THE ROLE OF KRÜPPEL-LIKE FACTORS IN EMBRYONIC STEM CELLS

JIANG JIANMING
(*M.Sc., Peking Union Medical College*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2009

ACKNOWLEDGEMENTS

It is a please to express my sincere thanks to all the people who helped me navigate my PhD studies. I am greatly indebted to my advisor, Dr. Huck-Hui Ng for his invaluable guidance, positive criticism, enlightening discussions and constructive suggestions throughout the candidature which immensely helped me in attaining the scientific and scholarly attitude of a researcher. I greatly admire his guidance and wish to express my sincere gratitude for his constant support, patience and supervision at every stage of my PhD life. I would like to appreciate my co-advisor Dr. Keh Chuang Chin for his kind help and guidance.

I am grateful to Dr. Jun Cai and Dr. Sheng Zhong for excellent bioinformatics analysis for ChIP-on-chip data. I am grateful to Dr. Guo-Qing Tong and Dr. Paul Robson for analysis of gene expression in early embryogenesis. I am grateful to Dr. Petra Kraus and Dr. Thomas Lufkin for excellent platform of animal work.

I would like to thank Yun-Shen Chan, Dr. Yuin-Han Loh, Dr. Ching-Aeng Lim, Dr. Bo Feng, Jia-Hui Ng, Jian-Chien Dominic Heng for the excellent collaboration for the projects. I am thankful to Dr. Ping Yuan, Dr. Junli Yan, Fang Fang, Na-Yu Chia, Kuee Theng Kuay, Kia-Ming Lee, Lai-Ping Yaw, Xiangling Ng, and Daoxun Lin for their supports and advices.

I am very thankful to Dr. Petra Kraus, Dr. Max Fun, Dr. Ching-Aeng Lim, Justin Lee Hong Tan, Guofeng Xu, Guoji Guo and Xinyi Lu for critical comments on my thesis.

I also appreciate National University of Singapore and Singapore Millennium Foundation for the scholarships for my PhD.

TABLE OF CONTENTS

SUMMARY

Embryonic stem (ES) cells are unique in their ability to self-renew indefinitely and maintain pluripotency. These properties require transcription factors that form the unique transcriptional regulatory network to specify the gene expression program of ES cells. It has been possible to redirect the highly differentiated state of somatic cells back to a pluripotent state with a combination of four transcription factors: Klf4 is one of the reprogramming factors required, in conjunction with Oct4, Sox2 and c-Myc. Maintenance of self-renewal and pluripotency of ES cells requires Oct4, Sox2 and c-Myc, but Klf4 is dispensable.

In this project, we show that the three Krüppel-like factors: Klf2, Klf4 and Klf5, are required for the self-renewal of ES cells. Individual Klf is dispensable for maintenance of the undifferentiated state of ES cells. However, simultaneous depletion of Klf2, Klf4 and Klf5 by RNA interference (RNAi) leads to ES cell differentiation. Any of the three Klf RNAi-immune cDNAs can rescue the differentiation phenotype, strongly suggesting the functional redundancy among the three Klfs.

The mechanisms of redundancy among Klf2, Klf4 and Klf5 are investigated in ES cells. Chromatin immunoprecipitation coupled to microarray assay (ChIP-on-chip) reveals the binding patterns of the three Klfs were strikingly similar and they also shared a significant portion of common binding loci *in vivo*, indicating Klf2, Klf4 and Klf5 are collaborated to regulate common target genes. The three Klfs target

to enhancer region of *Nanog* where the Klf motif is identified. Enhance reporter assay coupled with motif mutagenesis and Klf RNAi show that the intact Klf binding motif and the three Klfs are required for the enhancer activity in ES cells.

Klf2, Klf4 and Klf5 share many common targets of Nanog by comparing Klfs and Nanog ChIP-on-chip data, suggesting a close functional relationship between these factors. Expression analysis after triple RNAi of the Klfs shows that they regulate key pluripotency genes, such as *Nanog*, *Esrrb*, and *Tcl1*.

Functional redundancy between Klf2, Klf4 and Klf5 is revealed by factor-induced reprogramming assay. Klf2 or Klf5 is able to replace the Klf4 in generating reprogrammed cells together with other reprogramming factors Oct4, Sox2 and c-Myc. Klf2 reprogrammed cells are incorporated extensively in the chimeric embryo, displaying pluripotency.

Our study provides new insight into how the core Klf circuitry integrates into ES cell transcriptional network to specify gene expression program that is required for maintaining the pluripotent state of ES cells.

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS

1. Bo Feng*, **Jianming Jiang***, Petra Kraus, Jia-Hui Ng, Jian-Chien Dominic Heng, Yun-Shen Chan, Lai-Ping Yaw, Weiwei Zhang,Yuin-Han Loh, Jianyong Han, Vinsensius B. Vega,Valere Cacheux-Rataboul, Bing Lim,Thomas Lufkin, Huck-Hui Ng. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. ***Nature Cell Biology***. 2009,11:197-203* *equal first authors*

2. **Jianming Jiang**, Huck-Hui Ng. TGFbeta and SMADs talk to NANOG in human embryonic stem cells. ***Cell Stem Cel****l*. 2008, 3:127-8.

3. Xi Chen, Han Xu, Ping Yuan, Fang Fang, Mikael Huss, Vinsensius B. Vega, Eleanor Wong, Yuriy L. Orlov, Weiwei Zhang, **Jianming Jiang**, Yuin-Han Loh, Hock Chuan Yeo, Zhen Xuan Yeo, Vipin Narang,Kunde Ramamoorthy Govindarajan, Bernard Leong, Atif Shahab, Yijun Ruan, Guillaume Bourque,Wing-Kin Sung, Neil D. Clarke, Chia-Lin Wei, and Huck-Hui Ng. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. ***Cell***. 2008, 133:1106-17.

4. **Jianming Jiang**, Yun-Shen Chan, Yuin-Han Loh, Jun Cai, Guo-Qing Tong, Ching-Aeng Lim, Paul Robson, Sheng Zhong & Huck-Hui Ng. A core Klf circuitry regulates self-renewal of embryonic stem cells. ***Nature Cell Biology.*** 2008, 10: 353 - 360

5. Junli Yan*, **Jianming Jiang*** Ching Aeng Lim, Qiang Wu, Huck-Hui Ng, Keh-chuang Chin. BLIMP1 regulates cell growth through repression of p53 transcription. ***Proceedings of the National Academy of Sciences US***. 2007, 104:1841-6. ** equal first authors*

LIST OF ABBREVIATIONS

aa	amino acid
AP	alkaline phosphatase
APC	adenomatous polyposis coli
bFGF	basic filbroblast growh factor
BIO	6-bromoindirubin-3'-oxime
bioChIP	biotinylation mediated chromatin immunoprecipitation
BMP	bone morphogenetic protein
ChIP	chromatin immunoprecipitation
ChIP-on-chip	chromatin immunoprecipitation coupled with microarray
CNS	central nervous system
CR	conserved region
DA	dopamine
DE	distal enhancer
Dppa4	developmental pluripotency associated 4
Dnmt	DNA methyltransferase
EB	embryoid body
EC cells	embryonic carcinoma cells
EG cells	embryonic germ cells
EpiSCs	epiblast stem cells
ES cells	embryonic stem cells
EMSAs	electrophoretic mobility shift assays
Eras	ES cell-expressed Ras
Esrrb	estrogen-related receptor beta
FAB-SCs	bFGF, Activin, and BIO-derived stem cells
Fbxo15	F-box protein 15
Fgf4	fibroblast growth factor 4
Fox	fokhead box protein
GCNF	germ cell nuclear factor
gp130	glycoprotein 130
GSK	glycogen synthesis kinase
GST	glutathione S-transferase
ICM	inner cell mass
Id	inhibitor of differentiation
iPS	induced pluripotent stem
Klf	krüppel-like factor
LIF	leukaemia inhibitor factor
LRH	liver receptor homolog
MAPK	mitogen-activated protein kinase
MEFs	mouse embryonic fibroblasts
MTL	multiple transcription factor binding loci
mTOR	mammalian target of rapamycin

Nr0b1	nuclear receptor subfamily 0, group B, member 1
Oct4	octamer-binding transcription factor 4
PcG	polycomb-group
PE	proximal promoter
PGCs	primordial germ cells
PI3Ks	phosphoinositide 3-kinases
PICh	proteomics of isolated chromatin segments
POUs	POU-specific domain
POUh	POU homeodomain
RA	retinoic acid
RISC	RNA-induced silencing complex
RTK	receptor tyrosin kinase
SCF	stem cell factor
SCID	severe combined immunodeficiency disease
SCNT	somatic cell nuclear transfer
shRNA	short hairpin RNA
Sox2	SRY (sex determining region Y)-box 2
Tbx	T-box protein
Tcf	T-cell factor/lymphoid enhancer factor
TE	trophectoderm
TGF	transforming growth factor
TS cells	trophoblast stem cells
trxG	trithorax-group
Xist	inactive X specific transcripts
Zfp42	zinc finger protein 42
Zfp143	zinc finger protein 143
Zic3	zinc finger protein of the cerebellum 3

CHAPTER I

INTRODUCTION

CHAPTER I. INTRODUCTION

Stem cells can be identified and derived from both embryonic and adult tissues in most multi-cellular organisms where they are important for normal embryo development and cell regeneration. Two unique characteristics distinguish stem cells from other types of cells. First, stem cells are unspecialized cells that renew themselves for long periods by self-replication. Second, under certain conditions, they can be induced to become specialized cells with specific functions such as the dopamine (DA)-producing neurons or the beating heart cells. Thus they are thought to be holding great promise for regenerative medicine.

Stem cells in mammals are mainly classified into four groups according to their differentiation potentials: totipotent stem cells, pluripotent stem cells, multipotent stem cells, and unipotent stem cells. A totipotent cell can develop into a whole organism. Pluripotent cells can give rise to almost all cell types of a mammal, but can not form the trophoblast that is necessary for the embryo to grow in uterus. A multipotent cell can differentiate into many, but not all of specialized cell types. Only one type of cells can be derived from unipotent stem cells.

Mouse embryonic stem (ES) cells are pluripotent cells with various applications. For example, they are used as *in vitro* models to study mammalian developmental processes and cell therapy in regenerative medicine. They are widely used for making transgenic, knockout or knockin mice. Mouse ES cells also serve as a good model to study the

mechanisms of self-renewal and pluripotency maintenance of stem cells. Moreover, huge amounts of work go into the exploration of their therapeutic potential.

Mouse ES cells derived from inner cell mass (ICM) of blastocyst stage embryos can be cultured in dishes by using various combinations of feeder cells, conditioned medium, cytokines, growth factors, hormones and fetal calf serum. The external stimuli which play important roles in maintaining the ES cells in an undifferentiated state have been explored, such as LIF, BMP and Wnt signaling pathways. Mouse ES cell maintenance also require intact intrinsic transcriptional regulatory network to specify the gene expression program in ES cells. The core intrinsic transcription regulatory network in mouse and human ES cells has been revealed by genomic-wide location analysis of the key transcription factors: Oct4, Sox2 and Nanog. Recent studies show that external signaling pathways can be integrated with the core transcriptional regulatory network in mouse ES cells and these cells have an innate programme for self-renewal that does not require extrinsic instructions which highlight the critical roles of the intrinsic transcription regulatory network for ES cell self-renewal and pluripotency. However, in despite of the progress in recent research, the detailed mechanisms of how ES cells maintain their unique properties are still vague.

Reprogramming is a process of inducing highly differentiated cells to reverse their developmental potential to the pluripotent state. Somatic differentiated cells can be reprogrammed by somatic cell nuclear transfer (SCNT) or fusion with pluripotent stem cells. However, these methods are not applicable due to low efficiency, ethical problems

in obtaining human embryonic tissues and defects of fusion cells. These successful reprogramming approaches demonstrate that the epigenetic state of terminally differentiated cell is not permanently fixed but can revert into an embryonic state that is capable of directing development of a new organism. It is proposed that there are some key factors which are able to redirect the cell fate in the oocyte or embryonic stem cells. Recently, a breakthrough approach shows that mouse and human somatic cells can be reprogrammed into a pluripotent state through retroviral introduction of four transcription factors: Oct4, Sox2, Klf4 and c-Myc. This new reprogramming approach sheds light on the clinical application of stem cells which has been impeded by lack of efficient and straightforward ways to obtain the patient-specific stem cells.

1.1 Pluripotent stem cells

Pluripotent stem cells can be derived from teratocarcinomas, preimplantation embryos, postimplantation embryos and germ cells. They are unique in their ability to undergo prolonged symmetrical self-renewal and maintain capacity in differentiation. In addition, pluripotent stem cells can also be artificially obtained through reprogramming approaches, such as SCNT and cell fusion. These pluripotent stem cells share similar properties, despite the fact that they are isolated and derived from different sources and by different methods.

The stringent definitions of pluripotent stem cells are characterized by several factors: indefinite self-renewal; capacity for differentiation to specialized cells of three germ layers; extensive contribution to primary chimeras; and germline transmission. Most stringently, pluripotency can be measured using tetraploid embryo complementation: a standard achieved by only a limited subset of mouse ES cells[1].

1.1.1 Mouse embryonic carcinoma cells

Mouse embryonic carcinoma (EC) cells are initially derived from the spontaneous malignant teratomas (referred to as teratocarcinomas) which contain a mixture of differentiated cells and undifferentiated cells in mouse testes[2]. Follow-up studies show that teratocarcinomas can be obtained from transplantation of pregastrulation stage embryos[3-5]. The origin of teratocarcinoma-forming cells within pregastrulation stage embryos is shown to be the epiblast[6].

Mouse EC cells are able to proliferate indefinitely *in vitro* and form teratocarcinomas in nude mice. When introduced into mouse embryos, they sometimes incorporate into embryos and contribute to different tissues, displaying pluripotency[7, 8]. Mouse EC cells are almost always aneupoid which prevents germline transmission[9].

The pilot work in mouse EC cells contributes to the derivation of the mouse ES cells. The supporting layer of fibroblast and the conditioned medium were found to favor the maintenance of the EC cell lines[10].

1.1.2 Mouse embryonic stem cells

In 1981, two groups independently established ES cell lines directly from the inner cell mass (ICM) of mouse blastocysts[11, 12]. Mouse ES cells have also been derived from cleavage stage embryos and even from individual blastomeres of two- to eight-cell stage embryos[13, 14]. When reintroduced into mouse blastocysts, these cells were found to contribute extensively into different tissues and give rise to the chimeras. More importantly, these cells contribute to the germ lineage and transmit into the next generation.

Mouse ES cells are unique in their ability to self-renew indefinitely and maintain the pluripotency. They can be induced to differentiate into diverse types of specialized cells[15]; even germ cells by embryoid body (EB)-mediated differentiation or adhesion differentiation in culture dishes[16-18]. Mouse ES cells have become a good model and a

generic resource for the study of cell differentiation processes. Mouse ES cells also serve as a promising and powerful tool for regenerative medicine.

Mouse ES cell genomes can easily be modified by homologous recombination or insertion. The generation of knockout, knockin and transgenic mice utilizes cultured ES cells to target genes *in vitro*. These genetic modified mice are invaluable animal models for studying functions of genes or regulatory elements in genetic diseases and developmental processes[19].

1.1.3 Mouse embryonic germ cells and pluripotent spermatogonial stem cells

The germ cell lineage is a specific cell population that undergoes a series of differentiation steps before giving rise to either eggs or sperms. Primordial germ cells (PGCs) are the embryonic precursors of the germ cell lineage, which are restricted to form only sperms or eggs following their specification from pluripotent epiblast cells. PGCs can be induced to dedifferentiate or be reprogrammed into pluripotent embryonic germ (EG) cells *in vitro* when exposed to exogenous signaling molecules: LIF, bFGF and SCF. Mouse EG cell lines, derived from PGCs immediately before the onset of migration at 8.0 and 8.5 days post coitum (dpc), from germ cells migrating in the hind gut endoderm at 9.5 dpc, or after entry into the genital ridges at 11.5 and 12.5 dpc, have many properties of ES cells[20, 21].

Similar to ES cells, EG cells can be maintained at a self-renewal state in culture. They are capable to give rise to chimeras and transmit into the germline[22]. However, EG cells

retain the erased imprinting pattern acquired during the process of germ cell development, which account for the skeletal abnormalities seen in some chimeras from EG cells derived from germ cells at 11.5 and 12.5 dpc[23, 24]. The embryos, derived by transferring the nuclei of 15 dpc male germ cells into enucleated mouse eggs, survive to 9.5 dpc and show gross growth retardation. The imprinted genes examined exhibit hypomethylation at all sites [25].

Recently, pluripotent spermatogonial stem cells or adult germline stem cells have been derived from spermatogonial cells of adult mouse and human testis. These cells are able to differentiate into various types of specialized cells of all three germ layers when grown under differentiation conditions. When injected into an early blastocyst, mouse pluripotent spermatogonial stem cells contribute to the development of various organs and show germline transmission[26, 27].

1.1.4 Mouse EpiSCs and FAB stem cells

Pluripotent stem cells can be derived from the late epiblast layer of post-implantation mouse and rat embryos by using chemically defined, activin-containing culture medium that is sufficient for long-term maintenance of human ES cells. These cells are referred to as EpiSCs (post-implantation epiblast-derived stem cells). The EpiSC lines are distinct from mouse ES cells in their epigenetic state and the signals controlling their differentiation. Furthermore, EpiSCs and human ES cells share patterns of gene expression and signaling responses that normally function in the epiblast. However,

EpiSCs fail to contribute to chimera formation when injected into recipient blastocysts[28, 29].

A novel stem cell line is derived from murine blastocyst embryos in theculture conditions previously applied to derivation of EpiSCs from epiblast stage embryos. These cells are termed as FAB-SCs for bFGF, Activin, and BIO-derived stem cells. FAB-SCs express common molecular markers but fail to pass hallmark tests of pluripotent differentiation. However, FAB-SCs can be stimulated with LIF and BMP4 and induced into pluripotent cells[30].

1.1.5 Pluripotent stem cells obtained from other species

Development of mouse ES cell lines has great impact on research fields, thus significant efforts have been invested in the derivation of human pluripotent stem cell lines. Three counterpart pluripotent cell types have been established from human tissue: human EC cells, human EG cells, and human ES cells[31-34].

Human EC cell lines are derived from the undifferentiated stem cell component of germ cell tumors. The EC lines can grow continuously in culture and also produce specialized cells of all three germ layers *in vitro* or different tissues through teratoma formation. The EC cells are usually aneuploid and carrying an abnormal karyotype[35], thus prevent them from medical applications. Human EG lines are derived from primordial germ cells in the genital ridges of the developing embryos, typically at 5 to 11 weeks after fertilization in humans, and are also shown to be pluripotent[33].

Human ES cells are derived from inner cell mass (ICM) of human blastocysts. They can be maintained in culture under special conditions, and form derivatives of all three germ layers. However, human ES cells are different from mouse ES cells in many aspects. Human ES cells require the supplement of basic fibroblast growth factor (bFGF) and transforming growth factor b (TGFb)/activin[36]. Mouse ES cells remain undifferentiated in the presence of leukemia inhibitory factor (LIF) and bone morphogenetic protein (BMP)[37]. The colony morphology of human ES cells is different from that of mouse ES cells. The former cells form flat, sheet-like colonies, while the latter cells form round and island-like ones. Human ES cells grow slower than mouse ES cells and are able to be induced into trophectoderm linage cells while mouse ES cells normally can not[38, 39]. Human ES cells share some, but have distinct gene expression profiles and regulatory networks compared to mouse ES cells[40].

Rat models have phenotypic characteristics that are relevant to particular human conditions[41]. However, traditional methods fail to derive authentic rat ES cells with stringent characteristics of pluripotent stem cells[42, 43]. Recently, Rat ES cells have been efficiently derived in the presence of chemical compounds that specifically inhibit GSK3, MEK, and FGF receptor tyrosine kinases. These rat ES cells fulfill all the stringent criteria for the definition of pluripotent stem cells[44, 45]. Most importantly, germline transmittable rat ES cells will be a useful tool to manipulate the rat genome.

1.1.6 Pluripotent stem cells derived by reprogramming approaches

Pluripotent stem cells can also be artificially obtained from reprogramming approaches, such as SCNT, cell fusion or introduction of reprogramming factors. It has long been believed that developmental processes and differentiation are in one direction and irreversible for animals. This view is debated by studies which show that the enucleated frog eggs are able to develop into normal swimming tadpoles and frogs when transplanted nuclei from differentiated frog cells [46-48]. These results conclude that the process of cell differentiation can be fully reversed. The next major breakthrough in this field came with the production of an adult sheep (Dolly) by transferring the nuclei of cultured mammary gland cells derived from an adult sheep to enucleated sheep eggs[49]. Later on, SCNT or nuclear reprogramming has been successfully achieved in different mammalian species, suggesting that this procedure might work with humans[50]. SCNT or nuclear reprogramming has been shown to produce mouse or rhesus macaque blastocysts from fully differentiated cells and successfully isolate ES cell lines from them[50, 51].

Pluripotent ES cells such as EC cells, ES cells, and EG cells have a unique property of being able to carry out nuclear reprogramming of somatic nuclei, as shown after cell fusion[52-54]. The resultant hybrid cells are tetraploid and exhibit characteristics of normal pluripotent cells, including apparent immortality in culture and colony morphology. The pluripotency of the hybrid cells fused with mouse ES cells and somatic cells is demonstrated *in vivo*, since they contribute to all three primary germ layers of chimeric embryos[55].

A surprising advance in this field comes when it is discovered that viral transfection of four genes (Oct4, Sox2, c-Myc, and Klf4) into mouse embryonic or adult fibroblast cells lead to the appearance of some cells displaying ES cell characteristics[56]. These cells are referred to as induced pluripotent stem (iPS) cells and have been tested by all the pluripotency criteria[57, 58], demonstrating iPS cells are authentic pluripotent stem cells.

1.2 Extrinsic factors required for mouse ES cells

Mouse ES and MEF cells are used to elucidate the mechanisms of ES cell self-renewal, pluripotency and somatic cell reprogramming in this study. Mouse ES cells have many advantages over other types of stem cells as they are the first established authentic pluripotent cells which have been widely used as models both *in vitro* and *in vivo*. The pluripotency of reprogrammed cells can be validated by the different kinds of assays established for mouse ES cells.

The unique mouse ES cell properties are routinely maintained by extra-cellular signals and intra-cellular regulators. Through binding to cell-membrane receptors, extra-cellular factors can induce nucleus-directed signaling pathways to control genes expression. Two main important signal factors maintain the normal growth of ES cell in the absence of serum. One is leukaemia inhibitor factor (LIF), and the other is bone morphogenetic protein (BMP).

1.2.1 LIF-Stat3 signaling pathway

Experience from culturing mouse EC cells indicates that feeder cells or conditional medium can favor the maintenance of the undifferentiated cell state. Isolated ES cells can

be cocultured with mouse MEF feeders in the presence of serum. The conditioned medium can support the normal growth of ES cells in the absence of feeders indicating that fibroblasts are producing signals inhibiting ES cell differentiation. The active component, LIF is subsequently identified from the conditioned medium[59, 60].

It further has been demonstrated that LIF binds to its receptor (LIFR) which has a cytoplasmic tail to activate the signaling pathway. As next step in the signaling cascade, the LIF-LIFR complex recruits signal transducer glycoprotein 130 (gp130) to form a trimetric complex which binds and activates tyrosine kinase Jak. Subsequently, Stat3 is recruited to the receptor complex and phosphorylated by Jak. The activated Stat3 will homodimerize and then translocate into nucleus to regulate its downstream target genes[61].

The intact LIF-Stat3 signaling pathway is required for ES cell growth under normal culturing conditions. LIF-deficient fibroblasts are reported to be incapable of supporting ES cell self-renewal[62]. Recruitment and activation of Stat3 is essential for self-renewal of ES cells and expression of an inhibitory Stat3 mutant in ES cells forces differentiation[61, 63, 64]. ES cells heterozygous for the *Stat3* can only be isolated from E14 cells, a cell line least dependent on LIF for self-renewal[65].

Although the LIF-Stat3 signaling pathways are critical for maintenance of ES cells, *in vivo* knockout studies indicate that most components of the pathways are not important at the preimplantation stage when mouse ES cells are derived. Female mice lacking functional *LIF* gene are fertile, but their blastocysts fail to implant and do not develop,

indicating that expression of *LIF* in mice is essential for implantation[62]. Disruption of *Stat3* causes early, postimplantation embryonic lethality between E6.5 and 7.5[66]. Targeted mutation of *LIFR* causes perinatal death and embryos homozygous for the *gp130* mutation progressively die between 12.5 dpc and term[67-69]. *Jak1*-/- mice are runted at birth, fail to nurse, and die perinatally. *Jak2-/-* embryos are anemic and die around 12.5 dpc. *Jak3* deficient mice have profound reductions in thymocytes and severe B cell and T cell lymphopenia similar to severe combined immunodeficiency disease (SCID)[70-74]. These results demonstrate that signaling requirements for embryo development *in vivo* and ES cell growth *in vitro* are different.

1.2.2 BMP-Smad signaling pathway

In serum-free medium, LIF alone is insufficient to prevent mouse ES cell differentiation. Another growth factor BMP4 alone can drive ES cells into non-neural differentiation. However, mouse ES cells can be sustained in combination with LIF and BMP in the absence of serum[37].

BMPs belong to the transforming growth factor beta (TGF-β) family which controls a plethora of cellular responses in animal development. In general, a TGF-β ligand acts through heterodimers of type I and type II receptor serine/threonine kinases on the cell surface. Receptor II will phosphorylate the receptor I kinase domain, which then propagates the signal through phosphorylation of the Smad proteins[75].

Alk3 is the only type I BMP receptor detectable in the pluripotent ICM and early epiblast[76]. Although *Alk3*-/- mutants survive to E6.5–7.0, *Alk3* null ES cells can not be

obtained from blastocysts, highlighting the importance of BMP signaling through Alk3 in maintaining pluripotency[77].

Ying *et al.* showed that BMPs induce the expression of *Id* (inhibitor of differentiation) through the Smad pathway. Id, an inhibitor of basic helix-loop-helix transcription factors, can replace BMPs in promoting mouse ES cell proliferation in the presence of LIF alone[37]. Qi *et al.* suggested that BMPs might also act through inhibition of the mitogen-activated protein kinase (MAPK) pathways independent of Smads[77]. These findings suggest that the self-renewal of mouse ES cells is achieved by a delicate balance between the two cytokines, LIF and BMP.

1.2.3 Wnt signaling pathway

For cells without Wnt signal, β-catenin is degraded through interactions with a complex containing Axin, Adenomatous Polyposis Coli (APC), and glycogen synthase kinase-3β (Gsk-3). The canonical Wnt pathway is initiated upon Wnt protein binding to the Frizzled/LRP receptor complex on the cell surface. Dishevelled, a key componet of the Wnt receptor complex inhibits a complex including Axin, Gsk-3 and APC. Inactivation of Gsk-3 results in stabilization and accumulation of β-catenin in the cytoplasm and nucleus. β-catenin in collaboration of T cell specific factors (Tcfs) regulates Wnt downstream targets. Wnt signaling can also be activated by direct, intracellular inhibition of Gsk-3 function using specific inhibitors[78].

Wnt signaling components have been shown to regulate self-renewal of ES cells. Activation of Wnt signaling, by genetically eliminating the function of the negative regulator APC, promotes the undifferentiated phenotype of mouse ES cells[79]. Activation of the Wnt pathway by 6-bromoindirubin-3'-oxime (BIO), a specific pharmacological inhibitor of Gsk-3, maintains the undifferentiated phenotype in mouse or human ES cells and sustains expression of ES-specific transcription factors *Oct-4*, *Rex-1* and *Nanog*[80]. Mouse ES cells deleted of GSK-3α/β display hyperactivated Wnt/β-catenin signaling and are severely compromised in their ability to differentiate[81]. Tcf3, a Wnt controlled transcription factor, represses *Nanog* and *Oct4* which are required for mouse ES cell self-renewal and pluripotency[82, 83].

1.2.4 Ras signaling pathway

Ras-MAPK

Growth factors such as EGF, FGF will initiate the signaling pathway through binding to receptor tyrosin kinases (RTKs) resulting in receptor dimerization, autophosphorylation and activation. This will cause the activation of a number of intracellular signaling cascades, including the Ras-MAPK pathway. Ras proteins belong to a superfamily of low molecular weight GTP-binding proteins whose regulation depends on their binding to GTP or GDP. Activated Ras leads to activation of a MAP kinase signaling cascade, including MAP kinase, MAP kinase kinase (MKK), and MAP kinase kinase kinase (MKKK).

Much evidence has indicated that the Ras-Mek-Erk pathway plays important roles in mouse ES proliferation and differentiation. Ectopic expression of Ras is able to drive ES cell differentiation. Conditional activation of a mutant form of Ras in ES cells *in vitro* enables derivation of trophoblast stem (TS) cell lines which contribute to chimerize placental tissues *in vivo*[84]. Mek inhibitor is shown to facilitate the derivation of the ES cell lines from mouse embryos, especially those of recalcitrant mouse strains[85]. *Fgf4* null embryos developmentally arrest shortly after implantation[86]. Deletion of *Fgf4*, a main stimulus of Ras-Mek-Erk pathway in ES cells, leads to compromised differentiation capacity into different cell lineages. *Erk2*-null ES cells fail to undergo either neural or mesodermal differentiation in adherent culture and retain expression of pluripotency markers[87]. These results suggest Fgf4 and Ras-Erk signaling cascade play a critical role in the decision between self-renewal and differentiation.

Recently, Ying *et al.* showed that the derivation, propagation and pluripotency of ES cells is independent of extrinsic stimuli. Suppressing activities of mitogen-activated protein kinase (MAPK) and glycogen synthase kinase 3 (GSK) through small molecular inhibitors will enable ES cell self-renewal without LIF and BMP4 signaling[88]. *Stat3* null ES cells and recalcitrant-derived ES lines, which can not be isolated and maintained in traditional methods, are obtained from the defined medium bypassing the cytokine signaling[88]. More importantly, rat ES cells, another rodent pluripotent cells, have been successfully derived from rat preimplantation embryos using the same conditions. The rat ES cells have been shown to be able to contribute comprehensively to primary chimeras and transmit into the germline[44, 45]. These results suggest that “ES cells are a basal cell

state that is intrinsically self-maintaining if shielded effectively from inductive differentiation stimuli"[88].

Eras-PI3K pathway

Takahashi *et al.* reported that mouse ES cells specifically express ES cell-expressed Ras (*Eras*) which interacts with phospphatidylinositol-3-OH kinase (PI3K) but does not with Raf. Eras null ES cells are pluripotent but show significantly retarded growth and less tumorigenicity[89].

Phosphoinositide 3-kinases (PI3-kinases or PI3Ks) are a family of related enzymes that are capable of phosphorylating the 3 position hydroxyl group of the inositol ring of phosphatidylinositol (PtdIns). PI3Ks consist of p110 catalytic subunit and adaptor subunit complexes and are regulated by growth factors through receptor tyrosine kinases. Activated PI3Ks phosphorylate PtdIns(3,4)P2 and generate PtdIns(3,4,5)P3 as a second messenger. A serine/threonine kinase, Akt, binds to PIP3 through its pleckstrin homology (PH) domain and is translocated to the inner cell membrane, where it is phosphorylated and activated by another serine/threonine kinase PDK1. Activated Akt regulates the function of numerous substrates, such as Mdm2, IKK, and mammalian target of rapamycin (mTOR), and plays import roles in various cellular responses[90]. As a strong candidate for the downstream targets of the ERas/PI3K signaling in ES cells, mTOR is shown to be essential for both ES cell proliferation and mouse development[91].

1.3 Intrinsic factors required for mouse ES cells

The external signaling pathways such as LIF and BMP4 finally lead to activation of the transcription factors Stat3 and Smad1 which will translocate into the nucleus and result in the transcriptional regulation of their downstream targets. Apart from the signaling pathways related transcription factors such as Stat3 and Smad1, another set of transcription factors that are important for ES cell growth have been discovered, such as Oct4, Sox2 and Nanog.

1.3.1 Oct4

Oct4 (also known as Oct3) encoded by *Pou5f1* belongs to POU (**P**it-1, **O**ct-1 and Oct-2, **u**nc-86) transcription factor family which shares a conserved POU domain. The POU domain is a bipartite DNA-binding structure[92]. The POU domain of Oct4 has two subdomains, one is the N-terminal specific region (POUs), and the other is C-terminal homeodomain (POUh). Both subdomains are tethered by a helix-turn-helix structure. Both POU-specific and POU-homeo domains are reported to be essential for binding activity to octamer sequence ATGCAAAT. Regions flanking the POU domain are shown to be not critical for DNA binding and show little sequence conservation. Both N-terminal and C-terminal regions of Oct4 have transactivation functions, and C-terminal domain is cell type specific and its specificity is mediated by the Oct4 POU domain[93-95].

Oct4 is initially screened out by analyzing various adult organs and different developmental stages of mouse embryos for the presence of octamer-binding proteins. Oct4 is found in unfertilized oocytes, epiblast, primordial germ cells *in vivo* and pluripotent ES cells, EC cells and EG cells *in vitro*[96, 97].

Oct4 is maternally expressed in unfertilized eggs. The initial expression begins during cleavage stage. At the first cell fate decision stage, Oct4 is downreguated in the trophectoderm and strictly maintained in the ICM. Expression of *Oct4* remains in the epibast of pre- and post-implantation embryos, then strictly in the migratory primordial germ cells[98].

Oct4 knockout studies indicate that "the activity of *Oct4* is essential for the identity of the pluripotential founder cell population". *Oct4*-deficient embryos develop to form developmentally compromised blastocysts whose ICM cells fail to acquire the potential to differentiate into multiple lineages. These embryos restrict to differentiation and only give rise to the extraembryonic trophoblast cells *in vitro*[99]. Conditional gene targeting of *Oct4* is performed for primordial germ cells and results show that loss of Oct4 function leads to apoptosis of primordial germ cells[100]. *Oct4* gene ablation is carried out for many other somatic tissues which results in no abnormalities in homeostasis or regenerative capacity[101]. These results indicate the critical functions of Oct4 in its exclusively expressed cells.

Expression of *Oct4* is associated with an undifferentiated phenotype of ES cells. A 2-fold induction of *Oct4* causes ES cell differentiation into primitive endoderm and mesoderm. On the other hand, repression of *Oct4* induces ES cell differentiation to trophectoderm[102]. The fine-tuning regulation of *Oct4* is required for maintaining the level of its expression in pluripotent cell population.

The first differentiation event result in the separation of the early embryo blastocyst into into two different types: trophectoderm and inner cell mass. Oct4 functions to suppress differentiation of embryonic cells into trophectodem where *Oct4* is silenced. It is reported that a transcription factor Cdx2 is responsible for the suppression of *Oct4* and normal development of trophectodem lineage[103, 104].

To understand the specific expression pattern of *Oct4*, the regulation regions of Oct4 has been revealed by many studies. Two distinct enhancer elements are found in the upstream of the *Oct4* promoter by an *Oct4* transgenic reporter study. One is the distal enhancer (DE) which drives *Oct4* in the morula, ICM and primordial germ cells. The other is the proximal enhancer (PE) which activates *Oct4* in epiblast[105]. The detailed comparison between different species of *Oct4* promoter shows that there are 4 highly conserved regions (CR) 1-4[106] (**Figure 1.1**). Many transcription factors are found to target the conserved regions. Several members of the nuclear receptor family, including LRH-1, Sf-1, Gcnf, regulate *Oct4* expression by binding to its enhancer and promoter regions[107-109]. *Oct4* expression is regulated by Oct4/Sox2 complex by targeting to the CR4 region in ES cells[110, 111]. Sall4 is shown to bind to the highly conserved regulatory region of the *Oct4* distal enhancer and activate *Oct4* expression *in vivo* and *in vitro*[112]. Global mapping studies show that Nanog, Smad1 and Stat3 might regulate *Oct4* by binding to conserved regions[40, 113]. Apart from transcription factors, epigenetic regulators are shown to regulate the *Oct4* expression. Epigenetic silencing of *Oct4* at the promoter region is initiated by a pronounced increase in H3K9 methylation mediated by G9a and followed by *de novo*

DNA methylation by the enzymes Dnmt3a/3b[114]. Although many transcription factors and epigenetic modifier are found in regulation of *Oct4*, the detailed mechanisms on how they are precisely controlled *in vivo* and reactivated in reprogramming still need to be addressed.

As a transcription factor, Oct4 also regulates downstream genes that are important for self-renewal and pluripotency. Cdx2, a critical transcription factor required for the development of a functional trophectoderm and TS cell self-renewal, is repressed by Oct4 in pluripotential founder cell *in vivo* and ES cells *in vitro*[103, 104]. Fgf4, a growth factor controlled by Oct4, is essential for normal trophobast development, TS cell self-renewal, and lineage specification commitment[87, 99]. These results provide some detailed roles of Oct4 in the first cell specification event. Oct4 also targets to *Nanog*, *Sox2*, *Fbx015*, *Rex1* including itself to establish the pluripotent cell identity[110, 115, 116].

