Investigating the Regulation of Distinct Epigenetic States in Human Embryonic Stem Cells

by

Matias Ilmari Autio

A thesis submitted to Imperial College London

for the degree of Doctor of Philosophy

Department of Surgery and Cancer

Imperial College London

November 2012

Statement of Originality

The experiments described in this thesis were conducted by the author, unless acknowledged in the text below.

In Chapter 3 the immunofluorescence, flow cytometry and alkaline phosphatase characterisation of mouse and human embryonic stem cells and DNA replication timing analysis of mouse embryonic stem cells; and in Chapter 4 the DNA replication timing analysis comparing human embryonic stem cells grown in two different laboratories, immunofluorescence, flow cytometry and alkaline phosphatase characterisation of human embryonic stem cells grown in chemically defined medium, embryoid body assay and analysis and the chromatin immunoprecipitation analysis were carried out by Dr. Sandra Pinho.

Abstract

Accumulating evidence suggest that pluripotency – the ability to generate all somatic cell types – is not a fixed state. Pluripotent cell populations encompass a heterogeneous group of cells with different phenotypic and functional properties in equilibrium. At least two phases of pluripotency, immature and primed for differentiation, have been identified during early mouse development, and these are typified by the two distinct stem cell populations – mouse embryonic (mES) and epiblast-derived (mEpiS) stem cells – isolated from epiblast layers of pre- and post-implantation embryo, respectively. Many lines of human ES (hES) cells have also been established *in vitro* from the inner cell mass of pre-implantation blastocysts yet the precise *in vivo* lineage affiliation of these cells remains largely unresolved. 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 during S-phase of the cell cycle across a panel of loci, as an indicator of chromatin accessibility. This replication timing profiling was notably capable of discriminating pluripotent mES cells from cells with a more restricted differentiation capacity.

This study sought to address whether distinct pluripotent states could be reliably discriminated at the chromatin level. In particular, the replication timing profiles of a number of hES cell lines were characterised and compared to those of mES and mEpiS cell lines derived from different genetic backgrounds. Profiles of undifferentiated H1, H7 and H9 hES cell lines typically harboured an increased proportion of late-replicating loci during S-phase when compared to mES cells, which were confirmed to have a steady and mostly early-replicating profile regardless of their genetic background. 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 highlight that hES and mEpiS cells most likely share a common unstable epigenetic state or transitional state primed on the verge of differentiation.

The epigenetic state of hES cell lines was further interrogated by analysing cells grown under two different culture conditions, both however similarly relying on TGF-β/Activin signalling. Results demonstrated that hES cells could adopt distinct yet reversible epigenetic signatures while retaining their pluripotency. In particular, extensive and dynamic shifts of replication timing, from late-to-early, were consistently observed at many target loci in hES cells as well as in human induced pluripotent cells (iPS) cells, upon increased SMAD2/3-associated

P300/CBP histone acetyltransferase (HAT) activity. This was accompanied by fluctuations in the expression of *NANOG* and *REX1* (also known as *ZFP42*), and a change in hES cell's functional properties, as judged by their responsiveness to differentiation-inducing signals. Interestingly, inhibiting P300/CBP HAT activity by curcumin treatment in undifferentiated hES cells was sufficient to revert back to a late-replicating profile associated with a histone hypoacetylated state. Stable knockdown of *P300* in hES cell cultures, however, resulted in a gradual loss of the pluripotent identity through differentiation or apoptosis, preventing further analysis of effects on replication timing. Collectively, these data strongly support the view that different but interchangeable pluripotent states exist within hES cell cultures and suggest a role for P300/CBP HAT activity in determining distinct epigenetic states in hES cells.

As mentioned above, *REX1* was also significantly upregulated in H1 hES cells upon culture condition change. *REX1* is a developmental stage-specific marker that is expressed in the ICM of both mouse and human embryos, as well as in pluripotent mES, hES and iPS cells. However, its function in hES cells remains largely unclear. Human ES cells overexpressing *REX1* were here generated to investigate the role of *REX1* in regulating hES cell pluripotent identity. These cells expressed similar levels of key pluripotency markers than their normal counterparts and remained capable of self-renewing and differentiating into the three germ layers *in vitro*. Interestingly, however, upon withdrawal of exogenous Activin A, *REX1* overexpressing hES cells in contrast to control cells retained a comparatively high level of *OCT4* and stained positive for alkaline phosphatase, a known marker of undifferentiated hES cells. Taken together, these results point to a possible role for *REX1* in sustaining hES cell self-renewal ability.

Acknowledgements

Firstly I would like to thank my supervisor Véronique Azuara for the opportunity to work on my PhD project as well as for the support she has provided and the constructive feedback during the preparation of this manuscript.

I am indebted to several past and present members of the Epigenetics and Development group. Firstly, I am extremely grateful for Dr. Sandra Pinho for introducing me to the world of hES cells and for patiently teaching me the skills needed in working with them, not to mention for all the numerous hours shared in the lab and the cell culture room. I am also very grateful for Dr. Fabrice Lavial for sharing with me his skills on molecular cloning as well as several other lab techniques and for all the times when a sympathetic ear was needed. I would also like to thank Dr. Anne Helness, Miss Rute Tomaz and last, but by no means least, Dr. Michelle Percharde for sharing with me the ups and downs of a PhD.

In addition, I am very grateful for our external collaborators Dr. Ludovic Vallier, Miss Stephanie Brown and Miss Candy C. Cho, from the University of Cambridge, for generously sharing with us their data and cell lines as well as their valuable assistance and advice on transfection of hES cells.

I also offer my thanks to all my friends and colleagues on the 1st floor and at the IRDB in general for making it an enjoyable place to work at.

I want to also thank my dearest Wei Lynn, for always being there with me through these past years; even though in reality we have been separated by oceans and thousands of miles. Looking forwards to the day when the miles between us are there no longer.

Lopuksi haluan kiittää perhettäni, vanhempiani Riitta-Liisaa ja Seppoa, sekä sisaruksiani Minna-Riittaa, Miikkaa ja Mantaa, ja heidän perheitään, sekä kummejani Ritvaa ja Mattia. Ilman teidän tukeanne ja kannustustanne ei perheen kuopus olisi jaksanut ponnistaa näin pitkälle. Rakkaat kiitokset.

Table of Contents

Statement of Originality 2

Abstract 3

Acknowledgements 5

Table of Contents 6

Figures and Tables 11

Abbreviations 13

Chapter 1 Introduction 18

1.1 Levels of epigenetic gene regulation 19

1.1.1 DNA methylation 19

1.1.2 Chromatin structure 23

1.1.3 Histone modifications 25

1.1.3.1 Histone acetylation 26

1.1.3.2 Histone methylation 28

1.1.4 Spatial constraints – nuclear organisation 32

1.1.5 Temporal constraints – DNA replication timing 34

1.2 Early mammalian development and pluripotency 38

1.2.1 Early embryo development 38

1.2.2 Embryo-derived stem cells 40

1.2.2.1 Mouse embryo-derived stem cells 40

1.2.2.2 Mouse and human embryonic stem cells 41

1.2.3 Extrinsic signalling pathways that regulate the pluripotent state *in vitro* 44

1.2.4 Transcriptional networks that control pluripotency 47

1.2.5 Epigenetic regulation of pluripotency 47

1.2.6 The *in vivo* equivalence of mouse and human ES cells – a developmental paradox 48

1.2.7 Reprogramming to pluripotency 50
1.3 Aims of this study 52
Chapter 2 Materials and Methods 54
2.1 Materials 54
2.1.1 Primary antibodies 54
2.1.2 Secondary antibodies 56
2.1.3 Cell lines 57
2.2 Methods 59
2.2.1 Cell culture 59
2.2.2 Cloning and delivery of DNA into hES cells 61
2.2.2.1 Engineering of protein expression constructs 61
2.2.2.2 Preparation and selection of shRNA constructs 62
2.2.2.3 Plasmid construction and bacterial transformation 63
2.2.2.4 DNA delivery into hES cells 65
2.2.3 Small interfering RNA induced knockdown in hES cells 65
2.2.4 Human ES cell differentiation 66
2.2.4.1 Embryoid body formation 66
2.2.4.2 Neural progenitor cell differentiation 66
2.2.5 Replication timing analysis 67
2.2.5.1 BrdU labelling and cell cycle fractionation by flow cytometry 67
2.2.5.2 Isolation of newly-replicated DNA by immunoprecipitation 67
2.2.6 Reverse transcription polymerase chain reaction 68
2.2.6.1 RNA extraction and cDNA synthesis 68
2.2.6.2 Primer design and testing for quantitative PCR 68
2.2.6.3 Real-time quantitative PCR analysis 69
2.2.6.4 Semi-quantitative PCR analysis 70
2.2.7 Immunofluorescence analysis and alkaline phosphatase activity 70

2.2.8 Fluorescence activated cell sorting analysis 71

2.2.9 Protein analysis by western blotting 71

2.2.10 Colony formation assay 72

2.2.11 Chromatin immunoprecipitation analysis 72

2.2.11.1 ChIP for histone modifications 72

2.2.11.2 ChIP for P300 73

2.2.12 Analysis of DNA content and BrdU incorporation via acid denaturation of DNA 74

Chapter 3 Characterisation of replication timing profiles in hES cell lines alongside mES and mEpiS cell lines 76

3.1 Introduction 76

3.2 Epigenetic profiling of hES, mES and mEpiS using a PCR-based DNA replication timing assay 76

3.3 Mouse ES cells have conserved replication timing profiles between cell lines from different genetic backgrounds 81

3.4 Replication timing profiles are distinct between pluripotent mouse and human ES cells 82

3.5 Temporal comparison of hES and mEpiS cell's replication profiles suggests a common unstable or transitional state on the verge of differentiation 87

3.6 Conclusions 90

Chapter 4 Distinct epigenetic states within pluripotent human ES cells and characterisation of their regulation 91

4.1 Introduction 91

4.2 Replication timing profiles of undifferentiated pluripotent human cell populations undergo extensive alterations upon different cell culture conditions 92

4.2.1 Distinct replication timing profiles of H9 hES cells grown in two different laboratories 92

4.2.2 Dynamic modulations of replication timing profiles in H1 hES cells in response to cell culture conditions 93

4.2.3 Fully reprogrammed human iPS cells share the same level of epigenetic plasticity as hES cell 96

4.3 Phenotypic and functional characterisation of H1 hES cells grown under CM and CDM^{activin} conditions 98

4.3.1 H1 hES cells cultured in CDM^{activin} conditions retain their undifferentiated state 98

4.3.2 H1 CM and CDM^{activin} form embryoid bodies containing tri-germ layer differentiation 100

4.3.3 H1 CDM^{activin} conditions favour an ICM-like phenotype within hES cell cultures 102

4.4 Investigating the molecular mechanisms underlying distinct epigenetic states in hES cells 103

4.4.1 Increased TGF-β/Activin signalling in CDM^{activin} conditions 103

4.4.2 Changes in TGF-β/Activin signalling impact on the expression levels of P300 histone acetyltransferase in hES cells 104

4.4.3 Global increase of P300-mediated histone H3 acetylation in H1 CDM^{activin} cell cultures 108

4.4.4 Recruitment of active P300 complexes at target loci is enhanced in response to CDM^{activin} growth conditions 108

4.4.5 Stable knockdown of P300 in H1 hES cells leads to differentiation and gradual loss of cultures 110

4.4.6 Transient knockdown of P300 by RNA interference does not impact on the replication timing profile of H1 CDM^{activin} hES cells 112

4.4.7 Inhibition of P300/CBP histone acetyltransferase activity leads to early to late replication timing changes in H1 CDM^{activin} hES cells 114

4.5 Conclusions 116

Chapter 5 Characterisation of *REX1/ZFP42* overexpressing human ES cells 118

5.1 Introduction 118

5.2 Establishment of *REX1* overexpressing human ES cell clones 120

5.3 Examination of the growth characteristics of *REX1* overexpressing hES cell clones 122

5.4 H9 Rex1OE cells form embryoid bodies containing cells derived from all three germ-layers 124

5.5 H9 Rex1OE cells are able to maintain hall-marks of pluripotent cells in culture without Activin A supplementation 126

5.6 Conclusions 130

Chapter 6 Discussion 132

6.1 Human and mouse ES cells have distinct replication timing profiles 133

6.2 Human ES and mouse EpiS cells share a common variable epigenetic state as defined at the replication timing level 134

6.3 Distinct and reversible epigenetic states within hES cell cultures 136

6.4 Investigating the molecular pathways underlying distinct pluripotent states in hES cells 137

6.5 Effects of *REX1* overexpression in hES cells 142

References 144

Appendix I Primer sequences used 178

Appendix II Additional data 185

Figures and Tables

Figure 1.1 Epigenetic levels of gene regulation.....20
Figure 1.2 Post-translational modifications of human histones.....26
Figure 1.3 Replication timing foci patters during early, middle and late S-phase of cell cycle. 35
Figure 1.4 Schematic representation of early stages of mouse and human embryo development. 39
Figure 2.1 Transgenes and expression vector used in this study 61
Figure 2.2 Vector used for stable expression of shRNA 62
Figure 2.3 Selection of effective P300 shRNAs.....63
Figure 3.1 Phenotypic similarities and differences between undifferentiated mES, hES and mEpiS cells. 77
Figure 3.2 Schematic representation of DNA replication timing analysis. 80
Figure 3.3 DNA replication timing analysis controls.....81
Figure 3.4 Independently derived mES cell lines from distinct genetic backgrounds have similar DNA replication timing profiles.....83
Figure 3.5 Pluripotent mouse and human ES cells have distinct DNA replication timing profiles. 84
Figure 3.6 Human ES cell cultures analysed by replication timing display similar expression of pluripotency-associated markers.....87
Figure 3.7 DNA replication timing variability between established hES cell lines.....87
Figure 3.8 Mouse EpiS and hES cells have variable DNA replication timing profiles across many genes and similar cell cycle kinetics 89
Figure 4.1 Human ES cells grown in different cell culture conditions and different laboratories exhibit different replication timing profiles.....92
Figure 4.2 H1 hES cells undergo dramatic late to early shifts of replication timing profiles when transferred from CM to CDM^{activin} conditions. 93
Figure 4.3 Individual loci undergo shifts of replication timing upon growth culture condition change.94
Figure 4.4 Loci shifting their replication timing in CDM^{activin} are shared between different human ES cell lines. 95
Figure 4.5 Fully reprogrammed human iPS cells display similar shifts of replication timing at candidate loci as observed in hES cells upon change of culture conditions. 97
Figure 4.6 Human ES cells transferred from CM to CDM^{activin} conditions retain hallmarks of undifferentiation.....99
Figure 4.7 Human ES cells from both CM and CDM^{activin} conditions are able to form derivatives of all three germ lineages in an in vitro assay of differentiation. 101
Figure 4.8 H1 hES cells cultured in CM and CDM^{activin} have distinct expression patterns of key marker genes.....102
Figure 4.9 H1 cells cultured in CDM^{activin} conditions have increased levels of phosphorylated SMAD2/3 complexes when compared to cells cultured in CM. 104

Figure 4.10 Histone acetyltransferase P300 expression is significantly increased in H1 hES cells cultured in CDM^{activin} conditions as compared to CM conditions. 105
Figure 4.11 Levels of P300 in hES cells respond to SMAD2 activity. 106
Figure 4.12 12 P300 protein levels are maintained and global acetylation levels are increased at P300/CBP preferred target residues with increasing passages in CDM^{activin} conditions. 107
Figure 4.13 Loci changing replication timing from late in CM to early in CDM^{activin} hES cells harbour higher levels of P300 and histones acetylation at their promoter regions in CDM^{activin} hES cells. 109
Figure 4.14 Stable knockdown of P300 leads to differentiation or loss of cultures in H9 hES cells. 111
Figure 4.15 Transient siRNA induced knockdown of P300 does not affect the replication timing profiles of H1 CDM^{activin} cultured cells.....113
Figure 4.16 Inhibition of P300/CBP histone acetyltransferase activity by Curcumin leads to a reduction of histone acetylation levels and induces early to late replication timing changes in H1 CDM^{activin} cells.....115
Figure 5.1 Establishment of REX1 overexpressing (Rex1OE) hES cell clones. 119
Figure 5.2 REX1 overexpressing cells retain expression of pluripotency markers in culture. 121
Figure 5.3 Rex1OE hES cells retain normal morphology and stain positive for alkaline phosphatase activity.....122
Figure 5.5 H9 control and Rex1OE cells have similar cell cycle profiles. 123
Figure 5.4 REX1 overexpression does not inhibit hES cell self-renewal ability 123
Figure 5.6 H9 Rex1OE hES cells are able to differentiate in vitro into derivatives of the three germ layers.....126
Figure 5.7 H9 Rex1OE hES cells can be serially passaged in absence of Activin A supplementation. ..127
Figure 5.8 Rex1OE hES cells retain intermediate OCT4 expression levels in prolonged culture without Activin A. 130
Figure 6.1 Summary for observed interplay between TGF-β signalling and P300 HAT activity in the regulation of distinct epigenetic states in hES cells. 141
Figure A1 H7 undifferentiated hES cells can be clearly distinguished by replication timing analysis from neural progenitors (NP) H7-derived.....185
Figure A2 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. 187
Figure A3 Comparative replication timing analysis of mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines.....188
Figure A4 Human ES cells replication timing variability is not associated with gene expression or histone acetylation specific patterns. 189
Figure A5 Human ES cells replication timing variability is not associated with DNA methylation specific patters.....190
Figure A6 Transient increase in the expression levels of Nanog in H1 CDM^{activin} hES cells 191

Figure A7 Differential gene expression profiles of TGF-β members in H1 CM and CDM^{activin} hES cell populations. 192
Figure A8 Replication timing variability is not observed across human CD34+ haematopoietic stem cells. 192
Figure A9 Cell culture adaptation and karyotype abnormalities are not associated with earlier replication timing profiles in hES cells. 193
Figure A10 Human ES cells cultured in CDM^{activin} differentiate more slowly by EBs differentiation than hES cells cultured in CM 194
Figure A11 Inhibition of P300/CBP histone acetyltransferase activity by small molecule inhibitor C646 does not affect the replication timing profiles of H1 CDM^{activin} cultured cells. 195

Table 1 - Candidate genes selected for the replication timing analysis. 78
Table A1. Primers for the amplification of human or mouse transcripts by real time quantitative PCR 178
Table A2. Primers used for the amplification of human transcripts by semi-quantitative RT-PCR 180
Table A3. Primers for the selective amplification of human or mouse genomic regions by real time quantitative PCR 180
Table A4 - Genomic features of replication timing analysed genes 186

Abbreviations

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
$CDM^{activin}$	chemically defined medium supplemented by Activin A
ChIP	chromatin immunoprecipitation
CM	conditioned medium
C(T)	threshold cycle
CpG	cytosine-guanine dinucleotide
DAPI	4,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
Dnmt	DNA (cytosine-5) methyltransferase
EB	embryoid body
EDTA	ethylene-dinitrilo tetraacetic acid
EC	embryonic carcinoma
EG	embryonic germ
ES	embryonic stem
ExEc	extraembryonic ectoderm
ExEn	extraembryonic endoderm
F-Rex1OE	FLAG- *REX1/ZFP42* overexpressing
FACS	fluorescence-activated cell sorting
FBS	foetal bovine serum
Fig.	Figure
FITC	fluorescein isothiocyanate
FGF	fibroblast growth factor
g	gram
GC	guanine-cytosine

GFP	green fluorescent protein
H	histone
HAT	histone acetyltransferase
HDAC	histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hES	Human embryonic stem
HRP	horseradish peroxidase
HSC	haematopoietic stem cell
ICM	inner cell mass
Id	inhibitor of differentiation
IF	immunofluorescence
IGF	insulin-like growth factor
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
LB	Lysogeny Broth
m	milli
M	molar
Mb	mega base pair
me	methylated

MeDIP	methylated DNA immunoprecipitation
MEF	mouse embryonic fibroblast
mEpiS	mouse epiblast stem
mES	mouse embryonic stem
min	minute
mRNA	messenger RNA
NEAA	non-essential amino acids
NP	neural progenitor
NP-40	nonidet P-40
P	phosphorylated
PBS	phosphate buffered saline
PcG	polycomb group
PCR	polymerase chain reaction
PEV	position effect variegation
PFA	paraformaldehyde
PGC	primordial germ cell
PHD	plant homeodomain
PI	propidium iodide
PICMI	post-ICM intermediate
PK	proteinase K
Pol	polymerase
PRC	polycomb repressive complex
PrE	primitive endoderm
qRT-PCR	real-time quantitative PCR
Rex1OE	*REX1/ZFP42* overexpressing
RNA	ribonucleic acid
RNase	ribonuclease

rpm	rotations per minute
RT	room temperature
RT-PCR	reverse transcription PCR
SCNT	somatic cell nuclear transfer
scr	non-targeting scramble shRNA
SDS	sodium dodecyl sulphate
SET	Su(var)-Enhancer of zeste-Trithorax
shRNA	short hairpin loop RNA
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
TF	transcription factor
Tra	tumour rejection antigen
Tris	(hydroxymethyl)methylamine
TrxG	trithorax group
TS	trophoblast stem
TSS	transcriptional start site
U	units
VCT	villous cytotrophoblast
WB	western blot
XEN	extra-embryonic endoderm stem

Chapter 1 Introduction

During mammalian development a single zygote, formed by the fusion of two totally differentiated cells, the sperm and the egg, gives rise to more than 200 different somatic cell types of the mammalian body. The developmental process is unidirectional where cells progressively lose developmental potential and acquire specialised functions pertinent to their lineage (Hochedlinger et al. 2009). Even though development has been studied intensely for decades, if not even centuries, deciphering the details of how genetically identical cells diverge into cells with different identities during development remains one of the most interesting and important questions in biology (Fisher et al. 2012).

Experiments by John Gurdon in 1962 demonstrated that a nucleus of differentiated gut epithelial cell from *Xenopus laevis* tadpole could give rise to a fully developed adult frog when inserted into an enucleated oocyte (Gurdon 1962). These results undisputedly proved that somatic cells retain all the genetic information necessary to direct development. In accordance, cloning of Dolly the sheep (first mammal ever cloned) (Wilmut et al. 1997) and subsequently of laboratory mice from terminally differentiated lymphocytes (Hochedlinger et al. 2002), demonstrated the principle of nuclear equivalence; that specialised cells of metazoans possess a gene pool identical to that of the zygote nucleus.

The landmark work describing the completed human genome sequence (Lander et al. 2001; McPherson et al. 2001) has led to current sequencing technologies, which allow the sequencing of whole-genomes from relatively small samples of cells. However, having the genome sequence alone does not allow deciphering what makes, for example, a skin cell different from a nerve cell or a blood cell. It is now known that superimposed upon the deoxyribose nucleic acid (DNA) sequence exist layers of heritable information, which allow for the selective expression of only a fraction of the genes encoded in the genome in any given cell type. This carefully regulated control via epigenetic mechanisms imposes heritable cellular memories that define both specific cell lineages and cell types during development (Hemberger et al. 2009). Different levels of epigenetic regulation include changes in higher order chromatin structure, methylation of the DNA cytosine residues, histone modifications, DNA replication timing, regulation by small interfering RNAs and the spatial positioning of the genomic region within the nucleus, moreover, all these different levels of epigenetic regulation are subject to cross-talk. Hence, normal development requires careful control of both genetic and epigenetic programs in all cells. This is to induce appropriate transcriptional and epigenetic outcomes and to repress inappropriate developmental

programs in response to signalling cues within the developing embryo, resulting in a fully developed organism.

1.1 Levels of epigenetic gene regulation

The term 'epigenetic' was originally coined by C. Waddington in 1942. His original definition for the term was: the causal mechanisms by which the genes of a genotype bring about a phenotype (Waddington 1942). The modern usage of the term refers to heritable traits (over rounds of cell division and in some cases over generations) that occur without any changes to the underlying DNA sequence (Probst et al. 2009). Epigenetic information provides a form of memory that allows the maintenance of specific gene expression patterns in any given cell type. During development differentiating cells acquire epigenetic landscapes distinct from those of pluripotent cells, as well as from those established in cells differentiating into alternative lineage paths (Reik 2007). Disruption of the epigenetic regulation during development, or normal cell division can lead to disease (Probst et al. 2009). The mechanisms that regulate this epigenetic information are now being revealed, and these include DNA methylation, modification of histones and chromatin structure, regulation by non-coding ribonucleic acids (RNAs) and higher order genome organisation [Figure (Fig.) 1.1]. It is thought that DNA methyltransferases, histone-modifying enzymes, chromatin-binding proteins, non-coding RNAs and many other regulators act together with sequence-specific transcriptional factors to establish and convey gene expression patterns during development, ensuring the stability of the 'committed' state.

1.1.1 DNA methylation

DNA methylation is found in the genomes of organisms ranging from prokaryotes to eukaryotes. In prokaryotes methylation occurs at both cytosine and adenine bases, whereas in eukaryotes methylation seems to occur only at cytosine bases. Specifically the methylation takes place at the positive five of the cytosine ring of cytosine-guanine dinucleotides (CpG) [5-methylcytosine (5mC)] (Klose et al. 2006). The enzymes responsible for catalysing the methylation reaction in mammals are three essential DNA (cytosine-5) methyltransferases (Dnmt1, Dnmt3a and Dnmt3b) (Goll et al. 2005). Mice lacking Dnmt1 or Dnmt3b die by embryonic day 10.5 (Li et al. 1992; Okano et al. 1999). Those lacking Dnmt3a sometimes survive, but suffer serious malformation especially at germ cells and die within weeks of their birth (Okano et al. 1999). Two other related molecules also exist. Dnmt2, is highly conserved across species and highly expressed in many cell types, however it only shows very weak methyltransferase activity *in vitro* and genetic studies show no phenotype

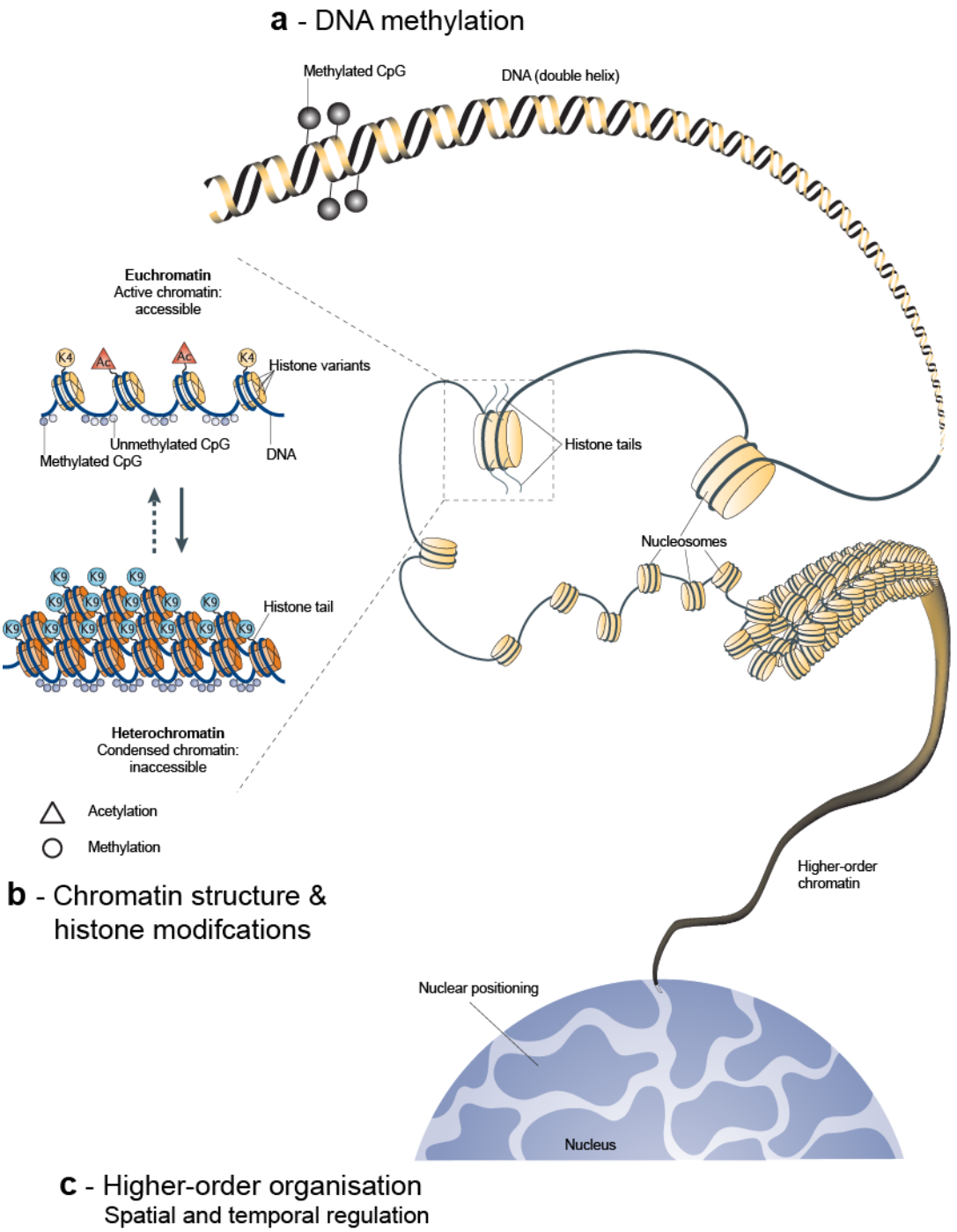


Figure 1.1 Epigenetic levels of gene regulation.

Schematic representation of different levels of genome organisation that have an impact on gene activity. (**a**) The DNA double helix carries the genetic information necessary for the development and function of an organism. Beyond the sequence, the most direct mechanism of epigenetic regulation is DNA methylation at the Cytosine (C) residue dinucleotide CpG (see section 1.1.1). (**b**) DNA is wrapped around nucleosomes, made up from different histones, in a 'beads on a string' structure which folds to form higher order chromatin domains (see section 1.1.2). Specific amino acid residues within the histone tails are targets for a number of covalent post-translational modifications, including for example acetylation (triangle) and methylation (circle) as well as phosphorylation and ubiquitination (see section 1.1.3). (**c**) Higher order chromatin is compartmentalised within the nucleus into specific chromatin domains and organised into functional nuclear compartments. These three-dimensional domains as well as temporal constraints (DNA replication timing) add to the levels of regulation of the genome (see sections 1.1.4 and 1.1.5). Taken together all the above mechanisms provide epigenetic cues to regulate the expression of genes in any given cell type and elucidate how a single genome can give rise to such a wide variety of different cell types Figure adapted from (Sparmann et al. 2006; Probst et al. 2009; Gaspar-Maia et al. 2011).

upon disruption (Okano et al. 1998; Hermann et al. 2003; Goll et al. 2005; Klose et al. 2006) and hence its role remains unclear. Dnmt3L is a Dnmt-related protein that does not have intrinsic DNA methyltransferase activity, however it physically associates with Dnmt3a and Dnmt3b to modulate their activity (Suetake et al. 2004).

CpG methylation, especially at the promoter regions of genes, is associated with stable transcriptional repression. In addition to this DNA methylation has important roles in X-chromosome inactivation, genomic imprinting and development as well as development of cancer and other diseases (Bird 2002; Reik 2007; Cedar et al. 2012). During mitotic cell division methylation patterns are maintained through methyltransferase activity by Dnmt1 (Li et al. 1992), which happens in concert with DNA replication as Dnmt1 is perpetually located at replication foci (Leonhardt et al. 1992). Dnmt1 is known to have a high substrate preference for hemi-methylated DNA (Gruenbaum et al. 1982; Chuang et al. 1997). As the DNA is replicated, synthesis of the new DNA generates hemi-methylated strands, which have symmetrically disposed methyl groups on CpG dinucleotides. These strands are identified and methylated by Dnmt1, thus allowing the maintenance of the methylation state (Gruenbaum et al. 1982; Cedar et al. 2012). Dnmt3a and Dnmt3b have been shown to be mainly responsible for *de novo* methylation, and are essential for establishing methylation patterns during development (Okano et al. 1999).

During mouse development the overall methylation patterns change dramatically (Santos et al. 2002; Reik 2007; Cedar et al. 2012). In early embryo development parental methylation patterns are lost, first actively from the paternal genome (Mayer et al. 2000) and then likely passively due to exclusion of Dnmt1 to the cytoplasm (Carlson et al. 1992; Howell et al. 2001). Imprinted genes, such as *Igf2r*, are protected from the global demethylation in early embryo development and retain their methylation status in the adult organism, apart from the germ cells (Birger et al. 1999; Cedar et al. 2012). At about the time of implantation the whole genome undergoes dramatic *de novo* methylation likely mediated by Dnmt3a and -3b, this however does not occur at CpG-islands in the promoters (Okano et al. 1999; Cedar et al. 2012). Post-implantation additional changes in methylation occur tissue and gene specifically at the promoter CpG-islands (Cedar et al. 2012). A well-characterised example of a locus that is regulated by DNA methylation is the embryonic stem (ES) cell-specific factor *Oct4*, which is also expressed in the early embryo, and undergoes stable repression during cell differentiation (Feldman et al. 2006; Yeo et al. 2007).

There are two general mechanisms by which DNA methylation acts to induce gene repression. Firstly, modification of methyl-CpGs can directly inhibit the association of some

DNA binding factors with their cognate DNA recognition sequences (Watt et al. 1988); and secondly, methyl-CpGs can be bound by proteins of a family that share the methyl CpG-binding domain, which can, in turn, act to repress transcription themselves and recruit other repressive proteins (Boyes et al. 1991; Hendrich et al. 1998; Klose et al. 2006). An example of the former is how the CCCTC-binding factor (CTCF) is unable to bind the imprinting control region of the *Igf2/H19* when the CpGs are methylated (Holmgren et al. 2001). Recent genome-wide studies have found high correlation with unmethylated DNA and accessibility of the genome as assessed by DNaseI footprinting (Neph et al. 2012). Repression by methyl binding proteins has also been shown to involve histone deacetylases (HDACs) and histone methyl transferases (Nan et al. 1998; Klose et al. 2006). Changes in histone acetylation and/or methylation at a specific chromatin region with DNA methylations may be brought about via the recruitment of histone modifying proteins by methyl binding proteins. This link between DNA methylation and chromatin conformation could provide a route for maintenance of chromatin regions during mitosis (Hashimshony et al. 2003; Cedar et al. 2012). Because the underlying DNA methylation pattern is autonomously copied following replication, it is possible that this modification pattern may play a role through the recruitment of histone modifying enzymes by methyl binding proteins, in the maintenance of histone modification patterns.

