab5694
Gata4	Santa Cruz	sc-25310
Nestin	R&D	mab2736
Patz1	Santa Cruz	sc-292109
c-Myc	Santa Cruz	sc-788
P16	Abcam	ab51243
H3ac	Millipore	06-599
H3K4me3	Abcam	ab8580
H3K27me3	Abcam	ab6002
HP1α	Abcam	ab77256
H3K36me3	Abcam	ab9050
Hdac3	Santa Cruz	sc-11417
Sin3a	Santa Cruz	sc-994
Hdac2	Santa Cruz	sc-7899
Klf4	Santa Cruz	sc-292109
Secondary antibodies		
Anti-Mouse	GE Healthcare Life Sciences	RPN4201
Anti-Rabbit	GE Healthcare Life Sciences	NA934V
Anti-Goat	Santa Cruz	Sc-2020
Alexa Fluor® conjugated secondary antibodies		
Anti-Mouse Alexa Fluor® 594	Life Technologies	A11032
Anti-Rabbit Alexa Fluor® 594	Life Technologies	A11037
Anti-Goat Alexa Fluor® 594	Life Technologies	A11058
Anti-Mouse Alexa Fluor® 488	Life Technologies	A21202
Anti-Rabbit Alexa Fluor® 488	Life Technologies	A21206
Anti-Goat Alexa Fluor® 488	Life Technologies	A11055

Appendix 4. 50 binding sites with top-ranked peak heights in Zfp322a ChIP-seq analysis

Transcript ID	Gene Symbol	Distance to TSS	Zfp322a: total reads in region	Peak Height
NM_008828	Pgk1	0	766	412
NR_030494	Mir715	0	3048	193
NM_007697	Chl1	138156	263	159
NM_001166549	Eif4enif1	-2744	619	117
NM_001145676	2210408I21Rik	303161	150	111
NM_001081436	Ino80d	0	133	102
NM_009340.1	Tcp10a	244752	135	87
NM_001103165	Pcbp2	0	118	81
NM_001044751	Hsd11b1	14086	108	81
NR_028428	2610005L07Rik	0	240	79
NM_176996	Smo	12222	259	79
NR_030708	6820431F20Rik	0	193	78
NM_027293	Dopey2	77011	119	78
NM_027290	Mcm10	2107	158	72
NM_001038607	Kcnh1	263008	85	69
NM_001033425	Zscan10	-1133	100	68
NM_030207	Sfi1	23320	1364	68
NM_001040397	Filip1l	37840	113	67
NM_001080548	Usp6nl	21705	267	66
NR_027956	1700052K11Rik	0	118	65
NM_145990	Cdk5rap2	31754	106	65
NM_148413	Myo3a	293741	282	64
NM_013846	Ror2	15762	81	64
NM_172383	Tmem125	-3283	284	62
NM_130452	Bbox1	11673	72	60
NM_001159953	Pde1c	128292	71	58
NM_015764	Greb1	52616	75	57
NM_008480	Lama1	44042	71	56
NM_183151	Mid1	105169	2089	55
NM_019574	Patz1	0	222	55
NM_153599	Cdk8	29306	103	55
NM_053011	Lrp1b	756966	60	54
NR_003518	Pisd-ps3	1097	1815	52
NM_181595	Ppp1r9a	20440	97	52
NM_001193266	Mdga2	162555	90	51
NM_031881	Nedd4l	80407	54	51
NM_018797	Plxnc1	72016	52	50
NM_001033266	Gm525	7659	78	50
NM_009439	Psmd3	0	62	49
NM_015820	Hs6st3	148662	53	49
NM_027600	4921504E06Rik	44158	54	48

NM_001111107	Zfp322a	11378	172	48
NM_018821	Socs6	0	62	47
NM_146241	Trhde	63572	49	47
NM_001145676	2210408I21Rik	169384	48	47
NM_025341	Abhd6	3163	49	46
NM_177393	Nalcn	175104	47	46
NR_003519	Pisd-ps2	0	385	45
NM_021377	Sorcs1	367575	49	45
NM_053122	Immp2l	491001	51	45
NM_007525	Bard1	23206	49	44
NM_001001295	Dis3l	0	116	43