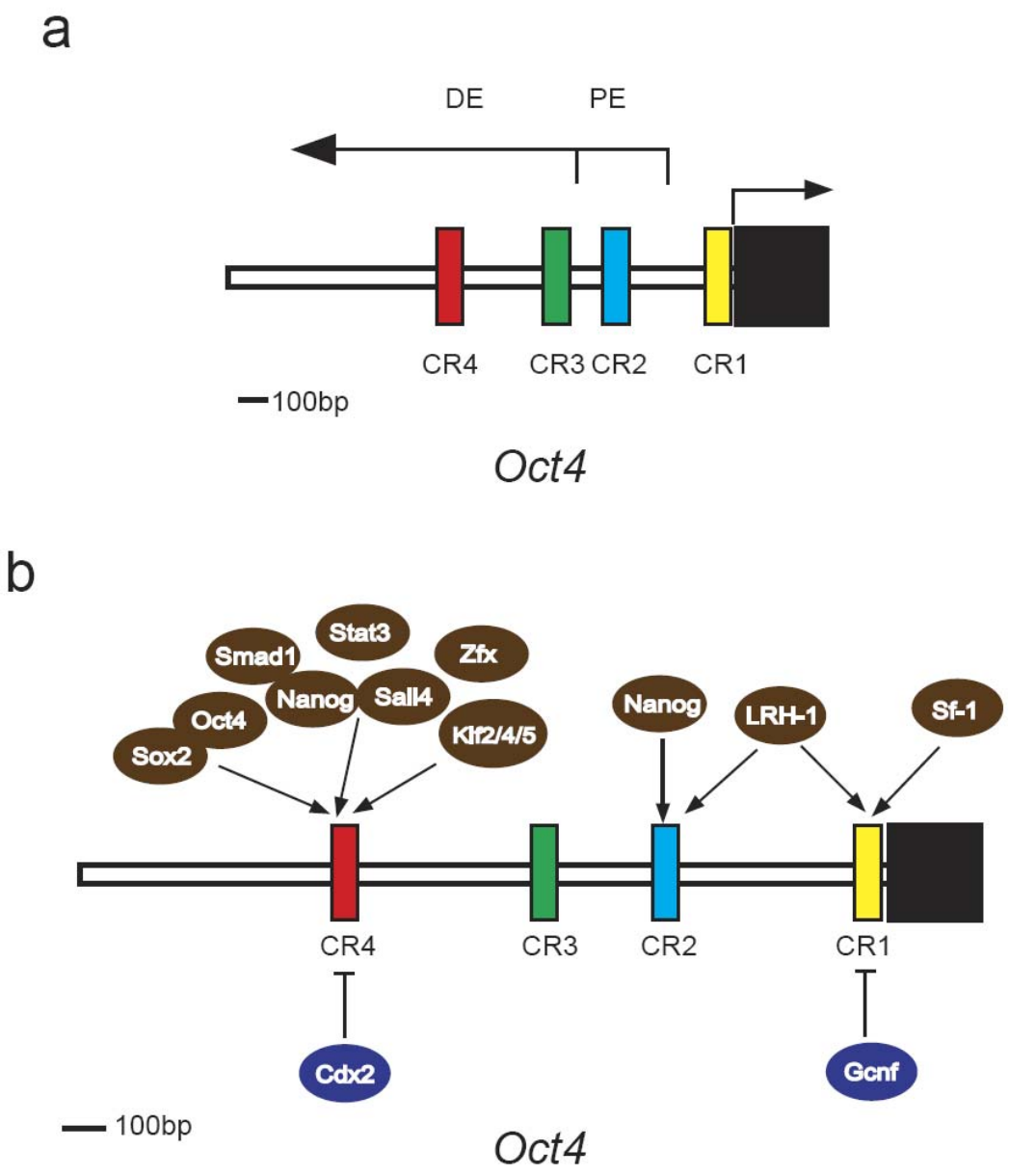


Figure 1.1 A schematic diagram illustrating the regulatory elements and transcriptional regulators of *Oct4*

(a) *Oct4* promoter contains important regulatory regions. Distal enhancer (DE) and proximal enhancer (PD) are two enhancer regions. Four highly conserved regions (CR1-4) are shown as colored boxes: yellow, blue, green and red.

(b) Regulatory regions of *Oct4* receive multiple inputs. Oct4, Sox2, Smad1, Stat3, Nanog, Sall4, Cdx2 bind to CR4 region. CR2 region is targeted by Nanog and LRH-1. LRH-1, Sf-1 and Gcnf occupy CR1 region.

1.3.2 Sox2

Sox2 (SRY (sex determining region Y)-box 2) belongs to the Sox transcription factor family which harbors a HMG specific DNA binding domain. Mouse *Sox2* is expressed in the pluripotent cells, in the extraembryonic ectoderm, and in germ cells. After gastrulation, *Sox2* is expressed in the neural tube from the earliest stages of its development (neural plate)[117, 118]. *Sox2* is expressed highly in the neuroepithelium of the developing central nervous system (CNS) and in adult neural stem cells[119]. These results indicate possible functions of Sox2 in pluripotent lineage, in neural development and in homeostasis of the adult CNS.

Avilion *et al.* reported that *Sox2* null blastcysts contain a relative normal ICM, but fail to maintain an epiblast and are unable to give rise to the *Sox2* null ES cell lines[120]. The expression pattern of *Oct4* and *Sox2* are overlapped in pri- and post-implantation epiblast, germ cells and ES cells. Both of them are essential for derivation of ES cell lines from the blastocysts. These results suggest that Sox2 and Oct4 function together to maintain the pluripotency.

Sox2 is shown to bind to heptamer CATTGTA and form a ternary complex with Oct-4 protein on *Fgf-4* enhancer[121, 122]. Numerous Sox2-Oct4 target genes are found to have Oct4 octamer and sox2 heptamer separated by either 0 or 3 bp. Crystal structure analysis and modeling reveal that Oc4 and Sox2 are able to dimerize onto different enhancers in distinct conformation arrangement[123]. Interestingly, Sox2 and Oct4 are shown to bind to

Sox-Oct elements of *Oct4*, *Sox* and *Nanog* as a binary complex and regulate their expression in mouse and human ES cells[110, 115].

The functions and regulatory mechanisms of Sox2 in ES cells are revealed by using inducible Sox2-null mouse ES cells. Suppression of Sox2 in ES cells leads to differentiation into trophectoderm-like cells, as *Oct4* null cells or suppression of *Oct4* [99, 103, 124]. However, the expression of Sox-Oct4 enhancer dependent genes including *Oct4*, *Sox2*, and *Nanog* is not reduced dramatically at an early time course, suggesting that Sox2 is dispensable for the activation of these Sox-Oct enhancers. Oct4 is shown to restore the loss of pluripotency in Sox2-null ES cells, indicating that the essential function of Sox2 is to stabilize ES cells in a pluripotent state by regulation of *Oct4* expression [124].

1.3.3 Nanog

Nanog is a homeodomain-containing transcription factor. Two independent groups discovered the important roles of Nanog for ES cell-renewal[125, 126]. One isolated Nanog through a functional screening for molecules that are capable of directing ES cell self-renewal in the absence of LIF[125]. The other group applied digital differential display to compare EST libraries from mouse ES cells and those from various somatic tissues[126]. Both demonstrate that ectopic expression of *Nanog* relieves ES cell self-renewal from dependence on LIF signaling.

Nanog protein contains 3 domains, N-terminal domain (ND), homodomain (HD) and C-terminal domain (CD) (**Figure 1.2a**). The homodomain is responsible for the specific DNA recognition. ND and CD are shown to have transactivation activities[127]. Nanog CD is separated by a WR domain into CD1 and CD2 domain. The WR domain in Nanog is 10 pentapeptide repeats starting with a tryptophan (W) residue. The WR domain is identified to be in charge of the Nanog dimerization. Nanog dimerization is required for Nanog tansactivation and interaction with other pluripotency network proteins[128, 129].

Nanog is initially expressed after the 32-cell stage of early embryos, and is later restricted in pre- and post-implantation epiblast, and in primordial germ cells[126, 130]. Mitsui *et al*. showed that *Nanog* ablation i*n vivo* causes a failure in the specification of early embryo pluripotent cells, which adopt a differentiated visceral/parietal endodermal fate. *Nanog* is highly expressed in undifferentiated ES cells. Disruption of *Nanog* in ES cell specifies cell fate exclusively into extraembryonic endoderm lineages[126]. However, Chambers *et al*. observed that Nanog protein and its expression fluctuates in mouse ES cells. Although *Nanog* null ES cells are prone to differentiation and expand slowly, they are able to self-renew, contribute to different germ layers and be recruited to the germ line. Primordial germ cells fail to mature on reaching the genital ridge without *Nanog*[131]. These results argue against Nanog serving an essential role in conjunction with Oct4 and Sox2 in the transcriptional machinery of pluripotency[40], while indicate that Nanog is required for cell state transitions during germ cell development and for cell state reversions in ES cell cultures.

Independent studies show that depletion of *Nanog* by RNAi impairs ES self-renewal and induces cell differentiation, while overexpression of *Nanog* strengthens ES self-renewal independent of LIF signaling pathway[40, 125, 126, 132]. These results highlight the critical role of Nanog for ES cell maintenance. Expression of *Nanog* is tightly regulated in ES cells (**Figure 1.2b)**. Oct4 and Sox2 bind to the *Nanog* proximal promoter and positively regulate its expression[115]. As an interacting partner of Oct4, Zfp143 regulate *Nanog* expression through modulation of Oct4 binding[133]. Foxd3, a member of the forkhead family of transcriptional regulators, positively regulates *Nanog* expression by modulating the *Nanog* promoter[134]. Nanog also binds strongly to its own distal enhancer region. Sall4, a member of spalt-like protein family, is shown to interact with Nanog and co-occupy many Nanog targets including *Pou5f1*, *Sox2*, *Nanog* and *Sall4*[135]. *Nanog* expression is up-regulated in mouse ES cells by the binding of *T* (Brachyury) and *Stat3* to an enhancer element[136]. Smad1 is also shown to be a partner of Nanog and is enriched at the *Nanog* distal enhancer[113, 136]. Loss of *Tcf3* in ES cells impairs the ability of these cells to differentiation by relieving the repression of *Nanog*[82, 83, 137]. Lin *et al*. reported that the tumor suppressor, p53, binds to the promoter of *Nanog* and suppresses its expression after DNA damage[138].

Post-transcriptional and Post-translational regulation can influence the stability and translation efficiency of mRNAs as well as the activity and interactions of proteins. MicroRNA miR-134, upregulated during ES cell differentiation, is found to specifically target the coding region of *Nanog* mRNA and attenuate its translation[139, 140]. Nanog appeares to be a phosphoprotein[141], however, the role and regulatory mechanisms for

Nanog phosphorylation need to be determined. A recent study uncovered a novel mode of regulating Nanog at the posttranslational level. Caspase-3, a major component of the programmed cell death pathway, is shown to be involved in inducing Nanog cleavage in differentiating ES cells[142]**.** These different regulating mechanisms may allow ES cells to respond rapidly to disassemble the pluripotency regulatory network upon differentiation**.**

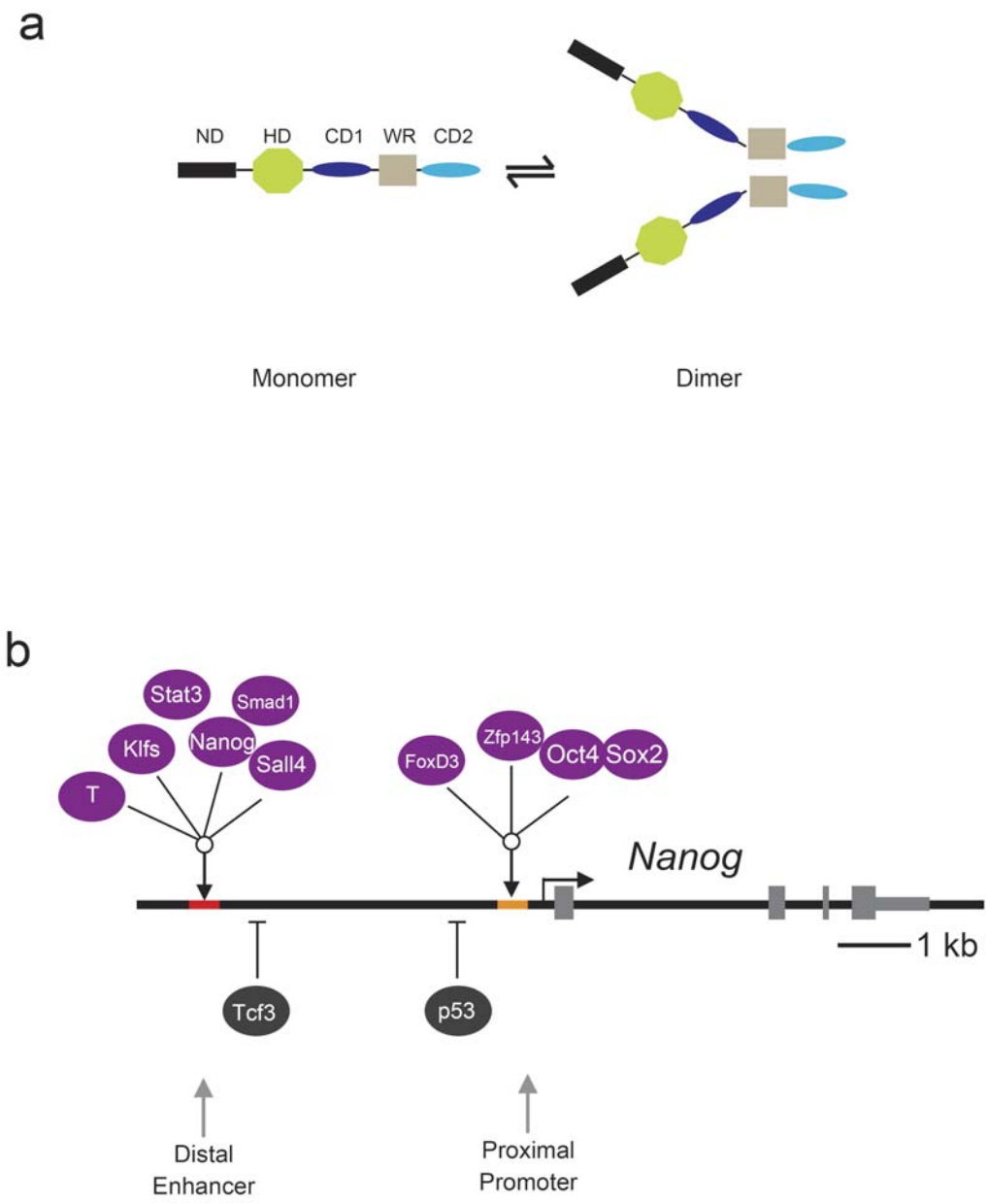


Figure 1.2 A schematic diagram illustrating monomer and dimer of Nanog and regulatory elements and transcriptional regulators of *Nanog*
(a) Nanog domains. Nanog protein contains 3 domains, N-terminal domain (ND), homodomain (HD) and C-terminal domain (CD). CD is separated by a WR domain into CD1 and CD2 domain. Different domains are indicated by colored boxes. Nanog dimerization depends on WR domain.
(b) Regulatory regions of *Nanog* receive multiple inputs. Two key regulatory regions (enhancer and proximal promoter) are shown as red and orange bars. T, Nanog, Sall4, Stat3 and Smad1 bind to *Nanog* enhancer. Oct4, Sox2, Foxd3 and Zfp143 occupy the *Nanog* proximal promoter.

1.3.4 Klf transcription factor

Krüppel-like factors (Klfs) are named after the Drosophila embryonic pattern regulator Krüppel. There are 17 Krüppel-like factor members identified in mammals so far. These Klf members contain conserved DNA binding domains and occupy the similar DNA binding motifs[143]. The Klf transcription factors play diverse roles in proliferation, differentiation, development and cell death. Many Klf members have been shown to have critical roles in mouse development by gene targeting approach (Table 1.1). Klf1 null embryos die of anaemia during fetal liver erythropoiesis and show the molecular and haematological features of beta-globin deficiency, found in beta-thalassaemia[144]. Klf2 -/- embryos die between E12.5 and E14.5 from severe intra-embryonic and intra-amniotic hemorrhaging coupled with specific defects in blood vessel morphology. In addition, single-positive T cell quiescence and survival also require the intact functions of Klf2[145]. Klf3 knockout mice have less white adipose tissue, and their fat pads contain smaller and fewer cells[146]. Klf4-/- mice die shortly after birth due to loss of skin barrier function and demonstrate a 90% decrease in the number of goblet cells in the colon[147, 148]. Klf5 -/- embryos die before E8.5 and Klf5 +/- mice show diminished levels of arterial-wall thickening, angiogenesis, cardiac hypertrophy and interstitial fibrosis in response to injury[149]. Klf6 -/- embryos die by E12.5 and are associated with markedly reduced hematopoiesis and poorly organized yolk sac vascularization[150]. Loss of Klf7 activity in mice leads to neonatal lethality and a complex phenotype which is associated with deficits in neurite outgrowth and axonal misprojection at selected anatomical locations of the nervous system[151]. Klf9 null mice have a normal life span but females show reduced fertility, due to a failure of normal uterine development and embryo implantation[152, 153].

Klf10 -/- mice have increased osteoblast numbers with no increase in bone formation parameters and Klf10 -/- null mouse derived osteoblasts are defective in mineralization and in support of osteoclast differentiation *in vitro*[154]. Klf13 -/- mice have reduced numbers of circulating erythrocytes and a larger spleen[155]. The Klf15 null mice are viable, fertile and have a normal lifespan but are prone to induce cardiac hypertrophy[156]. Klf11 null mice do not show evident pathological defects[157]. There is no knockout data available for the rest of Klf transcription factors (Klf8, Klf12, Klf14, Klf16, and Klf17).

Klf4 is shown to upregulate *p21* and suppress *p53* gene directly in mouse fibroblast cells[158, 159]. MEF cells null for the *Klf4* gene are genetically unstable[160]. These results support a crucial role for Klf4 in regulating cell proliferation and maintaining genetic stability.

Recent reports on the reprogramming of somatic cells to pluripotent cells highlight a critical role of Klf4 in remodeling cell fate[56]. Klf4 is found to be induced by LIF signaling pathway in mouse ES cells, and gain of function assay shows that Klf4 can inhibit EB induced differentiation[161]. Paradoxically, Klf4 alone is not important for the self-renewal state of mouse ES cells although it can directly regulate *Lefty1* gene expression in mouse ES cells[162]. Analysis of published expression profiles of LIF withdrawal or retinoic acid (RA) treatment induced differentiation shows that Klf2, Klf4 and Klf5 are found to be highly expressed in undifferentiated ES cells and downregulated during cell differentiation [132, 163]. However, the functions of Klf4 and other family members for ES cell self-renewal are not well studied.

Table 1.1 The phenotypes of the Klf knockout mice

Allele Symbol Gene; Name	Synonyms	Primary knockout phenotype
Klf1; Krüppel-like factor 1	EKLF	Lethal beta-thalassemia at E14.5
Klf2; Krüppel-like factor 2	LKLF	Lethal cardiac failure at E12.5 to E14.5
Klf3; Krüppel-like factor 3	BKLF, Tef-2	Viable with less white adipose tissue
Klf4; Krüppel-like factor 4	EZF, Gklf, Zie	Peri-natal failure of skin barrier
Klf5; Krüppel-like factor 5	Bteb2, CKLF, IKLF	Early embryonic death before E6.5
Klf6; Krüppel-like factor 6	BCD1, Copeb, CPBP, FM2, FM6, Ierepo1, Ierepo3, Zf9	Lethal failure of erythropoiesis and yolk sac vascularisation at E12.5
Klf7; Krüppel -like factor 7	UKLF	Severe neurological defects
Klf8; Krüppel-like factor 8	BKLF3, ZNF74	-
Klf9; Krüppel-like factor 9	BTEB-1, Bteb1, Klf9	Viable with impaired uterine development and defective embryo implantation
Klf10; Krüppel-like factor 10	EGR[a], Egral, Gdnfif, mGIF, Tieg1	Viable with bone defects, impaired skeletal development
Klf11; Krüppel-like factor 11	Tieg2, Tieg2b, Tieg3	Viable with no phenotype observed
Klf12; Krüppel-like factor 12	AP-2rep	-
Klf13; Krüppel-like factor 13	Bteb3, FKLF-2, NSLP1, RFLAT-1	Viable with defects in T cell differentiation
Klf14; Krüppel-like factor 14	BTEB5	-
Klf15; Krüppel-like factor 15	CKLF, KKLF	Viable with increased susceptibility to cardiac hypertrophy
Klf16; Krüppel-like factor 16	BTEB4, DRRF	-
Klf17; Krüppel-like factor 17	Zfp393	-

- : no available information

1.3.5 Other factors

Besides the transcription factors mentioned above, more factors have been found to be involved in ES cell self-renewal and pluripotency. A RNAi knockdown screening approach demonstrates that depletion of Oct4, Sox2, Nanog, Tbx3, Esrrb, Tcl1, Dppa4 and Mm.34388 can induce ES cell differentiation[132]. Suppression of *Nr0b1* (Dax1) expression, either by knockdown or inducible knockout method, induces ES cell differentiation[164]. Zic3 is reported to play an important role in the maintenance of pluripotency by preventing endodermal lineage specification in ES cells[165]. c-Myc has been implicated in promoting the self-renewal of ES cells under the LIF-Stat3 pathway[166]. Knockout of *Ronin* compromises the growth of ES cells while forced expression of *Ronin* allows ES cells to proliferate without differentiation in the absence of LIF[167].

1.3.6 Epigenetic regulators

Recently, epigenetic regulators have been reported to play important roles in ES cell maintenance and differentiation. DNA methlylation is involved in adding a methyl group to DNA without altering the DNA sequences by DNA methytransferases. DNA methlylation plays important roles in imprinting, X-chromosome inactivation, and suppression of repetitive elements. In mammals, DNA methyltransferase 1(Dnmt1) is important in the maintenance of the the DNA methylation pattern during DNA replication. Dnmt3a and Dnmt3b are the *de novo* methyltransferases that can build up DNA methylation patterns required in normal developmental processes[168, 169]. Although Dnmt1 mutant ES cells are viable and show no obvious abnormalities with respect to growth rate or morphology, Dnmt1 deficiency leads to enhanced microsatellite instability in ES cells[170, 171]. Dnmt1-dificient embyroid bodies aberrantly express *Xist* and *Xist*-linked gene,

and die when induced to differentiation[172]. Later-passage hypomethylated Dnmt (3a-/-, 3b-/-) ES cells are unable to form teratomas *in vivo*[173, 174] and differentiate upon LIF withdrawal due to low levels of DNA methylation[175].

In ES cells, the promoters of a subset of genes are repressed but become induced during differentiation. These regions have been shown to be marked with "bivalent" histone modifications that correlate with both activation (H3K4me3) and inactivation (H3K27me3)[176]. These markers are the results of the activities of Trithorax-group (trxG) and Polycomb-group (PcG) proteins. PcG proteins comprise at least two kinds of complex PRC1and PRC2 which work together to carry out their repressive effect. PcG proteins are critical for normal embryo development as deletion in mice of PRC2 genes, including *Ezh2*, *Eed*, *Suz12*, are shown to be embryonic lethal[177-179]. ES cells can not be derived from blastocysts without the PRC2 component *Ezh2*[177]. *Eed* and *Suz12* null ES cells strongly tend to differentiate or are impaired in proper differentiation[180, 181]. Interestingly, H2AZ, a histone H2A variant, is identified to be enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment[182]. These results highlight the critical functions of DNA methlylation, histone modifications and histone variants in ES cell differentiation processes.

1.3.7 Transcriptional regulatory network

The transcriptional regulatory network comprises a specific group of transcription factors and their co-regulators that collaborate to produce specific expression program for certain cells. To study a complex transcriptional regulatory network, genome-wide location

analysis such as ChIP-on-chip and ChIP-PET, is carried out for transcription factor Oct4, Sox2 and Nanog in mouse and human ES cells. The results show that Oct4, Sox2 and Nanog share a substantial portion of their target genes and auto-regulate each other as well as synergistically regulate downstream target genes, hence forming a tight regulatory network[40, 183].

Apart from the core regulators, more transcription factors have been found to concertedly build up the ES cell regulatory network by genomic-wide location analysis. Kim *et al*. applied *in vivo* biotinylation mediated ChIP (bioChIP) to genome-wide locating (bioChIP-chip) of a set of factors associated with pluripotency of mES cell including somatic cell reprogramming factors (Oct4, Sox2, Klf4, and c-Myc) and others (Nanog, Dax1, Rex1, Zpf281, and Nac1). Their target genes are found to be grouped into two classes: promoters bound by few factors are often inactive or silenced, whereas promoters bound by multiple factors tend to be active in the undifferentiated state[184]. Independently, Chen *et al*. mapped the binding sites of 13 transcription factors (Nanog, Oct4, STAT3, Smad1, Sox2, Zfx, c-Myc, n-Myc, Klf4, Esrrb, Tcfcp2l1, E2f1, and CTCF) and 2 transcription co-regulators (p300 and Suz12) in ES cells by using chromatin immunoprecipitation (ChIP) coupled with ultra-high-throughput DNA sequencing (ChIP-seq). The selective factors are known to play different roles in ES cell biology as components of the LIF and BMP signaling pathways, self-renewal and pluripotency regulators, and key reprogramming factors. Analysis reveals there are highly dense binding loci or multiple transcription factor binding loci (MTL) in the ES cells. The cluster analysis shows MTL can be grouped into two different classes: one is the Oct4-

Sox2-Nang cluster, the other is the Myc cluster. Transcriptional co-activator p300 is specifically recruited to the Oct4-Sox2-Nanog cluster but not to the Myc cluster, depending on Oct4, Sox2 and Nanog. Fragments from a subset of the Oct4-Sox2-Nanog cluster are shown to have ES specific enhancer activities. This data demonstrates that these MTL genomic loci as well as multiple transcriptional regulators might function together as specific enhanceosomes in ES cells. Interestingly, Stat3 and Smad1, the key components of the LIF and BMP signaling pathways respectively, co-occupy many binding targets with the core transcription factors Oct4, Sox2 and Nanog, indicating that the two key signaling pathways are integrated to the Oct4, Sox2, and Nanog transcriptional network through Smad1 and Stat3[113]. This data provides a framework for understanding the transcriptional regulatory networks and suggest a more comprehensive and complex view of the pluripotency network in ES cells than prior work.

Besides mapping of the transcriptional regulatory network by genomic-wide location analysis, analysis among the different networks will provide us more clues for understanding the unique networks in stem cells. For example, there is only a small portion of Oct4 or Nanog target genes overlap between mouse and human ES cells[40]. One possible explanation for the limited overlap between the mouse and human is that there might be different between species. Independent groups applied genome-scale location analysis to identify Oct4 targets in mouse ES cells, EpiSCs and human ES cells. The data confirm the result showing limited overlap of Oct4 targets between mouse ES cells and human ES cells, while showing a 7-fold greater overlap of Oct4 targets between human ES cells and EpiSCs[29]. Although the EpiSCs express core pluripotency genes *Oct4*, *Sox2*

and *Nanog*, they fail to colonize the embryo. Ectopic expression of *Klf4* can reprogram EpiSCs into iPS cells. These EpiSC derived induced pluripotent stem (Epi-iPS) cells produce high-contribution chimeras and show germline transmission[185]. Considering the similar condition used in culturing human and EpiSCs and regulatory network between them, human ES cells are more akin to EpiSCs than mouse ES cells.

Recent progresses indicate that transcriptional regulatory network in ES cells is much more complicated than the current models. First, *Nanog* deficient ES cells can still incorporate into all different lineages including early stage of primordial germ cells. Second, the differentiation phenotype of Sox2 knockout ES cells can be rescued by *Oct4* expression. Third, ES cells have an innate program for self-renewal that does not require extrinsic signal instruction when the differentiation pathways are suppressed. Fourth, more regulators have been found to play an important role in ES maintenance, such as Zic3, Esrrb, Zfp143.

1.3.8 Methods to study the network

Genetic screen by RNAi

RNAi is a valuable research tool which can specifically and robustly down-regulate gene expression. RNAi-mediated pathways are induced by transfection of synthesized small interfering RNAs (siRNAs) or introduction of vectors expressing short hairpin RNAs (shRNAs) into cells. shRNA will be processed by Dicer and then loaded into RNA-induced silencing complex (RISC) to induce sequence specific mRNA degradation or translational suppression[186]. RNAi can also be used for large-scale screenings that

systematically shut down each gene in the cell. For example, Oct4, Sox2, Nanog, Tbx3, Esrrb, Tcl1, Dppa4, Mm.34388 and Tip60-p400 are identified to be important for ES cell self-renewal and pluripotency by RNAi screening approach[132, 187].

ChIP

Chromatin immunoprecipitation (ChIP) is an approach to determine the chromatin context of a genomic region or the location of DNA binding sites on the genome for a particular protein *in vivo*. The development of tools for genome-wide location analysis of *in vivo* protein-DNA interactions, such as ChIP-on-chip, ChIP-PET or ChIP-seq, now allow global views of transcription factor binding and chromatin context. Genome-wide location analysis is able to identify regulatory regions, reveal the chromatin context at regulatory regions, and establish a framework for understanding the transcriptional regulatory networks.

DNA Microarray

Microarray technology allows us to examine many genes at once and determine the dynamics of the gene expression for many samples. For example, microarray-based gene expression profiling can be used to identify genes whose expression is changed in response to specific gene knockdown or overexpression.

Proteomics

Proteomics is the large-scale study of protein structures and functions. Proteomics can be used to determine post-translationally modified proteins and establish protein-protein

interactions in protein complexes. For instance, a protein interaction network in ES cells is mapped by using combination tagging and proteomics. Target proteins such as Nanog, Oct4 and Rex1 are fused with an N-terminal Flag epitope and a short peptide tag that serves as a substrate for *in vivo* biotinylation. Interaction partners have been successfully identified by a combination of single step purification and tandem purification, followed by mass spectrometry[188]. Recently, a new proteomics of isolated chromatin segments (PICh) can address the full set of proteins that regulate specific loci. A specific nucleic acid probe can be used to isolate genomic DNA with its associated proteins in sufficient quantity and purity to allow identification of the bound proteins[189].

1.4 Reprogramming of somatic cells

During animal development, cell differentiation is thought to be an irreversible process and specified cell types are non-switchable between each other. However, researchers found that cell fate is actually reversible and exchangeable in some cases.

A good example of converting the cell fate is to switch a range of non-muscle cell types into muscle cells by ectopic expression of a transcription factor MyoD which is a master regulator for myogenesis[190, 191]. Switching cell fate has also been successfully achieved in the hematopoietic system by overexpressing key transcription factors, such as Gata-1, Gata-3 and C/EBP[192].

The dominant effect of transcription factors in redirect cell differentiation is best exemplified by the reprogramming of somatic cells into pluripotent cells by defined

transcription factors: Oct4, Sox2, Klf4 and c-Myc[56]. Recently, exocrine cells of pancreas can be converted to the endocrine beta cells *in vivo* by introducing three transcription factors Pdx1, Ngn3 and MafA which is required for beta pancreas differentiation[193].

These results indicate that key transcription factors have unique functions to remodel the specific transcriptional regulatory network and rebuild the epigenetic landmark. Most importantly, these exogenous transcription factors are no longer required to maintain the network. Apart from transcription factors, epigenetic modifiers also play important roles in the remodeling processes. Inhibition of DNA methylation and histone modification are able to dramatically increase the reprogramming efficiency or substitute for some transcription factors[194, 195].

1.5 Purpose and scope

As a potent reprogramming factor, the roles of Klf4 in ES cells are obscure. The purpose of this thesis is to elucidate the functions of Klf4 and Klf family members in maintaining the unique property of ES cells by functional genomics. More specifically, the objectives of this thesis are to:

- investigate the biological functions of Klfs in ES cells by RNAi knockdown.
- identify the binding loci of Klfs in ES cells by ChIP-on-chip.
- understand the molecular basis of the redundancy by EMSA and luciferase enhancer reporter assay.
- measure changes of the gene expression controlled by Klfs
- test the reprogramming ability of Klfs

This project aims to decipher the functions of Klfs in mouse ES cells by using advanced functional genomics approaches. Combinatorial information provided in this study will contribute to a better understanding of ES cell biology and reprogramming.

- The functional results obtained should be valuable to the stem cell field as roles of this family are not well studies.
- The results of genomic-wide location analysis will refine the transcriptional regulatory network in the ES cells.
- Study of the binding motifs and their related activities in the binding element will provide the detailed mechanisms of Klfs in gene regulation.

CHAPTER II

MATERIALS AND METHODS

CHAPTER II MATERIALS AND METHODS

2.1 Cell culture and transfection

Feeder-free E14 mouse ES cells were cultured at 37°C with 5% CO_2. All cells were maintained on gelatin-coated dishes in Dulbecco's modified Eagle medium (DMEM; GIBCO), supplemented with 15% heat-inactivated fetal bovine serum (FBS; GIBCO), 0.055 mM β-mercaptoethanol (GIBCO), 2 mM L-glutamine, 0.1 mM MEM nonessential amino acid, 5,000 units/ml penicillin/streptomycin and 1,000 units/ml of LIF (Chemicon) as described previously. Transfection of shRNA and over-expression plasmids was performed using Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions. Briefly, 1.5 μg of plasmids were transfected into ES cells on 60mm plates for RNA and protein extraction. Detection of alkaline phosphatase, which is indicative of the nondifferentiated state of ES cells, was carried out using a commercial ES Cell Characterization Kit from Chemicon. For RNAi-ChIP assay, 12 μg of plasmids were transfected into ES cells on 150mm plates. Puromycin (Sigma) selection was introduced 1 day after transfection at 0.8 μg/ml, and maintained for 2 to 6 days prior to harvesting.

2.2 RNA extraction, reverse transcription and quantitative real-time PCR

Total RNA was extracted using TRIzol Reagent (Invitrogen) and purified with the RNAeasy Mini Kit (Qiagen). Reverse transcription was performed using SuperScript II Kit (Invitrogen). DNA contamination was removed by DNase (Ambion) treatment, and the RNA was further purified by an RNAeasy column (Qiagen). Quantitative PCR analyses were performed in real time using an ABI PRISM 7900 Sequence Detection System and SYBR Green Master Mix as described. For all the primers used, each gave a

single product of the right size. In all our controls lacking reverse transcriptase, no signal was detected.

2.3 ChIP assay

ChIP assays with feeder-free mouse ES cells were carried out as previously described. Briefly, cells were cross-linked with 1% formaldehyde for 10 min at room temperature and formaldehyde was then inactivated by the addition of 125 mM glycine. Chromatin extracts containing DNA fragments with an average size of 500 bp were immunoprecipitated using antibodies. For all ChIP experiments, quantitative PCR analyses were performed in real-time using the ABI PRISM 7900 sequence detection system and SYBR green master mix as previously described. Relative occupancy values were calculated by determining the apparent immunoprecipitation efficiency (ratios of the amount of immunoprecipitated DNA to that of the input sample) and normalized to the level observed.

2.4 ChIP-on-chip assay

We carried out three biological replicate ChIP experiments to obtain ChIP enriched DNA. Amplification of the linker-ligated ChIP DNA using ligation mediated-PCR and labeling of DNA samples for ChIP-chip were followed by NimbleGen protocol. Briefly, each of test and reference DNA samples (1 μg) was denatured in the presence of 5′ Cy3-or Cy5-labeled random nonamers (TriLink Biotechnologies) and incubated with 100U (exo-) Klenow fragment (New England Biolabs) and dNTP mixture [6 mM each in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.4; Invitrogen)] for 2 h at 37°C. Reactions were

terminated by addition of 0.5 M EDTA, pH 8.0, precipitated with isopropanol, and resuspended in water. Then, 13 μg Cy5-labeled ChIP sample and 13 μg Cy3-labeled total sample were mixed, dried down, and resuspended in 40 μL hybridization buffer (NimbleGen Systems). After denaturation, hybridization was carried out in a MAUI® (MicroArray User Interface) Hybridization System (BioMicro Systems, Salt Lake City, UT, USA) for 18 h at 42°C. The arrays were washed using wash buffer system (NimbleGen Systems), dried by centrifugation, and scanned at 5 μm resolution using the GenePix® 4000B scanner (Axon Instruments). Fluorescence intensity raw data were obtained from scanned images of arrays using NimbleScan™ 2.0 extraction software (NimbleGen Systems). For each spot on the array, log2-ratios of the Cy5-labeled test sample versus the Cy3-labeled reference sample were calculated. Then, the biweight mean of this log2 ratio was subtracted from each point; this procedure is approximately equivalent to mean-normalization of each channel. Please refer to Supplementary Methods for data analysis.