It is thought that the global demethylation seen in early mouse embryo development is likely occurring passively through the exclusion of Dnmt1 from the nucleus (Carlson et al. 1992). Interestingly a recent study reported 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). Despite intensive research, the mechanism for active erasure of the methylation mark remains unclear. Current evidence suggests that methyl groups attach through a carbon–carbon bond to the cytosine base and therefore might not be able to be directly removed (Morgan et al. 2005; Cedar et al. 2012). One possible path for active demethylation is via mismatch repair. The DNA deaminases Aid and Apobec1 have been shown *in vitro* to deaminate 5-methylcytosine in DNA to thymine, which results in T-G mismatches that can be repaired by the base-excision repair pathway (Morgan et al. 2004). Aid and Apobec1 are highly expressed by oocytes, stem cells and germ cells (Morgan et al. 2004) and interestingly recent studies with reprogramming of somatic cells to iPS cells have shown that AID is required for active demethylation of *OCT4* and *NANOG* loci during the process of reprogramming

(Bhutani et al. 2010). The recent discoveries that mammalian cytosines can be hydroxymethylated to 5-hydroxymethylcytosine (5hmC) by the ten-eleven translocation (TET) family (Kriaucionis et al. 2009; Tahiliani et al. 2009), demonstrate a new pathway of active DNA demethylation and have generated substantial excitement in the field (Bhutani et al. 2011). The 5mC to 5hmC modification is not a functional equivalent of 5mC to cytosine modification, which is linked to gene derepression. In contrast, 5hmC was shown in mES cells to be enriched at so-called bivalent loci, which carry both activating and repressive histone modifications and are thought to be 'poised' for future activation (Pastor et al. 2011). Furthermore, the activity of TET family proteins and the 5hmC modification have been linked to the expression of key regulators of pluripotent mES cells as well as transcriptionally active and repressed loci in ES cells (Ficz et al. 2011; Wu et al. 2011), highlighting complex and dynamic nature of DNA methylation.

1.1.2 Chromatin structure

Genomic DNA in the eukaryotic nucleus is complexed with core histones and other chromosomal proteins to form the chromatin (Fig. 1.1). The compaction of the genome is necessary for the DNA to fit in a cell's nucleus, because for example the human genome would stretch to about two meters in length in a single double helix thread. The fundamental repeating unit in the chromosome is the nucleosome, which consists of 146 base pairs (bp) of DNA wrapped in 1.7 super helical turns around a histone octamer, consisting of an H3-H4 tetramer and two H2A-H2B dimers, forming a 'beads-on-a-string' fibre with a diameter of 11 nm (Kornberg 1977; Richmond et al. 1984; Luger et al. 1997). While the length of DNA engaged in nucleosomes is always the same, the length of the spacer between nucleosomes varies from several to about 80 bp, making the whole length of DNA associated with each nucleosome complex throughout all eukaryotic genomes 200 ± 40 bp (Kornberg 1977; McGhee et al. 1980). The binding of the linker histone (H1 or H5) (Bates et al. 1981; An et al. 1998) allows the formation of the condensed 30 nm chromatin fibre (Robinson et al. 2006); this can compact DNA up to 30-50 fold (Luger et al. 1997; Felsenfeld et al. 2003; Bell et al. 2011) and can be considered as the second structural level of DNA organization (Li et al. 2011). Chromatin folding to higher level conformations remains much less understood (Felsenfeld et al. 2003; Li et al. 2011). Histone H1 has been shown to be required for metaphase chromosome compaction (Maresca et al. 2006), however other necessary players remain to be defined.

The positioning of the nucleosome has been shown to be non-random and has a functional significance in several cellular processes including transcriptional regulation, chromosome segregation, DNA replication and DNA repair (Henikoff 2008; Schones et al. 2008). Many chromatin remodelling ATPase complexes, such as the SWI/SNF, ISWI, CDH and INO80 complexes, which can move, destabilise, eject, or restructure nucleosomes have been shown to be involved in these processes (Clapier et al. 2009). The properties of histones that make up the nucleosome also have an effect on nucleosome dynamics. In addition to the four canonical histones numerous histone variants have been identified, such as H2A.Z and H3.3. H2A.Z has been shown to stabilize the histone octamer (Park et al. 2004), interestingly H2A.Z has also been shown to be enriched in a subset of genes and necessary for lineage commitment in mES cells (Creyghton et al. 2008). H3.3 is constitutively expressed and deposited into chromatin throughout the cell cycle, unlike the canonical H3, which is integrated to chromatin during DNA replication (Ahmad et al. 2002a; Ahmad et al. 2002b). Interestingly nucleosomes containing H3.3 have been shown to be highly unstable regardless of whether they are partnered with H2A or H2A.Z, this could be linked with maintenance of accessible chromatin structures as H3.3 is present at promoters and enhancers of actively transcribed genes (Jin et al. 2007).

In eukaryotes the genome is packaged into the chromatin template. This packaging inherently restricts accessibility of the DNA for regulatory proteins (Bell et al. 2011). Early studies using staining and light microscopy in 1928 by Emil Heitz defined two types of interphase chromatin; less condensed euchromatin and the more condensed heterochromatin (Dillon et al. 2002). Euchromatin is lightly packed and therefore accessible for transcription factors (TFs) and RNA polymerases, and is enriched in active genes and loci replicating early in the S-phase (Gilbert et al. 2004; Holmquist et al. 2006; Shahbazian et al. 2007). Genome-wide analysis of DNA binding by 119 different TFs in 72 different cell lines has indicated that most TFs are bound to regions depleted of nucleosomes (Wang et al. 2012). Heterochromatin, on the other hand, is thought to favour gene silencing (Holmquist et al. 2006). This has been demonstrated by position effect variegation (PEV), where an otherwise active gene locus can become sporadically silenced due to positioning close to heterochromatin. This phenomenon was first identified in *Drosophila*, where positioning of an eye colour gene close to centromeric heterochromatin can give rise to colourless phenotype (Muller et al. 1930), and has been reported in other organisms from yeast (Gottschling et al. 1990; Aparicio et al. 1991) to mammals (Azuara et al. 2003; Hiragami-Hamada et al. 2009). At least two types of heterochromatin have been distinguished. Firstly,

facultative heterochromatin, which is developmentally regulated and can be formed at loci that in other cell-types are found in euchromatic form (Arney et al. 2004). One obvious example of facultative heterochromatin formation is the inactivation of one of the two X-chromosomes in early embryonic development (Heard 2005). The other type is constitutive heterochromatin that is generally silenced in all cell types and includes major and minor satellite DNA repeats clustered around mouse centromeres, other repetitive DNA elements and transposons (Arney et al. 2004).

1.1.3 Histone modifications

Highly conserved histone proteins make up the octameric nucleosome to which the DNA wraps around (Fig. 1.1). It has been known ever since the 1960's that histones could be post-translationally modified (Allfrey et al. 1964), however this was considered to be artifactual by many and only recently the functional significance of these modifications has started to be elucidated (Kouzarides 2007). The high-resolution X-ray structure of the nucleosome indicated that the N-terminal (and C-terminal in the case of H2A and H2B) tails of histones protrude out from the nucleosome and are thus able to interact with other nucleosomes and affect the chromatin structure, as well as recruit regulatory proteins (Luger et al. 1997; Margueron et al. 2005; Bannister et al. 2011). These N- and C-terminal histone tails, as well as some residues in the globular domains, can carry many different types of post-translational modifications such as acetylation, methylation, phosphorylation, ubiquitination, sumoylation, biotinylation, citrullination and ADP-ribosylation (Fig. 1.2) all of which can, in turn, modulate the interaction between nucleosomes and with regulatory proteins and thus have functions in gene regulation (Bhaumik et al. 2007; Kouzarides 2007; Latham et al. 2007; Bannister et al. 2011). The most extensively studied of these modifications are acetylation and methylation, which were also the modifications that were discovered first (Allfrey et al. 1964).

The different types of covalent histone modifications act in concert to influence a multitude of cellular processes including transcription, replication, DNA repair and cell cycle procession (Bhaumik et al. 2007). Many different modifications can occur in the same histone simultaneously and different types of modifications occur in specific patterns related to gene functions. Acetylation of specific lysine residues both on histone H3 and H4 N-terminal tails, such as, H3K9ac and H4K16ac, is highly correlated with accessible chromatin and actively transcribed genes, whereas lysine methylation can correlate with either active (H3K4me2/3) or repressed (H3K9me3 or H3K27me3) regions of the chromatin (Bernstein et

al. 2007; Kouzarides 2007). The different types of modifications can also regulate one another both in *cis*- on the same histone and in *trans*- between different histones or nucleosomes (Latham et al. 2007; Lee et al. 2010). It has been suggested that together the different histone modifications on the genome form their own heritable 'code', which is read by effector molecules that drive changes in chromatin conformation and gene transcription leading to distinct downstream effects (Jenuwein et al. 2001; Margueron et al. 2005; Rando 2012).

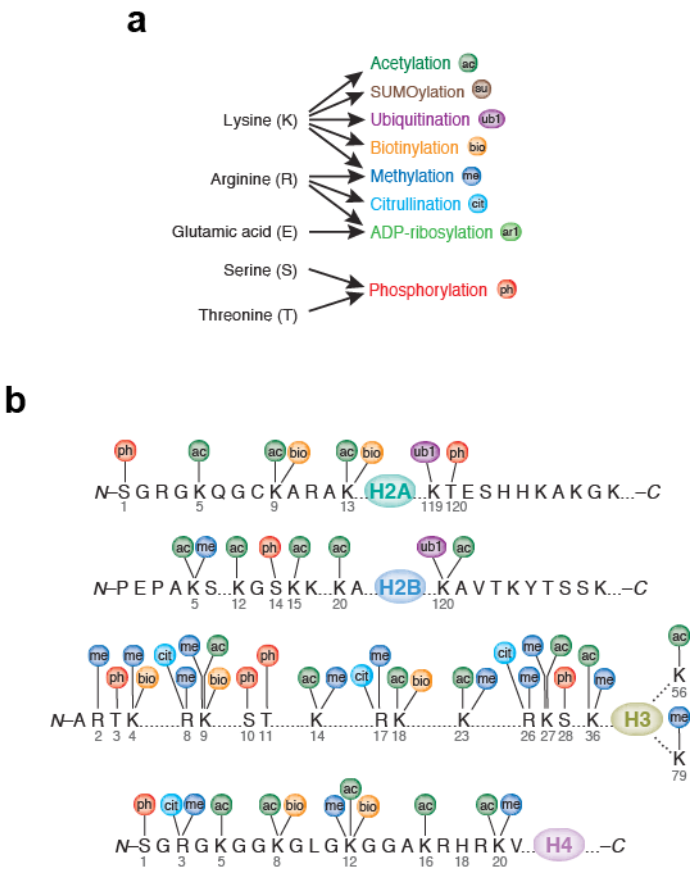


Figure 1.2 Post-translational modifications of human histones.
(**a**) Known post-translational histone modifications and the amino acid residues they modify. (**b**) Different covalent post-translational modifications of the four core histones (H2A, H2B, H3 and H4) have been identified and are here represented above the corresponding amino acid residue in the N-terminal tails of histones. 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. 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. Figure adapted from (Bhaumik et al. 2007; Latham et al. 2007).

1.1.3.1 Histone acetylation

Levels of histone acetylation have been shown to correlate with transcriptionally active and inactive regions of the genome. Hypoacetylation of all four nucleosomal histones has been consistently associated with transcriptionally inactive heterochromatin (Jeppesen et al. 1993; Maison et al. 2002). Conversely hyperacetylation of histones correlates highly with chromatin accessibility and transcriptional activity (Pogo et al. 1966; Hebbes et al. 1988; Shogren-Knaak et al. 2006). These correlations have been confirmed more recently by genome-wide studies in numerous different cell types (Koch et al. 2007; Wang et al. 2008; Wang et al. 2009b; Djebali et al. 2012; Dunham et al. 2012; Sanyal et al. 2012).

Acetylation of histones can occur on lysine residues on all the four-histone types (Fig. 1.2), the most extensively characterised are modifications on H3 and H4 N-terminal tails.

Acetylation of a lysine residue neutralises the positive charge of the target residue; this has the potential to weaken the interactions between the histones and DNA (Bannister et al. 2011) as well as between different nucleosomes (Shogren-Knaak et al. 2006). High levels of acetylation have been shown to inhibit higher order chromatin folding and dramatically enhance transcription in *in vitro* assays (Tse et al. 1998). More detailed studies have shown that H4K16 acetylation on the N-terminal tail directly affects the folding of the nucleosomes by disrupting the electrostatic interaction between the N-terminal tail of H4 and an acidic patch of the H2A/H2B dimer on an adjacent nucleosome, thus preventing the formation of compact 30 nm fibres (Shogren-Knaak et al. 2006; Robinson et al. 2008; Bell et al. 2011). Apart from the histone tails, lysine acetylation has also been found in the globular histone core. H3K56 acetylation has been found not to directly affect chromatin compaction, however it has been found to increase the transient partial unwrapping of DNA from the histone core, also known as DNA breathing (Neumann et al. 2009). Moreover, H3K56ac does not directly correlate with transcription per se, but instead has roles related to DNA repair and genome stability (Masumoto et al. 2005; Bhaumik et al. 2007). Interestingly, H3K56ac has also been found to be enriched at active regulators of human ES identity as well as inactive, lineage-specific genes, poised for activation upon differentiation (Xie et al. 2009). In addition to these effects, histone acetylation can also enable binding of factors associated with gene activation or chromatin modelling. The chromatin remodelling complex SWI/SNF, which mobilises nucleosomes and has an important role in transcriptional activation, has been shown to bind to acetylated histones through its bromodomain (Hassan et al. 2002).

Acetylation of lysine residues in histones is highly dynamic and regulated by two families of enzymes with opposing activity, histone acetyl transferases (HATs) and HDACs (Bannister et al. 2011). Interestingly, despite the positive correlation found between histone acetylation and transcriptional activity high stable acetylation levels are not strictly important for transcription, but instead a rapid turnover of high acetylation and deacetylation is key to this process. Turnover of histone acetylation can indeed be very rapid (down to one to five minutes), and remarkably the domains with the fastest turnover are associated with active transcription, whereas domains with slow turnover (up to 30-60 minutes) are not (Covault et al. 1980; Waterborg 2002). Genome-wide studies, have recently demonstrated transient binding by HATs and HDACs and subsequent dynamic cycle of acetylation and deacetylation at active genes, as well as genes poised for activation, however no detectable levels of HATs or HDACs were found at inactive genes (Wang et al. 2009b). High levels of HAT and HDAC binding were found at active genes whereas at poised genes much lower levels were

present. This could provide an advantage by allowing rapid upregulation of genes that are primed, in response to, for example, hormonal cues or growth factor signalling (Wang et al. 2009b).

The nuclear HATs that acetylate histones within the genome can be grouped into three families: Gcn5-related N-acetyltransferases, the MYST family (MOZ, Ybf2/Sas3, Sas2 and Tip60) and EP300 and its paralogue CBP (Hodawadekar et al. 2007; Lee et al. 2007). The first histone modifying enzyme identified was Gcn5 (Brownell et al. 1996). It was known from studies in yeast as a transcriptional activator and hence the link that histone acetylation directly affects transcription was established. In addition to the transcriptional co-activator function of all HATs, some such as PCAF and P300/CBP have also been shown to target non-histone proteins for acetylation, including p53 and MyoD (Hodawadekar et al. 2007; Lee et al. 2007). HDACs have been identified as transcriptional co-repressors and can be divided into 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 et al. 2008). Within the sirtuin family, there are seven members in humans (SIRT1– 7, related to yeast Sir2; also known as class III HDACs) (Haigis et al. 2006). Due to the important roles that HATs and HDACs have in regulating gene expression their misregulation has been linked to tumourigenesis in numerous studies (Bhaumik et al. 2007; Sharma et al. 2010; Bannister et al. 2011), and thus targeting their activity could provide new avenues in the fight against cancer. Indeed, HDAC inhibitors have been found to be effective in inhibiting tumour growth and, for example, suberoyllanidine hydroxamic acid, which is an HDAC inhibitor, has been approved for use in the clinic for treatment of T cell cutaneous lymphoma (Sharma et al. 2010).

1.1.3.2 Histone methylation

Unlike acetylation, histone methylation does not alter the charge of histone proteins and thus histone methylation does not seem to directly affect chromatin structure. Histone methylation does, however, allow binding of numerous effector proteins with various consequences relating to chromatin structure and gene regulation (Bannister et al. 2005; Taverna et al. 2007; Ng et al. 2009; Bannister et al. 2011). Histone methylation mainly occurs at lysine and arginine residues (methylation of histidine and proline has also been reported) of histone H3 and H4 N-terminal tails (Fig. 1.2), and it has been linked with several different cellular processes including DNA repair, replication, transcriptional activation and repression (Kouzarides 2007; Wolf 2009). Distinctly from acetylation, the histone methylation modification has added layers of complexity due to the fact that lysine residues may be

mono-, di- or tri-methylated and arginine residues may be mono-, symmetrically or asymmetrically di-methylated (Bannister et al. 2011). The different outcomes that histone methylation is associated with depend on the site and extent of methylation [unlike acetylation, which broadly correlates with transcriptional activation (Kouzarides 2007)].

Histone methylation at arginine residues is catalysed by several protein arginine methyltransferases (PRMTs). In humans, 11 PRMT proteins have been identified, based on protein sequences and conserved catalytic domain, and enzymatic activity has been confirmed with eight of them (Copeland et al. 2009; Wolf 2009). Histone arginine methylation has been shown to occur on H3R2, H3R8, H3R17 and H3R26 as well as on H4R3 (Ng et al. 2009). Transcriptional activation has been associated with methylation of H4R3 mediated by PRMT1 (Strahl et al. 2001; Wang et al. 2001) as well as with H3R2, H3R17 and H3R26 mediated by PRMT4 (also known as CARM1) (Schurter et al. 2001; Bauer et al. 2002). Interestingly, however, methylation of H3R8 and H4R3 by PRMT5 has been shown to be associated with transcriptional repression (Fabbrizio et al. 2002; Pal et al. 2003), demonstrating the complex nature of regulation by arginine methylation. Recent studies have shown that H3 arginine methylation by PRMT4 has a specific role in determining embryonic and extraembryonic lineage commitment by blastomeres in the early embryo as well as in mES cells (Torres-Padilla et al. 2007; Wu et al. 2009).

Histone lysine methyltransferases (HKMTs) can catalyse methylation on at least histone H3 lysine 4, 9, 14, 23, 27, 36 and 79 as well as histone H4 lysine 12 and 20 (Fig. 1.2). Six methylation marks have been most extensively studied and of these three have been implicated in activation of transcription: H3K4, H3K36 and H3K79; and three have been implicated with transcriptional repression: H3K9, H3K27 and H4K20 (Kouzarides 2007; Ng et al. 2009). Genome-wide studies have shown that H3K4 methylation is very closely associated with transcription start sites (TSS) of actively transcribed genes and H3K36 and H3K79 methylation are present downstream of the TSS (i.e. in the gene body) of actively transcribed genes (Koch et al. 2007; Wang et al. 2008; Djebali et al. 2012; Dunham et al. 2012; Sanyal et al. 2012; Thurman et al. 2012). The same studies have also seen the exclusion of H3K9, H3K27 and H4K20 methylation from regions of active transcription. H3K9 and H4K20 methylation are predominantly found in regions of constitutive heterochromatin (Peters et al. 2003; Martin et al. 2005) and H3K27 methylation at facultative heterochromatin sites, such as the inactive X-chromosome (Peters et al. 2003; Rougeulle et al. 2004). However, increased layers of complexity have been added to the interplay of histone lysine methylation and gene regulation with reports that indicate association of

H3K9me1, H3K27me1, H3K79me1 and H4K20me1 with actively transcribed genes, and H3K9me3, H3K27me3 and H3K79me3 with repressed genes (Barski et al. 2007). Interestingly, domains that possess both activating H3K4me3 and the repressive H3K27me3 marks have been found in mES cells (Azuara et al. 2006; Bernstein et al. 2006; Voigt et al. 2012). These bivalent domains in pluripotent cells are thought to allow for low level of activity while keeping the lineage differentiation options open for the cell. Once lineage commitment has occurred through differentiation the loci in question retain either the active or repressive marks according to their specific lineage and transcriptional status (Azuara et al. 2006; Bernstein et al. 2006; Voigt et al. 2012).

Unlike HATs, which often have numerous target residues, histone methyl transferases are typically more specific with regards to their target residues, resulting in a very large family of mammalian HKMTs. All human HKMTs, apart from the DOT1-like histone methyltransferase, contain a catalytic core formed by the approximately 130 amino acid SET domain [named after *D. melanogaster Su(var)3-9, Enhancer of zeste (E(z))* and *trithorax (trx)*] (Bhaumik et al. 2007; Kouzarides 2007; Ng et al. 2009). The first HKMT to be identified was Suv39h1, which targets H3K9 (Rea et al. 2000). SET domain-containing HKMTs, in general, target residues on the H3 and H4 histone tails. Other mammalian HKMTs that target H3K9 include Suv39h2, G9A and ESET. PR-Set7, Suv4-20h1 and Suv4-20h2, all catalyse histone H4K20 methylation (Shilatifard 2006; Sims et al. 2006; Bhaumik et al. 2007). H3K27 and H3K4 trimethylation are catalysed by the Polycomb group protein Ezh2 and the Trithorax/MLL complex respectively (Shilatifard 2006; Copeland et al. 2009).

Both Polycomb group (PcG) and Trithorax group (TrxG) proteins have been implicated in methylation of histone lysine residues. PcG and TrxG proteins were discovered in *Drosophila* as gene mutations at these loci affect the spatially restricted expression pattern of *Hox* genes and developmental patterning was affected (Lewis 1978). PcG and TrxG proteins have been shown to play a key role in defining and maintaining epigenetic methylation patterns that define cell identity during development (Ringrose et al. 2004; Sparmann et al. 2006). It is now clear that PcG and TrxG proteins regulate many genes in addition to the *Hox* genes and have an effect on several cellular processes, including proliferation, genomic imprinting, cancer, X-inactivation and pluripotency (Heard 2005; Azuara et al. 2006; Bernstein et al. 2006; Boyer et al. 2006; Schwartz et al. 2006; Schuettengruber et al. 2007). PcG proteins have been shown to act to repress loci and conversely TrxG proteins to induce activation of loci (Schuettengruber et al. 2007).

TrxG can be separated into two main classes. One class includes SET domain factors such as Trx, Ash1 and the vertebrate MLL, and their associated proteins. The second class includes components of ATP-dependent chromatin remodelling complexes such as SWI/SNF and NURF (Schuettengruber et al. 2007). As indicated previously, TrxG proteins have been shown to methylate H3K4, an activation associated modification (Shilatifard 2006; Ruthenburg et al. 2007). The basal MLL protein has been shown to be recruited to H3K4 by site-specific transcription factors and basal transcriptional machinery as well as direct recognition of histone modifications (Ruthenburg et al. 2007). Evidence from *Drosophila* has also shown that Ash1 can be recruited to specific loci by long-non coding RNA rising from the locus in question (Sanchez-Elsner et al. 2006). It has also been shown that, for example, the NURF complex can bind to H3K4me3 via a plant homeodomain (PHD) finger domain of the Nurf-301 protein (Li et al. 2006; Wysocka et al. 2006).

Contrasting the activity of TrxG proteins, the PcG proteins are associated with repressed chromatin, and the H3K27 tri-methylation mark. Two main protein complexes have been identified in humans: Polycomb repressive complex 1 (PRC1) and PRC2, a third complex, PhoRC, is known in *Drosophila*, however it remains unclear whether this complex also exists in humans (Schuettengruber et al. 2007). PRC2 complex contains four core components: EZH2, EED, SUZ12 and RbAp46/48 (Cao et al. 2004). PRC2 acts by catalysing the tri-methylation of H3K27 via the SET-domain protein EZH2 (Cao et al. 2004). This mark is then recognised by PRC1, which in humans is comprised of RING1A/B, BMI-1, Polycomb and polyhomeotic proteins as well as some other associated factors (Schuettengruber et al. 2007). RING1A/B in cooperation with BMI-1 has been shown to catalyse ubiquitination of lysine 119 of histone H2A, another repressive histone modification (de Napoles et al. 2004; Wang et al. 2004). The action of PRC2 and PRC1 has been shown to prevent active transcription of a number of developmental genes in mES cells, highlighting the collaborative action of PRC2 and PRC1 in maintaining silent chromatin (Jørgensen et al. 2006; Stock et al. 2007).

Methylation marks on histones are read and bound by numerous different effector molecules, for example, the already mentioned NURF complex and PRC1 complex. There are more distinct domain types recognising lysine methylation than any other modification. These include PHD fingers and the so called 'royal' family of domains, comprising chromodomains, Tudor, PWWP and MBT domains (Bannister et al. 2011). Known examples of effector protein binding to methyl histone marks include, for example, the binding of ING2 protein to the H3K4me3 mark. Following DNA damage ING2 tethers a repressive HDAC

complex to active proliferation specific genes downregulating their activity (Shi et al. 2006). Another important example is the binding of HP1 via its N-terminal chromodomain to H3K9me3 repressive mark, which is important for the overall structure of heterochromatin (Bannister et al. 2001; Lachner et al. 2001). The methylation marks can also be directly bound by enzymes with histone demethylase activity such as the case of the tandem Tudor domain containing JMJD2A histone demethylase binding to H3K4me3 (Huang et al. 2006).

Methyl groups can be removed from histone residues by histone demethylase enzymes, such as JMJD2A. The first histone lysine demethylase to be discovered was the lysine demethylase 1 (LSD1), which had previously been identified as a member of several HDAC complexes and catalyses demethylation of mono- and di-methyl H3K4 (Shi et al. 2004). The mechanism by which LSD1 (and the related protein LSD2) initiates demethylation prevents them from catalysing tri-methylated substrates (Klose et al. 2007). In contrast, another class of lysine demethylase enzymes, which includes JMJD2A, containing a JmjC domain as their catalytic core are able to demethylate tri-methyl groups. Members of this large family of proteins utilise iron and α-ketoglutarate as co-factors in an oxidative reaction for their demethylase activity (Tsukada et al. 2006). The JmjC containing proteins are able to demethylate H3K9, H3K27 and/or H3K36 residues (Klose et al. 2007) as well as H3K4 as mentioned above. Methylation at arginine residues can undergo either deamination, which converts the arginine to citrulline, or, as discovered more recently, demethylation, catalysed by PADI4 and JMJD6 enzymes, respectively (Nakashima et al. 2002; Wang et al. 2004; Chang et al. 2007; Klose et al. 2007).

1.1.4 Spatial constraints – nuclear organisation

A striking feature of nuclear architecture is the existence of distinct structural and functional compartments and the non-homogenous distribution of components of the transcription machinery (Lanctot et al. 2007). It is known that, within the nuclei of cells, individual chromosomes occupy distinct positions referred to as chromosome territories (Cremer et al. 2006; De et al. 2010). As a result of different compaction levels, different chromosome segments adopt a complex organisation and topography within their chromosome (Dostie et al. 2012). One of the widely conserved features of chromosomal positioning is the polarized distribution of gene-poor and gene-rich regions of chromatin to the nuclear periphery and nuclear interior, respectively (Deniaud et al. 2009; Ferrai et al. 2010).

Chromosome territories are dynamic structures and individual chromosomes can be repositioned in relation to nuclear structures and other chromosome territories. Increasing

evidence has linked repositioning of genomic regions in nuclear space to regulation of gene expression (Gasser 2002; Ferrai et al. 2010). For instance, live-imaging experiments using induced activation of an integrated array of lac operators with viral activators in mammalian cells have shown that activation of the locus leads to movement from the periphery to the interior of the nucleus, and furthermore the movement was dependent on nuclear actin and myosin suggesting the involvement of an active process (Tumbar et al. 2001; Chuang et al. 2006). Importantly, chromosomes are distributed tissue-specifically with respect to their position relative to each other and the nuclear architecture; this suggests a role for chromatin repositioning during development and cell differentiation (Parada et al. 2004; Laster et al. 2010). Interestingly, inhibition of RNA polymerase II (PolII) driven transcription leads to a reduction in relocation of chromosome domains outside their native chromosome territories as well as in intermingling of chromosomal domains between different chromosomes (Mahy et al. 2002; Branco et al. 2006). This suggests that the events are at least partly driven by the process of transcription (Fraser et al. 2007).

Recently it has become evident that the nuclear lamins have an important role in the structural organisation and function of chromatin (Dechat et al. 2008). Extensive genome-wide analysis of regions of chromosomes that interact with the nuclear lamina has mapped lamina-associated domains (LADs) of 0.1 to 10 megabase pair (Mb) in size at more than 1300 specific sites throughout the genome (Guelen et al. 2008). The data from LAD analysis correlates with the previously seen polarized distribution of gene-poor regions to the nuclear periphery and gene-rich to the interior. The mean gene density within LADs was found to be approximately half of that outside LADs (Guelen et al. 2008). LADs are also highly enriched in chromatin modifications associated with heterochromatin, such as H3K27me3 and H3K9me2, and have very low levels of active transcription marker H3K4me2 and PolII (Guelen et al. 2008). Association of heterochromatin with the nuclear periphery has also been previously reported, such as with the inactive X which is often observed at the nuclear periphery (Burke et al. 2002). Interestingly, mutation in human nuclear lamin A has been found to result in reduced amounts of constitutive and facultative heterochromatin associated with H3K9 and H3K27 methylation (Shumaker et al. 2006).

Analysis of LADs in neuronal differentiation of mES cells has recently provided evidence for the dynamic relocation of chromatin domains in response to changing cellular identity (Peric-Hupkes et al. 2010). Specifically, loci linked to neuronal identity such as *Pcdh9* were released from the nuclear lamina upon differentiation from mES cells to neuronal progenitor cells, and conversely ES cell specific genes, such as *Oct4*, *Nanog*, *Klf4* and *Rex1/Zfp42*

became associated with the nuclear lamina (Peric-Hupkes et al. 2010). Interestingly, evidence also suggests that loci can be released from the nuclear lamina upon initial cellular differentiation, but remain repressed until further differentiation, thus these loci can be thought to have been 'unlocked' for future expression by movement away from the nuclear lamina (Peric-Hupkes et al. 2010). Association with the nuclear periphery has also been show to correlate with late DNA replication timing of specific loci, furthermore activation of a locus upon differentiation and migration from the periphery to the interior is also accompanied by a shift of replication timing from late to earlier in the S-phase (Hiratani et al. 2008; Peric-Hupkes et al. 2010)

1.1.5 Temporal constraints – DNA replication timing

Replication of DNA precisely once during a cell cycle is fundamental to all biological systems, and the basic mechanism ensuring exact and complete duplication of the genome is highly conserved across all kingdoms (Gilbert 2010; Masai et al. 2010). Unlike in prokaryotes, which replicate their entire genome from a single origin of replication, eukaryotic cells replicate their genome by initiating bidirectional DNA replication from multiple origins of replication spread across the genome on average at 100 (kilobase pair) kb from each other (Huberman et al. 1968; Goren et al. 2003). The rate of new DNA synthesis in eukaryotes is approximately 1-3 kb/min, hence if all the replication origins were firing simultaneously the entire genome could be duplicated in approximately one hour (Hiratani et al. 2009b; Méndez 2009). However, this is not the case, and activation of replication origins occurs in a defined sequential order during the S-phase of the cell cycle, with some origins firing early and some late in S-phase (Camargo et al. 1982). The patterns of replication timing are inherited by the daughter cells during mitosis and are conserved in a particular cell type; they do, however, undergo changes during development and differentiation relating to activation or silencing of specific loci (Cimbora et al. 2001; Hiratani et al. 2004; Perry et al. 2004; Hiratani et al. 2008; Hiratani et al. 2010).

Eukaryotic replication origins were first identified from yeast. In *S. cerevisiae* a specific segment of chromosomal DNA was determined to confer autonomous replication when inserted into a plasmid; this sequence was defined as an autonomously replicating sequence (ARS) (Stinchcomb et al. 1979). Further investigation in *S. cerevisiae* revealed that the ARS contain sequence specific origins of replication that could be disrupted by induced mutations (Van Houten et al. 1990). The time of firing varies between different origins of replication, however no sequence correlation has been found between late or early

replicating ARSs in comparative sequence analyses (Wintersberger 2000). In metazoans no sequence specificity has been found for origins of replication and it has been reported that circular DNA that contains matrix attachment sites or nuclear localisation signals can replicate in transformed mammalian cell lines without any apparent sequence specificity (Krysan et al. 1993; Schaarschmidt et al. 2004). The larger genome size and the plasticity required for a large repertoire of different transcriptional programmes found in different specialised cells in metazoans seem the most plausible explanations for the differences observed in *S. cerevisiae* and metazoan replication origin specification (Mechali 2001).

Metazoan replication occurs mostly in large domains that are replicated at the same time in the cell cycle (Schübeler et al. 2002; Woodfine et al. 2004; Karnani et al. 2007; Farkash-Amar et al. 2008; Hiratani et al. 2008). Domains of distinct replication timing were first discovered by pulse labelling cells with the nucleotide analogue 5-bromo-2-deoxyuridine (BrdU) at different times during S-phase and visualizing its incorporation into discrete areas of condensed mitotic chromatin (Latt 1973; Stubblefield 1975). When these domains were compared with Giemsa-stained banding, which roughly demarcate regions of euchromatin and heterochromatin along the length of a chromosome, a general overlap was revealed of light-staining R bands with early-replicating domains and dark staining G-bands with late-replicating domains (Dutrillaux et al. 1976). Replication timing analysis by microarrays confirmed these correlations in high-resolution maps of human chromosomes (Woodfine et al. 2004; Costantini et al. 2008). When DNA replication dynamics are observed with immunofluorescence by detecting the BrdU incorporation into newly synthesised DNA, replication foci appear in a punctate pattern that changes over the progression of the S-phase. During early S-phase numerous small foci are distributed all across the nucleus. In middle S-phase foci are preferentially assembled around the nucleoli and nuclear periphery, while late S-phase is characterized by larger clusters of foci that seem to correspond to heterochromatic regions (Fig. 1.3) (O'Keefe et al. 1992; Dimitrova et al. 1999).