Appendix 5. Representative enriched gene ontology terms for ChIP-Seq targets

Function groups	GO terms	No. of genes	p-value
Gene expression	gene expression	916	9.54E-53
	transcription, DNA-dependent	590	2.27E-23
	translation	133	1.71E-12
	posttranscriptional regulation of gene expression	97	9.42E-07
	transcription from RNA polymerase II promoter	247	0.008354
	translational initiation	30	0.036402
Chromatin structures	chromosome organization	175	2.25E-13
	chromatin modification	108	4.21E-05
	chromatin assembly or disassembly	35	0.000161
	nucleosome organization	31	0.000107
	histone modification	76	0.001964
Cell death	programmed cell death	332	9.29E-12
	cell death	342	5.09E-11
	apoptotic process	328	1.42E-11
	anti-apoptosis	58	0.020638
Cell proliferation	cell cycle	334	4.03E-31
	mitotic cell cycle	148	4.60E-10
	cell division	119	7.60E-08
	cell cycle arrest	75	9.01E-08
	M phase	125	3.17E-08
	chromosome segregation	51	1.20E-06
	DNA replication	72	3.53E-06
	cell cycle checkpoint	48	5.32E-06
DNA repair	response to DNA damage stimulus	148	5.09E-15
	DNA repair	101	2.58E-09
	signal transduction in response to DNA damage	32	0.003629
	DNA integrity checkpoint	28	0.008639
	DNA damage checkpoint	27	0.011558
Protein modifications	protein modification process	586	3.60E-42
	protein modification by small protein conjugation or removal	150	7.19E-19
	protein ubiquitination	119	4.17E-15
	protein phosphorylation	281	6.98E-13
	protein glycosylation	35	0.010251
RNA metabolic processes	RNA processing	175	5.14E-21
	RNA splicing	89	2.13E-12
	RNA localization	37	4.96E-06
	RNA transport	36	2.78E-06
Cellular localization and transport	cellular component organization	804	4.05E-37
	establishment of localization in cell	735	1.31E-18
	organelle localization	47	1.30E-05
	cytoskeleton organization	163	0.000149

	membrane organization	78	0.001318
Development	developmental process	723	0.002337
	anatomical structure development	633	0.045452
Macromolecular complex assembly	macromolecular complex assembly	216	9.99E-11
	protein complex assembly	172	1.81E-06
	cellular macromolecular complex assembly	108	0.000128
	protein oligomerization	87	0.024279
Signal transduction	regulation of signaling	357	0.000192
	signal transduction by p53 class mediator	27	0.011558
	small GTPase mediated signal transduction	115	0.036488

Appendix 6. Gene ontology of overlapping genes in the gene expression microarray analysis of *Zfp322a* and *Oct4* RNAi

Function Groups	Gene Ontology terms	No. of genes	Corrected P-value
Development	developmental process	196	1.33E-10
	single-organism developmental process	165	3.48E-09
	anatomical structure development	175	3.79E-09
	system development	149	5.56E-07
	regulation of developmental process	84	3.51E-06
	muscle structure development	41	1.02E-06
	in utero embryonic development	35	1.51E-06
	chordate embryonic development	45	7.82E-06
	embryo development ending in birth or egg hatching	45	1.11E-05
	tissue development	73	2.58E-05
	embryo development	60	2.77E-05
	organ development	112	6.97E-05
	heart development	33	8.54E-05
	muscle cell differentiation	30	8.87E-05
	cardiovascular system development	50	9.04E-05
	circulatory system development	50	9.04E-05
	striated muscle cell differentiation	26	9.95E-05
	striated muscle tissue development	31	0.00012
	regulation of multicellular organismal development	64	0.00208
	muscle organ development	28	0.00404
	heart morphogenesis	19	0.00509
	muscle tissue development	31	0.00032
	positive regulation of developmental process	47	0.00038
	negative regulation of developmental process	37	0.00519
	tissue morphogenesis	34	0.00619
	generation of neurons	53	0.01051
	adrenal gland development	6	0.01626
	nervous system development	71	0.01757
	gonad development	16	0.02882
	epithelium development	36	0.03017
	cell migration	44	0.03138
	neurogenesis	54	0.03235
	gland development	21	0.03238
	mammary gland epithelium development	10	0.03408
	anatomical structure formation involved in morphogenesis	71	0.04907
Cell Development and differentiation	cellular developmental process	143	2.56E-09
	cell differentiation	135	1.01E-08
	cell development	80	1.11E-05
	regulation of cell differentiation	66	6.17E-06
	positive regulation of cell differentiation	37	0.00092