2.5 Generation of antibodies

cDNA encoding the N-terminus of mouse Klf2 (amino-acids 90-255), Klf4 (amino-acids 1-320) or Klf5 (amino-acids 1-320) was cloned into pET42b (Novagen) for the production of GST-fusion protein with 6 His tagged at the C-terminus. The recombinant proteins were produced in BL21 after IPTG induction and purified with GSH-sepharose (Pharmacia) and Ni-NTA-sepharose (Qiagen) column. The purified antigens were then used to immunize rabbits for polyclonal antibodies production.

2.6 Plasmids construction

shRNA constructs were designed as previously described. Two shRNA constructs each for *Klf2*, *Klf4* and *Klf5* were designed to target 19 base-pair gene-specific regions. Oligonucleotides were cloned into pSuper-puro (BglII and HindIII sites; Oligoengine). pSuperpuro constructs were used to express shRNA against *luciferase* (Firefly) or *GFP* as controls. ClaI-XhoI sites were used to insert different *H1*-shRNA cassette from other shRNA constructs digested by BstBI-XhoI to make double or triple RNAi constructs. The H1 promoter cassette in pLVTH lentivirus vector was replaced by the H1-shRNA cassette excised from pSUPER-puro, generating LV-shRNA plasmid. The coding regions of *Klf2*, *Klf4* and *Klf5* were amplified by RT-PCR, and cloned into pCAG-3HA-hyg. For reporter assays in ES cells, DNA constructs used in reporter assays were cloned into the pGL3 plasmid (Promega) (BamH I and Sal I sites) downstream of a *luciferase* gene driven by *Pou5f1 (Oct4)* minimal promoter.

2.7 Custom design genomic tilting arrays

Custom genomic tilting arrays were manufactured by NimbleGen. This microarray design contained probes interrogating 402 genes and 4 gene clusters at a 45 bp tiling interval. Sequences derived from 20 miRNA genes; 7 imprinted loci; 7 germ cell marker genes; 1 oocyte gene; 155 genes expressed in ES cells; 129 genes expressed in differentiated cells; 79 control genes and 8 cell cycle genes were present on this microarray. The single array contains more than 360,000 50mer probes selected from 57 megabases of mouse sequence (May 2004, mm5 build) and we surveyed 50 kb surrounding each gene. See appendix for a list of genes and their coordinates.

2.8 ChIP-on-chip data analysis

Identification of binding regions

For each probe, a ratio between its signal from transcription factor specific antibody and that from the control antibody was taken and subsequently $\log_2$-transformed. Tamalpais program[196] was applied to identify the peak regions on each of the triplicate ChIP-on-chip samples with a p-value cutoff of 0.05. Tamalpais computes the p-value for a peak by calibrating, to a geometric distribution, the number of consecutive probes whose log ratios are among top 5% largest log ratios of all the probes on the array. A binding region of a transcription factor is further determined by requiring two out of the three ChIP-on-chip replicates to have peaks at the same chromosomal location.

Significance analysis of overlaps of binding regions

260, 896 and 720 binding regions were computed for Klf2, Klf4 and Klf5, respectively. The Venn diagram shows that there are all together 1167 non-redundant binding regions and 205 overlaps of binding regions that bind all of the three transcription factors. A Monte Carlo simulation was performed to assess the significance of the overlapping binding regions. The simulation randomly drew 260, 896 and 720 objects from a total of 1,167 objects for 100,000 times. A p-value less than 0.00001 was derived from the proportion of the random draws whose overlaps were larger than the size of the actual overlap, 205.

2.9 DNA microarray

mRNAs derived from triple knockdown shRNA and *Luc* shRNA treated ES cells were reverse transcribed, labeled and analyzed using Illumina microarray platform (Sentrix Mouse-6 Expresion BeadChip v1.1). Arrays were processed as per manufacturer's instructions. Three biological repeats of profiles (each for control and triple knockdown) were used to generate statistically significant gene lists. The microarray data were analyzed by SAM. The thresholds for the differentially expressed genes were (I) more than 1.5 fold change and (II) q-value of less than 0.05.

2.10 Electrophoretic mobility shift assays (EMSAs)

Recombinant proteins of the Klf DNA binding domain (GST tagged) were used for the gel shift assays. The DNA binding domains were cloned into the pET42b (Novagen) vector. The proteins were expressed and purified with GSH-sepharose beads (Amersham). Eluents were dialyzed against a dialysis buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM $ZnCl_2$ and 10% glycerol) at 4°C for 6 h. Proteins were stored at -80°C. Concentrations of proteins were verified with the Biorad protein measurement assay. For EMSA, DNA oligonucleotides (Proligo) labeled with biotin at the 5' end of the sense strands were annealed with the antisense strands in the annealing buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA) and purified with agarose gel DNA extraction kit (Qiagen). DNA concentrations were determined by the NanoDrop ND-1000 spectrophotometer. The gel shift assays were performed using a LightShift Chemiluminescent EMSA kit (Pierce Biotechnologies). 100 ng of protein were added to a 5 µl reaction mixture (final) containing 1 µg of poly(dI-dC) (Amersham), 1 ng of biotinlabeled oligonucleotide in the binding buffer (12 mM HEPES, pH 7.9, 12%

glycerol, 60 mM KCl, 0.25 mM EDTA, 1 mM DTT). Binding reaction mixtures were incubated for 20 min at room temperature. Binding reaction mixtures were resolved on pre-run 6% native polyacrylamide gels in 0.5X Tris-buffered EDTA. Gels were transferred to Biodyne B nylon membranes (Pierce Biotechnologies) using Western blot techniques and detected using chemiluminescence.

2.11 Luciferase reporter assay

For reporter activity, 300 ng of reporter constructs and 0.5 ng *Renilla* luciferase (pRLCMV; Promega) were co-transfected into the ES cells. Firefly and *Renilla* luciferase activities were measured 48 h after transfection with the Dual Luciferase System (Promega) using a Centro LB960 96-well luminometer (Berthold Technologies). For RNAi luciferase assays, 200 ng of shRNA plasmids was co-transfected with the reporter and *Renilla* luciferase constructs. Puromycin selection was introduced 12 h after transfection at 1.0 μg/ml. Firefly and *Renilla* luciferase activity were measured 48-60 h after transfection.

2.12 Apoptosis assay

Annexin V staining was used to investigate the levels of apoptosis. ES cells were transfected with LV-shRNA (GFP) constructs for 72 h before they were harvested and labeled Annexin V-PE Apoptosis Detection Kit (BD Pharmingen; catalog number 559763) for cytometric assay. As positive control, ES cells were treated with 2 μg/ml camptothecin (Sigma-Aldrich) for 12 h to trigger apoptosis. Cells were trypsinized and

single cell suspensions were stained with Annexin V-PE according to the manufacture's instructions.

2.13 Embryo collection, RNA isolation, reverse transcription, and real-time PCR analysis

Embryos were collected from B6CBAF1 females (age 6–8 weeks) that had been superovulated with 5 IU PMSG (Calbiochem) followed by injection of 5 IU hCG (Sigma) 48 hrs later and after successful mating with ICR males. MII oocytes were collected 48 h post-PMSG injection. Embryos dissected at the appropriate time post-fertilization were collected and pooled in M2 medium prior to RNA extraction. Pools of 150 oocytes or embryos were collected at each stage. Total RNA was extracted from pooled embryos using the PicoPure RNA isolation kit (Arcturus Bioscience). cDNA synthesis was at 37°C for 2 h using the high capacity cDNA archive kit (Applied Biosystems). The subsequent PCR reaction mixes were loaded onto low density array (LDA) plates and run on a 7900HT Sequence Detection System (Applied Biosystems). 4.7 embryo equivalents were loaded into each port with a port leading to 48 reaction wells. TaqMan real-time primer probe sets (*Klf2*-Mm00500486_g1; *Klf4*-Mm00516104_m1; *Klf5*-Mm00456521_m1; *Oct4*-Mm00658129_gH; *Cdx2*-Mm00432449_m1; *Tbp*-Mm00446973_m1) were aliquoted into the reaction wells of the LDA by the manufacturer (Applied Biosystems). Two technical replicates for each gene at each developmental time point were generated. Threshold cycle (Ct) values were calculated from the system's software (SDS 2.2) and used as a direct measure of gene expression.

The equation (36-Ct) POWER1.9 was used to determine the fold-change above background with the background level of detection set at a Ct of 36.

2.14 Retrovirus packaging and infection.

Coding sequences of Klfs were amplified from mouse ES cells by RT‑PCR and cloned into MMLV-based pMXs retroviral vector. The retroviruses were generated as described previously. To induce iPS cells, equal amounts of viruses encoding different factors were applied to mouse embryonic fibroblasts (MEFs) at 50–70% confluence in DMEM containing 15% FBS and 6 ng/L polybrene. After 24 h, the medium was replaced with fresh MEF medium and on the following day (2 dpi), the cells were split as 1:6 to 1:20 on MEF feeders. The culture was then maintained for 11–24 days in mouse ES-cell culture medium.

Table 2.1 Quantitative realtime RCR primers

Gene	Sense	Antisense
Klf2	ATGGCGCTCAGCGAGCCTATCTT	CAATCCCATGGAGAGGATGAAGTC
Klf4	GAACTCACACAGGCGAGAAACCTTA	CTCTTCATGTGTAAGGCAAGGTGGT
Klf5	GCCTCAGTGGTAGACCAGTTCTTCA	CTCTGGTGGCTGAAAATGGTAACAG
Pou5f	CATACTGTGGACCTCAGGTTGGACT	AGTTGGTTCCACCTTCTCCAACTTC
Nano	GGTTTGCCTAGTTCTGAGGAAGCAT	TCAGGACTTGAGAGCTTTTGTTTGG
Esrrb	GCCTCAAAGTGGGGATGCTGAAGGAAGGTG	GCCAATTCACAGAGAGTGGTCAGGGCCTTG
Tbx3	TCCACAAACTGAAACTCACCAACAA	CAGTAACGGCGATGAATTCTGTTTC
Nr5a2	GCCATGTCTCAGGTGATCCAAGCAATGCCC	GGGTACTCAGACTTGATGGCCCGACTAGGA
Zfp42	CAGAGAAGCCAGAGGGCGGTGTGTA	GCTTCTCTCCGGTGTGCACCAGAAA
Jmjd1	TGCCTTCCAGGTTTGATGATCTGATGG	CCCACATAAACCATGACATTGGCTGCA
Nes	CAGAGAGGGGACCTGGAACATGAAT	CCTGGCCACTGATATCAAAGGTGTCT
Fgf5	GAGAGTGGTACGTGGCCCTGAACAAGAGAG	CTTCAGTCTGTACTTCACTGGGCTGGGACT
En1	GTGGTCAAGACTGACTCACAGCAAC	GTGATATAGCGGTTTGCCTGGAACT
Gata4	AAGCTCCATGGGGTTCCCAGGCCTCTTGCAAT	TGAATGTCTGGGACATGGAGCTGCTGTGCC
Gata6	TGTGCAATGCATGCGGTCTCTACAGCA	TTCATAGCAAGTGGTCGAGGCACCC

Table 2.2 RNAi sequences

Gene	Target sequence 1	Target sequence 2
Klf2	GGCAAGACCTACACCAAGA	GTGAGAAGCCTTATCATTG
Klf4	GAGGAACTCTCTCACATGA	GACCACCTTGCCTTACACA
Klf5	GACCATGCGTAACACAGAT	GTACCAGCTGTTGAATACA
Klf10	GACTGGAAGTCTCATTTCA	
Nanog	GAACTATTCTTGCTTACAA	

Table 2.3 Antibodies used for ChIP

Factor	Source	Catalog Number	Anitgen
Klf2	custom made		amino-acids 90-255
Klf4	custom made		amino-acids 1-320
Klf5	custom made		amino-acids 1-320
Nanog	Cosmo Bio	REC-RCAB0002PE	
p53	Santa Cruz	Sc-100	
HA	Santa Cruz	Sc-7392	
GST	Santa Cruz	Sc-459	

CHAPTER III

RESULTS

CHAPTER III RESULTS

Reprogramming is a process of inducing highly differentiated cells to reverse their developmental potential to attain the pluripotent state. Ectopic expression of four transcription factors (Oct4, Sox2, c-Myc and Klf4) are sufficient to convert fibroblasts to induced pluripotent stem (iPS) cells that have many characteristics of ES cells[56]. These iPS cells are genetically and epigenetically similar to ES cells and express ES-cell-specific markers. In a proof-of-principle experiment, iPS cells combined with gene and cell therapy were used to correct the defective human sickle haemoglobin allele in mice[197]. Hence, the derivation of specialized cell types from autologous iPS cells could be developed for potential therapeutic applications in different disease models.

Transcription factors Oct4, Sox2, c-Myc and Klf4 are also referred to as reprogramming factors. Oct4 and Sox2 have been shown to be critical for the self-renewal state of ES cells and c-Myc has also been implicated in promoting the self-renewal of ES cells[40, 102, 124, 132, 166]. Klf4 belongs to the Krüppel-like factor (Klf) family of evolutionarily conserved zinc finger transcription factors that regulate numerous biological processes, including proliferation, differentiation, development and apoptosis[198]. The functions of Klf4 and other family members in ES cells are largely unknown.

3.1 Klf2, 4 and 5 are required for the maintenance of mouse ES cells

3.1.1 Individual Klf is not essential for the mouse ES growth

Mouse ES cells are able to grow in culture dishes by using a combination of feeders, cytokine, growth factors or serum. Undifferentiated mouse ES cells show typical compact

and colony-like morphology and express specific alkaline phosphatase (AP). Upon differentiation, ES cells will lose typical morphology and stain poorly for alkaline phosphatase. To examine the functional roles of Klf4 in ES cells, RNAi was used to deplete the endogenous expression of the factor. Two independent short-hairpin RNA (shRNA) constructs targeting different regions of the *Klf4* transcript were used to make sure that the knockdown effects were specific. The knockdown effect by RNAi can be tested by RT-PCR. The *Klf4* transcript after knockdown was reduced to 20 percent or less compared to that in control ES cells (**Figure 3.1a**). Alkaline phosphatase staining indicated by red color and observations of morphological change were used to evaluate whether the ES cells underwent differentiation. Consistent with a previous report[162], no change in morphology and alkaline phosphatase staining in the ES cells was observed, indicating that Klf4 alone is not required for the maintenance of undifferentiated state of ES cells (**Figure 3.1b**).

Considering Klf4 is one of the Krüppel-like factor (Klf) family members, we hypothesized that other Klf transcription factors enriched in undifferentiated ES cells may play a redundant role with Klf4. To this end, published time-course expression profiling of retinoic acid (RA)-treated ES cells were analyzed to identify other Klfs that are upregulated in undifferentiated ES cell state[132]. In addition to Klf4, two other Klfs (Klf2 and Klf5) were enriched in ES cells and downregulated upon RA induced differentiation (**Figure 3.1c**). To address the roles of Klf2 (also known as lung Krüppel-like factor) and Klf5 (also known as intestine-enriched Krüppel-like factor), their expression was depleted by two independent RNAi respectively (**Figure 3.1d)**. As no

change in morphology and alkaline phosphatase staining were observed (**Figure 3.1e)**, we conclude that individual Klfs are not important for the self-renewal of ES cells.

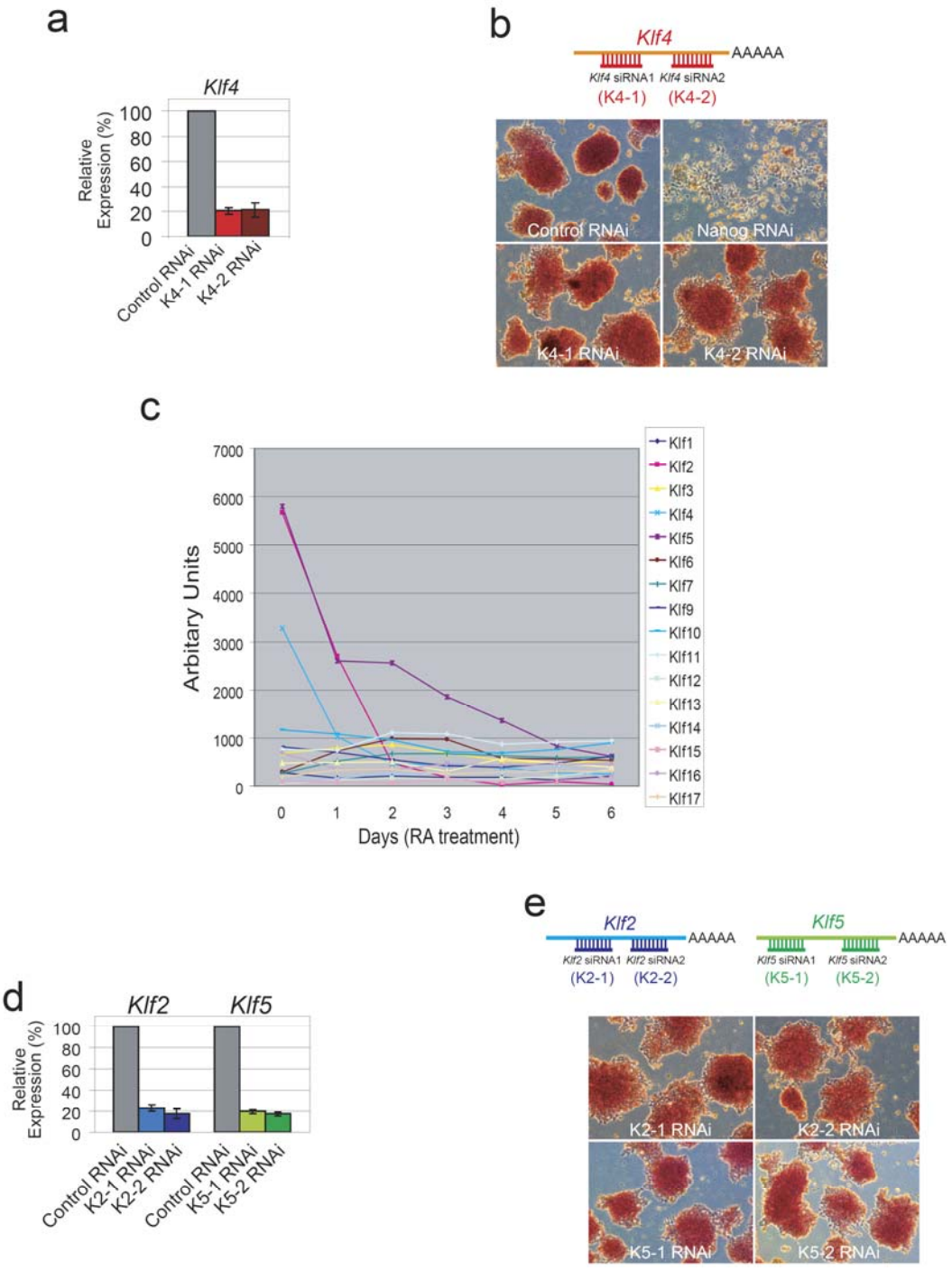
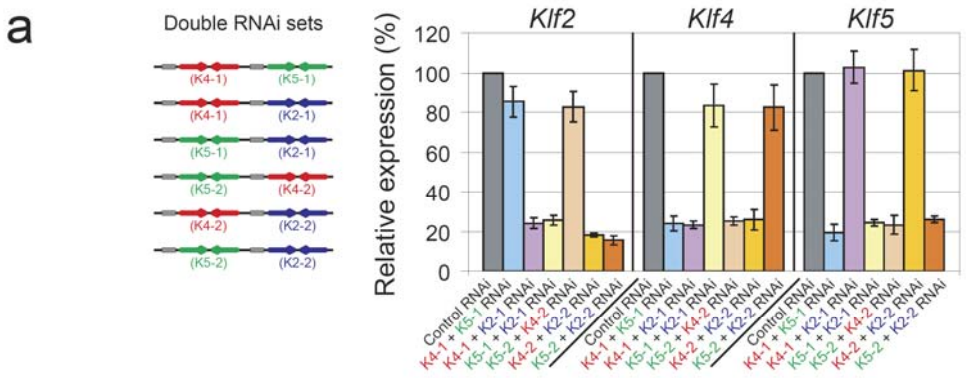

a
Klf4
Relative Expression (%)
100
80
60
40
20
0
Control RNAi
K4-1 RNAi
K4-2 RNAi
b
Klf4
AAAAA
Klf4 siRNA1
Klf4 siRNA2
(K4-1)
(K4-2)
Control RNAi
Nanog RNAi
K4-1 RNAi
K4-2 RNAi
c
Arbitary Units
7000
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
Days (RA treatment)
Klf1
Klf2
Klf3
Klf4
Klf5
Klf6
Klf7
Klf9
Klf10
Klf11
Klf12
Klf13
Klf14
Klf15
Klf16
Klf17
d
Klf2
Klf5
Relative Expression (%)
100
80
60
40
20
0
Control RNAi
K2-1 RNAi
K2-2 RNAi
Control RNAi
K5-1 RNAi
K5-2 RNAi
e
Klf2
AAAAA
Klf2 siRNA1
Klf2 siRNA2
(K2-1)
(K2-2)
Klf5
AAAAA
Klf5 siRNA1
Klf5 siRNA2
(K5-1)
(K5-2)
K2-1 RNAi
K2-2 RNAi
K5-1 RNAi
K5-2 RNAi

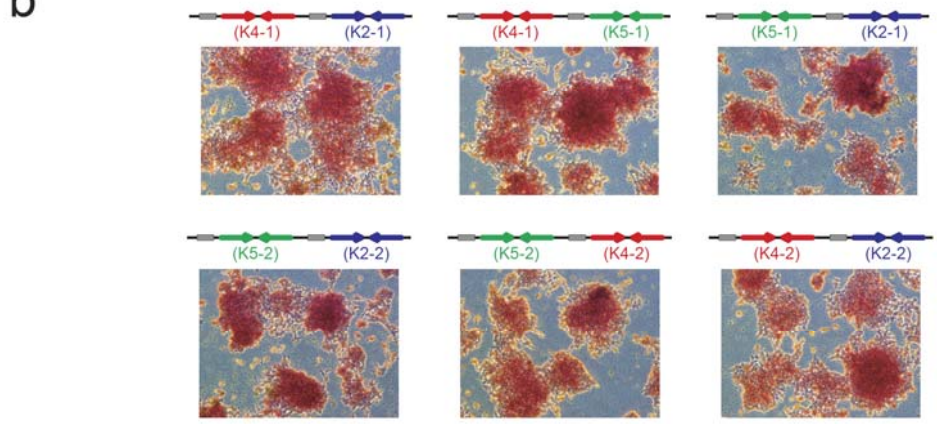
Figure 3.1 Individual Klf2, Klf4 and Klf5 is not essential for the mouse ES growth
(a) Quantitative real-time PCR analysis of *Klf4* expression after knockdown using two shRNA constructs (K4-1 and K4-2) targeting different regions. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Error bars represent the s.e.m. of three technical replicates.
(b) Depletion of Klf4 by RNAi. Typical colony morphology of ES cells with positive alkaline phosphatase staining (red) was maintained after Klf4 knockdown using two shRNA constructs targeting different regions of *Klf4*. Flattened fibroblast-like cells were formed after Nanog depletion. In control *Luc* shRNA transfected cells, normal undifferentiated phenotype with distinct ES cell colonies was maintained.
(c) *Klf2*, *Klf4* and *Klf5* are down-regulated upon ES cell differentiation. Expression of 16 members of the *Klf* family at different days after RA induced differentiation. The microarray expression data is obtained from the literature.
(d) Quantitative real-time PCR analysis of *Klf2* and *Klf5* expression after knockdown using two shRNA constructs (K2-1 and K2-2 for Klf2, and K5-1 and K5-2 for Klf5) targeting different regions of the respective genes. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Error bars represent the s.e.m. of three technical replicates.
(e) Depletion of Klf2 and Klf5 by RNAi. Normal undifferentiated ES colonies with positive alkaline phosphatase staining were maintained after knockdown using two shRNA constructs targeting different regions of *Klf2* or *Klf5* transcripts, respectively.

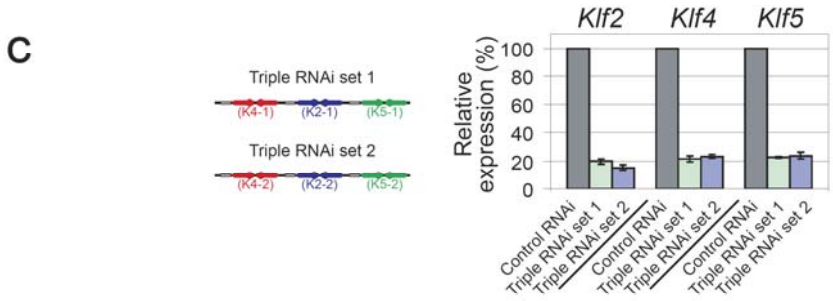
3.1.2 Klf2, Klf4 and Klf5 are required for the maintenance of ES cells

Next, it would be interested to examine whether Klf2 or Klf5 can replace functions of Klf4 in ES cells. Systematic pair-wise double knockdown constructs which targeting two different transcripts were made by cloning two different RNAi cassettes into one shRNA vector. Each double knockdown construct are capable to knockdown two different *Klf* transcripts at the same time. As there are two independent shRNA for each Klf factor, six double knockdown constructs were cloned to ensure that the knockdown effects were specific (**Figure 3.2a**). The results revealed no change in morphology and alkaline phosphatase staining upon depletion of any two of the Klfs (**Figure 3.2b**).

To further investigate whether any of Klf2, Klf4 or Klf5 alone is sufficient for the ES growth, triple knockdown using two independent constructs with each construct targeting the three Klf transcripts at different regions was then performed. RT-PCR result showed that RNAi effects were not compromised when three different factors simultaneously depleted (**Figure 3.2c**). Remarkably, the colony morphology of the ES cells was lost and they resembled fibroblasts. The alkaline phosphatase staining was also lost, indicative of differentiation (**Figure 3.2d**). Thus, depletion of Klf2, Klf4 and Klf5 led to differentiation and consequently disrupted the self-renewal of ES cells.


a
Double RNAi sets
(K4-1)
(K5-1)
(K4-1)
(K2-1)
(K5-1)
(K2-1)
(K5-2)
(K4-2)
(K4-2)
(K2-2)
(K5-2)
(K2-2)
Klf2
Klf4
Klf5
Relative expression (%)
120
100
80
60
40
20
0
Control RNAi
K4-1 + K5-1 RNAi
K4-1 + K2-1 RNAi
K5-1 + K2-1 RNAi
K5-2 + K4-2 RNAi
K4-2 + K2-2 RNAi
K5-2 + K2-2 RNAi


b
(K4-1)
(K2-1)
(K4-1)
(K5-1)
(K5-1)
(K2-1)
(K5-2)
(K2-2)
(K5-2)
(K4-2)
(K4-2)
(K2-2)


c
Triple RNAi set 1
(K4-1)
(K2-1)
(K5-1)
Triple RNAi set 2
(K4-2)
(K2-2)
(K5-2)
Klf2
Klf4
Klf5
Relative expression (%)
100
80
60
40
20
0
Control RNAi
Triple RNAi set 1
Triple RNAi set 2

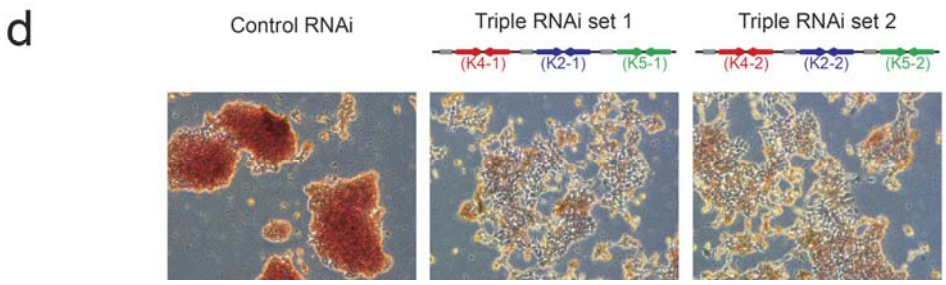

d
Control RNAi
Triple RNAi set 1
(K4-1)
(K2-1)
(K5-1)
Triple RNAi set 2
(K4-2)
(K2-2)
(K5-2)

Figure 3.2 Klf2, Klf4 and Klf5 are required for the maintenance of ES cells
(a) Depletion of two Klfs by combinatorial double RNAi. Quantitative real-time PCR analysis of *Klf2*, *Kfl4* and *Klf5* expression after knockdown using shRNA constructs targeting two Klfs of Klf2, Klf4 and Klf5. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Error bars represent the s.e.m. of three technical replicates.
(b) Concurrent knockdown of two Klfs using constructs targeting the respective Klfs. Undifferentiated ES colonies with positive alkaline phosphatase staining were maintained after knockdown using shRNA constructs targeting two of *Klf2*, *Klf4* and *Klf5*.
(c) Depletion of the three Klfs by triple RNAi. Quantitative real-time PCR analysis of *Klf2*, *Kfl4* and *Klf5* expression after knockdown using two sets of shRNA constructs targeting *Klf2*, *Klf4* and *Klf5*. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Error bars represent the s.e.m. of three technical replicates.
(d) Triple knockdown of Klf2, Klf4 and Klf5 with two constructs targeting different segments of the Klf transcripts. Differentiated cells with negative alkaline phosphatase staining were formed after knockdown using two sets of shRNA constructs targeting *Klf2*, *Klf4* and *Klf5*.

3.1.3 Specificity of Klf RNAi

RNAi is widely used for functional genomic studies and screens in a number of organisms[199-202]. However, recent studies have shown that siRNA-mediated silencing may be less specific than was originally believed[203, 204]. Several assays were performed to confirm the specificity of our RNAi experiments.

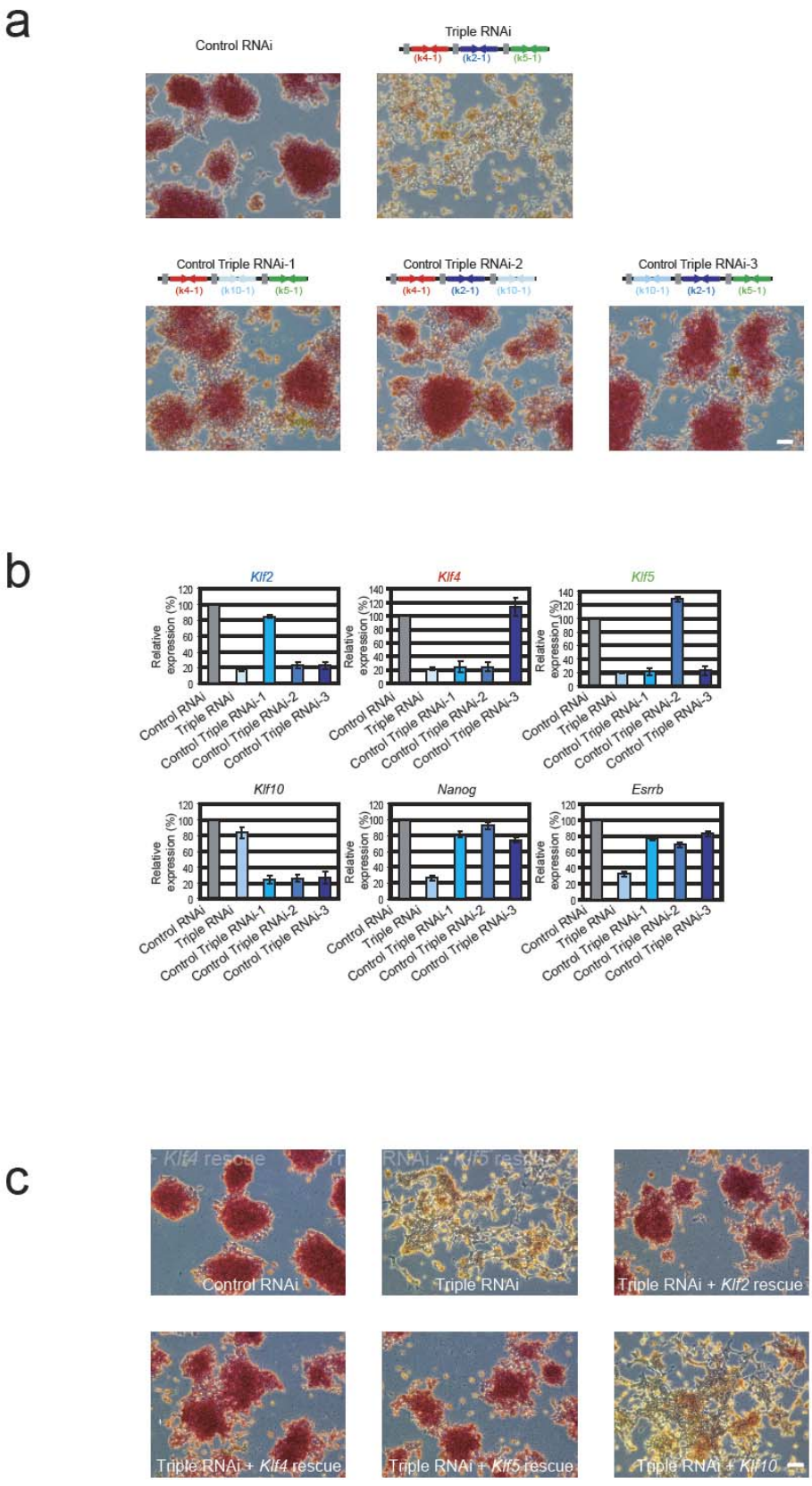
To rule out the possibility that differentiation phenotype by Klf triple RNAi is derived from over-saturating RNAi pathways, control triple RNAi constructs were made by replacing one Klf RNAi cassette with Klf10 RNAi in Klf triple RNAi construct. As one of Krüppel-like factor family members, Klf10 is expressed in ES cell according to the RT-PCR. The results revealed no change in morphology and alkaline phosphatase staining upon depletion of any two of the Klfs together with Klf10 (**Figure 3.3a**). The

knockdown effects of control triple RNAi constructs were confirmed by RT-PCR (**Figure 3.3b**).

For the RNAi experiments, off-target effects can not be excluded even by using independent RNAi constructs. "The key control for an RNAi experiment, which should be provided whenever technically feasible, is rescue of the observed RNAi phenotype by artificial expression of a RNAi-resistant form of the targeted mRNA"[205]. Rescue experiment remains the gold standard control for specificity in RNAi experiment, especially for multiple RNAi. To further confirm the results of the knockdown experiments, RNAi-resistant constructs encoding *Klf2*, *Klf4* or *Klf5* was made by silent mutating the RNAi target sequences. The results showed that RNAi-resistant *Klf2*, *Klf4* or *Klf5* cDNA was able to rescue the differentiation phenotype induced by Klf triple knockdown (**Figure 3.3c)**. However, *Klf10* was unable to rescue this phenotype. The concurrent requirement for Klf2, Klf4 and Klf5 in the maintenance of self-renewal of ES cells strongly suggests that they are functionally redundant.

Virus-based RNAi vector systems have some advantages over transient vector systems or RNAi duplex in terms of stable expression of shRNA and high transduction efficiency[206]. To further confirm the results of Klf triple RNAi effect in ES cells, triple RNAi cassette was conferred to a lentivirus-based RNAi vector containing a constitutive marker *GFP*. The high virus transduction efficiency was shown as most ES cells express GFP after virus infection. The ES cells were flatten and fibroblast-like upon two independent virus RNAi transduction, indicative of differentiation (**Figure 3.3d)**.

