

Regulation of pluripotent states in human embryonic stem cells

by

Sandra Isabel da Rocha Lourenço de Pinho

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Institute of Reproductive and Developmental Biology
Imperial College School of Medicine

November of 2009

STATEMENT OF ORIGINALITY

All experiments included in this thesis were performed by me Sandra Pinho unless otherwise stated in the text.

ABSTRACT

A growing array of mouse and human pluripotent stem cell lines has been derived from the early embryo as well as from adult cells reprogrammed by ectopic expression of transcription factors – i.e. induced pluripotent stem (iPS) cells. These cell lines share the expression of key pluripotency markers and are able to self-renew and to generate differentiated progenies when induced. Their relationship to each other and whether they correspond to different pluripotent states with distinct developmental capacities and affiliations *in vivo* remains unclear, however. Profiling chromatin in a particular cell line has proven to be a valuable signature for cell identity and developmental stage. One approach has been to assay the timing of DNA replication across a panel of loci, as an indicator of chromatin accessibility. Of interest, this replication timing profiling was capable of discriminating pluripotent mouse ES (mES) cells from cells with a more restricted differentiation capacity.

In this study, I have addressed whether distinct pluripotent states could be reliably discriminated at the chromatin level. In particular, I characterised the replication timing profile of a number of human ES (hES) cell lines alongside mES and mouse epiblast-derived stem (mEpiS) cell lines. I showed that mES cells have a steady and mostly early-replicating profile, regardless of their genetic background. In contrast, the profile of undifferentiated H1, H7 and H9 hES cell lines harboured an increased proportion of late-replicating loci during S-phase. Moreover, hES cell replication profile greatly varied between cultures and cell lines; a level of replication timing variability also observed among mEpiS cells, as opposed to mES cells. These results highlighted that hES and mEpiS cells share a common unstable or transitional state: primed on the verge of differentiation.

This view was, however, further challenged by exploring how hES cell cultures could be modulated towards an ES-like *versus* epiblast-like state under different conditions. In particular, extensive and dynamic shifts of replication timing, from late to early, were consistently observed at many target loci in hES and hiPS cells upon increased Smad2/3 and p300 histone acetyltransferase activity. Importantly, these alterations were reversible and associated with differential gene expression profiles and functional properties of hES cells. Collectively, these data revealed the existence of distinct but interchangeable pluripotent hES cell states and proposed a key role for TGF- β /Activin signalling and the HAT p300 in modulating the balance between a naive *versus* primed state in hES cell cultures.

ACKNOWLEDGEMENTS

I would first like to give my sincere thank you to Veronique for giving me the opportunity to work in her lab, for her teaching, support and help on the demanding challenges of a PhD. It was very exciting to be part of the Epigenetics and Development group foundations.

I would like to thank Olivia for being the perfect “desk and lab-neighbour” in the world, for all her friendship, advice and help (especially in filling up paperwork). I would additionally like to thank past and present members of the Epigenetics and Development Group (Anne, Fabrice, Kavita and Patrick) for being great colleagues, for all the scientific (and not that scientific) coffee discussions and for making work in the lab so much fun. In particular, I would like to specifically thank Matias for all his help in the lab, lost weekends in tissue culture room and for being a good colleague and a “stubborn” friend. I would like to thanks our collaborators: Neil Harrison and Peter Andrews (University of Sheffield, UK), Catherine Flores (Imperial College, London), Ludovic Vallier (University of Cambridge, UK) and especially to Wei Cui (Imperial College, London) for all her help and teaching with the human embryonic stem cells. Thanks also to everyone in IRDB first floor for being great colleagues and for creating a friendly atmosphere to work.

A special acknowledgement goes to the GABBA PhD Program for giving me the privilege of becoming one of their students, for the scientific training and above all for the freedom of choice. I am indebted to all my friends and teachers from the 8th edition for an amazing year of learning and scientific discussions. This work would not be possible without the generous financial support from the funding bodies FCT, Portugal, and the IOG Trust Fund, UK. I also want to thank all my housemates and friends in London that made this city a fantastic place to live and so difficult to leave.

My very special thank you goes to Filipe for always being present, for believing in me, for his patience, and for the constant help and support throughout these years. It is a privilege to have you by my side.

Finalmente agradeço aos meus pais Ana Maria e Manuel, às minhas irmãs Paula e Rosa e aos meus avós pelo amor e apoio incondicional que sempre me deram e por me ajudarem a concretizar os meus sonhos, mesmo quando isso implica a minha ausência. Obrigada.

TABLE OF CONTENTS

STATEMENT OF ORIGINALITY	2
ABSTRACT	3
ACKNOWLEDGEMENTS	4
TABLE OF CONTENTS	5
FIGURES AND TABLES	8
ABBREVIATIONS	11
CHAPTER 1. INTRODUCTION	15
1.1. Epigenetic levels of gene regulation and epigenetic inheritance	16
1.1.1. Chromatin structure	16
1.1.2. DNA methylation	19
1.1.3. Covalent histone modifications and their impact on chromatin dynamics	21
1.1.4. Spatial constraints – nuclear organisation	27
1.1.5. Temporal constraints – DNA replication timing	29
1.2. Early mammalian development	33
1.2.1. Early embryo development and pluripotency	33
1.2.2. Embryo-derived stem cells	34
1.2.3. Signalling pathways required for ES cells self-renewal and pluripotency	39
1.2.4. Transcriptional networks that control pluripotency	42
1.2.5. Epigenetic control of pluripotency	42
1.2.6. Reprogramming to pluripotency	43
1.3. Aims of this study	47
CHAPTER 2. MATERIALS AND METHODS	48
2.1. Materials	48
2.1.1. Primary antibodies	48
2.1.2. Secondary antibodies	49
2.1.3. Cell lines	50
2.2. Methods	52
2.2.1. Cell culture	52
2.2.2. Human ES cells differentiation	54
2.2.3. Replication Timing analysis	56
2.2.4. Reverse transcription (RT) polymerase chain reaction (PCR)	57
2.2.5. Genomic context analysis	59
2.2.6. Immunofluorescence analysis and alkaline phosphatase activity	59

2.2.7. Fluorescence activated cell sorting analysis	60
2.2.8. Western blot analysis	61
2.2.9. Chromatin immunoprecipitation analysis	62
2.2.10. Methylated DNA immunoprecipitation analysis	64
2.2.11. Cell fusion and experimental heterokaryons	64
2.2.12. Analysis of DNA content and BrdU incorporation following acid denaturation of DNA	65

CHAPTER 3. CHARACTERIZATION OF THE CHROMATIN STATUS OF HUMAN ES CELLS ALONGSIDE MOUSE ES AND EPI S CELLS

3.1. Replication timing analysis of hES, mES and mEpiS cells using a candidate based approach	67
3.2. Replication timing profiles are conserved between independent mES cell lines	72
3.3. Pluripotent mouse and human ES cells have distinct replication timing profiles	74
3.4. Human ES cell replication profiles can be reliably discriminated from that of cells with more restricted differentiation capacity	76
3.4.1. Replication timing comparison of hES and hES-derived neural progenitor cells	76
3.4.2. Replication timing comparison of human ES and CD34+ haematopoietic stem cells	80
3.5. Analysis of the gene expression status and chromosomal context of replication timing variable genes	81
3.5.1. Gene expression and histone acetylation analysis of genes shifting replication timing across H1, H7 and H9 hES cells	81
3.5.2. Genomic features analysis of genes shifting replication timing across H1, H7 and H9 hES cells	84
3.5.3. DNA methylation analysis at promoters of replication timing variable genes across H1, H7 and H9 hES cells	86
3.6. Human ES and mEpiS cells share a common unstable or transitional state: primed on the verge of differentiation	87

CHAPTER 4. IDENTIFICATION AND CHARACTERIZATION OF DISTINCT PLURIPOTENT STATES WITHIN HUMAN ES CELL CULTURES: NAIVE VS PRIMED

4.1. Undifferentiated hES cell replication timing profile is prone to extensive alterations upon different cell culture conditions	91
4.1.1. Replication profile of H9 hES cells grown in two different laboratories	91
4.1.2. Dynamic equilibrium and heterogeneity of hES cells in distinct epigenetic states	92
4.1.3. Fully reprogrammed hiPS cell replication profiles matches hES cells dramatic late to early shifts in response to alterations in culture conditions	95
4.2. Phenotypic and functional characterization of distinct populations of undifferentiated hES cells in CM and CDM ^{activin} growth conditions	97
4.2.1. CDM ^{activin} conditions maintain H1 hES cells in an undifferentiated state supporting their self-renewal	97
4.2.2. H1 CM and CDM ^{activin} form embryoid bodies containing tri-germ layer differentiation	98
4.2.3. H1 hES cells in CM and CDM ^{activin} conditions have different kinetics of <i>in vitro</i> differentiation	100
4.2.4. CDM ^{activin} conditions tip the balance towards a more stable, ICM-like cell phenotype within hES cell cultures	102
4.3. Generation of inter-species heterokaryons by cellular fusion	104

4.3.1. Mouse EpiS cells are not able to reprogram human B-lymphocytes for pluripotency in contrast to mES cells	106
4.3.2. Both H1 CM and CDM ^{activin} hES cells are able to reprogram mouse B-lymphocytes for pluripotency	108
CHAPTER 5. EPIGENETIC AND MOLECULAR MECHANISMS UNDERLYING	
DISTINCT PLURIPOTENT STATES WITHIN HUMAN ES CELLS	110
5.1. A possible role for TGF- β /Activin signalling in defining a hES cell naive <i>versus</i> primed state in culture	110
5.2. Shift from early to late replication is accompanied by an increased recruitment of active p300 at the promoter of target genes in CDM ^{activin} cells	111
5.2.1. Global increase of p300-mediated histone H3 acetylation in CDM ^{activin} hES cell cultures	111
5.2.2. Recruitment of active p300 complexes at target loci is enhanced in response to CDM ^{activin} growth conditions	113
5.3. Inhibition of p300 acetyltransferase activity leads to early to late replication timing alterations in H1 CDM ^{activin} hES cells	115
CHAPTER 6. DISCUSSION	118
6.1. Human and mouse ES cells have distinct replication timing profiles	118
6.2. Replication timing variability is a hallmark of hES cells in CM conditions and mEpiS cells	120
6.3. Identification and characterization of distinct pluripotent states within hES cell cultures: naive vs primed	122
6.4. Investigating the mechanisms that specify distinct pluripotent states within human ES cell cultures: a pivotal role for TGF- β /Activin signalling and the histone acetyltransferase p300	126
REFERENCES	129
APPENDIX I – PRIMER SEQUENCES	151
APPENDIX II – COMPLETE DATA SETS	160

FIGURES AND TABLES

Figure 1.1. Epigenetic levels of gene regulation.	17
Figure 1.2. Post-translational modifications of human histones.	22
Figure 1.3. Replication timing foci patterns during early, middle and late S-phase of cell cycle.	30
Figure 1.4. Schematic representation of the early stages of mouse development and derivation of stem cell types.	34
Figure 3.1. Characterization of the undifferentiated state of hES cells alongside mES and mEpiS cells.	68
Table 1. Candidate genes selected for the replication timing analysis.	69
Figure 3.2. The principle of replication timing analysis	71
Figure 3.3. Replication timing analysis controls.	72
Figure 3.4. Independently derived mES cell lines from distinct genetic backgrounds have similar replication timing profiles.	73
Figure 3.5. Pluripotent mouse and human ES cells have distinct replication timing profiles.	74
Figure 3.6. Replication timing variability between established hES cell lines.	75
Figure 3.7. Human ES cell cultures analysed by replication timing display homogeneous expression of pluripotency-associated markers.	77
Figure 3.8. H7 undifferentiated hES cells can be clearly distinguished by replication timing analysis from neural progenitors (NP) H7-derived.	79
Figure 3.9. Replication timing variability is not observed across human CD34+ haematopoietic stem cells.	80
Figure 3.10. Human ES cells replication timing variability is not associated with gene expression or histone acetylation specific patterns.	83
Table 2. Genomic features of replication timing analysed genes.	85
Figure 3.11. Human ES cells replication timing variability is not associated with DNA methylation specific patterns.	86
Figure 3.12. Mouse EpiS and hES cells have identical cell cycle kinetics and variable replication timing profiles across many genes.	88
Figure 4.1. Human ES cells grown under different culture conditions in different laboratories show dramatic replication timing late to early shifts.	92
Figure 4.2. H1 hES cells grown under CM and CDM ^{activin} conditions show dramatic replication timing late to early shifts.	93
Figure 4.3. Loci subjected to replication timing shift towards early S-phase in CDM ^{activin} conditions overlap between different hES cell lines.	94
Figure 4.4. Fully reprogrammed hiPS cells replication profile also display dramatic late to early shifts in response to alterations in culture conditions.	96
Figure 4.5. Human ES cells transferred from standard CM cell culture conditions to CDM ^{activin} conditions retain their undifferentiated characteristics.	97

Figure 4.6. Human ES cells in CDM ^{activin} conditions show a slight increase in the percentage of cells in S and G2 compared to cells grown under standard CM conditions.	98
Figure 4.7. Human ES cells in CM and CDM ^{activin} conditions are able to differentiate <i>in vitro</i> into derivatives of the three germ layers.	99
Figure 4.8. Human ES cells cultured in CDM ^{activin} differentiate more slowly by EBs differentiation than hES cells cultured in CM.	100
Figure 4.9. Human ES cells cultured in CDM ^{activin} differentiate more slowly into neuroectoderm derivatives than hES cells cultured in CM.	101
Figure 4.10. Differential gene expression profiles in H1 CM and CDM ^{activin} hES cell populations.	102
Figure 4.11. Transient increase in the expression levels of Nanog in H1 CDM ^{activin} hES cells.	103
Figure 4.12. Generation of inter-species heterokaryons by cell fusion.	105
Figure 4.13. Mouse EpiS cells are not able to reprogram human B-lymphocytes for pluripotency by cellular fusion in contrast to mES cells.	107
Figure 4.14. Both H1 CM and CDM ^{activin} hES cells are able to reprogram mouse B-lymphocytes for pluripotency by cellular fusion.	109
Figure 5.1. H1 CDM ^{activin} hES cells are characterized by higher levels of activated Smad2/3 complexes in comparison to H1 CM cell cultures.	111
Figure 5.2. The histone acetyltransferase p300 is highly expressed in H1 CDM ^{activin} hES cells in contrast to H1 CM cells.	112
Figure 5.3. Loci changing replication timing from late in CM to early in CDM ^{activin} hES cells have higher levels of p300 and histones acetylation at their promoter regions in CDM ^{activin} hES cells.	114
Figure 5.4. Inhibition of p300 activity in H1 CDM ^{activin} hES cells by curcumin leads to a reduction in the levels of histone acetylation without inducing cell differentiation.	116
Figure 5.5. Inhibition of p300 acetyltransferase activity in curcumin-treated H1 CDM ^{activin} hES cells induces replication timing changes towards a CM-like profile.	117
Figure 6.1. Proposed model for the maintenance of distinct pluripotent states within hES cell cultures	123
Table A1. Primers for the amplification of human or mouse transcripts by real time quantitative PCR	151
Table A2. Primers used for the amplification of human transcripts by semi-quantitative RT-PCR	154
Table A3. Primers for the selective amplification of human or mouse genomic regions by real time quantitative PCR	155
Figure A1. Comparison of replication timing profiles between mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines.	160
Figure A2. Comparative replication timing analysis of mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines.	161
Figure A3. Replication timing analysis of H1 hES cells re-swapped to CM conditions after growing in CDM ^{activin} .	162
Figure A4. Human ES cell populations grown under CDM ^{activin} conditions are more homogeneous than cells grown under standard CM conditions.	162

Figure A5. Differential gene expression of mES and mEpiS cell populations. 163

Figure A6. Cell culture adaptation and karyotype abnormalities are not associated with earlier replication timing profiles in hES cells. 163

ABBREVIATIONS

α-MEM	minimal essential medium
μ	micro
AP	alkaline phosphatase
APC	allophycocyanin
AT	adenine-thymine
ATP	adenosine tri-phosphate
BMP	bone morphogenetic protein
bp	base pair
BrdU	5-bromo-2-deoxyuridine
BSA	bovine serum albumin
cDNA	complementary DNA
CDM	chemically defined medium
ChIP	chromatin immunoprecipitation
CM	conditioned medium
C(T)	threshold cycle
CpG	cytosine-guanine dinucleotide
CTD	carboxy-terminal domain
DAPI	4,6-diamidino-2-phenylindole
DiD	1,1'-dioctadecyl-3, 3, 3', 3' tetramethylindodicarbocyanine
Dil	1,1'-dioctadecyl-3, 3, 3', 3'-tetramethylindodicarbocyanine perchlorate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
Dnmt	DNA methyltransferase
dpc	days post-coitum
DTT	dithiothreitol
EB	embryoid body
EDTA	ethylene-dinitrilo tetraacetic acid
EC	embryonic carcinoma
EG	embryonic germ
EGF	epidermal growth factor
ES	embryonic stem

FACS	fluorescence-activated cell sorter
FBS	foetal bovine serum
FISH	fluorescent <i>in situ</i> hybridisation
FITC	fluorescein isothiocyanate
FGF	fibroblast growth factor
g	gram
GC	guanine-cytosine
GFP	green fluorescent protein
GMEM	Glasgow's modified Eagle's medium
Gy	gray
H	histone
HAT	histone acetyltransferase
hB	Epstein-Barr virus-transformed human B lymphocytes
HDAC	histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hES	Human embryonic stem
HMT	histone methyltransferase
HRP	horseradish peroxidase
HP	haematopoietic progenitor
ICM	inner cell mass
IF	immunofluorescence
IL	interleukin
IMDM	Iscoe's modified Dulbecco's medium
IP	immunoprecipitation
iPS	induced pluripotent stem
ISWI	imitation switch family of chromatin remodelling ATPases
K	lysine
Kb	kilobase pair
kDa	kiloDalton
KSR	knockout serum replacement
L	litre
LIF	leukaemia inhibitory factor
LINE	long interspersed nuclear element
m	milli
M	molar
Mb	mega base pair

MBD	methyl-CpG binding domain
me	methylated
MeDIP	methylated DNA immunoprecipitation
MEF	mouse embryonic fibroblast
mEpiS	mouse epiblast stem
mES	mouse embryonic stem
MgCl ₂	magnesium chloride
min	minute
mRNA	messenger RNA
NaCl	sodium chloride
NEAA	non-essential amino acids
NGS	normal goat serum
NP	neural progenitor
NP-40	nonidet P-40
P	phosphorylated
PBS	phosphate buffered saline
PcG	polycomb group
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEV	position effect variegation
PFA	paraformaldehyde
PGC	primordial germ cell
PI	propidium iodide
PK	proteinase K
PMSF	phenylmethylsulphonyl fluoride
Pol	polymerase
PRC	polycomb repressive complex
PRE	polycomb response element
PrE	primitive endoderm
qPCR	real-time quantitative PCR
RIPA	radio immuno precipitation assay
RNA	ribonucleic acid
RNAse	ribonuclease
rpm	rotations per minute
RPMI	Roswell Park Memorial Institute medium
RT-PCR	reverse transcription PCR

<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i> , budding yeast
SDS	sodium dodecyl sulphate
SET	Su(var)-Enhancer of zeste-Trithorax
SSEA	stage-specific embryonic antigen
SWI/SNF	(mating-type) switch / sucrose non-fermenting family of chromatin remodelling ATPases
TBS-T	tris-buffered saline with tween20
TGF	transforming growth factor
TE	trophectoderm
Tra	tumor rejection antigen
Tris	(hydroxymethyl)methylamine
TrxG	trithorax group
TS	trophoblast stem
TSS	transcriptional start site
U	units
WB	western blot
XEN	extra-embryonic endoderm stem

CHAPTER 1. INTRODUCTION

Mammalian development is a remarkable process underlined by a progressive loss of cell developmental potential and the acquisition of lineage-specific functions. This process begins with the fusion of germ cells, two highly differentiated cells, leading to the formation of a single totipotent zygotic cell and ends with the establishment of more than 200 specialised cell types of the mammalian body. Elucidating the mechanisms by which genetically identical cells acquire different cell identities during development and how those are stably maintained has been a major challenge in developmental biology for many years.

With the completion of the human genome sequencing (Lander et al., 2001; McPherson et al., 2001) it became clear that sequences are only the tip of the iceberg for understanding how genetic information is interpreted. Superimposed upon the DNA sequence exist several layers of heritable information that we are only beginning to understand. The vast majority of cells in a multicellular organism share the same genetic information, yet a multitude of specialised cell types each using a unique subset of genes is generated throughout development. Current evidence points to the importance of a close coordination between genetic and epigenetic programs in the control of cellular determination during development. Epigenetic factors include modifications in chromatin structure, methylation of the DNA cytosine residues, DNA replication timing, regulation by small interfering RNAs and the spatial positioning of the genomic region within the nucleus. These genetic and epigenetic programs are, in turn, regulated by signaling molecules, which together with interactions among neighboring cells induce appropriate and differential transcriptional and epigenetic responses that are essential for cell fate determination. In addition, epigenetic mechanisms contribute to the repression of inappropriate developmental programs in time and space while ensuring heritability of existing or newly acquired phenotypic states. These extrinsic and intrinsic regulators determine the developmental origin and subsequent maintenance of lineage commitment throughout rounds of cell division (Hochedlinger and Plath, 2009; Surani et al., 2007).

1.1. Epigenetic levels of gene regulation and epigenetic inheritance

The term “epigenetic” was initially introduced by Conrad Waddington in 1942 (Waddington, 1942) to describe how genes of a genotype bring about a phenotype. Currently, the most widely accepted definition designates epigenetics as the study of heritable changes in gene function that occur without alterations to the DNA sequence (Russo et al., 1996). Epigenetic inheritance refers to the transmission of information beyond the DNA sequence during cell division and from one generation to the next. Inheritance of epigenetic information is crucial for maintaining differential gene expression patterns in differentiation, development and disease (Haig, 2004). The key players in epigenetic inheritance are now beginning to be revealed. These include heritable but potentially reversible modifications in chromatin structure, methylation of DNA, regulation by small interfering RNAs and higher-order genome organisation (Figure 1.1). Histone-modifying enzymes, DNA methyltransferases, chromatin-binding proteins, non-coding RNAs and many other regulators may act together with sequence-specific transcriptional factors to establish and convey gene expression patterns during development, ensuring the stability of the “committed” state. Uncoupling the chromatin changes that are “locked” and stably propagated throughout cell division from those that are transiently regulated constitutes a big challenge to define the epigenetic value of chromatin (Turner, 2002).

1.1.1. Chromatin structure

In the nuclei of all eukaryotic cells, the genomic information is divided between a set of different chromosomes. The genome is large, each human cell contains around three meters of DNA that needs to be highly compacted and ordered to fit within the nucleus. This is achieved by the interaction of DNA with histone and non-histone proteins to form a dynamic polymer called chromatin (Figure 1.1). The dynamic organization of chromatin is dependent on the higher order structuring of nucleosomes, the basic repeating unit of chromatin. The nucleosome includes the compact octameric histone complex consisting of two copies of each of the four core histones H2A, H2B, H3 and H4 wrapped by 147 basepairs (bp) of DNA (Luger et al., 1997). In addition to the canonical histones, several histone variants have been identified for example, the histone variant H3.3, which differs in only four amino acids from H3, but has different biological properties. While the canonical H3 is synthesised exclusively in dividing cells

and integrated into chromatin during DNA replication, H3.3 is constitutively expressed and deposited into chromatin throughout the cell cycle (Ahmad and Henikoff, 2002a, b). It has been proposed that the replication-independent histone replacement can be involved in the propagation of active chromatin states. Several variants have also been described for histone H2A. For example, H2AZ has emerged as a key regulator of chromatin function and necessary for embryonic stem (ES) cell differentiation (Creyghton et al., 2008).

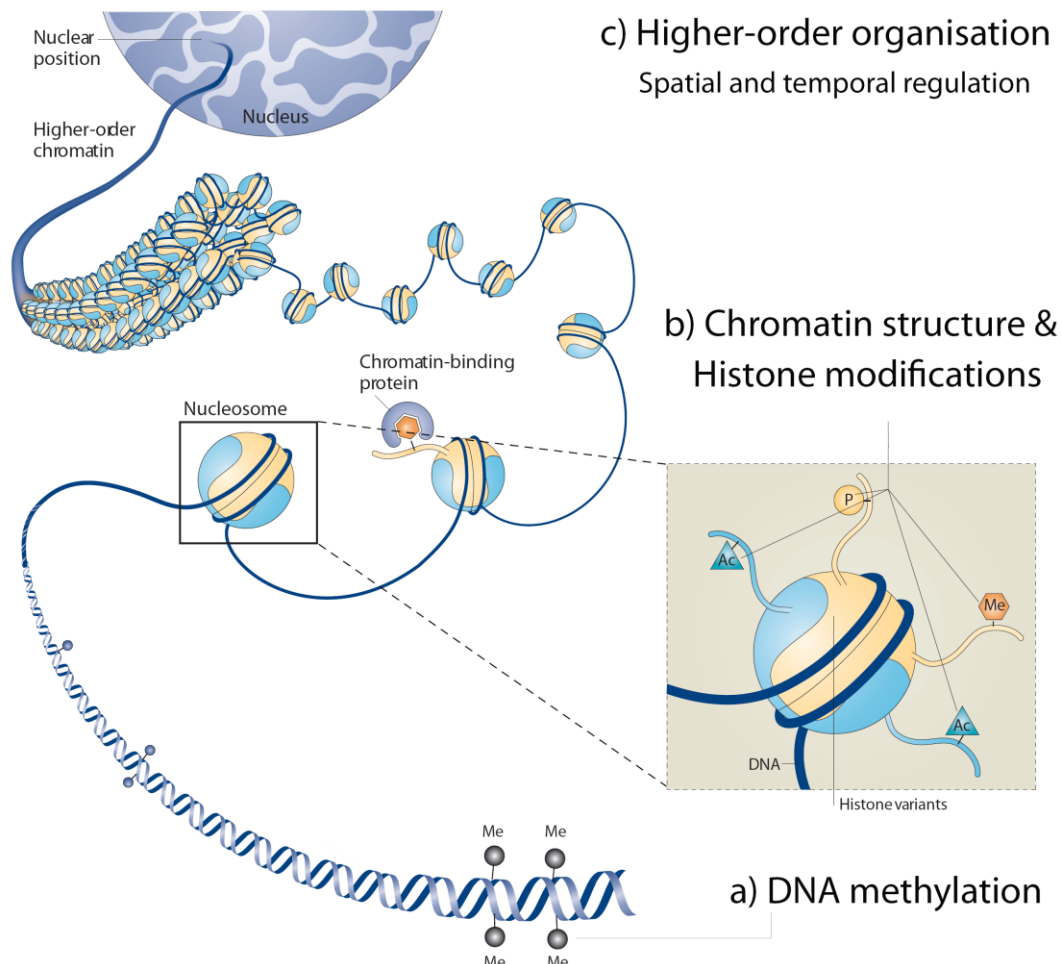


Figure 1.1. Epigenetic levels of gene regulation. Schematic representation of different levels of genome organisation that have an impact on gene regulation and epigenetic inheritance. **(a)** The DNA double helix carries the genetic information used for development and functioning of all known living organisms. DNA itself is covalently modified by methylation (Me) of cytosine residues, providing the most direct epigenetic mechanism of gene regulation. **(b)** DNA is wrapped around histones to form a “beads on a string” structure that is folded into higher order chromatin domains. Histones are also post-translationally modified at specific amino acids within their histone tails for example, by acetylation (Ac), methylation (Me) and phosphorylation (P), and each mark constitutes a signal that is read alone or in combination with other modifications on the same or neighbouring histones as a “histone code”. The nucleosome particles that comprise chromatin are represented as spheres and DNA as blue lines. **(c)** Nucleosomal arrays are packed into several levels of condensed, higher-order chromatin domains and organised in functional nuclear compartments. These three-dimensional structures along with temporal constraints (DNA replication timing) provide additional levels of regulation. Together these mechanisms provide some molecular clues on how a single genome can produce such a diversity of cell types. Adapted from (Probst et al., 2009; Sparmann and van Lohuizen, 2006).

While the length of DNA wrapped around the nucleosome is constant, the length of DNA separating nucleosomes varies between 10 and 80 bp. Although nucleosome mapping experiments did not reveal any strongly preferred DNA sequence consensus for nucleosome assembly it was suggested that nucleosome positioning is non-random and has a functional significance in several cellular processes including transcriptional regulation, chromosome segregation, DNA replication and DNA repair (Henikoff, 2008; Li et al., 2007; Schones et al., 2008). Several chromatin remodeling complexes (remodelers) that alter the positioning and/or nucleosome structure were identified. Many proteins of these complexes are ATPases that use energy of ATP hydrolysis to move, destabilise, eject, or restructure nucleosomes, including the SWI/SNF, ISWI, CDH and INO80 complexes (Clapier and Cairns, 2009). Together with additional chromatin associated proteins, including histone H1, nucleosomal arrays are packed into several levels of condensed, higher-order chromatin structures, which are organised in specific nuclear compartments (Figure 1.1) (Maresca and Heald, 2006; Tremethick, 2007). This folding can compact DNA by a factor of 30-40 fold (Widom, 1989).

The packaging of eukaryotic DNA must meet two requirements - DNA must be sufficiently compacted to fit inside the nucleus, and yet the DNA of transcriptionally competent or active genes must remain accessible to transcription factors and the RNA polymerase complex. These opposing requirements can be met by varying the mode of packaging. Early studies in 1928 by Emil Heitz suggested that the interphase chromatin is present in two distinguishable states: a less condensed form, euchromatin, and a more condensed form, heterochromatin, which stained intensely with specific dyes (Heitz, 1928). Euchromatin is the lightly packaged form of chromatin. Because of its unfolded structure, euchromatin is highly accessible to transcription factors and RNA polymerases. It is enriched for active genes replicating earlier in the S-phase of cell cycle than heterochromatin (Holmquist and Ashley, 2006; Shahbazian and Grunstein, 2007). In contrast, the higher-order structures of the tightly packed heterochromatin are thought to favor gene silencing (Holmquist and Ashley, 2006), as was first demonstrated by the phenomenon of position effect variegation (PEV). In this phenomenon, the integration of genes that are normally found in euchromatin, close to heterochromatic regions, often results in sporadic gene silencing, giving rise to a mosaic pattern of expression within a given tissue. This heritable silencing phenomenon was first identified in *Drosophila* where chromosomal rearrangements that resulted in the positioning of an eye color gene into a close proximity to centromeric heterochromatin resulted in a colorless (silencing) eye phenotype (Muller and Altenburg, 1930). Some chromosomal regions developmentally regulated have

long been known to be “naturally” present in both euchromatic and heterochromatic forms depending on the cell type. Such regions were termed facultative heterochromatin and one of the most extensive examples of facultative heterochromatin formation is the inactivation of one of the two X-chromosomes in early embryonic development (Heard, 2005). As opposed to facultative heterochromatin, the constitutive heterochromatin is generally silenced in all cell types and includes centromeres, telomeres, pericentric regions, repetitive DNA elements and transposons (Arney and Fisher, 2004).

1.1.2. DNA methylation

DNA methylation is an epigenetic repression pathway, which, in vertebrates, occurs only at cytosines in the context of CpG dinucleotides (Figure 1.1). It is catalyzed by three essential DNA (cytosine-5) methyltransferases (Dnmt1, Dnmt3a and Dnmt3b) (Goll and Bestor, 2005). Dnmt1- or Dnmt3b-deficient mouse embryos die by embryonic day 10.5 and Dnmt3a-deficient mice are born occasionally but suffer serious malformations especially at the germ cell level and die within weeks (Goll and Bestor, 2005). CpG methylation, especially within the promoter region of genes, is associated with stable transcriptional repression and plays important roles for X-chromosome inactivation, genomic imprinting, cell differentiation and carcinogenesis (Bird, 2002; Farthing et al., 2008; Jones and Baylin, 2002; Reik, 2007). Methylation patterns are propagated in mitotic cells by maintenance methylation activity that takes place in concert with DNA replication. The methyltransferase Dnmt1 performs this maintenance function (Bestor et al., 1988; Li et al., 1992). Dnmt1 is the most abundant DNA methyltransferase present in mammalian cells and has a substrate preference for hemimethylated DNA *in vitro* and is usually associated with replication foci in the S-phase of dividing cells (Chuang et al., 1997; Leonhardt et al., 1992).

During mammalian development, the genome goes through cycles of genome wide reprogramming that dramatically alter the overall patterns of DNA methylation (Reik, 2007). In early embryonic development parental methylation is lost, presumably due to the exclusion of Dnmt1 from the nucleus (Carlson et al., 1992) but methylation levels are restored at the blastocyst stage (Santos et al., 2002). Methylation patterns that are established at imprinted *loci* at this stage, such as the *Igf2/H19* locus, are preserved in all cells of the adult organism, except for the germ line (Edwards et al., 2007). The purpose of methylation reprogramming in preimplantation embryos is not well understood; one possible explanation is that demethylation in the early embryo is

needed for the parental genomes to lose their epigenetic marks so that the genome can return to totipotency (Reik et al., 2001). Genome-wide hypomethylation at the morula stage is then followed by lineage specific *de novo* methylation beginning at the blastocyst stage (Santos et al., 2002), presumably carried out by Dnmt3a and Dnmt3b (Okano et al., 1999). It is possible that this *de novo* methylation leads to epigenetic silencing of key promoters during early development. Indeed some regulators of pluripotency are hypomethylated in ES cells but become methylated upon differentiation in both mouse and human. A well characterised example of a locus regulated by DNA methylation is the ES cell-specific factor *Oct4*, which undergoes stable repression with cell differentiation (Farthing et al., 2008; Feldman et al., 2006; Gidekel and Bergman, 2002; Yeo et al., 2007).

DNA methylation is thought to interfere with gene transcription by at least two independent mechanisms. Firstly, methyl-CpGs can directly prevent transcription factors from binding to their recognition sites on the DNA (Bird and Wolffe, 1999). Secondly, methyl-CpGs can recruit a protein family sharing the methyl CpG-binding domain (MBD), which was shown to indirectly induce silencing by further recruiting co-repressor complexes to methylated DNA (Boyes and Bird, 1991), including histone deacetylases or methyltransferases (Bird, 2002). The links between DNA and histone methylation are thought to be important for propagating (hetero-) chromatin states through mitosis. Accordingly, it was suggested that DNA methylation is not a primary silencing mechanism, but rather one that makes silencing more stable and potentially irreversible (Jones and Takai, 2001; Li, 2002; Pannell et al., 2000) apart from the exceptional occasions when DNA methylation is broadly erased (upon fertilisation, during germ cell development and/or somatic cell reprogramming) (Reik, 2007). However, this view has been recently challenged with evidence that DNA methylation can be dynamically regulated by active CpG demethylation. Cyclical methylation and demethylation of CpG dinucleotides, with a periodicity of 1-2 hours, was shown to be a characteristic of the oestrogen (E2)-responsive *pS2* gene, in human cells (Kangaspeska et al., 2008). The DNA deaminases *Aid* and *Apobec1* have been shown *in vitro* to deaminate 5-methylcytosine in DNA to thymine (Morgan et al., 2004); this results in T-G mismatches, which can be repaired by the base-excision repair pathway. Indeed, *Aid* and *Apobec1* are highly expressed by oocytes, stem cells and germ cells (Morgan et al., 2004). Another DNA-damage-responsive gene, *Gadd45* (growth arrest and DNA-damage-inducible 45), has also been implicated in DNA demethylation (Barreto et al., 2007; Rai et al., 2008) but despite intensive research, the mechanism for active erasure of the methylation mark remains largely unknown and controversial (Jin et al., 2008).

1.1.3. Covalent histone modifications and their impact on chromatin dynamics

Histones are highly evolutionarily conserved proteins with a flexible amino-terminal tail and a characteristic histone fold, a globular domain that mediates interactions between histones to form the nucleosome scaffold (Figure 1.1). Structural studies confirm that the histone N-terminal tails protrude outward from the nucleosome and can thus influence interactions with other nucleosomes, as well as with regulatory factors (Luger et al., 1997). The N-terminal tail of histones, as well as more recently defined positions in the globular domain, can carry post-translational modifications such as acetylation (of lysine residues), phosphorylation (serine and threonine), ubiquitination (lysine), methylation (lysine and arginine), sumoylation (lysine) and ADP ribosylation (Figure 1.2) (Kouzarides, 2007). Acetylation and methylation are two of the most extensively studied post-translational histone modifications. These modifications were first discovered in the 1960s but were considered by many to be artifactual (Allfrey et al., 1964). Only in the last decade has it become clear that covalent histone modifications have crucial regulatory functions in gene regulation by altering DNA-histone interactions within and between nucleosomes and, thus, affect higher-order chromatin structures (Kouzarides, 2007). In this regard, histone covalent modifications can collaborate to influence a multitude of cellular processes, including transcription, DNA replication timing, DNA repair and cell cycle progression (Bhaumik et al., 2007). Note that a large number of residues within histones can be simultaneously modified, with different types of modifications occurring in distinct patterns. For instance lysine acetylation correlates with chromatin accessibility and transcriptional activity that is the case of the acetylation of lysine 9 or 14 (H3K9ac or H3K14ac) while lysine methylation can have different effects depending on which residue is modified. The di- and trimethylation of lysine 4 (H3K4me2/3) are marks of active chromatin, conversely other lysine methylations such as H3K9me3 or H3K27me3 correlates with a repressed chromatin state (Bernstein et al., 2007; Peterson and Laniel, 2004). Moreover, these modifications regulate one another, providing regulatory cross-talk. That is, one modification can influence the occurrence of one or more subsequent modifications either in *cis* - on the same histone molecule, or in *trans* - between histone molecules or across nucleosomes (Campos and Reinberg, 2009). Histone modifications may even form an informational “code” that is read by the binding of effector proteins, driving the transition between transcriptionally active and silent chromatin states as well as distinct

downstream biological processes (Jenuwein and Allis, 2001; Latham and Dent, 2007; Strahl and Allis, 2000; Turner, 2000).

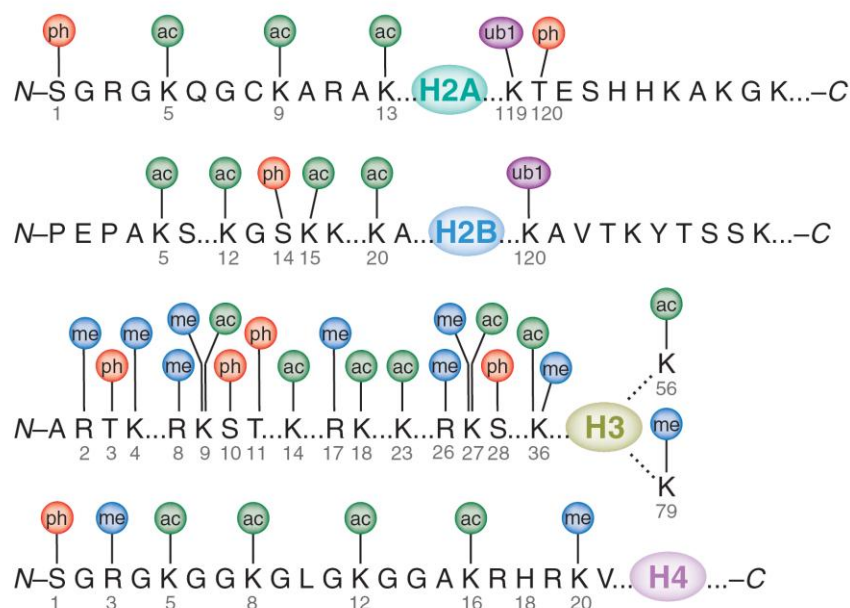


Figure 1.2. Post-translational modifications of human histones. Different covalent post-translational modifications of the four core histone (H2A, H2B, H3 and H4) have been identified and are here represented above the corresponding amino acid residue in the N-terminal tails of histones. These include acetylation (ac), phosphorylation by (ph), ubiquitylation (ub1) and methylation (me). Globular domains of each core histones are represented as ovals. Acetylation of lysines 9 or 14 of histone H3 (H3K9ac or H3K14ac) or dimethylation of lysine 4 (H3K4me2) are generally associated with open genomic regions and transcription, whereas others, such as H3K9me3 or H3K27me3, are indicative of a repressed chromatin state. Covalent histone modifications, alone or in combination, can influence nucleosome mobility and function as a platform for the recruitment of effector proteins (from Bhaumik et al. 2007).

1.1.3.1. Histone acetylation

Histone acetylation plays key roles in modulating chromatin structure and function (Shahbazian and Grunstein, 2007). Transcriptionally active chromatin was initially shown to be enriched in acetylated histones (Hebbes et al., 1988; Pogo et al., 1966), whereas heterochromatic regions of the genome were shown to be hypoacetylated (Jeppesen and Turner, 1993; Maison et al., 2002). These associations have been more recently explored and confirmed by numerous large-scale studies mapping histone modifications across the genome (Birney et al., 2007; Koch et al., 2007; Wang et al., 2009). Histone acetylation is a modification that neutralizes the positive charge of the target lysine. It can occur at specific lysines on the four core histones (Figure 1.2), however, modifications of histone H3 and H4 have been more extensively characterized. It has been proposed that histone acetylation promotes a “open” chromatin configuration by decreasing interactions between nucleosomes,

releasing the histone tails from the linker DNA, reducing fiber-fiber interactions and thus facilitating transcription machinery accessibility (Shahbazian and Grunstein, 2007). In addition, acetylation may promote transcription by providing binding sites for proteins involved in gene activation, such as those containing bromodomains. The chromatin remodelling complex SWI/SNF, which mobilizes nucleosomes and is important for transcriptional activation, is an example of a protein complex that binds to acetylated histones through a bromodomain (Hassan et al., 2002).

The acetylation state of a given chromatin locus is controlled by two classes of antagonizing histone modifying enzymes, histone acetyltransferases (HATs) and deacetylases (HDACs), which add or remove acetyl groups to/from target histones, respectively (Wade et al., 1997). Despite the general correlation between histone acetylation and gene transcription, it is not a stably increased level of acetylation that is important for transcription but rather a high level of acetylation, coupled with rapid deacetylation. It has been known for some time that the turnover of histone acetylation can be quite rapid. Remarkably, the fastest rates of acetylation turnover (with half-lives between 1-5 min) are associated with transcriptionally active regions, whereas slower rates of turnover (with half-lives between 30-60 min) do not correlate with transcriptional activity (Covault and Chalkley, 1980; Waterborg, 2002). Recent genome-wide studies showed that the dynamic cycle of acetylation and deacetylation by the transient binding of HATs and HDACs at promoters occupied by H3K4 methylation, may poise inactive genes for future activation (Wang et al., 2009). While this makes continuous demands on cell energy resources, it also offers an advantage to primed genes whose expression may need to be up- or downregulated relatively quickly, in respond to hormonal cues or growth factor stimulation, for example.

Three major groups of HATs have been identified as transcriptional coactivators: Gcn5-related N-acetyltransferases (GNATs), E1A-associated protein of 300 kDa (p300)/CREB-binding protein (CBP), and MYST proteins (Lee and Workman, 2007; Yang and Seto, 2008a). In contrast, histone deacetylation is generally associated with transcriptional repression, and HDACs have been identified as transcriptional corepressors. Known HDACs are divided into the Rpd3/Hda1 and sirtuin families. In humans, the former contains HDAC1, -2, -3, and -8 (class I, similar to yeast Rpd3); HDAC4, -5, -6, -7, -9, and -10 (class II, homologs of yeast Hda1); and HDAC11 (class IV) (Yang and Seto, 2008b). Within the sirtuin family, there are seven members in humans (SIRT1– 7, related to yeast Sir2; also known as class III HDACs) (Haigis and Guarente, 2006). Similar to HATs, HDACs have critical functions in many cellular pathways, and their misregulation has been linked with tumorigenesis. Thus chemicals

that inhibit HDAC activity are currently among the most promising drugs in anticancer therapies (Bolden et al., 2006).

Although histone acetylation is generally associated with actively transcribed genes, there are some reported exceptions to this rule. For example, acetylation of lysine 56 on histone H3 (H3K56ac) was shown to be dispensable for transcription but required for genome stability functions like DNA replication and repair (Li et al., 2008; Masumoto et al., 2005). More recently, H3K56ac was linked to the core transcriptional network of human ES cells, being enriched at active regulators that contribute to ES cell pluripotency and self-renewal, but also at inactive, lineage-specific genes that are poised for activation upon differentiation (Xie et al., 2009).

1.1.3.2. Histone methylation

Histone methylation at lysine and arginine residues in the N-terminal tails of histones H3 and H4 (Figure 1.2) has been linked to a number of cellular processes including DNA repair, replication, transcriptional activation and repression. Arginine residues can accept one or two methyl groups while lysine residues can be modified with one, two or three methyl groups. The enzymes that catalyse these modifications are known as histone methyltransferases and are either lysine (PKMTs) or arginine (PRMTs) specific. In contrast to acetylation (which broadly correlates with transcriptional activation), histone methylation is associated with different outcomes depending on the site and also the extent of methylation (Kouzarides, 2007).

There are at least eight human PRMTs for which some level of methyltransferase activity has been shown and implicated in both negative and positive regulation of transcription. These proteins share a canonical sequence domain that is associated with the binding sites for cofactor and substrate (arginine), yet the sequence conservation among these proteins is low. Estimates of the total number of PRMTs that are encoded by the human genome vary, depending on the method of sequence alignment and the level of alignment stringency that is applied (Copeland et al., 2009). Two PRMTs, PRMT1 and CARM1 were shown to be associated with transcriptional activation by methylating H4R3 (Strahl et al., 2001; Wang et al., 2001) or H3R2, H3R17 and H3R26 (Bauer et al., 2002; Schurter et al., 2001) respectively. Contrasting with this, PRMT5, that can methylate H3R8 and H4R3, was shown to act as a transcriptional repressor (Fabrizio et al., 2002; Pal et al., 2003). *In vivo*, CARM1 has also been described as having an important role in the choice between embryonic and

extraembryonic commitment of blastomeres in early mouse development (Torres-Padilla et al., 2007).

Methylation of lysine residues is known to occur on histone H3 at lysines 4, 9, 27, 36 and 79 and on histone H4 at lysine 20 (Figure 1.2). Histone H3 lysine 4, 36 and 79 methylation has been correlated with transcriptional activity being mostly found in euchromatic regions, whereas methylation of H3K9 and H4K20 are found predominantly but not exclusively at constitutive heterochromatin (Martin and Zhang, 2005; Peters et al., 2003). Methylated H3K27, in turn, is enriched at facultative heterochromatin, including the inactive X-chromosome (Peters et al., 2003; Rougeulle et al., 2004). Several proteins have been shown to specifically recognise and interact with methylated lysines (including chromodomain-, tudor domain- and WD40-repeat domain proteins), giving rise to diverse outcomes in terms of transcription (Martin and Zhang, 2005). Generally, proteins that associate with methylated H3K4 are involved in transcriptional activity (Pray-Grant et al., 2005; Wysocka et al., 2005), whereas those that bind methylated H3K9 or H3K27 are implicated in transcriptional repression (Fischle et al., 2003; Thiagalingam et al., 2003).

With the exception of DOT1-like histone H3 methyltransferase, all human PKMTs contain a ~130 amino acid domain, referred to as the SET domain, which constitutes the catalytic domain of these enzymes (Schneider et al., 2002). SET domain-containing enzymes can catalyze methylation of specific lysines on histones H3 and H4. The mammalian versions of H3K9 methyltransferases include Suv39h1, Suv39h2, G9A and ESET. PR-Set7 and two other SET domain proteins, Suv4-20h1 and Suv4-20h2, catalyze histone H4K20 methylation (Bhaumik et al., 2007; Shilatifard, 2006; Sims and Reinberg, 2006). Enhancer of zeste homologue 2 (Ezh2) is a SET domain protein that forms the catalytic subunit of the protein core of polycomb repressive complex 2 (PRC2). Ezh2 is a PKMT that catalyses the methylation H3K27. Although Ezh2 contains the catalytic active site, all members of PRC2 complex are required for full PKMT activity (see below). The methylation of H3K4 is catalysed by the Trithorax/MLL complex (Copeland et al., 2009; Lachner and Jenuwein, 2002; Milne et al., 2002).

Trithorax (TrxG) and Polycomb (PcG) group proteins were discovered in *Drosophila* as genes that, when mutated, would affect the spatially restricted expression pattern of *Hox* genes and developmental patterning (Lewis, 1978). These proteins are part of a widely conserved “cell memory” system that stabilises cell identity by restricting or maintaining transcription patterns set in the first stage of embryonic life, throughout development and in adulthood (Ringrose and Paro, 2004; Simon and Kingston, 2009). It is now accepted that PcG proteins regulate many genes in addition

to the *Hox* genes and several cellular processes, including proliferation, X-inactivation and pluripotency by regulating developmental genes such as transcription factors and signalling proteins (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Heard, 2005; Lee et al., 2006; Schwartz et al., 2006; Spivakov and Fisher, 2007). PcG proteins are involved in gene repression and silencing, whereas TrxG proteins counteract repressive PcG functions and maintain gene activity (Orlando and Paro, 1995; Pirrotta, 1998).

As previously mentioned, TrxG protein complexes were found to methylate H3K4 (Milne et al., 2002), a modification typically associated with active chromatin (Ruthenburg et al., 2007). The recruitment of the mammalian TrxG protein MLL is facilitated by basal transcriptional machinery as well as sequence-specific transcription factors (Ruthenburg et al., 2007). H3K4 methylation can also be catalyzed by several MLL unrelated methyltransferases and the recruitment of one of them, Ash1, was shown to be mediated by non-coding RNAs (Sanchez-Elsner et al., 2006). H3K4 methylation is also known to recruit transcriptional activators and nucleosome remodelling factors, such as the HAT-containing complexes SAGA and the ISWI-containing complex NURF (Pray-Grant et al., 2005; Wysocka et al., 2006).

The methylation of H3K27 is associated with PcG proteins, a family of multiple proteins involved in the epigenetic maintenance of the transcriptionally silent state. Progress over the past decade has defined two main protein complexes, Polycomb repressive complex 1 (PRC1) and PRC2, with fundamental roles in PcG-mediated silencing. The mammalian PRC2 complex contains the histone methyltransferase Ezh2 and also two other proteins, Eed (embryonic ectoderm development) and Suz12 (suppressor of zeste-12) that are essential for the catalytic activity of the complex (Cao and Zhang, 2004). This complex by catalysing the trimethylation of H3K27, creates a binding site for PRC1, that comprises Ring1A/Ring1B, Bmi-1, Polycomb and Polyhomeotic proteins (Wang et al., 2004a). In addition, Ring1A/B in cooperation with Bmi-1 were shown to act as ubiquitin ligases catalysing the ubiquitylation of lysine 119 of H2A (de Napoles et al., 2004), an event that was implicated in preventing active transcription of a number of developmental genes in ES cells (Jorgensen et al., 2006; Stock et al., 2007). This suggests that PRC1 and PRC2 collaborate in maintaining silent chromatin.

Histone methylation was regarded as a more permanent mark compared to other histone modifications such as acetylation or phosphorylation (Bannister et al., 2002). But with the discovery of novel histone arginine and lysine demethylases it is now considered a dynamic modification. Arginine demethylation is catalyzed by deiminases (Bannister et al., 2002). Members of the peptidyl arginine deiminase family,

including PADI4 deiminate arginine residues by converting them into citrulline (Nakashima et al., 2002; Wang et al., 2004b). The first histone demethylase to be identified was the lysine-specific demethylase 1 (LSD1), a protein previously found to be part of several HDAC complexes, that catalyses mono (me1)- or dimethylated (me2) H3K4 demethylation by an oxidative reaction (Shi et al., 2004). LSD1 has now been demonstrated to associate with both transcriptional repressors and activators. It demethylates H3K4me2/me1 when associated with the Co-REST complex at neuronal genes, or H3K9me2/me1 when associated with the androgen receptor (Metzger et al., 2005; Shi et al., 2004). Unlike LSD1, a structurally unrelated family of enzymes has been found capable of removing both mono-, di- and trimethyl marks (Tsukada et al., 2006). These proteins share a JmjC conserved motif and are thought to demethylate histones through a hydroxylation reaction. The methylation marks removed by these family of demethylases include H3K9, H3K27 and H3K36 (Klose and Zhang, 2007). Of interest, both H3K27 demethylases, JMJD3 and UTX, were found to be recruited with the H3K4 methylation machinery (Trithorax complex), resulting in the concerted demethylation of H3K27 (repression mark) and methylation of H3K4 (activation mark) at homeotic (*Hox*) loci during developmental activation (Agger et al., 2007; Lan et al., 2007). Conversely, functional interplays between PRC2 complexes (that methylate H3K27) and the H3K4 demethylase Rbp2 were also recently demonstrated in mouse ES cells upon differentiation (Pasini et al., 2008).

1.1.4. Spatial constraints – nuclear organisation

The nucleus is a highly organized entity with complex substructures that in addition to the regulation provided by DNA and histone modifications have the capacity to influence a variety of cellular processes, including gene expression (Dillon, 2008).

A striking feature of nuclear architecture is the non-homogeneous distribution in the nucleoplasm of components of the machinery required for gene transcription or repression (Iborra et al., 1996; Jackson, 1995; Negre et al., 2006). The positioning of a gene in the eukaryotic nucleus is also constrained by a variety of different parameters. The most basic constraint is the location of the gene as part of the linear sequence that forms each chromosome. Individual chromosomes form distinct territories in the nucleus (Figure 1.1). As a result of different compaction levels, different chromosome segments can adopt a complex organisation and topography within their chromosome territory (Croft et al., 1999; Edelman et al., 2001). For example, a polarised intranuclear distribution of gene-rich and gene-poor chromosomal segments has been

shown to be an evolutionarily conserved principle of nuclear organisation (Cremer et al., 2003; Tanabe et al., 2002). Gene-rich regions tend to be oriented towards the nuclear interior, whereas gene-poor regions tend to be oriented towards the periphery (Foster and Bridger, 2005). Although these territories were thought to be relatively discrete and separate, recent data has provided evidence of extensive intermingling between chromosomes (Branco and Pombo, 2006), which would make it easier for genes on different chromosomes to come into contact with one another. Importantly, chromosomes can also be dynamically repositioned during development and cell differentiation, implying that nuclear organisation vary between cell types (Parada et al., 2004). One example of such is the clustering of heterochromatin in distinct foci in the nucleus of mammalian cells, which can undergo morphological changes during differentiation. For instance, when ES cells differentiate towards neural progenitor cells, the number of heterochromatic foci per nucleus decreased while the average cluster size increased, suggesting a global reorganisation of chromatin upon differentiation (Williams et al., 2006).

Another parameter that can affect genes positioning in the nucleus is folding of the DNA into higher-order chromatin by trapping them within large aggregates of condensed heterochromatin. There is increasing evidence that gene repositioning to specific nuclear compartments is important for regulating gene expression (Gasser, 2002). In many cell types, heterochromatin has been shown to localize preferentially to the nuclear periphery, which raises the possibility that proximity to the nuclear envelope facilitates silencing. For example, the inactive X-chromosome is often observed at the nuclear periphery of female cells (Burke and Stewart, 2002). A mutation of the human nuclear lamin A protein has also been found to result in reduced amounts of constitutive and facultative heterochromatin and associated H3K9 and H3K27 methylation (Lammerding et al., 2005; Shumaker et al., 2006). However, there is not always a strict correlation between movement away from the periphery and gene activation. Indeed, the nuclear pores are important sites of transcription in yeast and genome-wide analysis has shown that nuclear pore-associated proteins bind preferentially to active genes (Casolari et al., 2004).

