The Role of the Histone Demethylase Jmjd2c/Kdm4c in Mouse Embryonic Stem Cell Pluripotency

by

Rute Alexandra da Costa Tomaz

Thesis submitted towards fulfilment of the degree of

Doctor of Philosophy

Institute of Reproductive and Developmental Biology

Faculty of Medicine

Imperial College London

Statement of Originality

All experiments included in this thesis were performed by the author unless otherwise stated throughout the text and figure legends.

Notably, Dr Cynthia Fisher (Wellcome Trust Institute, Cambridge, UK) generated the knockout embryonic stem cells (ESCs) used in this work. Lauren Weavers and Donja Karimlou contributed to this project during the completion of their Master degrees under my joint supervision with Dr Véronique Azuara, generating novel ESC lines and performing some of the experiments described in Chapter III. Moreover, other past laboratory members (Sergio German, Dr Anne Helness and Dr Tony Bou-Kheir) performed specific chromatin immunoprecipitation experiments described in Chapters III and V, and Dr Matias Autio derived the mouse embryonic fibroblast cultures utilised in Chapter IV. Jennifer Harman conducted the mesoderm differentiation experiments described in Chapter IV, under the supervision of Dr Helle Jørgensen (Cambridge University, UK). Dr Lauriane Fritsch (Paris VII University, France) preformed the co-immunoprecipitation experiments outlined in Chapter V, and Dr Ignacio del Vale Torres (NIMR Francis Crick Institute, London) performed the initial data processing of raw ChIP-sequencing reads generated during this project.

Copyright Declaration

Abstract

The development of a whole organism from a single fertilised egg is a tightly orchestrated process at the molecular level. Signalling pathways stimulate the expression of transcription factors that in turn coordinate the activation, repression or priming of genes through multi-protein complex assembly and chromatin remodelling at regulatory regions during key developmental transitions.

This project focuses on the role of the chromatin modifier Jmjd2c/Kdm4c in embryonic stem cell (ESC) pluripotency. Jmjd2c is a member of Jmjd2 H3K9-demethylase family, which is highly expressed in ESCs and during early mouse development, and was proposed to contribute to ESC identity. Here, *Jmjd2c*-depleted ESCs were generated and found to sustain an undifferentiated state. In contrast, these cells fail to execute multi-lineage differentiation, as evidenced by their inability to fully activate appropriate gene expression programs, but readily adopt an extra-embryonic endoderm-like phenotype under appropriate conditions. Moreover, *Jmjd2c*-knockout ESCs can transit to an early epiblast stage and be stably converted into epiblast stem cells (cEpiSCs) that harbour an immature state as highlighted by a lack of transcriptional gene priming for germ layer markers.

In order to unfold the mechanistic basis for the observed differentiation defect, genome-wide distribution of Jmjd2c was analysed by ChIP-sequencing. This revealed prominent binding to the TSS of active and bivalently marked genes, concordant with recent reports. Remarkably, acquisition of multi-lineage priming in ESCs was accompanied by Jmjd2c recruitment at poised enhancers of lineage-affiliated genes. Surprisingly, these regions were co-occupied by the H3K9-methyltransferase G9a/Ehmt2, which physically interacts with Jmjd2c and Mediator, within chromatin-bound complexes. Critically, the absence of Jmjd2c in cEpiSCs was sufficient to destabilize the assembly of enhancer protein complexes, suggesting a scaffolding role independent of its H3K9-demethylase activity. Collectively, this study uncovered a novel role for Jmjd2c in regulating gene expression at exit of the pluripotent state.

Acknowledgements

I would first like to thank my supervisor Dr Véronique Azuara for the excellent supervision and endless encouragement at every step of the way. Her motivation certainly led this project to its full potential giving me an invaluable experience for which I am grateful. I would also like to thank my co-supervisor Dr Tristan Rodriguez for the fruitful discussions and advice, which were much appreciated. I am grateful to the Fundação para a Ciência e a Tecnologia who supported the costs of my PhD, without whom this would have not been possible.

I am thankful to all the collaborators that have in any way helped pushing this project forward with their skills and expertise: Jennifer Harman and Dr Helle Jørgensen for the extraordinary and demanding mesoderm differentiation experiments, Dr Lauriane Fritsch for the performing the immunoprecipitations, Dr Cynthia Fisher for generating the cell lines here used, Dr Ignacio del Vale Torres and Dr Kathy Niakan for their contribution in the sequencing analysis and help in the setup of XEN cell differentiation assays. I also thank Van Nguyen and Dr Luca Magnani for teaching me how to generate sequencing libraries, and Mike Dekker and Dr Raymond Poot for kindly hosting me in their lab in Rotterdam and showing me how to perform pull down assays, and all other laboratories which contributed with materials or reagents. I am in debt to Emma Bell and Ed Curry for introducing me to the incredible world of bioinformatics, and to Donja Karimlou and Lauren Weavers who I had the pleasure to supervise during their Master degrees. I am extremely thankful to all the first floor colleagues that I have crossed paths with, who welcomed me in the lab (you have not been forgotten), and shared their knowledge and companionship. I am also very glad for having had the luck of being surrounded by beloved “pals” across all floors, who have made every day of this PhD a very joyful one.

Finally, I am forever grateful for the incredible support of my loving partner and flatmates, and last but not least, for the constant encouragement and unconditional love from my adored family, friends and “sestras” back home, who I have missed dearly. Thank you.

Table of Contents

List of Figures and Tables

Figures

Chapter I

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

Appendix A - Figures

Tables

Chapter I

Appendix B - Tables

Abbreviations

2i	2 inhibitors
5hmC	5-hydroxymethylated cytosine
5mC	5-methylated cytosine
AP	Alkaline phosphatase
AVE	Anterior visceral endoderm
atRA	All trans retinoic acid
bFGF	Basic fibroblast growth factor
bp	Base pair
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
°C	Celsius
CAG	Chicken beta-actin
cDNA	Complementary DNA
cEpiSC	Converted epiblast stem cell
ChIP	Chromatin immunoprecipitation
Co-IP	Co-immunoprecipitation
CpG	Cytosine-guanine dinucleotide
Ct	Cycle threshold
CTD	C-terminal domain
cXEN	Converted extra-embryonic endoderm
DAPI	4',6-diamidino-2-phenylindole
DE	Definitive endoderm
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DNMT	DNA methyltransferase
dNTP	Deoxynucleoside triphosphate
Dox	Doxycycline
DPBS	Dulbecco's phosphate-buffered saline
DTT	Dithiothreitol
DUB	Deubiquitylases
E	Embryonic day
EB	Embryoid body
ECL	Enhanced chemiluminescence
Ecto	Ectoderm
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene Glycol Tetraacetic Acid
EG	Embryonic germ
En2	Engrailed 2
Epi	Epiblast
EpiLC	Epiblast-like stem cell
EpiSC	Epiblast stem cell
EPC	Ectoplacental cone
ERK	Extracellular signal-regulated kinase
ESC	Embryonic stem cell
EtOH	Ethanol
EUCOMM	European conditional mouse mutagenesis
ExE	Extra embryonic ectoderm
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FDR	False discovery rate
Fgf	Fibroblast growth factor

Fgfr	Fibroblast growth factor receptor
Fig	Figure
Flp	Flippase
FRT	Flippase recognition target
GEO	Gene Expression Omnibus
GDF	growth and differentiation factors
GFP	Green fluorescent protein
GLP	G9a-like protein
GMEM	Glasgow modified eagle medium
GREAT	Genomic regions enrichment of annotations tool
GSK	Glycogen synthase kinase
H	Hour
H1.4K26me3	Tri-methylation of lysine 26 of histone 1 variant 4
H3K27ac	Acetylation of lysine 27 of histone H3
H3K27me3	Tri-methylation of lysine 27 of histone H3
H3K4me1/2/3	Mono/di/tri-methylation of lysine 4 of histone H3
H3K9me2/3	Di/tri-methylation of lysine 9 of histone H3
HAT	Histone acetyltransferase
HDAC	Histone deacetyltransferase
HD-EB	Hanging-drop Embryoid body
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOMER	Hypergeometric Optimization of Motif EnRichment
HRP	Horseradish peroxidase
ICM	Inner cell mass
Id	Inhibitor of differentiation
IF	Immunofluorescence
IgG	Immunoglobulin

IKMC	International knockout mouse consortium
iMEF	Irradiated mouse embryonic fibroblast
iPSC	Induced pluripotent stem cell
IRES	Internal ribosome entry site
JAK	Janus kinase
kDa	Kilo Dalton
KDM	Histone lysine demethylase
KMT	Histone lysine methyltransferase
KO	knockdown
KOSR	KnockOut serum replacer
lacZ	β-galactosidase
LB	Luria-Bertani broth
LIF	Leukemia inhibitory factor
LPM	Lateral plate mesoderm
LRP	Low-density lipoprotein receptor protein
Luc	Luciferase
MACS	Model-based Analysis of ChIP-Seq
MAPK/MEK	Mitogen-activated protein kinase
Meso	Mesoderm
MEF	Mouse embryonic fibroblast
MLL	Mixed-Lineage Leukemia
mRNA	Messenger RNA
MP	Mesoderm progenitors
MTL	multiple transcription factor-binding loci
NP-40	Nonidet P-40
NuRD	Nucleosome Remodeling Deacetylase
OSKM	Oct4-Sox2-Klf4-cMyc

OSN	Oct4-Sox2-Nanog
P	P-value
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Parietal endoderm
PFA	Paraformaldehyde
PGC	Primordial germ cell
PHD	Plant homeodomain
PM	Paraxial mesoderm
PMSF	Phenylmethanesulfonylfluoride
polyA	Poly adenylation
POU	Pit-Oct-Unc
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive compelx 2
PrE	Primitive endoderm
Puro	Puromycin
PVDF	Polyvinylidene fluoride
qPCR	Quantitative PCR
RAF	Rapidly Accelerated Fibrosarcoma
RIPA	Radio Immuno Precipitation Assay
RNA	Ribonucleic acid
RNAi	RNA interference
RNAPII	RNA polymerase II
RT-PCR	Reverse transcription PCR
sA	Splice acceptor
SCNT	Somatic cellular nuclear transfer
SD	Standard deviation

SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
shRNA	Small hairpin RNA
SEM	Standard error of the mean
Ser2/5P	Phosphorylated serine 2/5
siRNA	Small interfering RNA
SN	Supernatant
SRY	sex-determining region Y
STAT	Signal Transducer and Activator of Transcription
SV40	Simian virus 40
TBE	Tris-borate-EDTA
TBS-T	Tris-buffered saline with 0.1% Tween-20
TE	Trophectoderm
TET	Ten-eleven translocation
TGF	Transforming growth factor
TrxG	Trithorax group
TS	Trophoblast stem (cell)
TSS	Transcription start site
USCS	University of California Santa Cruz
v/v	Volume/volume
w/v	Weight/volume
XEN	Extra-embryonic endoderm cell

Chapter I - Introduction

I.1 Preface

For the purpose of this work, the mouse organism was used as a model system. In particular, the similarity between mouse and human early embryonic development are significant, making this a useful model to address several question regarding the regulation of this process. Generating a whole organism from a single fertilised egg is a remarkable process. Thus, at a molecular level this process must be perfectly well-orchestrated in order to ensure proper generation of specializes tissues.

This project focuses on the specific role of the chromatin modifier Jmjd2c in early embryonic development, using mouse embryo-derived stem cells and *in vitro* differentiation assays as tools to dissect the impact of this protein in the early cell fate choices. In this Chapter, the mouse embryonic development will be described with a specific focus on the pre-implantation and early post-implantation stages. Furthermore, the ability to derive stem cell lines from the mouse blastocysts and the knowledge we have thus far gained on the complex extrinsic signaling cues and gene regulatory mechanisms that coordinate cell fate choices, using *in vivo* and *in vitro* systems, will be discussed. Finally, for the scope of this project, attention will be given to our current understanding on the mechanisms of epigenetic regulation and long-range interactions that govern and safeguard gene expression upon embryonic lineage commitment.

I.2 The mouse as a model system

Over the past century, the mouse has been recognised as a leading mammalian model system for the study of molecular and cellular processes relevant to human biology. Because of the high degree of similarly between genetic and physiological circuitries, mice are a prime model for the study of the immune, nervous, cardiovascular, endocrine, skeletal, reproductive and further complex physiological processes shared among mammals. Moreover, with the completion of the full mapping of the Human and Mouse genomes (Lander, 2011; McPherson et al., 2001; Mouse Genome Sequencing et al., 2002; Waterston et al., 2002), we now know that despite 75 million

years since the divergence of the human and mouse lineages, about 99% of the mouse genes share a homolog in the human genome.

The rapid advances in sequencing technologies have had an enormous impact on the in-depth sequencing and decoding of genomic information. Still, we are far from knowing how all genes operate and which processes they regulate in a physiological scenery. A great effort to answer these question has been attained by genetic manipulation and phenotyping of mouse models (Bult et al., 2013).

Here, the isolation and propagation of pluripotent mouse embryonic stem cells (ESCs) in the early 1980s has triggered a major advance in mouse genetics with the emergence of gene targeting technology (Evans and Kaufman, 1981; Thomas and Capecchi, 1987). Thus, international consortia have ever since been established employing genetic engineering tools that permit deleting (knockout) or inserting altered versions of the same gene (knockin), for interrogation of virtually any gene function in whole-organisms (Skarnes et al., 2011).

Beyond being easily manipulated, the mouse has physiological processes highly similar to humans and a short breeding cycle. Thus, studying the mouse embryonic development overcomes ethical and technical obstacles encountered in the study of human embryology. Mammalian development is an extraordinary well regulated phenomena where the fusion of two highly specialized cells – egg and sperm – generate a complex multicellular organism. The study of early mouse development and derived stem cells, which has deserved much attention in the past decades, has provided instrumental insights into the molecular and cellular circuitry behind the earliest cell fate choices, a knowledge which has been invaluable not just for developmental biology, but also for regenerative medicine.

I.3 - Early mammalian development

1.3.1 Mouse pre-implantation embryo development

In mammals, fertilization occurring between two haploid gametes – egg and sperm – forms a diploid zygote. This totipotent cell contains the genetic information from both entities and holds

the ability to give rise to all cells of the developing organism. Through a series of sequential cell divisions that characterize the early stages of embryonic development up to uterine implantation, the blastocyst is formed, comprising three distinct cell populations: the trophectoderm (TE) that will contribute to the placenta formation, the primitive endoderm (PrE) that will give rise to the yolk sac, and the primitive ectoderm, or epiblast (Epi), comprising pluripotent progenitors that will generate the proper embryo (Fig.1.1A) (Bedzhov et al., 2014).

The specification and formation of these distinct blastocyst lineages is accomplished after the two major events of cell fate decisions occurring at embryonic day (E) 2.5 and E3.5. First, two rounds of asymmetric cell divisions between 8 to 16, and 16 to 32-cell stage, lead to the segregation of two distinct populations in morula-stage embryos depending on their location: polar cells located on the outside will go on to form the TE, whereas inner apolar cells will form the inner cell mass (ICM; Fig.1.1A) (Morris et al., 2010). This process is highly dependent on cell polarity, position and cell–cell signalling, being also influenced by molecular heterogeneity between blastomeres (Bedzhov et al., 2014). Although epigenetic asymmetry arises early on between maternal and paternal gametes, its role in pre-patterning of blastomeres still remains to be elucidated (Takaoka and Hamada, 2014). By the 32-cell stage the embryo cavitates with full specification of the TE, the orientation of which will define the embryonic–abembryonic axis of the embryo. The second event of cell decision occurs at the 32 to 64-cell transition where the ICM further segregates into the PrE and the epiblast by E4.5 (Fig.1.1A). Cell lineage tracing techniques have unveiled that before this segregation, ICM blastomeres are already engaged into either PrE or Epi, and this bias occurs in a heterogeneous “salt and pepper” manner, concomitant with acquisition of exclusive expression of PrE and Epi-specific markers Gata6 and Nanog, respectively (Chazaud et al., 2006; Rossant et al., 2003). Moreover, this bias is also thought to be influenced by both the timing of internalization within the morula (Krupa et al., 2014), and the expression of Fgfr2 or Fgf4, with cells expressing Fgf4 correlating with cells internalized earlier, and hence, Nanog-expressing epiblast progenitors, and Fgfr2 cells internalized later corresponding to Gata6-expressing PrE progenitors (Fig.1.1B) (Morris et al., 2013).

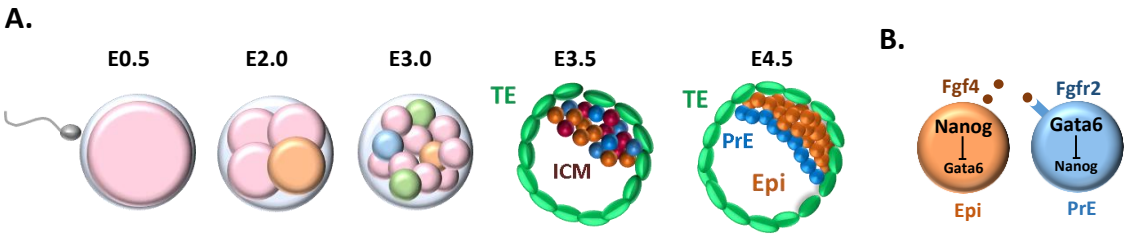


Figure 1.1 Pre-implantation development of the mouse embryo

A. After fertilization, the zygote forms and divides through the early cleavage stages of mammalian development up to the morula (E3.0). Embryonic and extra-embryonic lineages cells are specified by two cell fate decisions. First, waves of cell divisions segregate inside (pink) and outside (green) layers of cells with outside cells later giving rise to extra-embryonic trophectoderm (TE), while inside cells tend to preferentially form the pluripotent inner cell mass (ICM). Secondly, at the early blastocyst stage (E3.5) the ICM segregated into the extra-embryonic primitive endoderm (PrE; blue) and the epiblast (Epi; orange) progenitors, with further maturation and compartmentalization of these lineages at the late-blastocyst stage (E4.5), just before implantation. Cell fate decisions are influenced by their inner-outer location and molecular heterogeneity between individual blastomeres since the 4-cell stage (E2.0) as represented by the different colours. **B.** Epiblast progenitor (orange) cells, which are internalized earlier, express Fgf4 and Nanog, while PrE progenitors (blue) cells, internalized later, express higher levels of Fgf4 receptor Fgfr2 and Gata6. Nanog and Gata6 reciprocally inhibit their expression, with Fgf signalling also contributing the maintenance of Gata6 by inhibition of Nanog-mediated repression. Adapted from (Bedzhov et al., 2014).

As the embryo development progresses, the developmental potency of cells becomes restricted. Early cleavage state embryos are totipotent, as they are able to form both embryonic and extra-embryonic lineages of the developing embryo. As the TE and PrE emerge, these extra-embryonic cells have become multipotent as they are limited to form the placenta and yolk sac, respectively, where the pluripotent epiblast will give rise to the three germ layers (endoderm, mesoderm and ectoderm) and germ cells upon gastrulation, ultimately forming the proper embryo. However, despite early commitment, a small percentage of ICM cells isolated from the early blastocyst can still contribute to TE when aggregated into morula, showing that some level of plasticity is retained at this stage (Morgani and Brickman, 2014). Upon the second lineage

decision cell plasticity in the ICM/epiblast is then lost (Boroviak and Nichols, 2014). ICM plasticity has been suggested to provide a back-up mechanism for preserving the appropriate balance between the number of ICM and PrE cells, vital for subsequent post-implantation development (Bedzhov et al., 2014).

1.3.2 Mouse early post-implantation development

Upon implantation, the blastocyst attaches the zona pelucida where the TE invasion merges the blastocyst to the maternal uterine wall, forming the ectoplacental cone (EPC). Embryo growth and morphogenesis arise upon successful implantation, being a crucial step for proper development. Thus, following implantation, the epiblast undergoes a major rearrangement with the emergence of a single layered epithelium that forms an organized a rosette-like cylinder, enclosing a pro-amniotic cavity, in a manner highly influenced by polarization (Bedzhov and Zernicka-Goetz, 2014). The TE-derived extra-embryonic ectoderm (ExE) and the Epi become surrounded by PrE-derived parietal endoderm (PE) and visceral endoderm (VE) layer (Fig.1.2). The VE has a crucial role in anterior-posterior patterning of the embryo. In particular, a sub-population emerging at the tip of the egg cylinder migrates to the distal side of the embryo to form the anterior visceral endoderm (AVE), influencing the patterning of the newly formed primitive streak – the site of gastrulation, which emerges on its opposite side (Stower and Srinivas, 2014; Tam and Loebel, 2007). Correct formation and placement of the primitive streak is dependent on reciprocal signalling cues between Epi, ExE, and AVE, respectively. BMP4 signalling from the ExE stimulates gene expression programs and Wnt3a production in the posterior epiblast, in addition contributing to primitive germ cell (PGC) specification (Gunesdogan et al., 2014). Conversely, Wnt-mediated Nodal signalling from the posterior epiblast safeguards BMP4 signalling activity from the ExE. Additionally, asymmetry in Nodal signalling guides AVE migration towards the site of expression of its antagonists, which is necessary for pre-streak positioning (Stower and Srinivas, 2014).

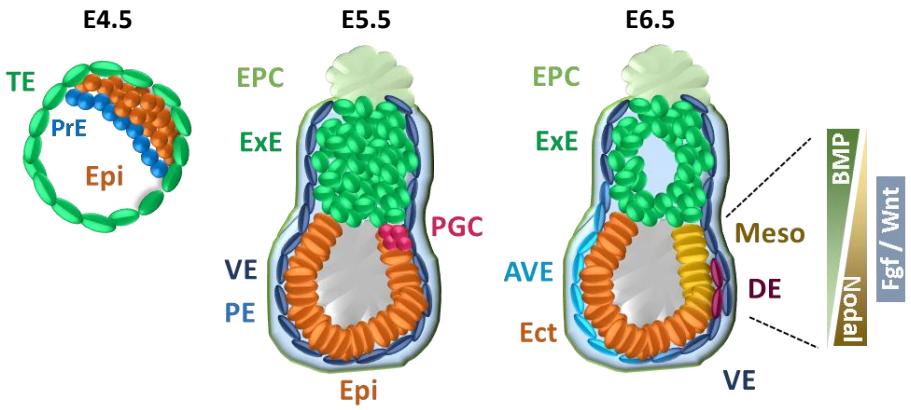


Figure 1.2 Early post-implantation development of the mouse embryo

Before implantation (E4.5), three distinct lineages in the early embryo have been segregated: trophectoderm (TE, green), primitive endoderm (PrE, blue) and epiblast (Epi, orange). Upon implantation (E.5.5), the polar TE (side attached to the Epi), invades the maternal tissues forming an ectoplacental cone (EPC, light green) and extra-embryonic ectoderm (ExE, green). The PrE colonizes the mural TE (opposite side of Epi) and segregates into parietal endoderm (PE, light blue) and visceral endoderm (VE, dark blue), surrounding the epiblast, primordial germ cells (PGC, pink) and the pro-amniotic cavity (grey). Upon gastrulation, the anterior visceral endoderm (AVE) is established, and epiblast cells are recruited to the posterior region (primitive streak), where they will form mesoderm (Meso, brown) and definitive endoderm (DE, purple). Epiblast cells retained at the anterior side will form the ectoderm (Ect, orange). Signalling gradients (BMP, Nodal) and activity of Fgf and Wnt pathways will influence the emergence of subsequent sub-populations in the primitive strike (Arnold and Robertson, 2009; Mendjan et al., 2014; Tam and Loebel, 2007).

Upon gastrulation, epiblast cells ingress into the primitive streak, where they develop the mesoderm layer and additionally are integrated into the surrounding VE layer as definitive endoderm (DE). Thus, prior to this point, epiblast cells transiently exist in a mesendoderm state, having the potential to become either mesoderm or endoderm. Cells that are retained in the epiblast anterior site give rise to the ectoderm, and later on the surface ectoderm, which includes the epidermis and the neural tissues (Tam and Loebel, 2007).

Mesoderm cells are subsequently compartmentalised depending on the time and location of ingression within the strike. The earliest and most posterior compartment forms the extra-embryonic mesoderm (yolk sac vasculature and blood), followed by the lateral plate (circulatory system and gut wall), paraxial (somites) and cardiac mesoderm in the intermediate and anterior regions. Finally, at the anterior tip, the axial mesoderm (notochord) is formed as well as the DE (gut tube and organs). Notably, the emergence and specification of definitive endoderm and mesoderm sub-populations is regulated by Fgf and Wnt signalling, as well as by reciprocal gradients of BMP and Nodal along the anterior-posterior axis of the streak (Fig.1.2) (Arnold & Robertson, 2009; Tam and Loebel, 2007).

I.4 Derivation of stem cell populations

1.4.1 Isolation of stem cell populations from the early embryo

A milestone in the developmental biology field has been the ability to isolate cell populations from the early blastocyst and stably culture these *in vitro*, while maintaining the developmental potency of their cognate lineages (Rossant, 2008). Blastocyst-derived stem cells provide powerful tools not only to study the precise molecular mechanisms underlying lineage commitment in the early embryo, but also hold great promise for the regenerative medicine and disease modeling fields.

Since the 1950s, sporadically acquired teratocarcinomas in mice have been studied in regards to their transplant ability. This property was attributed to the presence of embryonal carcinoma (EC) cells, a proliferative undifferentiated group of cells which not only could regenerate all of the cell types forming teratocarcinomas, but also self-renew *in vitro*. The capture and identification of culture conditions (serum and a feeder layer of mitotically inactivated fibroblasts) under which this population could retain its pluripotency *in vitro* was a critical landmark for the later derivation of embryonic stem cells (ESCs) (Kahan and Ephrussi, 1970; Martin and Evans, 1975). In the early 1980s, two groups independently derived ESCs from pre-implantation blastocysts. These cells could be expanded indefinitely while retaining their ability

to self-renew and to contribute to the germline in addition to forming all somatic cell types when re-introduced back into an embryo environment (Fig.1.3) (Evans and Kaufman, 1981; Martin, 1981). Later on, their self-renewal requirements for serum and feeders were found to be substitutable by BMP4 and Leukemia inhibitory factors (LIF), revealing these as key signaling pathways in promoting self-renewal and blocking differentiation cues (Niwa et al., 1998; Ying et al., 2003a).

Although ESCs are commonly derived from the ICM of early blastocysts, they have also been successfully established from any pre-implantation stage and even from single blastomeres (Chung et al., 2006; Tesar, 2005), acquiring very similar characteristics. This suggested that the derivation process allows cells to progress to a common developmental stage in which *in vitro* self-renewal can arise. Moreover, when challenged to differentiate either *in vivo* or *in vitro*, ESCs proved capable of also contributing to extra-embryonic lineages, hence suggesting that ESC cultures might be functionally heterogeneous as they encompass individual cells that are primed for an Epi or PrE fate (Morgani and Brickman, 2014). Consequently, the true *in vivo* counterpart of cultured ESCs has been and remains an object of debate (Fig.1.3) (Plusa and Hadjantonakis, 2014).

More recently, derivation of pluripotent stem cell lines from the epiblast of post-implantation stage embryos has also been achieved (Fig.1.3) (Brons et al., 2007; Tesar et al., 2007). Although epiblast stem cells (EpiSC) are self-renewing and possess similar pluripotent properties of ESCs, they are unable to contribute to chimaeras when introduced into a host blastocyst (Guo et al., 2009), unless when implanted in the post-implantation stage (Huang et al., 2012). Moreover, signaling pathway requirements for the maintenance of their self-renewal ability are different from ESCs, involving Activin/Nodal and Fgf signaling as opposed to BMP4 and LIF signaling pathways (Shen, 2007). Thus, it is considered that ESCs represent an immature or so-called naïve state of pluripotency, which resembles that of the pre-implantation epiblast and is characterized by the ability to give rise to all somatic cell types and germline competency. Conversely, EpiSCs are considered to represent a more advanced state of pluripotency, which is primed for

differentiation, as seen in the post-implantation epiblast lineage *in vivo* (Huang et al., 2012). Similarly to ESCs, derivation of EpiSCs from post-implantation E5.5 to E8.5 allows cells to progress to a common self-renewing stage *in vitro* (Kojima et al., 2014).

In the late 1990s, the first human ESCs were derived from pre-implantation embryos donated through in vitro fertilization (IVF) programs (Thomson et al., 1998). Like mouse ESCs, hESCs can contribute to teratoma formation when injected into mice, but due to ethical and practical reasons their ability to form chimeras cannot be tested. Yet, striking differences in morphology, growth factor requirement and epigenetic profiles between hESCs and mESCs later revealed that hESCs might in fact be more closely related to the primed mouse EpiSC state (Fig.1.3) , and hence might have a different embryonic origin from mouse ESCs (Tesar et al., 2007). These differences prompt the still ongoing quest for discovery of conditions that allow derivation and/or conversion of hESCs into a "naïve" state of pluripotency (Fonseca et al., 2015).

Another pluripotent stem cell population has been derived from the mouse embryo. The Embryonic Germ (EG) cells can be established from primordial germ cells (PGCs) in the developing gonads of E8.5 to E11.5 post-implantation embryos and re-acquire the expression of pluripotency genes and similarly contribute to chimera formation (Labosky et al., 1994; Matsui et al., 1992; Resnick et al., 1992).

In addition to pluripotent stem cell lines derived from the ICM/epiblast, extra-embryonic stem cell lines have also been successfully established. Here, extra-embryonic endoderm (XEN) stem cells derived from the PrE and trophoblast stem (TS) cells derived from the TE, have a restricted developmental potency as they can only contribute to the parietal endoderm or trophoblast lineages, respectively, when injected back into the embryo (Fig.1.3) (Kunath et al., 2005; Tanaka et al., 1998). Unlike ESCs, TS cells rely on Fgf and Nodal signaling for their self-renewal, whereas XEN require Fgf signaling for derivation only (Rossant, 2008).

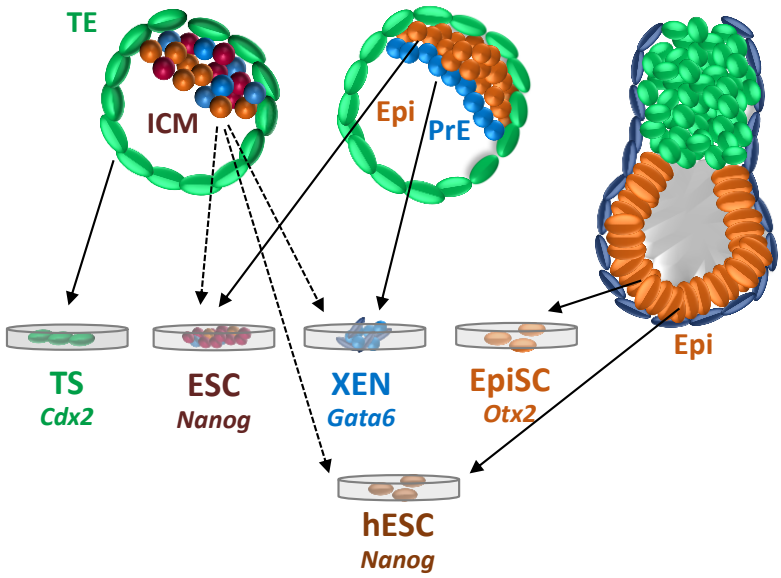


Figure 1.3 Stem cell populations derived from blastocyst-stage embryos

Schematic representation of the association between derived stem cell cultures and their *in vivo* counterparts. Three stem cells lines can be derived from the mouse pre-implantation blastocyst outgrowths: trophoblast stem cells (TS) fare derived from the trophectoderm (TE), embryonic stem cells (ESC) and extra-embryonic endoderm (XEN) are derived from the inner cell mass (dotted line) and represent the epiblast (Epi) and the primitive endoderm (PrE) lineages, respectively (full line). Epiblast stem cells (EpiSCs) are derived from post-implantation epiblasts. Human embryonic stem cells (hESC) are also derived from the ICM of pre-implantation blastocysts (dotted line) but are thought to represent the post-implantation epiblast in culture (full line). Markers indicated underneath are transcription factors associated with each of the respective lineages.

1.4.2 Pluripotent stem cell heterogeneity and plasticity

The establishment of embryonic and extra-embryonic cell lines have provided powerful tools for the dissection of the molecular mechanisms underlying lineage choices in the early embryo. In particular, the plasticity of ESCs has been exploited as an *in vitro* system to uncover regulatory switches between pluripotent naïve and primed states, different differentiated (ectoderm, endoderm and mesoderm) states, or even across embryonic and extra-embryonic fates (e.g. ESC conversion into TS or XEN cell states).

Standard conditions (serum and LIF) utilised to isolate and culture undifferentiated ESCs derived from the ICM have been proposed to support distinct cell sub-populations based on the observed fluctuations in the expression of a variety of pluripotency-associated factors (Torres-Padilla and Chambers, 2014). In particular, oscillations in the expression of *Rex1* associated with *Oct4* defined an ICM-like (*Rex1*+) versus an early primitive ectoderm-like (*Rex1*-) sub-population in culture. Moreover, the later could differentiate more efficiently to somatic lineages and showed poor ability to contribute to chimera formation (Toyooka et al., 2008). Similarly, Nanog's expression is found to be highly heterogeneous in the presence of serum and to associate with a naïve pluripotency state (Chambers et al., 2003; Kalmar et al., 2009). Cells with low levels of Nanog correlate with moderate expression of primitive-endoderm markers such as *Gata4* and *Hex1* (Canham et al., 2010; Singh et al., 2007) and the epiblast marker *Fgf5* (Kalmar et al., 2009). Additionally, *Nanog*-low cells also exhibit a more pronounced differentiation potential when exposed to differentiation cues *in vitro*.

Moreover, fluctuations in the expression of non-pluripotency markers, which include lineage-affiliated markers, has also been identified in culture (Fig.1.4). One example is the aforementioned extra-embryonic associated factor *Hex1*, in which sub-populations with higher levels of expression were shown to increase the contribution into extra-embryonic endoderm when introduced back in early blastocysts (Canham et al., 2010). Another example is *Otx2* – a transcription factor essential for brain development, which was found to be required to maintain the ESC metastable state, by promoting commitment to differentiation and essential for ESC transition into EpiSCs (Acampora et al., 2013).

Heterogeneity has long been defined as a hallmark of ESCs, where the constant switch between metastable states might provide a mechanism that safeguards the rapid response to differentiation cues while retaining the self-renewal ability. However, the degree of heterogeneity has been proposed to be associated with the uncontrolled signalling environment that might be provided by the serum in culture. Thus, efforts to identify novel chemically defined derivation and culture conditions have led to the discovery of 2 key inhibitors (2i) for Erk and Gsk3

pathways, which robustly block differentiation signals and lock ESCs in a so-called "ground" state of pluripotency (Ying et al., 2008). This translates into a more homogeneous colony morphology, transcriptome and epigenetic features, with reduced transcriptional gene priming for lineage-affiliated genes (Fig.1.4). Nevertheless, despite higher homogeneity in the expression of pluripotency markers, ESCs cultured in 2i still express heterogeneous levels of extra-embryonic markers such as *Hex1*, *Gata6* and even *Cdx2*, and gain enhanced capacity of contributing to extra-embryonic tissues (Fig.1.4) (Canham et al., 2010; Morgani et al., 2013). Thus, heterogeneity is likely to be not only an extrinsic factor, but also an intrinsic feature of uncommitted ESCs, the extent of which can be manipulated by the culture conditions.

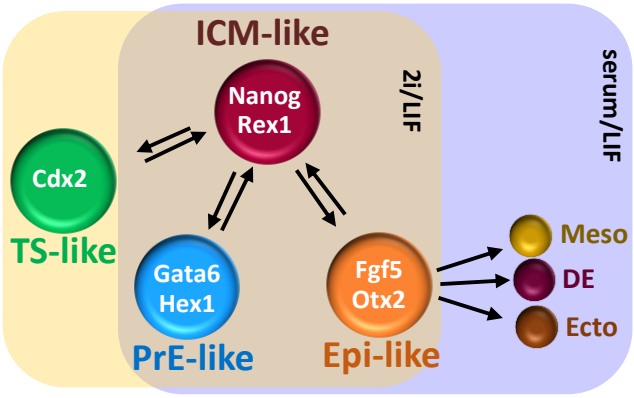


Figure 1.4 Equilibrium of ESCs sub-states in different culture conditions

Schematic representation of potential cell sub-types that coexist in 2i/LIF (yellow shade) and serum/LIF ESC cultures (purple shade). ICM-like "naïve" ESCs are represented in purple expressing high levels of pluripotency-associated factors *Nanog* and *Rex1*. Totipotent cells expressing low levels of TS marker *Cdx2* in 2i/LIF are represented in green. PrE precursors are indicated in blue, expressing low levels of markers such as *Hex1* and *Gata6*. Epi precursors are indicates in orange, expressing low levels of post-implantation Epi markers of *Fgf5* and *Eomes*. Moreover, in serum/LIF conditions, background expression of lineage-affiliated genes is also detected Adapted from Morgani et al., 2013.

Cell heterogeneity is also a defining feature of "primed" post-implantation derived EpiSC cultures. Sub-states have been identified by the heterogeneous levels of the pluripotency marker Oct4. Here, using an engineered EpiSC population expressing a GFP-coupled *Oct4* transgene, it

was found that a sub-population expresses high levels of *Oct4*, which correlated with a transcriptome and epigenetic features that resembled those of most immature (E5.5) post-implantation epiblast, notably devoid of transcriptional gene priming for germ layer markers. Remarkably, this minor fraction could readily contribute to chimera formation and could revert back to an ESC state more efficiently. EpiSCs expressing lower levels of *Oct4*-GFP were found to have a transcriptional profile aligned with later stage of epiblast maturation, and concordantly, could not form chimeras when injected into a host blastocyst. As seen in ESC cultures, these two EpiSC sub-populations can be interconverted (Han et al., 2010), and the degree of *Oct4* heterogeneity in culture can be attenuated by addition of Fgf4 in culture (Joo et al., 2014).

Furthermore, PGCs derived from post-implantation epiblasts are unipotent in the sense that their developmental fate is uniquely biased towards the formation of germ cells. However, and as mentioned above, these cells can also generate in culture pluripotent ES-like cells (Matsui et al., 1992). ESCs converted into an early epiblast-like cell (EpiLC) state can also generate functional PG-like cells in response to BMP4 and cytokines, whereas this process proved inefficient when using embryo-derived EpiSC cultures (Hayashi et al., 2011). Thus, the ability to generate pluripotent PGC-like cells from EpiLCs corroborates with the ability of immature EpiSCs to revert back into naïve ESCs in contrast to more advanced EpiSCs, as identified by high and low *Oct4*, respectively (Han et al., 2010).

ESCs and EpiSCs have different signalling requirements and these cell populations cannot spontaneously interconvert in culture. However, manipulation of signalling pathways can induce ESCs to differentiate into EpiSC cultures. ESCs cultures under the effect of Activin/Nodal and Fgf signalling, readily acquire an EpiSC phenotype and transcriptome (Guo et al., 2009; Jouneau et al., 2012) , and can be stably propagated. However, culturing EpiSCs in ground-state ESCs conditions does not revert them back to an ICM-like state, unless by forced expression of the pluripotency marker *Klf4* (Guo et al., 2009). Notably, the difficulty in reverting EpiSCs into ICM-like state is a feature attributed to the epigenetic barrier that separates these two pluripotent states. Nevertheless, EpiSCs were found able to revert into an ESC-like state without genetic

manipulation though at low frequency, by derivation and/or propagation on feeders and exposure to LIF-Stat signalling and serum (Fig.1.5) (Bao et al., 2009; Veillard et al., 2014), or by treatment with a cocktail of Lsd1, Alk5, Mek, Fgf and Gsk3 inhibitors (Zhou et al., 2010).

1.4.3 Generation of extra-embryonic stem cell populations

Mouse ESCs can also be directed to generate extra-embryonic lineages *in vitro* by genetic manipulation. Notably, over-expression of key TE or PrE-associated transcription factors, or down-regulation of transcription factors that safeguard the pluripotent state, can induce lineage switching between ESCs and TS or XEN cells lines. Notably, the pluripotency triad Oct4, Nanog and Sox2 was proposed to be at the heart of a multi-factor gene network that cooperates to sustain the pluripotent state by stimulating expression of downstream self-renewal genes while repressing differentiation-associated genes in ESCs (Yeo and Ng, 2013). Depletion of *Oct4* in the embryo causes malformation of the pluripotent ICM, with concomitant differentiation into TE tissues (Nichols et al., 1998). Similarly, depletion of *Oct4* in ESCs impairs the undifferentiated state and favors the formation of TS-like cells under appropriate culture conditions, due to de-repression of the TE marker *Cdx2* (Niwa et al., 2000). Concordantly, over-expression of *Cdx2* in ESCs concomitantly inhibits *Oct4* expression allowing the production of self-renewing TS-like cells, concordant with their mutual exclusive expression in TE versus ICM in the blastocyst (Niwa et al., 2005). Additionally, Sox2 cooperates with Oct4 to inhibit TE differentiation, by maintaining appropriate levels of *Oct4*. Concordantly, *Sox2*-depleted blastocysts cannot generate a pluripotent ICM, and depleted ESCs acquire a TS-like phenotype (Avilion et al., 2003; Masui et al., 2007). While Oct4/Sox2 and Cdx2 associate with and specify the ICM and TE, Nanog and Gata6 specifically mark Epi and PrE lineages, respectively (Chazaud et al., 2006). Differentiation into PrE is prevented by Nanog via transcriptional repression of *Gata6* and other PrE-associated markers (Lavial et al., 2012; Singh et al., 2007). Conversely, PrE differentiation in ESCs can be accomplished by over-expression of *Gata6* or deletion of *Nanog*, both *in vitro* and *in vivo* (Fujikura et al., 2002; Shimosato et al., 2007). However, although *Nanog* depleted pre-implantation

embryos do not form a pluripotent ICM, depleted ESCs can be established and maintain their pluripotency, highlighting a role in acquisition rather than maintenance of pluripotency (Chambers et al., 2003; Mitsui et al., 2003).

Without gene manipulation, conversion of ESCs into a TS and XEN-like state can also be accomplished and/or inhibited by exposure to extrinsic signalling factors. In particular, LIF and BMP4 signalling molecules cooperate to support ESC self-renewal state by inhibiting trophoblast and neural differentiation (Qi et al., 2004; Ying et al., 2003a). Thus, in defined conditions, removal of LIF and treatment with BMP4 promotes direct activation of *Cdx2*, with concomitant activation of other TS cell markers (Fig.1.5) (Hayashi et al., 2010). Moreover, retinoic acid treatment of ESCs in monolayer (Artus et al., 2011; Capo-Chichi et al., 2005; Soprano et al., 2007) or cell aggregation into embryoid bodies in the absence of LIF has been shown to promote the emergence of PrE-like derivatives (Fig.1.5) (Coucouvanis et al., 1995). However, generation of a self-renewing XEN-like population from ESCs can only be achieved upon combined retinoic acid and Activin treatment (Fig.1.5) (Niakan et al., 2010; Niakan et al., 2013).

1.4.4 Multi-lineage differentiation of pluripotent stem cells

The remarkable ability of ESCs to derive a broad spectrum of cell types from the three germ layers has been extensively studied and continuously optimized (Keller, 2005). In particular, ESCs culture in suspension in the absence of LIF has been exploited as an *in vitro* system that mimics early segregation events in the mouse blastocyst and promotes the emergence of multi-lineage progenitors. These progenitors can then be further induced into mature cell types and/or propagated in monolayer systems. Moreover, direct differentiation of ESC in monolayer and on specific extra-cellular matrixes offer an alternative and efficient platform to differentiate these cells into specific and relevant lineages with key identified signalling pathways, as summarised in Fig.1.5. Overall, generation of *in vitro* systems that model transitions between pluripotent and somatic states provide important tools to investigate the molecular events underlying not only

the shut-down of the pluripotency network, but also the signalling requirements and specific gene function upon lineage commitment.

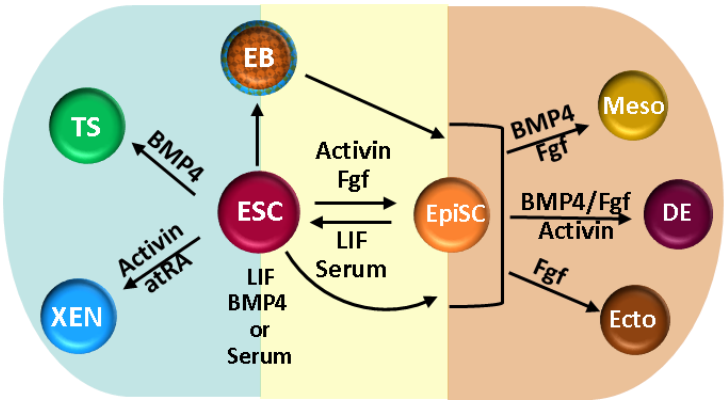


Figure 1.5 *In vitro* differentiation of pluripotent stem cells

Schematic representation of signalling pathways that can be manipulated in order to direct the fate of ESCs into extra-embryonic (blue shade), epiblast (yellow shade) and epiblast-derived lineages (orange shade) as depicted. ESCs rely on LIF and BMP4 (or serum) for the maintenance of their identity. Removal of LIF and BMP4 treatment directs ESCs towards a TS-like state, whereas treatment with all-trans retinoic acid (atRA) and Activin induces formation of self-renewing XEN cells. Embryo-derived EpiSCs sustain their pluripotent state through Nodal/Activin and Fgf signalling, hence culturing ESCs under these conditions promotes their conversion into an EpiSC-like state. Moreover, EpiSCs can sporadically revert back to an ESC state by exposure to LIF and serum on feeders. Culture of ESCs in suspension in the absence of LIF promotes emergence of progenitors of both embryonic lineages and extra-embryonic PrE, mimicking the early segregation events in the blastocyst. ESCs, EpiSCs and embryoid bodies (EB) can be further induced to differentiate into cell types representative of the three germ layers that arise upon gastrulation in the post-implantation epiblast. Mesoderm (Meso) and definitive endoderm (DE) can be induced by BMP4 and Fgf signalling, but DE requires additional Activin. Induction of ectoderm (Ect), in particular neuroectoderm commitment, is associated with active Fgf signalling. Adapted from (Keller, 2005).

1.4.5 Somatic reprogramming to pluripotency

For decades, cell differentiation was assumed to proceed in a one-way route (Waddington, 1957). However, this view has been challenged by the discovery that somatic cells are more

versatile than previously thought, and can be manipulated to re-acquire pluripotency through a process now commonly known as reprogramming.

Pioneer experiments in the 1950s showed that using somatic cellular nuclear transfer (SCNT) (Gurdon, 1962), where the genetic information from a somatic cell is transferred into an unfertilized and enucleated oocyte, a full organism could be *de novo* generated from an adult, differentiated cell (Fig.1.6A). These cloning experiments initially performed in frogs thus uncovered that adult cells indeed retained the genetic information necessary to develop an entire organism, and that unlike previously hypothesised, differentiation did not occur with progressive loss of genetic information during development, which we now know is attributed to epigenetic regulation. It was only in the late 1990s that SCNT could then be replicated in mammals with the cloning of "Dolly" the sheep (Wilmut et al., 1997). It these pioneer experiments, it became apparent that the factors required for reprogramming were present in the host cell. Supporting evidence came from cell fusion experiments, where two cells are fused together in order to generate a multinucleated heterokarion (or hybrid if cell division occurs) (Fig.1.6B). Here, the fast dividing cell (dominant) induces changes in the gene expression profile of the second cell (recessive). This technique has since been applied to a number of somatic cell types reprogrammed by fusion with ESCs (Pereira et al., 2008; Tada et al., 2001; Ying et al., 2002).

Altering the fate of a cell by forced expression of tissue-specific transcription factors has long been explored, but it was not until 2006 that the same procedure was found to be applicable to the reversion of somatic cells to a pluripotent state. In a ground-breaking study, mouse embryonic and adult fibroblasts were transduced with a combination of retroviral vectors encoding four key ESC transcription factors: *Oct4*, *Sox2*, *Klf4* and *c-Myc*, commonly referred as to OSKM or Yamanaka factors (Takahashi and Yamanaka, 2006). This successfully induced a stable self-renewal state with acquisition of a multi-lineage developmental potency similar to ESCs (Fig.1.6C). This new type of induced pluripotent stem cells (iPSC) has since been exploited for its potential in autologous transplantation in regenerative medicine, disease modelling and drug screening studies. Moreover, the iPSC technology has since been continuously optimized for identification

of faster, safer and more efficient approaches, which include combination of signalling and epigenetic modifiers (Ohnuki and Takahashi, 2015).

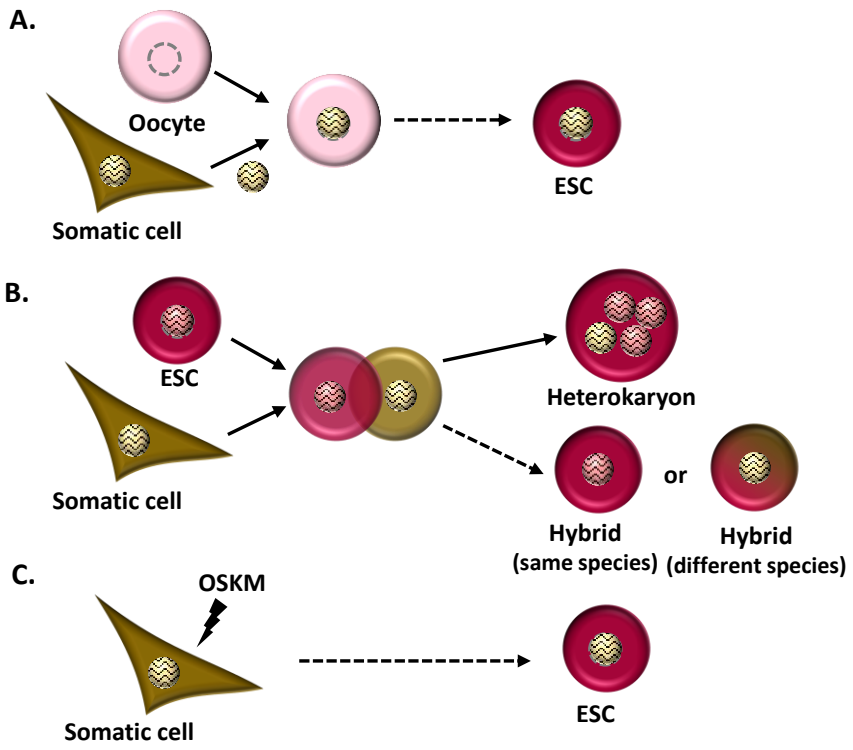


Figure 1.6 Multiple routes to reprogramming

A. Nuclear transfer: the nucleus of a somatic cell is transplanted into an enucleated oocyte where it is reprogrammed. Upon transfer back to the uterus of the host species a full organism can be generated, or ESCs can be derived from the blastocyst. **B.** Cell fusion: two distinct cell types (eg. an ESC and a somatic cell) are fused together. Without proliferation, a multinucleated heterokaryon is formed, and the nuclear ratio in the fusion reprograms the heterokarion to the desired cell type. With cell proliferation (dashed line), the cells become hybrids and, on division, nuclei fusion occurs. If both cells are from the same species, their karyotype will remain euploid, but if they belong to different species the hybrid cells become aneuploid. **C.** Transcription factors: by introduction of the four genes *Oct4* (O), *Sox2* (S), *Klf4* (K) and *c-Myc* (M) the pluripotent state can be induced in somatic cells, generating induced pluripotent stem cells (iPSC). Adapted from (Yamanaka and Blau, 2010).

I.5 Regulation of the pluripotent state

The ease of manipulating embryo-derived pluripotent stem cell lines in culture has provided an extraordinary tool to decipher the molecular mechanisms that govern self-renewal and the remarkable long-term maintenance of multi-lineage developmental potency. Over the past decade, accumulating evidence highlighted that not one, but multiple factors, cooperate to sustain the unique ESC identity. As will be described in this section, the integrated action of multiple signalling pathways with complex gene regulatory networks, elegantly reinforced by chromatin and DNA modifiers, safeguard the self-renewing state while priming gene expression for efficient response to differentiation stimuli.

1.5.1 Extrinsic signalling pathways

Embryo development, as well as *in vitro* culture of ESCs, requires the action of extrinsic growth factors to preserve the pluripotent state. In turn, growth factors influence different intrinsic signalling pathways that will ultimately regulate the transcription factor network that sustains ESC identity.

1.5.1.1 LIF-mediated signalling pathways

Originally, ESC culture was found to rely on serum and a layer of feeder cells that consisted of mitotically inactivated mouse embryonic fibroblasts (MEF). It was later uncovered that the key component produced by MEFs was LIF, a member of the Interleukin-6 family cytokines, which sustained pluripotency by inhibiting differentiation through the JAK/STAT signalling pathway (Niwa et al., 1998). The LIF-mediated activation of this pathway occurs through the heterodimerization of the LIF receptor with the membrane protein gp130. This process leads to trans-phosphorylation and activation of JAK1 tyrosine kinase (Janus kinase), which phosphorylates gp130 and subsequently promotes STAT3 (Signal Transducers and Activators of Transcription 3) recruitment (Yoshida et al., 1994). Thus, STAT3 is in turn phosphorylated by JAK1 and forms a homodimer, which translocates into the nucleus where it potentiates ESC-

specific gene expression (Niwa et al., 1998; Raz et al., 1999). Notably, STAT3 binds genomic regions co-occupied by the key transcription factors Nanog, Oct4 and Sox2 in the vicinity of genes specifically expressed in ESCs (Chen et al., 2008; Kidder et al., 2008). Removal of LIF, or knockdown of STAT3 target genes, results into the activation of a variety of endodermal and mesodermal-affiliated genes, proposing that the LIF/STAT3 action is essential for the silencing of differentiation markers via activation of key ESC-specific genes (Bourillot et al., 2009). Further evidence that STAT3 is a key player in this pathway comes from the observation that constitutive STAT3 activation confers LIF independency in ESCs (Matsuda et al., 1999).

Moreover, two other pathways are induced upon LIF stimuli: PI3K and MAPK pathways. Activation of PI3K (PI3-kinase) is achieved through association between JAK1 and the subunit p54 of Pi3K, leading to activation of the serine/threonine kinase AKT (protein kinase B). In turn, AKT stimulates a number of other pathways, involved in cell cycle regulation (e.g. mTOR) and metabolism, as well inhibiting the WNT signalling via GSK3 (glycogen synthase kinase 3) (Vanhaesebroeck et al., 2001). Similarly to STAT3, inhibition of Pi3K leads to ESC differentiation, and LIF unresponsiveness, whereas constitutive activation of AKT sustains self-renewal without requirement of LIF (Paling et al., 2004). Activation of the MAPK/ERK pathway occurs through recruitment of the phosphatase SHP2 to the gp130, which in turn activates RAS/RAF components of this pathway (Schiemann et al., 1997). Activation of this pathway is linked to induction of differentiation, however, this is counterbalanced by activation of SOCS3 by STAT3, which down-regulates ERK signalling by competition for the anchoring site of gp130 (Lehmann et al., 2003). Ultimately, the balance of the competing pathways activated by LIF sustains the pluripotent state in mouse ESCs.

1.5.1.2 TGF-β-mediated signalling pathways

The TGF-β superfamily is composed by a wide range of molecules, including TGF-β, Activin, Nodal, GDFs (growth and differentiation factors) and BMPs (bone morphogenetic proteins),

which regulate multiple cellular processes. The canonical pathway occurs via activation of SMAD proteins, leading to their nuclear translocation and target gene activation (Wrana et al., 1992).

Binding of BMP ligands to type II receptors leads to activation of type I receptors by phosphorylation. In turn, type I receptors phosphorylate SMAD1, 5 and 8, which stimulate gene expression. In particular, BMP4, has been found to substitute serum in the maintenance of pluripotency by inhibiting neural differentiation via activation of the Id genes, and inhibition of the MAPK/ERK pathway (Qi et al., 2004; Ying et al., 2003a). Additionally, it promotes proliferation by stimulation of WNT expression (Lee et al., 2009). Moreover, BMP4 also plays a role during differentiation into the mesoderm and hematopoietic lineages, and, together with Activin, induces cardiac development (Laflamme et al., 2007; Sadlon et al., 2004). This cooperative action is linked with the activity of SMADs in regulating *Oct4* expression, which is involved in cardiac lineage specification (Zeineddine et al., 2006). However, BMP4-alone induces extra-embryonic differentiation in human ESCs and mouse EpiSC (Brons et al., 2007; Vallier et al., 2009b; Vallier et al., 2009c), while high BPM4 concentrations in mouse EpiSCs self-renewing conditions can induce differentiation into germ cells (Hayashi and Surani, 2009; Tesar et al., 2007).

Activin and Nodal signalling comprise another branch of the canonical TGF-β pathway, which has been widely associated with the self-renewal maintenance of primed, LIF-irresponsive, mouse EpiSCs and human ESCs. This occurs via SMAD2/3-mediated activation of *Nanog* and, consequently, inhibition of neuronal differentiation (Vallier et al., 2009a). In mouse ESCs, however, it appears to have no impact on the undifferentiated identity but instead was found to be associated with cell proliferation (Ogawa et al., 2007).

1.5.1.3 FGF-mediated signalling pathways

FGF (Fibroblast growth factor) signalling has a key role in lineage segregation in the early mouse blastocyst. In particular, Fgf4 is produced by ICM cells and ICM-derived ESCs, and its expression is enhanced by the pluripotency-associated factors Oct4 and Sox2 (Yuan et al., 1995).

In the blastocyst, Fgf4 action is required for proper trophectoderm and primitive endoderm formation, where its receptor Fgfr2 is expressed (Feldman et al., 1995; Haffner-Krausz et al., 1999). Concordantly, addition of exogenous Fgf4 is required for derivation and/or maintenance of extra-embryonic derived stem cells *in vitro* (Kunath et al., 2005; Tanaka et al., 1998). Although ESCs produce Fgf4 in an autocrine fashion, this is dispensable for sustaining their self-renewal state, but instead promotes lineage commitment (Kunath et al., 2007; Stavridis et al., 2007)

FGF signalling occurs upon ligand-receptor interactions, leading to auto-phosphorylation of tyrosine residues of Fgfr2. In turn, this activates the mediators Grb2 and, consequently, the Pi3K/AKT and MEK/ERK pathways, which will contribute to induction of self-renewal and differentiation of mouse ESCs, respectively. Hence, chemical inhibition of MEK with PD0325901 supports the self-renewal state, but is insufficient to sustain it unless combined with a GSK3 inhibitor (Ying et al., 2008).

In mouse EpiSCs and human ESCs, maintenance of the self-renewal state requires active FGF signalling by addition of exogenous Fgf2 (Brons et al., 2007). However, and unlike human ESCs, FGF signalling in mouse EpiSCs is not essential for SMAD2/3-mediated activation of *Nanog*, but is rather required to prevent spontaneous neuronal differentiation and reversion to an ESC-like state (Greber et al., 2010).

1.5.1.4 WNT-mediated signalling pathways

The WNT signalling pathway has key roles in a wide range of cellular processes like proliferation, adhesion, morphology and migration (Bejsovec, 2005). It encompasses more than 30 extracellular ligands that can either operate via a canonical pathway, which involves β-catenin-mediated gene expression, or via a β-catenin-independent non-canonical pathway, which involves activation of many other signalling cascades (Widelitz, 2005). Upon ligand binding, activation of the two receptors Frizzled and LRP (low-density lipoprotein receptor protein) leads to inhibition of the GSK3-mediated destruction of β-catenin. Accumulated β-catenin translocates into the nucleus and promotes the expression of pluripotency factors including *Oct4*, *Nanog*,

Tfcp2l1 and *Esrrb*, through degradation of the repressor Tcf3 (Wray et al., 2011). Here, *Esrrb* is considered a key downstream target of GSK3 pathway as depletion of *Esrrb* abrogates the effect of Tcf3 inhibition and impairs ESC self-renewal (Martello et al., 2012). A parallel effect of Tcf3 degradation is the activation of *Nr5a2*, which stabilizes the expression of *Oct4*, *Nanog* and *Tbx3* (Fuhrmann et al., 2001).

In ESC cultures, the accumulation of β-catenin can be promoted by chemical inhibition of the GSK3 pathway with CHIR99021, but alone is not sufficient to promote self-renewal. Thus, activation of the β-catenin signalling by GSK3 inhibition, in combination with inhibition of MEK-induced differentiation (2 inhibitors, 2i) can together bypass the requirement for the LIF/STAT3 signalling pathway and moreover, has been proposed to promote a robust ESC self-renewal "ground" state (Ying et al., 2008). In mouse EpiSCs and human ESCs, however, activation of the WNT signalling is associated with differentiation into mesoderm, when coupled with BMP4 (Kurek et al., 2015).

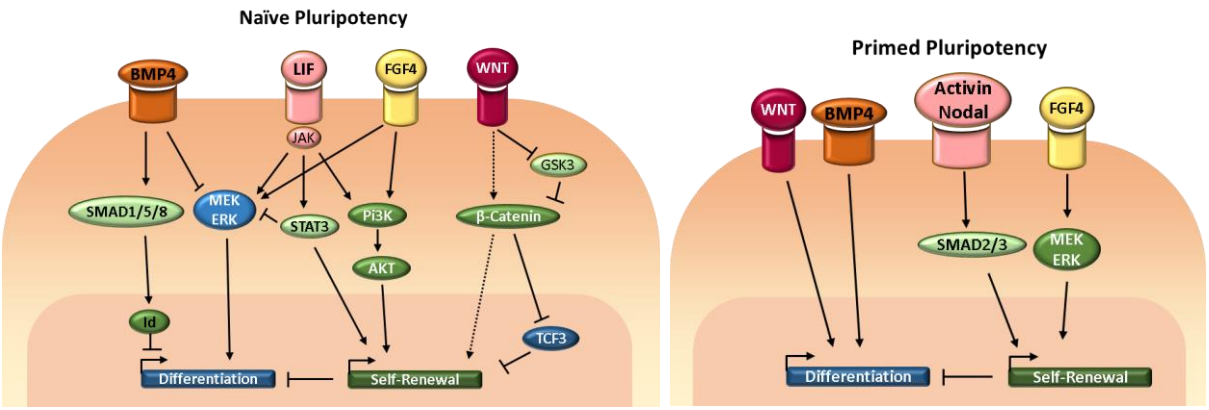


Figure 1.7 Self-renewal signalling pathways in naïve and primed pluripotent states

Summary of signalling events that directly (full line) or indirectly (dotted line) are involved in activating the expression of self-renewal (green) and blocking expression of differentiation-associated (blue) genes in naïve (mouse ESC) and primed (mouse EpiSC) pluripotent states. In ESCs BMP4 (TGF-β), LIF and WNT signalling cascades together sustain self-renewal, while MEK/ERK promotes differentiation. In EpiSCs, Activin/Nodal and FGF signalling are essential for self-renewal, while induction of BMP4 or WNT pathways promotes differentiation.

In summary, mouse ESCs and EpiSCs (and human ESCs) require different signals for the maintenance of their pluripotency, with the same signalling pathways, remarkably, having opposite roles in ESCs versus EpiSCs. However, and despite intrinsic differences, the combinatory action of signalling pathways in both naïve and primed pluripotent cell populations is required to support the expression of pluripotency genes and prevent that of differentiation genes, hence ultimately sustaining the undifferentiated state.

1.5.2 Transcription factors network

Ultimately, the cell's pluripotent state reflects the action of multiple signalling cues on the downstream gene expression programs. Transcription factors (TF) are the key mediators in this process, which will coordinate the activation, repression or even priming of transcription, trough recruitment of multi-protein complexes and alteration of the chromatin environments. Gene function studies in ESCs have allowed thus far the identification of a critical network of TFs that robustly confer their unique pluripotent identity.

1.5.2.1 Core transcription factors

The triad Oct4-Sox2-Nanog is at the core of the transcriptional regulation in ESCs as revealed by their co-binding to the same genome-wide DNA target sites (Kim et al., 2008). Oct4 is a member of the Pit-Oct-Unc (POU) family, and is encoded by the *Pou5f1* gene. It was one of the first transcription factors identified as crucial in early embryo development. It is expressed from the 8-cell stage and its expression becomes restricted to the pluripotent ICM prior to implantation, and remains on in the epiblast after implantation before being confined to the posterior epiblast and PGCs (Osorno et al., 2012; Radzisheuskaya et al., 2013; Scholer, 1991). Concordantly, *Oct4* is highly expressed in embryo-derived ESCs, EpiSCs and EGs, and its expression is reduced when cells are induced to lose their pluripotent identity (Nichols et al., 1998). Fine-tuning of *Oct4* levels is an essential requirement for the pluripotent state in ESCs, as reduction of *Oct4* expression is sufficient to trigger differentiation into TS-like cells, by lack of Oct4-dependend inhibition of the

TE markers *Cdx2* (Niwa et al., 2000; Niwa et al., 2005). Conversely, its over-expression induces differentiation of ESCs into primitive endoderm and mesoderm fates (Niwa et al., 2000), proposing that *Oct4* may function in a dynamic context-dependent manner.

Oct4 regulates the expression of a broad range of genes through direct DNA binding. In particular, it has been found to recognise an octamer-sox motif (Oct-Sox) in the DNA, where it cooperates with Sox2 in regulating gene activation (Tomioka et al., 2002). Evidence for Oct4 versatility comes from the identification of novel cell state-dependent protein partners. Upon induction of a PrE-like state in ESCs by enforced *Sox17* expression, Oct4 is re-distributed into regulatory sites of endodermal genes, switching interaction between Sox2 to Sox17 (Aksoy et al., 2013). Similarly, upon exit of the naïve ESC state into a primed EpiSC-like state, Oct4 binding is also re-distributed into Otx2 or Zic motifs, at regulatory regions of genes associated with the post-implantation epiblast (Buecker et al., 2014; Yang et al., 2014). Notably, some level of Oct4 binding is already present at a selection of post-implantation genes in the naïve state such as *Fgf5* or *Brachyury*, possibly associated with lineage priming functions.