The triple knockdown cells showed reduced proliferation rates compared to control RNAi cells by counting of cell number (**Figure 3.3e)**. To test whether triple knockdown cells undergo apoptosis, early marker annexin-V was measured after Klf triple RNAi. The results showed that triple knockdown cells did not undergo apoptosis as the number of annexin-V positive cells did not increase compared to that in control cells (**Figure 3.3f)**.


a
Control RNAi
Triple RNAi
(k4-1)
(k2-1)
(k5-1)
Control Triple RNAi-1
Control Triple RNAi-2
Control Triple RNAi-3
(k10-1)
b
Klf2
Klf4
Klf5
Klf10
Nanog
Esrrb
Relative expression (%)
Control RNAi
Triple RNAi
Control Triple RNAi-1
Control Triple RNAi-2
Control Triple RNAi-3
c
Control RNAi
Triple RNAi
Triple RNAi + Klf2 rescue
Triple RNAi + Klf4 rescue
Triple RNAi + Klf5 rescue
Triple RNAi + Klf10 rescue

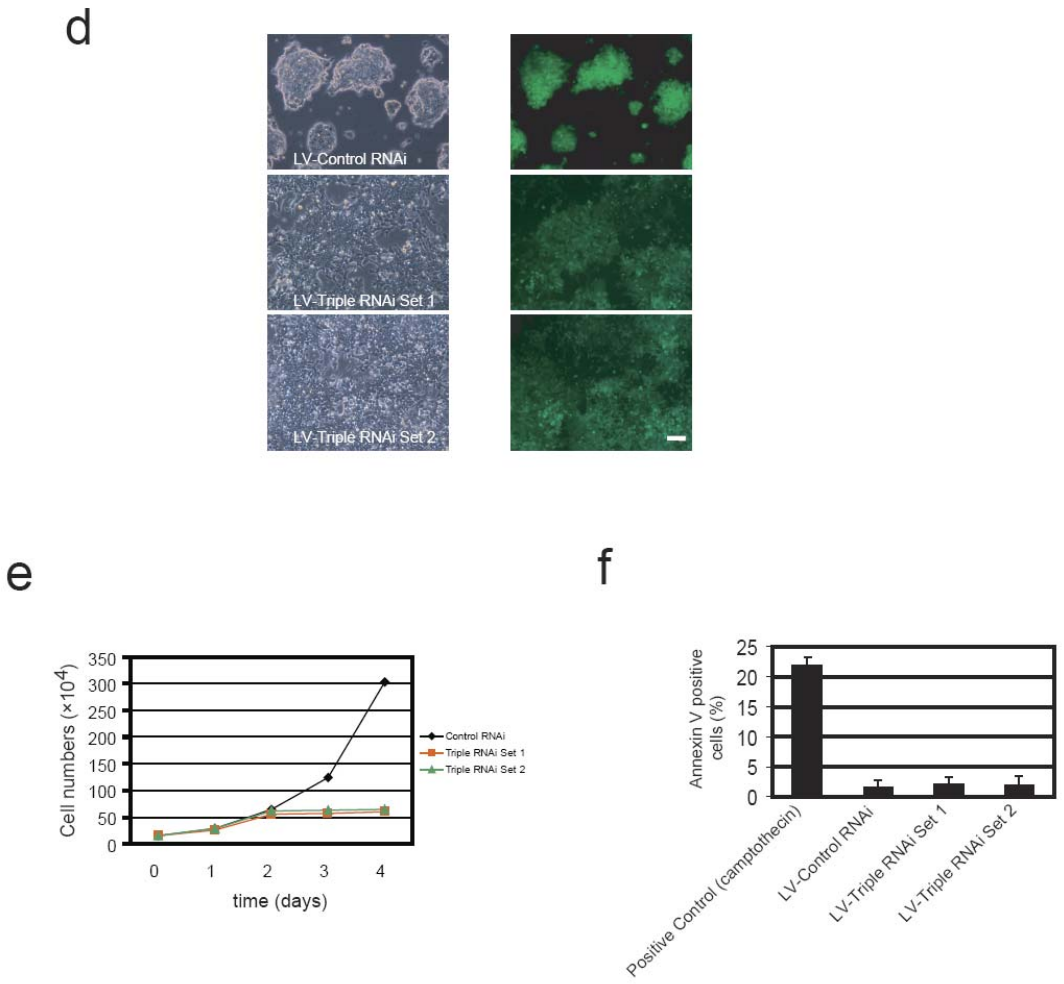


Figure 3.3 Specificity of Klf knockdown by RNAi
(a) Differentiated cells with reduced alkaline phosphatase staining were formed after depletion of Klf2, Klf4 and Klf5. Undifferentiated ES colonies with positive alkaline phosphatase staining were maintained after knockdown using shRNA constructs targeting Klf10 in conjunction with any other two Klfs. The scale bar represents 100 μm.
(b) Quantitative real-time PCR analysis of gene expression after triple Klf knockdown. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Error bars represent the s.e.m. of three technical replicates.
(c) Rescue of triple knockdown phenotype with coexpression of RNAiresistant cDNAs of *Klf2*, *Klf4* or *Klf5*. Typical colony morphology of ES cells with positive alkaline phosphatase staining (red) was restored in triple knockdown cells with coexpression of RNAi-resistant cDNA constructs of *Klf2*, *Klf4* or *Klf5*. Klf10 was unable to rescue the differentiation induced by depletion of Klf2, Klf4 and Klf5. The scale bars represent 100 μm in all panels.

(d) Triple knockdown of Klf2, Klf4 and Klf5 with two sets of LV-shRNA constructs targeting different segments of the Klf transcripts. Flattened fibroblast-like cells were formed after transduction of lentivirus produced by Klfs triple RNAi constructs. For control lentivirus transducted cells, normal undifferentiated phenotype with distinct ES cell colonies was maintained. GFP signal was detected by fluorescent microscope. The scale bar represents 100 μm.
(e) Proliferation rate of ES cells transfected with control, Klf triple RNAi set 1 or Klf triple RNAi set 2 constructs was determined by counting cells recovered at different days. Numbers of cells counted is average of 3 different wells from 2 independent experiments. The ES cells were seeded at 1.5×10^5 cells in six-well dishes and cells from each well were harvested and counted on 4 consecutive days.
(f) Apoptosis assay. Quantification of Annexin V positive cells after knockdown using two sets of LV-shRNA constructs targeting Klf2, Klf4 and Klf5. Percentage of Annexin V positive cells were detected by FACS sorting of GFP positive cells.

3.2 Mapping of Klf2, Klf4, Klf5 and Nanog binding loci by ChIP-on-Chip

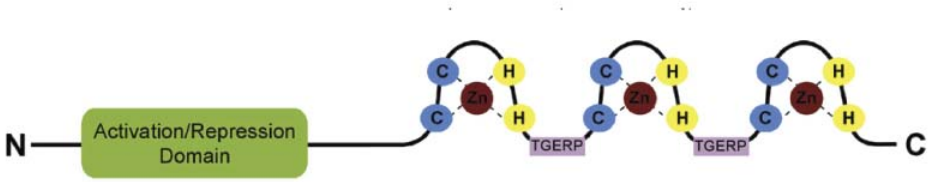
The concurrent requirement for Klf2, Klf4 and Klf5 in the maintenance of self-renewal of ES cells strongly suggests that they are functionally redundant. To identify the *in vivo* targets of these Klfs, chromatin immunoprecipitation (ChIP) was used to determine the binding loci and profiles of these proteins.

3.2.1 Characterization of antibodies raised against Klf2, Klf4 and Klf5

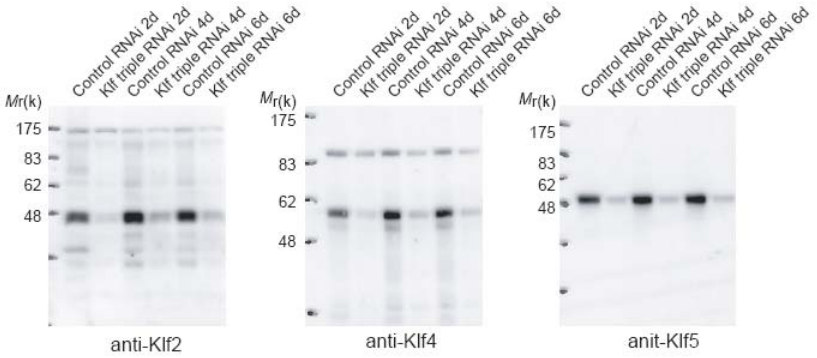
Currently, the chromatin immunoprecipitation technique is the most precise method to identify specific proteins or chromatin modifications associated with a region of the genome or, conversely, identification of regions of the genome associated with specific proteins and chromatin modifications *in vivo*. Specific ChIP-grade antibody is critical for the success of Chromatin IP. ChIP-grade antibodies must be highly specific and recognize the epitope in the chromatin context and under fixed conditions. As Klf2, Klf4 and Klf5 belong to the same family, it is crucial to characterize the antibodies.

Polyclonal antibodies were generated against the divergent amino-termini (**Figure 3.4a)**. Western blot was used to characterize the antibodies and results showed that each Klf antibody can recognize a distinct band with correct size in the ES cell extracts. More importantly, the signal of the bands could be depleted by Klf RNAi (**Figure 3.4b)**. To further test whether there are cross-react activities among the three Klf antibodies, immunoprecipitation were performed. HA tagged Klf2, Klf4 or Klf5 was overexpressed in 293T cells respectively, and cell extracts were subjected to carry out immunoprecipitation. The results showed that each Klf protein was specifically recognized and pulled down by its own antibody but not other Klf anitbodeis or control GST antibody (**Figure 3.4c)**. Taken together, our raised Klf antibodies can specifically recognize and pull down their respective Klf proteins.

a


N
Activation/Repression Domain
C
H
Zn
TGERP
C

b


Control RNAi 2d
Klf triple RNAi 2d
Control RNAi 4d
Klf triple RNAi 4d
Control RNAi 6d
Klf triple RNAi 6d
Mr(k)
175
83
62
48
anti-Klf2
anti-Klf4
anit-Klf5

c

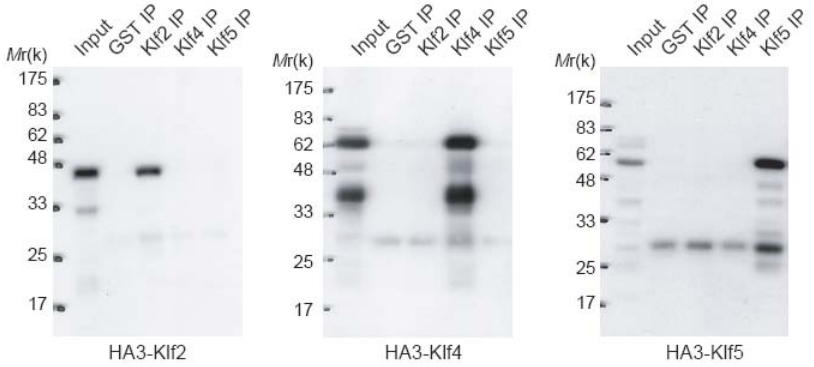

Input
GST IP
Klf2 IP
Klf4 IP
Klf5 IP
Mr(k)
175
83
62
48
33
25
17
HA3-Klf2
HA3-Klf4
HA3-Klf5

Figure 3.4 Characterization of antibodies raised against Klf2, Klf4 and Klf5
(a) Schematic representation of the Klf molecule (adapted from Richard P, 2007). Shown are the three C-terminal C2H2 zinc fingers, each chelating a single zinc ion. The fingers are linked by the "TGERP"-like motif, which assists in binding to target DNA. The activation/repression domain is found at the N-terminus of the molecule.
(b) Protein changes at different days after knockdown of Klf2, Klf4 and Klf5. Western blot analyses of ES cell extracts were carried out using anti-Klf2, anti-Klf4 and anti-Klf5 antibodies. The ES cells were either transfected with control or triple knockdown RNAi constructs and extracts were harvested at different days.
(c) Characterization of antibodies raised against Klf2, Klf4 and Klf5 by immunoprecipitation. 3HA-Klf2, 3HA-Klf4 and 3HA-Klf5 were overexpressed in 293T cells. The 293T cell lysates were harvested 48 h after transfection. Cell lysates were subjected to immunoprecipitation using anti-GST, anti-Klf2, anti-Klf4 or anti-Klf5. The immunoprecipitates were analyzed by Western blotting with anti-HA monoclonal antibody.

3.2.2 ChIP-on-chip assay for Klf2, Klf4, Klf5

ChIP-on-chip assay is a useful tool for genome-wide mapping of *in vivo* protein–DNA interactions, and allowing global views of transcription factor binding and chromatin context. ChIP-on-chip can help to identify regulatory regions, unravel the chromatin features at specific types of regulatory regions, and establish how transcription factors and chromatin modifications are coordinated to control gene expression in the genome.

To obtain the global views of Klf transcription factor binding in living cells, ChIP-on-chip were performed by using ES cells extracts. The ChIP-enriched DNA or input genomic DNA was amplified and the Cy5- and Cy3-labelled DNA was hybridized onto custom-designed genomic tiling microarrays. This microarray design contained probes interrogating 402 genes and four gene clusters at a 45 bp tiling interval. Probe Sequences were derived from 155 genes expressed in ES cells, 129 genes expressed in differentiated cells and 122 "control" genes. More than 360,000 50mer probes were selected from 57 megabases of mouse sequence.

Interestingly, the ChIP-on-chip analysis revealed that Klf2, Klf4 and Klf5 bound to genes encoding key regulators such as *Pou5f1* (encoding for Oct4), *Sox2*, *Nanog*, *Esrrb*, *Tcl1*, *Sall4*, *Tcf3* and *Mycn*[40, 99, 120, 132, 162, 166, 207] (**Figure 3.5)**. The binding profiles at genes compared between Klf2, Klf4 and Klf5 are remarkably similar. Control ChIP against p53, a transcription factor highly expressed in mouse ES cells, showed a different binding profile.

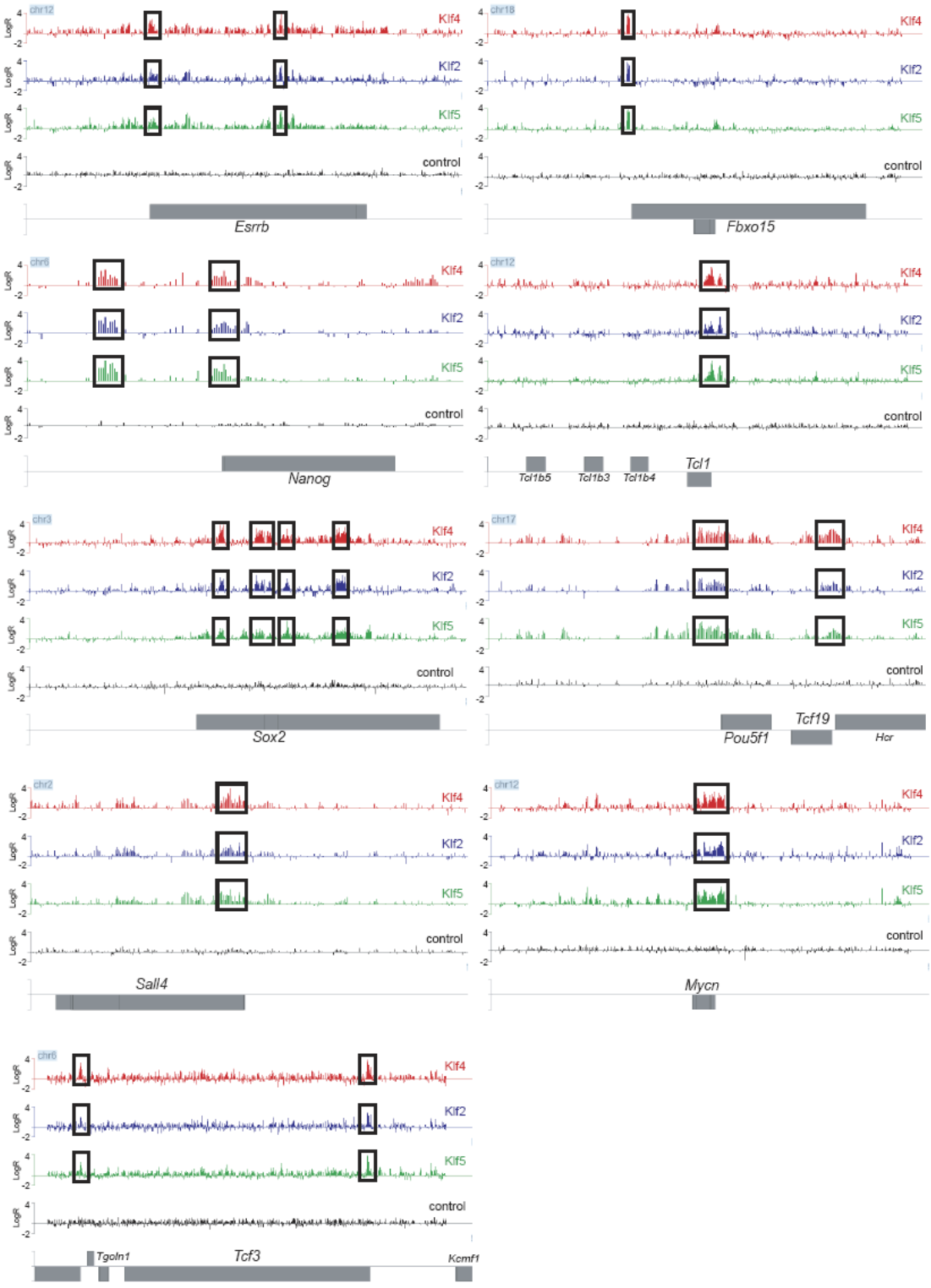

chr12
Klf4
Klf2
Klf5
control
LogR
Esrrb
chr18
Fbxo15
chr6
Nanog
chr12
Tcl1b5
Tcl1b3
Tcl1b4
Tcl1
chr3
Sox2
chr17
Pou5f1
Tcf19
Hcr
chr2
Sall4
Mycn
Tgoln1
Tcf3
Kcmf1

Figure 3.5 Identification of *in vivo* binding sites of Klf4, Klf2 and Klf5 by ChIP-on-chip.
Binding profiles of Klf4, Klf2 and Klf5 at *Esrrb*, *Fbxo15*, *Nanog*, *Tcl1*, *Sox2*, *Pou5f1*, *Sall4*, *Mycn* and *Tcf3* loci. The ChIP-on-chip result for each transcription factor is represented by a different colour (Klf4, red; Klf2, blue; Klf5, green; control p53, black). y axis shows the $\log_2$ ratio (LogR). Screen shots of the SignalMap browser showing Klf4, Klf2 and Klf5 binding at *Esrrb*, *Fbxo15*, *Nanog*, *Tcl1*, *Sox2*, *Pou5f1*, *Sall4*, *Mycn* and *Tcf3* genomic loci are shown.

3.2.3 Validation of Klf2, Klf4, Klf5 ChIP-on-chip

To validate the signals from ChIP-on-chip experiments, different assays were carried out to verify the binding of Klf to the regions of genome.

Conventional ChIP–qPCR was performed and the binding of Klf proteins at *Esrrb*, *Fbxo15*, *Nanog* and *Tcl1* loci were validated by real-time PCR analysis (**Figure 3.6a)**. The binding profiles at *Nanog* and *Fbxo15* promoter regions were also validated by scanning their promoter regions (**Figure 3.6b)** (Table 3.1**)**. The data showed that the binding patterns of three Klfs were prominently similar to the profiles obtained from the ChIP-on-chip.

To rule out the possibilities of non-specific binding by raised antibodies, ChIP using the epitope HA antibody was performed to profile the binding patterns of epitope tagged Klfs expressed in ES cells. The profiles of ectopically expressed HA–Klf2, HA–Klf4 and HA–Klf5 at the *Nanog* promoter, and three other loci, were identical to those obtained from ChIP against endogenous proteins. Control ChIP using anti-HA antibody with ES cell extract did not show enrichment (**Figure 3.6c)**.

To demonstrate further that the ChIP enrichment was specific, each Klf was depleted by RNAi and the occupancy of all three Klf proteins was analysed by ChIP. The binding of each Klf was reduced by shRNA constructs targeting the corresponding *Klf* transcript, indicating that the ChIP signal for each Klf was specific (**Figure 3.6d)**. Most importantly, the data showed that the other Klf proteins were bound to the target sites when one of them was depleted by RNAi. Therefore, the depletion of one Klf can be compensated for by the binding of other Klfs to the common target sites.

Table 3.1 ChIP primers for *Nanog* and *Fbxo15*

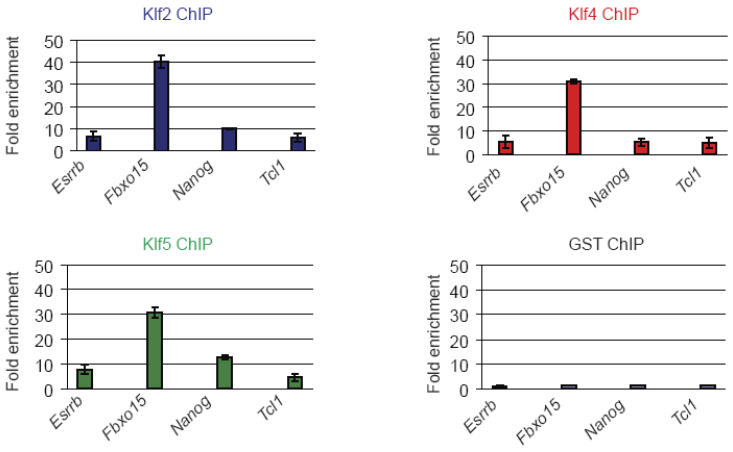
Nanog

Label	Sense	Antisense
1	AACATTTAAAGTGTGAAAGAATGACTCCC	AAATTTCATTGATAGACTCCTGGTTACCAT
2	CCTGGCTGGTGATTCTCCTTCAAGGGCTAA	CATCATCAGTATCAGGGGCAGCTATTGTGT
3	TGGTGATTCTCCTTCAAGGGCTAAGATTAC	TCTTTTGCCAGCATCATCAGTATCAGGG
4	GCCTCAGGAAGAATGAAGCTC	AAAAACTCAGAAAAGCATAACATGAA
5	CCCGGAGGACTCCAAGGCTAGCG	CGGTTTGAATAGGGAGGAGGGCGTCT
6	GCCACCTCTTCGCTCGGATCTTTCAC	TTGCCTCTGGGTCCACGGAGTCA
7	AGCTCTTACAATTCCTCTCCCGGACGGTT	AGCCCAGCAAACTACCTTCACTAGGCCAAA
8	CTTGAGTGTGGCAAAAGTATGTAACTGGGA	GGGGCCCATCGCTAACATATTGAAG
9	GGGCCCCGTGCTTTAGATTTTAGGCTT	ACAAGTGTGTTTTACCCGCTGAGCATTCTG
10	TCCGAGTACTTAACTCAACATACCATTCTTCGT	CCTGATGTGATTCTTGTCATTAGGCAAAG
11	GCATAAACCTTGATATTTTGAACGGCCTATT	ACAGATGGACTAAAGCCCCTAAGTAGAAATCAT
12	AAAATCCATGCAGGCAAAACACCAAT	CATAAACTAATGTGGCTGGTTGACCTCTGTC
13	CCCAAACCAAAAAGAGCCATTCAAGCTTAGTCT	GCCGATCAGTCCTTGTGCTCACAGATAAGT
14	GGCCTCTACATTAATTTAAACACTCCTTAA	CAAGTCTGAAGAAAGAGCCTGTGTC
15	CCAGGTTTCCCAATGTGAAGAGCAAGC	GGCGATCTCTAGTGGGAAGTTTCAGGTCAA
16	TAGATCAGAGGATGCCCCCTAAGCTTTCCC	CTCCTACCCTACCCACCCCCTATTCTCCC
17	GTGGGAAGGCTGCGGCTCACTT	AGTGTGATGGCGAGGGAAGGGATTTCT
18	CTTACAAGGGTCTGCTACTGAGATGCTCTGCAC	GGACACCCGCTTATGTTAATGACTACTACGGTTCA
19	ACCGGTGATACGTTGGCCTTCTAGTCTGAA	TGGGGTGCTCATTCCAAGCTAGGATGTTAG

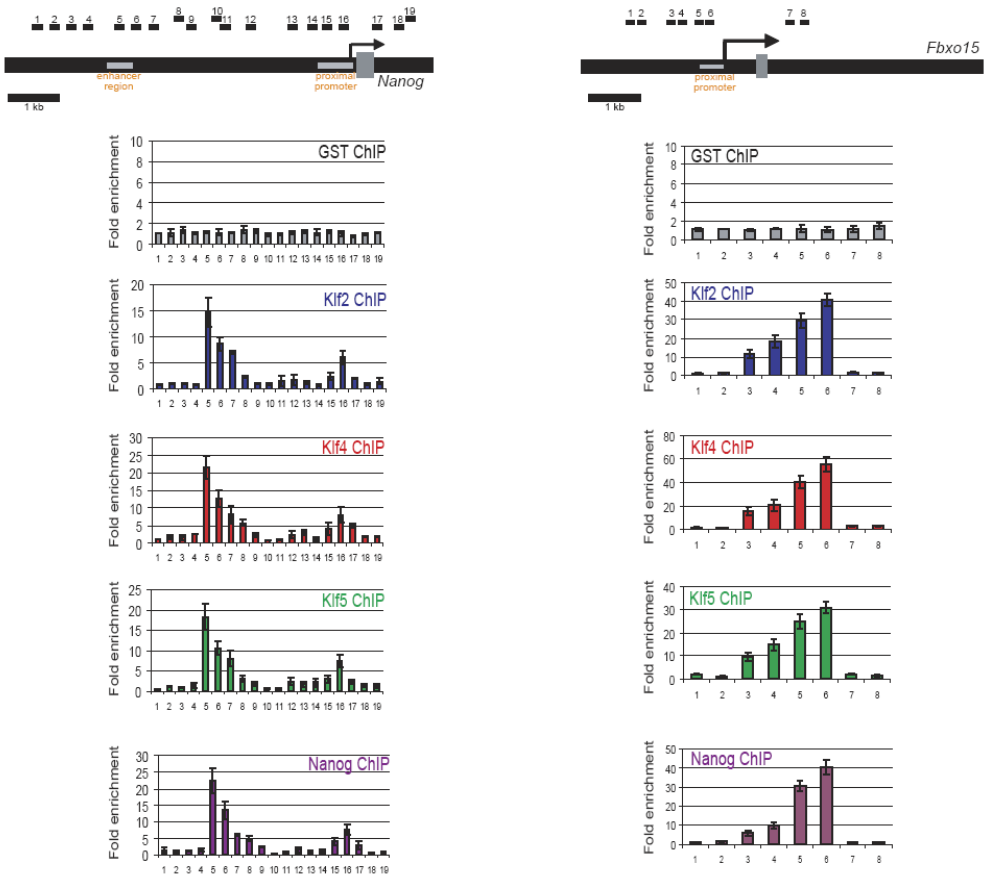
Fbxo15

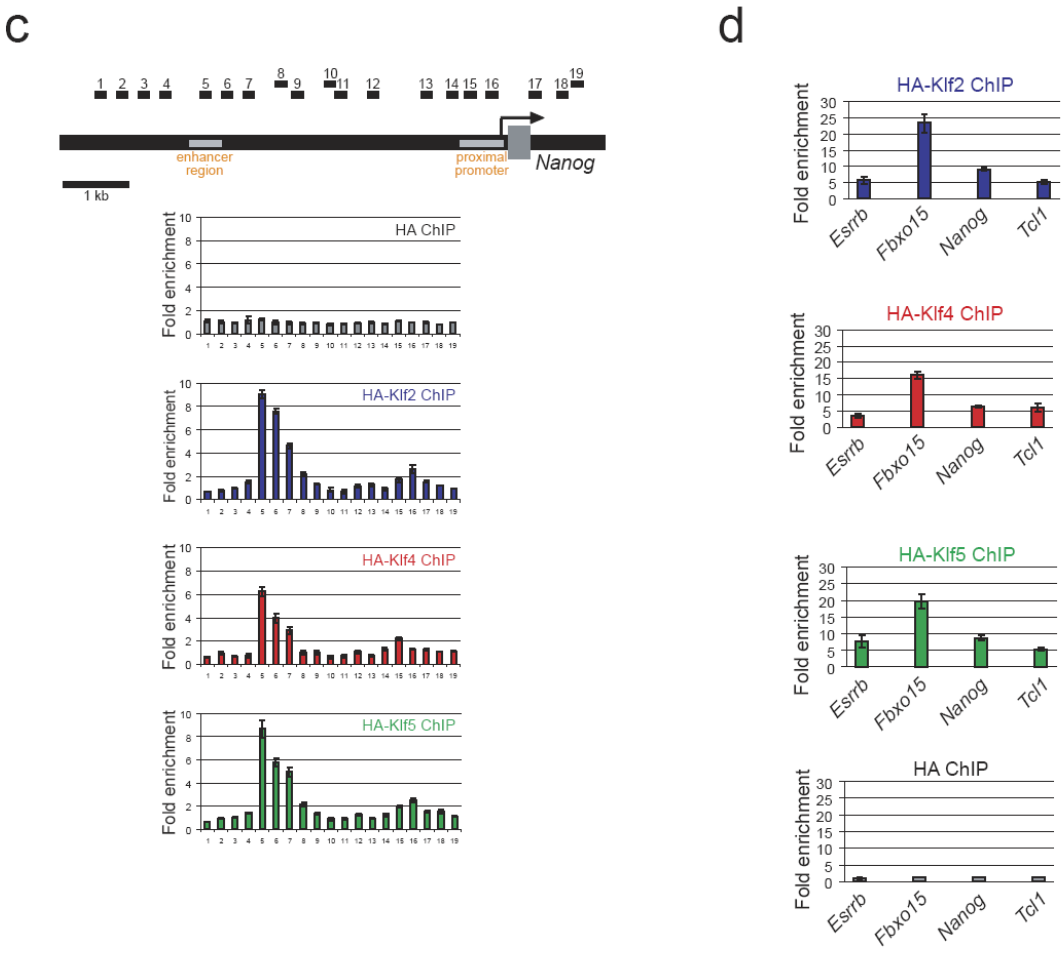
Label	Sense	Antisense
1	ATCTTTATTGGTGGTCCTTGGATTAAC	GCTGCTTTGCTATTTTCACACACTATTC
2	AGCAGACAGGCTTGACTGTTTTTCTATACC	CACACAGCAACAACAGAGAGCCTAGCAAT
3	GGTCTAGGAGGAGGCAGAGTTGGTTGTTA	CTTATGCAGGATTGGTCATACCTATGATTT
4	TTCTGTTATGATTGAAAGGTTTGGGTTCTA	GCCGTGCAATAAGTCAAACTGTATCCTC
5	CCCCTGTAAATTCACTCAGAGGATACAGT	TGGTCCACCTTAACAGTTTATCTAACAAGC
6	GCCCTTAGTTCCCAGATGTGCTTTATCATA	AAAGTAGGGTGAGGCAGTTTGCTTAGCAG
7	CGATGGGCTGTGATCATTTCTAGAGTAGGC	CTCACTCAGCTTCCACCCCTTGTTGCTC
8	TTGAGAATGAGTGGAGGCAGGGAGTAGGTC	GGGGACAAATGGTGAGGACAGTAAAGA

a


Klf2 ChIP
Klf4 ChIP
Klf5 ChIP
GST ChIP
Fold enrichment
Esrrb
Fbxo15
Nanog
Tcl1

b


enhancer region
proximal promoter
Nanog
Fbxo15
1 kb
GST ChIP
Klf2 ChIP
Klf4 ChIP
Klf5 ChIP
Nanog ChIP
Fold enrichment


c
d
enhancer region
proximal promoter
Nanog
1 kb
HA ChIP
HA-Klf2 ChIP
HA-Klf4 ChIP
HA-Klf5 ChIP
Fold enrichment
Esrrb
Fbxo15
Nanog
Tcl1

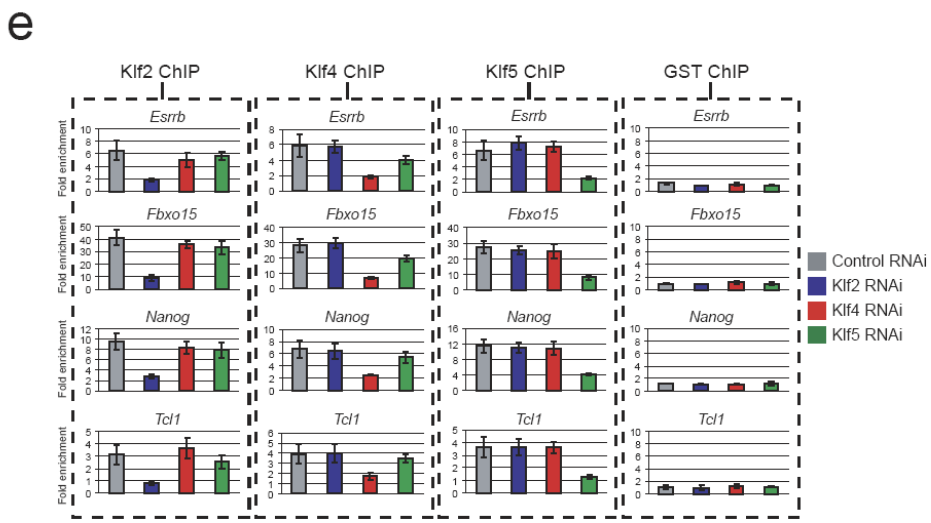

e
Klf2 ChIP
Klf4 ChIP
Klf5 ChIP
GST ChIP
Esrrb
Fbxo15
Nanog
Tcl1
Fold enrichment
Control RNAi
Klf2 RNAi
Klf4 RNAi
Klf5 RNAi

Figure 3.6 Validation of Klf2, Klf4, Klf5 ChIP-on-chip
(a) Klf2, Klf4 and Klf5 occupy *Esrrb*, *Fbxo15*, *Nanog*, *Tcl1* loci in ES cells. ChIP assay was performed using using anti-Klf2 (blue), anti-Klf4 (red), anti-Klf5 (green) antibodies with extract derived from ES cells. Control ChIP was performed with anti-GST antibody. Error bars represent the s.e.m. of three technical replicates.
(b) Validation of Klf2, Klf4 and Klf5 binding. Klf2, Klf4, Klf5 and Nanog bind to *Nanog* promoter region. ChIP assay was performed using anti-Klf2 (blue), anti-Klf4 (red), anti-Klf5 (green) and anti-Nanog (purple) antibodies with extracts derived from ES cells. Schematic showing the location of the amplicons (black bars labeled 1-19) used to detect ChIP-enriched fragment over the 8.2 kb *Nanog* promoter. Amplicons are numbered in order relative to their sites along the gene. The grey box represents an exon. Fold enrichment is the relative abundance of DNA fragments detected by real-time PCR at amplified region over a control amplified region. GST antibody was used as a ChIP control (grey). Error bars represent the s.e.m. of three technical replicates.
Klf2, Klf4, Klf5 and Nanog bind to *Fbxo15* promoter region. ChIP assay was performed using anti-Klf2 (blue), anti-Klf4 (red), anti-Klf5 (green) and anti-Nanog (purple) antibody with extract derived from ES cells. Schematic showing the location of the amplicons (black bars labeled 1-8) used to detect ChIP-enriched fragment over the 2.3 kb *Fbxo15* promoter. Amplicons are numbered in order relative to their sites along the gene. The grey box represents an exon. Fold enrichment is the relative abundance of DNA fragments detected by real-time PCR at amplified region over a control amplified region. GST antibody was used as a ChIP control (grey). Error bars represent the s.e.m. of three technical replicates
(c) 3HA-Klf2, Klf4 and Klf5 occupy *Nanog* promoter region in ES cells. ChIP assay was performed using anti-HA antibody with extract derived from ES cells ectopically expressed 3HA-Klf2 (blue), 3HA-Klf4 (red) or 3HA-Klf5 (green). These proteins contain 3 copies of HA tag at the amino terminus. Control ChIP (grey) was performed using anti-HA antibody with extract derived from mouse ES cells. Error bars represent the s.e.m. of three technical replicates.
(d) 3HA-Klf2, Klf4 and Klf5 occupy *Esrrb, Fbxo15*, *Nanog*, *Tcl1* loci in ES cells. ChIP assay was performed using anti-HA antibody with extract derived from ES cells ectopically expressed 3HA-Klf2 (blue), 3HA-Klf4 (red) and 3HA-Klf5 (green). Control ChIP (grey) was performed with anti-HA antibody using mouse ES cells. Error bars represent the s.e.m. of three technical replicates.
(e) Redundancy of Klfs at target sites. ChIP assays were performed using anti-Klf2, anti-Klf4 and anti-Klf5 antibodies with extracts derived from ES cells transfected with control RNAi construct (grey bar) or ES cells transfected with Klf RNAi constructs (Klf2 RNAi, blue; Klf4 RNAi, red; Klf5 RNAi, green). Fold enrichment represents the abundance of enriched DNA fragments over a control region. GST ChIP served as mock ChIP. Data are presented as the mean ± s.e.m. and derived from three independent experiments (n = 3).

3.2.4 ChIP-on-chip assay for Nanog

Nanog was identified as an important transcription factor that is able to sustain pluripotency in mouse ES cells, even in the absence of LIF[125, 126]. As the binding profiles of the Klf proteins resembled that of Nanog at several known targets (such as *Nanog* and *Fbxo15*)[40] (**Figure 3.6)**, a Nanog ChIP-on-chip experiment was performed to compare the target sites of these four transcription factors.

The data showed that many genes encoding key regulators bound by Klfs such as *Pou5f1*, *Sox2*, *Nanog*, *Esrrb*, *Mycn* were also bound by Nanog. The ChIP-on-chip profiles of Klfs at the *Nanog*, *Esrrb*, *Fbxo15*, *Mycn, Pou5f1* loci were identical to those obtained from Nanog (**Figure 3.7a)**. The Klfs and Nanog co-occupy many common binding loci, and the data provide insight into the integration of the core Klf and Nanog circuitries.