Regulation of gene transcription	regulation of gene expression	160	2.28E-11
	negative regulation of biological process	159	2.33E-11
	regulation of transcription, DNA-dependent	138	8.68E-11
	positive regulation of transcription, DNA-dependent	68	1.29E-07
	transcription, DNA-dependent	138	6.81E-10
	positive regulation of gene expression	76	2.41E-09
	gene expression	177	1.17E-08
	regulation of transcription from RNA polymerase II promoter	74	4.99E-06
	transcription from RNA polymerase II promoter	74	2.38E-05
	positive regulation of transcription from RNA polymerase II promoter	46	0.00147
	negative regulation of sequence-specific DNA binding transcription factor activity	14	0.00787
	negative regulation of transcription, DNA-dependent	47	0.01225
Cell proliferation	regulation of cell proliferation	72	5.22E-08
	cell proliferation	81	2.10E-07
	interphase	21	0.00169
	cell cycle	61	0.00407
	fibroblast proliferation	11	0.0041
	regulation of fibroblast proliferation	11	0.00349
	interphase of mitotic cell cycle	20	0.00369
	G1/S transition of mitotic cell cycle	14	0.00635
	negative regulation of cell proliferation	36	0.00018
	regulation of cell cycle	36	0.04021
Cell death	apoptotic process	84	9.15E-08
	programmed cell death	84	1.78E-07
	death	87	3.21E-07
	cell death	86	5.69E-07
	regulation of apoptotic process	71	9.85E-07
	regulation of programmed cell death	71	1.59E-06
	regulation of cell death	72	3.66E-06
	intrinsic apoptotic signaling pathway	15	0.00069
	intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	8	0.00087
	positive regulation of cell death	34	0.00097
	positive regulation of apoptotic process	32	0.00231
	positive regulation of programmed cell death	32	0.00295
	negative regulation of apoptotic process	39	0.00379
	negative regulation of programmed cell death	39	0.0054
	apoptotic signaling pathway	18	0.01203
	regulation of execution phase of apoptosis	18	0.01203
	execution phase of apoptosis	19	0.01433
	negative regulation of cell death	39	0.02187
Cellular response to stress	cellular response to stress	65	2.31E-07
	response to stress	110	1.66E-06
	response to DNA damage stimulus	37	0.00016
	response to ionizing radiation	14	0.0007
	response to radiation	23	0.00288

	cellular response to ionizing radiation	8	0.00658
	response to X-ray	7	0.00733
	intrinsic apoptotic signaling pathway in response to DNA damage	9	0.00712
	response to abiotic stimulus	39	0.00821
	response to topologically incorrect protein	11	0.00756
	response to unfolded protein	10	0.01918
Chromosome remodelling	chromosome organization	49	3.66E-07
	chromatin modification	36	1.20E-05
	chromatin organization	38	3.82E-05
	chromatin remodeling	12	0.00085
Signal transductions	intracellular signal transduction	85	0.00027
	signal transduction by p53 class mediator	12	0.00257
	response to transforming growth factor beta stimulus	16	0.00259
	cellular response to transforming growth factor beta stimulus	16	0.00259
	regulation of signal transduction	79	0.00289
	transforming growth factor beta receptor signaling pathway	15	0.00313
	regulation of response to stimulus	95	0.00404
	regulation of signaling	87	0.00655
	regulation of cell communication	87	0.00743
	response to growth factor stimulus	23	0.01366
	cellular response to growth factor stimulus	22	0.03141
	positive regulation of cell communication	45	0.04692
	positive regulation of signal transduction	43	0.03831
	positive regulation of signaling	45	0.04424
Protein modifications	macromolecule modification	111	6.59E-05
	cellular protein modification process	108	6.74E-05
	protein modification process	108	6.74E-05
	peptidyl-amino acid modification	40	0.00084
	phosphorylation	69	0.04452
Cellular localization and transport	organelle organization	106	6.62E-08
	cellular component assembly	72	0.00069
	protein complex subunit organization	49	0.00598
	macromolecular complex subunit organization	57	0.00392
	cellular component disassembly	16	0.00973
	cellular component organization	183	3.21E-13
	localization	163	0.00201