In addition to Sox2, genome-wide Oct4 binding sites in ESCs encompass motif sequences for other pluripotency factors (Buecker et al., 2014), which integrate the so-called "Oct4-centric module". The definition of such module was based on the chromatin co-occupancy of proteins at target chromatin sites, encompassing factors such as Nanog, Sox2, Esrrb, Klf4, Smad1, Stat3, Dax1 or Tcfp2l1 (Chen et al., 2008; Kim et al., 2008). This analysis further uncovered a second module centred on c-Myc ("Myc-centric module"), which involves factors such as n-Myc, E2f1, Zfx and Rex1, and appears to control different sets of genes that are associated with protein metabolism and cell cycle (Chen et al., 2008; Hu et al., 2009). Furthermore, Oct4 physically interacts with a set of pluripotency factors, consistent with their identified co-binding in the DNA. Since abolishing Oct4 strikingly impairs the recruitment of these factors, it has been proposed to play a pioneer role in the assembly of multi-protein complexes at the chromatin (van den Berg et al., 2010).

Sox2 [SRY (sex determining region Y)-box 2] is a member of the Sox family of transcription factors that, like *Oct4*, is detected in the pluripotent ICM, where it cooperates with *Oct4* in TE/ICM segregation (Avilion et al., 2003; Masui et al., 2007). However, after implantation the expression of *Oct4* and *Sox2* become oppositely restricted to the posterior (mesoderm, endoderm) and anterior (ectoderm) epiblast, respectively. Concordantly, *Sox2* is expressed in ESCs, EpiSCs, PGCs/EGs, and embryonic and adult neural stem cells (D'Amour and Gage, 2003; Ellis et al., 2004).

As mentioned above, Oct4 and Sox2 co-regulate pluripotency genes through binding to a common Oct–Sox motif at regulatory regions. Similarly to Oct4, upon induction of neuronal fates, Sox2 was found to switch partners to activate neuronal genes. Specifically, it has been found to co-occupy neuronal-specific regulatory sites with another member of the POU family – Brn2 (*Pou3f2*). However, unlike Oct4, Sox2 recognises the same DNA sequence in different cell states, suggesting that the re-distribution of binding at target sites is driven by its specific co-partners in a context-dependent manner (Lodato et al., 2013) .

Nanog is a homeobox transcription factor that, alongside *Oct4* and *Sox2*, is expressed in the pluripotent ICM, but becomes restricted to the epiblast progenitor blastomeres in late blastocysts by counteracting *Gata6*-driven PrE specification (Chazaud et al., 2006). *Nanog* expression is down-regulated before implantation and is re-activated in the posterior primitive-streak until early somitogenesis (Osorno et al., 2012). Similarly to *Oct4* and *Sox2*, it is expressed in ESCs, EpiSCs and PGCs/EGs, having an essential role in the maintenance of pluripotency as highlighted by the loss of ESC self-renewal upon conditional depletion, and the acquisition of LIF-independence upon over-expression (Chambers et al., 2003; Mitsui et al., 2003). Based on this property the gene was named after the legendary Tir na n'Og - the Celtic land of the ever young. Nanog inhibits the expression of the PrE markers *Gata4* and *Gata6*, while promoting the expression of many pluripotency genes in the Oct4-centric module (Chen et al., 2008; Kim et al., 2008). Unlike *Oct4*, the expression levels of *Nanog* (and *Sox2*) diminish in the post-implantation epiblast and in EpiSCs. Concordantly, forced expression of *Nanog* in EpiSCs can revert them back

to an ESC state, whereas over-expressing in ESCs prevents differentiation (Osorno and Chambers, 2011).

Similarly, the expression of a set of naïve pluripotency markers such as *Esrrb*, *Klf4*, *Klf5*, *Tbx3* and *Rex1* also become down-regulated in EpiSCs (Festuccia et al., 2012; Osorno and Chambers, 2011). Given the equal dynamics of *Nanog*, reduction of this set of pluripotency markers may be a consequence of Nanog decrease. In agreement, metastable culture conditions (Serum/LIF) stimulate high heterogeneity of *Nanog* expression, promoted by auto-regulatory loop of Nanog (Navarro et al., 2012). Heterogeneous expression of the downstream targets *Esrrb*, *Klf4* and *Tbx3* is also detected in such conditions, suggesting that these fluctuations are directly influenced by *Nanog* expression itself (Osorno and Chambers, 2011). Notably, over-expression of these Nanog's gene targets (*Klf4*, *Tbx3* and *Esrrb*) is sufficient to confer LIF-independence bypassing the requirement for Nanog (Festuccia et al., 2012; Niwa et al., 2009). In particular, LIF-independency mediated by enforced *Esrrb* expression was found to strictly require the recruitment of its essential co-activator *Ncoa3* in ESCs (Percharde et al., 2012). Additionally, forced expression of *Esrrb* or *Klf4* also enables reversion of EpiSCs into an ESC-like state (Festuccia et al., 2012; Guo et al., 2009).

LIF-independency studies furthermore highlighted key connections between LIF/STAT3 signalling and the core transcriptional machinery. Notably, LIF was found to activate a parallel circuitry where the JAK-STAT3 branch activates the expression of *Klf4*, and the Pi3K–AKT pathway preferentially induces the expression of *Tbx3*. *Klf4* and *Tbx3* stimulate *Sox2* and *Nanog*, and contribute to the maintenance of *Oct4* (Niwa et al., 2009). Moreover, *Esrrb* is a key downstream target of the canonical WNT signalling pathway and stimulates the expression of *Oct4*, *Nanog*, *Klf4* and *Esrrb* itself, though recognition of ERRE motifs at these loci (Martello et al., 2012; Percharde et al., 2012). In turn, the transcription of *Esrrb*, *Tbx3* and *Klf4* is positively regulated by the core Oct4-Sox2-Nanog transcription factors, which may altogether reinforce the stability of the pluripotent state. In post-implantation, the lower expression of *Nanog* is instead

regulated by the Nodal signalling pathway and directly sustained by Oct4 and Smad2 (Sun et al., 2014). Altogether, the core network of transcription factors functions in the auto-regulation of their expression via an interconnected regulatory loop and by activating expression of genes required for the ESC self-renewal, while expressing differentiation-associated genes.

1.5.2.2 Transcription factor clusters at "enhanceosomes"

Transcription factors (TF) bind proximal DNA elements, usually associated with promoter regions in the vicinity of a gene, but also bind to distal regions that can be located up to several kb away from the gene itself. These distal elements are typically involved in potentiating gene expression and are termed enhancers. Pioneer mapping of TF binding sites through chromatin-immunoprecipitation techniques, coupled with high-throughput sequencing (ChIP-seq) uncovered that multiple TFs tend to bind to the same DNA regions, termed multiple transcription factor-binding loci (MTL) (Kim et al., 2008; Chen et al., 2008). Moreover, most of these MTLs comprise enhancer regions that are co-occupied by transcriptional co-factors such as p300 and Mediator complexes, and furthermore display enhancer activity when cloned into luciferase reporters (Chen et al., 2008; Kagey et al., 2010). Thus, MTL regions are often referred as to "enhanceosomes" where higher order multi-protein complexes are assembled, being able to modify the chromatin environment and hence allowing the recruitment of the basal transcriptional machinery.

Enhancer activity implies the coordinated recruitment of co-factors by TFs. These molecules do not necessarily have a DNA-binding ability, but are able to further recruit chromatin modifiers. In particular, the p300 and Mediator co-activator complexes are large multi-subunit complexes that are able to simultaneously interact with several TFs (Malik and Roeder, 2005). Of relevance, p300 possesses chromatin modifying activity, being able to acetylate the lysine 27 of histone 3 - a modification associated with permissive chromatin environment for TF binding (Tie et al., 2009). Moreover, the Mediator complex has a key role in mammalian transcription and associates

with Cohesin-loading factor Nipbl at Oct4-Sox2-Nanog bound sites. In turn, Nipbl loads cohesin ring proteins at enhancer sites contributing to DNA-DNA interactions between enhancer and promoters (Kagey et al., 2010). Cohesin is a multi-meric complex that has been firstly identified for its role in sister chromatid cohesion, but has later been found to promote DNA double-strand break repair and to regulate gene expression in a manner that involves enclosing the DNA within its ring-shaped structure (Nasmyth and Haering, 2009).

The recruitment of given factors at enhancers regions occurs in a cell type specific manner. In ESCs, a subset of large enhancers regions are densely occupied by Oct4, Sox2, Nanog, Klf4, Esrrb and by Mediator, and potentiate the strong expression of pluripotency and self-renewal-associated genes. These regions were recently termed "super" enhancers, and differ from "typical" enhancers that are typically smaller and less densely occupied by TFs. Super enhancers are sensitive to the reduction levels of subunits of the Mediator complex, similarly to a reduction of Oct4, highlighting this TF as a key component of the enhanceosome and cell identity (Kagey et al., 2010). Thus, robust assembly of enhancer complexes is a crucial factor in the maintenance of a cell's identity due to their roles in bridging enhancers and promoters, shaping the chromatin landscape and allowing the recruitment of the transcriptional machinery for proper expression.

1.5.3 Epigenetic gene regulation

The genetic information identically contained within each cell is subjected to multiple layers of regulation during development to ensure that the appropriate gene expression programs are in place at the right time and in a cell type specific manner. Thus, the ability to establish and inherit suitable transcriptional states over cell divisions is vital for a cell's identity. In the nucleus of eukaryotic cells, DNA is wrapped around four core histone proteins (H2A, H2B, H3, H4) shaping a nucleosome unit. The resulting DNA-protein polymers form the chromatin template. Additional histone proteins – the histone protein H1 and its variants, play key roles in the assembly of higher order chromatin structures. Chromatin compaction is necessary to accommodate the extended DNA molecule into a cell's nucleus, and moreover to regulate multiple processes such as

chromosome segregation, safeguarding the genome integrity, and regulating gene expression and DNA replication (Chen and Dent, 2014; Lukas et al., 2011; Rosa and Shaw, 2013). Thus, chromatin compaction is dynamically regulated in response to environmental cues and occurs at multiple levels including post-translational modifications (PMT) of histone proteins, ATP-dependent chromatin remodelling, additional histone variants, DNA modifications and non-coding RNAs (Fig.1.8). Modifications in histones and in the DNA, without affecting its sequence, are mostly reversible and termed epigenetic modifications (Chen and Dent, 2014).

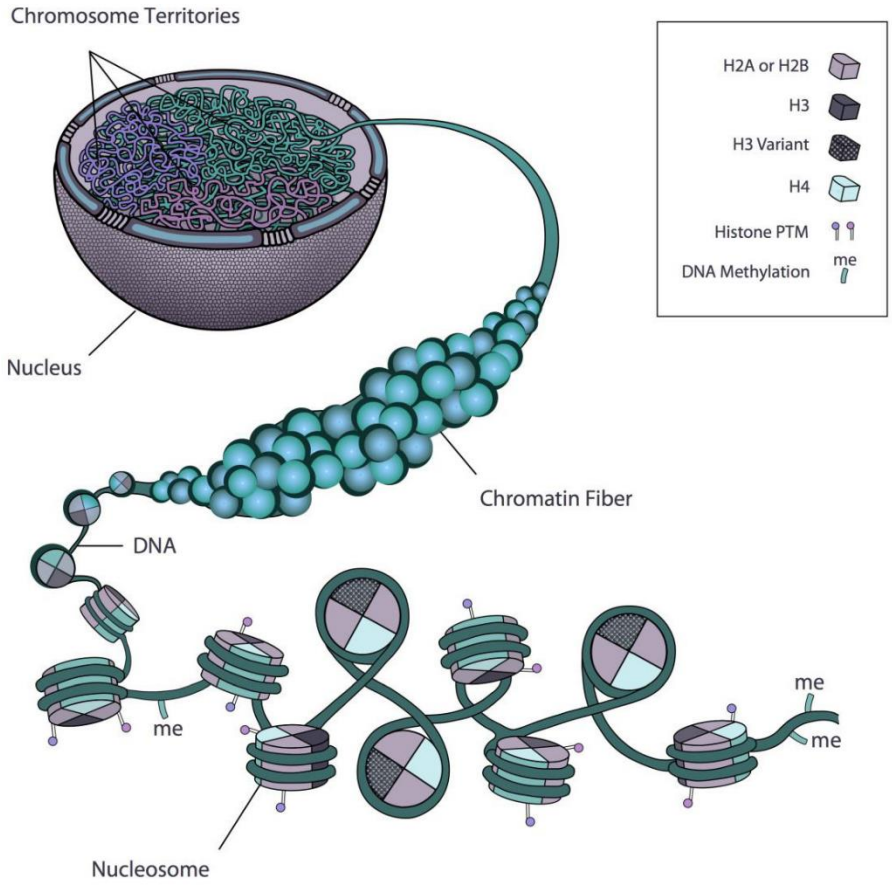


Figure 1.8 Chromatin organization in eukaryotes

Schematic representation of the levels of genome organisation in eukaryotic cells. DNA is wrapped around compact histone complexes containing a pair of each of the four core histones H2A, H2B, H3 and H4. Condensation of nucleosomes forms a compact chromatin fibre, which can be dynamically organised into functional nuclear territories. Histones can be post-translationally modified (PTM) and DNA can be methylated (me). Figure from (Rosa and Shaw, 2013).

Histone PMTs include acetylation, methylation and ubiquitination, and are regulated through the enzymatic action of histone acetyltransferases (HATs) and deacetylases (HDACs), lysine methyltransferases (KMTs) and demethylases (KDMs), and ubiquitylases (E1, E2 and E3 enzymes) and deubiquitylases (DUBs), respectively. In the other hand, DNA modification encompass different forms of cytosine methylation. DNA can be methylated by DNA methyltransferases (DNMTs), or demethylated by a passive mechanism, which involves inhibition of the "maintenance" DNMTs upon replication, or through an active multi-step mechanism that involves oxidation into intermediate states by the TET enzymes (Branco et al., 2012; Carey et al., 2011; Chen and Dent, 2014). Ultimately, the coordinated action of histone and DNA modifiers influences the degree of chromatin compaction, which can be either less compact, and thus more accessible – euchromatin, or more compact and inaccessible – heterochromatin. In turn, chromatin conformation will further allow recruitment of regulatory factors and enable transcription at euchromatin regions, or impede the access of binding factors and disable transcription at heterochromatin regions.

1.5.3.1 Chromatin dynamics during early development

The maternal and paternal genomes exhibit striking asymmetries in their chromatin organization in the zygote and early cleavage stage embryos (Burton and Torres-Padilla, 2014). Following fertilization, the two parental epigenomes must go through global reprogramming events in order to accommodate the gene expression programs necessary for the development of the embryo. The paternal chromatin is tightly packed, but upon fertilization is de-condensed while undergoing active genome-wide DNA demethylation. In contrast, the maternal genome undergoes a passive DNA demethylation process. The only genomic regions that escape demethylation are differentially methylated regions encompassing imprinted genes in either the paternal or maternal DNA (Burton and Torres-Padilla, 2014; Hemberger et al., 2009). *De novo* DNA methylation starts at the blastocyst stage, coinciding with the first lineage segregation event between ICM and TE. Notably, global DNA methylation levels are asymmetric between the two

lineages, being higher in the ICM, and this asymmetry persists throughout development (Rossant et al., 1986; Santos et al., 2002). Subsequently, a new wave of *de novo* DNA methylation in the epiblast marks the post-implantation stage concomitant with an increase in Dnmt3b expression levels, and the occurrence of random X chromosome inactivation (Hirasawa and Sasaki, 2009; Reik, 2007).

Moreover, the prevalence of histone modifications and/or associated histone modifiers has been implicated in lineage specification. One of these is the methyltransferase Carm1, which methylates H3R2, H3R17 and H3R26 residues; a set of modifications commonly associated with gene activation. High levels of Carm1 protein in individual blastomeres of cleavage-stage embryos introduce an early bias that preferentially leads to cell allocation within the ICM of newly formed blastocysts (Torres-Padilla et al., 2007). Another example is the histone methyltransferase Setdb1, which deposits the repressive marks H3K9me2 and H3K9me3 and acts as a co-repressor with Oct4 to *in vivo* silence the expression of TE-associated genes such as *Cdx2* in ICM cells. Concordantly, depletion of *Setb1* in ESCs results into the de-repression of TE-associated genes and differentiation into TS-like cell progenitors (Dodge et al., 2004; Yuan et al., 2009). Similarly, the nucleosome remodelling and deacetylation (NuRD) complex was involved in ICM/TE segregation. Depletion of the component *Mbd3* in early embryos is lethal, and in ESCs allows conversion into a TS-like state, displaying aberrant expression of the TE markers *Elf5* and *Eomes* (Kaji et al., 2006; Latos et al., 2012).

As described earlier in this Chapter, following implantation the epiblast give rises to the three germ layers. As seen in the pre-implantation embryos, several chromatin remodelers have also been associated with this process of lineage specification. Particularly, proteins of the Polycomb Repressor Complexes (PRC) are known to play an essential role in gastrulation. Depletion of members of the PRC2 complex such as Ezh2, Eed or Suz12 results into embryonic lethality shortly after implantation (Faust et al., 1995; O'Carroll et al., 2001; Pasini et al., 2004). Equally, components of the PRC1 complex such as Ring1b, Rybp and L3mbtl2 have been found to have a

key function at this developmental transition (Pirity et al., 2005; Qin et al., 2012b; Voncken et al., 2003). Notably, constitutive depletion of PRC1 and PRC2 members in ESCs results in inappropriate gene expression of differentiation-associated genes, without impact on the self-renewal ability of these cells, consistent with a key role for these complexes in the repression of lineage-affiliated genes during development (Azuara et al., 2006; Boyer et al., 2006; Pasini et al., 2004).

Critically, the expression of pluripotency genes also needs to be silenced during development. Persistent expression of *Oct4* and *Nanog* due to a lack of appropriate repression, leads to growth retardation and embryonic lethality. One key repressive factor that has been involved in the silencing of these genes is the H3K9me2-methyltransferase G9a (Ehmt2) (Tachibana et al., 2002). This protein is typically forms a heteromeric complex with the highly similar G9a-like protein (GLP; Ehmt1) and silencing each of the two proteins considerably reduces the levels of H3K9me1/2 (Collins and Cheng, 2010). ESCs lacking G9a protein expression present impaired differentiation due to the inability to silence *Oct4*. Although H3K9me2 is a histone modification typically associated with euchromatin, it is strictly necessary for the further formation of heterochromatin by H3K9me3 deposition and *de novo* Dnmt3b-mediated DNA methylation at the *Oct4* locus during differentiation (Feldman et al., 2006).

1.5.3.2 ESC-specific chromatin landscaping

ESCs harbour distinct chromatin features from adult cells. Notably, undifferentiated ESCs are characterized by an open and mostly accessible chromatin conformation with disperse chromatin fibres spread throughout most of the nuclei and no significant foci of chromatin compaction (Fig.1.9)(Ahmed et al., 2010). This is consistent with a widespread enrichment for H3K4 methylatiom (and concomitant H3K27me3 as discussed below), H3K9 acetylation and DNA hypomethylation, reflecting an active genome configuration (Azuara et al., 2006; Meissner et al., 2008). As differentiation progresses, the chromatin acquires a more compacted configuration (Fig.1.9). Staining of repressive heterochromatin marks such as H3K9me3, and the associated

protein HP1 reveals a higher density of heterochromatin foci emerging in differentiated cells as opposed to ESCs. Thus, in the undifferentiated state, HP1 and linker histones are hyperdynamic, as they loosely bind to chromatin (Meshorer et al., 2006). Moreover, the repressive mark H3K27me3 also changes from focal distribution in undifferentiated cells into expanded block that encompass intergenic regions and silent genes in differentiated cells (Pauler et al., 2009). Contrarily, euchromatin marks such as H3K9me2 display local rather than global changes, especially at genic regions (Lienert et al., 2011).

Concordantly, differences in gene accessibility between undifferentiated and differentiated cells are furthermore mirrored by different DNA replication timing at subsets of genes associated with development. In ESCs, lineage-affiliates genes display an early S-phase replication timing, which reflects an open chromatin environment despite these genes not being expressed. On the other hand, in differentiated cells, genes that are affiliated with different (or inappropriate) lineages display a late-replication timing, consistent with heterochomatin formation and/or stable silencing (Azuara et al., 2006; Hiratani et al., 2004; Hiratani et al., 2010; Hiratani et al., 2008; Perry et al., 2004).

ESC-specific chromatin configuration *in vitro* replicates the correspondent state of their *in vivo* early embryo counterpart, and is directly linked with lineage commitment (Fig.1.9). Chromatin in the pluripotent ICM of E3.5 embryos appears highly dispersed, and this configuration is established as early as the 8-cell embryo stage. Upon lineage segregation, chromatin compaction emerges in the blastomeres restricted to TE and PrE. Thus, higher degree of compaction can also be detected in embryo-derived XEN and TS cells as compared to ESCs (Ahmed et al., 2010).

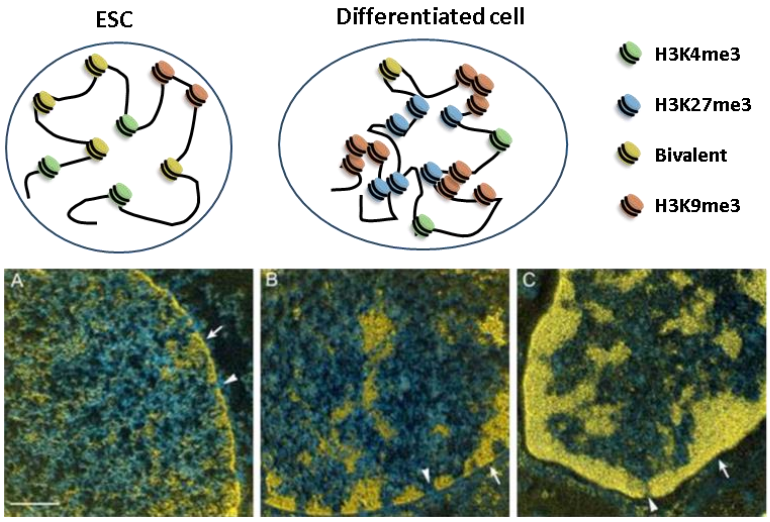


Figure 1.9 Genome organization in ESC and Differentiated cells

Schematic representation of the level of genome compaction as visualized in the nucleus of ESCs and differentiated cells. ESCs are characterized by dispersed chromatin fibres, poorly enriched in repressive foci of H3K9me3. Differentiated/somatic cells display a higher degree of chromatin compaction with blocks of condensed chromatin rich in H3K9me3 and H3K27me3. Adapted from (Chen and Dent, 2014). Bottom panel shows electron spectroscopic imaging of E3.5 epiblast (A), lymphocyte (B), and a spleen lymphocyte (C), showing the acquisition of chromatin compaction near the nuclear envelope in more differentiated cells (arrows). Bottom panel image from (Rapkin et al., 2012).

1.5.3.3 Bivalent chromatin domains

Accessible chromatin state provides a permissive environment for active transcription a seen across a large number of ESC-specific and differentiation-associated genes in ESCs. Thus, repressive mechanisms must be put in place in order to maintain lineage-affiliated genes silent, yet ready for activation during development. The repression of development genes is imposed through the action of the above-mentioned Polycomb Repressive Complexes (PRC), originally discovered as gene repressors in *Drosophila melanogaster* (Schuettengruber et al., 2007). The two PRC groups have distinct repressive enzymatic functions at gene promoter regions. PRC2 contains the enzyme Ezh2, which catalyses tri-methylation of H3K27 (H3K27me3). In turn, the Cbx proteins of the PRC1 recognize H3K27me3 and subsequently the enzymes Ring1a and Ring1b catalyse the monoubiquitylation of H2AK119 (H2AK119ub1). Conversely, the SET1A/B and MLL

family of proteins catalyse the methylation of H3K4 (H3K4me2/3) at promoters of genes, typically associated with active transcription (Clouaire et al., 2012; Hu et al., 2013). Notably, both active and repressive marks, together with PRC1 and PRC2 complexes, have been found to simultaneously occupy a large number of inactive, lineage-associated gene promoters in ESCs, thus termed bivalent chromatin structure (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006b). This distinctive chromatin configuration has also been identified in ICM and early epiblast cells of the embryo (Alder et al., 2010; Rugg-Gunn et al., 2010) and has been proposed to hold transcription in a poised state, allowing rapid activation (or continued repression) of genes upon differentiation stimuli as monitored in ESCs (Azuara et al., 2006; Boyer et al., 2006; Pasini et al., 2004).

Accumulating evidence from genome-wide transcriptional and epigenetic profiling has shown that during differentiation, the bivalent chromatin structure is resolved to cell-specific gene expression programs, but can also be acquired *de novo* at subsets of genes. Comparing ESCs, MEFs, and ESC-derived neuronal progenitor cells (NPC) it was found that most of the bivalently marked promoters become either activated by losing the PRC-mediated repressive H3K27me3, or repressed by losing the H3K4me3 mark. Yet, 4% (MEFs) and 8% (NPC) retained the bivalent marks at the promoters of genes of related lineages (Fig.1.10) (Mikkelsen et al., 2007). In a subsequent study comparing ESCs and terminally differentiated neurons, it was independently found that, while bivalency was resolved at a large proportion of promoters, it was also regained at subsets of previously unmarked regions (Mohn et al., 2008).

Moreover, during differentiation additional repressive marks such as DNA methylation and H3K9 methylation have also been found to be *de novo* acquired at bivalently marked regions to reinforce the silencing of lineage-inappropriate genes (Alder et al., 2010; Senner et al., 2012). Active and bivalent domains strongly correlate with CpG islands (Bernstein et al., 2006; Mohn et al., 2008), which lack DNA methylation and H3K9me2/3 deposition in ESCs, consistently with their mutual exclusion with H3K4me2/3 marks (Lienert et al., 2011; Mikkelsen et al., 2007; Mohn

et al., 2008). This is also in concordance with the presence of CpG dense DNA sequences that are likely involved in the recruitment at active and bivalent promoters of H3K4 methyltransferases SET1A/B and MLL2, respectively (Clouaire et al., 2012; Hu et al., 2013), two activating complexes known to specifically recognize unmethylated CpG islands via their DNA-binding domains (Birke et al., 2002; Lee et al., 2001). Yet these regions were selectively found to be targeted by Suv39h1-mediated H3K9 methylation and *de novo* DNA methylation without losing their activating H3K4 methylated marks upon ESC conversion into TS-like phenotype. This resulted into the formation of trivalent chromatin structure associated with the silencing of lineage-appropriate genes in extra-embryonic cells as furthermore confirmed *in vivo* (Alder et al., 2010).

Another characteristic feature of bivalent promoters is the assembly of a poised configuration of RNA Polymerase II (RNAPII) complexes (Brookes and Pombo, 2009; Stock et al., 2007). Gene expression in somatic cells typically relies on promoter recognition by a pre-initiation complex (PIC), and enhancer-promoter bridging by the Mediator complex, which, in turn, facilitates the recruitment of hypophosphorylated form of RNAPII. Targeted phosphorylation at Ser5 of the C-terminal domain (CTD) releases RNAPII from the interaction with the Mediator complex, keeping RNAPII arrested at pausing sites (Core and Lis, 2008). Phosphorylation on Ser2 of the CTD further promotes the stable assembly of the elongation complex, necessary for the progression of transcription (Marshall et al., 1996). Ser2 phosphorylation further induces elongation through the recruitment of H3K36-methyltransferases for chromatin remodelling as well as regulating splicing by recruitment of polyadenylation factors (Krogan et al., 2003; Proudfoot et al., 2002). Comparatively, bivalent promoters in ESCs were found to be selectively occupied by poised Ser5 phosphorylated RNAPII with no detectable Ser2 phosphorylated or hypophosphorylated forms at these sites (Brookes and Pombo, 2009; Stock et al., 2007). The lack of transcriptional elongation, and hence, Ser2P RNAP at PRC-repressed promoters mirrors the low or absent levels of expression of lineage-affiliated genes. However, a subset of PRC targets, with functions involved in development and

metabolism, can be dynamically switch between repressed and active states within an ESC population due to either cellular heterogeneity or allelic differences (Brookes et al., 2012)

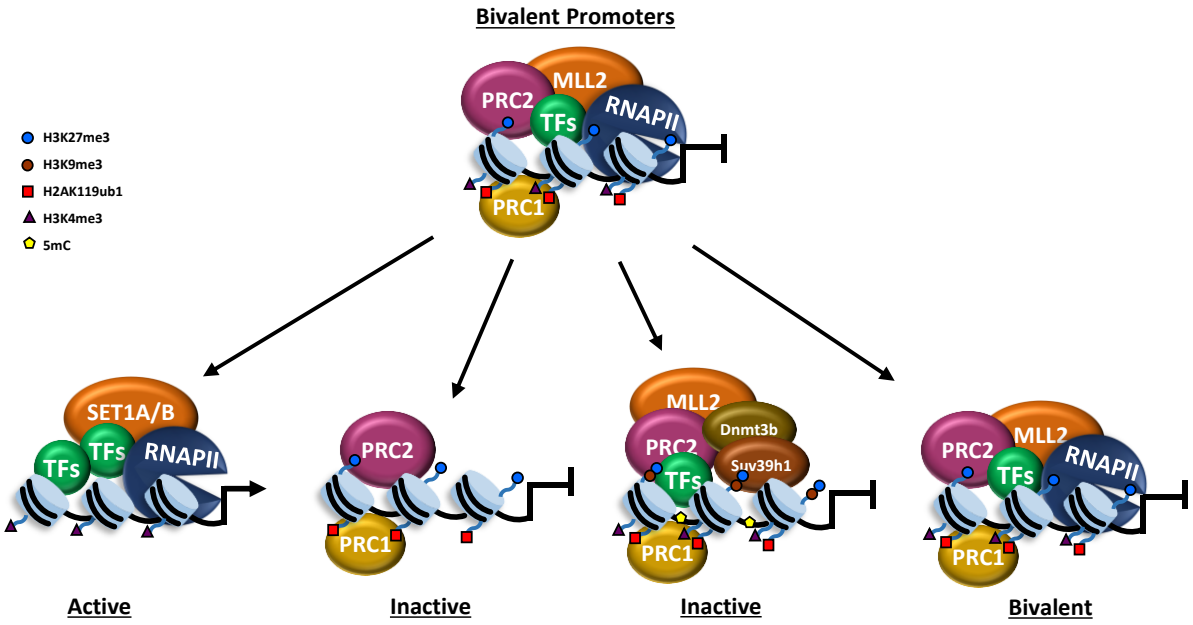


Figure 1.10 Possible outcomes for bivalent promoters upon differentiation of ESCs

Schematic representation of the bivalent chromatin structures that characterize ESCs. Promoters of lineage-affiliated genes are co-occupied by PRC1, PRC2, MLL2 and enriched in active (H3K4me2/3) and inactive (H3K27me3, H2AK119ub1) histone marks, and a poised conformation of RNAPII (closed). Upon differentiation, promoters can lose the PRC-mediated repression, RNAPII becomes active (open) and transcription progresses. Lineage-inappropriate genes can be silenced by loss of active mark, or by further acquisition of repressive marks (H3K9me3, 5mC) without losing the H3K4me2/3 decoration. A small subset of genes might retain the bivalent configuration in lineage-committed progenitors and derivatives.

1.5.3.4 Pre-patterning of enhancers

Enhancers play a key role in driving strong cell-specific gene expression, and are able to activate genes located several kilobases (kb) apart. Typically, enhancers are bound by lineage-specific TFs, which in turn recruit co-activators, chromatin remodellers and proteins involved in mediating long-range interactions with the transcriptional machinery (Calo and Wysocka, 2013). Nucleosomes flanking TF binding sites at enhancer regions are commonly enriched in H3K4me1 and H3K27ac, which together with DNAse I hypersensitivity typically mark active enhancers

(Creyghton et al., 2010; Gross and Garrard, 1988; Zentner et al., 2011). Furthermore, H3K4me1 has been found to precede nucleosomal depletion and H3K27ac deposition occurring upon enhancer activation, thus being associated with pre-marking of development-associated enhancers in ESCs (Creyghton et al., 2010; Zentner et al., 2011). Conversely, H3K4me3 is predominantly found at promoters, whereas H3K4me2 can be found at both enhancers and promoters (Pekowska et al., 2011). Binding of pioneer TFs to inactive enhancers has been proposed to trigger chromatin remodelling by nucleosomal repositioning, recruitment of additional TFs, co-factors and chromatin modifiers that together are required for gene activation (Fig.1.11). Given the association of H3K4me1/2 with poised enhancer regions, this has been proposed to facilitate TF binding and nucleosomal mobility, as well as potentially protecting these sites from acquisition of DNA methylation (Ooi et al., 2007; Zaret and Carroll, 2011). However, it is not clear whether pre-marking of enhancers with H3K4me1 is strictly required to precede the binding of pioneer TFs. Indeed, several TFs including Foxa2, Gata6, Oct4, Sox2 and Klf4 have been proposed to directly access and bind DNA motifs within compact chromatin, in order to facilitate the recruitment of additional factors, thus having a key role in cell-specific gene activation (Soufi et al., 2012; Xu et al., 2011). Furthermore, poised enhancers in human ESCs have been found to be already looped to their target promoters, whereas in differentiated cell types, enhancer-promoter looping events is specifically associated with gene expression (Sanyal et al., 2012).

H3K27ac marking active enhancers is catalysed by the ubiquitously expressed proteins p300 and CBP, which can be recruited via distinct TFs in a tissue-specific manner (Bonn et al., 2012; Creyghton et al., 2010; Goodman and Smolik, 2000; Zentner et al., 2011). Notably, H3K27ac is specifically associated with active, but not with disengaged or poised enhancers, even though binding of p300 can be detected before H3K27ac deposition and activation of poised enhancers (Rada-Iglesias et al., 2011). In addition, a subset of H3K4me1-marked poised enhancers lacking H3K27ac is instead marked by PRC2-mediated H3K27me3 (Rada-Iglesias et al., 2011; Zentner et al., 2011).

Another feature of poised enhancers is the presence of an intermediate form of cytosine methylation termed hydroxymethyl-cytosine (5hmC), mediated by the TET enzymes (Serandour et al., 2011; Stroud et al., 2011). Unlike 5mC, 5hmC positively correlates with active gene expression and is most abundantly found at both poised and active enhancers rather than at CpG-dense promoter regions (Yu et al., 2013). This proposes that priming of enhancers might involve an active DNA demethylation mechanism mediated by the TET proteins.

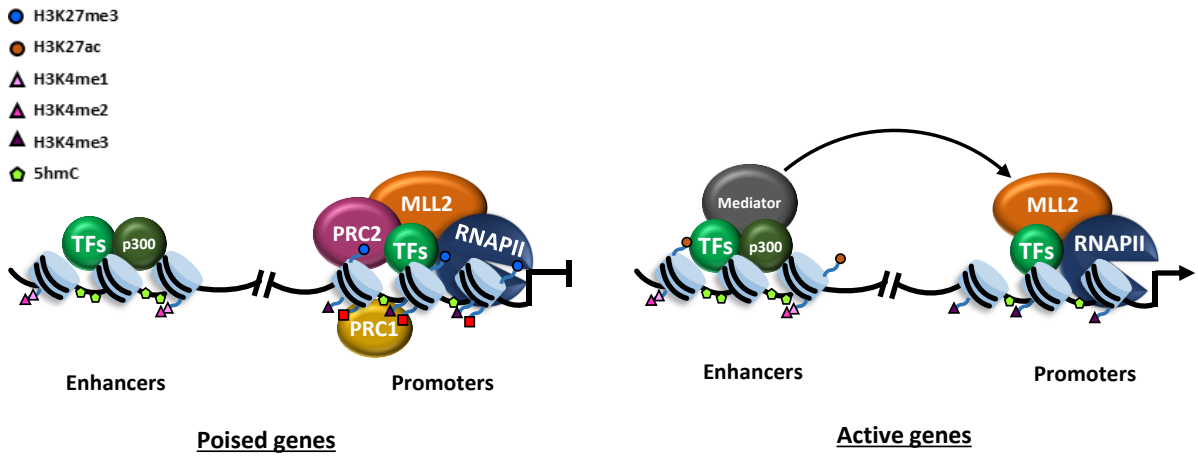


Figure 1.11 Configuration of poised and active gene regulatory regions in ESCs

Schematic representation of chromatin structures that characterize poised and active genes in ESCs. Promoters of poised lineage-affiliated genes are bivalently marked by H3K4 and H3K27me2/3 marks, repressed by PRC1 and PRC2 and occupied by poised (ser5 phosphorylated) RNAPII complexes (closed), whereas enhancers are marked by H3K4me1/2 and harbour 5hmC. Putative pioneer TFs and p300 binding might mark poised enhancers prior to activation. Active promoters are free of the repressive action of PRC1 and PRC2, assemble active RNAPII complexes (open) and are active via long-range interaction of distal enhancers enriched in H3K4me1/2, H3K27ac, TF binding and co-activators such as p300 and Mediator.

Altogether, the tight regulation of chromatin decompaction, TF binding, protection from silencing epigenetic marks, and long-range DNA crosstalk in ESCs, might altogether facilitate rapid and efficient activation of enhancers and promoters of lineage-affiliated genes upon differentiation.

I.6 Histone demethylases and their role in ESCs

1.6.1 Families of histone demethylases

Up until a decade ago, histone methylation was considered to be an irreversible mark. However, this idea was challenged by the discovery of the first lysine-specific demethylase (LSD1/KDM1), which was found to demethylate both di- and mono-methylated K4 (H3K4me2/1) (Shi et al., 2004). Another class of histone demethylase enzymes containing a Jumonji C-terminal (JmjC) domain, was later identified and shown to be able to demethylate all three states of lysine-methylation (Klose et al., 2006a). These two classes of enzymes demethylate histones via different mechanisms. The first utilizes flavin adenine dinucleotide (FAD) as a co-factor for amine oxidation, whereas JmjC-domain containing enzymes directly hydrolyse the methyl group utilising Fe(II) and αKG as co-factors (Shi et al., 2004; Tsukada et al., 2006). Several JmjC-domain containing proteins have since been identified and grouped into different sub-classes based on additional protein domains and substrate specificity, recognising histone substrates that embrace methylated lysine 4, 9, 27 and 36 of histone H3 (Table 1).

Family	Family member	Other Names	Histone Substrates
KDM1	KDM1	LSD1; AOF2; BHC110	H3K4me2/1; H3K9me2/1
KDM2	KDM2A	JHDM1A; FBXL11	H3K36me2/1
	KDM2B	JHDM1B; FBXL10	
KDM3	KDM3A	JHDM2A; JMJD1A; TSGA	H3K9me2/1
	KDM3B	JHDM2B; JMJD1B	
KDM4	KDM4A	JMJD2A; JHDM3A	H3K9me3/2; H3K36me3/2
	KDM4B	JMJD2B; JHDM3B	
	KDM4C	JMJD2C; GASC1; JHDM3C	
	KDM4D	JMJD2D; JHDM3C	
KDM5	KDM5A	RBP2; JARID1A	H3K4me3/2
	KDM5B	PLU-1; JARID1B	
	KDM5C	SMCX; JARID1C	
	KDM5D	SMCY; JARID1D	
KDM6	KDM6A	UTX	H3K27me3/2
	KDM6B	JMJD3	

Table 1. Histone demethylase nomenclature and histone substrates.

Adapted from (Nottke et al., 2009)

Accumulating evidence has shown that the concerted action of lysine methyltransferases (KMT) and demethylases (KDM) is crucial in the tight regulation of gene expression in ESCs and upon differentiation. As described earlier, promoters of lineage-affiliated genes are enriched in MLL2-mediated active (H3K4me3) and PRC2-mediated repressive (H3K27me3) marks, which upon differentiation resolve into monovalent chromatin domains for either activation or repression. During this process, the H3K27me3 demethylase Utx/Kdm6a was found to associate with the KMT complex MLL2, and hence promote activation of bivalent promoters upon differentiation together with Jmjd3/Kdm6b (Agger et al., 2007; Cho et al., 2007; Lee et al., 2007). Strikingly, the similarity of defective development found in *Xenopus* and *Zebrafish* embryos upon depletion of either H3K4 or H3K27 methylation suggests that the removal of H3K27me3 is functionally coordinated with the addition of H3K4me3 (Lan et al., 2007; Wysocka et al., 2005). Moreover, the H3K4me3 KDM Jarid1a/Kdm5a has also been found to extensively overlap and physically interact with the PRC2 at the chromatin of ESCs. Depletion of PRC2 impairs Jarid1a binding to promoters, suggesting that PRC2 recruits the H3K4me3-demethylase, and that are then both lost upon differentiation-induced gene activation (Pasini et al., 2008). Thus, the co-existence of opposing enzymatic activities occurring at the same complex is likely to be a vital mechanism for balancing H3K4 and H3K27 methylation levels, and ensure robust gene regulation during development.

Additional roles for, KDMs independent of their demethylase enzymatic activity have been most recently associated with PRC-mediated repression. The H3K36-demethylase Fbxl10/Kdm2b was, for instance, shown to specifically recognize unmethylated CpG islands where it recruits non-canonical PRC1 complexes, a process found to be necessary for the deposition of the repressive H2AK119ub1 mark at developmental-associated gene promoters. This was, however, not linked with removal of H3K36me3, a mark normally found at gene bodies of transcriptionally active genes (Farcas et al., 2012; Wu et al., 2013).

1.6.2 The Jmjd2/Kdm4 family of histone demethylases

The human JMJD2/KDM4 gene family was identified and characterized in 2004 (Katoh and Katoh, 2004). Structurally, JMJD2A, 2B and 2C proteins comprise one JmjN (N-terminal), one JmjC (C-terminal), two PHD (plant homeobox domain), and two Tudor domains (Fig.1.12). In contrast, JMJD2D, E and F are smaller proteins that only contain the two JmjN and JmjC domains. JMJD2A-D have orthologs in the mouse, whereas JMJD2D is likely to have originated the two additional human genes, JMJD2E and JMJD2F, through local retrotransposition (Yang et al., 2000). These proteins are highly conserved from yeast to human, suggesting a ubiquitously important role in gene regulation (Katoh and Katoh, 2004; Klose et al., 2006a). JMJD2A-C proteins have so far been found to demethylate H3K9me2/3, H3K36me2/3 and H1.4K26me3, whereas JMJD2D was proposed to demethylate these same histone modifications except H3K36me3 (Chen et al., 2006; Klose et al., 2006b; Shin and Janknecht, 2007; Trojer et al., 2009; Whetstine et al., 2006). This function does not rely on the PHD or Tudor domains, but requires both the JmjC and JmjN domains. Instead, the later domains are required for the retention of JMJD2 proteins in the nucleus, as deletion of these domains results into the relocation of JMJD2 to the cytoplasm (Shin and Janknecht, 2007).

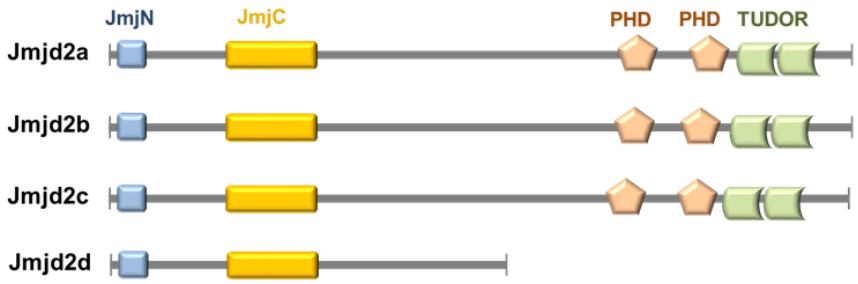


Figure 1.12 Members of the JMJD2 family in Human/Mouse.

Members of the JMJD2 family contain two jumonji catalytic domains: JmjN (Jumonji N-terminal; blue) and a JmjC (Jumonji C-terminal; yellow) domain. All members of this group apart from Jmjd2d contain additional C-terminal domains such as PHD (orange) and Tudor (green) domains, which were proposed to be involved in protein targeting (adapted from Klose et al., 2006a).

The identified recognition of activating and repressive histone marks suggests wide-ranging roles for this family of histone demethylases in gene activation. Thus, JMJD2 proteins have been found to act as co-regulators of hormone and hypoxia-mediated gene response, in addition to roles in DNA repair and chromosome stability (Berry and Janknecht, 2013; Labbe et al., 2013). Moreover, JMJD2 members have been associated to histone demethylase-mediated regulation of gene expression in cellular differentiation processes such as adypogenesis, osteogenesis and muscle differentiation (Lu et al., 2012; Verrier et al., 2011; Ye et al., 2012). Not surprisingly, abnormal expression of JMJD2 members has been extensively linked with deregulated gene expression in cancer, making them potential targets for the development of anti-cancer drugs (Hamada et al., 2010; Yu et al., 2012). In addition to histone modifications, JMJD2A-C members have also been found to demethylate tri-methylated lysine containing peptides from WIZ, CDYL1, CSB and G9a proteins at least as tested *in vitro* (Ponnaluri et al., 2009). Whether these events also occur *in vivo* and furthermore what is the biological significance of non-histone demethylation by the JMJD2 family are questions that remain to be investigated.

Although less studied, the Jmjd2 histone demethylases were also proposed to play a role in early embryonic mouse development. *In silico* expression analysis of the mouse Jmjd2 gene family revealed that *Jmjd2a*, *Jmjd2b* and *Jmjd2c* might be preferentially expressed in cleavage-stage embryos, with *Jmjd2c* being also highly expressed in undifferentiated ESCs (Katoh and Katoh, 2007). In the early embryos, *Jmjd2c* was found to be expressed from the 2-cell stage, reaching its peak of expression at the 4-cell stage, subsequently being detected though at lower levels throughout pre-implantation development (Wang et al., 2010). Consistently, dsRNA-knockdown of *Jmjd2c* in oocytes and induced parthenogenesis led to early developmental arrest at the transition from 8-cell to morula stage in depleted embryos as compared to controls (Wang et al., 2010). Moreover, this was accompanied by a reduction in the expression levels of Oct4, Nanog and other self-renewal genes as seen both at the mRNA and protein level in *Jmjd2c*-depleted 4-cell stage embryos (Wang et al., 2010). Specific knockdown of *Jmjd2c* in mouse ESCs using shRNA

vectors similarly suggested an essential role for Jmjd2c in sustaining ESC self-renewal. *Jmjd2c* was found to be a direct target of Oct4 and depletion of this demethylase resulted into a loss of the ESC undifferentiated state via deregulation of *Nanog* expression that strictly correlated with an increase in H3K9me3 enrichment levels at its promoter region (Loh et al., 2007).

I.7 Hypothesis and Aims

As highlighted in this Chapter, tight regulation of the pluripotent state is achieved by coordinated regulatory networks integrating cell signaling cues with the action of transcription factors and epigenetic remodellers to ultimately execute the appropriate gene expression programs that sustain ESC self-renewal, while silencing lineage-affiliated genes. Notably, the transcription of differentiation-associated factors is kept in a poised state, allowing for efficient activation upon differentiation stimuli. Thus, balancing both activating and repressor signals that feed onto the activity of promoters and enhancers of differentiation markers is a crucial mechanism in safeguarding gene priming in the ESC state.

Previous work in our lab has identified that ESC-specific and somatic affiliated genes are stably silenced by acquisition of additional epigenetic repressor layers in extra-embryonic stem cells and *in vivo*. In particular, Suv39h1-mediated H3K9me3 coupled with *de novo* DNA methylation, is acquired at a subset of bivalent promoters upon ESC conversion into a TS-like state, which can be achieved upon doxycycline-dependent Oct4 removal (Fig.1.13) (Alder et al., 2010). Here, the H3K9 KMT *Suv39h1* was identified to be dynamically up-regulated, with the H3K9 KDM *Jmjd2c* being oppositely extinguished upon trophoblast lineage commitment (Fig1.13), consistent with *Jmjd2c* expression being under direct control of *Oct4* (Loh et al., 2007).

The discovery that a set of bivalent promoters is targeted by Suv39h1-mediated H3K9me3 and hence harboring a repressed trivalent (H3K27me3/H3K9me3/H3K4me3) conformation in TS cells suggested that in ESCs, an active mechanism of H3K9 demethylation might be operating to prevent poised promoters from acquiring additional repressive layers.

Given the rapid and specific down-regulation of *Jmjd2c* upon extra-embryonic lineage commitment (Alder et al., 2010), and the previously reported requirement for *Jmjd2c* in maintaining an ESC identity (Loh et al., 2007), Jmjd2c was here proposed to play a possible role in actively safeguarding bivalent chromatin structure and lineage gene priming both in the ESCs and in the early embryo.

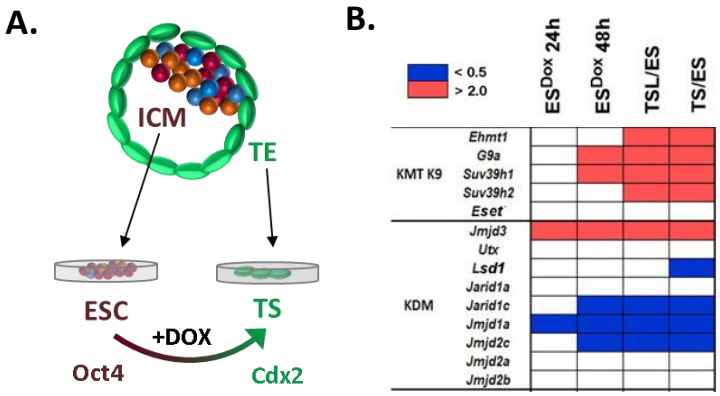


Figure 1.13 Gene expression profiles of chromatin regulators upon ES to TS-like conversion

A. ESCs are derived from the pluripotent ICM, and TS cells from the extra-embryonic TE. Conversion of ESC into a TS-like Cdx2-expressing cell population can be achieved by conditional depletion of the expression of *Oct4* transgene via doxycycline (DOX) treatment in the ZHBTc4 ESC line. **B.** Expression of chromatin modifiers during ESC to TS-like (TSL) conversion, and in embryo-derived TS cells. Blue and red represent a decrease in expression of at least 0.5 fold, and an increase of at least 2 fold respectively as compared to ESCs. *Jmjd2c* is specifically down-regulated in this process unlike *Jmjd2a* or *Jmjd2b*. Figure **B** from (Alder et al., 2010).

Up until the beginning of this project, little was known about the molecular mechanisms by which Jmjd2c function is conferred in pluripotent ESCs. Thus, the overall aim of this project was to elucidate the role of the Jmjd2c histone demethylase in undifferentiated ESCs and during the early steps of differentiation, taking advantage of newly generated constitutive *Jmjd2c*-knockout mouse ESCs as primary tools. Preliminary work in our laboratory had shown that these cells could be maintained in an undifferentiated state in culture (German S, MSc Thesis 2010),

hence providing an notable opportunity for further investigating how Jmjd2c functions in pluripotent ESCs.

In order to assess whether Jmjd2c indeed safeguards the bivalent chromatin structures of developmental-associated genes in pluripotent ESCs, specific steps of investigation were undertaken. First, careful validation of the established knockout and characterization of the phenotype of *Jmjd2c*-depleted ESCs was conducted, in addition to validating the demethylase activity of Jmjd2c in ESCs. Secondly, the impact of *Jmjd2c* depletion was thoroughly validated in differentiation, with particular focus on the early steps of pluripotency exit. Finally, the chromatin binding profile of Jmjd2c was explored, and the consequence of Jmjd2c depletion in regulatory regions of lineage-affiliated markers was identified.

Chapter II – Materials and Methods

II.1 Materials

2.1.1 Antibodies

Primary Antibody	Host species	Catalogue Number	Supplier	Concentration/Quantity			
				Western Blot	IF	ChIP	Co-IP
Dab2	rabbit	sc13982	SantaCruz	-	1:250	-	-
E-Cadherin	rat	U3254	Sigma	-	1:500	-	-
Esrrb	mouse	H6705	R&D Systems	1:1000	-	-	-
FLAG M2	mouse	F1804	Sigma	1:1000	1:500	10µg	-
G9a	rabbit	3306	Cell Signalling	1:1000	-	0.24µg	5 µg
G9a	mouse	PP-A8620A-00	R&D Systems	-	-	-	5 µg
Gata4	rabbit	sc-9053	SantaCruz	-	1:250	-	-
Gata6	goat	AF1700	R&D Systems	-	1:500	-	-
GFP	rabbit	ab290	Abcam	-	1:1000	-	-
GLP	mouse	PP-B0422-00	R&D Systems	1:1000	-	-	5µg
H3	rabbit	ab1791	Abcam	1:50K	-	-	-
H3K27ac	rabbit	ab4729	Abcam	1:100K	-	5µg	-
H3K27me3	rabbit	07-449	Millipore	1:100K	-	5µg	-
H3K36me3	mouse	mAb-183-050	Diagenode	1:10K	-	-	-
H3K4me2	rabbit	07-030	Millipore	1:50K	-	5µg	-
H3K9me2	mouse	ab1220	Abcam	1:20K	-	2µg	-
H3K9me2	mouse	mAb-154-050	Diagenode	1:50K	-	-	-
H3k9me3	rabbit	07-442	Millipore	1:1000	-	-	-
H3K9me3	rabbit	ab8898	Abcam	1:100K	-	-	-
H3K9me3	mouse	mAb-153-050	Diagenode	1:100K	-	-	-
Jmjd2b	rabbit	A301-478A	Bethyl Labs	1:2000	-	-	-
Jmjd2c	rabbit	-	H. Ng's lab	1:2000	-	-	-
Jmjd2c	goat	sc-104949	SantaCruz	1:2000	-	-	-
Med1	rabbit	A300-793A	Bethyl Labs	1:1000	-	8µg	5 µg
Nanog	rabbit	RC-AB002PF	Reprocell	1:2000	-	-	-
Oct4	goat	sc-8628	SantaCruz	1:1000	-	5µg	-
Oct4	mouse	sc-5279	SantaCruz	-	1:200	-	-
p300	rabbit	sc-585	SantaCruz	1:1000	-	5µg	-
Ring1b	mouse	39663	Active Motif	-	-	10µg	-
Smc1	rabbit	A300-055A	Bethyl Labs	1:1000	-	8µg	-
Sox7	goat	AF2766	R&D Systems	-	1:500	-	-
Suz12	rabbit	pAb-029-050	Diagenode	-	-	10µg	-
V5	mouse	46-0705	Invitrogen	1:2000	1:1000	-	-
α-Tubulin	mouse	T6074	Sigma-Aldrich	1:20K	-	-	-

Secondary Antibodies	Catalogue Number	Supplier	Concentration	
			Western Blot	IF
Alexa488 Goat Anti-Mouse IgG	A11034	ThermoFisher	-	1:500
Alexa488 Goat Anti-Rabbit IgG	A11029	ThermoFisher	-	1:500
Alexa488 Donkey Anti-Goat IgG	A11055	ThermoFisher	-	1:500
Alexa488 Goat Anti-Rat IgG	A11006	ThermoFisher	-	1:500
Alexa633 Goat Anti-Mouse IgG	A21052	ThermoFisher	-	1:500
Alexa633 Goat Anti-Rabbit IgG	A21071	ThermoFisher	-	1:500
Goat Anti-Mouse IgG-HRP	sc-2004	SantaCruz	1:2000	-
Goat Anti-Rabbit IgG-HRP	sc-2005	SantaCruz	1:5000	-
Donkey Anti-Goat IgG-HRP	sc-2020	SantaCruz	1:5000	-

Normal IgG	Catalogue Number	Supplier	Amount	
			ChIP	Co-IP
Normal Mouse IgG	sc-2025	Santa Cruz	2.5µg	5µg
Normal Rabbit IgG	sc-2027	Santa Cruz	2.5µg	5µg

2.1.2 Plasmids

Expression vectors	Source
pPyCAG-IRES-Puro(IP)	I.Chambers's lab
pPyCAG-IP- Jmjd2c	This work
pPyCAG-IP-Flag(2x)V5	R. Poot's lab
pPyCAG-IP-Flag(2x)V5-Jmjd2c	This Work
pPyCAG-IP-GFP	J. Nichols' lab
pPyCAG-IP-dsRed*MST	J Nichols' lab

Short Hairpin RNA vectors	Source
pLKO.1 shLuciferase	H. Ng's lab
pLKO.1 shJmjd2c1	H. Ng's lab
pLKO.1 shJmjd2c2	H. Ng's lab

2.1.3 Antibiotics

Bacterial Selection	Working Concentration	Catalogue Number	Source
Ampicillin	100µg/ml	10193433	FisherScientific

Eukaryotic Selection	Working Concentration	Catalogue Number	Source
Geneticin (G418)	100µg/ml	P11-012	PAA
Hygromicin	25µg/ml	P02-015	PAA
Puromycin	1µg/mL	P7255	Sigma-Aldrich

2.1.4 Enzymes

Enzyme	Catalogue Number	Source
RNase Out	10777-019	Lifetechnologies
Superscript II Reverse Transcriptase	18064014	Lifetechnologies
Superscript III Reverse Transcriptase	18080044	Lifetechnologies
KOD HotStart DNA Polymerase	71086-5	Millipore
BamHI	R3136	New England Biolabs
PacI	R0547	New England Biolabs
SalI	R3138	New England Biolabs
XhoI	R0146	New England Biolabs
Benzonase	70664	Novagen
Micrococcal nuclease	N3755	Sigma
Proteinase K	P2308	Sigma-Aldrich
Ribonuclease A	R6513	Sigma-Aldrich
SYBR Green Taq JumpStart Polymerase	S4438	Sigma-Aldrich
SYBR Green Taq Kicqstart Polymerase	KCQS02	Sigma-Aldrich

2.1.5 Kits

Kit	Catalogue Number	Source
Amersham ECL Prime Western Blotting Detection Reagent	25080060	GE Healthcare
EpiMark 5-hmC and 5-mC Analysis Kit	E3317S	New England Biolabs
NEBNext® Multiplex Oligos for Illumina® (Primers Set 1)	E7335	New England Biolabs
NEBNext® Ultra™ DNA Library Prep Kit for Illumina	E7370	New England Biolabs
Quick Ligation Kit	M2200S	New England Biolabs
HiSpeed Plasmid Maxi Kit	12663	Qiagen
RNAeasy mini	74106	Qiagen
RNase-Free DNase Set (50)	79254	Qiagen
Leukocyte Alkaline Phosphatase Kit	86R-1KT	Sigma-Aldrich

2.1.6 Inhibitors

Protease Inhibitors	Working Dilution	Catalogue Number	Source
Phenylmethylsulfonyl fluoride (0.1M)	1:1000	93482	Sigma-Aldrich
Complete Protease inhibitor tablets (50x)	1:50	11697498001	Roche

Phosphatase Inhibitors	Working Concentration	Catalogue Number	Source
Sodium Fluoride (NaF)	5mM	201154	Sigma
Sodium Orthovanadate (Na_3VO_4)	2mM	S6508	Sigma

Signalling Pathway Inhibitors	Working Concentration	Catalogue Number	Source
CHIR99021 - GSK3 inhibitor	3µM	130-103-926	Miltenyi Biotech
PD0325901 - MEK inhibitor	1µM	04-0006	Stemgent
LY294002 – Pi3K inhibitor	2-10nM	L9908	Sigma-Aldrich
LDN193189 – BMP receptor inhibitor	250nM	SML0559	Sigma-Aldrich

2.1.7 Growth factors media supplements

Supplement	Working Dilution	Catalogue Number	Source
Activin A	20ng/ml	-	In house
bFgf / Fgf2	12-20ng/ml	100-18B	Peptrotech
Transferrin	15µg/ml	652202	Roche
LIF	Batch-dependent	-	In house
Mono	450µM	M6145	Sigma
Insulin	7µg/ml	1376497	Roche
polyvinyl alcohol (PVA)	0.1mg/mL	P8136	Sigma-Aldrich
Heparin	1µg/mL	H3393	Sigma-Aldrich
BMP4	10-50 ng/ml	314-BP-010	R&D
Wnt3a	2ng/ml	315-20	Peprotech
N2	1x	17502-048	Lifetechnologies
B27 (-vitamin A)	1x	12587-010	Lifetechnologies
B27	1x	17504-044	Lifetechnologies

2.1.8 Reagents

Reagent	Source
Acetone	Fisher Scientific
Acrylamide/Bis Solution	Biorad
Agarose	Sigma-Aldrich

All Trans Retinoic Acid	Sigma-Aldrich
Ammonium Persulfate (APS)	Sigma-Aldrich
Bovine Serum Albumin (BSA)	PAA
Deoxycholic acid (DOC)	Sigma-Aldrich
Deoxynucleotide Triphosphate (dNTPs) 10mM	Lifetechnologies
DEPC-Treated Water	Ambion
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich
DL-Dithiothreitol (DTT)	Sigma-Aldrich
DL-Dithiothreitol (DTT) *for molecular biology*	Lifetechnologies
Dyalisis Tubing	ThermoScientific
Dynabeads Protein G-coupled	Lifetechnologies
Ethanol	Fisher Scientific
Ethidium Bromide	Sigma-Aldrich
Ethylene Diamine Tetraacetic Acid (EDTA)	Sigma-Aldrich
Ethylene Glycol Tetraacetic Acid (EGTA)	Sigma-Aldrich
FLAG coated sepharose beads	Sigma-Aldrich
FLAG tripeptide	Sigma-Aldrich
Formaldehyde Solution, 37%	Sigma-Aldrich
Glycerol	Sigma-Aldrich
Glycine	Sigma-Aldrich
Glycogen Blue (GlycoBlue) carrier	Ambion
Sulphuric acid (H_2SO_4)	Sigma-Aldrich
HEPES	Sigma-Aldrich
Horse Serum	Sigma-Aldrich
Isopropanol	Fisher Scientific
Laemmli Sample Buffer	BioRad
LB	MP Biomedicals
Lipofectamine 2000 Transfection Reagent	Lifetechnologies
Lithium Chloride (LiCl)	Sigma-Aldrich
Magnesium Sulphate (MgSO2)	Novagen
Methanol	Fisher Scientific
MgCl2	Lifetechnologies
Nonidet P-40 (NP-40)	Sigma-Aldrich
Normal Goat Serum	Sigma-Aldrich
Oligonucleotides	Lifetechnologies
Paraformaldehyde (PFA)	Sigma-Aldrich
Phenol-Chloroform	Sigma-Aldrich
Phosphate-buffered Saline (PBS) capsules	Fisher
Potassium Chloride (KCL)	Sigma-Aldrich
Protein A-Sepharose	Sigma-Aldrich
PVDF Membrane	GE Healthcare

Random Hexamers	Lifetechnologies
SDS-PAGE Resolving Buffer	Biorad
SDS-PAGE Stacking Buffer	BioRad
Skimmed Milk Powder	Fluka
Sodium Azide (NaN_3)	Sigma-Aldrich
Sodium Chloride (NaCl)	VWR
Sodium Deoxycholate (Na-Doc)	Sigma-Aldrich
Sodium Dodecyl Sulfate (SDS) Solution 10%	Fluka
Sodium Hidrogenocarbonate ($NaHCO_3$)	Sigma-Aldrich
Spermidine	Sigma-Aldrich
Spermine tetrahydrochloride	Sigma-Aldrich
Sucrose	Sigma-Aldrich
Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Trichloroacetic acid (TCA)	Sigma-Aldrich
Tris Base	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tween 20	Sigma-Aldrich
Vectashield with DAPI	Vector Laboratories
β-Mercaptoethanol	Sigma-Aldrich

2.1.9 Solutions

Alkaline phosphatase staining:

AP fixation solution	25.5%(v/v) Citrate solution 66.3%(v/v) Acetone 8.2%(v/v) Formaldehyde 37%
AP staining solution	1.9%(v/v) Na-Nitrite 1.9%(v/v)FRV-alkaline 94.3%(v/v) dH2O 1.9%(v/v) Naphthol Alkaline

Genomic DNA extraction:

Lysis Buffer	200mM NaCl 100mM Tris-HCL pH 7.4

	5mM EDTA 0.2%(w/v) SDS 0.5mg/ml Proteinase K (added fresh)

Immunofluorescent staining:

PFA 4%	4% (w/v) Paraformaldehyde in PBS 2mM NaOH
Blocking Buffer	90%(v/v) PBS 10%(v/v) serum (same species as the secondary antibody)
Permeabilization Buffer	0.4% TritonX-100 in Blocking Buffer

Western blot:

RIPA Buffer (Total Protein Extraction)	50mM Tris pH 8.0 1mM EDTA 0.5mM EGTA 1% (v/v) Triton X-100 0.1% (w/v) Na-deoxycholate 140mM NaCl Freshly supplemented with protease and phosphatase inhibitors.
Hypotonic lysis buffer (Histone extraction)	10 mM Tris–Cl pH 8.0 1 mM KCl 1.5 mM MgCl2 1 mM DTT Freshly supplemented with protease and phosphatase inhibitors.
8% Resolving gel	375mM Tris HCl pH 8.8 8% (w/v) Acrylamide/Bis 0.1% (w/v) SDS

	0.05% (w/v) Ammonium Persulfate 0.005% (v/v) TEMED
13% Resolving gel	375mM Tris HCl pH 8.8 13% (w/v) Acrylamide/Bis 0.1% (w/v) SDS 0.05% (w/v) Ammonium Persulfate 0.005% (v/v) TEMED
4% Stacking gel	125mM Tris HCl pH 6.8 4% (w/v) Acrylamide/Bis 0.1% (w/v) SDS 0.05% (w/v) Ammonium Persulfate 0.01% (v/v) TEMED
SDS Running Buffer	25mM Tris base 190mM Glycine 0.1% (w/v) SDS
Transfer Buffer	25 mM Tris, 192 mM glycine 0.01% (w/v) SDS 10% (v/v) methanol for proteins >50KDa or 20% (v/v) methanol for proteins <50KDa
TBS-T	130mM NaCl 20mM Tris pH 7.6 0.1% (v/v) Tween-20

FLAG pull down:

Buffer A	10 mM Hepes 7.6 1.5 mM MgCl2 10 mM KCl Freshly supplemented with 0.5 mM DTT and 1x complete protease inhibitor tablets

Buffer C	20 mM Hepes pH 7.6 420 mM NaCl 1.5 mM MgCl2 0.2 mM EDTA 20% (v/v) glycerol Freshly supplemented with 0.5 mM DTT and 1x complete protease inhibitor tablets
Buffer D	20 mM Hepes pH 7.6 100 mM KCl 1.5 mM MgCl2 0.2 mM EDTA 20% (v/v) glycerol Freshly supplemented with 0.5 mM DTT, 1x complete protease inhibitor tablets and 0.5µM sodium Meta bisulfite
Buffer C100	20 mM Hepes pH 7.6 100 mM KCl 1.5 mM MgCl2 0.2 mM EDTA 20% (v/v) glycerol 0.02% (v/v) NP-40 Freshly supplemented with 0.5 mM DTT and 1x complete protease inhibitor tablets

Protein co-immunoprecipitation Protocol A:

Hypotonic Buffer	10 mM Tris-HCl pH 7.6 1.5 mM MgCl2 10 mM KCl
Buffer SR	50 mM HEPES pH 7.0 0.25 mM EDTA 10 mM KCl 70% (w/v) sucrose

	0.15mM spermine
	0.15mM spermidine
High Salt Buffer	900 mM NaCl
	20 mM Tris pH 7.65
	25%(v/v) glycerol
	1.5 mM MgCl2
	0.2 mM EDTA

Protein co-immunoprecipitation Protocol B:

Buffer A	20 mM HEPES pH 7.0
	0.15 mM EDTA
	0.15 mM EGTA
	10 mM KCl
	0.15mM spermine
	0.15mM spermidine
Buffer SR	50 mM HEPES pH 7.0
	0.25 mM EDTA
	10 mM KCl
	70% (w/v) sucrose
	0.15mM spermine
	0.15mM spermidine
Buffer Sucrose	20mM Tris pH 7.65
	60 mM NaCl
	15 mM KCl
	0.34 M sucrose
	0.15mM spermine
	0.15mM spermidine
Wash Buffer	50 mM Tris pH 7.65
	150 mM NaCl
	0.5%(v/v) Triton X-100

TENG	20mM Tris pH 7.65 0.1 mM EDTA 150 mM NaCl 10%(v/v) glycerol 0.01% (v/v) NP-40

Protein co-immunoprecipitation Protocol C:

Cell Lysis Buffer	10 mM Tris pH 7.9 10 mM NaCl 0.1 mM EDTA pH 8 0.1 mM EGTA
Nuclei Lysis Buffer	20 mM Tris pH 7.9 400 mM NaCl 1 mM EDTA pH 8 1 mM EGTA
IP Buffer	50 mM Tris pH 7.9 150 mM NaCl 1 mM EDTA pH 8 1 mM EGTA

ChIP Protocol A:

Lysis Buffer	25mM HEPES pH 7.9 1.5 mM MgCl2 10mM KCl 0.1%(v/v) NP-40
IP Buffer	50mM HEPES pH 7.9 140mM NaCl 1mM EDTA 1%(v/v) Triton X-100 0.1%(w/v) Na-Doc 0.1%(w/v) SDS

Wash Buffer 1	50mM HEPES pH 7.9 500mM NaCl 1mM EDTA 1%(v/v) Triton X-100 0.1%(w/v) Na-Doc 0.1%(w/v) SDS
Wash Buffer 2	20mM Tris pH 8 1mM EDTA 250mM LiCl 0.5%(v/v) NP-40 0.5%(w/v) Na-Doc
TE Buffer	10mM Tris pH 8 1mM EDTA pH 8
Elution Buffer	50 mM Tris pH 8 1 mM EDTA pH 8 1% (w/v) SDS

ChIP Protocol B:

Lysis Buffer 1	50 mM HEPES pH 7.5 140 mM NaCl 1 mM EDTA 10% (v/v) glycerol 0.5% (v/v) NP-40 0.25% (v/v) Triton X-100
Lysis Buffer 2	200 mM NaCl 1 mM EDTA 0.5 mM EGTA 10 mM Tris pH8
Lysis Buffer 3	50 mM Tris pH 8

	10 mM EDTA 150mM NaCl 0.1 % (w/v) SDS 0.95 % (v/v) NP40 0.1% (w/v) Na-Doc
Wash Buffer 1	150mM NaCl 20mM Tris pH 8 2mM EDTA 0.1% (w/v) SDS 1% (v/v) TritonX-100
Wash Buffer 2	2mM EDTA 20mM Tris pH8 500mM NaCl 0.1% (w/v) SDS 1% (v/v) TritonX-100
Wash Buffer 3	0.25M LiCl 1% (v/v) NP-40 1 % (w/v) NaDoc 1mM EDTA 10mM Tris pH8
TE Buffer	50mM Tris-HCL pH8 1mM EDTA pH8 50mM NaCl
Elution Buffer	50mM Tris-HCL pH8 1mM EDTA pH8 1% (w/v) SDS

ChIP Protocol C:

Lysis Buffer	100mM NaCl

	50mM Tris-Cl pH8.1 5mM EDTA pH8.0 0.02% (v/v) NaN3 0.5% (w/v) SDS
IP Buffer	1V Lysis Buffer + 0.5V Dilution Buffer
Dilution Buffer	100mM Tris-HCl pH8.6 100mM NaCl 5mM EDTA pH8.0 0.2% (w/v) NaN3 5% (v/v) Triton X-100
Wash Buffer 1	150mM NaCl 20mM Tris-HCl pH8.1 5mM EDTA pH8.0 5.2% (w/v) Sucrose 0.02% (w/v) NaN3 1% (v/v) Triton X-100 0.2% (w/v) SDS
Wash Buffer 2	0.5% (w/v) DOC 1mM EDTA 250mM LiCl 10mM Tris-HCl pH8.0 0.02% (w/v) NaN3 0.5% (v/v) NP-40
Wash Buffer 3	0.1% (w/v) DOC 1mM EDTA 50mM HEPES pH7.5 500mM NaCl 1% (v/v) Triton X-100 0.02% (w/v) NaN3
TE Buffer	10mM Tris-HCl pH 8

	1mM EDTA
Elution Buffer	1% (w/v) SDS 0.1M $NaHCO_3$

All buffers were freshly supplemented with protease and phosphatase inhibitors.