ChIP-on-chip analysis also revealed that Klf2, Klf4 and Klf5 bound to certain genes involved in cell cycle regulation, DNA replication and cell proliferation such as *Cdkn1a*, *Pcna*, *Tcl1*, *and Mdm4* (**Figure 3.7b)**. It is of interest to note that *Cdkn1a* (*p21*) is also a known target of Klf4 in NIH3T3 and MEF cells[158]. These results suggest that Klfs have unique functions in regulation of cell growth in ES cells and differentiated cells.

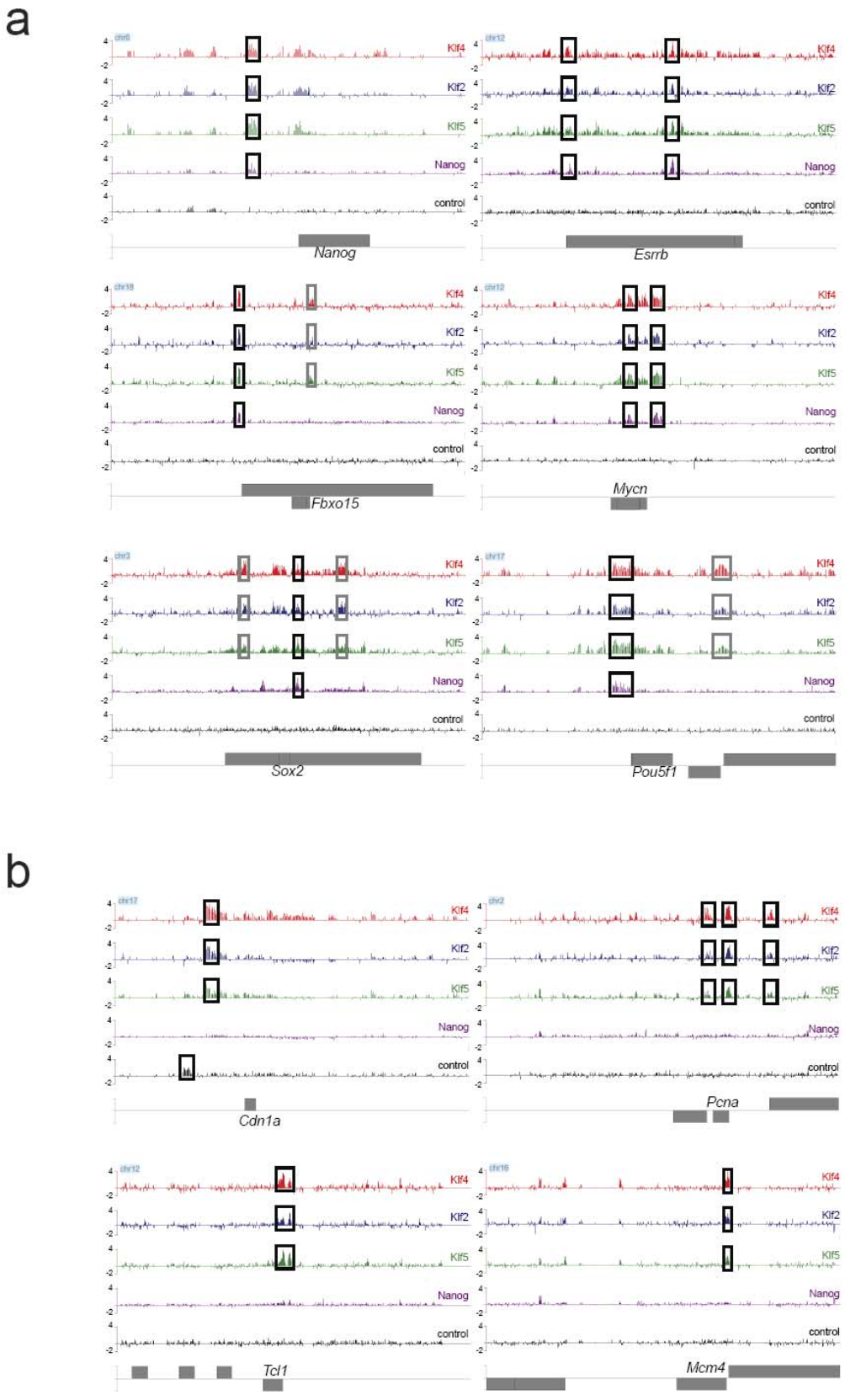

a
Klf4
Klf2
Klf5
Nanog
control
Nanog
Esrrb
Fbxo15
Mycn
Sox2
Pou5f1
b
Cdn1a
Pcna
Tcl1
Mcm4

Figure 3.7 ChIP-on-chip assay for Nanog
(a) Co-occupied sites for Klf2, Klf4, Klf5 and Nanog at *Nanog*, *Esrrb*, *Fbxo15, Mycn*, *Sox2* and *Pou5f1* genes. Screen shots of the SignalMap (NimbleGen) browser showing Klf4 (red), Klf2 (blue), Klf5 (green) and Nanog (purple) binding at *Nanog*, *Esrrb*, *Fbxo15, Mycn*, *Sox2* and *Pou5f1* loci. Y axis is Log_2Ratio. The peaks of Klf4, Klf2, Klf5 and Nanog binding are highlighted by open box. Control ChIP-on-chip (black) was done using p53 ChIP derived DNA. Genes in mouse genome are depicted as gray boxes.

(b) Klf2, Klf4 and Klf5 specific sites at *Cdn1a*, *Pcna, Mcm4* and *Tcl1* genes. Screen shots of the SignalMap (NimbleGen) browser showing Klf4, Klf2 and Klf5 binding at *Cdkn1a*, *Pcna*, *Mcm4* and *Tcl1* loci. Y axis is Log_2Ratio. The peaks of Klf4, Klf2 and Klf5 binding are highlighted by open box. Genes in mouse genome are depicted as gray boxes.

3.2.5 Analysis of ChIP-on-chip data for Klf2, Klf4, Klf5 and Nanog

The profiles of the three Klf proteins were strikingly similar. The overlaps between the genomic loci bound by all three factors were determined to be highly significant, with a *P* value of less than 0.00001. Thus, these Klfs were extensively colocalized to specific genomic regions (**Figure 3.8)**. We also noted that the Venn diagram did not overlap for a fraction of the Klf4 and Klf5 binding loci. Although the exact nature of this difference is not clear, it is possible that Klf4 and Klf5 also have divergent functions and exhibit different binding properties at certain genomic locations. Nevertheless, our results suggest that the three Klfs converge and work together to regulating common targets**.**

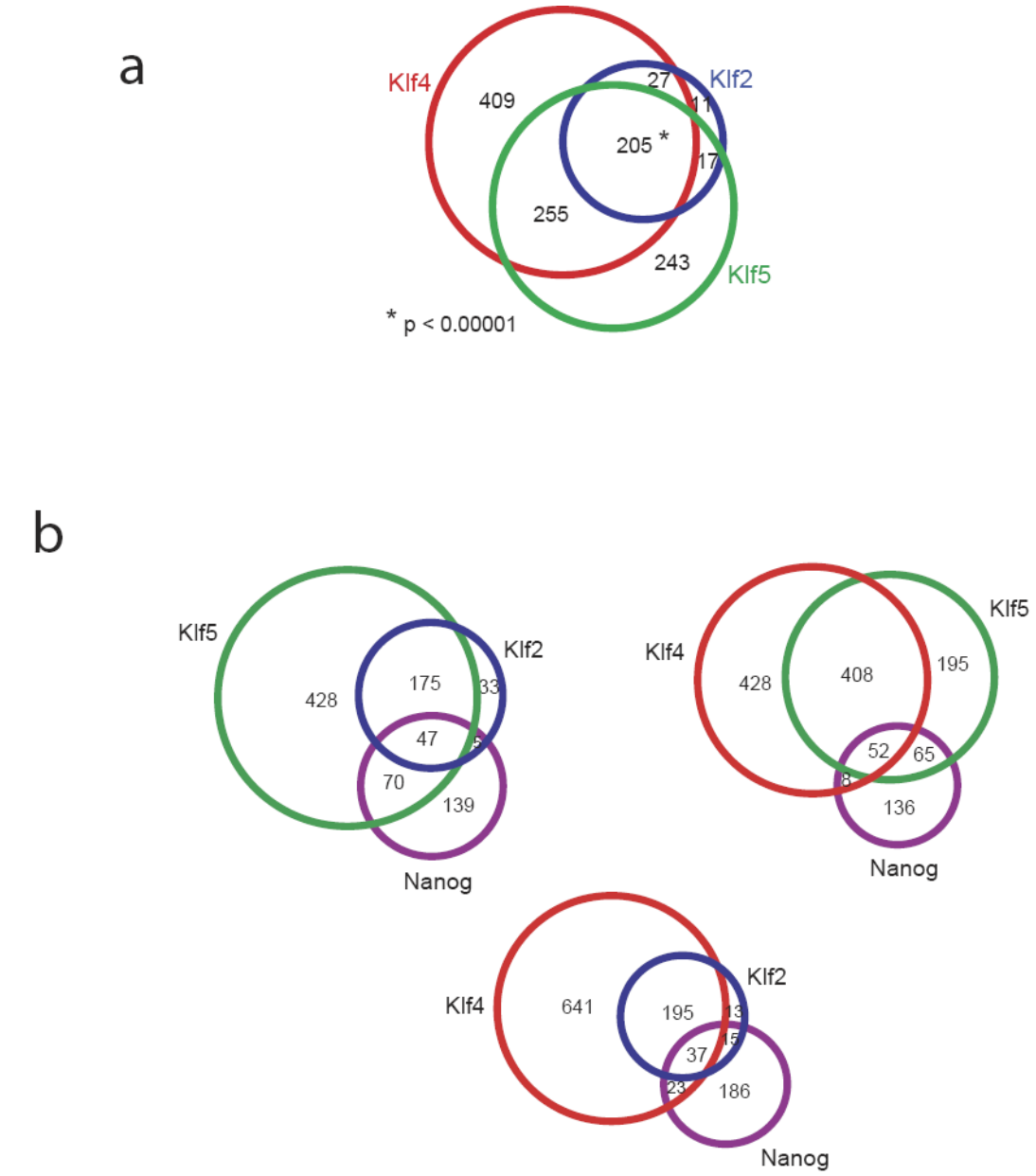


Figure 3.8 Comparison between Klf and Nanog binding profiles
(a) Venn diagram of the number of genomic loci bound by Klf2, Klf4 and Klf5. The diagram shows the overlap of Klf2 (blue), Klf4 (red) and Klf5 (green) binding loci in mouse ES cells. Of all the binding sites, 205 are the common genomic loci bound by Klf2, Klf4, and Klf5. The P value for the significance of the overlap is less than 0.00001 (asterisk), as determined by the Monte Carlo permutation test.
(b) Comparison between Klf and Nanog binding. Venn diagrams of the number of genomic loci bound by Klf2, Klf4, Klf5 and Nanog. These Venn diagrams show the overlap among of Klf2 (blue), Klf4 (red), Klf5 (green) and Nanog (purple) binding loci.

3.3 Klf2, 4 and 5 bind to similar regions in the *Nanog* promoter

To understand the molecular basis of the redundancy, DNA binding and transcriptional activation properties of the three Klfs were examined. Two Klf-binding loci were identified upstream of the *Nanog* coding region. We next investigated the role of Klf proteins in regulating the distal Klf-bound site that harbours enhancer activity. A series of 6 probes were used to scan a 140 bp peak region by EMSA for direct interaction with the Klf proteins (Table 3.2). The DNA binding domains of each Klf were expressed and purified as recombinant proteins. EMSA showed that a probe was bound by all three Klf proteins (**Figure 3.9a)**.

Every Klf family member harbors an evolutionary conserved three-zinc finger DNA binding domain at the C-terminal protein. It has been shown that Klf transcription factors bind specifically to GC and GT box with their DNA binding domain[208]. A putative Klf motif CCCACCC was found in the positive probe. To determine the critical nucleotides required for *in vitro* protein-DNA interactions, a series of point mutations replacing the CACCC sequence were introduced into the probe. EMSA revealed that mutagenesis of the first three nucleotides would abolish the protein-DNA interactions, while replacing the last two nucleotides would dramatically reduce their interactions (**Figure 3.9b)**. A slight difference of binding affinities to the mutated probe may account for the different binding circuitries among the three Klfs *in vivo*. Taken together, the data showed that the DNA binding properties of the Klfs are highly similar *in vitro*.

To address whether this site is functional in the context of an enhancer, *Nanog* enhancer genomic region including the Klf motif was cloned into the downstream of a Luciferase reporter containing a *Pou5f1* minimal promoter. Point mutations were introduced into the Klf motif and wild-type or mutant enhancer constructs were transfected into ES cells. Luciferase assays showed that wild-type *Nanog* enhancer region contained enhancer activity and the mutations of Klf motif abolished *Nanog* enhancer activity to the basal levels, indicating the intact Klf motif is required for the function of *Nanog* enhancer.

To reveal which Klf transcription factor is critical for *Nanog* enhancer activity, the *Nanog* enhancer construct was then cotransfected with shRNA expression constructs targeting each Klf. Knockdown of single Klf proteins did not affect the *Nanog* enhancer activity. However, depletion of Klf2, Klf4 and Klf5 by triple knockdown abolished the *Nanog* enhancer activity (**Figure 3.9c)**. Taken together, the data showed that the *Nanog* enhancer was dependent on the Klf motif and the three Klf proteins for its activity. Importantly, the downregulation of *Nanog* enhancer activity could be reversed by coexpression of RNAi-resistant *Klf2*, *Klf4* or *Klf5* cDNAs (**Figure 3.9d)**.

The enhancer assay was extended to examine ten other Klf-bound regions that are not found at proximal promoters. Fragments of 400–500 bp were cloned downstream of a *luciferase* reporter and transfected in ES cells. Seven constructs showed enhancer activities with at least fivefold more activity than a control reporter without insert. Six out of the seven enhancers were dependent on Klf2, Klf4 and Klf5, as triple knockdown reduced their activities. Again, the coexpression of RNAi-resistant *Klf2*, *Klf4* or *Klf5*

cDNAs rescued the downregulation of enhancer activities (**Data not shown**). Taken together, these experiments showed that the Klf2-, Klf4- and Klf5-occupied sites identified by ChIP may confer enhancer activity.

Table 3.2 Probes for screening Klf binding site on *Nanog* enhancer

Label	Sequence
A	biotin - agtccccgctccttttcagcactaaccatacaagttcatc
B	biotin - actaaccatacaagttcatcttttactcttgttttgactg
C	biotin - ttttactcttgttttgactgctaaccacccagaggaccca
D	biotin - ctaaccacccagaggacccacttaacattcctttccccac
E	biotin - cttaacattcctttccccacccacacagctagttccaacc
F	biotin - ccacacagctagttccaaccaggtctaaggcatctagacg

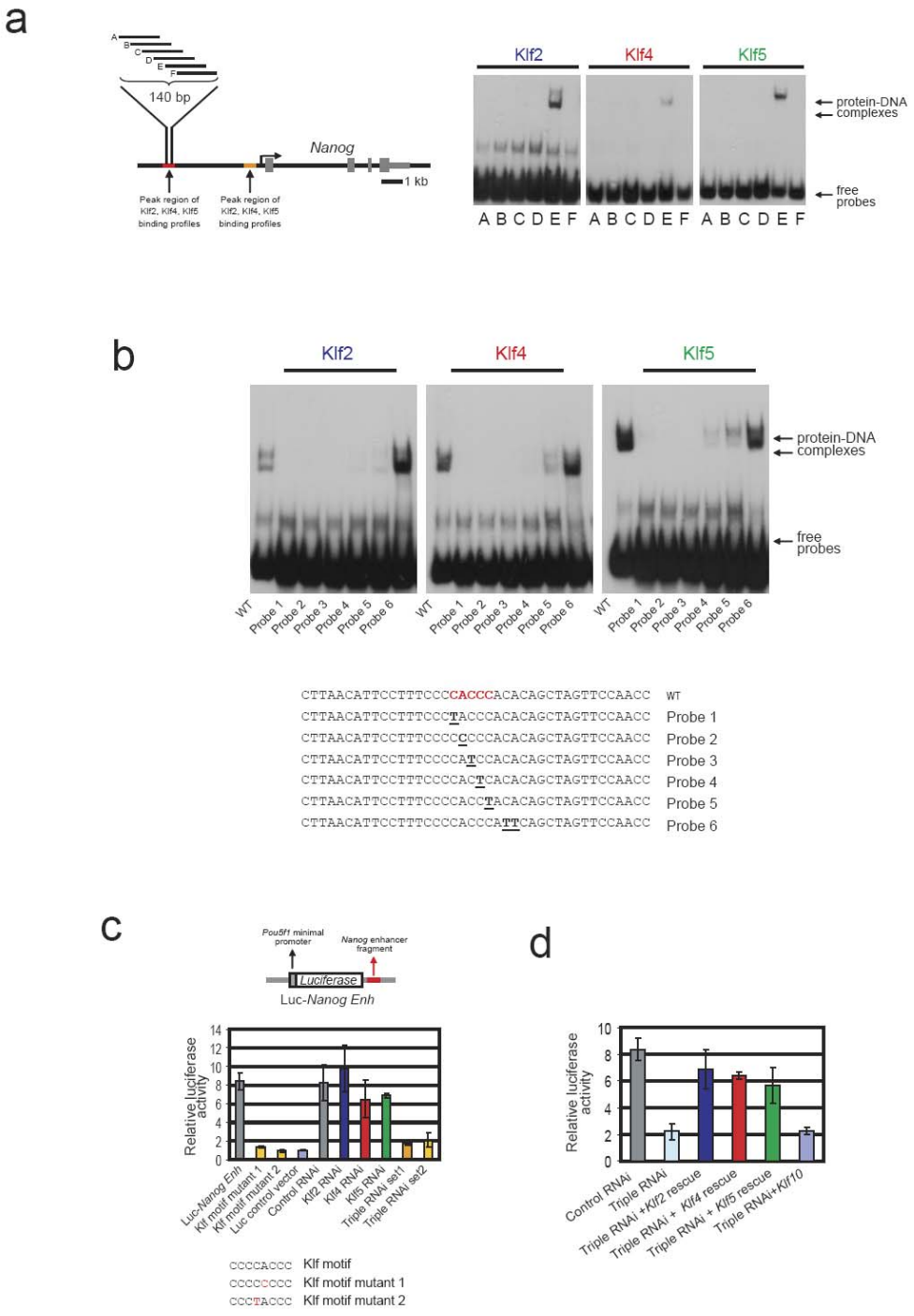

a
A
B
C
D
E
F
140 bp
Nanog
1 kb
Peak region of Klf2, Klf4, Klf5 binding profiles
Peak region of Klf2, Klf4, Klf5 binding profiles
Klf2
Klf4
Klf5
protein-DNA complexes
free probes
A B C D E F A B C D E F A B C D E F
b
Klf2
Klf4
Klf5
protein-DNA complexes
free probes
WT
Probe 1
Probe 2
Probe 3
Probe 4
Probe 5
Probe 6
CTTAACATTCCTTTCCCCACACCACACAGCTAGTTCCAACC WT
CTTAACATTCCTTTCCCTACCCACACAGCTAGTTCCAACC Probe 1
CTTAACATTCCTTTCCCCCCCACACAGCTAGTTCCAACC Probe 2
CTTAACATTCCTTTCCCCATCCACACAGCTAGTTCCAACC Probe 3
CTTAACATTCCTTTCCCCACTCACACAGCTAGTTCCAACC Probe 4
CTTAACATTCCTTTCCCCACCTACACAGCTAGTTCCAACC Probe 5
CTTAACATTCCTTTCCCCACCCATTCAGCTAGTTCCAACC Probe 6
c
Pou5f1 minimal promoter
Nanog enhancer fragment
Luciferase
Luc-Nanog Enh
Relative luciferase activity
Luc-Nanog Enh
Klf motif mutant 1
Klf motif mutant 2
Luc control vector
Control RNAi
Klf2 RNAi
Klf4 RNAi
Klf5 RNAi
Triple RNAi set1
Triple RNAi set2
CCCCACCC Klf motif
CCCCCCCC Klf motif mutant 1
CCCTACCC Klf motif mutant 2
d
Relative luciferase activity
Control RNAi
Triple RNAi
Triple RNAi +Klf2 rescue
Triple RNAi + Klf4 rescue
Triple RNAi + Klf5 rescue
Triple RNAi+Klf10

Figure 3.9 Klf2, 4 and 5 bind to similar regions in the *Nanog* promoter
(a) Klf2, Klf4 and Klf5 directly bind to *Nanog* enhancer region (red bar). EMSA was used to map the binding motif in the peak region of the binding profile. Double stranded probes (A–F) spanning 140 bp of the enhancer region were tested for DNA–protein interaction with the DNA binding domains of Klf2, Klf4 and Klf5.
(b) Mutagenesis of Klf motif. A conserved motif CACCC was identified in the screen. A series of single or double point mutations was used to study the binding interaction of DNA binding domain of Klf2, Klf4 and Klf5.
(c) *Nanog* enhancer is dependent on the Klf motif and the three Klfs for its activity. Schematic of *Nanog* enhancer luciferase reporter construct used to measure the effect of motif mutations and *Klf* shRNAs. The effect of Klf motif mutations on *Nanog* enhancer was tested by transfecting constructs point mutation at the Klf motif in mouse ES cells. Effect of *Klf* shRNAs on *Nanog* enhancer activity was tested by cotransfecting each RNAi construct along with the Luc-*Nanog* enhancer construct into mouse ES cells. Luciferase activity was measured at day 2 after transfection. Data are presented as the mean ± s.e.m. and derived from three independent experiments ($n = 3$).
(d) Rescue of *Nanog* enhancer activity with coexpression of RNAi-resistant *Klf2*, *Klf4* or *Klf5* cDNAs. RNAi-resistant *Klf2*, *Klf4*, *Klf5* or *Klf10* expression constructs were cotransfected with triple *Klf2*, *Klf4* and *Klf5* knockdown and *Nanog* enhancer constructs into mouse ES cells. Data are presented as the mean ± s.e.m. and derived from three independent experiments ($n = 3$).

3.4 Gene regulation by Klf2, 4 and 5

As concurrent depletion of the three Klf proteins by RNAi led to differentiation, DNA microarray experiments were performed to capture the transcriptome dynamics at different times after transfection of shRNA expression constructs. At day 2, the morphology of the ES cells depleted of Klf2, Klf4 and Klf5 was similar to ES cells transfected with control plasmid expressing shRNA against *luciferase*; these cells also stained positive for alkaline phosphatase. Despite maintaining these ES cell characteristics, the levels of ES cell-associated transcripts (such as *Nanog*, *Tcl1*, *BMP4* and *Stra8*) were already reduced. Expression of the ectoderm markers *Fgf5*, *Nes* and *Cxcl12* was induced at this early time point. The data suggest that Klf2, Klf4 and Klf5 may suppress differentiation along specific lineages. At day 4 and day 6, the Klf-depleted ES cells were already differentiated, with distinct morphology and loss of alkaline

phosphatase staining. At day 6, a strong induction of the differentiation markers *T* and *Cdx2* was observed in addition to a continual increase in *Fgf5*, *Nes* and *Msx1* (**Figure 3.10a)**.

Interestingly, the ChIP-on-chip data revealed a weak signal of Klf2, Klf4 and Klf5 occupancies at the *Fgf5* and *Nes* promoters. This binding was confirmed using ChIP–qPCR assays (Table 3.3). Expression analysis showed that triple knockdown of the Klf proteins induced higher expression than single or double knockdown (**Figure 3.10b)**. However, endoderm markers (*Gata4* and *Gata6*) were not induced, indicating that the cells differentiate into specific lineages.

To further understand the gene regulation by Klf, the expression of selected genes was analyzed by RT-PCR. The expression of ES upregulated genes *Nanog*, *Tbx3* and *Esrrb* was restored and induction of lineage marker *Nes* and *Fgf5* was repressed by coexpression of RNAi-resistant *Klf2*, *Klf4* or *Klf5* cDNAs in Klf triple knockdown experiment (**Figure 3.10c)**. Triple knockdown caused more change in expression compared to single or double knockdown (**Figure 3.10d)**. Thus, the data shows that Klf2, Klf4 and Klf5 can positively regulate the expression of many self-renewal genes and potentially repress lineage specific genes.

Recent works have generated pluripotent cell-lines from the epiblast of mouse embryos[28, 29]. These epiblast stem cells (EpiSCs) are pluripotent and express ES cell markers such as Oct4, Sox2 and Nanog. However, the EpiSCs are deficient in a subset of well known ES

cell-associated transcripts such as *Zfp42*, *Dazl*, *Stra8* and *Pecam1* and express a high level of *Fgf5*. The expression pattern of these five genes in EpiSCs is consistent with their profiles obtained from the triple Klf depleted ES cells. The microarray profiling data for different EpiSCs were obtained from published literature and we plotted the expression values of 30 genes in ES cells, EpiSCs and triple Klfs knockdown cells as heatmaps (**Figure 3.10e)**. Interestingly, the expression of *Klf2*, *Klf4* and *Klf5* was down-regulated in EpiSCs compared to ES cells. In addition to *Zfp42*, *Dazl*, *Stra8* and *Pecam1*, the expression of *Tbx3*, *Esrrb*, *Tcl1*, *Nr5a2*, and *Piwil2* was also found to be down-regulated in EpiSCs. The levels of these transcripts were also reduced after depletion of the three Klfs in ES cells.

Global expression profiling of ES cells and different EpiSCs reveal a group of ES cell-specific transcripts, which are up-regulated in ES cells but not EpiSCs. A statistical analysis was performed to test whether the genes down-regulated by triple knockdown are associated with the ES cell-specific genes (defined as genes up-regulated in ES cells but not in EpiSC). In this test for association, the genes up-regulated in ES cells were divided into two groups: genes commonly up-regulated in both pluripotent cells and ES cell-specific genes. The former were genes being differentially up-regulated in both ES cells and EpiSCs as compared to differentiated cells, while the latter were genes up-regulated only in ES cells, including the three Klfs. Depletion of the three Klfs downregulated 11.75% of ES cell-specific genes, but merely 3.75% of the ES cell-EpiSC common genes (**Figure 3.10f)**. A p-value of 5e-20 was derived from Fisher's Exact Test, indicating a very strong association between triple knockdown suppressed genes and ES

cell-specific genes. The data indicate that *Klf2*, *Klf4* and *Klf5* are downregulated with a set of genes in EpiSCs. These genes are also downregulated upon depletion of the three *Klfs* in ES cells. We suggests that Klfs are not generally required for the maintenance of self-renewal of pluripotent cells such as EpiSCs and their roles may be specific to ES cells. The role of these Klf proteins in EpiSCs has yet to be examined.

It is also of interest that the expression of *Klf2*, *Klf4* and *Klf5* are induced during early embryogenesis, as measured by Taqman real-time PCR assays (**Figure 3.10g)**. The *in vivo* function of these three Klf proteins during early embryogenesis remains to be demonstrated. In particular, given the observation that EpiSCs are deficient in these three proteins, it is not clear whether these or other Klf-family members have a critical role in regulating pluripotency.

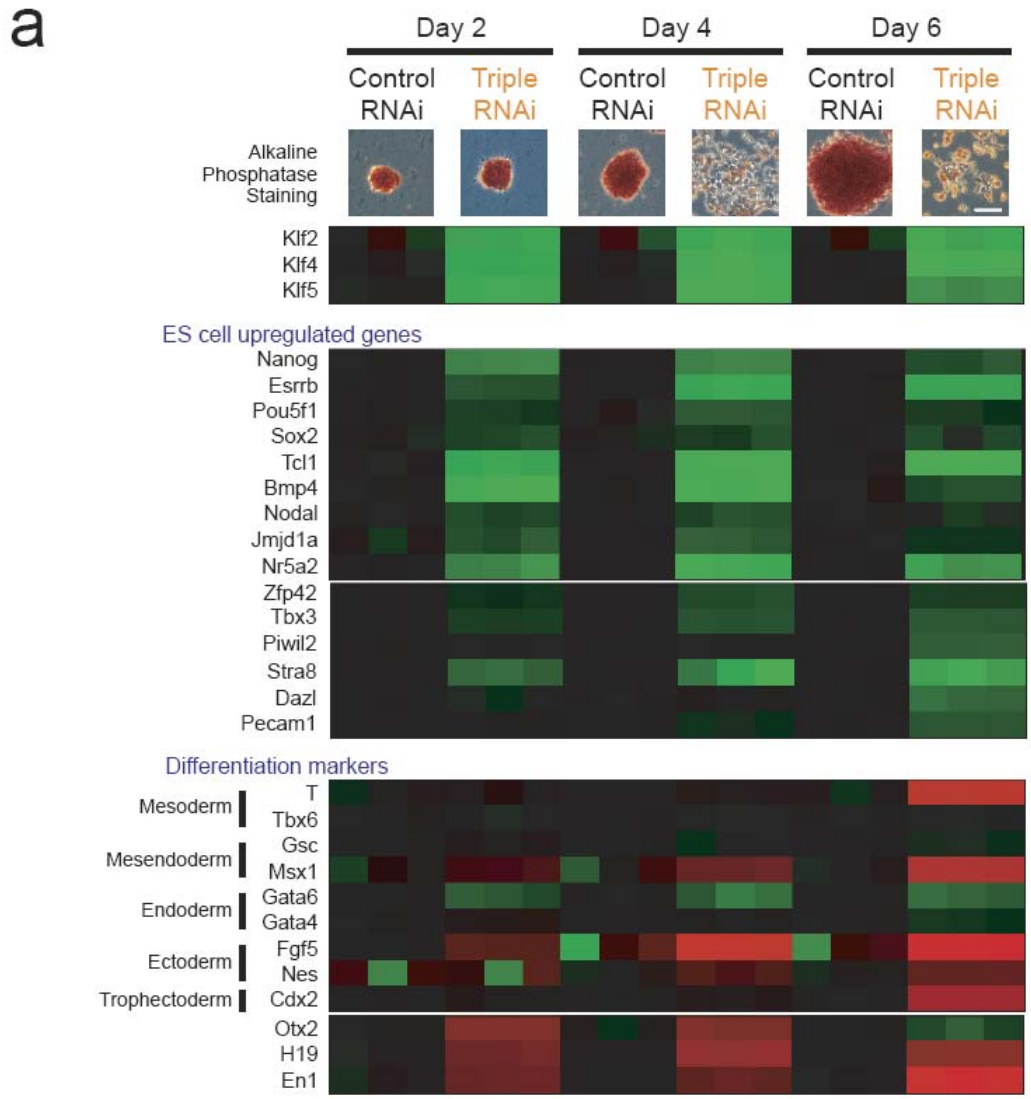
Table 3.3 ChIP primers for *Nes* and *Fgf5*

Nes

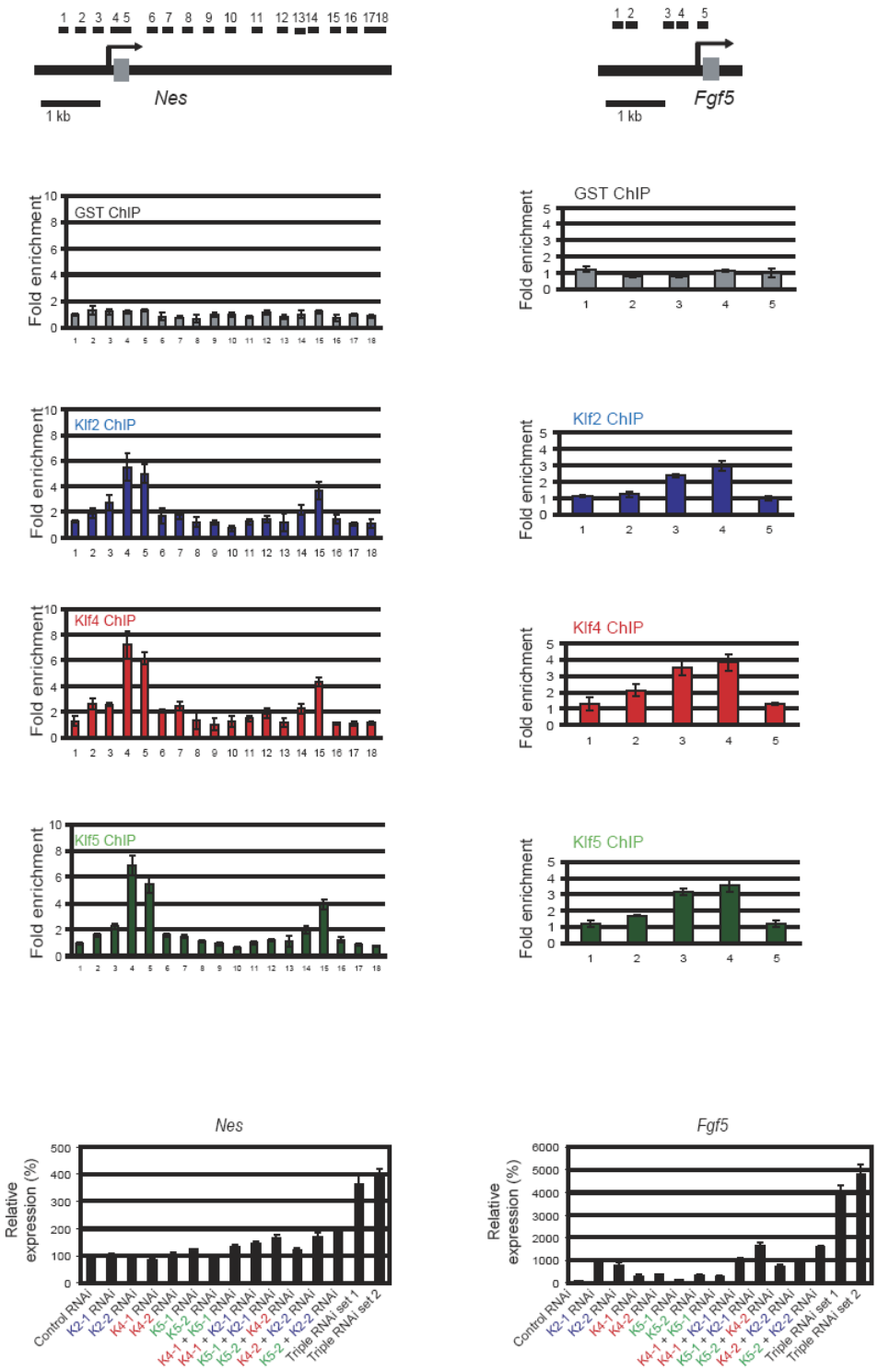
Label	Sense	Antisense
1	TCTACAGCTTCATTTCTTTCTTCGCCAAAC	CCGCACCAGGAAAAGGAGACTAGCAAGG
2	TTACTACCCAGGCTGGTGTCAATCCTAC	AAACGGACGGTGCAGTGTTTTGTGTATAAC
3	TGCACCGTCCGTTTTTCCAACAGTTCACGA	AGACAGCCCAGAACACCCTAAGATGCAC
4	TGATGCAGGGACCCGGGCTGTGTGTTGCAC	GAAAGGGAGTAGCGGCGACAGTGACCCAG
5	GCTTCCGCTGGGTCACTGTCGCCGCTACT	CCACATCTGAAAAGATTCTTCCCCGACGCA
6	GCTCGGGAGAGTCGCTTAGAGGTGCAGCA	CCTAGGGCTTGGTTTTACCAGGGACAGGGT
7	GATGGATGGAATTCTGAGACCTCCTAGCAC	GGGCACCTTCTGATTGACAACTGCATAGT
8	GCCCAGACAGAAGTCCCGACGTGACCAC	AGGAGAGGAGATGGGTTACACAGCGGAAGA
9	GGAACAGGACTCTGCCCTACAAAGATGTGA	CCCACCTTCCAGGATCTGAGCGATCTGAC
10	ATGTTTTTGTCTGTGATGGAGGCTTATG	TTTCAAAGAACTCGAGAGCAGGTGACTGTA
11	AGCCCGGGCCTCAGCTGTGTCTGGGGTAAG	CTCAACAGTGGGACTCCAGCCGGGACCTCT
12	CCAGAGAAGGGAAGCCAGTGAGTGTCTAGG	ATCAGTTCACAAGGGATTCGGAGCTAGTC
13	AAGGGTCCAGGCAGCTCTGAGGAATGTAAG	GCTAATTATTGCCTTTTGTCCACACTACCA
14	TAGTGTGGACAAAAGGCAATAATTAGCA	TATACCCACAGATATATACTTGGTGAACCC
15	GACTGTGTGGCTTTCATACGACAATGGTCA	CTTCCAGGTGTCTGCAAGCGAGAGTTC
16	CTCTCCTTCCACATCTTCCCTAAGCA	AAACACTGTAAGTATGTCCTGACTTCACCC
17	GCTGAAGCTGCATTTCCTTGGGATACCAG	TGGGCTCTGACCTCTGCATTTTTAGGATAG
18	ATCCTTCTCTTCCCCCTTGCCTAATACCCT	GGAATGGGGGACATCCTGGGCTCTGACCTC