The corollary to large-scale chromatin condensation is that escape from regions of condensed chromatin could be involved in regulating transcriptional activation (Dillon, 2008). The mouse *Hoxb* locus provides an example of such chromatin looping. Fluorescent in situ hybridization (FISH) analysis showed how *Hoxb1* gene selectively “loops out” from its chromosome territory, moving towards the center of the nucleus, in *Hoxb1*-expressing cells during retinoic acid-induced differentiation of ES cells into neural progenitors (Chambeyron and Bickmore, 2004). Conversely, chromatin loops

can spatially separate genomic regions and contribute to gene silencing (Horike et al., 2005).

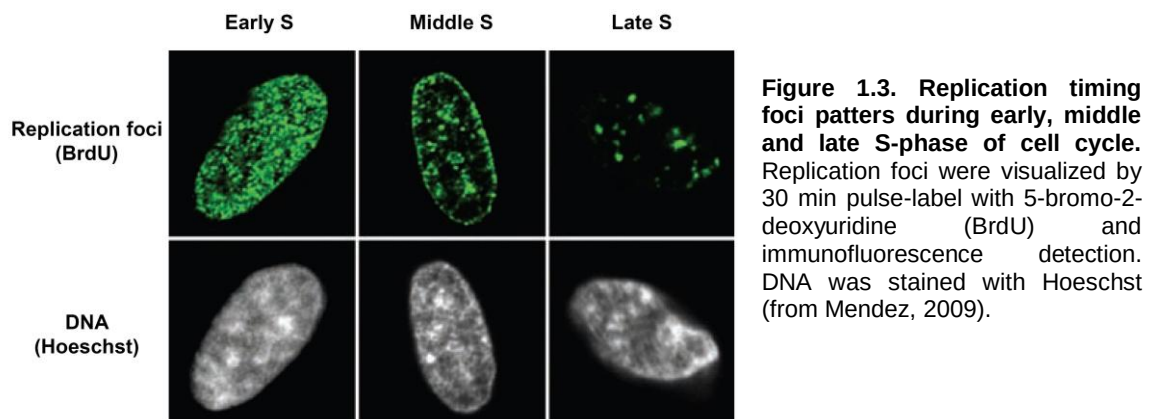
1.1.5. Temporal constraints – DNA replication timing

The eukaryotic cell cycle consists of an orderly series of events in which cells grow, replicate their chromosomes, and segregate the duplicated genome to the daughter cells (Aladjem, 2007). DNA replication is central to this process and occurs by a complex series of events involving the activation of thousands of replication initiation zones that are spaced every 100 Kb on average (Huberman and Riggs, 1968). Activation of replication origins does not occur synchronously. Usually this process occurs in a defined sequential order during the S-phase of the cell cycle - some origins fire early and some fire late in S-phase (Camargo and Cervenka, 1982). This pattern is clonally inherited through subsequent cell divisions, but it can change at genes undergoing developmental activation or silencing (Cimbora and Groudine, 2001; Perry et al., 2004).

Replication origins in *S. cerevisiae* are known to be DNA sequence-specific (Van Houten and Newlon, 1990). These sequences are designated autonomously replicating sequences (ARSs) and when inserted into a plasmid are able to confer autonomous replication (Stinchcomb et al., 1979). Although the time of firing during S-phase has been defined for many ARSs, comparative sequence analyses at these sites has revealed no clear differences between early and late firing origins (Wintersberger, 2000). In metazoans, no consensus motifs that direct DNA replication from specific sites have been found and current evidence indicates that origin recognition complexes do not rely on sequences to be targeted to origins. The increase in genome size and the diversification of transcriptional programmes required for cellular specialisation in metazoans seem the most plausible explanations for the differences observed in *S. cerevisiae* and metazoan replication origin specification (Mechali, 2001).

The time at which a locus replicates during S-phase (locus replication timing) is likely to be regulated at two levels; *locally* by the regulation of origin activity as described above and more *globally* by the mechanisms that determine the replication timing of entire chromosomal domains (Mendez, 2009). Chromosome replication domains were originally detected in the 1970s by pulse-labelling cells with nucleotide analogue 5-bromo-2-deoxyuridine (BrdU) at different times during S-phase and visualizing its incorporation into discrete areas of condensed mitotic chromosomes

(Latt, 1973; Stubblefield, 1975). It was then noted that the banding patterns observed after Giemsa staining overall coincide with early and late replicating domains; the dark G-bands tend to be late replicating, whereas early-replicating regions map to reverse (R) light bands (Dutrillaux et al., 1976). These correlations were further confirmed by replication timing high resolution maps of human chromosomes using microarray technology (Costantini and Bernardi, 2008; Woodfine et al., 2004). The dynamics of DNA replication can also be monitored by the detection of BrdU incorporation onto newly synthesized DNA during S-phase. Replication foci appear as intranuclear punctuated structures, which form different patterns as the S-phase progresses. In early S-phase many small foci are distributed all over the nuclei, in middle S foci are preferentially assembled around the nucleoli and nuclear periphery while late S is characterized by clusters of foci that correspond to heterochromatic regions (Figure 1.3) (Dimitrova and Gilbert, 1999; O'Keefe et al., 1992).



The temporal order of gene replication during S-phase has been proposed to be important for maintaining gene expression patterns (Gilbert, 2002). Early replication is generally associated with accessible chromatin carrying acetylated histones (Lin et al., 2003; Perry et al., 2004; Vogelauer et al., 2002), whereas late replication is a characteristic of constitutive heterochromatin and some facultative heterochromatin, formed through gene repression (Gilbert, 2002; Peters and Schubeler, 2005). In accordance, housekeeping genes constitutively replicate early in S-phase, whereas many tissue-specific genes are developmentally regulated so that they replicate later in most cell types but earlier in the tissue of expression (Goren and Cedar, 2003). In the case of imprinted genes in which only one of the two parental alleles is active, the expressed allele replicates early and the repressed allele replicates late (Simon et al., 1999). However, there are also many inactive genes that replicate early. For example genes involved in stress response and apoptosis, which are not expressed during normal proliferation but need to be accessible for rapid transcription in special

circumstances. These results suggest that a correlation may exist between early replication timing and the potential for gene expression rather than actual expression. A small number of late replicating genes with strong promoters that may render regulatory elements accessible was also shown to be expressed (Hiratani and Gilbert, 2009; Mendez, 2009).

As previously mentioned, genome-wide studies in *Drosophila* and mammalian cells suggested that replication timing might also correlate with its genomic context. More specifically, GC-rich, gene-dense regions tend to replicate early in S-phase, while AT-rich, gene-poor regions with high LINE (linear interspersed nuclear element) content tend to replicate later (Hiratani et al., 2004; Schubeler et al., 2002; Woodfine et al., 2004). The nature of such correlation between replication timing and genomic features remains unclear, however. Interestingly, candidate-based and genome-wide approaches have begun to define possible cross-tables between replication timing and histone marking. In particular, early-replicating domains map to regions enriched in histone modifications that are characteristic of active chromatin (Hiratani et al., 2008; Karnani et al., 2007; Perry et al., 2004). The presence of repressive chromatin marks in late replication timing domains was conversely reported in human (Karnani et al., 2007) but not in mouse replication timing maps (Hiratani et al., 2008), although the reason for this discrepancy is not clear. Interestingly, the chromatin modification that correlates most strongly with early replication timing is histone acetylation (Mendez, 2009). In particular, Cedar and colleagues demonstrated how the recruitment of HATs to the human β -globin locus in a non-erythroid cell line would change its replication timing from late to early. Conversely, tethering an HDAC to the locus in an erythroid cell line, where β -globin is normally expressed, was sufficient to change its replication timing from early to late (Goren et al., 2008). A correlation between histone acetylation and replication timing was further supported by the fact that the use of HDAC inhibitors (sodium butyrate or trichostatin A) was sufficient to remove the replication imprinting signals in two loci (*Wt1* and *Igf2*) on human chromosome 11p (Bickmore and Carothers, 1995). It was hypothesised that histone acetylation renders replication origins more efficiently or specifically orchestrates earlier firing leading to earlier conclusion of locus replication (Aparicio et al., 2004; Vogelauer et al., 2002). Besides histone acetylation, the influence of other DNA or histone modifications on replication timing determination is less certain and may be restricted to specific chromosome structures (Jorgensen et al., 2007; Mendez, 2009).

Single-cell microinjection experiments in mammalian cells further suggested that replication timing might, in turn, play a key role in maintaining heritable epigenetic states through mitosis. A reporter gene microinjected into early S-phase nuclei was

preferentially assembled into nucleosomes with acetylated histones and expressed with higher efficiency than the same DNA microinjected in late-S nuclei, which was packed into deacetylated histone nucleosomes (Zhang et al., 2002). This result, which strongly suggests that the timing of replication can influence on chromatin structure, has been supported by the recent demonstration that DNA sequences packaged in nucleosomes characteristic of late replication (i.e. containing deacetylated histones) can be reassembled with acetylated histones if replicated in early S-phase, and vice versa (Lande-Diner et al., 2009). The presence at replication forks of different sets of chromatin assembly factors and histone-modifying enzymes in early or late S-phase is consistent with this function (McNairn and Gilbert, 2003). Although these different sets of experiments strongly suggest a link between replication timing and chromatin structure, how this relationship is precisely regulated remains unclear. So far, indirect evidence have supported two rather intuitive models in which open chromatin dictates early replication or early replication dictates open chromatin. Combining these models creates an attractive scenario in which DNA replication provides a means to inherit chromatin states that, in turn, regulate replication timing in the subsequent cell cycle, so that once a timing program is established, it forms a self-reinforcing auto-regulatory loop that is very stable for many generations (Hiratani and Gilbert, 2009).

Replication timing is also subject to developmental control. A study comparing the replication timing of a panel of approximately 50 tissue-specific genes in undifferentiated ES cells and ES-derived neural progenitors revealed that neural commitment results in dynamic changes in the replication timing of many genes. The replication of several ES-associated genes was delayed upon neural induction, alongside with an advance in replication of a number of neural-associated genes (Hiratani et al., 2004; Perry et al., 2004). Moreover, these replication timing “shifts” from early (in ES cells) to late S-phase (in neural progenitors), was coupled to significant deacetylation of loci, for example, at the *Rex1* promoter region (Perry et al., 2004). In a subsequent study, the replication timing of a panel of loci encoding transcription factors associated with lineage specification was analysed in pluripotent ES cells and in differentiated cell types such as haematopoietic progenitors (HP) and T lymphocytes (T) (Azuara et al., 2006). In ES cells, many loci were early-replicating, whereas in HP and T many genes replicated later. This difference in the replication timing of neural-associated genes was detected despite the fact that these genes were not expressed by ES, HP or T cells. Thus, these changes were not merely reflecting differences in gene activity. Additionally, the replication timing analysis of a subset of genes in two different embryonic carcinoma (EC) cell lines (F9 and P19) with a more restricted differentiation capacity than ES cells, showed that EC replication profiles were distinct

from that of ES cells, reflecting their distinct potentials. Together, these data suggested that replication timing profiles can be used to identify the lineage potential and developmental stage of a given cell population.

1.2. Early mammalian development

1.2.1. Early embryo development and pluripotency

The combination of the oocyte and sperm in the zygote represents the climax of cellular potency. The zygote is the only unequivocally totipotent cell in the life cycle with potential to form every cell type of the organism (including specialised extraembryonic cells) (Hemberger et al., 2009). Progressively, throughout development specific programmes of gene expression are established and early progenitor cells gradually become more restricted in their potential while they commit to different cellular lineages (Morgan et al., 2005).

The access to human embryos and *in vivo* tissues from postimplantation stages is clearly very limited. Consequently the steps leading towards the earliest lineage decisions in mammals have been best studied in mice (Pera and Trounson, 2004). Following fertilization, the mouse zygote undergoes a series of sequential cleavage divisions during the first two days of development to produce a 4-cell embryo that progresses to an 8-cell embryo. At this point the cells (or blastomeres) maximise cell-cell contact, giving rise to a compacted morula (Figure 1.4). Asymmetrical cell divisions at the morula stage along a basolateral cleavage plane accompanies the first differentiation event in the mouse embryo that gives rise to an internal cavity, the blastocoel and two visibly distinct subpopulations - the inner cell mass (ICM) surrounded by an epithelial cell layer that will become allocated to the extraembryonic trophectoderm (TE) lineage (Arnold and Robertson, 2009). This typical embryonic structure characterises the early blastocyst stage that corresponds to embryonic day 3.5 (E3.5) of development. Cell fate studies (mostly using mouse chimeras) have shown that ICM cells are pluripotent, giving rise to all tissues of the foetus. In contrast, TE cells are more restricted in their potential (multipotent) and solely contribute to placenta formation (Rossant, 2001). Around the time of implantation in the late blastocyst, the ICM further differentiates into the epiblast or primitive ectoderm and another extraembryonic lineage, the hypoblast or primitive endoderm (PrE) (Figure 1.4) (Ralston and Rossant, 2005). Epiblast cells are also pluripotent, giving rise to the three germ layers (ectoderm, mesoderm and endoderm) upon gastrulation as well as

primordial germ (PGC) cells. Cells from the PrE (like the TE) are multipotent, being able to exclusively differentiate into the visceral and parietal endoderm layers that will line the yolk sac (Bielinska et al., 1999; Rossant, 2001). Besides providing growth support and protection within the uterus, the extraembryonic TE and PrE lineages are sources of signals to the embryonic lineages to promote correct patterning and differentiation (Rossant and Tam, 2004). The germ cells represent the end point of a differentiation cascade and can be considered highly specialised. At the same time, however, they have the capacity to return to totipotency (as a consequence of germ cell-specific epigenetic reprogramming) when the gametes come together in the zygote (Hemberger et al., 2009).

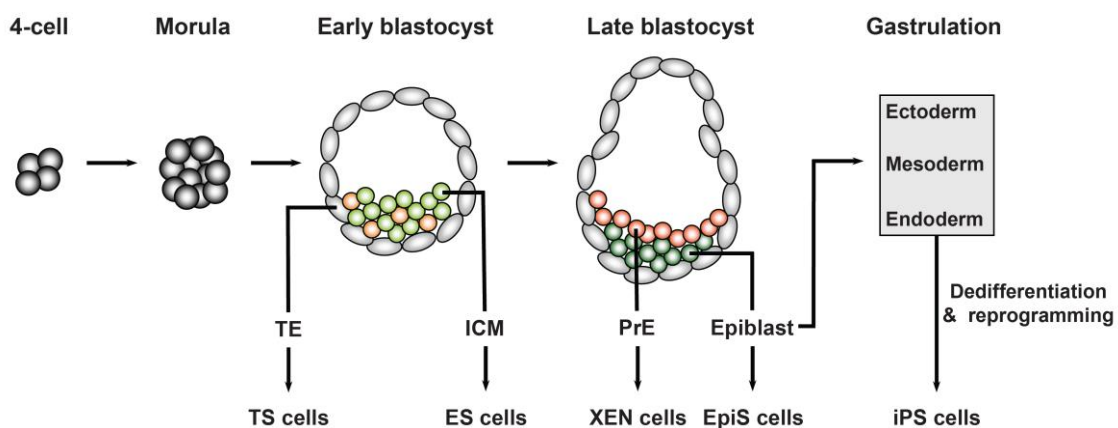


Figure 1.4. Schematic representation of the early stages of mouse development and derivation of stem cell types. Following fertilization the mouse zygote undergoes two rounds of cleavage division that leads to the 4-cell stage embryo. The next round of cells division leads to the eight-cell embryo or early morula where the cells become more compact. Subsequent divisions lead to the generation of distinct inside and outside cell populations and formation of an internal cavity, the blastocoel. Two distinct lineages are apparent in the early blastocyst stage: the trophoblast (TE) and the inner cell mass (ICM) from where the trophoblast stem (TS) cells and the embryonic stem (ES) cells are derived, respectively. The TE is also present in the late blastocyst, but the ICM has meanwhile segregated into two lineages: the epiblast and the primitive endoderm (PrE). Epiblast stem (EpiS) cells can be derived from the epiblast immediately after implantation and extraembryonic endoderm (XEN) cells from the PrE. The TE and PrE lineages will give rise to extraembryonic tissues that support the embryo, while the epiblast will give rise to the three germ layers (ectoderm, mesoderm and endoderm) during gastrulation that will differentiate into all cell types of the adult organism. Recent breakthroughs in dedifferentiation techniques allow the reprogramming of fully differentiated somatic cells into induced pluripotent stem (iPS) cells, which closely resemble ES cells in their developmental potency.

1.2.2. Embryo-derived stem cells

1.2.2.1. Mouse embryo-derived stem cell

In mammalian blastocysts, three distinct cell lineages can be discriminated before implantation: the ICM, TE and PrE lineages. Different populations of stem cell lines were successfully derived from each of these lineages, namely ES cells from the

ICM (Evans and Kaufman, 1981; Martin, 1981), trophoblast stem (TS) cells from the TE (Tanaka et al., 1998) and extraembryonic endoderm (XEN) cells from the PrE (Kunath et al., 2005) (Figure 1.4). All of these cells share a similar characteristic: they can self-renew without differentiating in culture while faithfully retaining the developmental competence of their cognate lineage. Even though they share this property, all these stem cells are remarkably different in terms of their morphology, gene expression patterns and growth factor requirements that typify the lineage from where they are derived (Rossant, 2008). Pluripotent mouse ES cells contribute to all tissues of the foetus, but poorly to PrE and rarely to TE lineages (Beddington and Robertson, 1989); TS cells only contribute to the trophoblast lineages of the placenta (Tanaka et al., 1998) and XEN cells contribute exclusively to the parietal and visceral endoderm lineages (Kunath et al., 2005). Although, the molecular mechanisms underlying self-renewal and lineage restriction are not fully understood, several factors have been clearly identified as key regulators of cell fate in the early embryo. The expression or combination of key lineage-specific transcription factors, such as *Oct4/Sox2/Nanog*, *Cdx2/Eomes/Gata3* and *Gata6/Gata4/Sox7* was shown to be determinant in the maintenance of ES, TS and XEN cells, respectively (Home et al., 2009; Ralston and Rossant, 2005; Rossant, 2008).

Recently, pluripotent epiblast stem (EpiS) cells have been established from the epiblast lineage, isolated from E5.5 to E6.5 postimplantation mouse embryos (Figure 1.4) (Brons et al., 2007; Tesar et al., 2007). Although EpiS cells like ES cells are able to differentiate into the three embryonic germ layers, they have difficulties in contributing to the germ line and forming chimeras after introduction into blastocyst (Brons et al., 2007; Tesar et al., 2007). Embryonic germ (EG) cells can also be derived from early PGC cells between E8.5 and E11.5 postimplantation embryos (Matsui et al., 1992; Resnick et al., 1992) and likewise ES cells, upon blastocyst injection they can contribute extensively to chimeric mice including germ cells (Labosky et al., 1994; Stewart et al., 1994). More recently, pluripotent stem cells were also derived from both neonatal and adult mouse testis (Guan et al., 2006; Kanatsu-Shinohara et al., 2004), suggesting that pluripotency can be maintained during germ cell development.

1.2.2.2. Mouse and human embryonic stem (ES) cells

Pluripotent cells are present only transiently in embryos *in vivo*, as they quickly differentiate into various somatic cells through development. However, it is possible to isolate *ex vivo* pluripotent mouse (mES) and human (hES) ES cells from the ICM under

defined growing conditions that promote proliferation in the absence of differentiation (Yu and Thomson, 2008).

Mouse ES cells were first isolated in 1981 from blastocyst stage embryos (Evans and Kaufman, 1981; Martin, 1981) and exhibit two remarkable features in culture. First, under the appropriate conditions, they can be propagated indefinitely as a stable self-renewing population maintaining their cell identity. This immortalized phenotype allows ES cells to be cultured over extended periods of time. Upon differentiation, this feature is lost and progeny succumb to cellular aging mechanisms (Hayflick limit) as all other non-transformed primary cells. A second feature is that during extended culture, mES cells clonally derived from a single cell retain their pluripotency and differentiate into the same range of cell types as those formed by the ICM *in vivo*, when injected into blastocyst (Martin, 1981). Mouse ES cells have been derived from cleavage stage embryos and even from individual blastomeres of 2- to 8-cell stage embryos (Chung et al., 2006; Wakayama et al., 2007).

The derivation of hES cells provides a unique opportunity to study early human development and is thought to hold great promise for regenerative medicine (Pera and Trounson, 2004; Reubinoff et al., 2000; Thomson et al., 1998). There was a considerable delay however, between the derivation of mES cells in 1981 and the derivation of hES cells in 1998 (Thomson et al., 1998), in spite of several earlier attempts. This delay was primarily due to species-specific ES cell differences and suboptimal human embryo culture conditions (Yu and Thomson, 2008). For example, the isolation of ICMs from human blastocysts had been previously reported (Bongso et al., 1994), but subsequent culture in media supplemented with leukemia inhibitory factor (LIF) and serum, conditions that allow the derivation of mES cells, resulted in cells differentiation. Improvements in the culture conditions for human IVF embryos (Gardner et al., 1998) and experience in growing non-human primates cell lines (Thomson et al., 1995; Thomson et al., 1996) permitted the subsequent derivation of hES cells in 1998. A common alternate assay for ES cells potency, particularly for hES cells where embryo transfer is not permitted for research purposes, is to inject ES cells into immunocompromised mice where they form mixed cell tumours known as teratomas. Another approach is to grow the cells in non-adherent conditions leading to the formation of three-dimensional structures called embryoid bodies (EBs) (Vallier and Pedersen, 2005). Like mES cells, hES cells are capable of maintaining their developmental potential to contribute to derivatives of all three germ layers, even after prolonged undifferentiated proliferation and/or clonal derivation (Amit et al., 2000). In addition, and in contrast to mES cells (Beddington and Robertson, 1989), hES cells can give rise to extraembryonic lineages including TE (Thomson et al., 1998; Xu et al.,

2002) and PrE-like cells (Pera et al., 2004). Similarly to mES cells, hES cells have been derived from morula, later blastocyst stage embryos (Stojkovic et al., 2004; Strelchenko et al., 2004), single blastomeres (Klimanskaya et al., 2006), and parthenogenetic embryos (Lin et al., 2007; Mai et al., 2007; Revazova et al., 2007). However, it is not known yet how similar or distinct pluripotent hES cell lines are between them, especially in terms of their developmental potential. Recently some studies using directed differentiation methods or EBs, suggested that independently-derived hES cells may greatly differ in their developmental potential, with some cell lines showing marked differences in their differentiation propensity (Kim et al., 2007; Osafune et al., 2008). Gene expression profiling also revealed that distinct hES cell lines can display differences in the expression of specific genes (Abeyta et al., 2004; Allegrucci and Young, 2007; Skottman et al., 2006), although being clearly distinguishable from other cell types. In addition, epigenetic profiling showed that hES cells can display some degree of epigenetic variability and instability across lines and cultures (Allegrucci et al., 2007; Maitra et al., 2005).

Nevertheless, mouse and hES cells share the expression of high levels of alkaline phosphatase (Adewumi et al., 2007) and key transcription factors that were identified as being important in the maintenance of the ICM/ES cell pluripotency. These include the POU-family transcription factor Oct4 (Nichols et al., 1998; Palmieri et al., 1994), a homeodomain DNA-binding protein called Nanog (Chambers et al., 2003; Mitsui et al., 2003) and the SOX-family transcription factor Sox2 (Avilion et al., 2003). Embryos lacking *Oct4* do not form ICM, consisting exclusively of TE tissue (Nichols et al., 1998). Similarly, the *in vitro* deletion of *Oct4* leads to the conversion of ES to TS-like cells (Niwa et al., 2000). Mouse ES cells deficient for *Nanog* are still able to self-renew, but show an increased propensity to differentiate towards XEN-like cells and cannot give rise to mature germ cells (Chambers et al., 2003; Chambers et al., 2007). Functional studies further suggested similar roles for these transcription factors in maintaining the pluripotent stem cell state of hES cells. For example, reduced expression of *Oct4* in hES cells by genetic manipulation also results in TE differentiation and *Nanog* gene expression inhibition also causes hES cells differentiation to extraembryonic cell lineages (Hyslop et al., 2005; Zaehres et al., 2005). The function of Sox2 in early embryonic development is more difficult to study due to the presence of maternally-derived Sox2 protein, however after implantation (as the maternal contribution is largely lost), Sox2-null embryos generate exclusively TE and PrE lineages (Avilion et al., 2003). Similarly Sox2 deletion in ES cells also leads to their conversion in TS-like cells (Masui et al., 2007). Sox2 was shown to heterodimerise with Oct4, acting synergistically by regulating the expression of stem

cell-associated genes (Chew et al., 2005; Okumura-Nakanishi et al., 2005; Tomioka et al., 2002). Recently, it was suggested that Sox2 may stabilise pluripotent ES cells by maintaining the requisite level of Oct4 expression (Masui et al., 2007).

Mounting evidence suggests that ES cells in culture may not all be identical but may exist in different, interconvertible substates that have significant consequences for their behaviour and ability to differentiate. In particular, undifferentiated mES cell cultures are likely to contain different interconvertible subpopulation of cells that either correspond to an ICM-like or epiblast-like identity (Enver et al., 2009). Expression levels of *Rex1* (an ICM marker) (Toyooka et al., 2008), *Nanog* (Chambers et al., 2007) and *Stella* (an ICM and germ cell lineage marker) (Hayashi et al., 2008) are highly variable in mES cells, contrasting with the apparent homogeneity of *Oct4* and *Sox2*. Although detailed studies of the dynamics of expression of these markers in hES cells have yet to be carried out, the patterns of cell surface marker expression present on pluripotent hES cells also points to the existence of interconvertible substates of the stem cell compartment (Enver et al., 2005; Stewart et al., 2006) a feature also evident in several adult stem cell systems (Booth and Potten, 2000; Hu et al., 1997; Jones et al., 2007).

1.2.2.3. Embryonic equivalence of mouse and human ES cells – a developmental paradox

Despite the similarities of the intrinsic control of mouse and human pluripotent ES cells by Oct4, Nanog and Sox2, several phenotypic differences between mES and hES cells have been reported particularly in terms of cells morphology and cell-surface markers (Pera and Trounson, 2004). Mouse ES cells grow as three-dimensional dome-shaped colonies whereas hES cells grow in colonies as thin and flat layers. The embryonic markers defined by the antibodies SSEA-1, -3 and -4 against cell surface glycolipids are expressed differently in undifferentiated mES (SSEA1+) and hES (SSEA3/4+) cells. Also the cell-surface proteoglycan recognised by several monoclonal antibodies, including TRA-1-60 and TRA-1-81 are reactive with hES cells but not detected in mES cells (Thomson et al., 1998). The question of whether differences between mES and hES cells are related to species differences or relate to differences in the stages of embryonic development from which they arise, remains open (Lovell-Badge, 2007). On this line of thoughts, one study showed that ICM cells of human blastocysts and hES cells share the same pattern of antigen expression. SSEA3, SSEA4 and TRA-1-60 were all found to be expressed in the ICM of *in vitro* human preimplantation blastocysts cultured, although a limited number of embryos were

analysed. Note that these markers are not expressed in the mouse ICM, highlighting possible intrinsic species-specific differences in the expression of these markers as seen in mES and hES cells (Henderson et al., 2002). Consistent with this idea, mouse and human pluripotent cell lines harbour the same phenotypic differences whether they originate directly from embryos (ES cells) or from PGC cells through teratocarcinogenesis (embryonic carcinoma (EC) cells) (Pera et al., 2000). Thus, human EC cells resemble hES cells, mouse EC stem cells resemble mES cells, and the two cell-types show the same inter-species differences. An alternative view would be that hES cells are in a pre-epiblast stage of commitment, where all lineage pathways including TE are still open, whereas mES cells represent the already lineage-restricted epiblast progenitor of the blastocyst. This could relate to the slower pace of lineage restriction in the early stages of human development compared with mouse (Pera and Trounson, 2004). Most recently, however, two groups have been able to derive several pluripotent cell lines (EpiS cells) directly from the early postimplantation epiblast in the mouse (Brons et al., 2007; Tesar et al., 2007). Like mES and hES cells, these novel pluripotent cell lines express the three key pluripotency transcription factors, Oct4, Sox2, and Nanog, yet are distinct from mES cells in their genetic and epigenetic state as well as the signals driving their differentiation. In contrast, mEpiS cells share a number of features with hES cells, leading to the suggestion that hES cells might in fact be an *in vitro* equivalent of early postimplantation epiblast cells. For example, mEpiS cells derivation failed in the presence of LIF and/or BMP4, the two factors required for the derivation and self-renewal of mES cells. In contrast, similar to hES, FGF and Activin/Nodal signalling appear critical for EpiS cell derivation and BMP4 induces differentiation towards TE. EpiS cells were able to generate tissues from all three germ layers *in vitro* and to form teratomas. Data mining of the deposited gene expression arrays for mEpiS cells shows that they do express pluripotent markers but also seem to be expressing higher levels of mesoderm and definitive endoderm transcripts as compared to mES cells. This suggests that mEpiS cells, and by extension hES cells, may be in an unusual transition state as compared to mES cells: poised on the verge of differentiation into multiple directions (Rossant, 2008).

1.2.3. Signalling pathways required for ES cells self-renewal and pluripotency

Although hES and mES cells share a number of intrinsic regulators, the extrinsic signals that drive ES cell self-renewal and pluripotency have undergone

significant species-specific deviation (Pera and Trounson, 2004). The undifferentiated state of mES cells can be maintained in culture by foetal calf serum in combination with LIF, a cytokine shown to sustain ES cell growth and self-renewal (Smith et al., 1988; Williams et al., 1988). LIF is known to act mainly via the JAK-STAT3 signalling pathway (Niwa et al., 1998; Yoshida et al., 1994). The activation of this pathway involves heterodimerisation of LIF receptor with the gp130 membrane protein (Yoshida et al., 1994) that leads to the activation of JAK tyrosine kinase and subsequent recruitment and activation of STAT3, that translocates to the nucleus and regulates the expression of genes required for “stemness” (Niwa et al., 1998; Raz et al., 1999). Withdrawal of LIF triggers differentiation to a mixture of lineages (Nichols et al., 2001) and the provision of activated STAT3 renders mES cells independent of LIF (Matsuda et al., 1999). In serum-free medium, LIF alone is insufficient to prevent mES cell differentiation, but in combination with bone morphogenetic protein (BMP; a member of the TGF- β superfamily), mES cells self-renewal is sustained (Ying et al., 2003). BMPs induce the expression of inhibitor of differentiation (Id) proteins through the activation of the Smad 1/5/8 pathway. However, studies indicate that BMPs might also act independently of Smads, through inhibition of the MAPK pathways. The latter is supported by the facts that ES cells can be derived from blastocysts lacking Smad4 (the common partner for all Smads) (Sirard et al., 1998) and that inhibition of p38 MAPK allowed derivation of ES cells from blastocysts lacking BMP type I receptor Alk-3, which were previously refractory to ES cell derivation (Qi et al., 2004). Recently, it was shown that mES cells can be efficiently derived and maintained in minimal conditions without extrinsic requirements for growth factors or cytokines (Ying et al., 2008). This work showed that ES cells can be maintained in their self-renewing “ground state” by blocking differentiation-inducing signalling using a cocktail of inhibitors (for FGF receptor tyrosine kinases, ERK and GSK3 pathway). In normal development, however, there is no apparent requirement for LIF, gp130 or STAT3 prior to gastrulation, this seems to be a specific requirement of embryos in diapause (Nichols et al., 2001). Also, homozygous Alk-3 mutant mouse embryos can develop normally to early postimplantation stage (Mishina et al., 1995). Thus, the pluripotent ICM/epiblast could use alternative signalling pathways for undifferentiated proliferation.

Mitotically inactivated fibroblast feeder layers and serum-containing medium were used in initial attempts to derive hES cells, essentially the same conditions used for the derivation of mES cells prior to the identification of LIF (Thomson et al., 1998). However, it now appears largely to be a lucky coincidence that fibroblast feeder layers support both mES and hES cells, as the specific factors used to sustain mES cells do not support human. In contrast to mES cells, LIF signalling is not active in

undifferentiated hES cells (Daheron et al., 2004; Humphrey et al., 2004) and BMPs induce TE differentiation, either in the presence (Gerami-Naini et al., 2004; Pera and Trounson, 2004) or absence (Xu et al., 2002) of serum. Conversely, the MEK kinase pathway that promotes differentiation in mES cells (Nichols et al., 2001), is activated by fibroblast growth factor 2 (FGF2) and required to maintain the undifferentiated state of hES cells (Amit et al., 2000). At higher concentrations, FGF2 allows feeder independent growth of hES cells cultured in commercial serum replacement (Levenstein et al., 2006; Wang et al., 2005; Xu et al., 2005). The mechanism through which these high concentrations of FGF2 exert their functions is not fully understood, although one of the effects is suppression of BMP signalling (Xu et al., 2005) present in the serum and the widely used commercially available serum replacement. FGF2 is also able to up-regulate the expression of TGF- β ligands in both feeder cells and hES cells, which in turn, could promote hES cell self-renewal (Greber et al., 2007). Several studies suggest that FGF and TGF- β /Activin/Nodal signalling interactions are important for maintenance of hES cell pluripotency (Amit et al., 2004; Noggle et al., 2005; Wang et al., 2005; Xu et al., 2005), as recently described for mEpiS cells (Brons et al., 2007; Tesar et al., 2007). Both TGF- β and Activin A have strong positive effects on undifferentiated proliferation of hES cells in the presence of low or modest concentrations of FGF2, and based on inhibitor studies, it has been suggested that this signalling pathway is essential for hES cell self-renewal and pluripotency (Beattie et al., 2005; James et al., 2005; Vallier et al., 2005; Xiao et al., 2006). The TGF- β /Activin signalling is transmitted by the phosphorylation of Smad2/3 (Schmieder and Hill, 2007). Inhibition of the TGF- β /Activin receptors by the chemical inhibitor SB431542 reduces Smad2/3 phosphorylation and induces differentiation (James et al., 2005; Vallier et al., 2005; Xiao et al., 2006). Recently, it was shown that Smad2/3 bind directly to the Nanog proximal promoter in hES cells, modulating its expression (Vallier et al., 2009a; Xu et al., 2008). In contrast, BMP-responsive Smad1 also binds to the Nanog promoter (Xu et al., 2008). These suggest that an intricate balance of Smad2/3 and Smad1/5/8 could modulate the expression of *Nanog* and potentially determine the choice between undifferentiated and lineage-committed fates. Although other growth factors have been reported to have a positive effect on hES cell growth including Wnt (Sato et al., 2004), IGF1 (Bendall et al., 2007), heregulin (Wang et al., 2007), sphingosine-1-phosphate (S1P), and PDGF (Pebay et al., 2005), there are clearly additional important pathways yet to be identified. Defining the molecular mechanisms of hES cells self-renewal and pluripotency will enable growth of hES cells under defined conditions and aid controlled differentiation of cells into specified lineages, thus providing cells suitable for therapeutic applications.

1.2.4. Transcriptional networks that control pluripotency

A network of transcription factors including *Oct4*, *Nanog* and *Sox2* have been shown to be crucial for development of pluripotent cells in the preimplantation embryo and for maintenance of pluripotency of human (Boyer et al., 2005) and mouse (Ivanova et al., 2006; Loh et al., 2006; Mathur et al., 2008) ES cells. To gain more insights into these transcriptional regulatory networks operating in ES cells, it was used chromatin immunoprecipitation (ChIP) assays to map the genome-wide binding sites for *Oct4*, *Nanog* and *Sox2* in human and mouse. Despite the low overlap between human and mouse datasets, these authors found that *Oct4* and *Nanog* (and in human ES cells, also *Sox2*) co-occupy a large number of developmental genes, including both transcriptionally active and inactive genes in ES cells. Interestingly, among the targeted genes *Oct4*, *Nanog* and *Sox2* were shown to be able to regulate themselves. Other active targets were transcription factors, chromatin remodelling and components of signalling pathways identified to be important for ES cell biology, as for example, *Cripto* and *Lefty2*, two members of the TGF- β pathway. In contrast, a large portion of the inactive targets identified encode developmentally important transcription factors, such as genes involved in the specification of ectoderm, endoderm, mesoderm and extra-embryonic lineages. These findings have been collectively taken to suggest that *Oct4*, *Sox2* and *Nanog* are as so-called “master regulators” of the pluripotent stem cell identity (Spivakov and Fisher, 2007). However, the network that control pluripotency is much more complex and many other players are currently being discovered (Dejosez et al., 2008; Kim et al., 2008).

1.2.5. Epigenetic control of pluripotency

The chromatin features of pluripotent ES cells seem to reflect a generalised “open” and transcriptionally “permissive” state (Meshorer and Misteli, 2006). In contrast, ES cell differentiation resulted in an increase in global levels of the heterochromatic mark H3K9me3 and a decrease of acetylated H3 and H4 (Lee et al., 2004; Meshorer et al., 2006). ES cells have also been shown to have a characteristic replication timing profile (as discussed above in section 1.1.5.) in which many non-expressed tissue-specific regulator genes replicate early in S-phase (Azuara et al., 2006; Hiratani et al., 2008; Perry et al., 2004). A series of studies have shown that tissue-specific genes that were thought to be in a repressed state in pluripotent cells might in fact be poised for transcription in ES cells. Interestingly, the promoters of many

inactive genes are marked in ES cells by both repressive (H3K27me3) and active (H3K4me2 or H3K4me3) modifications - so-called bivalent domain (Azuara et al., 2006; Bernstein et al., 2006). It has been suggested that bivalent chromatin may allow tissue-specific regulator genes to remain silent in pluripotent cells, but primed for rapid activation in response to developmental cues (Giadrossi et al., 2007; Spivakov and Fisher, 2007). Interestingly, this gene priming for transcription has been shown to correlate at the molecular level with the presence of poised RNA polymerase II complexes at bivalent promoter genes in ES cells (Stock et al., 2007). Although with some exceptions, in ES-derived neural and haematopoietic progenitors, bivalent chromatin profiles at many developmental genes were found to be resolved as lineage-appropriate genes became activated and depleted for H3K27me3, while lineage-inappropriate genes lost histone acetylation and H3K4me3 (Bernstein et al., 2006; Mikkelsen et al., 2007). In addition, ChIP studies in both mouse and human ES cells revealed that members of the PRC2 complex (Suz12 and Eed) bind to the promoters of many developmental genes that are not actively transcribed in ES cells, but are important for the regulation of lineage specification of extraembryonic lineages or in later stages of development (Boyer et al., 2006; Jorgensen et al., 2006; Lee et al., 2006). Remarkably, Suz12 occupied many of non-expressed genes targeted by transcription factors Oct4, Sox2 and Nanog, suggesting that these regulators may specifically recruit PcG proteins to their repressed genes (Lee et al., 2006). In addition, many developmental PcG target genes become derepressed upon genetic disruption of the PRC2 components *Eed* or *Suz12* (Azuara et al., 2006; Boyer et al., 2006; Lee et al., 2006) or the PRC1 component *Ring1B* (Jorgensen et al., 2006; Stock et al., 2007). Altogether these findings demonstrated a pivotal role for PcG-mediated silencing in stably maintaining an ES cell pluripotent state in culture.

1.2.6. Reprogramming to pluripotency

Epigenetic mechanisms ensure that once committed, the differentiated state of a cell is normally stable and can be inherited throughout cell divisions. However, under certain conditions cell fate can be modified or reversed. This exceptional change in cellular identity leading to the establishment of a novel gene expression program is designated reprogramming (Jaenisch and Young, 2008).

Reprogramming is a phenomenon of particular interest by giving the opportunity to revert developmental programs, including the reprogramming of somatic cells to an embryonic state. This provides an exciting approach to address the functional

significance of epigenetic mechanisms implicated in cellular differentiation and establishment of pluripotency. In addition, reprogramming somatic cells to become pluripotent is an important goal in cell replacement therapy since it is an opportunity to generate patient-specific ES-derived cells as grafts (Hochedlinger and Plath, 2009; Jaenisch and Young, 2008). The increasing knowledge of how the undifferentiated state is maintained in mouse and human ES cells provided valuable clues for the understanding how re-orchestration of pluripotency occurs during reprogramming (Jaenisch and Young, 2008). Different experimental reprogramming strategies have been created including somatic cell nuclear transfer (SCNT) (Gurdon, 1962), cell fusion (Miller and Ruddle, 1976) and the forced expression of specific factors (Takahashi and Yamanaka, 2006).

The transfer of a nucleus from a somatic cell to an enucleated oocyte can reprogram the differentiated state to that of a zygote, potentially generating a new organism with the genetic information of the donor nucleus (a clone). The cloning of Dolly demonstrated that trans-acting material in mammalian oocytes is sufficient to change the epigenetic status of a differentiated nucleus to a totipotent state (Wilmut et al., 1997). This discovery completely changed the mindset of developmental biologists who previously thought this was impossible in mammals (McGrath and Solter, 1984). Immediately following the derivation of hES cells, it was suggested that SCNT might be used to create patient-specific stem cells. Yet applying SCNT to human material has proven challenging, and has only very recently been achieved in non-human primates (Byrne et al., 2007) with very low efficiency that would make it impractical for widespread human clinical use.

Experimental cell fusion between different cell types has also been used to study the plasticity of differentiated cells. Cell fusion creates cells harbouring two or more nuclei sharing the same cytoplasm, termed heterokaryon if the fused cells originate from different tissues or species, or homokaryons if the parental cells belong to the same cell-type. If heterokaryons are maintained in culture for several days, hybrids will form, resulting from fusion between nuclei and subsequent cell division (Kikyo and Wolffe, 2000). In most cell hybrids, the phenotype of the less differentiated cell is dominant over the phenotype of the more differentiated fusion partner. Consistent with this, fusion of mouse myoblasts with human B-lymphocytes leads to the activation of muscle regulatory factors in human B-lymphocyte nuclei accompanied by declining expression of several lymphocyte-specific genes (Terranova et al., 2006). Spontaneous and experimental cell fusion between ES and somatic cells has also shown that the phenotype of ES cells is dominant, imposing an ES-specific gene expression pattern in the nuclei of the less differentiated cell type (Tada et al., 2001;

Ying et al., 2002). The ability to reprogram somatic cells seems to be conserved in hES cells (Cowan et al., 2005; Yu et al., 2006) and analysis of mouse and human hybrid cell transcriptome revealed global patterns of gene expression reminiscent of ES cells (Ambrosi et al., 2007; Cowan et al., 2005). Other stem cell populations such as EG and EC cells have also been described to reprogram somatic cells *in vitro* (Miller and Ruddle, 1976; Tada et al., 1997) towards their identity. The reason why undifferentiated cell-types usually dominate over more differentiated fusion partners remains unknown. Interestingly, the principle of phenotypic exclusion seen during normal development also applies to hybrid cells. When two differentiated cell types are fused, hybrid cells have the potential to express either one or the other phenotype, but not both (Fougere and Weiss, 1978). Cell fusion experiments might not be the best source of cells for therapy, due to the low efficiency and requirement to remove the ES-cell nucleus from the heterokaryon in order to generate diploid customised cells for transplantation therapy (Pralong et al., 2005). However, it provides a powerful *in vitro* system to test the contribution of individual factors to the reprogramming of somatic cells towards pluripotency. Recently, using genetically engineered mES cells, it was demonstrated that the successful reprogramming of human lymphocytes is independent of Sox2. In contrast, there is a distinct requirement for Oct4 in the establishment but not the maintenance of the reprogrammed state (Pereira et al., 2008).

In 2006, Takahashi and Yamanaka reported the reprogramming of mouse fibroblasts to a pluripotent-like state by viral transfection of four transcription factors, Oct4, Sox2, kruppel-like factor 4 (Klf4) and c-Myc (Takahashi and Yamanaka, 2006). These factors had been selected as being involved in maintaining pluripotency and reprogrammed cells designated induced pluripotent stem (iPS) cells (Figure 1.4). Since this ground-breaking report other combinations of factors that have been successfully used for the derivation of iPS cells include the oestrogen-related receptor- β (Esrrb) in conjunction with Oct4 and Nanog in mouse fibroblasts (Feng et al., 2009a), and Oct4, Sox2, Nanog and Lin-28 in human fibroblasts (Yu et al., 2007). Since Takahashi and Yamanaka initial report in 2006 the reprogramming field is in constant revise, iPS cell lines have been generated from human (Takahashi et al., 2007; Yu et al., 2007), monkey (Liu et al., 2008) and rat (Li et al., 2009b; Liao et al., 2009) somatic sources, as well as from multiple cell types including disease-specific iPS cells for research studies (Park et al., 2008b), using various vehicles such as doxycycline-inducible viral vectors, adenoviral and plasmid vectors, and excisable genetic inserts (transposon-based and Cre-based), to deliver the exogenous genes (Colman and Dreesen, 2009). Furthermore, recombinant transcription factor proteins (Zhou *et al*, 2009), and small

chemicals that substitute for specific transcription factors and/or enhance the efficiency of the reprogramming process have been identified (Feng et al., 2009b).

Mouse and human iPS cells are remarkably similar to mouse and human ES cells, respectively. Although the initial mouse iPS cells did not contribute to the germline in chimeras, subsequent modification of the procedure to select iPS cells based on the reactivation of *Oct4* or *Nanog* promoter resulted in generation of postnatal chimeras and contribution to the germ line (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). More recently, it was shown that iPS cells were able to generate adult mice through tetraploid complementation, the most stringent test for developmental potency (Boland et al., 2009). In addition, iPS cells showed epigenetic features of ES cells, including the reactivation of somatically silenced X-chromosome, DNA demethylation of pluripotency-associated genes and acquisition of chromatin structure that closely resemble ES cells (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). Human iPS cells are morphologically similar to hES cells, express typical cell surface antigens and genes, differentiate into multiple lineages *in vitro*, and form teratomas containing differentiated derivatives of all three primary germ layers when injected into immunocompromised mice. Interestingly, generation of human iPS cells was dependent on FGF2 (Park et al., 2008c) rather than LIF (required to generate mouse iPS). Derivation of human iPS cells in mES cell media with LIF was fruitless (Takahashi et al., 2007).

These pioneering studies have illustrated the importance of several factors in reprogramming somatic cells to pluripotency. However, iPS cells are generated at very low efficiency and different types of reprogrammed cells (full and partially) are usually obtained (Chan et al., 2009). A better understanding of the reprogramming mechanisms and the true lineage identity of pluripotent mouse and human ES cells as well as the development of new and more efficient defined cell culture conditions should improve the efficiency of derivation and maintenance of iPS cells with full potential and clinically safe.

1.3. Aims of this study

A growing panel of mouse and human pluripotent stem cell lines has been derived from the early embryo. Their relationship to each other and whether they correspond to different pluripotent states with distinct developmental potencies and affiliations *in vivo* remains unclear, however. The overall aim of this project is to define the developmental status of embryo-derived pluripotent cell lines using a chromatin-based approach. Here I will take advantage of replication-timing assay as a tractable approach to investigate how chromatin shutdown contributes to lineage restriction and cell fate identity through development.

Initially I will focus in the comparison of the replication timing status of a large panel of key developmental genes in several hES cell lines alongside mES and mEpiS cells to determine whether distinct pluripotent states can be delineated at this level. Further I will explore how hES cell cultures can be modulated towards a self-renewing (ICM-like) *versus* primed for cell differentiation (epiblast-like) state under different growth conditions. Finally, I will investigate the molecular and epigenetic mechanisms that underlie hES cell self-renewal and differentiation, and evaluate the influence of TGF- β /Activin signalling pathway on favouring a self-renewing and homogeneous state within hES cell cultures. This is with the ultimate aim of defining optimal conditions to stably maintain hES cells and fully competent human iPS cells generated by reprogramming.

This study might contribute in the advance of our understanding of early development but is also important for developing stem cell-based therapies, as responses to differentiation protocols are likely to vary depending on the initial position of cells in a developmental hierarchy.

CHAPTER 2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Primary antibodies

SSEA1	Allophycocyanin (APC) –conjugated mouse monoclonal anti-SSEA1 (FAB2155A; R&D Systems Inc. Minneapolis, MN). Used for immunofluorescence (IF) and fluorescence activated cell sorting (FACS) at 1:200 dilution.
SSEA4	Mouse monoclonal anti-SSEA4 (MC-813-70; Developmental Hybridoma Studies Bank, Iowa City, IA). Used for IF and FACS at 1:3 dilution.
TRA-1-60	Mouse monoclonal anti-TRA-1-60 (MAB4360; Chemicon/Milipore, Temecula, CA). Used for IF and FACS at 1:12 dilution.
TRA-1-81	Mouse monoclonal anti-TRA-1-81 (MAB4381; Chemicon). Used for FACS at 1:20 dilution.
Nanog	Rabbit polyclonal anti-human Nanog (ab21624; Abcam Ltd., Cambridge, UK). Used for IF at 1:500 and western blot (WB) at 1:1000 dilution. Rabbit polyclonal anti-mouse Nanog (REC-RCAB0002P-F; Cosmo Bio Co., Ltd., Japan) used for IF at 1:500, FACS at 1 µg/1x10 ⁶ cells and WB at 1:1000 dilution.
Oct4	Goat polyclonal anti-Oct4 (N19 sc-8628; Santa Cruz Biotechnology Inc., Santa Cruz, CA). Used for IF at 1:100, FACS at 1 µg/1x10 ⁶ cells and WB at 1:1000 dilution.
Sox2	Rabbit polyclonal anti-Sox2 (ab15830; Abcam). Used for IF at 1:1000 and WB at 1:1500 dilution.
Nestin	Mouse monoclonal anti-Nestin (MAB5326; Chemicon/Milipore). Used for IF at 1:500 dilution.
Pax6	Rabbit polyclonal anti-Pax6 (AB5409; Chemicon/Milipore). Used for IF at 1:1000 dilution.
Musashi	Rabbit polyclonal anti-Musashi (AB5977; Chemicon/Milipore). Used for IF at 1:200 dilution.
H3	Rabbit polyclonal anti-histone 3 carboxy terminal (ab1791; Abcam). Used for WB at 1:2 500 and ChIP at 5 µg/IP dilution.
H3Ac	Rabbit polyclonal anti-histone 3 acetylation (17-615; Upstate Biotechnology /Milipore, Lake Placid, NY). Used for WB at 1:10 000 and ChIP at 5 µg/IP dilution.
H3K9Ac	Rabbit polyclonal anti-histone 3 lysine 9 acetylation (07-352; Upstate). Used for WB at 1:2 500 and ChIP at 5 µg/IP dilution.
H3K14Ac	Rabbit polyclonal anti-histone 3 lysine 14 acetylation (06-911; Upstate). Used for ChIP at 5 µg/IP dilution.
H3K18Ac	Rabbit polyclonal anti-histone 3 lysine 18 acetylation (07-354; Upstate). Used for WB at 1:10 000 and ChIP at 5 µg/IP dilution.
H3K56Ac	Rabbit monoclonal anti-histone 3 lysine 56 acetylation (2134-1; Epitomics, Inc., Burlingame, CA). Used for WB at 1:750 and ChIP at 10 µg/IP dilution.

H3K27me3	Rabbit polyclonal anti-histone 3 lysine 27 trimethylation (07-449; Upstate). Used for WB at 1:2500 dilution.
H3K4me2	Rabbit polyclonal anti-histone 3 lysine 4 dimethylation (07-030; Upstate). Used for WB at 1:5000 dilution.
H4	Rabbit monoclonal anti-histone 4 pan (05-858; Upstate). Used for ChIP at 5 µg/IP dilution.
H4Ac	Rabbit polyclonal anti-histone 4 acetylation (06-866; Upstate). Used for ChIP at 5 µg/IP dilution.
BrdU	Mouse monoclonal anti-BrdU antibody (347580; Becton Dickinson Biosciences, San Jose, CA). Used for replication timing at 80 µl/reaction dilution.
Methyl-C	Mouse monoclonal anti-5-methylcytidine (33D3; Eurogentec, Belgium). Used for methylated DNA immunoprecipitation (MeDIP) at 10 µl/reaction dilution.
IgG	Rabbit anti-IgG (whole molecule) fraction of antiserum antibody (M-7023; Sigma Aldrich Company Ltd., Gillingham, UK). Used for ChIP at 5 µg/IP and replication timing 35 µg/reaction dilution.
β-actin	Rabbit polyclonal anti-β-actin antibody (ab8227; Abcam). Used for WB at 1:5000 dilution.
P-Smad2/3	Rabbit monoclonal anti-Phospho Smad2/3 (serines 465/467) (3108; Cell Signalling Technology, Inc., Boston, MA). Used for WB at 1:1000 and ChIP at 10 µl/IP dilution.
Smad2/3	Rabbit polyclonal anti-Smad2/3 (3102; Cell Signalling Technology). Used for WB at 1:1000 and ChIP at 10 µl/IP dilution.
p300/CBP	Rabbit polyclonal anti-P300/CBP (C-20 sc-585; Santa Cruz). Used for WB at 1:1000 and ChIP at 2 µg/IP dilution.

2.1.2. Secondary antibodies

Alexa 488	Goat anti-mouse IgG (A11029; Molecular Probes Invitrogen, Ltd., Paisley, UK). Used for IF at 1:500 and FACS at 1:1000 dilution.
Alexa 488	Goat anti-rabbit IgG (A11034; Molecular Probes). Used for IF at 1:500 and FACS at 1:1000 dilution.
Alexa 568	Goat anti-mouse IgG (A11031; Molecular Probes). Used for IF at 1:500 dilution.
Alexa 568	Goat anti-rabbit IgG (A11036; Molecular Probes). Used for IF at 1:500 dilution.
FITC	Goat anti-mouse IgM (F9259; Sigma). Used for IF at 1:200 and FACS at 1:500 dilution.
FITC	Goat anti-rat IgM (F6258; Sigma). Used for IF at 1:100 dilution and FACS at 1:500 dilution.
HRP-conjugated	Goat anti-mouse IgG HRP (sc-2005, Santa Cruz). Used for WB at 1:2000 dilution.
HRP-conjugated	Goat anti-rabbit IgG HRP (sc-2004, Santa Cruz). Used for WB at 1:5000 dilution.