2.1.10 Media

Bacterial media and agar

LB Agar	1% (w/v) Bactotryptone 0.5% (w/v) Yeast Extract 0.5% (w/v) NaCl 0.1% (w/v) Glucose 1.5% (w/v) Bactoagar
LB Broth	1% (w/v) Bactotryptone 0.5% (w/v) Yeast Extract 0.5% (w/v) NaCl 0.1% (w/v) Glucose
Low-Salt 2X LB Broth	2% (w/v) LB-Broth-Lennox 1% (w/v) Peptone 0.5% (w/v) Yeast Extract
SOC Media (Invitrogen)	2% (w/v) Bactotryptone 0.5% (w/v) Yeast 10mM NaCl 2.5mM KCl 10mM $MgCl_2$ 10mM $MgSO_4$ 20mM Glucose

Mammalian tissue culture media

Embryonic stem cell media:

JM8 Media	Knock-out DMEM supplemented with: 10% (v/v) Batch tested FBS 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 25U/mL Penicillin 25µg/mL Streptomycin Optimized dilution of LIF
E14 Media	GMEM- BHK21 supplemented with: 10% (v/v) Batch tested FBS 0.1mM 2-Mercaptoethanol 0.24% (w/v) Sodium Bicarbonate 1mM Sodium Pyruvate 0.1mM Non-essential Amino Acids 2mM L-Glutamine 25U/mL Penicillin 25µg/mL Streptomycin Optimized dilution of LIF
2i/LIF Media	1:1 mixture of Neurobasal and DMEM-F12 media supplemented with: 0.5% (v/v) N-2 supplement 1% (v/v) B-27 supplement 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 50U/mL Penicillin 50µg/mL Streptomycin Optimized dilution of LIF Freshly supplemented with 3µM CHIR99021 and 1µM PD0325901.

Neuronal differentiation media:

NDM Rich Media	1:1 mixture of Neurobasal and DMEM-F12 media supplemented with: 1% (v/v) N-2 supplement 2% (v/v) B-27 supplement 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 50U/mL Penicillin 50µg/mL Streptomycin

Epiblast stem cell media:

N2B27 Media	1:1 mixture of Neurobasal and DMEM-F12 media supplemented with: 0.5% (v/v) N-2 supplement 1% (v/v) B-27 minus supplement vitamin A 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 50U/mL Penicillin 50µg/mL Streptomycin Freshly supplemented with 20ng/ml Activin A and 12ng/ml bFGF
Rich Media	1:1 mixture of Neurobasal and DMEM-F12 media supplemented with: 1% (v/v) N-2 supplement 2% (v/v) B-27 minus supplement vitamin A 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 50U/mL Penicillin 50µg/mL Streptomycin Freshly supplemented with 20ng/ml Activin A and 12ng/ml bFGF

CDM-PVA Media	1:1 mixture of IMDM and Ham's F12 medium supplemented with: 1mg/ml Poly(vinyl alcohol) 1% (v/v) Chemically Defined Lipid Concentrate 15µg/ml transferrin 450µM Monothioglycerol 7µg/ml Insulin Freshly supplemented with 20ng/ml Activin A and 12ng/ml bFGF
CDM-BSA Media	1:1 mixture of IMDM and Ham's F12 medium supplemented with: 5mg/ml BSA 1% (v/v) Chemically Defined Lipid Concentrate 15µg/ml transferrin 450µM Monothioglycerol 7µg/ml Insulin Freshly supplemented with 20ng/ml Activin A and 12ng/ml bFGF
Coating Media	DMEM supplemented with: 10% (v/v) FBS 25U/mL Penicillin 25µg/mL Streptomycin

Mesoderm differentiation media:

FlyB	CDM-PVA Media freshly supplemented with: 20ng/µl bFgf 10nM LY294002 10ng/ml BMP4
FB40	CDM-PVA Media freshly supplemented with: 20ng/µl bFgf 50ng/ml BMP4

FlyWLDN	CDM-PVA Media freshly supplemented with: 20ng/ul bFgf 2nM LY294002 2ng/ml Wnt3a 250nM LDN

XEN cell media:

Standard Media	RPMI supplemented with: 15% (v/v) FBS 0.1mM 2-Mercaptoethanol 1mM Sodium Pyruvate 2mM L-Glutamine 25U/ml Penicillin 25µg/ml Streptomycin
Derivation media	Advanced RPMI supplemented with: 15% (v/v) Biosera FBS 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 25U/ml Penicillin 25µg/mL Streptomycin Freshly supplemented with 10ng/mL Activin A, 0.01µM atRA, 24 ng/mL bFGF, 1 µg/mL Heparin

MEF/COS-7 media:

	DMEM supplemented with: 10% (v/v) FBS 0.1mM 2-Mercaptoethanol 2mM L-Glutamine 25U/mL Penicillin 25µg/mL Streptomycin

All reagents used in mammalian cell culture media, unless previously stated, were purchased from Lifetechnologies (Gibco).

2.1.11 Strains and cell lines

Bacterial strains	Genotype	Information
TAM1	*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 (*ara-leu*)7697. *gal*U *gal*K *rps*L *end*A1 *nup*G	Chemically competent Echerichia Coli.

Mammalian Cell lines	Parental Line	Genetic Background	Information
E14	E14Tg2A	129/Ola	Feeder-independent *Hprt*-deficient mouse ES cell line derived from Lesch-Nyhan embryos (Morita et al. 2000).
WT	JM8.N4	C57BL/6N	Feeder-independent wildtype mouse ES cell line derived from the subclone N4 of the JM8 parental line (Skarnes et al. 2011)
E2$^{-/-}$	JM8.N4	C57BL/6N	Feeder-independent Jmjd2c homozygous knock-out mouse ES cell clone derived from the WT JM8.N4 (Skarnes et al. 2011)
E3$^{-/-}$	JM8.N4	C57BL/6N	Feeder-independent Jmjd2c homozygous knock-out mouse ES cell clone derived from the WT JM8.N4 (Skarnes et al. 2011)
XEN	IM8A1	PO	Feeder-independent extra-embryonic (XEN) stem cells hemizygous for the ROSA26 β-geo transgene (Kunath et al. 2005)
MEF	-	CF-1	Primary culture of mouse embryonic fibroblasts derived from CF-1 mice.
COS-7	**COS-7**	CV-1	Derived from monkey kidney tissue cell line CV-1 by transformation with an origin defective mutant of SV40 which codes for wild type T antigen (Gluzman, 1981)
HeLa	**-**	Human	Immortalized cell line derived from cervical cancer (Schercher et al., 1953)

II.2 Methods

2.2.1 Cell culture

Unless stated, all cell lines used in this work were previously derived from mouse embryos and were maintained in culture at 37°C and 5% CO_2 in their respective media as described in I.1. In general, all stem cell lines were not used beyond 45 passages. Cells and media were manipulated in a laminar flow cabinet under aseptic conditions, except for isolation of cell colonies. All cell lines used in experiments were routinely tested negative for the presence of Mycoplasma infections.

2.2.1.1 Embryonic stem cell culture

The wild-type (WT) JM8.N4 and the two *Jmjd2c*-knockout (*Jmjd2c*-KO) embryonic stem cell clones (E2, E3) derived from the same parental line were routinely cultured in feeder-free conditions, in plates pre-coated with 0.1% gelatine and in FBS-containing medium (JM8 media). Geneticin (100µg/ml) and Hygromycin B (25µg/ml) were routinely added to the culture medium in the Jmjd2c-KO cell lines for selection. Drug selection was removed prior to any differentiation experiments or lysate collection. The E14Tg2A (E14) embryonic cell line was also routinely cultured in feeder-free conditions, in plates pre-coated with 0.1% gelatin and in FBS-containing medium (E14 media). Genetically modified cell lines derived from any of these ESC lines were maintained in culture in the same conditions as the parental cell line, with the presence of Puromycin (1µg/mL) for selection.

Where mentioned, ESC cultures were adapted to serum-free conditions and cultured in 2i/LIF media. Cells were seeded in their respective FBS-containing medium in plates pre-coated with 0.2% gelatin and medium was changed to 2i/LIF medium after 24 hours of seeding. Cells were expanded in these conditions for at least 7 days prior to any experiment or lysate collection.

For passaging, ESCs were dissociated into single cells with Trypsin for 3 minutes at 37°C, after being washed with DPBS. Trypsin was inactivated by dilution and cells suspensions were collected and centrifuged for 5 minutes at 1000rpm (FBS-containing media) or 1200rpm (2i/LIF

media). Cells were seeded in a 1:5 – 1:10 ratio every other day or every 3 days. The media was replenished every day.

2.2.1.2 Epiblast cell culture

The parental epiblast cell line EpiSC^{Oct4GIP} was routinely cultured in Epiblast N2B27 media with Puromycin (1µg/mL), whereas a second parental epiblast cell line was routinely cultured in Epiblast CDM-BSA media. Both cell lines were cultured in plates pre-coated with 0.1% gelatin and Epiblast Coating Media. For passaging, cells were washed and incubated with DPBS for 2 minutes at 37°C. The detached clumps of cells were collected and centrifuged for 5 minutes at 1000rpm. Clumps were gently dissociated to into smaller clumps by pipetting and seeded in a 1:5 – 1:10 ratio every 3 days. The media was replenished every other day.

2.2.1.3 XEN cell culture and differentiation

The IM8A1 extra-embryonic (XEN) cell line was routinely cultured in XEN media in the absence of coating. For passaging, cells were dissociated through the same process as embryonic stem cells and seeded in a 1:20 ratio every 3 days. Media was replenished every other day.

For visceral endoderm differentiation, XEN cells were seeded in a 1:50 ratio and cultured in the presence of BMP4 (10ng/ml) for 4 days.

2.2.1.4 Derivation of mouse embryonic fibroblast

Primary mouse embryonic fibroblasts (MEFs) were derived from E13.5 embryos. Once the embryos were removed from the uteri of pregnant mice, the head and internal organs were removed, and the remaining tissue incubated with trypsin for 10 minutes at 37°C. The tissue was then dissociated by passing through 18G needles 5 times then incubated for a further 5 min at 37°C. Trypsin was inactivated by dilution with MEF media and cells were centrifuged at 1000rpm for 5 minutes. Cell pellets were resuspended in MEF media (10ml per embryo) and plated in

100mm dishes (1 plate per embryo). Adherent fibroblasts were cultured until confluent and frozen until further use.

For mitotic inactivation, MEFs were expanded up to 3 passages and dissociated with Trypsin when 80% confluent. Cell suspensions were exposed to 40Gγ of γ-radiation. Irradiated MEFs (iMEFs) were used as feeders in specific experiments.

2.2.1.5 Cryopreservation

All cell lines were routinely frozen in their respective media supplemented with 10% DMSO and 50% FBS (only for cells growing in FBS-containing media) and stored in cryogenic tubes which were transferred into a Mr Frosty Freezing Container and kept at -80°C for up to a month before moved indefinitely into N_2 liquid. All cell lines were thaw by placing the cryogenic tube in a 37°C water bath for 2 minutes, and subsequent dilution of the cell suspension by 10 times with plain media. Cells were centrifuged for 5 minutes at 950rpm and seeded in their respective conditions. Media was replenished the next morning to eliminate dead cells.

2.2.1.6 Production of Leukaemia Inhibitory Factor

The Leukaemia Inhibitory Factor (LIF), a key component of mouse ES cell media, was prepared in-house. The monkey kidney cell line COS-7 was used as a host for the production of this component. COS-7 cells were cultured in COS-7 media for one passage, after which cells were expanded in 20x10cm plates in 10ml of E14 media in the absence of LIF. When cells reached 70% confluency 24μg of LIF-expressing vector were delivered per plate following the procedures described in 2.3.3. After 24h the media of each plate was collected and replenished with fresh 10ml of E14 media (minus LIF) per plate. After 24h the media was again collected and the 2 batches mixed together, filtered, aliquoted and stored at -20°C. The optimal dilution of each batch produced was determined by titration and comparison with the previous batch and a commercially available batch of LIF. Cells plated at low density were cultured in the presence of

the three batches of LIF and their undifferentiated state evaluated following Alkaline Phosphatase Staining (Section 2.2.2.2).

2.2.2 Phenotypic characterization of embryonic stem cells

2.2.2.1 Growth curve

To access the growth rate of ESC lines, 500,000 cells (per well) were plated in FBS-containing medium supplemented with LIF, in the absence of drug selection (Jmjd2c-KO ESCs), in a 6-well pre-coated with 0.1%gelatin. After 48h the cells were dissociated to single cell level with Trypsin and the total number of cells counted using a haemocytometer. Subsequently, the initial number of cells (500.000) was re-plated in a new well. The same procedure was repeated over 8 passages and the media replenished in between. The total number of cells at each passage was determined by multiplying by the previous passage.

2.2.2.2 Self-renewal assay

To access the self-renewal ability of an ESC population, 1000 cells (per well) were seeded in 6-well plates, in duplicate. The medium (plus LIF and minus LIF; without drug selection) was changed every day and after five days the colonies were assessed for alkaline phosphatase activity following the manufacturer's instruction. Briefly, the cells were washed with DPBS and fixed for 30 seconds with 600µl of AP fixation solution, washed with distilled water (dH_2O) for 45 sec and stained with 600µl of AP staining solution for up to 20 min in the dark. Finally, stained cells were washed with dH_2O and allowed to dry at room temperature. Colony morphology was scored as differentiated (<20% stained), mixed (>20%-90% stained), undifferentiated (>90% stained).

2.2.3 Differentiation of embryonic stem cells

2.2.3.1 Embryoid body formation

JM8 and E14 ES cells growing routinely in FBS-containing media were dissociated with Trypsin and 100,000 cells (per well) were thoroughly re-suspended in 4 ml (per well) of JM8/E14 media with 5% FBS and no LIF and transferred into a 6-well ultra-low attachment plate (Corning, USA). Half of the media was replenished every day. EB aggregates were collected at the indicated time points for RNA extraction.

For hanging drop EB (HD-EB) assay, approximately 500 cells were used to form single drops of 30µl which were deposited in an inverted lid of a 10cm plate. Nearly 30 drops were added in each lid, and 3 to 4 plates were used per time point of RNA collection. The bottom of the plate was filled with about 10ml of DPBS to maintain humidity levels within the plate. Lids were swiftly turned and placed on the plates, allowing the individual cells to aggregate at the bottom of the drops.

2.2.3.2 Retinoic acid treatment

For all trans retinoic-acid (atRA) induced differentiation, JM8 ESCs growing routinely in JM8 media were dissociated and 80,000 cells (per well) seeded in 12-well plates in routine conditions for 24h. Media was then changed into JM8 media with 5% FBS with 1µM of atRA and no LIF to promote differentiation, and replenished every day for up to 4 days. Cells were collected at the indicated time point for RNA extraction. For staining, 20,000 cells (per well) were plated in 8-well chamber wells, and treated for 4 days under the same conditions mentioned above.

2.2.3.3 XEN cell conversion

JM8 ESCs growing routinely in FBS-containing were dissociated with Trypsin and resuspended in XEN media. Around 30,000 cells (per well) were plated in a 6-well plate pre-coated with 0.1% gelatin. After 24h the media was replaced with XEN derivation media (containing atRA/Activin A/bFGF/Heparin).

The derivation media was replenished again, allowing a total of 48h of derivation, after which the differentiated cells were dissociated with Trypsin and re-plated again in derivation media into a 6-well plate previously coated with iMEFs. The next day, the medium was changed into standard XEN media supplemented with bFGF (24ng/mL) and Heparin (1µg/mL). Approximately 10 days after the beginning of the experiment, converted XEN (cXEN) colonies were picked under the microscope by scraping the boundaries of the colonies with an 18G needle. Detached colonies were collected and dissociated by pipetting several times with a P1000. Colonies were seeded in XEN media supplemented with bFGF and Heparin in a 24-well plate pre-coated with iMEFs. From here on, the cXEN populations were split when confluent into 6-well plates pre-coated with 0.1% gelatin until all the cell population resembled XEN cells. RNA was extracted from 3 different passages of established cXEN cell lines.

2.2.3.4 Epiblast cell conversion

ESCs growing routinely in serum-free conditions (2i/LIF) for at least 7 days dissociated with Trypsin and 50,000 cells (per well) were plated in 6-well plates previously coated with Fibronectin (10ug/ml) for at least 1 hour, in 2i/LIF media. After 24h the following procedures were taken:

- For short Epiblast-like (EpiLC) conversion, the media was changed to Epiblast N2B27 Media supplemented with 1%KOSR (Buecker et al, 2014). RNA, protein and chromatin were extracted after 24h and/or 48h.
- For transient EpiLC conversion, the media was changed to Epiblast Rich Media supplemented with Activin A (20ng/ml) and bFGF (12ng/ml) and replenished every day. RNA was extracted every day up to day 3.
- For stable EpiSC conversion, the cells were cultured in Epiblast Rich Media for 7 to 10 days days, after which cells were detached by incubation on DPBS for 5 minutes at 37°C and gentle scraping with a tip. Cell clumps were plated in a 1:3 ratio in Fibronectin-coated wells with Epiblast Rich Media for up to 5 passages. From here on, cells were passaged by

incubation on DPBS for 2 minutes at 37°C and gentle tapping on the plate to specifically detach EpiS-like colonies. Converted EpiSC (cEpiSC) cultures were routinely seeded in Epiblast N2B27 Media in a 1:5 – 1:10 ratios in 6-well plates previously coated with Epiblast Coating Media. After at least 12 passages in these conditions cells were used for subsequent experiments. RNA, protein and chromatin were extracted from at least 3 different passages in stable conditions.

2.2.3.5 Neuronal differentiation

ES cells growing routinely in serum-free conditions (2i/LIF) for at least 7 days dissociated with Trypsin and 50.000 cells (per well) were plated in 6-well plates previously coated with 0.2% gelatin, in 2i/LIF media. After 24h the media was changed to Neuronal Rich Media, which was replenished every day up to day 5. At day 5, differentiated cells were fixed in the plate and stained as described in Section 2.2.7.1.

2.2.3.6 Mesoderm differentiation

Converted EpiSCs routinely cultured in Epiblast N2B27 media, and the parental EpiSC line routinely cultured in Epiblast CDM-BSA media were dissociated by incubation on DPBS, and seeded in their respective media at a 1:6 ratio in a 6-well plate previously coated with Epiblast Coating media. After 24h the media was changed into FlyB media. After 36h in FlyB conditions, the media was changed into either FB40 or FlyWLDN, for lateral plate mesoderm and paraxial mesoderm differentiation, respectively. The media was replenished every day up to 4 days in both conditions, after which the cells were harvested for FACS analysis (section 2.2.7.2), or fixed and stained (section 2.2.7.1). RNA was extracted at the indicated time points.

2.2.4 Cloning and DNA delivery

2.2.4.1 Generation of plasmids

For the generation of the vectors pPyCAG-IP-Jmjd2c and pPyCAG-IP-FLAG(x2)V5-Jmjd2c, the Jmjd2c full-length transcript was amplified from the cDNA of WT ESCs with primers containing restriction sequences recognised by the enzymes containing a restriction site in each of the vectors used (Primer sequences in Appendix B). In particular, restriction sites for XhoI and PacI were inserted in the 5' and 3' sites of Jmjd2c by PCR, and used to ligate this insert into the pPyCAG-IP vector. Similarly, the restriction sites for BamHI and SalI were inserted in the 5' and 3' sites of Jmjd2c by PCR, and used to ligate this insert into the pPyCAG-IP-FLAG(x2)V5 vector. For amplification of Jmjd2c with the indicated restriction sites, a PCR reaction was prepared with the following components: 1X KOD buffer, 1mM $MgSO_2$, 0.2mM dNTP mix, 0.3µM each primer, 1U KOD Hotstart Polymerase and approximately 5ng template cDNA. The following cycling conditions were used: 95°C for 2 minutes (denaturation), 35 cycles of 95°C for 20 seconds (denaturation) and 58°C for 25 seconds (annealing), and finally 70°C for 60 seconds (extension); followed by a final elongation step of 70°C for 10 minutes. A sample of the amplified product were separated in a 1% (w/v) agarose gel containing 0.01 g/mL Ethidium Bromide by electrophoresis 1X TBE buffer, over 1h at 100V.

Insert and vector DNA were digested with the respective restriction endonucleases. Reactions were performed in a total volume of 50µL containing no more than 10% (v/v) of enzyme, and 1x of the recommended NEB buffer (10x), with addition of 1x BSA. Digestions were performed at 37°C for 2-3 hours followed by an inactivation step of 65°C, for 20 min. Digested PCR products were purified using the QIAquick PCR Purification Kit. Digested products were separated by electrophoresis, excised from the gel and purified with QIAquick PCR Gel Extraction Kit.

Inserts were ligated into the corresponding vectors with the NEB Quick Ligation Kit. 50ng linearized vector DNA was incubated with 1uL Quick T4 DNA Ligase, 1X Quick Ligation Buffer and a 3-fold molar excess of insert DNA, and in a total volume of 20µL. Reactions were performed for

5 minutes at room temperature then incubated on ice for 5 minutes more. Prior to plasmid amplification, vectors were purified with the QIAquick PCR Purification kit.

2.2.4.2 Plasmid amplification

The following procedure was used for amplification of any of the plasmids used in this work. Per reaction, 50µL of chemically competent TAM1 cells were thawed on ice briefly. Between 1 to 5µL of DNA ligation, plasmid, or shRNA vector, were pipetted into each vial and mixed gently, then incubated on ice for 30 minutes. DNA was inserted by heat-shock, upon incubation of bacteria at 42°C for 30 seconds. Transformed bacteria were incubated again on ice for 2 minutes, after which 250µL of SOC media were added per vial, and incubated for 1 hour at 37°C with rotation. Bacteria were spread onto pre-warmed LB-agar plates containing 50µg/mL Ampicillin and plates were incubated overnight at 37°C. Newly formed single colonies were picked and further expanded in liquid cultures in 2X-low salt LB-broth (shRNA pLKO.1 plasmids) or LB-broth (all other plasmids) containing 100µg/mL Ampicillin, overnight at 37°C with agitation. Transformed bacteria cultures were pelleted by centrifugation at 4,500rpm for 15 min, and DNA was collected and purified using the HiSpeed Plasmid Maxi Kit, and eluted in 500-1000µL sterile TE buffer and stored at 4°C.

2.2.4.3 Transfection of embryonic stem cells

All transfected ESC lines were routinely cultured in FBS-containing media. When necessary, cell density was optimized for each cell line, type of transfection and surface area. The DNA/Lipofectamine2000 ratio was kept constant but absolute quantities adjusted according to the surface area.

Transient transfections were used to validate newly generated expression vectors and to validate the efficiency of shRNA vectors, whereas stable transfections were used to generate new cell lines and perform long-term knockdown experiments.

- For transient transfections: 100.000 to 200.000cells/cm^2 were seeded in the routinely used media depleted of antibiotics (penicillin/streptomycin). Between 4 to 6 hours after plating, the transfection mix was prepared as follows. In general, per well of a 6-well plate, 1µg of plasmid DNA was diluted in 200µl of OPTIMEM and 3µl of Lipofectamine2000 in 200µl OPTIMEM. The 2 solutions were mixed together in a ratio of 1:1 and incubated at room temperature for 20 minutes. Transfection mix was added to each well drop-wise, and homogenised by gentle movement of the plate. The next day, the media was changed into routine ESC media (with antibiotics), and cells lysates collected 48h post-transfection for RNA or protein extraction.

- For stable transfection: 50.000 to 100.000cells/cm^2 were seeded and transfected as described above. Additionally, 24h post-transfection 1µg/ml Puromycin was added for selection and these culture conditions maintained for 4 days (for knockdown experiments) or indefinitely (for stable cells lines). In the latter, individual cell clones were isolated from the pool of transfected cells 5 to 10 days post-transfection, by scraping and aspirating each clone with a pipette. Individual clones were transferred into 96-well plates and expanded for validation and further use.

2.2.5 Quantitative reverse transcriptase PCR

2.2.5.1 RNA extraction

RNA extraction was performed using RNeasy Mini kit following the manufacturer's instructions. For RNA extraction of adherent cultures, 350ul of RLT buffer supplemented with 1% (v/v) 2-Mercaptoethanol, were added to each well of cells washed with DPBS and the lysates collected and frozen at -80°C until extraction. For RNA extraction of suspension cultures (EBs), cells were collected and pelleted by centrifugation, before the same procedure above was applied.

In brief, compete lysis was achieved by pipetting and vortexing the cell suspension in RLT. The lysates were mixed with 1 volume of 70%(v/v) ethanol and transferred into a MinElute spin column. Columns were centrifuged at 10.000rpm for 1 minute and flow-through discarded.

Columns were washed with 350µl of Buffer BW1 and contaminant DNA was removed by incubation with 10µl DNAse I diluted in 70µl of buffer RDD, at room temperature for 20 minutes. After a second wash with 350µl of Buffer BW1, columns were washed twice with 500µl Buffer RPE. RNA was eluted from the columns with 30µl RNase-free water. Concentration was measured in a NanoDrop 2000 spectrophotometer and RNA samples stored at -80°C.

2.2.5.2 cDNA synthesis

Between 0.5 and 1µg of total RNA were reverse transcribed using Superscript First-Strand Synthesis kit. RNA was mixed with 1µl of 10 mM dNTP mix and 1 µl of oligo (dT)12-18 and diluted in RNase-free water to a final volume of 13 µl and incubated at 65°C for 5 minutes. Subsequently, 2 µl of 0.1 M DTT, 4µl of 5X first strand buffer, and 1 µl of RnaseOUT were added and the mixture at 42°C. After 1 minute 1µl of 200U/µl Superscript II was added to each reaction and incubated at 42°C for 1hour and at 70°C for 15 min. Each newly synthesised cDNA sample was diluted 20 times with RNase-free and stored a -20°C.

2.2.5.3 Quantitative real-time PCR analysis

Quantitative real-time polymerase chain reactions (qPCR) were carried out in either a total volume of 20µl with 50% (v/v) Jumpstart Sybr-Green PCR Mastermix and 1% ROX, or in a total volume of 10µl with 50% (v/v) KicQstart Sybr-Green PCR Mastermix (incuding ROX). To each reaction 300nM of primers and 5% (v/v) of cDNA/DNA were added. A control reaction without cDNA/DNA was included to check for the formation of primer-dimers, and all reactions were performed in duplicate in the same plate. Cycling conditions were as follows: 95°C for 15 minutes (denaturation), 40 cycles of 94°C for 15seconds (denaturation) and 60°C for 30seconds (annealing), and finally 72°C for 30 seconds (extension).

The quantification of the abundance of target sequences was determined by using the ΔC_T and $\Delta\Delta C_T$ method (Livak and Schmittgen, 2001; Pfaffl, 2001) in which the C_T represents the number of cycles in the reaction necessary to achieve a defined level of fluorescence in the

exponential phase of amplification. For gene expression analysis, this was either normalized to the C_T average of 2 housekeeping genes (*S17* and *L19*) and represented as mRNA levels, or normalized to both the housekeeping genes and a control sample and represented as fold enrichment over control.

For detection of protein binding by chromatin immunoprecipitation, the C_T of the target region was normalized to 1% of input DNA and to a background (typically intergenic) region and represented as fold enrichment over background.

2.2.5.4. Primer design and validation

Primers for RT-qPCR were design with the following criteria: amplicon size of 100-200bp, primer GC content of 45-55%, primer melting temperature 57-63°C , low 3' stability (no more than 2 C'2 and/or G's in the last 5bp of the primer) and primer size of 18-25bp. Moreover, for expression analysis primers were designed in different exons. Target sequences were obtained in the Ensembl and the UCSB genome browsers and the primers were aligned to the genome using the BLAT function of the NCBI.

For each pair of primers designed, the efficiency in qRT-PCR was tested in 2-fold dilutions of DNA/cDNA (Pfaffl, 2001). Primers were considered efficient, and therefore suitable for the $\Delta\Delta C_T$ method, when yielding a good linear fit (R^2<0.98) in Ct vs. logarithm of the dilution factor plot and the efficiency calculated was between 1.8 and 2.2 (Efficiency=$10^{(1/slope)}$).

2.2.6 Analysis of DNA methylation

2.2.6.1 Genomic DNA extraction

Genomic DNA was extracted from adherent cells expanded in 10cm dishes. Cells were dissociated with Trypsin and washed with DPBS. Cell pellets were resuspended in 0.5-1ml DNA Lysis Buffer with 0.5mg/ml Proteinase K and incubate for 16h at 55°C. An equal volume of Phenol-chloroform was added to the sample, followed by vortexing and centrifugation for 5 minutes at 14000rpm. The organic (top) layer was collected and mixed with 50% volume of isopropanol.

Samples were incubated at -20°C for at least 2 hours, after which were centrifuged at 4°C for 20minutes at 14000rpm. DNA pellets were washed with 70% (v/v) ethanol and centrifuged again for 10 minutes. Pellets were allowed to dry at room temperature until all ethanol had evaporated and finally resuspended in 0.5 – 1ml of TE buffer with 20ug/ml RNAse A. Samples were stored at -20°C until further use.

2.2.6.2 Restriction analysis

The levels of DNA methylation and hydroxymethylation at target sequences were evaluated with the EpiMark 5-hmC and 5-mC Analysis Kit, following the manufacturer's instructions. First, 5hmC sites were glycosylated (5ghmC) by adding 15U of T4-β-glucosyltransferase, 80μM UDP-Glucose and 1%(v/v) NEBuffer 4 to 3μg of genomic DNA. The reactions were incubated at 37°C for 16h. One third of the sample was further incubated with 100U of MspI, one third with 50U of HpaII and another third with the equivalent volume of H_2O. The 3 reactions were incubated at 37°C for 6 hours, followed by the addition of 20μg of Proteinase K and incubation at 40°C for 30 minutes. The Proteinase K was finally inactivated at 95°C for 10 minutes and samples diluted 1:3 in nuclease-free H_2O for qPCR analysis. Primers for this assay were designed around target sequences containing one (or more) CCGG, which is recognized by both restriction enzymes. However, digestion by HpaII is blocked by any modification (5mC, 5hmC or 5-ghmC), whereas MspI is only blocked by 5ghmC. Therefore, HpaII is only able to digest unmethylated DNA, whereas MspI is only able to digest 5mC DNA. Amplification of HpaII-resistant products is proportional to the levels of 5mC plus 5ghmC DNA, whereas amplification of MspI-resistant products is proportional to the levels of 5ghmC (and thus 5hmC) DNA alone. Thus, the levels of total DNA were determined by the amplification levels in the un-digested fraction, the 5hmC levels corresponded to the levels quantified in MspI-resistant 5mC DNA, and the 5mC levels were determined by subtracting the level quantified in HpaII-resistant (5mC + 5ghmC) fraction by the levels of MspI-resistant (5hmC) DNA. All reactions were normalized to a control region (without a restriction site) and represented as percentage relative to total DNA (100%).

2.2.7 Protein analysis

2.2.7.1 Immunofluorescent staining

Cells were cultured on 0.2% gelatin-coated glass coverslips or plastic 8-well chambers. When at the desired confluency, cells were washed in PBS and fixed in 4% paraformaldehyde (PFA) for 10 minutes. Fixed cells were permeabilised with 0.4% Triton X-100 in blocking buffer (Section 2.1.9), for 30 minutes at room temperature. After washing with PBS, the coverslips were incubated in blocking buffer for 30 minutes. Finally, samples were incubated with primary antibody diluted in blocking buffer overnight at 4°C in a humid chamber. Coverslips were washed in PBS and incubated with the fluorophore-conjugated secondary antibodies diluted in blocking buffer for 1 hour at room temperature. Coverslips were washed in PBS and mounted in Vectashield with 1μg/ml DAPI. Samples were visualised using a SP5 Leica laser-scanning confocal microscope or an inverted fluorescent/brightfield microscope with an attached Nikon camera.

2.2.7.2 Flow cytometry analysis

Cells were harvested using cell dissociation buffer (Gibco) and 10.000 cells were blocked in PBS containing 2%(v/v) FBS, stained for Flk-1 with 5μg/mL of anti-Flk-1-Biotin (avas12 a1 clone, eBioscience), followed by 0.4 μg/mL APC-conjugated streptavidin (Biolegend) or Pdgfrα with 2μg/mL anti-Pdgfrα-PE (clone APA5, eBioscience) and analysed on an Accuri C6 Flow cytometer.

2.2.7.3 Protein western blot

Whole cell extracts were obtained by lysing the cells cultured in 6-well plates with cold RIPA buffer supplemented with protease and phosphatase inhibitors. Lysates were incubated for 20 minutes on a rotating wheel at 4°C, centrifuged at 14,000rpm for 20 minutes and the supernatant collected. The protein extracts were quantified by Bradford assay and stored at -20°C.

The appropriate amount of protein lysates (5-30μg for whole extracts) was diluted 3:1 with Laemmli sample buffer (4x) with 5%(v/v) β-mercaptoethanol and incubated at 95˚C for 5 minutes. Lysates were resolved on an 8% SDS polyacrylamide gel electrophoresis (SDS-PAGE) gel

and subsequently transferred into a methanol-activated Polyvinylidene fluoride (PDVF) membrane using wet-transfer method at 90V for 1 hour at 4˚C. Blots were blocked with 5% skimmed milk in Tris-Buffered Saline with 0.5% Tween20 (TBS-T) for 1 hour at room temperature. After washing with TBS-T, the appropriate amount of primary antibody was diluted in 5% skimmed milk in TBS-T and incubated overnight t 4˚C. Membranes were washed with TBS-T and then incubated with the appropriate HRP-conjugated secondary antibody diluted in 5% skimmed milk in TBS-T for 1 hour at room temperature. After washing with TBS-T, detection of the proteins was achieved by treating the blots with enhanced chemiluminescence (ECL) western blotting detection reagents and exposing in the ImagQuant LAS4000.

2.2.7.4 Histone western blot

Histones were acid extracted from adherent cells (Shechter et al., 2007), which were expanded in 10cm dishes. Cells were dissociated with Trypsin and centrifuged at 1200rpm for 4 minutes. After washing with DPBS, cell pellets were resuspended in 1ml of hypotonic lysis buffer and incubated on ice for 30 minutes. Lysates were centrifuged for 10,000rpm for 10 minutes at 4°C. The intact nuclei were pelleted by spinning in cooled tabletop centrifuge: 10,000rpm, 10 minutes at 4 °C. Pelleted nuclei were resuspended in 400 µl 0.4N H_2SO_4 and incubated on a wheel overnight. Nuclear debris were removed by centrifuging samples for 10 minutes at 14.000rpm. Histones were precipitated with 33%(v/v) TCA for 30 minutes on ice, followed by a centrifugation step at 14,000rpm for 10 minutes at 4°C. Pellets were washed with ice-cold acetone, and centrifuged again. After air drying, the pellets were resuspended in 100µl dH2O, quantified by Bradford assay and stored at -20°C.

1 or 2µg of histones were diluted 3:1 with Laemmli sample buffer (4x) with 5%(v/v) β-mercaptoethanol and incubated at 95˚C for 5 minutes. Histone lysates were resolved on a 13% SDS-PAGE gel and subsequently transferred into a methanol-activated PVDF using a semi-dry method at 20V for 1 hour at room temperature. The remaining steps were as described above (Section 2.2.7.3). Relative quantification of band intensity was performed using ImageJ software.

2.2.8 Protein-Protein interaction analysis

2.2.8.1 FLAG pull down and mass spectrometry

Cells were expanded into 25x 15cm plates, which were washed with PBS twice and scraped in ice-cold PBS supplemented with protease inhibitors. Cell suspensions were centrifuged at 1100rpm for 5minutes at 4˚C. Cell pellets were resuspended in 5 volumes of Buffer A and incubated on ice for 10 minutes. Lysates were centrifuged at 3000rpm for 10minutes and pellets resuspended in 2 volumes of Buffer A. Suspensions were subject to mechanical lysis with loose pestle in a dounce homogeneizer. After a centrifugation step at 3000rpm for 10 minutes, nuclear extracts were ressuspended in 1.5 pellet volume of Buffer C and subject to mechanical lysis with a tight pestle in a dounce homogeneizer, and further incubated at 4˚C for 30 minutes. Insoluble fractions were removed by centrifugation. Nuclear (soluble) fraction was dialyzed twice against 1L of Buffer D in a 7kDa dialysis bag (Thermoscientific). FLAG protein and bound complexes were purified by incubation with 60µl anti-FLAG M2 beads and 0.15U/µl Benzonase. Beads were washed 5 times in Buffer C100 and complexes eluted with 0.2 mg/ml FLAG tripeptide four times. Elution of the FLAG-tagged protein was validated by western blot. Positive fractions were pooled together and the proteins precipitated for 30 minutes on ice with 0.25%(v/v) of TCA and 0.1%(w/v) DOC. Protein pellets were washed with ice-cold acetone and allowed to dry at room temperature. Finally samples were dissolved in 50µl sample buffer and incubated at 95 ˚C for 10 minutes. Samples were run on a SDS-PAGE gel up to 2/3 of the size and the gel and stained with Colloidial Coomassie kit for 2 hours.

Mass spectrometric analysis was performed as van der Berg et al., 2010. Briefly, gel lanes were cut with an automatic gel slicer and reduced. Peptides were separated on a capillary system coupled to a mass spectrometer. A Mascot search algorithm was used for identification of peptides against the NCBI database and emPAI scores were attributed to each identified peptide. The Mascot score cut-off for a positive protein here used was 40, given the high variability between experiments. Moreover, only peptides with a score 3-fold higher in the test sample versus the control sample were retained as significant. The list of proteins obtained in the three

experiments was overlapped, and a new list was generated containing only proteins identified in 2 out 3, or 3 out of 3, experiments. For the purpose of this work, cellular location of the identified proteins was manually annotated and only nuclear proteins were considered as biologically significant.

2.2.8.2 Protein co-immunoprecipitation

Co-IP of nuclear extracts from HeLa cells was performed with Co-IP protocol A (Section 2.1.9), as described in Fritsch et al., 2010. Briefly, HeLa-S3 cell lines expressing a HA-FLAG-G9a were generated, together with control cell line expressing an empty vector. Cells were expanded, washed in DPBS, scraped and pelleted. Pellets were resuspended in the hypotonic buffer and disrupted with 20 strokes of a tight-fitting Dounce homogenizer. Subsequently, a sucrose, spermidine and spermine-containing buffer (Buffer SR) was added to limit nuclei leak. The cytosolic fraction was separated from nuclei (pellet) after centrifugation for 7 minutes at 9000rpm, at 4°C. The nuclear soluble fraction was obtained upon incubation of the pellet in a high salt buffer for 30 minutes at 4°C followed by centrifugation at 10,000 rpm. The pellet corresponded to chromatin fraction and the supernatant to the nuclear soluble fraction. The chromatin fraction was resuspended and digested with micrococcal nuclease (Mnase). Nuclear soluble and chromatin fractions were then centrifuged at 3,200 rpm for 1 h at 4 °C. G9a-bound complexes were purified by IP with anti-FLAG antibody bound beads overnight on a rotating wheel at 4°C. Complexes were eluted with a FLAG peptide and further purified on anti-HA antibody-conjugated agarose beads, and subsequently eluted with the HA peptide. Double-immunopurified complexes were resolved on 4-12% SDS-PAGE bis Tris acrylamide gradient gel and immunoblotted with the indicated antibodies.

Co-IPs in nuclear extracts of $ESC^{control}$ and $ESC^{FV\text{-}Jmjd2c\text{-}WT}$ were performed with the Co-IP Protocol B (Section 2.1.9). Cells were expanded, washed and scraped with ice-cold PBS supplemented with protease inhibitors. Cells were centrifuged at 1300rpm for 5minutes, at 4°C and pellets were snap-frozen in dry ice until extraction. For protein extraction, the pellets were 3

times the volume (V) of the pellet in Buffer A. NP-40 was added to a final concentration of 1% (v/v) and lysates homogenised by inversion. 8V/9 of Buffer SR were added and lysates again homogenised by inversion. The nuclei were pelleted at 2000x g for 5 minutes at 4°C and resuspended in 1V of Sucrose Buffer. High Salt Buffer was added to the nuclei pellet drop by drop and incubated on ice for 30 minutes. 1V/3 of Sucrose Buffer was added, together with 1mM Cacl2 and 0.00025U/µl of Mnase. The reaction was incubated at 37°C for 10 minutes, and stopped by the addition of 4mM EDTA. Nuclei suspensions were briefly sonicated in a bioruptor (Diagenode) and lysates were centrifuged at 86.600 RCF for 30 minutes at 4°C. The protein concentration of the nuclear extract (supernatant) was quantified by BCA assay. The required amount was pre-cleared with 10 µl protein G-coupled sepharose beads, previously washed twice with Washing Buffer, for 2h at 4°C on rotating wheel. The supernatant was recovered after spinning and 500 to 1000µg of nuclear extract were used per IP. After diluting in a total volume of 500µl with TEGN buffer, 2.5 to 5µg of the antibody of interest or 5µg of the corresponding normal IgG as negative control, were added and IP's were incubated overnight on a rotating wheel at 4°C. Lipids and debris were removed from the IPs after centrifugation and 7.5µl of protein G beads previously blocked with 0.3% (w/v) BSA and 0.6µg/µl of ssDNA, were added and further incubated for 2 hours. IPs were washed 5 times with Wash Buffer. The protein complexes were eluted from the beads by incubation at 95°C for 10 minutes in 7µl of Wash Buffer and 7µl of Laemmli Buffer with 5%(v/v) β-mercaptoethanol. Proteins were detected following western blot (Section 2.2.7.3).

Co-IP in nuclear extracts of $ESC^{control}$ shown in Fig.5.7C was performed with Co-IP Protocol C (2.1.9). Cells were washed and scraped with DPBS and protease inhibitors, and pelleted at1300rpm for 5 minutes at 4°C. Cell pellets were resuspend in Cell Lysis Buffer and incubate on ice for 10 to 25 minutes. Cells were further lysed with a dounce homogneizer with a loose pestle, 10% (v/v) NP-40 were added to a final concentration of 0.5% and cells vortexed for 5 seconds, and incubate on ice 3 minutes. Nulei were pelleted at 4000rpm for 5 minutes at 4°C and resuspend in Nuclei Lysis Buffer and incubated for 10 minutes on ice. Pellets were further homogenised with a tight pestle in a douncer homogeneiser. Samples were sonicated in a

Diagenode bioruptor for 10 minutes with a cycle of 30 seconds ON/30 seconds OFF for 10min at high intensity. Nuclear extracts were centrifuge at 14,000rpm for 20 minutes at 4°C and the supernatant kept. The appropriate amount was precleared with 10 µl of magnetic Dynabeads for 2h at 4°C on rotating wheel. About 1mg of protein was used per IP and diluted in IP Buffer with the indicated antibodies. IPs were performed overnight on rotating wheel at 4°C. About 20 µl of Dynabeads were used to capture the IP complexes, which were subsequently washed in IP Buffer and separated by Western Blot after being eluted from the beads by incubation at 95°C for 10 minutes in IP Buffer and Laemmli Buffer with 5% (v/v) β-mercaptoethanol.

2.2.9 Protein-DNA interaction analysis

2.2.9.1 Chromatin immunoprecipitation

In general, all chromatin immunoprecipitation (ChIP) experiments followed the procedure described below with minor changes between as outlined at the end of this section. Solutions used in each protocol (A, B or C) are described in Section 2.1.9. Cells were expanded in 1 to 3 175cm^2 flasks and cross-linked in either 2mM DSG in fixation mix for 45 minutes followed by 1% formaldehyde for 10 minutes at 37°C (double fixed), or uniquely in the later conditions (single fixed). The reaction was then quenched with 125mM glycine, and the cells washed in PBS and scraped from the plates. Cell lysis was performed with the indicated lysis buffers for each protocol. This procedure was accompanied by douce homogeneizer with a loose pestle when cell lysis appeared incomplete. The nuclei were pelleted by centrifugation at 1200rpm and resuspended in the respective IP buffers (Section 2.1.9) and sonicated in a Diagenode bioruptor. Sonication cycles were optimized for each cell line and protocol to produce fragments with sizes between 200 and 500bp. Following a rapid de-crosslinking at 65°C, the fragment size of sonicated DNA was confirmed by separating a small sample in a 1%(w/v) agarose gel containing 0.01 g/mL Ethidium Bromide by electrophoresis in 1X TBE buffer, over 1h30 at 90V. DNA was measured by spectrophotometry and/or protein measured by BCA assay. The amounts of DNA or protein indicated below was pre-cleared with sepharose beads or magnetic Dynabeads for 2h at 4°C on

rotating wheel. IPs were also performed on rotating wheel overnight at 4°C with the indicated amount of primary antibodies. When indicated, normal IgG of the same species as the primary antibody was used as negative control. The antibody-protein complexes were further incubated with the indicated amount of beads for 3h. The unbound chromatin was removed by sequential washes with the indicated wash buffers for each protocol, followed by a final wash with TE buffer. For magnetic Dynabeads, protein-DNA complexes were isolated by incubation on a magnetic rack for 1 minute, whereas for sepharose beads these were separated by centrifugation at 1300 rpm for 3 minutes. Chromatin was eluted from the beads with addition of 250µl of elution buffer supplemented with 0.04µg/µl RNase A at in a shaker at 65°C overnight. Samples were subsequently diluted 1:2 in TE Buffer and treated with 0.2µg/µl Proteinase K at 55°C for 2 hours. DNA was purified by addition of equal volume of phenol/chloroform extraction. DNA was then precipitated overnight at -20°C with 20-30µg GlycoBlue carrier, 1/10 volumes of of NaAc 3M and 2 volumes of 100% ethanol. Pellets were centrifuged, washed in 70%(v/v) ethanol and finally resuspended in TE buffer and used for qPCR (Section 2.2.5.3) with primers listed in Appendix B, or for generation of libraries for sequencing.

Target protein	**ChIP Protocol**	**Fixation**	**DNA amount**	**Protein amount**	**Antibody amount**	**Volume of Beads**	**Beads coating**	**Type of Beads**
Suz12 Ring1b IgG	A	Single	600µg 600µg 600µg	-	10µg 10µg 10µg	100µl 100µl 100µl	Protein A	Sepharose
H3K9me2	B	Double	20µg	-	2µg	20µl	Protein G	Magnetic
Oct4 H3K27ac p300 FLAG Med1	C	Single	-	400µg 500µg 500µg 500µg 800µg	5µg 5µg 5µg 5µg 8µg	20µl 40µl 40µl 40µl 40µl	Protein G	Magnetic
G9a Smc1a	C	Double	-	250µg 800µg	5µg 8µg	40µl 40µl	Protein G	Magnetic

2.2.9.2 Generation of libraries for sequencing

Libraries were prepared from ChIP DNA using the NEBNext® Ultra™ DNA Library Prep Kit and Multiplex Oligos. Briefly, 5ng of eluted DNA were incubated with 3 µl End Prep Enzyme Mix and 6.5 µl of End Repair Reaction Buffer, in a 65 µl reaction volume. Reactions were incubated at 20°C for 30 minutes, and to 65°C for 30 minutes, for dA tailing and repair of the 3' end. Subsequently, a sequence adaptor was ligated by incubation with 15 µl Blunt/TA Ligase Master Mix, 1 µl of Ligation Enhancer and 2.5 µl of the NEBNext Adaptor (1.5µM) in a total reaction volume of 83.5 µl. Following a 15 minute incubation at 20°C for 15 minutes, 3 µl of USER Enzyme were added. Ligation reactions were incubate at 37°C for 15 minutes.

For selection of inserts containing 200bp size, 3.5 µl DEPC-treated dH2O were added to the ligation reaction for a 100 µl total volume, followed by 55 µl of AMPure XP Beads. Following a 10 minute incubation at room temperature, the unwanted large fragments bound to the Beads were removed by separation in a magnetic stand. Separated DNA was further incubated with 25 µl of AMPure XP Beads to the supernatant, for another 10 minutes at room temperature. Bound DNA was kept after magnetic separation and washed 3 times with 85% (v/v) ethanol. DNA was eluted with 25 µl of 10 mM Tris-HCl pH8, from which 20 µl were taken for further amplification by PCR with 1x NEBNext High Fidelity Master Mix (2X), 2.5 µl of each individual Index Primers and 2.5 µl of the Universal PCR Primer in a total volume of 50 µl. Combination of index primer in each lane are outlines below. PCR cycling were as follows: 98°C for 30 seconds, 15 cycles of 98°C for 10 seconds, 65°C for 30 seconds and 72°C for 30 seconds, and finally an extension step of 72°C for 5 minutes. PCR products were separated by incubation with 40 µl of AMPure XP Beads for 10 minutes at room temperature and washed twice with 85% (v/v) ethanol. DNA was eluted with 33 µl 10 mM Tris-HCl and size distribution was analysed in an Agilent Bioanalyzer and sequenced in an Illumina platform. A sample was taken for validation of a known target region by qPCR.

Condition	Sample	Rep	Index	Lane	Barcode
2i	input	rep1	4	A	TGACCA
2i	IP	rep1	5	A	ACAGTG
2i	IP	rep3	6	A	GCCAAT
FBS	input	rep1	7	A	CAGATC
FBS	IP	rep1	12	A	CTTGTA
EpiLC	input	rep1	4	B	TGACCA
EpiLC	IP	rep1	5	B	ACAGTG
EpiLC	IP	rep2+3	6	B	GCCAAT
FBS	input	rep3	7	B	CAGATC
FBS	IP	rep3	12	B	CTTGTA

2.2.9.3 Bioinformatic analysis of ChIP-seq datasets

Quality of the generated ChIP-seq reads was assessed using the FASTQC program (Babraham Bioinformatics). Reads were aligned to the UCSC mouse mm9 reference genomes using Bowtie2 (Langmead and Salzberg, 2012), and only reads with one match were retained. Aligned reads (BAM) were converted into SAM files, and further transformed into a BED format to subsequently generate bigwig files for visualization in the IGV browser (Robinson et al., 2011). Bigwig files of each pair or replicates were correlated with a wigcorrelate, which showed high correlation in all comparisons. Thus, individual BAM files for each replicate were merged with the second replicate and used in peak calling. List of enriched peaks were obtained using MACS2 with an FDR <0.0001 by calling peaks relative to the respective input control samples (background). Gene annotation and motif enrichment analysis were performed using HOMER (Heinz et al., 2010). Peak overlaps were identified with the R package ChiPpeakAnno (Zhu et al., 2010). Density plots and heatmaps were generated with deepTools (Ramirez et al., 2014). Snapshots were taken from the IGV (Robinson et al., 2011) and UCSC browsers (Kent et al., 2002).

The accession numbers of the published ChIP-seq datasets used in the generation of heatmaps, peak overlaps and snapshots are listed below. As indicted, treated files were obtained from the ESC ChIP-seq compendium (Martello et al., 2012), from the Gene Expression Omnibus (GEO) repository or the Cistome pipeline (Qin et al., 2012a; Sun et al., 2013). When files were

unavailable, bigwig or BED files were generated using the pipeline here described earlier, with the specified parameters indicates in each report.

Project	**Reference**	**Factor**	**Source of files**
GSE11724	(Marson et al., 2008)	Oct4	Compendium
GSE22562	(Kagey et al., 2010)	Med1	Compendium
GSE22562	(Kagey et al., 2010)	Smc1a	Compendium
GSE24165	(Creyghton et al., 2010)	p300	Compendium
GSE27827	(Lienert et al., 2010)	H3K4me2	GEO
GSE57092	(Bulut-Karslioglu et al., 2014)	Suv39h1	GEO
GSE48519	(Hon et al., 2014)	H3K4me1	GEO
GSE48519	(Hon et al., 2014)	H3K27ac	GEO
GSE48519	(Hon et al., 2014)	H3K4me3	GEO
GSE48519	(Hon et al., 2014)	H3K9me3	GEO
GSE43229	(Das et al., 2014)	H3K9me2	generated
GSE46536	(Mozzetta et al., 2012)	G9a	generated
GSE56138	(Buecker et al., 2014)	Oct4 - 2i	Cistrome
GSE56138	(Buecker et al., 2014)	H3K27ac - 2i	Cistrome
GSE56138	(Buecker et al., 2014)	H3K4me1 - 2i	Cistrome
GSE56138	(Buecker et al., 2014)	Oct4 - EpiLC	Cistrome
GSE56138	(Buecker et al., 2014)	H3K27ac – EpiLC	Cistrome
GSE56138	(Buecker et al., 2014)	H3K4me1 – EpiLC	Cistrome

Chapter III – Results: Assessing the impact of *Jmjd2c*-depletion in ESC self-renewal and pluripotency

III.1 Introduction

Jmjd2c has been identified as a member of the Jmjd2 family of histone demethylases specifically targeting the histone marks H3K9me2/3 and H3K36me3 (Katoh and Katoh, 2004; Klose et al., 2006b; Whetstine et al., 2006). *In silico* expression analysis of *Jmjd2* gene family members in early mouse embryogenesis revealed that *Jmjd2a*, *Jmjd2b* and *Jmjd2c* may be preferentially expressed in cleavage-stage embryos, with *Jmjd2c* being also highly expressed in undifferentiated ESCs (Katoh and Katoh, 2007; Wang et al., 2010). Jmjd2c was furthermore proposed to play a key role in maintaining ESC identity. Specific knock-down of *Jmjd2c* using shRNA vectors resulted in the rapid down-regulation of *Nanog* expression, and this effect was strictly correlated with an increase in H3K9me3 enrichment levels at Nanog gene promoter region (Loh et al., 2007). Double-stranded RNA-knockdown of *Jmjd2c* in oocytes and induction of parthogenesis resulted in early developmental arrest during the transition from 8-cell to morula stages (Wang et al., 2010). Additionally, this was accompanied by a reduction in the expression levels of *Oct4* and *Nanog* in 4-cell stage embryos, corroborating with Jmjd2c-dependent regulation of pluripotency-associated gene expression in ESCs (Wang et al., 2010). Altogether, these observations pointed to a critical role for Jmjd2c in the early developing embryo and in pluripotent stem cells; a role that remained to be fully explored at the start of this project.

In this chapter, the impact of *Jmjd2c* depletion on ESC self-renewal and differentiation potential was explored in-depth taking advantaged of newly generated *Jmjd2c*-knockout ESCs, which were obtained from a large-scale knockout (KO) consortia (Bradley et al., 2012; Skarnes et al., 2011). Here, the approaches used to generate these cells are described, as well as the steps taken towards the validation of their KO status. A detailed characterization of the phenotype and functional abilities of *Jmjd2c*-depleted ESCs was conducted using well-established methodologies. In particular, *Jmjd2c*-depleted ESC pluripotency –i.e their ability to differentiate into the three germ layers, was investigated upon formation of embryoid bodies using both *Jmjd2c*-KO and *Jmjd2c*-knockdown ESC clones.

III.2 Results

3.2.1 Generation of *Jmjd2c*-knockout ESC lines

ESCs lines carrying a single *Jmjd2c* targeted allele and their wild-type (WT) counterparts, both from C57BL/6 mouse genetic background, were obtained from the EUCOMM/IKMC repository pipeline through collaboration with Dr Cynthia Fisher and Dr William Skarnes (Wellcome Trust Sanger Institute, Cambridge) (Bradley et al., 2012; Skarnes et al., 2011). The KO first allele (tm1a) was generated by gene trapping through the insertion of a cassette encoding lacZ/β-galactosidase and a neomycin resistance gene, upstream of a "critical exon" common to all predicted *Jmjd2c* transcripts (Fig.3.1A). The generation of homozygous knockout cell lines was then carried out by Dr Cynthia Fisher (Wellcome Trust Sanger Institute, Cambridge) through homologous recombination of a gene-trap cassette into the second allele (tm2), encoding the hygromycin resistance gene and GFP (Fig.3.1.A). Both cassettes are promoterless, hence relying on the active transcription of *Jmjd2c*, and contain an En2 splice acceptor (SA) signal and the SV40 polyadenylation (pA) sequences to ensure transcription is spliced into the cassette and is stopped at the inserted sites (Fig.3.1A).

Validation of the knockout was performed by Dr Cynthia Fisher (Wellcome Trust Sanger Institute, Cambridge) at the DNA level using a Long-Range PCR approach with specific primer pairs (Fig.3.1A; Table B-I – Appendix B), hence showing correct insertion of the trapping cassettes in both alleles. A 5kb and a 6.3kb fragment corresponding to the first and second mutated alleles, respectively, were successfully amplified using genomic DNA extracted from the homozygous *Jmjd2c*-KO ESC ($ESC^{Jmjd2c\text{-}KO}$) clones E2 and E3, as opposed to WT ESCs (ESC^{WT}) (Fig.3.1B). Additionally, the expression of GFP from the second allele was confirmed in E2 and E3 mutant clones by immunofluorescence ($ESC^{KO\text{-}E2}$, $ESC^{KO\text{-}E3}$; Fig.3.1C). GFP expression, alongside with drug resistance to neomycin and hygromycin in culture, further indicated correct transcription of the inserted cassettes in homozygous KO clones.

The depletion of Jmjd2c was validated both at the mRNA and protein level in the two independent *Jmjd2c*-KO ESC lines. For this, specific primer pairs were designed spanning exons

located upstream (exons 5 and 7; Fig3.1D) and downstream (exons 14 and 15, exons 18 and 20; Fig.3.1.D) of the "critical exon" and hence the trapping cassettes. *Jmjd2c* mRNA levels detected when using upstream primer pairs were unchanged in mutant cells relative to WT ESCs. In contrast, transcript levels spanning exons located downstream of the critical exon were significantly reduced by about 90% (Fig.3.1D). The residual expression of full-length transcript could arise due to alternative splicing around the inserted cassettes, but this however did not result in translation of a full-length Jmjd2c protein, as described below.

To examine whether the targeted trapping resulted in the abolishment of full-length Jmjd2c protein, western blot analysis was performed on total protein lysates extracted from WT and *Jmjd2c*-KO ESC clones E2 and E3 using a previously described (Loh et al., 2007) anti-Jmjd2c antibody (kindly provided by Professor Huck-Hui Ng, Genome Institute Singapore). As expected, high levels of Jmjd2c protein were detected in the WT ESC line, whereas expression of full-length Jmjd2c was completely abolished in *Jmjd2c*-KO ESC lines (Fig.3.1E), hence confirming that these cells are indeed constitutively depleted for Jmjd2c. Moreover, Jmjd2c depletion did not affect Jmjd2b, as Jmjd2b protein expression remained unchanged between the same WT and KO ESC lines (Fig.3.1E). The insertion of gene trapping cassettes is predicted to result in the expression of putative truncated proteins, however since the trapping cassettes are inserted upstream from the PHD and Tudor domains, this is not expected to be located in the nucleus (Pedersen et al., 2014; Shin and Janknecht, 2007). Altogether, these results validate the successful generation of homozygous *Jmjd2c*-knockout ESC clones.