Fgf5

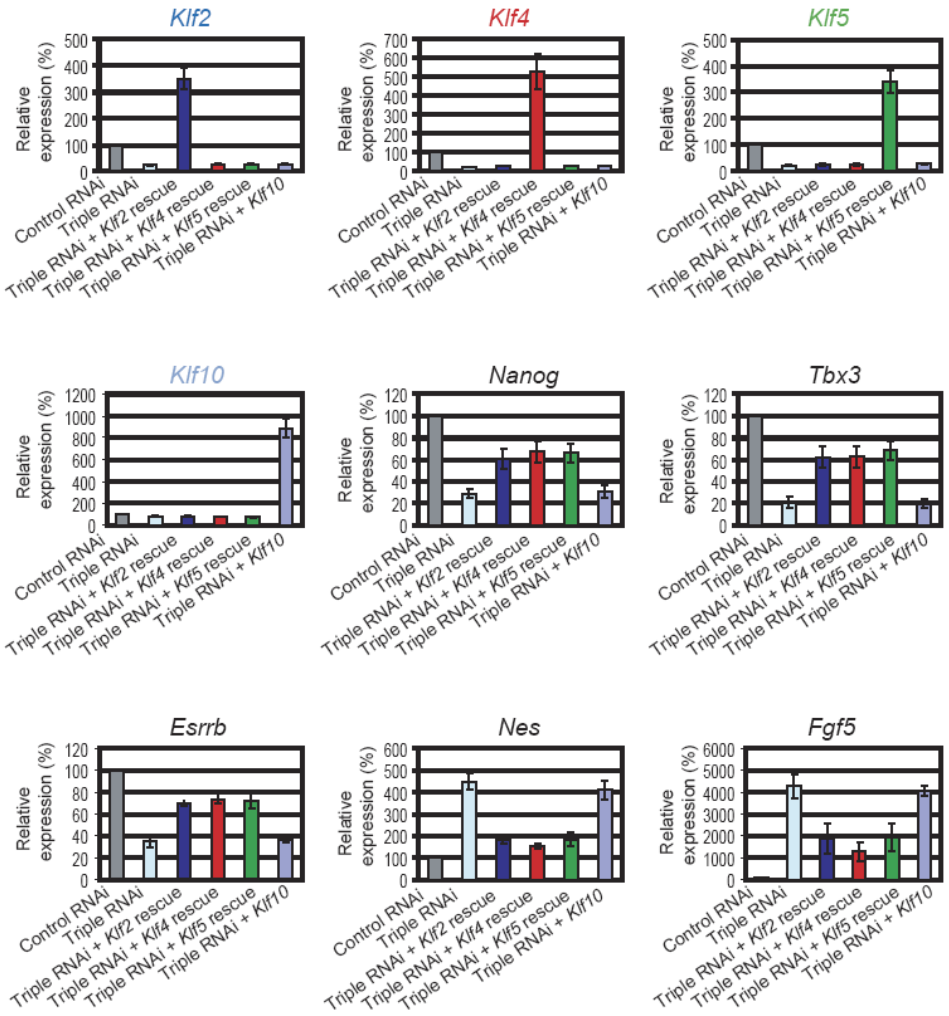
Label	Sense	Antisense
1	TTCTCAGATAAATAACAGCCCCAGCATTAC	CCCATGATCCGTTGCCTGGAGGAGTTAGCC
2	GGCTCTCTAGTGGAGAAATAAACGCACACC	TGTGTCCCAGCCAACTCACTCTTTACCCAG
3	GGCGGGTGAGGGGAAGCTTCGCAGGCGTGC	GCATCTTGCAGGGCTGGCTCCGGGCTCCGT
4	AGATCACTGGCGTTATAAATATCCCGGTGC	CGGGGGTCCTAAGTGCATCTTGCAGGGCT
5	TGATCCACAGCGCTTGGGCTCACGGGGAGA	CCTTGGCTGCCCGGAGAAGCTGCGACTGGT


a
Day 2
Day 4
Day 6
Control RNAi
Triple RNAi
Control RNAi
Triple RNAi
Control RNAi
Triple RNAi
Alkaline Phosphatase Staining
Klf2
Klf4
Klf5
ES cell upregulated genes
Nanog
Esrrb
Pou5f1
Sox2
Tcl1
Bmp4
Nodal
Jmjd1a
Nr5a2
Zfp42
Tbx3
Piwil2
Stra8
Dazl
Pecam1
Differentiation markers
Mesoderm
T
Tbx6
Mesendoderm
Gsc
Msx1
Endoderm
Gata6
Gata4
Ectoderm
Fgf5
Nes
Trophectoderm
Cdx2
Otx2
H19
En1

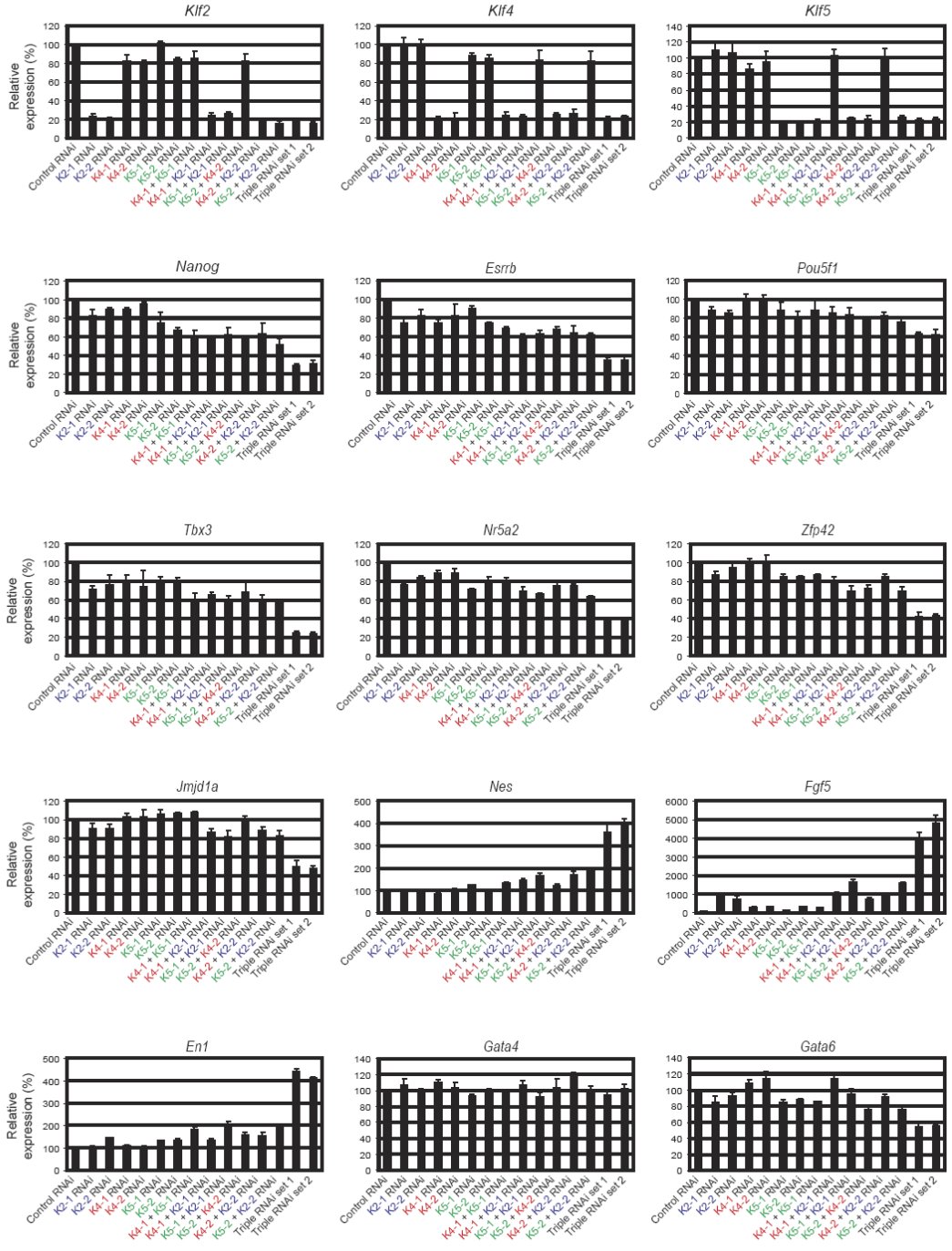
b

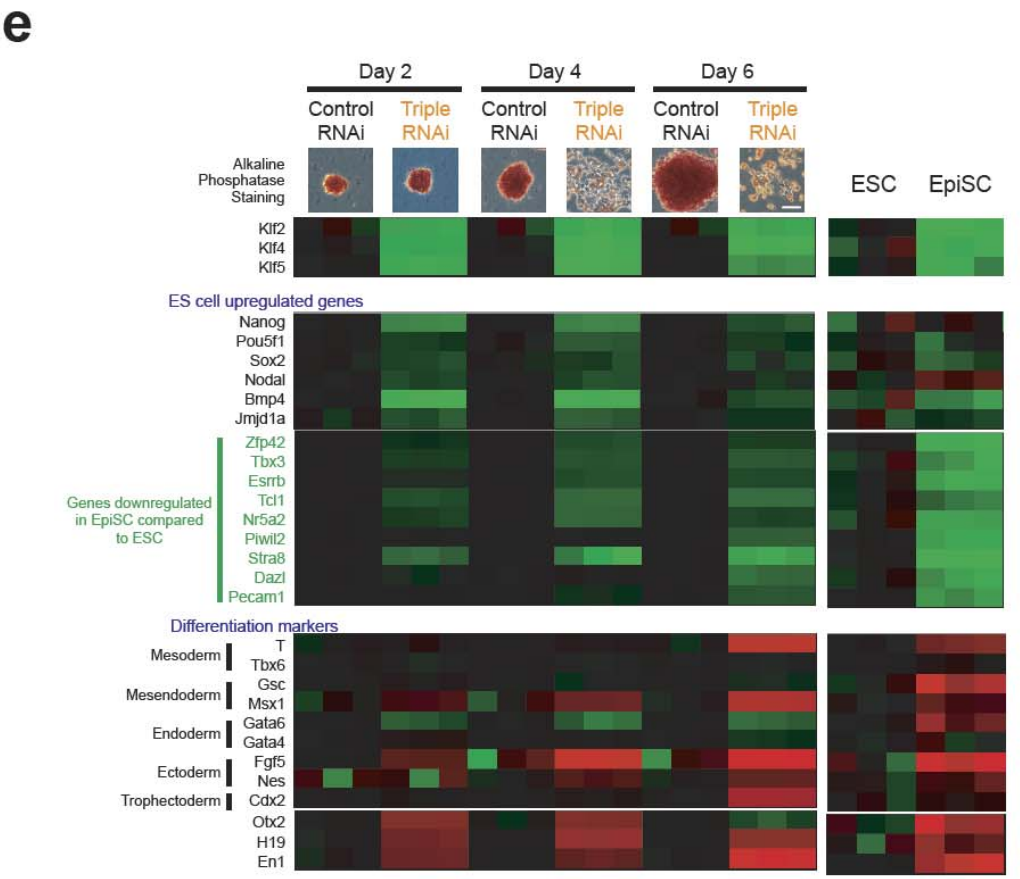

Nes
1 kb
Fgf5
1 kb
GST ChIP
Klf2 ChIP
Klf4 ChIP
Klf5 ChIP
Fold enrichment
Nes
Fgf5
Relative expression (%)
Control RNAi
K2-1 RNAi
K2-2 RNAi
K4-1 RNAi
K4-2 RNAi
K5-1 RNAi
K5-2 RNAi
K4-1 + K5-1 RNAi
K4-1 + K2-1 RNAi
K5-1 + K2-1 RNAi
K4-2 + K2-2 RNAi
K5-2 + K2-2 RNAi
Triple RNAi set 1
Triple RNAi set 2

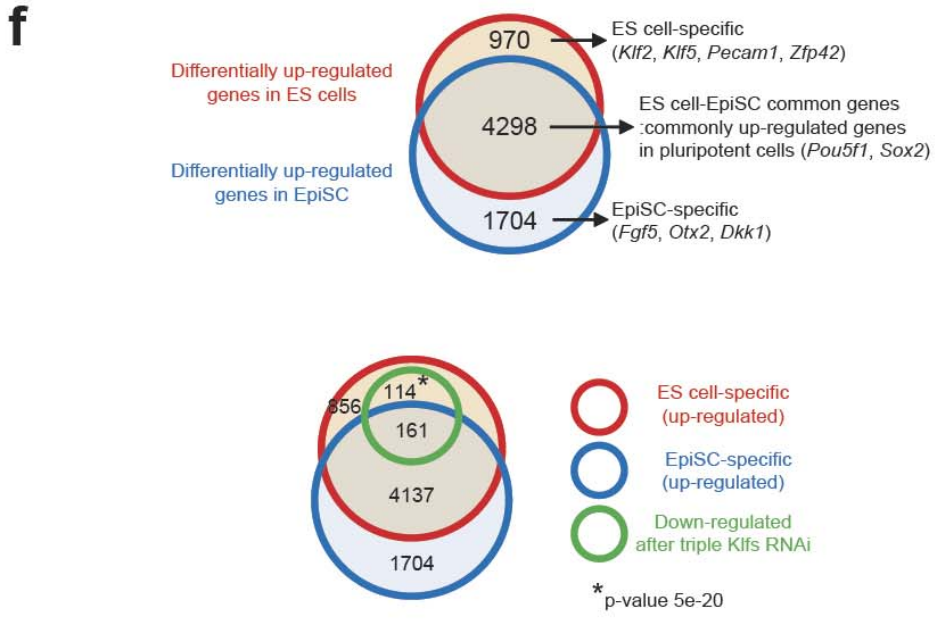
C


Klf2
Klf4
Klf5
Klf10
Nanog
Tbx3
Esrrb
Nes
Fgf5
Relative expression (%)
Control RNAi
Triple RNAi
Triple RNAi + Klf2 rescue
Triple RNAi + Klf4 rescue
Triple RNAi + Klf5 rescue
Triple RNAi + Klf10

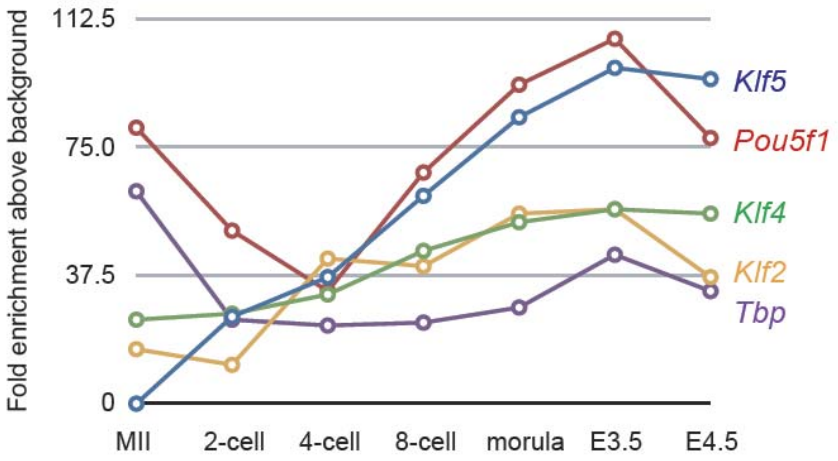
d


Klf2
Klf4
Klf5
Nanog
Esrrb
Pou5f1
Tbx3
Nr5a2
Zfp42
Jmjd1a
Nes
Fgf5
En1
Gata4
Gata6
Relative expression (%)
Control RNAi
K2-1 RNAi
K2-2 RNAi
K4-1 RNAi
K4-2 RNAi
K5-1 RNAi
K5-2 RNAi
K4-1 + K5-1 RNAi
K4-1 + K2-1 RNAi
K5-1 + K2-1 RNAi
K5-2 + K4-2 RNAi
K4-2 + K2-2 RNAi
K5-2 + K2-2 RNAi
Triple RNAi set 1
Triple RNAi set 2


e
Day 2
Day 4
Day 6
Control RNAi
Triple RNAi
Control RNAi
Triple RNAi
Control RNAi
Triple RNAi
Alkaline Phosphatase Staining
ESC
EpiSC
Klf2
Klf4
Klf5
ES cell upregulated genes
Nanog
Pou5f1
Sox2
Nodal
Bmp4
Jmjd1a
Genes downregulated in EpiSC compared to ESC
Zfp42
Tbx3
Esrrb
Tcl1
Nr5a2
Piwil2
Stra8
Dazl
Pecam1
Differentiation markers
Mesoderm
T
Tbx6
Mesendoderm
Gsc
Msx1
Endoderm
Gata6
Gata4
Ectoderm
Fgf5
Nes
Trophectoderm
Cdx2
Otx2
H19
En1


f
Differentially up-regulated genes in ES cells
Differentially up-regulated genes in EpiSC
970
4298
1704
ES cell-specific (Klf2, Klf5, Pecam1, Zfp42)
ES cell-EpiSC common genes :commonly up-regulated genes in pluripotent cells (Pou5f1, Sox2)
EpiSC-specific (Fgf5, Otx2, Dkk1)
856
114*
161
4137
1704
ES cell-specific (up-regulated)
EpiSC-specific (up-regulated)
Down-regulated after triple Klfs RNAi
*p-value 5e-20

g


112.5
75.0
37.5
0
Fold enrichment above background
MII
2-cell
4-cell
8-cell
morula
E3.5
E4.5
Klf5
Pou5f1
Klf4
Klf2
Tbp

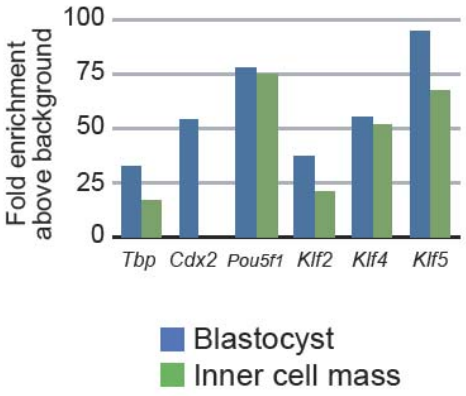

100
75
50
25
0
Fold enrichment above background
Tbp
Cdx2
Pou5f1
Klf2
Klf4
Klf5
Blastocyst
Inner cell mass

Figure 3.10 Gene regulation by Klf2, Klf4 and Klf5
(a) Time course microarray analyses were performed to measure gene expression changes at different days after *Klf2*, *Klf4* and *Klf5* triple knockdown. The morphology and alkaline phosphatase staining are shown for each time point for both the control and Klf depleted cells. The scale bar represents 100 μm. Note that no difference in morphology and alkaline phosphatase staining was detected at day 2, indicating that the ES cells were undifferentiated. Microarray heatmaps depicting expression changes of selected ES cell-associated and differentiation marker genes at different days are shown. Red indicates increased expression compared to control samples, whereas green means decreased expression. The genes expression levels were mean centred to show their relative change. (Done by Yuin-Han Loh and Ching-Aeng Lim)
(b) Klf2, Klf4, Klf5 bind to *Nes and Fgf5* promoter regions and regulate their expression. ChIP assay was performed using anti-Klf2 (blue), anti-Klf4 (red), anti-Klf5 (green) antibody with extracts derived from ES cells. Schematic showing the location of the amplicons (black bars labeled 1-18) used to detect ChIP enriched fragment over the 6 kb *Nes* promoter and the location of the amplicons (black bars labeled1-5) used to detect ChIP-enriched fragment over the 1.8 kb *Fgf5* promoter. Amplicons are numbered in order relative to their sites along the gene. The grey box represents an exon. Fold enrichment is the relative abundance of DNA fragments detected by realtime PCR at amplified region over a control amplified region. GST antibody was used as a ChIP control (grey). Error bars represent the s.e.m. of three technical replicates. Expression of *Nes* and *Fgf5* is induced upon triple knockdown. Gene expression changes after single, double and triple *Klfs* RNAi. Quantitative real-time PCR analysis of gene expression changes after single, double and triple knockdown of *Klf2*, *Klf4* and *Klf5*. The levels of the transcripts were normalized against control *Luc* shRNA transfection. The RNA was harvested from the cells at two days after transfection. Error bars represent the s.e.m. of three replicates.
(c) Gene expression analysis after *Klf2*, *Klf4* and *Klf5* rescue. RNAi-resistant *Klf2*, *Klf4*, *Klf5* or *Klf10* expression constructs were cotransfected with triple *Klf2*, *Klf4* and *Klf5* knockdown construct into mouse ES cells. The expression of *Nanog*, *Tbx3*, *Esrrb*, *Nes* and *Fgf5* was determined by quantitative real-time PCR analysis. The levels of the transcripts were normalized against control *Luc* shRNA transfection. Data are presented as the mean ± s.e.m. and derived from independent experiments ($n = 3$).
(d) Gene expression changes after single, double and triple *Klfs* knockdown. Quantitative real-time PCR analysis of gene expression changes after single, double and triple knockdown of *Klf2*, *Klf4* and *Klf5*. The levels of the transcripts were normalized against control *Luc* shRNA transfection. The RNA was harvested from the cells at two days after transfection. Error bars represent the s.e.m. of three technical replicates. *Nanog*, *Esrrb*, *Pou5f1*, *Tbx3*, *Nr5a2*, *Zfp42* and *Jmjd1a* are up-regulated in ES cells. *Nes*, *Fgf5*, *En1*, *Gata4* and *Gata6* are down-regulated in ES cells.
(e) Comparison between gene expression profiles of triple knockdown ES cells, ES cells and EpiSCs. The microarray profiling data for different EpiSCs were obtained from published literature and we plotted the expression values of 30 genes in ES cells, EpiSCs and triple Klfs knockdown cells as heatmaps. The scale bar represents 100 μm. Microarray heatmaps (left panel) depicting expression changes of selected ES cell-associated and differentiation marker genes at different days are shown. Microarray heatmaps (right panel) depicting expression profile changes between triple knockdown

ES cells (day 2) and epiblast stem cell are shown. genes expression levels were mean centered to show their relative change. (Done by Ching-Aeng Lim)

(f) The Venn diagram shows the overlap of differentially up-regulated genes in ES cells and EpiSCs. DNA microarray datasets for ES cells and EpiSCs were obtained from literature (GEO dataset GSE7866). To obtain differentially up-regulated genes in ES cells (red circle) or EpiSCs (blue circle), we compared these microarray data with those obtained from mouse embryonic fibroblast (MEF) (GSM189664, GSM189667 of GEO dataset GSE7815). Statistical analysis of microarray (SAM) was used. Genes with q-value of less than 0.01 and more than 1.5 fold change were selected. This analysis reveals genes specifically upregulated in ES cells (denotes as ES cell-specific) and EpiSCs (denotes as EpiSC-specific). The overlap represents genes commonly up-regulated in both pluripotent cells. A statistical test was performed to test whether the genes down-regulated by triple knockdown are associated with the ES cell-specific genes (defined as genes up-regulated in ES cells but not in EpiSC). In this test for association, the genes up-regulated in ES cells were divided into two groups: genes commonly up-regulated in both pluripotent cells and ES cell-specific genes. The former were genes being differentially up-regulated in both ES cells and EpiSCs as compared to differentiated cells, while the latter were genes up-regulated in ES cells only, including the three Klfs. (Done by Ching-Aeng Lim)

(g) Expression within the preimplantation embryo. Embryo equivalent amounts of total RNA from freshly harvested embryos at the developmental stages indicated were analyzed for *Klf2*, *Klf4*, and *Klf5* transcript levels by reverse transcription and TaqMan real-time PCR. TaqMan realtime PCR analysis of their expression through preimplantation development indicate *Klf2* and *Klf4* are present as maternal transcripts and detected throughout the preimplantation period. *Klf5* expression is not detected in the mature oocyte but levels in the 2-cell embryo indicate dramatic and early transcriptional activation of this gene. *Pou5f1* and *Tbp* were included as reference genes. Values from technical duplicates are expressed as fold change above background with background set at a threshold cycle (Ct) value of 36 and a decrease in 1 Ct calculated as a 1.9-fold increase in expression. E3.5 and E4.5 are blastocysts at 3.5 and 4.5 days post-fertilization, respectively. Real-time PCR comparison of expression levels between embryo-equivalent amounts of total RNA from whole blastocysts and immunosurgically dissected inner cell masses (ICM), both at 4.5 days post-fertilization. All three Klfs are expressed in the ICM. Note the lack of detection within the ICM for the TE-specific *Cdx2* and virtually equivalent amounts of *Pou5f1* between whole blastocysts and ICMs for this ICM-specific transcript. (Done by Guo-Qing Tong)

3.5 Klf2, Klf5 are reprogramming factors

3.5.1 Reprogrammed cells induced by Klf2, 4, 5

To test whether Klf2 and Klf5 also function as reprogramming factors, the standard reprogramming assay were performed by using fibroblasts harbouring an endogenous *Pou5f1-GFP* reporter[209]. Oct4 encoded by *Pou5f1* is a marker that is silenced in MEFs or other differentiated cells and only induced at the later stage of reprogramming process. Reprogramming of MEFs is monitored by the reactivation of the silenced *Pou5f1-GFP* reporter (**Figure 3.11a**). After infection with the retroviruses, the number of GFP positive colonies on 14 dpi (days post infection) were quantified. Consistent with our findings on the functional redundancy between Klf2, Klf4 and Klf5 in ES cells, Klf2 and Klf5 in conjunctions with other three reprogramming factors can replace Klf4 in inducing GFP positive colonies that resemble ES cells and the similar result was also demonstrated in another independent study[210] (**Figure 3.11b**).

Klf2 showed relative higher reprogramming ability compared to Klf4 according to the number of GFP positive colonies. However, the number of GFP positive colonies generated using Klf5 was much lower than Klf4 by introduction together with Oct4, Sox2 and c-Myc. The similar results are obtained from an independent study[210]. The reason for the lower efficiency of reprogramming by Klf5 compared to Klf4 is not clear. Amino-acid sequence alignment of DNA binding domains has revealed greater structural similarities between Klf2 and Klf4 (**Figure 3.11c**), and may explain the stronger functional redundancies of these two family members in reprogramming. We also noted that the Venn diagram did not overlap for a fraction of the Klf4 and Klf5 binding loci. It

is possible that Klf4 and Klf5 may exhibit different binding properties at certain genomic locations and have divergent functions. Our RNAi knockdown, ChIP-on-chip, reporter assay and EMSA results strongly indicate that Klf2, 4, 5 show a redundant role in regulation of ES cell self-renewal. This redundant role also implies their function in reprogramming.

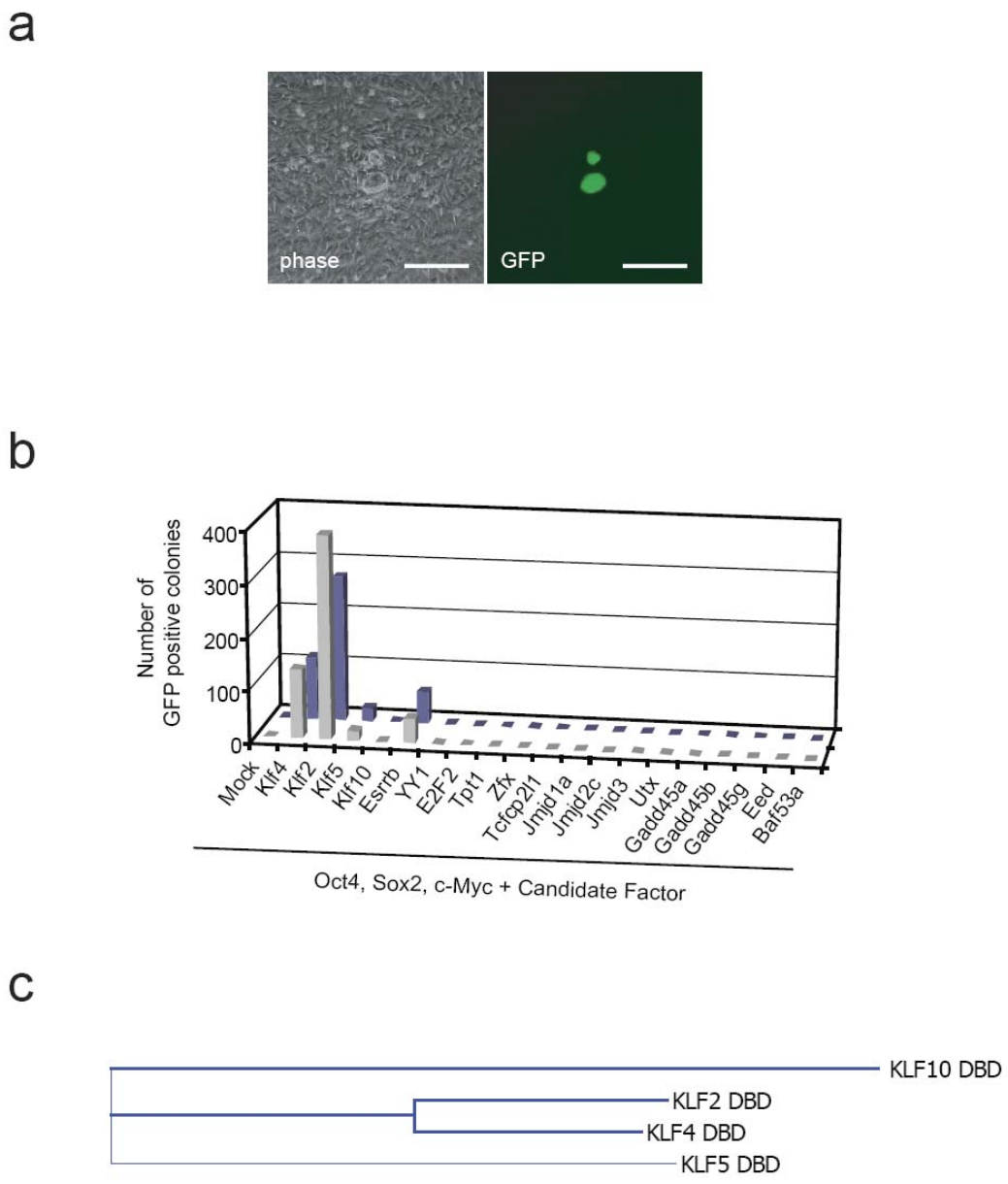


Figure 3.11 Identification of Klf2 and Klf5 as a reprogramming factor
(a) Reprogrammed cells induced by defined factors from the MEFs harboring an integrated *Pou5f1- GFP* reporter. Bright field image is shown. GFP positive signal indicates that the expression of endogenous *Pou5f1* was restored specifically in the reprogrammed cells, but not in the surrounding fibroblasts. The scale bars represent 200 μm.
(b) Candidate factors were screened by co-transduction with Oct4, Sox2 and c-Myc in *Pou5f1-GFP* MEFs. At 14 dpi, the number of GFP-positive colonies was recorded and used as an indicator of reprogramming efficiency. The results of two independent experiments are shown.
(c) DNA binding domain similarity tree of Klf2, 4, 5 and10. DNA binding domains of Klf2, Klf4 and Klf5 were analysis by Software AlignX.

3.5.2 Validation of Klf2 reprogrammed cells

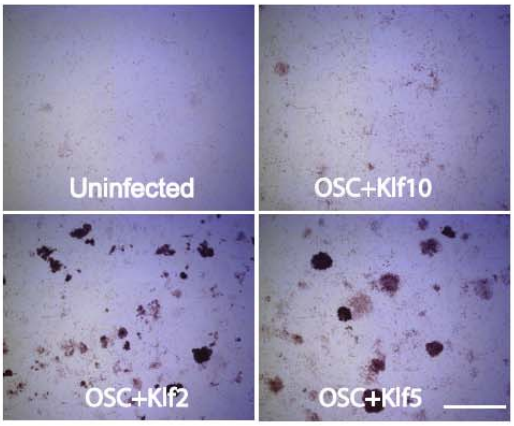
MEFs with *GFP* under the control of a constitutively active *actin* promoter were used for the standard reprogramming assay. Constitutive expression of GFP allows us to trace these cells in the microinjection assays. Klf2 or Klf5, with Oct4, Sox2 and c−Myc, was able to induce the formation of ES-like cells that stained positive for alkaline phosphatase (**Figure 3.12a**). However Klf10, another Klf family member could not produce the ES-like cells.

As it has been demonstrated that c-Myc is not essential for reprogramming[210, 211], *actin-GFP* MEFs were transduced with Oct4, Sox2 and Klf2 (OSK2), and ES-like cells could be obtained. The established OSK2 line was expressing similar *Oct4*, *Sox2*, *Nanog* and *Zfp42* levels compared to one ES cell line (**Figure 3.12b**). Nakagawa *et al.* showed that Klf2 can replace Klf4 in inducing iPS cells which are able to give rise to teratomas in nude mice[210].

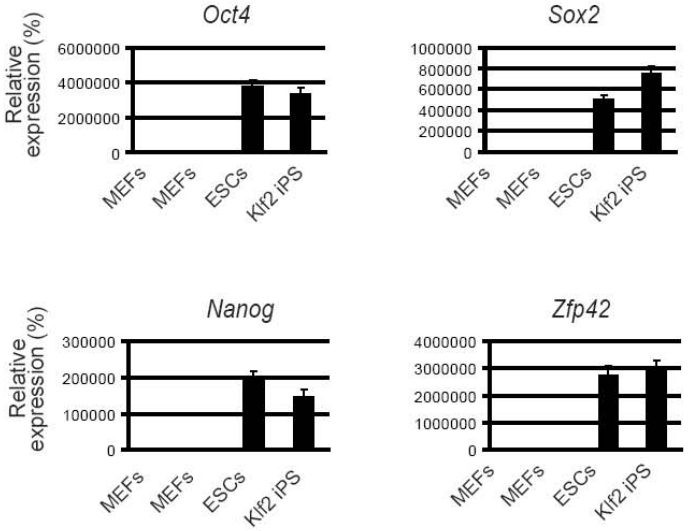
To further validate the pluripotency of Klf2 reprogrammed cells, *in vivo* test was carried out. After injection into 8-cell embryos, the OSK2 reprogrammed cells contributed extensively to the resulting chimeric mouse embryos, as the E9.5 embryos showed widespread incorporation of GFP-labelled cells (**Figure 3.12c**). Immunohistochemical analysis demonstrated that chimeric embryo contained GFP positive cells which were widely and extensively distributed among all tissues and organs and were represented in tissues derived from all three major germ layers (ectoderm, mesoderm and endoderm) of the developing embryo (**Figure 3.12d**).

Further studies need to be carried out to determine whether these OSK2 reprogrammed cells can contribute to the germ lineage and show germline transmission. Taken together, these results demonstrate some Klf family proteins are capable of inducing iPS cells from somatic MEFs.

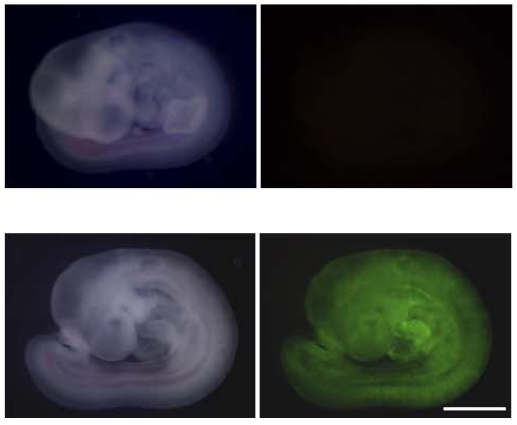
a


Uninfected
OSC+Klf10
OSC+Klf2
OSC+Klf5

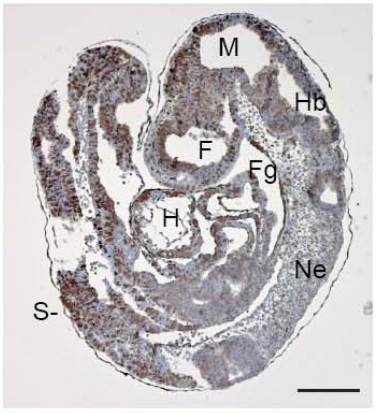
b


Oct4
Sox2
Nanog
Zfp42
Relative expression (%)
6000000
4000000
2000000
0
1000000
800000
600000
400000
200000
300000
100000
MEFs
MEFs
ESCs
Klf2 iPS

c



d


M
Hb
F
Fg
H
Ne
S-

Fiugre 3.12 Klf2 reprograms MEFs with Oct4 and Sox2.
(a) Alkaline phosphatase expression in Klf2 and Klf5 reprogrammed cells.
(b) Quantitative real-time PCR analysis of gene expression of *Oct4*, *Sox2*, *Nanog* and *Zfp42* in Klf2 reprogrammed cells. The levels of the transcripts were normalized against two control MEFs (129 and CD1). Error bars represent the s.e.m. of three technical replicates.
(c) Embryos derived from the 8-cell stage embryos injected with Klf2 reprogrammed cells are shown. The embryos were collected at E9.5. GFP fluorescence images for the embryos shown (right two panels). The embryos were observed directly under stero-microcope for GFP expression. Wild-type embryo shows no GFP fluorescence (upper panes).
(d) Immunohistochemical analysis of parasagittal sections to show the distribution of GFP-positive cells in embryo. The section was probed with anti-GFP antibody (B-2, Santa Cruz) and counterstained with haematoxylin. Abbreviations: F, forebrain; Fg, foregut diverticum; H, heart, Hb, hindbrain; M, midbrain; Ne, neuroepithelium; O, otic vesicle; S, somite. (Done by Lai-ping Yaw)

3.6 Integration of the core Klf circuitry with the Nanog transcriptional regulatory network

Klf2, Klf4 and Klf5 are shown to be functionally redundant and they exhibit similar binding properties in ES cells. Collectively, the bound targets of these Klf proteins constitute the core Klf circuitry. The Klfs and Nanog co-occupy many common binding loci, and the data provide insight into the integration of the core Klf and Nanog circuitries (**Figure 3.13a**).