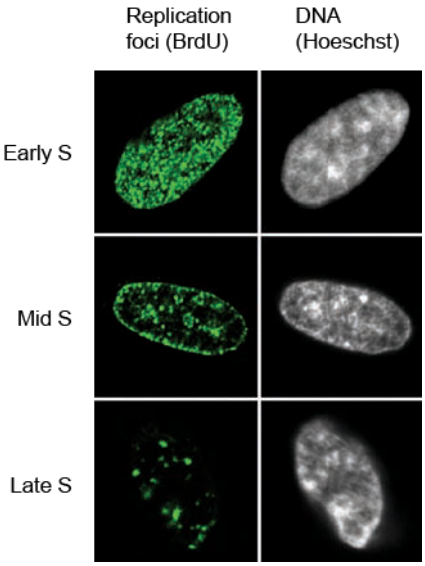


Figure 1.3 Replication timing foci patters during early, middle and late S-phase of cell cycle.
Replication foci were visualized by 30 min pulse-label with 5-bromo-2-deoxyuridine (BrdU) and immunofluorescence detection. DNA was stained with Hoechst. Adapted from (Méndez 2009).

In every multi-cellular system examined, early replication has been shown to correlate strongly with transcription; this has been statistically confirmed in *Drosophila*, human and mouse cells (Schübeler et al. 2002; Woodfine et al. 2004; Karnani et al. 2007; Farkash-Amar et al. 2008; Hiratani et al. 2008; Hiratani et al. 2010). Early replication has been shown to also correlate to accessible chromatin that carries acetylated histones (Vogelauer et al. 2002; Lin et al. 2003; Perry et al. 2004; Ryba et al. 2010) as well as H3K4me3, H3K36me3 and H4K20me1 (known markers of active transcription) (Hiratani et al. 2008; Ryba et al. 2010). Late replication on the other hand has been shown to generally correlate with repression of transcription, however, small number of late replicating (less than 5% of total in mouse) genes with CpG-island rich, strong promoters are expressed and thus late replication does not preclude expression (Schübeler et al. 2002; Woodfine et al. 2004; Karnani et al. 2007; Farkash-Amar et al. 2008; Hiratani et al. 2008; Hiratani et al. 2010). Although late replication is a characteristic of constitutive heterochromatin and some facultative heterochromatin (Gilbert 2002; Peters et al. 2005), no direct correlation to late replication was found with H3K9me3, H3K27me3 or H4K20me3 repressive histone marks (Hiratani et al. 2008; Ryba et al. 2010). Interestingly, however, data from a mammalian model system of PEV provides direct evidence that the local chromosome environment is important in determining timing of replication (Azuara et al. 2003). This study demonstrated that in transgenic fibroblasts when integrated within late-replicating heterochromatin the replication timing of a λ5 transgene changed to roughly coincident with that of the neighbouring γ-satellite repeats, in contrast to the early replication found in normal fibroblasts, notably the λ5 gene is not expressed in either the normal or transgenic fibroblasts. Many expressed genes do replicate early and examples of correlation of replication timing with gene expression include house-keeping genes that constitutively replicate early in S-phase whereas many tissue-specific genes only replicate early in their tissue of expression (Goren et al. 2003). Furthermore, in the case of imprinted genes, the active allele replicates early and the inactive late during S-phase (Simon et al. 1999). However, there are also many inactive genes that replicate early. These include, for example, genes involved in stress response and apoptosis, which are not expressed during normal proliferation but need to be accessible for rapid transcription when necessary, as well as lineage-specific genes that are not expressed in mES cells, but are upregulated upon differentiation to their cognate lineage. It is therefore suggested that a correlation exists between early replication timing and the potential for gene expression rather than actual expression (Azuara et al. 2006; Hiratani et al. 2008; Hiratani et al. 2009a; Hiratani et al. 2009b; Méndez 2009).

Of all the histone modifications examined, acetylation seems to correlate with early replication timing most strongly (Méndez 2009). Recent experiments have shown that recruitment of HATs to the human β-GLOBIN locus in a non-erythroid cell line, where β-GLOBIN is normally not expressed, would change its replication timing from late to early and conversely tethering an HDAC to the β-GLOBIN 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). Use of HDAC inhibitors (trichostatin A and sodium butyrate) was sufficient to remove replication imprinting signals from two loci *WT1* and *IGF2* (Bickmore et al. 1995); this fact further supports the correlation between histone acetylation and replication timing. Histone acetylation has been hypothesised to render replication origins firing more efficient or earlier at the affected domains (Vogelauer et al. 2002; Aparicio et al. 2004). Influence of other histone modifications remains unclear and may be restricted to specific chromatin structures, such as repetitive satellite DNA that underwent replication timing changes in mES cells lacking for example Dnmt1 (Jørgensen et al. 2007; Méndez 2009). Replication timing has, however, been found to be correlated with genomic features in the actual DNA sequence, such as gene density, guanine-cytosine (GC)-content and long interspersed nuclear element (LINE) density. Locus specific and genome-wide studies have shown that GC-rich and gene dense regions tend to replicate early in S-phase, while AT-rich, gene-poor regions with high LINE density tend to replicate late in S-phase (Schübeler et al. 2002; Hiratani et al. 2004; Woodfine et al. 2004; Hiratani et al. 2008; Ryba et al. 2010).

Further experimental evidence has suggested that the association with acetylation and early replication is not just a passive element and that replication timing might play a role in maintaining heritable epigenetic states through mitosis. A reporter gene microinjected into early S-phase nuclei was assembled into nucleosomes with acetylated histones and expressed with higher efficiency than the same reporter gene microinjected into nuclei of cells in late S-phase, which was preferentially assembled into deacetylated nucleosomes (Zhang et al. 2002). This result strongly suggests that the timing of replication has the capacity to influence chromatin structure. Support for this hypothesis comes from recent studies where it was shown that DNA sequences originally packaged into nucleosomes containing deacetylated histones during late S-phase become reassembled with acetylated histones after undergoing replication in early S-phase, and vice versa (Lande-Diner et al. 2009). Consistent with this function different sets of chromatin assembly factors and histone-modifying enzymes are present at the replication fork in early and late S-phase (McNairn et al. 2003). The experimental evidence suggests a link between replication timing

and chromatin structure, however precisely how this relationship operates remains unresolved. It could be hypothesised that 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 et al. 2009a).

During development replication timing is subject to changes across different loci. Studies comparing the replication timing of a panel of approximately 50 tissue-specific candidate loci in mES and ES-derived neuronal progenitor cells demonstrated that commitment to the neural lineage results in shifts of replication timing for many loci. Specifically it was observed that many ES-associated genes shifted their replication timing towards late S-phase and concurrently the replication timing of neural lineage associated genes was advanced towards early S-phase (Hiratani et al. 2004; Perry et al. 2004). Furthermore, significant deacetylation was observed at loci that underwent a shift from early to late replication upon differentiation from ES to neural progenitor cells (Perry et al. 2004). A subsequent candidate based study compared the replication timing of a panel of lineage specification associated loci in pluripotent mES cells as well as haematopoietic stem cells (HSC) and T lymphocytes (two differentiated cell types) (Azuara et al. 2006). Many lineage-specific loci were found replicating early in mES cells and later in HSC and T cells. This difference was observed despite the fact that these neural-associated loci were not expressed in any of the three cell types analysed. The same study also used replication timing analysis to distinguish mES cells from embryonic carcinoma (EC) cells that possess a more restricted differentiation capacity, demonstrating that the replication timing profiles were reflective of the potential of the cells. Complementing these findings, further genome-wide analysis confirmed similar and extensive shifts in replication timing upon stem cell differentiation, as seen using candidate based approaches (Hiratani et al. 2008). Taken together, these results indicate that replication timing profiling can be used to identify the developmental stage and lineage potential of a given cell population.

1.2 Early mammalian development and pluripotency

1.2.1 Early embryo development

The fertilised zygote has the potential to give rise, via intermediate cell types, to all the cells necessary to form an adult organism. During normal development cells get progressively more specialised and more restricted in their potential, through the numerous cell divisions and directed differentiation to specialised lineages. The limited availability and restrictions

on the use of human embryos and *in vivo* tissues from post-implantation stages mean that the cellular and molecular mechanisms governing early human development remain poorly understood (Niakan et al. 2012). The understanding of mammalian embryology, and also of pluripotent stem cells, is based mainly on studies in the mouse (Pera et al. 2004a), however

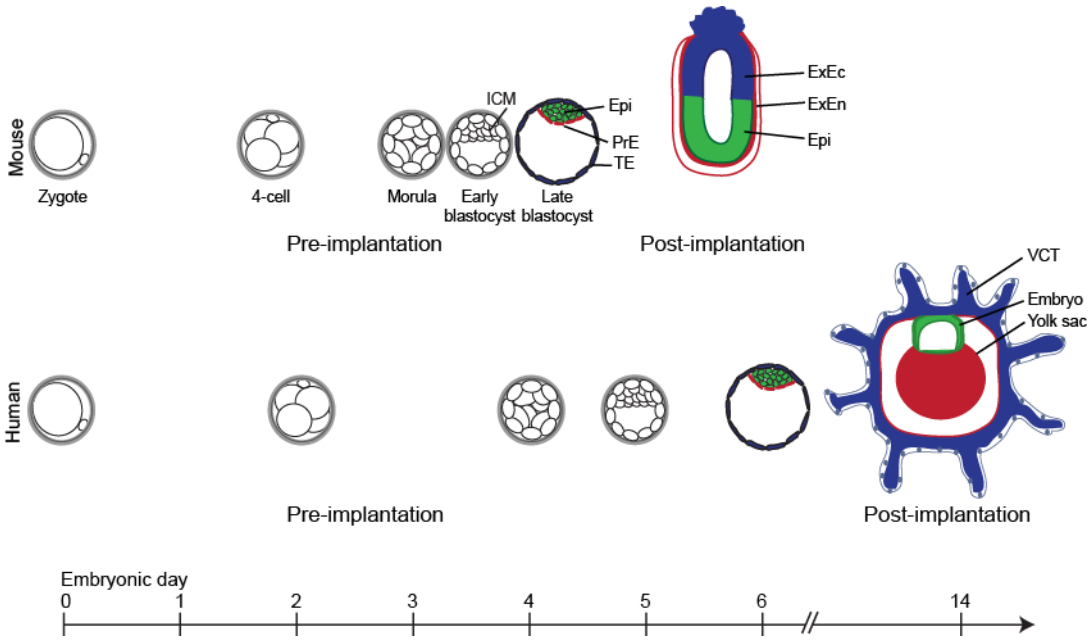


Figure 1.4 Schematic representation of early stages of mouse and human embryo development.
Prior to implantation, both human and mouse embryos similarly undergo cell divisions culminating in the development of a blastocyst comprising a discernible inner cell mass (ICM) and trophoectoderm (TE). The timing of compaction and blastocyst formation differs significantly between mouse and human embryos, with human embryos showing delayed development compared with mouse embryos. The mouse blastocyst forms between days 3 and 4, whereas human blastocysts form between days 5 and 6. Both human and mouse pre-implantation blastocysts comprise an outer layer of TE cells, which form the trophoblast lineage of the placenta, and an ICM that segregates into epiblast (Epi) and primitive endoderm (PrE) layers. Epiblast cells eventually give rise to all the tissues of the future foetus, whereas the PrE gives rise to extra-embryonic endoderm (ExEn) cells that will form the yolk sac. In the mouse, the TE gives rise to a proliferative stem cell pool of extra-embryonic ectoderm (ExEc) cells. By contrast, human TE gives rise to placental villous cytotrophoblast (VCT) cells. Figure adapted from (Niakan et al. 2012).

species-specific differences may limit the extrapolation of some findings to human embryo development (Niakan et al. 2012).

In mouse development, the fertilized zygote undergoes a series of cleavage divisions to give rise to a 4-cell embryo at day 2 of embryonic development. The embryo then develops further and at embryonic day 3 it forms a compacted morula (Fig. 1.4). Asymmetrical cell divisions at the morula stage along a basolateral cleavage plane give rise to two distinct subpopulations: the inner cell mass (ICM) and trophoectoderm (TE). Coincident with the segregation of the TE and ICM, the embryo cavitates to form the blastocyst (Arnold et al. 2009). Cell fate studies using mouse chimaeras have shown that the ICM cells are pluripotent and able to give rise to all somatic cell types (Rossant 2008). TE cells are

multipotent, only being able to contribute to placental tissues, including the extra-embryonic ectoderm (ExEc) (Arnold et al. 2009; Niakan et al. 2012). The early blastocyst arises typically at embryonic day 3.5 and matures to the late blastocyst with further segregation of ICM cells and the formation of the epiblast (Epi) and primitive endoderm (PrE) (also known as hypoblast) (Fig. 1.4) at embryonic day 4.5 just prior to implantation (Ralston et al. 2005; Rossant et al. 2009). The cells of the epiblast layer retain pluripotency and can contribute to all lineages of the adult organism, as shown by microinjection experiments (Gardner 1998). Epiblast cells in the pre-implantation blastocyst mature into primed epiblast of the post-implantation embryo, which upon gastrulation gives rise to the three germ layers (ectoderm, mesoderm and endoderm) as well as primordial germ cells (PGCs) (Arnold et al. 2009). Similarly to the TE, the cells from the PrE are multipotent, giving rise to the parietal and visceral endoderm layers lining the yolk sac (Bielinska et al. 1999; Arnold et al. 2009). In addition to forming the tissues supporting the developing embryo, the cells of the TE and PrE lineages also provide signals that direct the patterning and differentiation of the embryonic lineages (Rossant et al. 2009). Human and mouse embryos appear morphologically similar during pre-implantation development, however there are significant differences in developmental timing, which are likely caused by key molecular differences (Fig. 1.4) (Niakan et al. 2012). Upon implantation the Epi, PrE and TE cells in human embryos give rise to the embryo, yolk sac and villous cytotrophoblast, respectively.

1.2.2 Embryo-derived stem cells

1.2.2.1 Mouse embryo-derived stem cells

A number of different types of stem cells have been derived from the mouse embryo. In the pre-implantation blastocyst three lineages can be distinguished, namely the Epi, PrE and TE (Fig. 1.4). Stem cells have been derived from all of the three lineages, namely ES cells from the Epi (Evans et al. 1981; Martin 1981; Brook et al. 1997; Nichols et al. 2009a), trophoblast stem (TS) cells from the TE (Tanaka et al. 1998) and extraembryonic endoderm stem (XEN) cells from the PrE (Kunath et al. 2005). These three different cell types share the defining features of stem cells, i.e. being able to self-renew without differentiating and thus retaining their developmental potential. However, they display distinct morphology, gene expression patterns and growth factor requirements, reflecting their lineage origins (Rossant 2008). All three lines recapitulate the lineage of their cognate blastocyst precursors when injected into blastocysts to generate chimeric animals (Yamanaka et al. 2006). Pluripotent ES cells are able to contribute to all the somatic lineages of the developing foetus, but only poorly to PrE

and rarely to TE lineages (Beddington et al. 1989). TS cells are only able to contribute to trophoblast lineages (Tanaka et al. 1998) and XEN cells are only able to contribute to the extra-embryonic endoderm lineages (Kunath et al. 2005). Key transcriptional factors that regulate blastocyst lineage allocation as well as the cell identity of embryo-derived stem cells have been identified and these include *Oct4/Sox2/Nanog*, *Cdx2/Eomes/Gata3* and *Gata6/Gata4/Sox17* in Epi & ES cells, TE & TS cells and PrE & XEN cells, respectively (Ralston et al. 2005; Rossant 2008; Home et al. 2009; Morris et al. 2010).

Stem cells have also been isolated from the post-implantation mouse embryo. Embryonic germ (EG) cells were derived from early PGC cells between embryonic day 8.5 and 11.5 (Matsui et al. 1992; Resnick et al. 1992). EG cells are pluripotent and thus contribute to chimaeras, including germline transmission, when injected into blastocysts (Matsui et al. 1992; Labosky et al. 1994). More recently, stem cells have also been derived from the Epi layer of post-implantation embryonic day 5.5 to 6.5 mouse embryos. These cells, named epiblast stem (EpiS) cells, are pluripotent and can form derivatives of all three embryonic germ layers *in vitro* and *in vivo* when injected into immunodeficient mice to form teratocarcinomas. However, EpiS cells contribute very poorly, if at all, to chimaeras and the germ line when injected into blastocysts (Brons et al. 2007; Tesar et al. 2007). Interestingly, recent studies have demonstrated that mES cells can be converted to mEpiS cells and vice versa by genetic manipulation and modification of the culture environment, highlighting the plasticity of the pluripotent state *in vitro* (Bao et al. 2009; Guo et al. 2009; Hanna et al. 2009).

1.2.2.2 Mouse and human embryonic stem cells

During mammalian development, the pluripotent state exists only transiently *in vivo* and is lost as the embryo develops and the cells differentiate to somatic lineages. It is, however, possible to isolate both mouse and human pluripotent ES cells *ex vivo* under defined growth conditions that promote proliferation of the pluripotent cells and inhibit differentiation (Yu et al. 2008).

Mouse ES cells were first derived in 1981 from ICM explants from the early mouse blastocysts that were grown in serum-containing media and feeder cell layers (Evans et al. 1981; Martin 1981). Under appropriate culture conditions mES cells are able to propagate indefinitely in a self-renewing state, while retaining their pluripotency. If injected back into a blastocyst, they are able to contribute to all tissues of the adult organism including the germ line (Bradley et al. 1984). Further studies have shown that mES cells can be derived from

single blastomeres of 2- to 8-cell stage mouse embryos (Eistetter 1989; Chung et al. 2006; Wakayama et al. 2007). It is thought that during the derivation process the isolated blastomeres undergo cell division creating a 'mini-blastocyst' environment, from which mES can arise (Nichols et al. 2011). The epiblast of pre-implantation embryos has been identified as the source of mouse ES cells, and it is thought that all epiblast cells may transiently have the capacity to become ES cells (Brook et al. 1997; Nichols et al. 2009a; Nichols et al. 2011).

Derivation of the first human ES cells in 1998 was accomplished by isolating ICM explants from cultured blastocyst embryos and growing them in serum containing media and feeder cells, in a similar manner to mES cells (Thomson et al. 1998). The considerable delay between the successful derivation of mES cells and hES cells is likely due to species-specific ES cell differences and/or suboptimal human embryo culture media (Yu et al. 2008). Several previous attempts indeed failed to derive hES cells. For example, isolation of human blastocysts was previously reported in 1994 (Bongso et al. 1994), but their subsequent culture in media supplemented with leukaemia inhibitory factor (LIF) and serum - conditions that allow the derivation and propagation of mES cells - resulted in cell differentiation. Improvements in human embryo culture (Gardner et al. 1998) and previous experience with non-human primate cell lines (Thomson et al. 1995; Thomson et al. 1996) enabled the derivation of hES cells in 1998.

Human ES cells maintain the developmental potential to give rise to derivatives of all three germ layers even after prolonged undifferentiated proliferation as well as clonal derivation (Amit et al. 2000). For obvious ethical reasons the ultimate test for cell potency, chimaera formation, is not possible with hES cells. The most stringent and commonly used alternative is the injection of cells to immunodeficient mice and the formation of teratocarcinomas (Thomson et al. 1998). Another alternative is *in vitro* differentiation in non-adherent conditions, which promote the formation of three-dimensional embryoid bodies (EBs) that mimic early embryo differentiation (Vallier et al. 2005b). In addition to being able to give rise to the three germ layers, and in contrast to mES cells (Beddington et al. 1989), hES cells are also able to give rise to extraembryonic lineages including TE (Itskovitz-Eldor et al. 2000; Reubinoff et al. 2000; Xu et al. 2002) as well as PrE –like cells (Pera et al. 2004a).

In the same way as mES cells, human ES 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). Whether any differences, especially with regards to developmental potential, exist between the hES cell lines derived from these various sources, is yet to be

determined. Recent results have demonstrated that independently derived hES cell lines may differ significantly in their differentiation propensity as tested by spontaneous differentiation with EB formation or directed differentiation methods (Osafune et al. 2008). Furthermore, recent studies have clearly reported variability between hES cell populations at the transcriptional level (Skottman et al. 2005; Wei et al. 2005; Adewumi et al. 2007; Guenther et al. 2010; Narva et al. 2010; Newman et al. 2010; Amps et al. 2011) and at the epigenetic level (Guenther et al. 2010; Bruck et al. 2011; Nazor et al. 2012).

Although the evidence mentioned above suggests that mES and hES cells differ in their developmental potency, they do nevertheless share high levels of alkaline phosphatase activity (Adewumi et al. 2007) and high expression of key TFs known to regulate the pluripotent state, which include the POU-family TF Oct4 (Palmieri et al. 1994; Nichols et al. 1998), homeodomain DNA-binding protein Nanog (Chambers et al. 2003; Mitsui et al. 2003) and the SOX-family TF Sox2 (Avilion et al. 2003; Fong et al. 2008). Mouse embryos negative for *Oct4* are unable to form an ICM and consist only of TE tissue (Nichols et al. 1998). Consistently, when *Oct4* is deleted from mES cells, they convert into TS-like cells (Niwa et al. 2000). In hES cells knock-down of *OCT4* expression levels leads to differentiation towards TE phenotype (Zaehres et al. 2005), suggesting similar function in both mES and hES cells. Mouse ES cells deficient for *Nanog* are still able to self-renew, however they show an increased propensity to differentiate towards XEN-like cells and are unable to give rise to mature germ cells (Chambers et al. 2003; Chambers et al. 2007). In hES cells downregulation of *NANOG* causes differentiation to extraembryonic lineages (Hyslop et al. 2005), again suggesting conserved function between mES and hES cells. The presence of maternally derived Sox2 has made deciphering its role more difficult. The maternal contribution is lost largely by the time of implantation, and after this *Sox2* null mouse embryos generate exclusively TE and PrE lineages (Avilion et al. 2003). Similarly *Sox2* deletion in mES cells also leads to their conversion in TS-like cells (Masui et al. 2007). The role of SOX2 also seems to be conserved in hES cells, as knock down of *SOX2* leads to differentiation towards a TE lineage (Fong et al. 2008).

In the light of recent evidence it has become clear that mES cell cultures are heterogeneous and contain cells in different interconvertible sub-states, which have significant consequences on their behaviour and differentiation ability. Expression levels of *Rex1* (an ICM related marker) (Toyooka et al. 2008), *Nanog* (Chambers et al. 2007) and *Stella* (an ICM and germ cell lineage related marker) (Hayashi et al. 2008) have been observed to be highly variable within mES cultures, contrasting the apparent homogeneity of *Oct4* and *Sox2*

expression levels. In hES cell cultures *NANOG* has also been shown to have variable expression levels, which also correlate with levels of, for example, *OCT4* expression in the same cells (Fischer et al. 2010). Furthermore, variable patterns of cell surface marker expression in pluripotent hES cells also points to the existence of interconvertible sub-states of the stem cell compartment (Enver et al. 2005; Stewart et al. 2006a). The cell surface marker stage specific embryonic antigen 3 (SSEA3) was used to isolate $SSEA3^+$ and $SSEA3^-$ hES sub-populations from pluripotent hES cultures, and subsequent analysis indicated that these sub-populations have different phenotypes with $SSEA3^+$ cells being truly pluripotent and $SSEA3^-$ possibly representing culture adapted cells that, although not truly pluripotent, retain multilineage differentiation potential (Enver et al. 2005). Moreover, similar interconvertible sub-states have also been identified in adult stem cell populations (Hu et al. 1997; Booth et al. 2000; Jones et al. 2007).

1.2.3 Extrinsic signalling pathways that regulate the pluripotent state *in vitro*

Despite the fact that both mES and hES cells are pluripotent and are controlled by the same key transcription factors (Oct4, Sox2 and Nanog) there are numerous differences between the two pluripotent cell populations. Phenotypically mES cell colonies are round and dome-shaped whereas hES cells grow in colonies that are flatter and often display more distinct colony borders (Pera et al. 2004b). The initial derivation of both mouse and human ES cells was done in serum containing media with mitotically inactivated fibroblast feeder cells (Evans et al. 1981; Martin 1981; Thomson et al. 1998). The fact that fibroblast feeder layers support both mouse and human ES cells, however, now appears largely to be a lucky coincidence, as the specific factors used to sustain mouse ES cells do not support human ES cells (Yu et al. 2008).

Undifferentiated mES cells can be maintained self-renewing in media containing foetal calf serum (FCS) supplemented with the cytokine LIF, which was identified from media conditioned with feeder cells (Smith et al. 1988; Williams et al. 1988). LIF and its related cytokines act via the gp130 and LIF receptors (Yoshida et al. 1994). Binding of LIF leads to the heterodimerisation of gp130 and LIF receptors and the subsequent activation of the JAK-STAT3 pathway (Yoshida et al. 1994; Niwa et al. 1998). Activated Stat3 translocates to the nucleus and regulates the levels of genes essential for mES pluripotent identity (Niwa et al. 1998; Raz et al. 1999). Mouse ES cells that have been transfected with constitutively active Stat3 are rendered LIF independent (Matsuda et al. 1999), whereas withdrawal of LIF

supplementation from normal mES cultured with media containing FCS leads to differentiation to a mixture of lineages (Nichols et al. 2001). When mES cells are cultured in serum free media, LIF alone is not sufficient to prevent differentiation, however in combination with supplementation by bone morphogenetic protein [BMP; a member of the transforming growth factor-β (TGF-β) family] the pluripotent identity of mES cells is sustained (Ying et al. 2003). BMPs induce expression of inhibitor of differentiation (Id) proteins via the activation of Smad1/5/8 pathway. The LIF/Stat3 and BMP/Id signalling act in combination to prevent differentiation towards mesoderm & endoderm and neuroectoderm respectively and thus maintain undifferentiated mES cells (Ying et al. 2003). More recently it has been shown that mES cells can be efficiently derived and maintained in an undifferentiated state in serum free medium by use of small molecules to inhibit specific kinases. The combination of two specific inhibitors (2i), PD184352 and CHIR99021, which target MEK/Erk and Gsk3 signalling pathways, respectively (Ying et al. 2008) allows for mES culture in the absence of cytokine supplementation. In essence, the inhibition of MEK/Erk signalling in the developing mouse embryo abrogates the pro-differentiation activities that are promoted during normal development (Nichols et al. 2009a). Gsk3 inhibition enhances the expansion of the mES cultures, but the mechanism for this remains unclear, although it is known to lead to the activation of the Wnt/β-catenin pathway (Sato et al. 2004; Ying et al. 2008). It has been suggested that mES cells maintained in serum-free conditions with 2i might represent the 'ground-state' of pluripotency that exists transiently in the epiblast of pre-implantation blastocysts (Ying et al. 2008; Nichols et al. 2009a; Nichols et al. 2009b; Nichols et al. 2011).

Human ES cells are not responsive to LIF signalling and media containing serum and LIF fails to support undifferentiated hES cultures (Thomson et al. 1998; Daheron et al. 2004). In addition to this, BMPs induce hES cultures to differentiate towards TE (Xu et al. 2002; Gerami-Naini et al. 2004; Pera et al. 2004a). In contrast to mES cells, fibroblast growth factor (FGF) and TGF-β/Activin signalling are central to the maintenance of hES cell undifferentiated phenotype (Yu et al. 2008). Initial experiments indicated that supplementation by FGF2 was necessary for extended culture and maintenance of undifferentiated hES colonies on feeder cells (Amit et al. 2000). Further research revealed that used at an increased concentration FGF2 allows for feeder-independent culture of undifferentiated hES cells in serum replacement medium (Wang et al. 2005; Xu et al. 2005; Levenstein et al. 2006). The mechanisms, by which FGF2 action is conferred in hES cultures, are not yet fully understood (Pera et al. 2010). Nevertheless, several modes of action have

been proposed for this growth factor. FGF2 was, for example, shown to suppress BMP activity present in serum and serum replacement medium that can induce differentiation in hES cells (Xu et al. 2005). FGF2 has also been reported to promote the expression of members of the TGF-β family in hES cells, and also feeder cells, which can in turn enhance self-renewal (Greber et al. 2007). To promote hES self-renewal, FGF2 signalling has also been shown to co-operate with insulin-like growth factor (IGF) produced in a paracrine manner by fibroblast-like cells within the hES cell cultures (Bendall et al. 2007). IGF signalling cannot, however, substitute FGF2 signalling to support self-renewal in hES cells, and it has been shown that withdrawal of FGF2 promotes apoptosis in addition to differentiation in hES cultures (Wang et al. 2009a).

Both TGF-β and Activin 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. 2005a; Xiao et al. 2006). Signalling from the TGF-β/Activin pathway acts via phosphorylated SMAD2/3 (Schmierer et al. 2007). Inhibiting TGF-β/Activin signalling in hES cells with a chemical inhibitor SB431542 leads to a reduction in the levels of phosphorylated SMAD2/3 and induces differentiation (James et al. 2005; Vallier et al. 2005a; Xiao et al. 2006). Recent evidence has shown that SMAD2/3 directly binds the *NANOG* proximal promoter in hES cells and regulates the levels of its expression (Xu et al. 2008; Vallier et al. 2009a). Other signalling pathways have been implicated in the regulation of hES self-renewal, however remain less well defined. Previously it has been reported that Wnt/β-catenin signalling would be beneficial for hES self-renewal (Sato et al. 2004), however recent data from experiments inhibiting the Wnt/β-catenin pathway in hES showed no effect on self-renewal (Davidson et al. 2012). The PI3K/Akt pathway has recently also been implicated in maintenance of hES pluripotency. PI3K/Akt signalling has been shown to modulate the activity of TGF-β/Activin signalling in promotion of self-renewal in crosstalk with MEK/Erk and Wnt/β-catenin signalling (Chen et al. 2012; Singh et al. 2012). Hence, coordinated yet complicated interactions between signalling networks are likely to govern the balance between self-renewal and differentiation in hES cells. Defining the exact mechanisms regulating these decisions will be of great importance for possible future applications in regenerative medicine.

Mouse EpiS cells derived from post-implantation embryos are maintained self-renewing in a pluripotent state by Fgf2 and Activin A supplementation (Brons et al. 2007; Tesar et al.

2007); the key signalling pathways also implicated in maintenance of hES pluripotent state. TGF-β/Activin signalling promotes the expression of *Nanog* via the SMAD2/3 pathway, which, in turn, maintains a self-renewing mEpiS state (Greber et al. 2010), in a similar manner to hES cells. Moreover, blocking TGF-β/Activin signalling induces differentiation towards the neural lineage in mEpiS cells and in hES cells (Vallier et al. 2009a; Greber et al. 2010). Interestingly, Fgf2 signalling in mEpiS cells has a role in maintaining the EpiS-identity, as blocking Fgf2 signalling in mEpiS cells leads to the spontaneous generation of mES-like colonies (Greber et al. 2010). The effect of Fgf2 signalling maintaining a mEpiS-identity can be overcome, however, by constitutive activation of JAK-STAT3 pathway, which allows for reversion of mEpiS cells to mES cell like naïve pluripotent identity even in the presence of Fgf2 and Activin A supplementation (van Oosten et al. 2012).

1.2.4 Transcriptional networks that control pluripotency

In both mouse and human ES cells the key transcription factors *Oct4, Sox2* and *Nanog* are at the centre of a regulatory network crucial for the maintenance of pluripotency (Boyer et al. 2005; Loh et al. 2006; Kim et al. 2008; Mathur et al. 2008; Kunarso et al. 2010). The same network has also been shown to be critical for the development of the transiently pluripotent cell population *in vivo*. Data from genome-wide studies demonstrated that Oct4, Sox2 and Nanog bind together at their own promoters forming interconnected auto-regulatory loops that maintain the expression of these key factors in ES cells. Furthermore the three factors often co-occupy common target genes including important self-renewal regulators that are strongly expressed in ES cells as well as somatic lineage-specific genes that are inactive but poised for future activation in ES cells (Jaenisch et al. 2008; Young 2011). These findings together suggest that *Oct4*, *Sox2* and *Nanog* are so-called 'master regulators' of the pluripotent stem cell identity (Spivakov et al. 2007). However, it is important to note that the pluripotency network is very complex and involves an extending number of new effector proteins, such as nuclear receptor Esrrb and co-activator Ncoa3 (Zhang et al. 2008; Percharde et al. 2012).

1.2.5 Epigenetic regulation of pluripotency

Recent studies have indicated that ES cell chromatin exists in an 'open' state, which is transcriptionally 'permissive' and is proposed to allow the transcriptional programme to switch rapidly upon differentiation (Meshorer et al. 2006; Efroni et al. 2008; Gaspar-Maia et al. 2009; Gaspar-Maia et al. 2011). Interestingly, genome-wide analysis of the repressive H3K9me2 mark indicates that differentiated cells harbour large (up to 4.9 Mb) regions of K9

methylation covering between 10% and 46% of the genome depending on the specific tissue, in contrast, pluripotent ES cells were shown to only have H3K9me2 enriched only through 4% of the genome (Wen et al. 2009). Analysis of tissue-specific genes in pluripotent ES cells has revealed that many of them are not actively transcribed or repressed, but poised for transcription. Promoters of such poised genes carry both repressive (H3K27me3) and active (H3K4me3) histone modifications, forming so called bivalent chromatin domains (Azuara et al. 2006; Bernstein et al. 2006). These domains are thought to prime tissue-specific genes for future activation (or repression) upon ES cell differentiation, yet preclude their premature expression through PRC2-mediated repression (Giadrossi et al. 2007; Spivakov et al. 2007). Consistent with gene priming, bivalently marked promoters have also been shown to assemble RNA polymerase II complexes, preferentially harbouring a poised conformation that is enforced by the presence of PRC1 complexes (Stock et al. 2007; Brookes et al. 2012). 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* (Jørgensen et al. 2006; Stock et al. 2007), many developmental PcG target genes become derepressed, demonstrating the importance of polycomb-mediated gene silencing in ES cells. It has also been shown that bivalent domains generally become resolved at lineage-specific genes upon differentiation of ES cells, with lineage-specific genes retaining the active H3K4me3 mark and losing the repressive H3K27me3 in their cognate lineage and vice versa in other lineages (Bernstein et al. 2006; Mikkelsen et al. 2007).

1.2.6 The *in vivo* equivalence of mouse and human ES cells – a developmental paradox

Both mouse and human ES cells can be derived from ICM explants of pre-implantation embryos, and both cell populations are pluripotent and under the regulation of the key transcription factors Oct4, Sox2 and Nanog. However, clear differences exist between these two types of pluripotent cells (Rossant 2011). As mentioned before, mES and hES cells grow in morphologically different colonies, with mES cells growing as three-dimensional, dome-shaped colonies whereas hES cells grow in colonies as thin and flat layers. The two cell types also differ in their expression of cell-surface markers, with SSEA1 and SSEA3/4 being specifically expressed in undifferentiated mES and hES cells, respectively (Pera et al. 2004a; Pera et al. 2010). Further evidence of the divergent nature of mouse and human pluripotent ES cells comes from the low overlap of target genes regulated by the Oct4/Sox2/Nanog transcriptional network (Boyer et al. 2005; Loh et al. 2006), which suggests that the pluripotent states differ between mES and hES cells. It has to be noted, however, that these

studies were performed with different technology platforms, which is likely to explain at least some of the lack of overlap (Rossant 2008).