2.1.3. Cell lines

H1 (WA-01)	Male human ES cell line (Thomson et al., 1998); grown in collaboration with Dr Wei Cui (IRDB, Imperial College London, UK).
H1-BH1	Male human ES cell line clonally derived from H1 cells. Cell number passages with ≥ 100 are karyotypically abnormal. Cells were grown and kindly provided by Dr Peter Andrews's laboratory (Centre for Stem Cell Biology, University of Sheffield, UK).
H7 (WA-07)	Female human ES cell line (Thomson et al., 1998); grown in collaboration with Dr Wei Cui (IRDB, Imperial College London, UK).
H9 (WA-09)	Female human ES cell line (Thomson et al., 1998); grown in collaboration with Dr Wei Cui (IRDB, Imperial College London, UK).
iPS 40	Human iPS cell line (Vallier et al., 2009b) derived from neonatal foreskin fibroblasts by retroviral infection and expression of four reprogramming factors (Oct4, Sox2, c-Myc and Klf4). These cells were obtained from Dr Ludovic Vallier (Laboratory for regenerative medicine, University of Cambridge, UK).
CD34+	Human CD34+ stem cells isolated from peripheral blood of 3 healthy volunteers following informed consent. These cells were kindly provided by Dr Catherine Flores (Department of Haematology, Imperial College London, UK).
NP7	Human neural progenitor cells derived from H7 ES cells (Gerrard et al., 2005).
ReNcell VM	Human immortalized neural stem cell line derived from a 10-week fetal brain (SCC008; Chemicon/Milipore).
hB	Human Epstein-Barr Virus transformed adult B cell clones (Eden et al., 2002) obtained from Dr Filipe Pereira (MRC Clinical Sciences Centre, London, UK).
Oct4-GFP B	Abelson transformed mouse B cell line derived from the transgenic mice GOF18 Δ PE (Palmieri et al., 1994); C5BL/6 genetic background. This cell line was kindly provided by Dr Filipe Pereira (MRC Clinical Sciences Centre, London, UK).
E14tg2A	Hprt deficient mouse ES cell line derived from Lesch-Nyhan embryos (Hooper et al., 1987); feeder-independent; 129 genetic background.
B6	Mouse ES cell line (Kontgen et al., 1993); feeder-dependent; C57BL/6 genetic background; a gift from Dr Colin Steward (National Cancer Institute, Frederick, Maryland, USA).
Dc7	Mouse ES cell line newly derived; feeder-independent; 129/ <i>castaneus</i> mixed genetic background; a gift from Dr Tatyana Nesterova (MRC Clinical Sciences Centre, London, UK).
Oct4GiP	Mouse ES cell line derived from embryos carrying an Oct4GiP (eGFPiresPuro) transgene (Ying et al., 2002) feeder-independent; MF1/129 genetic background; a gift from Dr Jennifer Nichols (Wellcome Trust Centre for Stem Cell Research, University of Cambridge, UK).
EpiS 129	Mouse EpiS cell line derived from the dissection of the late-epiblast layer from pre-gastrula E6.5 embryos (Vallier et al., 2009a); 129 genetic background; a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).
EpiS B6/CBA	Mouse EpiS cell line derived from the dissection of the late-epiblast layer from pre-gastrula E5.75 embryos (Brons et al., 2007); B6/CBA F1 genetic background; a gift from

	Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).
EpiS NOD	Mouse EpiS cell line derived from the dissection of the late-epiblast layer from pre-gastrula E6.5 embryos (Brons et al., 2007); NOD genetic background; a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).
EpiS Oct4GiP	Mouse EpiS cell line derived from E5.75 embryos carrying an Oct4GiP (eGFPiresPuro) transgene (Ying et al., 2002); feeder-independent; MF1/129 genetic background; a gift from Dr Jennifer Nichols (Wellcome Trust Centre for Stem Cell Research, University of Cambridge, UK).
MEFs	Mouse embryonic fibroblasts isolated from wild type CF-1 embryos (E13.5).

2.2. Methods

2.2.1. Cell culture

All tissue culture reagents used were from Invitrogen (Invitrogen Ltd, Paisley, UK), unless otherwise stated. H1, H7 and H9 hES cell lines were cultured under feeder-free conditions on matrigel-coated (Becton Dickinson) 6 well plates as previously described (Xu et al., 2001). According to these conditions, cells were maintained in an undifferentiated state in mitotically inactivated primary MEF cells-conditioned medium (CM) [KnockoutTM Dulbecco's Modified Eagle Medium (DMEM) medium supplemented with 20% (v/v) KnockoutTM Serum Replacement (KSR), 1 mM L-glutamine, non-essential amino acids (NEAA), 0.1 mM β -mercaptoethanol and antibiotics (50 U/ml Penicillin/Streptomycin)] supplemented with 8 ng/ml basic fibroblast growth factor (FGF2) (PeproTech EC Ltd., London, UK). CM was prepared by incubating hES cells medium on flasks (0.5 ml/cm²) with 40 Gy irradiated MEFs at 56 000 cells/cm² during 24 hours. The CM was filtered (0.2 μ m) and stored at -20°C. MEFs were cultured in high-glucose DMEM supplemented with 10% foetal bovine serum (FBS; Globepharm, Esher, UK), 2 mM L-glutamine and antibiotics, grown until 80-85% confluence, dissociated with trypsin-EDTA before irradiation and plated onto 0.5% gelatine-coated surfaces (Sigma).

For feeder and Serum Replacer free culture, H1 and H9 hES cell lines were transferred from CM conditions to a chemically defined medium (Johansson and Wiles, 1995; Wiles and Johansson, 1999) supplemented with 10 ng/ml activin (R&D Systems) and 12 ng/ml FGF2 (PeproTech) as previously described (Vallier et al., 2005) and here designated as CDM^{activin}. The composition of CDM was 50% Iscove's Modified Dulbecco's Medium (IMDM) plus 50% F12 NUT-MIX, supplemented with 7 μ g/ml of insulin (Roche Ltd., West Sussex, UK), 15 μ g/ml of transferrin (Roche), 450 μ M of monothioglycerol (Sigma), 1% chemically defined lipid concentrate100x, 5 mg/ml of batch tested bovine serum albumin (BSA) fraction V (Sigma) and antibiotics. To allow hES cells adhesion in CDM^{activin}, plates were pre-coated with adhesion media (KnockoutTM DMEM medium supplemented with 10% batch tested FBS and 2 mM L-glutamine), for 24 hours at 37°C and then washed twice in phosphate buffered saline (PBS) to eliminate any serum. A range of concentrations (10, 25 and 50 μ M) of curcumin (a gift from Dr Matthew J Fuchter, Department of Chemistry, Imperial College London, UK) solution in dimethyl sulfoxide (DMSO) (Sigma) was added to CDM^{activin} media of H1 hES cells when indicated. Medium was daily changed and cells routinely

passed at a 1:3 split ratio by treatment with 200 U/ml collagenase type IV in Knockout™ DMEM followed by mechanical dissociation with 5 ml pipette in both CM and CDM^{activin} conditions. Nomenclature for culture conditions is according to the following convention: cell line, number of passages in CM culture, plus additional passages in CDM^{activin} conditions, plus number of passages of cells returned to CM here labelled as re-swapped (RS) cells, if applied. Thus, H1p40(CDM6;RS5) would be H1 grown for 51 passages overall, 40 initial passages were in CM, passages 41-46 were in CDM^{activin} and passages 47-51 were again in CM. Since hES cells are susceptible to genetic anomalies on prolonged passages (Draper et al., 2004), unless otherwise stated only hES cells earlier passages (p30-p80) were used for these experiments.

Multiple passages of H1-BH1 consisting of early and karyotypically abnormal late adapted cultures separated by at least 100 passages (early-passage hES duplicates ranged from p11-p26, and late-passage hES duplicates ranged from p145-p162) were cultured as previously described (Baker et al., 2007). Briefly cells were maintained on 600 cells/cm² MF-1 mitomycin C (Sigma) inactivated MEFs, in Knockout™ DMEM medium supplemented with 10% KSR and 4 ng/ml of FGF2. Cells were harvested for passaging by scraping with glass beads after brief incubation with a solution of 1% collagenase type IV in DMEM/F12. Early passage cells were typically passaged using 1:2 to 1:3 split ratios, every 5-6 days; later passage, adapted cells grew more robustly and were split at 1:5 to 1:10 split ratios.

Human iPS 40 cells (Vallier et al., 2009a) were grown under CM or CDM^{activin} conditions as described for hES cells.

Human CD34⁺ hematopoietic stem cells were isolated by Dr Catherine Flores from blood samples of three healthy volunteers after informed consent and under the local research ethics committee approval. Briefly, peripheral blood mononuclear cells (PBMCs) obtained by leukapheresis processed samples were isolated with the help of Lymphoprep (Axix-Shield; Oslo, Norway) gradient. CD34⁺ cells were isolated from the pool of PBMCs using the MiniMacs CD34⁺ isolation kit (MACS Miltenyi Biotec; Camberley, UK) as detailed in the manufacturer protocol. Isolated CD34⁺ cells were plated on 35 mm² Petri dishes with α -minimal essential medium (α -MEM) supplemented with 30 %FBS and cytokines (R&D Systems) [20 ng/ml stem cell factor, 1 ng/ml granulocyte-macrophage colony stimulating factor, 5 ng/ml interleukin (IL)-3 and 100 ng/ml granulocyte colony-stimulating factor].

EBV-transformed hB clones were maintained in RPMI-1640 supplemented with 10% FBS, 2 mM L-glutamine and antibiotics. The Oct4-GFP B-cell line was grown in

RPMI-1640 supplemented with 20% FBS, 1x NEAA, 2 mM L-glutamine, 50 μ M β -mercaptoethanol, antibiotics and IL-7 (5 ng/ml; R&D systems).

Human ReNcell VM cells were cultured in N2B27 medium (1:1 mix of DMEM/F12 supplemented with N2 and Neurobasal medium supplemented with B27) with 20 ng/ml FGF2 and 20 ng/ml epidermal growth factor (EGF) (R&D Systems) onto poly-L-lysine/laminin (Sigma) coated culture dishes. ReNcell VM cells were plated at 1×10^5 cells/cm² in 6 well plates and expanded by disassociation into single cells with TrypLE express.

Mouse ES cells were grown and maintained in an undifferentiated state either on 0.1% gelatin-coated surfaces (E14tg2A and Oct4GiP) or on irradiated MEF feeder layers (B6 and Dc7), and dissociated with trypsin-EDTA. With the exception of the newly derived Dc7 cell line, all mES cells were grown in Minimal Eagle medium (GMEM-BHK 21) supplemented with 15% batch tested FBS (Sigma), 1x NEAA, 1 mM sodium pyruvate, 0.075% sodium bicarbonate, 0.1 mM β -mercaptoethanol, 2 mM L-glutamine, antibiotics, and 1000 U/ml of LIF (ESGRO-LIF, Chemicon/Millipore). Dc7 cells were cultured and analyzed in derivation medium with DMEM supplemented with 5% batch tested FBS (Sigma), 15% KSR, 1x NEAA, 50 μ M β -mercaptoethanol, 2 mM L-glutamine and 2000 U/ml LIF. Differentiated cells could be eliminated from Oct4GiP cultures by puromycin (1 μ g/ml) selection for expression of the Oct4GiP transgene.

Mouse EpiS cell lines (129, B6/CBA, NOD and Oct4GiP) were cultured as previously described (Brons et al., 2007) in CDM^{activin} media supplemented with 20 ng/ml activin and 12 ng/ml FGF2 on pre-coated adhesion medium plates in the same way that hES cells in CDM^{activin} conditions. Cells were propagated as small clumps by scraping with 5 ml pipette after incubation with 1% collagenase type IV in KnockoutTM DMEM using 1:5 to 1:15 split ratios depending on proliferation ratio. Differentiated cells could be eliminated from Oct4GiP mEpiS cultures by puromycin (1 μ g/ml) selection for expression of the Oct4GiP transgene. Unless otherwise stated all cells were maintained at 37°C and 5% CO₂.

2.2.2. Human ES cells differentiation

2.2.2.1. Embryoid body formation

Human ES cells differentiation was achieved by promoting embryoid body (EB) outgrowths from hES cells CM and CDM^{activin} conditions, as previously described (Osafune et al., 2008). Briefly, 2 wells from a 6 well plate of hES cells at 80-90%

confluence were rinsed with PBS and treated with 1 mg/ml collagenase type IV in Knockout™ DMEM for 10 min at 37 °C. The collagenase was rinsed away with PBS and replaced with undifferentiation CM or CDM^{activin} medium (as described above). Cells were then scraped off with a cell scraper, dissociated by pipetting, and distributed into ultra-low attachment 10 cm plates (Corning Incorporated, NY, USA) containing Knockout™ DMEM supplemented with 20% KSR, 0.1 mM NEAA, 2 mM L-glutamine, 500 U/ml Penicillin/Streptomycin and 0.55 mM β-mercaptoethanol. On day 7, EBs were transferred back to 0.1% gelatinized-tissue culture plates, which allow the cells to adhere and further differentiate into different cell lineages. Media was changed every 3 days and EBs were collected for real-time quantitative PCR analysis at culture day 3, 6, 9 and 30. Primer pairs used for this analysis are listed in Appendix I, Table A1.

2.2.2.1. Neural progenitor cell differentiation

Neural differentiation of hES cells was achieved using two different methodologies as previously described (Gerrard et al., 2005) and (Chambers et al., 2009) with slight modifications.

Human neural progenitor (NP) cells derived from undifferentiated H7 cells from CM conditions were derived according to Gerrard et al. Briefly, confluent hES cells were split with 0.5 mM EDTA/PBS in 1:5 ratios into culture dishes coated with poly-L-lysine/laminin and cultured in N2B27 medium (as described above), supplemented with 100 ng/ml mouse recombinant noggin (R&D Systems). Cells were consistently split at 1:3 ratios by collagenase into small clumps, similar to hES cells CM culture. After 24 days in culture cells were disassociated into single cells by TrypLE express, and noggin was withdrawn from N2B27 medium, while 20 ng/ml FGF2 and 20 ng/ml EGF were added after an extra 8 days in culture. NP cells were plated at 1×10^5 cells/cm² in 6 well culture plates. Media was changed every 2 days. The neural progenitor phenotype was confirmed by immunostaining and semi-quantitative PCR (as described in the Results section). Primer pairs used for this analysis are listed in Appendix I, Table A2.

Chambers et al., neural differentiation protocol was applied to H1 hES cells grown under CM and CDM^{activin} conditions. Briefly, undifferentiated hES cells expanded in their regular cell culture conditions for 2 days, were incubated with N2B27 medium supplemented with 10 μM SB4341542 (TOCRIS/Sigma), 500 ng/ml noggin and 10 ng/ml FGF2. Media was daily changed, and cells collected for real-time quantitative

PCR analysis at culture day 2, 4 and 6. Primer pairs used for this analysis are listed in Appendix I, Table A1.

2.2.3. Replication Timing analysis

2.2.3.1. BrdU labelling and cell cycle fractionation by flow cytometry

All reagents used in this analysis were from Sigma unless otherwise stated. The replication timing of individual loci was assessed using a previously described method (Azuara, 2006). Exponentially growing non-synchronized cells were incubated with 50 μ M 5-bromo-2-deoxyuridine (BrdU) for 30 min (mES and mEpiS cells) or 90 min (hES, iPS 40, NPs and CD34+ cells) at 37°C, washed in cold PBS (Invitrogen) without calcium and magnesium, fixed in 70% ethanol and stored at 4°C in the dark for at least 24 hours. Ethanol was removed by washing the cells twice with PBS and collected by centrifugation at 1700 rpm for 10 min, 4°C. BrdU-labelled cells were resuspended in staining buffer (40 mM Tris pH 7.4, 0.8% NaCl, 21 mM MgCl₂ and 0.05% NP-40) with 50 μ g/ml propidium iodide (PI) and 1 mg/ml RNase A, carefully disaggregated with the help of a 40 μ m cell strainer (Becton Dickinson) and allowed to stain on ice, in the dark for 20 min prior to separation on a FACS Vantage flow cytometer, Diva. Equal numbers of cells (ranging from 20 000 to 50 000) were collected for each of six fractions (G1, S1, S2, S3, S4 and G2) into Lysis buffer I [1 M NaCl, 10 mM EDTA, 50 mM Tris-HCL pH 8.0, 0.5% SDS, 0.2 mg/ml proteinase K (PK), 0.25 mg/ml salmon sperm DNA (Invitrogen)] and stored at -20°C following incubation for 2 hours at 50°C.

2.2.3.2. Isolation of newly-replicated DNA by immunoprecipitation

Equal amounts of *Drosophila melanogaster* S2 BrdU-labelled DNA were added to each sorted cell cycle fraction (7.5 ng per 10⁴ sorted cells). DNA was then purified by phenol, chloroform/isoamylalcohol extraction and ethanol precipitation. Precipitated DNA was dissolved in 480 μ l of Tris-EDTA (TE) with 0.2 mg salmon sperm DNA (Invitrogen) at 37°C for 1 hour in the dark. Samples were then sonicated 45 sec on a Vibra-Cell™ VC130 (Sonics & Materials, Inc, Newtown, USA) (130W, 30Hz) using a microtip ultrasonic probe in order to generate fragments of an average size of 700 bp. After denaturation for 5 min at 95°C and cooling down on ice for 2 min, it was added 50 μ l of Adjusting buffer [10 mM sodium phosphate (pH 7), 0.14 M NaCl, and 0.05% Triton X-100] to each fraction and incubated with 80 μ l anti-BrdU antibody for 30 min at room

temperature with constant rotation. Secondary antibody (rabbit anti-mouse IgG) was added (35 µg) and incubated at room temperature with constant rotation for additional 30 min. DNA-protein complexes were collected by centrifugation at 15 300 rpm for 20 min, 4°C, washed once with Washing buffer (10 mM sodium phosphate buffer pH 7.0, 0.14 M NaCl, 0.05% Triton X-100) and resuspended in 200 µl Lysis buffer II (50 mM Tris-HCL pH 8, 10 mM EDTA, 0.5% SDS, 0.25 mg/ml PK) and incubated 2 hour, 55°C. BrdU-labelled DNA was then extracted by phenol/chloroform/isoamyl alcohol and ethanol precipitation before resuspension in Tris-EDTA (TE; 10 mM Tris-Hcl pH 8.0, 1 mM EDTA) at a concentration of 150/200 cell equivalents per microlitre. The abundance of newly replicated DNA in each cell-cycle fraction was determined by real-time quantitative PCR. Primer pairs used for this analysis are listed in Appendix I, Table A3.

2.2.4. Reverse transcription (RT) polymerase chain reaction (PCR)

2.2.4.1. RNA extraction and cDNA synthesis

RNA extraction was performed using RNeasy protect mini kit (Qiagen, Valencia, CA) and residual DNA eliminated with RNase-free DNase set (Qiagen) according to manufacturer's instructions. RNA was then reverse transcribed using the Superscript™ first-strand synthesis system. 1.5 µg of total RNA was diluted in RNase free water to a final volume of 11 µl and supplemented with 1 µl of 10 mM dNTP mix (Invitrogen) and 1 µl of oligo (dT)₁₂₋₁₈ (Invitrogen). The mixture was incubated at 65°C for 5 min and put on ice for 1 min, when 1 µl of 0.1 M DTT, 4 µl of 5X first strand buffer, 1 µl of RnaseOUT (Invitrogen) and 1 µl of 200 U/µl Superscript III were added. A reaction mixture without the enzyme was also set up as a control (designated "-RT"). The mixture was then incubated at 50°C for 1 hour and at 70°C for 15 min in order to stop the reaction. 180 µl of RNase-free water was then added to each cDNA sample. cDNAs of interest were then detected by semi-quantitative PCR or real-time PCR.

2.2.4.2. Semi-quantitative PCR analysis (RT-PCR)

Semi-quantitative PCR analysis was performed in 50 µl reaction volume with 1.25 U of HotStarTaq™ DNA polymerase (Qiagen) using 0.4 µM specific primers, with the following program: 95°C for 15 min, then cycles of 94°C for 15 sec, 60°C for 30

sec, 72°C for 30 sec. PCR products were separated by electrophoresis on 1.5% agarose (Invitrogen) gels and visualised by ethidium bromide staining. *β-actin* was used to normalize input. The primers and conditions used for this analysis can be found in Appendix I, Table A2. For each reaction of PCR amplification, conditions were maintained within the dynamic range.

2.2.4.3. Real-time quantitative PCR analysis (qPCR)

Real-time PCR analysis was performed in 30 µl reaction volume with 2X Sybr-Green PCR Mastermix (Qiagen), 300 nM primers and 2 µl of template. A reaction without DNA was included to control for the formation of primer dimers and each measurement was performed in duplicate. PCR reactions were carried out on a DNA Engine Opticon™ II using Opticon Monitor 3 software (MJ Research Inc., Waltham, MA), running the following program: an initial denaturing step at 95°C for 15 min, 40 cycles of denaturation at 94°C for 15 sec, annealing at 60°C for 30 sec, elongation at 72°C for 30 sec at which point the fluorescence was read at 72°C, 75°C, 78°C, 80°C and 83°C. The melting curve was determined from 70°C to 90°C, at 0.1°C intervals. For analysis of the qPCR data the relative abundance of sequences was calculated using the $\Delta C(T)$ method (Pfaffl et al., 2001). Gene expression data was normalized to the average of two housekeeping genes: *Gapdh* and *Hprt* for human data and for mouse *Actin* and *L19* as previously described (Vandesompele et al., 2002). Primer sequences can be found in Appendix I, Table A1.

2.2.4.4. Primer design and testing for real-time quantitative PCR

Replication timing, ChIP and methylated DNA immunoprecipitation (MeDIP) primers were designed with the Primer Express software (Applied Biosystems, Warrington, UK) while gene expression primers designed with PerlPrimer software (Marshall, 2004) using sequences from the mouse and human Ensemble Genome Browser (<http://www.ensembl.org/index.html>) database. The following design criteria were used: amplicon size of 100-200 bp, GC content of 40-60%, melting temperature 55-65°C. All primer pairs were tested for predicted products using *in silico* PCR tools such as PUNS (<http://okeylabimac.med.utoronto.ca/PUNS>) or the UCSC genome browser (<http://genome.ucsc.edu>). Primers yielding a single predicted PCR product of the correct size were ordered from Sigma Genosys.

The efficiency of amplification for each primer pair was determined by qPCR with sequential dilutions of cDNA or genomic DNA (see section 2.2.10.1.) and primers

yielding poor linear fits of the C(T) versus logarithm of concentration ($R^2 < 0.99$) or efficiencies lower than 1.8 or greater than 2.2 were discarded. When species specificity was required (i.e. analysis in inter-species heterokaryons), primers were designed and tested to specifically amplify the human or the mouse sequences but not the orthologs.

2.2.5. Genomic context analysis

The GC content, long interspersed nuclear elements (LINE) density and gene density genomic features of the replication timing analyzed genes was obtained from the University of California, Santa Cruz, genome browser website (<http://genome.ucsc.edu/>), based on the May 2004 human genome assembly. For GC content and LINE density calculations it was analyzed, for each gene of interest a 200 Kb surrounding DNA region (100 Kb upstream and downstream of the transcription start position). As previously described (Hiratani et al., 2004), regions with < 43% GCs are considered poor for this parameter and regions with > 6% LINE are highly dense. Gene density was defined as the number of transcriptional units in 1 Mb (500 Kb upstream and downstream of the middle position of the gene). Regions with ≤ 10 genes were considered poor.

2.2.6. Immunofluorescence analysis and alkaline phosphatase activity

Undifferentiated hES, mES and mEpiS cells were cultured on coated glass coverslips according to their growth conditions and NP cells on poly-L-lysine/laminin coated Thermanox® coverslips (Nalge-Nunc Inc., Rochester, NY) for at least 24 hours. Coverslips were removed, washed in PBS and fixed in 4% paraformaldehyde (PFA) in PBS for 15 min. Fixed samples were washed in PBS and, for intracellular staining, permeabilised with 0.4% Triton X-100 for 15 min at room temperature. Samples were incubated sequentially in blocking solution [2.5% BSA, 0.05% Tween20, 10% Normal Goat Serum (Sigma) in PBS] for 30 min and in primary antibody diluted in blocking solution (2 h at room temperature) in a humid chamber. Coverslips were washed in wash buffer (0.2% BSA, 0.05% Tween20 in PBS; 3x5 min) and incubated with secondary antibodies coupled with appropriated fluorophores diluted in blocking solution for 45 min. Finally, cells were washed twice in wash buffer (5 min), once in PBS (3 min) and mounted in Vectashield (Vector) with DAPI (1 $\mu\text{g/ml}$). For SSEA1 staining, cells were incubated with APC-coupled anti-SSEA1 antibody, washed twice in

wash buffer (5 min), once in PBS (3 min) and mounted in Vectashield (Vector) with DAPI (1 µg/ml). A list of primary and secondary antibodies used for IF as well as quantities and suppliers is shown in section 2.1.1. and 2.1.2.

Alkaline Phosphatase staining was carried out on cells grown for 2 days under their standard culture conditions, using a kit from Sigma (86R-1KT) according to the instructions.

Samples were visualised using a Nikon Eclipse TE2000-U microscope and images processed using QCapture Pro software and Adobe Photoshop CS2, or a TCS SP5 Leica laser-scanning confocal microscope using the Leica Confocal software and Adobe Photoshop CS2. Microscope settings and laser power were kept constant between the controls and samples.

2.2.7. Fluorescence activated cell sorting analysis

The analysis of cell surface antigens (SSEA4, SSEA3, SSEA1, Tra-1-60 and Tra-1-81) was performed as follows: cells were harvested by trypsinisation and 1.5×10^6 cells were washed twice in FACS buffer (PBS, 2% FBS), resuspended in primary antibody diluted in FACS buffer and incubated 30 min at 4°C. The cells were then washed twice and incubated for an additional 30 min with a secondary antibody also diluted in FACS buffer. Finally, cells were washed twice and then resuspended in 100 µl of FACS buffer for analysis. For SSEA1 staining, cells were incubated with APC-coupled anti-SSEA1 antibody, washed twice and resuspended directly in 100 µl of FACS buffer. For intracellular staining (Oct4 and Nanog) cells were harvested by trypsinisation, washed in FACS buffer and 1.5×10^6 cells fixed (10 min, 37°C) with 0.1% PFA in PBS. Cells were then washed and permeabilised (30 min, 4°C) with cold 90% methanol. After washing once, cells were incubated for 30 min at room temperature with primary antibody diluted in FACS buffer with 0.1% Triton X-100. Cells were again washed twice before incubating for 30 min with secondary antibody diluted in FACS buffer with 0.5% Triton X-100. The cells were finally washed twice and resuspended in 600 µl of FACS buffer before analysis on a FACScalibur flow cytometer (BD Biosciences) with CellQuest software. The profile of stained cells was compared to unstained cells and cells stained with the secondary antibody only. A list of primary and secondary antibodies used for FACS as well as quantities and suppliers is shown in section 2.1.1. and 2.1.2.

For cell fusion efficiency (see section 2.2.11.) assessment it was used the cell labelling solutions Vibrant™ 1,1'-dioctadecyl-3, 3', 3', 3' tetramethylindodicarbocyanine

(DiD) and 1,1'-dioctadecyl-3, 3', 3'-tetramethylindocarbocyanine perchlorate (DiI) (Molecular Probes; emission spectra: DiI λ =565nm, DiD λ =665nm). Human B-lymphocytes were labelled with DiI while mES and mEpiS cells with DiD using the FL2 and FL4 FACS detection channels, respectively.

2.2.8. Western blot analysis

Whole cell extracts were prepared by direct lysis of cells with cold Radio Immuno Precipitation Assay (RIPA) buffer [150 mM NaCl, 50 mM Tris pH 8.0, 0.5% Deoxycholic acid, 0.1% sodium dodecyl sulfate (SDS), 1% NP-40 plus protease inhibitors]. Cells were scrapped, incubated on ice for 20 min and centrifugated at 13 000 rpm for 20 min at 4°C and the supernatant collected. For western blot analysis of modified histones, histone proteins were isolated from whole cells by acid extraction. 10×10^7 cells were pelleted and resuspended in 1 ml PBS (4°C), centrifuged (500 g for 5 min) and the supernatant removed. Cell pellets were resuspended in 180 μ l of ice cold lysis buffer (100 mM HEPES pH 7.9, 1.5 mM $MgCl_2$ 10 mM KCl, 0.5 mM DTT and 1.5 mM PMSF), 20 μ l of 2 M HCl and incubated on ice for 30 min. Following acid lysis the solution was centrifuged 11 000 g for 10 min at 4°C, the supernatant of acid soluble proteins collected and sequentially dialyzed against 0.1 M acetic acid (twice for 1 hour) and water (1 hour, 3 hours and overnight, respectively). The protein solution was quantified and stored at -80°C.

10-20 μ g of protein sample was diluted 1:1 with Laemmli sample buffer (BioRad) with 5% β -mercaptoethanol and heated at 95°C for 5 min. Cell lysates were resolved on 8-14% SDS polyacrylamide gel electrophoresis (SDS-PAGE) gels with 11 μ l Novex Sharp protein ladder (Invitrogen). Resolved acrylamide gels were blotted to Polyvinylidene fluoride (PDVF) membranes (Chemicon/Milipore) using a wet-blot method at 200 mA for 1 hour in transfer buffer (48 mM Trizma base, 39 mM glycine, 0.037% SDS and 20% methanol). The membranes were then blocked with 5% skimmed milk (Sigma) in Tris-Buffered Saline with 0.5% Tween20 (TBS-T) for 1 hour at room temperature, followed by primary antibody incubation diluted in blocking buffer or 5% BSA in TBS-T, overnight at 4°C, with agitation. Membranes were washed, to remove any excess antibody, for 10 min three times in TBS-T under agitation before incubating with the appropriate HRP-conjugated secondary antibody in blocking buffer, for 1 hour at room temperature. Excess secondary antibody was then removed by washing the membrane for 10 min three times in TBS-T buffer under agitation. The blot was treated by enhanced chemiluminescence western blotting detection reagents

(Amersham Pharmacia Biotech, Little Chalfont, UK) for HRP-conjugated antibody detection before being exposed to X-ray films (CL-XPosure Film, Thermo Scientific, UK) and using the OPTIMAX IGP developer. A list of primary and secondary antibodies used for WB as well as quantities and suppliers is shown in sections 2.1.1. and 2.1.2.

2.2.9. Chromatin immunoprecipitation analysis

2.2.9.1. Histone modifications ChIP

All reagents used were from Sigma unless stated otherwise. Approximately 0.5×10^8 exponentially growing cells were harvested, washed twice in PBS and cross linked with 1% PFA for 10 min at room temperature. After quenching of PFA with glycine at a final concentration of 125 mM, cells were washed twice with ice-cold PBS and lysed in 1 ml of ice cold lysis buffer (50 mM Tris-HCL pH 8.1, 10 mM EDTA pH 8.0, 1% SDS) with protease inhibitor cocktail (Roche). Using a Bioruptor™ 200 (Diagenode, SA, Liège, Belgium), the chromatin was sonicated to an average size of 300 to 1000 bp as analysed on 1.5% agarose gels. Insoluble proteins were discarded after centrifugation of the lysate at 15 000 rpm for 15 min at 18°C and DNA concentration quantified using a NanoDrop® ND-1000. 100-150 µg of fragmented chromatin was diluted 1/10 in dilution buffer (1% Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCL pH 8.1) with protease inhibitors and incubated at 4°C for 2 hours, with 30 µl of blocked protein A-sepharose beads to pre-clear the solution of any non-specific binding. Beads were overnight blocked at 4°C in dilution buffer supplemented with 10 µl of BSA (New England Biolabs, 10 mg/ml) and 4 µl salmon sperm DNA (10 mg/ml) and then resuspended in dilution and lysis buffer in a 10:1 ratio. Pre-cleared chromatin was then immunoprecipitated (4°C, overnight) with primary antibodies for histone's modification, a negative anti-IgG control and an anti-H3 or -H4 total antibody for data normalization, on a spinning wheel. A list of antibodies used for ChIP as well as quantities and suppliers is shown in section 2.1.1. Immune-complexes were collected by adding 30 µl of blocked beads once again and left immunoprecipitating 2 hours on a spinning wheel, at 4°C. Unbound chromatin was removed by washing four times in 1 ml of cold wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris.Cl pH 8.1 and protease inhibitors) and once in final high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris.Cl pH 8.1 and protease inhibitors). After adding elution buffer (1% SDS, 0.1 M NaHCO₃) to elute the immunocomplexes from the beads, these were treated with 100 µg/ml RNase A and

500 µg/ml PK, 2 hours at 37°C and 6 hours of incubation at 65°C to reverse the cross-links. Finally, the DNA was sequentially extracted with phenol/chloroform/Isoamylalcohol and precipitated in 50% isopropanol containing 5 mM NaAc and 20 µg of glycogen carrier (Glycoblue, Ambion). After purification, DNA was resuspended in 100 µl of TE buffer. Histone's modification levels were normalized against total H3 or H4 detected, and the ratio of DNA bound to modified-H3 or -H4 to total H3 or H4 was denoted as "relative abundance". Quantification of precipitated DNA was performed using qPCR amplification and the primers listed on Appendix I, Table A3.

2.2.9.2. p300 and Smad2/3 ChIP

This protocol used has previously been described in detail by (Stock et al., 2007) with minor modifications. Briefly, cells were treated with 1% formaldehyde (37°C, 10 min) and quenched with 125 mM glycine. Cells were washed in ice-cold PBS and lysed at 4°C for 10 min in swelling buffer (25 mM HEPES pH 7.9, 1.5 mM MgCl₂, 10 mM KCl and 0.1% NP-40). Cells were scraped from plates, nuclei extracted by homogenization, resuspended in sonication buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate and 0.1% SDS) and sonicated to produce chromatin fragments with less than 1.6 kb in length. 500-600 µg of solubilised chromatin was pre-cleared with blocked Protein-A (rabbit antibodies) or Protein-G-sepharose beads (mouse antibodies) and incubated overnight at 4°C with primary antibody on a rotating wheel. Smad2/3 ChIP was performed with a mixture of P-Smad2/3 and total Smad2/3 antibody. A list of antibodies used for ChIP as well as quantities and suppliers is shown in section 2.1.1. Antibody-chromatin complexes were collected with blocked beads for 3 hours at 4°C before sequential washes: 1x sonication buffer, 1x sonication buffer containing 500 mM NaCl, 1x with 20 mM Tris pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40 and 0.5% Na-deoxycholate, and 2x TE buffer. All buffers were supplemented with 5 mM NaF, 2 mM Na₃VO₄, 1 mM PMSF, and protease inhibitor cocktail. Immunoprecipitated chromatin was eluted and DNA recovered by phenol/chloroform/Isoamylalcohol extraction and ethanol precipitation as previously described in 2.2.9.1. Final DNA concentrations were determined by PicoGreen fluorimetry (Molecular Probes, Invitrogen) and the amount of DNA precipitated by each antibody was normalized against the total input material. Quantification of precipitated DNA was performed using qPCR amplification and the primers listed on Appendix I, Table A3.

2.2.10. Methylated DNA immunoprecipitation analysis

2.2.10.1. Genomic DNA extraction

MeDIP was carried out in hES genomic DNA as previously described (Weber et al., 2005). Genomic DNA was isolated from cells by overnight incubation in 500 µl of Lysis buffer (200 mM NaCl, 100 mM Tris-HCl pH 7.4, 5 mM EDTA, 0.2% SDS, 250 µg PK) at 55°C and subsequent phenol/chlorophorm/isoamylalcohol and chlorophorm/isoamylalcohol extractions. DNA was then precipitated with 50% isopropanol, followed by 70% ethanol wash, and resuspension in TE with 20 µg/ml RNaseA. DNA concentration was determined using a NanoDrop® ND-1000.

2.2.10.2. Sonication and immunoprecipitation of methylated DNA

Genomic DNA was sonicated (VC 130 PB, Sonics & Materials, Inc) to produce random fragments ranging in size from 300 to 1000 bp. 5 µg of fragmented DNA was used for the MeDIP assay. After denaturation for 10 min at 95°C and cooling down on ice for 2 min, samples were immunoprecipitated for 2 hours at 4°C with 10 µl of Methyl-C antibody in a final volume of 500 µl of immunoprecipitation buffer (10 mM Na-Phosphate pH 7.0, 0.14 M NaCl, and 0.05% Triton X-100) with constant rotation. Chromatin-antibody complexes were then precipitated with 30 µl of Dynabeads® with M-280 sheep antibody against mouse IgG (DynaL Biotech, Oslo, Norway) for 2 hours at 4°C and washed three times with 700 µl of immunoprecipitation buffer. Beads were then treated with 70 µg of PK for 3 hours at 50°C and methylated DNA was recovered by phenol-chloroform extraction followed by ethanol precipitation at -20°C and resuspended in 60 µl of TE buffer. The abundance of methylated DNA was determined by qPCR amplification with 25 ng of input DNA and 2 µl of the immunoprecipitated methylated DNA. Primer pairs used in this analysis are listed in Appendix I, Table A3.

2.2.11. Cell fusion and experimental heterokaryons

2.2.11.1. Mouse ES and EpiS-Human B-lymphocyte heterokaryons

Heterokaryons were generated by fusing either mES (Oct4GiP) or mEpiS (Oct4GiP) cells and human B-lymphocytes using 50% polyethylene glycol, pH 7.4 (PEG 1500, Roche) as previously described (Pereira and Fisher, 2009; Pereira et al.,

2008). Mouse ES cells or EpiS cells were labelled with DiD and hB-lymphocytes with Dil cell labelling solutions (Molecular Probes), respectively. Cells were resuspended at 1×10^6 cells/ml in DMEM and labelled with 5 μ l/ml of dye at 37°C, 15 min. Mouse ES/EpiS and hB cells were then mixed at 1:1 ratio and washed twice in PBS. The supernatant was completely removed and 1 ml of PEG at 37°C was added to the pellet of cells over 60 sec and incubated at 37°C for 90 sec with constant stirring. Then, 4 ml of serum-free medium (DMEM) were carefully added over a period of 3 min, followed by 10 ml of DMEM and incubation at 37°C for 3 min. After centrifugation (1350 rpm, 5 min), the pellet was allowed to swell in mES/mEpiS cell culture media for 3 min. Cell mixtures were then resuspended and cultured under conditions promoting the maintenance of undifferentiated mES or mEpiS cells at 0.5×10^6 cells/cm². To eliminate unfused hB cells Ouabain (10^{-5} M) (g-strophanthin, Sigma) was added to the medium. Proliferating mES or mEpiS cells were eliminated by the addition of 10^{-5} M Ara-C (Cytosine-D arabino furanoside; Sigma) 6 hours after fusion and then removed after 16 hours. Fused cells were cultured under conditions promoting the maintenance of undifferentiated mouse stem cells and heterokaryons collected for qPCR analysis at day 0, 1, 2 and 3. Cell fusion efficiencies were determined by FACS analysis at day 0. Primer pairs used in this analysis are listed in Appendix I, Table A1.

2.2.11.2. Human ES CM and CDM^{activin}-Mouse B-lymphocyte heterokaryons

Undifferentiated H1 hES from CM or CDM^{activin} culture conditions were fused with mouse B-lymphocytes (Oct4-GFP B) as described above. Cells were mixed for fusion at 1:1 ratio and cultured under conditions promoting the maintenance of undifferentiated hES cells in CM or CDM^{activin} conditions at 0.5×10^6 cells/cm². Proliferating hES cells were eliminated by the addition of Ouabain (10^{-5} M) 6 hours after cell fusion. Heterokaryons were collected for qPCR analysis at day 0, 1, 2 and 3. Primer pairs used in this analysis are listed in Appendix I, Table A1.

2.2.12. Analysis of DNA content and BrdU incorporation following acid denaturation of DNA

All reagents used in this analysis were from Sigma unless otherwise stated. Exponentially growing non-synchronized H1 hES CM and CDM^{activin} cells were incubated with 30 μ g/ml BrdU for 60 min at 37°C. Cells were collected from plates by trypsinization, pooled with the floating cells in the medium and resuspended in room

temperature PBS at 1×10^6 cells/ml. With a Pasteur pipette 1 ml aliquots of the cell suspension were transferred into tubes with 10 ml of ice-cold 70% ethanol and fixed overnight at 4°C, in the dark. Cells were collected by centrifugation (1200 rpm, 5 min) and pellets incubated in 250 µl of 2 M HCl for 20 min at room temperature. After this step it was added 5 ml of phosphate/citric acid buffer, pH 7.4 (182 mM Na_2HPO_4 and 9 mM citric acid) to the suspension, cells were centrifugated and washed in another 5 ml of phosphate/citric acid buffer, pH 7.4. Cells were resuspended in 100 µl antibody dilution buffer (0.1% Triton X-100, 1 % BSA in PBS) with 0.5 µg mouse anti-BrdU antibody, gently vortexed and incubated 30 min, room temperature. Cells were then washed once with 5 ml antibody dilution buffer and resuspended in 100 µl antibody dilution buffer with 1 µg Alexa 488 goat anti-mouse antibody, gently vortexed and incubated 60 min at room temperature and 5 min on ice. After washing once, cells were finally stained in the dark with 2 ml PI staining solution (100 µg PI, 1 mg RNase A and 10 ml PBS) for 30 min prior to FACS analysis. Cells that were not incubated with BrdU and cells incubated only with secondary antibody were used as a negative control.

CHAPTER 3. CHARACTERIZATION OF THE CHROMATIN STATUS OF HUMAN ES CELLS ALONGSIDE MOUSE ES AND EPI S CELLS

3.1. Replication timing analysis of hES, mES and mEpiS cells using a candidate based approach

Profiling chromatin in a particular cell type has proven to be a valuable signature for cell identity and developmental stage. One approach has been to assay the timing of DNA replication across a panel of loci, as an indicator of chromatin accessibility. Remarkably, this replication timing profiling was capable of discriminating pluripotent mES cells from cells with a more restricted differentiation capacity (Azuara et al., 2006; Hiratani et al., 2008; Perry et al., 2004). Here, DNA replication timing will be used to examine and compare human and mouse stem cell populations derived from the early embryo, this includes well-established hES (H1, H7 and H9), mES (E14tg2A, B6 and Dc7) and mEpiS (129, B6/CBA and NOD) cell lines derived from different genetic backgrounds (see section 2.1.3. for cell line details). As illustrated in Figure 3.1a, all three pluripotent stem cell types analysed (hES, mES and mEpiS) share the expression of key pluripotency markers (Nanog and Oct4) and can be propagated and maintained in an undifferentiated state *in vitro* for an unlimited period of time under appropriate conditions. Despite the similarities, several differences between hES, mES and mEpiS cells have been reported, in particular in terms of cells morphology and cell-surface markers (Pera and Trounson, 2004). Undifferentiated mES cells typically grow in small, compact, domed colonies while in contrast, hES and mEpiS cell colonies are larger, flat and cells grow as a monolayer (Figure 3.1b). The expression of the cell surface marker SSEA1 is restricted to mES and mEpiS cells while AP activity can only be detected in mES and hES cells (Figure 3.1a and c). Importantly, the extrinsic signals that drive pluripotent mouse and human cells self-renewal and pluripotency have undergone significant deviations. The undifferentiated state of mES cells depends on LIF/BMP signalling, while hES and mEpiS cells rely on Activin/FGF (Yu and Thomson, 2008). Altogether, this raises a developmental paradox, where the true lineage affiliation of hES cells is questionable. Are the differences

between mES and hES cells related to species divergence or to differences in the stages of embryonic development from which they arise? In this study, I will further investigate this developmental paradox, by investigating the cell identity and chromatin status of mouse and human pluripotent cell lines, using DNA replication timing.

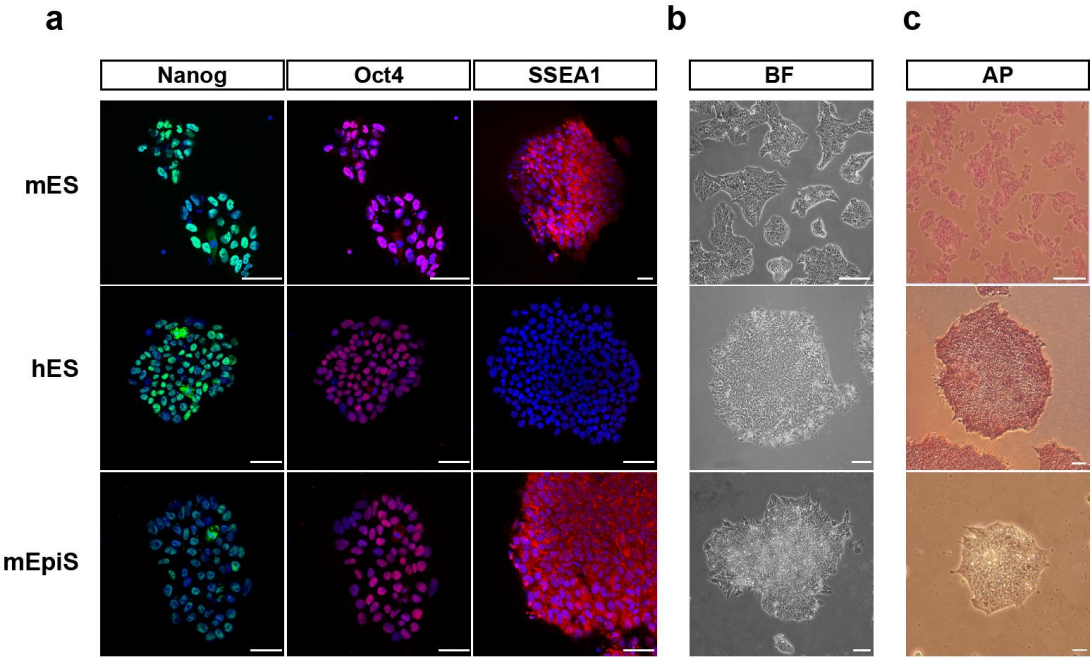


Figure 3.1. Characterization of the undifferentiated state of hES cells alongside mES and mEpiS cells. (a) Mouse ES (E14tg2A line, top panel), hES (H1 line, middle panel) and mEpiS (129 line, bottom panel) cells were analysed by immunofluorescence for the expression of Nanog (green), Oct4 (red) and SSEA1 (red). Nanog and Oct4 are highly expressed in mES, hES and mEpiS cells, while SSEA1 is not expressed in hES cells. Nuclei are shown by DAPI staining (blue). (b) Bright field (BF) pictures of representative undifferentiated mES, hES and mEpiS cell colonies. Mouse ES cells typically grow in small, compact, domed colonies, while hES and mEpiS cell colonies are larger and flat growing as a monolayer. (c) Alkaline phosphatase (AP) enzymatic activity (pink) is detected in both hES and mES cells and absent in mEpiS cells. Scale bars, 50 μ m.

For this analysis, it was selected a panel of 45 key regulatory genes mainly encoding human and mouse homologs, which are selectively expressed in embryonic (pluripotent) and extraembryonic (trophectoderm) tissues of the early embryo or are involved in specifying ectoderm, mesoderm, endoderm or germ-line lineages (Table 1). The POU transcription factor Oct4 (Nichols et al., 1998; Palmieri et al., 1994; Reubinoff et al., 2000; Rosner et al., 1990; Scholer et al., 1990) and the homeobox gene Nanog (Chambers et al., 2003; Clark et al., 2004; Mitsui et al., 2003) are examples of key pluripotency-associated genes in both human and mouse stem cell populations.

Table 1 - Candidate genes selected for the replication timing analysis.

	Loci	Chromosome location	
		Mouse	Human
Controls	<i>α-globin</i>	11 16.0 cM	16p13.3
	<i>Amylase</i>	3 50.0 cM	1p21
Inner Cell Mass	<i>Cripto</i>	9 60.0 cM	3p21.31
	<i>Dnmt3b</i>	2 A2-A3	20q11.2
	<i>Dppa4</i>	16 B5	3q13.13
	<i>Gbx2</i>	1 65.0 cM	2q37
	<i>Gdf3</i>	6 60.6 cM	12p13.1
	<i>Nanog</i>	6 F2	12p13.31
	<i>Oct4/Pou5f1</i>	17 19.23 cM	6p21.31
	<i>Rex1/Zfp42</i>	8 24.0 cM	4q35.2
	<i>Sox2</i>	3 15.0 cM	3q26.3-q27
	<i>Tert</i>	13 43.0 cM	5p15.33
	<i>Tle1</i>	4	9q21.32
Epiblast	<i>Fgf5</i>	5 55.0 cM	4q21
	<i>Prce/Espl1</i>	15 F3	12q
	<i>Psc1/Rbm27</i>	18 B3	5q32
Trophectoderm	<i>Cdx2</i>	5 82.0 cM	13q12.3
	<i>Cga</i>	4 10.5 cM	6q12-q21
	<i>Eomes</i>	9 67.0 cM	3p21.3-p21.2
	<i>Esxl1</i>	X 57.0 cM	Xq22.1
Germ cells	<i>Dazl</i>	17 25.6 cM	3p24.3
	<i>Sycp1</i>	3 A3	1p13-p12
	<i>Tekt1</i>	11 B4	17p13.2
Somatic lineages	<i>Bmp4</i>	14 15.0 cM	14q22-q23
	<i>Cxcr4</i>	1 67.4 cM	2q21
	<i>Ebf</i>	11 20.0 cM	5q34
	<i>Foxa2</i>	2 84.0 cM	20p11
	<i>Gata4</i>	14 28.0 cM	8p23.1-p22
	<i>Gata6</i>	18 3.0 cM	18q11.1-q11.2
	<i>Hoxb1</i>	11 56.08 cM	17q21.3
	<i>Ikaros</i>	11 6.0 cM	7p13-p11.1
	<i>Mash1</i>	10 C1	12q22-q23
	<i>Mixl1</i>	1 H4	1q42.12
	<i>Myf5</i>	10 59.0 cM	12q21
	<i>Nestin</i>	3 42.5 cM	1q23.1
	<i>NeuroD</i>	2 46.0 cM	2q32
	<i>Neurog1</i>	13 35.0 cM	5q23-q31
	<i>Pax6</i>	2 58.0 cM	11p13
	<i>Pdx1</i>	5 82.0 cM	13q12.1
	<i>Sox1</i>	8 4.0 cM	13q34
	<i>Sox17</i>	1 7.0 cM	8q11.23
	<i>Sox3</i>	X 24.5 cM	Xq27.1
	<i>Sox7</i>	14 C3	8
	<i>Wnt3</i>	11 63.0 cM	17q21
	<i>Zic1</i>	9 61.0 cM	3q24

The chromosomal location for each locus is according to the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/>).

Other loci showing differentiated gene expression between ICM/ES cells (*Rex1* and *Gbx2*) (Chapman et al., 1997; Rogers et al., 1991) and epiblast/EpiS cells (*Fgf5*, *Psc1* and *Prce*) (Brons et al., 2007; Haub and Goldfarb, 1991; Hebert et al., 1991; Pelton et al., 2002; Tesar et al., 2007) were also included in this analysis. Note that the selection of these markers could only be based on expression analysis of mouse embryos, as equivalent information is not available in human embryos. Genes involved in the trophoctoderm lineage specification in both human and mouse embryos include the caudal-related gene *Cdx2* (Beck et al., 1995; Hay et al., 2004) and the T-box gene *Eomes* (Babaie et al., 2007; Ciruna and Rossant, 1999; Russ et al., 2000). While *Dazl* (Ruggiu et al., 1997) and *Sycp1* (Meuwissen et al., 1992) are typical examples of selected germ cell markers. 22 additional genes implicated in the specification of somatic lineages were also included among which are the ectoderm, neuronal-specific genes, *Sox1* (Collignon et al., 1996; Malas et al., 1997) and *Sox3* (Collignon et al., 1996; Stevanovic et al., 1993) and the mesoderm, myogenic or haematopoietic genes *Myf5* (Braun et al., 1989; Rudnicki et al., 1993) and *Ikaros* (Georgopoulos et al., 1992), respectively. The zinc-finger transcription factor *Gata6* (Koutsourakis et al., 1999; Morrissey et al., 1998) is known to be involved in PrE development while the pancreatic *Pdx1* (Ahlgren et al., 1996; Stoffers et al., 1997) is an example of definitive endoderm marker. The full list of selected loci as well as their chromosomal locations on mouse and human genomes is summarized in Table 1.

The replication timing of individual loci was assessed using a previously described method (Azura, 2006; Azura et al., 2003; Hansen et al., 1993); an outline of this assay is shown in Figure 3.2. Briefly, non-synchronised exponentially growing cells were pulse-labelled with 5-bromo-2-deoxyuridine (BrdU), a thymidine analogue that is incorporated by DNA polymerase into newly replicating DNA. The cells are then stained with the DNA-binding dye propidium iodide (PI) and fractionated by flow cytometry into six fractions according to DNA content. The first fraction includes cells in G0 and G1, while S-phase is represented by four fractions (designated as S1 to S4) and the G2 fraction includes cells that have completed DNA replication. A representative cell cycle profile for mES cells is shown in Figure 3.2, where the non-overlapping gates used to sort the six cell cycle fractions, are indicated. After DNA isolation and sonication, newly synthesised BrdU-labelled DNA, from an equal number of sorted cells, is purified by immunoprecipitation with anti-BrdU antibodies. Locus-specific sequences are then quantified by real-time quantitative PCR (qPCR) using primers amplifying the regions around 600 bp upstream of the transcription start sites of the selected genes. The relative abundance of newly synthesized locus-specific DNA present in each sorted cell cycle fraction was represented as a percentage of the

total. Importantly, the replication timing of each 45 candidate genes and controls were analysed in parallel using the same cell cycle sorted, BrdU-labelled DNA template.

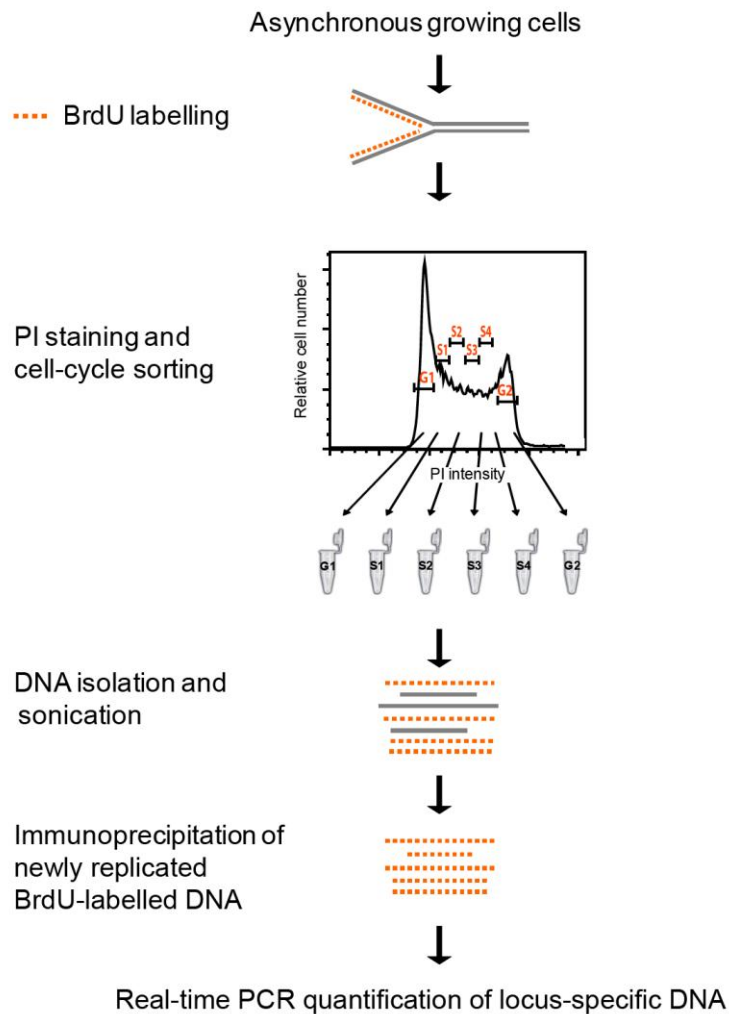


Figure 3.2. The principle of replication timing analysis.

Asynchronously growing cells are pulse-labelled with 5-bromo-2-deoxyuridine (BrdU) which is incorporated into replicating DNA and sorted according to DNA content into six cell cycle fractions (designated as G1, S1-S4 and G2) using propidium iodide (PI) staining. DNA is then isolated from each of the sorted fractions and sheared by sonication. Newly synthesised BrdU-labelled DNA is then recovered by immunoprecipitation using anti-BrdU antibodies and the relative abundance of locus-specific sequences analysed by qPCR.

Controls for this analysis included the *α -globin* locus, a constitutively early-replicating gene (Smith and Higgs, 1999) and pancreatic *Amylase*, a tissue-specific gene, which consistently replicates late in S-phase in several cell types (Azuara et al., 2003; Hatton et al., 1988; Mileham and Brown, 1996). In agreement with previously published data, the early replicating control *α -globin* showed peak abundance in S1/S2 fractions isolated from hES, mES and mEpiS cells, whereas the late replicating control *Amylase* was abundant in S3 and peaked in the S4 fractions (Figure 3.3). To ensure equivalent precipitation of BrdU-labelled DNA, each cell cycle fraction was spiked with a constant amount of *Drosophila* S2 BrdU-labelled DNA prior to immunoprecipitation. The detection of similar levels of *Drosophila EMC* PCR products in each fraction confirmed uniform recovery of BrdU-labelled DNA (Figure 3.3, right panel) as assessed in each experiment.