Klf2, Klf4 and Klf5 are coordinately upregulated in ES cells, thereby permitting the expression of a subset of ES cell-specific genes. As each Klf was observed to bind to their respective gene, the coordinated regulation of Klf proteins may, in part, be mediated through autoregulatory loops commonly found among key regulators in ES cells (**Figure 3.13b**). The data also show that these regulators form a highly interconnected transcriptional regulatory network: one example is the extensive cotargeting of the two

regulatory loci of Nanog by transcription factors important to maintain ES cells (**Figure 3.13c**). Although our study only investigated the roles of Klf2, Klf4 and Klf5, it is possible that other members of the Klf family may play similar roles in ES cells.

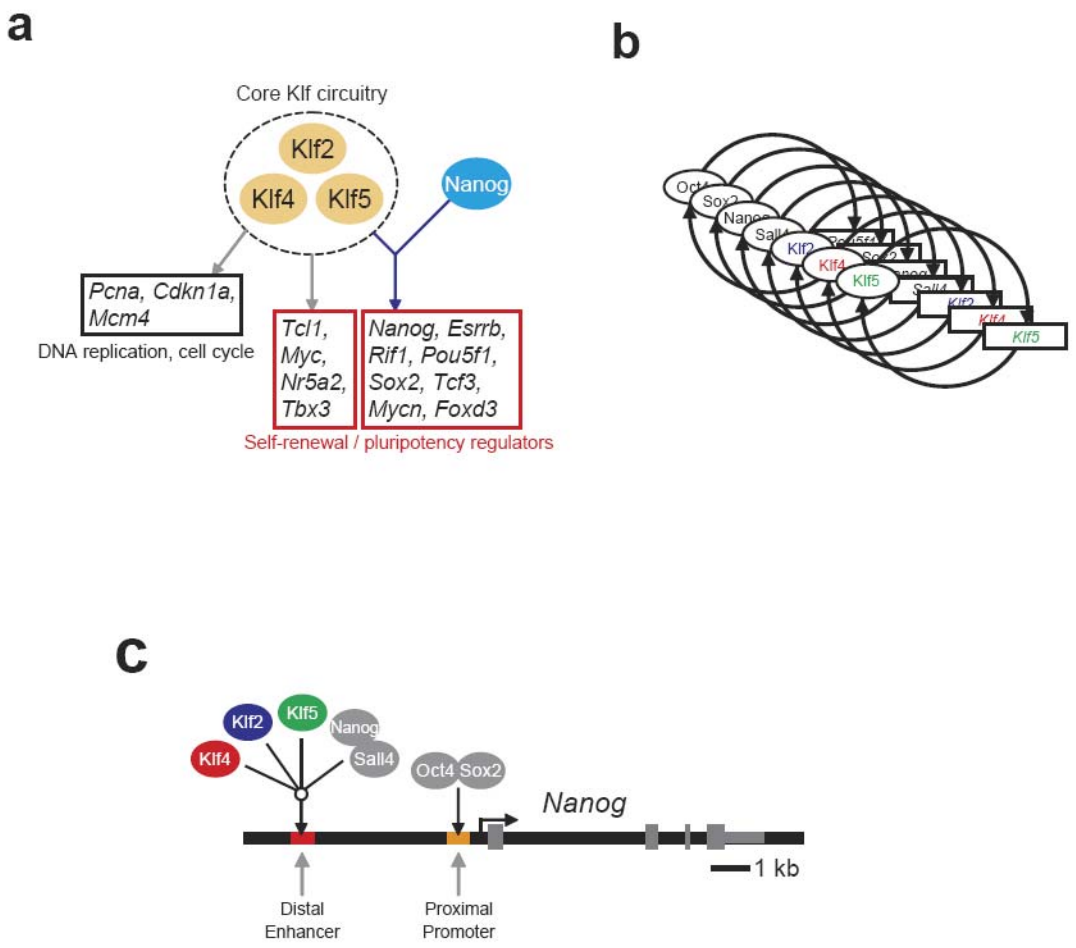


Figure 3.13 transcriptional regulatory network in ES cells
(a) Schematic representation of a model for the transcriptional regulatory network targeted by Klf2, Klf4, Klf5 and Nanog. Klf2, Klf4 and Klf5 target genes involved in DNA replication and selfrenewal are shown. Common targets of Klfs and Nanog are genes required for self-renewal and maintenance of pluripotency of ES cells (red boxes).
(b) Autoregulatory loops formed by different transcription regulators in ES cells. The genes are presented in the open boxes and transcription factors are represented by ovals. Oct4, Sox2, Nanog and Sall4 have previously been shown to have autoregulatory configurations.
(c) Regulatory regions of *Nanog* receive multiple inputs. Two key regulatory regions (enhancer and proximal promoter) are shown as red and orange bars, respectively. Klf2, Klf4 and Klf5 direct bind to *Nanog* enhancer. The *Nanog* enhancer is also targeted by Nanog and Sall4. Oct4 and Sox2 occupy the *Nanog* proximal promoter.

CHAPTER IV

DISCUSSION

CHAPTER IV DISCUSSION

4.1 Klf4, Klf2 and Klf5 have diverse functions in cell proliferation and differentiation and animal development

The critical functions of Klf4, Klf2 and Klf5 are indicated in many studies. Klf4 can serve as a tumor suppressor or an oncogene in a context-dependent manner[212]. Forced expression of Klf4 in cultured cells results in the inhibition of DNA synthesis and cell cycle processes[158, 213]. Klf4 null mice die shortly after birth and show an impaired differentiation in the skin and in the colon[147, 148]. Loss of Klf4 causes altered proliferation and differentiation and precancerous changes in the adult stomach of mice[214]. Within the hematopoietic system, Klf4 is a critical regulator in the transcriptional network controlling monocyte differentiation[215]. These results indicate that Klf4 plays a crucial role in cell proliferation and differentiation.

Klf4 is shown to be a downstream target of LIF signaling pathway in ES cells as its expression is rapidly induced by supplement of cytokine LIF and the inactivation of Stat3 markedly decreases *Klf4* expression[161, 185, 216]. Ectopic expression of K*lf4* in ES cells promotes their capacity to self-renew based on secondary embryoid body (EB) formation and restored expression of *Oct4*[161]. Interestingly, *Klf4* overexpression supports ES growth in the absence of LIF[162, 216]. This function may link to Nanog as overexpression of *Nanog* enables ES cells to undergo self-renewal in the absence of LIF[125, 126].

Klf5 is shown to be a mediator of mouse fibroblast cell proliferation in the various *in vivo* and *in vitro* tests, and transforms the cells possibly through the H-Ras signaling

pathway[217, 218]. In acute lymphoblastic leukemia, KLF5 interacts with tumor suppressor p53 in regulating the expression of the inhibitor-of-apoptosis gene *survivin*, which may play a role in pathological process of cancer[219]. The expression of *KLF5* is lower and deletion of *KLF5* is often detected in various breast cancer cell lines[220]. *KLF5* is frequently deleted and down-regulated but rarely mutated in prostate cancer[221]. These results indicate Klf5 play roles in tumor progresses. Klf5 also play diverse roles in the development of many tissues. Klf5 is shown to play a pivotal role in the pathogenesis of cardiovascular disease as *Klf5*+/- mice show diminished levels of arterial-wall thickening, angiogenesis, cardiac hypertrophy and interstitial fibrosis in response to injury[149]. Neonatal heterozygous *Klf5* knockout mice exhibit a marked deficiency in white adipose tissue development, suggesting that Klf5 is also required for adipogenesis[222]. Transgenic mice conditionally deleted Klf5 in respiratory epithelial cells die of respiratory distress immediately after birth, indicating Klf5 is required for perinatal lung morphogenesis and function[223].

Klf5 null embryos fail to implant due to defects in trophectoderm development and die before E6.5 days[149, 224]. The expression of *Klf5* initiates at the 2 cell-stage, and Klf5 protein is detected in both ICM and trophoblast. Cdx2, an important transcription factor required for trophoblast development, is downregulated in *Klf5* knockout embryos which probably accounts for the abnormal phenotype in early embryo development. *Klf5* -/- ES cells can not be derived from *Klf5* knockout ICM cells directly, but can be established from Klf5 heterozygous ES cells. Overexpression of *Klf5* can overcome the requirement

of mouse ES cells for LIF[224]. These results indicate Klf5 have specific functions in early embryo development and ES cell self-renewal.

Klf2 is expressed in lung, endothelial cells and lymphocytes, and is essential for normal blood-vessel integrity and lung development[145, 225, 226]. Deficiency in *Klf2* leads to a massive loss of the peripheral T-cell pool, suggesting Klf2 regulates T-cell quiescence and survival[145]. *Klf2*-deficient thymocytes also show impaired expression of several receptors required for thymocyte emigration and peripheral trafficking, indicating Klf2 is essential for T-cell trafficking[227]. *Klf2* null ES cells grow normal in spite of their derived MEF cells having adipogenesis defects[228]. Klf2 also is shown to support LIF-independent self-renewal[162].

Although Klf2, Klf4 and Klf5 have the diverse roles in cell proliferation and differentiation, and animal development, their detailed roles in ES cells are not appreciated.

4.2 Redundancy between family members

To study the roles of Klf members in the stem cells, we use a combination of different assays. We show that Klf4 alone is dispensable for the self-renewal state of ES cells, while a group of Klfs, including Klf2, Klf4 and Klf5, are required for maintaining the undifferentiating state of ES cells. Simultaneously triple depletion of these three Klfs leads to differentiation, and this phenotype can be rescued by any of the three Klfs, strongly indicating a redundant role among these Klfs in ES cells. To dissect the detailed

mechanisms under the phenotype, ChIP-on-chip analysis is performed and reveals that the binding profiles of Klf2, Klf4 and Klf5 are strikingly similar. In another genomic-wide mapping study, about 41.9% of Klf4 binding loci with multiple transcription factor binding loci (MTL) are associated with binding of core transcription factors Oct4, Sox2 and Nanog[113]. For the three Klf common binding locus in Nanog enhancer region, Klf2, Klf4 and Klf5 bind to the same motif and all three Klfs are required for the enhancer activities. The expression of *Klf2*, *Klf4* and *Klf5* is enriched in ES cells and highly associated with that of many important self-renewal and pluripotent regulators, such as *Nanog*, *Esrrb*, *Tcl1*, and *Tbx3*. These results highlight that Klf2, Klf4 and Klf5 are required to maintain the unique transcriptional regulatory network and expression program for ES cells.

Our results demonstrate the functional redundancy between Klf2, Klf4 and Klf5 in ES cells. There are some clues indicating their possible redundant functions. The expression of all three *Klf*s is highly upregulated in ES cells, and downregulated upon the differentiation. The expression of three *Klf*s is induced during the early development and shows a similar pattern to *Oct4* induction. The DNA binding domains of these three Klfs are conserved and highly similar which indicate they can bind to similar DNA sequence. All three Klfs are shown to support LIF-independent self-renewal of ES cells when ectopicly expressed[162, 224].

The family members have been shown to play redundant roles for each other in some cases. For instance, c-*Myc* and n-*Myc* genes have broad and divergent expression patterns

in development. When c-Myc coding sequences are replaced with n-myc coding sequences, n-Myc is capable of replacing most of the essential c-Myc functions required for embryonic development and for proliferation of differentiated cells[229]. The different physiological roles of c-Myc and n-myc relate primarily to differences in transcriptional regulation of the genes and their spatial and temporal expression, rather than distinct biochemical properties of the n-Myc and c-Myc proteins.

The redundancy issues are not be widely explored in the stem cell field. The core circuitry of stem cell intrinsic transcriptional regulatory network are shown to be indispensable as Oct4, Sox2 and Nanog play unique roles in maintaining the self-renewal and pluripotent state of ES cells. Knockout or depletion of any of them causes early embryo development defect *in vivo* or impairment of self-renewal and pluripotency *in vitro*[40, 99, 102, 120, 124, 126]. The Oct or Sox family has several members respectively. However, their possible redundancy roles in the early development and ES cells have not been appreciated. Apart from the core trinity factors, other factors also have been shown to guard the unique properties of ES cells, including Zfp143, Esrrb, and Sall4. The study of redundant roles of family members, especially through loss-of-functions, is hindered by many factors, such as the difficulty of raising the reagents, time-consuming experiments and unknown functional correlation between different family members.

The redundant function between different Klf family members was also tested by factor-induced reprogramming experiments whose results showed that Klf1, Klf2 or Klf5 can replace Klf4 in inducing iPS cells in the conjunction with other three reprogramming

factors, Oct4, Sox2 and c-Myc[210]. The redundant functions are largely explored, especially in reprogramming. c-Myc can be replaced by n-Myc or l-myc[210, 230]. Sox2 is shown to be substituted by several family members, such as Sox1, Sox3[210]. We independently show that Klf4 can be replaced by Klf2 or Klf5 in a reprogramming assay[231]. However, all these reprogramming experiments are gain-of-function analysis which means that one of family members will sufficiently accomplish the required biological functions.

It is of interest that the expression of *Klf2*, *Klf4* and *Klf5* is induced during early embryogenesis, as measured by a Taqman real-time PCR assay. The *in vivo* function of these three Klf proteins during early embryogenesis remains to be demonstrated. The multiple knockout of Klf2, Klf4 and Klf5 experiment will provide us more valuable information on their functions. It is interesting to know whether these three Klf can functionally replace each other *in vivo*. The knockin studies can be performed to elucidate this issue.

The functional redundancy can also be satisfied by different family members. When introduced into *Sox2* deficient ES cells, Oct4 is shown to rescue the *Sox2* null caused differentiation phenotype[124]. A possible explanation is the key role of Sox2 is to maintain the expression of *Oct4* in pluripotent ES cells. Nanog overexpression can substitute for the depletion of *Esrrb*, *Tbx3, Tcl1,* or *Dppa4* probably through compensatory upregulation of other self-renewal components[132]. ES cells with overexpression of *Ronin* are capable of forming self-renewal colonies after downregulation of Oct4. These results

suggest that complex interconnected transcriptional regulatory networks or protein interaction networks are integrated to form a unique framework in maintaining the self-renewal of ES cells. The loss of some components can be substituted by compensational expression of functionally redundant family members or other regulators.

4.3 The reprogramming process and reprogramming factors

Successful reprogramming approaches such as SCNT and cell fusion indicate that oocytes or embryonic stem cells harbor key factors that can induce pluripotency. It seems that these factors also play important roles in the maintenance of pluripotency. Based on this hypothesis, 24 candidate factors were selected for testing in a reprogramming assay. The results show that retrovirus introduction of four transcription factors, Oct4, Sox2, c-Myc and Klf4, is sufficient to induce ES-like cells from somatic mouse fibroblast cells[56]. Further studies confirm that these induced pluripotent stem (iPS) cells are similar to ES cell and fulfilled the entire criteria tests for authentic pluripotent cells[58, 232, 233]. Apart from fibroblasts, B lymphocytes, hepatocytes and gastric epithelial cells can also be converted to iPS cells using the same four transcription factors[234, 235]. iPS cells are tested to be able to correct the genetic disease in mice and give rise to specialized cell types, indicating their promising practical applications in regenerative medicine[197].

Among the four reprogramming factors, c-Myc is shown to be a downstream target of LIF signaling pathways to promote the self-renewal of ES cells[166]. However, c-Myc is shown to be dispensable in forming the iPS cells and the omission of the c-Myc retrovirus significantly reduced the risk of tumorigenicity in chimeras[210]. Oct4 and Sox2

are critical for the pluripotent state of cells *in vitro* and *in vivo* and are indispensable in generating iPS cells from somatic cells. Neural stem cells endogenously express *Sox2*, *c-Myc*, and *Klf4* as well as several intermediate reprogramming markers, thus Oct4 alone is sufficient to produce the pluripotent stem cells despite the efficiency being much lower compared with the two factors (Oct4-Klf4 or Oct4-cMyc)[236]. However, Oct4, Sox2, and Klf4 are still the indispensable factors for inducing authentic pluripotency from mouse somatic cells.

Human iPS cells can be derived by transduction of Oct4, Sox2, Klf4 and c-Myc or Oct4, Sox2, Nanog and Lin28 into human somatic cells[232, 237]. In fact, Oct4 and Sox2 are sufficient to generate human iPS cells and other factors are able to increase the efficiency of induction. However, it is argued that human iPS cells and human ES cells are equal to their mouse counterparts due to the lack of stringent tests of pluripotency *in vivo* for human cells. Evidence has been shown that human ES cells are more akin to mouse EpiSCs which are derived from post-implantation embryos. EpiSCs differ from ES cells by upregulation of the postimplantation markers *Fgf5* and *T* [28, 29]. Female EpiSCs, but not ES cells, contain a silent X chromosome which is a definite epigenetic distinction between early and late epiblast. Recently, Guo *et al* reported that when cultured in EpiSCs conditions, mouse ES cells can readily differentiate to EpiSCs. Interestingly, when introduced with Klf4 and cultured in ground stage conditions, EpiSCs can generate iPS cells which are referred to as EpiSCs-iPS cells. These cells are able to generate high-contribution chimaeras that yield germline transmission, indicative of authentic pluripotency[185]. These results indicate the Klf4 can reprogram the cells into a full

pluripotent state. The detailed roles of Klf4 to build up the unique transcriptional regulatory network and reset the epigenetic landmarks specific for ES cells still need to be explored. It is also of interest to know whether Klf2 and Klf5 show similar functions.

The detailed mechanisms of reprogramming factors that build up pluripotency in somatic cells are not clear. It seems that the reprogramming process by defined transcription factors is different from traditional approaches such as SCNT and cell fusion. During iPS cell generation, exogenous reprogramming factors are required for about 10 days. Upregulaiton of alkaline phosphatase and stage-specific embryonic antigen 1 (SSEA1) occur at early time points. Expression of endogenous *Oct4*, *Sox2*, *Nanog* genes and reactivation of the silent X chromosome, marking fully reprogrammed cells, is only observed late in the process[238, 239]. These results suggest that transcription factor-induced reprogramming is a gradual process. It is shown that reprogramming of the somatic cell genome begins during the first cell cycle after the fusion of a somatic cell with a pluripotent cell. Reactivation of *Oct4*-GFP transgene expression is detected on day 2 (around 44–48 hours) after fusion of embryonic stem (ES) cells with neurosphere cells[240, 241]. The transfer of the genome of a differentiated cell into an unfertilized oocyte transforms the transcriptional program of the somatic cell into an embryonic transcriptional program in a remarkably short time period. This may come from the difference that cell fusion and SCNT can take advantage of the presence of all players in a complex network, whereas factor-induced reprogramming must start up to induce other genes to form a complete transcriptional regulatory network[242].

4.4 Klf and Esrrb are functionally related

According to the global expression analysis of Klfs RNAi, many important pluripotent genes are shown to be regulated by Klfs, such as *Nanog, Esrrb,* and *Tcl1*. The rescue experiment shows that Esrrb is able to rescue the differentiation phenotype induced by Klf triple RNAi on ES cells. *Esrrb* is a direct downstream target of Klf2, Klf4 and Klf5 as these three Klfs bind to *Esrrb* gene in ChIP assays and expression of *Esrrb* is correlated with those of *Klf*s. Esrrb also positively regulates many key self-renewal regulators such *Sox2*, *Nanog*, *Tcl1*, *Tbx3* which are also targets of Klfs. Interestingly, expression of *Klf4* and *Klf5* decreases with the depletion of *Esrrb*. Esrrb binds to *Klf4* and are responsible for the activity of the regulatory element[231]. These results strongly indicate there is a reciprocal regulation between Esrrb and Klf transcription factors in ES cells which may explain the substitutionary functions between Klf4 and Esrrb in reprogramming.

Genomic location analysis showns that Esrrb shares many of its targets with those of Klf4 which indicates Esrrb may directly substitute the functions of Klfs in ES cells. However, the exact function of Klfs and Esrrb in the reprogramming should be investigated further. To dissect the functions of Klfs and Esrrb, the knockout *Klf2, Klf4*, *Klf5* or *Esrrb* cells need to be examined in the process of reprogramming. Due to our previous results showing redundant role of Klf2, Kl4 and Klf5, other Klf family members also need to be considered.

In the reprogramming screening experiments, we showed that Klf4 can be substituted by Esrrb in inducing reprogrammed cells. The Esrrb reprogrammed cells has been confirmed to be pluripotent as ES cells by all the stringent tests *in intro* and *in vivo* analysis. To our knowledge, this is first time shown in the mouse system that exogenous Klf4 is not required for the somatic cell reprogramming.

CHAPTER V

CONCLUSIONS

CHAPTER V CONCLUSIONS

Mouse ES cells belong to pluripotent stem cells which are invaluable tools and models to study diseases, development and regenerative therapy[19]. Recently, the successful derivation of pluripotent cells by defined factor-induced reprogramming from somatic cells makes the generation of patient-specific pluripotent stem cells in the near future look very promsing[56, 232]. The elucidation of functions of reprogramming factors in ES cells will provide more insights for our understanding of maintenance and recapture of pluripotency. To reveal the role of Klf4 and Klf family members in maintaining unique property of ES cells, different assays were carried out.

This project has revealed the critical roles of Klf2, Klf4 and Klf5 in the maintenance of ES cell self-renewal. RNAi was performed to deplete its expression in ES cells to dissect the functions of Klf4. No change in morphology and alkaline phosphatase staining was observed after Klf4 knockdown, indicating that Klf4 alone is not required for the normal ES cell growth. Two other Klf family members, Klf2 and Klf5, were found to be enriched in ES cells and downregulated during differentiation. Depletion of either Klf2 or Klf5 also showed no effect on ES cell self-renewal. Double knockdown of any two of the three Klfs did not induce ES cell differentiation. Simultaneously triple depletion of all three Klfs induced the loss of typical ES cell morphology and specific alkaline phosphatase staining, a characteristic of ES cell differentiation. The differentiation phenotype induced by triple Klf knockdown could be rescued by any RNAi-resistant Klf2, Klf4 and Klf5 cDNA, indicating the three Klfs are playing a redundant role in maintaining the undifferentiated ES cell state. Although overexpession of Klf2, Klf4 or Klf5 has been

shown to suppress the differentiation and support the LIF-independent ES self-renewal, we were the first to demonstrate that the three endogenous Klf family members, Klf2, Klf4 and Klf5, are critical for ES cell self-renewal.

To investigate the molecular mechanisms of redundancy for Klf2, Klf4 and Klf5 in ES cells, ChIP-on-chip was carried out by using the specific Klf antibodies and custom-designed genomic tiling microarrays. ChIP-on-chip results showed that the binding patterns of Klf2, Klf4, and Klf5 were prominently similar and these three Klfs shared a significant portion of common binding loci, indicating that these three Klfs collaborate to regulate common target genes. Klf2, Klf4 and Klf5 were shown to co-occupy many self-renewal and pluripotent regulators, such as *Nanog*, *Esrrb*, and *Tcl1*. More importantly, the other two Klfs were bound to the target sites when one of them was depleted. Klfs also have a portion of co-occupied binding sites with Nanog, suggesting the integration of the core Klfs and Nanog transcriptional regulatory networks.

To further understand the molecular basis of redundancy, the DNA binding and transcriptional activation properties of the three Klfs were validated by examining the *Nanog* enhancer region. EMSA results showed there contains a Klf motif in *Nanog* enhancer region and the same motif was bound by recombinant Klf2, Klf4 or Klf5 DNA binding domain. The intact binding motif and three Klf proteins were required for *Nanog* enhancer activity tested by reporter assay. Consistent with the differentiation phenotype induced by triple Klf RNAi, DNA microarray showed that many self-renewal and pluripotency regulators were downregulated and some of the lineage markers were

upregulated upon the suppression of the expression of the three *Klf* genes. Any of the three Klf factors were able to rescue the expression of selected markers. Taken together, these results highlight that Klf2, Klf4 and Klf5 form a close alliance to sustain the ES cell state. The impact of losing one of them is masked by the other two family members.

The functional redundancy among Klf2, Klf4 and Klf5 was also confirmed by a factor-induced reprogramming assay. Consistent with an independent observation[210], Klf2 and Klf5 were able to substitute Klf4 to generate reprogrammed cells in the conjunction with the other three reprogramming factors, Oct4, Sox2 and c-Myc. Klf2 reprogrammed cells were incorporated extensively in the chimeric embryo. These results indicate these three Klfs share similar functions in the reprogramming process.

BIBLIOGRAPHY

1. Daley, G.Q. *et al.* Broader implications of defining standards for the pluripotency of iPSCs. *Cell Stem Cell* **4**, 200-201; author reply 202 (2009).

2. Finch, B.W. & Ephrussi, B. Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture in vitro and their extinction upon hybridization with cells of permanent lines. *Proc Natl Acad Sci U S A* **57**, 615-621 (1967).

3. Stevens, L.C. The development of teratomas from intratesticular grafts of tubal mouse eggs. *J Embryol Exp Morphol* **20**, 329-341 (1968).

4. Stevens, L.C. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol* **21**, 364-382 (1970).

5. Solter, D., Skreb, N. & Damjanov, I. Extrauterine growth of mouse egg-cylinders results in malignant teratoma. *Nature* **227**, 503-504 (1970).

6. Diwan, S.B. & Stevens, L.C. Development of teratomas from the ectoderm of mouse egg cylinders. *J Natl Cancer Inst* **57**, 937-942 (1976).

7. Brinster, R.L. The effect of cells transferred into the mouse blastocyst on subsequent development. *J Exp Med* **140**, 1049-1056 (1974).

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

9. Smith, A.G. Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* **17**, 435-462 (2001).

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

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

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

13. Chung, Y. *et al.* Embryonic and extraembryonic stem cell lines derived from single mouse blastomeres. *Nature* **439**, 216-219 (2006).

14. Wakayama, S. *et al.* Efficient establishment of mouse embryonic stem cell lines from single blastomeres and polar bodies. *Stem Cells* **25**, 986-993 (2007).

15. Murry, C.E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661-680 (2008).

16. Hubner, K. *et al.* Derivation of oocytes from mouse embryonic stem cells. *Science* **300**, 1251-1256 (2003).

17. Geijsen, N. *et al.* Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* **427**, 148-154 (2004).

18. Nayernia, K. *et al.* In vitro-differentiated embryonic stem cells give rise to male gametes that can generate offspring mice. *Dev Cell* **11**, 125-132 (2006).

19. Capecchi, M.R. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* **6**, 507-512 (2005).

20. Matsui, Y., Zsebo, K. & Hogan, B.L. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841-847 (1992).

21. Durcova-Hills, G., Ainscough, J. & McLaren, A. Pluripotential stem cells derived from migrating primordial germ cells. *Differentiation* **68**, 220-226 (2001).

22. McLaren, A. & Durcova-Hills, G. Germ cells and pluripotent stem cells in the mouse. *Reprod Fertil Dev* **13**, 661-664 (2001).

23. Labosky, P.A., Barlow, D.P. & Hogan, B.L. Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. *Development* **120**, 3197-3204 (1994).

24. Tada, T. *et al.* Epigenotype switching of imprintable loci in embryonic germ cells. *Dev Genes Evol* **207**, 551-561 (1998).

25. Kato, Y. *et al.* Developmental potential of mouse primordial germ cells. *Development* **126**, 1823-1832 (1999).

26. Guan, K. *et al.* Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* **440**, 1199-1203 (2006).

27. Conrad, S. *et al.* Generation of pluripotent stem cells from adult human testis. *Nature* **456**, 344-349 (2008).

28. Brons, I.G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195 (2007).

29. Tesar, P.J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199 (2007).

30. Chou, Y.F. *et al.* The growth factor environment defines distinct pluripotent ground states in novel blastocyst-derived stem cells. *Cell* **135**, 449-461 (2008).

31. Andrews, P.W., Bronson, D.L., Benham, F., Strickland, S. & Knowles, B.B. A comparative study of eight cell lines derived from human testicular teratocarcinoma. *Int J Cancer* **26**, 269-280 (1980).

32. Pera, M.F., Cooper, S., Mills, J. & Parrington, J.M. Isolation and characterization of a multipotent clone of human embryonal carcinoma cells. *Differentiation* **42**, 10-23 (1989).

33. Shamblott, M.J. *et al.* Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc Natl Acad Sci U S A* **95**, 13726-13731 (1998).

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

35. Andrews, P.W. *et al.* Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro. *Lab Invest* **50**, 147-162 (1984).

36. Rao, B.M. & Zandstra, P.W. Culture development for human embryonic stem cell propagation: molecular aspects and challenges. *Curr Opin Biotechnol* **16**, 568-576 (2005).

37. Ying, Q.L., Nichols, J., Chambers, I. & Smith, A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-292 (2003).

38. Gerami-Naini, B. *et al.* Trophoblast differentiation in embryoid bodies derived from human embryonic stem cells. *Endocrinology* **145**, 1517-1524 (2004).

39. Odorico, J.S., Kaufman, D.S. & Thomson, J.A. Multilineage differentiation from human embryonic stem cell lines. *Stem Cells* **19**, 193-204 (2001).

40. Loh, Y.H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440 (2006).

41. Jacob, H.J. & Kwitek, A.E. Rat genetics: attaching physiology and pharmacology to the genome. *Nat Rev Genet* **3**, 33-42 (2002).

42. Brenin, D. *et al.* Rat embryonic stem cells: a progress report. *Transplant Proc* **29**, 1761-1765 (1997).

43. Buehr, M. *et al.* Rapid loss of Oct-4 and pluripotency in cultured rodent blastocysts and derivative cell lines. *Biol Reprod* **68**, 222-229 (2003).

44. Buehr, M. *et al.* Capture of authentic embryonic stem cells from rat blastocysts. *Cell* **135**, 1287-1298 (2008).

45. Li, P. *et al.* Germline competent embryonic stem cells derived from rat blastocysts. *Cell* **135**, 1299-1310 (2008).

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

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

48. Gurdon, J.B. & Uehlinger, V. "Fertile" intestine nuclei. *Nature* **210**, 1240-1241 (1966).

49. Campbell, K.H., McWhir, J., Ritchie, W.A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380**, 64-66 (1996).

50. Hochedlinger, K. & Jaenisch, R. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415**, 1035-1038 (2002).

51. Byrne, J.A. *et al.* Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* **450**, 497-502 (2007).

52. Miller, R.A. & Ruddle, F.H. Teratocarcinoma X friend erythroleukemia cell hybrids resemble their pluripotent embryonal carcinoma parent. *Dev Biol* **56**, 157-173 (1977).

53. Tada, M., Tada, T., Lefebvre, L., Barton, S.C. & Surani, M.A. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J* **16**, 6510-6520 (1997).

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

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

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

57. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317 (2007).

58. Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-324 (2007).

59. Smith, A.G. *et al.* Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-690 (1988).

60. Williams, R.L. *et al.* Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687 (1988).

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

62. Stewart, C.L. *et al.* Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* **359**, 76-79 (1992).

63. Boeuf, H., Hauss, C., Graeve, F.D., Baran, N. & Kedinger, C. Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. *J Cell Biol* **138**, 1207-1217 (1997).

64. Ernst, M., Novak, U., Nicholson, S.E., Layton, J.E. & Dunn, A.R. The carboxyl-terminal domains of gp130-related cytokine receptors are necessary for suppressing embryonic stem cell differentiation. Involvement of STAT3. *J Biol Chem* **274**, 9729-9737 (1999).

65. Raz, R., Lee, C.K., Cannizzaro, L.A., d'Eustachio, P. & Levy, D.E. Essential role of STAT3 for embryonic stem cell pluripotency. *Proc Natl Acad Sci U S A* **96**, 2846-2851 (1999).

66. Takeda, K. *et al.* Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc Natl Acad Sci U S A* **94**, 3801-3804 (1997).

67. Li, M., Sendtner, M. & Smith, A. Essential function of LIF receptor in motor neurons. *Nature* **378**, 724-727 (1995).

68. Ware, C.B. *et al.* Targeted disruption of the low-affinity leukemia inhibitory factor receptor gene causes placental, skeletal, neural and metabolic defects and results in perinatal death. *Development* **121**, 1283-1299 (1995).

69. Yoshida, K. *et al.* Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc Natl Acad Sci U S A* **93**, 407-411 (1996).

70. Rodig, S.J. *et al.* Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell* **93**, 373-383 (1998).

71. Neubauer, H. *et al.* Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell* **93**, 397-409 (1998).

72. Parganas, E. *et al.* Jak2 is essential for signaling through a variety of cytokine receptors. *Cell* **93**, 385-395 (1998).

73. Thomis, D.C., Gurniak, C.B., Tivol, E., Sharpe, A.H. & Berg, L.J. Defects in B lymphocyte maturation and T lymphocyte activation in mice lacking Jak3. *Science* **270**, 794-797 (1995).

74. Nosaka, T. *et al.* Defective lymphoid development in mice lacking Jak3. *Science* **270**, 800-802 (1995).

75. Shi, Y. & Massague, J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**, 685-700 (2003).

76. Mishina, Y., Suzuki, A., Ueno, N. & Behringer, R.R. Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes Dev* **9**, 3027-3037 (1995).

77. Qi, X. *et al.* BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc Natl Acad Sci U S A* **101**, 6027-6032 (2004).

78. Logan, C.Y. & Nusse, R. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol* **20**, 781-810 (2004).

79. Kielman, M.F. *et al.* Apc modulates embryonic stem-cell differentiation by controlling the dosage of beta-catenin signaling. *Nat Genet* **32**, 594-605 (2002).

80. Sato, N., Meijer, L., Skaltsounis, L., Greengard, P. & Brivanlou, A.H. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med* **10**, 55-63 (2004).

81. Doble, B.W., Patel, S., Wood, G.A., Kockeritz, L.K. & Woodgett, J.R. Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. *Dev Cell* **12**, 957-971 (2007).

82. Pereira, L., Yi, F. & Merrill, B.J. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol* **26**, 7479-7491 (2006).

83. Tam, W.L. *et al.* T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. *Stem Cells* **26**, 2019-2031 (2008).

84. Lu, C.W. *et al.* Ras-MAPK signaling promotes trophectoderm formation from embryonic stem cells and mouse embryos. *Nat Genet* **40**, 921-926 (2008).

85. Batlle-Morera, L., Smith, A. & Nichols, J. Parameters influencing derivation of embryonic stem cells from murine embryos. *Genesis* **46**, 758-767 (2008).

86. Wilder, P.J. *et al.* Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Dev Biol* **192**, 614-629 (1997).

87. Kunath, T. *et al.* 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 (2007).

88. Ying, Q.L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523 (2008).

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

90. Takahashi, K., Murakami, M. & Yamanaka, S. Role of the phosphoinositide 3-kinase pathway in mouse embryonic stem (ES) cells. *Biochem Soc Trans* **33**, 1522-1525 (2005).

91. Murakami, M. *et al.* mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. *Mol Cell Biol* **24**, 6710-6718 (2004).

92. Sturm, R.A. & Herr, W. The POU domain is a bipartite DNA-binding structure. *Nature* **336**, 601-604 (1988).

93. Imagawa, M., Miyamoto, A., Shirakawa, M., Hamada, H. & Muramatsu, M. Stringent integrity requirements for both trans-activation and DNA-binding in a trans-activator, Oct3. *Nucleic Acids Res* **19**, 4503-4508 (1991).

94. Scholer, H.R., Ciesiolka, T. & Gruss, P. A nexus between Oct-4 and E1A: implications for gene regulation in embryonic stem cells. *Cell* **66**, 291-304 (1991).

95. Brehm, A., Ohbo, K. & Scholer, H. The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain. *Mol Cell Biol* **17**, 154-162 (1997).

96. Scholer, H.R., Balling, R., Hatzopoulos, A.K., Suzuki, N. & Gruss, P. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *EMBO J* **8**, 2551-2557 (1989).

97. Scholer, H.R., Hatzopoulos, A.K., Balling, R., Suzuki, N. & Gruss, P. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. *EMBO J* **8**, 2543-2550 (1989).

98. Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K. & Gruss, P. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**, 435-439 (1990).

99. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391 (1998).