One interesting difference between mES and hES cells is that hES cells are able to give rise to extra-embryonic tissues (Itskovitz-Eldor et al. 2000; Reubinoff et al. 2000; Xu et al. 2002) whereas this has not been reported for mES cells. One possible explanation for this discrepancy could be that hES cells are in a pre-epiblast stage of commitment, where all lineage options including TE are still open, whereas mES cells might 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 (See Fig. 1.4) (Pera et al. 2004a; Niakan et al. 2012). However, the recent derivation of mEpiS cells, a pluripotent stem cell population derived from post-implantation mouse embryos (Brons et al. 2007; Tesar et al. 2007), which have properties closely resembling hES cells, contradicts such interpretation. Both hES and mEpiS cells are maintained in a pluripotent state with FGF and TGF-β/Activin signalling, in contrast to the LIF/BMP signalling required for mES pluripotency (Brons et al. 2007; Tesar et al. 2007) and interestingly recent genome-wide DNA replication timing studies have indicated that hES replication timing profiles align more closely to those of mEpiS than of mES cells (Ryba et al. 2010). However, there are many crucial differences between hES and mEpiS cells (reviewed in Pera and Tam, 2010). For example, hES cells express genes such as *DPPA3, KLF4* and *REX1*, which are not expressed in mEpiS cells and conversely hES cells lack expression of *FGF5,* a marker of mEpiS cells (Pera et al. 2010).

As already mentioned before mES cell cultures are heterogeneous and contain cells in different interconvertible sub-states. Mouse ES cells expressing high levels of *Nanog* were found to have higher self-renewal capabilities as compared to cells with low *Nanog* expression, which are prone to differentiate (Chambers et al. 2007). In similar analysis, mES cells with high expression of *Stella* were found correlate with an ICM-like phenotype in contrast to *Stella*-negative cells, which were more like the post-implantation epiblast derived mEpiS cells (Hayashi et al. 2008). Recently it has been shown that mEpiS cells can also be isolated from pre-implantation mouse embryos alongside mES cells when derived in specific culture conditions in the absence of any growth factor additives (Najm et al. 2011). In addition to this, early and late pre-gastrulation epiblast like stages could be delineated within mEpiS cell cultures, using an *Oct4*-green fluorescent protein (GFP) reporter, forming different subpopulations (Han et al. 2010). Notably, the *Oct4*-GFP positive early epiblast-like cells were able to contribute to chimaeras upon injection to blastocyst in contrast to the late

epiblast-like cells. These data strongly suggest that the pluripotent state in mouse cells exists in a continuum of phenotypically and functionally distinct states.

One hypothesis to explain the similarity between mEpiS and hES cells could be that during the process of derivation from the ICM explant hES cells, would continue to develop and the cells isolated preferentially adopt a primed mEpiS-like pluripotent state in culture, rather than ground-state mES-like pluripotency (Nichols et al. 2011). Recent studies tracking the derivation process of hES cells, have identified a post-ICM intermediate (PICMI), a transient morphologically distinct epiblast-like structure, which is *OCT4* and *NANOG* positive and is both necessary and sufficient for hES derivation (O'Leary et al. 2012). This study, however, did not find a correlation with the PICMI and mES-like naïve pluripotent state. Human ES cells with properties similar to mES cells have been derived via the ectopic induction of *OCT4, KLF2* and *KLF4* (Hanna et al. 2010a), however these cells are metastable and remain dependent on transgene expression. The true embryological equivalence of hES cells to mES or mEpiS cells pluripotent state remains largely unresolved and further research is needed to determine whether similar continuum of pluripotency exists in humans as has been identified in mouse.

1.2.7 Reprogramming to pluripotency

The gene expression patterns that give each cell type its identity are governed by epigenetic mechanisms, which during development allow for the stable inheritance of a differentiated phenotype in lineage committed cells. Under certain conditions, differentiated cells can, however, revert into a less differentiated state, a process termed 'nuclear reprogramming' (Hochedlinger et al. 2009). Reprogramming cells to pluripotency offers unprecedented potential for disease research, drug screening, toxicology and regenerative medicine via the possibility of generating disease specific cell models and possible patient specific cells for grafting (Yamanaka 2009). Different experimental strategies for nuclear reprogramming have been created including somatic cell nuclear transfer (SCNT) (Gurdon 1962), fusion of different cell types (Miller et al. 1976) and induction by specific factors (Takahashi et al. 2006).

Initial experiments on SCNT were developed to assess whether the nuclei of differentiated cells retain all the genetic information required for the development of a whole organism (Hochedlinger et al. 2009). These were also the first experiments of nuclear reprogramming. The transfer of a differentiated cell nucleus into an enucleated oocyte was able to reprogram the differentiated nucleus to the state of a zygote, indicating that the oocyte

contains factors able to induce reprogramming. Initial experiments were conducted in amphibians (Gurdon 1962), and only much later this was also repeated in mammals (Wilmut et al. 1997; Hochedlinger et al. 2002). Only very recently SCNT has been achieved in non-human primates (Byrne et al. 2007), indicating the challenging nature of this technique. Similar successful SCNT has yet to be reported for human cells, however it was recently reported that injection of a somatic genome into a human oocyte that was not enucleated allows for embryonic development up to the blastocyst stage through the formation of a triploid embryo (Noggle et al. 2011). Due to the very low efficiency and numerous technical challenges as well as the ethical questions with regards to the use of human oocytes and development of cloned embryos, SCNT is unlikely to be practical for widespread clinical use.

Nuclear reprogramming can also occur upon cell fusion, as first achieved using mouse EC cells and thymocytes (Miller et al. 1976). 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. If heterokaryons are maintained in culture for several days, hybrids will form upon cell division and nuclear fusion (Kikyo et al. 2000). In most cell hybrids, the phenotype of the less differentiated cell is dominant over the phenotype of the more differentiated fusion partner. This was evident in the first fusions between mouse EC cells and thymocytes and further confirmed with fusions between mouse myoblasts and human lymphocytes, where the myoblast phenotype was dominant, and with fusions between ES cells and various somatic cell types with both mouse and human ES cells (Miller et al. 1976; Tada et al. 2001; Ying et al. 2002; Cowan et al. 2005; Terranova et al. 2005; Yu et al. 2006). The reason why the less differentiated phenotype dominates in heterokaryons remains largely unknown. Cell fusion experiments and heterokaryon formation, however, provide a powerful *in vitro* system to further study the contribution of specific factors to the reprogramming process (Piccolo et al. 2011). Nevertheless, this approach is also unlikely to be of use for customised cell therapy due to low efficiency and the tetraploidy of the reprogrammed cells (Jaenisch et al. 2008).

Ground-breaking experiments in 2006 demonstrated that mouse fibroblasts can be reprogrammed to become iPS cells by viral transfection of four TFs, *Oct4*, *Sox2*, *Klf4* and *c-Myc* (Takahashi et al. 2006). This was then repeated with human fibroblasts using the same four factors (Takahashi et al. 2007), as well as with a different set of factors that include *OCT4*, *SOX2*, *NANOG* and *LIN-28* (Yu et al. 2007). Induced pluripotent stem cells have since been derived from many different cell types including keratinocytes (Aasen et al. 2008), peripheral blood cells (Loh et al. 2009) and pancreatic islet beta cells (Bar-Nur et al. 2011).

Induced pluripotent stem cells offer an opportunity to study many diseases in cell models and hence disease-specific iPS cells have been generated for many disorders (Park et al. 2008a). Numerous improvements in the delivery methods of the reprogramming factors, and identification of additional factors or small molecules that improve reprogramming efficiency have been made to the technology as reviewed in (González et al. 2011). Interestingly, it has been shown, for example, that reprogramming can also be induced by delivery of recombinant proteins (Zhou et al. 2009).

Mouse iPS cells are maintained by LIF/BMP signalling (Takahashi et al. 2006) and by contrast human iPS cells by FGF2/Activin signalling (Park et al. 2008b). Interestingly, derivation of induced EpiS cells was recently reported by introducing *Oct4*, *Sox2*, *Klf4* and *c-Myc* and derivation under mEpiS culture conditions (Han et al. 2011) This suggests that distinct pluripotent states can be reached by reprogramming as found in 'native' developmental context. The equivalence of ES cells and iPS cells remains, nevertheless, a question of debate. Induced pluripotent cells are very similar to ES cells (Hanna et al. 2010b), and have been conclusively shown to be pluripotent via the generation of chimaeras and germline transmission, as well as generation of adult mice via tetraploid complementation (Maherali et al. 2007; Okita et al. 2007; Wernig et al. 2007; Boland et al. 2009). However, the somatic cell type used in reprogramming protocol influences the molecular and functional properties of the generated iPS cells with somatic specific DNA methylation and histone modification patterns being retained in iPS cells, referred to as 'epigenetic memory' (Polo et al. 2010; Ohi et al. 2011). Recent reports have highlighted the existence of somatic coding mutations, copy number variation and hotspots of aberrant epigenomic reprogramming in human iPS cells (Gore et al. 2011; Hussein et al. 2011; Lister et al. 2011), however whole genome sequencing of mouse and human iPS clones and their parental cells suggests that background mutations in parental cells account for most of the genetic heterogeneity found in iPS cells (Abyzov et al. 2012; Young et al. 2012). The differences between iPS and ES cells, as well as those among iPS cells, might hence have consequences for the use of these cells in research, disease modelling and therapeutics (Robinton et al. 2012), and this needs to be taken into consideration when research is taken forwards.

1.3 Aims of this study

A growing panel of different mouse and human pluripotent stem cell lines has been derived from the developing early embryo as well as via somatic cell reprogramming. The exact relationship between each other of these lines and whether they correspond to different

pluripotent states with distinct developmental potencies and affiliations *in vivo* remain, however, unclear. The overall aim of this project is to assess the developmental status of embryo-derived pluripotent cell lines using a chromatin-based approach. Here a DNA replication timing assay is used as a tractable approach to investigate whether and how variations in a cell's chromatin status can reliably define different pluripotent cell identities.

For this purpose, a comparison of the replication timing status of a large panel of key developmental genes in several hES and hiPS cell lines alongside mES and mEpiS cells is initially conducted to determine whether distinct pluripotent states can indeed be delineated at this level. The impact of different growth culture conditions on modulating hES cell's epigenetic and pluripotent identity will then be investigated and the underlying molecular and epigenetic mechanisms explored further. In particular, the role of TGF-β/Activin signalling axis in favouring a self-renewing and epigenetically stable state as well as the role of the pluripotent marker REX1 will be subsequently assessed in this thesis.

This study might advance current understanding of early development and could also be beneficial for developing stem cell-based therapies for regenerative medicine, as responses to directed differentiation are likely to vary depending on the initial position of cells in the developmental hierarchy.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Primary antibodies

β-actin Rabbit polyclonal anti-β-actin antibody (ab8227; Abcam Ltd., Cambridge, UK). Used for western blot (WB) at 1:5000 dilution.

BrdU Mouse monoclonal anti-BrdU antibody (347580; Becton Dickinson Biosciences, San Jose, CA). Used for replication timing at 80 µl/reaction dilution and for analysis of DNA content at 5 µg/ml.

H3 Rabbit polyclonal anti-histone 3 carboxy terminal (ab1791; Abcam). Used for WB at 1:2500 and chromatin immunoprecipitation (ChIP) at 5 µg/immunoprecipitation (IP) dilution.

H3Ac Rabbit polyclonal anti-histone 3 acetylation (17-615; Upstate Biotechnology /Millipore, 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.

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.

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.

Methyl-C Mouse monoclonal anti-5-methylcytidine (33D3; Eurogenetec, Belgium). Used for methylated DNA immunoprecipitation (MeDIP) at 10 μl/reaction dilution.

Musashi Rabbit polyclonal anti-Musashi (AB5977; Chemicon/Millipore). Used for immunofluorescence (IF) at 1:200 dilution.

Nanog Rabbit polyclonal anti-human Nanog (ab21624; Abcam). Used for IF at 1:500 and WB at 1:1000 dilution. Rabbit polyclonal anti-mouse Nanog (REC-RCAB0002P-F; Cosmo Bio Co., Ltd., Japan) used for IF at 1:500, fluorescence-activated cell sorting (FACS) at 1 μg/1x10^6 cells and WB at 1:1000 dilution.

Nestin Mouse monoclonal anti-Nestin (MAB5326; Chemicon/Millipore). Used for IF at 1:500 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.

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.

Pax6 Rabbit polyclonal anti-Pax6 (AB5409; Chemicon/Millipore). Used for IF at 1:1000 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.

Sirt1 Rabbit monoclonal anti-Sirt1 (ab32441; Abcam). Used for WB at 1:1000 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.

Sox2 Rabbit polyclonal anti-Sox2 (ab15830; Abcam). Used for IF at 1:1000 and WB at 1:1500 dilution.

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-Tumour rejection antigen (TRA)-1-60 (MAB4360; Chemicon/Millipore, 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.

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

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).

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).

Dc7 Mouse ES cell line newly derived; feeder-independent; 129/*castaneus* mixed genetic background; a gift from Dr Tatyana Nesterova (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.

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).

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).

H9 *GFP* OE Female human ES cell line (Thomson et al. 1998) with constitutive expression of *GFP* (Vallier et al. 2009a), a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).

H9 *NANOG* KD Female human ES cell line (Thomson et al. 1998) with constitutive knockdown of *NANOG* (Vallier et al. 2009a), a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).

H9 *NANOG* OE Female human ES cell line (Thomson et al. 1998) with constitutive expression of *NANOG* (Vallier et al. 2009a), a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).

H9 shSCR Female human ES cell line (Thomson et al. 1998) expressing non-targeting scramble shRNA (Vallier et al. 2009a; Brown et al. 2011), a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, UK).

H9 *SMAD2* KD Female human ES cell line (Thomson et al. 1998) with stable knockdown of *SMAD2* (Brown et al. 2011), a gift from Dr Ludovic Vallier (Department of Surgery and Laboratory for Regenerative Medicine, University of Cambridge, 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).

NP7 Human neural progenitor cells derived from H7 ES cells (Gerrard et al. 2005a).

MEFs Mouse embryonic fibroblasts isolated from wild type CF-1 embryos (E13.5).

ReNcell VM Human immortalized neural stem cell line derived from a 10-week foetal brain (SCC008; Chemicon/Millipore).

2.2 Methods

2.2.1 Cell culture

All tissue culture reagents used were from LifeTechnologies (Paisley, UK) and all plastics from Corning Inc. (NY, USA) unless otherwise stated. H1, H7 and H9 hES cell lines were cultured under feeder-free conditions on matrigel-coated 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 mouse embryonic fibroblast (MEF) cells-conditioned medium (CM) [Knockout™ Dulbecco's Modified Eagle Medium (DMEM) medium supplemented with 20% (v/v) Knockout™ Serum Replacement (KSR), 1 mM L-glutamine, non-essential amino acids (NEAA) and 0.1 mM β-mercaptoethanol] supplemented with 8 ng/ml basic fibroblast growth factor (FGF2) (R&D Systems Europe Ltd, Abingdon, UK). CM was prepared by incubating hES cells medium on flasks (0.5 ml/cm2) with 40 grays irradiated MEFs at 56 000 cells/cm2 during 24 hours. The CM was filtered (0.2 μm) and stored at -20°C. MEFs were cultured in high-glucose DMEM supplemented with 10% foetal bovine serum (FBS; Globepharm, Esher, UK) and 2 mM L-glutamine, grown until 80-85% confluence, dissociated with trypsin ethylene-dinitrilo tetraacetic acid (EDTA) (0.05%/0.02%) (PAA Laboratories, Yeovil, UK) before irradiation and plated onto 0.1% gelatine-coated surfaces (Sigma-Aldrich, Gillingham, UK).

For feeder and Serum Replacer free culture, H1 and H9 hES cell lines were transferred from CM conditions to a chemically defined medium (CDM) (Johansson et al. 1995; Wiles et al. 1999) supplemented with 10 ng/ml Activin A (R&D Systems Europe Ltd) and 12 ng/ml FGF2 as previously described (Vallier et al. 2005a) ($CDM^{activin}$). The composition of CDM was 50% Iscove's Modified Dulbecco's Medium plus 50% F12 NUT-MIX, supplemented with 7 μg/ml of insulin (Roche Ltd, West Sussex, UK), 15 μg/ml of transferrin (Roche Ltd), 450 μM of monothioglycerol (Sigma-Aldrich), 1% chemically defined lipid concentrate 100x and 5 mg/ml of batch tested bovine serum albumin (BSA) fraction V (PAA Laboratories). To allow hES cells adhesion in CDM, plates were pre-coated with 0.1% gelatine followed by a coating with adhesion media (Knockout™ 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 μM) of

curcumin (a gift from Dr Matthew J Fuchter, Department of Chemistry, Imperial College London, UK) or c646 (R&D Systems Europe Ltd) dissolved in dimethyl sulfoxide (DMSO) (Sigma-Aldrich) was added to CM or CDM media of H1 hES cells when indicated. Medium was changed daily and cells routinely passaged at a 1:3 split ratio by treatment with 200 units (U)/ml collagenase type IV in Knockout™ DMEM followed by mechanical dissociation with 5 ml pipette in both CM and CDM conditions. Nomenclature for culture conditions is according to the following convention: cell line, number of passages in CM culture, plus additional passages in CDM 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 and passages 47-51 were again in CM. Since hES cells are susceptible to genetic anomalies on prolonged passages (Draper et al. 2004), where possible only hES cells earlier passages (p30-p80) were used for these experiments.

Human iPS 40 cells (Vallier et al. 2009b) were grown under CM or CDM conditions as described for hES cells.

Hela cells were cultured in high glucose DMEM supplemented with 10% FBS and passaged using trypsin EDTA.

Mouse ES cells were grown and maintained in an undifferentiated state either on 0.1% gelatin-coated surfaces (E14tg2A) 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 analysed in derivation medium with DMEM supplemented with 5% batch tested FBS (Sigma), 15% KSR, 1x NEAA, 50 µM β-mercaptoethanol, 2 mM L-glutamine and 2000 U/ml LIF.

Mouse EpiS cell lines (129, B6/CBA and NOD) were cultured as previously described (Brons et al. 2007) in CDM media supplemented with 20 ng/ml Activin A and 12 ng/ml FGF2 on pre-coated adhesion medium plates in the same way that hES cells in CDM conditions. Cells were propagated as small clumps by scraping with 5 ml pipette after incubation with 1% collagenase type IV in Knockout™ DMEM using 1:5 to 1:15 split ratios depending on proliferation ratio. Unless otherwise stated all cells were maintained at 37°C and 5% CO2.

2.2.2 Cloning and delivery of DNA into hES cells

2.2.2.1 Engineering of protein expression constructs

The *REX1* open reading frame (ORF) was amplified from hES complementary DNA (cDNA) (prepared according to section 2.2.5.1.) by reverse transcription polymerase chain reaction (RT-PCR) using linker primers 5'- ATTCCCTCGAGATGAGCCAGCAACTG -3' and 5'-CGCGGTTAATTAACTACTTTCCCTCTTG -3', which introduced flanking restrictions sites for XhoI and PacI (underlined) respectively) using *PFUTurbo* (Agilent Technologies, La Jolla, CA). Briefly the RT-PCR reaction contained 2 µl cDNA, 1 µl forward primer, 1 µl reverse primer, 1 µl 10nM dNTP, 5 µl 10x buffer, 1 µl *PFUTurbo* and dH_2O up to 50 µl. The reaction was heated to 95°C for 4 min, then kept at 95°C for 1 min, 30 sec at 60°C, 1 min 15 sec at 72°C then cycled back to step two for 40 times, then incubated at 72°C for 10 min. The amplified *REX1* with the restriction sites (Fig. 2.1a) was ligated into pGEM-T Easy vector (Promega Corporation, Madison, WI) and sequence was verified

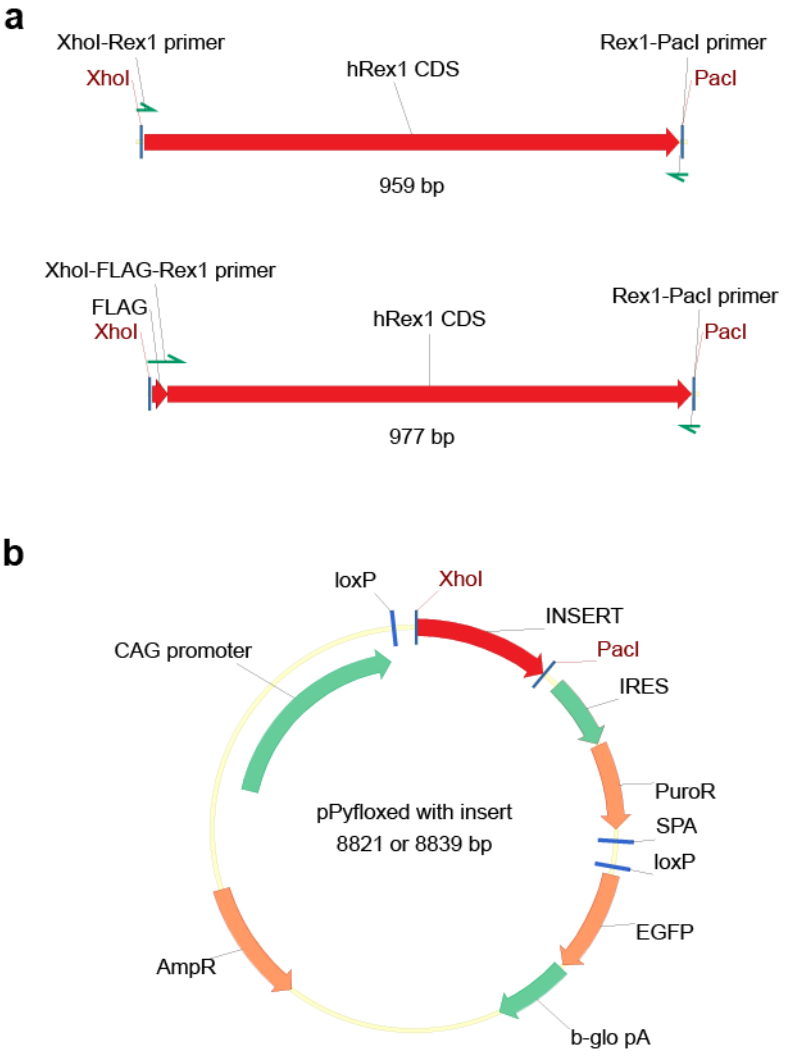


Figure 2.1 Transgenes and expression vector used in this study
(**a**) Schematic representation of human *REX1* coding sequence cloned for this study. Indicated are the restriction sites and primer pairs used for PCR amplification as well as FLAG-tag added in one construct. (**b**) Vector map of pPyfloxed expression vector with insert. Expression of the transgene is driven by the CAG promoter. Indicated are loxP sites, which allow the creation of revertant clones by Cre-Lox excision, restriction sites, internal ribosomal entry site (IRES), puromycin resistance gene (puroR), synthetic polyA site, enhanced green fluorescent protein (EGPP) coding sequence, which will be active in revertant clones, beta-globin polyA site (b-glo pA) and ampicillin resistance gene (AmpR).

(MRC Clinical Sciences Centre sequencing facility). To generate a eukaryotic expression construct the *REX1* was then subcloned into pPyfloxed expression vector (a kind gift Prof Ian Chambers) using XhoI–PacI sites (Fig. 2.1b). Details of cloning procedures are described in section 2.2.2.3. In the pPyfloxed vector expression was driven by the CAG-promoter and selection enabled by the puromycin resistance gene.

The FLAG-Rex1 construct was RT-PCR amplified from the pPyfloxed-*REX1* plasmid using linker primers 5'-CTCGAG**ATGGACTACAAGGACGACGATGACAAG**AGCCAGCAACTGAAGAAA CGGGCA-3' and 5'- CGCGGTTAATTAACTACTTTCCCTCTTG -3', which introduced flanking restrictions sites for XhoI and PacI (underlined) respectively and the FLAG sequence (bold) using *Phusion* DNA polymerase (New England Biolabs, Ipswich, MA). Briefly the RT-PCR reaction contained 2 μl cDNA, 2.5 μl forward primer, 2.5 μl reverse primer, 1 μl 10nM dNTP, 10 μl 5x GC buffer, 0.5 μl *Phusion*, 1.5 μl DMSO and dH_2O up to 50 μl. The reaction was heated to 98°C for 30 sec, then kept at 98°C for 10 sec, 30 sec at 68°C, 30 sec at 72°C then cycled back to step two for 35 times, then incubated at 72°C for 10 min. The amplified FLAG-*REX1* was ligated into pGEM-T Easy vector and subcloned into pPyfloxed expression vector as described above for *REX1* (Fig. 2.1a and b).

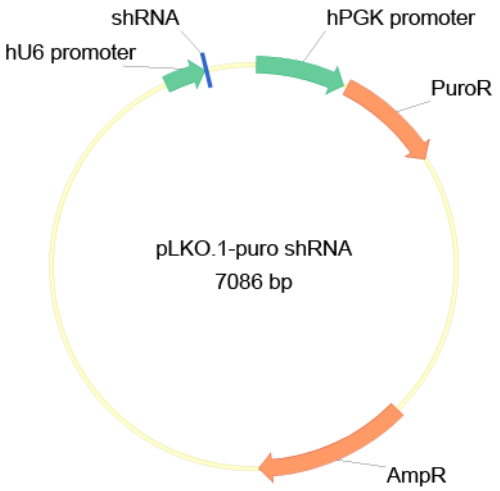


Figure 2.2 Vector used for stable expression of shRNA
Vector map of pLKO.1-puro shRNA expression vector. Expression of the shRNA is driven by hU6 promoter. Indicated are human phosphoglycerate kinase eukaryotic promoter (hPGK), puromycin resistance gene (puroR) and ampicillin resistance gene (AmpR).

2.2.2.2 Preparation and selection of shRNA constructs

Five individual MISSION® short hairpin loop RNA (shRNA) constructs targeting human *P300* and a non-targeting control shRNA were ordered from Sigma-Aldrich in bacterial glycerol stocks. The bacteria were picked according to manufacturer's instructions and plated onto Lysogeny Broth (LB) agar [LB-Broth Lennox 20 g/l (BD) and L-agar 15 g/l (Sigma) in dH_2O] with 100 μg/μl ampicillin and incubated overnight. Individual colonies were picked into low salt LB [LB-Broth-Lennox 10 g/l, Peptone 5 g/l and Yeast Extract 2.5 g/l (all from BD) in dH_2O] supplemented with 100 μg/μl ampicillin and incubated overnight at 37°C under agitation. Plasmid DNA was extracted from exponentially growing cultures with Purelink Plasmid Mini and Midiprep kits (LifeTechnologies) according to manufacturer's instructions. Final DNA

product was resuspended in double distilled H_2O. Figure 2.2 shows a schematic representation of the pLKO.1-puro shRNA vector used for the expression of the interfering RNAs.

All five P300 shRNAs were tested by transforming Hela cells with each individual shRNA and comparing them to the non-targeting shRNA, using Lipofectamine 2000 (LifeTechnologies), according to manufacturer's instructions. Analysis of knockdown of P300 expression was done with evaluating gene expression levels by real-time quantitative PCR (qRT-PCR) and protein expression levels by western blotting, samples were collected 48 hours after transfection. Results of these analyses are summarised in Figure 2.3. Knockdown of *P300* was most efficient with shRNAs 3 and 4 with approximately 70% reduction of relative gene expression and nearly complete knockdown of protein expression as compared to the non-targeting scramble shRNA (scr) transformed sample and these were selected. ShRNA 3 5'-CCGGCAATTCCGAGACATCTTGAGACTCGAGTCTCAAGATGTCTCGGAATTGTTTTTG -3' and shRNA 4 5'- CCGGTACACTAGAGACACCTTGTATCTCGAGATACAAGGTGTCTCTAGTGTA TTTTTG -3' and the scr 5'-CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT -3' were used in the experiments with hES cells. Expression of the shRNA was driven by hU6 promoter and selection of clones was enabled by puromycin resistance gene (Fig. 2.2).

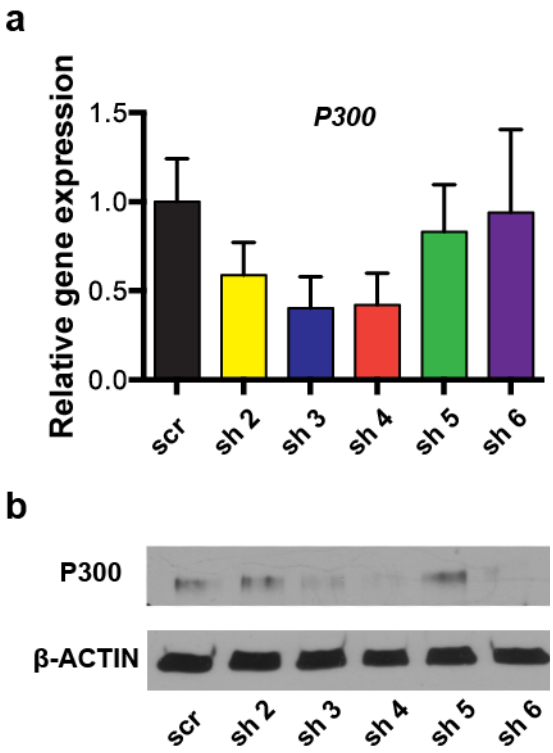


Figure 2.3 Selection of effective *P300* shRNAs.
P300 specific shRNAs (sh2, sh3, sh4, sh5 and sh6) were transfected in HELA cells alongside non-targeting scramble shRNA. Samples were analysed for the efficiency of *P300* knockdown by (**a**) qRT-PCR analysis and (**b**) WB analysis.

2.2.2.3 Plasmid construction and bacterial transformation

All reagents used were from New England Biolabs, unless otherwise stated. PCR-amplified fragments were run on a 1% agarose (Sigma-Aldrich) gel along with a 2-log ladder. Correct sized fragments were isolated using QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions. Isolated DNA fragments were TA-tailed by incubating different amounts of DNA with 1µl of Taq-polymerase (LifeTechnologies), 1µl Taq PCR –buffer

(LifeTechnologies), 0.3 μl of 50 mM $MgCl_2$ and 0.2 μl of 10 mM dNTP (LifeTechnologies) in 10 μl reaction volume with dH_2O for 20 min at 70°C. After TA-tailing various amounts of the DNA fragments were ligated with pGEM-T Easy vector in a 20 μl reaction with 10 μl of 2x rapid ligation buffer (Promega) and T4 DNA ligase (Promega). Ligations were incubated overnight at 4°C.

The ligation products were used to transform RapidTrans™ TAM1 competent *E.coli* (Active Motif) according to manufacturer's instructions. Briefly, 5 μl of plasmid product was added to 50 μl of bacteria and mixed gently. Transformations were incubated on ice for 30 min, then heat shocked by immersing in 42°C water for 30 s and then immediately returned to ice for 2 min. 250 μl of SOC-medium was added to the reaction and the bacteria were incubated shaking at 37°C for one hour. 40 μl of X-Gal (to enable blue/white screening of colonies by β-galatosidase activity) was added to each reaction and an aliquot of the transformed bacteria was plated onto LB-agar plates supplemented with 100 μg/μl ampicillin. Plates were incubated overnight at 37°C and the next day individual white colonies were picked into LB-broth (LB-Broth Lennox 20 g/l in dH_2O) supplemented with 100 μg/μl ampicillin and incubated overnight at 37°C under agitation. Plasmid DNA was extracted from exponentially growing cultures with Purelink Plasmid Mini and Midiprep kits (LifeTechnologies) according to manufacturer's instructions. Final DNA product was resuspended in double distilled H_2O and insertion of correct DNA fragment was confirmed by sequencing.

4 μg of insert (*REX1* or FLAG-*REX1*) and 4 μg of vector (pPyfloxed) were digested in a 20 μl reaction with 1 μl of *PacI*, 1μl of *XhoI* and 2 μl of 10x buffer, incubated at 37°C overnight. Digestion products were ran on a 1% agarose gel with the 2-log ladder and bands of correct sizes were isolated with QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions. Digested pPyfloxed vector was dephosphorylated with Antarctic phosphatase in a 40 μl reaction with 4 μl of 4x buffer and 1 μl of phosphatase incubated at 37°C for 30 min. 50 ng of vector and three-molar excess of insert were ligated in a 20 μl reaction with 2 μl of 10x buffer and 1 μl of T4 ligase (400 U) incubated overnight at 16°C.

The ligation products were used to transform RapidTrans™ TAM1 competent *E.coli* (Active Motif) according to manufacturer's instructions and cultures were expanded as described above for pGEM-T Easy ligations, with the exception of no X-Gal added and blue/white screening done. Plasmid DNA was extracted from exponentially growing cultures with Purelink Plasmid Mini and Midiprep kits (LifeTechnologies) according to manufacturer's

instructions. Final DNA product was resuspended in double distilled H_2O and insertion of correct DNA fragment was confirmed by sequencing.

2.2.2.4 DNA delivery into hES cells

H1 and H9 hES cells were transfected with the appropriate overexpression on shRNA constructs using Lipofectamine 2000 according to manufacturer's instructions with some adjustments. Briefly, hES cells were seeded in high density of small colonies onto six wells of a six well plate. Cells were maintained in CDM media without antibiotics for two days after which they were transfected. 4 μg of plasmid DNA was diluted in 250 μl of OPTIMEM media (LifeTechnologies) and 10 μl of Lipofectamine was diluted in 250 μl of OPTIMEM, per each well transfected. Lipofectamine mix was incubated at room temperature (RT) for 5 min before mixing carefully with the plasmid DNA mix and further incubated for 20 min at RT. Medium on cells to be transfected was changed to 1 ml of OPTIMEM without any supplements. After the incubation Lipofectamine/DNA mix was added at a dropwise manner to the wells. The plates were then briefly shaken and placed back into the cell incubator. The following day media was changed back to $CDM^{activin}$ and then maintained in the media until approximately day seven after transfection. Once the cells had reached confluence at around day seven, selection was started by adding 1 μg/ml of puromycin in the $CDM^{activin}$ culture medium. Cells were maintained under selection for a further two to three weeks at which point individual colonies were picked and seeded onto 12 –well plates to establish lines. Overexpression or knockdown was confirmed by qRT-PCR and positive clones were used for downstream analysis.