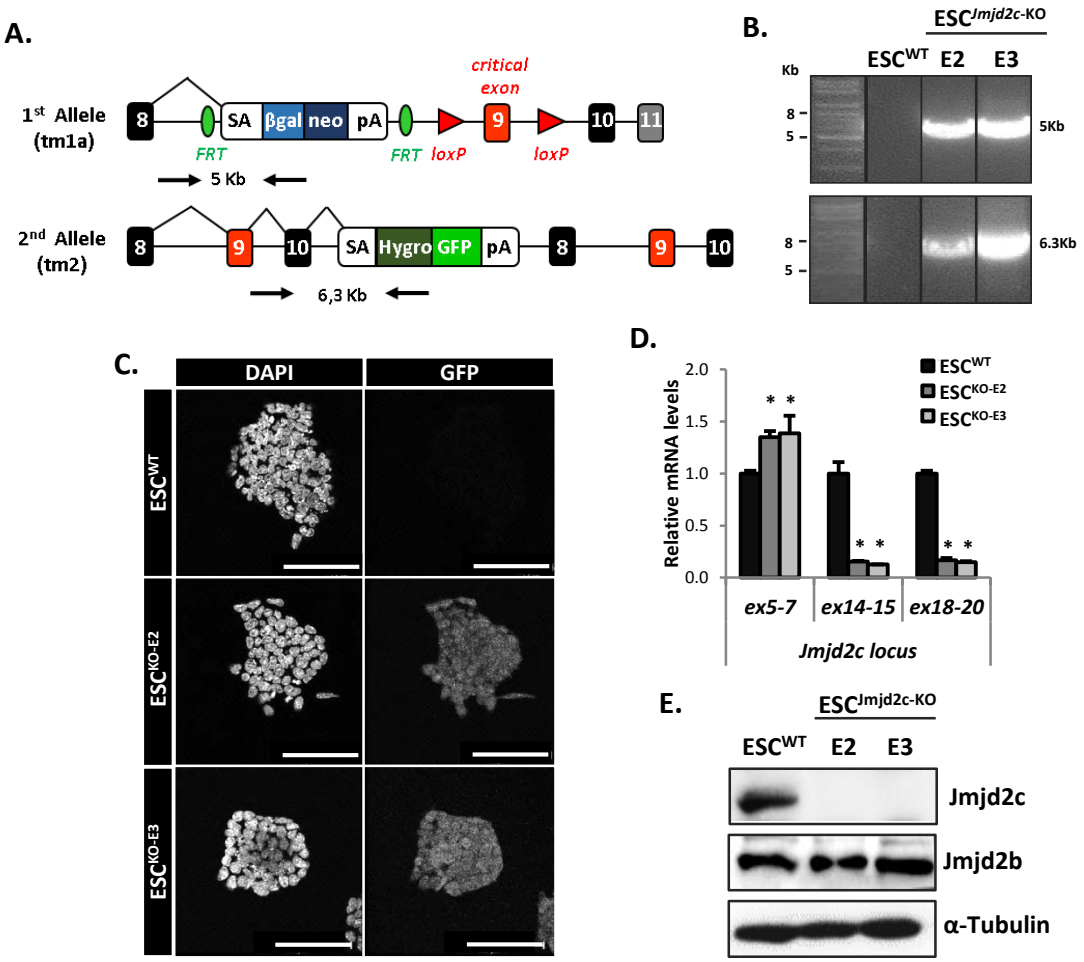


Figure 3.1 Validation of *Jmjd2c*-knockout in ESCs

A. Schematic representation of the double knockout strategy. The first and second mutated alleles carry a cassette containing the neomycin-βgalactosidase and hygromycin-GFP resistance/expression genes, to allow for efficient selection and identification. In the first allele, the FRT sites were inserted to allow for optional reversion of homozygous to heterozygous by removal of the cassette with FLP recombinase. The loxP sites were inserted in the vicinity of the critical exon 9 to allow for the optional creation of a conditional null allele with Cre recombinase. The splice acceptor (SA) site in the cassettes promotes incorporation of the insert into the RNA transcript after RNA splicing, whereas the polyA (pA) signal site ensures that RNA polymerase transcription is stopped. Arrows indicate the location of the primers used for genotyping and the expected PCR band sizes is indicated for each primer pair. **B.** Genotyping of both alleles was achieved by Long Range PCR on the JM8 wild-type (ESC^{WT}) cells, and homozygous *Jmjd2c*-knockout ($ESC^{Jmjd2c\text{-}KO}$) mutant clones (E2, E3). Performed in collaboration with Dr Cynthia Fisher. **C.** Immunofluorescent labelling of GFP in ESC^{WT} and $ESC^{Jmjd2c\text{-}KO}$ clones (KO-E2, KO-E3). DAPI staining was used to visualise nuclei. Scale,

100 µm. **D.** Transcript levels of *Jmjd2c* were quantified using primers located upstream (exon5-7) and downstream (exon14-5, exon18-20) of the critical exon. Levels were normalized to 2 housekeeping genes (*S17* and *L19*), and expressed relative to ESC^{WT}. Error bars represent SEM of 3 biological replicates. $P<0.05$ (*), Mann-Whitney *U* test. **E.** Western Blot for Jmjd2c and Jmjd2b and on whole cell extracts from ESC^{WT} and $ESC^{Jmjd2c\text{-}KO}$ clones as indicated. α-Tubulin was used as a loading control.

3.2.2 *Jmjd2c*-KO ESC clones remain undifferentiated in culture

The resulting Jmjd2c-KO ESC lines were able to be routinely maintained in culture in the presence of leukemia inhibitory factor (LIF) and serum, suggesting that the constitutive depletion of this protein was not detrimental to the maintenance of an undifferentiated ESC state, unlike previously reported upon transient knockdown (Loh et al., 2007). Although the targeted trapping approaches used for the generation of this knockout relies solely on having an active Jmjd2c promoter, *Jmjd2c*-KO ESCs were maintained under selection with G418/geneticin and hygromycin, to prevent any cross-contamination with their WT counterparts, cultured in parallel. This selection was however removed prior to any functional experiment to avoid influencing their cell growth rate, and hence the outcome of the experiments.

To fully evaluate the undifferentiated state of *Jmjd2c*-KO ESCs, the growth rate was initially assessed by scoring the number of cells cumulating in routine culture conditions over eight consecutive passages, and found to be similar between WT and KO ESC populations (Fig.3.2A). Next, Alkaline Phosphatase (AP) activity was used as an indicator of undifferentiated state in colony forming efficiency assays. For this, *Jmjd2c*-KO and WT cells were plated at low density, cultured in the presence or in the absence of LIF for 5 days, and then scored as undifferentiated, mixed or differentiated colonies based on the observed AP staining intensity and colony morphology (Fig.3.2B). In the presence of LIF, the three cell lines (WT, E2 and E3) studied demonstrated similar self-renewing abilities regardless of *Jmjd2c* expression (Fig.3.2B). In all cases, over 60% of colonies were scored as undifferentiated (Undiff), showing a typical round and compact morphology and strong AP staining, whereas nearly 30% and 10% of colonies were

scored as mixed and differentiated respectively, with bright AP-positive centres surrounded by a ring of flattened, differentiated cells (Mixed) or displaying a sparse morphology and little or no AP staining (Diff) (Fig.3.2B). As expected, in the absence of LIF, 80-90% of colonies stained negative for AP, indicative of spontaneous differentiation (Fig.3.2B). Noteworthy, increased percentages of AP-positive colonies (Undiff and Mixed) were observed in *Jmjd2c*-KO ESC clones (10-20%) as compared to their WT counterparts (5-10%), suggesting that a higher proportion of *Jmjd2c*-depleted colonies could retained an undifferentiated state in the absence of LIF.

Finally, the expression of pluripotency-associated markers was investigated through three approaches: immunofluorescent staining, western blotting and quantitative expression analysis (RT-qPCR). *Nanog* is typically expressed in a heterogeneous manner in ESCs grown under routine (LIF, serum) culture conditions as used in our laboratory (Chambers et al., 2007). Indeed we detected a strong and heterogeneous expression pattern for Nanog by immunofluorescence staining in both WT and *Jmjd2c*-KO ESC lines (Fig.3.2C). Moreover, strong Oct4 signal was also detected in all of the three cell lines (WT, E2 and E3) analysed, though Oct4 expression appeared more heterogeneous in *Jmjd2c*-KO ESC relative to WT ESCs (Fig.3.2C). The levels of Nanog, Oct4 and Esrrb protein were also assessed by western blot in WT and KO ESC clones (Fig.3.2D), showing variable expression patterns across these cell lines. Here, Nanog levels were reduced in KO clone E2, whereas the levels of Esrrb were conversely increased in the same clone. Oct4 appeared slightly reduced in the KO clone E3 as compared to WT ESCs (Fig.3.2D).

Moreover, the transcript levels of the pluripotency-associated markers *Nanog*, *Oct4*, *Sox2*, *Klf4*, *Esrrb*, and *Tbx3* were also quantified in the RNA from WT and *Jmjd2c*-KO ESCs, showing an overall reduction of about 20-50% of expression of all markers as seen in both KO clones E2 and E3 (Fig.3.2D). Altogether, the analysis of pluripotency markers in Jmjd2c-KO ESC by three different approached revealed a high degree of heterogeneity, reiterating a role for Jmjd2c in the stability of ESC-specific gene expression.

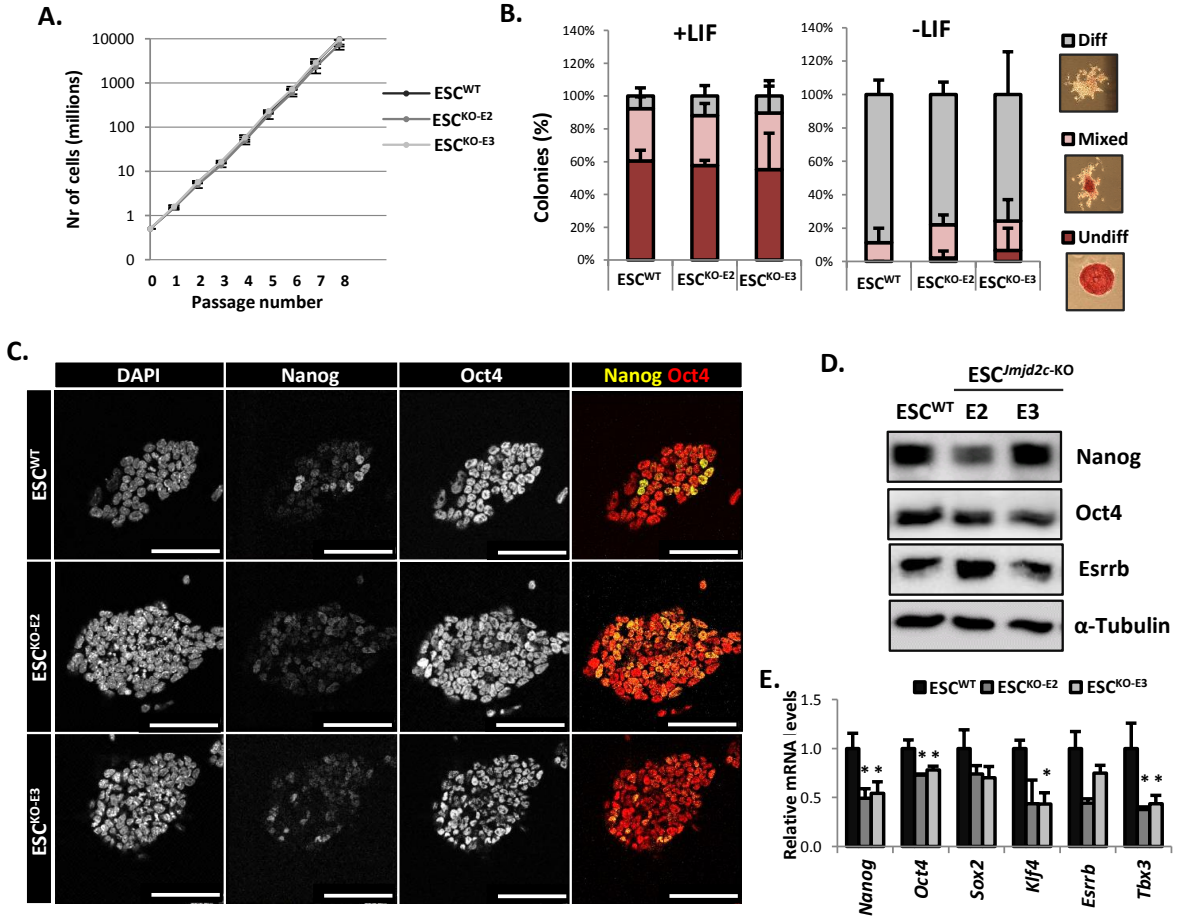


Figure 3.2 *Jmjd2c*-knockout ESCs retain normal self-renewal ability

A. Growth curve of WT (ESC^{WT}) and *Jmjd2c*-KO (ESC$^{Jmjd2c\text{-}KO}$) clones E2 (KO-E2) and E3 (KO-E3) in self-renewing conditions over 8 passages (16 days). Plot represents an average of the cell number scored in 3 independent experiments. Error bars indicate ±SEM. **B.** Percentage of colony types scored in cells growing with (+) or without (-) LIF for 5 days. Colonies were counted and scored as undifferentiated (Undiff), mixed or differentiated (Diff) according to the intensity of alkaline phosphatase (AP) staining. Plots represent percentage of colonies counted in two biological replicates plated in duplicate. Error bars indicate ±SEM. **C.** Immunofluorescent labelling of Nanog and Oct4 in ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clones. DAPI staining was used to visualise nuclei. Scale, 100 µm. **D.** Western blot for Nanog, Oct4 and Esrrb in whole cell lysates from ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clones as indicated. α-Tubulin was used as a loading control. **E.** Transcript levels of pluripotency-associated markers *Nanog, Oct4, Sox2, Klf4, Esrrb* and *Tbx3* in ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clones. Data were normalized to 2 housekeeping genes (*S17* and *L19*), and expressed relative to ESC^{WT}. Error bars represent ±SEM of 3 biological replicates. $P<0.05$ (*), Mann-Whitney *U* test.

Altogether, these results suggest that that constitutively *Jmjd2c*-depleted ESCs retain normal self-renewal abilities despite lessened and more heterogeneous expression levels of pluripotency-associated genes.

Another key feature of ESC identity is their unique epigenetic signature at the promoters of differentiation-associated genes, which is characterized by the simultaneous presence of permissive (H3K4me2/3) and repressive (H3K27me3) histone modifications (Fig.3.3A) (Azuara et al., 2006; Bernstein et al., 2006). These so-called bivalent chromatin domains are thought to prime genes for future activation (or repression) upon differentiation. Their premature expression is prevented by the action of the Polycomb Repressive Complexes 1 and 2, which maintain these genes repressed by tri-methylating H3K27 (PRC2) and mono-ubiquitinating H2AK119 (PRC1; Fig.3.3A).

To evaluate whether the depletion of Jmjd2c had an impact on PRC's action or recruitment, as reported for the histone demethylase Kdm2b (Farcas et al., 2012; Wu et al., 2013), and hence on the maintenance of a transcriptional repressive state at differentiation-associated genes, the transcript levels of a selected panel of genes was evaluated by RT-qPCR (Fig.3.3B). This panel included markers associated with the trophoectoderm (*Cdx2, Hand1*), mesoderm (*Brachyury/T, Mixl1*), endoderm (*Gata6, Foxa2*) and neuroectoderm (*Math1, Pax3*) lineages, and none of them were found to be de-repressed in *Jmjd2c*-KO clones compared to WT ESCs (Fig.3.3B). Moreover, the enrichment levels for members of PRC2 (Suz12) and PRC1 (Ring1B) complexes was evaluated using Chromatin Immunoprecipitation (ChIP) and qPCR, at the bivalent promoters of differentiation-associated genes: *Cdx2*, *Hand1*, *Flk-1*, *Mixl1*, *Gata4, Sox7, Pax3* and *Math1* alongside a control active gene (*Actin*) and a non PRC-bound gene region (*Gata1*). Likewise, a similar pattern of enrichment was detected in WT cells, and the two *Jmjd2c*-KO clones E2 and E3, indicating that PRC recruitment and action was not affected by *Jmjd2c* depletion in ESCs (Fig.3.3C).

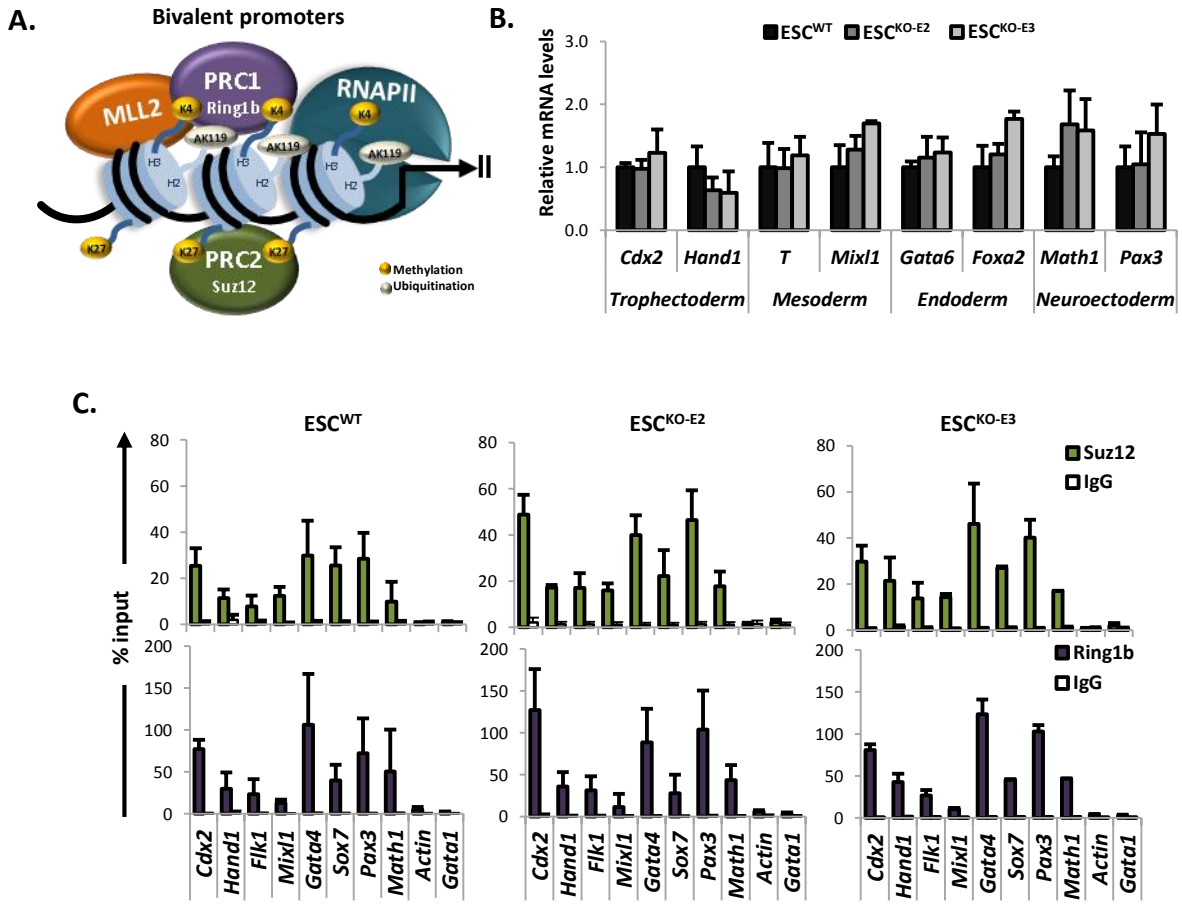


Figure 3.3 PRC-mediated repression of bivalent genes is preserved in *Jmjd2c*-knockout ESCs

A. Schematic representation of bivalent chromatin structure harboring H3K4me3 (active), H3K27me3 (inactive) and H2AK119Ub histone marks, deposited by Mll2, PRC2 and PRC1, respectively, and a poised conformation of RNA Polymerase II (RNAP) complexes. **B.** Transcript levels of a selection of genes associated with each of the indicated lineages in WT (ESC^{WT}) and KO ($ESC^{Jmjd2c\text{-}KO}$) clones E2 (KO-E2) and E3 (KO-E3), cultured under self-renewing conditions. Data was normalised to 2 housekeeping genes (*S17* and *L19*) and expressed relative to ESC^{WT}. Error bars represent ±SEM of 3 biological replicates. **C.** Binding of PRC2 and PRC1 components, Suz12 (green) and Ring1b (purple), at the promoters of differentiation-associated genes were assessed by ChIP and quantitative PCR of the immunoprecipitated DNA fragments in ESC^{WT} and $ESC^{Jmjd2c\text{-}KO}$ clones. Background levels (IgG binding) are shown as white bars. Negative controls (not occupied by PRC1 and PRC2) included *Actin* and *Gata1*. Enrichment is expressed relative to input. Error bars represent ±SD of two independent ChIP assays. The experiment in **C.** was performed by Sergio German and Anne Helness.

In summary, these data demonstrate that Jmjd2c-KO ESC lines retain an undifferentiated state in culture despite unstable expression of pluripotency-associated genes. The epigenetic signature of Jmjd2c-KO ESC lines appeared unaffected, and differentiation-associated genes remained repressed and bound by PRC1 and PRC2 complexes as seen in WT ESCs.

3.2.3. Depletion of Jmjd2c in ESCs correlates with an increase in bulk H3K9 methylation levels

Jmjd2c has been previously identified as an H3K9me2/me3 and H3K36me3 histone demethylase in various cell types (Katoh and Katoh, 2004; Klose et al., 2006b; Whetstine et al., 2006). Transient knockdown of *Jmjd2c* in ESCs was, however, shown to result in an increase of H3K9me3 levels alone, with no effect on H3K9me2 and H3K36me3 (Loh et al., 2007). Hence, the effect of constitutive *Jmjd2c* depletion in the bulk levels of these histone modifications was re-visited using acid-extracted histone lysates from WT and *Jmjd2c*-KO ESC clones. Here, an initial titration of the amount of histone lysates used was conducted for H3K9me2 and H3K9me3 together with H3K4me2 as control (Fig.A-I – Appendix A). With 2ug of histone lysates, an increase in the bulk levels of H3K9me2 was consistently identified in *Jmjd2c*-KO as compared to WT ESCs, whereas only a milder effect and no effect were observed when analysing H3K9me3 and H3K36me3 levels, respectively (Fig3.4A; Fig.A-I – Appendix A), which was validated with different antibody sources (Fig.A-I – Appendix A). Moreover, no changes were detected in the bulk levels of control (H3K4me2 and H3K27me3) histone modifications whose deposition is not linked with Jmjd2c activity in ESCs (Fig.3.4A). To ensure that the identified increase in the levels of H3K9me2 and H3K9me3 strictly correlated with the absence of Jmjd2c in the KO ESCs, the mRNA expression levels of other H3K9-demethylases (*Jmjd2a*, *Jmjd2b*, *Jmjd1a* and *Lsd1*; Fig.3.4B) and H3K9-methytransferases (*G9a*, *GLP*, *Suv39h1*, *Suv39h2* and *Eset*; Fig.3.4C) were quantified and found to be similar to WT levels for most genes, despite slight variations in clone E3.

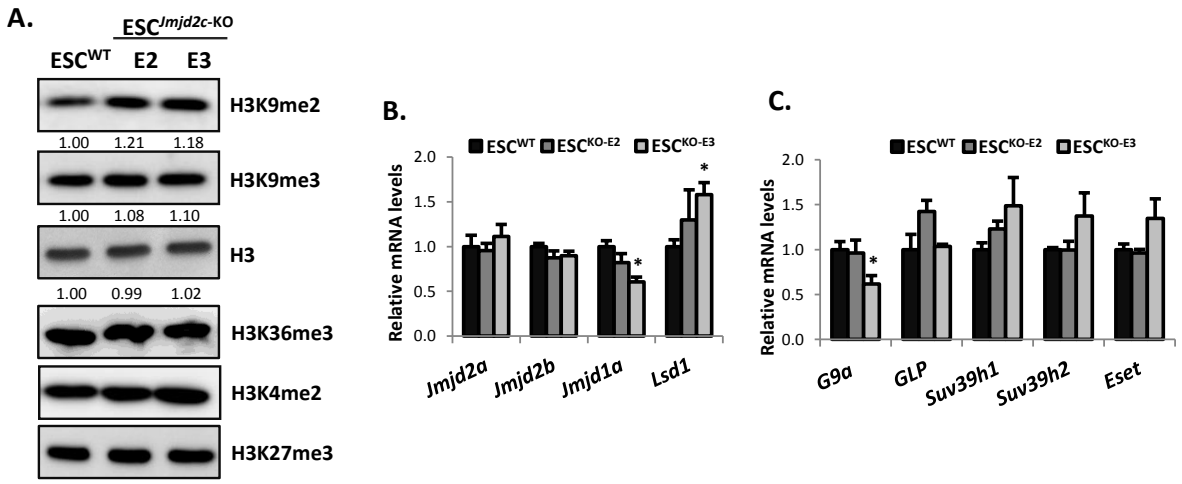


Figure 3.4 Depletion of Jmjd2c results in increased bulk levels of di- and tri-methylation of H3K9

A. Western blots showing the bulk levels of H3K9me2, H3K9me3, H3K36me3, H3K4me2, H3K27me3 and total H3 in acid-extracted histone lysates from WT (ESC^{WT}) and KO ($ESC^{Jmjd2c\text{-}KO}$) clones E2 (KO-E2) and E3 (KO-E3) ESCs. Results were produced in at least 3 biological replicates and differences validated with alternative antibody sources. Band quantification is shown below relative to WT. **B.** Transcript levels of a selection of H3K9-demethylases (left panel) and H3K9-methyltransferases (right panel) in ESC^{WT} and $ESC^{Jmjd2c\text{-}KO}$ clones growing in self-renewing conditions. Data was normalised to 2 housekeeping genes (*S17* and *L19*) and expressed relative to ESC^{WT}. Error bars represent ±SEM of 3 biological replicates. $P<0.05$ (*), Mann-Whitney *U* test.

To further demonstrate the link between Jmjd2c and H3K9-demethylation in ESCs, a rescue ESC line was generated by stably transfecting a vector carrying wildtype *Jmjd2c* cDNA into *Jmjd2c*-KO clone E3 ($ESC^{KO\text{-}rescue}$; Fig.3.5). As a control, an empty vector was transfected in a parallel culture of the same clone ($ESC^{KO\text{-}control}$). Validation of the rescued phenotype was carried out at the protein level by western blot (Fig.3.5A). Additionally, the protein expression levels of the pluripotency markers Nanog, Oct4 and Esrrb was also assessed and found to be higher for Nanog and, to some extent, for Esrrb in the rescued KO clone E3 (Fig.3.5A). A significant up-regulation of *Nanog* was also confirmed at the transcript levels, whereas the levels of *Oct4*, *Klf4* and *Esrrb* remained comparatively unchanged (Fig.3.5B). This observation corroborates with a key role for Jmjd2c in regulating *Nanog* gene expression in ESCs (Loh et al., 2007). Nanog and Oct4 protein

expression was also evaluated by immunofluorescence. Rescuing the expression of Jmjd2c did not impact on the mosaic expression pattern of Nanog (Fig.3.5C), but appeared to restore a more homogeneous Oct4 expression pattern as compared to control KO ESCs (Fig.3.5C) and WT ESCs (Fig3.2C).

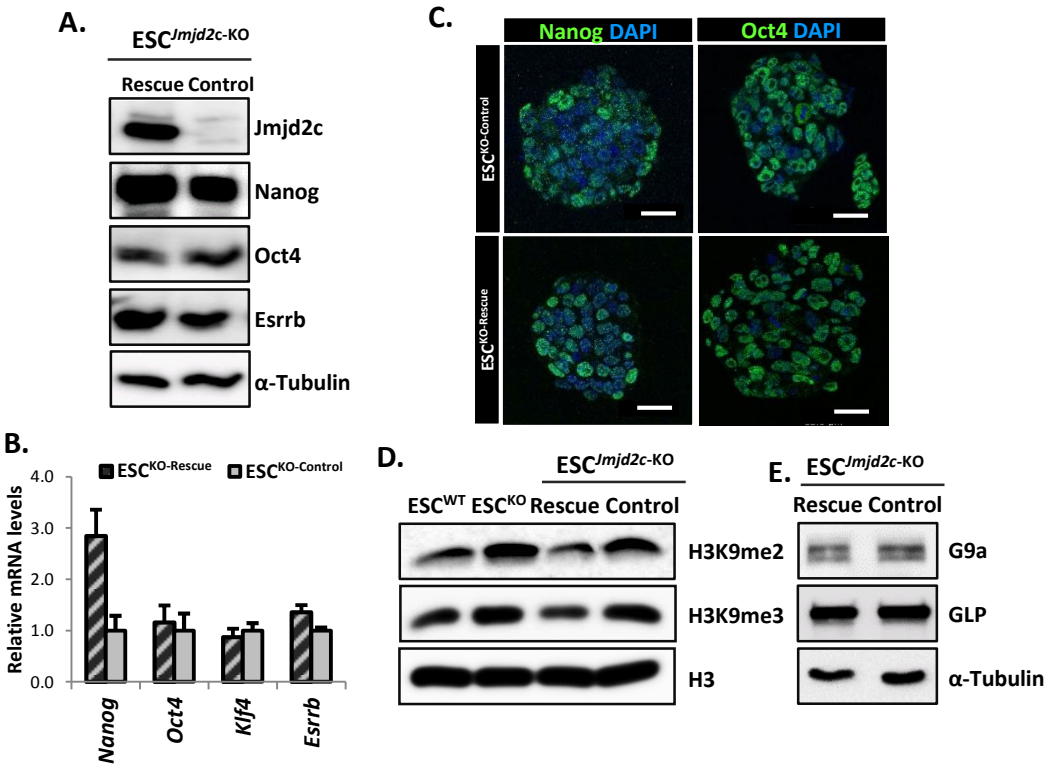


Figure 3.5 Rescue of *Jmjd2c* expression in knockout ESCs

A. Western blot for Jmjd2c, Nanog, Oct4 and Esrrb in ESC$^{Jmjd2c\text{-}KO}$ clone E3 transfected with full-length Jmjd2c cDNA (Rescue) or with an empty vector (Control). α-Tubulin was used as a loading control. **B.** Transcript levels of pluripotency markers *Nanog, Oct4, Klf4,* and *Esrrb* in rescued (ESC$^{KO\text{-}Rescue}$) and control (ESC$^{KO\text{-}Control}$) clones. Data were normalized to 2 housekeeping genes (*S17* and *L19*), and expressed relative to control. Error bars represent ±SEM of 3 biological replicates. **C.** Immunofluorescent labelling of Nanog and Oct4, in rescue and control KO ESCs. DAPI staining was used to visualise nuclei. Scale, 30µm. **D.** Western blots showing the bulk levels of H3K9me2, H3K9me3 and total H3 in acid-extracted histone lysates from ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clone E3, and E3 rescue and control KO ESCs. **E.** Western blots for G9a and GLP in Jmjd2c Rescue and Control ESCs. α-Tubulin was used as a loading control. The rescue and control ESC lines and results in figure **C** were produced by Lauren Weavers.

Finally, the impact of re-introducing *Jmjd2c* on the levels of H3K9me2/3 was investigated on histone lysates from rescued and control KO ESCs, showing restoration to WT levels in the rescued cells (Fig.3.5D). Importantly, this occurred without any obvious alteration in the protein levels of the known H3K9me2-methytransferases G9a and GLP in ESCs (Fig.3.5E).

Overall, these experiments demonstrate that constitutive depletion of *Jmjd2c* in ESCs results in a noticeable increase of H3K9me2 with a milder effect on H3K9me3 levels, hence inferring a specific role for Jmdj2c as an H3K9me2, in addition to H3K9me3 demethylase in ESCs, as previously reported upon transient knockdown of *Jmjd2c* (Loh et al., 2007).

3.2.4. Jmjd2c is required for efficient somatic ESC differentiation

Pluripotent ESCs can be induced to differentiate into the three germ layers, when cultured under permissive conditions. To address whether depleting Jmjd2c in ESCs impacts on ESC differentiation, WT and *Jmjd2c*-KO clone E3 ESCs were induced to form embryoid bodies (EB). This involves culturing ESCs in suspension in the absence of LIF and in reduced serum level over 9 days, during which cell aggregates will form allowing the spontaneous emergence of cells derived from the three embryonic germ layers. EB morphology was recorded on a daily basis (Fig.3.6A), and samples were collected over different time-points for subsequent mRNA expression analysis (Fig.3.6B). In particular, the kinetics of induction of early lineage-specific markers for epiblast (*Fgf5*) and mesoderm (*Brachyury* and *Mixl1*), and late markers for endoderm (*Foxa2*, *Gata6*, *Gata4*) and neuroectoderm (*Mash1*, *Pax3* and *Pax6*) was investigated. Additionally, the expression of the pluripotency-associated factors *Nanog* and *Oct4*, and *Jmjd2c* itself, were also profiled during EB formation.

Under these conditions WT ESCs formed large compact cell aggregates (Fig.3.6A; upper panel), whereas *Jmjd2c*-KO ESCs formed smaller aggregates, indicative of compromised differentiation and/or increased cell death (Fig.3.6A; lower panel).

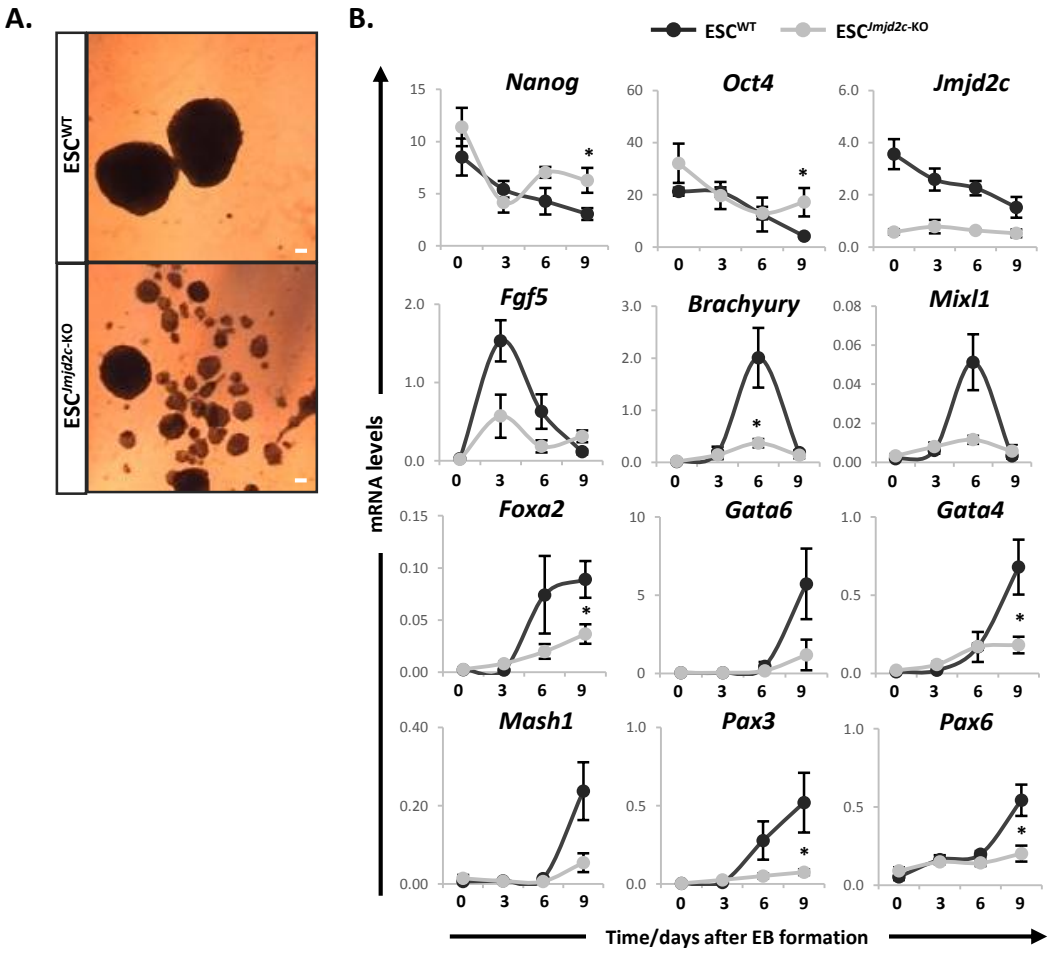


Figure 3.6 Embryoid body (EB) formation is impaired in *Jmjd2c*-knockout ESCs

A. ESCs were cultured in suspension and induced to differentiate upon EBs formation in the absence of LIF and reduced serum level. Representative phase-contrast images of day 9 EBs are shown. Bars, 100µm. **B.** Transcript levels for pluripotency-associated (*Nanog*, *Oct4*), Jmjd2c, epiblast (*Fgf5*), mesoderm (*Brachyury*, *Mixl1*), endoderm (*Foxa2, Gata6, Gata4*), and neuroectoderm (*Mash1, Pax3, Pax6*) markers in WT and *Jmjd2c*-KO clone E3 derived EBs collected at day 0, 3, 5, 6 and 9 of differentiation. Values were multiplied by 100 and plotted as the average of 4 experiments, normalized to housekeeping genes (S17, L19). Error bars represent ±SEM of at least 3 experiments. Data represented as day 6 is an average of day 5 (n=2) and day 6 (n=2). $P<0.05$ (*), Mann-Whitney *U* test performed at peak time points in WT for each differentiation marker.

Both cell populations showed progressive reduction in the expression of *Nanog* and *Oct4*. Partial expression of these pluripotency factors was, however, inappropriately retained at later

time-points in *Jmjc2c*-KO EBs only (Fig.3.6B). Interestingly, *Jmjd2c* mRNA levels remained high in early WT EBs together with *Oct4*, then progressively declined to near background levels as found in *Jmjc2c*-KO cells (Fig.3.6B), suggesting that Jmjd2c activity may be required during the early steps of ESC differentiation. The expression of the early epiblast-specific marker *Fgf5* transiently peaked at day 3 during WT EBs formation and decreased back to basal levels at day 9 (Fig.3.6A). In *Jmjd2c*-KO EBs *Fgf5* induction was about 3-fold lower at day 3 relative to WT EBs, and its expression was also retained at day 9 as seen for both *Oct4* and *Nanog* (Fig.3.6B). Similarly, the expression of the early mesoderm-specific markers *Brachyury* and *Mixl1* peaked at day 6 in WT EBs, being reduced by about 10 and 5-fold, respectively, in Jmjd2c-KO EBs (Fig.3.6B). The expression of the late endoderm- and neuroectoderm-specific markers only peaked at day 9 in WT EBs, and was found to be again severely impaired or reduced in *Jmjd2c*-KO EBs (Fig.3B).

These striking impairment of the induction of differentiation markers in *Jmjd2c*-KO EBs suggests that Jmjd2c might be essential for the proper execution of multi-lineage pathways, with a possible blockage during the early steps of ESC differentiation.

To further explore this and confirm the requirement of Jmjd2c for proper ESC differentiation, an alternative differentiation assay was used: monolayer differentiation of ESCs into neuronal precursors. In this assay, ESCs are cultured at low density in serum-free N2B27 media in the absence of LIF, for a period of 5 days, during which typical neuronal rosette-like structures emerge (Ying et al., 2003b). Initial optimization of this assay revealed that there was a high level of cellular stress when swapping the media conditions from serum-based ESC media to serum-free N2B27 media. In order to overcome this effect and ensure that both cell populations were differentiated in optimal conditions, the WT and *Jmjd2c*-KO clone E3 ESC cultures were adapted into serum-free media by being cultured in the so-called 2i conditions: N2B27 supplemented with LIF, the Mek inhibitor PD0325901 and the Gsk3 inhibitor CHIR99021 (Ying et al., 2008). These ESCs culture conditions have been reported to promote the acquisition of a more naïve or ground-

state of pluripotency, with lower or absent gene priming for differentiation, and acquisition of dome-shaped colonies (Fig.3.7A), as opposed to the metastable state cells exhibited in serum cultures. Morphology changes were monitored for induction of neuronal precursors over the course of differentiation (Fig.3.7B) and the expression of the neuronal-specific marker *Nestin* was assessed at day 5 by immunofluorescence staining (Fig.3.7C). During this process, a transient epiblast(EpiS)-like population has been proposed to emerge at days 2/3, which has been identified by the expression of the epiblast marker *Fgf5* (Abranches et al., 2009). Here, both cell populations adopted an EpiS-like morphology by day 2 (Fig.3.7B), with no striking morphological differences between them. By day 5 neuronal progenitor colonies emerged in WT cultures revealed by the appearance of neural tube-like and rosette-like structures (Fig.3.7B; Fig.A-II – Appendix A).

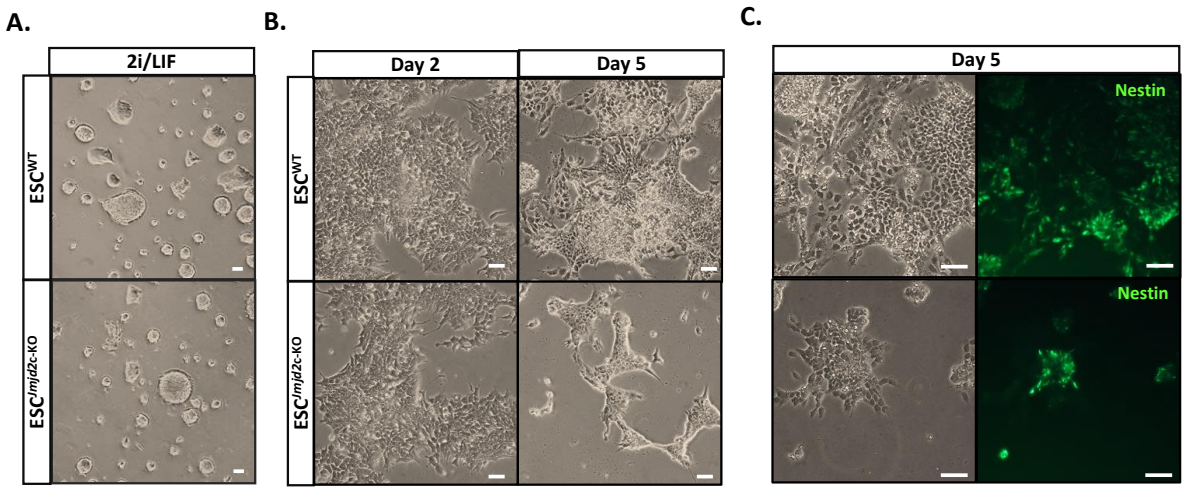


Figure 3.7 Inefficient generation of neuronal progenitors in *Jmjd2c*-depleted ESCs

A. WT (ESC^{WT}) and KO ($ESC^{Jmjd2c\text{-}KO}$) clone E3 were cultured in serum-free 2i/LIF conditions prior to differentiation. **B.** ESCs were seeded at low density and cultured in neuronal differentiation media for 5 days. Phase-contrast images showing the morphology at day 2 and day 5 in these conditions are shown. **C.** Phase-contrast and fluorescent signals for the staining of the neuronal marker Nestin in day 5 neuronal precursors Bars, 100µm.

In comparison, a lower incidence of *Jmjd2c*-KO neural-like colony formation was observed at this late time point (Fig.3.7B; Fig.A-II – Appendix A) the nature of which was further confirmed

by immunofluorescence using the early neuronal marker Nestin (Fig.3.7C).Notably, the few colonies present at this late stage of differentiation were harbouring a neural-like phenotype and showed positive Nestin signals indicating that differentiation towards neuronal progenitors was possible, although strikingly less efficient.

Altogether, these two experimental approaches uncovered a key role for Jmjd2c in the proper execution of ESC differentiation. Moreover, the persistence of *Fgf5* expression upon EB formation, and the possible blockage of differentiation after an EpiS-like step under neuronal conditions in *Jmjd2c*-KO ESCs, suggest that Jmjd2c function might be essential at the onset of ESC differentiation, possibly during the transition to an early *Fgf5*-expressing epiblast-like stage.

3.2.5 Stable knockdown of *Jmjd2c* mimics *Jmjd2c*-knockout phenotype

Transient depletion of *Jmjd2c* was previously shown to be detrimental for the maintenance of ESC integrity (Loh et al., 2007). However, under constitutive depletion of *Jmjd2c* using gene-trap and homologous recombination approaches, ESCs were able to sustain a self-renewing and undifferentiated state, as demonstrated in this Chapter. To re-assess the effect of knocking down *Jmjd2c* expression using shRNA vectors, the possibility to stably generate undifferentiated *Jmjd2c* knockdown (KD) ESC clones was carefully investigated upon colony picking and prolonged drug selection. ESCs from a different mouse background (129) than *Jmjd2c*-KO (B6) cells were chosen to be independently transfected with two previously validated *Jmjd2c* shRNA vectors (shRNA1 and shRNA2) and a control shRNA against Luciferase (shLuc), kindly provided by Professor Huck Ng (Genome Institute of Singapore; Loh et al., 2007). The knockdown efficiency of these vectors was firstly tested upon transient transfection followed by 4 days of puromycin selection. As previously reported, this ensued a notable incidence of spontaneous differentiation as indicated by the presence of AP-negative colonies in pools of transfected cells with either shRNA1 or shRNA2 vector in contrast to control cultures (Fig.3.8A). Concordantly, the expression of *Jmjd2c* upon transient knockdown was significantly reduced by about 50% with both vectors (Fig.3.8.B).

Nevertheless, AP-positive and puromycin-resistant colonies could still be identified following *Jmjd2c* knockdown, which suggested that stable ESC cell lines could be established. To achieve this, ESCs were transfected with either shRNA1 or shRNA-Luc vectors, and after four days of puromycin selection individual ESC clones were manually picked and seeded in order to be further expanded under puromycin selection. Three ESC clones were selected for characterization, and the knockdown efficiency in these clones validated by western blot (Fig.3.8C) and mRNA quantification (F.3.8D). Here, a knockdown efficiency of about 80% was achieved, with no dramatic impact on the expression of other *Jmjd2* family members (Fig.3.8D). Moreover, the expression of pluripotency-associated markers *Nanog*, *Oct4*, *Sox2*, *Klf4*, *Esrrb* and *Tbx3* was assessed in three *Jmjd2c* KD clones and found to be heterogeneous or unstable (Fig.3.8D). Concordantly with the effect observed in constitutive *Jmjd2c*-KO, shJmjd2c-clone1 presented down-regulation of all pluripotency factors, clone2 showed down-regulation of *Esrrb* and *Tbx3*, whereas clone3 did not show any difference (Fig.3.8D). Notably, the mild decrease in expression of pluripotency markers in KD clones 1 and 2 mirrored the mild down-regulation detected in the expression of *Jmjd2b* (Fig.3.8D). This is concordant with a recently described role for Jmjd2b in the regulation of these markers (Das et al., 2014), further discussed in Chapter VII. Importantly, all *Jmjd2c* KD and control clones exhibited similar self-renewal abilities in the presence of LIF, with around 80% of colonies showing a strong-AP staining and undifferentiated morphology, reiterating that Jmjd2c is dispensable for ESC self-renewal (Fig.3.8E).

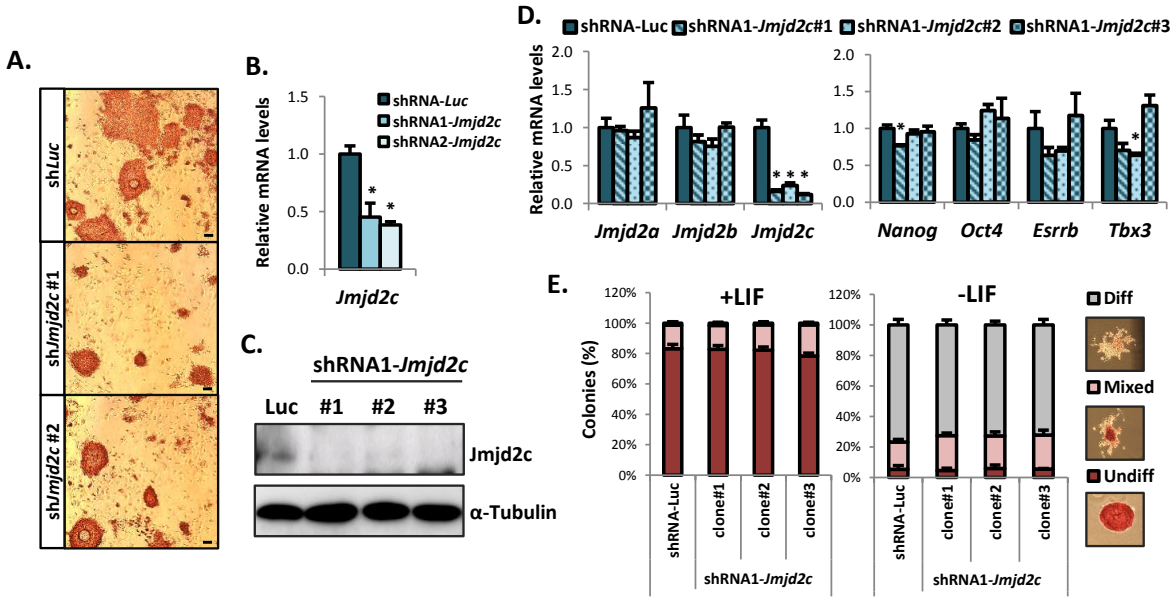


Figure 3.8 Generation of stable *Jmjd2c*-knockdown ESC lines

A. E14 ESCs cells were transfected with either of two different Jmjd2c shRNA vectors (shRNA1 and 2) or a control vector for Luciferase (shRNA-Luc). Alkaline phosphatase (AP) staining was preformed 4 days after puromycin selection. Bars, 100µm. **B.** *Jmjd2c* mRNA levels upon transfection with Jmjd2c shRNA1, 2 or shRNA-Luc, 4 days after puromycin selection. Expression levels were normalised to housekeeping genes (*L19* and *S17*) and represented as the average of 3 experiments ±SEM, normalised to the control vector. P<0.05 (*), Mann-Whitney *U* test. **C.** Western blot for Jmjd2c in 3 stable clones picked from ESCs transfected with *Jmjd2c* shRNA1 and a control shRNA-Luc cell line. **D.** Transcript levels of *Jmjd2* family members and pluripotency factors in 3 stable clones picked from *Jmjd2c* shRNA1-transfected ESCs. Levels were normalised to housekeeping genes (*L19* and *S17*) and expressed relative to control (shRNA-Luc) as average ±SEM of 3 biological replicates. P<0.05 (*), Mann-Whitney *U* test. **E.** Percentage of colonies scored as undifferentiated (Undiff), mixed or differentiated (Diff) according to the intensity of AP staining of 3 stable shRNA1-*Jmjd2c* transfected ESC clones, and a control shRNA-Luc ESC clone. Cells were grown with (+) or without (-) LIF for 5 days. Plots represent percentage of colonies counted in 3 biological replicates. Error bars represent ±SEM. The results in figures **A** and **B** were produced by Lauren Weavers, whereas the results in figures **D** and **E** were produced by Donja Karimlou.

Constitutive knockout of *Jmjd2c* resulted in impaired differentiation upon EB formation (Section 3.2.4), presumably at the onset of ESC differentiation, as suggested by maintenance of *Oct4* and *Fgf5* low expression levels at the later stages of EB induction. Here, the differentiation abilities of stable *Jmjd2c*-KD ESCs upon EB formation was subsequently investigated, with a focus on early epiblast and primitive streak markers. Upon EB formation, no striking morphological differences were observed between the control shLuc and shJmjd2-clone1 (Fig.3.9A). Progressive down-regulation of the pluripotency factors *Nanog* and *Oct4* was measured in both shRNA-Luc and shRNA1-*Jmjd2c* EBs, and a decrease in the expression of *Jmjd2c* was consistently identified in shRNA-Luc EBs (Fig.3.9B).

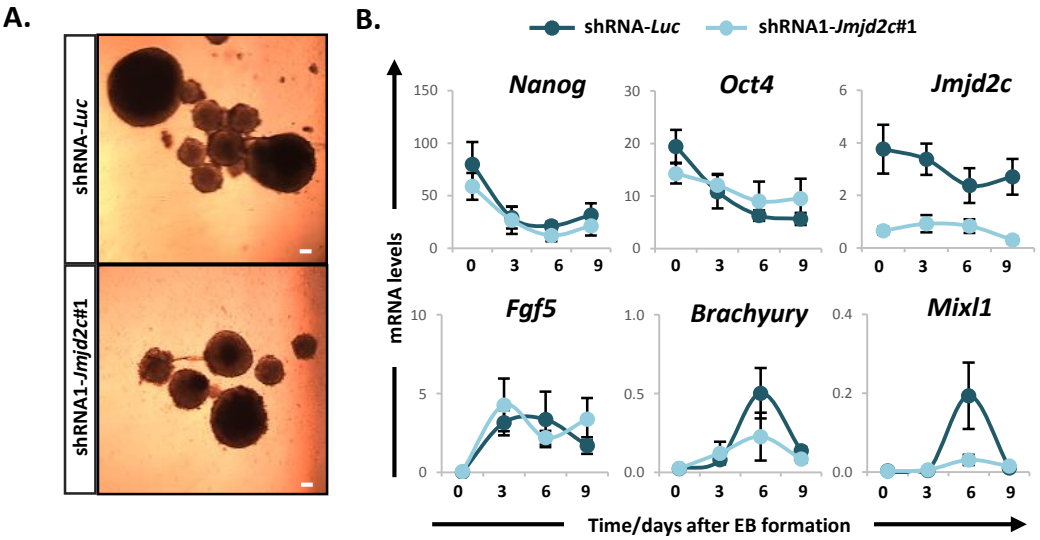


Figure 3.9 Embryoid body (EB) formation is compromised in stable *Jmjd2c*-knockdown ESCs

A. *Jmjd2c*-knockdown clone 1 and the control shRNA-Luc ESCs were cultured in suspension and induced to differentiate upon EBs formation in the absence of LIF and reduced serum levels. Representative phase-contrast images of day 9 EBs are shown. Bars, 100µm. **B.** Transcript levels for pluripotency-associated (*Nanog*, *Oct4*), *Jmjd2c*, and early differentiation markers for epiblast (*Fgf5*), mesoderm (*Brachyury, Mixl1*) lineages in shRNA1-Jmjd2c clone 1 and shRNA Luc EBs collected at day 0, 3, 6 and 9 of differentiation. Levels were multiplied by 100 and plotted as the average of at least 4 experiments, normalized to housekeeping genes (*S17* and *L19*), ±SEM. The results in this figure were produced by Donja Karimlou.

Moreover, the expression of the epiblast marker *Fgf5* peaked at day 3 in both EB cultures (Fig.3.9B). Strikingly, however, the induction of mesodermal markers *Brachyury* and *Mixl1* was greatly diminished in shRNA1-*Jmjd2c* EBs (Fig.3.9B), as seen in *Jmjd2c*-KO EBs (Fig.3.6B). Similarly to *Jmjd2c-KO* EBs (Fig.3.6B), the expression of *Oct4* and *Fgf5* was also retained higher levels at day 9 (Fig.3.9.B). Altogether, these results recapitulate the impaired differentiation ability observed in constitutive *Jmjd2c*-KO ESCs and reinforce the specific requirement for Jmjd2c at the onset of differentiation.

III.3 Summary and Conclusions

In summary, *Jmjd2c*-knockout ESC model was successfully generated via gene-trap and homologous recombination approaches. These *Jmjd2c*-depleted ESCs retain an undifferentiated state in culture with similar growth rate than their WT counterparts, and no obvious de-repression of differentiation-associated genes known to be occupied by PRC complexes. Despite retaining normal self-renewal ability, the transcriptional activity of pluripotency-associated factors in *Jmjd2c*-KO ESCs was found to be attenuated with factors such as *Nanog* and *Esrrb* being variably down-regulated in independent KO clones and cultures. Moreover, abolishing *Jmjd2c* expression resulted in an increase in the bulk levels of H3K9me2 and a milder increase in the levels of H3K9me3, highlighting Jmjd2c as a key H3K9me2 and H3K9me3 demethylase in ESCs, unlike previously reported. This effect in H3K9-methylation levels, together with the down-regulation of *Nanog* and *Esrrb*, could be rescued by re-introducing *Jmjd2c* in KO ESCs.

Importantly, the differentiation abilities of *Jmjd2c*-KO were here further examined and found to be consistently impaired upon embryoid body formation and upon lineage-specific induction by neuronal differentiation, indicating an important role for Jmjd2c in the proper execution of multi-lineage pathways. Interestingly, the partial retention of *Oct4*, *Nanog* and *Fgf5* expression upon EB formation in the absence of Jmjd2c points to a possible blockage at an early epiblast-like stage of ESC differentiation. Finally, the normal self-renewal capacity and reduced differentiation

potential, the identified key features of Jmjd2c-KO ESCs phenotype, could be mimicked using an alternative ESC model of stable *Jmjd2c* depletion via RNA interference.

The maintenance of an undifferentiated identity in *Jmjd2c*-depleted ESCs contrasted with previous reports showing that acute differentiation occurred upon transient knockdown of Jmjd2c with shRNA vectors. Here, the generation of stably depleted cell populations upon prolonged culture might have allowed for the selection of clones in which functional compensation by other demethylases is put in place, hence safeguarding robust expression of ESC-specific genes. Nevertheless, in both strategies (knockout and knockdown) an impaired ability to differentiate into somatic cell types was detected, uncovering a specific requirement for Jmjd2c in this process which cannot be functionally compensated in an *in vitro* system.

Chapter IV – Results: Identification of a pivotal role for Jmjd2c at the onset of differentiation

IV.1 Introduction

The experiments described in Chapter III demonstrate that, in the absence of Jmjd2c, ESCs grow normally, and furthermore sustain an undifferentiated state in prolonged culture. However, the ability of *Jmjd2c*-KO ESCs to differentiate towards epiblast-derived germ layers was severely compromised, as assessed upon embryoid body formation. Upon blastocyst injection, ICM-derived mouse ESCs contribute mostly to the epiblast of chimeric embryos and, although at a lesser frequency, can also colonize the yolk sac of the primitive endoderm (PrE) (Beddington and Robertson, 1989). Under self-renewing conditions, a subset of ESCs expresses PrE factors such as *Gata4, Gata6, Sox17* and *Hex* at variable levels (Canham et al., 2010; Niakan et al., 2010), indicative of priming for this extra-embryonic lineage. Moreover, ESCs can be converted into PrE-like cells upon LIF withdrawal, atRA treatment (Artus et al., 2010; Capo-Chichi et al., 2005) and forced expression of the PrE-specific transcription factors *Gata4* or *Gata6* (Fujikura et al., 2002; Shimosato et al., 2007; Wamaitha et al., 2015). In this chapter, to further explore the differentiation abilities of *Jmjd2c*-depleted ESCs, their capacity of generating extra-embryonic PrE-like cells was tested using different approaches. To additionally identify the specific stage at which embryonic (epiblast) differentiation is impeded in *Jmjd2c*-KO ESCs, these cells were stably converted into EpiSCs, and their phenotypic and functional properties further characterized.

IV.2 Results

4.2.1 *Jmjd2c*-KO ESCs swiftly adopt a primitive endoderm-like phenotype upon retinoic acid treatment

In order to test the ability of *Jmjd2c*-KO ESCs to differentiate towards a PrE-like state, cells were transiently treated with all-trans retinoic acid (atRA), previously shown to induce the expression of PrE markers (Capo-Chichi et al., 2005). For this, ESCs routinely cultured in serum/LIF conditions were cultured in monolayer over 4 days using reduced serum level and 1μM atRA, in the absence of LIF. The expression levels of *Nanog*, *Jmjd2c*, epiblast and PrE markers were quantified in RNA samples collected every 24h. As expected, *Nanog* expression was down-

regulated upon atRA treatment (Fig.4.1A), although substantial *Nanog* transcript levels were sustained 3-4 days post-treatment in WT, but not in *Jmjd2c*-KO cultures, possibly due to a small proportion of ESC retaining an undifferentiated state (Fig.4.1A). Conversely, the expression of PrE markers *Gata6*, *Gata4* and *Dab2* gradually increased upon atRA treatment, yet the kinetics and induction levels of these genes were strikingly higher in *Jmjd2c*-KO relative to WT cells (Fig.4.1A). Interestingly, and in contrast to the gradual decreased of *Jmjd2c* transcript observed upon EB differentiation, *Jmjd2c* expression was swiftly reduced in treated WT ESC cultures (Fig.4.1A). As a control, the expression of the epiblast-specific marker *Fgf5* was also assessed, which remained low in both cell populations (Fig.4.1A), reiterating that atRA-induced cells preferentially adopt an extra-embryonic identity. This observation was furthermore evidenced by dramatic changes in cell morphology with a loss of typical ESC colonies in culture and the emergence of large epithelial-like cells in both populations (Fig.4.1B). Additionally, WT and *Jmjd2c*-KO ESCs were fixed and stained for Oct4, Gata6 along with the cell membrane protein phalloidin prior to and after 4 days of atRA treatment (Fig.4.1C). This confirmed a complete loss of the pluripotency Oct4 protein in both treated cell populations, as opposed to untreated (Fig.4.1C). A positive staining for the PrE-associated Gata6 marker was then detected in most (if not all) cells upon atRA treatment (Fig.4.1C). Phalloidin staining here clearly highlighted cell morphology changes between self-renewing and atRA-treated cultures (Fig.4.1C).

These observations indicate that Jmjd2c is not required for atRA-induced PrE differentiation, and furthermore that *Jmjd2c*-depleted ESCs might have a higher propensity to differentiate towards the extra-embryonic lineages, as compared to WT ESCs.

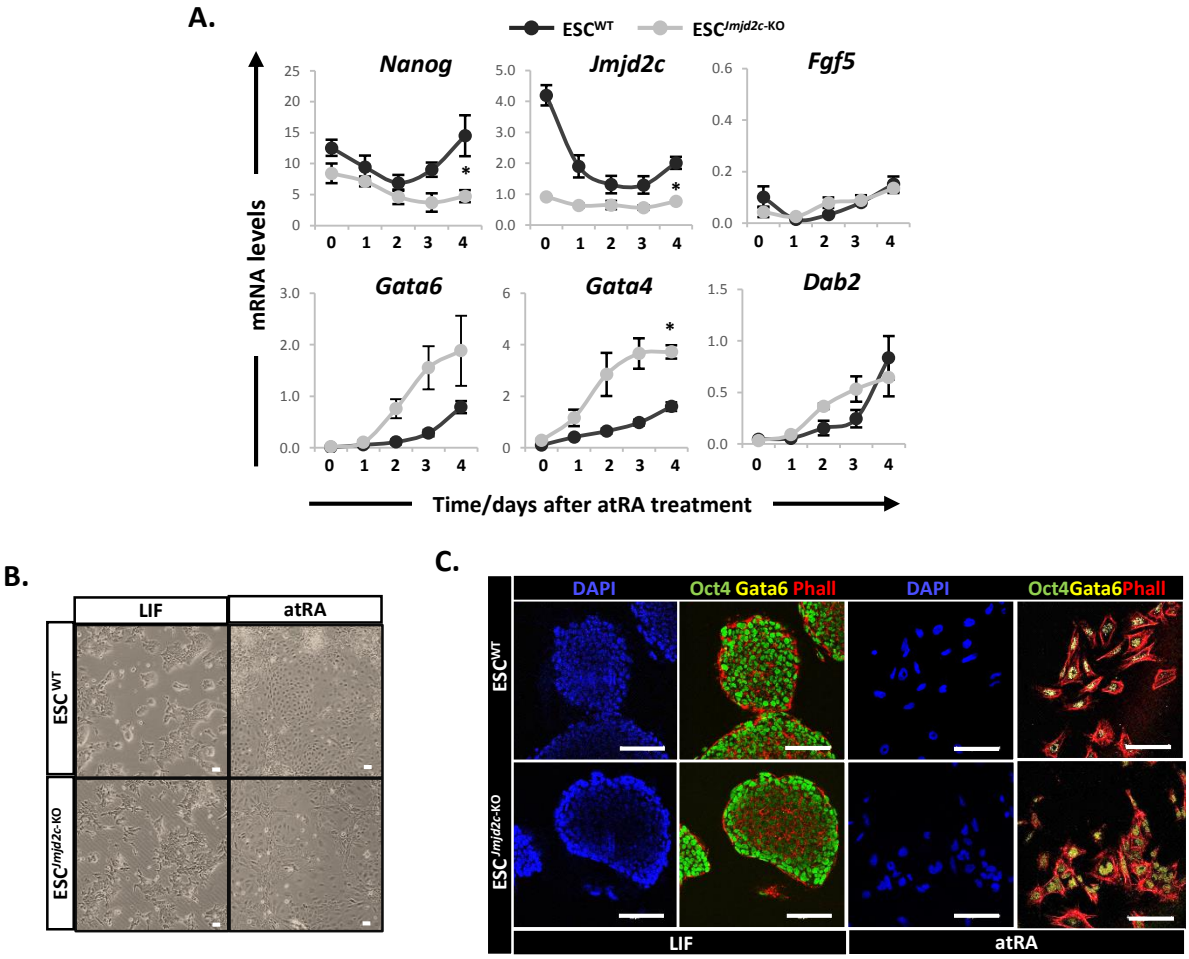


Figure 4.1 All-trans retinoic acid-induced differentiation to primitive endoderm (PrE)-like is enhanced in *Jmjd2c*-knockout ESCs

A. Transcript levels of *Nanog*, *Jmjd2c*, *Fgf5* (epiblast), and *Gata6*, *Gata4* and *Dab2* (PrE) markers upon treatment with 1μM all-trans retinoic acid (atRA) and LIF withdrawal in WT(ESC^{WT}) and KO (ESC$^{Jmjd2c\text{-}KO}$) clone E3 cultures for 4 days. Data were multiplied by 100, normalized to housekeeping genes (*S17* and *L19*) and represented as average of 3 experiments ±SEM.P<0.05 (*), Mann-Whitney *U* test at final time point. **B.** Phase-contrast images of ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clone E3 growing in serum/LIF conditions (LIF) and after 4 days of atRA treatment (atRA). Bars, 100μm. **C.** Immunofluorescence staining for Oct4 (green), Gata6 (yellow) and Phalloidin (red) in ESC^{WT} and ESC$^{Jmjd2c\text{-}KO}$ clone E3 growing in serum/LIF conditions (+LIF) and after 4 days of atRA treatment (atRA). DAPI staining was used to visualize nuclei. Bars, 100μm.