100. Kehler, J. *et al.* Oct4 is required for primordial germ cell survival. *EMBO Rep* **5**, 1078-1083 (2004).

101. Lengner, C.J. *et al.* Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* **1**, 403-415 (2007).

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

103. Niwa, H. *et al.* Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917-929 (2005).

104. Strumpf, D. *et al.* Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**, 2093-2102 (2005).

105. Yeom, Y.I. *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-894 (1996).

106. Nordhoff, V. *et al.* Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm Genome* **12**, 309-317 (2001).

107. Barnea, E. & Bergman, Y. Synergy of SF1 and RAR in activation of Oct-3/4 promoter. *J Biol Chem* **275**, 6608-6619 (2000).

108. Fuhrmann, G. *et al.* Mouse germline restriction of Oct4 expression by germ cell nuclear factor. *Dev Cell* **1**, 377-387 (2001).

109. Gu, P. *et al.* Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol Cell Biol* **25**, 3492-3505 (2005).

110. Chew, J.L. *et al.* Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* **25**, 6031-6046 (2005).

111. Okumura-Nakanishi, S., Saito, M., Niwa, H. & Ishikawa, F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem* **280**, 5307-5317 (2005).

112. Zhang, J. *et al.* Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* **8**, 1114-1123 (2006).

113. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* (2008).

114. Feldman, N. *et al.* G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**, 188-194 (2006).

115. Rodda, D.J. *et al.* Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* **280**, 24731-24737 (2005).

116. Tokuzawa, Y. *et al.* Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. *Mol Cell Biol* **23**, 2699-2708 (2003).

117. Wood, H.B. & Episkopou, V. Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* **86**, 197-201 (1999).

118. Zappone, M.V. *et al.* Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* **127**, 2367-2382 (2000).

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

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

121. Yuan, H., Corbi, N., Basilico, C. & Dailey, L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* **9**, 2635-2645 (1995).

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

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

124. Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell. Biol.* **9**, 625-635 (2007).

125. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655 (2003).

126. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642 (2003).

127. Pan, G.J. & Pei, D.Q. Identification of two distinct transactivation domains in the pluripotency sustaining factor nanog. *Cell Res* **13**, 499-502 (2003).

128. Wang, J., Levasseur, D.N. & Orkin, S.H. Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A* **105**, 6326-6331 (2008).

129. Mullin, N.P. *et al.* The pluripotency rheostat Nanog functions as a dimer. *Biochem J* **411**, 227-231 (2008).

130. Hart, A.H., Hartley, L., Ibrahim, M. & Robb, L. Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev Dyn* **230**, 187-198 (2004).

131. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-1234 (2007).

132. Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-538 (2006).

133. Chen, X., Fang, F., Liou, Y.C. & Ng, H.H. Zfp143 regulates Nanog through modulation of Oct4 binding. *Stem Cells* **26**, 2759-2767 (2008).

134. Pan, G., Li, J., Zhou, Y., Zheng, H. & Pei, D. A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J* **20**, 1730-1732 (2006).

135. Wu, Q. *et al.* Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *J. Biol. Chem.* **281**, 24090-24094 (2006).

136. Suzuki, A. *et al.* Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. *Proc Natl Acad Sci U S A* **103**, 10294-10299 (2006).

137. Yi, F., Pereira, L. & Merrill, B.J. Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. *Stem Cells* **26**, 1951-1960 (2008).

138. Lin, T. *et al.* p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat Cell Biol* **7**, 165-171 (2005).

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

140. Tay, Y.M. *et al.* MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells* **26**, 17-29 (2008).

141. Yates, A. & Chambers, I. The homeodomain protein Nanog and pluripotency in mouse embryonic stem cells. *Biochem Soc Trans* **33**, 1518-1521 (2005).

142. Fujita, J. *et al.* Caspase activity mediates the differentiation of embryonic stem cells. *Cell Stem Cell* **2**, 595-601 (2008).

143. Pearson, R., Fleetwood, J., Eaton, S., Crossley, M. & Bao, S. Kruppel-like transcription factors: a functional family. *Int J Biochem Cell Biol* **40**, 1996-2001 (2008).

144. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318-322 (1995).

145. Kuo, C.T. *et al.* The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. *Genes Dev.* **11**, 2996-3006 (1997).

146. Sue, N. *et al.* Targeted disruption of the basic Kruppel-like factor gene (Klf3) reveals a role in adipogenesis. *Mol Cell Biol* **28**, 3967-3978 (2008).

147. Segre, J.A., Bauer, C. & Fuchs, E. Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat. Genet.* **22**, 356-360 (1999).

148. Katz, J.P. *et al.* The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. *Development* **129**, 2619-2628 (2002).

149. Shindo, T. *et al.* Kruppel-like zinc-finger transcription factor KLF5/BTEB2 is a target for angiotensin II signaling and an essential regulator of cardiovascular remodeling. *Nat. Med.* **8**, 856-863 (2002).

150. Matsumoto, N. *et al.* Developmental regulation of yolk sac hematopoiesis by Kruppel-like factor 6. *Blood* **107**, 1357-1365 (2006).

151. Laub, F. *et al.* Transcription factor KLF7 is important for neuronal morphogenesis in selected regions of the nervous system. *Mol Cell Biol* **25**, 5699-5711 (2005).

152. Morita, M. *et al.* Functional analysis of basic transcription element binding protein by gene targeting technology. *Mol Cell Biol* **23**, 2489-2500 (2003).

153. Simmen, R.C. *et al.* Subfertility, uterine hypoplasia, and partial progesterone resistance in mice lacking the Kruppel-like factor 9/basic transcription element-binding protein-1 (Bteb1) gene. *J Biol Chem* **279**, 29286-29294 (2004).

154. Subramaniam, M. *et al.* TIEG1 null mouse-derived osteoblasts are defective in mineralization and in support of osteoclast differentiation in vitro. *Mol Cell Biol* **25**, 1191-1199 (2005).

155. Gordon, A.R. *et al.* Splenomegaly and modified erythropoiesis in KLF13-/- mice. *J Biol Chem* **283**, 11897-11904 (2008).

156. Fisch, S. *et al.* Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. *Proc Natl Acad Sci U S A* **104**, 7074-7079 (2007).

157. Song, C.Z., Gavriilidis, G., Asano, H. & Stamatoyannopoulos, G. Functional study of transcription factor KLF11 by targeted gene inactivation. *Blood Cells Mol Dis* **34**, 53-59 (2005).

158. Zhang, W. *et al.* The gut-enriched Kruppel-like factor (Kruppel-like factor 4) mediates the transactivating effect of p53 on the p21WAF1/Cip1 promoter. *J Biol Chem* **275**, 18391-18398 (2000).

159. Rowland, B.D., Bernards, R. & Peeper, D.S. The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nat. Cell. Biol.* **7**, 1074-1082 (2005).

160. Hagos, E.G., Ghaleb, A.M., Dalton, W.B., Bialkowska, A.B. & Yang, V.W. Mouse embryonic fibroblasts null for the Kruppel-like factor 4 gene are genetically unstable. *Oncogene* (2009).

161. Li, Y. *et al.* Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* **105**, 635-637 (2005).

162. Nakatake, Y. *et al.* Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol. Cell. Biol.* **26**, 7772-7782 (2006).

163. Hailesellasse Sene, K. *et al.* Gene function in early mouse embryonic stem cell differentiation. *BMC Genomics* **8**, 85 (2007).

164. Niakan, K.K. *et al.* Novel role for the orphan nuclear receptor Dax1 in embryogenesis, different from steroidogenesis. *Mol Genet Metab* **88**, 261-271 (2006).

165. Lim, L.S. *et al.* Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Mol Biol Cell* **18**, 1348-1358 (2007).

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

167. Dejosez, M. *et al.* Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell* **133**, 1162-1174 (2008).

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

169. Chen, T. & Li, E. Establishment and maintenance of DNA methylation patterns in mammals. *Curr Top Microbiol Immunol* **301**, 179-201 (2006).

170. Li, E., Bestor, T.H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926 (1992).

171. Kim, M., Trinh, B.N., Long, T.I., Oghamian, S. & Laird, P.W. Dnmt1 deficiency leads to enhanced microsatellite instability in mouse embryonic stem cells. *Nucleic Acids Res* **32**, 5742-5749 (2004).

172. Panning, B. & Jaenisch, R. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* **10**, 1991-2002 (1996).

173. Okano, M., Bell, D.W., Haber, D.A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).

174. Chen, T., Ueda, Y., Dodge, J.E., Wang, Z. & Li, E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol* **23**, 5594-5605 (2003).

175. Jackson, M. *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol* **24**, 8862-8871 (2004).

176. Bernstein, B.E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).

177. O'Carroll, D. *et al.* The polycomb-group gene Ezh2 is required for early mouse development. *Mol Cell Biol* **21**, 4330-4336 (2001).

178. Wang, J. *et al.* Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* **28**, 371-375 (2001).

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

180. Boyer, L.A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353 (2006).

181. Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M. & Helin, K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* **27**, 3769-3779 (2007).

182. Creyghton, M.P. *et al.* H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* **135**, 649-661 (2008).

183. Boyer, L.A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956 (2005).

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

185. Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* (2009).

186. Siomi, H. & Siomi, M.C. On the road to reading the RNA-interference code. *Nature* **457**, 396-404 (2009).

187. Fazzio, T.G., Huff, J.T. & Panning, B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134**, 162-174 (2008).

188. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368 (2006).

189. Dejardin, J. & Kingston, R.E. Purification of proteins associated with specific genomic Loci. *Cell* **136**, 175-186 (2009).

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

191. Weintraub, H. *et al.* Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc Natl Acad Sci U S A* **86**, 5434-5438 (1989).

192. Orkin, S.H. & Zon, L.I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631-644 (2008).

193. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D.A. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* **455**, 627-632 (2008).

194. Shi, Y. *et al.* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* **3**, 568-574 (2008).

195. Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**, 1269-1275 (2008).

196. Bieda, M., Xu, X., Singer, M.A., Green, R. & Farnham, P.J. Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. *Genome Res* **16**, 595-605 (2006).

197. Hanna, J. *et al.* Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **318**, 1920-1923 (2007).

198. McConnell, B.B., Ghaleb, A.M., Nandan, M.O. & Yang, V.W. The diverse functions of Kruppel-like factors 4 and 5 in epithelial biology and pathobiology. *Bioessays* **29**, 549-557 (2007).

199. Hao, L. *et al.* Drosophila RNAi screen identifies host genes important for influenza virus replication. *Nature* **454**, 890-893 (2008).

200. Guo, Y. *et al.* Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. *Nature* **453**, 657-661 (2008).

201. Westbrook, T.F. *et al.* SCFbeta-TRCP controls oncogenic transformation and neural differentiation through REST degradation. *Nature* **452**, 370-374 (2008).

202. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature* **434**, 462-469 (2005).

203. Svoboda, P. Off-targeting and other non-specific effects of RNAi experiments in mammalian cells. *Curr Opin Mol Ther* **9**, 248-257 (2007).

204. Ma, Y., Creanga, A., Lum, L. & Beachy, P.A. Prevalence of off-target effects in Drosophila RNA interference screens. *Nature* **443**, 359-363 (2006).

205. Cullen, B.R. Enhancing and confirming the specificity of RNAi experiments. *Nat Methods* **3**, 677-681 (2006).

206. Li, C.X. *et al.* Delivery of RNA interference. *Cell Cycle* **5**, 2103-2109 (2006).

207. Matoba, R. *et al.* Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS ONE.* **1**, e26 (2006).

208. Suske, G., Bruford, E. & Philipsen, S. Mammalian SP/KLF transcription factors: bring in the family. *Genomics* **85**, 551-556 (2005).

209. Szabo, P.E., Hubner, K., Scholer, H. & Mann, J.R. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech. Dev.* **115**, 157-160 (2002).

210. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* **26**, 101-106 (2008).

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

212. Rowland, B.D. & Peeper, D.S. KLF4, p21 and context-dependent opposing forces in cancer. *Nat. Rev. Cancer.* **6**, 11-23 (2006).

213. Shields, J.M., Christy, R.J. & Yang, V.W. Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J Biol Chem* **271**, 20009-20017 (1996).

214. Katz, J.P. *et al.* Loss of Klf4 in mice causes altered proliferation and differentiation and precancerous changes in the adult stomach. *Gastroenterology* **128**, 935-945 (2005).

215. Feinberg, M.W. *et al.* The Kruppel-like factor KLF4 is a critical regulator of monocyte differentiation. *EMBO J* **26**, 4138-4148 (2007).

216. Yamanaka, S. Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell* **1**, 39-49 (2007).

217. Sun, R., Chen, X. & Yang, V.W. Intestinal-enriched Kruppel-like factor (Kruppel-like factor 5) is a positive regulator of cellular proliferation. *J Biol Chem* **276**, 6897-6900 (2001).

218. Nandan, M.O. *et al.* Kruppel-like factor 5 mediates the transforming activity of oncogenic H-Ras. *Oncogene* **23**, 3404-3413 (2004).

219. Zhu, N. *et al.* KLF5 Interacts with p53 in regulating survivin expression in acute lymphoblastic leukemia. *J Biol Chem* **281**, 14711-14718 (2006).

220. Chen, C., Bhalala, H.V., Qiao, H. & Dong, J.T. A possible tumor suppressor role of the KLF5 transcription factor in human breast cancer. *Oncogene* **21**, 6567-6572 (2002).

221. Chen, C., Bhalala, H.V., Vessella, R.L. & Dong, J.T. KLF5 is frequently deleted and down-regulated but rarely mutated in prostate cancer. *Prostate* **55**, 81-88 (2003).

222. Oishi, Y. *et al.* Kruppel-like transcription factor KLF5 is a key regulator of adipocyte differentiation. *Cell Metab.* **1**, 27-39 (2005).

223. Wan, H. *et al.* Kruppel-like factor 5 is required for perinatal lung morphogenesis and function. *Development* **135**, 2563-2572 (2008).

224. Ema, M. *et al.* Kruppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs. *Cell Stem Cell* **3**, 555-567 (2008).

225. Wani, M.A., Means, R.T., Jr. & Lingrel, J.B. Loss of LKLF function results in embryonic lethality in mice. *Transgenic Res* **7**, 229-238 (1998).

226. Wani, M.A., Wert, S.E. & Lingrel, J.B. Lung Kruppel-like factor, a zinc finger transcription factor, is essential for normal lung development. *J. Biol. Chem.* **274**, 21180-21185 (1999).

227. Carlson, C.M. *et al.* Kruppel-like factor 2 regulates thymocyte and T-cell migration. *Nature* **442**, 299-302 (2006).

228. Wu, J., Srinivasan, S.V., Neumann, J.C. & Lingrel, J.B. The KLF2 transcription factor does not affect the formation of preadipocytes but inhibits their differentiation into adipocytes. *Biochemistry* **44**, 11098-11105 (2005).

229. Malynn, B.A. *et al.* N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation. *Genes Dev* **14**, 1390-1399 (2000).

230. Blelloch, R., Venere, M., Yen, J. & Ramalho-Santos, M. Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell* **1**, 245-247 (2007).

231. Feng, B. *et al.* Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* **11**, 197-203 (2009).

232. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).

233. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55-71 (2007).

234. Aoi, T. *et al.* Generation of Pluripotent Stem Cells from Adult Mouse Liver and Stomach Cells. *Science* (2008).

235. Hanna, J. *et al.* Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency. *Cell* **133**, 250-264 (2008).

236. Kim, J.B. *et al.* Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**, 411-419 (2009).

237. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920 (2007).

238. Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151-159 (2008).

239. Stadtfeld, M., Maherali, N., Breault, D.T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* **2**, 230-240 (2008).

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

241. Han, D.W. *et al.* Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. *Stem Cells* **26**, 445-454 (2008).

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

APPENDIX

APPENDIX

List of genes and their coordinates on the custom designed microarray

Chromosome	mm5 coordinates	Gene(1)	Gene(2)
chr1	chr1:4453577-4559061	Sox17	
chr1	chr1:15917437-16055202	Terf1	
chr1	chr1:52332580-52575344	Stat1	Stat4
chr1	chr1:63046252-63258435	Nrp2	
chr1	chr1:64380327-64566045	Klf7	
chr1	chr1:72682489-72873479	Xrcc5	
chr1	chr1:73208897-73359173	Igfbp2	Igfbp5
chr1	chr1:75608366-75712623	Tuba4	
chr1	chr1:78507003-78701663	Pax3	
chr1	chr1:89786448-89889662	Gbx2	
chr1	chr1:118338416-118495209	Tcfcp2l1	
chr1	chr1:120326028-120431347	En1	
chr1	chr1:132022652-132122652	miR135b	
chr1	chr1:134400020-134573652	Jarid1b	
chr1	chr1:136462843-136566187	Zfp281	
chr1	chr1:136682175-136901392	Nr5a2	
chr1	chr1:151381520-151513029	Rnf2	
chr1	chr1:167669829-167924518	Lmx1a	
chr1	chr1:179719335-179839315	Elys	
chr1	chr1:180716861-180820845	Mixl1	
chr1	chr1:180918897-181025242	Lefty2	
chr1	chr1:189937294-190086198	Prox1	
chr1	chr1:194880944-194980944	miR29b-2,29c	
chr2	chr2:9758958-9878902	Gata3	
chr2	chr2:25969030-26132803	Btbd14a	
chr2	chr2:29972065-30081814	Set	
chr2	chr2:38261575-38379579	Lhx2	
chr2	chr2:38636905-38936411	Gcnf	
chr2	chr2:51982823-52132047	Rif1	
chr2	chr2:57010885-57127288	Nurr/Nr4a2	
chr2	chr2:71373611-71478146	Dlx1	
chr2	chr2:72780814-72923763	Sp3	
chr2	chr2:105518674-105640073	Pax6	
chr2	chr2:113778865-113884388	actin alpha	
chr2	chr2:120929828-121102278	Trp53bp1	
chr2	chr2:121846710-121952105	B2m	
chr2	chr2:126252166-126387265	Slc27a2	
chr2	chr2:130987118-131087118	miR103-2	
chr2	chr2:131948310-132052188	Pcna	
chr2	chr2:133252031-133360738	Bmp2	

chr2	chr2:145924800-146026594	Insm1	
chr2	chr2:146880109-146988816	Nkx2.2	
chr2	chr2:147745484-147849530	Foxa2	
chr2	chr2:152156958-152262443	Tcf15	
chr2	chr2:152749605-152850755	Id1	
chr2	chr2:153662806-153801075	Dnmt3b	
chr2	chr2:154572987-154683201	E2f1	
chr2	chr2:157573480-157675797	Nnat	
chr2	chr2:157750811-158004963	Ctnnbl1	
chr2	chr2:164087501-164268872	Stk4	
chr2	chr2:165668411-165884743	Eya2	
chr2	chr2:167505758-167655922	Zfp313	Snai1
chr2	chr2:168761454-168863088	Sall4	
chr2	chr2:172562732-172671727	Tcfap2c	
chr2	chr2:174284552-174384552	miR296	
chr2	chr2:179963340-180072880	Gata5	
chr2	chr2:180296171-180440307	Dido1	
chr2	chr2:180532282-180632282	miR124a3	
chr2	chr2:181307962-181409744	Sox18	
chr3	chr3:17397574-17497574	miR124a2	
chr3	chr3:26811411-26968571	Ect2	
chr3	chr3:32066873-32197695	Wig1	
chr3	chr3:32441932-32560177	Actl6a	Baf53a-pending
chr3	chr3:34402927-34505331	Sox2	
chr3	chr3:36316610-36423682	Cyclin A2	
chr3	chr3:37172037-37430829	Spaf-pending	
chr3	chr3:41258650-41418305	Phf17	
chr3	chr3:52302398-52482354	Foxo1	
chr3	chr3:54825115-54940261	Smad8	
chr3	chr3:88410785-88520099	Nestin	
chr3	chr3:89793593-89904993	Shc1	
chr3	chr3:90864641-90966901	Nf45	
chr3	chr3:95268308-95401938	Eset	
chr3	chr3:99127398-99327786	Tbox15	
chr3	chr3:116007205-116111813	Edg1	
chr3	chr3:127975356-128165308	miR302	Ngn2
chr3	chr3:129695993-129801603	Pitx2	
chr3	chr3:131643385-131857096	Lef1	
chr3	chr3:138372195-138474516	H2A histone	
chr4	chr4:54722188-54827228	Klf4	
chr4	chr4:56029036-56183287	Catnal1	
chr4	chr4:72476250-72658007	Jmjd2c	
chr4	chr4:84881156-84993015	Adrp	
chr4	chr4:93329330-93432462	Jun	
chr4	chr4:95862141-96493100	Nfia	

chr4	chr4:97936574-98038897	FoxD3	
chr4	chr4:98045677-98209393	Itgb3bp	
chr4	chr4:98225534-98383039	Pgm2	
chr4	chr4:115391997-115510472	Ptch2	
chr4	chr4:116433957-116575647	Jmjd2a	
chr4	chr4:121316414-121422793	Lmyc1	
chr4	chr4:132004453-132105199	Mkrn3	
chr4	chr4:132251917-132392374	Rps6ka1	
chr4	chr4:132407956-132523442	Lin28	
chr4	chr4:134551837-134702729	Id3	E2f2
chr4	chr4:134958574-135110725	Aof2	
chr4	chr4:144356021-144806143	Rex2	
chr5	chr5:4760451-5436916	Pftk1	
chr5	chr5:25154786-26162667	Dpp6	
chr5	chr5:26805325-26915572	Shh	
chr5	chr5:27887082-27987662	Dppa5	
chr5	chr5:30967903-31142020	Yes	
chr5	chr5:32734344-32851366	Rnf4	
chr5	chr5:52166672-52332152	Rbpsuh	
chr5	chr5:65648456-65751938	Phox2b	
chr5	chr5:73721071-73928966	Gsh2	Pdgfra
chr5	chr5:75769147-75897339	Hod	
chr5	chr5:75951660-76068600	Rest	
chr5	chr5:89251072-89369184	Afp	
chr5	chr5:95523937-95645077	Fgf5	
chr5	chr5:101754477-101860387	Spp1	
chr5	chr5:105457444-105599471	Mtf2	
chr5	chr5:112309606-112432182	Tcf1	
chr5	chr5:112844811-112951001	SIRT4	
chr5	chr5:117106216-117218017	Tbx3	
chr5	chr5:117271565-117420164	Tbx5	
chr5	chr5:117873981-117984167	Lhx5	
chr5	chr5:140333848-140437454	Beta-actin	
chr5	chr5:144544959-144682045	Pdx1	Cdx2
chr6	chr6:6539624-6658041	Dlx6	Dlx5
chr6	chr6:28824844-28924844	miR129-1	
chr6	chr6:30533166-30733385	Mest	
chr6	chr6:30890068-30990068	miR29a,29b-1	
chr6	chr6:39275681-39398292	Mkrn1	
chr6	chr6:52052653-52301941	whole Hoxa	
chr6	chr6:56739854-56886810	Fkbp9	
chr6	chr6:65097702-65269284	Etl1	
chr6	chr6:67097607-67200124	Gadd45a	
chr6	chr6:71766586-71910113	Jmjd1a	
chr6	chr6:72805447-73066310	Tcf3	

chr6	chr6:88129862-88237773	Cnbp1
chr6	chr6:88478258-88591605	Gata2
chr6	chr6:89171092-89286482	Mcm2
chr6	chr6:97510667-97632275	Ube1c
chr6	chr6:100559568-100717648	Rybp
chr6	chr6:109000683-109106767	Stra13
chr6	chr6:113512025-113626609	Rosa26
chr6	chr6:120713755-120868801	Jarid1a
chr6	chr6:122899306-123021209	Edr1
chr6	chr6:123133071-123434300	Nanog cluster
chr6	chr6:125728777-125832529	AK002273
chr6	chr6:127710276-127831972	Ccnd2
chr6	chr6:128968032-129079234	FoxM1
chr6	chr6:137524744-137797510	Eps8
chr6	chr6:143932355-144410309	Sox5
chr6	chr6:148331390-148638842	Rps4y2
chr7	chr7:3260406-3399839	Nalp4c
chr7	chr7:17685464-17807436	SIRT2
chr7	chr7:23486833-23589548	Cebpa
chr7	chr7:32321968-32431551	Prmt1
chr7	chr7:32561111-32679291	Tead2
chr7	chr7:47037038-47159141	SNRPN-Snurf
chr7	chr7:49275003-49376576	Ndn
chr7	chr7:57342749-57449900	Nr2f2
chr7	chr7:67487870-67668101	Blm
chr7	chr7:77078806-77205096	Eed
chr7	chr7:91310192-91411687	Hbb
chr7	chr7:97281577-97415309	Zfp143
chr7	chr7:126478778-126597310	Bnip3
chr7	chr7:127545941-127647197	Utf1
chr7	chr7:128503267-128621436	SIRT3
chr7	chr7:130228405-130439374	Igf2/H19
chr7	chr7:130621598-130723999	Ascl2
chr7	chr7:132516903-132622712	Fgf4
chr7_random	chr7_random:7029230-7130282	Zfp296
chr7_random	chr7_random:10040039-	ZEC
chr7_random	chr7_random:17561914-	Zfp110
chr7_random	chr7_random:20793562-	Peg3
chr7_random	chr7_random:43614382-	Tif1b
chr8	chr8:12361556-12462746	Sox1
chr8	chr8:21079639-21200812	Slc25a15
chr8	chr8:24543666-24669438	Ash2l
chr8	chr8:42363594-42475497	Zfp42
chr8	chr8:56055840-56159374	Hand2
chr8	chr8:69687685-69789333	JunD1

chr8	chr8:71301373-71403587	Klf2	
chr8	chr8:76732518-76894301	Ednra	
chr8	chr8:77498417-77600761	Pou4f2	
chr8	chr8:78411507-78572446	Smad1	
chr8	chr8:79825439-79941942	Smarca5	
chr8	chr8:83893115-84096623	Nac1	Nfix
chr8	chr8:84202476-84304243	JunB	
chr8	chr8:88292659-88406860	Sall1	
chr8	chr8:91070548-91173261	Irx3	
chr8	chr8:104900005-105046440	Ctcf	
chr8	chr8:106161852-106284099	Tex292	
chr8	chr8:107977057-108321141	Atbf1	
chr8	chr8:115910107-116174268	Cdyl2	
chr8	chr8:122828683-122939128	Tubb3	
chr8	chr8:123135303-123265303	Rab4a	
chr8	chr8:127730554-127974781	Nrp1	
chr9	chr9:20784672-20929794	Dnmt1	
chr9	chr9:21456980-21577004	Carm1	
chr9	chr9:23117838-23486710	Bmper	
chr9	chr9:32360361-32578018	Fli1	
chr9	chr9:57946016-58089739	Arid3b	
chr9	chr9:63833000-64044100	Smad3	
chr9	chr9:64141097-64308600	Smad6	
chr9	chr9:65885235-66085293	Zfp609	
chr9	chr9:69954918-70058009	Foxb1	
chr9	chr9:70611922-70725763	Cyclin B2	
chr9	chr9:72929790-73116983	Nedd4	
chr9	chr9:75145574-75273070	ONECUT1	
chr9	chr9:76060578-76283573	Bmp5	
chr9	chr9:106045068-106145068	miR135a1	
chr9	chr9:110068086-110282805	BAF155	
chr9	chr9:110893759-111000223	Tdgf1	
chr9	chr9:118398496-118505650	Eomes	
chr9	chr9:120878941-121005960	beta-catenin	
chr9	chr9:122880595-122988536	Zfp105	
chr10	chr10:19005947-19108012	Olig3	
chr10	chr10:21668568-21773755	Sgk	
chr10	chr10:42419342-42541008	Nr2e1	
chr10	chr10:44308196-44429651	Blimp1	
chr10	chr10:61128850-61236076	Nodal	
chr10	chr10:62298001-62418147	Ddx21	
chr10	chr10:63033272-63153303	SIRT1	
chr10	chr10:66933001-67110421	Jmjd1c	Nrbf2
chr10	chr10:69058799-69170623	Cdc2a	
chr10	chr10:77883251-77997150	DNMT3L	

chr10	chr10:79777175-79914231	Arid3a	Wdr18
chr10	chr10:82357052-82471278	Nfyb	
chr10	chr10:87189599-87290827	Ascl1	
chr10	chr10:88372950-88480841	Spic	
chr10	chr10:91789944-91889944	miR135a2	
chr10	chr10:95110933-95238734	Socs2	
chr10	chr10:102665136-102786170	Cart1	
chr10	chr10:107106895-107209915	Myf5	
chr10	chr10:107961900-108143330	Pawr	
chr10	chr10:120001945-120217367	Hmga2	
chr10	chr10:128136462-128254518	Tebp-pending	
chr10	chr10:128314825-128429698	Timeless	
chr11	chr11:4783546-4892852	NFH	
chr11	chr11:4914689-5043865	EWS	
chr11	chr11:7053174-7158624	Igfbp3	
chr11	chr11:18726000-18964236	Meis1	
chr11	chr11:20047172-20172425	Rab1	
chr11	chr11:21840346-21947131	Otx1	
chr11	chr11:35413368-35513368	miR103-1	
chr11	chr11:46198188-46318996	Dppa1	
chr11	chr11:51842343-51972310	Tcf7	
chr11	chr11:57391240-57494674	Hand1	
chr11	chr11:68961076-69076231	Jmjd3	
chr11	chr11:69142944-69319177	Trp53	Sox15
chr11	chr11:72003925-72114105	Mybbp1a	
chr11	chr11:75026318-75126318	miR22	
chr11	chr11:86135778-86235778	miR21	
chr11	chr11:87037951-87207487	Tex14	
chr11	chr11:88620603-88741416	Trim25	
chr11	chr11:94761574-94866604	Dlx4	
chr11	chr11:95583855-95833085	Igf2bp1	Ndp52
chr11	chr11:95901890-96083789	whole Hoxb	
chr11	chr11:98662388-98773764	Igfbp4	
chr11	chr11:100509359-100660510	Stat3	
chr11	chr11:101110890-101272966	Brca1	
chr11	chr11:101390874-101615477	Etv4	Meox1
chr11	chr11:102508519-102618161	Gfap	
chr11	chr11:106276006-106436276	Pecam	
chr11	chr11:112403353-112508878	Sox9	
chr11	chr11:117587215-117690291	Socs3	
chr11	chr11:120767272-120869440	Tex19	
chr12	chr12:3650324-3827880	Dnmt3a	
chr12	chr12:13010106-13113542	Nmyc1	
chr12	chr12:20007157-20109424	Id2	
chr12	chr12:22248251-22355747	Sox11	

chr12	chr12:48528098-49342163	Npas3	
chr12	chr12:51674690-51779635	Titf1	
chr12	chr12:68328216-68441676	Six4	
chr12	chr12:80748307-80851673	Fos	
chr12	chr12:81636141-81899428	Esrrb	
chr12	chr12:99133736-99235763	Gsc	
chr12	chr12:99879273-99985226	Tcl1	
chr12	chr12:103496666-103620182	YY1	
chr12	chr12:104245775-104389564	Gtl2	miR127, 136, 337
chr12	chr12:106368657-106471567	Ckb	
chr13	chr13:5600424-5706588	Klf6	
chr13	chr13:14769517-15133250	Gli3	
chr13	chr13:26423894-26529931	Prl3d1	
chr13	chr13:28278726-28383380	Sox4	
chr13	chr13:29239454-29415745	E2F3	
chr13	chr13:35046542-35360566	Cdyl	
chr13	chr13:40109182-40223950	Tcfap2a	
chr13	chr13:44145406-44434990	Jarid2	
chr13	chr13:46487908-46628938	AOF1	
chr13	chr13:47707354-47809960	Id4	
chr13	chr13:50177175-50277608	Histone H2A	
chr13	chr13:55301697-55403361	Neurog1	
chr13	chr13:55758625-55897352	Smad5	
chr13	chr13:74738340-74847658	Nr2f1	
chr13	chr13:80319428-80419428	miR9-2	
chr13	chr13:91552374-91660422	Otp	
chr13	chr13:109488347-109642277	Ddx4	
chr13	chr13:113208357-113318783	Isl1	
chr14	chr14:22986942-23088903	Hesx1	
chr14	chr14:30407535-30557024	Bmpr1a	
chr14	chr14:30668726-30840915	Wapl	
chr14	chr14:38995072-39097484	Bmp4	
chr14	chr14:39419546-39533567	Gmfb	
chr14	chr14:41281847-41391167	Otx2	
chr14	chr14:43433573-43537614	Mm.343880	
chr14	chr14:44696329-44810617	Zfp219	
chr14	chr14:46292529-46440620	Slc7a7	
chr14	chr14:48330736-48491472	Rnf17	
chr14	chr14:48847550-48994916	Zfp198	
chr14	chr14:55137504-55283853	Gata4	
chr14	chr14:55882889-55989917	Sox7	
chr14	chr14:56527472-56627472	miR124a1	
chr14	chr14:62098884-62255498	Piwil2	
chr14	chr14:89907436-90390711	Dach1	
chr14	chr14:91472441-91587148	Klf5	

chr14	chr14:114762314-114866705	Zic2	
chr15	chr15:10420640-10668160	Rai14	
chr15	chr15:37241795-37474132	Tcfcp2l3	
chr15	chr15:38322369-38428606	Klf10	
chr15	chr15:62150297-62255314	Myc	
chr15	chr15:76944649-77046213	FoxH1	
chr15	chr15:80170049-80310237	Cbx7	
chr15	chr15:93650232-93802057	Yaf2	
chr15	chr15:100639511-100772685	Atf1	
chr15	chr15:102837955-102964000	Sp1	
chr15	chr15:103351435-103573764	whole Hoxc	
chr16	chr16:15349099-15462234	Mcm4	
chr16	chr16:22029998-22150096	Sfrs10	
chr16	chr16:22164205-22322399	Etv5	
chr16	chr16:24625200-24725200	miR28	
chr16	chr16:48137444-48273389	Dppa4	Dppa2
chr16	chr16:85287109-85606017	App	
chr16	chr16:91609775-91756171	Oligo2	Oligo1
chr16	chr16:96375440-96493458	Ets2	
chr17	chr17:7792410-7900483	brachyury	
chr17	chr17:12037798-12288727	Igf2r/Air	
chr17	chr17:22145756-22327344	Zn206	Cldn6
chr17	chr17:23216411-23328674	Zfp598	
chr17	chr17:25376258-25479203	Nkx2.5	
chr17	chr17:27649155-27750510	p21	
chr17	chr17:30671414-30818280	Notch3	
chr17	chr17:33689346-33801186	Bat1a	
chr17	chr17:33965652-34076734	Pou5f1	Tcf19
chr17	chr17:35475987-35583584	Zfp57	
chr17	chr17:45708714-45814812	Cyclin D3	
chr17	chr17:48408531-48522729	Dazl	
chr17	chr17:54471290-54731109	miR7b	Jmjd2b
chr17	chr17:69223035-69330557	Tgif	
chr17	chr17:86221544-86373483	Msh2	
chr18	chr18:11226033-11357301	Gata6	
chr18	chr18:16909144-17228997	Cdh2	
chr18	chr18:21017614-21126518	Ttr	
chr18	chr18:35214792-35328913	Jmjd1b	
chr18	chr18:35508924-35744775	Catna1	
chr18	chr18:69819951-70259952	Tcf4	
chr18	chr18:71046188-71204011	Mbd2	
chr18	chr18:74115634-74280361	Smad4	
chr18	chr18:75848753-75977151	Smad7	
chr18	chr18:76302444-76888598	Zbtb36	Smad2
chr18	chr18:77435591-77625178	Miz1	

chr18	chr18:85397012-85543379	Fbox15
chr19	chr19:5376638-5487037	RelA
chr19	chr19:6932089-7038651	coREST
chr19	chr19:13676671-13924397	Tle4
chr19	chr19:18165595-18323312	Rorb
chr19	chr19:22364224-22488426	Klf9
chr19	chr19:24730514-24929136	Dmrt1
chr19	chr19:38107560-38244847	Hells
chr19	chr19:42899051-43002549	Nkx2.3
chr19	chr19:44523008-44625453	Lbx1
chr19	chr19:56164644-56307162	Tdrd1
chrX	chrX:5877689-5992230	Syp
chrX	chrX:6198239-6306794	Gata1
chrX	chrX:29801083-29905847	Ehox
chrX	chrX:45462380-45595902	Hprt
chrX	chrX:49633109-49738369	Zic3
chrX	chrX:52408425-52509552	Sox3
chrX	chrX:64495744-64639927	Zfp185
chrX	chrX:65535214-65693977	MeCP2
chrX	chrX:77692687-77796846	Nr0b1
chrX	chrX:85576160-85724643	Zfx
chrX	chrX:92783268-92902166	Nono
chrX	chrX:93601021-93705348	Cited1
chrX	chrX:94765470-94929759	Xist/Tsix
chrX	chrX:126171135-126272600	Rex3
chrX	chrX:127072559-127177333	Esx1
chrX	chrX:142070723-142211219	Jarid1c
chrUn_random	chrUn_random:86135778-	miR290-295
chrUn_random	chrUn_random:95380040-	Obox5, Obox2