2.2.3 Small interfering RNA induced knockdown in hES cells

ON-TARGETPlus non-targeting control pool (siSCR) small interfering RNA (siRNA) and ON-TARGETPlus siP300 pool (siP300) were purchased from ThermoScientific (Cambridge, UK). Sequences for the four siRNAs in the siP300 pool are as follows: 1; 5'-CGACAGGGAUGCAGUAACA-3', 2; 5'-GUUCAAUAAUGCCUGGUUA-3', 3; 5'-GACAAGGGAUAAUGCCUAA-3' and 4; 5'-GGACUACCCUAUCAAGUAA-3'. The siRNAs were resuspended at 100 μM for stock solution in deoxyribonuclease (DNase)/ribonuclease (RNase) free dH_2O. For P300 knockdown hES cells were double transfected with siP300 and compared to cells transfected with siSCR. Briefly, for transfection in 6-well plates and for each transfected well, siRNA was diluted at a 25 nm final concentration into 15 μl of dH_2O (total final transfection volume is 1.5 ml). 12 μl of HiPerfect transfection reagent (Qiagen) was mixed with 300 μl of OPTIMEM media (LifeTechnologies). After this the diluted siRNA

and the media with the transfection reagent were added into the well and left incubating for approximately 30 min at RT. Following the 30 min incubation, hES cells to be transfected were plated in small clumps on top of the siRNA/medium in 1.2 ml of OPTIMEM used to passage the cells. Cells were then returned to the incubator and left overnight when the medium was changed to $CDM^{activin}$. On the second day after the first transfection the cells were re-transfected with siRNA. Media on the cells was changed to 1.2 ml of OPTIMEM. 25 nM final concentration of siRNA and 12 μl of HiPerfect were diluted in 300 μl of OPTIMEM. After incubation at RT for 30 min the siRNA mixture was added dropwise onto the cells. Cells were returned to the incubator and medium changed to $CDM^{activin}$ the following day. All analysis for the siRNA transfections was done after day four of treatment.

2.2.4 Human ES cell differentiation

2.2.4.1 Embryoid body formation

Human ES cells differentiation was achieved by promoting embryoid body (EB) outgrowths from hES cells CM and CDM conditions, as previously described (Xu et al. 2001; Osafune et al. 2008). Briefly, 1 well 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 EB media containing Knockout™ DMEM supplemented with 20% FBS, 0.1 mM NEAA, 2 mM L-glutamine and 0.55 mM β-mercaptoethanol. Cells were then scraped off with a cell scraper, dissociated by pipetting, and distributed into 1 well of an ultra-low attachment 6 well plate (Corning Inc.). On day 10, EBs were transferred back to 0.1% gelatinized-tissue culture plates, which allowed 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.

2.2.4.2 Neural progenitor cell differentiation

Human ES cells were differentiated to neural progenitor cells according an established protocol (Smith et al. 2008), with slight modifications. Briefly hES cells were maintained in CDM conditions as described above. Following 48h after splitting the cells neural differentiation was induced by culturing the cells in CDM media supplemented with 100 ng/ml of Noggin (R&D Systems Europe Ltd), 10 μM SB431542 (Sigma-Aldrich) and 12 ng/ml of FGF2. Media was changed daily and cells collected for real-time quantitative PCR on days 0, 2, 4, 6, 8 and 10. Primer pairs used for this analysis are listed in Appendix I, Table A1.

Human neural progenitor (NP) cells derived from undifferentiated H7 cells from CM conditions were derived according to Gerrard et al. (Gerrard et al. 2005a). 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 epidermal growth factor were added after an extra 8 days in culture. NP cells were plated at $1x10^5$ cells/cm^2 in 6 well culture plates. Media was changed every 2 days.

2.2.5 Replication timing analysis

2.2.5.1 BrdU labelling and cell cycle fractionation by flow cytometry

All reagents used in this analysis were from Sigma-Aldrich 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 µM 5-bromo-2-deoxyuridine (BrdU) for 90 min at 37°C, washed in cold PBS (PAA laboratories) 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 rotations per minute (rpm) for 10 min, 4°C. BrdU-labelled cells were resuspended in staining buffer (40 mM (hydroxymethyl)methylamine (Tris) pH 7.4, 0.8% NaCl, 21 mM MgCl2 and 0.05% NP-40) with 50 µg/ml propidium iodide (PI) and 1 mg/ml RNase A, carefully disaggregated with the help of a 40 µm cell strainer (Becton Dickinson, Oxford, UK) 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% sodium dodecyl sulphate (SDS), 0.2 mg/ml proteinase K (PK), 0.25 mg/ml salmon sperm DNA (LifeTechnologies)] and stored at -20°C following incubation for 2 hours at 50°C.

2.2.5.2 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 10^4 sorted cells). DNA was then purified by phenol, chloroform/isoamylalcohol extraction and ethanol precipitation. Precipitated DNA was dissolved in 480 µl of Tris-EDTA (TE) with 0.2 mg salmon sperm DNA (LifeTechnologies) at

37°C for 1 hour in the dark. Samples were then sonicated 45 sec on a Vibra-Cell™ VC130 (Sonics & Material 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, 50 µl of Adjusting buffer [10 mM sodium phosphate (pH 7), 0.14 M NaCl, and 0.05% Triton X-100] was added to each fraction and incubated with 80 µl 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 200 cells per microliter. 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.

2.2.6 Reverse transcription polymerase chain reaction

2.2.6.1 RNA extraction and cDNA synthesis

RNA extraction was performed using RNeasy protect mini kit (Qiagen, Valencia, CA) and residual DNA was eliminated using the RNase-free DNase kit (Qiagen) according to manufacturer's instructions. RNA was then reverse transcribed using Superscript™ First-Strand Synthesis system (LifeTechnologies). 0.5 - 1 µg of total RNA was supplemented with 1 µl of 10 mM dNTP mix (LifeTechnologies) and 1 µl of oligo (dT)12-18 and diluted in RNase free water to a final volume of 12 µl. The mixture was incubated at 65°C for 5 min and put on ice for 2 min, when 2 µl of 0.1 M dithiothreiol, 4 µl of 5X first strand buffer, 1 µl of RnaseOUT (LifeTechnologies) were added before being further incubated at 42°C for 2 min and 1 µl of 200 U/µl Superscript II was added. A reaction mixture without the enzyme was also set up as a control (designated "-RT"). The mixture was incubated at 42°C for 1 h and at 70°C for 15 min to stop the reaction. 380 µl of RNase-free water was added and cDNAs of interest were then detected by quantitative real-time PCR.

2.2.6.2 Primer design and testing for quantitative PCR

Primers for ChIP, replication timing and gene expression analysis were designed with the NCBI (National Centre for Biotechnology Information) primer blast search engine

(http://www.ncbi.nlm.nih.gov/tools/primer-blast/). The following design criteria were used: amplicon size of 100-250 bp, GC content of 45-55%, melting temperature 57-63°C and primer size 15-27 bp. Additionally for gene expression analysis one of the primers was required to span an exon-exon junction. All primer pairs were tested for predicted products within the mouse and human genome using *in silico* PCR UCSC genome browser (http://genome.ucsc.edu). Primers yielding a single predicted PCR product of the correct size were ordered from Sigma Genosys (the primers sequences are shown in their 5' to 3' orientation). The efficiency of amplification for each primer pair was tested by qRT-PCR analysis with sequential 2-fold dilutions of genomic DNA or cDNA (Pfaffl 2001). Primers yielding poor linear fits of the threshold cycle [C(T)] versus logarithm of concentration ($R2<0.98$) or efficiencies lower than 1.8 or greater than 2.2 were discarded. All primers are listed in Appendix I, Table A1.

2.2.6.3 Real-time quantitative PCR analysis

Real-time PCR analysis was carried out in 20 µl reaction volume with Jumpstart Sybr-Green PCR Mastermix (Sigma-Aldrich), 300nM 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 performed on a DNA Engine Opticon™ using Opticon Monitor 3 software (Bio-Rad Laboratories Inc., Hercules, CA), 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 and 83°C. The melting curve was determined from 70°C to 90°C, at 0.2°C intervals. The analysis of the qRT-PCR data was performed with the Opticon Monitor 3 software and the relative abundance of sequences was calculated using the ΔΔC(T) method (Livak et al. 2001; Pfaffl 2001). When the amplification efficiency is close to 2, the relative amount of PCR products between reactions 1 and 2 can be calculated as $2^{-\Delta C(T)1/2-\Delta C(T)2}$, being C(T) the threshold cycle at which fluorescence due to PCR products becomes detectable above background. Gene expression data was normalized to the average of two housekeeping genes as previously described (Vandesompele et al. 2002). Briefly, the relative expression levels of target genes were normalised to the average expression levels of *HPRT* and *GAPDH* in each specific sample. These housekeeping genes were confirmed not to significantly change in expression between the tested samples.

2.2.6.4 Semi-quantitative PCR analysis

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.

2.2.7 Immunofluorescence analysis and alkaline phosphatase activity

Undifferentiated hES cells were cultured on coated glass coverslips according to their growth conditions for at least 24 hours. Coverslips were removed, washed in PBS and fixed in 4% paraformaldehyde (PFA) (Sigma) 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 overnight at 4°C in a humid chamber. Coverslips were washed in wash buffer (0.2% BSA, 0.05% Tween20 in PBS; 3 x 5 min) and incubated with secondary antibodies coupled with appropriate 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 4,6-diamidino-2-phenylindole (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 at least 2 days under their standard culture conditions, using a kit from Sigma (86R-1KT) according to the manufacturer's instructions.

Samples were visualised using a Nikon Eclipse TE2000-U microscope and images processed using QCapture Pro software, or Nikon Eclipse TS100 and Nikon D5100 digital camera or a TCS SP5 Leica laser-scanning confocal microscope using the Leica Confocal software. All images were processed using Adobe Photoshop CS5. Microscope settings and laser power were kept constant between the controls and samples.

2.2.8 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 $5x10^6$ cells were washed twice in FACS buffer (PBS, 2% FBS, 2 mM EDTA), resuspended in 50 μl of blocking buffer (PBS, 2% FBS, 2 mM EDTA and 40% normal goat serum) and incubated for 15 min on ice to block no-specific binding. Primary antibody was then added at an appropriate dilution and incubated 30 min at 4ºC. The cells were then washed and resuspended in 100 μl of FACS buffer with appropriate secondary antibody and incubated for an additional 30 min in the dark on ice. Finally, cells were washed and then resuspended in 1 ml of PBS with 2 mM EDTA 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 $1.5x10^6$ 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 or to cells stained with an isotype control antibody. 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.

2.2.9 Protein analysis by western blotting

Whole cell extracts were prepared by direct lysis of cells with cold Radio Immuno Precipitation Assay buffer [150 mM NaCl, 50 mM Tris pH 8.0, 0.5% Deoxycholic acid, 0.1% sodium dodecyl sulphate (SDS), 1% NP-40 plus protease inhibitors]. Cells were scraped from the plates, incubated under rotation at 4°C for 20 min and centrifuged at 13 000 rpm for 20 min at 4°C and the supernatant collected. The protein solution was quantified by Pierce BCA Protein assay (ThermoScientific) and stored at -80°C.

20-40 μg 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 5 μl PageRuler™ protein ladder (ThermoScientific). Resolved acrylamide gels were blotted to Polyvinylidene fluoride (PVDF)

membranes (Millipore) using either a wet-transfer method for large proteins [> 100 kiloDalton (kDa)] in a transfer tank at 100 V for 1 hour in transfer buffer (48 mM Trizma base, 39 mM glycine, 0.037% SDS and 10% methanol) in 4°C or semi-dry transfer method for smaller proteins (< 100 kDa) between transfer plates soaked 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) or 5% BSA 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 with 5% milk or BSA, 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 either Pierce ECL western blotting substrate (ThermoScientific) or ECL prime western blotting detection reagents (GE Healthcare Life Sciences, Little Chalfont, UK) for HRP-conjugated antibody detection before being exposed to X-ray films (CL-XPosure Film, ThermoScientific) and using the OPTIMAX IGP developer, or by using GE LAS4000 chemiluminescence imager. 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.

2.2.10 Colony formation assay

H1 and H9 hES cells were trypsinised to a single cells suspension and seeded on 12-well plates at a density of 5000 cells per cm^2 in either normal $CDM^{activin}$ or in CDM media without Activin A supplementation. After eight days of culture the cells were fixed and assayed for alkaline phosphatase activity as described in section 2.2.7. The resulting colonies were classified as undifferentiated, mixed or differentiated (See Fig. 5.5 for examples), and the amount of each class in the plates was counted and represented as a proportion of total number of colonies.

2.2.11 Chromatin immunoprecipitation analysis

2.2.11.1 ChIP for histone modifications

All reagents used were from Sigma unless stated otherwise. Approximately $0.5\text{-}1x10^8$ exponentially growing cells were harvested, washed twice in PBS (PAA) 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 tablets (Roche). Using a Bioruptor^{TM} 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 14 000 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 blocked overnight 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) (LifeTechnologies) 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 normalisation, under rotation. 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 under rotation, 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-HCl 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-HCl pH 8.1 and protease inhibitors). After adding elution buffer (1% SDS, 0.1 M $NaHCO_3$) 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 normalised 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 qRT-PCR amplification and the primers listed on Appendix I, Table A3.

2.2.11.2 ChIP for P300

The protocol used with minor modifications has previously been described in detail by (Stock et al. 2007). 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 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (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. 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 phenylmethylsulphonyl fluoride, 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 normalised against the total input material. Quantification of precipitated DNA was performed using qRT-PCR amplification and the primers listed on Appendix I, Table A3.

2.2.12 Analysis of DNA content and BrdU incorporation via acid denaturation of DNA

All reagents used in this analysis were from Sigma unless otherwise stated. Exponentially growing non-synchronized H9 cells were incubated with 30 µg/ml BrdU for 60 min at 37°C. Cells were collected from plates by trypsinisation, pooled with the floating cells in the medium and resuspended in room temperature PBS at $1x10^6$ 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 5 ml of phosphate/citric acid buffer, pH 7.4 (182 mM Na_2HPO_4 and 9 mM citric acid) was added to the suspension, cells were centrifuged 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 Characterisation of replication timing profiles in hES cell lines alongside mES and mEpiS cell lines

3.1 Introduction

Chromatin profiles are increasingly being used as a valuable signature for defining the cell identity and developmental state of a given cell population. One approach has been to assay the timing of DNA replication across a panel of candidate loci, as an indicator of chromatin accessibility. Acetylated histones, which are associated with accessible chromatin, have been demonstrated to be highly correlated with early replication, and conversely late replication has been correlated with inaccessible heterochromatin (Gilbert 2002; Vogelauer et al. 2002; Peters et al. 2005; Ryba et al. 2010). Remarkably, this replication timing profiling allowed the distinction of pluripotent mES cells from cells with a more restricted differentiation capacity, including mouse EC cells (Perry et al. 2004; Azuara et al. 2006; Hiratani et al. 2008).

Despite being called embryonic stem cells, hES cells differ from mES cells in a number of ways, and have been suggested to be more closely affiliated with mEpiS cells. *In vivo* determination of hES lineage potential is not possible, due to ethical reasons, and their exact developmental affiliation remains unclear. As mentioned above chromatin profiling has been successfully used to distinguish closely related pluripotent mouse cell populations with different differentiation potentials. Here, human and mouse pluripotent stem cell populations derived from the early embryo, namely hES (H1, H7 & 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 details of cell lines used), are analysed and compared to each other using DNA replication timing profiling. This is with the aim of determining the embryonic affiliation of hES cells.

3.2 Epigenetic profiling of hES, mES and mEpiS using a PCR-based DNA replication timing assay

Human ES, mES and mEpiS cells are all types of pluripotent cells that can be propagated undifferentiated for an unlimited period of time under appropriate conditions and commonly share the expression of key pluripotency markers such as *Oct4, Sox2* and *Nanog* (Yu et al. 2008). As Figure 3.1a illustrates Oct4 and Nanog expression was confirmed in the three cell types analysed. Despite these similarities, several differences between mES, hES and mEpiS cells have been reported especially with regards to surface antigen expression

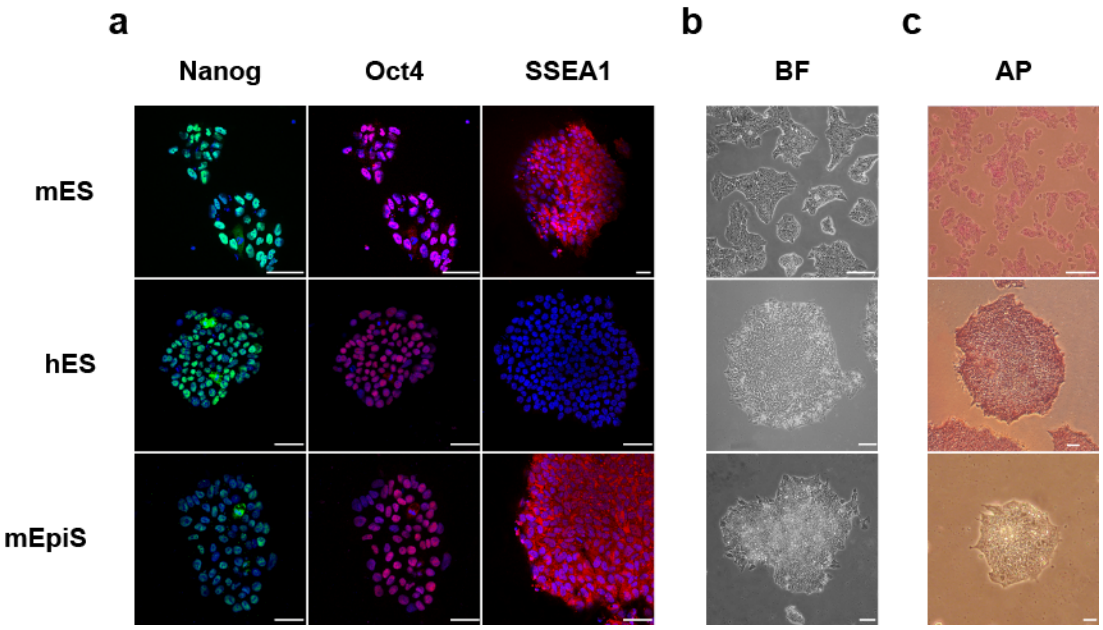


Figure 3.1 Phenotypic similarities and differences between undifferentiated mES, hES and mEpiS cells.
(**a**) Mouse ES (E14tg2A line, top panel), hES (H1 line, middle panel) and mEpiS (129 line, bottom panel) cells were analysed by immunofluorescence for the expression of Nanog (green), Oct4 (red) and SSEA1 (red). Nanog and Oct4 are highly expressed in mES, hES and mEpiS cells, while SSEA1 is not expressed in hES cells. Images overlaid with DAPI staining (blue) showing nuclei. (**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**) Both hES and mES cells stain positive for alkaline phosphatase (AP) enzymatic activity (pink), however mEpiS cells do not. Scale bars 50 µm.

and cell morphology (Pera et al. 2004b; Pera et al. 2010). As can be seen in Figure 3.1b, undifferentiated mES cells grow in small and domed colonies. In contrast hES and mEpiS form larger flattened colonies that grow as a monolayer. Expression of the cell surface marker SSEA1 is restricted to undifferentiated mES and mEpiS cells, whereas only mES and hES cells stain positive for alkaline phosphatase (AP) activity (Fig. 3.1a & c). Moreover, a key difference between mES and hES cells is the extrinsic signalling that enables their self-renewal. Undifferentiated mES cells require LIF/BMP signalling to remain self-renewing, whereas hES cells do not respond to LIF/BMP signalling but instead require signals through the Activin/FGF pathway. Interestingly undifferentiated mEpiS cells also strictly depend on Activin/FGF signalling (Yu et al. 2008). Taken together, this raises the question of the true lineage affiliation of hES cells, a developmental paradox. Are the differences observed between mES and hES cells due to species divergence or directly reflect the embryonic development stage from which the cells arise? Using DNA replication timing, this study sought to revisit this developmental paradox, by investigating the cell and chromatin status of mES, hES and mEpiS cells.

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

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

A panel of 45 key transcriptional regulatory genes (mostly encoding human and mouse homologs) was selected for this analysis. These genes are selectively expressed in embryonic (pluripotent) and extra-embryonic (trophectoderm) tissues of the early embryo or are involved in specifying ectoderm, mesoderm, endoderm or germ-line lineages (Table 1). The POU TF Oct4 (Nichols et al. 1998; Reubinoff et al. 2000; Brons et al. 2007; Tesar et al. 2007) and the homeobox gene Nanog (Chambers et al. 2003; Clark et al. 2004; Brons et al. 2007; Tesar et al. 2007) are key pluripotency-associated genes in both human and mouse stem cell populations. Other loci showing differential gene expression between ICM/ES cells (*Rex1* and *Gbx2*) (Rogers et al. 1991; Chapman et al. 1997) and epiblast/EpiS cells (*Fgf5, Psc1* and *Prce)* (Haub et al. 1991; Hebert et al. 1991; Pelton et al. 2002; Brons et al. 2007; 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 was not available in human embryos. Genes involved in the trophectoderm 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* (Ciruna et al. 1999; Russ et al. 2000; Babaie et al. 2007). 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* (Stevanovlć et al. 1993; Collignon et al. 1996) 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 TF *Gata6* (Morrisey et al. 1998; Koutsourakis et al. 1999) 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 candidate loci was assessed using a previously described method (Hansen et al. 1993; Azuara et al. 2003; Azuara 2006) a schematic representation 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 selectively incorporated into newly replicating DNA. Following this the cells were stained with the DNA-binding dye propidium iodide (PI) and fractionated by flow cytometry into six cell cycle fractions according to DNA content. The first fraction included cells in G0 and G1, while S-phase was represented by four fractions (designated as S1 to S4) and the G2 fraction included cells that had completed DNA replication. A representative cell cycle profile for

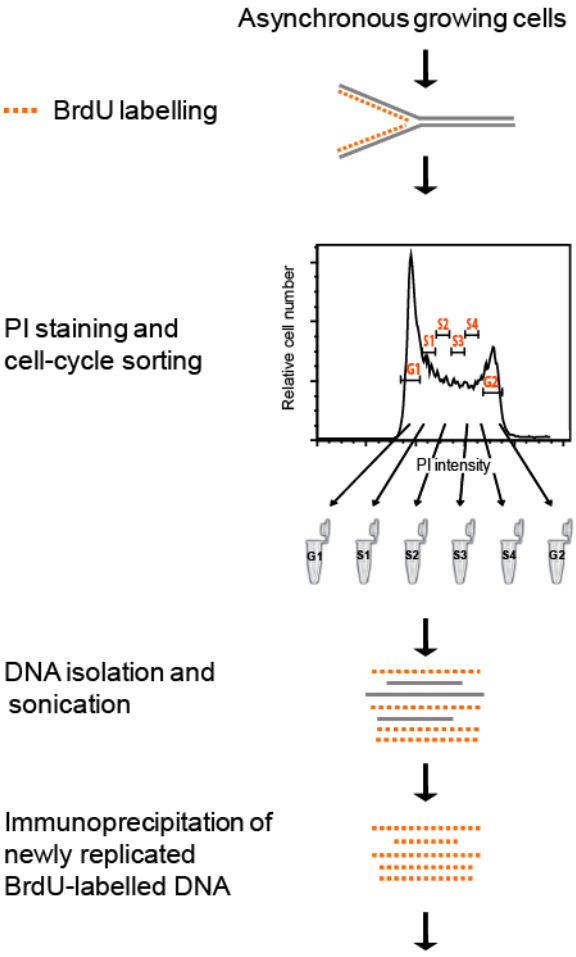


Figure 3.2 Schematic representation of DNA replication timing analysis.
Asynchronously growing cells are pulse-labelled with 5-bromo-2-deoxyuridine (BrdU), which is incorporated into replicating DNA. Cells are sorted according to DNA content into six cell cycle fractions (designated as G1, S1-S4 and 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 recovered by immunoprecipitation using anti-BrdU antibodies and the relative abundance of locus-specific sequences analysed by qRT-PCR.

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, was purified by immunoprecipitation with anti-BrdU antibodies. Locus-specific sequences were then quantified by real-time quantitative PCR (qRT-PCR) 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.

Controls for replication timing analysis included the *α-globin* locus, a constitutively early-replicating gene (Smith et al. 1999) and pancreatic *Amylase*, a tissue-specific gene, which was consistently shown to replicate late in S-phase in several different cell types (Hatton et al. 1988; Mileham et al. 1996; Azuara et al. 2003). 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 (Fig. 3.3). To ensure equivalent precipitation

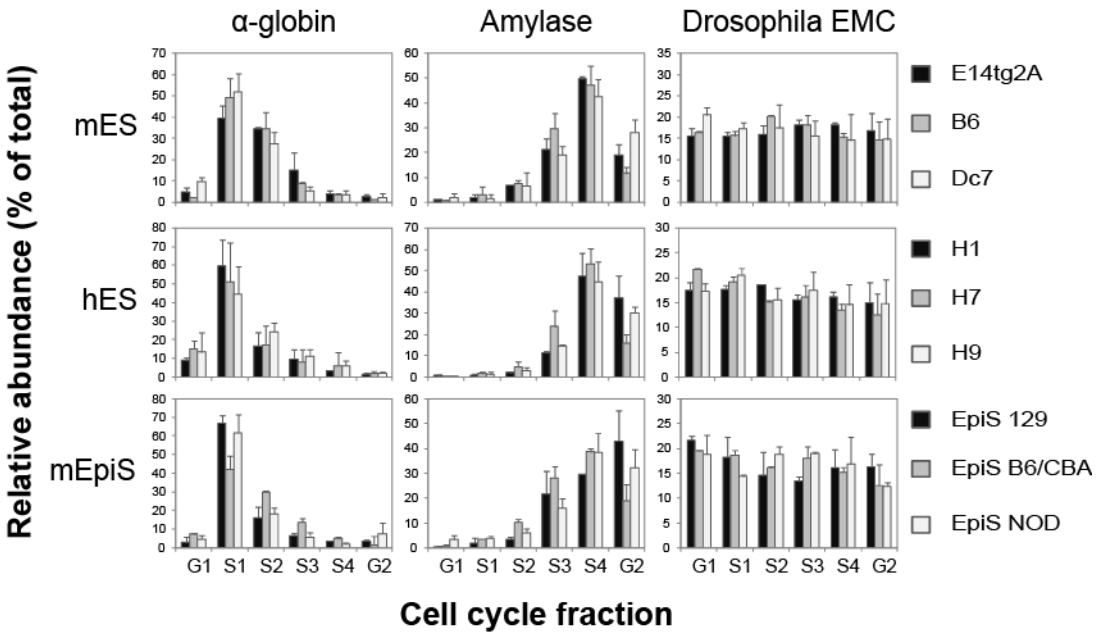


Figure 3.3 DNA replication timing analysis controls.
Histograms showing the relative abundance of PCR products for early and late replicating controls (*α-globin* and *Amylase*, respectively) in each cell cycle fraction for mES (E14tg2A, B6 and Dc7), hES (H1, H7 and H9) and mEpiS (129, B6/CBA and NOD) 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.

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 (Fig. 3.3, right panel) as assessed in each experiment.

3.3 Mouse ES cells have conserved replication timing profiles between cell lines from different genetic backgrounds

Replication timing profile was classified for each candidate locus according to its peak abundance in the different cell cycle fractions from G_1/S_1 (early), S_2 (middle-early), S_2 and S_3 (middle), S_3 (middle-late) and S_4/G_2 (late). For ease of comparison the replication timing data is represented as a gradation of colours from early (green) to late (red) as shown in figure 3.4a. Three mES cell lines were used for analysis of their replication timing profiles. These lines were all independently derived from inbred mouse lines from distinct genetic backgrounds. The E14tg2A line was derived from a 129 strain in 1987 (Hooper et al. 1987). Line B6 was derived in 1993 from a C57BL/6 strain (Kontgen et al. 1993). More recently the Dc7 line was derived from a mixed background of 129/*castaneus* hybrid mouse (2007, T. Nesterova, unpublished). All three mES cell lines were maintained in media containing LIF and serum. E14tg2A was cultured on gelatin-coated wells, whereas B6 and Dc7 lines

required co-culture with MEFs as feeder cells. In addition to this the Dc7 line was cultured in media with a slightly different composition to the others (see section 2.2.1 for details).

Although all three mES cell lines have distinct genetic backgrounds, were derived at different times and further maintained in different culture conditions, their the replication timing profiles were highly conserved across many loci with only a few exceptions *Bmp4* and *Gata6,* which replicated earlier during the S-phase in E14tg2A and DC7 cells respectively (Fig. 3.4b). Results from this analysis confirmed previous reports (Azuara et al. 2006; Hiratani et al. 2008) showing that in mES cells majority of genes replicate early (green) or middle-early (light green) during the S-phase (73-76%, see Fig. 3.4c & d) and is in accordance with the most recent genome-wide analysis (Hiratani et al. 2008), which demonstrated that replication timing occurs in megabase level domains and is conserved across mES cells from distinct backgrounds. Loci replicating early included genes associated with pluripotency and thus actively transcribed in mES cells (e.g. *Oct4*, *Nanog*, *Sox2* and *Rex1*), but also several loci only expressed in specific somatic lineages (e.g. *Hoxb1*, *Nestin* and *Pdx1*), which are silent in ES cells. Hence, these results reiterate that many lineage-specifiers in mES cells harbour a broadly accessible chromatin structure being primed for future gene activation upon differentiation (Azuara et al. 2006; Bernstein et al. 2006; Hiratani et al. 2008). In contrast, other lineage-specific genes, such as *Foxa2*, *Myf5*, *NeuroD*, *Sox3* and *Zic1* were consistently replicating middle-late (orange) or late (red) during the S-phase in all three cell lines analysed in agreement with lineage induction in mES cells being likely regulated by selective locus derepression (Perry et al. 2004; Williams et al. 2006).

3.4 Replication timing profiles are distinct between pluripotent mouse and human ES cells

Next, replication timing analysis was extended to several hES cell lines. The same panel of selected loci (most of which are localised in syntenic regions of the genome between mouse and human) was analysed in H1, H7 and H9 hES cells. These cell lines were all derived at the same time in 1998 (Thomson et al. 1998) and for this analysis were maintained and expanded in widely used feeder-free cell culture conditions (Xu et al. 2001) on matrigel coated plates in MEF-conditioned media (CM), supplemented with Fgf2 (as described in section 2.2.1).

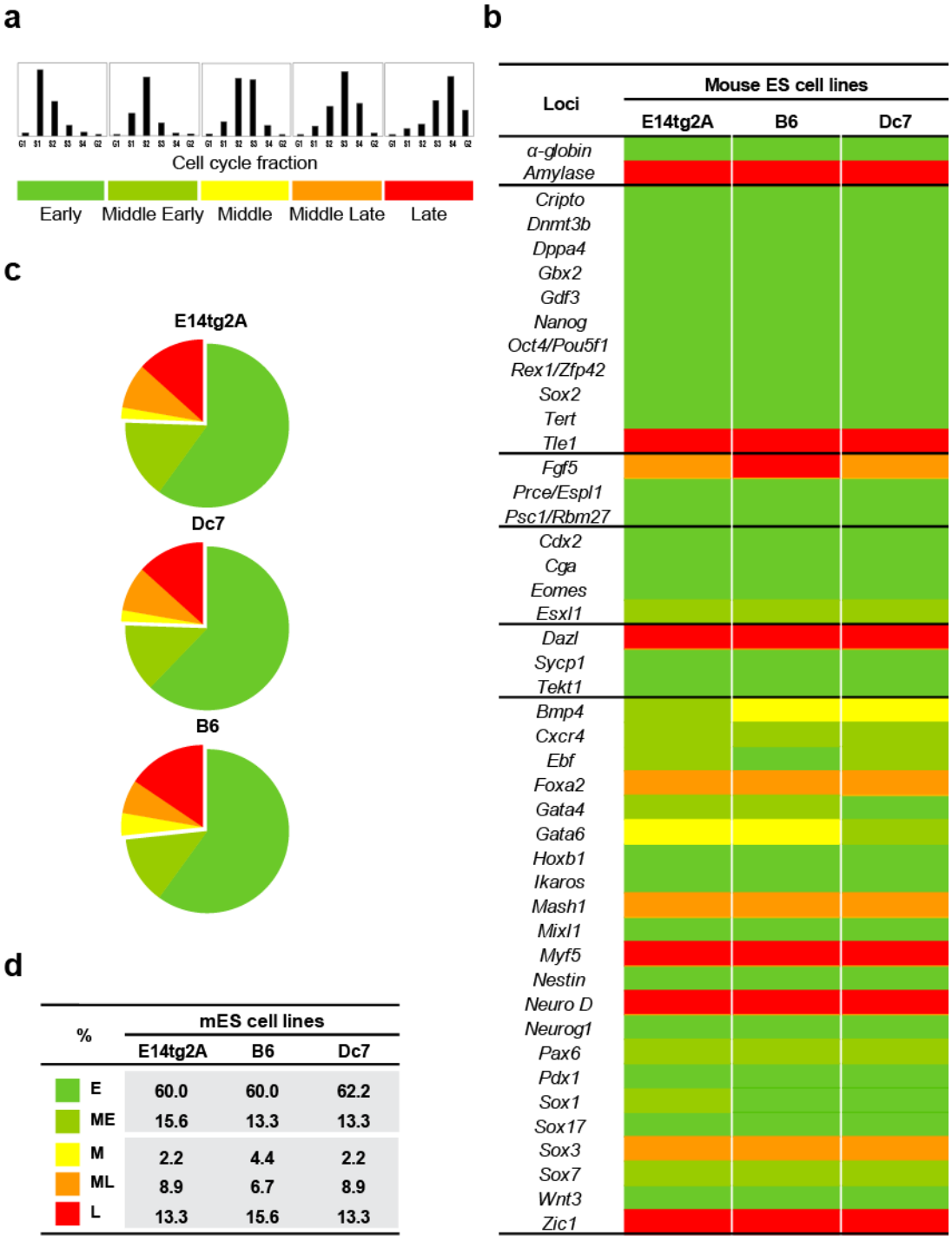
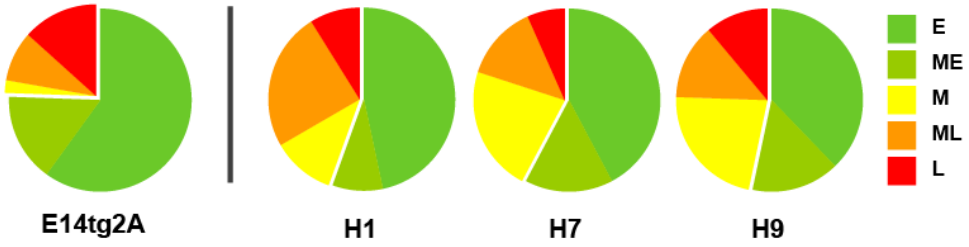


Figure 3.4 Independently derived mES cell lines from distinct genetic backgrounds have similar DNA replication timing profiles.