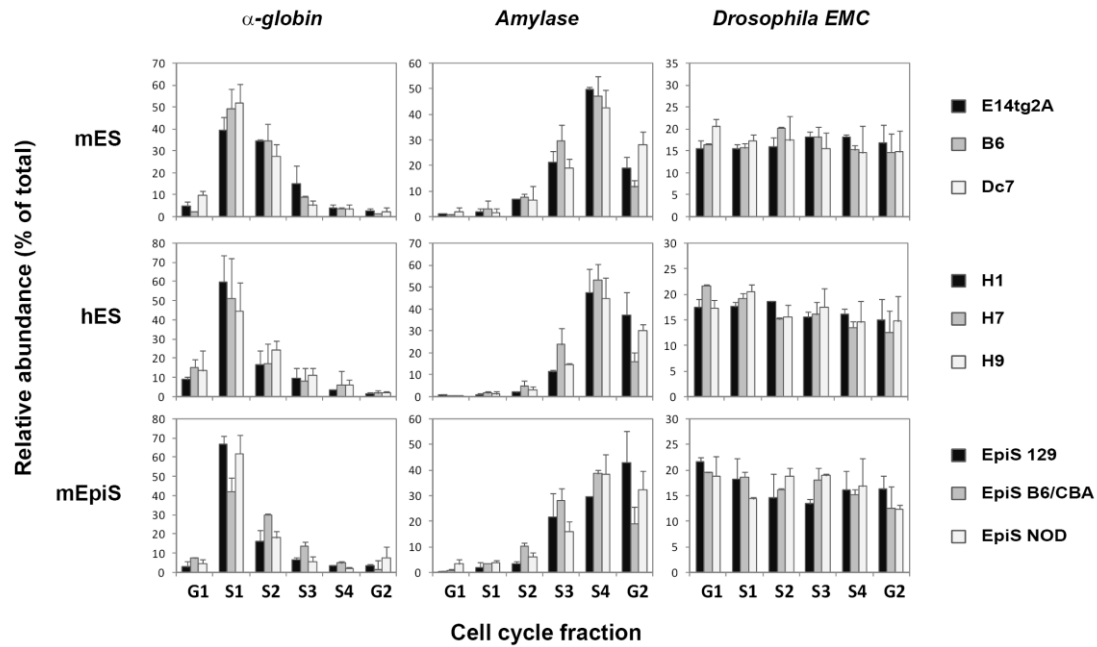


Figure 3.3. Replication timing analysis controls. Histograms showing the relative abundance of PCR products for early and late replicating controls (α -globin and Amylase, respectively) in each cell cycle fraction for mES (E14tg2A, B6 and Dc7), hES (H1, H7 and H9) and mEpiS (129, B6/CBA and NOD) cells. The *Drosophila melanogaster EMC* gene provides an additional internal control to ensure uniform recovery of BrdU-labelled DNA. Mean and standard deviation of at least two independent experiments are shown for each cell type analysed.

3.2. Replication timing profiles are conserved between independent mES cell lines

Replication timing of each candidate gene was classified according to its peak of abundance in G1/S1 (early), S2 (middle-early), S2 and S3 (middle), S3 (middle-late) and S4/G2 (late) and shown as a gradation of colours from green (earliest) to red (latest) (see examples of replication timing assignment in Figure 3.4a).

Three mES cell lines independently established from different mouse inbred strains were analyzed by DNA replication timing. Line E14tg2A was derived 22 years (Hooper et al., 1987) ago from a 129 mouse strain, while B6 was derived 16 years ago (Kontgen et al., 1993) from a C57BL/6 mouse strain. In contrast, Dc7 was more recently derived (2 years ago; Nesterova T., unpublished) from a mixed background 129/castaneus hybrid mouse and is therefore genetically polymorphic. All the three mES cell lines were consistently cultured in the presence of LIF and serum; E14tg2A cells were grown on gelatin-coated surfaces while B6 and Dc7 required MEFs as a substrate for optimal propagation. Note that Dc7 cells were also maintained in a slightly different medium composition, as described in section 2.2.1.

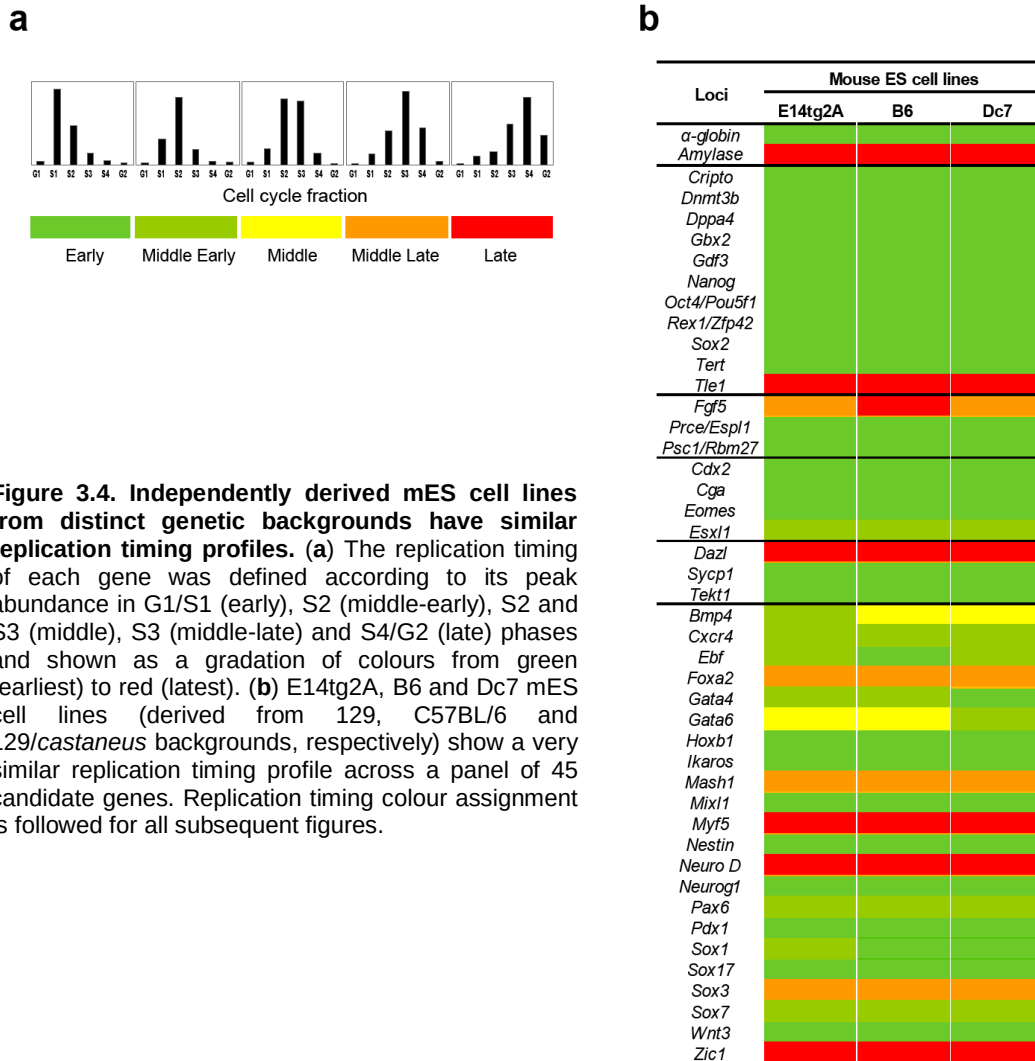


Figure 3.4. Independently derived mES cell lines from distinct genetic backgrounds have similar replication timing profiles. (a) The replication timing of each gene was defined according to its peak abundance in G1/S1 (early), S2 (middle-early), S2 and S3 (middle), S3 (middle-late) and S4/G2 (late) phases and shown as a gradation of colours from green (earliest) to red (latest). **(b)** E14tg2A, B6 and Dc7 mES cell lines (derived from 129, C57BL/6 and 129/castaneus backgrounds, respectively) show a very similar replication timing profile across a panel of 45 candidate genes. Replication timing colour assignment is followed for all subsequent figures.

Despite the disparate genetic, temporal and culture condition histories of these three cell lines, their replication profiles were remarkably consistent across many loci, with few exceptions (*Bmp4* and *Gata6*) (Figure 3.4b). In agreement with previously published results (Azuara, 2006; Hiratani et al., 2008), I confirmed that mES cells have a largely early-replicating profile with a majority of genes replicating early (green) or middle-early (light green) S-phase (73 - 76%, see Appendix II, FigureA1). This included a number of key pluripotency-associated genes highly expressed in mES cells (such as *Oct4*, *Rex1* and *Nanog*). However, many other genes that are selectively expressed in somatic lineages (such as *Ikaros*, *Neurog1* and *Sox17*) but silent in ES cells were also found to be early replicating, consistent with a broadly accessible chromatin structure in ES cells (Azuara et al., 2006). In contrast, *Foxa2*, *Mash1*, *Myf5*, *NeuroD*, *Sox3* and *Zic1* replicated steadily in middle-late (orange) or late (red) S-phase, across the three cell lines, further highlighting that lineage induction in ES cells is likely to be regulated by selective gene derepression (Perry et al., 2004; Williams et al., 2006).

3.3. Pluripotent mouse and human ES cells have distinct replication timing profiles

The replication timing of the same panel of loci examined in mES cells (most of them localized in syntenic regions between mouse and human genomes) were next analysed in H1, H7 and H9 hES cell lines. All these cell lines were simultaneously derived 11 years ago (Thomson et al., 1998) and expanded for this analysis using hES cell standard culture conditions (Xu et al., 2001) on matrigel-coated plates in MEFs-conditioned media (CM), supplemented with FGF2, as previously described in section 2.2.1. To facilitate a comparison between mouse and human cell lines, the relative proportion of genes replicating in early, middle-early, middle, middle-late or late S-phase was here summarised as pie-charts shown in Figure 3.5a and Appendix II, Figure A1. Although many genes analysed in the three hES cell lines consistently replicated early or middle-early (53-58%), the replication profiles of hES cells were significantly distinct from that of mES cells. A profound difference was the increased proportion of genes replicating in middle or second half of S-phase in hES cells (42-47%) when compared to mES cells (24-27%) (Figure 3.5b). Despite the similar temporal and culture condition histories of these three hES cell lines, their replication timing profiles were also significantly different among cell lines, with variable proportion of loci replicating in middle, middle-late and late fractions (Figure 3.5).

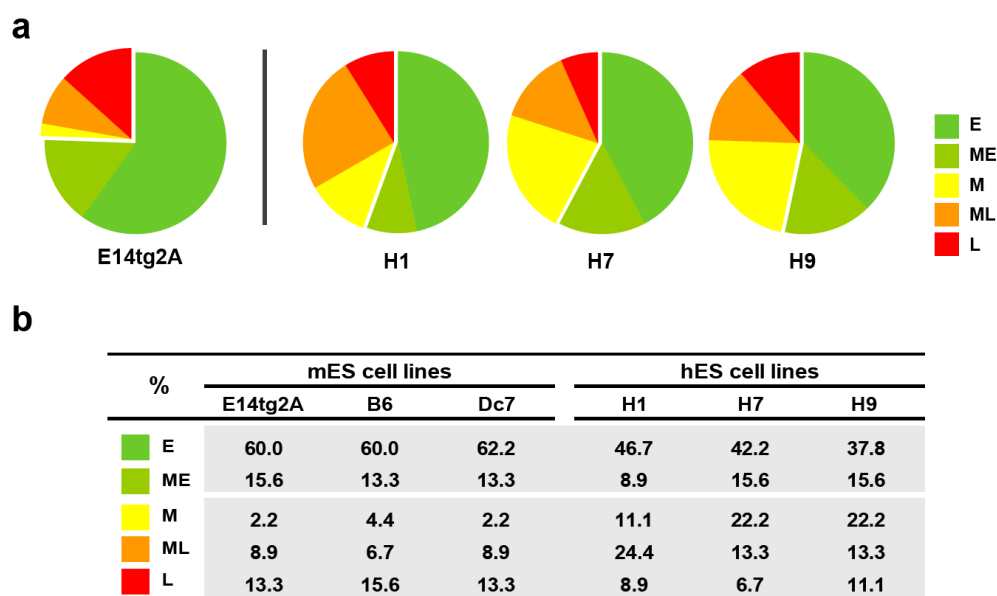


Figure 3.5. Pluripotent mouse and human ES cells have distinct replication timing profiles. (a) The proportion of genes that replicate early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) in S-phase in mES cells (E14tg2A) and in three hES cell lines (H1, H7 and H9) are represented as a pie-chart. (b) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of total as assessed in mES (E14tg2A, B6 and Dc7) and hES (H1, H7 and H9) cells.

At the locus level, 11 out of the 45 genes analysed (*Dppa4*, *Tert*, *Tle1*, *Sycp1*, *Bmp4*, *Ebf*, *Foxa2*, *Gata6*, *Sox17*, *Sox3* and *Zic1*) showed variable replication kinetics between H1, H7 and H9 hES cell lines (Figure 3.6a). Moreover, among these genes, *Tert* replication timing profiles also varied between cultures of the same cell line. *Dppa4* and *Ebf* showed the most dramatic replication timing changes from early to late, as seen in only one of the lines or across the three cell lines, respectively (Figure 3.6b). Other genes such as *Gata6* and *Tle1* displayed more subtle differences (from middle to middle-late). In contrast, *Nanog* and *NeuroD* are shown as examples of genes that consistently replicated in H1, H7 and H9 cells, in early and late S-phase, respectively (Figure 3.6c).

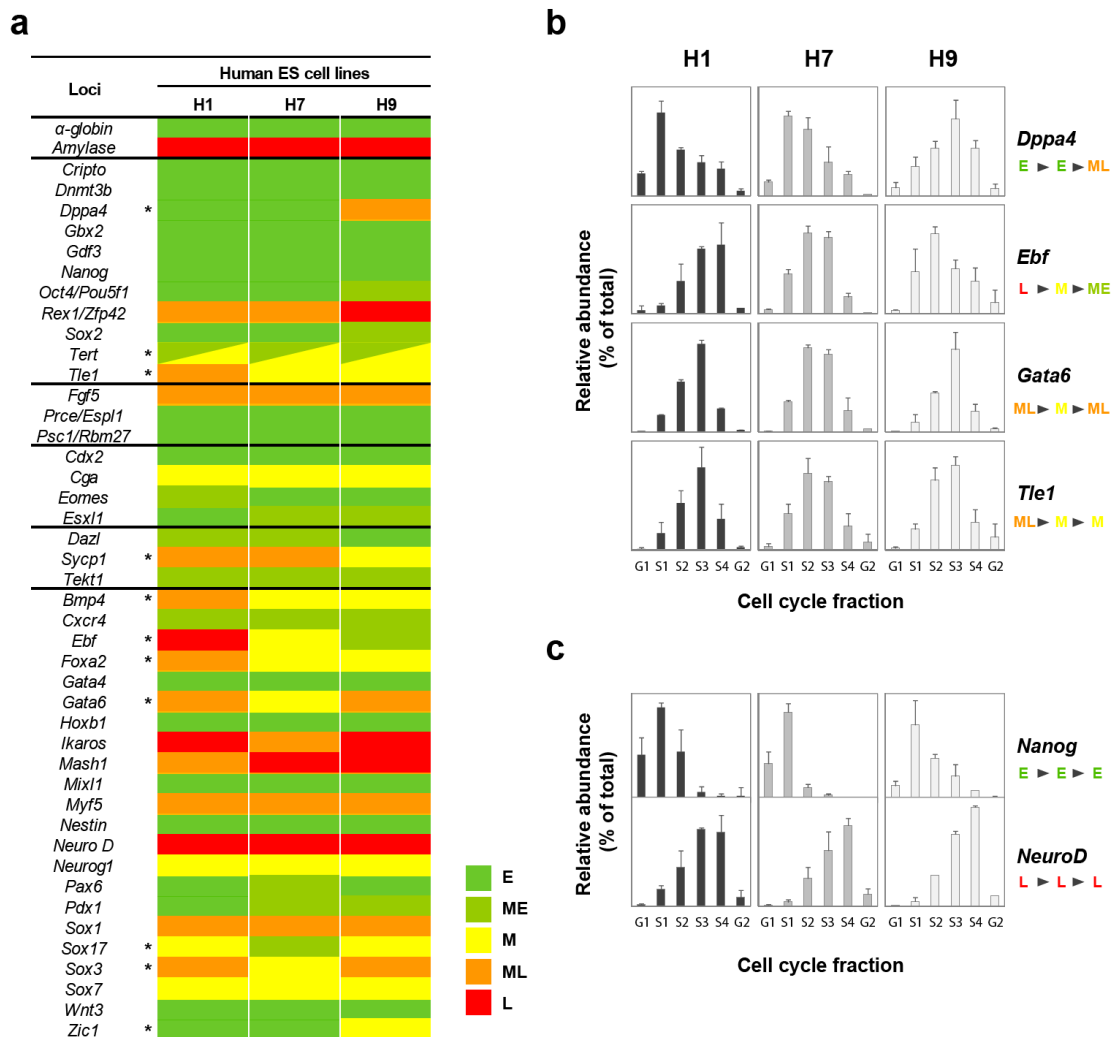


Figure 3.6. Replication timing variability between established hES cell lines. (a) The replication timing of individual genes was compared between three hES cell lines (H1, H7 and H9). Genes showing differential replication kinetics are marked by asterisks. (b) Histograms showing the relative abundance of locus-specific DNA within each cell cycle fraction (G1, S1-S4 and G2 phases). Examples of loci that switch dramatically their replication timing, as seen in only one of the lines (*Dppa4*) or across the three lines (*Ebf*), or those who show slight temporal delays (*Gata6* and *Tle1*) or (c) not changing at all their replication timing (*Nanog* and *NeuroD*), as assessed by qPCR in H1 (black), H7 (grey) and H9 (white) hES cell lines. Mean and standard deviation of two or more experiments are shown for each cell type analysed.

The observation that hES cells showed an increased proportion of late replicating loci, a feature that is normally found in differentiated cells (Azuara et al., 2006), as well as high levels of variability across lines and cultures, was puzzling and prompted us to carefully verify the undifferentiated state of these cultures. To do so, I analysed the expression patterns of the pluripotency transcription factors Oct4, Nanog and Sox2, alongside the specific surface markers SSEA4, Tra-1-81 and Tra-1-60 by immunofluorescence and/or FACScan analysis. As shown in Figure 3.7a and b, hES cell cultures displayed homogeneous expression of Oct4, Nanog and Sox2 as seen across the three cell lines analysed. Tra1-60, Tra1-81 and SSEA4 were also highly expressed in H1, H7 and H9 cells, while SSEA1 (a marker of differentiation) was virtually undetectable. AP activity was also strongly detected in more than 90% hES cell colonies (Figure 3.7c). These observations exclude that high level of spontaneous differentiation within hES cell cultures can simply explain the higher proportion of late replication and variability observed in hES cell profiles, in contrast to that of mES cells.

3.4. Human ES cell replication profiles can be reliably discriminated from that of cells with more restricted differentiation capacity

3.4.1. Replication timing comparison of hES and hES-derived neural progenitor cells

To investigate whether pluripotent hES cells have a distinct chromatin profile from that of cells with a more restricted differentiation capacity, as previously demonstrated in mES cells (Azuara et al., 2006; Hiratani et al., 2008; Perry et al., 2004), we compared the replication timing of 24 candidate genes in undifferentiated hES cells (H7 line) and H7-derived neural progenitors (NP7). Neural cells were induced using an *in vitro* adherent culture system, where inhibition of the BMP signalling by its antagonist Noggin at the initial stage of differentiation is sufficient to produce highly enriched and expandable populations of NP cells (Gerrard et al., 2005). Note that FGF2 and EGF growth factors were continuously added to NP cultures in order to keep them cycling, as requirement for the replication timing experiment. The replication profiles of NP7 cells were determined alongside a well-characterized NP cell line (ReNcell VM) derived from a 10-week human fetal brain, as a reference.

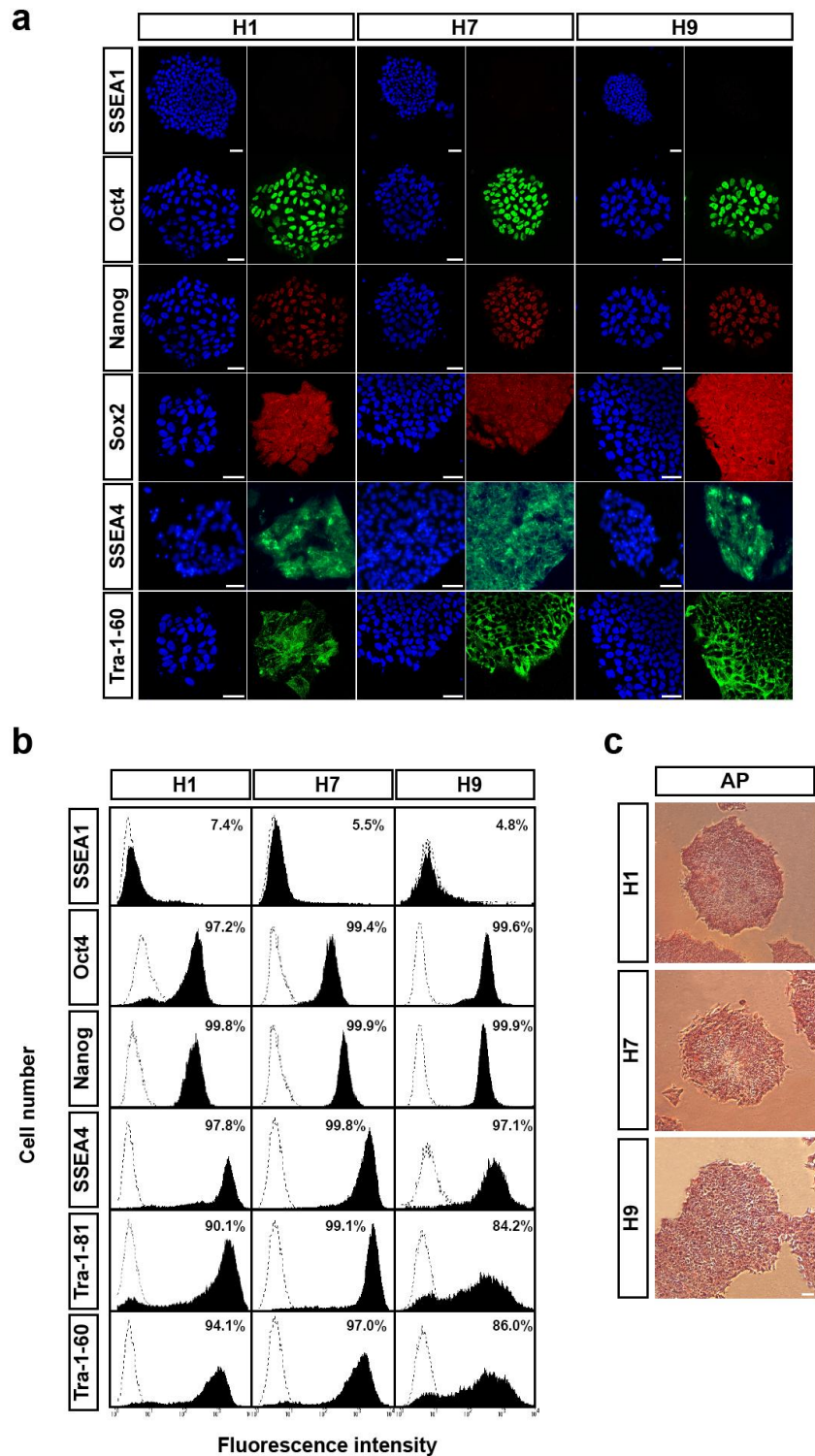


Figure 3.7. Human ES cell cultures analysed by replication timing display homogeneous expression of pluripotency-associated markers. H1, H7 and H9 hES cells were analyzed by (a) immunofluorescence and (b) flow cytometry for the hES differentiation surface marker SSEA1 (red) and the characteristic pluripotency and undifferentiation markers Oct4 (green), Nanog (red), Sox2 (red, immunofluorescence), SSEA4 (green), Tra-1-60 (green) and Tra-1-81 (FACS). Nuclei are shown by DAPI staining (blue). FACS analysis showed that 84.2-99.9% hES cells expressed markers of undifferentiation, while only 4.8-7.4% cells were positive for SSEA1, as indicated in the histograms. The profile of stained cells (black) was compared to cells stained with the secondary antibody only (dotted line). (c) Alkaline phosphatase (AP) activity (pink) was also strongly detected in all hES cell colonies. Scale bars, 50 μ m.

As illustrated in Figure 3.8a, after >60 days of differentiation, a homogeneous population of hES-derived NP cells displaying a typical bipolar morphology is observed. In contrast to undifferentiated H7 cells, hES-derived NPs showed by immunofluorescence a drastic downregulation of Oct4 and a robust staining for Sox2 and other neural precursor markers, including the intermediate filament Nestin (Hockfield and McKay, 1985; Lendahl et al., 1990), Musashi (Kaneko et al., 2000; Sakakibara et al., 1996) and Pax6 (Walther and Gruss, 1991) confirming their neural identity and the absence of contaminating undifferentiated cells (Figure 3.8b). Efficient neural induction was further confirmed in NP7 cells by RT-PCR (Figure 3.8c), which showed the upregulation of *Nestin*, *Sox1* and *Pax6* mRNA levels and a dramatic reduction in the expression of the pluripotency-associated markers *Oct4* and *Nanog*. As expected *Sox2* showed consistent levels in both hES and NP cell populations (Figure 3.8b and c) (Graham et al., 2003), while the endodermal and mesodermal markers, *Gata6* (Morrissey et al., 1996) and *T* (Edwards et al., 1996) were undetectable (Figure 3.8c).

As shown in Figure 3.8d, undifferentiated H7 hES cells have a distinct profile from that of NP7 and ReNcell VM cells. In particular, neuroectodermal markers either remain early (*Nestin* and *Pax6*) or late (*Mash1*) or undergo a shift towards early replication (*NeuroD*, *Sox1* and *Sox3*) in NP7 as compared to H7 cells, suggesting an increase in the chromatin accessibility and transcriptional permissiveness of these genes. In contrast, trophoblast (*Eomes*), endodermal (*Foxa2*, *Gata6* and *Sox17*) and mesodermal (*Bmp4*, *Ebf*, *Ikaros* and *Myf5*) markers, all undergo a shift towards late replication in NP7 cells, in accordance with ReNcell VM neural stem cell profile. This suggests that epigenetic changes associated with decreased chromatin accessibility at lineage inappropriate genes in neural committed cells have taken place, as previously seen in mES cells (Azura et al., 2006; Hiratani et al., 2008; Perry et al., 2004). The pluripotency-associated markers (*Nanog*, *Oct4* and *Sox2*) remained early replicating in H7 hES cells upon neural differentiation, while (*Cripto*, *Dppa4* and *Tle1*) showed a dramatic delay of their replication timing in NP7 cells, consisted with their shut-down upon differentiation and loss of pluripotency. *Tert* was the only exception, its instable middle-early/middle replication profile across cultures become consolidated as a middle-early locus in NP7 differentiated cells, probably associated with the reactivation of *Tert* in late NP cultures (M. Sheldon, unpublished results). Altogether, these results suggest that pluripotent hES cells can be reliably discriminated from cells with a more restricted developmental capacity by DNA replication timing.

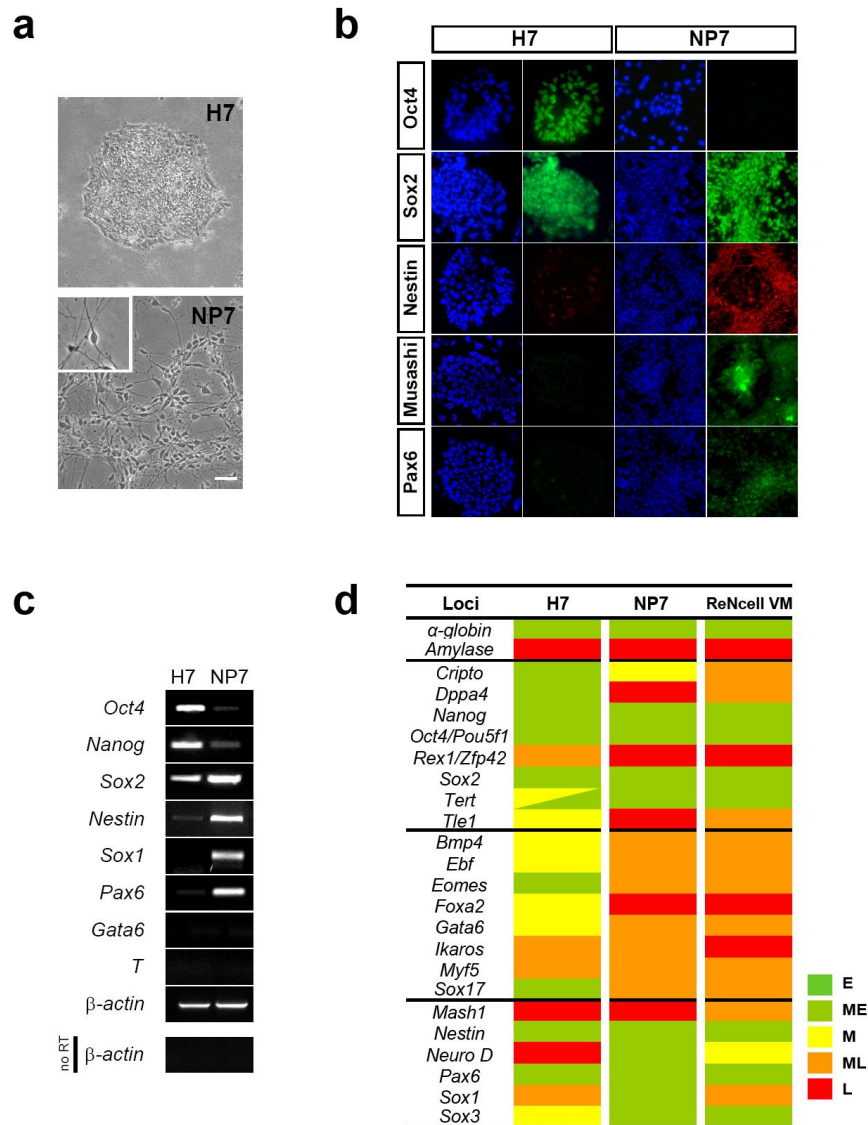


Figure 3.8. H7 undifferentiated hES cells can be clearly distinguished by replication timing analysis from neural progenitors (NP) H7-derived. (a) Phase-contrast images of undifferentiated H7 hES cells (top panel) grown on matrigel-coated plates in conditioned media and NP7 cells with typical bipolar morphology (bottom panel) after >60 days of differentiation, induced as previously described (Gerrard et al., 2005). Scale bar, 75 μ m. (b) Undifferentiated H7 hES and NP7-derived cells were analysed for the expression of Oct4, Sox2, Nestin, Musashi and Pax6 by immunofluorescence staining. Nuclei were stained with DAPI. Magnification, 200X (c) RT-PCR analysis of the expression of ES-associated genes (*Oct4*, *Nanog* and *Sox2*), neural determinants (*Nestin*, *Sox1* and *Pax6*), endodermal and mesodermal markers (*Gata6* and *T*, respectively) in H7 hES and NP7 cells. β -actin was used to standardise input. (d) Replication timing comparison between H7 and NP7-derived cells alongside the well-characterized ReNcell VM human NP cell line derived from a 10-week human fetal brain, used as a reference for neural differentiation. The loci selected for this analysis included key pluripotency markers (*Cripto*, *Dppa4*, *Nanog*, *Oct4*, *Rex1*, *Sox2*, *Tert* and *Tle1*), trophoblast (*Eomes*), endoderm (*Foxa2*, *Gata6* and *Sox17*), mesoderm (*Bmp4*, *Ebf*, *Ikaros* and *Myf5*) and neuroectoderm (*Mash1*, *Nestin*, *NeuroD*, *Pax6*, *Sox1* and *Sox3*) markers.

3.4.2. Replication timing comparison of human ES and CD34+ haematopoietic stem cells

The observation that hES cell lines display variable replication timing profiles, prompted us to ask if such level of plasticity is an exclusive feature of undifferentiated hES cells or could be due to the unique genetic identity of embryos from which hES cells are derived. If this is the case, other human cell types derived from the “outbred” human population should also display variability in their profiles. Here, the replication timing of the 11 loci targets of variability across hES cell lines, alongside other loci that were described as steady, was analyzed in human CD34+ haematopoietic stem cells isolated from peripheral blood of three healthy volunteers (Figure 3.9).

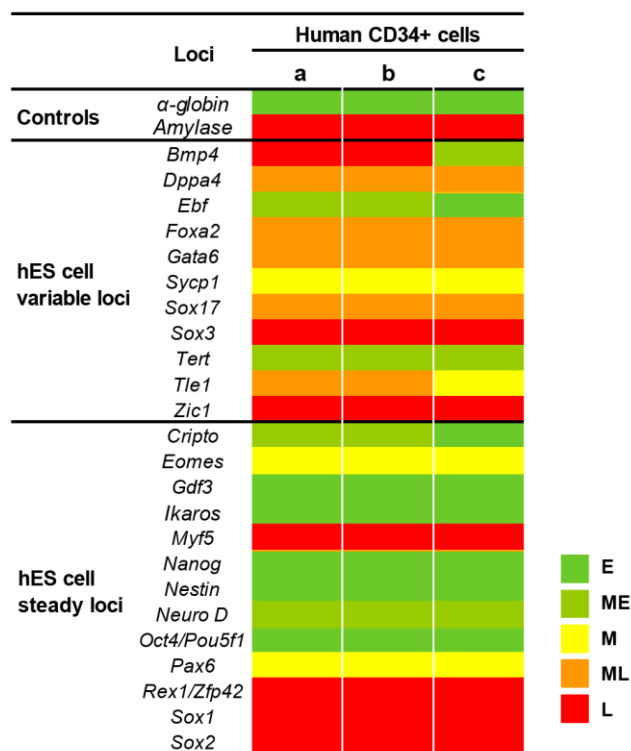


Figure 3.9. Replication timing variability is not observed across human CD34+ haematopoietic stem cells. The replication timing of variable loci across different hES cell lines was analysed in CD34+ cells isolated from peripheral blood of three different healthy volunteers (here represented as **a**, **b** and **c**), alongside loci that showed steady replication timing. All the loci that were shown to be steady were maintained, while all other targets of variability in hES cells replicated at similar times across the three different sources of adult stem cells, with the exception of *Bmp4* and *Tle1* that were earlier in **c**.

The replication timing of variable genes was shown to be remarkably similar (9 out of 11 genes) across the three different populations of CD34+ cells, with the exception of *Bmp4* and *Tle1* (Figure 3.9). Consistently, all the analyzed genes that previously were described as steady remained constant across samples (Figure 3.9). Interestingly, some genes replicated at later times in S-phase in all the three CD34+ samples than previously seen in hES cells (Figure 3.6), that is the case of *Eomes*, *Pax6* and *Sox2* consistent with a progressive decline of chromatin accessibility in more lineage committed cells. In contrast, other genes replicate earlier in CD34+ cells than in hES cells. These differences are lineage and transcriptional dependent, as highlighted

by *Ebf* and *Ikaros*, two known haematopoietic genes (Medina and Singh, 2005), and *Neuro D* (Graf et al., 2001) that are all expressed and early replicating in CD34+ cells in contrast to hES cells. These observations further corroborate that hES cells have a distinct profile from cells with a more restricted developmental potential and that replication timing variability is a hallmark of undifferentiated hES cells under standard CM cultures, independently of the genetic variation across individuals.

3.5. Analysis of the gene expression status and chromosomal context of replication timing variable genes

3.5.1. Gene expression and histone acetylation analysis of genes shifting replication timing across H1, H7 and H9 hES cells

Changes on DNA replication timing are thought to reflect variations in chromatin accessibility and transcriptional permissiveness (Donaldson, 2005). Although the precise relationship between DNA replication, chromatin structure and transcription is complex and remains largely unresolved, previous data in mES cells (Azucara et al., 2006; Jorgensen et al., 2007; Perry et al., 2004) together with studies in other cell types (Lande-Diner et al., 2009; Lin et al., 2003; Schubeler et al., 2004; Zhang et al., 2002) suggested a positive correlation between histone acetylation and early-replicating active genes and between deacetylated histones and late-replicating, less active genes. Here it was investigated if the differential replication kinetics observed in hES cell lines at specific genes (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sox17*, *Sox3*, *Sycp1*, *Tert*, *Tle1* and *Zic1*) may reflect changes at the transcriptional and/or histone acetylation levels (Figure 3.10).

Chromatin immunoprecipitation (ChIP) assays were performed to evaluate the levels of acetylated histones at the promoters of individual loci, in H1, H7 and H9 hES cell lines. Crosslinked chromatin isolated from hES cells was immunoprecipitated with antibodies against H3K9ac, H3K14ac, H3ac, H4ac and total H3 and H4 or with a rabbit anti-mouse IgG antiserum as a negative control (Figure 3.10c). DNA eluted after immunoprecipitation was analysed by qPCR using primers amplifying sequences in promoter regions looking at both steady (*Oct4*, *Nanog*, *Sox2*, *Rex1*, *Myf5* and *Eomes*)

and variable (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sycp1*, *Sox17*, *Sox3*, *Tert*, *Tle1* and *Zic1*) replicating genes in H1, H7 and H9 hES cells. Primers for pericentric heterochromatin (*satellite repeats*) were also included, providing a negative control for active histone marks immunoprecipitations. To reduce potential bias that might arise from different nucleosomal densities, the data for each gene was expressed relative to the enrichment obtained with the antibodies against total histone H3 or H4 (Figure 3.10).

These ChIP data was compared with the DNA replication timing and expression status of selected genes across the three hES cell lines (Figure 3.10a). Highly expressed and early replicating genes across H1, H7 and H9 hES cells (such as *Oct4*, *Nanog* and *Sox2*) showed a correlation with higher levels of histone acetylation, as previously described for mES cells (Azuara et al., 2006; Jorgensen et al., 2007; Perry et al., 2004). In contrast, silent and late replication genes (such as *Myf5*) showed only background levels as seen at *satellite repeats* control region. Accordingly, *Rex1*, an ES-specific late replicating gene showing lower levels of expression (relatively to *Oct4* and *Nanog*) and *Eomes*, an early replicating but silent lineage-specific gene showed low but detectable levels of acetylated histones. These data, from steady replicating loci, are consistent with histone acetylation being correlated with early replication and/or active expression as previously reported. In contrast, no straightforward correlation between replication timing variability, histone acetylation patterns and/or expression could be drawn when analysing variable replicating genes across H1, H7 and H9 cells. *Dppa4*, *Tert* and *Tle1*, are three variable replicating genes expressed in the three hES cell lines and showing high levels of both H3 and H4 acetylated histones, yet no correlation with their replication timing status is observed. Although the replication timing pattern of *Dppa4* (early in H1 and H7 and late in H9 cells) may reflect its variable expression (up to three-fold lower in H9 when compared to H1 cells, Figure 3.10a), no correlation was observed at the histone acetylation level since only H7 seems to have higher levels of enrichment for H3K9ac, H3K14ac and H3ac in contrast to H1 and H9 hES cells (Figure 3.10b). Other genes expressed at low levels (*Bmp4*, *Foxa2* and *Gata6*) and replicating in the second half of S-phase show however, different patterns; *Bmp4* has substantial levels of acetylated histones in all three hES cell lines, relatively to *Foxa2* or *Gata6*. Other silent genes (*Ebf*, *Sycp1*, *Sox17*, *Sox3* and *Zic1*) also show variable or absent patterns of acetylated histones across H1, H7 and H9 cells, regardless of their replication timing status. Altogether, these data demonstrate that replication timing variability is not directly underlined by gene expression variability at these loci, across hES cell lines, or merely explained by distinct histone acetylation patterns.

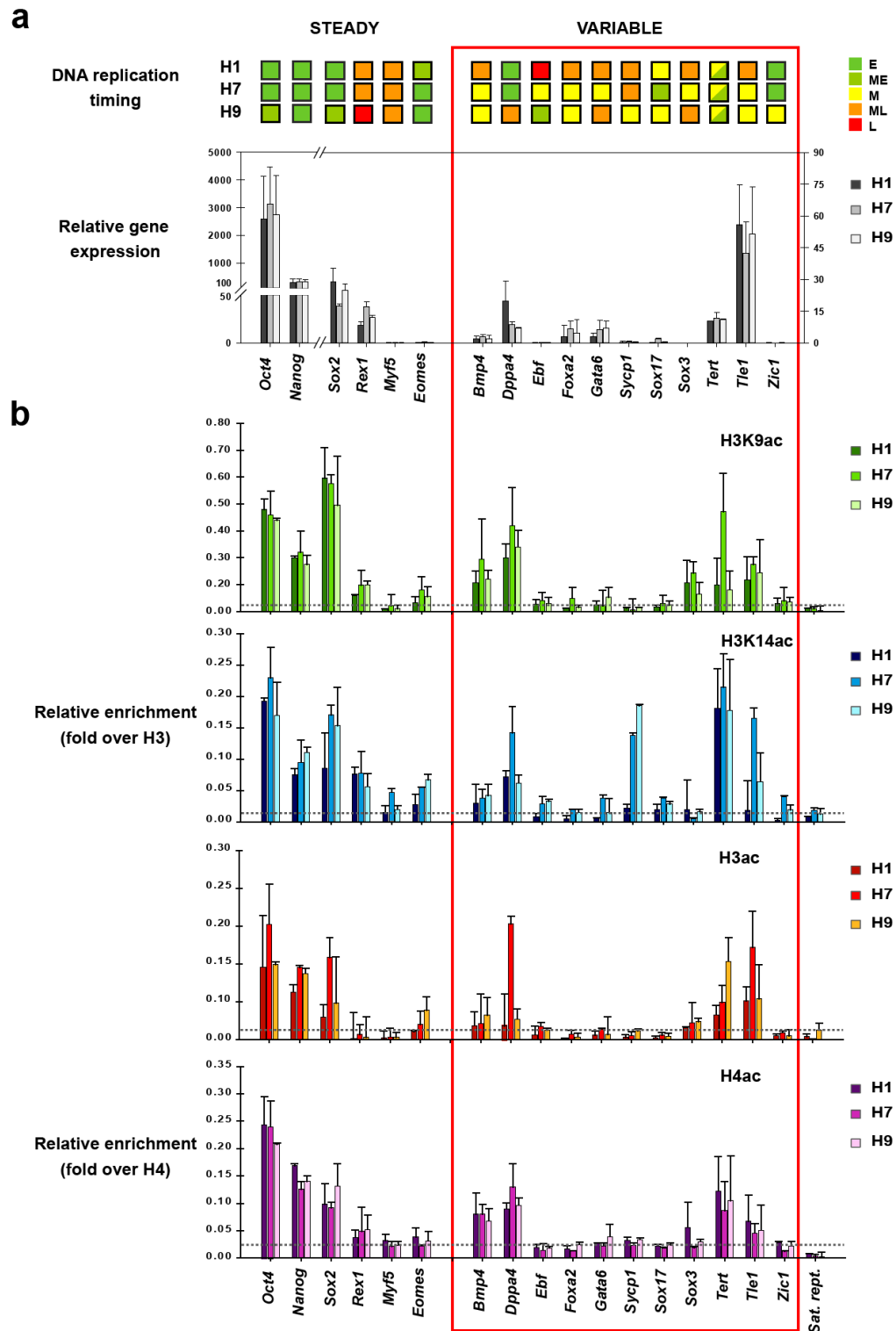


Figure 3.10. Human ES cells replication timing variability is not associated with gene expression or histone acetylation specific patterns. (a) Replication timing status and expression of variable loci (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sycp1*, *Sox17*, *Sox3*, *Tert*, *Tle1* and *Zic1*) and early (*Oct4*, *Nanog*, *Sox2* and *Eomes*) or late (*Rex1* and *Myf5*) steady control loci between H1, H7 and H9 hES cell lines. The replication timing of each locus is indicated in a colour coded way as previously designated. The relative expression of variable and steady replicating loci was assessed by qPCR and compared between H1 (black), H7 (grey) and H9 (white) hES cell lines. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*). Values shown are the mean of three independent experiments and error bars indicate standard deviation. (b) Relative enrichments for acetylated histones (H3K9ac, H3K14ac, H3ac and H4ac) were assessed in H1, H7 and H9 hES cells using chromatin immunoprecipitation and qPCR. The colour code for different histone marks is indicated in each panel.

Primers for satellite repeats (*Sat. rept.*) were used as a negative control for active histone marks immunoprecipitations. To eliminate bias for different nucleosome density, data was normalized to enrichments obtained with antibodies recognizing total histone H3 or H4. Error bars indicate standard deviations of at least two independent experiments. Background levels (from control IgG antibodies) are shown as dashed grey lines. The red open box highlights the variable replicating cohort of genes across H1, H7 and H9.

3.5.2. Genomic features analysis of genes shifting replication timing across H1, H7 and H9 hES cells

Work from I. Hiratani and colleagues suggested that genes located within low GC content (<43% GCs) and high LINE density (>6% LINE) regions might be more likely to shift their replication timing during development, irrespectively of their expression status (Hiratani et al., 2004). To investigate whether genes showing differential replication kinetics in hES cell lines (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sox17*, *Sox3*, *Sycp1*, *Tert*, *Tle1* and *Zic1*) may also share common chromosomal features, GC content, LINE repeat and gene density were analysed in the vicinity of each of those loci (Table 2), (see section 2.2.5 for details). As previously outlined in Table 1, these genes are located in different chromosomes throughout the genome or very far apart from each other when located in the same chromosome. Although a substantial proportion of shifting genes (8 out of 11) were scored as low GC content and high LINE density (black boxes), other genes that remain early (such as *Pax6*) or late replicating (such as *Ikaros*) across H1, H7 and H9 cells showed similar genomic features. Hence, no clear correlation could be established between hES cell replication timing variability and chromosomal location or genomic features at target genes. Interestingly, however, all genes shifting replication timing, apart from *Tert* were located in gene-poor genomic regions (Table 2, grey boxes). Although this is not a unique feature of variable loci in hES cells, it implies that changes in replication timing as observed at those genes are unlikely to be influenced by the transcriptional status of neighbouring genes.

Table 2 - Genomic features of replication timing analysed genes

	Loci	GC content (%) ^a	LINE density (%) ^b	Gene density ^c
hES cell variable loci	<i>Bmp4</i>	39.7	23.2	2.0
	<i>Dppa4</i>	41.6	17.9	5.0
	<i>Ebf</i>	39.2	14.2	5.0
	<i>Foxa2</i>	42.9	15.3	4.0
	<i>Gata6</i>	43.9	9.4	2.0
	<i>Sox17</i>	40.8	24.7	4.0
	<i>Sox3</i>	40.3	25.2	3.0
	<i>Sycp1</i>	36.6	31.9	10.0
	<i>Tert</i>	55.1	10.2	15.0
	<i>Tle1</i>	43.3	8.9	3.0
	<i>Zic1</i>	38.8	10.0	1.0
hES cell steady loci	<i>Amylase</i>	35.4	19.5	4.0
	<i>Cdx2</i>	44.1	15.1	9.0
	<i>Cga</i>	38.7	17.8	9.0
	<i>Cripto</i>	44.6	17.9	19.0
	<i>Cxcr4</i>	42.6	13.9	5.0
	<i>Dazl</i>	39.8	22.2	5.0
	<i>Dnmt3b</i>	47.5	6.0	17.0
	<i>Eomes</i>	40.5	19.1	2.0
	<i>Esxl1</i>	41.1	15.1	7.0
	<i>Fgf5</i>	38.5	22.9	3.0
	<i>Gata4</i>	47.2	15.1	22.0
	<i>Gbx2</i>	47.8	12.8	7.0
	<i>Gdf3</i>	43.6	14.3	18.0
	<i>Hoxb1</i>	48.5	4.1	23.0
	<i>Ikaros</i>	40.9	17.9	6.0
	<i>Mash1</i>	39.9	26.5	3.0
	<i>Mixl1</i>	42.5	15.5	14.0
	<i>Myf5</i>	35.5	15.9	6.0
	<i>Nanog</i>	44.9	7.0	19.0
	<i>Nestin</i>	51.1	7.6	32.0
	<i>NeuroD</i>	38.3	24.2	6.0
	<i>Neurog1</i>	47.0	15.3	9.0
	<i>Oct4/Pou5f1</i>	47.1	13.2	40.0
	<i>Pax6</i>	41.6	14.8	6.0
	<i>Pdx1</i>	44.3	18.8	11.0
	<i>Prce/Espl1</i>	50.8	10.5	31.0
	<i>Psc1/Rbm27</i>	39.0	4.6	9.0
	<i>Rex1/Zfp42</i>	41.4	20.8	2.0
	<i>Sox1</i>	47.4	13.1	1.0
	<i>Sox2</i>	39.5	9.5	0.0
	<i>Sox7</i>	46.6	15.7	7.0
	<i>Tekt1</i>	42.1	39.5	34.0
	<i>Wnt3</i>	49.0	7.3	12.0
	<i>α-globin</i>	52.1	6.5	53.0

a & b - Genes displaying a GC content <43% and a LINE density >6% are highlighted in black.

c - Light grey boxes indicate gene poor regions which contain ≤10 genes in the surrounding 500 kb.

3.5.3. DNA methylation analysis at promoters of replication timing variable genes across H1, H7 and H9 hES cells

Different studies have recently reported the existence of variability at the DNA methylation level across different hES cell lines, mostly at random sites across the genome (Allegrucci et al., 2007; Maitra et al., 2005). Based on these studies we decided to investigate whether replication timing variability observed in H1, H7 and H9 hES cells at specific genes (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sox17*, *Sox3*, *Sycp1*, *Tert*, *Tle1* and *Zic1*; Figure 3.11, red open box) may be underlined by DNA methylation variability.

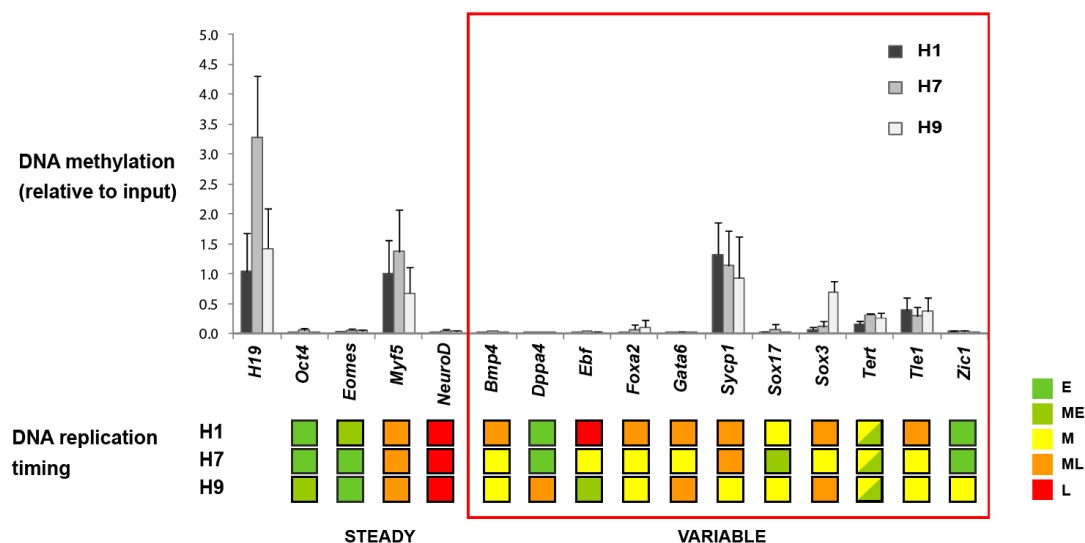


Figure 3.11. Human ES cells replication timing variability is not associated with DNA methylation specific patterns. Abundance of CpG DNA methylation at each promoter region of replication timing variable loci (*Bmp4*, *Dppa4*, *Ebf*, *Foxa2*, *Gata6*, *Sycp1*, *Sox17*, *Sox3*, *Tert*, *Tle1* and *Zic1*) highlighted by the red open box, between H1 (black), H7 (grey) and H9 (white) hES cells. DNA methylation was assessed using methylated DNA immunoprecipitation (MeDIP) (Weber et al., 2005) and qPCR analysis. Error bars indicate the standard deviation of three independent experiments. Early (*Oct4* and *Eomes*) and late (*Myf5* and *NeuroD*) steady replicating loci across H1, H7 and H9 hES cells were also analyzed. Controls for methylated (*H19-Igf2* imprinting control region) and unmethylated (*Oct4*) regions were included. The enrichment indicates the ratio of methylated DNA over input DNA. The replication timing of each locus (steady and variable) in H1, H7 and H9 cells is indicated by a colour code as previously designated.

For this analysis I used, a previously described assay, methylated DNA immunoprecipitation (MeDIP) (Weber et al., 2005) that allows the generation of DNA methylation profiles across large panel of genes. The principle of this assay relies on the immunoprecipitation of methylated genomic DNA using an antibody that specifically recognizes 5-methylcytidine (see section 2.2.10. for details). The methylated imprinted

control region *H19-Igf2* was used as a positive control in the experiments. As shown in Figure 3.11, high levels of DNA methylation were only observed at the promoter of *Myf5* (steady replicating) and *Sycp1* (variable replicating), both genes replicate in the second half of S-phase in H1, H7 and H9 cells. In contrast, *Neuro D* (steady replicating) or *Foxa2* and *Gata6* (variable replicating) that also replicate in late or middle/middle-late S-phase were not significantly enriched relatively to control *H19*, excluding a direct link between DNA methylation and late replication at single-copy genes. This contrasts with previous reports showing a key role for DNA methylation in maintaining a late replicating status of repetitive sequences, including major satellite repeats, in mouse cell lines (Jorgensen et al., 2007; Selig et al., 1988). While *Sox3* showed variable enrichment levels across the three hES cell lines (undetectable in H1 and H7 but high in H9 cells), this did not correlate with its replication timing profile in each of these cell lines, showing no clear relationship between DNA methylation and replication timing variability.

Altogether, these data argue against a direct role for DNA methylation in regulating the timing of DNA replication at specific loci in hES cells. This is consistent with previous studies showing that the replication timing of individual loci did not change in DNA methyltransferase-deficient mES cells (Gribnau et al., 2003; Jorgensen et al., 2007).

3.6. Human ES and mEpiS cells share a common unstable or transitional state: primed on the verge of differentiation

As described above, hES cell replication timing profiles greatly vary between cell lines and cultures, a level of variability not observed among mES cell lines derived from different genetic backgrounds. In undifferentiated hES cells, replication timing variability was shown to be unrelated to genetic variation or any specific genomic features. Such a level of plasticity within these cultures could however indicate an unstable or transitional state: poised on the verge of differentiation into multiple directions. Interestingly, it has been proposed that hES cells more closely resemble mEpiS cells than mES cells, two stem cell populations representing the pluripotent progenitors in mouse gastrulating embryos (mEpiS) and blastocysts (mES), respectively (Brons et al., 2007; Tesar et al., 2007). In favour of this idea, *Rex1* - an ICM-specific marker that is shutdown in the epiblast (Rogers et al., 1991), was found to

replicate early during S-phase in mES cells but consistently late in H1, H7 and H9 hES cell lines grown under standard CM conditions, (Figure 3.4 and 3.6). This prompted us to further characterize the level of similarities and/or differences between hES, mES and mEpiS cells at the chromatin level. In particular, I asked whether higher levels of replication timing variability could also be observed in mEpiS cells, in contrast to mES cells. To do so, I analysed the replication timing profiles of three independently derived mEpiS cell lines from different genetic backgrounds (129, B6/CBA and NOD) and compared them with those previously obtained from mES and hES cell lines (Figure 3.12 and Appendix II, Figure A2 for full analysis).