4.2.2 Jmjd2c is dispensable for ESC conversion into extra-embryonic endoderm stem cells

To further demonstrate the ability of *Jmjd2c*-depleted cells to stably adopt a PrE-like phenotype, a newly developed protocol for the establishment of self-renewing extra-embryonic endoderm stem (XEN) cells from ESCs was used (Niakan et al., 2013) During this process, ESCs were initially treated with low dose of atRA and Activin for 2 days (Fig.4.2A), after which cells were passaged onto irradiated MEFs (iMEFs), allowing the emergence of highly refractile and compact colonies in both cultures over the following 7 days (highlighted in Fig.4.2B). These colonies were carefully picked and re-plated onto iMEFs for an additional 2-day period, after which cells were expanded on gelatinised plates using standard XEN cell medium supplemented with serum and a combination of growth factors (Fgf2 and Heparin) (Niakan et al., 2013). Both WT and *Jmjd2c*-KO converted XEN (cXEN) cell populations acquired and faithfully retained a typical XEN-like morphology with a mixture of round refractile and flat epithelial-like cells over serial passages (Fig.4.2C). These cXEN cell populations, alongside embryo-derived XEN cells used as control, were stained for the XEN markers Dab2 and Gata4, showing strong positive signals in both WT and *Jmjd2c*-KO converted cultures (Fig.4.2C). Furthermore, expression levels of *Nanog*, *Jmjd2c* and several typical XEN cell markers (*Gata6*, *Gata4*, *Sox17*, *Lama1*, *Sparc* and *Sox7*), previously found to be induced upon XEN cell conversion (Cho et al., 2012), was quantified by RT-qPCR in ESCs and cXEN populations and compared to embryo-derived XEN cells (Fig.4.2D). As expected, *Nanog* transcript was undetectable in converted XEN cell lines and embryo-derived XEN cells (Fig.4.2D). The expression of *Jmjd2c* was reduced to background levels in WT cXEN as also seen in embryo-derived XEN cells, further suggesting that *Jmjd2c* is dispensable in this extra-embryonic lineage (Fig.4.2D). Lastly, the expression of all XEN cell markers analysed was successfully induced, and stably expressed in both cXEN cell populations to similar levels expressed by the IM8A-XEN cell line, with the exception of *Gata4*, which consistently harboured lower levels in WT and *Jmjd2c*-KO cXEN cells (Fig.4.2D).

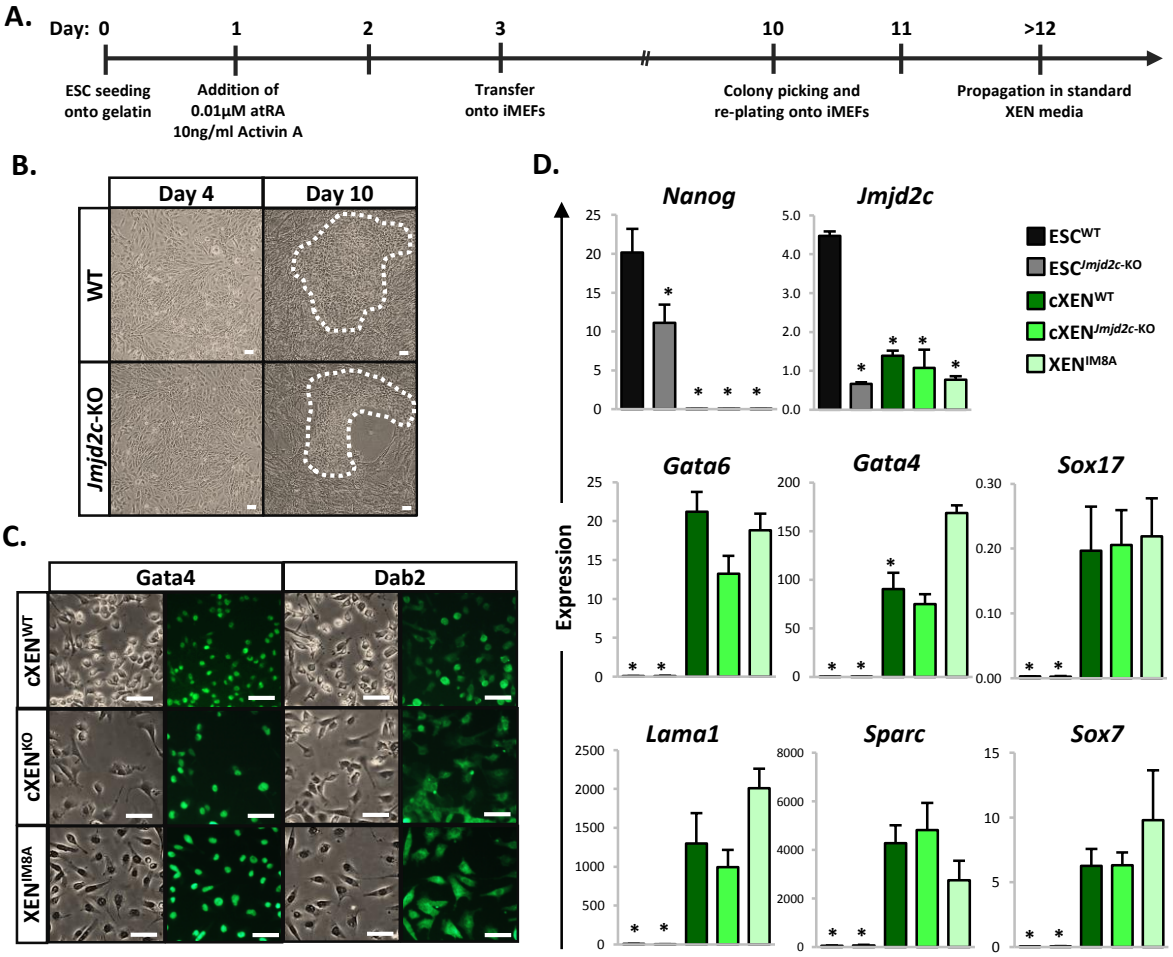


Figure 4.2 Conversion of *Jmjd2c-knockout* ESCs into extra-embryonic endoderm (XEN) stem cells

A. Timeline of ESC-to-cXEN derivation protocol (adapted from Cho et al., 2012). After day 12 established cell lines were routinely cultured on gelatinised plates and passaged about 15 times before staining and expression analysis. **B.** Representative phase-contrast images upon conversion of WT (ESC^{WT}) and KO ($ESC^{Jmjd2c\text{-}KO}$) clone E3 into XEN cells ($cXEN^{WT}$ and $cXEN^{KO}$, respectively) at the indicated time points. Emerging cXEN colonies are depicted with a white dotted line. **C.** Phase-contrast and fluorescent signals detected for the staining of Dab2 and Gata4 in stable ESC-derived $cXEN^{WT}$ and $cXEN^{KO}$ and an embryo-derived XEN cell line. Bars, 100µm. **D.** Transcript levels of *Nanog*, *Jmjd2c* and XEN-associated markers (*Gata6*, *Gata4, Sox17, Lama1, Sparc* and *Sox7*) in WT and *Jmjd2c*-KO ESCs, cXEN, and embryo-derived XEN cells. Data were normalized to housekeeping genes (*S17* and *L19*) and represented as average of 3 biological replicates ±SEM. P<0.05 (*), Mann-Whitney *U* test versus ESC^{WT} (*Nanog, Jmjd2c*) or *versus* XEN^{IM8A} (all the other markers).

Collectively, these experiments establish that *Jmjd2c*-KO ESCs, alike their WT counterparts, can efficiently differentiate towards the primitive endoderm lineage *in vitro*, where *Jmjd2c* is functionally expendable and expressed at greatly reduced levels relative to ESCs.

4.2.3 *Jmjd2c*-depleted ESCs adopt an immature epiblast stem cell identity upon activin and fibroblast growth factor induction

Experiments presented in Chapter III demonstrate that differentiation of *Jmjd2c*-KO ESCs is impaired upon EB formation and neuronal induction (Section 3.2.4). Moreover, *Jmjd2c*-KO ESCs were capable of transiting through an epiblast (EpiS)-like state upon neuronal induction, as seen in WT cultures, but failed to efficiently differentiate into neural-like progenitors (Fig.3.7B). Noticeably, *Jmjd2c*-depleted cells induced *Fgf5* upon EB induction. Yet this early epiblast marker remained inappropriately expressed, to some extent, in late EB differentiation time-points (Fig. 3.6B). Altogether, these observations point to a possible differentiation blockage at an EpiS-like state in the absence of Jmjd2c.

To explore this hypothesis, the ability of *Jmjd2c*-KO ESCs to adopt an EpiS-like phenotype was tested by culturing ESCs in serum-free media supplemented with activin and basic fibroblast growth factor (bFgf/Fgf2) (Guo et al., 2009). Transient ESC-to-EpiS-like conversion was initially optimized in WT ESCs by testing different plating conditions and coating as depicted in Fig.A-III (Appendix A). Similarly to neuronal progenitor induction (Section 3.2.4), this optimization revealed a high incidence of cell death when swapping media from serum-based to serum-free N2B27 conditions, with consequently fewer colonies surviving and/or exhibiting a clear EpiS-like morphology after 4 days of conversion (Fig.A-III – Appendix A). Moreover, stronger induction of *Fgf5* was detected when ESCs were initially grown in 2i/LIF compared to serum/LIF prior to conversion (Fig.A-III – Appendix A). Therefore, both WT and *Jmjd2c*-KO ESCs were first adapted to serum-free 2i/LIF for at least 7 days, and furthermore, plated on fibronectin-coated plates in these same conditions prior to EpiSC induction in the presence of activin and bFgf. The emergence of EpiS-like colonies occurred within a week under these conditions, after which cells were

serially propagated over 10 passages to generate stably converted EpiSC (cEpiSC) populations. Both WT and *Jmjd2c*-KO ESCs were equally capable of acquiring an EpiS-like cell morphology that closely resembled that of embryo-derived EpiSCs grown in the same conditions, and this morphology was retained in long-term culture (Fig.4.3A). Expression of Oct4 was also confirmed in both cEpiSC cultures by immunofluorescence staining, and compared to embryo-derived EpiSC cell line carrying an Oct4 eGFP-Ires-Puro (GiP) transgene (Ying et al., 2002), stably maintained under puromycin selection (Fig.4.3B). *Oct4* expression was also quantified and confirmed at the transcript level, together with *Nanog*, the epiblast markers *Fgf5*, *Otx2* and *Dnmt3b* in early time points of EpiSC induction and in converted cells, as well as the expression of differentiation markers *Brachyury/T* and *Foxa2* and, *Jmjd2c* itself (Fig.4.3C). The expression of Oct4 transcript was maintained at high levels in both cEpiSC cultures, although at a lower level in *Jmjd2c*-KO cEpiSC (Fig.4.3C). As expected, *Nanog* was down-regulated by day 2 and maintained low in stable cultures (Fig.4.3C). The Epiblast markers Fgf5 and Otx2 were greatly induced in the first 3 days of conversion, alongside with Dnmt3b (Fig.4.3C), previously linked with the known wave of DNA methylation that occurs between these 2 stages (Veillard et al., 2014).

Although conversion of *Jmjd2c*-null ESCs appeared successful, closer examination revealed clear phenotypic differences between WT and *Jmjd2c*-KO cEpiSCs. In culture, *Jmjd2c*-KO cEpiSCs colonies displayed a flat and round morphology, with well-defined borders, as opposed to the WT cEpiSC colonies which displayed some degree of metastability as revealed by spontaneous differentiation at the edge of the colonies. Notably, the initial burst of expression of early epiblast markers *Fgf5*, *Otx2* and *Dnmt3b* in WT ESCs occurred after two days of EpiSC induction, but this appeared consistently delayed in *Jmjd2c*-KO cultures, where expression peaked at day 3 rather than day 2 (Fig.4.3C). Moreover, in stable cultures the expression of these markers were stably maintained at lower levels, with the exception of *Fgf5*, which was maintained highly expressed in WT but not in *Jmjd2c*-KO cEpiSC (Fig.4.3C). Consistent with the observed morphological differences for WT and *Jmjd2c*-KO cEpiSC colonies, the mesoderm marker *Brachyury/T* and endoderm marker *Foxa2* acquired a low degree of expression in stable WT, but not KO, cEpiSC

cultures, concordant with the acquisition of some degree of gene priming (Fig.4.3C). Furthermore, the expression of *Jmjd2c* remained high throughout ES-to-EpiS-like conversion of WT ESCs, decreasing by half in stable cEpiSC cultures (Fig.4.3C).

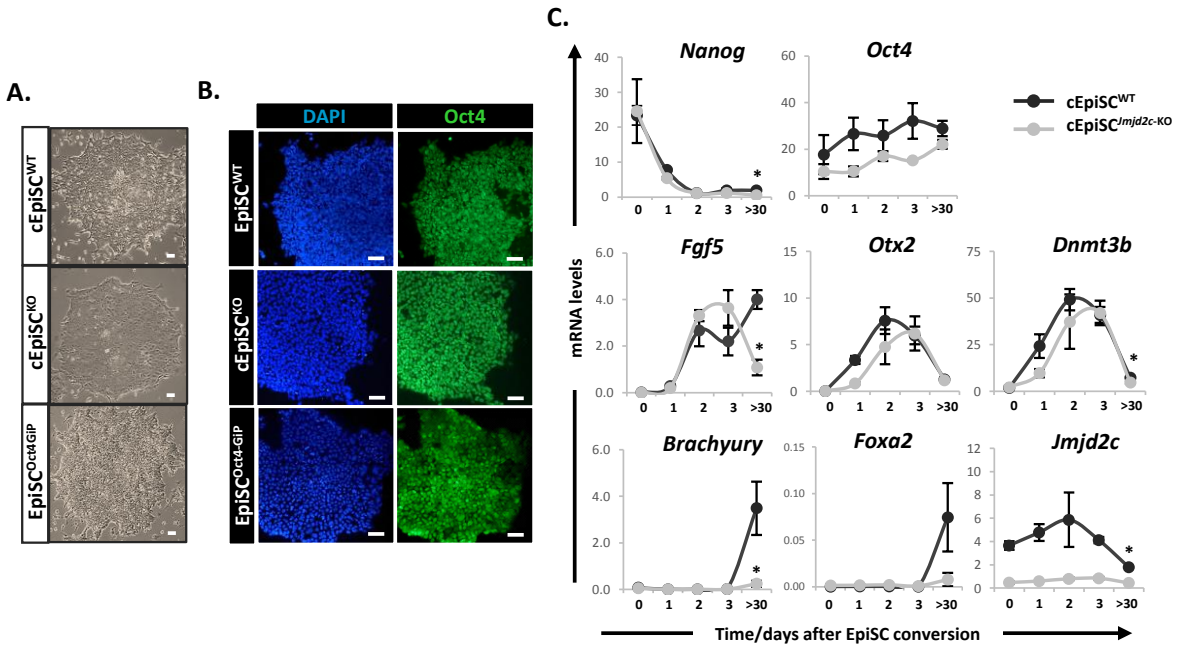


Figure 4.3 Conversion of *Jmjd2c*-knockout ESCs into epiblast stem cells (EpiSC) cells

A. Phase-contrast images of stable EpiSC cell lines derived from WT (cEpiSC^{WT}) and KO (ESC^{KO}) clone E3 ESCs, and embryo-derived EpiSC$^{Oct4\text{-}GiP}$. **B.** Fluorescent signals detected for the staining of Oct4 in stable cEpiSC^{WT} and cEpiSC^{KO} and embryo-derived EpiSC$^{Oct4\text{-}GiP}$ in which GFP expression is coupled with Oct4. DAPI staining was used to visualize nuclei. Bars, 100µm. **C.** Transcript levels of ESC (*Nanog*), Epiblast (*Oct4, Fgf5, Otx2*, and *Dnmt3b*), differentiation markers (*Brachyury* and *Foxa2*), and *Jmjd2c* during conversion of ESC^{WT} and ESC$^{Jmjd2\text{-}KO}$ into cEpiSCs (day 0 to 3) and in stably converted cEpiSCs (day >30). Data were multiplied by 100, normalized to housekeeping genes (*S17* and *L19*) and represented as average of at least 3 biological replicates ±SEM. P<0.05 (*), Mann-Whitney *U* test in stable cEpiSC.

Together, these results indicate that *Jmjd2c*-KO ESCs could successfully be converted into stable EpiSCs that, however, expressed lower levels of *Fgf5* and no expression of differentiation markers, most likely indicative of an immature EpiSC phenotype.

4.2.4 Skewed cell fate differentiation towards extra-embryonic visceral endoderm-like identity in *Jmjd2c*-KO cEpiSCs.

Generating stably converted EpiSCs (cEpiSCs) lines from both WT and *Jmjd2c*-KO ESCs formally establish that, in the absence of Jmjd2c, ESCs can transit to an EpiSC-like state *in vitro*. These cell lines were also crucial to further pinpoint the stage at which epiblast differentiation is blocked in *Jmjd2c*-depleted cells. In collaboration with Jennifer Harman and Dr. Helle Jørgensen (Cambridge University), the behavior of *Jmjd2c*-KO cEpiSCs was here studied upon induction of the mesoderm lineage. Firstly, cEpiSCs were cultured in chemically defined media (CDM)(Brons et al., 2007) supplemented with bFgf/Fgf2, BMP4 and the PI3K inhibitor LY294002 (FlyB conditions) for 36h to induce differentiation into mesoderm progenitors (MP) that typically express the primitive streak and early mesodermal marker *Brachyury* (Fig.4.4A). Secondly, FlyB-treated cEpiSCs were subsequently differentiated into distinct mesoderm subtypes: lateral plate mesoderm (LPM) cells with Fgf2 and BMP4 (FB40 conditions) and, paraxial mesoderm (PM) with Fgf2, Wnt3a, and inhibitors for both PI3K and BMP (LY294002 and LDN193189; FlyWLDN conditions) (Fig.4.4A). The efficiency of differentiation was scored by monitoring the emergence of Flk-1 (LPM) and Pdgfrα (PM) expressing (+) cells as detected by FACS analysis, alongside with embryo-derived EpiSCs differentiated under the same conditions (Fig.4.4B). In these control (EpiSC) cultures, differentiation towards LPM was generally more efficient than PM with over 55 % of Flk-1+ cells emerging under FB40 conditions, and about 28% Pdgfrα+ cells under FlyWLDN conditions (Fig.4.4B). Similar efficiency was obtained when differentiating WT cEpiSC, with about 46% Flk-1+ and 37% Pdgfrα+ emerging in FB40 and FlyWLDN, respectively (Fig.4.4B). Remarkably, the efficiency of differentiation was greatly reduced in *Jmjd2c*-KO cEpiSC, with approximately 14% and 6% cells staining positive for Flk-1 and Pdgfrα, respectively (Fig.4.5B). Moreover, the morphology adopted by WT and *Jmjd2c*-KO differentiated cEpiSCs was also strikingly different (Fig.4.4C). After 4 days of differentiation, WT cEpiSCs appeared as small elongated cells (Fig.4.4C), similarly to differentiated embryo-derived EpiSCs (Fig.A-IV – Appendix

A). On the contrary, differentiated *Jmjd2c*-KO cEpiSC formed a monolayer of adhesive and polarized epithelial-like cells regardless of culture conditions (Fig.4.4C).

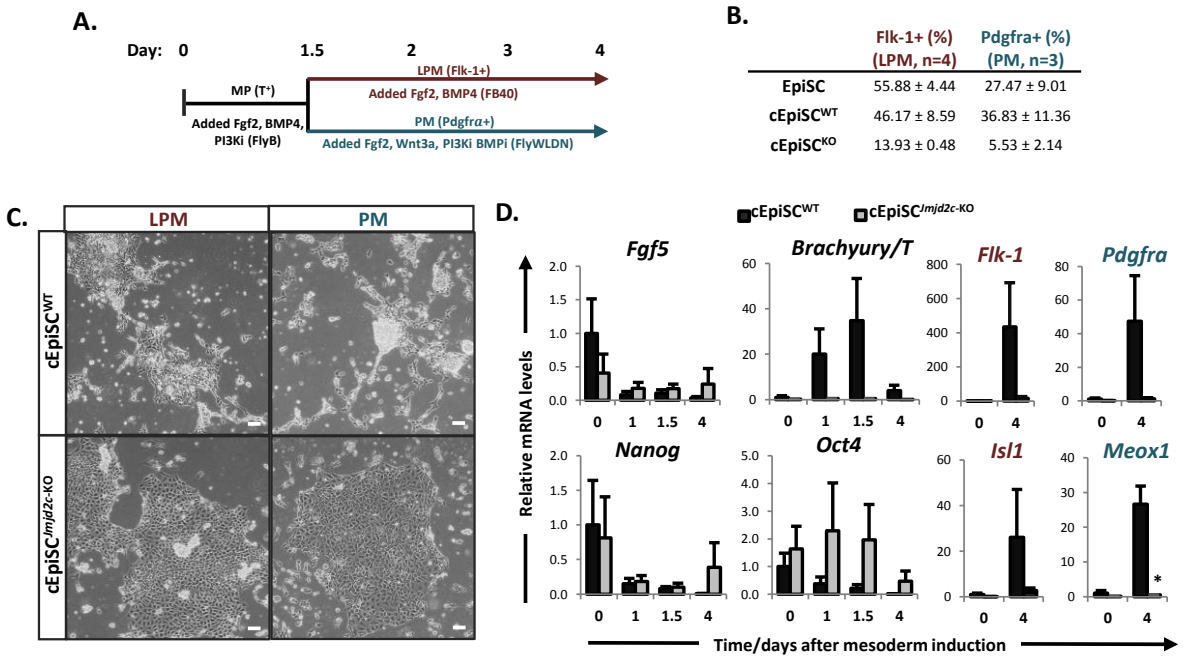


Figure 4.4 Induction of mesoderm is impaired in *Jmjd2c*-knockout cEpiSCs

A. Time line of Mesoderm induction. EpiSC are initially cultured in FlyB conditions to induce the emergence of early mesoderm progenitors (MP). Subsequently, cells are cultured for 3 days in either FB40 conditions to induce lateral plate mesoderm (LPM) differentiation, or in FlyWLDN to induce paraxial mesoderm (PM) differentiation. **B.** Average percentage (±SEM) of Flk-1 and Pdgfrα-positive cells detected by FACS analysis in embryo-derived EpiSC and WT (cEpiSC^{WT}) and KO (ESC$^{Jmjd2c\text{-}KO}$) cEpiSCs differentiated in FB40 (PM) and FlyWLDN (LPM) conditions, respectively. **C.** Representative phase-contrast images of day 4 differentiated WT and *Jmjd2c*-KO cEpiSCs populations, in FB40 (PM) or FlyWLDN (LPM) conditions. Bars, 100µm. **D.** Transcript levels of *Fgf5*, *Brachyury/T*, *Nanog, Oct4* in cEpiSC (day 0) differentiated with FlyB for 1 and 1.5 days, and FB40 at day 4. Transcript levels of *Flk-1*, and *Isl1* and *Pdgfra* and *Meox1*, in cEpiSC prior and after 4 day LPM and PM differentiation, respectively. Data were normalized to housekeeping genes (*Hmbs, Ywak* and *Gapdh*) and represented as average of at least 3 independent experiments ±SEM. P<0.05 (*), Mann-Whitney *U* test. Results in this figure were produced by Jennifer Harman.

Furthermore, the expression levels of epiblast and mesoderm markers were quantified in WT and *Jmjd2c*-KO cEpiSCs treated with FlyB conditions for 24h and 36h, and at the end-points

of LPM and PM differentiation pathways (Fig. 4.4D). Here, *Fgf5* and *Nanog* were promptly down-regulated upon early mesoderm progenitor induction in both cell populations (Fig.4.4D). However, while the expression of *Brachyury* peaked 36h after FlyB treatment in WT cEpiSCs, this failed to be induced in *Jmjd2c*-KO cEpiSCs (Fig.4.4D). Conversely, the expression of *Oct4* was down-regulated at the onset of differentiation in WT cEpiSCs, but was maintained at heterogeneous levels during FlyB treatment in *Jmjd2*c-KO cEpiSCs, indicating that these cells might not have acquired an early mesodermal progenitor phenotype (Fig.4.5D). Consequently, the induction of LPM markers (*Flk-1* and *Isl1*), and PM markers (*Pdgfra* and *Meox1*) was dramatically reduced upon FB40 and FlyWLDN treatment respectively in *Jmjd2c*-KO cEpiSCs relative to WT cEpiSCs and embryo-derived EpiSCs, as assessed by RT-qPCR (Fig.4.4D). These results demonstrate that mesodermal induction in cEpiSCs is inhibited at an early stage (MP) in the absence of Jmjd2c.

Notably, the epithelial-like cell morphology acquired by *Jmjd2c*-KO cEpiSCs upon mesodermal induction closely resembled that of either definitive endoderm (mesendoderm derivatives) or visceral endoderm (extra-embryonic derivatives) (Yasunaga et al., 2005). To clarify the identity of cells similarly emerging under FB40 and FlyWLDN treatments of *Jmjd2c*-KO cEpiSCs, the expression of visceral and parietal endoderm markers (*Lrp2* and *Sparc*) and definitive endoderm (*Foxa2* and *Cxcr4*) was further quantified at the end time-points of differentiation (Fig.4.5A). Remarkably, the expression of both *Lrp2* and *Sparc* was greatly induced in differentiated *Jmjd2c*-KO cEpiSC (Fig.4.5A). In contrast, *Foxa2* expression was poorly induced in differentiated WT cEpiSCs, as expected, and not induced in *Jmjd2c*-KO cEpiSC. The expression of the definitive endoderm marker *Cxcr4* was also poorly induced, although at similar levels in both cell populations (Fig4.5A), suggesting that *Jmjd2c*-KO cEpiSCs preferentially adopt an extra-embryonic endoderm-like phenotype. In line with this, differentiated *Jmjd2c*-KO cEpiScs also stained positive for the adherent junction marker E-Cadherin, clearly outlining the retained polarized epithelium-like cell morphology (Fig. 4.5B), as seen upon BMP4 treatment of XEN cells,

which was previously reported to induce visceral endoderm differentiation in extra-embryonic stem cells (Artus et al., 2012)

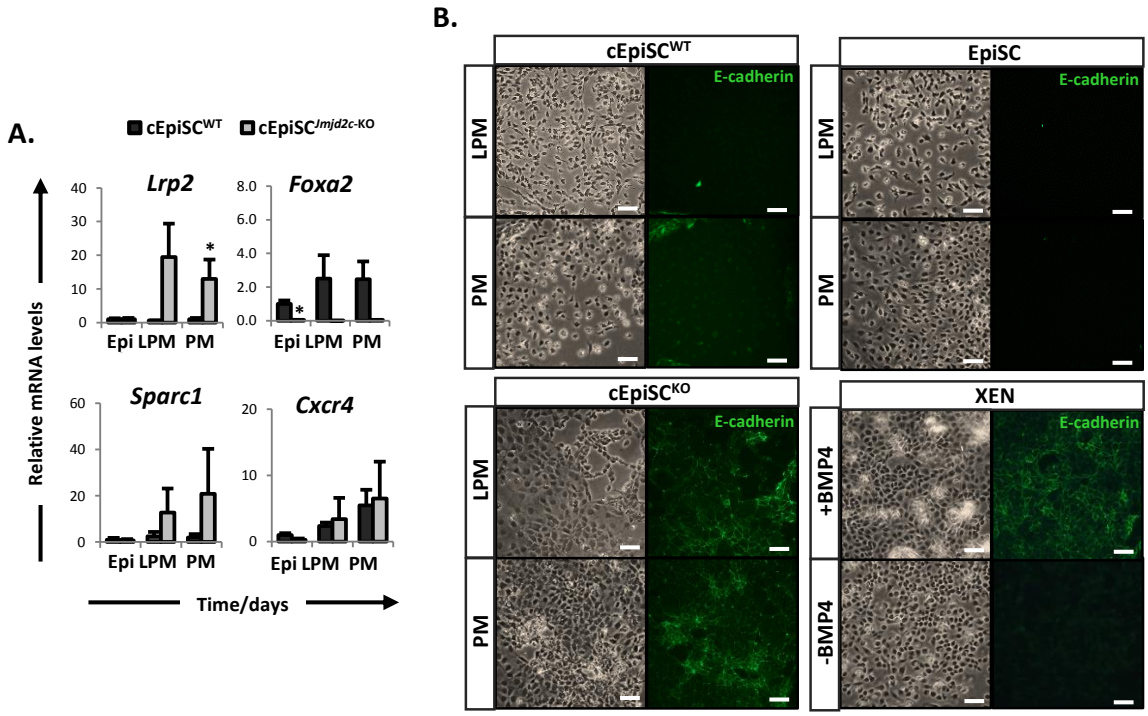


Figure 4.5 Characterization of skewed differentiation towards extra-embryonic endoderm-like phenotype of *Jmjd2c*-knockout cEpiSC

A. Transcript levels of visceral and parietal endoderm (*Lrp2, Sparc*) and endoderm (*Foxa2, Cxcr4*) in WT (cEpiSC^{WT}) and KO (cEpiSC$^{Jmjd2c\text{-}KO}$) cEpiSC before (day0, Epi) and after differentiation in FB40 conditions (LPM) and FlyWLDN condition (PM) at day 4. Data were normalized to housekeeping genes (*Hmbs, Ywak, Gpdh*) and represented as average of at least 3 independent experiments ±SEM. P<0.05 (*), Mann-Whitey *U* test. **B.** Phase-contrast images and fluorescent signals detected for the staining of E-cadherin in cEpiSC^{WT} , cEpiSC$^{Jmjd2c\text{-}KO}$ and embryo-derived EpiSC at day 4 of lateral plate mesoderm (LPM) and paraxial mesoderm (PM) differentiation, and in BMP4-treated and non-treated embryo-derived XEN cells. Bars, 100µm.

Taken together, these experiments demonstrate that, in the absence of Jmjd2c, epiblast stem cell fate determination is skewed towards an extra-embryonic endoderm-like identity as tested in permissive conditions for mesodermal differentiation.

4.2.5 Exploring the impact of *Jmjd2c* depletion in early cell fate commitment

Experiments described above unveiled that Jmjd2c is not required for the generation and maintenance of extra-embryonic endoderm and epiblast stem cells. Moreover, *Jmjd2c*-depleted cells exhibit a higher propensity to differentiate towards extra-embryonic endoderm lineages, as notably suggested by the rapid induction of PrE markers upon atRA treatment of *Jmjd2c*-KO ESCs (Fig.4.1), and by skewed differentiation into extra-embryonic endoderm-like cells upon cEpiSC mesoderm induction (Fig.4.5). To further test whether Jmjd2c might impact on the balance between PrE and epiblast lineage segregation, WT and *Jmjd2c*-KO ESCs were induced to differentiate upon hanging-drop embryoid body (HD-EB) formation. Unlike classic embryoid body assays, static suspension of EBs formed by low number of cells here facilitates the synchronised emergence of PrE- and epiblast-like populations in a 'salt-and-pepper' manner, which are subsequently segregated into an organized outer (PrE-like) layer and an inner (epiblast-like) population (Fig.4.6A), hence mimicking the early PrE/epiblast segregation events seen in the developing blastocyst (Lavial et al., 2012; Rula et al., 2007). A comprehensive optimization was initially preformed for drop volume, cell number, and timing of differentiation. About 500 ESCs were used per drop to induce the formation small aggregates in the presence of reduced serum level and in the absence of LIF, over 5 days. Expression of epiblast (*Oct4*, *Fgf5*) and PrE (*Gata4*, *Gata6*) markers was quantified over the course of the experiment by RT-qPCR (Fig.4.6B). Additionally, day 5 HD-EBs were fixed and stained for Oct4 and Gata4, in order to delineate the formation of the two inner (epiblast-like) and outer (PrE-like) EB layers, respectively (Fig.4.6C). Despite careful optimization this approach proved to be technically challenging and only two independent experiments out of 10 performed provided a sufficient amount of material for transcript quantification (Fig.4.6B). Here, overall down-regulation of *Oct4* was observed in WT HD-EBs, in contrast to *Jmjd2c*-KO HD-EBs that retained partial Oct4 expression at day 5 of differentiation, as previously found in late *Jmjd2c*-KO EBs (Section 3.2.4). *Fgf5* expression was induced from day 3 onwards and the early PrE marker *Gata6* highly up-regulated by day 5 in both cell populations. (Fig.4.6B). Interestingly, the expression of the late PrE

maker *Gata4* was also prominently induced but in *Jmjd2c*-KO HD-EBs only (Fig.4.6B). *Gata4* is a known Gata6 target gene, indicative of PrE maturation rather than specification (Frankenberg et al., 2011), hence suggesting that PrE maturation might be favoured or accelerated in HD-EBs in the absence of Jmjd2c. However, in the few successful Gata4 staining experiments performed on whole HD-EBs, no striking prevalence for the formation of an outer layer could be identified in *Jmjd2c*-KO HD-EBs (Fig.4.6C).

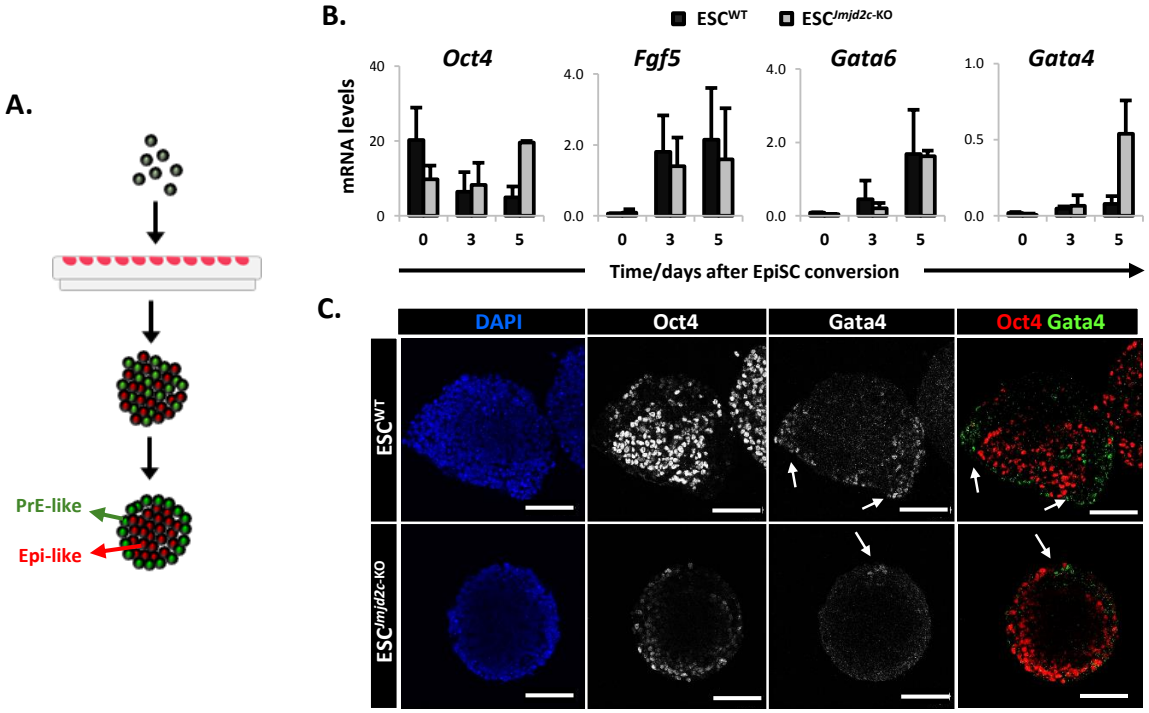


Figure 4.6 Induction of hanging-drop embryoid bodies (HD-EBs) is insufficient to uncover a bias in cell fate commitment in *Jmjd2c*-knockout ESCs

A. Schematic representation of hanging-drop embryoid body (HD-EB) assay, mimicking the early segregation events in the blastocyst. Formation of cell aggregates from low number of cells in suspension promotes the synchronised emergence of primitive-endoderm (PrE)- and epiblast(Epi)-like populations in a 'salt-and-pepper' manner, which are subsequently segregated into and organized outer (green) an inner (red) layer, respectively. **B.** Transcript levels of epiblast (*Oct4, Fgf5*) and PrE markers (*Gata6, Gata4*), during induction of HD-EBs from WT (ESC^{WT}) and KO ($ESC^{Jmjd2c\text{-}KO}$) ESCs. Data were normalized to housekeeping genes (*S17, L19*) and represented as average of 2 independent experiments ±SD. **C.** Immunofluorescence labelling of Oct4 and Gata6, for identification of Epi-like and PrE-like (arrows) layers respectively in HD-

EBs induced from WT and *Jmjd2c-KO* ESCs at day 6. Bars, 100µm. Results in **B** and the HD-EBs used to generate the results in **C** were produced by Donja Karimlou.

The unforeseen variability between experiments suggested that this approach might not be fully reliable to uncover a bias towards the extra-embryonic PrE in *Jmjd2c*-KO ESCs. Moreover, the quality of immunofluorescent staining could have been greatly improved by alternative approaches such as staining of sectioned agarose-embedded EBs (Lavial et al., 2012).

As a different approach to test whether *Jmjd2c*-KO ESCs might preferentially be allocated into the outer PrE-like layer as opposed to the inner epiblast-like layer, labelled WT and *Jmjd2c*-KO ESCs were mixed in a 50:50 ratio, and induced to form chimeric HD-EBs (Lavial et al., 2012). For this, WT ESCs were first labelled with a red fluorescence protein by transfection with a piggyBac vector carrying a CAG promoter driven dsRed transgene (CAG-dsRed*MST). Moreover, to ensure that the level of GFP produced by the genetrap cassette in *Jmjd2c*-KO ESCs would not be a limitation, and to guarantee that both ESC lines would be subjected to the same manipulations, KO ESCs were also transfected with a piggyBac vector carrying a CAG promoter driven GFP (CAG-GFP). These vectors were kindly provided by Dr Jennifer Nichols (Cambridge University). Thus, both cell populations were cultured for at least 4 passages under puromycin selection before induction of HD-EBs, and the levels of fluorescence produced by these vectors was confirmed under ESC culture conditions (Fig.4.7A). An equal ratio of WT (red) and KO (GFP) ESCs was mixed and 500 cells were used per drop in this assay (Fig.4.7A). The distribution of GFP and Red fluorescence was determined by confocal microscopy up to 7 days after HD-EB formation (Fig.4.7B). Both Green and Red signals appeared however, evenly distributed throughout the analysed aggregates, showing no striking evidence for preferential assignment of *Jmjd2c*-KO ESCs to the PrE layer (Fig.4.7B).

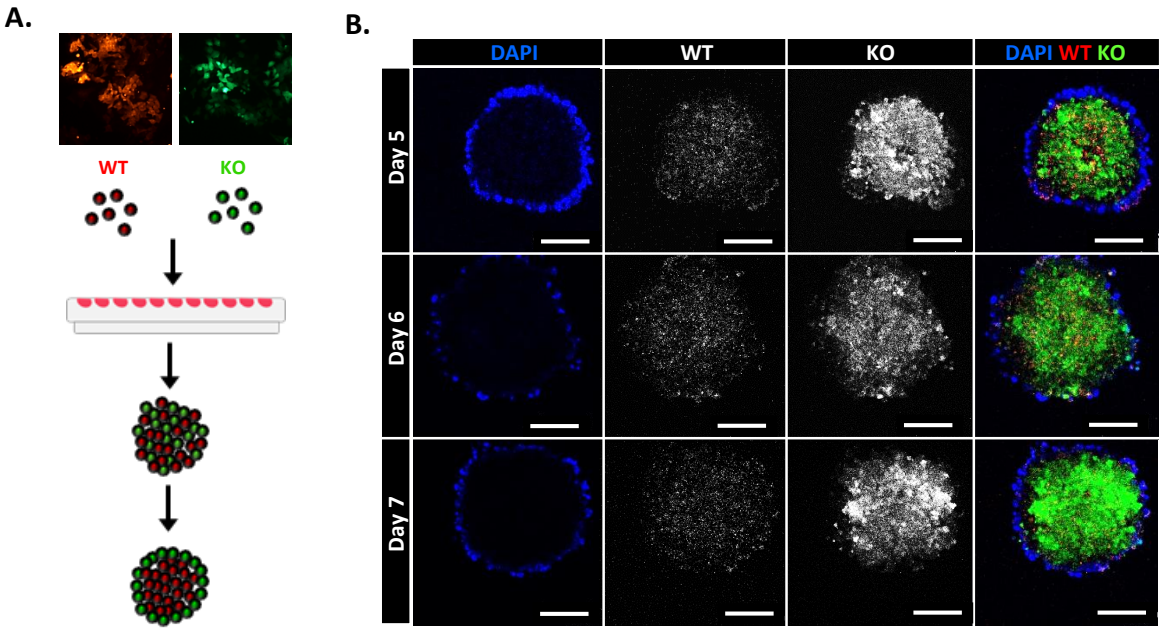


Figure 4.7 Generation of chimeric hanging-drop embryoid bodies

A. Schematic representation of generation chimeric hanging-drop embryoid body (HD-EB) with equal number of labelled WT (dsRed) and *Jmjd2c*-KO (GFP) ESCs. **B.** Immunofluorescence signals for WT (red) and *Jmjd2c*-KO (green) ESCs upon HD-EB formation at day 5, 6, and 7, showing no preferential location of labelled cells in the outer (PrE) and inner (epiblast) parts of chimeric HD-EBs. Bar, 100µm.

IV.3 Summary and Conclusions

Collectively, experiments described in Chapters III and IV uncover that *Jmjd2c*-KO ESCs fail to execute multi-lineage, epiblast-derived differentiation upon EB formation. Instead, these cells were found to preferentially adopt an extra-embryonic endoderm-like phenotype during the early steps of germ layer differentiation as further unveiled upon mesodermal induction of converted *Jmjd2c*-KO EpiSCs. This corroborates with a prominent and rapid induction of a PrE-like phenotype upon atRA treatment of *Jmjd2c*-depleted ESCs. Moreover, these cells also proved capable to stably convert into self-renewing XEN cell-like populations. However, no evidence could be obtained to suggest that Jmjd2c might directly impact in PrE and epiblast lineage segregation in the early embryo, as investigated here *in vitro* upon hanging-drop EB formation.

Moreover, the persistent expression of the Epiblast marker *Fgf5* upon EB induction, and the EpiS-like phenotype adopted upon neuronal progenitor induction, uncovered in Chapter III, suggested that *Jmjd2c*-KO differentiation was blocked at an epiblast-like stage. To validate this assumption, the ability of Jmjd2c-KO ESCs to be converted into EpiSCs was tested. Interestingly, converted EpiSCs were established, acquiring however an immature phenotype in the absence of Jmjd2c. Testing the differentiation capacity of this "immature" cell population upon mesoderm induction uncovered a differentiation bias towards an extra-embryonic endoderm like phenotype.

Altogether, these experiments established that Jmjd2c is strictly required for proper somatic differentiation after an epiblast-like stage, but dispensable for the emergence of an extra-embryonic phenotype. The molecular basis for the impaired differentiation and the lack of transcriptional lineage priming in mutant cEpiSCs will be further deciphered in the following Chapter V and VI.

Chapter V – Results: Exploring the genome-wide binding profile of Jmjd2c with identification of novel protein partners

V.1 Introduction

Results described in Chapters III and IV uncovered a key requirement for Jmjd2c at the onset of pluripotent stem cell differentiation into epiblast-derived lineages, as opposed to extra-embryonic endoderm. Impaired differentiation was evident from the inability of *Jmjd2c*-depleted cells to fully activate the required gene expression programs upon somatic lineage induction. Moreover, *Jmjd2c*-KO ESCs exhibited lessened and variable expression levels of pluripotency-associated factors, although this did not impact on their self-renewal ability. Additionally, an increase in global levels of H3K9 di- and tri- methylation was observed in the absence of Jmjd2c in ESCs, a set of histone modifications which, together with *de novo* DNA methylation, have been specifically linked with silencing of inappropriate gene expression in ESCs and in extra-embryonic lineages (Alder et al., 2010; Senner et al., 2012; Yuan et al., 2009). Altogether, these observations led to the hypothesis that Jmjd2c might play a key role in gene transcriptional regulation at the exit of pluripotency and/or during epiblast lineage priming, possibly through its H3K9-demethylase activity. In particular, it is here proposed that Jmjd2c, could prevent the acquisition of repressive epigenetic marks at the promoter regions of pluripotency and poised differentiation-associated genes in pluripotent stem cells.

During the course of this project, two reports employing chromatin immunoprecipitation (ChIP) coupled with high-throughput DNA sequencing (ChIP-seq) identified the genome-wide DNA-binding sites for Jmjd2c in ESCs, hence revealing that Jmjd2c is predominantly located at CpG-dense regions encompassing H3K4me3-rich TSS of both active and silent (bivalent) genes in these cells (Das et al., 2014; Pedersen et al., 2014). These findings corroborate with the proposed molecular basis for *Jmjd2c*-KO pluripotent stem cell differentiation defect. It is worth noting, however, that the two studies employed different ESC model systems and culture conditions – i.e. 2i/LIF (Pedersen et al., 2014) and serum/LIF (Das et al., 2014) known to distinctively support a naïve and primed ESC state, respectively (Ying et al., 2008). Although these reports commonly focused on the prevalent location of Jmjd2c at TSS gene regions, careful interrogation of the

deposited ChIP-seq data revealed striking differences in Jmjd2c distribution, with a higher number of non-TSS Jmjd2c DNA-binding sites being found in ESC "priming" (serum/LIF) (Das et al., 2014) as compared to 2i/LIF conditions (Pedersen et al., 2014). Nonetheless, the fact that different experimental approaches (Biotin-tagging versus endogenous IP) were used does not exclude that these differences might be due to different ChIP efficiencies rather than a direct effect of culture conditions and, hence, distinct ESC states on Jmjd2c genomic distribution.

Hence, to further explore this and test whether Jmjd2c DNA-binding sites are indeed re-distributed at the onset of differentiation, ChIP-seq analysis was performed in ESCs expressing tagged-Jmjd2c molecule under naïve (2i/LIF) and priming (serum/LIF) culture conditions as well as upon induction of an early epiblast-like state as previously described (Buecker et al., 2014). Candidate binding sites were then validated by ChiP-qPCR in ESCs that express wild-type or mutant tagged-Jmjd2c protein lacking the two Tudor domains required for Jmjd2c recognition and recruitment at H3K4 methylated sites (Pedersen et al., 2014). Moreover, in this Chapter, different approaches were employed to identify putative protein partners acting in complex with Jmjd2c in the control of gene transcription in ESCs.

V.2 Results

5.2.1 Generation of FLAG-tagged Jmjd2c expressing ESC lines

To map the genome-wide DNA-binding sites of Jmjd2c, tagged Jmjd2c expressing ESC lines were established due to the lack of commercially available ChIP grade antibodies. Three ESC lines were generated by stable transfection with a FLAG-FLAG-V5 tagged full-length Jmjd2c cDNA (FV-Jmjd2c-WT; Fig5.1A) in serum/LIF conditions. Empty FLAG-FLAG-V5 vector was kindly provided by Dr Raymond Poot (Erasmus MC, Rotterdam). Moreover, a stop codon was introduced at exon 18 to create a deletion of the two downstream Jmjdc2's Tudor domains (FV-Jmjd2c-ΔT; Fig5.1A), both essential for targeting Jmjd2c to the nucleus (Shin and Janknecht, 2007) and for recognition of H3K4me2/3 histone modifications *in vitro* and hence its recruitment at H3K4 methylated sites

in ESCs (Pedersen et al., 2014; Shin and Janknecht, 2007). In particular, it is here proposed that Jmjd2c, could prevent the acquisition of repressive epigenetic marks at the promoter regions of pluripotency and poised differentiation-associated genes in pluripotent stem cells. As a control, ESCs expressing an empty version of the same construct were also generated (Control; Fig.5.1A).

The expression of the FLAG epitope was firstly validated by immunofluorescence in stable pools of transfected ESCs with FV-Jmjd2c-WT and control vectors (Fig.5.1B). Moreover, both cell populations were found to grow normally, and similarly maintained high levels of Oct4 protein expression (Fig.5.1B, bottom panel). Individual FV-Jmjd2c-WT ESC clones were picked and expanded, and the expression of Jmjd2c and the V5 epitope was confirmed by western blot on six established clones (Fig.5.1C), revealing variable expression levels of this construct. To avoid clone-dependent bias, two clones expressing similar levels of Jmjd2c (approximately 4-fold higher than endogenous levels) were selected and used in subsequent experiments. Hence, FV-Jmjd2c-WT clone 8 was used in ChIP-seq, whereas clone 3 was used for further validation by ChIP-qPCR at a panel of selected Jmjd2c-bound sites (Section 5.2.7). Moreover, both FV-Jmjd2c-WT clones 3 and 8 were exploited for FLAG pull-down studies to be described in section 5.2.8 of this Chapter. Similarly, ESCs were transfected with the mutant vector FV-Jmjd2c-ΔT and individual clones picked and expanded. Here, the expression of a 108KDa truncated Jmjd2c fragment was confirmed by western blot (Fig.5.1D-E), and FV-Jmjd2c-ΔT clone1 selected to be used in ChIP-qPCR validation (Section 5.2.7).

To investigate Jmjd2c binding sites in naïve and differentiating conditions, FV-Jmjd2c-WT ESC clone 8, established and routinely cultured in serum/LIF, was first adapted to serum-free 2i/LIF conditions for at least 7 days prior to chromatin collection for ChIP-seq, as indicated by the acquisition of a dome shape-like colony morphology (Fig.5.2A). These 2i/LIF cultures were then induced to differentiate into early epiblast-like cells (EpiLCs) upon Activin, bFgf and KSOR (knockout serum replacement) treatment for 48h (Buecker et al., 2014), which similarly to cells grown in serum/LIF displayed a more flattened colony morphology (Fig.5.2A).

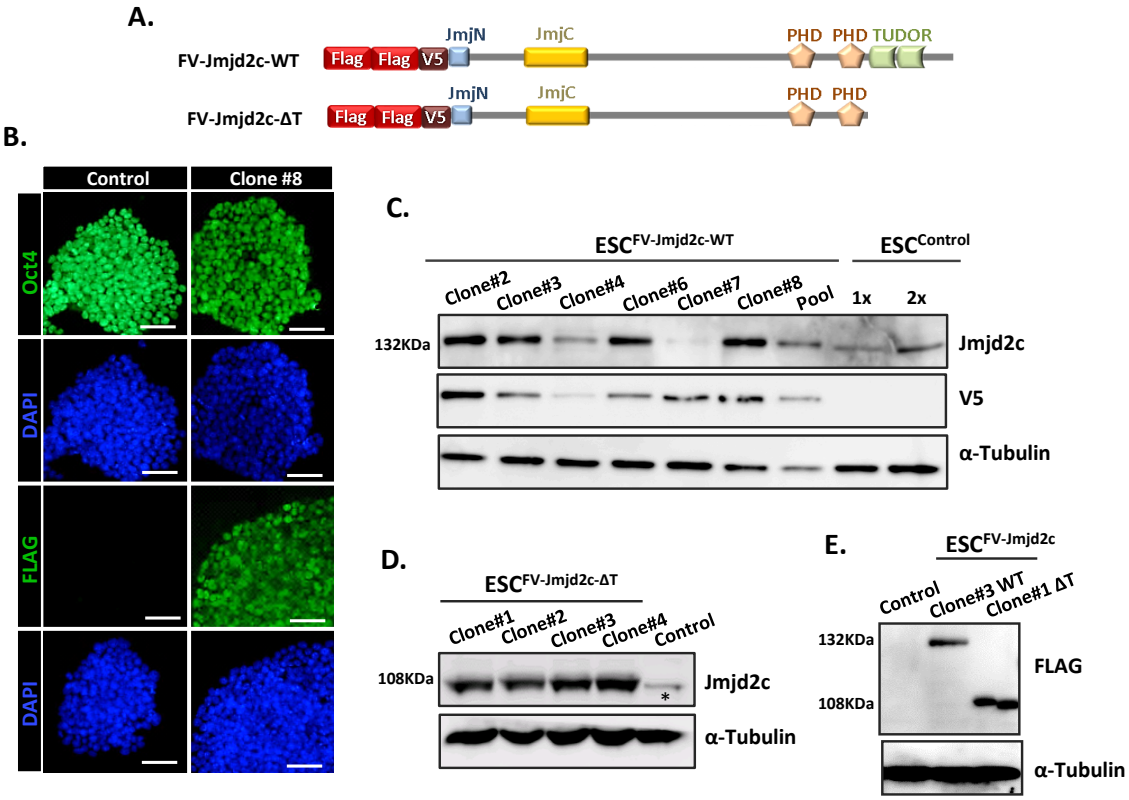


Figure 5.1 Characterization of FV-Jmjd2c expressing ESCs

A. Schematic representation of the generated FLAG(x2)-V5 tagged wild-type Jmjd2c construct (FV-Jmjd2c-WT) and mutated Jmjd2c construct where the Tudor domains were deleted (FV-Jmjd2c-ΔT). Both constructs were used to generate stable ESC lines. An empty version of the same construct (without Jmjd2c) was used to generate a control ESC line. **B.** Immunofluorescent labelling of DNA (DAPI), Oct4 protein and FLAG epitope in ESC cultures stably transfected with Control vector or F2V5-Jmjd2c-WT (Clone#8). Bars, 100μm. **C.** Western blot showing Jmjd2c protein levels, together with the expression of the V5 epitope in individual FV-Jmjd2c-WT ESC clones and pool of transfected cells. Endogenous levels of Jmjd2c were determined using protein lysates from control ESCs, which was loaded at two different amounts. Detection of α-Tubulin was used as a loading control. **D.** Western blot showing protein levels of mutant Jmjd2c in FV-Jmjd2c-ΔT ESC clones. Control ESC lysates were loaded as negative control, but an unspecific band was detected with similar size (*). α-Tubulin was used as a loading control. **E.** Western blot showing protein levels of the FLAG epitope in FV-Jmjd2c-WT clone 3 (132KDa) and FV-Jmjd2c-ΔT clone 1 (108KDa) ESC lines. α-Tubulin was used as a loading control.

These differentiation conditions were previously reported to promote genomic re-distribution of key transcription factors such as Oct4 and Otx2 at newly formed enhancer regions, and this was consistently accompanied by gene expression changes with the acquisition of a typical epiblast-like phenotype (Buecker et al., 2014). Accordingly, levels of induction of the Epiblast markers *Fgf5* and *Otx2* were found to be up-regulated in induced cultures, as compared to 2i/LIF or Serum/LIF ESCs (Fig.5.2.B). As expected, the levels of Nanog were conversely down-regulated upon induction of the EpiLC state, whereas the levels of Oct4 remained unchanged as expected (Fig.5.2B).

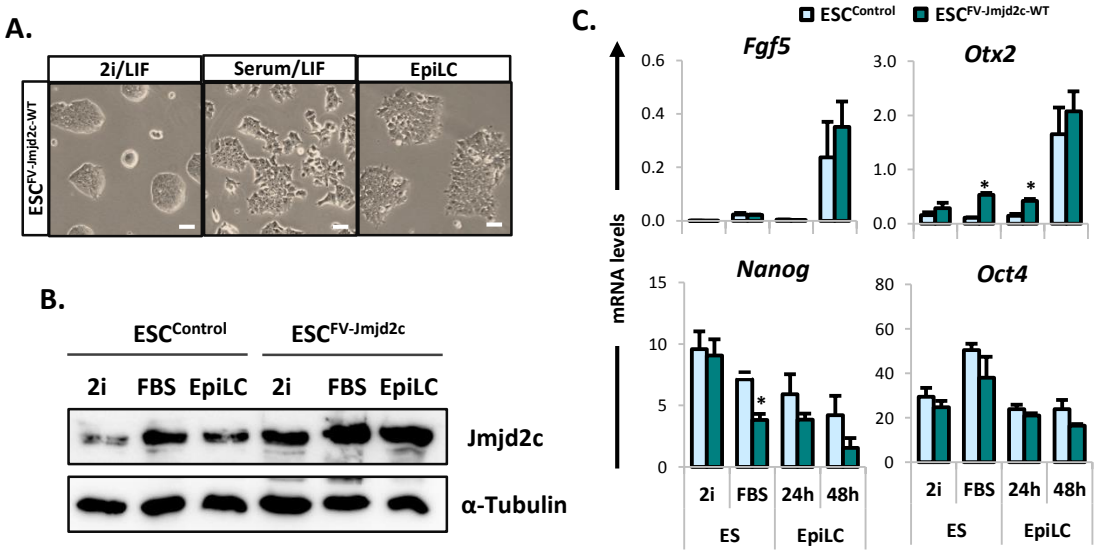


Figure 5.2 Generation of naïve and primed FV-Jmjd2c-WT ESC cultures

A. Morphology of FV-Jmjd2c-WT ESC clone 8 when routinely cultured in 2i/LIF, serum/LIF, and after 48h of Activin/Fgf2/KOSR induction (EpiLC). Bars, 100µm. **B.** Western blot showing Jmjd2c protein levels of in control and FV-Jmjd2c-WT ESCs cultured in 2i/LIF (2i), serum/LIF (FBS), and upon 48h in EpiLC (EpiLC) conditions. α-Tubulin was used as a loading control. **C.** Quantification of the mRNA levels of epiblast markers *Fgf5* and *Otx2*, and pluripotency markers *Nanog* and *Oct4*, in control and FV-Jmjd2c-WT ESCs routinely cultured in 2i/LIF, serum/LIF, and after 24h and 48h of EpiLC induction. Data were multiplied by 100 normalized to housekeeping genes (*S17* and *L19*) and represented as average of 3 biological replicates ±SEM. P<0.05 (*), Mann-Whitney *U* test.

Interestingly, Jmjd2c protein level appeared higher in ESC priming (serum/LIF) conditions than in naïve (2i/LIF) conditions, and remained high in epiblast-like cultures, as determined by western blot (Fig.5.2C). An increase of Jmjd2c expression with the acquisition of a primed pluripotent state reiterates a possible role for Jmjd2c at the onset of differentiation.

5.2.2 Genome-wide distribution of Jmjd2c during ESC priming for multi-lineage differentiation

Chromatin was extracted from 2 biological replicates of FV-Jmjd2c-WT ESC cultures respectively grown in serum/LIF and 2i/LIF conditions, or upon 48h induction into epiblast-like cells, and protein-bound DNA fragments immunoprecipitated with anti-FLAG antibody. To generate genome-wide maps of Jmjd2c-DNA interactions, DNA libraries were generated with the purified fragments and these were sequenced and analysed in collaboration with Dr Ignacio del-Vale Torres and Dr Kathy Niakan (NIMR Francis Crick Institute, London). Fragmented DNA samples were also collected prior to ChIP (input) for each culture condition, and used as background controls for sequencing. Reads were aligned to the reference mouse genome (mm9) and converted into a bigwig format for visualization in the Genome Browser. Given the high correlation score obtained for each pair of duplicates (Fig.A-V – Appendix A), data from both replicates was merged in BAM format, and then treated as one in order to identify significantly tagged-Jmjd2c enriched regions over inputs.

For this, a peak-finding algorithm was employed (MACS2) (Zhang et al., 2008) with a False Discovery Rate (FDR) lower than 0.0001, resulting in the identification of 20.377 and 45.485 peaks for Jmjd2c ChIP in 2i/LIF and serum/LIF respectively, and 22.985 peaks in differentiating (EpiLC) conditions. Peaks were annotated relative to the nearest TSS and the great majority annotated to protein-coding TSS in all conditions, with a small percentage annotated to non-coding RNAs, and under 1% to other such as small nuclear and nucleolar RNAs, ribosomal RNAs and pseudogenes (Fig.5.3A). Peaks were sorted into 2 groups according to their distance to the nearest TSS and assigned as “TSS peaks” if located within 1KB of the nearest TSS, and as “distal

peaks" if falling outside this interval. Concordant with previous reports, Jmjd2c peaks were consistently found within the defined TSS interval in 2i/LIF and serum LIF conditions (Fig.5.3B). However, the proportion of TSS peaks amongst total peaks detected was clearly altered between these conditions with a majority of Jmjd2c peaks in serum/LIF conditions being located outside the defined TSS interval and hence assigned as distal peaks (60% compared to 40% in 2i/LIF conditions). In contrast, analyzing the location of Jmjd2c peaks detected under differentiating (EpiLC) conditions did not revealed similar enrichment for distal sites (Fig. 5.3B), but instead retained a prevalence of TSS peaks as seen in 2i/LIF conditions.

Notably, a greater number of peaks identified in priming serum/LIF conditions, as opposed to naïve 2i/LIF conditions, supported the hypothesis that the binding pattern of Jmjd2c is changed upon priming. Hence, to unravel the extent of the re-distribution of Jmjd2c upon priming, the peaks identified in 2i/LIF conditions were separately compared to the priming and differentiating culture conditions here tested (Fig.5.3C). In both comparisons, a similar number of conserved TSS peaks were identified, with about 13,500 peaks overlapping between 2i/LIF versus serum/LIF, or 2i/LIF versus EpiLC (Fig.5.3C, left panel), which corresponded to the majority of the TSS peaks identified in all conditions. Furthermore, a proportion of distal peaks were conserved between 2i/LIF and serum/LIF conditions (4,700), and between 2i/LIF and EpiLC conditions (4,000), (Fig.5.3C). Notably, around 22,600 new peaks were acquired in the metastable serum/LIF state, while only about 2,700 peaks were acquired upon transition from a naive to an EpiLC state (Fig.5.3C, middle panel), highlighting the higher levels of re-distribution of Jmjd2c binding in the presence of serum. Comparison of the annotated genes associated with the total set of peaks identified in each condition revealed that the majority of peaks detected were in fact located within the same gene locus and, hence, associated with the same gene (Fig.5.3C). The notable different degree of re-distribution found in serum/LIF, as opposed to EpiLC conditions, proposed that the short-term conversion here applied might have been insufficient to fully promote a re-organization of Jmjd2c binding. Hence, the characterization of Jmjd2c peaks was focused on the serum/LIF dataset.

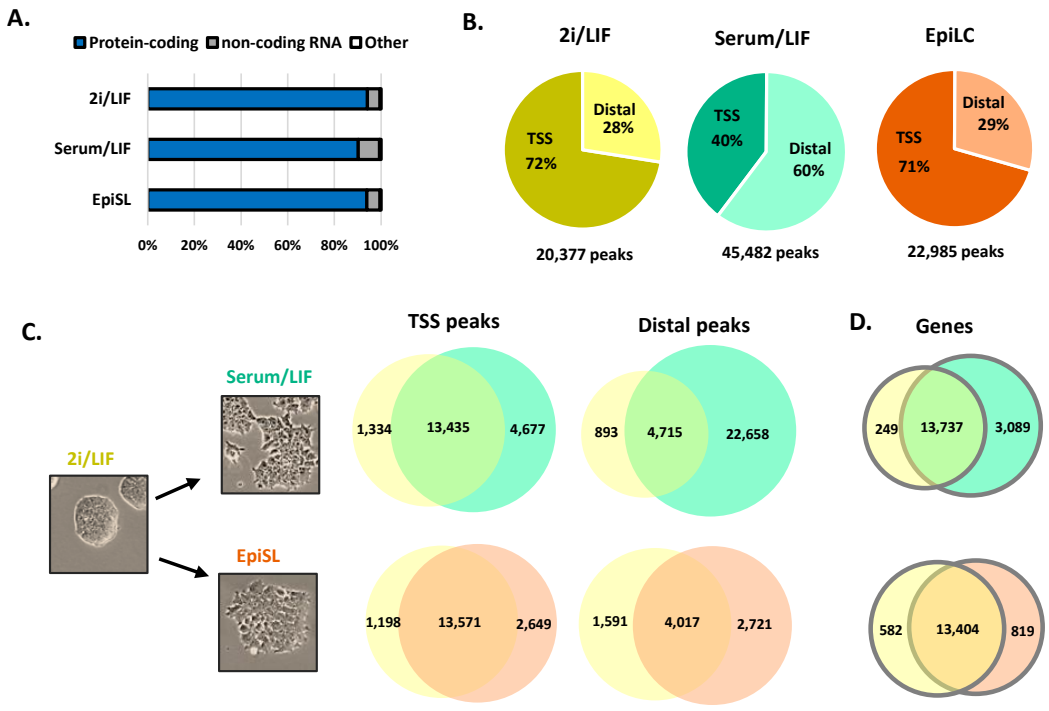


Figure 5.3 Genome-wide distribution of Jmjd2c in naïve and priming culture conditions

A. Class of genes annotated for the peaks identified in the three culture conditions tested (2i/LIF, serum/LIF, EpiLC). **B.** Percentage of peaks located within 1Kb of the TSS (TSS) or outside this interval (Distal) in the different culture conditions tested. **C.** Peak overlaps (minimum overlap = 1bp) for peaks identified as TSS or Distal between the three conditions: naïve (2i/LIF), primed (serum/LIF) and EpiLC (Activin/bFgf/KOSR). **D.** Overlap of gene names associated with all the peaks in 2i/LIF dataset against either serum/LIF or EpiLC datasets.

Taken together, these data suggest that Jmjd2c is preferentially found in the vicinity of a relatively large cohort of genes (about 13,500 genes). As previously reported, Jmjd2c peaks are consistently located at the TSS regions of these genes in both 2i/LIF and serum/LIF conditions. Yet additional enrichment peaks are observed in serum/LIF at distal sites that might overlap with distal regulatory regions associated with the same set of genes.

5.2.3 Jmjd2c TSS and distal peaks coincide with H3K4me3 and H3K4me1/me2 deposition

The association of Jmjd2c with CpG islands and H3K4me3-bound TSS was previously reported (Das et al., 2014; Pedersen et al., 2014). The identification of a large number of additional Jmjd2c distal binding sites in priming (serum/LIF) conditions, located more than 1kb apart from TSS regions, suggests that Jmjd2c might also be recruited at H3K4me1/me2 decorated sites possibly encompassing enhancer regulatory regions (Creyghton et al., 2010; Zentner et al., 2011). To determine whether distal Jmjd2c peaks preferentially coincide with H3K4me1/me2 enriched regions, rather than H3K4me3, heatmaps and average density plots were generated to examine the presence of each of these histone modifications across sorted Jmjd2c TSS and distal peaks identified in serum/LIF conditions (Fig.5.4A). As expected, a clear overlap between H3K4me3 enrichment at the Jmjd2c TSS peaks was observed (Fig.5.4A; green), with very little or no enrichment at distal regions (Fig.5.4A; blue). H3K4me2 marks occupied both sets of regions, whereas H3K4me1 deposition was selectively found at Jmjd2c-bound distal sites in contrast to TSS regions (Fig.5.4A). As hypothesized, these observations unveil novel association between Jmjd2c and H3K4me1/2-rich sites, in addition to H3K4me3-rich TSS as previously described.