(**a**) The replication timing of each candidate locus 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 highly uniform replication timing profiles across a panel of 45 candidate genes. Same assignment of colours for replication timing of candidate loci is followed in all subsequent figures. (**c**) 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). (**d**) 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).

A profound difference between mES and hES cells was the increased proportion of loci replicating in the later half of the S-phase in hES cells. To allow easy comparison of mES and hES replication timing profiles the relative proportion of loci replicating in early, middle-early, middle, middle-late or late S-phase was summarised as pie-charts shown in figure 3.5a. Although the majority of examined loci replicated early or middle-early (53-58%) in hES cells, a significant difference when compared to mES cell replication timing profiles was the increased proportion of loci replicating in middle, middle-late or late S-phase (24-27% and 42-47% in mES and hES cells respectively) (Fig. 3.5b).

a



b

%	mES cell lines			hES cell lines		
	E14tg2A	B6	Dc7	H1	H7	H9
E	60.0	60.0	62.2	46.7	42.2	37.8
ME	15.6	13.3	13.3	8.9	15.6	15.6
M	2.2	4.4	2.2	11.1	22.2	22.2
ML	8.9	6.7	8.9	24.4	13.3	13.3
L	13.3	15.6	13.3	8.9	6.7	11.1

Figure 3.5 Pluripotent mouse and human ES cells have distinct DNA 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 pie charts. (**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. For full comparison of mES and hES profiles see Appendix II, Fig. A2.

This data was surprising as increased proportion of late replicating loci was previously reported as a feature of differentiated cells (Azuara et al. 2006). To exclude the possibility of spontaneous differentiation as a possible cause for the observed late replication timing profiles, H1, H7 and H9 cultures were carefully analysed for the expression of key markers of pluripotency. TFs OCT4, NANOG and SOX2 were analysed alongside the specific surface markers SSEA4, TRA-1-81 and TRA-1-60 by immunofluorescence combined with FACScan

analysis. Figure 3.6a and b demonstrate that hES cell cultures in this study displayed similar expression levels of OCT4, NANOG and SOX2 as seen across the three cell lines. TRA1-60, TRA1-81 and SSEA4 were also highly expressed in H1, H7 and H9 cells, while SSEA1 (a marker of differentiation in hES cells) was virtually undetectable. The hES cell cultures furthermore stained positive for AP activity (Fig. 3.6c), confirming their undifferentiated status. Moreover, the replication timing profiles of undifferentiated H7 hES cells could be clearly distinguished from differentiated neuronal progenitor cells and established neuronal cells (Appendix II, Fig. A1).

Interestingly and in contrast to the highly conserved mES cell's replication timing profile, hES cells exhibited a significant degree of variability between their profiles, as seen among the three cell lines analysed and/or independent cultures for each cell line (Fig. 3.5 and 3.7, Appendix II Fig. A2 and A3). Eleven out of the 45 loci analysed (*DPPA4*, *TERT*, *TLE1*, *SYCP1*, *BMP4*, *EBF*, *FOXA2*, *GATA6*, *SOX17*, *SOX3* and *ZIC1*) were observed to have different replication timing profiles between H1, H7 and H9 hES cell lines (Fig. 3.7a). Furthermore the profile of the *TERT* locus varied even between cultures within the same cell line. The observed variations ranged from drastic differences, from early to late as seen for *DPPA4* and *EBF*, with *DPPA4* being selectively replicated in late S-phase in H9 cells only, while *EBF* replicated in late, middle and middle early in H1, H7 and H9 cells, respectively. Subtler, variations from middle to middle-late, were also seen, for example at the *GATA6* and *TLE1* loci (Fig. 3.7b). In contrast, *NANOG* and *NEUROD* are shown as examples of genes that consistently replicated in early and late S-phase, respectively (Fig. 3.7c). The loci with variable replication timing profiles, did not show any bias towards a specific lineage. Furthermore, no correlation was found with the variable replication timing and expression levels of the specific genes, or with histone acetylation or DNA methylation (Appendix II Fig. A4 and A5). However, the loci found variable in hES cells were all among gene poor regions of the genome, apart from *TERT* (Appendix II, Table A4). 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.

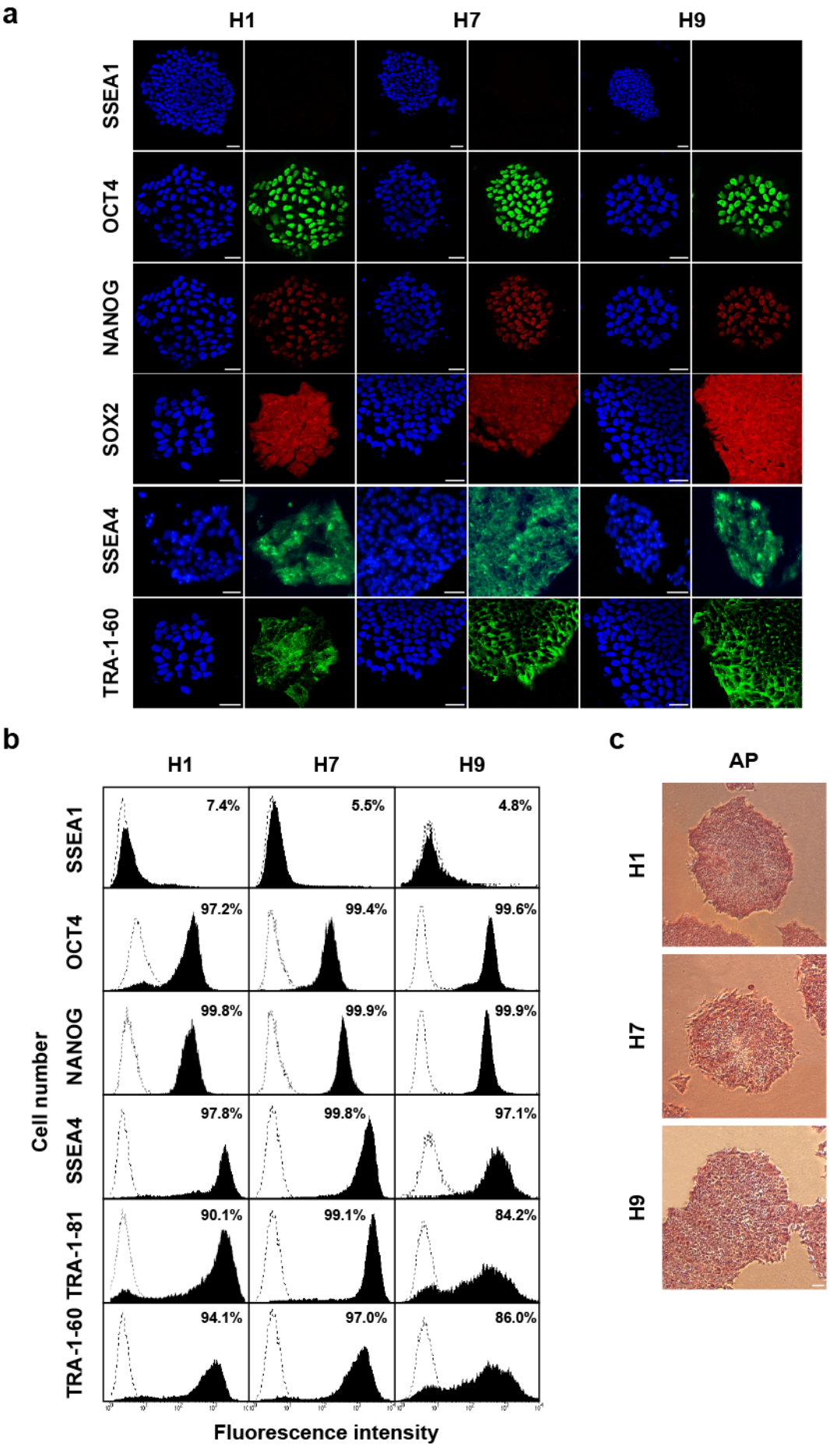

a
H1
H7
H9
SSEA1
OCT4
NANOG
SOX2
SSEA4
TRA-1-60
b
H1
H7
H9
SSEA1
OCT4
NANOG
SSEA4
TRA-1-81
TRA-1-60
7.4%
5.5%
4.8%
97.2%
99.4%
99.6%
99.8%
99.9%
99.9%
97.8%
99.8%
97.1%
90.1%
99.1%
84.2%
94.1%
97.0%
86.0%
Cell number
Fluorescence intensity
c
AP
H1
H7
H9

Figure 3.6 Human ES cell cultures analysed by replication timing display similar expression of pluripotency-associated markers.
H1, H7 and H9 hES cells were analysed by (**a**) immunofluorescence and (**b**) flow cytometry for the hES cell 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**) Positive staining for alkaline phosphatase (AP) activity (pink) was also strongly detected in all hES cell colonies. Scale bars 50 µm.

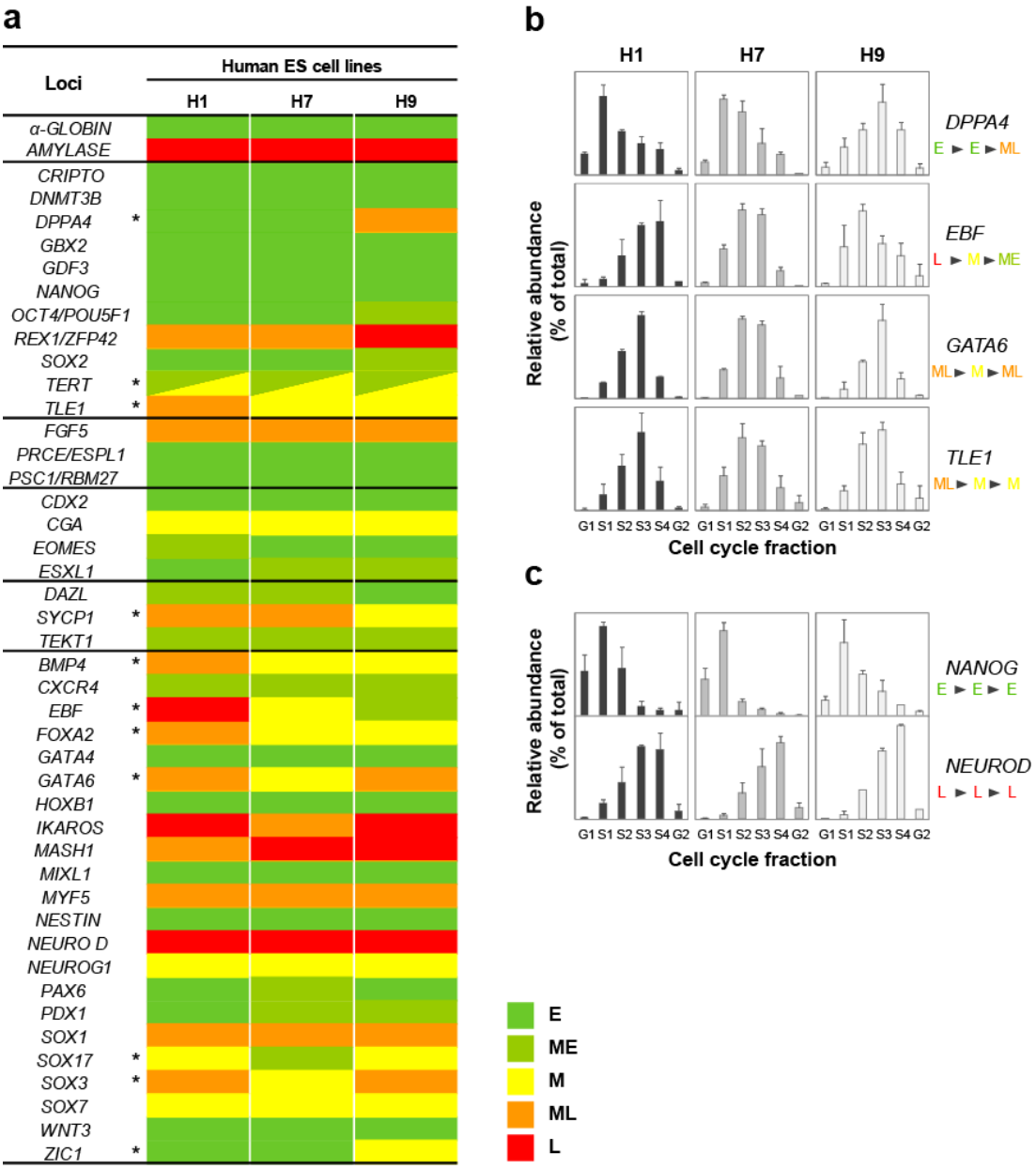


Figure 3.7 DNA replication timing variability between established hES cell lines.
(**a**) The replication timing of individual candidate loci was compared between three hES cell lines (H1, H7 and H9). Loci showing differential replication kinetics are marked by asterisks. (**b** & **c**) Histograms showing the relative abundance of locus-specific DNA within each cell cycle fraction (G1, S1-S4 and G2 phases). Examples of loci that switch their replication timing, (*DPPA4*, *EBF*, *GATA6* and *TLE1*) (**b**) or consistently replicate in early and late S-phase (*NANOG* and *NEUROD*), respectively (**c**). Are shown 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.

3.5 Temporal comparison of hES and mEpiS cell's replication profiles suggests a common unstable or transitional state on the verge of differentiation

Human ES cells have been proposed to more closely resemble mEpiS cells than mES cells, which are derived from gastrulating mouse embryos and blastocysts respectively (Brons et al. 2007; Tesar et al. 2007). This view was then further investigated by comparing the replication timing profiles of three independent mEpiS cell lines from different genetic backgrounds (NOD and B6/CBA (Brons et al. 2007) and 129 (Vallier et al. 2009a)) to those of mES and hES cells (Fig. 3.8a and Appendix II Fig. A3, for full analysis).

To first ascertain that mES and mEpiS could be reliably discriminated at the replication timing level, initial analysis was conducted, specifically looking at key ICM (*Dppa4*, *Gbx2*, and *Rex1*) (Rogers et al. 1991; Chapman et al. 1997; Bortvin et al. 2003; Maldonado-Saldivia et al. 2007) and epiblast-specific markers (*Fgf5*, *Prce* and *Psc1*) (Haub et al. 1991; Hebert et al. 1991; Pelton et al. 2002; Brons et al. 2007; Tesar et al. 2007) alongside *Nanog*, *Oct4* and *Sox2*, which are expressed in both ICM and epiblast. Consistently with transcriptional status *Rex1* and *Dppa4* were found to replicate early in mES cells and late in mEpiS cells, respectively. Conversely *Fgf5* was found replicating late in mES but early in mEpiS cells, where this gene is normally induced in a developmental-regulated manner (Brons et al. 2007; Tesar et al. 2007) (Fig. 3.8a). The replication timing profile of *Nanog*, *Oct4* and *Sox2* remained early in both mEpiS cells and mES cells, and furthermore profiles of *Gbx2*, *Prce* and *Psc1* were also early across these cell types, regardless of their gene expression patterns.

Whether hES are more closely relate more closely to a mES or a mEpiS-like pluripotent identity, could not be, however, determined on this set of markers alone. While *Rex1* was indeed found late replicating in both hES and mEpiS cell populations, *Fgf5* remained late in hES cells as seen in mES. *Dppa4* was variably replicated from early to middle late between the hES cell lines, as previously described. Nevertheless, it has to be noted that the selection of these markers relied on gene expression datasets in mouse embryos, as equivalent information was not available in human embryos. Interestingly, however, mEpiS and hES cells were observed to have remarkably 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 (Fig. 3.8b). The relatively high number of cells in S-phase in mES cells, compared to mEpiS and hES cells, is likely due to the faster cell cycle kinetics of mES cells. Furthermore a high level of variable loci was observed among the

replication timing profiles of the different mEpiS cell lines (22.2% variable loci), a level similar to what was previously seen with hES cells (24.4%) and in contrast to the homogeneous and steady replicating mES cell profiles (4.4% variable loci) (Fig. 3.8c). Together with previous analysis (Fig. 3.7 and Appendix II Fig. A3), these data suggest that

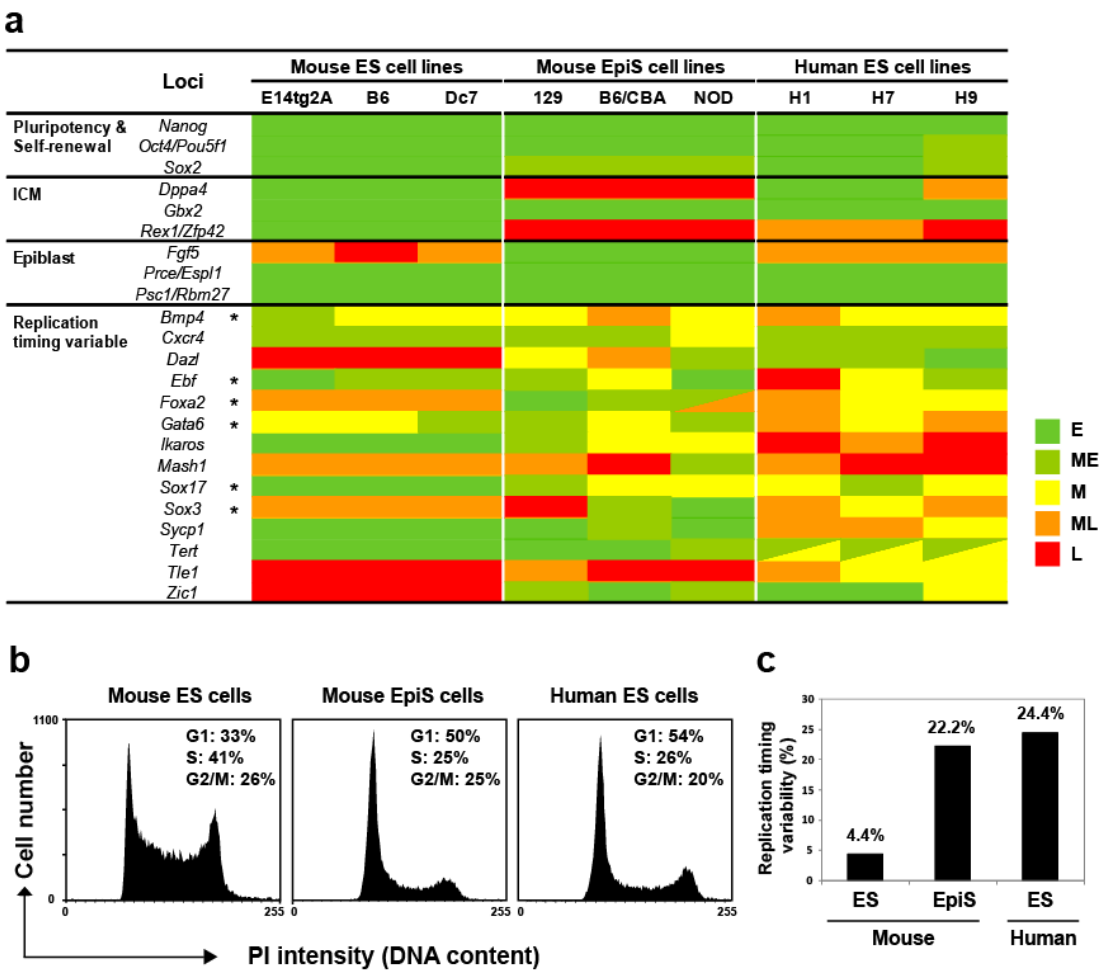


Figure 3.8 Mouse EpiS and hES cells have variable DNA replication timing profiles across many genes and similar cell cycle kinetics

(**a**) 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 variable in both mEpiS cell lines and hES cell lines are marked by asterisks. For the full list of loci analysed see Appendix II, Fig. A3. (**b**) Representative cell cycle profiles of mES (E14tg2A), mEpiS (129) and hES (H1) cells. Cells were stained with propidium iodide (PI) and DNA content analysed 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, which is notably different from the pattern observed with mES cells. (**c**) Percentage of replication timing variable loci between cell lines of the same population relative to the total 45 analysed 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.

variability might link to the pluripotent state related to post-implantation epiblast, and that hES cells preferentially adopt a mEpiS-like state in culture.

3.6 Conclusions

Results from this chapter highlight that mES cells have stable and broadly early replicating profiles, regardless of their genetic background and in accordance with data previously published (Azuara et al. 2006; Hiratani et al. 2008). Rather surprisingly replication timing analysis of three independent hES cell lines revealed considerable differences between mES and hES cells. Human ES cells had a notably higher proportion of candidate loci replicating in either middle or later S-phase as compared to mES cells. A shift toward late S-phase across many loci has previously been linked with cell differentiation (Azuara et al. 2006), however the undifferentiated status of hES cells was carefully assessed and confirmed and thus could not solely explain the observed differences. Moreover, hES cells also exhibited unprecedented variability in their replication timing profiles as seen between different cell lines and cultures. No correlation was found with the variable loci and gene expression or epigenetic modifications, however nearly all variable loci were within gene poor regions of the genome (See Appendix II Fig. A4 and Table A4). Analysis of the replication timing profiles of mEpiS cells alongside mES cells could discriminate between these closely related cell populations. Even though both cell types had mostly early replicating profiles, key marker loci were found to be consistently distinct between mES and mEpiS cells, such as *Rex1, Dppa4* and *Fgf5* (Appendix II Fig. A3). However, comparing hES replication timing profiles to those of mES and mEpiS cells, was unable to clarify the lineage affiliation of hES cells, with different markers aligning between mES and mEpiS cells. However, interestingly similar replication timing variability as in hES cells was also observed in mEpiS cell lines analysed. This similarity suggests that the variability might be linked to a shared pluripotent state between mEpiS and hES cells.

Chapter 4 Distinct epigenetic states within pluripotent human ES cells and characterisation of their regulation

4.1 Introduction

Phenotypic and functional heterogeneity within mES cultures has been previously reported (Chambers et al. 2007; Hayashi et al. 2008; Toyooka et al. 2008) within the mES cultures, delineating distinct pluripotent states from a more naïve to a more primed state (Enver et al. 2009; Nichols et al. 2009b). These different states are capable of interconvertion in response to different culture conditions (Chou et al. 2008; Bao et al. 2009; Guo et al. 2009; Hanna et al. 2009). For example, *Nanog* was seen to be expressed at varying levels within mES populations cultured in serum containing medium with LIF supplementation. Interestingly, *Nanog*^{low} cells were more prone for differentiation and *Nanog*^{high} cells more prone for self-renewal, however both populations were able to also give rise to each other (Chambers et al. 2007). *Rex1* was also observed to exhibit similar varied expression patterns in mES cells, as well as similar bias for differentiation and self-renewal in *Rex1*^{low} and *Rex1*^{high} cells, respectively, as seen with *Nanog* (Marks et al. 2012). Importantly, when mES cells were grown in 2i conditions, thought to promote a naïve pluripotent state, both *Nanog* and *Rex1* are expressed at a higher and more uniform level (Wray et al. 2010).

Heterogeneity has also been reported previously in hES cell cultures. *NANOG* expression levels have been show to be variable in hES cells (Fischer et al. 2010), as well as patterns of cell surface markers (Enver et al. 2005; Stewart et al. 2006b). These previous observations, and the observed epigenetic variability in replication timing profiles of H1, H7 and H9 cells could suggest the existence of sub states in hES pluripotent identity. Such states would likely be regulated by culture conditions, as seen in mES cells. With hES cells, supplementation by Activin A and/or Fgf2 has been shown to be essential for the maintenance of pluripotent ES cultures (Vallier et al. 2005a; Levenstein et al. 2006; Xiao et al. 2006), despite this a great variety of techniques to stably maintain pluripotent hES cultures have been developed and are used in different laboratories (Adewumi et al. 2007). To investigate the effects of the culture conditions on hES epigenetic states, here, replication timing analysis was conducted on hES cells grown in the different culture medium.

4.2 Replication timing profiles of undifferentiated pluripotent human cell populations undergo extensive alterations upon different cell culture conditions

4.2.1 Distinct replication timing profiles of H9 hES cells grown in two different laboratories

To directly assess the impact of culture environment on hES epigenetic states, H9 hES cells grown in our laboratory under CM conditions [matrigel-coated plates and MEF-CM with KSR and Fgf2 (Xu et al. 2001)] were compared to cells of the same cell line grown in CDM^{activin} conditions [FBS-coated plates and KSR-free defined medium supplemented with FGF2 and Activin A (Vallier et al. 2005a)] grown in Dr Ludovic Vallier's laboratory (Fig. 4.1a). Surprisingly the replication timing profiles of H9 cells grown in CDM^{activin} were dramatically different to those grown in CM conditions, with an increase in the proportion of candidate loci replicating in the early and middle-early S-phase. In H9 hES cells cultures maintained under CDM^{activin} conditions for more than 20 passages [H9p50(CDM20)], 75% of loci analysed replicated in early or middle early S-phase, as compared to H9-CM cultures (H9p35CM), where only 55% of loci were found to replicate in the first half of the S-phase (Fig. 4.1b).

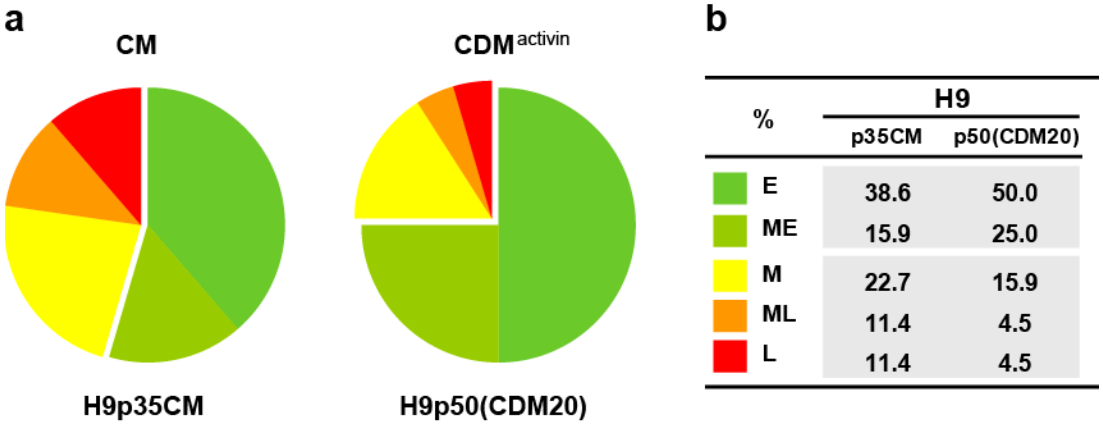
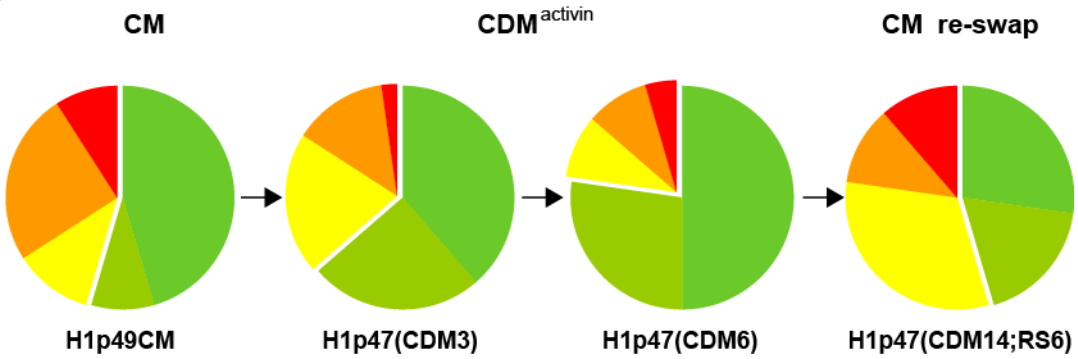


Figure 4.1 Human ES cells grown in different cell culture conditions and different laboratories exhibit different replication timing profiles
(**a**) The proportion of candidate loci replicating early (E), middle-early (ME), middle (M), middle-late (ML) and late (L) in S-phase are represented as pie charts. The replication timing profile of H9 hES cells grown in our laboratory, in feeder free cultures on matrigel-coated plates and in conditioned medium (CM), prepared with irradiated mouse embryonic fibroblasts, Knockout™ Serum Replacement (KSR) and fibroblast growth factor 2 (Fgf2) (Xu et al. 2001) (H9p35CM), was compared to that of cells from the same line grown on gelatin coated plates in KSR-free chemically defined medium (CDM) supplemented with Fgf2 and Activin A (CDM^{activin}) (Vallier et al. 2005a) kindly provided by Dr. Ludovic Vallier [H9p50(CDM20)]. (**b**) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for H9 hES cells grown under 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.

These results indicate that cells maintained in different culture conditions in different laboratories can adopt distinct epigenetic states.

a



b

%	H1			
	p49CM	p47(CDM3)	p47(CDM6)	p47(CDM14;RS6)
E	45.5	38.6	50.0	27.3
ME	9.1	25.0	27.3	18.2
M	11.4	20.5	9.1	31.8
ML	25.0	13.6	9.1	11.4
L	9.1	2.3	4.5	11.4

Figure 4.2 H1 hES cells undergo dramatic late to early shifts of replication timing profiles when transferred from CM to CDM^{activin} conditions.
Replication timing profiles of H1 cells grown in CM conditions (H1p49CM) were compared to H1 cells transferred to CDM conditions and grown for three [H1p47(CDM3)] or six passages [H1p47(CDM6)] or fourteen passages in CDM and the re-swapped back to CM growth conditions for six passages [H1p47(CDM14;RS6)]. (**b**) Proportion of candidate loci that replicate in each S-phase fraction, expressed as a percentage of the total for H1 hES cells grown under CM and CDM^{activin} conditions.

4.2.2 Dynamic modulations of replication timing profiles in H1 hES cells in response to cell culture conditions

To address whether these alterations were reproducible and reversible, replication timing analysis was conducted comparing with H1 hES cell line when transferred from CM conditions to CDM^{activin} conditions and subsequently reversed back to CM conditions. After three passages H1 cells were growing stably in CDM^{activin} conditions, with little or no evidence of spontaneous differentiation (see section 4.3.1) and already exhibited extensive replication timing alterations (Fig. 4.2a). The proportion of early and middle-early replicating loci increased from 55% in H1p49CM to 64% in H1p47(CDM3) and up to 77% in H1p47(CDM6) cells (Fig. 4.2b). At the individual locus level, *REX1* was amongst the first loci that promptly shifted their replication timing from late to early upon growth culture change

(Fig. 4.3a). Others such as *BMP4* progressively advanced their replication timing over six passages (Fig. 4.3a). To further explore changes in replication timing of the *REX1* locus in its genomic context, several additional sites spanning an approximately 2 Mb region around *REX1* in H1 CM and CDM^{activin} cell populations, were analysed. Late replication was observed across the whole region studied in H1 CM cells. Strikingly, however, in H1 CDM^{activin} cells a domain of early replication centred on the *REX1* locus was observed embedded within the late replicating region (Fig. 4.3b). This data could suggest that localised changes in the chromatin structure and/or histone modifications promote the observed shift of replication, by likely allowing firing of an origin of replication in H1 CDM^{activin} cells. Similar evidence of local regions of early replication at the *Rex1* locus have previously been reported in mES cells (Perry et al. 2004).

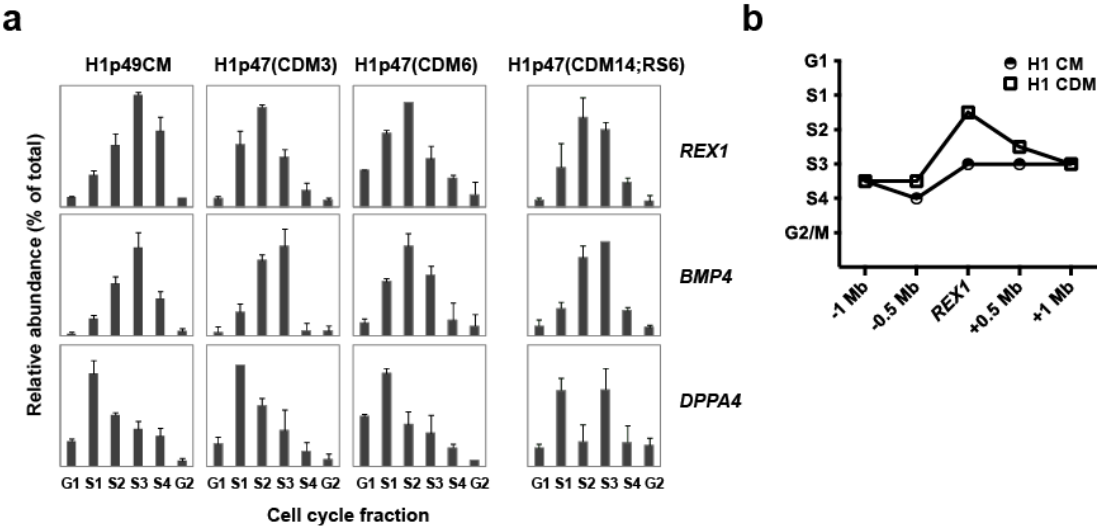


Figure 4.3 Individual loci undergo shifts of replication timing upon growth culture condition change.
(**a**) Replication timing shifts expressed as histograms of individual shifting loci demonstrating the relative abundance of DNA within each cell cycle fraction (G1, S1-S4 and G2 phases). Examples of loci shifting immediately (*REX1*) and gradually (*BMP4*) towards early replication timing profile (CM to CDM^{activin}) and revert back towards late profile upon re-swap (CDM^{activin} to CM). The replication timing profile of *DPPA4* is consistently early across cells grown in CM and CDM^{activin} but becomes unstable and heterogeneous within the same population in the re-swapped conditions, displaying bi-phasic peaks. Mean and standard deviation of at least two experiments are shown for each cell type analysed. (**b**) The observed shifts of replication timing occur at specific region in the loci studied. The *REX1* locus was analysed more carefully for replication timing with primers spanning 2 Mb around the locus. A line plot of the average peak replication timing abundance of two H1 CM and CDM^{activin} samples at the different primer pair positions. A shift towards an early replication (S1-S2) is observed at the *REX1* locus, in CDM^{activin} embedded within a region of late repletion. No similar 'peak' of early replication in CM samples where the whole region studied replicates in late S-phase.