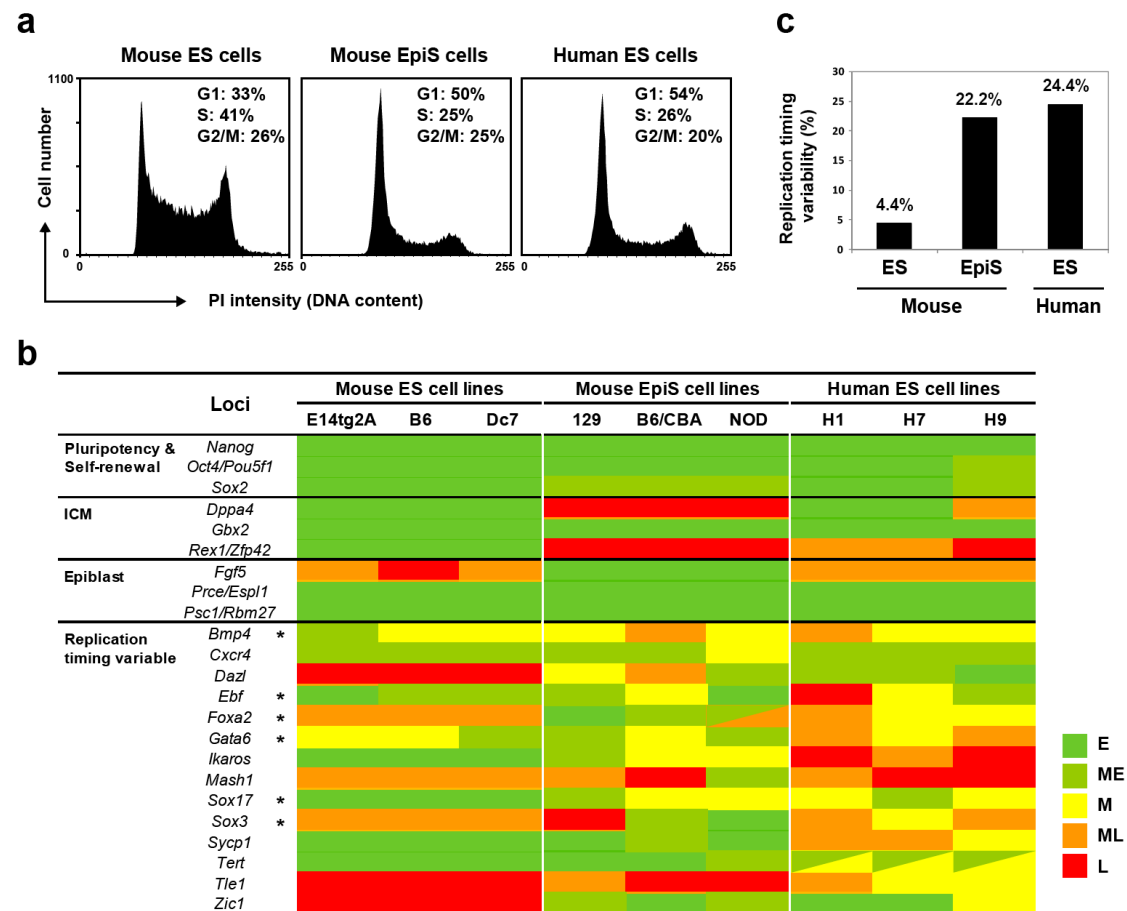


Figure 3.12. Mouse EpiS and hES cells have identical cell cycle kinetics and variable replication timing profiles across many genes. (a) Representative cell cycle profiles of mES (E14tg2A), mEpiS (129) and hES (H1) cells. Cells were stained with propidium iodide (PI) and DNA content analyzed on a FACScalibur flow cytometer. The percentages of cells into each phase of the cell cycle were determined using the WinMDI 2.9 software. The cell cycle profiles of mEpiS cells and hES cells (middle and right panels, respectively) show a similar pattern, with a lower number of cells seen in the S-phase when compared with mES cells. (b) Comparative replication timing analysis of mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7, and H9) cell lines. Selected loci shown for comparison are either associated with pluripotency and self-renewal (*Nanog*, *Oct4* and *Sox2*), selectively expressed in the inner cell mass (ICM) (*Dppa4*, *Gbx2* and *Rex1*), and the epiblast (*Fgf5*, *Prce* and *Psc1*) or showing variable replication timing across lines and cultures with specific roles during development (*Bmp4*, *Cxcr4*, *Dazl*, *Ebf*, *Foxa2*, *Gata6*, *Ikaros*, *Mash1*, *Sox17*, *Sox3*, *Sycp1*, *Tert*, *Tle1* and *Zic1*). Genes that are equally variable between mEpiS cell lines and also between hES cell lines are marked by asterisks. For full list of

genes analyzed see Appendix II, Figure A2. (c) Percentage of replication timing variable loci between cell lines of the same population relative to the total 45 analyzed loci in mES (4.4%), mEpiS (22.2%) and hES (24.4%) cell lines. Mouse EpiS and hES cells have in common high percentages of variable replication timing loci in contrast to mES cells. This analysis was performed in collaboration with Kavita Mehta and Matias Autio in our laboratory.

Initially it was assessed the replication timing of key ICM (*Dppa4*, *Gbx2* and *Rex1*) (Bortvin et al., 2003; Chapman et al., 1997; Maldonado-Saldivia et al., 2007; Rogers et al., 1991) and epiblast-specific markers (*Fgf5*, *Prce* and *Psc1*) (Brons et al., 2007; Haub and Goldfarb, 1991; Hebert et al., 1991; Pelton et al., 2002; Tesar et al., 2007) alongside *Nanog*, *Oct4* and *Sox2*, that are expressed in both pluripotent tissues. This analysis was further extended to loci previously studied in mES and hES cells, especially differentiation-associated or loci that were subjected to variable replication kinetics in human cells (see Figure 3.12).

During the course of this analysis, we observed that mEpiS and hES cells share similar cell cycle profiles, with a relatively lower number of cells in S-phase (25% and 26%, respectively) as compared to mES cells (41%), suggesting common growth properties (Figure 3.12a). Mouse EpiS cells showed a mostly early replication timing profile with 71-78% of loci analysed replicating in early or middle-early S-phase, a percentage similar to that found in mES cells (73-76%) (Appendix II, Figure A1). These two developmentally distinct but closely related stem cell populations could, however, be reliably discriminated using our replication timing-based approach. The ICM markers *Rex1* and *Dppa4* were found to replicate early and late during S-phase in mES and mEpiS cells, respectively consistent with dynamic changes in gene expression patterns. In contrast, the replication timing of *Fgf5*, an epiblast-specific marker, shifts from late in mES cells to early in mEpiS cells, underlying gene derepression (Figure 3.12b). Further differences were observed among lineage-specific genes, such as *Zic1* late replicating in mES cells and early in mEpiS cells.

As anticipated, *Nanog*, *Oct4* and *Sox2* were consistently expressed and early replicating in mouse (ES and EpiS) and hES cells. However, we could not confirm whether or not hES cells are indeed more closely related to an mES cell-like or mEpiS cell-like developmental state, based on ICM/epiblast marker replication timing profiles only. While *Rex1* was found late replicating in both hES and mEpiS cell populations, *Fgf5* remained late in hES cells as seen in mES. *Gbx2*, *Psc1* and *Prce* were early replicating in all mouse and human cell types analysed, regardless of their developmentally regulated gene expression patterns. However, the selection of these markers is here based on gene expression analysis in mouse embryos, as equivalent information is not available in human embryos yet. Interestingly, however, I observed a

high level of replication timing variability among mEpiS cell lines (22.2% variable loci), as previously seen in hES cells (24.4%) and in contrast to the homogeneous and steady replicating mES cell (4.4%) profiles (Figure 3.12c).

Altogether, these results suggest that both hES and mEpiS pluripotent cells share a common unstable or transitional state: primed on the verge of differentiation, opening more questions concerning the true nature and origin of hES cells during normal development as recently raised (Lovell-Badge, 2007).

CHAPTER 4. IDENTIFICATION AND CHARACTERIZATION OF DISTINCT PLURIPOTENT STATES WITHIN HUMAN ES CELL CULTURES: NAIVE VS PRIMED

4.1. Undifferentiated hES cell replication timing profile is prone to extensive alterations upon different cell culture conditions

4.1.1. Replication profile of H9 hES cells grown in two different laboratories

As recently shown, mES cultures can toggle between distinct pluripotent states from a naïve ICM-like to a more primed epiblast-like states (Nichols and Smith, 2009) that can interconvert in response to different culture conditions (Bao et al., 2009; Chou et al., 2008; Guo et al., 2009; Hanna et al., 2009). The presence of activin A and/or FGF2 was shown essential to maintain a hES pluripotent state (Levenstein et al., 2006; Vallier et al., 2005; Xiao et al., 2006) however, since 1998 when the first hES cell lines were reported (Thomson et al., 1998), a plethora of techniques to stably maintain hES cells in culture have been developed in different laboratories (Adewumi et al., 2007). Here I asked whether hES cell replication timing profiles could be directly and dynamically altered by changing cell culture conditions.

Initially the profiles of H9 hES cells grown in our laboratory using standard CM conditions [matrigel-coated plates and MEF-CM with KSR and FGF2 (Xu et al., 2001)] was compared with that of cultures of the same cell line that were, however, routinely maintained into CDM^{activin} conditions [FBS-coated plates and KSR-free defined medium supplemented with FGF2 and activin A (Vallier et al., 2005)], as provided by Dr Ludovic Vallier's laboratory (Figure 4.1a). Interestingly, the replication timing profile of H9 hES cells was dramatically altered, from late to early, when cultured under CDM^{activin} conditions. As shown in Figure 4.1b, the proportion of genes replicating early or middle-early was significantly increased in H9 cells grown for more than 20 passages in CDM^{activin} conditions [H9p50(CDM20)] when compared to H9 cells in CM (H9p35CM)

(from 55% to 75%). This suggested that hES cell culture can adopt different epigenetic states when grown under different culture conditions in different laboratories.

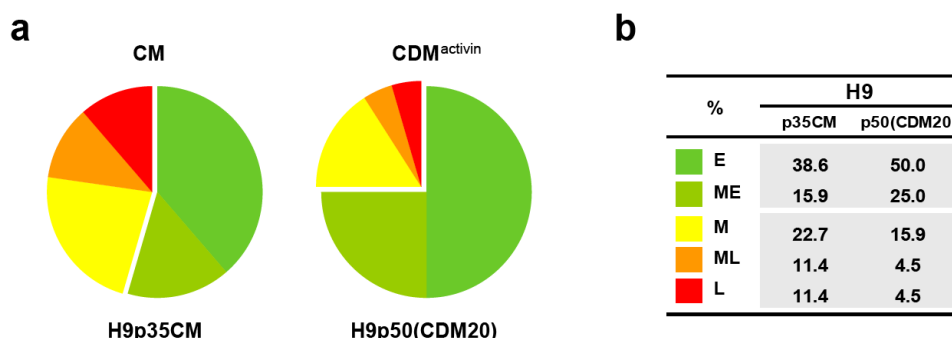


Figure 4.1. Human ES cells grown under different culture conditions in different laboratories show dramatic replication timing late to early shifts. The proportion of genes that replicate early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) in S-phase are represented as a pie-chart. **(a)** The replication timing of H9 hES cells grown in our laboratory under standard matrigel-coated plates and irradiated mouse embryonic fibroblast – conditioned medium (CM) with Knockout™ Serum Replacement (KSR) and fibroblast growth factor (FGF2) conditions (Xu et al., 2001) (H9p35CM) were compared with that of cells that have been cultured for more than 20 passages into FBS-coated plates, KSR-free chemically defined medium (CDM) supplemented with FGF2 and activin A (CDM^{activin}) (Vallier et al., 2005) as provided by Dr Ludovic Vallier's laboratory [H9p50(CDM20)]. **(b)** Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for H9 cells grown under standard CM and CDM^{activin} conditions. Note that nomenclature for culture conditions is according to the following convention: cell line, number of passages in CM culture, plus additional passages in CDM^{activin} conditions. This format is followed for all subsequent figures.

4.1.2. Dynamic equilibrium and heterogeneity of hES cells in distinct epigenetic states

To directly address whether this was a reproducible and dynamically reversible process, the behaviour of other hES cell lines (H1 and H7) grown in our laboratory in CM conditions and transferred to CDM^{activin} conditions was compared and backwards. We observed differences among the two hES cell lines in their response to CDM^{activin} conditions. While H1 hES cells rapidly settled into CDM^{activin} conditions when plating at high densities and gradually exposed to increasing concentrations of CDM^{activin} media (25, 50 and 75%), H7 cells could not be maintained in an undifferentiated state when transferred into these conditions, despite several attempts using variable conditions (i.e. different plating densities and/or in the presence of MEFs).

After 3 passages in CDM^{activin} H1 hES cells were readily stabilised, with little evidence of differentiation (see section 4.2.1). Remarkably, extensive replication timing alterations were already evident, with an increased proportion of early and middle-early replicating genes from 55% in H1p49CM to 64% in H1p47(CDM3) and up to 77% in H1p47(CDM6) cells (Figure 4.2). Some loci, such as *Rex1* promptly and dramatically

shifted their replication timing (from middle-late to middle-early). Progressive changes were also noticed as seen for *Bmp4* (from middle-late, middle to middle-early) over 6 passages (Figure 4.3a and b). Interestingly, there is a significant overlap (11 loci) between the loci that consistently advanced their replication timing towards early S-phase in H1 and H9 hES cell lines in CDM^{activin} conditions (*Rex1*, *Tert*, *Cga*, *Sycp1*, *Bmp4*, *Foxa2*, *Gata6*, *Neurog1*, *Sox1*, *Sox17* and *Sox3*) as highlighted by asterisks in Figure 4.3a and summarised by the Venn diagram in Figure 4.3c. Moreover, all the loci that previously were described as variable replicating across H1, H7 and H9 hES cell lines grown in CM become earlier under CDM^{activin} conditions in both H1 and H9 cells, providing valuable indicators for the epigenetic state of hES cells.

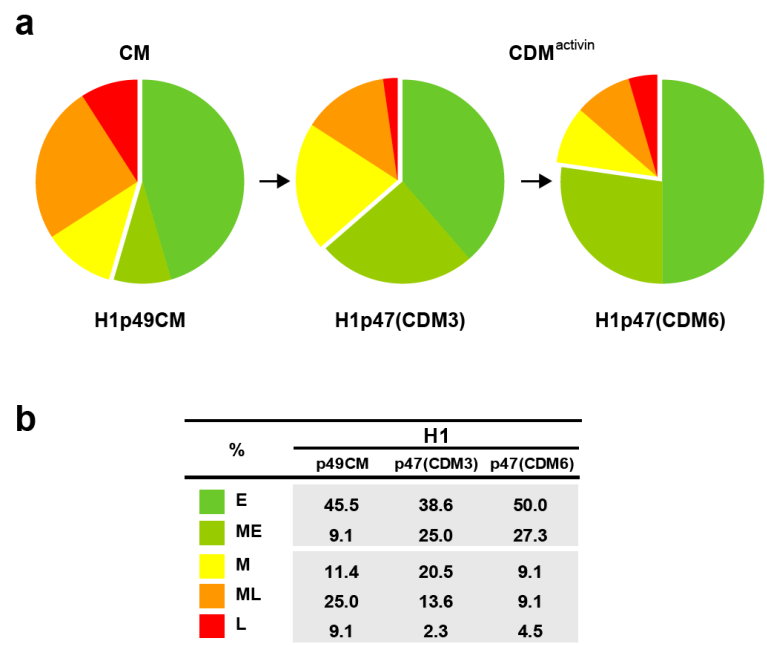


Figure 4.2. H1 hES cells grown under CM and CDM^{activin} conditions show dramatic replication timing late to early shifts. (a) The replication timing of H1 hES cells grown in standard CM (H1p49CM) conditions was compared to that of cells transferred to CDM^{activin} conditions after 3 [H1p47(CDM3)] and 6 [H1p47(CDM6)] passages. (c) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for H1 hES cells grown under standard CM and CDM^{activin} conditions.

Importantly, the modulation of hES cell replication profile was demonstrated to be a reversible process as returning the cells from CDM^{activin} to CM conditions was sufficient to revert to a more unstable late-replicating profile. The proportion of early and middle-early genes decreased to 46% in H1 re-swapped [H1p47(CDM14;RS6)] cells (Figure 4.3a and Appendix II, Figure A3), further illustrating the epigenetic plasticity of hES cells. Consistently, the replication profiles of both *Rex1* and *Bmp4* shifted back from middle-early to middle (Figure 4.3a and b). In contrast, *Dppa4*, a locus that was consistently early-replicating in H1 hES cells grown in either CM or CDM^{activin} conditions display a biphasic replication profile with two main peaks in S1 and S3 fractions (Figure 4.3a and b), underlying high level of variability in re-swapped cultures. Altogether these data challenge the view that hES cells most closely resemble mEpiS cells and suggest that distinct but interchangeable subpopulations may exist

within hES cell cultures as delineated by a stable, early-replicating (mES-like) *versus* unstable, late-replicating (mEpiS-like) profile.

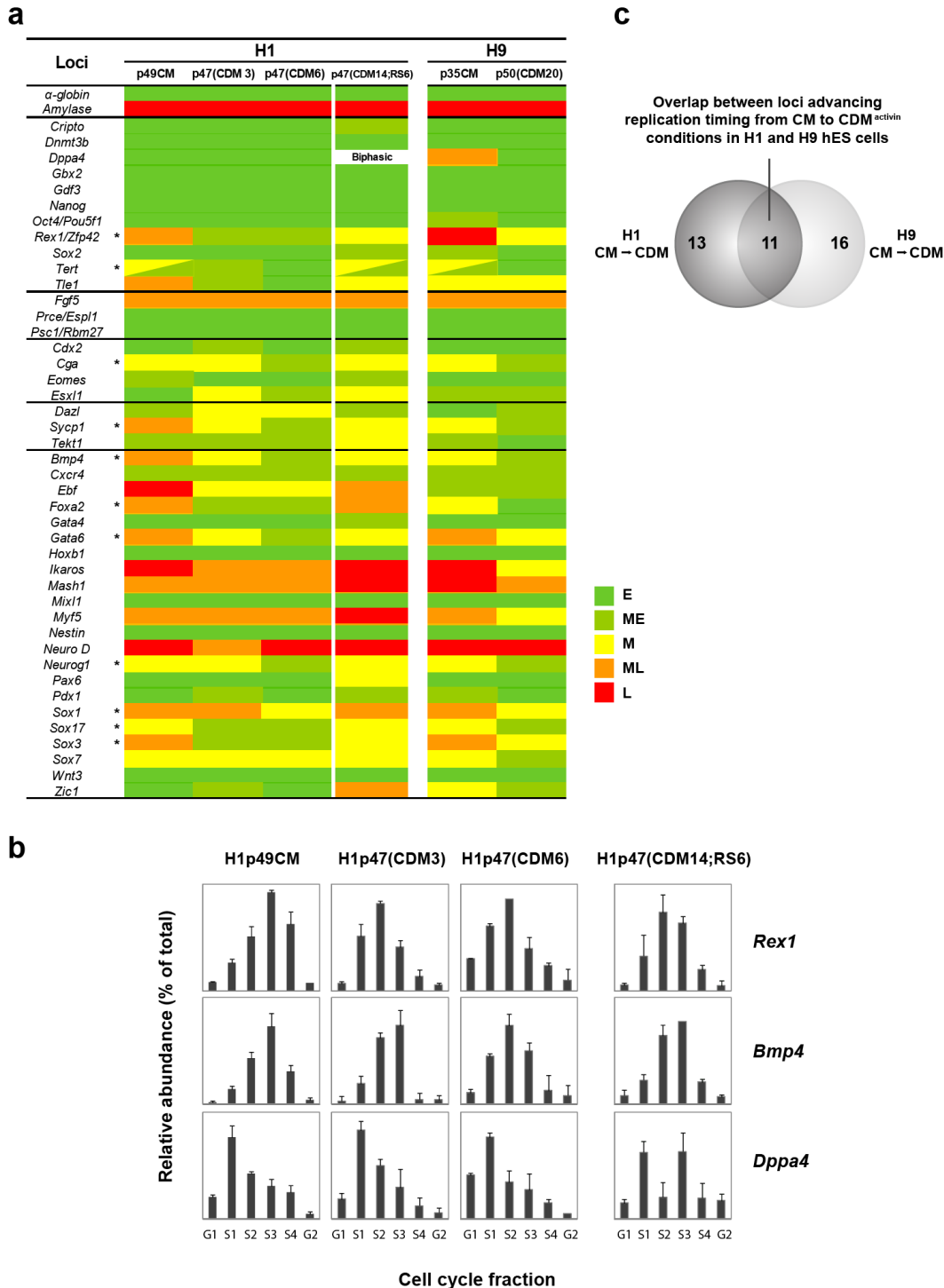


Figure 4.3. Loci subjected to replication timing shift towards early S-phase in CDM^{activin} conditions overlap between different hES cell lines. (a) The replication timing of individual loci was compared between H1 and H9 hES cells grown under standard CM culture conditions with that of cells that have been cultured for different number of passages into CDM^{activin}. Replication timing late to early shifts occur immediately in H1 cells after 3 passages [H1p47(CDM3)] and becomes even earlier after 6 passages [H1p47(CDM6)]. All these replication timing shifts can be reverted to some extent by re-swapping cells back to CM conditions [H1p47(CDM14;RS6)]. Loci that consistently advance their replication timing towards early across H1 and H9 hES cell lines when in CDM^{activin} conditions (*Rex1*, *Tert*, *Cga*, *Sycp1*, *Bmp4*, *Foxa2*, *Gata6*, *Neurog1*, *Sox1*, *Sox17* and *Sox3*) are marked by asterisks. (b) Histograms showing

the relative abundance of locus-specific DNA within each cell cycle fraction (G1, S1-S4 and G2 phases). Examples of loci that switch dramatically and immediately (*Rex1*) or progressively in time (*Bmp4*) their replication timing towards early as assessed by qPCR in H1 cells in CM and CDM^{activin} and are able to delay it back again when in re-swapped conditions. *Dppa4* is consistently early across cells grown in CM or CDM^{activin} conditions but becomes unstable and heterogeneous in terms of replication timing within the same population of cells in re-swapped conditions, displaying biphasic (S1 and S3) peaks. Mean and standard deviation of at least two experiments are shown for each cell type analysed. (c) Venn diagram illustrating the 11 common targets (overlap) between loci changing their replication timing towards early in CM to CDM^{activin} conditions across H1 (13 earlier loci) and H9 (16 earlier loci) hES cell lines.

4.1.3. Fully reprogrammed hiPS cell replication profiles matches hES cells dramatic late to early shifts in response to alterations in culture conditions

Human somatic cells can be reprogrammed into hiPS cells by the enforced expression of transcription factors (Takahashi et al., 2007; Yu et al., 2007) providing a long-sought strategy to generate patient and disease-specific pluripotent stem cells (Dimos et al., 2008; Park et al., 2008a). However, it is essential to completely understand the equivalence between hiPS and hES cells and clearly define appropriate cell culture conditions for the derivation and maintenance of hiPS cells with full developmental potential (Colman and Dreesen, 2009). Our results described above indicate that hES cell replication profiles can shift away from a mEpiS cell-like towards a mES cell-like profile in response to alterations in the culture conditions. Here, I analysed the effect of CM and CDM^{activin} conditions on hiPS cell replication profile to test whether these cells shown similar levels of epigenetic plasticity that hES cells. To do so, I took advantage of a previously established fully reprogrammed hiPS 40 cell line which was derived from neonatal foreskin fibroblasts by overexpression of Oct4, Sox2, c-Myc and Klf4 as described (Vallier et al., 2009b). Human iPS cells in CM conditions showed a profile very similar to the ones previously observed in H1, H7 and H9 hES cells, with a higher proportion of late replicating genes in contrast to mES cells (Figure 4.4a and b). Interestingly, hiPS cell replication profiles also shifted dramatically from late to early when cultured in CDM^{activin} conditions for more than 12 passages, with 57% and 77% of genes replicating in early or middle-early in CM and CDM^{activin} conditions, respectively (Figure 4.4b). This included *Rex1*, *Tert*, *Cga*, *Bmp4*, *Foxa2*, *Sox17* and *Sox3* loci previously seen to also shift their replication timing towards early in CDM^{activin} in H1 and H9 hES cells (Figure 4.4c). These results provide additional evidence that induced (iPS) and natural (ES) pluripotent cell lines share common properties and reiterate the impact of culture conditions in modulating or supporting different epigenetic states within pluripotent hES cell cultures. This is an important

consideration, as responses to differentiation protocols are likely to vary depending on where the cells stand in a developmental hierarchy (see section 4.3).

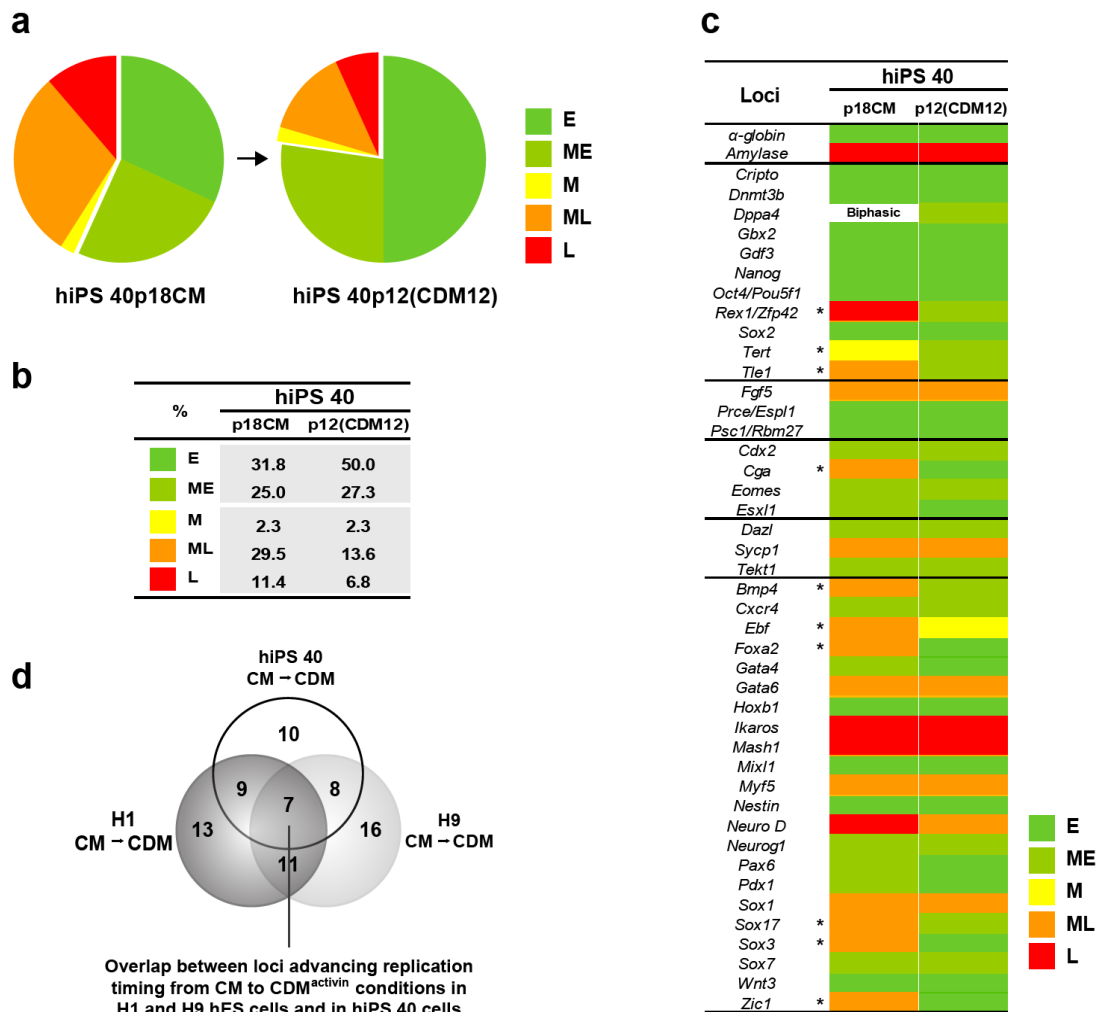


Figure 4.4. Fully reprogrammed hiPS cells replication profile also display dramatic late to early shifts in response to alterations in culture conditions. (a) The proportion of loci that replicate early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) in S-phase are represented as a pie-chart. The replication timing of fully reprogrammed hiPS 40 cells derived from neonatal foreskin fibroblasts by overexpression of four reprogramming factors (Oct4, Sox2, c-Myc and Klf4) (Vallier et al., 2009b) grown under CM conditions was compared to that of cells in CDM^{activin} conditions. (b) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for hiPS 40 cells grown under standard CM and CDM^{activin} conditions. (c) The replication timing of individual loci in hiPS 40 cells grown in CM or CDM^{activin} conditions is shown. Earlier replicating loci in CDM^{activin} conditions (*Rex1*, *Tert*, *Tle1*, *Cga*, *Bmp4*, *Ebf*, *Foxa2*, *Sox17*, *Sox3* and *Zic1*) are marked by asterisks. *Dppa4* is heterogeneous in terms of replication timing within the same population of cells in CM conditions, displaying biphasic (S1 and S4) peaks. (d) Venn diagram illustrating the 7 common targets (overlap) between loci changing their replication timing towards early in CM to CDM^{activin} conditions across H1 (13 earlier loci), H9 (16 earlier loci) hES cell lines and hiPS 40 (10 earlier loci) cells. These experiments were performed in collaboration with Matias Autio in our laboratory.

4.2. Phenotypic and functional characterization of distinct populations of undifferentiated hES cells in CM and CDM^{activin} growth conditions

4.2.1. CDM^{activin} conditions maintain H1 hES cells in an undifferentiated state supporting their self-renewal

Once acclimated, H1 hES cells grown in CDM^{activin} conditions formed homogeneous, tightly clustered colonies that adhered firmly to FBS-coated plates (Appendix II, Figure A4 and Figure 4.5a), with low or no sign of spontaneous differentiation, as sometimes observed in the edge of H1 cells cultured in CM. The undifferentiated state of CDM^{activin} converted H1 hES cells was further assessed by AP activity, IF and FACS analysis. As previously seen in H1 CM cells, H1 CDM^{activin} cells showed strong AP activity across colonies (Figure 4.5b)

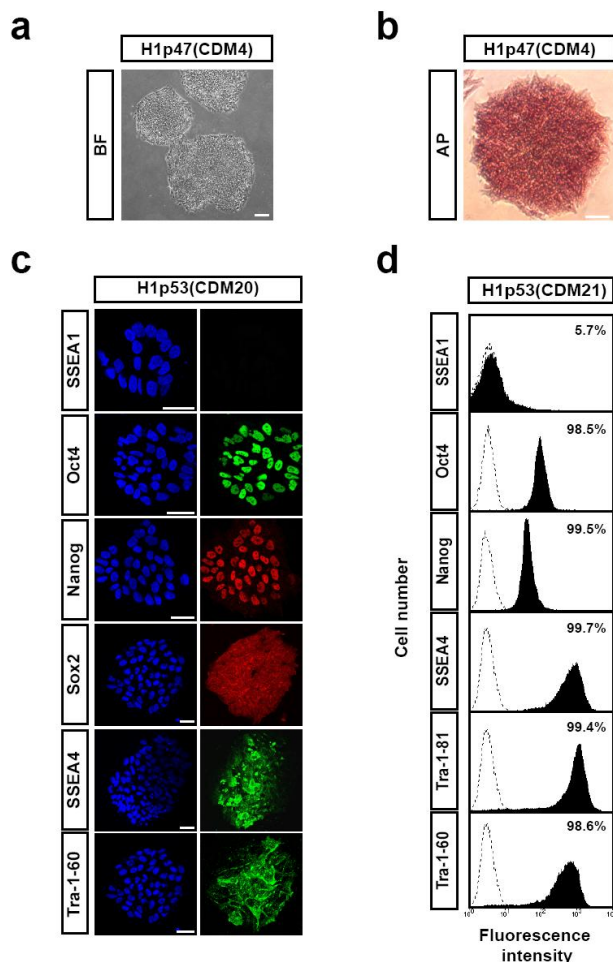


Figure 4.5. Human ES cells transferred from standard CM cell culture conditions to CDM^{activin} conditions retain their undifferentiated characteristics. (a) Bright field (BF) picture of representative undifferentiated H1 hES cell colonies and (b) detection of alkaline phosphatase (AP) activity (pink) in cells grown under CDM^{activin} conditions. Scale bars, 70 μ m. (c) Immunofluorescence and/or (d) flow cytometry analysis of H1 cells grown under CDM^{activin} conditions for the differentiation surface SSEA1 (red) marker and the pluripotency-associated markers Oct4 (green), Nanog (red), Sox2 (red, immunofluorescence), SSEA4 (green), Tra-1-60 (green) and Tra-1-81 (FACS). Nuclei are shown by DAPI staining (blue). FACS analysis showed that 98.5-99.7% hES cells expressed markers of undifferentiation, while only 5.7% of the cells were positive for SSEA1, as indicated in the histograms. The profile of stained cells (black) was compared to cells stained with the secondary antibody only (dotted line). Cells were exposed to CDM^{activin} conditions for ≥ 4 passages in all assays. Scale bars, 50 μ m.

The expression of undifferentiated hES cell markers like Oct4, Nanog, Sox2 and the specific surface markers SSEA4, Tra-1-60 and Tra-1-81 was highly and homogeneously detected in H1 CDM^{activin} cells, while SSEA1 (a differentiation marker) was absent (Figure 4.5c and d). Quantification by FACS analysis confirmed that 98.5% - 99.7% hES cells harboured an undifferentiation phenotype, while only 5.7% cells were positive for SSEA1.

As illustrated in Figure 4.6, H1 hES cells in CDM^{activin} are actively cycling, showing a slight increase in the percentage of cells in S and G2 phases of the cell cycle (45.1% and 26.8%, respectively) as compared to CM cells (40.8% and 19.4%, respectively).

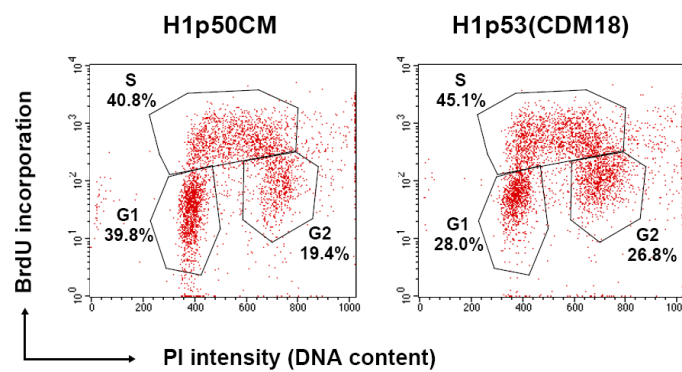


Figure 4.6. Human ES cells in CDM^{activin} conditions show a slight increase in the percentage of cells in S and G2 compared to cells grown under standard CM conditions. Bivariate distribution of H1 cells grown under CM (left) versus CDM^{activin} (right) conditions, showing incorporation of 5-bromo-2-deoxyuridine (BrdU) and propidium iodide (PI) staining intensity reflecting the DNA content. Numbers indicate percentages of cells in G1/S/G2. These findings are representative of two independent experiments.

4.2.2. H1 CM and CDM^{activin} form embryoid bodies containing tri-germ layer differentiation

Next, I tested the ability of H1 hES cells in CM and CDM^{activin} conditions to generate derivatives of all three germ layers using an *in vitro* embryoid body (EB) assay of development. EB formation is a simple and widely used method in which aggregates of pluripotent cells initiate a differentiation program that is reminiscent of early embryonic development (Doetschman et al., 1985; Leahy et al., 1999). In the context of the EB, molecular interactions that drive early embryonic development are recapitulated and cells differentiate to form ectoderm, endoderm, and mesoderm derivatives. Here, *in vitro* differentiation was achieved by growing hES cell colonies as EBs in non-adherent conditions for 7 days, then plating EBs back onto gelatin-coated plates and growing them for additional 23 days (Figure 4.7a).

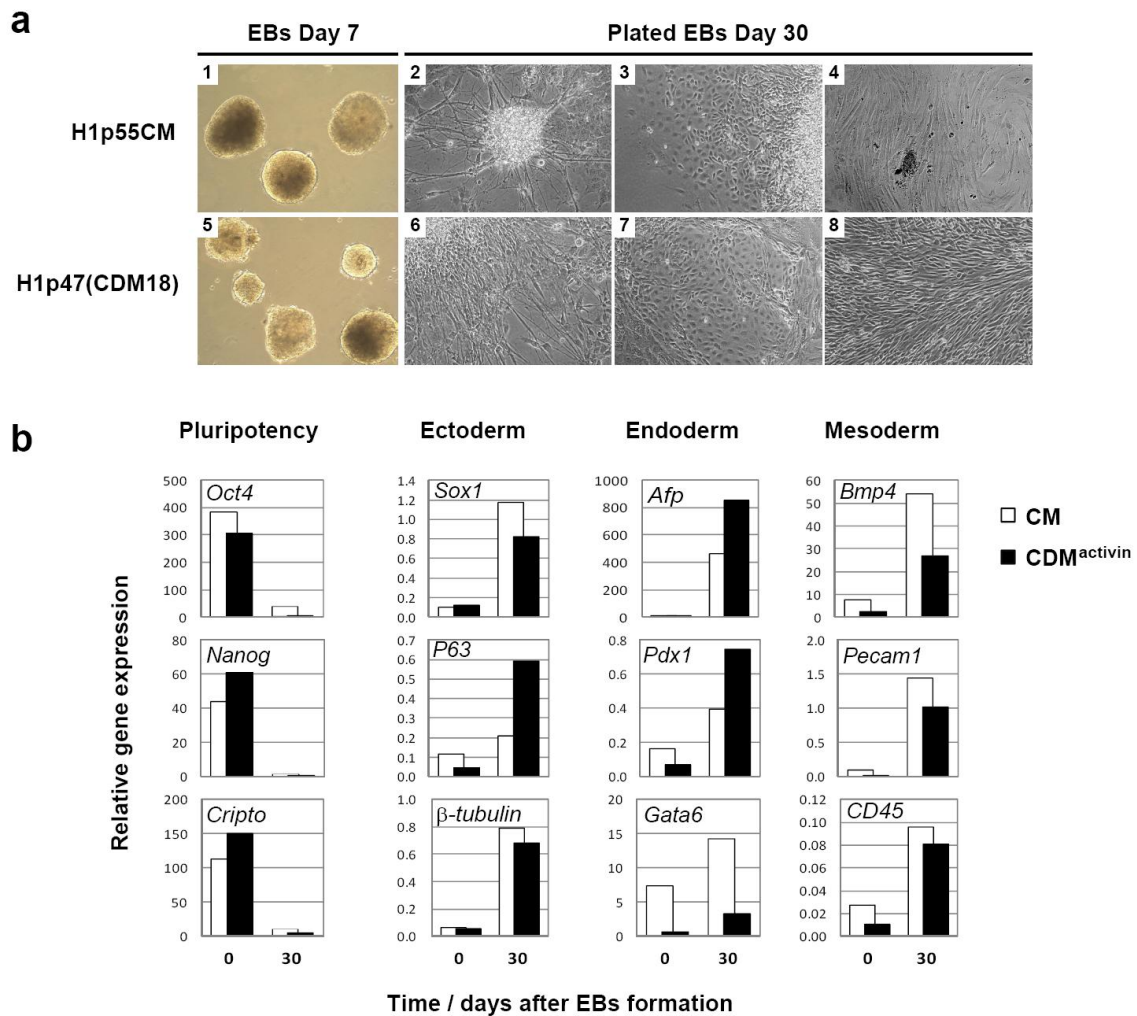


Figure 4.7. Human ES cells in CM and CDM^{activin} conditions are able to differentiate *in vitro* into derivatives of the three germ layers. (a) Morphology of EBs formed in non-adherent conditions for 7 days as well as cell types derived from EBs re-plated onto gelatine-coated surfaces for 23 days. After this period of time a wide variety of morphological cell types could be observed in the plates of both H1 CM and CDM^{activin}: neuronal (A2)-(A6), liver (A3)-(A7) and endothelial-like cells (A4)-(A8). Magnification, 100X. (b) Generation of derivatives of the three germ layers was further confirmed by qPCR analysis looking at *Sox1*, *P63*, β -*tubulin* (ectoderm), *Afp*, *Pdx1*, *Gata6* (endoderm), *Bmp4*, *Pecam1* and *CD45* (mesoderm) in both CM (white bars) and CDM^{activin} (black bars) cells. Expression of the pluripotency-associated markers *Oct4*, *Nanog* and *Cripto* is downregulated and almost absent after 30 days of differentiation. Graphs are a representative qPCR analysis of three independent experiments. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*).

At this stage, a wide variety of morphologically different cell types could be observed in the plates of both H1 CM and CDM^{activin} cells, likely indicating the presence of neuronal (Figure 4.7, panels 2 and 6), liver (Figure 4.7, panels 3 and 7) and endothelial-like cells (Figure 4.7, panels 4 and 8). Generation of derivatives of the three germ layers after long term differentiation (30 days) was further confirmed by qPCR analysis for the differentiation markers *Sox1*, *P63*, β -*tubulin* (ectoderm), *Afp*, *Pdx1*, *Gata6* (endoderm), *Bmp4*, *Pecam1* and *CD45* (mesoderm); all of these markers were steadily up-regulated in both CM and CDM^{activin} cells (Figure 4.7b). Expression of the pluripotency-associated

markers *Oct4*, *Nanog* and *Cripto* was, in contrast, dramatically downregulated or absent confirming that both CM and CDM^{activin} cells had lost their pluripotent status in culture conditions favoring differentiation. Altogether, these data demonstrate that both CM and CDM populations of hES cells can undergo differentiation into the three germ layers upon EBs formation and prolonged cultures.

4.2.3. H1 hES cells in CM and CDM^{activin} conditions have different kinetics of *in vitro* differentiation

Our replication timing studies above suggest that hES CM and CDM^{activin} cell cultures might represent distinct but interchangeable developmental states, with CM cells sharing features of primed (epiblast-like) cells for differentiation, as opposed to CDM^{activin} cells. To further test this hypothesis, I compared the early kinetic of differentiation of CM and CDM^{activin} H1 hES cells upon spontaneous differentiation (EBs) and lineage-specific induction (neuroectoderm) using a previous established protocol (Chambers et al., 2009). As shown in Figure 4.8 EBs derived from CM and CDM^{activin} cells were collected at four time points: undifferentiated hES cells - day 0 and EBs at culture day 3, 6 and 9 and analysed by qPCR.

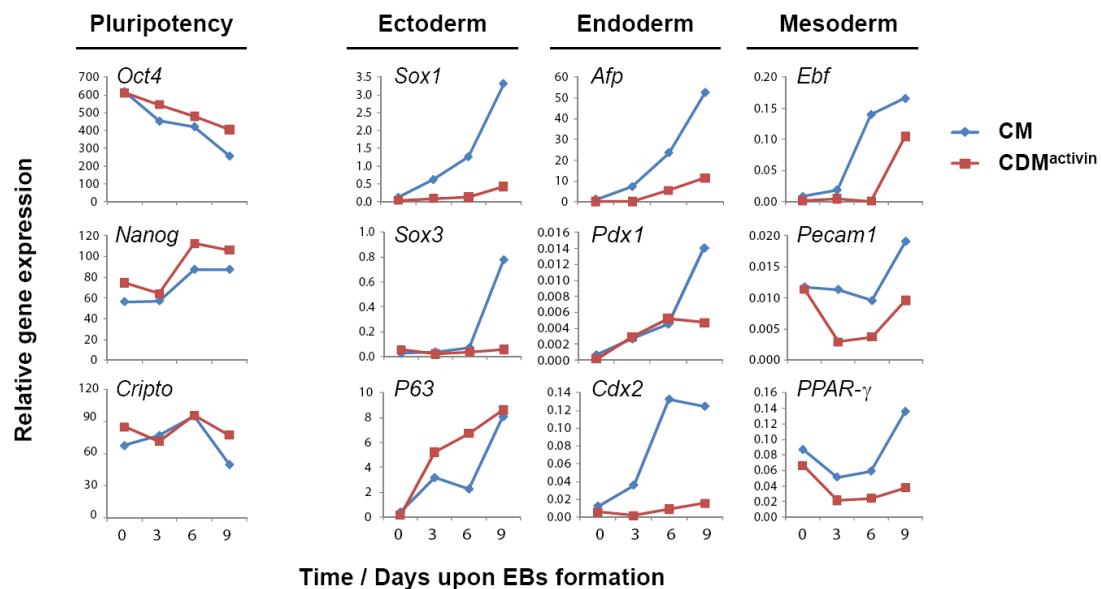


Figure 4.8. Human ES cells cultured in CDM^{activin} differentiate more slowly by EBs differentiation than hES cells cultured in CM. Time course analysis of the expression of pluripotency-associated (*Oct4*, *Nanog* and *Cripto*), ectoderm (*Sox1*, *Sox3* and *P63*), endoderm (*Afp*, *Pdx1* and *Cdx2*) and mesoderm (*Ebf*, *Pecam1* and *PPAR-γ*) markers, during an 9 day course of EB differentiation in H1 CM (H1p63CM; blue curves) and CDM^{activin} cells [H1p47(CDM19); red curves]. Graphs show a representative qPCR analysis of three independent experiments. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*).

The expression of markers associated with the undifferentiated state (*Oct4*, *Nanog* and *Cripto*) was similar in both H1 CM and CDM^{activin} EBs culture. As previously reported (Darr et al., 2006), *Nanog* and *Cripto* expression was transiently upregulated; *Oct4* expression was steadily downregulated in both cell populations. In contrast, we observe clear differences in the onset of ectoderm (*Sox1*, *Sox3* and *P63*), endoderm (*Afp*, *Pdx1* and *Cdx2*) and mesoderm (*Ebf*, *Pecam1* and *PPAR-γ*) associated markers, with an earlier induction of nearly all genes analysed in CM-derived EBs (Figure 4.8). Consistent trends were also observed when culturing CM and CDM^{activin} H1 hES cells in N2B27 media supplemented with 10 μM SB4341542 and 500 ng/ml noggin to induce neuroectoderm differentiation, as assessed at days 0, 2, 4 and 6 upon treatment (Figure 4.9).

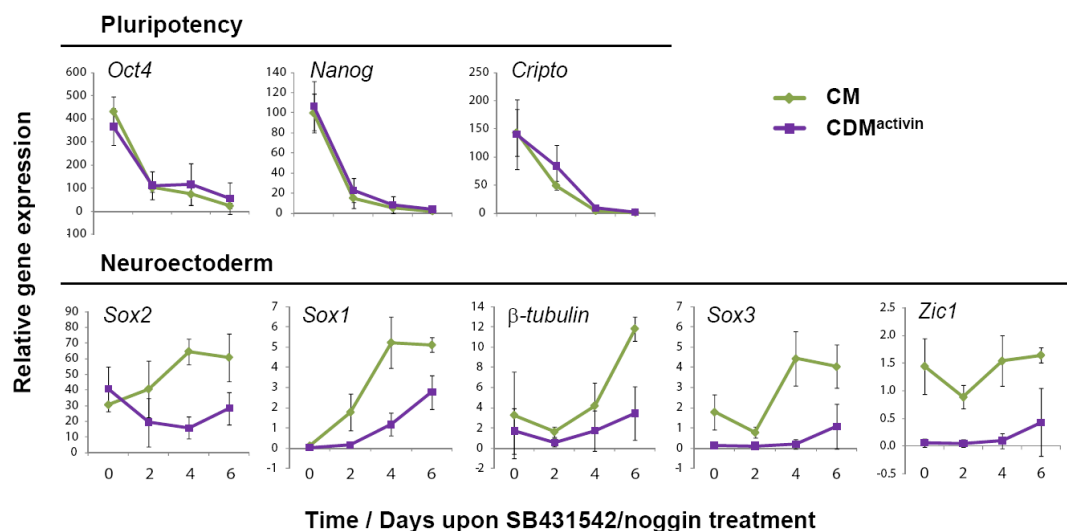


Figure 4.9. Human ES cells cultured in CDM^{activin} differentiate more slowly into neuroectoderm derivatives than hES cells cultured in CM. Time course analysis of the expression of pluripotency-associated (*Oct4*, *Nanog* and *Cripto*) and neuroectoderm (*Sox2*, *Sox1*, *β-tubulin*, *Sox3* and *Zic1*) markers, during a 6 day course of neuroectoderm differentiation as previously published (Chambers et al., 2009) in H1 CM (H1p65CM; green curves) and CDM^{activin} cells [H1p58(CDM16); purple curves]. Neuroectoderm differentiation was performed in N2B27 medium supplemented with 10 μM SB4341542 and 500 ng/ml noggin. Mean and standard deviation of three experiments are shown. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*).

The kinetic of *Oct4*, *Nanog* and *Cripto* downregulation was remarkably similar in induced H1 CM and CDM^{activin} cells. In contrast, the expression of neuroectoderm-specific genes, *Sox2*, *Sox1*, *β-tubulin*, *Sox3* and *Zic1* was enhanced in H1 CM as compared to CDM^{activin} cultures (Figure 4.9). Taken together, these data demonstrate that CM and CDM^{activin} hES cells possess functionally distinct properties, with CM cells being more sensitive to the signal inducer of differentiation as tested in these experiments.

4.2.4 CDM^{activin} conditions tip the balance towards a more stable, ICM-like cell phenotype within hES cell cultures

To further explore the similarities and differences among hES cells grown in CM and CDM^{activin} conditions, we next examined the transcription profiles of undifferentiated H1 CM and CDM^{activin} cultures as assessed at different passage numbers (3, 6, 14 and over 25 passages) post transfer (CM to CDM^{activin}) by qPCR analysis (Figure 4.10).

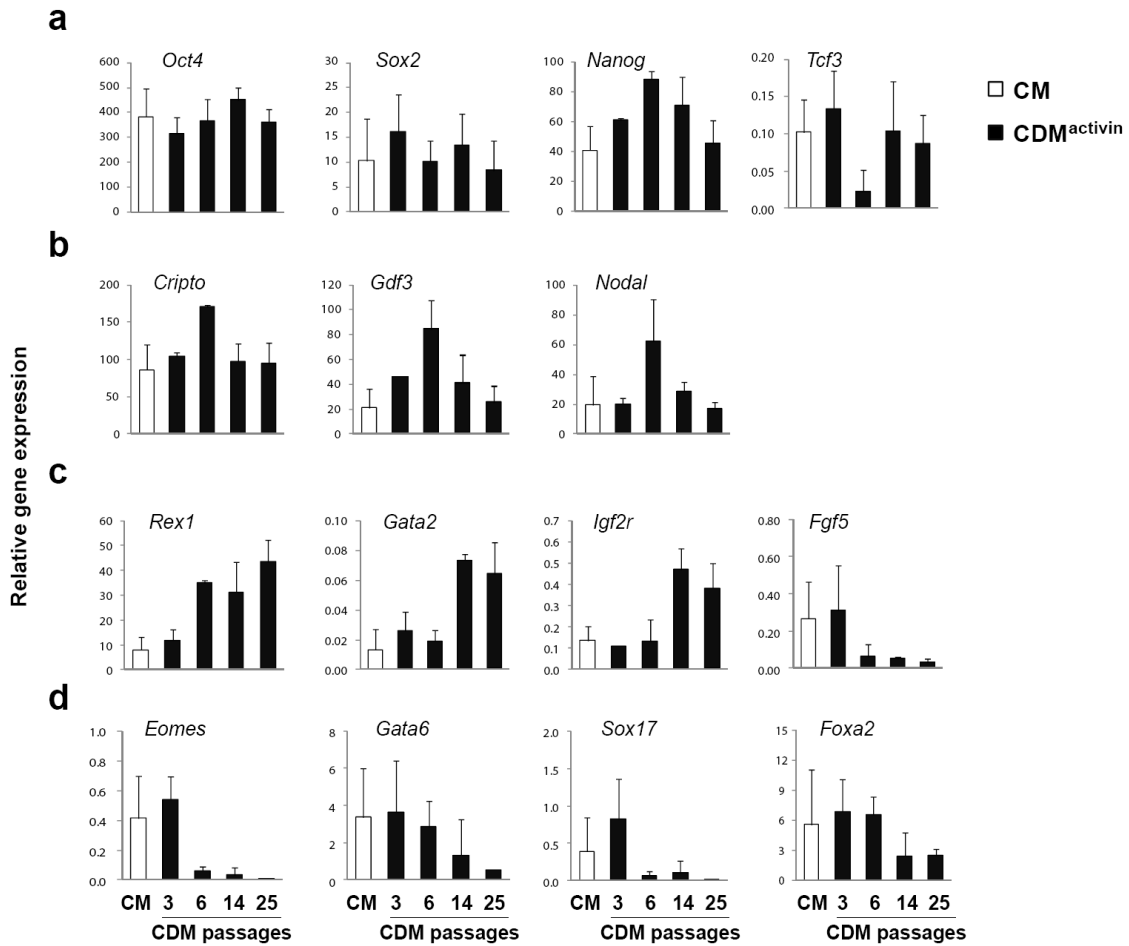


Figure 4.10. Differential gene expression profiles in H1 CM and CDM^{activin} hES cell populations. Real-time PCR analysis of selected genes in H1 CM (H1p49CM, white bars) and cells transferred and cultured for different passage numbers in CDM^{activin} conditions [H1p47(CDM3, 6, 14 and 25), black bars]. (a) Pluripotency-associated genes (*Oct4* and *Sox2*) stably expressed in CM and CDM^{activin} cells or showing a transient increase (*Nanog*) or downregulation (*Tcf3*) in CDM^{activin} cells (b) TGF-β superfamily members transiently increased (*Cripto*, *Gdf3* and *Nodal*) in CDM^{activin} cells. (c) CDM^{activin} strongly induced in H1 hES cells in CDM^{activin} conditions the expression of preimplantation-ICM markers (*Rex1*, *Gata2* and *Igf2r*) and downregulation of the postimplantation marker *Fgf5*, as well as the (d) downregulation of germ layer markers (*Eomes*, *Gata6*, *Sox17* and *Foxa2*). Mean and standard deviation of three experiments are shown. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*).

Here, we found that the expression levels of *Oct4* and *Sox2* remained constant across cell populations analysed (Figure 4.10a). In contrast, *Nanog* expression was transiently increased upon transfer to CDM^{activin} conditions; a trend also observed for three members of the TGF- β superfamily, *Cripto* (Minichiotti, 2005), *Gdf3* (Levine and Brivanlou, 2006) and *Nodal* (Besser, 2004) (Figure 4.10b). Transient increase in *Nanog* expression was further confirmed at the protein level by FACS analysis. As shown in Figure 4.11, we observed a clear shift in the mean fluorescence intensity of *Nanog* staining in early CDM^{activin} cultures as opposed to CM and long-term CDM cultures. Note that, *Oct4* levels remained steady in all three cell populations. Interestingly, the expression of *Tcf3* was, in contrast, transiently downregulated and thus showed a converse pattern with that of *Nanog*; an observation consistent with *Tcf3* being a putative repressor of *Nanog* as described in mES cells (Pereira et al., 2006; Yi et al., 2008).