5.2.4 G9a and Jmjd2c binding profiles intersect at Jmjd2c distal peaks in ESCs.

Jmdj2c was previously identified as a histone H3K9me2/me3 demethylase (Katoh and Katoh, 2004; Klose et al., 2006b; Whetstine et al., 2006). Accordingly, constitutive depletion of Jmjd2c protein in ESCs correlates with a marked increase in bulk levels of H3K9me2, with a milder effect on H3K9me3 (Section 3.2.3), proposing an active and prominent role for Jmjd2c in the removal of di-methyl groups from H3K9 residues in ESCs. Hence, it would be expected that no deposition of H3K9me2/me3 and/or enrichment for their cognate histone H3K9-methyltransferases are detected at Jmjd2c-bound regions in ESCs. To validate this, enrichment profiles for these histone modifications together with the H3K9me2 (G9a) and H3K9me3 (Suv39h1) methyltransferases were examined at Jmjd2c-bound TSS and distal regions using

average density plot analysis (Fig.5.4B). As hypothesized, no H3K9me2/3 enrichment was detected at either set of Jmjd2c peaks. Moreover, G9a was found mutually exclusive with Jmjd2c at TSS peaks (Fig.5.4B-C). This was in agreement with previously reported G9a enrichment patterns in ESCs where G9a peaks was seen closely surrounding, but not overlapping with the TSS of PRC2-bound genes (Mozzetta et al., 2014), strictly mirroring H3K9me2 enrichment patterns (Lienert et al., 2011).

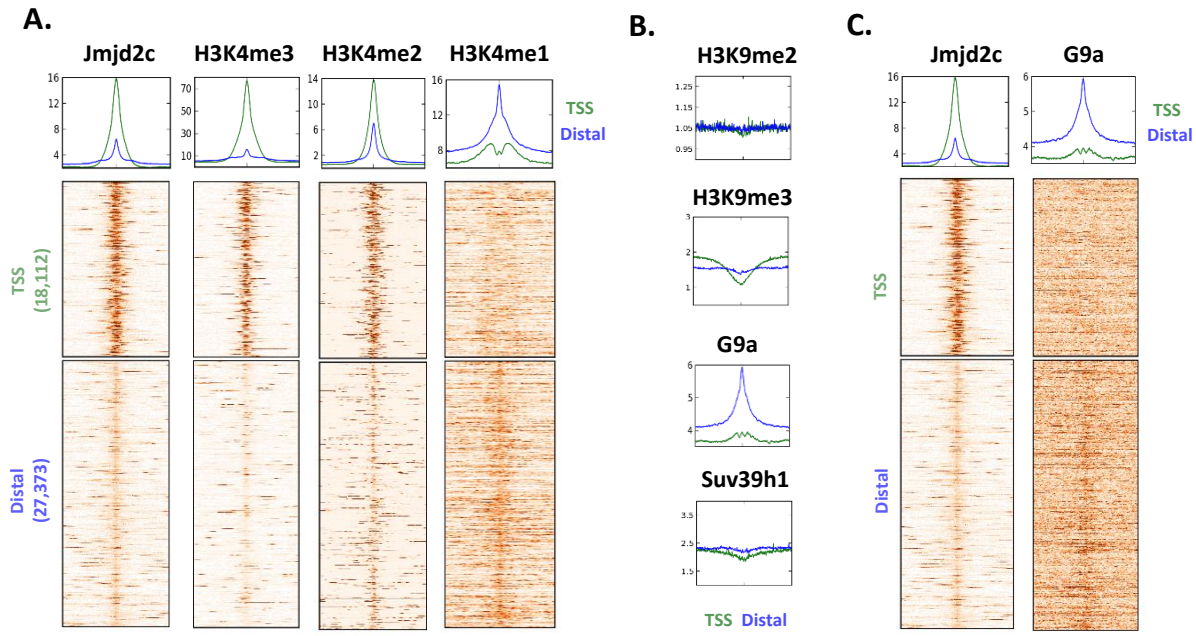


Figure 5.4 Jmjd2c associates with H3K4me1/2 and G9a at distal sites

A. Average density plots and heatmaps showing enrichment levels for Jmjd2c, H3K4me3, H3K4me2 and H3K4me1 across a ±5Kb window around the centre of Jmjd2c peaks regions defined as TSS (green) and distal (blue) in serum/LIF conditions. Scale was adjusted automatically to maximum intensity in each ChIP-seq data. **B.** Average density plots showing the enrichment levels of H3K9me2, H3K9me3, G9a, and Suvh39h1 across a 5Kb window around the centre of Jmjd2c peaks defined as TSS and Distal. Scale was adjusted for each ChIP-seq data individually. **C.** Average density plots and heatmaps for the level of enrichment of Jmjd2c and G9a cross a ±5Kb window around the centre of Jmjd2c TSS and Distal peaks. Scale was adjusted automatically to maximum intensity in each ChIP-seq data.

Unexpectedly, however, G9a and Jmjd2c were jointly detected at Jmjd2c-bound distal regions as further revealed using heat map analysis (Fig.5.4C). This contrasted with Suv39h1, which was

equally absent at both Jmjd2c-bound TSS and distal sites (Fig.5.5B, bottom panel). Taken together, these observations unexpectedly revealed that G9a and Jmjd2c are selectively co-enriched at H3K9me2-depleted, Jmj2c-bound distal sites in ESCs, suggesting that Jmjd2c might counteract the H3K9 methyltransferase activity of G9a at distal, H3K4me1-enriched regions rather than TSS gene regions.

5.2.5 Jmjd2c and G9a are co-enriched at both active and poised enhancers in ESCs

H3K4me1 was previously recognized as an epigenetic marker of enhancers, demarking both active enhancers when combined with H3K27ac, and inactive (poised) enhancers that might also carry the repressive H3K27me3 and/or H3K9me3 marks (Bonn et al., 2012; Creyghton et al., 2010; Zentner et al., 2011). To determine whether H3K4me1-enriched Jmjd2c distal peaks might encompass enhancer regions in ESCs, H3K27ac and p300 enrichment profiles were plotted against Jmjd2c at distal peaks (Fig.5.5A). While H3K27ac deposition was present at these sites, this mark distinctly delineated two groups of Jmjd2c-bound distal regions based on high (green) and low or absent (blue) H3K27ac enrichment levels (Fig.5.5A; left panel). Additionally, the p300 enzyme, which is responsible for depositing H3K27ac in ESCs, (Holmqvist and Mannervik, 2013) was consistently found enriched at Jmjd2c distal H3K27ac-high peaks, with lower enrichment at H3K27-low regions, as expected (Fig.5.5A; middle panel). Strikingly, the vast majority (85%) of H3K4me1-enriched Jmjd2c distal peaks showed low or no H3K27ac enrichment (Fig.5.5A; right panel), suggesting that Jmjd2c might preferentially bind to poised or inactive enhancers in ESCs when cultured in serum conditions. The remaining Jmjd2c peaks both exhibited high levels of H3K4me1 and H3K27ac enrichment, reiterating that Jmjd2c is also found at active ESC-specific enhancers presumably in conjunction with Jmjd2b as previously reported (Das et al., 2014).

The enrichment of Jmjd2c, together with G9a, and known enhancer-related factors including Oct4, Med1 and Smc1a (Kagey et al., 2010; Whyte et al., 2013) were furthermore plotted against the Jmjd2c-distal H3K27-high and low groups (Fig.5.5B). Remarkably, the level of G9a co-occupancy at distal regions clearly correlated with H3K27ac levels, with greater G9a binding

intensity being found at H3K27ac-high, and a reduced but still prominent enrichment detected at H3K27ac-low Jmjd2c sites (Fig.5.5B). A strong enrichment for Oct4 and the mediator component Med1 also correlated with high levels of H3K27ac at G9a and Jmjd2c co-bound sites, with a diminished but still noticeable binding being observed at nearly all H3K27ac-low sites (Fig5.5B). This suggested that G9a and Jmj2c might co-exist within protein complexes assembled at both active and poised enhancers in ESCs. In contrast to G9a, Oct4 and Med1, the cohesion factor Smc1 appeared similarly enriched at both H3K27ac-high and H3K27ac-low distal regions, although with a broader distribution in the first group (Fig.5.5B; right panel). Further validating these findings, Jmjd2c and G9a co-bound regions significantly overlapped with regions substantially occupied by Med1, p300 and/or Smc1a as represented in Venn diagrams (Fig.5.5C).

Taken together, these observations reveal that nearly all H3K4me1-enriched Jmjd2c distal sites are co-occupied by G9a alongside the enhancer-related factors Oct4, Med1 and Smc1, which show enrichment levels that strictly correlate with H3K27ac intensity. This, however, excludes Smc1a that appears evenly enriched at both active and poised Jmjd2c-bound enhancer regions.

5.2.6 Poised Jmjd2c-bound enhancers are associated with lineage-affiliated genes

Jmjd2c distal H3K27ac-high and low sites were next screened for known transcription factor binding motifs. Here, motifs for the ESC-associated nuclear receptors Esrrb and Nr5a2 were amongst the most significantly detected at H3K27ac highly enriched peaks (top panel; Fig.5.5D; Table B-VI – Appendix B). Conversely, motifs for a variety of differentiation-associated transcription factors such as Brachyury, Gata3 and Atoh1 were prominently identified at H3K27ac-low peaks, which also included the motif for the pioneer enhancer-associated factor Oct4 (bottom panel; Fig.5.5D, Table B-VII – Appendix B). Concordantly, GREAT gene ontology analysis revealed distinct biological processes associated with H3K27ac-high and H3K27ac-low Jmjd2c-bound regions. While H3K27ac highly enriched enhancers were mostly associated with genes involved in essential ESC functions (regulation of gene expression, metabolic process, translation) (Fig.5.5E; left panel), H3K27ac-low associated genes were predominantly involved

in developmental processes and differentiation pathways (germ layer formation and critical signalling pathways) (Fig.5.5E; right panel).

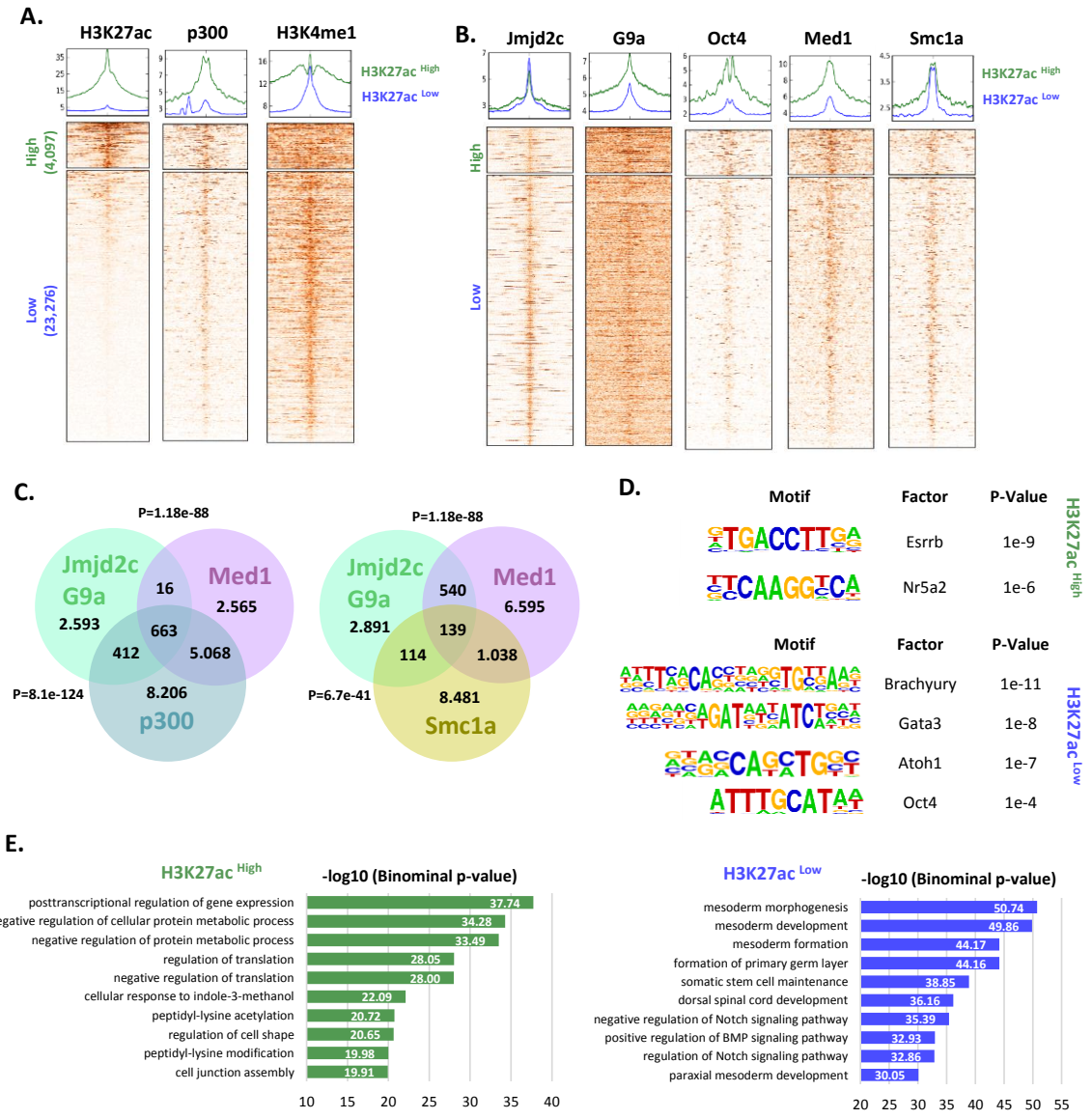


Figure 5.5 Jmjd2c and G9a co-occupy active and poised enhancers in ESCs

A. Jmjd2c distal peaks were sorted in H3K27ac-high and low with a kmeans cluster algorithm. Average density plots and heatmaps showing the level of enrichment for H3K27ac, p300 and H3K4me1 across a ±5Kb window around the centre of Jmjd2c distal peaks defined as high (green) and low (blue), in serum/LIF conditions. **B.** Average density plots and heatmaps showing the level of enrichment for Jmjd2c, G9a, Oct4, Med1 and Smc1a across a ±5Kb window around the centre of Jmjd2c distal peaks defined high (green) and low (blue). Heatmap and density plot scales were adjusted automatically to maximum intensity in each ChIP-seq data. **C.** Distal Jmjd2c peaks were intersected with and G9a peaks and overlapped with Med1 and

p300, or Med1 and Smc1a. A miminum overlap interval of 1bp and maximum gap of 500 bp were used. P-value obtained with a hypergeometric test. **D.** Representative motifs identified in H3K27ac-high peaks against low (top panel) or in H3K27ac-low peaks against high (bottom panel). **E.** Top 10 most significant biological functions of genes associated with H3K27ac-high (green) and H3K27ac-high low (blue) Jmjd2c peaks.

5.2.7 Validation of Jmjd2c binding at candidate regions

Mapping of genome-wide Jmjd2c binding sites here presented revealed that Jmjd2c is not only recruited at H3K4me3-rich TSSs of active and bivalent genes as previously reported (Das et al., 2014; Pedersen et al., 2014), but also at H3K4me1-rich, active and poised enhancer regions of the same genes, in ESCs grown in serum/LIF conditions. To further validate these findings at the locus level, ChIP-qPCR analysis was next performed in ESCs expressing wild type (FV-Jmjd2c-WT) or a mutant (FV-Jmjd2c-ΔT) form of Jmjd2c that lacks both Tudor domains (Section 5.2.1). For this, five candidate genes were selected encompassing two active genes in ESCs (*Esrrb, Klf4*), and three bivalently marked genes that are poised for future activation upon ESC differentiation (*Fgf5, Brachyury, Foxa2*). Levels of Jmjd2c binding were examined both at the TSS of these genes as well as at cognate enhancer regions previously identified based on Oct4, H3K27ac and H4K4me1 enrichment peaks in ESCs and induced epiblast-like (EpiLC) cells (Figure 5.6A) (Buecker et al., 2014). A high and low enrichment for Jmjd2c at the selected TSS and distal regions was detected in the generated ChIP-seq serum/LIF dataset (Fig.A-VI.A – Appendix A). Moreover, the enrichment of G9a together with the enhancer factors (Med1, p300, Oct4) and histone marks (H3K4me1 and H3K27ac) at these candidate regions was also verified in the publicly available ChIP-seq profiles (Fig.A-VI.A – Appendix A).

Jmjd2c enrichment was quantified following anti-FLAG ChIP on chromatin extracted from ESC clones that were stably transfected FV-Jmjd2c-WT, the mutated FV-Jmjd2c-ΔT, or the Control (empty) vector used to assess background levels (Section 5.2.1). PCR quantification of the recovered ChIP DNA revealed indeed a significant enrichment for Jmjd2c at the TSS of active and

bivalent genes, which is lost upon deletion of both Jmjd2c Tudor domains as expected (Fig. 5.6B, left panel). Furthermore, a relatively lower but significant enrichment for Jmjd2c was detected at the selected active and poised enhancer regions (Fig. 5.6B; right panel). Similarly, loss of Jmjd2c enrichment was observed in ESCs expressing FV-Jmjd2c-ΔT at the poised enhancer regions analyzed, confirming H3K4-dependent Jmjd2c DNA-binding. Unexpectedly, Jmjd2c appeared to be recruited to active enhancers in a Tudor-independent manner (Fig. 5.6B; right panel), possibly through interaction with Jmjd2b (Das et al., 2014).

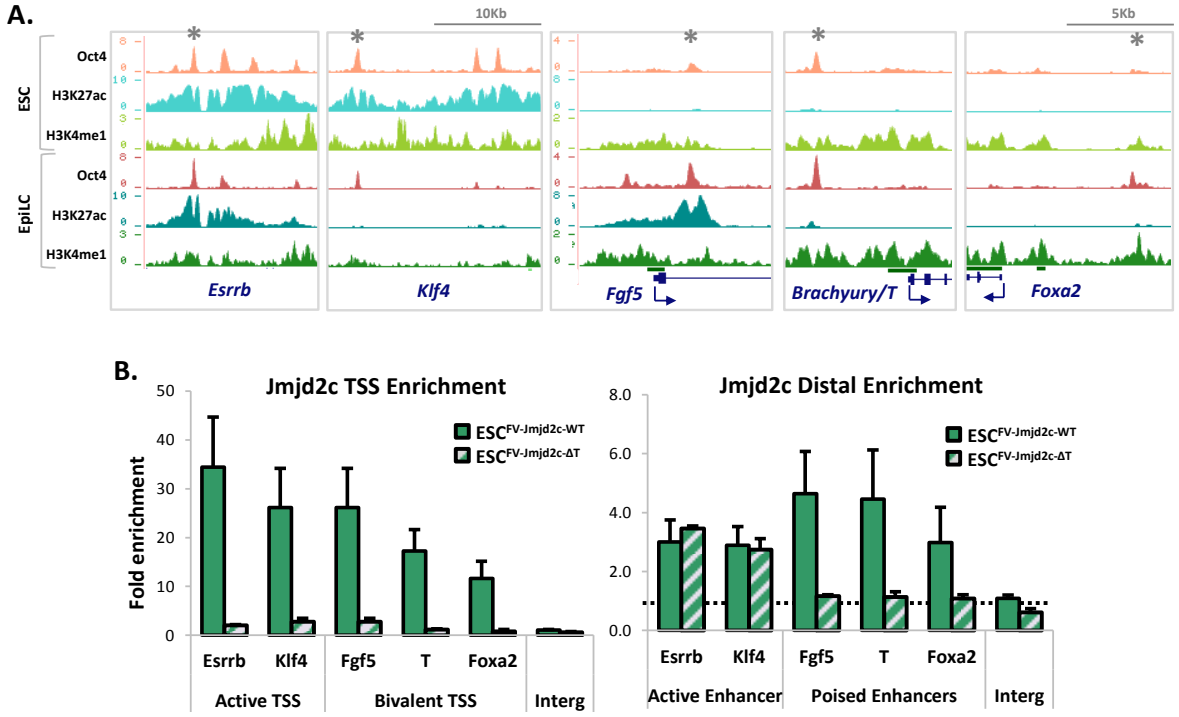


Figure 5.6 Jmjd2c binds to TSS of active and poised enhancers in ESCs

A. Snapshots showing the location of binding sites for Oct4 and enrichment levels for H3K27ac and H3K4me1 at enhancers in ESCs (2i/LIF) or in EpiLCs (Act/bFgf/KOSR), at *Esrrb, Klf4, Fgf5*, *Brachyury/T* and *Foxa2* loci, as previously described (Buecker et al., 2014). The position of selected enhancer primer pairs used in **B** are indicated by (*). **B.** Jmjd2c enrichment at the TSS of the selected active and bivalent genes, and enhancer regions as indicated. Levels were normalized to the control (empty vector) and represented as average ±SEM. An intergenic (Interg) region was included to confirm background levels. ChIP templates used in **C** were produced by Tony Bou-Kheir.

Overall, this analysis established that Jmjd2c is bound to poised enhancers of genes whose expression was here found to be impaired upon induction of differentiation (Section 3.2.4; Section 4.2.4), proposing a role for Jmjd2c in the activation of these enhancers.

5.2.8 Identification of Jmjd2c protein partners

As described above (Sections 5.2.4-5), Jmjd2c was found to co-localise with G9a alongside the known enhancer-associated factors p300, Oct4, Med1 and Smc1a at both active and poised enhancers, suggesting that Jmjd2c and G9a might be physically associated and/or be co-recruited at these sites within multi-protein complexes. To test whether Ga9 and Jmjd2c indeed interact and furthermore determine whether this interaction takes place in Med1-containing complexes, HA-tagged G9a was exogenously transfected in HeLa cells, and immunoprecipitated using anti-HA antibody in both soluble and chromatin nuclear fractions extracted from control and HA-G9a expressing cells, in collaboration with Dr Lauriane Fritsch (Paris VII University, France) (Fritsch et al., 2010). Immunoblotting using antibodies raised against Jmjd2c and Med1 as well as Cdyl – a control molecule known to interact with G9a, identified that G9a and Med1 are physically associated in the chromatin fractions of HA-FLAG-tagged (HF) G9a expressing cells (Fig.5.7A). Similarly, Jmjd2c was found to interact with G9a as detected in both soluble and chromatin fractions, though a stronger signal was seen in the later (Fig.5.7A).

Moreover, in nuclear extracts from ESC transfected with an empty vector (Control; Section 5.2.1), by immunoprecipitation of Med1 an interaction could be identified between Med1 and Jmjd2c, whereas in ESCs overexpressing FV-Jmjd2c-WT (Section 5.2.1) an interaction could also be detected between G9a, GLP and Jmjd2c (Fig.5.7B). Furthermore, in a separate experiment Med1 could also be pulled-down with G9a and with GLP in nuclear extracts from ESCs (Fig.5.7C), therefore establishing that in ESCs, Med1 can interact with Jmjd2c and with the G9a-GLP complex. Surprisingly, the antagonistic molecules G9a and Jmjd2c can also interact, proposing that Med1-Jmjd2c-G9a interaction might occur at the same complex.

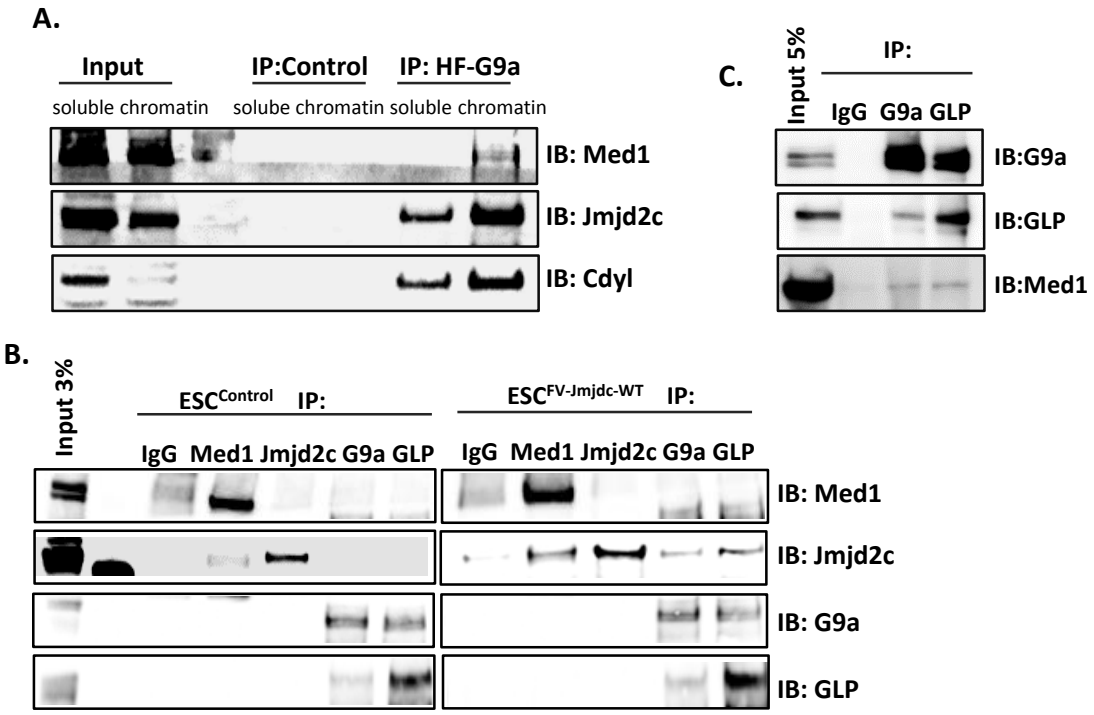


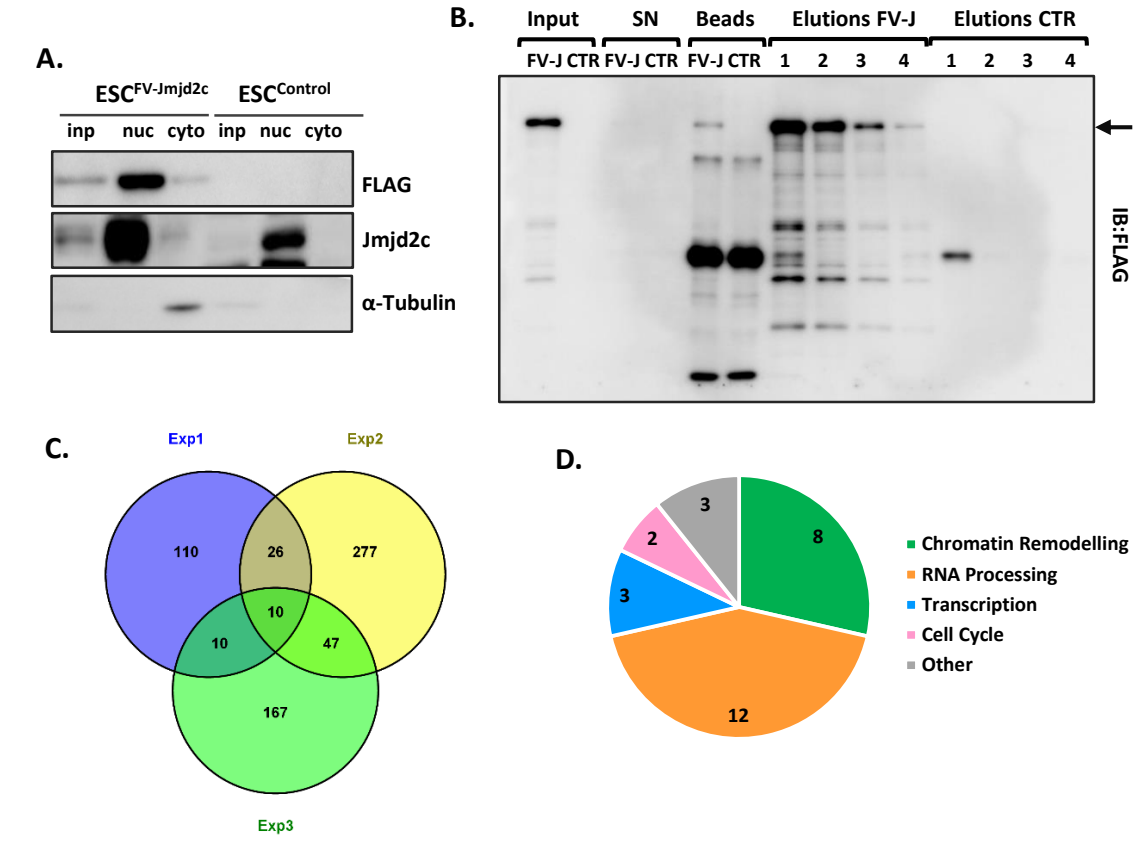
Figure 5.7 Jmjd2c interacts with G9a-GLP and Med1

A. HF-G9a was immunoprecipitated in in both the soluble and chromatin nuclear fractions of HeLa cells. Pulled proteins were visualised by western blot with antibodies for Med1, Jmjd2c and Cdyl. HeLa cells expressing an empty HF vector were used as control. **B.** Endogenous Med1, G9a and GLP were immunoprecipitated in nuclear extracts from ESCs expressing an empty vector (control) or the FV-Jmjd2c-WT construct. Interactions were visualized with antibodies against Med1, Jmjd2c, G9a and GLP. **C.** Endogenous G9a and GLP were immunoprecipitated in nuclear extracts of ESCs expressing an empty vector and immunobloted with G9a, GLP and Med1 antibodies. Results in **A** and **B** were generated by Dr Lauriane Fritsch.

To further explore the composition of Jmjd2c-containing protein complexes, FLAG pull-down was performed in two independently derived FV-Jmjd2c-WT clones (Section 5.2.1). Here, nuclear extracts rich in Jmjd2c, were isolated from both FV-Jmjd2c-WT and control ESC cultures (Fig.5.8A), and incubated with FLAG-bound sepharose beads. Bound protein complexes were then eluted from the beads, and the efficiency of pull-down and elution tested by anti-FLAG immunoblotting. This showed no residual FLAG signal in the supernatant (SN) fractions, and little residual signal associated with stripped beads. In contrast, a strong signal was detected in the

eluted fractions from FV-Jmjd2c-WT cells but not from control cells, as expected (Fig.5.8B). Eluted proteins were further precipitated, separated on a polyacrilamyde gel and then analysed by mass spectrometry in collaboration with Dr Raymond Poot (Erasmus MC, Rotterdam).

Identified proteins with a Mascot Score higher that 40, and at least 3-fold higher in FV-Jmjd2c-WT samples as compared to control pull-down in each of the three independent experiments performed were considered as positive hits (Fig.5.8C). The proteins identified in each experiment were overlapped, resulting in 93 proteins being identified as positive hits in at least two out of three experiments (Fig.5.8C). Cytoskeletal and cytoplasmic proteins were removed from this list, resulting in a total of 28 proteins with a nuclear location and roles associated with chromatin remodelling, RNA processing, cell cycle, DNA repair, mitosis and apoptosis (Fig.5.8D; Fig. A-VII – Appendix A). Chromatin modifier proteins identified included Jmjd2b - another member of the Jmjd2 histone demethylase family (Das et al., 2014), and Jmjd1a – a known functional H3K9me2 demethylase in ESCs (Loh et al., 2007), L3mbtl2 – a PRC1 protein found at H3K9me2 regions (Qin et al., 2012b), and the H3K9me2 methyltransferase GLP/Ehmt1 (Shinkai and Tachibana, 2011) (Fig.5.8E). However, none of the enhancer-related factors Med1, p300, Oct4, Smc1a, could be identified with this approach, possibly due to the complexity and size of enhancer-promoter bound complexes. Yet, Jmjd2c pull-down could identified novel interactors and the biological function of these complexes in ESCs remains to be further investigated.



E.

	Experiment 1						Experiment 2						Experiment 3					
	FV-Jmjd2c-cl#3			Control			FV-Jmjd2c-cl#3			Control			FV-Jmjd2c-cl#8			Control		
symbol	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.
Kdm4c	**5.95**	**3150**	**43**				**4.31**	**2315**	**39**				**12.4**	**3735**	**50**			
Kdm3a	0.02	101	1				0.02	48	1				0.19	360	6	0.08	173	3
Kdm4b	0.21	189	3										0.29	187	4			
*Ehmt1	0.13	238	4				0.08	178	5					65	2		68	1
L3mbtl2	0.04	87	1							0.04	67	2	0.04	54	1			
*Mbd3	8.61	830	12	9.76	930	13	0.76	245	5				0.57	116	3	0.4	98	2
Ppp1cc	1.14	395	7	1.84	592	9	0.21	122	4				0.61	234	4			
Prmt5	0.16	64	1				0.28	244	6									
Arid3a		110	1							0.12	106	3	0.82	102	2	0.25		
Leo1							0.15	182	5				0.1	68	2			
Rbpj							0.07	66	2				0.07	45	1			

Figure 5.8 Identification of Jmjd2c binding partners by pull down and mass spectrometry

A. Efficiency of nuclear extraction was confirmed by immunobloting nuclear (nuc) and cytoplasmic (cyto) extracts with antibodies against FLAG, Jmjd2c and α-tubulin. **B.** Elution of FLAG-bound complexes was validated by immunobloting the IP supernatant (SN), boiled beads (Beads), and eluted fractions (Elutions) with FLAG antibody. **C.** Overlap between positive hits identified in each of the three independent experiments. Mascot Score cut-off = 40; hits were considered positive if the score was 3-fold higher in FLAG compared to control samples. **D.** Function of proteins identified as positive hits in 2 out of 3 experiments with known nuclear locations. **E.** List of Jmjd2c-bound proteins with functions associated with chromatin remodelling and transcription.

V.3 Summary and Conclusions

Altogether, the experiments herein described unveil a novel link between Jmjd2c binding and enhancer regions in ESCs grown in 'priming' culture conditions. In this chapter genome-wide Jmjd2c DNA-binding sites were characterized using ChIP-sequencing in ESCs cultured in naïve (2i/LIF) and priming (serum/LIF) conditions. As previously reported, Jmjd2c occupies H3K4me3-rich TSS sites encompassing both active and poised/bivalent genes (Das et al., 2014; Pedersen et al., 2014). Here, this association appears consistently conserved in ESCs regardless of culture conditions used. Remarkably, alteration of the Jmjd2c binding profile is observed in serum/LIF, with the acquisition of additional Jmjd2c peaks at non-TSS (distal) sites. The approach here utilised involved overexpression of a FLAG-tagged version of Jmjd2c in WT ESCs, which could have resulted in non-specific binding of Jmjd2c. However, comparison of the resulting ChIP-seq profiles with a previously published dataset using an antibody against the endogenous protein in 2i/LIF conditions (Pedersen et al., 2014), revealed highly similar binding profiles (Fig.A-VI.B – Appendix A). This suggests that FLAG-Jmjd2c binding sites identified might indeed mimic the endogenous binding pattern.

Surprisingly, distal Jmjd2c-bound regions are co-occupied by G9a as further validating by co-immunoprecipitation analysis, and overlap with several known enhancer-associated factors such as p300, Oct4, Mediator/Med1 and Cohesin/Smc1a in ESCs (Kagey et al., 2010; Whyte et al., 2013). Moreover, these regions are distinctively decorated with the enhancer-associated H3K4me1 and H3K27ac marks together or H3K4me1 alone, most likely representing active and poised regulatory enhancer regions in ESCs. Taken together, these data point to a novel and selective role for Jmjd2c in delineating and/or regulating poised enhancer regions associated with somatic, lineage-affiliated genes in ESCs as further investigating upon Jmjd2c depletion in the following result Chapter.

Chapter VI – Results: Dissecting the molecular mechanism underlying the differentiation defect in *Jmjd2c*-depleted cells

VI.1 Introduction

PRC-bound, differentiation-associated genes in ESCs harbor bivalent chromatin structures with inactive (H3K27me3) and active (H3K4me2/3) marks at their promoter regions, which have been proposed to prevent untimely gene expression while keeping these genes primed for future activation (or repression) upon differentiation (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006b). Similarly, enhancer regions of differentiation-associated genes can also occur in a primed state typically harboring the active H3K4me1 mark alone, or in combination with repressive marks (H3K27me3, H3K9me3) (Bonn et al., 2012; Creyghton et al., 2010; Zentner et al., 2011). Upon differentiation, chromatin re-organization at regulatory regions ensures that the appropriate gene expression programs are activated, repressed or remain poised (Chen and Dent, 2014). In particular, acquisition of H3K9me3 marks and/or *de novo* DNA methylation at gene promoters are two mechanisms previously shown to be responsible for silencing inappropriate gene expression in extra-embryonic and neuronal lineages (Alder et al., 2010; Meissner et al., 2008; Senner et al., 2012).

In Chapter III, a noticeable increase in the bulk level of H3K9me2, and a moderate increase in H3K9me3, were detected in the absence of Jmjd2c in ESCs. Furthermore, when challenged to differentiate, *Jmjd2c*-knockout (KO) ESCs and cEpiSCs showed impaired abilities to up-regulate the appropriate gene expression programs upon somatic differentiation. Instead, *Jmjd2c*-depleted cells showed skewed differentiation abilities towards extra-embryonic endoderm-like phenotypes as tested upon retinoic acid induction in ESCs and under mesodermal permissive conditions in cEpiSCs. In Chapter IV, detailed analysis of Jmjd2c genome-wide DNA-binding sites showed strong binding at the TSS of active and poised (bivalent) genes as previously reported in two independent studies (Das et al., 2014; Pedersen et al., 2014). This was here confirmed in both naïve (2i/LIF) and primed (serum/LIF) ESC states, yet additional distal peaks were readily detected nearby the same set of genes in "priming" culture conditions only. Moreover, further bioinformatics analysis of newly generated (Chapter IV) and publicly available ChIP-seq datasets uncovered a co-enrichment for Jmjd2c and the H3K9me2 methyltransferase G9a at enhancer

regions that either harbor an active (H3K4me1, H3K27ac) or a poised (H3K4me1) epigenetic signature. These Jmjd2c and G9a co-bound regions furthermore overlap with known enhancer-associated factors including Oct4, p300, Med1 and Smc1a (Kagey et al., 2010; Whyte et al., 2013), as seen in primed (serum/LIF) ESC cultures.

Altogether, these observations propose a possible role for Jmjd2c in the protection of poised regulatory regions - promoters and/or enhancers - from acquisition of G9a-mediated H3K9me2 and *de novo* DNA methylation. To test this hypothesis, enrichment levels of H3K9me2, DNA methylation, G9a, and enhancer-related factors were assessed in *Jmjd2c*-depleted ESCs and cEpiSCs. Here, candidate gene loci were selected to include representative examples of epiblast (*Fgf5*), mesoderm (*Brachyury*) and endoderm (*Foxa2*) markers, who were found to be bound by Jmjd2c (Chapter V) and whose expression was found to be consistently impaired upon embryoid body formation, epiblast and/or mesodermal induction in *Jmjd2c*-KO pluripotent stem cells (Chapter III).

VI.2 Results

6.2.1 *Jmjd2c*-depleted ESCs do not harbour aberrant H3K9me2 and DNA methylation at gene regulatory regions.

H3K9me2 is enriched at low levels across large genomic regions, being mutually exclusive with H3K4me2-rich sites (Lienert et al., 2011) commonly associated with TSS but also enhancer regions(Chen and Dent, 2014) . To test whether, in the absence of Jmjd2, the silencing marks H3K9me2 and DNA methylation might be aberrantly acquired at Jmjd2c-bound TSS and enhancer regions upon lineage priming, H3K9me2 ChIP-qPCR analyses were performed using chromatin lysates extracted from both WT and *Jmjd2c*-KO ESCs and cEpiSCs. Candidate locations included poised TSS and enhancers associated with *Fgf5*, *Brachyury* and *Foxa2* loci where Jmjd2c binding was validated by ChiP-qPCR in FV-Jmjd2c-ESCs (Section 5.2.7). Primers amplifying control regions closely located to each of the three analyzed loci were additionally selected based on the absence of Jmjd2c binding at these sites as identified in ChIP-seq datasets. Moreover, the levels of

H3K9me2 at *Magea2* promoter region (Tachibana et al., 2002) were also used as a positive control. In these experiments, no increase in H3K9me2 enrichments was identified both at TSSs and enhancer regions of these genes in *Jmjd2c*-KO ESC and cEpiSCs relative to WT cells (Fig.6.1A). Moreover, a consistent decrease in H3K9me2 enrichment levels at these TSS and enhancer regions relative to their flanking regions was observed (Fig.6.1A), in agreement with mutually exclusive deposition of H3K4me2 and H3K9me2 marks at regulatory regions(Lienert et al., 2011). In contrast, H3K9me2 was substantially enriched at *Magea2* TSS in all ESC and cEpiSC populations examined, indicative of efficient ChIP recovery (Fig.6.1A, right panel). These data point to a role for Jmjd2c at bound target sites independent of its histone demethylase activity.

Additionally, the levels of DNA methylation were examined and quantified at the same genomic regions using a previously described enzymatic assay (Ficz et al., 2011). Briefly, pre-treated WT and *Jmjd2c*-KO cell genomic DNA was digested with either MspI or HpaII, which recognize the same sequence (CCGG). As a measure for total DNA amount an enzyme-free reaction with the same input DNA was performed in parallel. DNA was quantified with primers surrounding at least one CCGG site, where the amplification of HpaII-resistant and MspI-resistant DNA fractions was directly proportional to methylated (5mC) plus hydroxymethylated (5hmC) DNA, or 5hmC alone, respectively. Hence, levels of 5mC DNA were determined by subtracting the MspI-resistant DNA amount from the HpaII-resistant fraction, and represented as percentage of total (undigested) DNA (Fig6.1B). Note that primers for both H3K9me2 ChIP and DNA methylation assay were designed within regions that selectively acquire DNA methylation in XEN cells but not in embryo-derived EpiSCs (Senner et al., 2012)(Fig.A-VIII – Appendix A) as confirmed in these experiments (Fig. 6.1B, bottom panel). Conversely *Chrna3* promoter region was used as a positive control for detection of 5mC in WT and *Jmjd2c*-KO cEpiSCs, as well as in embryo-derived EpiSCs (Veillard et al., 2014) (Fig.6.1B; right panel).

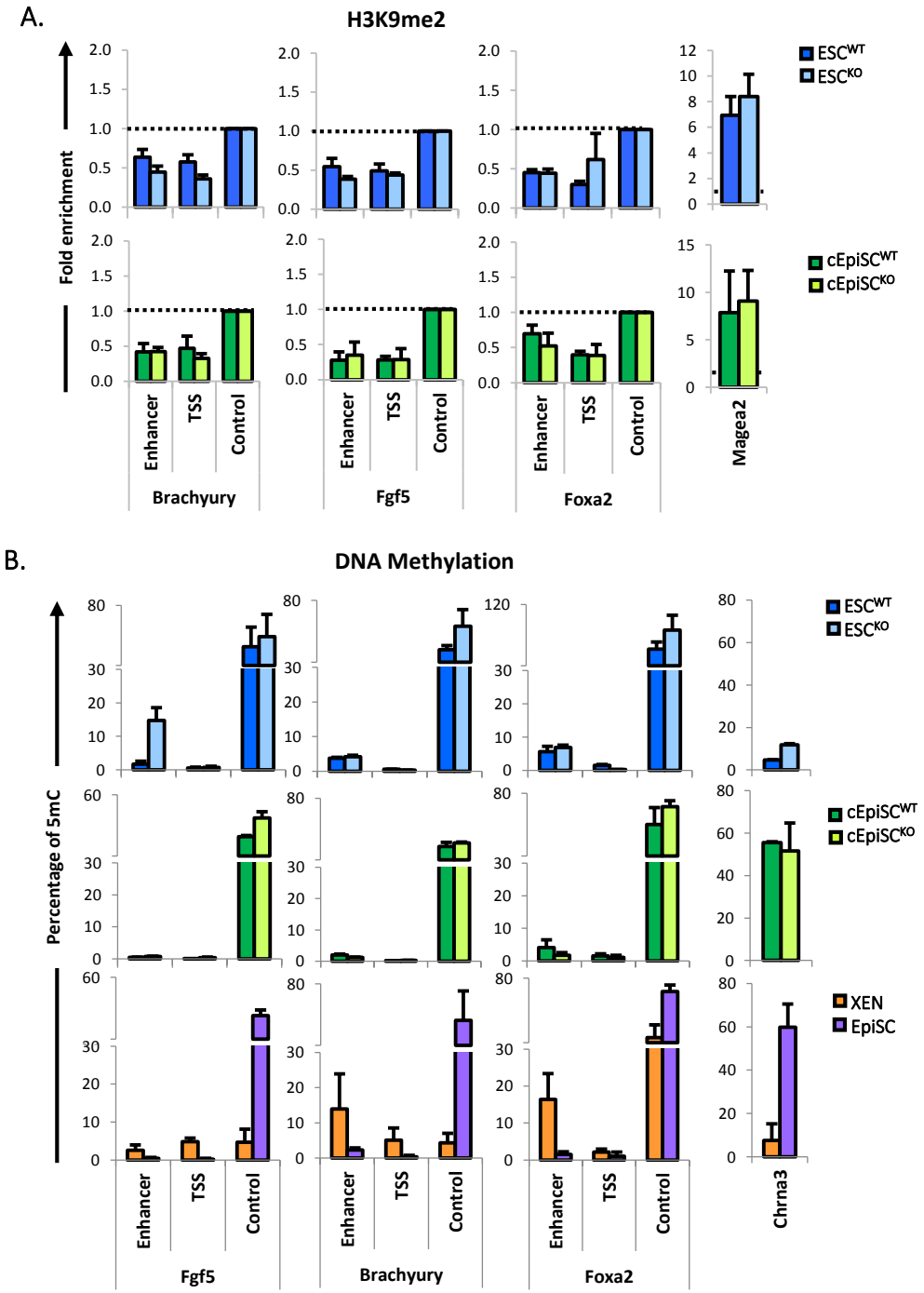


Figure 6.1 Levels of H3K9me2 and DNA methylation at TSS and enhancers of differentiation-associated genes are unchanged in *Jmjd2c*-depleted ESCs and cEpiSCs

A. ChIP-qPCR analysis showin H3K9me2 enrichment levels at enhancer, TSS and control (Jmjd2c-free) regions in the gene locui of *Fgf5*, *Brachyury/T* and *Foxa2*, in WT and *Jmjd2c*-KO ESCs (blue) and WT and *Jmjd2c*-KO cEpiSCs (green). *Magea2* was used as a positive control for enrichment of H3K9me2. Levels at each locus were normalized to the respective control regions (dotted lines), and levels of Magea2 normalized to a TSS H3K9-depleted region (dotted lines). Data was represented as average of 3

independent experiments ±SEM. **B.** DNA methylation levels were quantified by subtracting MspI-resistant (5hmC) from HpaII-resistant (5hmC+5mC) fractions. Levels were normalized to a control region without a MspI/HpaII restriction site and represented relative to total/undigested DNA (100%). The average of 2 experiments ± SD is plotted. Primers across all regions (enhancer and TSS) were selectively designed in regions known to be methylated in XEN cells (orange) (Fig.A-VIII – Appendix A), with an additional region (*Chrna3*) known to be methylated in EpiSC (purple).

Similarly to H3K9me2, higher levels of 5mC were found in the flanking regions of each of the studied loci. 5mC levels were, however, strongly reduced (or undetectable) at the TSS and enhancer regions of *Fgf5*, *Brachyury* and *Foxa2* (Fig.6.1B). This pattern appeared to be conserved between WT and *Jmjd2c*-KO ESC/cEpiSC populations, as also observed when additionally looking at 5hmC levels across the same regions (Fig.IX – Appendix A). While higher levels of 5mC were found at the *Fgf5* poised enhancer region *in Jmjd2c*-KO ESC, these were lost upon EpiSC induction (Fig.6.1B), and thus are unlikely to account for the reduced *Fgf5* expression observed in *Jmjd2c*-KO cEpiSCs.

6.2.2 Jmjd2c is required for the stable recruitment of G9a at active and poised enhancers in ESCS.

Mapping of genome-wide Jmjd2c DNA-binding sites in ESCs (Section 5.2.6) revealed that Jmjd2c is recruited at both active (H3K4me1, H3K27ac) and poised (H3K4me1) enhancer regions that are also co-bound by known enhancer proteins: Oct4, p300, Med1 and Smc1a. Remarkably, a concomitant binding of the antagonistic enzyme G9a at Jmjd2c distal/enhancer peaks was also identified. Nevertheless, no aberrant H3K9me2 levels were acquired at active and poised enhancers as well as at TSSs of associated genes in the absence of Jmjd2c. Interestingly, and in addition to G9a, Jmjd2c was also found to physically interact with Med1 in ESCs (Section 5.2.8), pointing to a functional role for Jmjd2c at enhancer regions independent of its H3K9 demethylase activity. To explore this further, the effect of Jmdj2c removal in the loading of enhancer protein complexes was here examined at both active (*Esrrb*, *Klf4*) and poised (*Fgf5*, *T, Foxa2*) enhancers

in ESCs, specifically looking by ChIP-qPCR at the enrichment levels of Oct4, p300, H3K27ac, Med1, the Cohesin factor Smc1a as well as G9a itself . A comparative analysis of Oct4, p300, H3K27ac and Med1 occupancies revealed higher enrichment levels at *Esrrb* and *Klf4* as compared to *Fgf5*, *T* and *Foxa2* enhancers, reflecting differences in the activity of these enhancers in ESCs. Interestingly, and as previously observed through interrogations of ChIP-seq datasets (Section 5.2.6), the Cohesin factor (Smc1a) appeared similarly recruited at both active and poised enhancers. These patterns remained unchanged at both sets of regions in the absence of Jmjd2c in ESCs.

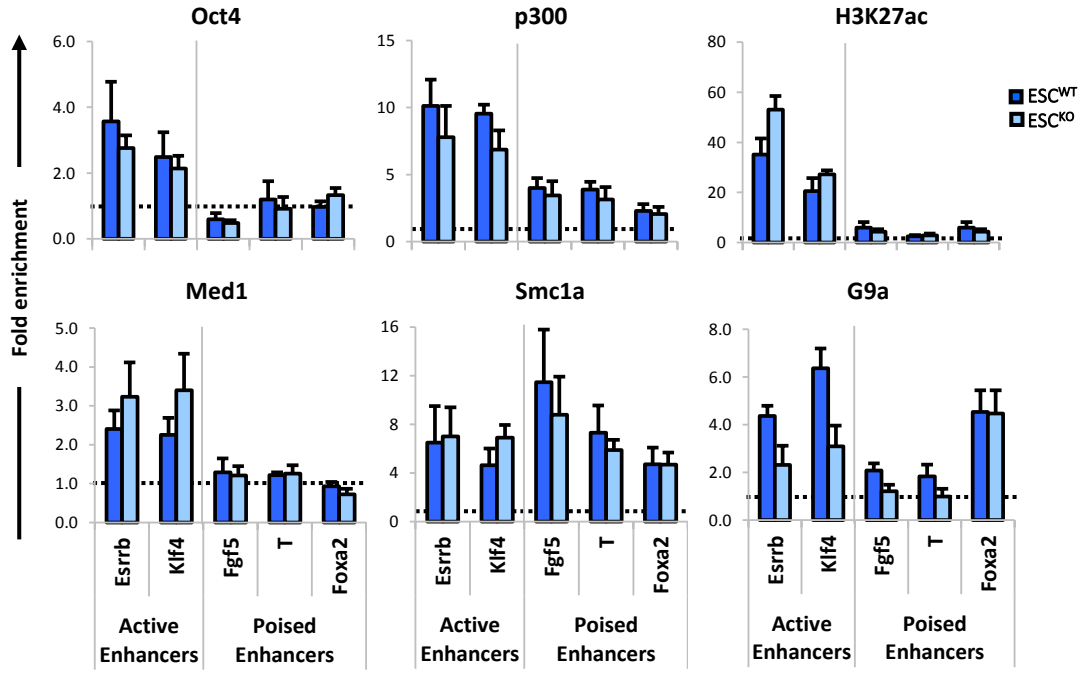


Figure 6.2 Formation of the enhancer complexes is retained in *Jmjd2c*-depleted ESCs

Enrichment levels for Oct4, p300, H3K27ac, Med1, Smc1a and G9a as assessed by ChIP-qPCR in WT and *Jmjd2c*-KO ESCs (ESC^{WT} and ESC^{KO}) at active (*Esrrb* and *Klf4*) and poised (*Fgf5*, *Brachuyry/T* and *Foxa2*) enhancers in ESCs. Levels at each region were normalized to an intergenic control region (dotted lines), where none of these factors bind to. Data was represented as average of 3 independent experiments ±SEM.

G9a has been previously shown to have a widespread distribution across the genome (Mozzetta et al., 2014), similarly to H3K9me2 (Lienert et al., 2011). Not surprisingly, G9a ChIP experiments proved to be technically challenging, and required a double chromatin cross-linking

approach, with titration of the chromatin amount used (Methods Section 2.2.9.1) as well as ChIP efficiency validations. Analysis of published G9a ChIP-seq profile revealed that G9a enrichment specifically accumulates at regions closely located to, but not overlapping, the TSS of a defined set of genes (Mozzetta et al., 2014). Thus, to validate the efficiency of anti-G9a ChIP, *Vstm2l* and *Npas4* were selected as G9a positive control targets regions as described. Conversely, the *Nnat* locus, where G9a was barely detected, was used as negative control genes (Mozzetta et al., 2014). Accordingly, G9a was found highly enriched at *Vstm2l* and *Npas4* and virtually absent at *Nnat* TSS regions in WT ESCs (Fig.A-X – Appendix A). Similarly, high G9a enrichment levels were detected at active enhancers in WT ESCs (Fig.6.2). G9a detection at poised enhancers was only marginal in these experiments. In contrast, in the absence of Jmjd2c, G9a binding appeared consistently reduced at *Esrrb*, *Klf4*, *Fgf5* and *T* enhancers, which were all demonstrated to be bound by Jmjd2c in ESCs (Section 5.2.7). These data indicate that Jmjd2c might be required for the stable recruitment of G9a at these sites in ESCs.

6.2.3 Jmjd2c facilitates the assembly of protein complexes at poised enhancers upon ESC-to-EpiSC transition.

In WT cEpiSCs, *Fgf5* was highly expressed relative to *Brachyury* and *Foxa2*, which only show low and very low gene expression, respectively (Section 4.2.3). In contrast, *Fgf5* mRNA levels were found to be reduced, and *Brachyury* and *Foxa2* undetectable in *Jmjd2c*-KO cEpiSCs. Regardless of transcriptional output, however, all enhancer regions analysed here were similarly occupied by the pioneer factor Oct4 in WT and KO cEpiSCs (Fig. 6.3; upper panel), indicative of active chromatin remodelling at these sites during ESC to EpiSC conversion. Indeed, during this transition, Oct4 binding was shown to undergo global re-distribution with increased loading at previously bound (poised) enhancers in ESCs, but also at newly defined enhancer regions in EpiLCs associated with expression of post-implantation epiblast markers (Buecker et al., 2014). Hence, these results also suggest that EpiSC-specific enhancer regions can be specified in the absence of Jmj2dc as further indicated by high H3K27ac and p300 enrichment levels, despite

lessened levels of p300 being detected at the *Fgf5* locus in *Jmjd2c*-KO cEpiSCs (Fig. 6.3; upper panel). In contrast, G9a and Med1 both failed to be properly recruited at *Fgf5* and *Brachyury* enhancers during ESC-to-EpiSC conversion as evidenced by G9a and Med1 enrichment levels close to background in *Jmjd2c*-KO as opposed to WT cEpiSCs, despite no changes being observed in the levels of this proteins (Fig.A-X – Appendix A). This corroborates with the impaired expression of these genes in the absence of Jmjd2c (Section 3.2.4; Section 4.2.4). Collectively, these data show that Jmjd2c is required for the proper assembly of enhancer protein complexes as seen at active (*Fgf5*) and poised (*Brachury* and *Foxa2*) enhancers during pluripotency maturation towards a primed, post-implantation EpiSC state.

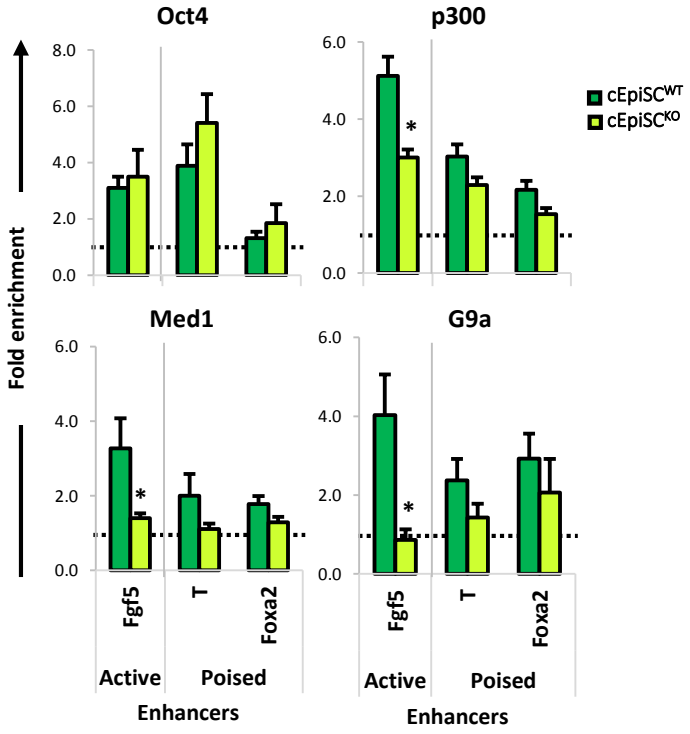


Figure 6.3 Assembly of enhancer complexes is destabilized in the absence of Jmjd2c in cEpiSC

Enrichment levels for Oct4, p300, Med1 and G9a as assessed by ChIP-qPCR in WT and *Jmjd2c*-KO cEpiSCs (cEpiSC^{WT} and cEpiSC^{KO}) at active (*Fgf5*) and poised (*Brachuyry/T* and *Foxa2*) enhancers in EpiSCs. Levels at each region were normalized to an intergenic control region (dotted lines), where none of these factors bind to. Data was represented as average of 3 independent experiments ±SEM. P<0.5(*), Mann-Whitney *U* test.

To further strengthen this conclusion, detailed analyses across the *Fgf5* locus was performed by assessing the recruitment of enhancer-associated factors at the poised enhancer constituent (PE) and two downstream EpiSC-specific sites (E1 and E3) as identified through Oct4, H3K27ac and H3K4me1 profiling in ESCs and EpiLCs (Fig.6.4A)(Buecker et al., 2014).

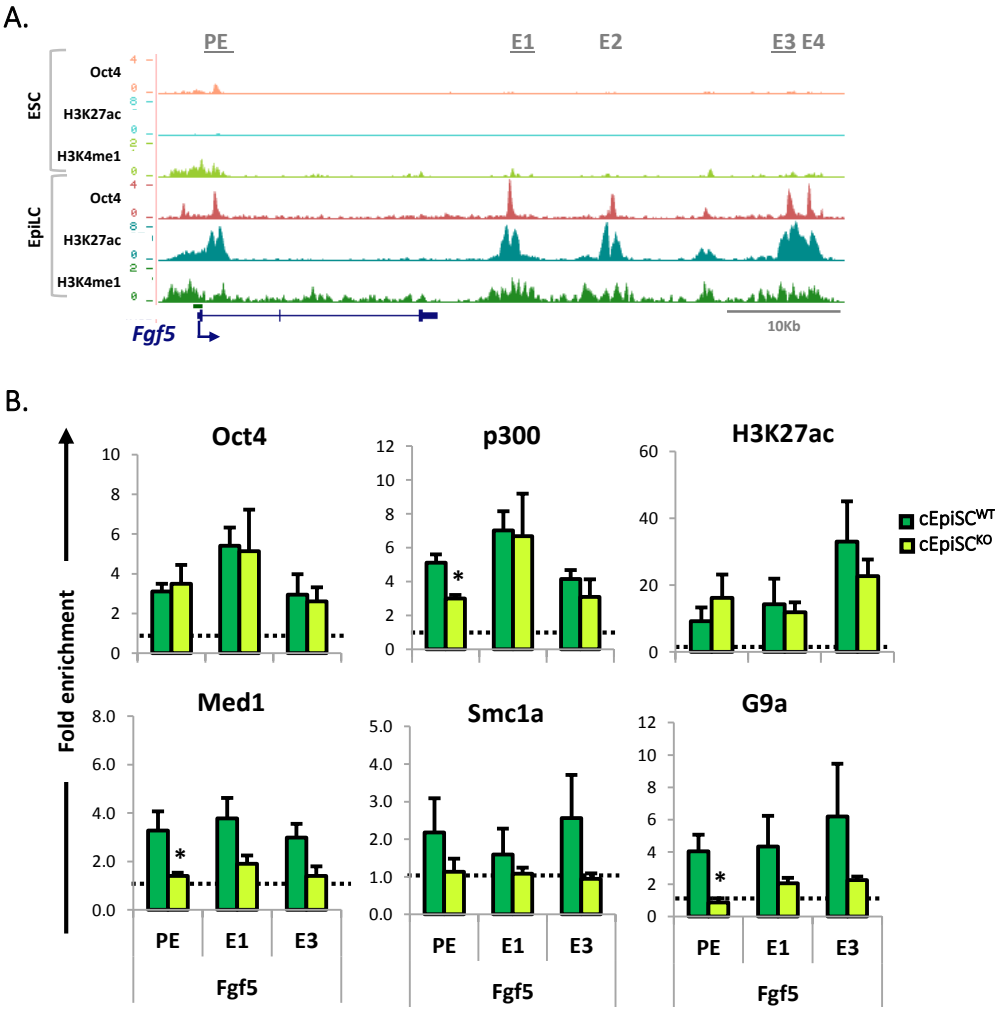


Figure 6.4 Recruitment of enhancer factors at EpiSC enhancers is impaired in the absence of Jmjd2c

A. Location of poised (PE) and EpiSC-specific enhancers (E1-4) acquired at the Fgf5 locus upon transition of ESC (2i/LIF) to a primed EpiLC (Act/Fgf/KOSR) state as previously defined on the basis of Oct4, H3K27ac and H3K4me1 ChIP-seq profiles (Buecker et al., 2014). **B.** Enrichment levels for Oct4, p300, Med1, Smc1a and G9a at PE, E1 and E3 Fgf5 enhancer contituents as determined by ChIP-qPCR in WT and *Jmjd2c*-KO cEpiSCs (cEpiSC^{WT} and cEpiSC^{KO}). Levels at each region were normalized to an intergenic control region (dotted lines), where none of these factors bind to. Data was represented as average of 3 independent experiments ±SEM. P<0.05 (*), Mann-Whitney *U* test.

In agreement with the data presented above, all three Fgf5 enhancer constituents showed evidence of epigenetic programming as indicated by equally high enrichment levels for Oct4, p300 and also H3K27ac deposition in WT and Jmjd2c-KO cEpiSCs (Fig. 6.4B; upper panel). These sites were furthermore confirmed as genomic targets for the enhancer and DNA looping factors Med1 and Smc1a, as well as G9a, in the presence of Jmjd2c. Yet Med1, Smc1a and G9a binding was consistently abrogated at both poised and newly defined EpiSC-specific enhancer constituents in the absence Jmjd2c (Fig. 6.4; bottom panel).

Collectively, these data uncover an important role for Jmjd2c in recruiting and/or facilitating the assembly of selective enhancer protein components as demonstrated for Med1, Smc1a and G9a but not for Oct4 and p300.

VI.3 Summary and Conclusions

In summary, the results described in this Chapter unveil a novel function for Jmjd2c in the recruitment and/or stable assembly of enhancer protein complexes upon lineage priming in pluripotent stem cells. Depletion of Jmjd2c was found to result in impaired gene activation upon differentiation into epiblast-derived lineages (Chapters III, IV). The strong binding of Jmjd2c at TSS regions, its concomitant detection with G9a at enhancer regions (Chapter V), together with the identified increase in bulk levels of H3K9me2 (Chapter III), pointed to a possible role for Jmjd2c in regulating and/or protecting TSS and enhancer sites against H3K9 methylation via its histone demethylase activity.

Here, aberrant acquisition of H3K9me2 (and *de novo* DNA methylation) at regulatory regions of poised (bivalent) genes in ESCs, in the absence of Jmjd2c, was hypothesised as a silencing mechanism underlying the identified defective expression upon differentiation. However, no increase in H3K9me2 deposition and DNA methylation was detected at the candidate loci studied both in *Jmjd2c*-knockout ESC and cEpiSC. Instead, defective loading of the enhancer and DNA-looping factors Med1 and Cohesin/Smc1a was selectively identified upon the transition from a

ESC state to a primed, post-implantation EpiSC-like state, as seen at the epiblast *Fgf5* and differentiation-associated gene loci *Brachyury* and *Foxa2* whose expression was found compromised in *Jmjd2c*-KO cEpiSC (Chapter IV).