Appendix 7. List of Oct4-interacting proteins that are altered upon *Zfp322a* RNAi

Gene Symbols	Fold Change in Microarray analysis
Down-regulated	
Tcfcp2l1	-1.8276629
Chd4	-1.464085696
Gatad2a	-1.484523571
Mta2	-1.337927555
Phc1	-1.222640278
Ep400	-1.681792831
Rcor2	-1.366040257
Wdr5	-1.347233577
Zfp462	-1.613283518
Ubp1	-1.231144413
Hcfc1r1	-1.292352831
Rbpj	-2.114036081
Esrrb	-1.328685814
Pml	-1.602139755
Ctbp2	-1.453972517
Klf5	-1.777685362
L1td1	-1.248330549
Msh2	-1.536875181
Smc1a	-1.366040257
Ogt	-1.879045498
Zcchc8	-1.569168196
0610010K14Rik	-1.790050142
Smarca4 (Brg1)	-1.394743666
Smarcc1	-2.345669898
Ssrp1	-1.319507911
Supt16h	-1.558329159
Cabin1	-1.443929196
Mitf	-1.301341855
Sp1	-1.328685814
Tcfeb	-1.222640278
Rfx2	-1.375541818
Lig3	-1.328685814
Parp1	-1.2397077
Top2a	-1.404444876
Xrcc5	-1.547564994
Xrcc6	-1.729074463
Rpa1	-1.366040257
Dhx9	-1.248330549
Hnrnpab	-1.292352831
Hnrnpu	-1.265756594

Cad	-1.474269217
Dnmt3l	-1.404444876
Myst2	-1.484523571
Ppp2r1a	-1.222640278
Trim24	-1.693490625
Trim33	-1.231144413
Nudc	-1.319507911
Emd	-1.274560627
Up-regulated	
Rad21	1.972465409
Yes1	1.681792831
Nr0b1 (Dax1)	1.986184991
Hdac2	1.394743666
Smarca5	1.526259209
Nfrkb	1.205807828
Asf1a	1.558329159
Ubn2	1.328685814
Ctbp1	1.231144413
Zbtb10	1.375541818
Zbtb2	1.356604327
Zbtb43	1.265756594
Zfhx3	1.602139755
Zfp219	1.214194884
Zic2	2.566851795
Acin1	1.214194884
Zmym2	1.536875181
Rybp	1.283425898
Sall4	1.214194884
Zfp219	1.214194884
Mga	1.310393404
Hells	1.283425898
Zfp143	1.231144413
Rif1	1.670175839
Akap8	1.265756594
Rbm14	1.433955248
Frg1	1.464085696
Sall3	2
Rpa3	1.328685814
Dnmt3a	1.385109468
Kpna2	1.231144413
Kpna3	1.635804117
Dnaja1	1.753211443

Appendix 8. List of publications

Ma H, Ow JR, Tan CP, Goh ZY, Feng B, Loh YH, Fedele M, Li H, Wu Q. The dosage of Patz1 modulates reprogramming process. **Sci Rep** 2014 (accepted)

Ma H, Ng HM, Teh X, Li H, Lee YH, Chong YM, Loh YH, Collins JJ, Feng B, Yang H and Wu Q. Zfp322a regulates mouse ES cell pluripotency and enhances reprogramming efficiency. **PLoS Genet** 2014 10(2): e1004038.

Ow JR*, **Ma H***, Jean A, Lee YH, Chong YM, Soong R, Fu XY, Yang H and Wu Q. Patz1 regulates embryonic stem cell identity. **Stem Cells Dev** 2014 23 (10):1062-1073. *Co-first author

Ma H*, Ow JR*, Chen X and Wu Q. With or without them: essential roles of cofactors in ES Cells. **J Stem Cell Res & Ther** 2012 S10:006. *Co-first author

Lee YH, **Ma H**, Tan TZ, Ng SS, Soong R, Mori S, Fu XY, Zernicka-Goetz M, Wu Q. Protein arginine methyltransferase 6 regulates embryonic stem cell identity. **Stem Cells Dev** 2012 21(14):2613-2622.