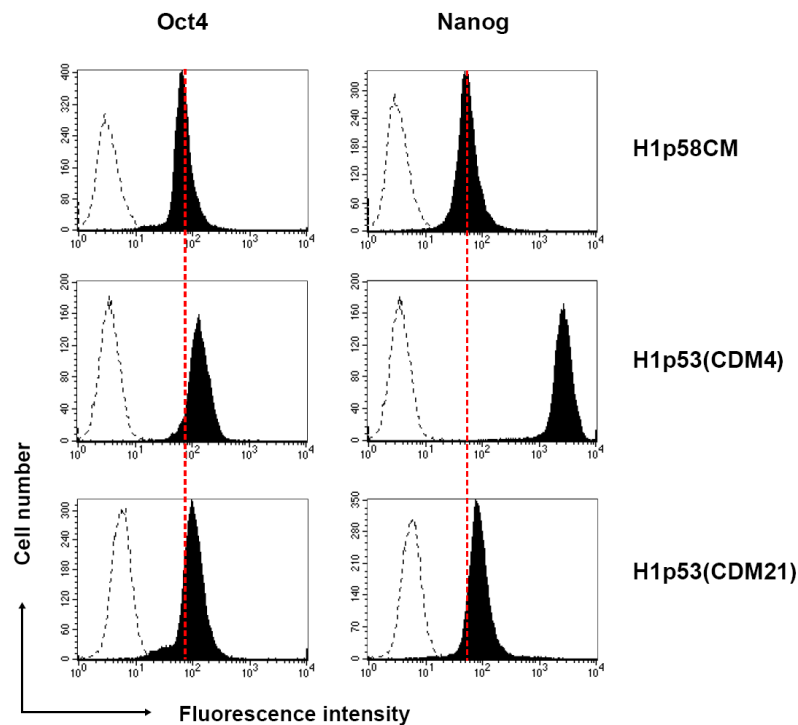


Figure 4.11. Transient increase in the expression levels of *Nanog* in H1 CDM^{activin} hES cells. Flow cytometry analysis of *Oct4* (left panel) and *Nanog* (right panel) levels in H1 CM (H1p58CM) and CDM^{activin} early [H1p53(CDM4)] and long term cultures of CDM^{activin} [H1p53(CDM21)]. The red dotted lines indicate the mean fluorescence intensity of *Oct4* and *Nanog* in H1p58CM cells. The profile of stained cells (black) was compared to cells stained with the secondary antibody only (dotted line). These findings are representative of two independent experiments.

Remarkably, CDM^{activin} conditions strongly and rapidly enhanced the expression of *Rex1* (Figure 4.10c), a landmark of pluripotency that is specifically expressed in the ICM of mouse and human blastocysts, but downregulated in the epiblast stage

(Appendix II, Figure A5) (Pelton et al., 2002; Reijo Pera et al., 2009; Rogers et al., 1991). A strong bias toward an ICM-like phenotype in our CDM^{activin} cultures was further confirmed by the steady induction of two other human ICM-specific markers, *Gata2* and *Igf2r* (Reijo Pera et al., 2009) and the extinction of *Fgf5* – a marker expressed in mEpiS cells but not in mES cells (Appendix II, Figure A5) (Pelton et al., 2002; Tesar et al., 2007). Similarly, transcripts of *Eomes*, *Gata6*, *Sox17* and *Foxa2* which are upregulated in epiblast and early germ layers (Appendix II, Figure A5) (Pelton et al., 2002; Tesar et al., 2007), were dramatically downregulated upon transfer to CDM^{activin} conditions, consistent with a global decrease of transcriptional noise. Altogether these data strongly suggest that hES cells can transit from a primed epiblast-like to a more naïve ICM-like state by changing growth culture conditions (CM to CDM^{activin}), a level of heterogeneity not previously described within normal hES cell cultures.

4.3. Generation of inter-species heterokaryons by cellular fusion

In normal development, cells transit in a unidirectional process from totipotent zygote, pluripotent ICM/epiblast progenitors to fully differentiated somatic cells. However, under certain conditions, cell identity can be modified or reversed, leading to the establishment of a novel gene expression program defined as a reprogramming event. Reprogramming towards pluripotency can be achieved in different ways including nuclear transfer, forced expression of one or more transcription factors (iPS cells) or by cellular fusion (Egli et al., 2008; Jaenisch and Young, 2008). Spontaneous and experimental cell fusion between ES and somatic cells has shown that the phenotype of ES cells is dominant, imposing an ES-specific gene expression pattern in the nuclei of more differentiated cell types (Tada et al., 2001; Ying et al., 2002). Other stem cell populations such as EG and EC cells have also been described to reprogram somatic cells (Miller and Ruddle, 1976; Tada et al., 1997) towards their own identity *in vitro* as well as other extraembryonic stem cell populations TS and XEN cells (J. Santos and F. Pereira, unpublished data). Here we asked whether hES cells in CM and CDM^{activin} conditions share with mES cells the potential to reprogram the gene expression program of somatic nuclei. Also, we further compared the reprogramming ability of CDM^{activin} versus CM hES cell cultures alongside mES and mEpiS cells, using a cell fusion based assay (Pereira and Fisher, 2009; Pereira et al., 2008).

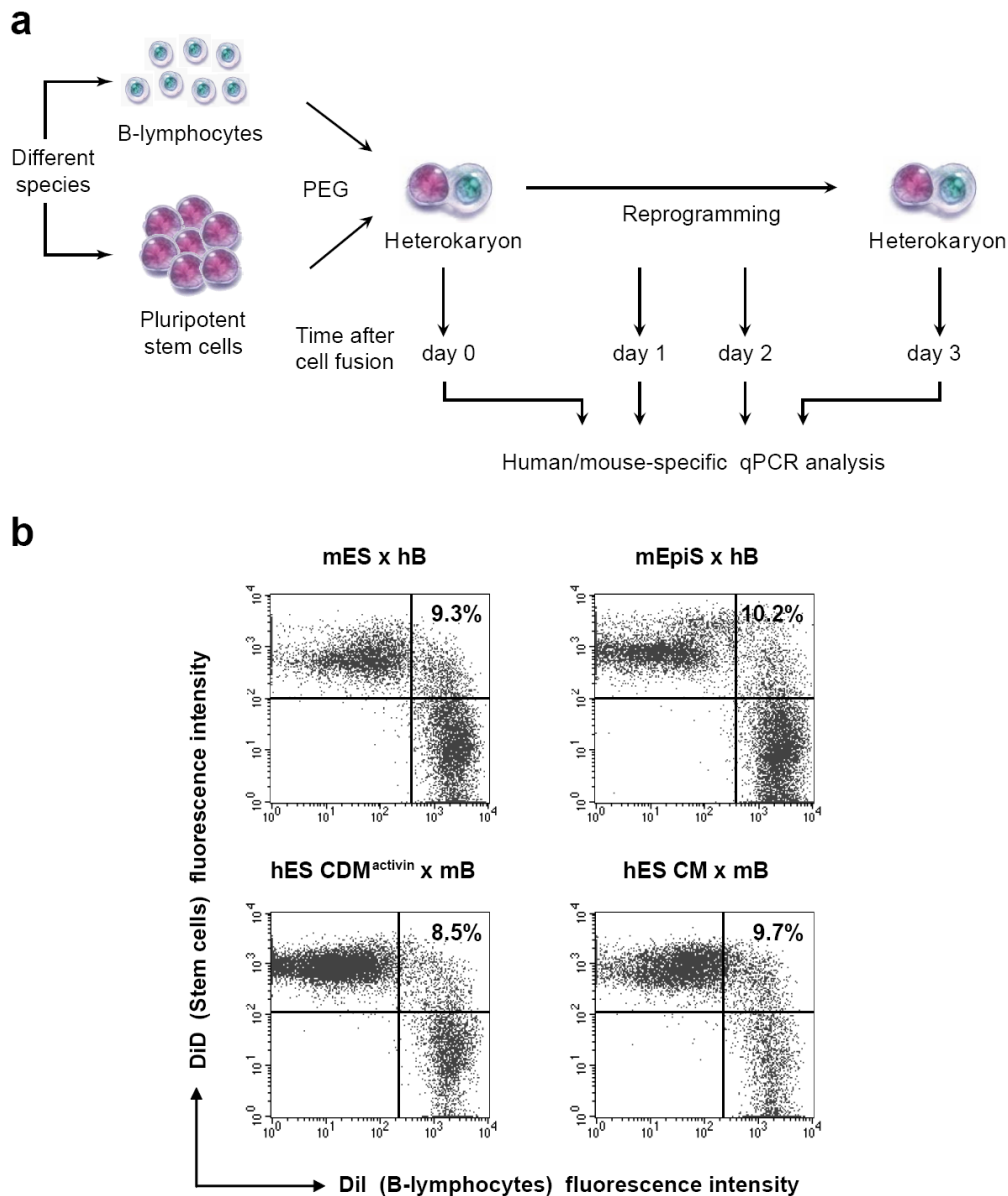


Figure 4.12. Generation of inter-species heterokaryons by cell fusion. (a) Schematic representation (adapted from Pereira et al., 2008) of the experimental strategy used to generate inter-species heterokaryons between B-lymphocytes and pluripotent stem cell lines derived from different species (human versus mouse and *vice versa*). Cell partners were fused in the presence of polyethylene glycol (PEG) and subsequent cells cultured up to three days under cell culture conditions promoting an undifferentiated state. Heterokaryons collected at different time points were analyzed by qPCR using primers that specifically amplify human or mouse transcripts. (b) Cell fusion efficiency was determined by FACS analysis. B-lymphocytes (hB and mB) were labelled with the cell membrane dye Dil and human [H1p70CM and H1p47(CDM30)] and mouse (Oct4Gip mES p31 and Oct4Gip mEpiS p32) stem cell populations with DiD. The resulting labelled cell populations were then fused and analyzed by FACS. Fusion efficiency evaluated as the percentage of double-labelled cells is indicated in each plot (upper right quadrant). These experiments were done in collaboration with Filipe Pereira.

As outlined in Figure 4.12a, pluripotent stem cells and B-lymphocytes from different species (human *versus* mouse and *vice versa*) were mixed in a 1:1 ratio, giving rise to inter-species heterokaryons (in which parental nuclei share the same cytoplasm but remain separated) by PEG-mediated cell fusion. Fused cells were

cultured under conditions promoting the maintenance of undifferentiated stem cells and heterokaryons collected at day 0 (no PEG), 1, 2 and 3 prior nuclei fusion, for subsequent qPCR analysis, using specific-species primers. Cell fusion efficiency was verified by FACS analysis, based on different cell membrane labelling of B-lymphocytes (DiI) and stem cells (DiD) prior to fusion. Figure 4.12b shows the FACS profiles obtained and the percentage of double-labelled cells (fusion efficiency) from mES x hB (upper left panel; 9.3%), mEpiS x hB (upper right panel; 10.2%), hES CDM^{activin} x mB (lower left panel; 8.5%) and hES CM x mB (lower right panel; 9.7%) fusions. These results indicate that all the population of cells analyzed (mouse and human) are able to be successfully fused with B-lymphocytes at similar efficiency rates.

4.3.1. Mouse EpiS cells are not able to reprogram human B-lymphocytes for pluripotency in contrast to mES cells

A panel of hES-associated genes (*Oct4*, *Nanog*, *Rex1*, *Cripto*, *Tert*, *Tle1* and *Dnmt3b*) was selected for this analysis. The kinetics of gene re-activation potentially induced at these loci upon heterokaryon formation was assessed at day 0 (before fusion), 1, 2 and 3 by qPCR using species-specific primers. Human *Hprt* housekeeping gene was also included in these experiments as internal control, showing consistent and robust expression levels in both mES x hB and mEpiS x hB fusion experiments (Figure 4.13a). As previously published, mES cells were able to reprogram hB lymphocytes for pluripotency as indicated by the rapid re-activation of human *Oct4*, *Nanog*, *Cripto*, *Tle1* and *Dnmt3b* in hB cell nuclei 2 days after heterokaryon formation and subsequent induction of *hRex1* and *hTert* one day later. Conversely, we observed a rapid extinction of *hCD19* expression – a specific lymphocyte marker, further indicating a loss of hB cell identity in fused somatic cells and the acquisition of a hES cell-like phenotype upon heterokaryon formation.

Surprisingly and in contrast to mES cells, when pluripotent mEpiS cells were used as fusion partners, *hOct4* and *hNanog* expression was not induced in hB cell nuclei as well as other hES-associated markers analysed. A modest induction of *hTle1* and *hDnmt3b* was seen at day 2 yet disappeared by day 3 (Figure 4.13a). Note that *Fgf5*, a putative human epiblast marker was neither induced in our mEpiS x hB (or mES x hB) fusion experiments (data not included in Figure 4.13). Notably, *CD19* expression remained relatively steady in mEpiS x hB as compared to mES x hB heterokaryons, further indicating that reprogramming may have failed to occur in hB cell nuclei fused with mEpiS cells.

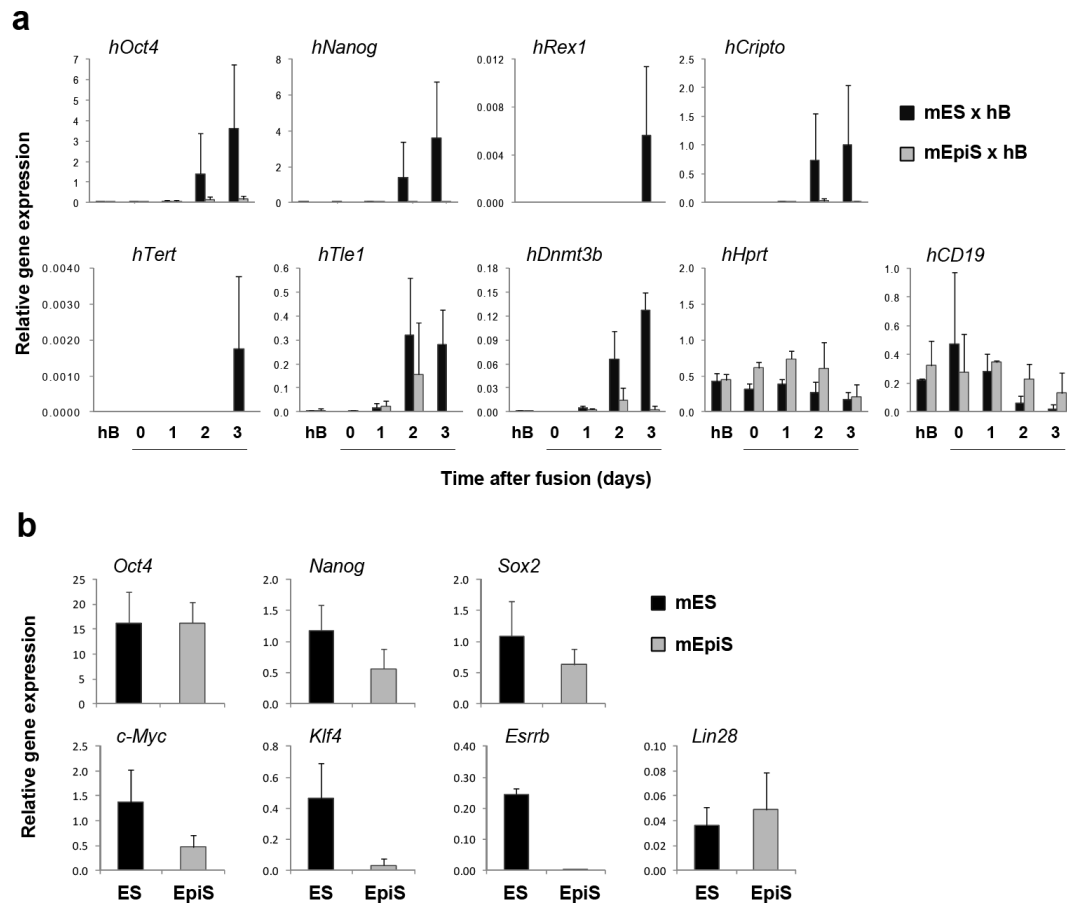


Figure 4.13. Mouse EpiS cells are not able to reprogram human B-lymphocytes for pluripotency by cellular fusion in contrast to mES cells. (a) Oct4Gip mES p31 (black bars) and Oct4Gip mEpiS p32 (grey bars) cell lines were fused to human B-lymphocytes (hB). The expression of hES-specific genes (*hOct4*, *hNanog*, *hRex1*, *hCripto*, *hTert*, *hTle1* and *hDnmt3b*) was assessed by qPCR analysis between day 0 and day 3 after cell fusion using human-specific primers. Human B-lymphocytes were included as a negative control for this analysis. The constitutively expressed gene *hHprt* remains detectable at similar levels at all time points in both mES and mEpiS fusions. Activation of embryonic genes in mES x hB fusions is accompanied by silencing of the lymphocyte-specific gene (*hCD19*). Data were normalised to *hGapdh* expression. Error bars indicate the standard deviation of three independent experiments. (b) Real-time PCR analysis of gene expression levels of reprogramming factors (*Oct4*, *Nanog*, *Sox2*, *c-Myc*, *Klf4*, *Esrrb* and *Lin28*) in Oct4Gip mES (black bars) and Oct4Gip mEpiS (grey bars) cells. Mean and standard deviation of three independent experiments are shown. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*). These experiments were performed in collaboration with Filipe Pereira.

In an attempt to explain the differences observed between mES and mEpiS cells in terms of reprogramming potential, we compared the expression levels of known reprogramming factors - Oct4, Sox2, c-Myc, Klf4 (Takahashi and Yamanaka, 2006), Esrrb (Feng et al., 2009a) Nanog and Lin28 (Yu et al., 2007) as assessed in both cell populations by qPCR (Figure 4.13b). As anticipated, *Oct4* was expressed at similar levels in mES and mEpiS cells, while *Nanog* and *Sox2* showed a reduced mRNA levels in mEpiS cells as compared to mES cells, consistent with previous reports (Brons et al., 2007; Tesar et al., 2007). Interestingly, mEpiS cells displayed a drastic reduction in the expression of *c-Myc*, *Klf4* and *Esrrb* (more than 2, 15 and 100-fold, respectively)

compared to mES cells (Figure 4.13b). No significant change in *Lin28* expression patterns was noticed in these experiments. Altogether these data demonstrate fundamental differences between pluripotent mouse stem cells derived from the early embryo. In contrast to mES cells, we found that mEpiS cells derived from gastrulating embryos were unable to reprogram the fate of hB-lymphocytes upon cell fusion, possibly reflecting the loss of key reprogramming factors expression in this cell type.

4.3.2. Both H1 CM and CDM^{activin} hES cells are able to reprogram mouse B-lymphocytes for pluripotency

The same methodology was used to generate heterokaryons between H1 hES cells grown in either CM or CDM^{activin} conditions and mouse B-lymphocytes. Reprogramming events were assessed by qPCR analysis of the mES-associated genes, *Oct4*, *Nanog*, *Sox2*, *Rex1*, *Cripto*, *Tert* and *Dnmt3b*. Mouse *Hprt* was also included in these experiments as internal control (Figure 4.14a). Here we found that both hES cell populations analysed were able to induce the reactivation of m*Oct4* in B-lymphocyte nuclei upon cell fusion (Figure 4.14a). However, a more detailed comparison between hES CM and hES CDM^{activin} x mB fusions revealed clear differences in terms of genes being reactivated in each cell fusion type. Strikingly, m*Nanog* was not reactivated in CM-type fusion in contrast to CDM^{activin}-type fusion. m*Sox2* and m*Rex1* showed a converse pattern, being strongly enhanced in B-cell nuclei fused with CM cells but not with CDM^{activin} cells. The expression of mouse *Oct4*, *Cripto*, *Tert* and *Dnmt3b* was detected at similar levels by day 3, although differences in the kinetic of gene reactivation were observed. Such differences could not be simply explained by differential *Nanog* or *Sox2* mRNA levels in the two hES cell populations analysed (Figure 4.14b). Note that long-term CDM^{activin} hES cell cultures were solely used for these fusion experiments. Similarly, we did not notice any significant difference in the expression patterns of other key reprogramming factors (*c-Myc*, *Klf4*, *Esrrb* and *Lin28*) between CM and CDM^{activin} hES cell cultures (Figure 4.14b). Note that we were not able to accurately verify whether the expression of the mB-lymphocyte-associated CD19 gene was shutdown upon reprogramming, as anticipated, in the absence of selective drugs to eliminate unfused mB-lymphocytes in these experiments (see section 2.2.11).

Collectively, these results suggest that hES cells share with mES cells the potential to reprogram somatic cells for pluripotency upon heterokaryon formation. This contrast with mEpiS cells that were found unable to reprogram hB cell nuclei within 3

days post cell fusion as tested in our experiments. Interestingly, we noticed clear differences in the set of pluripotency-associated genes reactivated in mB cell nuclei when fused with either CM or CDM^{activin} hES cells, further highlighting distinct properties between these two hES cell populations.

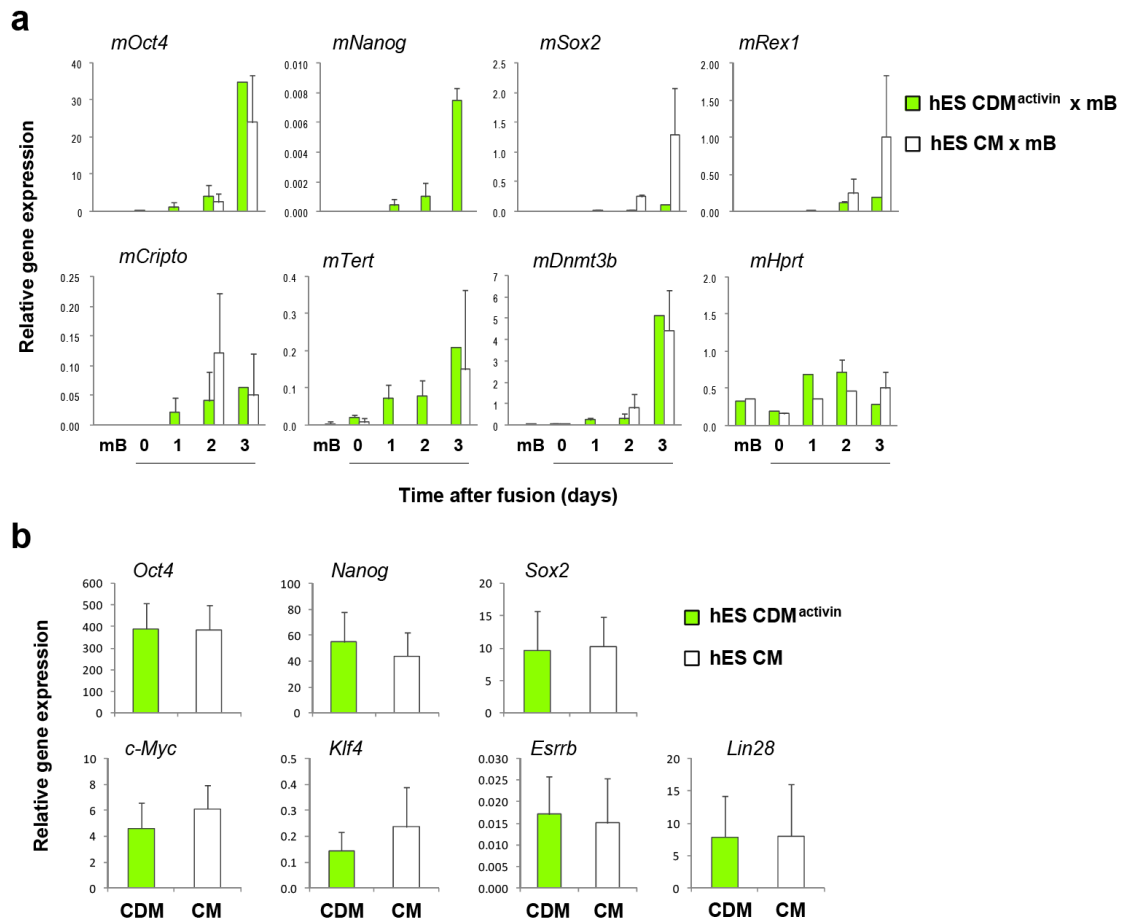


Figure 4.14. Both H1 CM and CDM^{activin} hES cells are able to reprogram mouse B-lymphocytes for pluripotency by cellular fusion. (a) H1p70CM (white bars) and H1p47(CDM30) (green bars) hES cells were fused to mouse B-lymphocytes (mB). The expression of mES-specific genes (*mOct4*, *mNanog*, *mSox2*, *mRex1*, *mCripto*, *mTert* and *mDnmt3b*) was assessed by qPCR analysis between day 0 and day 3 after cell fusion, using mouse-specific primers. Mouse B-lymphocytes were included as a negative control for this analysis. The constitutively expressed *mHprt* gene remains detectable at similar levels at all time points in both hES CM and hES CDM^{activin} fusions. Data were normalized to *mGapdh* expression. Error bars indicate the standard deviation of at least two independent experiments. **(b)** Real-time PCR analysis of gene expression levels of reprogramming factors (*Oct4*, *Nanog*, *Sox2*, *c-Myc*, *Klf4*, *Esrrb* and *Lin28*) in H1p70CM (white bars) and H1p47(CDM30) (green bars) hES cells. Mean and standard deviation of three independent experiments are shown. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*).

CHAPTER 5. EPIGENETIC AND MOLECULAR MECHANISMS UNDERLYING DISTINCT PLURIPOTENT STATES WITHIN HUMAN ES CELLS

5.1. A possible role for TGF- β /Activin signalling in defining a hES cell naive *versus* primed state in culture

The mechanisms controlling self-renewal *versus* lineage commitment in hES cells remain poorly understood. Nonetheless, several studies suggest that TGF- β superfamily ligands are involved in sustaining an undifferentiated and self-renewing hES cell state. In particular, the TGF- β /Activin signaling through its signal inducer Smad2/3 was proven to be essential for maintaining pluripotency in hES cells; while its inhibition leads to the upregulation of differentiation markers (James et al., 2005; Vallier et al., 2005; Xiao et al., 2006; Xu et al., 2008). Interestingly, we found that the acquisition of a more stable early-replicating epigenetic state in pluripotent CDM^{activin} hES cell cultures was accompanied by a transient upregulation of three key TGF- β -associated members, *Cripto*, *Gdf3* and *Nodal*, alongside *Nanog* (Figure 4.10).

To further explore whether differential levels of TGF- β /Activin signaling activation might directly influence the balance between a more naive and primed state for differentiation in hES cells, we compared the relative levels of phosphorylated Smad2/3 (P-Smad2/3) in CM and CDM^{activin} H1 cultures from different passage numbers (3, 5, 14 and over 25), by western blot analysis. As shown in Figure 5.1, clear differences in the levels of P-Smad2/3 detected in CM and CDM^{activin} cell cultures were observed. Active P-Smad2/3 levels become consistently higher in hES cells when transferred from CM to CDM^{activin} conditions. This increase is also mirrored by an increase in Nanog induction in CDM^{activin} cell cultures. This is consistent with a direct role for Smad2/3 in regulating Nanog expression in hES cells (Xu et al., 2008). While Nanog was only transiently upregulated, P-Smad2/3 levels remained steady in CDM^{activin} cultures, regardless of passage numbers. Note that Oct4 levels were kept constant across CM and CDM^{activin} cultures, indicating that both CM and CDM^{activin} hES cell populations retain an undifferentiated state, as previously verified (Figure 4.5 and 5.1). Interestingly, Nanog was also recently shown to interact directly with the Smad2/3

proteins, thus modulating the activity of TGF- β /Activin signaling pathway (Vallier et al., 2009a). This negative feedback loop, may in turn, regulate and perhaps reduce the levels of Nanog expression in our long-term CDM^{activin} cultures to levels seen in CM cultures. Unfortunately, our attempt to quantify the levels of Smad2/3 occupancy at the Nanog promoter in CM and early and late CDM^{activin} cultures by ChiP assay were unsuccessful and did not allowed verifying this hypothesis.

Altogether these results indicate that the activity of TGF- β /Activin signaling pathway can be dynamically altered in response to CDM^{activin} growth conditions, highlighting a possible role in modulating a self-renewing versus primed state within hES cell cultures.

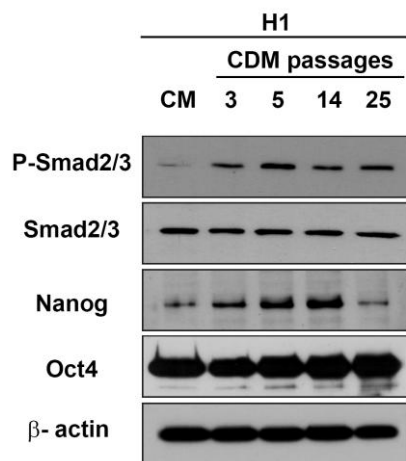


Figure 5.1. H1 CDM^{activin} hES cells are characterized by higher levels of activated Smad2/3 complexes in comparison to H1 CM cell cultures. Western blot analysis of H1 hES cells cultured in CM (H1p49CM) and CDM^{activin} for different passage numbers [H1p47(CDM3, 5, 14 and 25)]. Membranes were probed with antibodies specific for phosphorylated (P) Smad2/3, Smad2/3, Nanog, Oct4 and β -actin (as a control for protein loading).

5.2. Shift from early to late replication is accompanied by an increased recruitment of active p300 at the promoter of target genes in CDM^{activin} cells

5.2.1. Global increase of p300-mediated histone H3 acetylation in CDM^{activin} hES cell cultures

Activated Smad complexes (in particular Smad2/3) are thought to alter the chromatin states of target loci by recruitment of various histone acetyltransferases (HATs), including p300/CBP (Ross et al., 2006; Ross and Hill, 2008), Pcaf (Itoh et al., 2000; Kahata et al., 2004) and Gcn5 (Kahata et al., 2004) as well as the chromatin remodeler SWI/SNF ATPase, Brg1 (Ross et al., 2006) in a TGF- β dependent manner. To further investigate a possible link between Smad2/3-associated chromatin

remodeling activity and the establishment of early-replicating domains in hES cells grown in CDM^{activin} conditions, we initially surveyed dynamic changes in gene expression levels of different HATs (*p300*, *CBP*, *Pcaf*, *Gcn5*) and *Brg1* that are potentially recruited by activated Smad2/3 complexes in H1 CM and CDM^{activin} hES cells from different passage numbers (Figure 5.2a). Interestingly, *p300* showed a robust increase of more than 3-fold in CDM^{activin} as compared to CM cells; a level of induction maintained in CDM^{activin} cultures upon serial passages (Figure 5.2a). This was accompanied by a transient induction of both *CBP* and *Brg1*, two co-factors commonly associated with p300. In contrast, *Pcaf* was expressed at very low levels in all cell populations analysed while *Gcn5* mRNA levels remained unchanged in both CM and CDM^{activin} culture conditions.

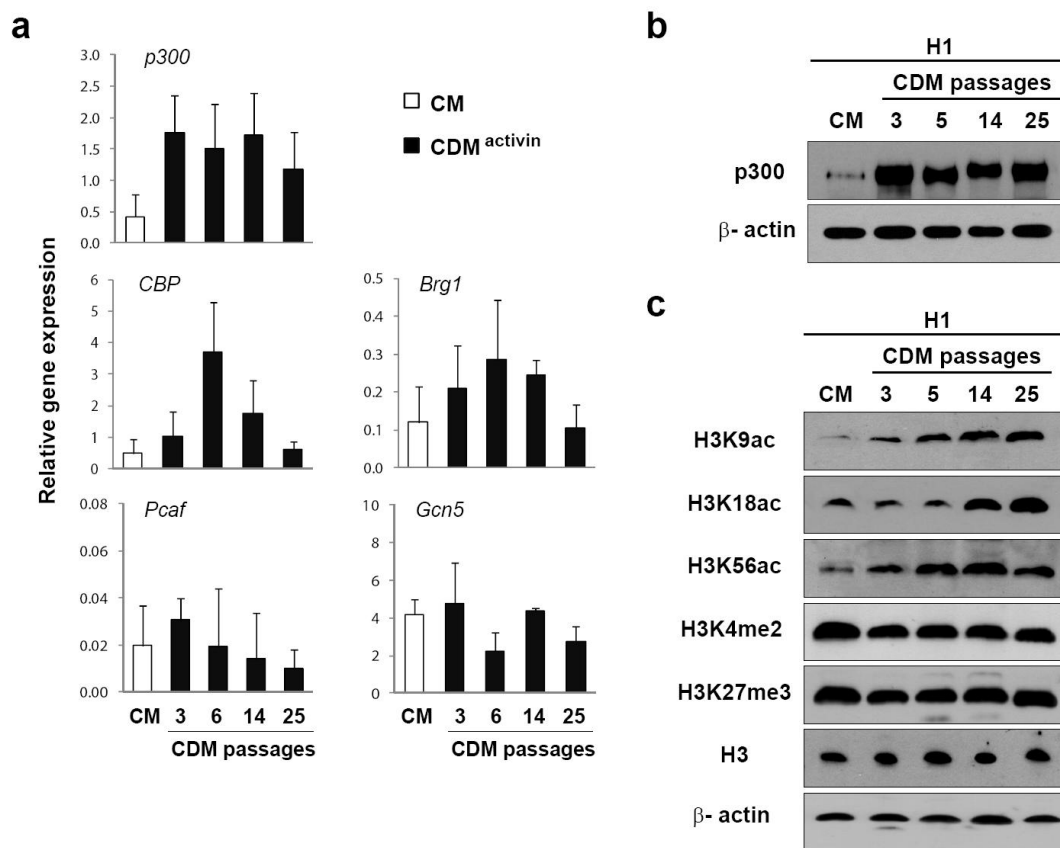


Figure 5.2. The histone acetyltransferase p300 is highly expressed in H1 CDM^{activin} hES cells in contrast to H1 CM cells. (a) Real-time PCR analysis of HATs recruited by activated Smad complexes, (*p300*, *CBP*, *Pcaf*, *Gcn5*) and *Brg1* in H1 CM (H1p49CM, white bars) and CDM^{activin} hES cells [H1p47(CDM3, 6, 14 and 25), black bars]. Mean and standard deviation of three experiments are shown. Data was normalized to the expression of two housekeeping genes (*Gapdh* and *Hprt*). (b) Western blot analysis of H1 hES cells cultured in CM (H1p49CM) and CDM^{activin} for different passage numbers [H1p47(CDM3, 5, 14 and 25). Membranes were probed with antibodies specific for p300 and β -actin (as a control for protein loading) or (c) with antibodies specific for p300 preferentially acetylated histone H3 targets (H3K9ac, H3K18ac and H3K56ac). Antibodies against two histone modifications independent of p300 activity - H3K4me2 and H3K27me3 were also used as controls. Equivalent protein loading is shown by the total H3 and β -actin detection.

The induction of p300 was further confirmed at the protein level by western blot analysis (Figure 5.2b) and shown to be accompanied by a global increase of its HAT activity (Figure 5.2c). Here we focused on p300 preferentially acetylated nucleosomal histone H3 (H3K9ac, H3K18ac and H3K56ac) (Das et al., 2009; Ross et al., 2006), showing a steady and gradual increase in histone acetylation levels in continued CDM^{activin} cultures as compared to H1 CM cells. In contrast, H3K4me2 and H3K27me3 levels, two histone modifications independent of p300 activity, remained constant between CM and CDM^{activin} cells. These results indicate that hES cells grown under CDM^{activin} conditions display high and steady levels of active Smad2/3 and the HAT p300, likely responsible for the global increase of histone H3 acetylation observed within these cultures.

5.2.2. Recruitment of active p300 complexes at target loci is enhanced in response to CDM^{activin} growth conditions

Histone acetylation has been shown to positively correlate with early replication timing (Mendez, 2009). To investigate whether p300 is directly recruited to loci changing replication timing, from late to early, CDM^{activin} hES cells as compared to CM cells, we examined by ChIP assay and qPCR analysis the levels of p300 occupancy at selected promoters as well as the levels of enrichment for associated acetylated histone H3 (H3K9ac, H3K18ac and H3K56ac) in H1 CM and CDM^{activin} hES cell cultures. As summarised in Figure 5.3a, *Cga*, *Foxa2*, *Gata6*, *Rex1*, *Sox17* and *Tert* are all targets for replication timing alterations (from late to early S-phase) in response to CDM^{activin} conditions. *Nanog* is here shown as an early-replicating control in both CM and CDM^{Activin} populations. Strikingly, and as seen in Figure 5.3b, p300 was consistently detected at higher levels in the promoter regions of these loci in CDM^{activin} hES cells, where they replicate in early S-phase as compared to CM cells. In contrast, no significant change in the levels of enrichment for p300 at *Nanog* promoter was noticed between the two cell populations. As anticipated from the presence of p300, the relative enrichment for acetylated H3K9, K18 and 56 at the promoter regions of *Cga*, *Foxa2*, *Gata6*, *Rex1* and *Sox17* was also found generally higher in H1 CDM^{activin} cells than in CM (Figure 5.3c). *Tert* did not however, show the same trend, yet replication timing was found to be highly variable (from middle and middle-early) across CM cultures of the same line, highlighting *Tert* as a complex exception. These studies clearly suggest a direct correlation between the establishment of early-replication profiling and p300 HAT recruitment and activity in hES cells regardless of the

transcriptional states of target loci – note that *Rex1* and *Tert* are expressed in hES cells while all other candidate loci analysed are silent. These results are in accordance with recent observations from Cedar and colleagues that HATs recruitment to the human β -globin locus in a non-erythroid cell line is sufficient to trigger a change of replication timing, from late to early. Conversely, tethering a histone deacetylase (HDAC) to this locus in an erythroid cell line (where its normally expressed) was sufficient to change its replication timing from early to late (Goren et al., 2008).

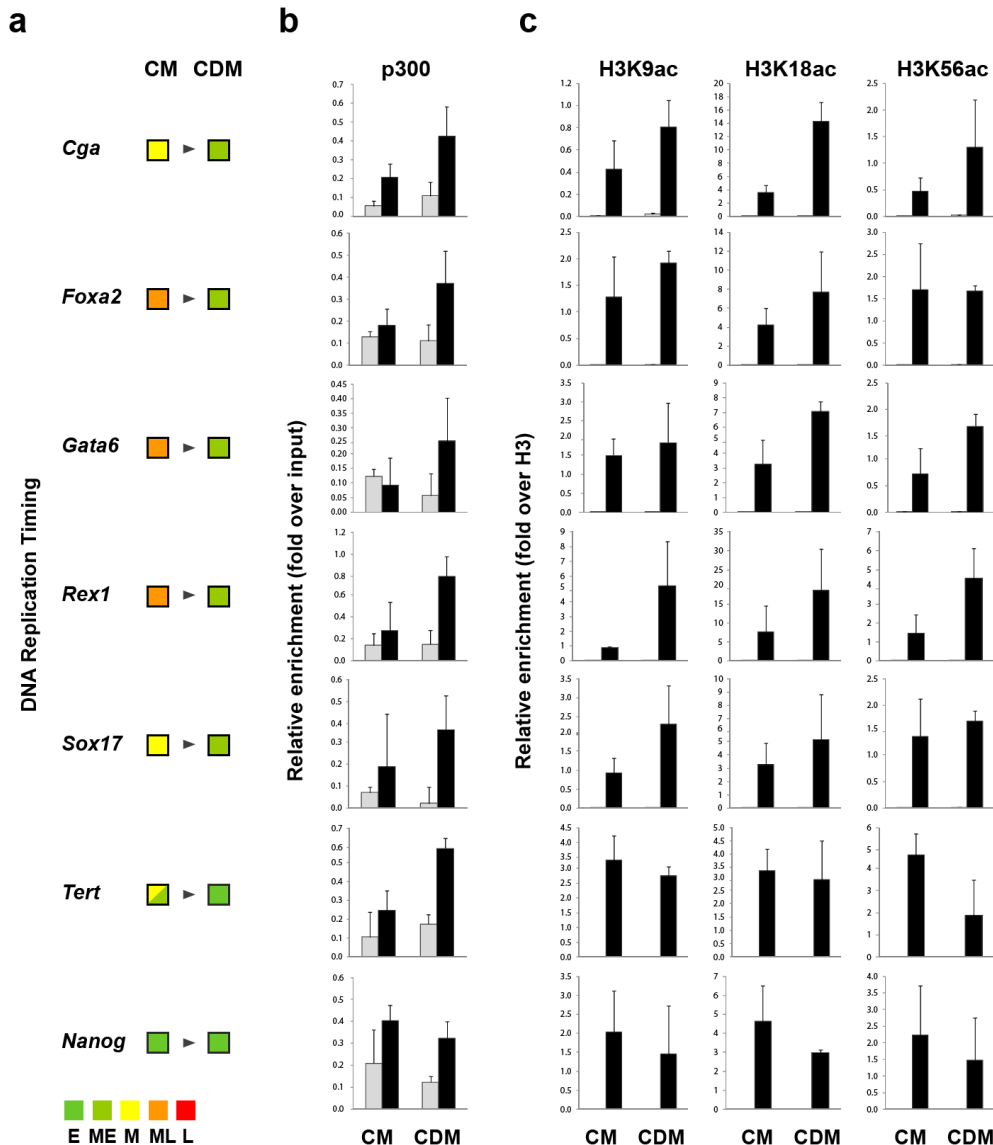


Figure 5.3. Loci changing replication timing from late in CM to early in CDM^{activin} hES cells have higher levels of p300 and histones acetylation at their promoter regions in CDM^{activin} hES cells. (a) Replication timing status of *Cga*, *Foxa2*, *Gata6*, *Rex1*, *Sox17*, *Tert* and *Nanog* was assessed in hES cells in CM and CDM^{activin} conditions. The replication timing of each locus is indicated in a colour coded way as previously designated. Relative enrichments for (b) p300 and (c) acetylated histone H3 (H3K9ac, H3K18ac and H3K56ac) was assessed in H1p67CM and H1p47(CDM23) hES cells using ChIP assay and qPCR analysis (black bars). Grey bars indicate background levels (control IgG). p300 enrichments is expressed relative to input and histones data normalised to total histone H3. Error bars indicate standard deviations of at least two independent experiments.

5.3. Inhibition of p300 acetyltransferase activity leads to early to late replication timing alterations in H1 CDM^{activin} hES cells

To further decipher the relationship between p300-mediated histone acetylation and maintenance of stable, early-replicating profiles in hES cell cultures, we examined whether inhibiting p300 activity would be sufficient to change the replication timing status (from early to late) of loci directly targeted by this enzyme. Here, the replication profile of H1 CDM^{activin} hES cells treated with curcumin (also known as diferuloylmethane), a specific inhibitor of p300/CBP HAT activity (Balasubramanyam et al., 2004) was compared to that of untreated CDM^{activin} and CM hES cells.

Initially it was tested by western blot a range of curcumin concentrations (10, 25 and 50 μ M) in H1 CDM^{activin} cells to assess the impact of the drug on global histone acetylation levels. As shown in Figure 5.4a, the acetylation of histone H3 (H3K9ac, H3K56ac and H3K18ac) by p300 was strongly inhibited after 24 hours treatment with 10 and 25 μ M curcumin. Higher concentration of curcumin (50 μ M) induced a high incidence of cell death, as previously reported (Balasubramanyam et al., 2004). H3K9 and H3K56 acetylation was completely abolished after treatment with 10 μ M curcumin and H3K18ac levels reduced yet still detected at both 10 and 25 μ M concentrations. In contrast, the presence of curcumin did not induce any change in H3K27me3 levels, confirming the specificity of this treatment (Figure 5.4a). In these experiments asynchronously dividing H1 hES cell cultures grown in CDM^{activin}, were treated for 4 days with the lowest concentration of curcumin able to produce an effect on histone acetylation levels (10 μ M). Culture media supplemented with curcumin was changed daily. In order to exclude the influence of spontaneous differentiation in these curcumin-treated cultures, the expression of undifferentiated ES-associated genes (*Oct4*, *Nanog* and *Rex1*) was verified by RT-PCR analysis at day 4 of treatment, showing similar levels in both H1 CDM^{activin} untreated and treated cells (DMSO and curcumin). Differentiated-associated markers (*Foxa2*, *Gata6*, *Eomes* and *Cdx2*) were also included and found absent or residual in all three conditions (Figure 5.4b). Homogeneous staining for Oct4 and Nanog were further observed by IF analysis, while the differentiated surface marker SSEA1 was absent in cell populations tested (Figure 5.4c).

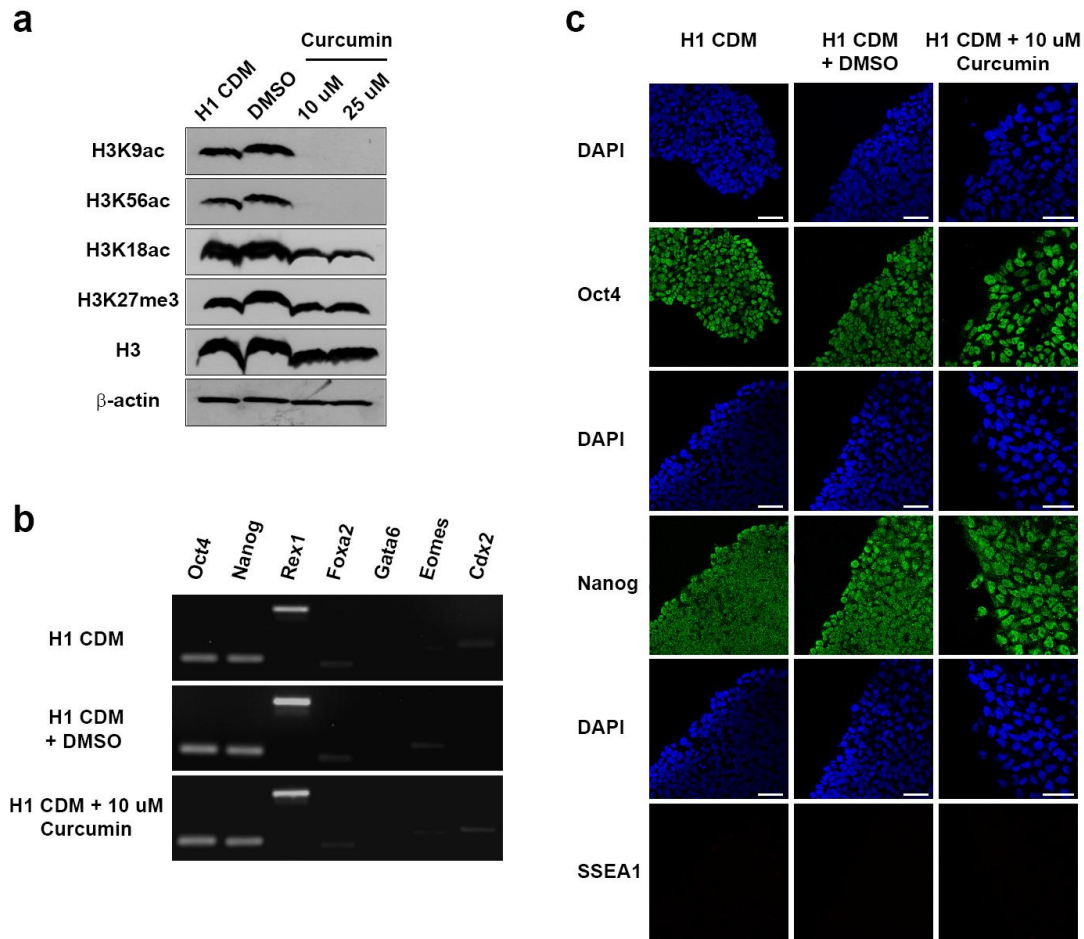


Figure 5.4. Inhibition of p300 activity in H1 CDM^{activin} hES cells by curcumin leads to a reduction in the levels of histone acetylation without inducing cell differentiation. (a) Western blot analysis of H1 CDM^{activin} hES cells treated during 24 hours with DMSO (solvent control), 10 and 25 μM of p300 inhibitor curcumin. Membranes were probed with antibodies specific for p300-mediated acetylated histone H3 (H3K9ac, H3K56ac and H3K18ac). Antibody against H3K27me3 was also used as a control for the specificity of curcumin p300 inhibition. Equivalent protein loading is shown by the total H3 and β-actin detection. (b) RT-PCR analysis of the expression of undifferentiated ES-associated (*Oct4*, *Nanog* and *Rex1*) and differentiated-associated (*Foxa2*, *Gata6*, *Eomes* and *Cdx2*) genes in H1 CDM^{activin} hES cells treated during 4 days with DMSO and 10 μM of curcumin. (c) Immunofluorescence analysis of H1 CDM^{activin} hES cells treated during 4 days with DMSO and 10 μM of curcumin for Oct4 (green), Nanog (green) and SSEA1 (red). Nuclei are shown by DAPI staining (blue). Scale bars, 70 μm. These experiments were performed in collaboration with Matias Autio in our laboratory.

Interestingly, we found that inhibiting p300 activity was sufficient to convert the replication timing profile of H1 CDM^{activin} cells into a CM-like profile when treated with curcumin. In particular, 12 out of 13 loci previously identified as targets for replication timing changes, from late to early in response to CDM^{activin} conditions (*Bmp4*, *Cga*, *Ebf*, *Foxa2*, *Gata6*, *Neurog1*, *Rex1*, *Sox1*, *Sox17*, *Sox3*, *Sycp1* and *Tle1*), replicated in the second half of S-phase upon p300 activity inhibition in H1 CDM^{activin} cell cultures (Figure 5.5). Most of the shifting loci started to display a replication timing similar to that previously observed in H1 CM cells, as for example *Foxa2*: middle-late in CM, middle-

early in CDM^{activin} and back to middle-late upon curcumin treatment. Importantly, loci that were consistently early (*Dppa4*, *Gata4*, *Gbx2*, *Hoxb1*, *Nanog*, *Oct4*, *Wnt3* and *Zic1*), middle (*Sox7*) or late (*Ikaros* and *Mash1*) replicating in both CM and CDM^{activin} cells were not affected by the global inhibition of p300 activity (Figure 5.5), highlighting the target specificity of p300 action. Altogether these results suggest a pivotal role for p300-mediated histone acetylation in defining the balance between hES cell naive *versus* primed states for differentiation as reflected at the DNA replication timing level.

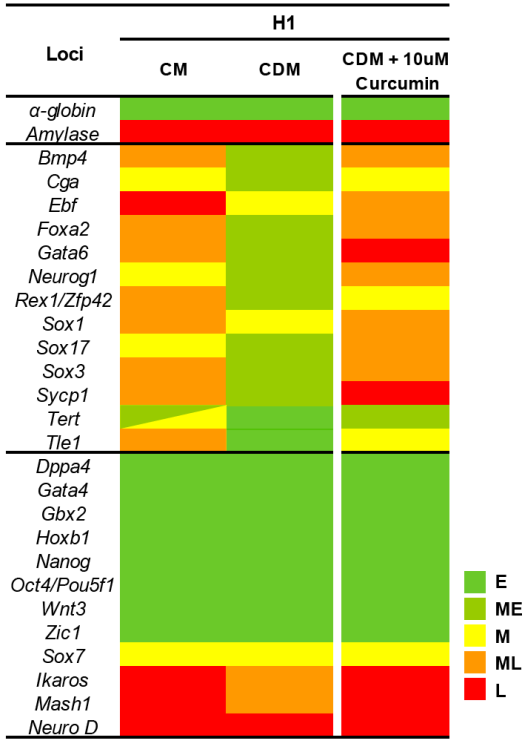


Figure 5.5. Inhibition of p300 acetyltransferase activity in curcumin-treated H1 CDM^{activin} hES cells induces replication timing changes towards a CM-like profile. (a) The replication timing of individual loci was compared between untreated H1 CM, CDM^{activin} and CDM^{activin} hES cells treated during 4 days with 10 μ M curcumin. Loci that changed replication timing from late to early in response to CDM^{activin} conditions (*Bmp4*, *Cga*, *Ebf*, *Foxa2*, *Gata6*, *Neurog1*, *Rex1*, *Sox1*, *Sox17*, *Sox3*, *Sycp1*, *Tert* and *Tle1*) delayed their replication timing towards late in curcumin-treated CDM^{activin} cultures, with the exception of *Tert*. Other loci consistently replicated early (*Dppa4*, *Gata4*, *Gbx2*, *Hoxb1*, *Nanog*, *Oct4*, *Wnt3* and *Zic1*), middle (*Sox7*) and late (*Ikaros*, *Mash1* and *Neuro D*) in untreated CM and CDM^{activin} cells as well as in CDM^{activin} cells treated during 4 days with 10 μ M curcumin.

CHAPTER 6. DISCUSSION

Pluripotent ES cells provide a unique opportunity for studying mammalian development and differentiation *in vitro*. A growing panel of mouse and human pluripotent stem cell lines has been derived from the early embryo at different developmental stages, as well as from adult cells reprogrammed by a variety of approaches. These cell lines share the expression of key pluripotency markers and are able to self-renew and to generate differentiated progenies when induced (Jaenisch and Young, 2008; Rossant, 2008). However, their relationship to each other and whether they correspond to different pluripotent states with distinct developmental potencies and affiliations *in vivo* is not clear. For example, hES cells, in contrast to mES cells, can spontaneously differentiate into extra-embryonic cell types despite being derived from the ICM of pre-implantation embryos. One possible explanation is that hES cells originate from a less committed (possibly pre-epiblast) stage while mES cells represent the lineage-restricted epiblast progenitors in the ICM. Yet such an explanation is hard to reconcile with most recent reports suggesting that hES cells closely resemble mEpiS cells - a pluripotent stem cell population derived from post-implantation embryos (Brons et al., 2007; Tesar et al., 2007). In this study, I have investigated whether distinct pluripotent states could be reliably delineated at the chromatin level by comparing the replication timing profiles of hES, mES and mEpiS cell populations. Importantly I showed that hES cells can transit from a primed epiblast-like to a naive ICM-like state by changing growth culture conditions, a level of heterogeneity not previously described within hES cell cultures. Early and steady replication timing profiles are shown to be a unique feature of the naive state while late and variable replication profiles characterize a more primed hES cell state. In addition, our findings strongly suggest a pivotal role for TGF- β /Activin signalling through Smad2/3 action and its downstream target p300, in modulating the balance between distinct hES cell states in culture (Figure 6.1).

6.1. Human and mouse ES cells have distinct replication timing profiles

In this study, I have compared the DNA replication timing status of a large panel of key developmental genes in several hES cell lines alongside mES and mEpiS cells to investigate whether we can reliably discriminate these distinct but closely related

developmental states using a chromatin-based approach. The candidate genes selected for this analysis mainly encode human and mouse homologs selectively expressed in embryonic and extraembryonic tissues of the early embryo or involved in somatic and germ-line lineages specification.

Here, I showed that the replication timing profile of independently derived mES cell lines from different genetic backgrounds is remarkably steady and consistent across cell lines. This contrasts with the recent demonstration that mES cells display considerable cell to cell heterogeneity in the expression levels of certain pluripotency-associated genes such as *Nanog*, *Rex1* and *Stella* (Chambers et al., 2007; Hayashi et al., 2008; Toyooka et al., 2008) suggesting that replication-timing profiles might be a more robust indicator of a pluripotent mES cell identity than transcriptional profiles. In addition, and in agreement with previous studies (Azuara et al., 2006; Hiratani et al., 2008), I confirmed that many genes in mES cells replicate early during S-phase, regardless of their transcriptional status, thus reflecting a broadly accessible chromatin structure. In fact, replication timing as an indicator of chromatin accessibility has been previously used to study epigenetic events underlying lineage commitment and pluripotency (Azuara et al., 2006; Hiratani and Gilbert, 2009; Kitsberg et al., 1993; Wutz and Jaenisch, 2000). In particular, replication timing profiling was capable of discriminating pluripotent mES cells from cells with a more restricted differentiation capacity, with many lineage-specifying genes becoming late replicating in unrelated cell types (i.e. neural genes in fully differentiated lymphocytes) (Azuara et al., 2006; Hiratani et al., 2008; Perry et al., 2004). These features, made of replication timing an attractive approach to characterize hES cells, where lineage affiliation and functional verification of embryonic potential *in vivo* is not possible, due to obvious ethical reasons.