Remarkably, a significant overlap (11 loci) could be observed 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 4.4a and summarised by the

Venn diagram in figure 4.4b. Furthermore, these loci were all previously shown to have variable replication kinetics across H1, H7 and H9 hES cell lines grown in CM yet are consistently replicating earlier during S-phase under CDM^{activin} conditions in both H1 and H9 cells. Notably, the loci that underwent replication timing shift upon culture condition change are all situated in gene poor regions.

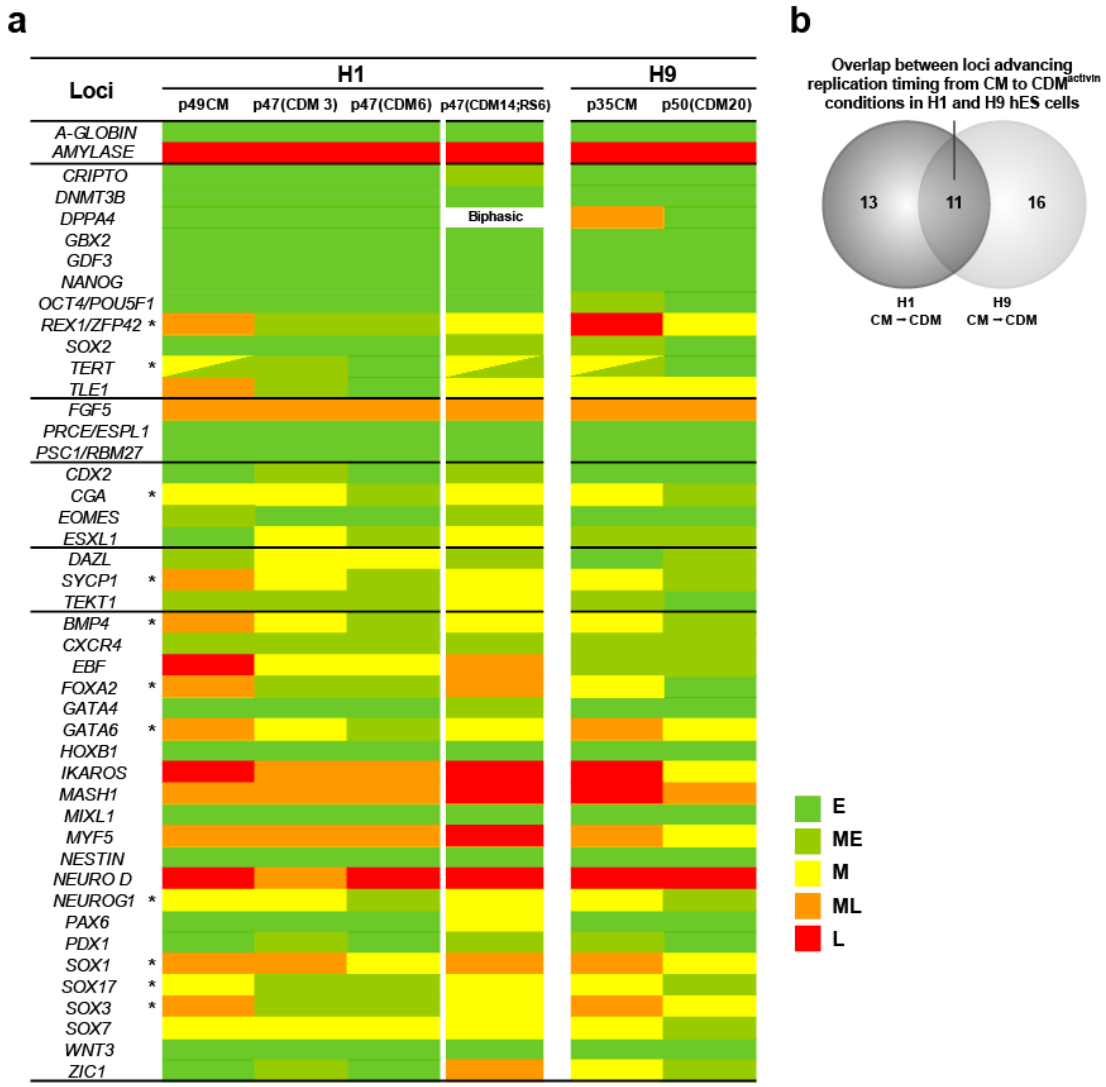


Figure 4.4 Loci shifting their replication timing in CDM^{activin} are shared between different human ES cell lines. (**a**) Replication timing of individual candidate loci was assessed in H1 and H9 hES cells grown in CM conditions and compared to that of cells cultured for different number of passages in CDM^{activin}. Late towards early shifts of replication timing in H1 cells were observed immediately after three passages in CDM^{activin} [H1p47(CDM3)] and became more pronounced after six passages [H1p47(CDM6)]. Re-swapping the H1 cells from CDM^{activin} back to CM conditions lead to some degree or reversal of the replication timing shift in all observed shifting loci [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**) 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.

Significantly, however, modulations of hES cell epigenetic state were demonstrated to be reversible as returning the cells from CDM^{activin} back to CM conditions was sufficient to revert their replication timing to a more unstable, late-replicating profile, as seen in figure 4.2a. The proportion of early and middle-early genes decreased from 77% to 46% in H1 re-swapped [H1p47(CDM14;RS6)] cells), uncovering the epigenetic plasticity of hES cells. Consistently both *REX1* and *BMP4* loci shifted back from middle-early to middle (Fig. 4.3a), as also seen for loci such as *CGA* and *NEUROG1* (Fig. 4.4a). Others, such as *FOXA2* and *SOX1* resumed a late replication profile, as seen in CM culture conditions (Fig. 4.4a). Note that, *DPPA4*, a locus that was consistently early-replicating in H1 hES cells grown in either CM or CDM^{activin} conditions displayed a biphasic replication timing profile with two main peaks in S1 and S3 fractions (Fig. 4.3a, lower panel), underlying the level of instability in re-swapped cultures. Altogether, these data demonstrate that interchangeable subpopulations may exist within hES cell cultures, as delineated by a stable, early-replicating *versus* unstable, late-replicating profiles.

4.2.3 Fully reprogrammed human iPS cells share the same level of epigenetic plasticity as hES cell

Reprogramming somatic cells to an induced pluripotent state holds a promise for applications in regenerative medicine and disease modelling through the generation of patient and disease specific pluripotent cells (Dimos et al. 2008; Park et al. 2008a; Yusa et al. 2011). Human iPS cells can be generated by the forced overexpression of key TFs (Takahashi et al. 2007; Yu et al. 2007). These cells share the hall marks of pluripotent cells, being able to self-renew and give rise to differentiated progenitor cells as well as expressing the key pluripotency factors *OCT4, SOX2* and *NANOG* (Takahashi et al. 2007; Yu et al. 2007). However, iPS cells have been shown to have differences between themselves as well as when compared to hES cells in epigenetic features such as DNA methylation (Newman et al. 2010; Polo et al. 2010; Lister et al. 2011; Nazor et al. 2012; Ruiz et al. 2012). 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 et al. 2009). To investigate whether similar epigenetic alterations, as observed in hES cells, also occur in hiPS cell cultures in response to changing growth conditions, the replication timing profile of a fully reprogrammed hiPS 40 cell line, which was derived from neonatal foreskin fibroblasts by overexpression of OCT4, SOX2, C-MYC and KLF4 (Vallier et al. 2009b), was similarly analysed when grown under CM and CDM^{activin} conditions.

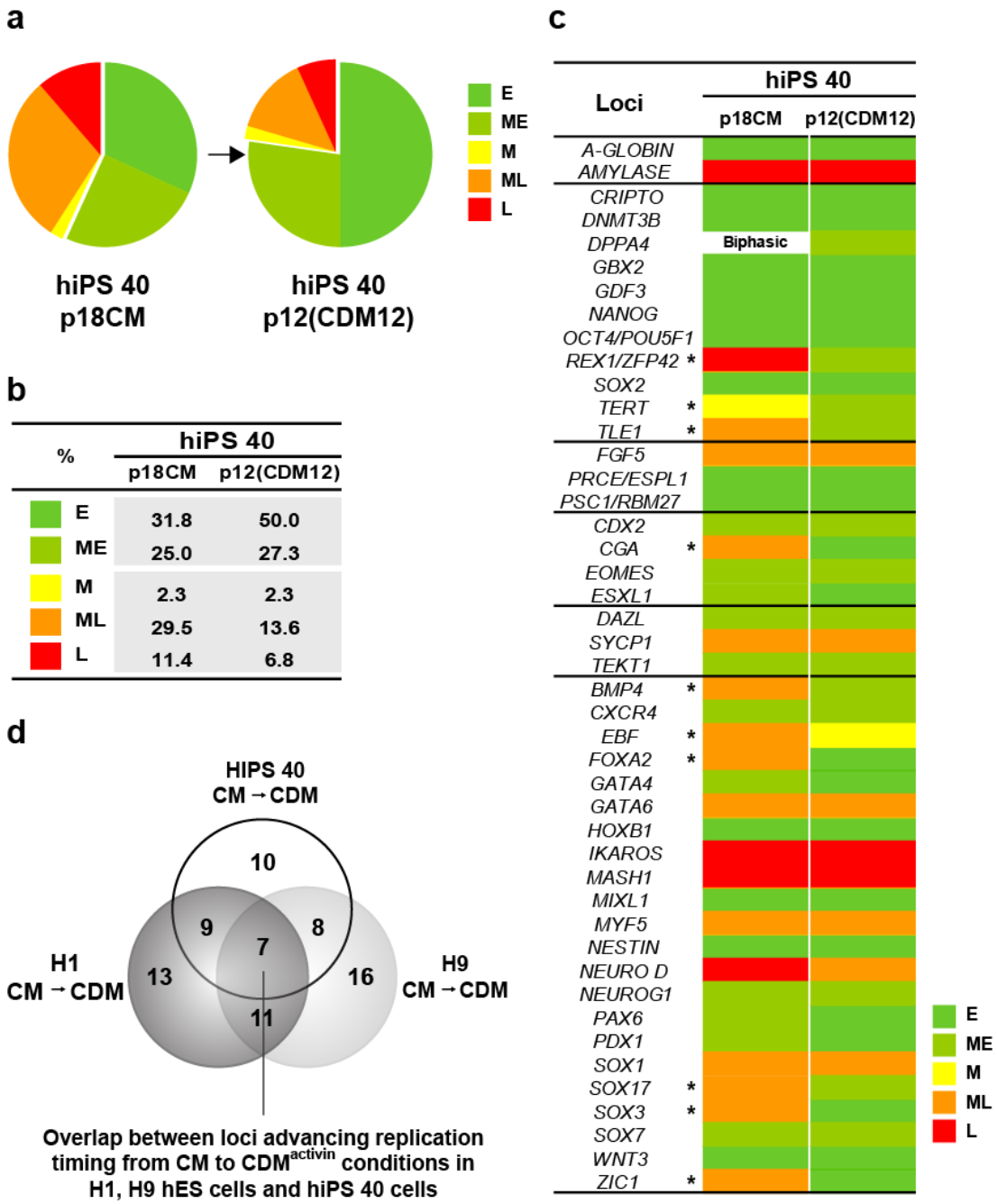


Figure 4.5 Fully reprogrammed human iPS cells display similar shifts of replication timing at candidate loci as observed in hES cells upon change of culture conditions.

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 grown in CDM^{activin} conditions. The proportion of candidate loci replicating early (E), middle-early (ME), middle (M), middle-late (ML) and late (L) in S-phase are represented (**a**) as pie charts or (**b**) as percentage of total candidate loci analysed. (**c**) The replication timing profiles of individual candidate loci assessed in hiPS 40 cells comparing CM (hiPS 40 p18CM) and CDM^{activin} [hiPS p12(CDM12)] cultured cells. Loci shifting their replication timing profile towards early in CDM^{activin} conditions (*REX1, TERT, TLE1, CGA, BMP4, EBF, FOXA2, SOX17*, *SOX3* and *ZIC1*) are marked by asterisks. *DPPA4* was 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 (*REX1, TERT, CGA, BMP4, FOXA2, SOX17* and *SOX3*) 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.

As previously shown for H1 and H9 hES cells, the hiPS cells showed contrasting replication timing profiles when grown under different growth conditions (Fig. 4.5a and b). Human iPS cell replication timing profiles shifted dramatically from late under CM conditions to a more early profile 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 (Fig. 4.5b). The loci shifting their replication timing 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 (Fig. 4.5c & d), these shared loci are of different types and do not show any clear bias towards a specific lineage. 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.

4.3 Phenotypic and functional characterisation of H1 hES cells grown under CM and CDM^{activin} conditions

4.3.1 H1 hES cells cultured in CDM^{activin} conditions retain their undifferentiated state

Following transfer from CM to CDM^{activin} culture conditions, H1 hES cells readily stabilised and once acclimatised grew in tightly packed flat colonies, with little or no evidence of spontaneous differentiation (as sometimes observed at the edges of H1 colonies cultured in CM conditions) (Fig. 4.6a). Further phenotypic analysis of H1 CDM^{activin} cells by AP, IF and FACS analysis confirmed the undifferentiated state of these cultures. When assayed for AP activity H1 CDM^{activin} cells exhibited strong positive staining (pink) across colonies (Fig. 4.6b). Expression of cell surface markers TRA-1-60, SSEA4 and TRA-1-81, associated with undifferentiated ES cells, was highly and homogeneously detected in H1 CDM^{activin} cells by IF and/or FACS staining, whereas SSEA1, a surface marker upregulated upon differentiation, was not observed by IF staining (Fig. 4.6c & d). Furthermore H1 hES cells stained positive for key pluripotency associated factors OCT4, NANOG and SOX2 by both IF (all) and FACS (OCT4 & NANOG) analysis (Fig. 4.6c &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. These results confirm that H1 hES cells can assume distinct epigenetic states while maintaining an undifferentiated state and are similar to the results obtained when analysing H1 cells grown in CM conditions (see Fig. 3.6).

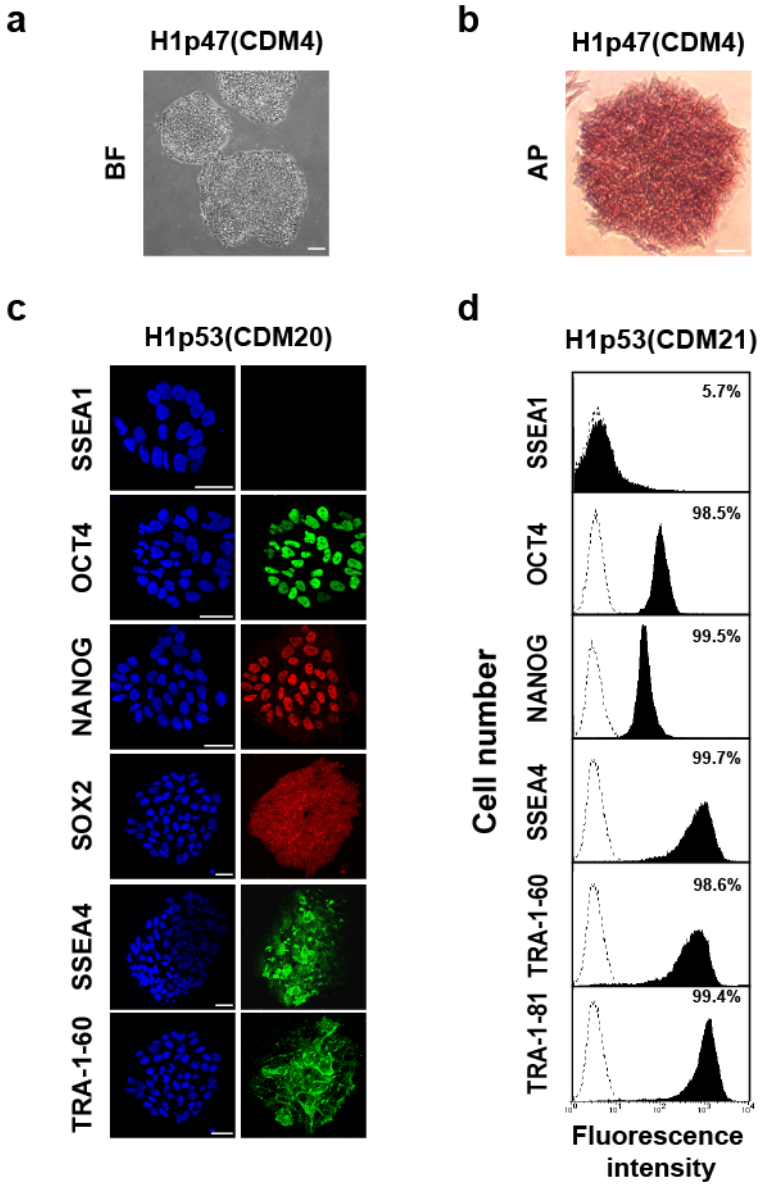


Figure 4.6 Human ES cells transferred from CM to CDM^{activin} conditions retain hallmarks of undifferentiation.

(**a**) Bright field (BF) picture of representative undifferentiated H1 hES cell colonies and (**b**) positive staining for alkaline phosphatase (AP) activity (pink) in cells grown under CDM^{activin} conditions. Scale bars, 70 µm. (**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 µm.

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

To determine whether H1 CM and CDM^{activin} cells are pluripotent, the cells were assayed using embryoid body (EB) differentiation assay. In this assay ES cells are induced to spontaneously form cell aggregates in suspension then allowed to differentiate into ectoderm, endoderm and mesoderm –like, a process that mimics early embryonic development (Doetschman et al. 1985; Leahy et al. 1999). Here, EBs were formed by growing H1 CM or CDM^{activin} cells in non-adherent conditions in differentiation media (Knockout™ DMEM supplemented with 20% FBS, 0.1 mM NEAA, 2 mM L-glutamine and 0.55 mM β-mercaptoethanol) for 10 days (Fig. 4.7a) and then plating the cells back onto gelatin-coated plates for further differentiation up to 20 days (Fig. 4.7b). After differentiation for up to 30 days, a wide variety of morphologically different cell types were observed in the culture plates, likely indicating the presence of neuronal-like cells with extended filaments (Fig. 4.7b 1 & 2), liver-like cells with well-defined cytoplasms and cell bodies (Fig. 4.7b 3 & 4) and endothelial-like cells with extended and flattened morphology (Fig. 4.7b 5 & 6). To further corroborate the morphological evidence, qRT-PCR analysis of differentiation markers was used. Relative gene expression levels of markers for ectoderm (*SOX1, P63* and *β-TUBULIN*) endoderm (*AFP, PDX1* and *GATA6*) and mesoderm (*BMP4, CD45* and *PECAM1*) were all upregulated in both H1 CM and CDM^{activin} samples after 30 days of differentiation (Fig. 4.7c). Furthermore expression levels of pluripotency markers (*OCT4, NANOG* and *CRIPTO*) were all reduced to minimal levels in both H1 CM and CDM^{activin} cells as compared to the undifferentiated samples, indicating a loss of pluripotent identity (Fig. 4.7c). Altogether, these results indicate that H1 CM and CDM^{activin} hES cells can undergo differentiation into derivatives of all the three germ layers.

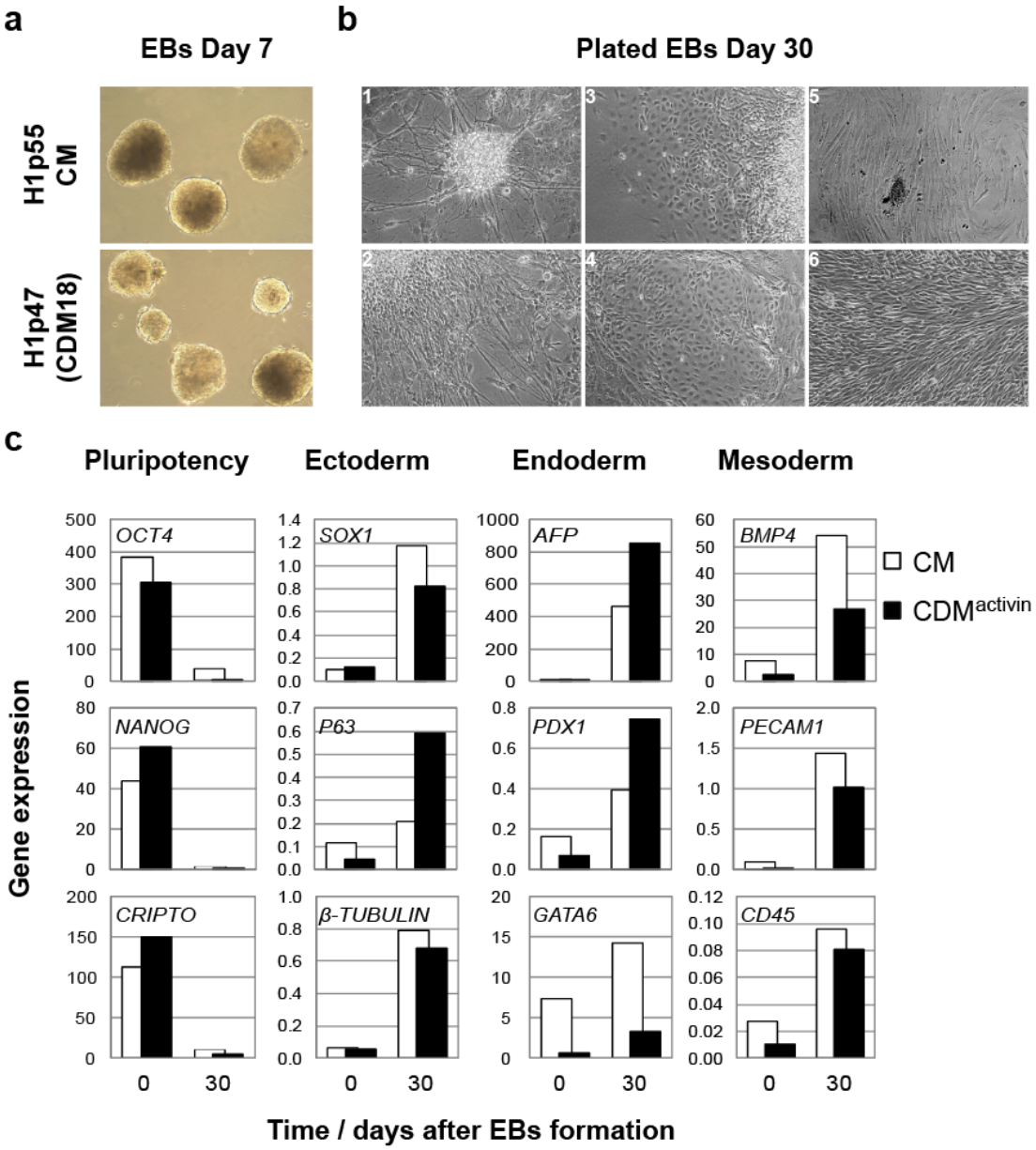


Figure 4.7 Human ES cells from both CM and CDM^{activin} conditions are able to form derivatives of all three germ lineages in an *in vitro* assay of differentiation.

(**a**) Morphology of embryoid bodies (EBs) formed in non-adherent conditions for 7 days as well as (**b**) 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. (**c**) Generation of derivatives of the three germ layers was further confirmed by qRT-PCR analysis looking at *SOX1, P63, B-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 qRT-PCR analysis of three independent experiments. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Primers used for this analysis are listed in Appendix I, Table A1

4.3.3 H1 CDM^{activin} conditions favour an ICM-like phenotype within hES cell cultures

To further highlight possible phenotypic similarities and/or differences among H1 hES cells grown in CM and CDM^{activin} conditions, the expression patterns of several pluripotency and early differentiation markers were analysed by qRT-PCR in H1 hES cells in CM and upon transfer into CDM^{activin} growth conditions at the indicated passage numbers (3, 6, 14 and 25) (Fig. 4.8). As shown in figure 4.8, *OCT4* and *SOX2* remained expressed largely at similar levels upon growth condition changes. In contrast *NANOG* expression was transiently, yet significantly upregulated upon transfer from CM to CDM^{activin} conditions (Fig. 4.8a),

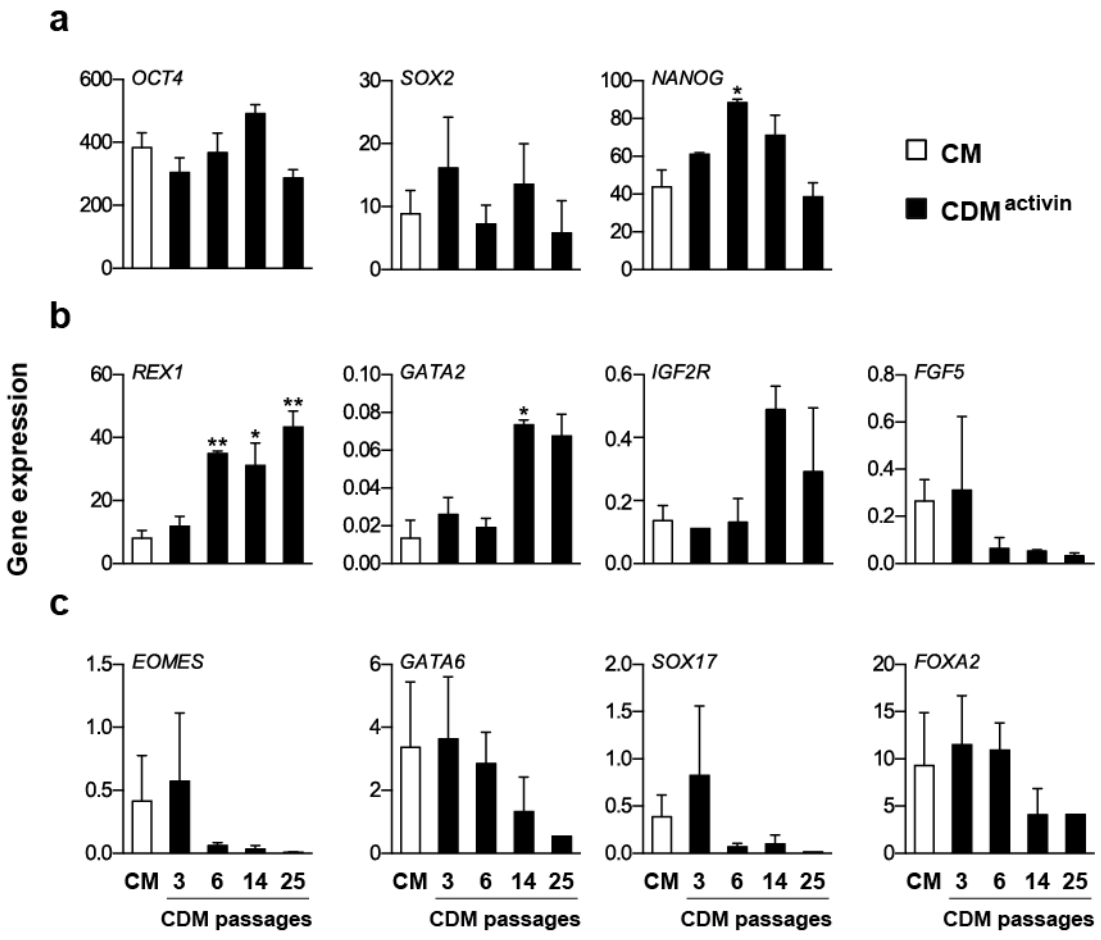


Figure 4.8 H1 hES cells cultured in CM and CDM^{activin} have distinct expression patterns of key marker genes. Real-time PCR analysis of selected genes in H1 CM (H1p49CM, white bars) and cells transferred into CDM^{activin} conditions and cultured for different passage numbers [H1p47(CDM3, 6, 14 and 25), black bars]. (**a**) Pluripotency-associated genes (*OCT4* and *SOX2*) stably expressed in CM and CDM^{activin} cells or showing a transient increase (*NANOG*) in CDM^{activin} cells. (**b**) Continued culture in CDM^{activin} strongly induced in H1 hES cells the expression of preimplantation-epiblast/ICM markers (*REX1, GATA2* and *IGF2R*) and downregulation of the postimplantation marker *FGF5*, as well as the (**c**) downregulation of germ layer markers (*EOMES, GATA6, SOX17* and *FOXA2*). Mean and standard error of at least two experiments are shown. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Un-paired two tailed t-test (* = $P < 0.05$ and ** = $P < 0.01$). Primers used for this analysis are listed in Appendix I, Table A1.

furthermore this transient increase was also observed at the protein level (Appendix II Fig. A6).

Remarkably, a robust and rapid induction of *REX1* messenger RNA (mRNA) levels was observed. *REX1* is selectively expressed in both mouse and human blastocysts (Rogers et al. 1991; Pelton et al. 2002; Reijo Pera et al. 2009) in early development (Fig. 4.8b). Similarly two other genes associated with human ICMs, *GATA2* and *IGF2R* (Reijo Pera et al. 2009), were also induced to some extent by culture in CDM^{activin} conditions highlighting a possible shift towards a more ICM-like phenotype (Fig. 4.8b). In addition the epiblast specific marker gene *FGF5* (Pelton et al. 2002; Tesar et al. 2007) was downregulated following transfer to CDM^{activin} conditions (Fig. 4.8b). Furthermore, markers of early differentiation (*EOMES*, *GATA6*, *SOX17* and *FOXA2*) were downregulated upon increasing passages in CDM^{activin} conditions as compared to expression levels observed in cells cultured in CM conditions (Fig. 4.8c). Taken together, these data support the view that hES cells can adopt distinct yet still pluripotent states in response to changing culture conditions, with CDM^{activin} conditions favouring a stable epigenetic state with a possibly ICM-like phenotype.

4.4 Investigating the molecular mechanisms underlying distinct epigenetic states in hES cells

4.4.1 Increased TGF-β/Activin signalling in CDM^{activin} conditions

Recent reports have indicated that the TGF-β superfamily ligands are important to sustain an undifferentiated state in hES cells. In particular TFG-β/Activin signalling through the signal inducer SMAD2/3 was shown to be strictly required for maintaining hES cell pluripotency (James et al. 2005; Vallier et al. 2005a; Xiao et al. 2006; Xu et al. 2008; Vallier et al. 2009a). *NANOG* has been identified as a direct target of TGF-β/Activin signalling (Xu et al. 2008; Vallier et al. 2009a) and interestingly was found transiently upregulated upon transfer to CDM^{activin} at the transcriptional level in H1 hES cells (Fig. 4.8a). The expression pattern of *NANOG* was mirrored by that of three TGF-β members (*CRIPTO*, *GDF3* and *NODAL*) (Appendix II Fig. A7). This directly points to a change in the cell signalling environment of hES cells when grown under CM conditions and transferred into CDM^{activin} conditions, and provides a plausible explanation for changes in *NANOG* expression levels.

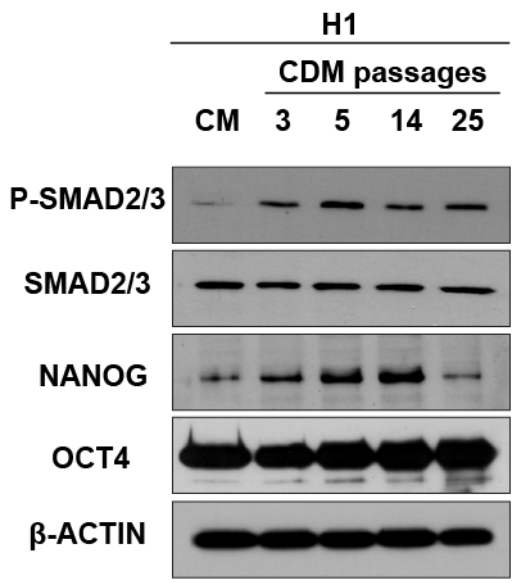


Figure 4.9 H1 cells cultured in CDM^{activin} conditions have increased levels of phosphorylated SMAD2/3 complexes when compared to cells cultured in CM.
Western blot analysis of H1 hES cells cultured in CM (H1p49CM) and CDM^{activin} conditions for increasing passages [H1p47(CDM3, 5, 14 and 25). Membranes were probed with antibodies specific for phosphorylated (P) - SMAD2/3, SMAD2/3, NANOG, OCT4 and β-ACTIN (as a control for protein loading).

To further explore whether differential levels of TGF-β/Activin signalling indeed operate on H1 hES cultures, levels of phosphorylated SMAD2/3 (P-SMAD2/3) were compared between CM and increasing passages (3, 5, 14 and over 25) of CDM^{activin} H1 hES cultures by western blot analysis (Fig. 4.9). Clear differences in the levels of P-SMAD2/3 detected in CM and CDM^{activin} cell cultures were observed. Active P-SMAD2/3 levels became consistently increased in hES cells when transferred from CM to CDM^{activin} conditions. This increase was also accompanied by a transient increase in the levels of NANOG protein in CDM^{activin} cell cultures. While the increase in NANOG levels was transient, levels of active P-SMAD2/3 remained high in CDM^{activin} conditions, regardless of passage number. Note that no change was observed in levels of Oct4 protein between CM and CDM^{activin} cultures. Of interest, NANOG was recently shown to interact directly with the SMAD2/3 proteins, thus modulating the activity of TGF-β/Activin signalling pathway (Vallier et al. 2009a). This negative feedback loop, could in turn, modulate the levels of *NANOG* expression and thus autoregulate its own expression levels in hES cell cultures. It would be interesting to quantify the levels of SMAD2/3 occupancy at the Nanog promoter in CM and early and late CDM^{activin} cultures by ChIP assay to further investigate this hypothesis.

4.4.2 Changes in TGF-β/Activin signalling impact on the expression levels of P300 histone acetyltransferase in hES cells

TGF-β/Activin signalling through its inducer SMAD2/3 was previously shown to recruit various histone acetyltransferases including P300/CBP (Pouponnot et al. 1998; Ross et al. 2006), 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. 2008) leading to local chances in chromatin accessibility. To further investigate a possible link between Smad2/3-associated

chromatin remodelling activity and the acquisition of an early-replicating profile in hES cells mRNA levels for each of these epigenetic modulators were assessed by qRT-PCR (Fig. 4.10). Interestingly, *P300* levels were significantly increased upon culture condition change and the level of induction was maintained in CDM^{activin} cultures as assessed over 25 passages (Fig. 4.10). In contrast, a transient induction of both *CBP* and *BRG1*, was observed, while *PCAF* and *GCN5* mRNA levels were very low in all cell populations and/or not significantly altered respectively.