Interestingly, the recruitment of the pioneer enhancer factor Oct4, and acquisition of H3K27ac/p300 – markers of active enhancers, were, in contrast, not impaired, suggesting that rather than the *de novo* establishment of enhancers *per se*, the assembly of the Mediator-Cohesin complex involved in looping events between enhancers and promoters (Kagey et al., 2010), might require the action of Jmjd2c. Furthermore, co-binding of Jmdj2c and G9a was identified at Med1-enriched regions, and these three proteins were furthermore found to physically interact in ESCs (Chapter V), suggesting that Jmj2dc-G9a interaction could take place in activating, Med1-enriched enhancer complexes. In line with this, in the absence of Jmjd2c, loading of G9a was also impaired at both active and poised enhancer regions as seen in ESCs and cEpiSCs. Altogether, these observations suggest that, rather than counteracting the deposition of G9a-mediated H3K9 methyaltion, Jmjd2c-G9a interaction at enhancers could instead be an essential feature of Med1 and Cohesin-mediated enhancer activity and DNA looping.

Chapter VII – Discussion

VII.1 Discussion

In this project, the role of Jmjd2c in embryonic stem cell (ESC) biology was explored using a *Jmjd2c*-knockout (KO) ESC model system. Depletion of Jmjd2c in the ESC state correlated with overall lessened expression of pluripotency markers. However, this was insufficient to impair ESC self-renewal capacity. Concordant with Jmjd2c previously reported H3K9-demethylase activity, *Jmjd2c*-KO ESCs harboured higher H3K9me2/3 bulk levels relative to wild-type (WT) cells. Interestingly, when challenged to differentiate upon embryoid body formation, *Jmjd2c*-depleted ESCs displayed impaired up-regulation of lineage-affiliated markers, and this seemingly resulted into a blockage at an epiblast-like stage of differentiation. Although *Jmjd2c*-KO ESCs could successfully generate self-renewing cEpiSCs, these epiblast cells, however, expressed reduced levels of *Fgf5* and failed to show any evidence of transcriptional priming for germ layer markers, as typically expected in primed pluripotent stem cells (Han et al., 2010). In contrast, differentiation into extra-embryonic endoderm was uncompromised and, furthermore emerged as a default pathway upon mesodermal induction in mutant cEpiSCs.

Dissection of the molecular mechanism underlying this phenotype uncovered a role for Jmjd2c in the transcriptional regulation of somatic, lineage-associated genes together with G9a. This action was found be independent of Jmjd2c H3K9-demethtylase activity but rather through the stabilization of activating protein complexes at poised and EpiSC-specific enhancer regions upon ESC differentiation.

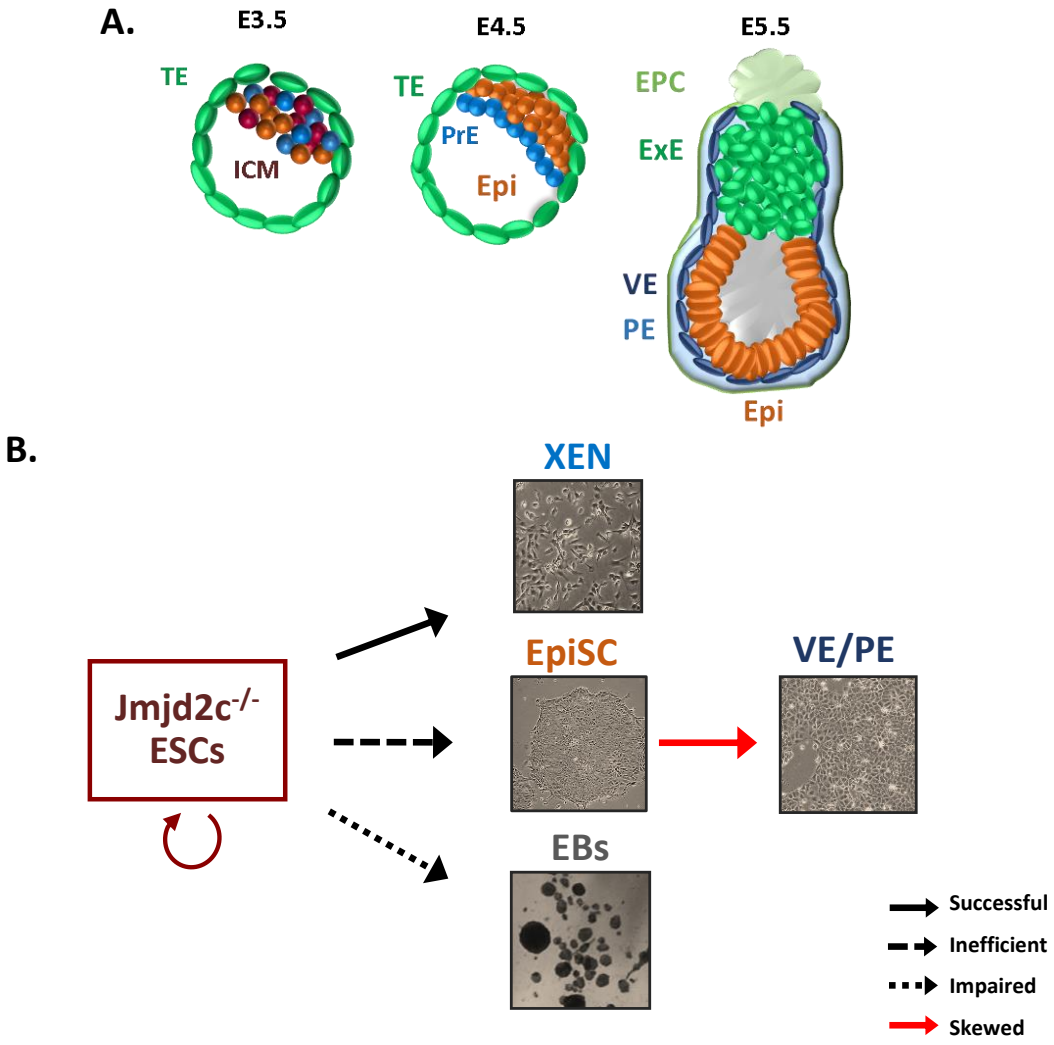


Figure 7.1 Summary of *Jmjd2c*-knockout ESCs phenotype

A. Schematic representation of early mouse embryonic development between pre-implantation (E.3.5-4.5) and post-implantation stages (E5.5), with highlighted key early embryonic and extra-embryonic lineage populations emerging at these stages: inner cell mass (ICM), Trophectoderm (TE), primitive endoderm (PrE), epiblast (Epi), visceral endoderm (VE), parietal endoderm (PE), extra-embryonic ectoderm (ExE) and ectoplacental cone (EPC). **B.** Summary representation of the different outcomes observed upon differentiation induction in *Jmjd2c*-knockout pluripotent stem cells.

7.1.1 Targeting of Jmjd2c in embryonic stem cells

Over decades the mouse has been exploited as a key model organism for the study of mammalian gene function and development. The isolation and propagation of pluripotent mouse ESCs has since provided a platform for *in vitro* study of molecular events occurring in the early

embryo. This has also laid the foundations for large-scale gene engineering, where virtually any gene function can be interrogated by targeted mutagenesis. The wide range of targeted mutations generated in ESCs by the EUCOMM/IKMC programs represents, up to now, the largest effort in providing tools for the scientific community to decipher gene function (Skarnes et al., 2011).

Hence, for the purpose of this work, genetically engineered *Jmjd2c*-knockout mouse ESCs obtained from the EUCOMM repository were used as a primary tool to uncover the role of Jmjd2c in ESC self-renewal state and beyond. Previous efforts to investigate the function of Jmjd2c in ESCs via transient RNAi approaches reported that knockdown of *Jmjd2c* mRNA leads to spontaneous differentiation, with concomitant repression of the ESC-specific gene network and de-repression of PRC2-bound differentiation markers, highlighting this as an essential marker of ESC identity (Das et al., 2014; Loh et al., 2007) . In striking contrast, ESCs stably depleted in Jmjd2c protein in this study maintain efficient self-renewal ability in culture, grow normally and express variable but high levels of pluripotency markers (Section 3.2.2). A concordant study using a conditional *Jmjd2c*-null ESC line generated from the same original EUCOMM cell line, similarly reported that these cells retain normal growth rate upon Jmjd2c depletion and maintain high expression levels of the pluripotency markers *Oct4*, *Sox2* and *Nanog* (Pedersen et al., 2014).

In RNAi systems, reagents such as shRNA or siRNAs are designed to target mRNA transcripts and are delivered to cells (Mohr et al., 2014), carrying a drug-resistant gene commonly used to select for successful transfection. This process ultimately relies on the specificity on the target sequence and efficiency of gene delivery. Here, gene trapping technology provided a stable and drug-free alternative system to interrogate the function of Jmjd2c in an *in vitro* embryo-derived ESC model. Although the gene trap system permitted some level of residual expression of Jmjd2c transcript, an occurrence previously found in genetrap systems (Pedersen et al., 2014; White et al., 2013), this appeared to robustly abolish expression of a full-length Jmjd2c protein (Fig.3.1). Moreover, a putative truncated Jmjd2c protein arising as a result of the inserted cassettes would lack both Jmjd2c PHD and Tudor domains and its nuclear localization signal leading to altered cellular location as experimentally demonstrated by combining mutagenesis and staining in

previous studies where Jmjd2c N-terminal domains were selectively deleted (Pedersen et al., 2014; Shin and Janknecht, 2007). Overall, the gene trap approach utilised to target Jmjd2c was validated and successfully generated a stable and robust knockout ESC model.

7.1.2 Possible redundancy between histone demethylases in embryonic stem cells and *in vivo*

The striking differences between the phenotype identified upon transient knockdown of *Jmjd2c* with shRNA vectors and in stable *Jmjd2c*-KO ESC lines used in this work, motivated a re-assessment of the state of ESCs upon shRNA knockdown. Thus, using *Jmjd2c* shRNA vectors previously validated and reported by Loh and colleagues (Loh et al., 2007), careful interrogation of the phenotype of these cells revealed that despite prominent spontaneous differentiation, alkaline phosphatase positive (undifferentiated) colonies with compact morphology could be readily identified in culture up to 4 days post-transfection and selection. Upon manual isolation and expansion of these colonies, a significant down-regulation of *Jmjd2c* transcript could be validated with concomitant maintenance of normal ESC self-renewal abilities (Fig.3.8). Notably, shRNA depleted ESCs displayed reduced expression of the pluripotency factors *Nanog*, *Esrrb* and *Tbx3*, concomitant with modest down-regulation of *Jmjd2b* (Fig.3.8D). Indeed, Jmjd2b was found to cooperate with Nanog in an interconnected regulatory loop, influencing the expression of Nanog's downstream targets such as *Klf4, Esrrb* and *Tbx3*. Moreover, the same study uncovered that both Jmjd2b and Jmjd2c co-occupy TSS regions of active genes, with Jmjd2b being also found at active ESC-specific enhancers (Das et al., 2014). In this work, Jmjd2c was also detected at a subset of distal regulatory regions that encompassed active (H3K4me1/H3K27ac-rich) enhancers in ESCs (Fig.5.4). Thus, it is possible that in active TSS and enhancers, Jmjd2c and Jmjd2b cooperate in a redundant manner to maintain the tight regulation of gene expression of self-renewal genes.

Concordantly, experimental evidence generated in this study suggested that Jmjd2c might in fact form a complex with the H3K9me2/3-demethylases such as Jmjd2b and Jmjd1a (Fig.5.9), further supporting that these could be able to cooperate and/or compensate for loss of function.

Further evidence for redundancy comes from the discrepancy of results in regards to the impact of depleting Jmjd2c in target histone modifications. One study documented an increase in H3K9me3, but not H3K9me2 bulk levels upon transient shRNA-mediated depletion of *Jmjd2c* in ESCs (Loh et al., 2007), whereas another study using a similar RNAi approach reported no obvious changes in global histone modifications (Das et al., 2014). Similarly, in constitutive Jmjd2c-null ESCs described by Pedersen and colleagues, no change in H3K9me2/3 levels were identified (Pedersen et al., 2014). Moreover, in this project, an initial titration of acid-extracted histone lysates and antibody dilutions was performed, allowing the detection of an increase in bulk levels of H3K9me2, with a modest increase in H3K9me3, in constitutive *Jmjd2c*-KO ESCs (Fig.3.4). Thus, it remains to be elucidated whether different levels of redundancy, or limitations on the detection of small changes of these histone modification due to antibody specificity and/or saturation of the detection method, explain the discrepancy of published results.

7.1.3 An essential role for Jmjd2c at the onset of embryonic stem cell differentiation

A key finding in this project was the identification of a requirement for Jmjd2c in epibast-derived lineage differentiation upon formation of embryoid bodies (EBs); a functional assay commonly used to investigate the developmental potential of ESCs *in vitro* (Fig.3.6; Fig.3.9). In this procedure, ESCs are cultured in suspension and in the absence of LIF to allow the generation of three-dimensional cellular aggregates that progress to form cystic structures comprising a mixture of differentiated cells from the three embryonic germ layers. Although heterogeneous, this system is able to recapitulate the early differentiation events occurring in early embryo.

When challenged to form EBs, *Jmjd2c*-depleted ESCs were unable to efficiently express the differentiation-associated markers tested from each of the three germ layers (mesoderm, endoderm and ectoderm), and concomitantly showed a reduced size of formed EBs as compared

to WT cells (Fig.3.6). Importantly, in both depletion systems tested (knockout and knockdown), weak but detectable expression levels of the early epiblast marker *Fgf5* were aberrantly retained in later differentiation time-points, similarly to *Oct4* (Fig.3.6; Fig.3.9). This suggested that differentiation of *Jmjd2c*-depleted ESCs could be impaired and possibly stalled *in vitro* at an early post-implantation equivalent epiblast-like stage of the mouse embryo. This was further supported by an impaired induction of neuronal progenitors in *Jmjd2c*-KO ESCs (Fig.3.8), in which the poor survival rate observed beyond the transiently reached epiblast-like state in normal differentiation pathway revealed that, under specific permissive conditions (Ying et al., 2003b), ESCs could not efficiently progress and acquire a neuronal–like phenotype.

In contrast, converted EpiSC populations could nevertheless be derived from *Jmjd2c*-KO ESCs upon provision of activin and fibroblast growth factor (Guo et al., 2009). Yet these cells strikingly proved unable to differentiate towards somatic lineages, as assessed in this study upon mesodermal induction (Fig.4.4). This confirmed that the observed differentiation defect in mutant cells arises from the emerging epiblast-like progenitors, as further revealed upon lack of lineage priming. Despite the ability of *Jmjd2c*-KO converted EpiSCs to be maintained in prolonged culture, these cells exhibited key phenotypic differences with their WT counterparts, suggestive of an underlying defect in transcriptional regulation. Indeed, these cells expressed diminished levels of *Fgf5*, and also displayed no primed expression of mesoderm (*Brachyury*) or endoderm (*Foxa2*) markers, as seen in WT cEpiSC (Fig.4.3). The absence of transcriptional gene priming for germ layer markers suggested that *Jmjd2c*-depleted cEpiSCs might have adopted an immature phenotype, failing to progress further as also evidenced by the lack of spontaneous differentiation typically occurring at the edge of WT EpiSC colonies. The existence of distinct EpiSCs have indeed been proposed based on the preferential use of distal (ESC-specific) versus proximal (EpiSC-specific) enhancers to drive Oct4 expression, concomitant with differential gene expression and developmental potency (Han et al., 2010). In this published report, gene expression profiling of high Oct4-GFP expressing EpiSCs was found to correlate with E5.5 embryos, and hence with an early (immature) stage of epiblast maturation, whereas profile of low or negative Oct4-GFP

EpiSCs resembled that of late-stage epiblast cells on the verge of differentiation. Moreover, early epiblast-like Oct4-high populations consistently express lower levels of *Fgf5* and *Brachyury* markers (Han et al., 2010). Based on these observations, it is proposed that *Jmjd2c*-KO cEpiSCs most likely represent a homogeneous, early epiblast stage.

7.1.4 Jmjd2c is expendable for the acquisition of extra-embryonic endoderm-like identity

The formation of the mouse blastocyst coincides with the first cell fate decisions to take place: the emergence of the pluripotent ICM and extra-embryonic trophoectoderm lineages. Around the time of implantation, the ICM further segregates into the extra-embryonic primitive endoderm (PrE) and the epiblast that will give rise to the embryo proper (Zernicka-Goetz et al., 2009). ESCs are derived from the pluripotent ICM of pre-implantation blastocysts, and in self-renewing culture conditions displayed some level of transcriptional priming for PrE markers (Artus et al., 2010; Canham et al., 2010; Capo-Chichi et al., 2005; Niakan et al., 2010), indicating that these cells retain the potential to differentiate towards extra-embryonic cells. Accordingly, conditions that allow the generation of extra-embryonic endoderm (XEN) stem cell populations from ESCs without gene manipulation have now been defined (Niakan et al., 2013).

In striking contrast with the outcome of EB induction experiments, *Jmjd2c*-KO ESCs could readily up-regulate the expression of PrE markers when treated with atRA in the absence of LIF, with faster kinetics of induction as seen for *Gata6*, *Gata4* and *Dab2* relative to WT cells (Fig.4.1.). Moreover, converted *Jmjd2c*-null XEN cells could be successfully generated, expressing XEN markers to similar levels than converted WT and embryo-derived XEN cells (Fig.4.2). These experiments suggested that Jmjd2c might be dispensable for the formation of extra-embryonic lineages, as further supported by the reduced or absence of Jmjd2c expression in embryo-derived XEN (Fig.4.2) and trophoblast stem cells (TS) cells (Alder et al., 2010), respectively. Hence, *Jmjd2c*-KO ESCs could similarly be expected to convert into TS-like cells (Ng et al., 2008), though this hypothesis remains to be tested.

Experiments described in this study further indicated that differentiation in *Jmjd2c*-KO pluripotent stem cells might be skewed towards extra-embryonic endoderm-like phenotypes. Here, *Jmjd2c*-KO cEpiSCs were induced to differentiate into early mesodermal progenitors by treatment with BMP4, Fgf4 and a Pi3K inhibitor in chemically defined media. At this stage, expression of *Brachyury* was found to be impaired in the absence of Jmjd2c as previously found upon EB-induced ESC differentiation. Moreover, following subsequent treatment with either, BMP4 and Fgf2, for lateral plate mesoderm induction, or with BMP4 and Pi3K inhibitors, Fgf2 and Wnt3a, for paraxial mesoderm induction (Fig.7.2), *Jmjd2c*-depleted cells consistently acquired a pronounced extra-embryonic endoderm-like identity, with only a small proportion of cells gaining expression of mesoderm markers and correct morphology (Fig.4.4). The extra-embryonic phenotype was identified by the polarized epithelial-like appearance with concomitant expression of visceral/parietal endoderm markers *Lrp2* and *Sparc*, similarly to BMP4-treated XEN cells (Fig.4.5).

Indeed, BMP4 treatment of XEN cells has been found to induce a VE identity (Artus et al., 2012; Paca et al., 2012; Soares et al., 2008) concordant with a role established for this signalling molecule in the regulation of anterior VE (Soares et al., 2008) and in proper development of extra-embryonic lineages in the early embryo (Graham et al., 2014). Furthermore, BMP4 treatment of embryo-derived EpiSCs has been shown to promote expression of PrE and trophectoderm markers (Brons et al., 2007; Vallier et al., 2009b) (Fig.7.1). Mesoderm differentiation of *Jmjd2c*-KO cEpiSCs was here induced by an initial treatment with BMP4 and selectively displayed a skewed differentiation into an extra-embryonic endoderm rather than a trophoblast fate indicated by the absence of *Cdx2* expression in day 4 differentiated cells (not shown). Interestingly while a low proportion of cells acquired expression of *Cxcr4* under mesodermal induction, marking the emergence of definitive endoderm-like cells, this occurred similarly in WT and *Jmjd2c*-KO cultures. Thus, these results suggest that absence of Jmjd2c could compromise the expression of somatic differentiation-associated factors, favouring the emergence extra-embryonic endoderm-like cells as a major default pathway. Of note, this same default pathway

was not seemingly occurring upon neuronal induction of *Jmjd2c*-KO ESCs (Fig.3.7), possibly due to the lack of permissiveness growth factors in the culture conditions used in this assay. Overall, this highlights key differences in the transcriptional regulation mechanisms in embryonic versus extra-embryonic lineages, which remain largely unexplored.

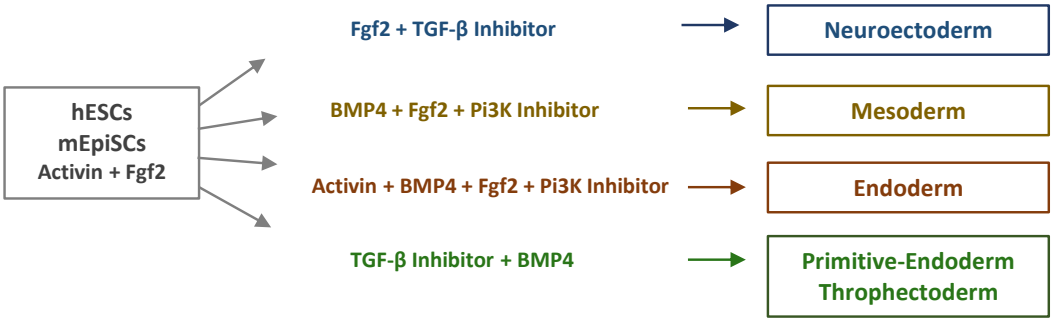


Figure 7.2 Signalling Pathways involved in differentiation of hESCs and mEpiSCs

Schematic representation of conditions used for in vitro differentiation of human embryonic stem cells (hESCs) and mouse Epiblast stem cells (mEpiSCs) into the indicated lineages (Bernardo et al., 2009; Vallier et al., 2009b).

7.1.5 Exploring a role for Jmjd2c in early cell fate decisions

The identification of a default extra-embryonic identity upon ineffective differentiation of *Jmjd2c*-depleted cells raised the question of whether the presence of Jmdj2c could be an essential factor in early cell fate decisions in the developing blastocyst. To explore this further, hanging-drop EBs (HD-EBs) were generated as an *in vitro* model to specifically study the impact of Jmjd2c depletion on the segregation of the epiblast and PrE lineages. Here, generation of HD-EBs from *Jmjd2c*-KO ESCs was hypothesised to favour the formation of a PrE-like outer layer. However, and despite enhanced expression of PrE transcripts as assessed by RT-qPCR (Fig.4.6B), when staining whole EBs against Oct4 and Gata4 no proper segregated outer layer could be identified in *Jmjd2*c-KO HD-EBs in these experiments (Fig.6.C). Staining agarose-embedded sectioned EBs could have, however, greatly improved the resolution of this confocal analysis and hence the readout of these experiments as previously described (Lavial et al., 2012). As an alternative approach, *Jmjd2c*-KO

and WT ESCs were labelled and co-cultured in suspension to induced the formation of chimeric HD-EBs, and cell allocation between outer versus inner layers investigated by immunofluorescence (Fig.4.7). Nevertheless, distribution of *Jmjd2c*-KO cells across the HD-EB did not show any preferential cell allocation bias into the PrE outer layer. Interestingly and concordantly with these preliminary analyses, *in vivo* expression data previously generated from single blastomeres harvested from E3.5 and E4.5 embryos revealed that indeed Jmjd2c transcript is expressed at variable levels being found in both PrE and epiblast blastomeres. However, consistently high levels of Jmjd2c correlated with E4.5 Epiblast blastomeres, the stage where expression of ICM/PrE factors such as Nanog/Gata6, respectively, is clearly segregated (Ohnishi 2014) (Fig.7.3).

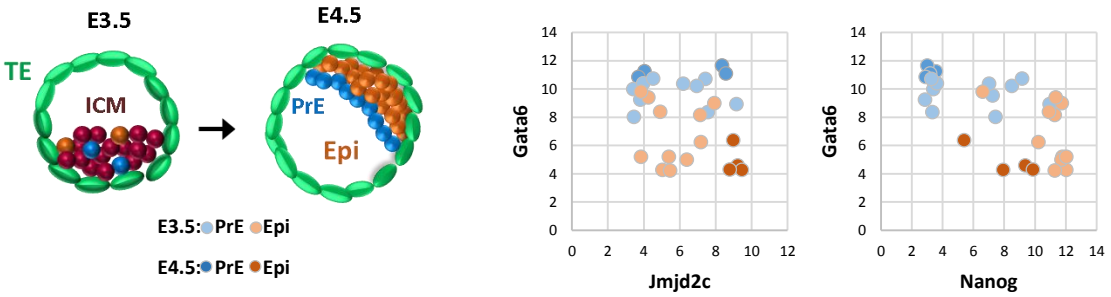


Figure 7.3 Single-cell *in vivo* expression in dissected ICMs of E3.5 and E4.5 mouse embryos

Single-cell expression analysis from single blastomeres harvested from the ICM of E3.5 and E4.5. Each dot in scatter plots represents mRNA levels detected in individual blastomere. The prospective blastomere identity, as depicted in the schematic representation of E3.5 and E4.5 embryos, was assigned based on the expression of the PrE *Gata4* marker as previously reported. Data from (Ohnishi et al., 2014).

Ultimately, the most stringent approach to uncover if presence of Jmjd2c would indeed impact on cell allocation would be by injecting labelled *Jmjd2c*-KO ESCs into morula embryos and subsequently implant these embryos into foster female host to examine the preferential distribution into extra-embryonic versus embryonic lineages of depleted cells at later embryonic stages. Nevertheless, Jmjd2c-null mice have been previously generated from inter-crossed heterozygous mice and were reported to be viable and fertile (Pedersen et al., 2014; White et al., 2013). Thus, abolishment of Jmjd2c function in whole-organism context does not appear to

impact on early embryo development as judged by the lack of gross phenotype in born mice, despite these studies lacking careful dissection of early development in Jmjd2c-null embryos. The evidence described in this project proposes a role for Jmjd2c as scaffolding molecule facilitating gene activation upon exit of the pluripotent state, as further discussed below. Hence, it is possible that functional redundancy takes place with other demethylases during embryo development in order to robustly sustain the expression of essential genes. Notably, cell competition has been described as a mechanism of elimination of "unfit" or "defective" cells during the post-implantation epiblast formation (Sancho et al., 2013). Thus, it still remains to be elucidated whether instead single-cell Jmjd2c depletion in an *in vivo* competitive environment would indeed impact on Epi/PrE allocation, or whether redundancy would also allow normal differentiation and development of injected mutant cells.

7.1.6 Genome-wide re-distribution of Jmjd2c between naïve and priming conditions

Comprehensive analysis of the pluripotent abilities of *Jmjd2c*-KO ESCs revealed inefficient lineage-associated gene activation upon induced differentiation, proposing a role for Jmjd2c in transcriptional regulation of poised genes. Previous reports described a predominant occupancy of Jmjd2c at H3K4me3-rich TSS sites of both poised and active genes in ESCs cultured under naïve (2i/LIF) condition (Pedersen et al., 2014) or priming (serum/LIF) conditions (Das et al., 2014). In priming conditions a larger number of non-TSS binding sites was also identified, though the function of Jmjd2c at these sites and its possible link with the acquisition of transcriptional gene primed remained to be fully investigated.

Therefore, Jmjd2c genome-wide binding sites were here directly compared in naïve and priming conditions using the same experimental setting, as well as upon induction of an epiblast-like (EpiLC) state (Buecker et al., 2014). It was hypothesised that upon lineage priming, Jmjd2c might be re-distributed into distal sites, possibly overlapping with EpiLC-specific enhancer regions. These sites indeed become dynamically targeted by the pioneer factor Oct4 at the exit of the pluripotent ground state, and this re-distribution of Oct4 binding is associated with co-

occupancy with the epiblast transcription factor Otx2 and acquisition of enhancer activity (Buecker et al., 2014; Yang et al., 2014). In this study, higher numbers of additional peaks were identified in priming (serum/LIF) ESC conditions as compared to the ESC ground-state also referred as naïve (2i/LIF) conditions (Fig.5.3B). Analysis of peak location revealed that nearly 13,500 TSS sites were similarly bound by Jmjd2c in both culture conditions, and that major re-distribution of Jmjd2c binding occurred at distal sites in serum/LIF conditions only (Fig.5.3). This was also accompanied by higher Jmjd2c protein expression being detected in priming conditions (Fig.5.2).

Induction of the EpiLC state did not, however, result into any significant genome-wide re-distribution of Jmjd2c binding as seen in serum/LIF conditions. Although the activation of epiblast markers was confirmed following 48h post-induction, this short time window may have not been sufficient for the acquisition of an homogenous EpiLC state in induced cultures and hence re-location of Jmjd2c binding sites in the ESC line and genetic background used in this study. Alternatively, the acquisition of distal binding sites could solely depend on intrinsic factors present in serum/LIF conditions.

Hence, Jmjd2c binding sites in efficiently induced EpiLC populations or embryo-derived EpiSCs still remains to be identified to fully elucidate the extent of Jmjd2c re-distribution to distal sites. Due to the lack of commercially available ChIP-grade antibodies raised against Jmjd2c, FLAG-tagged Jmjd2c expressing stable cEpiSCs would be best used for this purpose (Fig.5.2). This system, however, implies enforced expression of *Jmjd2c*, which may impact on the generation of self-renewing cEpiSCs as lower endogenous *Jmjd2c* expression levels are detected in EpiSCs relative to ESCs (Fig.4.3). In a preliminary attempt to generate stable cEpiSCs expressing FV-Jmjd2c construct consistently proved difficult, as the cells failed to acquire a proper EpiSC morphology instead displaying high incidence of cell differentiation. Alternatively, tagging of endogenous Jmjd2c with the recently developed CRISPR/Cas9 technology (Mali et al., 2013; Ratz et al., 2015) would overcome this effect and allow the generation of stable cEpiSC populations, in order to map Jmjd2c binding sites in an unbiased manner.

7.1.7 Jmjd2c associates with all H3K4 methylation states at TSS and distal genomic sites

Interestingly, the re-distribution of Jmjd2c in priming serum/LIF conditions selectively occurred at H3K4me1/me2-rich distal sites, in addition to H3K4me3-rich TSS sites. Concordantly, Jmjd2c has been proposed to recognise H3K4me3 and H3K4me2 via its Tudor domains, although a link with H3K4me1 had not been previously recognized (Pedersen et al., 2014). Moreover, a small proportion of distal sites were also highly co-enriched with H3K27ac, defining two distinct groups of Jmjd2c-distal peaks encompassing active (H3K27ac-high) enhancers, and poised (H3K27-ac low) enhancers (Fig.5.5). In line with this observation, Jmjd2c was recently identified in pull-down of histone modifications H3K4me3 and H3K4me1, and to a lesser extent H3K27ac in serum/LIF ESCs (Engelen et al., 2015). ChIP-qPCR at known active and poised enhancers furthermore unveiled that, although Jmjd2c binding could be confirmed at both sets of enhancers, recruitment at active enhancers appeared to occur independently of its Tudor domains (Fig.5.6). Thus, these observations extend the previously characterized link between Jmjd2c recruitment and H3K4me3 recognition, to all H3K4 methylation states (Fig.7.4), and furthermore proposed that Jmjd2c might be recruited through distinct mechanisms at poised and active enhancers in serum/LIF conditions.

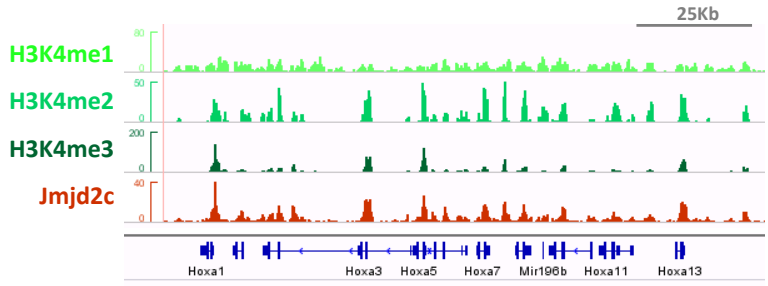


Figure 7.4 Jmjd2c associates with all H3K4-methylation states

Representative snapshot of Jmjd2c binding at H3K4me3/2/1-rich locations in a region spanning 100Kb of the *Hoxa*-cluster loci at chromosome 6, in serum/LIF conditions.

Epigenetic programming of enhancers is an important component of gene regulation during development. The pre-marking of enhancers by H3K4me1 and H3K4me2 has been proposed to facilitate the binding of pioneer transcription factors and hence, stimulate the establishment of an open chromatin environment (Zaret and Carroll, 2011). Thus, synergistic action of chromatin modifications and pioneer factor binding are likely to have a key role in the establishment of permissive conditions for transcription. Profiling of the epigenetic landscape of ground-state (2i/LIF) ESC cultures revealed reduced transcriptional gene priming as indicated by lower transcript levels of differentiation-associated genes and reduced enrichment for H3K27me3 at gene promoters, hence resulting into less bivalent chromatin domains being detected (Marks et al., 2012), a chromatin structure that was previously associated with poised gene expression in serum/LIF ESCs (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006b) . Thus, it is reasonable to envisage that in naïve conditions the pre-marking of enhancers by H3K4me1 and H3K4me2 could also be reduced. Altogether, the re-distribution of Jmjd2c between naïve and primed states could mimic the distribution of H3K4-methylation states, being prominently bound to H3K4me3-rich TSS in naïve conditions and further re-distributed to H3K4me1/2-rich distal sites upon enhancer pre-marking in the serum/LIF ESC state.

7.1.8 Jmjd2c and G9a are co-enriched at both active and poised enhancers in ESCs

Remarkably, the majority of Jmjd2c distal sites detected were associated with epigenetic signature typical of poised (H3K27ac-low) enhancers. These H3K27ac-low peaks were similarly co-occupied by reduced levels of Med1, Cohesin and Oct4, as compared to H3K27ac-high peaks. Moreover, H3K27ac-low, Jmjd2c-bound regions encompassed a great variety of sequence motifs significantly associated with differentiation-affiliated transcription factors, including also the pioneer factor Oct4 (Fig.5.5C; Table B-VII – Appendix B). Gene ontology analysis further confirmed that the nearest genes associated to H3K27ac-low peaks were functionally involved in embryonic development and differentiation processes (Fig.5.5D). This corroborates with a role for Jmjd2c in epiblast lineage priming.

An interesting finding that arose from the analysis of Jmjd2c-distal peaks was the co-occupancy of H3K9me2-demethylase G9a at these sites (Fig.5.4C). Strikingly, and in contrast to its well-established role in gene silencing, G9a highest enrichment levels were detected at H3K27ac-high peaks and hence active enhancer regions in ESCs (Fig. 5.5A), proposing that Jmjd2c and G9a co-enrichment might coincide with the formation of activating Mediator complexes at these sites. In line with this, Jmjd2c, G9a and Med1 interactions were confirmed exogenously in HeLa cells and endogenously in ESCs by co-immunoprecipitation (Fig.5.7). Yet Med1 failed to be identified upon FLAG-Jmjd2c pull-down in ESCs (Fig.5.8). While validation that these proteins form a trimeric complex is still missing, Jmjd2c, G9a and Med1 co-occupancy at the chromatin strongly suggests that they physically interact as part of the same complex at the chromatin. However, despite unlikely, the possibility that these are forming two distinct complexes in heterogeneous cell populations cannot be discarded.

7.1.9 Jmjd2c and G9a - counteracting roles in the modification of H3K9?

The co-occupancy of the antagonistic enzymes Jmjd2c and G9a at poised enhancer regions proposed a counteracting role in activation versus repression of these sites, possibly as a mechanism for fine-tuning primed gene expression. Similarly, histone acetyltransferases (HATs) and deacetylases (HDACs) were, for example, found to co-occupy both active and poised gene promoters in human T cells (Wang et al., 2009). In this published report, transient binding of HAT/HDAC at H3K4-methylated poised genes was demonstrated to be essential for keeping genes in a primed/inactive state through an active cycle of acetylation and deacetylation. However, in *Jmjd2c*-depleted ESCs and cEpiSCs, no acquisition of H3K9me2 and *de novo* DNA methylation - two silencing marks associated with gene repression in extra-embryonic tissues (Alder et al., 2010; Senner et al., 2012), were detected at regions co-bound by Jmjd2c and G9a (enhancers) or Jmjd2c alone (TSS) (Fig.6.1). Thus, despite the acquisition of enriched bulk levels of H3K9 methylated histones in the absence of Jmjd2c (Fig.3.4), these did not correlate with an increase at selected target regions. Here, reduced H3K9me2 enrichment levels at regulatory

regions versus flanking regions were consistently observed regardless of Jmjd2c expression, and furthermore corroborates with the previously reported widespread genomic location of H3K9me2 marks with the notable exception of H3K4me2-rich regions (Lienert et al., 2010). Notably, H3K4me2 and the H3K9me2-depositing enzyme G9a co-occur at the same sites. However, H3K4me2-rich sites appear to remain depleted of H3K9me2 in the absence of Jmjd2c, proposing histone-independent roles for G9a and Jmjd2c in H3K4me2-rich locations. Thus, it remains to be clarified how regulatory sites are indeed protected from acquisition of H3K9me2.

7.1.10 Jmjd2c-G9a association favours stable recruitment of enhancer-associated protein complexes

Unexpectedly *Jmjd2c*-depleted ESCs displayed reduced G9a enrichment levels at active and poised enhancer sites that appear to be normally co-bound by these two molecules in ESCs. This further highlights a cooperative rather than competitive interplay between Jmjd2c and G9a. Interestingly, impaired G9a binding was here observed at not only *Fgf5* poised enhancer in *Jmjd2c*-KO ESCs, but also at newly formed Fgf5 enhancer constituents as defined by Oct4 and p300 occupancy in mutant cEpiSCs (Fig.6.3). Strikingly, at this stage, loss of Jmjd2c led to compromised *Fgf5* expression in cEpiSCs that was mirrored by the destabilized G9a-Med1-Smc1a complex formation as seen at all Fgf5 enhancers. Taken together, these findings suggest that Jmjd2c and G9a could synergistically promote the recruitment and/or stable assembly of the Mediator-Cohesin complex at enhancers (Fig.7.5).

The Mediator complex as emerged as a central co-activator of transcription, which by interaction with key transcription factors, enables recruitment of RNA polymerase II to promoters (Conaway et al., 2005; Kornberg, 2005; Malik and Roeder, 2005) . Cooperation between the Mediator complex and the Cohesin ring has been shown to bridge interactions between transcription factors at enhancers and the transcription machinery at core promoters in ESCs (Kagey et al., 2010). Thus, cell-specific DNA looping formation between enhancer-bound transcription factors and promoters is a fundamental process in gene activation, and inefficient

loading or assembly of these factors could lead to impaired gene expression. Although the Mediator subunit Med1 and Cohesin subunit Smc1a genome-wide profiles have not yet been reported in EpiSCs, here it was found that Med1 and Smc1a remain similarly expressed in EpiSCs as in ESCs (Fig.A-X – Appendix A), implying that these proteins could play conserved regulatory roles in gene activation at both developmental stages. Hence, it would be important to validate genome-wide their re-location to newly formed enhancer regions upon pluripotency maturation that is the transition from an ESC to a primed EpiSC state.

Collectively, this study provides novel insights into the role of Jmjd2c-G9a as molecular scaffolding for the assembly of the Mediator-Cohesin complexes, hence potentially contributing to the formation of DNA loops between newly activated enhancers and promoters upon triggering of ESC differentiation. In this model, removal of Jmjd2c would both impact on the loading of G9a of the Mediator-Cohesin complexes leading to inefficient gene activation during somatic, lineage priming. To further validate this hypothesis, it would be interesting to assess the impact of depletion of Jmjd2c and/or G9a in assembly of DNA looping events between the candidate genes using chromosome conformation capture (3C) technology (Dekker et al., 2002).

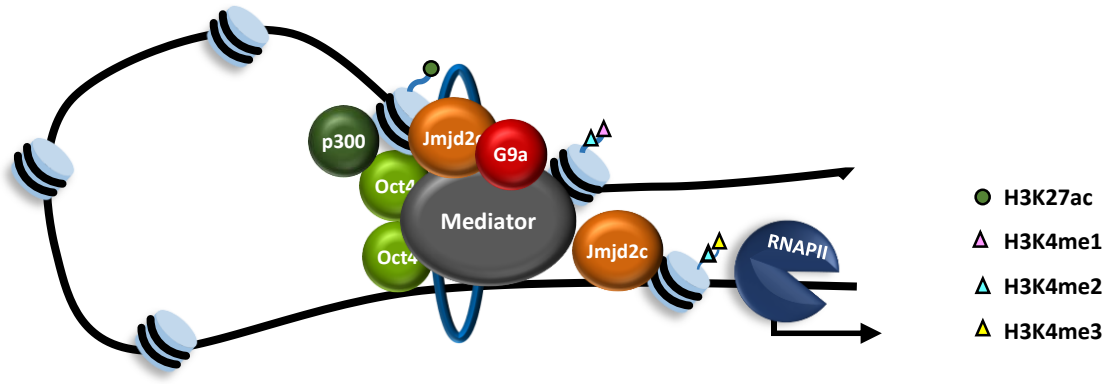


Figure 7.5 Model of assembly of active Jmjd2c-G9a centred enhancer protein complex

Proposed model for Jmjd2c and G9a action in the assembly of the Mediator-Cohesin Complex. Jmjd2c and G9a co-bind enhancer sites, promoting and/or stabilizing the recruitment of the Mediator complex and Cohesin (ring) necessary for the formation of DNA loops and potent activation of poised genes upon triggering of differentiation and exit of the pluripotent state.

7.1.11 Identification of novel roles for the chromatin modifiers Jmjd2c and G9a

A role for Jmjd2c as a co-activator via H3K9me3-demethylation of promoters has been reported in different cell contexts, where abolishing Jmjd2c led to impaired differentiation (Cascante et al., 2014; Lu et al., 2012; Wu et al., 2015). Moreover, Jmjd2c has also been described as a co-repressor through interaction with HDAC inhibiting PPARγ-mediated transcriptional pathways in preadipocyte differentiation (Lizcano et al., 2011) as well as SCR-1/NCOA1-dependent gene activation in breast cancer cells (Walsh et al., 2014), suggesting that Jmjd2c might endorse various functions that are either linked or not with its catalytic activity.

Conversely, the G9a protein exists in a heteromeric complex with GLP, and has been widely associated with gene repression via its H3K9me2-methylase activity (Collins and Cheng, 2010). However, a growing number of reports have documented that G9a can act as an activator in a context dependent-manner (Chaturvedi et al., 2009; Chaturvedi et al., 2012; Lee et al., 2006a; Lehnertz et al., 2010; Oh et al., 2014). Interestingly, an activating interaction between the Mediator complex and G9a had been previously identified in hematopoietic (erythroid) progenitor cells. Here, G9a was found to participate in two distinct protein complexes: as an activator with the Mediator complex and as a repressor with the H3K4me3-demethylase Jarid1a, both regulating the differential activation of globin genes (Chaturvedi et al., 2012). In particular, depletion of G9a was shown to de-stabilize the recruitment of Mediator complex at the β^{maj} globin (adult) locus, whose expression was then abrogated, while concomitantly re-activating the E^{y} globin (embryonic) gene.

In ESCs, an interplay between G9a/GLP and PRC2 has been shown to safeguard repression of developmental-associated genes (Mozzetta et al., 2014). In that context, depletion of G9a resulted into an up-regulation of many differentiation markers, including *Fgf5*. However, a large proportion of PRC2-bound genes were not de-repressed. These findings together with the findings described in this project suggest that G9a could also be part of two different complexes in ESCs: an activator complex with Jmjd2c-Mediator-Cohesin at active and poised enhancers, and a repressor complex with PRC2 at bivalent promoter regions. Thus depletion of G9a might

compromise the loading of both complexes, where the balance between activator and repressive actions might set the extent of repression versus de-repression of developmental associated genes.

The repressive function of G9a has been associated with its auto-methylation ability, which mediates the co-binding of other repressive factors such as HP1 or CDYL (Sampath et al., 2007), although the same link with recruitment of PRC2 has not been explored (Mozzetta et al., 2014). Interestingly, non-histone targets have been identified *in vitro* for the Jmjd2 family, and these included the tri-methylated sites of proteins such as WIZ, CDYL, CSB and the auto-methylation site of G9a (Ponnaluri et al., 2009). Here, the precise molecular mechanism underlying the identified Jmjd2c-G9a functional interplay remains to be fully uncovered. Yet one possibility could be that the integration of G9a into distinct complexes in ESCs relies on its methylation state. Thus, interaction of G9a with an activator complex could require demethylation by Jmjd2c, whereas its methylated form creates a platform for recruitment of repressive complexes (Fig.7.6). Investigating the methylated state of G9a in Med1-bound versus PRC2 or HP1-bound complexes would be fundamental to address this question.

Moreover, in a recent report it was uncovered that G9a induced the degradation of MyoD – a transcription factor involved in skeletal muscle differentiation, through methylation of this protein. In addition, Jmjd2c was found to demethylate G9a-mediated methylation resulting in stabilization of the MyoD protein, thus enhancing its transcriptional ability (Jung et al., 2015). Hence, another possibility would be that the same regulatory mechanism is necessary for stabilization of proteins that integrate the enhancer complexes in ESCs and at the onset of differentiation.

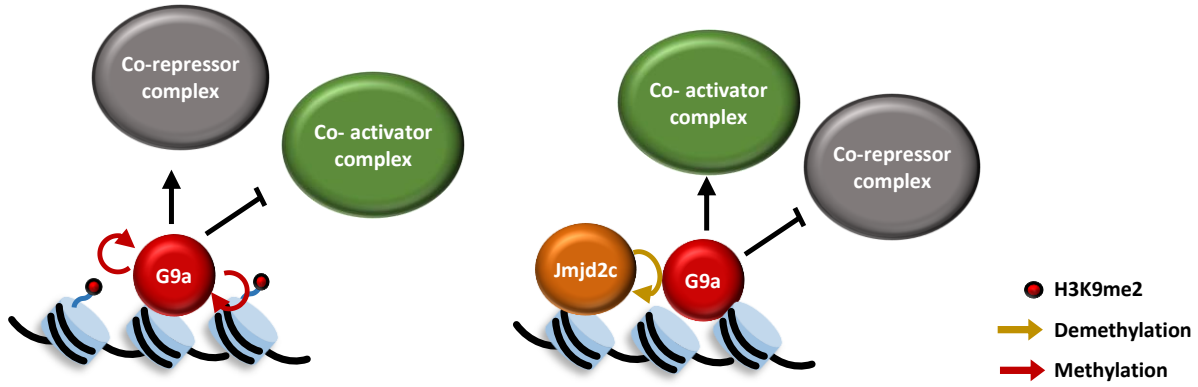


Figure 7.6 Hypothesised outcomes of different methylation states of G9a

In this model, G9a is proposed to be part of both co-activator and co-repressor complexes. Methylation activity of G9a (auto-methylation and H3K9me2 methylation) might provide a platform for recruitment of repressive complexes. Demethylation of G9a by Jmjd2c could allow G9a to bind activating complexes, hence precluding the recruitment of repressor complexes at regulatory regions.

References

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009). Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. PLoS One *4*, e6286.

Acampora, D., Di Giovannantonio, L.G., and Simeone, A. (2013). Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. Development *140*, 43-55.

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731-734.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. PLoS One *5*, e10531.

Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G.K., Teo, R., Leng Ng, C.K., Herath, W., Lili, S.*, et al.* (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. EMBO J *32*, 938-953.

Alder, O., Lavial, F., Helness, A., Brookes, E., Pinho, S., Chandrashekran, A., Arnaud, P., Pombo, A., O'Neill, L., and Azuara, V. (2010). Ring1B and Suv39h1 delineate distinct chromatin states at bivalent genes during early mouse lineage commitment. Development *137*, 2483-2492.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol *10*, 91-103.

Artus, J., Douvaras, P., Piliszek, A., Isern, J., Baron, M.H., and Hadjantonakis, A.K. (2012). BMP4 signaling directs primitive endoderm-derived XEN cells to an extraembryonic visceral endoderm identity. Dev Biol *361*, 245-262.

Artus, J., Panthier, J.J., and Hadjantonakis, A.K. (2010). A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. Development *137*, 3361-3372.

Artus, J., Piliszek, A., and Hadjantonakis, A.K. (2011). The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. Dev Biol *350*, 393-404.

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

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

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

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733-737.

Bedzhov, I., Graham, S.J., Leung, C.Y., and Zernicka-Goetz, M. (2014). Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. Philos Trans R Soc Lond B Biol Sci *369*.

Bedzhov, I., and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. Cell *156*, 1032-1044.

Bejsovec, A. (2005). Wnt pathway activation: new relations and locations. Cell *120*, 11-14.

Bernardo, A.S., Cho, C.H., Mason, S., Docherty, H.M., Pedersen, R.A., Vallier, L., and Docherty, K. (2009). Biphasic induction of Pdx1 in mouse and human embryonic stem cells can mimic development of pancreatic beta-cells. Stem Cells *27*, 341-351.

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

Berry, W.L., and Janknecht, R. (2013). KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. Cancer Res *73*, 2936-2942.

Birke, M., Schreiner, S., Garcia-Cuellar, M.P., Mahr, K., Titgemeyer, F., and Slany, R.K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res *30*, 958-965.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nat Genet *44*, 148-156.

Boroviak, T., and Nichols, J. (2014). The birth of embryonic pluripotency. Philos Trans R Soc Lond B Biol Sci *369*.

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

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

Bradley, A., Anastassiadis, K., Ayadi, A., Battey, J.F., Bell, C., Birling, M.C., Bottomley, J., Brown, S.D., Burger, A., Bult, C.J., *et al.* (2012). The mammalian gene function resource: the International Knockout Mouse Consortium. Mamm Genome *23*, 580-586.

Branco, M.R., Ficz, G., and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet *13*, 7-13.

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

Brookes, E., de Santiago, I., Hebenstreit, D., Morris, K.J., Carroll, T., Xie, S.Q., Stock, J.K., Heidemann, M., Eick, D., Nozaki, N., *et al.* (2012). Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell Stem Cell *10*, 157-170.

Brookes, E., and Pombo, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. EMBO Rep *10*, 1213-1219.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838-853.

Bult, C.J., Eppig, J.T., Blake, J.A., Kadin, J.A., Richardson, J.E., and Mouse Genome Database, G. (2013). The mouse genome database: genotypes, phenotypes, and models of human disease. Nucleic Acids Res *41*, D885-891.

Bulut-Karslioglu, A., De La Rosa-Velazquez, I.A., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galan, C., Winter, G.E., Engist, B., Gerle, B., *et al.* (2014). Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. Mol Cell *55*, 277-290.

Burton, A., and Torres-Padilla, M.E. (2014). Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. Nat Rev Mol Cell Biol *15*, 723-734.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol Cell *49*, 825-837.

Canham, M.A., Sharov, A.A., Ko, M.S., and Brickman, J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol *8*, e1000379.

Capo-Chichi, C.D., Rula, M.E., Smedberg, J.L., Vanderveer, L., Parmacek, M.S., Morrisey, E.E., Godwin, A.K., and Xu, X.X. (2005). Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. Dev Biol *286*, 574-586.

Carey, N., Marques, C.J., and Reik, W. (2011). DNA demethylases: a new epigenetic frontier in drug discovery. Drug Discov Today *16*, 683-690.

Cascante, A., Klum, S., Biswas, M., Antolin-Fontes, B., Barnabe-Heider, F., and Hermanson, O. (2014). Gene-specific methylation control of H3K9 and H3K36 on neurotrophic BDNF versus astroglial GFAP genes by KDM4A/C regulates neural stem cell differentiation. J Mol Biol *426*, 3467-3477.

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

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

Chaturvedi, C.P., Hosey, A.M., Palii, C., Perez-Iratxeta, C., Nakatani, Y., Ranish, J.A., Dilworth, F.J., and Brand, M. (2009). Dual role for the methyltransferase G9a in the maintenance of beta-globin gene transcription in adult erythroid cells. Proc Natl Acad Sci U S A *106*, 18303-18308.

Chaturvedi, C.P., Somasundaram, B., Singh, K., Carpenedo, R.L., Stanford, W.L., Dilworth, F.J., and Brand, M. (2012). Maintenance of gene silencing by the coordinate action of the H3K9 methyltransferase G9a/KMT1C and the H3K4 demethylase Jarid1a/KDM5A. Proc Natl Acad Sci U S A *109*, 18845-18850.

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

Chen, T., and Dent, S.Y. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nat Rev Genet *15*, 93-106.

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

Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T.G., Simpson, M., Mao, Q., Pan, C.H., Dai, S.*, et al.* (2006). Structural insights into histone demethylation by JMJD2 family members. Cell *125*, 691-702.

Cho, L.T., Wamaitha, S.E., Tsai, I.J., Artus, J., Sherwood, R.I., Pedersen, R.A., Hadjantonakis, A.K., and Niakan, K.K. (2012). Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. Development *139*, 2866-2877.

Cho, Y.W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G.R., Copeland, T.D., Kalkum, M.*, et al.* (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. J Biol Chem *282*, 20395-20406.

Chung, Y., Klimanskaya, I., Becker, S., Marh, J., Lu, S.J., Johnson, J., Meisner, L., and Lanza, R. (2006). Embryonic and extraembryonic stem cell lines derived from single mouse blastomeres. Nature *439*, 216-219.

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J.H., Skalnik, D., and Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev *26*, 1714-1728.

Collins, R., and Cheng, X. (2010). A case study in cross-talk: the histone lysine methyltransferases G9a and GLP. Nucleic Acids Res *38*, 3503-3511.

Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. Trends Biochem Sci *30*, 250-255.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791-1792.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

D'Amour, K.A., and Gage, F.H. (2003). Genetic and functional differences between multipotent neural and pluripotent embryonic stem cells. Proc Natl Acad Sci U S A *100 Suppl 1*, 11866-11872.

Das, P.P., Shao, Z., Beyaz, S., Apostolou, E., Pinello, L., De Los Angeles, A., O'Brien, K., Atsma, J.M., Fujiwara, Y., Nguyen, M.*, et al.* (2014). Distinct and combinatorial functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in mouse embryonic stem cell identity. Mol Cell *53*, 32-48.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. Mol Cell Biol *24*, 2478-2486.

Ellis, P., Fagan, B.M., Magness, S.T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M., and Pevny, L. (2004). SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. Dev Neurosci *26*, 148-165.

Engelen, E., Brandsma, J.H., Moen, M.J., Signorile, L., Dekkers, D.H., Demmers, J., Kockx, C.E., Ozgur, Z., van, I.W.F., van den Berg, D.L.*, et al.* (2015). Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. Nat Commun *6*, 7155.

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

Farcas, A.M., Blackledge, N.P., Sudbery, I., Long, H.K., McGouran, J.F., Rose, N.R., Lee, S., Sims, D., Cerase, A., Sheahan, T.W.*, et al.* (2012). KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife *1*, e00205.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development *121*, 273-285.

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

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

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

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398-402.

Fonseca, S.A., Costas, R.M., and Pereira, L.V. (2015). Searching for naive human pluripotent stem cells. World J Stem Cells *7*, 649-656.

Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and Chazaud, C. (2011). Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. Dev Cell *21*, 1005-1013.

Fritsch, L., Robin, P., Mathieu, J.R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. Mol Cell *37*, 46-56.

Fuhrmann, G., Chung, A.C., Jackson, K.J., Hummelke, G., Baniahmad, A., Sutter, J., Sylvester, I., Scholer, H.R., and Cooney, A.J. (2001). Mouse germline restriction of Oct4 expression by germ cell nuclear factor. Dev Cell *1*, 377-387.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. Genes Dev *16*, 784-789.

Goodman, R.H., and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. Genes Dev *14*, 1553-1577.

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

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

Gross, D.S., and Garrard, W.T. (1988). Nuclease hypersensitive sites in chromatin. Annu Rev Biochem *57*, 159-197.

Gunesdogan, U., Magnusdottir, E., and Surani, M.A. (2014). Primordial germ cell specification: a context-dependent cellular differentiation event [corrected]. Philos Trans R Soc Lond B Biol Sci *369*.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063-1069.

Gurdon, J.B. (1962). Adult frogs derived from the nuclei of single somatic cells. Dev Biol *4*, 256-273.

Haffner-Krausz, R., Gorivodsky, M., Chen, Y., and Lonai, P. (1999). Expression of Fgfr2 in the early mouse embryo indicates its involvement in preimplantation development. Mech Dev *85*, 167-172.

Hamada, S., Suzuki, T., Mino, K., Koseki, K., Oehme, F., Flamme, I., Ozasa, H., Itoh, Y., Ogasawara, D., Komaarashi, H.*, et al.* (2010). Design, synthesis, enzyme-inhibitory activity, and effect on human cancer cells of a novel series of jumonji domain-containing protein 2 histone demethylase inhibitors. J Med Chem *53*, 5629-5638.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Arauzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T.*, et al.* (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell *143*, 617-627.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519-532.

Hayashi, K., and Surani, M.A. (2009). Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro. Development *136*, 3549-3556.

Hayashi, Y., Furue, M.K., Tanaka, S., Hirose, M., Wakisaka, N., Danno, H., Ohnuma, K., Oeda, S., Aihara, Y., Shiota, K.*, et al.* (2010). BMP4 induction of trophoblast from mouse embryonic stem cells in defined culture conditions on laminin. In Vitro Cell Dev Biol Anim *46*, 416-430.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat Rev Mol Cell Biol *10*, 526-537.

Hirasawa, R., and Sasaki, H. (2009). Dynamic transition of Dnmt3b expression in mouse pre- and early post-implantation embryos. Gene Expr Patterns *9*, 27-30.

Hiratani, I., Leskovar, A., and Gilbert, D.M. (2004). Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores. Proc Natl Acad Sci U S A *101*, 16861-16866.

Hiratani, I., Ryba, T., Itoh, M., Rathjen, J., Kulik, M., Papp, B., Fussner, E., Bazett-Jones, D.P., Plath, K., Dalton, S.*, et al.* (2010). Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. Genome Res *20*, 155-169.

Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schubeler, D., and Gilbert, D.M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol *6*, e245.

Holmqvist, P.H., and Mannervik, M. (2013). Genomic occupancy of the transcriptional co-activators p300 and CBP. Transcription *4*, 18-23.

Hon, G.C., Song, C.X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y., Yen, C.A., Ye, Z., Mao, S.Q., Wang, B.A.*, et al.* (2014). 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. Mol Cell *56*, 286-297.

Hu, D., Garruss, A.S., Gao, X., Morgan, M.A., Cook, M., Smith, E.R., and Shilatifard, A. (2013). The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. Nat Struct Mol Biol *20*, 1093-1097.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev *23*, 837-848.

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

Joo, J.Y., Choi, H.W., Kim, M.J., Zaehres, H., Tapia, N., Stehling, M., Jung, K.S., Do, J.T., and Scholer, H.R. (2014). Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. Sci Rep *4*, 7477.

Jouneau, A., Ciaudo, C., Sismeiro, O., Brochard, V., Jouneau, L., Vandormael-Pournin, S., Coppee, J.Y., Zhou, Q., Heard, E., Antoniewski, C.*, et al.* (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. RNA *18*, 253-264.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kahan, B.W., and Ephrussi, B. (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J Natl Cancer Inst *44*, 1015-1036.

Kaji, K., Caballero, I.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol *8*, 285-292.

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

Katoh, M., and Katoh, M. (2004). Identification and characterization of JMJD2 family genes in silico. Int J Oncol *24*, 1623-1628.

Katoh, Y., and Katoh, M. (2007). Comparative integromics on JMJD2A, JMJD2B and JMJD2C: preferential expression of JMJD2C in undifferentiated ES cells. Int J Mol Med *20*, 269-273.

Keller, G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev *19*, 1129-1155.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006.

Kidder, B.L., Yang, J., and Palmer, S. (2008). Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. PLoS One *3*, e3932.

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

Klose, R.J., Kallin, E.M., and Zhang, Y. (2006a). JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet *7*, 715-727.

Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006b). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature *442*, 312-316.

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

Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. Trends Biochem Sci *30*, 235-239.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C.*, et al.* (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Krupa, M., Mazur, E., Szczepanska, K., Filimonow, K., Maleszewski, M., and Suwinska, A. (2014). Allocation of inner cells to epiblast vs primitive endoderm in the mouse embryo is biased but not determined by the round of asymmetric divisions (8-->16- and 16-->32-cells). Dev Biol *385*, 136-148.

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

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

Kurek, D., Neagu, A., Tastemel, M., Tuysuz, N., Lehmann, J., van de Werken, H.J., Philipsen, S., van der Linden, R., Maas, A., van, I.W.F.*, et al.* (2015). Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. Stem Cell Reports *4*, 114-128.

Labbe, R.M., Holowatyj, A., and Yang, Z.Q. (2013). Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential. Am J Transl Res *6*, 1-15.

Labosky, P.A., Barlow, D.P., and Hogan, B.L. (1994). Embryonic germ cell lines and their derivation from mouse primordial germ cells. Ciba Found Symp *182*, 157-168; discussion 168-178.

Laflamme, M.A., Chen, K.Y., Naumova, A.V., Muskheli, V., Fugate, J.A., Dupras, S.K., Reinecke, H., Xu, C., Hassanipour, M., Police, S.*, et al.* (2007). Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. Nat Biotechnol *25*, 1015-1024.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E.*, et al.* (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. Nature *449*, 689-694.

Lander, E.S. (2011). Initial impact of the sequencing of the human genome. Nature *470*, 187-197.

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

Latos, P.A., Helliwell, C., Mosaku, O., Dudzinska, D.A., Stubbs, B., Berdasco, M., Esteller, M., and Hendrich, B. (2012). NuRD-dependent DNA methylation prevents ES cells from accessing a trophectoderm fate. Biol Open *1*, 341-352.

Lavial, F., Bessonnard, S., Ohnishi, Y., Tsumura, A., Chandrashekran, A., Fenwick, M.A., Tomaz, R.A., Hosokawa, H., Nakayama, T., Chambers, I.*, et al.* (2012). Bmi1 facilitates primitive endoderm formation by stabilizing Gata6 during early mouse development. Genes Dev *26*, 1445-1458.

Lee, D.Y., Northrop, J.P., Kuo, M.H., and Stallcup, M.R. (2006a). Histone H3 lysine 9 methyltransferase G9a is a transcriptional coactivator for nuclear receptors. J Biol Chem *281*, 8476-8485.

Lee, J.H., Voo, K.S., and Skalnik, D.G. (2001). Identification and characterization of the DNA binding domain of CpG-binding protein. J Biol Chem *276*, 44669-44676.

Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science *318*, 447-450.

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

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006b). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lehmann, U., Schmitz, J., Weissenbach, M., Sobota, R.M., Hortner, M., Friederichs, K., Behrmann, I., Tsiaris, W., Sasaki, A., Schneider-Mergener, J.*, et al.* (2003). SHP2 and SOCS3 contribute to Tyr-759-dependent attenuation of interleukin-6 signaling through gp130. J Biol Chem *278*, 661-671.

Lehnertz, B., Northrop, J.P., Antignano, F., Burrows, K., Hadidi, S., Mullaly, S.C., Rossi, F.M., and Zaph, C. (2010). Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function. J Exp Med *207*, 915-922.

Lienert, F., Mohn, F., Tiwari, V.K., Baubec, T., Roloff, T.C., Gaidatzis, D., Stadler, M.B., and Schubeler, D. (2011). Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. PLoS Genet *7*, e1002090.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lizcano, F., Romero, C., and Vargas, D. (2011). Regulation of adipogenesis by nuclear receptor PPARgamma is modulated by the histone demethylase JMJD2C. Genet Mol Biol *34*, 19-24.

Lodato, M.A., Ng, C.W., Wamstad, J.A., Cheng, A.W., Thai, K.K., Fraenkel, E., Jaenisch, R., and Boyer, L.A. (2013). SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. PLoS Genet *9*, e1003288.

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

Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A.*, et al.* (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474-478.

Lukas, J., Lukas, C., and Bartek, J. (2011). More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. Nat Cell Biol *13*, 1161-1169.

Mali, P., Esvelt, K.M., and Church, G.M. (2013). Cas9 as a versatile tool for engineering biology. Nat Methods *10*, 957-963.

Malik, S., and Roeder, R.G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. Trends Biochem Sci *30*, 256-263.

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

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem *271*, 27176-27183.

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

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

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

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

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

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

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

McPherson, J.D., Marra, M., Hillier, L., Waterston, R.H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E.R., Wilson, R.K.*, et al.* (2001). A physical map of the human genome. Nature *409*, 934-941.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B.*, et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766-770.

Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjosukarso, D.W., Ng, Y., Moreau, T., and Pedersen, R.A. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. Cell Stem Cell *15*, 310-325.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell *10*, 105-116.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

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

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell *30*, 755-766.

Mohr, S.E., Smith, J.A., Shamu, C.E., Neumuller, R.A., and Perrimon, N. (2014). RNAi screening comes of age: improved techniques and complementary approaches. Nat Rev Mol Cell Biol *15*, 591-600.

Morgani, S.M., and Brickman, J.M. (2014). The molecular underpinnings of totipotency. Philos Trans R Soc Lond B Biol Sci *369*.

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

Morris, S.A., Graham, S.J., Jedrusik, A., and Zernicka-Goetz, M. (2013). The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. Open Biol *3*, 130104.

Morris, S.A., Teo, R.T., Li, H., Robson, P., Glover, D.M., and Zernicka-Goetz, M. (2010). Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. Proc Natl Acad Sci U S A *107*, 6364-6369.

Mouse Genome Sequencing, C., Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M.*, et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.

Mozzetta, C., Pontis, J., Fritsch, L., Robin, P., Portoso, M., Proux, C., Margueron, R., and Ait-Si-Ali, S. (2014). The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol Cell *53*, 277-289.

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu Rev Genet *43*, 525-558.

Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N.P., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M.*, et al.* (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. EMBO J *31*, 4547-4562.

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

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A.*, et al.* (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes Dev *24*, 312-326.

Niakan, K.K., Schrode, N., Cho, L.T., and Hadjantonakis, A.K. (2013). Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. Nat Protoc *8*, 1028-1041.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

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

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

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

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

Nottke, A., Colaiacovo, M.P., and Shi, Y. (2009). Developmental roles of the histone lysine demethylases. Development *136*, 879-889.

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

Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., and Miyazono, K. (2007). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. J Cell Sci *120*, 55-65.