Surprisingly, I observed that human ES cells have a significantly distinct chromatin profile from that of mES cells. A profound difference was the increased proportion of genes replicating later during S-phase as seen in H1, H7 and H9 hES cell lines; a feature that is normally found in differentiated cells (Azuara et al., 2006; Perry et al., 2004). Yet these hES cell cultures displayed high and homogeneous expression levels of undifferentiated markers (like Oct4 and Nanog), excluding spontaneous differentiation as the primary cause of late replication. Moreover, hES cell replication profiles could still be discriminated from that of cells with a more restricted developmental competence like hES-derived neuronal progenitor (NP) cells, as previously seen for mES cells (Hiratani et al., 2008; Perry et al., 2004). In particular, I found that many neural-specific genes (such as *NeuroD*, *Sox3* and *Sox1*) advanced their replication timing toward early S-phase upon neural induction in differentiating

hES cells, consistent with an increase of chromatin accessibility at these loci. In contrast, many non-neural genes, including ES-specific (*Cripto* and *Tle1*) and other somatic lineage-associated markers (such as *Bmp4*, *Foxa2*, *Sox17* and *Eomes*), shifted to their replication timing toward late S-phase in NP cells as compared to hES cells, suggesting that epigenetic changes associated with neural lineage commitment and loss of pluripotency have indeed taken place.

6.2. Replication timing variability is a hallmark of hES cells in CM conditions and mEpiS cells

Intriguingly, hES cell replication profiles were also found to significantly vary between cell lines and cultures; a level of variability not observed among mES cells that harbour a steady and mostly early-replicating profile. This variability could not be simply explained by distinct derivation histories since H1, H7 and H9 were all derived simultaneously in the same laboratory (Thomson et al., 1998) and expanded in the same culture conditions (CM) (Xu et al., 2001). Moreover, variability was observed at specific loci (*Dppa4*, *Tert*, *Tle1*, *Sycp1*, *Bmp4*, *Ebf*, *Foxa2*, *Gata6*, *Sox17*, *Sox3* and *Zic1*) and it is likely that this is a lower-limit estimate as inclusion of additional hES cell lines and genes could substantially increase the spectrum of observed replication timing variability. One possibility would be that such level of epigenetic variability solely reflects the considerable genetic variation that exists within the human population (in contrast to inbred mouse laboratory strains). In accordance, several reports have similarly revealed variability between hES cells at the transcriptional level (Abeyta et al., 2004; Adewumi et al., 2007; Skottman et al., 2006; Wei et al., 2005). This possibility was, however, ruled out by analysing the replication timing status of variable replicating loci (as defined in hES cells) in human CD34+ haematopoietic stem cells isolated from blood samples of three unrelated volunteers; these profiles were shown to be remarkably similar across samples, with no or little evidence of epigenetic variability. Moreover, a recent study comparing the replication timing of human and mouse homolog genes in lymphocytes showed a high conservation across species in terms of replication timing (Farkash-Amar et al., 2008). This study together with our observations, suggest that replication timing variability might be a unique feature of hES cell cultures and/or their *in vivo* equivalent in the human embryo.

Changes in DNA replication timing profiles are generally thought to reflect variations in chromatin accessibility and transcriptional permissiveness (Donaldson,

2005). ChIP data generated for different histone acetylation marks (H3K9ac, H3K14ac, H3ac and H4ac) were compared with the replication timing and expression status of selected genes across hES cell lines grown under the same culture conditions. Highly expressed and early replicating genes across H1, H7 and H9 hES cells (such as *Oct4*, *Nanog* and *Sox2*) showed a positive correlation with higher levels of histone acetylation, as previously described for mES cells (Azuara et al., 2006; Jorgensen et al., 2007; Perry et al., 2004). Accordingly, silent and late replicating genes (as for example *Myf5*) showed background levels for histone acetylation as seen at *satellite repeats* (a constitutive heterochromatic control region used in these experiments). In contrast, no straightforward correlation between replication timing variability, histone acetylation patterns and/or expression could be drawn when analysing variable replicating genes across H1, H7 and H9 cells. Complementary DNA methylation analysis did not reveal any correlation between DNA methylation and replication timing variability at specific loci in hES cells neither. This is in agreement with previous studies showing that the replication timing status of individual loci did not change in DNA methyltransferase-deficient mES cells (Gribnau et al., 2003; Jorgensen et al., 2007). No direct correlation was observed neither between chromosomal location or genomic features (such as GC content and LINE density) and replication timing variable targets, in contrast to previous studies (Hiratani et al., 2004). Interestingly, all variable genes (apart from *Tert*) were located in gene poor genomic regions. Although this is not a unique feature of variable loci in hES cells, it implies that the changes in replication timing of those genes are not influenced by the transcriptional status of neighbouring genes.

What is then the functional significance of replication timing variability? Such a level of plasticity within hES cell cultures could indicate an unstable or transitional state: primed on the verge of differentiation. This interpretation would be consistent with hES cells resembling most closely to mEpiS cells than mES cells, as previously suggested (Brons et al., 2007; Tesar et al., 2007). In this study, we observed that mEpiS and hES cells share similar cell cycle profiles, most likely reflecting their common growth requirements (Brons et al., 2007; Tesar et al., 2007). Similarly, we found that *Rex1* – an ICM-specific marker that is shutdown in the epiblast - replicates early during S-phase in mES cells but consistently late in all mEpiS and hES cell lines analysed. In contrast, the replication timing of *Fgf5*, an epiblast-specific marker, shifts from late in mES cells to early in mEpiS cells, underlying gene de-repression mechanisms yet remains late replicating in hES cells. Thus, establishing the replication status of ICM/epiblast-specific markers was not sufficient to clarify the developmental

status of hES cells. It should, however, be noted that the regulation of *Rex1* and *Fgf5* expression might be different between species. In fact, the selection of these markers was based on gene expression analysis in mouse embryos, in the absence of equivalent information in human embryos. Our replication timing analysis was further extended to other target loci previously analysed in mES and hES cells, i.e. differentiation-associated markers such as *Ebf*, *Ikaros*, *Foxa2*, *Sox17*, *Sox3* and *Zic1* that are subjected to variable replication timing profile in hES cells. Remarkably, hES cells (grown under MEF-CM conditions) and mEpiS cells were revealed to share a common 'unstable' state, as indicated by a high level of replication timing variability across a common set of loci, as opposed to mES cells. This observation is in agreement with previous suggestions that hES cells harbour an epiblast-like phenotype in culture, suggesting a pre-gastrulating developmental stage affiliation *in vivo*.

6.3. Identification and characterization of distinct pluripotent states within hES cell cultures: naïve vs primed

One of the most surprising results arising from this work is the demonstration that hES cell replication timing profile can be prone to extensive changes under different culture conditions. We found that H1 and H9 hES cells grown under standard MEF-CM conditions undergo dramatic epigenetic alterations when transferred into CDM^{activin} conditions, with many genes stably and consistently shifting their replication timing towards early S-phase in both cell lines. Moreover, the modulation of hES cell replication profiles was here demonstrated to be a reversible process as returning cells from CDM^{activin} to CM conditions was sufficient to revert to an unstable and late profile. A significant overlap between genes advancing their replication timing in H1 and H9 hES cells was observed, suggesting that a conserved mechanism is driving hES cell cultures towards a more stable state. Importantly, these results further challenge the view that hES cells closely resemble mEpiS cells, and suggest that distinct but interchangeable subpopulations may exist within hES cell cultures, as delineated by a stable, early-replicating (naïve) *versus* unstable, late-replicating (primed) profile (Figure 6.1). These results are reminiscent of recent studies showing that undifferentiation mES cell cultures contain distinct subpopulations that either correspond to a ICM-like or epiblast-like identity, that can vary in response to different culture conditions (Bao et al., 2009; Chou et al., 2008; Hanna et al., 2009; Nichols and Smith, 2009). Importantly,

we showed for the first time that hES cell cultures are equally metastable and can alternate between different pluripotent states when grown in CM and CDM^{activin} conditions. Our observations were further extended and confirmed in human iPS cells, providing additional evidence that induced (iPS) and natural (hES) pluripotent cells share common properties (Colman and Dreesen, 2009) and reiterating the impact of culture conditions in modulating or supporting different epigenetic states within pluripotent cell cultures. This is an important consideration as responses to differentiation protocols are likely to vary depending on where the cells stand in a developmental hierarchy. Interestingly, a recent study shows that reprogramming using a commercial CDM media supplemented with FGF2 and TGF- β , favors the formation of fully reprogrammed over partially reprogrammed colonies (Chan et al., 2009). It would be interesting to compare the replication profiles of different iPS cell lines derived from distinct methodologies.

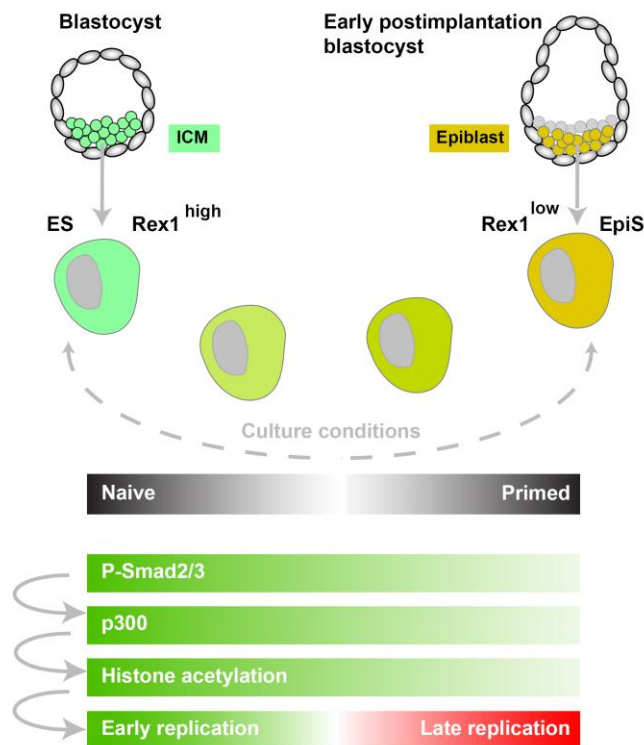


Figure 6.1. Proposed model for the maintenance of distinct pluripotent states within hES cell cultures. Human ES cell cultures contain distinct subpopulations of pluripotent cells that can transit from a naive ICM-like to a primed for differentiation epiblast-like state or *vice versa*. Oscillations between these subpopulations (indicated by the broken arrow) are induced by changing hES cell culture conditions. Human ES cells grown under standard MEF-CM share characteristics with primed cells for differentiation, and are notably specified by low expression levels for the ICM marker *Rex1*. In contrast, cells transferred to CDM^{activin} show characteristics of a naive state, with high levels of *Rex1* expression. Changes in culture conditions (from CM to CDM^{activin}) are here accompanied by increased levels of the TGF- β /Activin signal inducer P-Smad2/3. Activated Smad2/3 act through the HAT p300, that is also present at higher levels in CDM^{activin} cells, promoting an increase in histone acetylation levels at the promoters of target loci, shifting from late to early replication timing (in CDM^{activin}) and thus a general increase in chromatin accessibility across many lineage-specifying genes in hES cells.

The fact that hES cell replication profiles can be very rapidly and dynamically shifted upon culture condition changes (from CM to CDM^{activin} conditions and *vice versa*) excludes “cell culture adaptation” as a unique cause for these alterations. Cell adaptation is a common phenomenon observed in undifferentiated hES cell cultures, especially long-term cultures, often associated with karyotypic changes that provide some advantages in terms of clonogenic capacity and self-renewal properties of these cells (Baker et al., 2007; Draper et al., 2004; Enver et al., 2005). However, this phenomenon did not seem to influence on hES cell replication properties as verified by directly comparing the profiles of normal diploid H1-BH1 cells and adapted, long-term H1-BH1 cultures that were confirmed to be karyotypically abnormal; these cells were kindly provided by Professor Peter Andrews in Sheffield (Appendix II, Figure A6).

Phenotypically, hES cells grown in both CM and CDM^{activin} conditions display characteristics of undifferentiated and self-renewing cells, similarly showing strong staining for alkaline phosphatase, and high expression levels for the key pluripotency transcription factors, Oct4, Nanog and Sox2, and cell surface markers, Tra-1-60, Tra-1-81 and SSEA4. Moreover, both hES cell populations can undergo differentiation into derivatives of the three germ layers upon EBs formation and prolonged cultures. Yet CM and CDM^{activin} hES cells were demonstrated to possess functionally distinct properties, with CM cells being more responsive to differentiation signals than CDM^{activin} cells. This corroborates the view that CM cells might be already in a pre-gastrulating stage, while CDM^{activin} cells preferentially adopt an ICM-like phenotype. Such bias towards an ICM-like phenotype in cultures transferred to CDM^{activin} conditions was further confirmed by a rapid and dramatic increase in *Rex1* expression levels, as well as the up-regulation of *Gata2* and *Igf2r* - a set of markers recently found to be specifically expressed in ICM cells of the early human embryo (Reijo Pera et al., 2009). Conversely, many early germ layer markers such as *Eomes*, *Gata6*, *Sox17* and *Foxa2* that are commonly expressed at low levels in mEpiS but not in mES cells, were drastically downregulated in CDM^{activin} hES cell cultures, indicating a loss of transcriptional noise. These results suggest that CM conditions might either contain extrinsic differentiation stimuli, possibly secreted by the MEFs that prime cells for lineage commitment or on the contrary be deprived of adequate concentrations of signalling factors essential for the maintenance of a true embryonic ground state, as recently suggested for mES cells (Ying et al., 2008).

Several studies have previously shown that somatic (i.e, lymphocyte) nuclei can be reprogrammed to a pluripotent state when injected into *Xenopus* oocytes (Byrne et

al., 2003), upon nuclear transfer into mouse oocytes (Hochedlinger and Jaenisch, 2002), or through the generation of cell hybrids with pluripotent ES (Tada et al., 2001) and EG cells (Tada et al., 1997). Here, I investigated the potential of hES cells grown either in CM or CDM^{activin} conditions, alongside mEpiS and mES cells to reprogram the gene expression program of lymphocyte nuclei, using inter-species cell fusion assays (Pereira and Fisher, 2009). While mES cells were able to reprogram hB lymphocytes for pluripotency, as expected, re-activation of human pluripotency-associated genes, including *hOct4* and *hNanog* and extinction of *hCD19* – a lymphocyte-specific marker were not observed in our mEpiS x hB fusions, indicating that mEpiS cells might lack the potential to reprogram hB lymphocytes, at least in the conditions tested in our experiments. This result was surprising and could not be simply explained by differential fusion efficiencies between mES or mEpiS and hB cells. Moreover, other specialised cell types like fibroblasts (Massa et al., 2000) or extraembryonic stem cells (J. Santos and F. Pereira, personal communication) were previously found capable of reprogramming lymphocytes towards their own identity. This may relate to the limiting amounts of *c-Myc*, *Esrrb* and *Klf4*, three key reprogramming factors (Feng et al., 2009a; Takahashi and Yamanaka, 2006), present in mEpiS cells as compared to mES cells. Accordingly, a recent study showed that the forced expression of *Klf4* in mEpiS cells was sufficient to induce the formation of cells with characteristics of mES cells (Guo et al., 2009). Another possibility is that reprogramming inhibitor trans-acting factors present in mEpiS and lacking in mES cells may constrain gene activation in hB lymphocytes. Although the precise mechanisms controlling reprogramming are still not full understood, recent evidences suggest a central inhibitory role for senescence factors in reprogramming efficiency (Banito et al., 2009; Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009a; Marion et al., 2009; Utikal et al., 2009). Further studies comparing the regulation of senescence pathways in mES and mEpiS cells would bring more insights into this subject. A third possibility for the lack of reprogramming potential in mEpiS cells might correlate with cell cycle differences between mES and mEpiS cells. It was shown that in cell-hybrids, cells were most effective at reprogramming when they contained a high proportion of cells in the S- and G2/M- phases of the cell cycle. Accordingly, we showed that mES cells have a higher proportion of cells in S-phase than mEpiS cells.

Human ES cells in CM and CDM^{activin} share with mES cells the potential to reprogram somatic cells for pluripotency upon heterokaryon formation, as previously described (Cowan et al., 2005; Yu et al., 2006). This contrast with mEpiS cells that were unable to reprogram hB cell nuclei within three days after fusion. Importantly, these results demonstrate that hES cells, independently of their culture conditions are

clearly different from mEpiS cells. Interestingly, the set of pluripotency-associated genes reactivated in the mB cell nuclei when fused with either CM or CDM^{activin} hES cells was different. In particular, *mNanog* was not reactivated in CM in contrast to CDM^{activin} fusions while *mSox2* and *mRex1* were strongly enhanced in CM fusions but not CDM^{activin}, further highlighting distinct properties between these hES cell populations.

6.4. Investigating the mechanisms that specify distinct pluripotent states within human ES cell cultures: a pivotal role for TGF- β /Activin signalling and the histone acetyltransferase p300

Although the mechanisms controlling self-renewal *versus* lineage commitment remain poorly understood in hES cells, several studies demonstrated that TGF- β /Activin signaling acting through the signal inducer Smad2/3 is essential for maintaining pluripotency in human cell lines (James et al., 2005; Vallier et al., 2005; Xiao et al., 2006; Xu et al., 2008). In this study, we showed that increased levels of the active P-Smad2/3 forms underlie the acquisition of an ICM-like phenotype in pluripotent hES cell cultures when transferred from CM to CDM^{activin} conditions. Although both culture conditions similarly rely on the presence of the TGF- β member Activin A to sustain an undifferentiated hES cell state (note that FGF2 directly induces the secretion of activin A by MEFs used to prepare CM media (Beattie et al., 2005), this essential signaling pathway was found to be differentially activated in CM and CDM^{activin} conditions - two commonly used protocols to maintain hES cells *in vitro*. This is of major importance, considering that many different members of the TGF- β family, including Activin A, are known to either favor self-renewal or differentiation depending on their respective concentrations; the concentrations of these pleiotropic factors are indeed very likely to vary between CM and CDM^{activin} conditions (Sulzbacher et al., 2009). In agreement, follistatin, a known inhibitor of activin A activity (Nakamura et al., 1990) was shown to be highly abundant in MEF-CM media (Bendall et al., 2007).

Interestingly, the acquisition of a more stable, early-replicating pluripotent state in CDM^{activin} hES cells is also accompanied by a transient yet significant upregulation of *Nanog* levels. This is consistent with a direct role for Smad2/3 in regulating *Nanog* expression in hES cells (Vallier et al., 2009a; Xu et al., 2008). Note that *Nanog* was shown to directly regulate the expression of *Rex1* in mES cells (Shi et al., 2006). If

similar mechanisms are conserved in human cells, this could explain the rapid and steady increase in *Rex1* levels observed in hES cell cultures upon transfer from CM to CDM^{activin} conditions. Further ChIP studies to assess the levels of Nanog occupancy at the human *Rex1* promoter in both CM and CDM^{activin} cultures would be required to test this possibility. Moreover, our results suggest a possible role for Nanog in the establishment of an ICM-like state as defined in this study by a stable, early-replicating profile in CDM^{activin} hES cells. These results are in agreement with Nanog's role in stabilizing mES cells in culture, as well as during epigenetic reprogramming of pluripotent and germ cells *in vivo* (Chambers et al., 2007) or upon cell fusion *in vitro* (Silva et al., 2006). Once established, this new pluripotent state is thought to be stably propagated throughout cell divisions while negative feedback mechanisms are likely to regulate *Nanog* expression levels. In agreement, *Tcf3*, a putative repressor of *Nanog* in mES cells (Pereira et al., 2006; Yi et al., 2008) showed a converse pattern of gene expression with that of *Nanog* in hES cell cultures when transferred from CM to CDM^{activin} conditions. Of note, several members of the TGF- β family, including *Nodal* were also found first up-regulated yet down-regulated alongside *Nanog*, also suggesting the occurrence of negative feedback mechanisms that modulate their expression levels in hES cells. Accordingly, *Nodal* expression in undifferentiated hES cells was previously shown to require activation of Smad2/3 and is regulated by a negative feedback mechanism, involving the expression of the antagonists *Lefty-A* and *B*. (Besser, 2004).

The transition from early to late replication in CDM^{activin} hES cells was also accompanied by an increase in the expression and activity of a known Smad2/3 co-activator, the HAT p300 (Schmierer and Hill, 2007). Importantly, we showed that p300 is recruited at specific target loci shifting replication timing from late to early in CDM^{activin} cells, where it mediates high levels of histone H3 acetylation. Such increase of chromatin accessibility through p300 recruitment was shown to occur at both active (i.e. *Rex1*) and inactive genes (i.e. *Cga* and *Foxa2*) that commonly replicate in early S-phase, in accordance with a strong positive correlation between histone acetylation and early replication timing, as previously reported (Mendez, 2009). In particular, a recent study showed that the recruitment of HATs to the silent human β -globin locus in a non-erythroid cell line changed its replication timing from late to early. Conversely, tethering a HDAC to this locus in an erythroid cell line (where β -globin is expressed) was sufficient to change its replication timing from early to late (Goren et al., 2008). Similarly and most importantly, we demonstrated that inhibiting p300 activity by a specific inhibitor (curcumin) was sufficient to convert the replication timing profiles of

early CDM^{activin} cells into a late CM-like profile. Altogether our results strongly suggest a key role for the HAT p300 in regulating the balance between naïve versus primed states in hES cell cultures (Figure 6.1). Interestingly, this mirrors recent findings showing that the use of trichostatin A, an HDAC inhibitor was sufficient to bias mES cell cultures towards a Stella-positive, ICM-like state (Hayashi et al., 2008). On the same note, sodium butyrate, another HDAC inhibitor, was found capable of promoting both mouse and human ES cells self-renewal in minimum culture conditions; interestingly, both pluripotent cell types displayed a more closely related, ICM-like phenotype in these culture conditions (Ware et al., 2009). It would be of high interest to further explore how different HDACs and HATs may act together to regulate and balance distinct pluripotent states in hES cell cultures.

REFERENCES

- Abeyta, M.J., Clark, A.T., Rodriguez, R.T., Bodnar, M.S., Pera, R.A., and Firpo, M.T. (2004). Unique gene expression signatures of independently-derived human embryonic stem cell lines. *Hum Mol Genet* 13, 601-608.
- Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S., *et al.* (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* 25, 803-816.
- 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.
- Ahlgren, U., Jonsson, J., and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development* 122, 1409-1416.
- Ahmad, K., and Henikoff, S. (2002a). Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16477-16484.
- Ahmad, K., and Henikoff, S. (2002b). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* 9, 1191-1200.
- Aladjem, M.I. (2007). Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* 8, 588-600.
- Allegrucci, C., Wu, Y.Z., Thurston, A., Denning, C.N., Priddle, H., Mummery, C.L., Ward-van Oostwaard, D., Andrews, P.W., Stojkovic, M., Smith, N., *et al.* (2007). Restriction Landmark Genome Scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. *Hum Mol Genet*.
- Allegrucci, C., and Young, L.E. (2007). Differences between human embryonic stem cell lines. *Human reproduction update* 13, 103-120.
- Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* 51, 786-794.
- Ambrosi, D.J., Tanasijevic, B., Kaur, A., Obergfell, C., O'Neill, R.J., Krueger, W., and Rasmussen, T.P. (2007). Genome-wide reprogramming in hybrids of somatic cells and embryonic stem cells. *Stem Cells* 25, 1104-1113.
- Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Dev Biol* 227, 271-278.
- Amit, M., Shariki, C., Margulets, V., and Itskovitz-Eldor, J. (2004). Feeder layer- and serum-free culture of human embryonic stem cells. *Biol Reprod* 70, 837-845.
- Aparicio, J.G., Viggiani, C.J., Gibson, D.G., and Aparicio, O.M. (2004). The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in *Saccharomyces cerevisiae*. *Mol Cell Biol* 24, 4769-4780.
- Arney, K.L., and Fisher, A.G. (2004). Epigenetic aspects of differentiation. *J Cell Sci* 117, 4355-4363.
- Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nature reviews* 10, 91-103.
- 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. (2006). Profiling of DNA replication timing in unsynchronized cell populations. *Nature protocols* 1, 2171-2177.
- Azuara, V., Brown, K.E., Williams, R.R., Webb, N., Dillon, N., Festenstein, R., Buckle, V., Merkenschlager, M., and Fisher, A.G. (2003). Heritable gene silencing in

- lymphocytes delays chromatid resolution without affecting the timing of DNA replication. *Nat Cell Biol* 5, 668-674.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M., *et al.* (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-538.
- Babaie, Y., Herwig, R., Greber, B., Brink, T.C., Wruck, W., Groth, D., Lehrach, H., Burdon, T., and Adjaye, J. (2007). Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* 25, 500-510.
- Baker, D.E., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., and Andrews, P.W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol* 25, 207-215.
- Balasubramanyam, K., Varier, R.A., Altaf, M., Swaminathan, V., Siddappa, N.B., Ranga, U., and Kundu, T.K. (2004). Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription. *The Journal of biological chemistry* 279, 51163-51171.
- Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., *et al.* (2009). Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev.*
- Bannister, A.J., Schneider, R., and Kouzarides, T. (2002). Histone methylation: dynamic or static? *Cell* 109, 801-806.
- 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.*
- Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltby, N., Wu, W., Lyko, F., *et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675.
- Bauer, U.M., Daujat, S., Nielsen, S.J., Nightingale, K., and Kouzarides, T. (2002). Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO reports* 3, 39-44.
- Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells* 23, 489-495.
- Beck, F., Erler, T., Russell, A., and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* 204, 219-227.
- 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.
- Bendall, S.C., Stewart, M.H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bosse, M., *et al.* (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature* 448, 1015-1021.
- Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. *Cell* 128, 669-681.
- 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.
- Besser, D. (2004). Expression of nodal, lefty-a, and lefty-B in undifferentiated human embryonic stem cells requires activation of Smad2/3. *The Journal of biological chemistry* 279, 45076-45084.
- Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of molecular biology* 203, 971-983.

- Bhaumik, S.R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14, 1008-1016.
- Bickmore, W.A., and Carothers, A.D. (1995). Factors affecting the timing and imprinting of replication on a mammalian chromosome. *J Cell Sci* 108 (Pt 8), 2801-2809.
- Bielinska, M., Narita, N., and Wilson, D.B. (1999). Distinct roles for visceral endoderm during embryonic mouse development. *Int J Dev Biol* 43, 183-205.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.
- Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. *Cell* 99, 451-454.
- Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799-816.
- Boland, M.J., Hazen, J.L., Nazor, K.L., Rodriguez, A.R., Gifford, W., Martin, G., Kupriyanov, S., and Baldwin, K.K. (2009). Adult mice generated from induced pluripotent stem cells. *Nature* 461, 91-94.
- Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nat Rev Drug Discov* 5, 769-784.
- Bongso, A., Fong, C.Y., Ng, S.C., and Ratnam, S. (1994). Isolation and culture of inner cell mass cells from human blastocysts. *Human reproduction (Oxford, England)* 9, 2110-2117.
- Booth, C., and Potten, C.S. (2000). Gut instincts: thoughts on intestinal epithelial stem cells. *The Journal of clinical investigation* 105, 1493-1499.
- Bortvin, A., Eggan, K., Skaletsky, H., Akutsu, H., Berry, D.L., Yanagimachi, R., Page, D.C., and Jaenisch, R. (2003). Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* 130, 1673-1680.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947-956.
- 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.
- Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64, 1123-1134.
- Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E., and Arnold, H.H. (1989). A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *The EMBO journal* 8, 701-709.
- 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.
- Burke, B., and Stewart, C.L. (2002). Life at the edge: the nuclear envelope and human disease. *Nat Rev Mol Cell Biol* 3, 575-585.
- Byrne, J.A., Pedersen, D.A., Clepper, L.L., Nelson, M., Sanger, W.G., Gokhale, S., Wolf, D.P., and Mitalipov, S.M. (2007). Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* 450, 497-502.
- Byrne, J.A., Simonsson, S., Western, P.S., and Gurdon, J.B. (2003). Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. *Curr Biol* 13, 1206-1213.
- Camargo, M., and Cervenka, J. (1982). Patterns of DNA replication of human chromosomes. II. Replication map and replication model. *American journal of human genetics* 34, 757-780.
- Campos, E.I., and Reinberg, D. (2009). Histones: Annotating Chromatin. *Annual review of genetics*.
- Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* 15, 57-67.

- Carlson, L.L., Page, A.W., and Bestor, T.H. (1992). Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev* 6, 2536-2541.
- Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117, 427-439.
- 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.
- Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 27, 275-280.
- Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18, 1119-1130.
- Chan, E.M., Ratanasirintrao, S., Park, I.H., Manos, P.D., Loh, Y.H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T.A., *et al.* (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nat Biotechnol* 27, 1033-1037.
- Chapman, G., Remiszewski, J.L., Webb, G.C., Schulz, T.C., Bottema, C.D., and Rathjen, P.D. (1997). The mouse homeobox gene, Gbx2: genomic organization and expression in pluripotent cells in vitro and in vivo. *Genomics* 46, 223-233.
- Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25, 6031-6046.
- Chou, Y.F., Chen, H.H., Eijpe, M., Yabuuchi, A., Chenoweth, J.G., Tesar, P., Lu, J., McKay, R.D., and Geijsen, N. (2008). The growth factor environment defines distinct pluripotent ground states in novel blastocyst-derived stem cells. *Cell* 135, 449-461.
- Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277, 1996-2000.
- 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.
- Cimbora, D.M., and Groudine, M. (2001). The control of mammalian DNA replication: a brief history of space and timing. *Cell* 104, 643-646.
- Ciruna, B.G., and Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. *Mechanisms of development* 81, 199-203.
- Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annual review of biochemistry* 78, 273-304.
- Clark, A.T., Rodriguez, R.T., Bodnar, M.S., Abeyta, M.J., Cedars, M.I., Turek, P.J., Firpo, M.T., and Reijo Pera, R.A. (2004). Human STELLAR, NANOG, and GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. *Stem Cells* 22, 169-179.
- Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P.N., and Lovell-Badge, R. (1996). A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* 122, 509-520.
- Colman, A., and Dreesen, O. (2009). Induced pluripotent stem cells and the stability of the differentiated state. *EMBO reports* 10, 714-721.
- Copeland, R.A., Solomon, M.E., and Richon, V.M. (2009). Protein methyltransferases as a target class for drug discovery. *Nat Rev Drug Discov* 8, 724-732.

- Costantini, M., and Bernardi, G. (2008). Replication timing, chromosomal bands, and isochores. *Proc Natl Acad Sci U S A* 105, 3433-3437.
- Covault, J., and Chalkley, R. (1980). The identification of distinct populations of acetylated histone. *The Journal of biological chemistry* 255, 9110-9116.
- Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309, 1369-1373.
- Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *The Journal of cell biology* 162, 809-820.
- Creyghton, M.P., Markoulaki, S., Levine, S.S., Hanna, J., Lodato, M.A., Sha, K., Young, R.A., Jaenisch, R., and Boyer, L.A. (2008). H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* 135, 649-661.
- Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of cell biology* 145, 1119-1131.
- Daheron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Andrews, P.W., Itskovitz-Eldor, J., and Daley, G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem Cells* 22, 770-778.
- Darr, H., Mayshar, Y., and Benvenisty, N. (2006). Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features. *Development* 133, 1193-1201.
- Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459, 113-117.
- de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., *et al.* (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* 7, 663-676.
- Dejosez, M., Krumenacker, J.S., Zitur, L.J., Passeri, M., Chu, L.F., Songyang, Z., Thomson, J.A., and Zwaka, T.P. (2008). Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell* 133, 1162-1174.
- Dillon, N. (2008). The impact of gene location in the nucleus on transcriptional regulation. *Dev Cell* 15, 182-186.
- Dimitrova, D.S., and Gilbert, D.M. (1999). The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol Cell* 4, 983-993.
- Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., *et al.* (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321, 1218-1221.
- Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of embryology and experimental morphology* 87, 27-45.
- Donaldson, A.D. (2005). Shaping time: chromatin structure and the DNA replication programme. *Trends Genet* 21, 444-449.
- Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol* 22, 53-54.
- Dutrillaux, B., Couturier, J., Richer, C.L., and Viegas-Pequignot, E. (1976). Sequence of DNA replication in 277 R- and Q-bands of human chromosomes using a BrdU treatment. *Chromosoma* 58, 51-61.
- Edelmann, P., Bornfleth, H., Zink, D., Cremer, T., and Cremer, C. (2001). Morphology and dynamics of chromosome territories in living cells. *Biochimica et biophysica acta* 1551, M29-39.

- Eden, E.R., Patel, D.D., Sun, X.M., Burden, J.J., Themis, M., Edwards, M., Lee, P., Neuwirth, C., Naoumova, R.P., and Soutar, A.K. (2002). Restoration of LDL receptor function in cells from patients with autosomal recessive hypercholesterolemia by retroviral expression of ARH1. *The Journal of clinical investigation* 110, 1695-1702.
- Edwards, C.A., Rens, W., Clarke, O., Mungall, A.J., Hore, T., Graves, J.A., Dunham, I., Ferguson-Smith, A.C., and Ferguson-Smith, M.A. (2007). The evolution of imprinting: chromosomal mapping of orthologues of mammalian imprinted domains in monotreme and marsupial mammals. *BMC evolutionary biology* 7, 157.
- Edwards, Y.H., Putt, W., Lekoape, K.M., Stott, D., Fox, M., Hopkinson, D.A., and Sowden, J. (1996). The human homolog T of the mouse T(Brachyury) gene; gene structure, cDNA sequence, and assignment to chromosome 6q27. *Genome Res* 6, 226-233.
- Egli, D., Birkhoff, G., and Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. *Nature reviews* 9, 505-516.
- Enver, T., Pera, M., Peterson, C., and Andrews, P.W. (2009). Stem cell states, fates, and the rules of attraction. *Cell stem cell* 4, 387-397.
- Enver, T., Soneji, S., Joshi, C., Brown, J., Iborra, F., Orntoft, T., Thykjaer, T., Maltby, E., Smith, K., Dawud, R.A., *et al.* (2005). Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. *Hum Mol Genet* 14, 3129-3140.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154-156.
- Fabrizio, E., El Messaoudi, S., Polanowska, J., Paul, C., Cook, J.R., Lee, J.H., Negre, V., Rousset, M., Pestka, S., Le Cam, A., *et al.* (2002). Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO reports* 3, 641-645.
- Farkash-Amar, S., Lipson, D., Polten, A., Goren, A., Helmstetter, C., Yakhini, Z., and Simon, I. (2008). Global organization of replication time zones of the mouse genome. *Genome Res* 18, 1562-1570.
- Farthing, C.R., Ficiz, G., Ng, R.K., Chan, C.F., Andrews, S., Dean, W., Hemberger, M., and Reik, W. (2008). Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS genetics* 4, e1000116.
- 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.
- Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H., Han, J., *et al.* (2009a). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* 11, 197-203.
- Feng, B., Ng, J.H., Heng, J.C., and Ng, H.H. (2009b). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. *Cell stem cell* 4, 301-312.
- Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* 17, 1870-1881.
- Foster, H.A., and Bridger, J.M. (2005). The genome and the nucleus: a marriage made by evolution. *Genome organisation and nuclear architecture. Chromosoma* 114, 212-229.
- Fougere, C., and Weiss, M.C. (1978). Phenotypic exclusion in mouse melanoma-rat hepatoma hybrid cells: pigment and albumin production are not reexpressed simultaneously. *Cell* 15, 843-854.
- Gardner, D.K., Vella, P., Lane, M., Wagley, L., Schlenker, T., and Schoolcraft, W.B. (1998). Culture and transfer of human blastocysts increases implantation rates and reduces the need for multiple embryo transfers. *Fertility and sterility* 69, 84-88.
- Gasser, S.M. (2002). Visualizing chromatin dynamics in interphase nuclei. *Science* 296, 1412-1416.
- Georgopoulos, K., Moore, D.D., and Derfler, B. (1992). Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. *Science* 258, 808-812.

- Gerami-Naini, B., Dovzhenko, O.V., Durning, M., Wegner, F.H., Thomson, J.A., and Golos, T.G. (2004). Trophoblast differentiation in embryoid bodies derived from human embryonic stem cells. *Endocrinology* 145, 1517-1524.
- Gerrard, L., Rodgers, L., and Cui, W. (2005). Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling. *Stem Cells* 23, 1234-1241.
- Giadrossi, S., Dvorkina, M., and Fisher, A.G. (2007). Chromatin organization and differentiation in embryonic stem cell models. *Curr Opin Genet Dev* 17, 132-138.
- Gidekel, S., and Bergman, Y. (2002). A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element. *The Journal of biological chemistry* 277, 34521-34530.
- Gilbert, D.M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* 14, 377-383.
- Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annual review of biochemistry* 74, 481-514.
- Goren, A., and Cedar, H. (2003). Replicating by the clock. *Nature reviews* 4, 25-32.
- Goren, A., Tabib, A., Hecht, M., and Cedar, H. (2008). DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin. *Genes Dev* 22, 1319-1324.
- Graf, L., Heimfeld, S., and Torok-Storb, B. (2001). Comparison of gene expression in CD34+ cells from bone marrow and G-CSF-mobilized peripheral blood by high-density oligonucleotide array analysis. *Biol Blood Marrow Transplant* 7, 486-494.
- Graham, V., Khudyakov, J., Ellis, P., and Pevny, L. (2003). SOX2 functions to maintain neural progenitor identity. *Neuron* 39, 749-765.
- Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem Cells* 25, 455-464.
- Gribnau, J., Hochedlinger, K., Hata, K., Li, E., and Jaenisch, R. (2003). Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* 17, 759-773.
- Guan, K., Nayernia, K., Maier, L.S., Wagner, S., Dressel, R., Lee, J.H., Nolte, J., Wolf, F., Li, M., Engel, W., *et al.* (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* 440, 1199-1203.
- 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.
- Haig, D. (2004). The (dual) origin of epigenetics. *Cold Spring Harbor symposia on quantitative biology* 69, 67-70.
- Haigis, M.C., and Guarente, L.P. (2006). Mammalian sirtuins--emerging roles in physiology, aging, and calorie restriction. *Genes Dev* 20, 2913-2921.
- Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., *et al.* (2009). Metastable pluripotent states in NOD-mouse-derived ESCs. *Cell stem cell* 4, 513-524.
- Hansen, R.S., Canfield, T.K., Lamb, M.M., Gartler, S.M., and Laird, C.D. (1993). Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* 73, 1403-1409.
- Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* 111, 369-379.
- Hatton, K.S., Dhar, V., Brown, E.H., Iqbal, M.A., Stuart, S., Didamo, V.T., and Schildkraut, C.L. (1988). Replication program of active and inactive multigene families in mammalian cells. *Mol Cell Biol* 8, 2149-2158.
- Haub, O., and Goldfarb, M. (1991). Expression of the fibroblast growth factor-5 gene in the mouse embryo. *Development* 112, 397-406.

- Hay, D.C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells* 22, 225-235.
- Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell stem cell* 3, 391-401.
- Heard, E. (2005). Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr Opin Genet Dev* 15, 482-489.
- Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *The EMBO journal* 7, 1395-1402.
- Hebert, J.M., Boyle, M., and Martin, G.R. (1991). mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation. *Development* 112, 407-415.
- Heitz, E. (1928). "Das heterochromatin der moose." *Jahrb Wiss Bot* 69, 762-818.
- Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature reviews* 10, 526-537.
- Henderson, J.K., Draper, J.S., Baillie, H.S., Fishel, S., Thomson, J.A., Moore, H., and Andrews, P.W. (2002). Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells* 20, 329-337.
- Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* 9, 15-26.
- Hiratani, I., and Gilbert, D.M. (2009). Replication timing as an epigenetic mark. *Epigenetics* 4, 93-97.
- 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., 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 biology* 6, e245.
- Hochedlinger, K., and Jaenisch, R. (2002). Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* 415, 1035-1038.
- Hochedlinger, K., and Plath, K. (2009). Epigenetic reprogramming and induced pluripotency. *Development* 136, 509-523.
- Hockfield, S., and McKay, R.D. (1985). Identification of major cell classes in the developing mammalian nervous system. *J Neurosci* 5, 3310-3328.
- Holmquist, G.P., and Ashley, T. (2006). Chromosome organization and chromatin modification: influence on genome function and evolution. *Cytogenet Genome Res* 114, 96-125.
- Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M., and Paul, S. (2009). GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. *The Journal of biological chemistry* 284, 28729-28737.
- Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 460, 1132-1135.
- Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* 326, 292-295.
- Horiike, S., Cai, S., Miyano, M., Cheng, J.F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet* 37, 31-40.
- Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 11, 774-785.
- Huberman, J.A., and Riggs, A.D. (1968). On the mechanism of DNA replication in mammalian chromosomes. *Journal of molecular biology* 32, 327-341.

- Humphrey, R.K., Beattie, G.M., Lopez, A.D., Bucay, N., King, C.C., Firpo, M.T., Rose-John, S., and Hayek, A. (2004). Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. *Stem Cells* 22, 522-530.
- Hyslop, L., Stojkovic, M., Armstrong, L., Walter, T., Stojkovic, P., Przyborski, S., Herbert, M., Murdoch, A., Strachan, T., and Lako, M. (2005). Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages. *Stem Cells* 23, 1035-1043.
- Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci* 109 (Pt 6), 1427-1436.
- Itoh, S., Ericsson, J., Nishikawa, J., Heldin, C.H., and ten Dijke, P. (2000). The transcriptional co-activator P/CAF potentiates TGF-beta/Smad signaling. *Nucleic Acids Res* 28, 4291-4298.
- Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* 442, 533-538.
- Jackson, D.A. (1995). Nuclear organization: uniting replication foci, chromatin domains and chromosome structure. *Bioessays* 17, 587-591.
- Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567-582.
- James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* 132, 1273-1282.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074-1080.
- Jeppesen, P., and Turner, B.M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* 74, 281-289.
- Jin, S.G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS genetics* 4, e1000013.
- Johansson, B.M., and Wiles, M.V. (1995). Evidence for involvement of activin A and bone morphogenetic protein 4 in mammalian mesoderm and hematopoietic development. *Mol Cell Biol* 15, 141-151.
- Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3, 415-428.
- Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068-1070.
- Jones, P.H., Simons, B.D., and Watt, F.M. (2007). Sic transit gloria: farewell to the epidermal transit amplifying cell? *Cell stem cell* 1, 371-381.
- Jorgensen, H.F., Azuara, V., Amoils, S., Spivakov, M., Terry, A., Nesterova, T., Cobb, B.S., Ramsahoye, B., Merckenschlager, M., and Fisher, A.G. (2007). The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells. *Genome biology* 8, R169.
- Jorgensen, H.F., Giadrossi, S., Casanova, M., Endoh, M., Koseki, H., Brockdorff, N., and Fisher, A.G. (2006). Stem cells primed for action: polycomb repressive complexes restrain the expression of lineage-specific regulators in embryonic stem cells. *Cell cycle (Georgetown, Tex)* 5, 1411-1414.
- Kahata, K., Hayashi, M., Asaka, M., Hellman, U., Kitagawa, H., Yanagisawa, J., Kato, S., Imamura, T., and Miyazono, K. (2004). Regulation of transforming growth factor-beta and bone morphogenetic protein signalling by transcriptional coactivator GCN5. *Genes Cells* 9, 143-151.
- Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S., *et al.* (2004). Generation of pluripotent stem cells from neonatal mouse testis. *Cell* 119, 1001-1012.
- Kaneko, Y., Sakakibara, S., Imai, T., Suzuki, A., Nakamura, Y., Sawamoto, K., Ogawa, Y., Toyama, Y., Miyata, T., and Okano, H. (2000). Musashi1: an evolutionally

- conserved marker for CNS progenitor cells including neural stem cells. *Developmental neuroscience* 22, 139-153.
- Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). Transient cyclical methylation of promoter DNA. *Nature* 452, 112-115.
- Karnani, N., Taylor, C., Malhotra, A., and Dutta, A. (2007). Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas. *Genome Res* 17, 865-876.
- Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460, 1140-1144.
- Kikyo, N., and Wolffe, A.P. (2000). Reprogramming nuclei: insights from cloning, nuclear transfer and heterokaryons. *J Cell Sci* 113 (Pt 1), 11-20.
- 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.
- Kim, S.E., Kim, B.K., Gil, J.E., Kim, S.K., and Kim, J.H. (2007). Comparative analysis of the developmental competence of three human embryonic stem cell lines in vitro. *Molecules and cells* 23, 49-56.
- Kitsberg, D., Selig, S., Keshet, I., and Cedar, H. (1993). Replication structure of the human beta-globin gene domain. *Nature* 366, 588-590.
- Klimanskaya, I., Chung, Y., Becker, S., Lu, S.J., and Lanza, R. (2006). Human embryonic stem cell lines derived from single blastomeres. *Nature* 444, 481-485.
- Klose, R.J., and Zhang, Y. (2007). Regulation of histone methylation by demethylination and demethylation. *Nat Rev Mol Cell Biol* 8, 307-318.
- Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., *et al.* (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome research* 17, 691-707.
- Kontgen, F., Suss, G., Stewart, C., Steinmetz, M., and Bluethmann, H. (1993). Targeted disruption of the MHC class II Aa gene in C57BL/6 mice. *International immunology* 5, 957-964.
- Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. *Development* 126, 723-732.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- 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.
- Labosky, P.A., Barlow, D.P., and Hogan, B.L. (1994). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. *Development* 120, 3197-3204.
- Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. *Curr Opin Cell Biol* 14, 286-298.
- Lammerding, J., Hsiao, J., Schulze, P.C., Kozlov, S., Stewart, C.L., and Lee, R.T. (2005). Abnormal nuclear shape and impaired mechanotransduction in emerin-deficient cells. *J Cell Biol* 170, 781-791.
- 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.
- Lande-Diner, L., Zhang, J., and Cedar, H. (2009). Shifts in replication timing actively affect histone acetylation during nucleosome reassembly. *Mol Cell* 34, 767-774.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

- Latham, J.A., and Dent, S.Y. (2007). Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14, 1017-1024.
- Latt, S.A. (1973). Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. *Proc Natl Acad Sci U S A* 70, 3395-3399.
- Leahy, A., Xiong, J.W., Kuhnert, F., and Stuhlmann, H. (1999). Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation. *The Journal of experimental zoology* 284, 67-81.
- Lee, J.H., Hart, S.R., and Skalnik, D.G. (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* 38, 32-38.
- Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nature reviews* 8, 284-295.
- 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.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125, 301-313.
- Lendahl, U., Zimmerman, L.B., and McKay, R.D. (1990). CNS stem cells express a new class of intermediate filament protein. *Cell* 60, 585-595.
- Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865-873.
- Levenstein, M.E., Ludwig, T.E., Xu, R.H., Llanas, R.A., VanDenHeuvel-Kramer, K., Manning, D., and Thomson, J.A. (2006). Basic fibroblast growth factor support of human embryonic stem cell self-renewal. *Stem Cells* 24, 568-574.
- Levine, A.J., and Brivanlou, A.H. (2006). GDF3 at the crossroads of TGF-beta signaling. *Cell Cycle* 5, 1069-1073.
- Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 3, 662-673.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-926.
- Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009a). The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* 460, 1136-1139.
- Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* 134, 244-255.
- Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009b). Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. *Cell stem cell* 4, 16-19.
- Liao, J., Cui, C., Chen, S., Ren, J., Chen, J., Gao, Y., Li, H., Jia, N., Cheng, L., Xiao, H., *et al.* (2009). Generation of induced pluripotent stem cell lines from adult rat cells. *Cell stem cell* 4, 11-15.
- Lin, C.M., Fu, H., Martinovsky, M., Bouhassira, E., and Aladjem, M.I. (2003). Dynamic alterations of replication timing in mammalian cells. *Curr Biol* 13, 1019-1028.
- Lin, G., OuYang, Q., Zhou, X., Gu, Y., Yuan, D., Li, W., Liu, G., Liu, T., and Lu, G. (2007). A highly homozygous and parthenogenetic human embryonic stem cell line derived from a one-pronuclear oocyte following in vitro fertilization procedure. *Cell Res* 17, 999-1007.
- Liu, H., Zhu, F., Yong, J., Zhang, P., Hou, P., Li, H., Jiang, W., Cai, J., Liu, M., Cui, K., *et al.* (2008). Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts. *Cell stem cell* 3, 587-590.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., *et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38, 431-440.
- Lovell-Badge, R. (2007). Many ways to pluripotency. *Nature biotechnology* 25, 1114-1116.

- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., *et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* 1, 55-70.
- Mai, Q., Yu, Y., Li, T., Wang, L., Chen, M.J., Huang, S.Z., Zhou, C., and Zhou, Q. (2007). Derivation of human embryonic stem cell lines from parthenogenetic blastocysts. *Cell Res* 17, 1008-1019.
- Maison, C., Bailly, D., Peters, A.H., Quivy, J.P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 30, 329-334.
- Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassaei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., *et al.* (2005). Genomic alterations in cultured human embryonic stem cells. *Nat Genet* 37, 1099-1103.
- Malas, S., Duthie, S.M., Mohri, F., Lovell-Badge, R., and Episkopou, V. (1997). Cloning and mapping of the human SOX1: a highly conserved gene expressed in the developing brain. *Mamm Genome* 8, 866-868.
- Maldonado-Saldivia, J., van den Bergen, J., Krouskos, M., Gilchrist, M., Lee, C., Li, R., Sinclair, A.H., Surani, M.A., and Western, P.S. (2007). Dppa2 and Dppa4 are closely linked SAP motif genes restricted to pluripotent cells and the germ line. *Stem Cells* 25, 19-28.
- Maresca, T.J., and Heald, R. (2006). The long and the short of it: linker histone H1 is required for metaphase chromosome compaction. *Cell cycle (Georgetown, Tex)* 5, 589-591.
- Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 460, 1149-1153.
- Marshall, O.J. (2004). *PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR*. Bioinformatics (Oxford, England) 20, 2471-2472.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6, 838-849.
- 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.
- Massa, S., Junker, S., and Matthias, P. (2000). Molecular mechanisms of extinction: old findings and new ideas. *The international journal of biochemistry & cell biology* 32, 23-40.
- 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.
- Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* 436, 294-298.
- Mathur, D., Danford, T.W., Boyer, L.A., Young, R.A., Gifford, D.K., and Jaenisch, R. (2008). Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET. *Genome biology* 9, R126.
- 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.
- McGrath, J., and Solter, D. (1984). Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development in vitro. *Science* 226, 1317-1319.

- McNairn, A.J., and Gilbert, D.M. (2003). Epigenomic replication: linking epigenetics to DNA replication. *Bioessays* 25, 647-656.
- 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.
- Mechali, M. (2001). DNA replication origins: from sequence specificity to epigenetics. *Nat Rev Genet* 2, 640-645.
- Medina, K.L., and Singh, H. (2005). Genetic networks that regulate B lymphopoiesis. *Current opinion in hematology* 12, 203-209.
- Mendez, J. (2009). Temporal regulation of DNA replication in mammalian cells. *Critical reviews in biochemistry and molecular biology* 44, 343-351.
- Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 7, 540-546.
- 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.
- Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.
- Meuwissen, R.L., Offenberg, H.H., Dietrich, A.J., Riesewijk, A., van Iersel, M., and Heyting, C. (1992). A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *The EMBO journal* 11, 5091-5100.
- 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.
- Mileham, P., and Brown, S.D. (1996). YAC clone contigs covering 5 Mb of a repeat sequence island on the mouse X chromosome. *Mamm Genome* 7, 253-261.
- Miller, R.A., and Ruddle, F.H. (1976). Pluripotent teratocarcinoma-thymus somatic cell hybrids. *Cell* 9, 45-55.
- Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* 10, 1107-1117.
- Minichiotti, G. (2005). Nodal-dependant Cripto signaling in ES cells: from stem cells to tumor biology. *Oncogene* 24, 5668-5675.
- Mishina, Y., Suzuki, A., Ueno, N., and Behringer, R.R. (1995). Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes Dev* 9, 3027-3037.
- 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.
- Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *The Journal of biological chemistry* 279, 52353-52360.
- Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. *Hum Mol Genet* 14 Spec No 1, R47-58.
- Morrissey, E.E., Ip, H.S., Lu, M.M., and Parmacek, M.S. (1996). GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. *Dev Biol* 177, 309-322.
- Morrissey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev* 12, 3579-3590.
- Muller, H.J., and Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in *Drosophila*. *Genetics* 15, 283-311.
- Nakamura, T., Takio, K., Eto, Y., Shibai, H., Titani, K., and Sugino, H. (1990). Activin-binding protein from rat ovary is follistatin. *Science* 247, 836-838.

- Nakashima, K., Hagiwara, T., and Yamada, M. (2002). Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *The Journal of biological chemistry* 277, 49562-49568.
- Negre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during *Drosophila* development. *PLoS biology* 4, e170.
- Nichols, J., Chambers, I., Taga, T., and Smith, A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development* 128, 2333-2339.
- Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. *Cell stem cell* 4, 487-492.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., 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.
- Noggle, S.A., James, D., and Brivanlou, A.H. (2005). A molecular basis for human embryonic stem cell pluripotency. *Stem Cell Rev* 1, 111-118.
- O'Keefe, R.T., Henderson, S.C., and Spector, D.L. (1992). Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. *J Cell Biol* 116, 1095-1110.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.
- Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313-317.
- Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *The Journal of biological chemistry* 280, 5307-5317.
- Orlando, V., and Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr Opin Genet Dev* 5, 174-179.
- Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol* 26, 313-315.
- Pal, S., Yun, R., Datta, A., Lacomis, L., Erdjument-Bromage, H., Kumar, J., Tempst, P., and Sif, S. (2003). mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. *Mol Cell Biol* 23, 7475-7487.
- Palmieri, S.L., Peter, W., Hess, H., and Scholer, H.R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Dev Biol* 166, 259-267.
- Pannell, D., Osborne, C.S., Yao, S., Sukonnik, T., Pasceri, P., Karaiskakis, A., Okano, M., Li, E., Lipshitz, H.D., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. *Embo J* 19, 5884-5894.
- Parada, L.A., Sotiriou, S., and Misteli, T. (2004). Spatial genome organization. *Exp Cell Res* 296, 64-70.
- Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008a). Disease-specific induced pluripotent stem cells. *Cell* 134, 877-886.
- Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008b). Disease-Specific Induced Pluripotent Stem Cells. *Cell*.

- Park, I.H., Lerou, P.H., Zhao, R., Huo, H., and Daley, G.Q. (2008c). Generation of human-induced pluripotent stem cells. *Nat Protoc* 3, 1180-1186.
- 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.
- Pebay, A., Wong, R.C., Pitson, S.M., Wolvetang, E.J., Peh, G.S., Filipczyk, A., Koh, K.L., Tellis, I., Nguyen, L.T., and Pera, M.F. (2005). Essential roles of sphingosine-1-phosphate and platelet-derived growth factor in the maintenance of human embryonic stem cells. *Stem Cells* 23, 1541-1548.
- Pelton, T.A., Sharma, S., Schulz, T.C., Rathjen, J., and Rathjen, P.D. (2002). Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. *J Cell Sci* 115, 329-339.
- Pera, M.F., Andrade, J., Houssami, S., Reubinooff, B., Trounson, A., Stanley, E.G., Ward-van Oostwaard, D., and Mummery, C. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *J Cell Sci* 117, 1269-1280.
- Pera, M.F., Reubinooff, B., and Trounson, A. (2000). Human embryonic stem cells. *J Cell Sci* 113 (Pt 1), 5-10.
- Pera, M.F., and Trounson, A.O. (2004). Human embryonic stem cells: prospects for development. *Development* 131, 5515-5525.
- Pereira, C.F., and Fisher, A.G. (2009). Heterokaryon-based reprogramming for pluripotency. *Current protocols in stem cell biology Chapter 4*, Unit 4B 1.
- Pereira, C.F., Terranova, R., Ryan, N.K., Santos, J., Morris, K.J., Cui, W., Merckenschlager, M., and Fisher, A.G. (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS genetics* 4, e1000170.
- Pereira, L., Yi, F., and Merrill, B.J. (2006). Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol* 26, 7479-7491.
- Perry, P., Sauer, S., Billon, N., Richardson, W.D., Spivakov, M., Warnes, G., Livesey, F.J., Merckenschlager, 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.
- Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., *et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12, 1577-1589.
- Peters, A.H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. *Curr Opin Cell Biol* 17, 230-238.
- Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. *Curr Biol* 14, R546-551.
- Pfaffl, M.W., Lange, I.G., Daxenberger, A., and Meyer, H.H. (2001). Tissue-specific expression pattern of estrogen receptors (ER): quantification of ER alpha and ER beta mRNA with real-time RT-PCR. *Apmis* 109, 345-355.
- Pirrotta, V. (1998). Polycomb the genome: PcG, trxG, and chromatin silencing. *Cell* 93, 333-336.
- Pogo, B.G., Allfrey, V.G., and Mirsky, A.E. (1966). RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc Natl Acad Sci U S A* 55, 805-812.
- Pralong, D., Mrozik, K., Occhiodoro, F., Wijesundara, N., Sumer, H., Van Boxtel, A.L., Trounson, A., and Verma, P.J. (2005). A novel method for somatic cell nuclear transfer to mouse embryonic stem cells. *Cloning Stem Cells* 7, 265-271.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433, 434-438.
- Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nature reviews* 10, 192-206.

- 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.
- Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 135, 1201-1212.
- Ralston, A., and Rossant, J. (2005). Genetic regulation of stem cell origins in the mouse embryo. *Clin Genet* 68, 106-112.
- 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.
- Reijo Pera, R.A., DeJonge, C., Bossert, N., Yao, M., Hwa Yang, J.Y., Asadi, N.B., Wong, W., Wong, C., and Firpo, M.T. (2009). Gene expression profiles of human inner cell mass cells and embryonic stem cells. *Differentiation; research in biological diversity* 78, 18-23.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-432.
- Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* 293, 1089-1093.
- 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.
- Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat Biotechnol* 18, 399-404.
- Revazova, E.S., Turovets, N.A., Kochetkova, O.D., Kindarova, L.B., Kuzmichev, L.N., Janus, J.D., and Pryzhkova, M.V. (2007). Patient-specific stem cell lines derived from human parthenogenetic blastocysts. *Cloning Stem Cells* 9, 432-449.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* 38, 413-443.
- Rogers, M.B., Hosler, B.A., and Gudas, L.J. (1991). Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development* 113, 815-824.
- Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt, L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* 345, 686-692.
- Ross, S., Cheung, E., Petrakis, T.G., Howell, M., Kraus, W.L., and Hill, C.S. (2006). Smads orchestrate specific histone modifications and chromatin remodeling to activate transcription. *The EMBO journal* 25, 4490-4502.
- Ross, S., and Hill, C.S. (2008). How the Smads regulate transcription. *The international journal of biochemistry & cell biology* 40, 383-408.
- Rossant, J. (2001). Stem cells from the Mammalian blastocyst. *Stem Cells* 19, 477-482.
- Rossant, J. (2008). Stem cells and early lineage development. *Cell* 132, 527-531.
- Rossant, J., and Tam, P.P. (2004). Emerging asymmetry and embryonic patterning in early mouse development. *Dev Cell* 7, 155-164.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C.D., Reinberg, D., Avner, P., and Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* 24, 5475-5484.
- Rudnicki, M.A., Schnegelsberg, P.N., Stead, R.H., Braun, T., Arnold, H.H., and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* 75, 1351-1359.
- Ruggiu, M., Speed, R., Taggart, M., McKay, S.J., Kilanowski, F., Saunders, P., Dorin, J., and Cooke, H.J. (1997). The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* 389, 73-77.
- Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C., et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* 404, 95-99.

- Russo, V.E.A., Martienssen, R.A., and Riggs, A.D. (1996). Epigenetic mechanisms of gene regulation (Plainview, N.Y., Cold Spring Harbor Laboratory Press).
- Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* 25, 15-30.
- Sakakibara, S., Imai, T., Hamaguchi, K., Okabe, M., Aruga, J., Nakajima, K., Yasutomi, D., Nagata, T., Kurihara, Y., Uesugi, S., *et al.* (1996). Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. *Dev Biol* 176, 230-242.
- Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. (2006). Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* 311, 1118-1123.
- 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.
- Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med* 10, 55-63.
- Schmierer, B., and Hill, C.S. (2007). TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nature reviews* 8, 970-982.
- Schneider, R., Bannister, A.J., and Kouzarides, T. (2002). Unsafe SETs: histone lysine methyltransferases and cancer. *Trends in biochemical sciences* 27, 396-402.
- Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. *Nature* 344, 435-439.
- Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887-898.
- Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., *et al.* (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18, 1263-1271.
- Schubeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J., and Groudine, M. (2002). Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. *Nat Genet* 32, 438-442.
- Schurter, B.T., Koh, S.S., Chen, D., Bunick, G.J., Harp, J.M., Hanson, B.L., Henschen-Edman, A., Mackay, D.R., Stallcup, M.R., and Aswad, D.W. (2001). Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry* 40, 5747-5756.
- Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat Genet* 38, 700-705.
- Selig, S., Ariel, M., Goitein, R., Marcus, M., and Cedar, H. (1988). Regulation of mouse satellite DNA replication time. *The EMBO journal* 7, 419-426.
- Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annual review of biochemistry* 76, 75-100.
- Shi, W., Wang, H., Pan, G., Geng, Y., Guo, Y., and Pei, D. (2006). Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. *The Journal of biological chemistry* 281, 23319-23325.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, 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.
- Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annual review of biochemistry* 75, 243-269.
- Shumaker, D.K., Dechat, T., Kohlmaier, A., Adam, S.A., Bozovsky, M.R., Erdos, M.R., Eriksson, M., Goldman, A.E., Khuon, S., Collins, F.S., *et al.* (2006). Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* 103, 8703-8708.

- Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. *Nature* 441, 997-1001.
- Simon, I., Tenzen, T., Reubinoff, B.E., Hillman, D., McCarrey, J.R., and Cedar, H. (1999). Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* 401, 929-932.
- Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature reviews* 10, 697-708.
- Sims, R.J., 3rd, and Reinberg, D. (2006). Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev* 20, 2779-2786.
- Sirard, C., de la Pompa, J.L., Elia, A., Itie, A., Mirtsos, C., Cheung, A., Hahn, S., Wakeham, A., Schwartz, L., Kern, S.E., *et al.* (1998). The tumor suppressor gene *Smad4/Dpc4* is required for gastrulation and later for anterior development of the mouse embryo. *Genes Dev* 12, 107-119.
- Skottman, H., Stromberg, A.M., Matilainen, E., Inzunza, J., Hovatta, O., and Lahesmaa, R. (2006). Unique gene expression signature by human embryonic stem cells cultured under serum-free conditions correlates with their enhanced and prolonged growth in an undifferentiated stage. *Stem Cells* 24, 151-167.
- Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336, 688-690.
- Smith, Z.E., and Higgs, D.R. (1999). The pattern of replication at a human telomeric region (16p13.3): its relationship to chromosome structure and gene expression. *Hum Mol Genet* 8, 1373-1386.
- Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6, 846-856.
- Spivakov, M., and Fisher, A.G. (2007). Epigenetic signatures of stem-cell identity. *Nat Rev Genet* 8, 263-271.
- Stevanovic, M., Lovell-Badge, R., Collignon, J., and Goodfellow, P.N. (1993). SOX3 is an X-linked gene related to SRY. *Hum Mol Genet* 2, 2013-2018.
- Stewart, C.L., Gadi, I., and Bhatt, H. (1994). Stem cells from primordial germ cells can reenter the germ line. *Dev Biol* 161, 626-628.
- Stewart, M.H., Bosse, M., Chadwick, K., Menendez, P., Bendall, S.C., and Bhatia, M. (2006). Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment. *Nat Methods* 3, 807-815.
- Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterisation of a yeast chromosomal replicator. *Nature* 282, 39-43.
- 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.
- Stoffers, D.A., Zinkin, N.T., Stanojevic, V., Clarke, W.L., and Habener, J.F. (1997). Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nat Genet* 15, 106-110.
- Stojkovic, M., Lako, M., Stojkovic, P., Stewart, R., Przyborski, S., Armstrong, L., Evans, J., Herbert, M., Hyslop, L., Ahmad, S., *et al.* (2004). Derivation of human embryonic stem cells from day-8 blastocysts recovered after three-step in vitro culture. *Stem cells (Dayton, Ohio)* 22, 790-797.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R., *et al.* (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol* 11, 996-1000.
- Strelchenko, N., Verlinsky, O., Kukharensko, V., and Verlinsky, Y. (2004). Morula-derived human embryonic stem cells. *Reprod Biomed Online* 9, 623-629.

- Stubblefield, E. (1975). Analysis of the replication pattern of Chinese hamster chromosomes using 5-bromodeoxyuridine suppression of 33258 Hoechst fluorescence. *Chromosoma* 53, 209-221.
- Sulzbacher, S., Schroeder, I.S., Truong, T.T., and Wobus, A.M. (2009). Activin A-induced differentiation of embryonic stem cells into endoderm and pancreatic progenitors-the influence of differentiation factors and culture conditions. *Stem cell reviews and reports* 5, 159-173.
- Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. *Cell* 128, 747-762.
- Tada, M., Tada, T., Lefebvre, L., Barton, S.C., and Surani, M.A. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *The EMBO journal* 16, 6510-6520.
- 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., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.
- Tanabe, H., Muller, S., Neusser, M., von Hase, J., Calcagno, E., Cremer, M., Solovei, I., Cremer, C., and Cremer, T. (2002). Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci U S A* 99, 4424-4429.
- 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.
- Terranova, R., Pereira, C.F., Du Roure, C., Merckenschlager, M., and Fisher, A.G. (2006). Acquisition and extinction of gene expression programs are separable events in heterokaryon reprogramming. *J Cell Sci* 119, 2065-2072.
- 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.
- Thiagalingam, S., Cheng, K.H., Lee, H.J., Mineva, N., Thiagalingam, A., and Ponte, J.F. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. *Annals of the New York Academy of Sciences* 983, 84-100.
- 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.
- Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., and Hearn, J.P. (1995). Isolation of a primate embryonic stem cell line. *Proc Natl Acad Sci U S A* 92, 7844-7848.
- Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., and Hearn, J.P. (1996). Pluripotent cell lines derived from common marmoset (*Callithrix jacchus*) blastocysts. *Biol Reprod* 55, 254-259.
- 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., 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.
- Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* 128, 651-654.

- 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.
- Turner, B.M. (2000). Histone acetylation and an epigenetic code. *Bioessays* 22, 836-845.
- Turner, B.M. (2002). Cellular memory and the histone code. *Cell* 111, 285-291.
- Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 460, 1145-1148.
- Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* 118, 4495-4509.
- 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., and Pedersen, R.A. (2005). Human embryonic stem cells: an in vitro model to study mechanisms controlling pluripotency in early mammalian development. *Stem Cell Rev* 1, 119-130.
- Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A., *et al.* (2009b). Signalling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *Stem Cells*.
- Van Houten, J.V., and Newlon, C.S. (1990). Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol Cell Biol* 10, 3917-3925.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 3, RESEARCH0034.
- Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. *Mol Cell* 10, 1223-1233.
- Waddington, C.H. (1942). "The epigenotype". *Endeavour* 1, 18-20.
- Wade, P.A., Pruss, D., and Wolffe, A.P. (1997). Histone acetylation: chromatin in action. *Trends in biochemical sciences* 22, 128-132.
- Wakayama, S., Hikichi, T., Suetsugu, R., Sakaide, Y., Bui, H.T., Mizutani, E., and Wakayama, T. (2007). Efficient establishment of mouse embryonic stem cell lines from single blastomeres and polar bodies. *Stem Cells* 25, 986-993.
- Walther, C., and Gruss, P. (1991). Pax-6, a murine paired box gene, is expressed in the developing CNS. *Development* 113, 1435-1449.
- Wang, G., Zhang, H., Zhao, Y., Li, J., Cai, J., Wang, P., Meng, S., Feng, J., Miao, C., Ding, M., *et al.* (2005). Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. *Biochemical and biophysical research communications* 330, 934-942.
- Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., *et al.* (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* 293, 853-857.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004a). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873-878.
- Wang, L., Schulz, T.C., Sherrer, E.S., Dauphin, D.S., Shin, S., Nelson, A.M., Ware, C.B., Zhan, M., Song, C.Z., Chen, X., *et al.* (2007). Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood* 110, 4111-4119.
- Wang, Y., Wysocka, J., Sayegh, J., Lee, Y.H., Perlin, J.R., Leonelli, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y., *et al.* (2004b). Human PAD4 regulates histone arginine methylation levels via demethyliminination. *Science* 306, 279-283.

- 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.
- Ware, C.B., Wang, L., Mecham, B.H., Shen, L., Nelson, A.M., Bar, M., Lamba, D.A., Dauphin, D.S., Buckingham, B., Askari, B., *et al.* (2009). Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. *Cell stem cell* 4, 359-369.
- Waterborg, J.H. (2002). Dynamics of histone acetylation in vivo. A function for acetylation turnover? *Biochemistry and cell biology = Biochimie et biologie cellulaire* 80, 363-378.
- Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37, 853-862.
- Wei, C.L., Miura, T., Robson, P., Lim, S.K., Xu, X.Q., Lee, M.Y., Gupta, S., Stanton, L., Luo, Y., Schmitt, J., *et al.* (2005). Transcriptome profiling of human and murine ESCs identifies divergent paths required to maintain the stem cell state. *Stem Cells* 23, 166-185.
- Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318-324.
- Widom, J. (1989). Toward a unified model of chromatin folding. *Annual review of biophysics and biophysical chemistry* 18, 365-395.
- Wiles, M.V., and Johansson, B.M. (1999). Embryonic stem cell development in a chemically defined medium. *Experimental cell research* 247, 241-248.
- Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* 336, 684-687.
- Williams, R.R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jorgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., *et al.* (2006). Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J Cell Sci* 119, 132-140.
- Wilmot, 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.
- Wintersberger, E. (2000). Why is there late replication? *Chromosoma* 109, 300-307.
- Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I., and Carter, N.P. (2004). Replication timing of the human genome. *Human molecular genetics* 13, 191-202.
- Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 5, 695-705.
- 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.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86-90.
- Xiao, L., Yuan, X., and Sharkis, S.J. (2006). Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. *Stem Cells* 24, 1476-1486.
- Xie, W., Song, C., Young, N.L., Sperling, A.S., Xu, F., Sridharan, R., Conway, A.E., Garcia, B.A., Plath, K., Clark, A.T., *et al.* (2009). Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell* 33, 417-427.

- Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol* 19, 971-974.
- Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol* 20, 1261-1264.
- Xu, R.H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nature methods* 2, 185-190.
- Xu, R.H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., *et al.* (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell stem cell* 3, 196-206.
- Yang, X.J., and Seto, E. (2008a). Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol Cell* 31, 449-461.
- Yang, X.J., and Seto, E. (2008b). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nature reviews* 9, 206-218.
- Yeo, S., Jeong, S., Kim, J., Han, J.S., Han, Y.M., and Kang, Y.K. (2007). Characterization of DNA methylation change in stem cell marker genes during differentiation of human embryonic stem cells. *Biochem Biophys Res Commun* 359, 536-542.
- Yi, F., Pereira, L., and Merrill, B.J. (2008). Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. *Stem Cells* 26, 1951-1960.
- Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115, 281-292.
- Ying, Q.L., Nichols, J., Evans, E.P., and Smith, A.G. (2002). Changing potency by spontaneous fusion. *Nature* 416, 545-548.
- Ying, Q.L., Wray, J., Nichols, J., Battle-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. *Mechanisms of development* 45, 163-171.
- Yu, J., and Thomson, J.A. (2008). Pluripotent stem cell lines. *Genes Dev* 22, 1987-1997.
- Yu, J., Vodyanik, M.A., He, P., Slukvin, I., and Thomson, J.A. (2006). Human embryonic stem cells reprogram myeloid precursors following cell-cell fusion. *Stem Cells* 24, 168-176.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917-1920.
- Zaehres, H., Lensch, M.W., Daheron, L., Stewart, S.A., Itskovitz-Eldor, J., and Daley, G.Q. (2005). High-efficiency RNA interference in human embryonic stem cells. *Stem Cells* 23, 299-305.
- Zhang, J., Xu, F., Hashimshony, T., Keshet, I., and Cedar, H. (2002). Establishment of transcriptional competence in early and late S phase. *Nature* 420, 198-202.

APPENDIX I – PRIMER SEQUENCES

Table A1. Primers for the amplification of human or mouse transcripts by real time quantitative PCR

Human primers		
Species/Gene		Sequence 5'-3'
<i>hβ-tubulin</i>	S	CAGATGTTTCGATGCCAAGAA
	as	TGCTGTTCTTGCTCTGGATG
<i>hAfp</i>	S	AAATGCGTTTCTCGTTGCTT
	as	GCCACAGGCCAATAGTTTGT
<i>hBmp4</i>	S	TGGTCTTGAGTATCCTGAGCG
	as	GCTGAGGTTAAAGAGGAAACGA
<i>hBrg1</i>	S	TGGACTACAGCGACTCACTGA
	as	TCGATGGCCTTTGACGCAC
<i>hCBP</i>	S	GACAAGCGAAACCAACAAACC
	as	TGGGGTCTATGGGATTTGGGT
<i>hCD19</i>	S	GCTCAAGACGCTGGAAAGTATTATT
	as	GATAAGCCAAAGTCACAGCTGAGA
<i>hCD45</i>	S	CCCCATGAACGTTACCATTG
	as	GATAGTCTCCATTGTGAAAATAGGCC
<i>hCdx2</i>	S	CAGGACGAAAGACAAATATCGAGTG
	as	CCAGATTTTAACCTGCCTCTCAGA
<i>hc-Myc</i>	S	TTGTACCTG CAGGATCTGAG
	as	CTTGTTCTCCTCAGAGTCG
<i>hCripto</i>	S	AGAAGTGTTCCCTGTGTAAATGCTG
	as	CACGAGGTGCTCATCCATCA
<i>hDnmt3b</i>	S	GTCAAGCTACACACAGGACTTGACAG
	as	AGTTCGGACAGCTGGGCTTT
<i>hEbf</i>	S	GGAGATCGAGAGGACAGCGT
	as	GTCAATGAGGCGCACGTAGA
<i>hEomes</i>	S	CAAAGGCGCAAATAACAACAAC
	as	CCACGCCATCCTCTGTAACCTC
<i>hEsrrb</i>	S	TGAGGCCAGAGGTGATCCAG
	as	TCTTGATGAAGGAGCCGCAG
<i>hFgf5</i>	S	CAGCACCAAAGGCTCAGCTT
	as	CCTTGCTTCTAACCCATCATATCC
<i>hFoxa2</i>	S	GGGAGCGGTGAAGATGGA
	as	TCATGTTGCTCACGGAGGAGTA
<i>hGapdh</i>	S	TCTGCTCCTCCTGTTTCGACA
	as	AAAAGCAGCCCTGGTGACC
<i>hGata6</i>	S	ACCACCTTATGGCGCAGAAAC
	as	TTTTTCATAGCAAGTGGTCTGGG

<i>hGcn5</i>	S	GCTACATCAAGGACTACGAG
	as	GCTTCTTGATGATCTCTTTCTG
<i>hGdf3</i>	S	GTA CTTCGCTTTCTCCAGAC
	as	GCCAATGTCAACTGTTCCCTT
<i>hHprt</i>	S	TCCTTGGTCAGGCAGTATAATCC
	as	GTCAAGGGCATATCCTACAACAAA
<i>hKlf4</i>	S	CACATTAATGAGGCAGCCACC
	as	AAGTCGCTTCATGTGGGAGAG
<i>hLin28</i>	S	TGCGGGCATCTGTAAGTGG
	as	GGAACCCTTCCATGTGCAG
<i>hNanog</i>	S	CCAACATCCTGAACCTCAGCTAC
	as	GCCTTCTGCGTCACACCATT
<i>hNodal</i>	S	AGAAGCAGATGTCCAGGGTAGCT
	as	GAGAGGTTGGAGTAGAGCATAAGGA
<i>hOct4</i>	S	TCGAGAACCGAGTGAGAGGC
	as	CACACTCGGACCACATCCTTC
<i>hp300</i>	S	TCAGCCAAGCGGCCTAAAC
	as	TCACCACCATTGGTTAGTCCC
<i>hP63</i>	S	TTTCAGAGGCAATCCACACA
	as	ATGCATGCAAATGAGCTCTG
<i>hPcaf</i>	S	CGAATCGCCGTGAAGAAAGC
	as	GAGGGGTTAGGGTTTTTCCAG
<i>hPdx1</i>	S	CCTTTCCCATGGATGAAGTC
	as	GGAACTCCTTCTCCAGCTCTA
<i>hPecam1</i>	S	AACAGTGTTGACATGAAGAGCC
	as	TGTAAACAGCACGTCATCCTT
<i>hPPAR-γ</i>	S	GCTGGCCTCCTTGATGAATA
	as	TTGGGCTCCATAAAGTCACC
<i>hRex1</i>	S	GCGTACGCAAATTAAAGTCCAGA
	as	CAGCATCCTAAACAGCTCGCAGAAT
<i>hSox1</i>	S	AACACTTGAAGCCCAGATGGA
	as	GCAGGCTGAATTCGGTTCTC
<i>hSox17</i>	S	GGCGCAGCAGAATCCAGA
	as	CCACGACTTGCCCAGCAT
<i>hSox2</i>	S	CACACTGCCCCTCTCACACAT
	as	CATTTCCCTCGTTTTTCTTTGAA
<i>hSox3</i>	S	TTTGATAAGCCTACCCTTCC
	as	GTTCTTGAGTTCAGTCTCCA
<i>hTcf3</i>	S	GCCTCCCGACTCCTACAGT
	as	CGCTGACGTGTTCTCCTCG
<i>hTert</i>	S	GCCAGCATCATCAAACCCC
	as	CTGTCAAGGTAGAGACGTGGCTC
<i>hTle1</i>	S	TGTCTCCCAGCTCGACTGTCT
	as	AAGTACTGGCTTCCCCTCCC

<i>hZic1</i>	S	CTGGCTGTGGCAAGGTCTTC
	as	CAGCCCTCAAACCTCGCACTT
Mouse primers		
Species/Gene		Sequence 5'-3'
<i>mActin</i>	S	CGAGGCCCAGAGCAAGAG
	as	CGTCCCAGTTGGTAACAATGC
<i>mCD19</i>	S	GGAGAGCACCCGGTCAGA
	as	CCACACTGCTGACCTTGCAA
<i>mc-Myc</i>	S	TTCATCTGCGATCCTGACGAC
	as	AGGGGTCAATGCACTCGGA
<i>mCripto</i>	S	CACCAACCCAGGGTATCAGTT
	as	AGAGTTCTGTCCAGTGTCGTC
<i>mDnmt3b</i>	S	ACTGCCTGGAGTTCAGTAGGA
	as	CCCTGTCTGATGGAGTTCGAC
<i>mEomes</i>	S	CCTGGTGGTGTTTTGTTGTG
	as	TTTAATAGCACCGGGCACTC
<i>mEsrrb</i>	S	GCACCTGGGCTCTAGTTGC
	as	TACAGTCCTCGTAGCTCTTGC
<i>mFgf5</i>	S	TGTGTCTCAGGGGATTGTAGG
	as	AGCTGTTTTCTTGGAATCTCTC
<i>mFoxa2</i>	S	TAGCGGAGGCAAGAAGACC
	as	CTTAGGCCACCTCGCTTGT
<i>mGapdh</i>	S	TGCACCACCAACTGCTTAGC
	as	GGCATGGACTGTGGTCATGAG
<i>mGata6</i>	S	GACTCCTACTTCCTCTTCTTAATTCAGA
	as	ACCTGAATACTTGAGGTCACTGTTCTC
<i>mGdf3</i>	S	GGCCTCGCAGGACTTATGC
	as	TGGTCGCAGGTTATAGTAGGAC
<i>mHprt</i>	S	TTGAAATTCCAGACAAGTTTGTTGTTGG
	as	CCTGCTAATTTTACTGGCAACATCAACA
<i>mKlf4</i>	S	TCCTTTCCAACCTCGCTAACCC
	as	CGGATCGGATAGCTGAAGCTG
<i>ml-19</i>	S	TGATCTGCTGACGGAGTTG
	as	GGAAAAGAAGGTCTGGTTGGA
<i>mLin28</i>	S	TCACTGGCCCTGGTGGTGTGTT
	as	TGGACACGAGGCCACCATATGG
<i>mNanog</i>	S	CTTACAAGGGTCTGCTACTGA
	as	TCTGCTTCCTGGCAAGGACC
<i>mNodal</i>	S	ACCATGCCTACATCCAGAGC
	as	ATCAGCATTGTGGAATGCAA
<i>mOct4</i>	S	CGTGGAGACTTTGCAGCCTG
	as	GCTTGGCAAACCTGTTCTAGCTCCT
<i>mRex1</i>	S	CTCCTAGCCGCCTAGATTTCCA
	as	CGTGTCCCAGCTCTTAGTCCATT

<i>mSox17</i>	s	TTTAAATGGGAGGGAGGGTC
	as	TCAGATGTCTGGAGGTGCTG
<i>mSox2</i>	s	GAGTGGAACTTTTGTCCGAGA
	as	GAAGCGTGTACTTATCCTTCTTCAT
<i>mTcf3</i>	s	TTTTAAGGGAGGTGCGTATCCT
	as	GGGGCCATTTTCATCTGTAGGTA
<i>mTert</i>	s	TCAAGAGCAGTAGTCGCCAG
	as	TCTCGGGACAGGATAGCATCT
<i>mTle1</i>	s	TCTGAGTGGGCAGTCTCACTT
	as	GAAGAAGGGTCCTCGTTAGACA

s indicates the sense and as the anti-sense primer.

Table A2. Primers used for the amplification of human transcripts by semi-quantitative RT-PCR

Human primers				
<i>Species/Gene</i>	<i>Tm</i>	<i>n</i>		<i>Sequence 5'-3'</i>
<i>hβ-actin</i>	60°C	25	s	TCACCACCACGGCCGAGCG
			as	TCTCCTTCTGCATCCTGTCTG
<i>hGata6</i>	60°C	25	s	CCATGACTCCAACTTCCACC
			as	ACGGAGGACGTGACTTCGGC
<i>hNanog</i>	60°C	25	s	AGCCTCTACTCTTCTACCACC
			as	TCCAAAGCAGCCTCCAAGTC
<i>hNestin</i>	65°C	25	s	CAGCTGGCGCACCTCAAGATG
			as	AGGGAAGTTGGGCTCAGGACTG
<i>hOct4</i>	60°C	25	s	CTTGCTGCAGAAAGTGGGTGGAGGAA
			as	CTGCAGTGTGGGTTTCGGGCA
<i>hPax6</i>	60°C	25	s	AACAGACACAGCCCTCACAAACA
			as	CGGGAAGTTGAACTGGAAGTAC
<i>hSox1</i>	60°C	25	s	CAATGCGGGGAGGAGAAGTC
			as	CTCTGGACCAAAGTGTGGCG
<i>hSox2</i>	60 °C	25	s	CCCCCGGCGGCAATAGCA
			as	TCGGCGCCGGGAGATACAT
<i>hT</i>	63°C	30	s	GTGACCAAGAACGGCAGGAGG
			as	TGTTCCGATGAGCATAGGGGC

Tm indicates the annealing temperature of the primers and *n* the number of cycles of amplification used in PCR reactions. s indicates the sense and as the anti-sense primer.

Table A3. Primers for the selective amplification of human or mouse genomic regions by real time quantitative PCR

Human specific primers for replication timing, ChIP and MeDIP analysis		
<i>Species/Gene</i>		<i>Sequence 5'-3'</i>
h α -globin promoter	S	TTTGTCCCCGGACCTGCTGCC
	as	CCCCACGGTCGGCTGACAC
hAmylase promoter	S	GGTGAGTCTGTGTGGTCAGCAGTCTCT
	as	CCACGGTGCTCTGGTAGATAACGTAAG
hBmp4 promoter	S	CAGCTCCTTTCTCCATAACCTGA
	as	GGTACTAGAAAGCATGCACCGAC
hCdx2 promoter	S	AACAACCACTGCTCCTGTCTCC
	as	CGACCTGACACAGCTAAATATTCAA
hCga promoter	S	ATGGCTCCAAACAAAAATGACC
	as	TCCTCACCTGCTTTTATACCAGC
hCripto promoter	S	ACTTCAAGTCTGGAGCCCCC
	as	GCCTGAGGAGCTAGGTGTGTGT
hCxcr4 promoter	S	GAAGTCACTATGGGAAAAGATGGG
	as	GCTTGCTTTCTTCAGGAAATTCTG
hDazl promoter	S	TCCTGAATGTCAAACCTGGCATT
	as	CGCCATCACTTTTAGCTTCACA
hDnmt3b promoter	S	TGTGTGTCTCCGTTCCGGTT
	as	ATCAGAAGCCCTAAGCGGGA
hDppa4 promoter	S	AAATAGCTTTCTTCAAGCCGCC
	as	CGGAGACATTGGGAGATTGAG
hEbf promoter	S	ACTTGAAGTAGCAAAGCCGGCA
	as	TGTAATGATCACAGGCCGGTGG
hEomes promoter	S	AGGAAAAGAAAGTCACAGGCGA
	as	CATCGGTCAAGTTGACCACTTG
hEsx1 promoter	S	AGCATGTTCACCTGCACCA
	as	TATAGCGGTACGGTGCCCTCC
hFgf5 promoter	S	AAAAACAGCCTTCTCCCCGAAC
	as	CACTGCTTCTACGTCTTTCGGC
hFoxa2 promoter	S	CCCTTACTCAAGCTTCAATGCG
	as	TTCTTGTTCTGAAAACCTGGC
hGata4 promoter	S	CGGTAGCACTTGGGCATTTT
	as	AGCAGGCAAAGTCCAGGCT
hGata6 promoter	S	ACGACCTGAGCCGTAGCAT
	as	CCTAGCCCCTTCACAGCAAA
hGbx2 promoter	S	GGCAGGCAAAATGTGAATGAG
	as	GAAGCCGGCGTACTTATCTCC
hGdf3 promoter	S	AAAGCCATCGTGAGCCTCAT
	as	GACAGAGTGAAAAGGAGACATGCA

hHoxb1 promoter	S	CTAGCCCACCACCCTAAGACAA
	as	GGCTCCTGGGTTATTAGATTGG
hIgf2/hH19 ICR	S	ACATTACACGAGCATCCAGG
	as	GCTCTTTAGGTTTGGCGCAAT
hIkaros promoter	S	GAACAATGCGAGTGAGCAACTTCAGG
	as	GGCCGAGCGTCCCGCCCAGGCTG
hMash1 promoter	S	CAATTTCTAGGGTCACCGAGGA
	as	GAACGGTGTGGAGGTGAGGA
hMixl1 promoter	S	GTTTGACCAACCTGGTACAGGG
	as	GCAGGCGACAGAACTCTCTTCT
hMyf5 promoter	S	TGTCTGTAGAGGAAAGGCGGA
	as	ATATTCTTTCGGGAGGGAGGC
hNanog promoter	S	ACGGCCTCCCAATTTACTGG
	as	GGTTCAACAGGAATGGGATAAAAC
hNestin promoter	S	TTTCCAGGCAGCGTCTCTCTAG
	as	GGATGTGGCCAGGTTTTTGT
hNeuroD promoter	S	CTTCACTGCGTGCCTCAGTCT
	as	GAGAGATTAACCCTTTCAGGCG
hNeurog1 promoter	S	TTGTTGCGCCGGTACTTAAG
	as	GCCGGTCTCCTGAGTGATGT
hOct4/Pou5f1 promoter	S	TTGGGAGTTGAAAGTTGGGTGT
	as	AGGCTGGTCTTGAATTCCTGTC
hPax6 promoter	S	AAAGGACAGCACAGAACTTGC
	as	TCAAGTGAAGGTTTCTGGGTTG
hPdx1 promoter	S	GAATGCCAGAGTTTCGTGTGTTT
	as	CCCAGCTCTCTAGGTCAGAAGC
hPsc1/Rbm27 promoter	S	GTGAACGTATTGCACGTTGTGTTGC
	as	TCAGCCTGGTTCCTTCCTCTCGTTT
hPrce/Esp1 promoter	S	AGGAAGGAAGAGCGGAATCGT
	as	GGAAGGAGGCAAAGAATCACAT
hRex1/Zfp42 promoter	S	CTCATGTGATCCCCCGTCT
	as	TACGCGTGGGTGTAATCACATT
hSatellite repeats	S	ATCGAATGGAAATGAAAGGAGTCA
	as	GACCATTGGATGATTGCAGTCA
hSox1 promoter	S	TCTTTGGCAAGTGTTTGTGC
	as	GGAGGCAACGACAACAAAAA
hSox17 promoter	S	TATCAACGGTGTCTTTCGCATT
	as	AAGCCGAGGGTGTTTCAGTG
hSox2 promoter	S	CGTGAGAGAGTGTTGGCACCT
	as	TTGTTCTCCCGCTCATCCAC
hSox3 promoter	S	AATGGCCACAGAGCAGAGACA
	as	TGTGGAAAGTATGTGTGGGACG
hSox7 promoter	S	TGGCAGTCGGAGTATTAAGAACG
	as	TCAGAAAATCAACCCAACCAGG

hSycp1 promoter	s	GAGCCTGACAGTTTCCCTAGCA
	as	CTGGCAGTTCAACGGCTAAATA
hTek1 promoter	s	GAGTGACTTCTGGAGTGTGGCC
	as	GAGGGTTAAAAATCAGGAAGCCA
hTert promoter	s	TCACGTCCGGCATTCTGTT
	as	TCGAATCGGCCTAGGCTGTG
hTle1 promoter	s	CTGTTTTCCCCACCATCTCGT
	as	TCCTTCACCTTCGTGTGCTTCT
hWnt3 promoter	s	CGCGTGATTGACAGGCTG
	as	CAGCCGGGTTTGAGGATGT
hZic1 promoter	s	CTGTGTTTACTAGCCCCATCCC
	as	ATTCCTCTTCCTCCTTTGCC
<i>D. melanogaster EMC</i>	s	GGTTTCAAGTTGTGGTCCAAGTTCC
	as	GTTCGTTTCGGCTGGAGAGAGAT

Mouse specific primers for replication timing, ChIP and MeDIP analysis

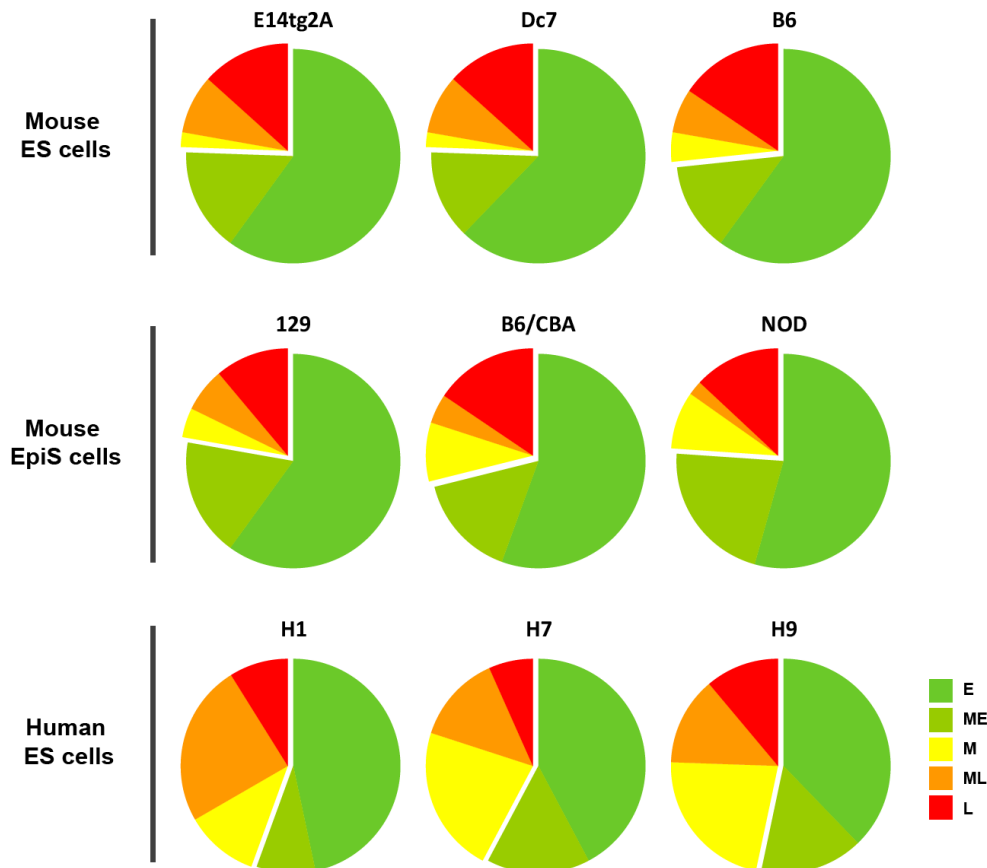
Species/Gene		Sequence 5'-3'
m α -globin promoter	s	CCACAAGCTGCGTGTGGAT
	as	ATGCCGCCTGCCAGGT
mAmylase promoter	s	CAAAGCAAAATGAAGTTCGTTCTGCTGC
	as	GAAGTTATCTTACCTGCACCCCTCCAAATC
mBmp4 promoter	s	TTGAAGAGGTATCTCCTTCGGG
	as	GATTGGTTTAAAGGGTGGCCT
mCdx2 promoter	s	ACCACCTTCTGCCTGAGAATGTAC
	as	CCTCCAATCACAGGTTCAAAGACT
mCga promoter	s	GCCATCCAATCACTGGGTAGA
	as	TCCCACTGTCTTTTATACCAGC
mCripto promoter	s	CTGTATAGGGTCAGCACTTCCAGCC
	as	CCCCGGAGGCACCCTTTACT
mCxcr4 promoter	s	CCTTCCTCCTTTGGCTAATCCT
	as	TCACTCTCAAACTGTGTGCGG
mDazl promoter	s	TGCCGCCCGTTATTAATTC
	as	CAAAAATGGCACCTCGAAGG
mDnmt3b promoter	s	CGAAAAGCCAATGCAGGGTTGCTGT
	as	GCACACACGCACATACAAGCACATTAC
mDppa4 promoter	s	TCCCAGATCTACAGTTGCCAGG
	as	GGCCAGAAATGCGCTTTGT
mEbf promoter	s	ATTGGCTGTGATAAGGCGGAAGGATGACG
	as	CCCCAGAGCTAAACACAGGCACACTAA
mEomes promoter	s	CCTCCTCTCACCCCAACAGAGCGAA
	as	AACCGATGTCTAGCTTGTGGTCACAGG
mEsx1 promoter	s	GCGAACTCGGAGCTGCTAAA
	as	TCCACGTGCTGTAACTGACTTTT

mFgf5 promoter	s	CATGTGTCCCAGCCAACTCACT
	as	TTCCCAGGCTCCAGATGGC
mFoxa2 promoter	s	CAGCCTAGACTTCTCTGAGATCCTCC
	as	TCAATAATGGGCCTTGTGGGA
mGata4 promoter	s	GATCTGAGGCTAGCAAGGCAT
	as	TTTGCTGCTCTGCCAAGGA
mGata6 promoter	s	AGCTCTTATTGTCCGCTAGGGC
	as	TGAAGAGAAGGAGGAGGGATCC
mGbx2 promoter	s	CGAACAGCTGCTGGAGCTGGAGAAAG
	as	GCACCTCGCTGAGTTTGAGGGCG
mGdf3 promoter	s	GCTCCCAGAAGTTATCTTATTGCC
	as	TCAGCAGTGTACAGCGTGAGTG
mHoxb1 promoter	s	CTTCCTAGTCATCCTTTTGTCCCA
	as	CTTATCCCAGAACCCCCATTCC
mIkars promoter	s	CCAGTTTCAGGGACTCGGCT
	as	TCGGGGAACACGGGACAC
mMash1 promoter	s	CCAGGCTGGAGCAAGGGA
	as	CGGTTGGCTTCGGGAGC
mMixl1 promoter	s	CACAGATACTCTAGGCAGGAGCAGC
	as	CCTGTCACCATCAAAGGCGC
hMyf5 promoter	s	GGAGATCCGTGCGTTAAGAATCC
	as	CGGTAGCAAGACATTAAAGTTCCGTA
mNanog promoter	s	CCCTCTGAGTTTGACCGGTGA
	as	CAAGCTAGGATGTTAGGTCTCCCTG
mNestin promoter	s	GGAGCACTGGCAGACTTCCCT
	as	GGTGTGGAAACCTCGTGAAGT
mNeuroD promoter	s	TTCAGGCTAGGACATAGACTCAGTGA
	as	CCCTTTGTGGCAGCAGAAGA
mNeurog1 promoter	s	TTGTTGCGCGCCGTACTTAAGG
	as	TGGTCTCCTGAGTGATGTCGCC
mOct4/Pou5f1 promoter	s	GGCTCTCCAGAGGATGGCTGAG
	as	TCGGATGCCCCATCGCA
mPax6 promoter	s	GGTCAGAGGTAATTATGTCACCGCG
	as	TTATTCCAGAGTTGAGCCGTGAGCA
mPdx1 promoter	s	GGACTACATCTTGAGTTGCAGGC
	as	CCCCCTCGCTATGTTTTTTGT
mPsc1/Rbm27 promoter	s	ACATAAACGCACCGCACGTTGT
	as	CCTGGTTACTTCCGCTGGTTTTT
mPrce/Esp1 promoter	s	GACAGCGCGGCAGATAAACTAC
	as	GCACCTCCTTTCAAATATCCCTAG
mRex1/Zfp42 promoter	s	TTTGCGGGAATCCAGCAGT
	as	CGTCCCATCGCCACTCTAGAC
mSox1 promoter	s	ACAAGAGGAGGCAGCGAACC
	as	TCGCAGGTGGAAAGTTTCTCC

mSox17 promoter	s	TGAGGCTCGCTGTAGAAGAGTG
	as	AGCTGTGCCCCGTAACCATTT
mSox2 promoter	s	CCATCCACCCTTATGTATCCAAG
	as	CGAAGGAAGTGGGTAAACAGCAC
mSox3 promoter	s	AGGAGGAAGCGGGAATCCAG
	as	AAACCTCCGCTGCCCCG
mSox7 promoter	s	TCCACAGTGCCAGTTTAGGGAA
	as	CAGTAAACGTGGGCAGCCAC
mSycp1 promoter	s	AAGCCATCTTTAGCACCCAGC
	as	TGGTTGTTAGTTGCCTGGTGG
mTekt1 promoter	s	GATAAGCAGATACACCCAGCTCATC
	as	CCTTCTGGTGTGCATGAAAACA
mTert promoter	s	GTCGCACCACAATAAAGCCTTAAC
	as	GCTGTGGTTGACACAGACAGTGA
mTle1 promoter	s	ATCCCTCCCCCAACAAGAAA
	as	TTGAAGTTTGCAGAAGCCTGTG
mWnt3 promoter	s	CATTTTCCTAGCCCCTGATCC
	as	TTTCTGTGATCCCTTGGCACTT
mZic1 promoter	s	AGGGAAAAGAAATCGTGGCATT
	as	CGCTTGACATCCCCTTTTGA

APPENDIX II – COMPLETE DATA SETS

a



b

%	mES cell lines			mEpiS cell lines			hES cell lines		
	E14tg2A	B6	Dc7	129	B6/CBA	NOD	H1	H7	H9
E	60.0	60.0	62.2	60.0	55.6	54.4	46.7	42.2	37.8
ME	15.6	13.3	13.3	17.8	15.6	21.7	8.9	15.6	15.6
M	2.2	4.4	2.2	4.4	8.9	8.7	11.1	22.2	22.2
ML	8.9	6.7	8.9	6.7	4.4	2.2	26.7	13.3	13.3
L	13.3	15.6	13.3	11.1	15.6	13.0	6.7	6.7	11.1

Figure A1. Comparison of replication timing profiles between mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines. (a) The proportion of genes that replicate early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) in S-phase are represented as a pie-chart. The late replicating set (from yellow to red) is separated from the early set (dark and light green) in mouse and human cells. (b) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines.

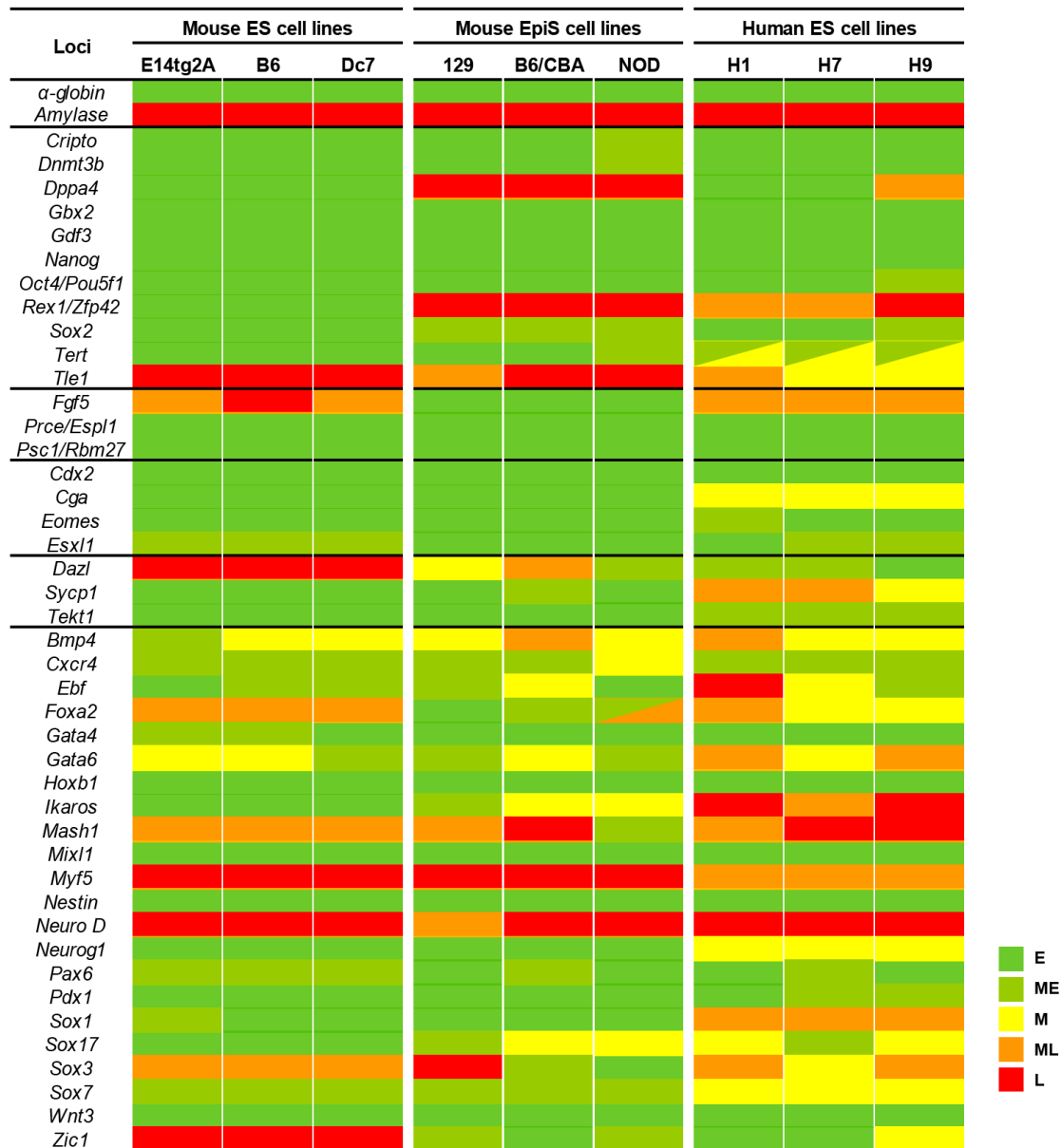


Figure A2. Comparative replication timing analysis of mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines. The replication timing of each gene was defined as early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) and shown as a gradation of colours from green (earliest) to red (latest).

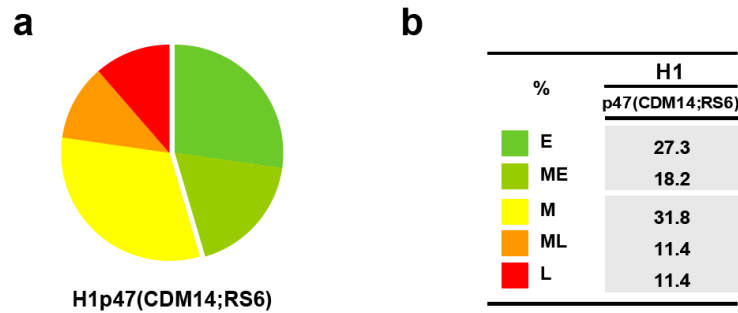


Figure A3. Replication timing analysis of H1 hES cells re-swapped to CM conditions after growing in CDM^{activin}. (a) The proportion of genes that replicate early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) in S-phase is represented as a pie-chart for H1 hES cells grown for 14 passages in CDM^{activin} and re-swapped back to CM conditions for 6 additional passages [H1p47(CDM14;RS6)]. (b) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for H1p47(CDM14;RS6).

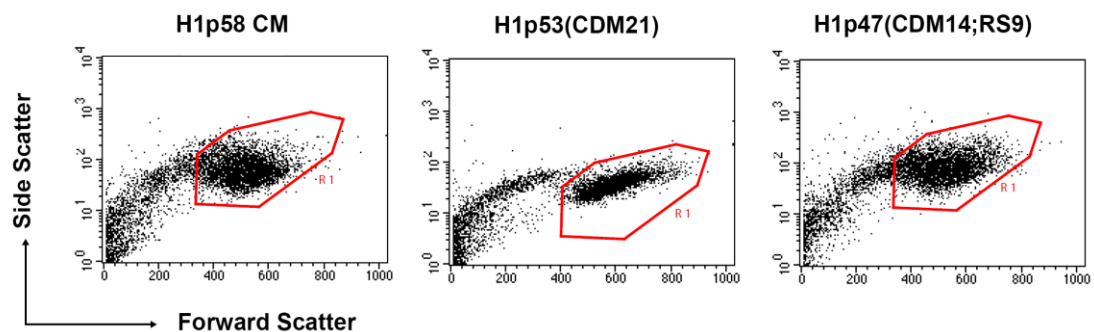


Figure A4. Human ES cell populations grown under CDM^{activin} conditions are more homogeneous than cells grown under standard CM conditions. Flow cytometry correlated measurements of cells size (forward scatter) and internal complexity (side scatter) of hES cells grown under standard CM (H1p58CM), transferred to CDM^{activin} conditions [H1p53(CDM21)] and backwards to CM conditions [H1p47(CDM14;RS9)]. When hES cells are transferred from CM to CDM^{activin} conditions the population of cells become more homogeneous as highlighted by the differential position of region 1 (R1) in red. This can be reverted by transferring cells back to CM growth conditions. These findings are representative of three independent experiments.

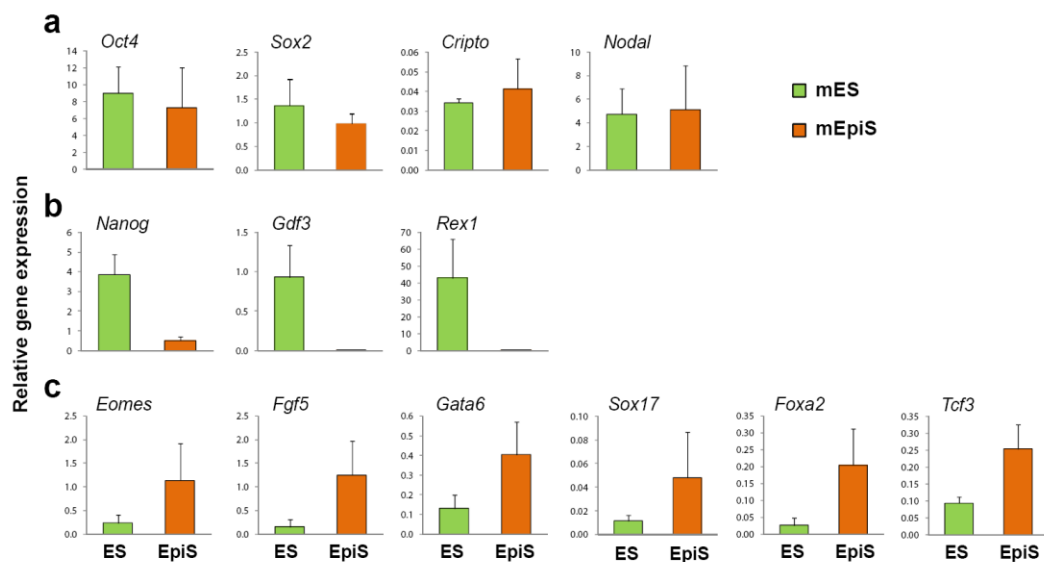


Figure A5. Differential gene expression of mES and mEpiS cell populations. Real-time PCR analysis of transcription profiles of key early developmental genes in mES (average of three mES cell lines – E14tg2A, B6 and Dc7; green bars) and mEpiS cells (average of three mEpiS cell lines – 129, B6/CBA and NOD; orange bars). (a) Stably expressed genes (*Oct4*, *Sox2*, *Cripto* and *Nodal*) across mES and mEpiS cells. (b) Genes that are downregulated in mEpiS cells in contrast to mES cells (*Nanog*, *Gdf3*, *Rex1* and *Stella*) and (c) upregulated in mEpiS cells (*Eomes*, *Fgf5*, *Gata6*, *Sox17*, *Foxa2* and *Tcf3*). Mean and standard deviation of three independent experiments are shown. Data was normalized to the expression of two housekeeping genes (*Actin* and *L19*).

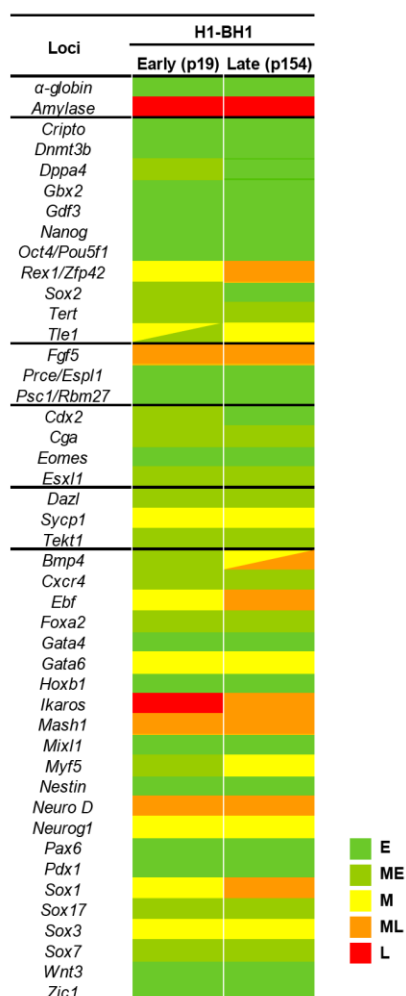


Figure A6. Cell culture adaptation and karyotype abnormalities are not associated with earlier replication timing profiles in hES cells. The replication timing of individual loci was assessed in H1-BH1 hES cells from early (P19) and late (P154) passage number cultures, as provided by Peter Andrews's laboratory. H1-BH1 late passage cells were previously referred as "culture adapted" and are karyotypically abnormal (Baker et al., 2007).