Oh, S.T., Kim, K.B., Chae, Y.C., Kang, J.Y., Hahn, Y., and Seo, S.B. (2014). H3K9 histone methyltransferase G9a-mediated transcriptional activation of p21. FEBS Lett *588*, 685-691.

Ohnuki, M., and Takahashi, K. (2015). Present and future challenges of induced pluripotent stem cells. Philos Trans R Soc Lond B Biol Sci *370*.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D.*, et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Osorno, R., and Chambers, I. (2011). Transcription factor heterogeneity and epiblast pluripotency. Philos Trans R Soc Lond B Biol Sci *366*, 2230-2237.

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Development *139*, 2288-2298.

Paca, A., Seguin, C.A., Clements, M., Ryczko, M., Rossant, J., Rodriguez, T.A., and Kunath, T. (2012). BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. Dev Biol *361*, 90-102.

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

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

Pasini, D., Hansen, K.H., Christensen, J., Agger, K., Cloos, P.A., and Helin, K. (2008). Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. Genes Dev *22*, 1345-1355.

Pauler, F.M., Sloane, M.A., Huang, R., Regha, K., Koerner, M.V., Tamir, I., Sommer, A., Aszodi, A., Jenuwein, T., and Barlow, D.P. (2009). H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. Genome Res *19*, 221-233.

Pedersen, M.T., Agger, K., Laugesen, A., Johansen, J.V., Cloos, P.A., Christensen, J., and Helin, K. (2014). The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development. Mol Cell Biol *34*, 1031-1045.

Pekowska, A., Benoukraf, T., Zacarias-Cabeza, J., Belhocine, M., Koch, F., Holota, H., Imbert, J., Andrau, J.C., Ferrier, P., and Spicuglia, S. (2011). H3K4 tri-methylation provides an epigenetic signature of active enhancers. EMBO J *30*, 4198-4210.

Percharde, M., Lavial, F., Ng, J.H., Kumar, V., Tomaz, R.A., Martin, N., Yeo, J.C., Gil, J., Prabhakar, S., Ng, H.H.*, et al.* (2012). Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. Genes Dev *26*, 2286-2298.

Pereira, C.F., Terranova, R., Ryan, N.K., Santos, J., Morris, K.J., Cui, W., Merkenschlager, M., and Fisher, A.G. (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. PLoS Genet *4*, e1000170.

Perry, P., Sauer, S., Billon, N., Richardson, W.D., Spivakov, M., Warnes, G., Livesey, F.J., Merkenschlager, M., Fisher, A.G., and Azuara, V. (2004). A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction. Cell Cycle *3*, 1645-1650.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45.

Pirity, M.K., Locker, J., and Schreiber-Agus, N. (2005). Rybp/DEDAF is required for early postimplantation and for central nervous system development. Mol Cell Biol *25*, 7193-7202.

Plusa, B., and Hadjantonakis, A.K. (2014). Embryonic stem cell identity grounded in the embryo. Nat Cell Biol *16*, 502-504.

Ponnaluri, V.K., Vavilala, D.T., Putty, S., Gutheil, W.G., and Mukherji, M. (2009). Identification of non-histone substrates for JMJD2A-C histone demethylases. Biochem Biophys Res Commun *390*, 280-284.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-512.

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

Qin, B., Zhou, M., Ge, Y., Taing, L., Liu, T., Wang, Q., Wang, S., Chen, J., Shen, L., Duan, X., *et al.* (2012a). CistromeMap: a knowledgebase and web server for ChIP-Seq and DNase-Seq studies in mouse and human. Bioinformatics *28*, 1411-1412.

Qin, J., Whyte, W.A., Anderssen, E., Apostolou, E., Chen, H.H., Akbarian, S., Bronson, R.T., Hochedlinger, K., Ramaswamy, S., Young, R.A*., et al.* (2012b). The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. Cell Stem Cell *11*, 319-332.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

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

Rapkin, L.M., Anchel, D.R., Li, R., and Bazett-Jones, D.P. (2012). A view of the chromatin landscape. Micron *43*, 150-158.

Ratz, M., Testa, I., Hell, S.W., and Jakobs, S. (2015). CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. Sci Rep *5*, 9592.

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

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425-432.

Resnick, J.L., Bixler, L.S., Cheng, L., and Donovan, P.J. (1992). Long-term proliferation of mouse primordial germ cells in culture. Nature *359*, 550-551.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Rosa, S., and Shaw, P. (2013). Insights into chromatin structure and dynamics in plants. Biology (Basel) *2*, 1378-1410.

Rossant, J. (2008). Stem cells and early lineage development. Cell *132*, 527-531.

Rossant, J., Chazaud, C., and Yamanaka, Y. (2003). Lineage allocation and asymmetries in the early mouse embryo. Philos Trans R Soc Lond B Biol Sci *358*, 1341-1348; discussion 1349.

Rossant, J., Sanford, J.P., Chapman, V.M., and Andrews, G.K. (1986). Undermethylation of structural gene sequences in extraembryonic lineages of the mouse. Dev Biol *117*, 567-573.

Rugg-Gunn, P.J., Cox, B.J., Ralston, A., and Rossant, J. (2010). Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. Proc Natl Acad Sci U S A *107*, 10783-10790.

Rula, M.E., Cai, K.Q., Moore, R., Yang, D.H., Staub, C.M., Capo-Chichi, C.D., Jablonski, S.A., Howe, P.H., Smith, E.R., and Xu, X.X. (2007). Cell autonomous sorting and surface positioning in the formation of primitive endoderm in embryoid bodies. Genesis *45*, 327-338.

Sadlon, T.J., Lewis, I.D., and D'Andrea, R.J. (2004). BMP4: its role in development of the hematopoietic system and potential as a hematopoietic growth factor. Stem Cells *22*, 457-474.

Sampath, S.C., Marazzi, I., Yap, K.L., Sampath, S.C., Krutchinsky, A.N., Mecklenbrauker, I., Viale, A., Rudensky, E., Zhou, M.M., Chait, B.T*., et al.* (2007). Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell *27*, 596-608.

Sancho, M., Di-Gregorio, A., George, N., Pozzi, S., Sanchez, J.M., Pernaute, B., and Rodriguez, T.A. (2013). Competitive interactions eliminate unfit embryonic stem cells at the onset of differentiation. Dev Cell *26*, 19-30.

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

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109-113.

Schiemann, W.P., Bartoe, J.L., and Nathanson, N.M. (1997). Box 3-independent signaling mechanisms are involved in leukemia inhibitory factor receptor alpha- and gp130-mediated stimulation of mitogen-activated protein kinase. Evidence for participation of multiple signaling pathways which converge at Ras. J Biol Chem *272*, 16631-16636.

Scholer, H.R. (1991). Octamania: the POU factors in murine development. Trends Genet *7*, 323-329.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745.

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

Serandour, A.A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Metivier, R.*, et al.* (2011). Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. Genome Res *21*, 555-565.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nat Protoc *2*, 1445-1457.

Shen, M.M. (2007). Nodal signaling: developmental roles and regulation. Development *134*, 1023-1034.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

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

Shin, S., and Janknecht, R. (2007). Diversity within the JMJD2 histone demethylase family. Biochem Biophys Res Commun *353*, 973-977.

Shinkai, Y., and Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. Genes Dev *25*, 781-788.

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

Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T.*, et al.* (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature *474*, 337-342.

Soares, M.L., Torres-Padilla, M.E., and Zernicka-Goetz, M. (2008). Bone morphogenetic protein 4 signaling regulates development of the anterior visceral endoderm in the mouse embryo. Dev Growth Differ *50*, 615-621.

Soprano, D.R., Teets, B.W., and Soprano, K.J. (2007). Role of retinoic acid in the differentiation of embryonal carcinoma and embryonic stem cells. Vitam Horm *75*, 69-95.

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

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

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

Stower, M.J., and Srinivas, S. (2014). Heading forwards: anterior visceral endoderm migration in patterning the mouse embryo. Philos Trans R Soc Lond B Biol Sci *369*.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol *12*, R54.

Sun, H., Qin, B., Liu, T., Wang, Q., Liu, J., Wang, J., Lin, X., Yang, Y., Taing, L., Rao, P.K.*, et al.* (2013). CistromeFinder for ChIP-seq and DNase-seq data reuse. Bioinformatics *29*, 1352-1354.

Sun, L.T., Yamaguchi, S., Hirano, K., Ichisaka, T., Kuroda, T., and Tada, T. (2014). Nanog co-regulated by Nodal/Smad2 and Oct4 is required for pluripotency in developing mouse epiblast. Dev Biol *392*, 182-192.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H.*, et al.* (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev *16*, 1779-1791.

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

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

Takaoka, K., and Hamada, H. (2014). Origin of cellular asymmetries in the pre-implantation mouse embryo: a hypothesis. Philos Trans R Soc Lond B Biol Sci *369*.

Tam, P.P., and Loebel, D.A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. Nat Rev Genet *8*, 368-381.

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

Tesar, P.J. (2005). Derivation of germ-line-competent embryonic stem cell lines from preblastocyst mouse embryos. Proc Natl Acad Sci U S A *102*, 8239-8244.

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

Thomas, K.R., and Capecchi, M.R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell *51*, 503-512.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131-3141.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res *30*, 3202-3213.

Torres-Padilla, M.E., and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. Development *141*, 2173-2181.

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

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

Trojer, P., Zhang, J., Yonezawa, M., Schmidt, A., Zheng, H., Jenuwein, T., and Reinberg, D. (2009). Dynamic Histone H1 Isotype 4 Methylation and Demethylation by Histone Lysine Methyltransferase G9a/KMT1C and the Jumonji Domain-containing JMJD2/KDM4 Proteins. J Biol Chem *284*, 8395-8405.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

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

Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A.*, et al.* (2009b). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. Stem Cells *27*, 2655-2666.

Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A.*, et al.* (2009c). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. PLoS One *4*, e6082.

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

Vanhaesebroeck, B., Leevers, S.J., Ahmadi, K., Timms, J., Katso, R., Driscoll, P.C., Woscholski, R., Parker, P.J., and Waterfield, M.D. (2001). Synthesis and function of 3-phosphorylated inositol lipids. Annu Rev Biochem *70*, 535-602.

Veillard, A.C., Marks, H., Bernardo, A.S., Jouneau, L., Laloe, D., Boulanger, L., Kaan, A., Brochard, V., Tosolini, M., Pedersen, R.*, et al.* (2014). Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. Stem Cells Dev *23*, 2014-2029.

Verrier, L., Escaffit, F., Chailleux, C., Trouche, D., and Vandromme, M. (2011). A new isoform of the histone demethylase JMJD2A/KDM4A is required for skeletal muscle differentiation. PLoS Genet *7*, e1001390.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci U S A *100*, 2468-2473.

Waddington, C.H. (1957). The strategy of the genes; a discussion of some aspects of theoretical biology (London,: Allen & Unwin).

Walsh, C.A., Bolger, J.C., Byrne, C., Cocchiglia, S., Hao, Y., Fagan, A., Qin, L., Cahalin, A., McCartan, D., McIlroy, M.*, et al.* (2014). Global gene repression by the steroid receptor coactivator SRC-1 promotes oncogenesis. Cancer Res *74*, 2533-2544.

Wamaitha, S.E., del Valle, I., Cho, L.T., Wei, Y., Fogarty, N.M., Blakeley, P., Sherwood, R.I., Ji, H., and Niakan, K.K. (2015). Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. Genes Dev *29*, 1239-1255.

Wang, J., Zhang, M., Zhang, Y., Kou, Z., Han, Z., Chen, D.Y., Sun, Q.Y., and Gao, S. (2010). The histone demethylase JMJD2C is stage-specifically expressed in preimplantation mouse embryos and is required for embryonic development. Biol Reprod *82*, 105-111.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell *138*, 1019-1031.

Waterston, R.H., Lander, E.S., and Sulston, J.E. (2002). On the sequencing of the human genome. Proc Natl Acad Sci U S A *99*, 3712-3716.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M.*, et al.* (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.

White, J.K., Gerdin, A.K., Karp, N.A., Ryder, E., Buljan, M., Bussell, J.N., Salisbury, J., Clare, S., Ingham, N.J., Podrini, C.*, et al.* (2013). Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell *154*, 452-464.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.

Widelitz, R. (2005). Wnt signaling through canonical and non-canonical pathways: recent progress. Growth Factors *23*, 111-116.

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

Wrana, J.L., Attisano, L., Carcamo, J., Zentella, A., Doody, J., Laiho, M., Wang, X.F., and Massague, J. (1992). TGF beta signals through a heteromeric protein kinase receptor complex. Cell *71*, 1003-1014.

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

Wu, L., Wary, K.K., Revskoy, S., Gao, X., Tsang, K., Komarova, Y.A., Rehman, J., and Malik, A.B. (2015). Histone Demethylases KDM4A and KDM4C Regulate Differentiation of Embryonic Stem Cells to Endothelial Cells. Stem Cell Reports *5*, 10-21.

Wu, X., Johansen, J.V., and Helin, K. (2013). Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Mol Cell *49*, 1134-1146.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Xu, C.R., Cole, P.A., Meyers, D.J., Kormish, J., Dent, S., and Zaret, K.S. (2011). Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. Science *332*, 963-966.

Yamanaka, S., and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. Nature *465*, 704-712.

Yang, S.H., Kalkan, T., Morissroe, C., Marks, H., Stunnenberg, H., Smith, A., and Sharrocks, A.D. (2014). Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. Cell Rep *7*, 1968-1981.

Yang, Z.Q., Imoto, I., Fukuda, Y., Pimkhaokham, A., Shimada, Y., Imamura, M., Sugano, S., Nakamura, Y., and Inazawa, J. (2000). Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. Cancer Res *60*, 4735-4739.

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

Ye, L., Fan, Z., Yu, B., Chang, J., Al Hezaimi, K., Zhou, X., Park, N.H., and Wang, C.Y. (2012). Histone demethylases KDM4B and KDM6B promotes osteogenic differentiation of human MSCs. Cell Stem Cell *11*, 50-61.

Yeo, J.C., and Ng, H.H. (2013). The transcriptional regulation of pluripotency. Cell Res *23*, 20-32.

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

Ying, Q.L., Nichols, J., Evans, E.P., and Smith, A.G. (2002). Changing potency by spontaneous fusion. Nature *416*, 545-548.

Ying, Q.L., Stavridis, M., Griffiths, D., Li, M., and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol *21*, 183-186.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T., and Kishimoto, T. (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. Mech Dev *45*, 163-171.

Yu, P., Xiao, S., Xin, X., Song, C.X., Huang, W., McDee, D., Tanaka, T., Wang, T., He, C., and Zhong, S. (2013). Spatiotemporal clustering of the epigenome reveals rules of dynamic gene regulation. Genome Res *23*, 352-364.

Yu, V., Fisch, T., Long, A.M., Tang, J., Lee, J.H., Hierl, M., Chen, H., Yakowec, P., Schwandner, R., and Emkey, R. (2012). High-throughput TR-FRET assays for identifying inhibitors of LSD1 and JMJD2C histone lysine demethylases. J Biomol Screen *17*, 27-38.

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

Yuan, P., Han, J., Guo, G., Orlov, Y.L., Huss, M., Loh, Y.H., Yaw, L.P., Robson, P., Lim, B., and Ng, H.H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev *23*, 2507-2520.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev *25*, 2227-2241.

Zeineddine, D., Papadimou, E., Chebli, K., Gineste, M., Liu, J., Grey, C., Thurig, S., Behfar, A., Wallace, V.A., Skerjanc, I.S.*, et al.* (2006). Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early heart development. Dev Cell *11*, 535-546.

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res *21*, 1273-1283.

Zernicka-Goetz, M., Morris, S.A., and Bruce, A.W. (2009). Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. Nat Rev Genet *10*, 467-477.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhou, H., Li, W., Zhu, S., Joo, J.Y., Do, J.T., Xiong, W., Kim, J.B., Zhang, K., Scholer, H.R., and Ding, S. (2010). Conversion of mouse epiblast stem cells to an earlier pluripotency state by small molecules. J Biol Chem *285*, 29676-29680.

Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.

Appendix A – Figures

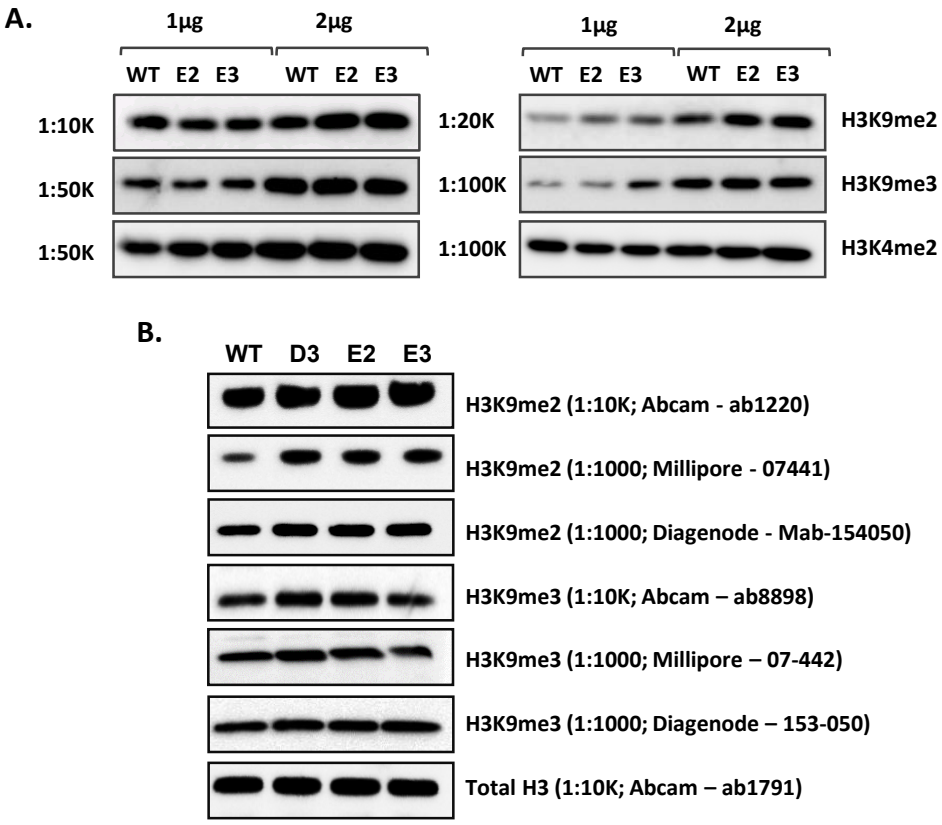


Figure A-I Titration of histone lysates and validation of antibody sources for detection of H3K9me2/3 levels in *Jmjd2c*-knockout ESCs

A. 1 and 2µg of acid-extracted histone lysates collected from WT and *Jmjd2c*-knockout clones E2 and E3 were used for detection of H3K9me2/3 levels by western blot (source of antibodies: Abcam). The two panels represent two different concentration so antibody as indicated. H3K4me2 were used as a loading control. **B**. Detection of H3K9me2 and H3K9me3 with three different antibody sources as indicated. Western blots were performed in 2µg of acid-extracted histone lysates collected from WT, the *Jmjd2c* heterozygous-knockout clone D3 and the *Jmjd2c*-knockout clones E2 and E3. Total H3 was used as a loading control.

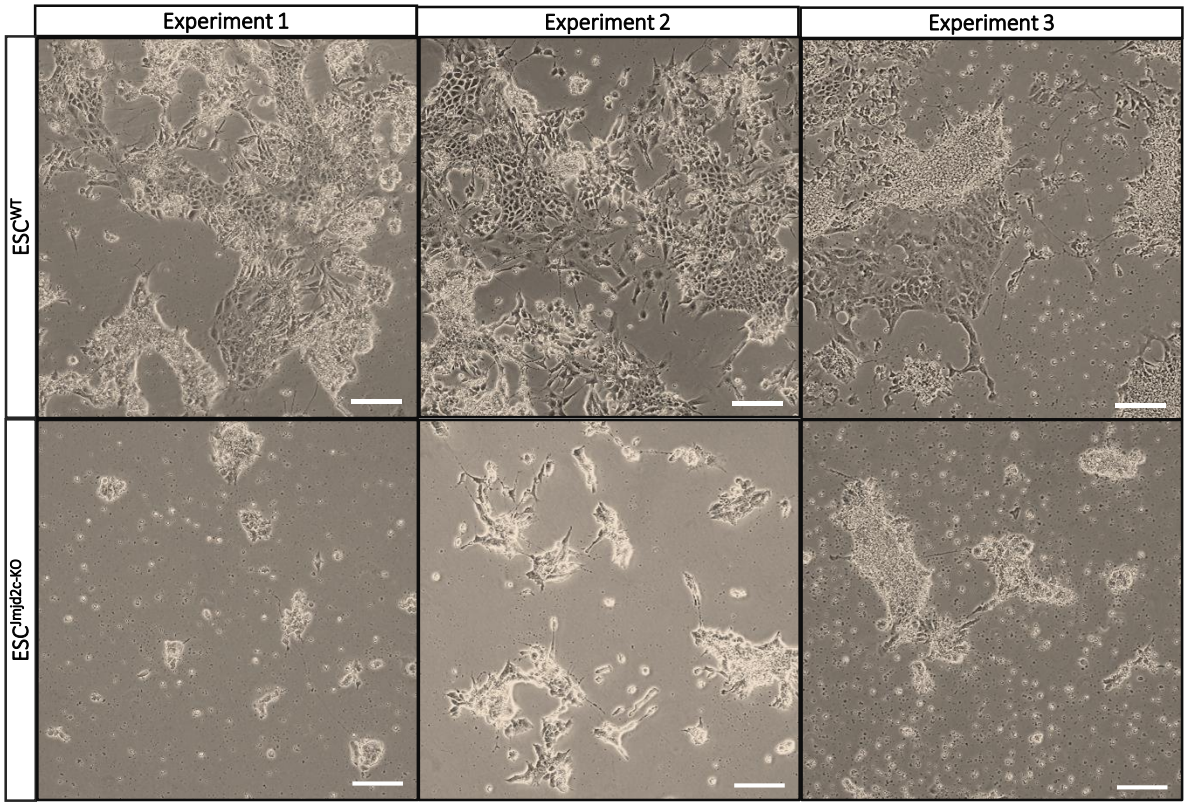


Figure A-II Morphology of neuronal progenitors at day 5

The three panels of phase-contrast images represent the cell morphology observed at the final time point (day 5) of three independent neuronal differentiation experiments performed in WT (ESC^{WT}) and KO ($ESC^{/mjd2c\text{-}KO}$) ESCs. Images labelled as experiment 2 correspond to different sections of the same well of the experiment shown in Fig.3.7B. Bars, 100µm.

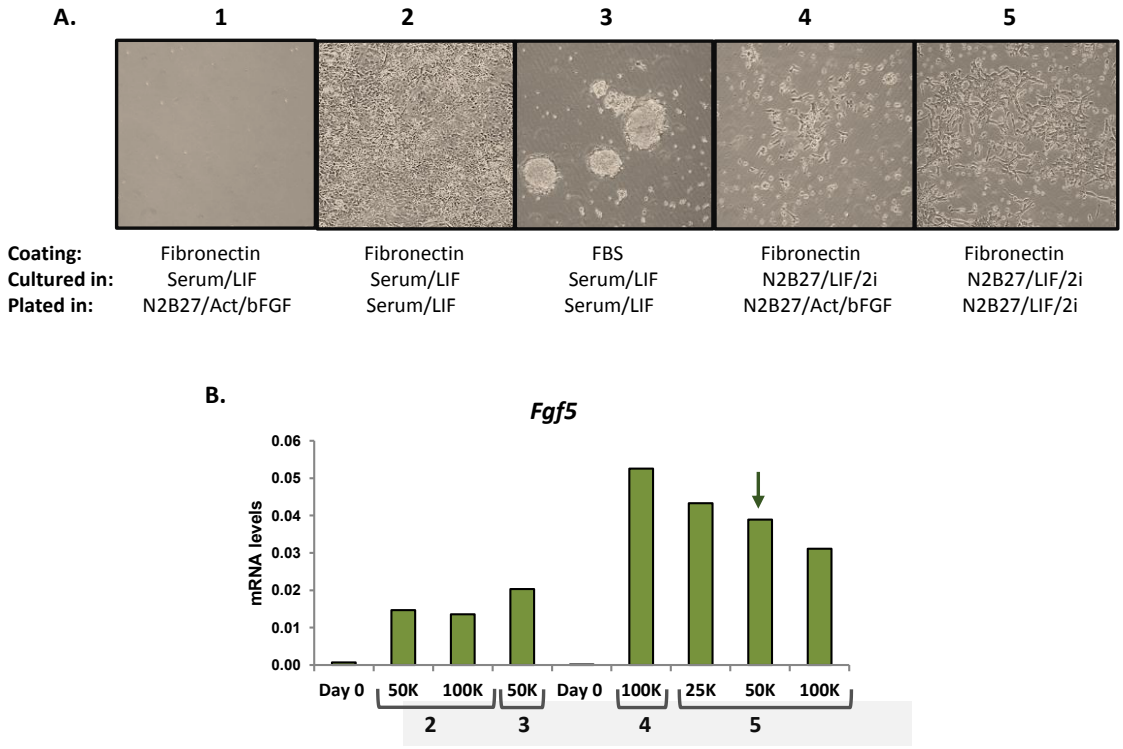


Figure A-III Optimization of ESC-to-EpiSC conversion

A. Representative images of WT ESCs after 4 days of ESC-to-EpiSC conversion, with the indicated conditions: WT ESCs were routinely cultured in serum/LIF or in 2i/LIF (N2B27/2i/LIF) conditions and plated in wells coated with either Fibronectin or FBS. Cells were plated in either the aforementioned medium routinely used in ESC culture for 24h before swapping into EpiSC conditions (N2B27/Act/bFGF), or directly in EpiSC conditions. **B.** *Fgf5* transcript levels quantified after 4 days of conversion in the different tested conditions as indicated below the graph (n=1). Where indicated, conversion experiments were tested with different starting cell numbers. The arrow indicates the conversion conditions selected to perform the experiments depicted in Fig.4.3.

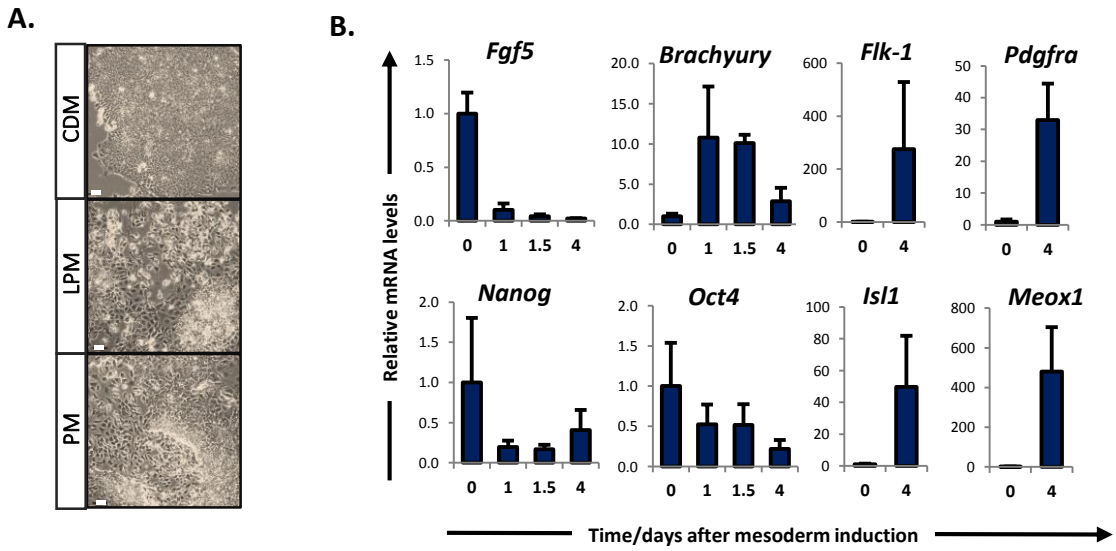


Figure A-IV Mesoderm induction in parental EpiSC

A. Representative phase-contrast images of embryo-derived EpiSC cultured in CDM-PVA and day 4 differentiated EpiSCs in FB40 (PM) or FlyWL (LPM) conditions. Bars, 100µm. **B.** Transcript levels of *Fgf5*, *Brachyury/T*, *Nanog, Oct4* in cEpiSC (day 0) differentiated with FlyB for 1 and 1.5 days, and FB40 at day 4. Transcript levels of *Flk-1,* and *Isl1;* and *Pdgfra* and *Meox1,* in EpiSC prior and after 4 day LPM and PM differentiation, respectively. Data were normalized to housekeeping genes (*Hmbs, Ywak* and *Gpdh*) and represented as average of at least 3 independent experiments ±SEM. P<0.01 (**), *t*-test. Results in this figure were produced by Jennifer Harman.

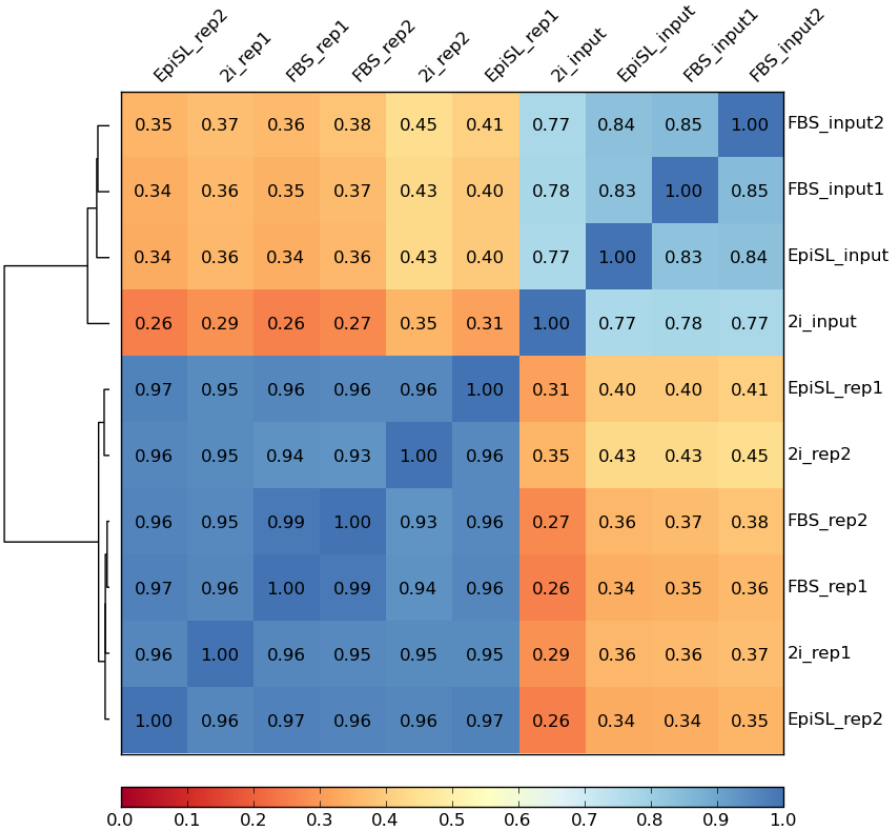


Figure A.V Correlation heatmap for ChIP-seq biological replicates

Pearson correlation was performed using the BAM files for each replicate and input files, with a genome bin size of 1000bp. Scale represents the degree of correlation with 0 being no correlation and 1 representing perfect correlation.

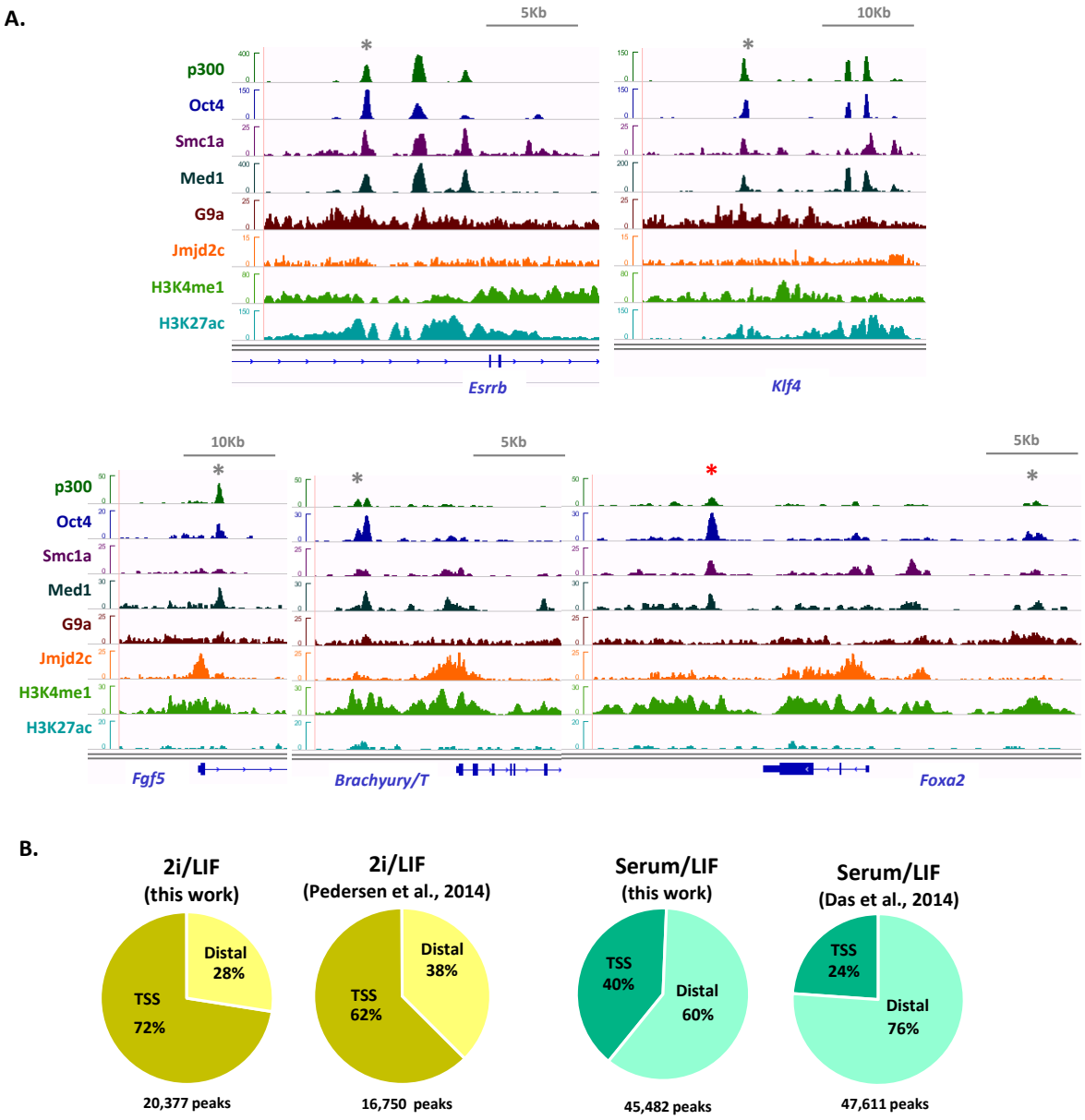


Figure A.VI Enrichment profiles across enhancers and peak location in published datasets

A. Enrichment profiles for p300, Oct4, Smc1a, Med1, G9a, H3K4me1, H3K27ac from published datasets (outlined in Section 2.2.9.3), and the Jmjd2c serum/LIF dataset generated in this work. Active and poised enhancers are distinguished by the presence of H3K27ac. Grey (*) indicates the candidate regions analysed in Fig.5.6B, whereas red (*) indicates a second poised enhancer regions noticeable at the *Foxa2* locus, which was not analysed in this work. **B.** Location of annotated peaks identified in this work and in published datasets. Peak calling in published datasets was performed with MACS 1.4 using default parameters.

A.

	Experiment 1						Experiment 2						Experiment 3					
	F2V5-Jmjd2c-cl#3			Control			F2V5-Jmjd2c-cl#3			Control			F2V5-Jmjd2c-cl#8			Control		
symbol	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.
Kdm4c	**5.95**	**3150**	**43**				**4.31**	**2315**	**39**				**12.4**	**3735**	**50**			
Kdm3a	0.02	101	1				0.02	48	1				0.19	360	6	0.08	173	3
Kdm4b	0.21	189	3										0.29	187	4			
* Ehmt1	0.13	238	4				0.08	178	5					65	2		68	1
L3mbtl2	0.04	87	1							0.04	67	2	0.04	54	1			
* Mbd3	8.61	830	12	9.76	930	13	0.76	245	5				0.57	116	3	0.4	98	2
Ppp1cc	1.14	395	7	1.84	592	9	0.21	122	4				0.61	234	4			
Prmt5	0.16	64	1				0.28	244	6									
Arid3a		110	1							0.12	106	3	0.82	102	2	0.25		
Leo1							0.15	182	5				0.1	68	2			
Rbpj							0.07	66	2				0.07	45	1			
Cpsf6	0.06	54	1	0.2			0.13	116	2				0.27	276	4	0.06	43	1
Cstf3				0.09	124	2	0.14	117	3				0.09	66	2			
Ddx1				0.14	163	3	0.24	186	5	0.04	41	1	0.04	52	1			
Ddx19a		56	1		86	1		155	2				0.7	414	9	0.22	136	3
Dkc1	0.21	102	2	0.21	112	2	0.13	74	2				0.13	103	3			
Fip1l1							0.18	166	3				0.25	200	4			
Pabpc1	0.75	733	11	1.92	1191	18	0.67	465	10		140	3	0.94	597	13	0.23	195	4
Plrg1	0.07	47	1	0.14	94	2	0.21	212	4	0.07	41	1	0.21	112	3			
Snrpb2	0.55	83	2				0.55	90	2				0.92	146	3			
Srsf12	0.12	59	1										0.26	99	2			
Thoc6				0.24	138	2	0.38	153	3				0.54	188	4			
Ctnnbl1	0.06	47	1				0.06	51	1				0.12	86	2	0.06	56	1
Pcnp	0.33	62	1	0.56	129	2	1.52	224	5	0.56	72	2						
Poldip2					351	5	0.39	119	3				0.11	64	1			
Msh6		548	8	0.02	67	1	0.27	452	10	0.02	75	2	0.1	263	5	0.02	53	1
Champ1	0.04	42	1										0.04	53	1			
Bclaf1							0.04	57	2				0.09	70	2			

B.

			Experiment 1						Experiment 2						Experiment 3					
			F2V5-Jmjd2c-cl#3			Control			F2V5-Jmjd2c-cl#3			Control			F2V5-Jmjd2c-cl#8			Control		
	Acession	symbol	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.	emPAI	Mascot	Pept.
*	E9Q5A3	Ehmt1	0.13	238	4				0.08	178	5					65	2		68	1
	Q5DW34	Ehmt1		238	4	0.13	240	4		178	5		410	12		65	2		68	1
	Z4YJZ7	Ehmt1								150	4	0.22	410	12	0.05	65	2	0.03	68	1
	Q9Z2D8	Mbd3	8.05	865	12	9.1	955	13	1.16	319	6	1.69	383	7		116	3		98	2
*	D3YTR4	Mbd3	8.61	830	12	9.76	930	13	0.76	245	5				0.57	116	3	0.4	98	2
	D3YTR5	Mbd3														116	3		98	2
	F6WZH1	Mbd3														116	3			

Figure A.VII List of peptides identified in FLAG-Jmjd2c pull-down

A. Full list of peptides recovered after Jmjd2c pull-down, with known location in the nucleus. **B.** Multiple hits identified for the proteins highlighted with (*) in **A**.

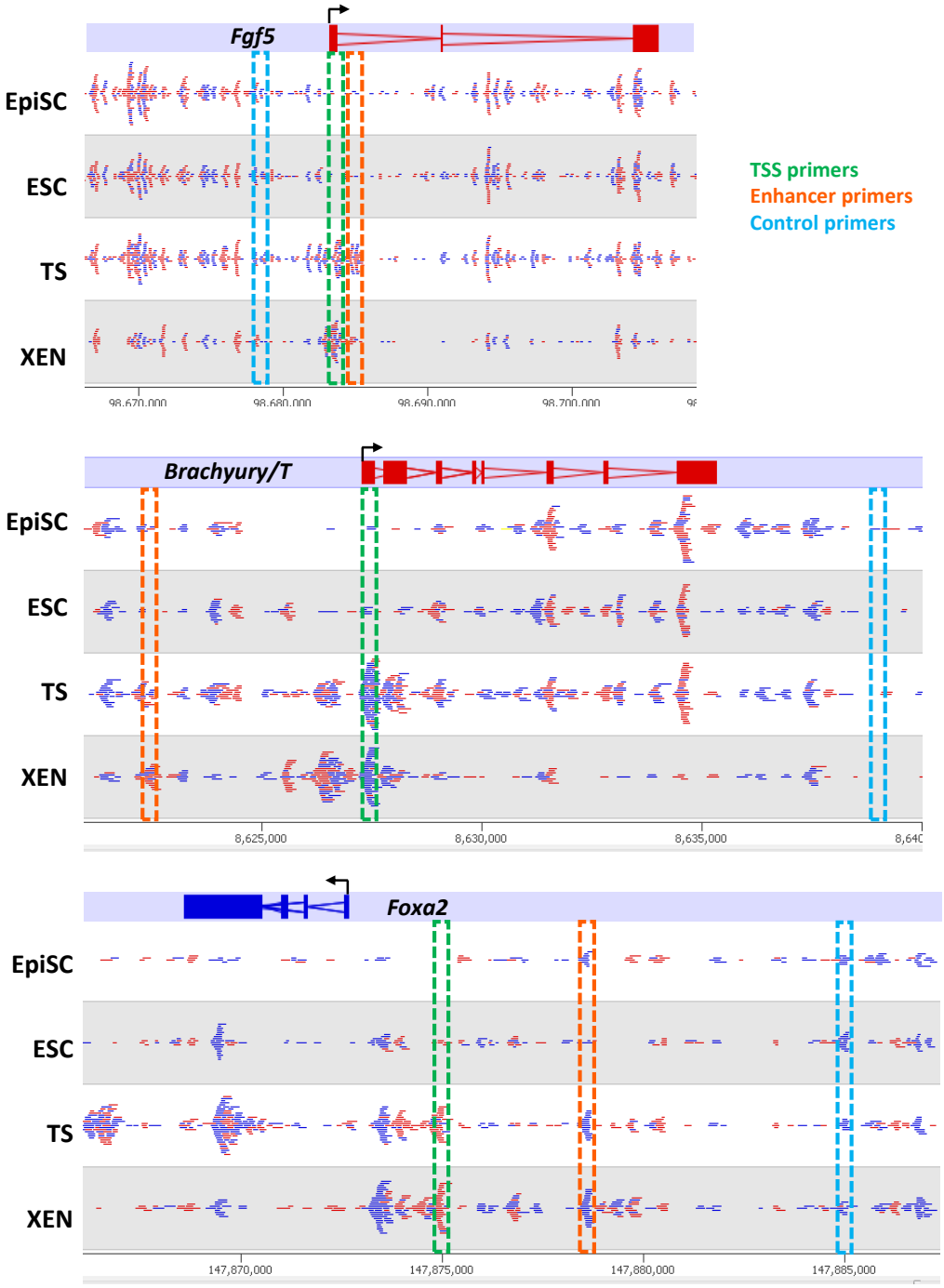


Figure A-VIII DNA methylation profiles of EpiSC, ESCs, TS and XEN cells and primer locations

Screenshots of methylation DNA immunoprecipitation (MeDIP)-seq profiles of EpiSC, ESC, TS and XEN cells across *Fgf5*, *Brachyury/T* and *Foxa2* gene loci. Horizontal red and blue lines show the number of reads indicative of methylation levels detected in the sense and antisense strands, respectively. Primers designed in TSS, enhancer or control regions are depicted in green, orange and blue. Data from Senner et al., 2012; screenshots obtained with the software SeqMonk (Babraham Institute, Cambridge).

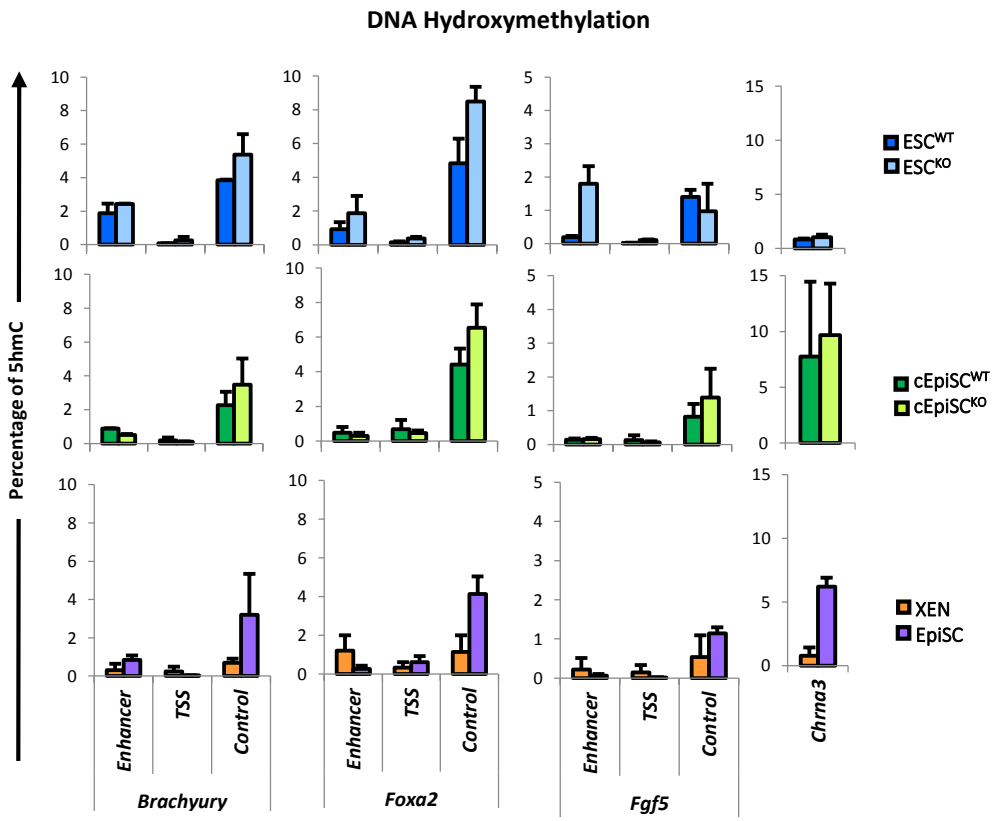


Figure A-IX Levels of DNA hydroxymethylation at TSS and enhancers of differentiation-associated genes

DNA hydroxymethylation correspond to the MspI-resistant fraction in WT and *Jmjd2c*-KO ESCs (ESC^{WT} and ESC^{KO}; blue) and WT and *Jmjd2c*-KO cEpiSCs ($cEpiSC^{WT}$ and $cEpiSC^{KO}$; green). Levels were normalized to a control region without an MspI/HpaII restriction site and represented relative to total/undigested DNA (100%). The average of 2 experiments ± SD is plotted. Regions represented (enhancer and TSS) are the same as described in Fig.6.1B.

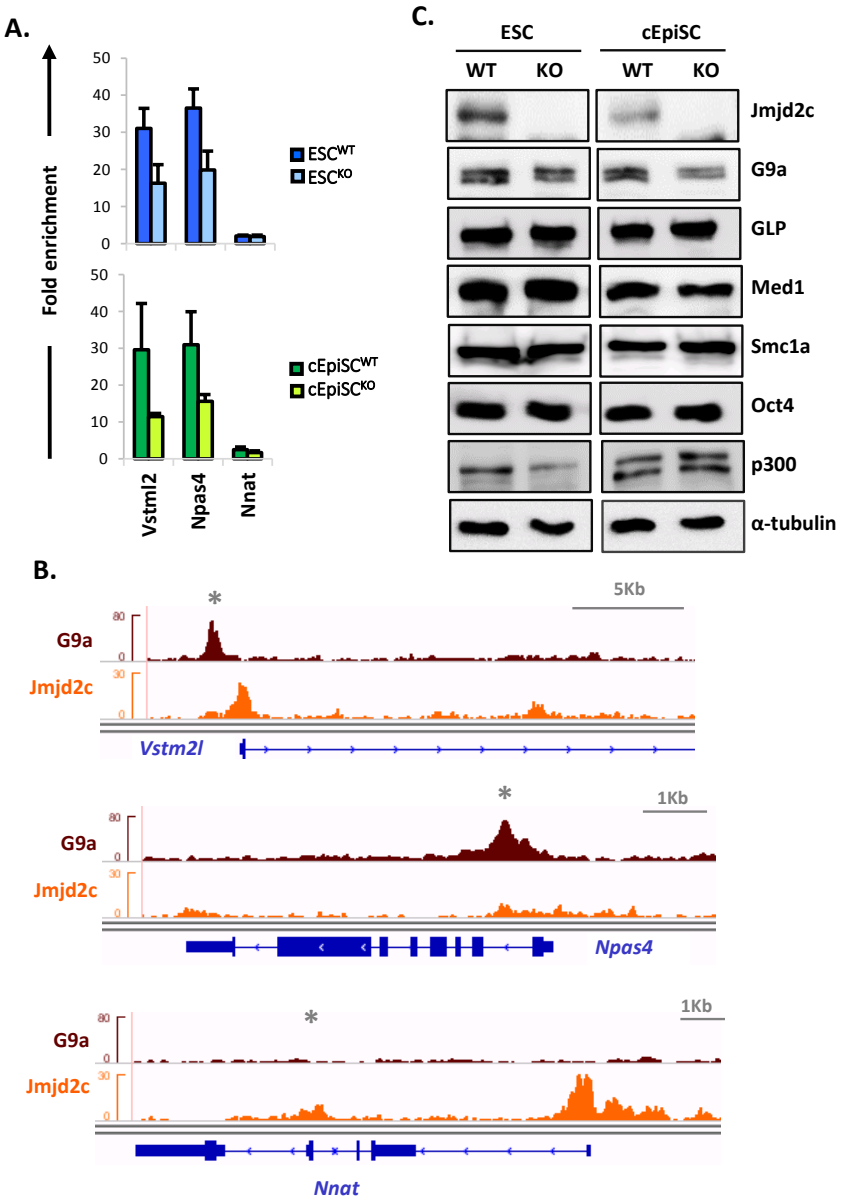


Figure A-X Validation of G9a ChIP efficiency and quantification of enhancer proteins in *Jmjd2c*-depleted pluripotent cells

A. Enrichment levels for G9a as assessed by ChIP-qPCR in in WT and *Jmjd2c*-KO ESCs (ESC^{WT} and ESC^{KO}; blue) and WT and *Jmjd2c*-KO cEpiSCs ($cEpiSC^{WT}$ and $cEpiSC^{KO}$; green) in positive controls regions enriched in G9a (*Vstm2l, Npas4*) and a negative control region not enriched in G9a (*Nnat*). Levels at each region were normalized to an intergenic control region. Data was represented as average of 3 independent experiments ±SEM. **B**. Snapshots of the ChIP-seq enrichment profiles for G9a and Jmjd2c at the regions assessed in A. Primer locations are indicated by (*). **C.** Western blot showing the levels of Jmjd2c, G9a, GLP, Med1, Smc1a, Oct4a and p300 in on whole cell extracts from WT and *Jmjd2c*-knockout ESC and cEpiSC cultures. α-Tubulin was used as a loading control.

Appendix B – Tables

Table B – I. Sequenced of primer used in genotyping of Jmjd2c-knockout

Name	Sequence (5' – 3')
tm1a-F	GAAGTAGGAGGAGAACCATAGTTCCAGAG
tm1a-R	GCCTCTTCGCTATTACGCCAGCTG
tm2-F	AAGAGGTAAGAAGCCCATCACATC
tm2-R	GCAGCTATTTACCCGCAGGA

Table B – II. Sequences of primers used for cloning of Jmjd2c

Name	Sequence (5' – 3')
XhoI-Jmjd2c-F	ATACATTCTCGAGATGGAGGTGGT
PacI-Jmjd2c3-R	CTCTTAATTAACTACTGTCTCTTCTGACACT
BamHI-Jmjd2c-F	ATTAGGATCCGAGGTGGTGGAGGTG
SalI-Jmjd2c-R	TAATGTCGACCTACTGTCTCTTCTGACACTT
SalI-Del-Tudor-R	TAATGTCGACCTACGCGTTTGAGTTGAC

Table B – III. Sequences of primers used for RT-qPCR

Gene	Forward (5' – 3')	Reverse (5' – 3')
Oct4	CGTGGAGACTTTGCAGCCTG	GCTTGGAAACTGTTCTAGCTCCT
*Oct4**	ACCTCAGGTTGGACTGGGCCTA	GCCTCGAAGCGACAGATGGT
Brachyury	CCGGTGCTGAAGGTAAATGT	CCTCCATTGAGCTTGTTGGT
*Brachyury**	TCAGCAAAGTCAAACTCACC	TCATTCTGGTAGGCAGTCAC
Cdx2	TCAACCTCGCCACAACCTTCCC	TGGCTCAGCCTGGGATTGCT
Cxcr4	CGGGATGAAAACGTCCATTT	ATGACCAGGATCACCAATCCA
Dab2	GGCAACAGGCTGAACCATTAGT	TTGGTGTCGATTTCAGAGTTTAGAT
Dnmt3b	TTCAGTGACCAGTCCTCAGACACGAA	TCAGAAGGCTGGAGACCTCCCTCTT
Eset	GAGGCCGGGGAGAGGCTGAA	TAGCTGGTCGGCGAAGCCCT
Esrrb	GCACCTGGGCTCTAGTTGC	TACAGTCCTCGTAGCTCTTGC
Fgf5	TGTGTCTCAGGGGATTGTAGG	AGCTGTTTTCTTGGAATCTCTC
*Fgf5**	ACGTTTTCTTCGTCTTCTGC	TTCCTACAATCCCCTGAGAC
Flk-1	GACGGAGAAGGAGTCTGTGC	TTTCTGTGTGCTGAGCTTGG
Foxa2	CCATCAGCCCCACAAAATG	CCAAGCTGCCTGGCATG
*Foxa2**	ACTGGAGCAGCTACTACGC	AGGATGACATGTTCATGGAG
G9a	TCACCCTGACTGACAATGAG	AGACAGGAACAACAGAACAC
Gapdh	AGA AGG TGG TGA AGC AGG CA	CGA AGG TGG AAG AGT GGG AG
Gata4	GAACACCCTGAGCAGGCCTC	TTTCTGGGAAACTGGAGCTGG
Gata6	GAGCTGGTGCTACCAAGAGG	TGCAAAAGCCCATCTCTTCT
GLP	CAGATGGAGAAACAAATGGGTCT	TTTGCTTCCCCACTTCTGTGT

Hmbs	ACTGGTGGAGTATGGAGTCTAGATGGC	GCCAGGCTGATGCCCAGGTT
Isl1	AGACCCTCTCAGTCCCTTG	CGTCATCTCCACTAGTTGCTC
Jmjd2a	TTCTGTGAATCCTGCGTCTG	TAGCTGTGCACGGTGAGAAC
Jmjd2b	CATGTGGAAGACCACGTTTG	AAGGGGATGCCGTACTTCTT
Jmjd2c(ex14-15)	CCAAATGCCTTCCTTGAAGA	CACAGCCATCCATCACAGAC
Jmjd2c(ex18-20)	GATGACTGGCCTTACGTGGT	CTTCACACAGTTTCGGCTCA
Jmjd2c(ex5-7)	ACACGGAGGACATGGATCTC	TCACAGCCTTGGGAGCTACT
Klf4	GTGCCCCGACTAACCGTTG	GTCGTTGAACTCCTCGGTCT
L19	TGATCTGCTGACGGAGTTG	GGAAAAGAAGGTCTGGTTGGA
Lrp2	CCTTGCCAAACCCTCTGAAAAT	CACAAGGTTTGCGGTGTCTTTA
LSD1	GATGATGTCTGGGAAACAGG	GGGCCAAGGTAGAATACAGAG
Math1	GGAGAAGCTTCGTTGCACGC	GGGACATCGCACTGCAATGG
Meox1	ACGGAGAAGAAATCATCCAG	GTAGTTGTGGTGGGCAAAC
Mixl1	GACAGACCATGTACCCAGAC	GCTTCAAACACCTAGCTTCAG
Nanog	CTTACAAGGGTCTGCTACTGA	TCTGCTTCCTGGCAAGGACC
Otx2	TATCTAAAGCAACCGCCTTACG	AAGTCCATACCCGAAGTGGTC
Pax3	GGGAACTGGAGGCATGTTTA	GTTTTCCGTCCCAGCAATTA
Pax6	ACACGCCCTGGTTGGTATC	CATCTGAGCTTCATCCGAGTC
Pdgfra	GGTCCCAACCTGTAATGAAG	GTAAATGGGACCTGACTTGG
S17	ATGACTTCCACACCAACAAGC	GCCAACTGTAGGCTGAGTGAC
Sox17	GGAATCCAACCAGCCCACTG	GGACACCACGGAGGAAATGG
Sox2	AAAACCACCAATCCCATCCA	CGAAGCGCCTAACGTACCAC
Sox7	CAAGGATGAGAGGAAACGTCTG	TCATCCACATAGGGTCTCTTCTG
Sparc	AGGGCCTGGATCTTCTTTCTC	CAAATTCTCCCATTTCCACCT
Suv39h1	TGTCAACCATAGTTGTGATCC	ATTCGGGTACTCTCCATGTC
Suv39h2	ATCTACGAATGCAACTCAAGGTG	CCACAGCCATTGCTAGTTCTAA
Tbx3	GAACCTACCTGTTCCCGGAAA	CAATGCCCAATGTCTCGAAAAC
Ywhaz	CGTTGTAGGAGCCCGTAGGTCAT	TCTGGTTGCGAAGCATTGGG

** Primers used in Fig.4.4D*

Table B – IV. Sequences of primers used for ChIP

Gene	Forward (5' – 3')	Reverse (5' – 3')	Region
Intergenic	AAGGGGCCTCTGCTTAAAA	AGAGCTCCATGGCAGGTAGA	Control
Vstm2l	GCACGGCTCACAGTACCTAAAGT	CAGCAAGCACGGTGTGTCA	TSS
Npas4	CTATGGCCATTTCAGCACCG	AGCTGTTCGACGTCCTGAAG	TSS
Nnat	CCCTACCCAACCCATCCTATC	CCACCGCGGCACTTTG	TSS
Esrrb	GGAAGTGGGAAGTGTTGCTAT	GAGCTCCAGATCCCCTACAC	Enhancer
Esrrb	CTCCCACTCCGCCTTCTC	GGCAAATCACGCGGAAGA	TSS
Klf4	AAGGAAGGCGTTCCAGATTT	TTGAGATCCCTGGTGAAAGG	TSS
Klf4	CGCCTGCCTGTACCTTCTAA	GGGCCTGGGAGAGATTGTAA	Enhancer
Foxa2	GCTATCACAGGGCCAGGAA	CTGTGGGCGAGCTAACAATG	Enhancer
Foxa2	GAGTAGAGGTGTCTGCGACCA	CAGAGGCCTAGGGAGAGACC	TSS
Brachyury	CTCTGTCTCAGTTTGCCATTCTC	TGTACCGACCACTGGAATCC	Enhancer
Brachyury	GGCCGACTTTGTTTCTTCCC	GAATGCCTCCTTCCTCCTCA	TSS
Fgf5	GTGCATGCATGGGACTGTAG	AACCCGACCTGAACCTAGTG	Enhancer
Fgf5	CTCACCAGTCGCAGCTTCT	GAAGGGCTCCACTGGAAACT	TSS
Fgf5	TTGGACTGCTGGGGATAGTG	CACGGGGTAAGGTGCTTCTA	Enhancer 1
Fgf5	GGGGCTTGGAGAGTTCAGAT	AAAGCTCACCTGGCAGTCTA	Enhancer 3

Table B – V. Sequences of primers used for DNA methylation assay and H3K9me2 ChIP

Gene	Forward (5' – 3')	Reverse (5' – 3')	Region
Foxa2	CAGTACCACCTGGGAGTGTTT	TGGATTCCCTAGGCATCTTC	Enhancer
Foxa2	GAGTAGAGGTGTCTGCGACCA	CAGAGGCCTAGGGAGAGACC	TSS
Foxa2	CCTCTGGGAAACAGAGGACA	TTTCTGCGGTCAAGGGTATT	Control
Brachyury	GGATTCCAGTGGTCGGTACA	AGGCTCTTGGCTCCAGCTAGT	Enhancer
Brachyury	ACCAAGAACGGCAGGTAGGT	GCTGTTAACTCAGCGGGAAG	TSS
Brachyury	CAAGCCTCCCTGCTGACTT	GGTTCAGCCACTGACACTCA	Control
Fgf5	GCCGACCAGTCTCTTTCATT	TTGCTTCTCCTTCTGCACCT	Enhancer
Fgf5	GGCTCGGAACATAGCAGTTT	GCCCAAAGGAATCTTGACC	TSS
Fgf5	GTGTCCCATGGTAACCTGCT	GCAGCCCTGTACAAGATCCA	Control
Magea2	TTGGTGGACAGGGAAGCTAGGGGA	CGCTCCAGAACAAAATGGCGCAGA	TSS
Chrna3	AGGAGGCAGATGCTGGAAG	TCCCAGTCCCTACCACTGAG	TSS

Table B – VI. Motifs identified in H3K27-high versus H3K27ac-low Jmjd2c peak dataset

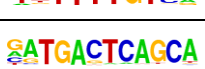
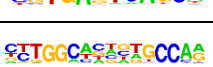
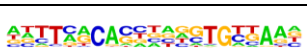
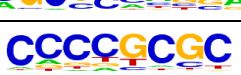
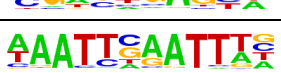
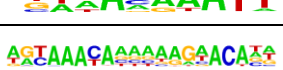
Motif	Name	P-value	# Target Sequences with Motif	% Targets Sequences with Motif	# Background Sequences with Motif	% Background Sequences with Motif
	Esrrb	1e-9	428.0	10.45%	1797.6	7.72%
	Nr5a2	1e-6	363.0	8.86%	1605.4	6.89%
	MITF	1e-4	434.0	10.59%	2055.4	8.82%
	ZFX	1e-3	798.0	19.48%	4016.0	17.24%
	ZNF711	1e-3	1064.0	25.97%	5485.0	23.54%
	KLF5	1e-3	883.0	21.55%	4530.7	19.45%
	BMAL1	1e-2	772.0	18.84%	3993.2	17.14%
	ELF1	1e-2	362.0	8.84%	1776.8	7.63%
	CRE	1e-2	93.0	2.27%	388.8	1.67%
	Nur77	1e-2	79.0	1.93%	322.0	1.38%
	PPARE	1e-2	549.0	13.40%	2783.2	11.95%
	Sox10	1e-2	691.0	16.87%	3557.6	15.27%
	Sox4	1e-2	373.0	9.10%	1840.9	7.90%
	Sox3	1e-2	708.0	17.28%	3653.7	15.68%
	Gata4	1e-2	329.0	8.03%	1617.1	6.94%
	Sox2	1e-2	366.0	8.93%	1817.3	7.80%
	NF-E2	1e-2	37.0	0.90%	133.6	0.57%
	NF1	1e-2	235.0	5.74%	1136.5	4.88%
	EKLF	1e-2	139.0	3.39%	644.8	2.77%

Table B – VII. Motifs identified in H3K27-low versus H3K27ac-high Jmjd2c peak dataset

Motif	Name	P-value	# Target Sequences with Motif	% Targets Sequences with Motif	# Background Sequences with Motif	% Background Sequences with Motif
	E2F1	1e-29	818.0	3.53%	87.8	2.34%
	REST	1e-21	43.0	0.19%	0.6	0.02%
	SFP1	1e-16	37.0	0.16%	0.9	0.02%

	NF1	1e-16	5508.0	23.75%	808.3	21.50%
	E2F7	1e-15	386.0	1.66%	40.9	1.09%
	MyoD	1e-12	2524.0	10.88%	355.7	9.46%
	Brachyury	1e-11	449.0	1.94%	52.3	1.39%
	Pbx3	1e-11	507.0	2.19%	60.9	1.62%
	MyoG	1e-9	3328.0	14.35%	488.3	12.99%
	Atf4	1e-9	369.0	1.59%	43.3	1.15%
	GATA3	1e-8	138.0	0.60%	13.7	0.36%
	Ascl1	1e-8	5028.0	21.68%	757.9	20.16%
	Atoh1	1e-7	3324.0	14.33%	494.1	13.14%
	Hoxc9	1e-7	606.0	2.61%	79.4	2.11%
	SCL	1e-6	12866.0	55.48%	2022.9	53.81%
	Nkx6.1	1e-6	3699.0	15.95%	555.7	14.78%
	Chop	1e-5	289.0	1.25%	35.5	0.95%
	Myf5	1e-5	1997.0	8.61%	292.7	7.79%
	PAX3	1e-5	238.0	1.03%	28.1	0.75%
	E2F4	1e-5	1508.0	6.50%	217.9	5.80%
	SUT1?	1e-5	19300.0	83.23%	3085.2	82.07%
	E2F6	1e-5	1740.0	7.50%	253.2	6.74%
	CES-1	1e-5	495.0	2.13%	65.8	1.75%
	Nkx2.5	1e-4	5165.0	22.27%	795.1	21.15%
	Oct4	1e-4	485.0	2.09%	65.0	1.73%
	CTCF	1e-4	572.0	2.47%	79.1	2.10%
	Olig2	1e-3	4779.0	20.61%	738.2	19.64%
	LIN-39	1e-3	1285.0	5.54%	188.8	5.02%
	FOXA1	1e-3	61.0	0.26%	6.8	0.18%
	PRDM14	1e-3	631.0	2.72%	89.2	2.37%

The top 30 results are shown in this table.