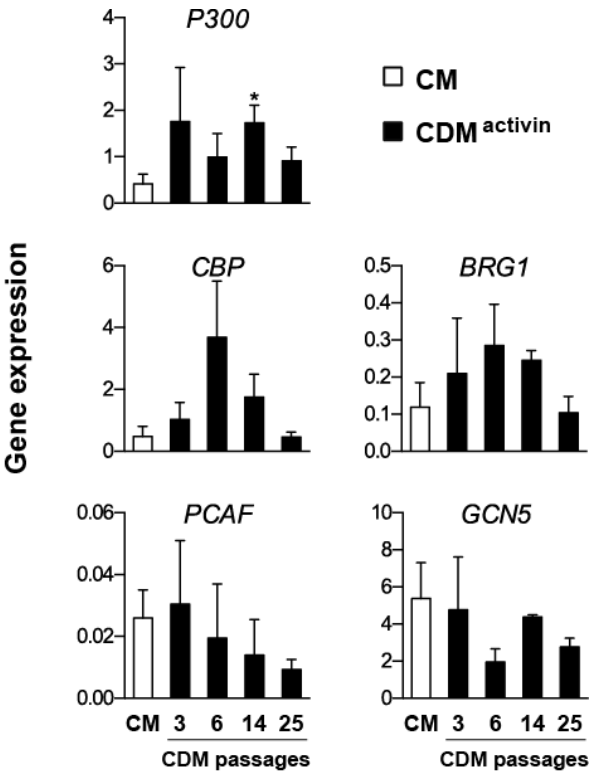


Figure 4.10 Histone acetyltransferase P300 expression is significantly increased in H1 hES cells cultured in CDM^{activin} conditions as compared to CM conditions.
(**a**) Real-time PCR analysis of histone acetyl transferases (*P300*, *CBP, PCAF,* and *GCN5*) and chromatin modifier *BRG1*, all known to be associated with phosphorylated SMAD2/3 complexes. H1 CM (H1p49CM, white bars) and CDM^{activin} hES cells [H1p47(CDM3, 6, 14 and 25), black bars] analysed. Mean and standard error of at least experiments are shown. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Primers used for this analysis are listed in Appendix I, Table A1. Un-paired two tailed t-test (* = P < 0.05).

Whether *P300* gene expression levels directly respond to TGF-β/Activin cell signalling activity was further tested in H1 hES cells cultured in CDM^{activin} conditions. Here, untreated cells were compared to cells starved of TGF-β/Activin signalling by treatment with SB431542, a commonly used TGF-β/Activin-receptor inhibitor, followed by restimulation with Activin A. Interestingly, inhibition of TGF-β/Activin signalling led to a significant reduction of approximately 30% in the expression levels of *P300* after 12 hours of SB431542 treatment. Upon restimulation, the expression of *P300* returned to levels seen in untreated cell samples (Fig. 4.11a). Levels of *NANOG*, a known target of TGF-β/Activin signalling (Xu et al. 2008), were also observed to effectively respond to TGF-β signalling inhibition (by approximately 50%) and restimulation as expected, however expression levels of *CBP* did not significantly vary in the same experiments (Fig. 4.11a).

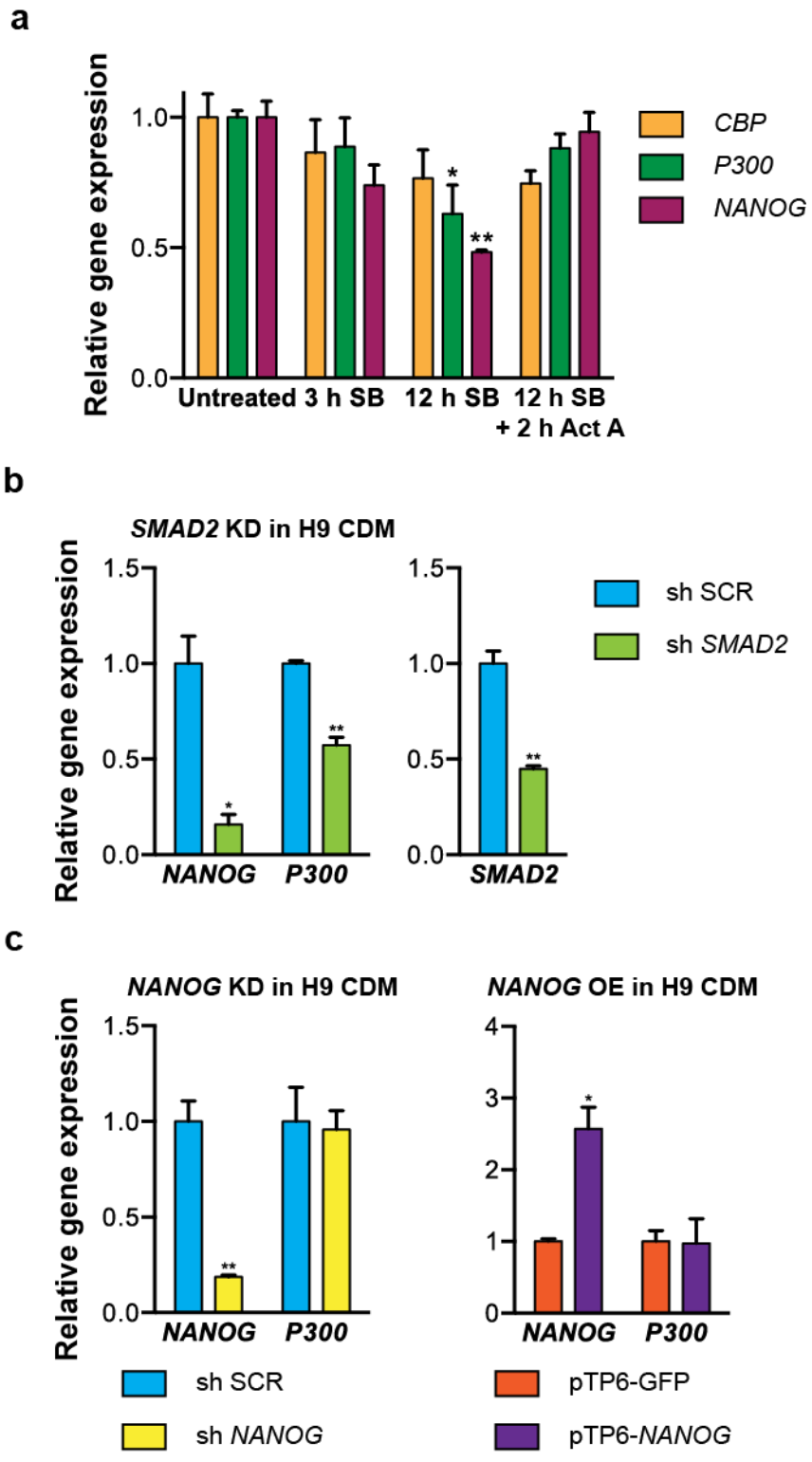


Figure 4.11 Levels of P300 in hES cells respond to SMAD2 activity.
Real-time PCR analysis of selected genes in H1 and H9 ES cells cultured in CDM^{activin} conditions. (**a**) H1 hES cells grown in CDM^{activin} conditions (Untreated) were compared to cells starved of Activin A supplementation and treated with TGF−β/Activin signalling inhibitor SB431542 for three (3 h SB) or twelve (12 h SB) hours or first starved for twelve hours and then washed with PBS and re-supplied with normal CDM^{activin} medium for two hours (12 h SB + 2 h Act A). Relative expression levels of *CBP*, *P300* and *NANOG* were assessed and are presented normalised to untreated samples. (**b**) Relative expression levels of *NANOG*, *P300* and *SMAD2* were

assessed in H9 hES cells with stable knockdown of *SMAD2* and compared to cells expressing non-targeting scramble shRNA, a kind gift from Dr. Ludovic Vallier (Brown et al. 2011). (**c**) Relative expression levels of *NANOG* and *P300* were assessed in H9 hES cells with stable knockdown of *NANOG* or overexpressing *NANOG* and compared to cells expressing non-targeting scramble shRNA or *GFP*, a kind gift from Dr. Ludovic Vallier (Vallier et al. 2009a). Data normalised to either clones expressing non-targeting scramble shRNA or overexpressing GFP. Mean and standard error of three independent experiments are presented. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Un-paired two tailed t-test (* = P < 0.05 and ** = P < 0.01). Primers used for this analysis are listed in Appendix I, Table A1.

Recent studies using ChIP-sequencing revealed the presence of binding sites for both NANOG (Kunarso et al. 2010) and SMAD2/3 (Brown et al. 2011) upstream and downstream of the human *P300* coding sequence respectively. The observed modulations in *P300* expression levels in response to changes in TGF-β/Activin signalling could be directly due to SMAD2/3 modulating *P300* levels or indirectly via increased *NANOG* levels. To investigate this question, previously established H9 hES cell lines with knockdown of *SMAD2* as well as overexpression and knockdown of *NANOG* were utilised. These were compared to cells expressing non-targeting scramble shRNA or overexpression vector with GFP sequence. All kind gifts from Dr Ludovic Vallier (Vallier et al. 2009a; Brown et al. 2011). H9 cells with a knockdown of *SMAD*2 also exhibited a significant reduction in the levels of both *NANOG* and *P300* (Fig. 4.11b). In contrast neither knockdown nor overexpression of *NANOG* in H9 CDM^{activin} cells affected the levels of *P300* expression (Fig. 4.11c). Altogether, these results indicate that levels of *P300* can be directly modulated in response to TGF-β/Activin signalling activity most likely via SMAD2/3 in hES cells.

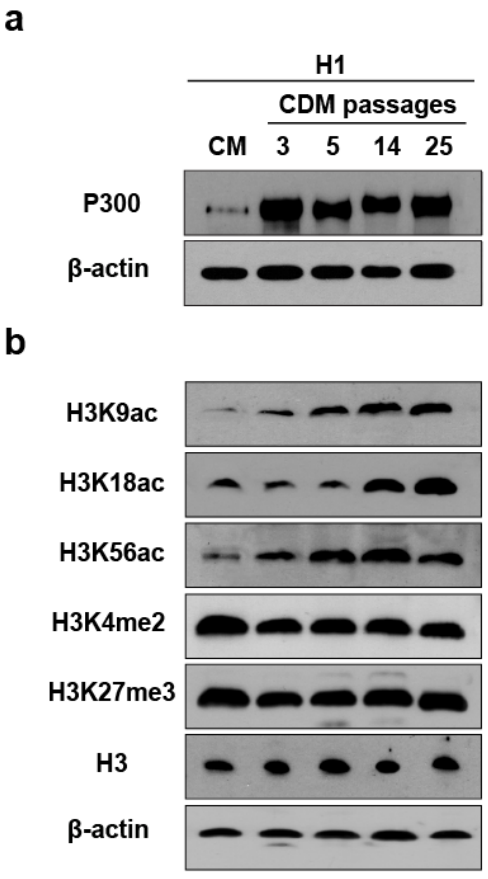


Figure 4.12 12 P300 protein levels are maintained and global acetylation levels are increased at P300/CBP preferred target residues with increasing passages in CDM^{activin} conditions.
(**a**) Western blot analysis of H1 hES cells cultured in CM (H1p49CM) and CDM^{activin} for increasing number of passages [H1p47(CDM3, 5, 14 and 25). Membranes were probed with antibodies specific for P300 and (**b**) for p300/CBP 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 loading controls.

4.4.3 Global increase of P300-mediated histone H3 acetylation in H1 CDM^{activin} cell cultures

P300 induction upon growth culture changes was confirmed in hES cells at the protein level by Western blot analysis (Fig. 4.12a) and was furthermore shown to be accompanied by a global increase of its HAT activity (Fig. 4.12b). For this analysis, P300 preferentially acetylated nucleosomal histone H3 residues (H3K9ac, H3K18ac and H3K56ac) (Ross et al. 2006; Das et al. 2009) were analysed, 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 unchanged between CM and CDM^{activin} cells. These results demonstrate that hES cells grown under CDM^{activin} conditions are characterised by an enhanced expression of SMAD2/3-associated HAT P300 and a global histone H3 hyperacetylated state as compared to cells in CM conditions.

4.4.4 Recruitment of active P300 complexes at target loci is enhanced in response to CDM^{activin} growth conditions

Histone acetylation positively correlates with early DNA replication timing (Méndez 2009). To investigate whether P300 is directly recruited to loci subject to late-to-early replication timing changes in hES cells, ChIP assay and qRT-PCR analysis were used. In particular, the levels of P300 occupancy were compared in H1 CM and CDM^{activin} hES cell cultures at selected promoters as well as the levels of enrichment for associated acetylated histone H3 (H3K9ac, H3K18ac and H3K56ac). Note that, *CGA, FOXA2, GATA6*, *REX1, SOX17* and *TERT* are all candidate targets for replication timing alterations in response to growth culture condition change (Fig. 4.13a). *NANOG* is here used as a control, being consistently early replicating in both CM and CDM^{activin} populations. Interestingly, and as shown in figure 4.13b, P300 was detected in the promoter regions of these genes at higher levels in CDM^{activin} hES cells, where these loci replicate in early S-phase as compared to in CM cells. In contrast, no significant change in the levels of enrichment for P300 at *NANOG* promoter was observed between the two studied cell populations. As anticipated from the detection 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 cells (Fig. 4.13c). The *TERT* locus 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

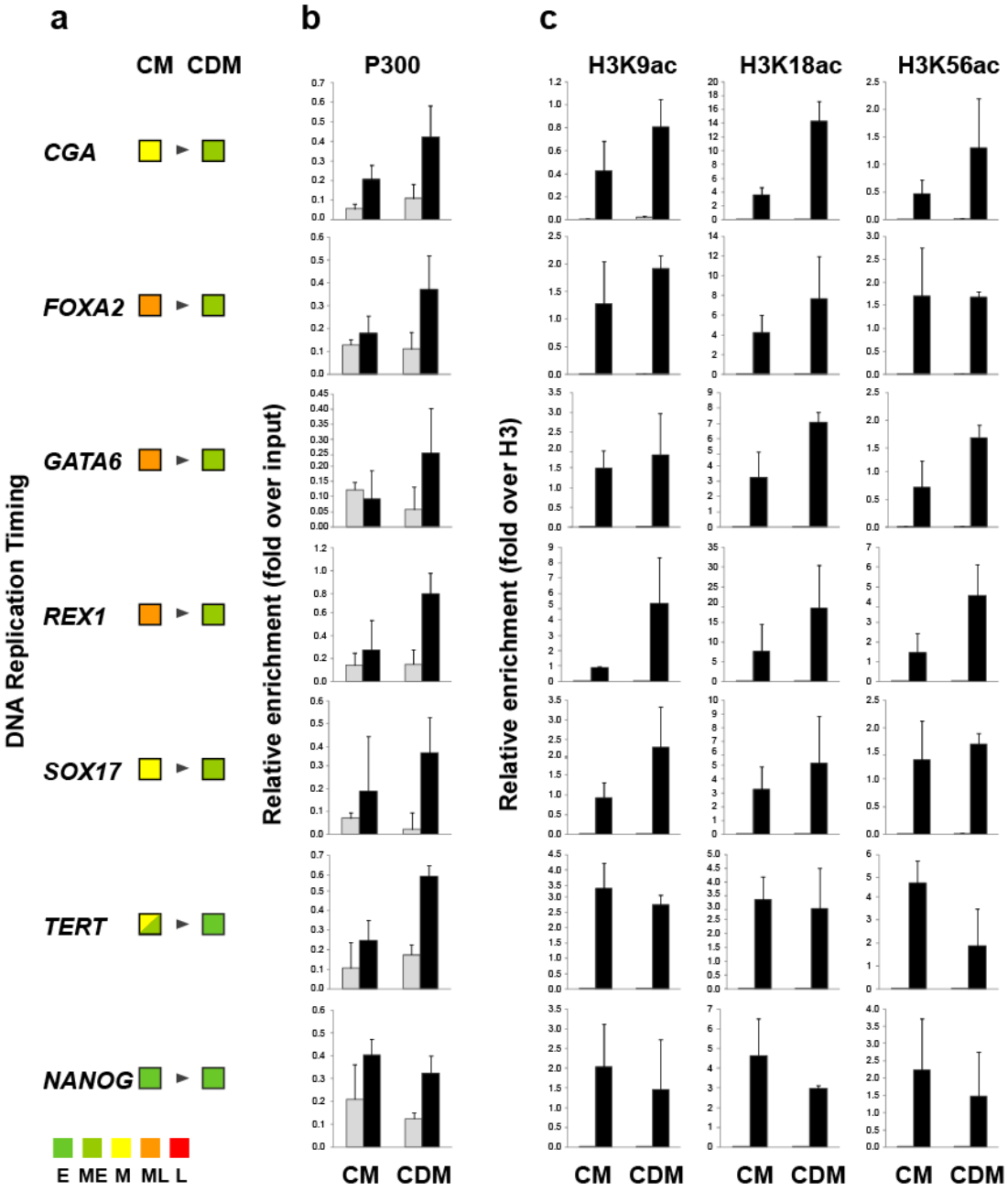


Figure 4.13 Loci changing replication timing from late in CM to early in CDM^{activin} hES cells harbour higher levels of P300 and histones acetylation at their promoter regions in CDM^{activin} hES cells.
(**a**) Replication timing status of *CGA, FOXA2, GATA6, REX1, SOX17, TERT* and *NANOG* was assessed 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 qRT-PCR analysis (black bars). Grey bars indicate background levels (control IgG). P300 enrichments are expressed relative to input and histones data normalised to total histone H3. Error bars indicate standard deviations of at least two independent experiments.

direct correlation between the establishment of early-replication profiles and P300 loci; *REX1* and *TERT* are two examples of genes that 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 showing that HATs recruitment to the human β-*GLOBIN* locus in a non-erythroid cell line (where this gene is not expressed) 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 it is normally expressed) was sufficient to change its replication timing from early to late (Goren et al. 2008).

4.4.5 Stable knockdown of P300 in H1 hES cells leads to differentiation and gradual loss of cultures

To further investigate the role of P300 in the maintenance of distinct epigenetic states in hES cells, *P300* expression was stably knocked down in hES cells using two independent puromycin-selectable shRNA (shP300 3 and shP300 4) vectors and further analysed alongside cell lines expressing a non-targeting shRNA control (shSCR) (see sections 2.2.2.2 and 2.2.2.4 for details). Briefly, here H9 hES cell cultures were transfected by specified shRNA vectors using Lipofectamine 2000. Approximately one week after transfection puromycin selection was applied to the cells. Cells were then cultured for approximately 2-3 weeks to allow for establishment of clones, at this point 12 individual clones were picked for each transfected vector for expansion (see Fig. 4.14a for scheme of transfection). Four clones transfected with shP300 3 and five of shP300 4 exhibited knockdown of *P300* expression levels and were used for analysis. Both P300 shRNAs were found to effectively reduce *P300* gene expression by approximately 60% as compared to shSCR controls with no downregulation of the closely related HAT *CBP* (Fig 4.14b). Note that *CBP* expression seemed instead to increase in *P300*-depleted hES cells, possibly pointing to compensatory effects (Fig. 4.14b). A dramatic loss of P300 was observed at the protein level by western blot analysis (Fig. 4.14c). In all cases the *P300* knockdown clones could not, however, be maintained and propagated beyond four or five passages after establishment of clones. The *P300* knockdown clones became either overcome by differentiated cells or seemingly stopped dividing leading to a loss of undifferentiated cultures, whereas clones with the non-targeting shSCR remained positive for the undifferentiation associated AP activity and continued to grow in normal hES colonies (Fig. 4.14d). Further qRT-PCR analysis revealed a reduction in the levels of *OCT4* and *NANOG* in the *P300* knockdown clones from the beginning suggesting a bias towards differentiation in the shP300 clones (Fig. 4.14e). Taken together, these data demonstrate that stable P300-depletion is detrimental to hES cells, as previously shown for e.g. human fibroblasts (Prieur et al. 2011).

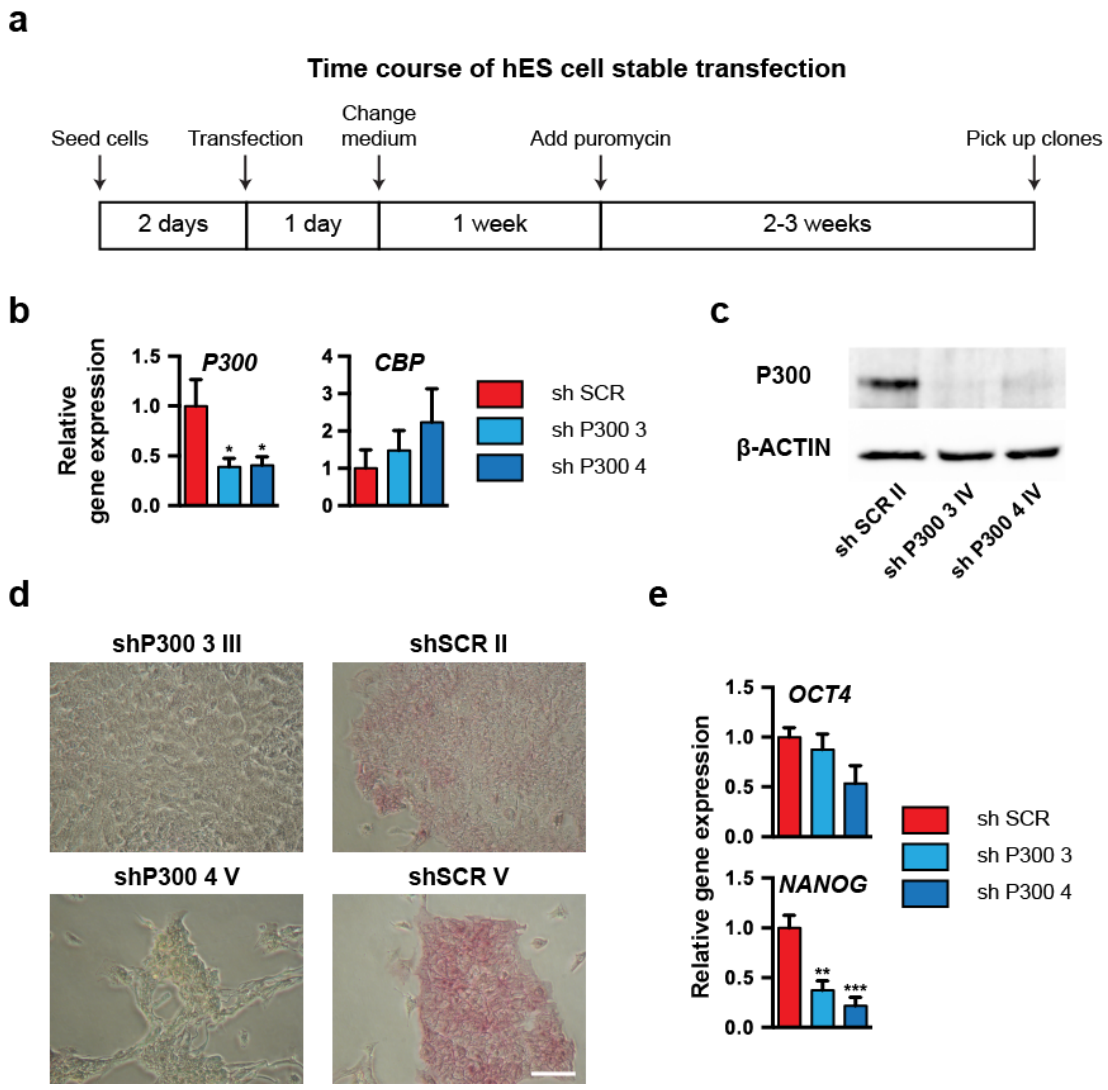


Figure 4.14 Stable knockdown of *P300* leads to differentiation or loss of cultures in H9 hES cells.
Clones expressing either non-targeting shRNA (shSCR) or one of two independent shRNAs targeting *P300* (shP300 3 and shP300 4) were established in H9 hES cells. (**a**) Schematic representation of hES stable transfection time-course. (**b**) Real-time PCR analysis confirms knockdown of *P300*. Levels of *P300* and *CBP* were assessed in shP300 3 and shP300 4 clones and compared to the shSCR clones. (**c**) Western blot analysis of H9 hES cells expressing shSCR, shP300 3 or shP300 4 cultured in CDM^{activin} conditions. Membranes were probed with antibodies specific for P300 and β-ACTIN (as a control for protein loading). (**d**) Representative images of shP300 3, shP300 4 and two different shSCR colonies stained for alkaline phosphatase (AP) activity cultured for 4, 5 or 7 passages after establishment respectively. Positive staining (pink) for AP activity is evident in colonies with normal hES morphology for shSCR clones, but is however completely absent in both shP300 3 and shP300 4 clones. Scale bar 25 μm. (**e**) Real-time PCR analysis of *OCT4* and *NANOG* in shP300 3 and shP300 4 clones and compared to the shSCR clones. For qRT-PCR analyses mean and standard error of three independent shSCR, four independent shP300 3 and five independent shP300 4 clones averaged are presented. Data was normalised to shSCR and to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Un-paired two tailed t-test (* = $P < 0.05$, ** = $P < 0.01$ and *** = $P < 0.001$). Primers used for this analysis are listed in Appendix I, Table A1. All the P300 knockdown and scramble control samples analysed here were collected between two to five passages after picking of the clones. The cells were transfected at passage H9p81 in CDM^{activin} conditions.

4.4.6 Transient knockdown of P300 by RNA interference does not impact on the replication timing profile of H1 CDM^{activin} hES cells

P300 proved to be strictly necessary for the maintenance of self-renewing hES cell cultures. To further explore the role of P300 in defining distinct epigenetic states within hES cell cultures, hES cells were transiently targeted with either a pool of siRNAs targeting *P300* (siP300) or non-targeting siRNA control (siSCR). A double transfection protocol (day 0 and day 2) was applied to H1 CDM^{activin} cells (Fig. 4.15a). The extent of the knockdown was then validated at day 4 by qRT-PCR (Fig. 4.15b), showing that *P300* expression was effectively reduced by approximately 60%, with no significant effect on *CBP*, *SIRT1* [an HDAC known to target same residues as P300 (Das et al. 2009)], *OCT4* or *NANOG* mRNA levels compared control cells. Note that the time of sample collection no obvious differences in colony morphology or cell death were observed between siSCR and siP300 cultures (Fig. 4.15c). Western blot analysis confirmed a partial loss of P300 at the protein level, while in contrast SIRT1 or OCT4 expression remained unchanged, confirming undifferentiated status of *P300*-depleted hES cultures. However despite the clear reduction of P300 in the siP300 samples, only a slight decrease in the bulk levels of P300 associated acetylated histone H3 (H3K9ac, H3K18ac and H3K56ac) was observed (Fig. 4.15d). When siSCR and siP300 H1 CDM^{activin} cells were next assayed by DNA replication timing analysis at day 4, post initial transfection, both cell populations were found to retain a mostly early replicating profile (Fig. 4.15e). No alterations could be observed at loci previously identified as targets for early-to-late replication timing changes, upon growth culture condition change (Fig. 4.15e; top panel) as well as at loci that consistently replicated within the same time windows during S-phase in both CM and CDM^{activin} cells (Fig. 4.15e; bottom panel). Taken together, these results indicate that transient knockdown of P300 is not sufficient to induce replication timing changes in H1 CDM^{activin} cells. This could possibly be due to unknown compensatory effects or due to insufficient time for epigenetic alterations to occur.

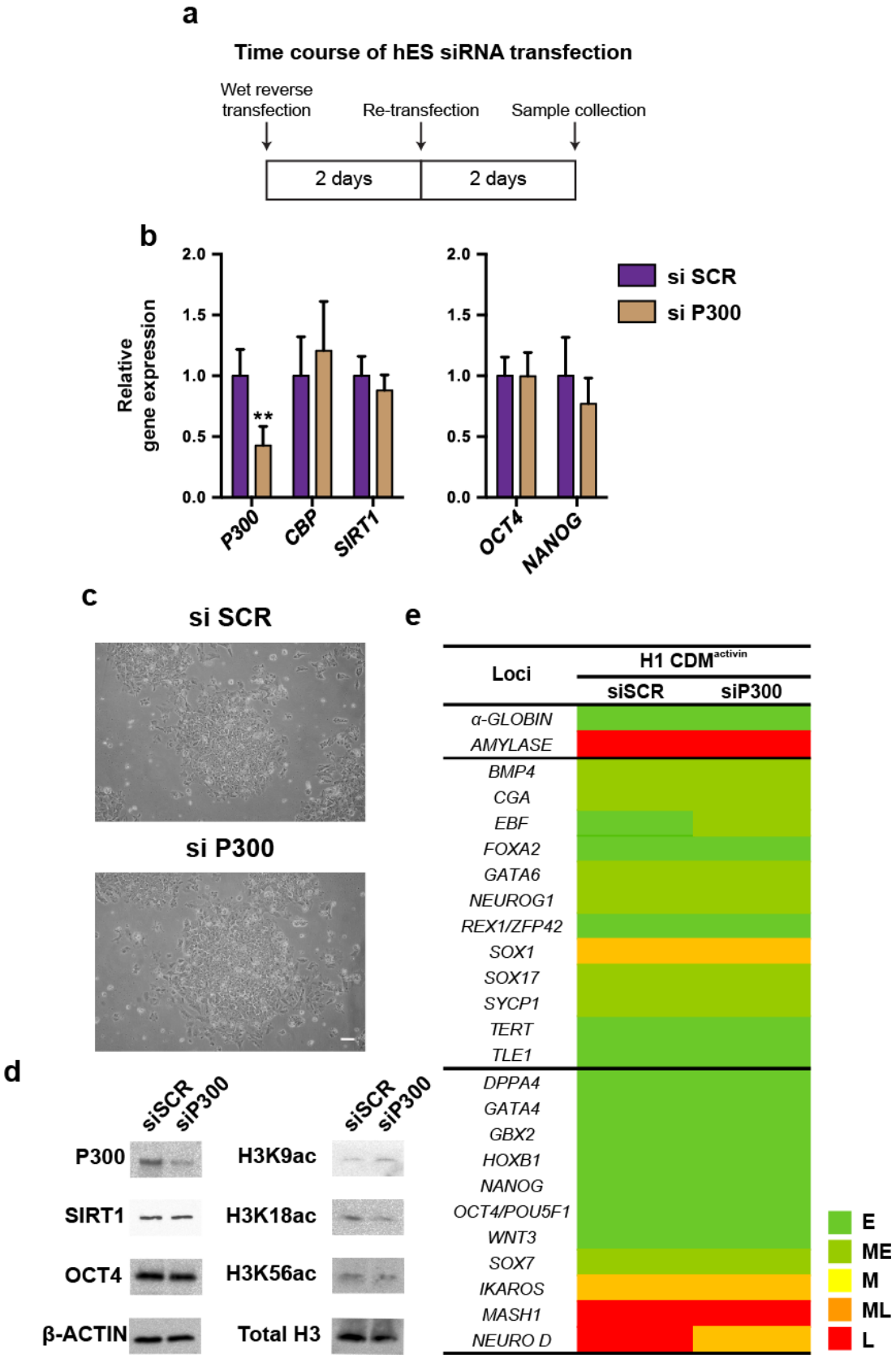


Figure 4.15 Transient siRNA induced knockdown of *P300* does not affect the replication timing profiles of H1 CDM^{activin} cultured cells.

H1 hES cells cultured in CDM^{activin} conditions were double transfected with either non-targeting siRNA pool (siSCR) or siRNA pool targeting P300 (siP300), first by reverse wet-transfection at day 0 and a second time at

day 2 by normal transfection. (**a**) Schematic representation of hES siRNA double-transfection. (**b**) Double knockdown by siP300 induces approximately 60% reduction of *P300* expression levels as compared to siSCR-transfected controls. Quantitative RT-PCR analysis of *P300* and the closely related histone acetyltransferase *CBP* and histone deacetylase *SIRT1*, known to act on the same histone targets as P300, and of pluripotency markers *OCT4* and *NANOG*. Mean and standard error of three independent experiments are presented. Data was normalised to the levels of siSCR samples and to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Un-paired two tailed t-test (** = P < 0.01). Primers used for this analysis are listed in Appendix I, Table A1. (**c**) Representative bright field images of H1 hES colonies transfected with siSCR and siP300. Scale bar 10 μm. (**d**) Western blot analysis of siSCR and siP300 transfected H1 CDM^{activin} cell populations and analysed after four days of culture. Membranes were probed with antibodies specific for P300/CBP, SIRT1, OCT4 and β-actin (as a control for protein loading) or with antibodies specific for p300/CBP preferentially acetylated histone H3 targets (H3K9ac, H3K18ac and H3K56ac) and total H3. Equivalent protein loading is shown by the total H3 and β-actin detection. (**e**) The replication timing of individual loci compared between siSCR and siP300 transfected H1 CDM^{activin} cell populations and analysed after four days of culture. No effect on replication timing was observed in 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*) or other loci that 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}. Data representative of a single experiment with a technical duplicate.

4.4.7 Inhibition of P300/CBP histone acetyltransferase activity leads to early to late replication timing changes in H1 CDM^{activin} hES cells

To assess whether inhibiting P300 HAT activity might be a valuable, alternative approach, the specific P300/CBP inhibitor curcumin (also known as diferuloylmethane), a major curcumanoid in the spice turmeric, (Balasubramanyam et al. 2004) was used. A range of curcumin concentrations (10, 25 and 50 μM) were first tested in H1 CDM^{activin} cells to assess the impact of the drug on global histone H3 acetylation levels upon 24 hour treatment by Western blot analysis. Acetylation levels for all histone H3 residues analysed were abrogated (H3K9ac and H3K56ac) or significantly reduced (H3K18ac) in curcumin-treated CDM^{activin} hES cells with 10 and 25 μM curcumin, thereby confirming P300-dependent histone H3 acetylation in hES cells. 50 μM concentration of curcumin resulted, however, in a high incidence of cell death, as previously reported (Balasubramanyam et al. 2004). In contrast, treatment with curcumin did not induce any change in H3K27me3 levels, an unrelated histone modification, further highlighting the specificity of this treatment (Fig. 4.16a). To exclude any incidence of differentiation in curcumin-treated H1 CDM^{activin} cells, the expression of undifferentiated ES-associated genes (*OCT4, NANOG* and *REX1*) was verified by qRT-PCR analysis following four days of 10 μM treatment, showing similar levels in both H1 CDM^{activin} untreated and treated cells (DMSO and curcumin). Differentiation-associated markers (*FOXA2, GATA6, EOMES* and *CDX2*) were also analysed and found not to be significantly different between the three conditions (Fig. 4.16b). Cells treated for four days with 10 μM curcumin were also assayed by IF, revealing homogeneous staining for OCT4 and NANOG, while the differentiation associated surface marker SSEA1, was absent in all cell populations tested (Fig. 4.16c).

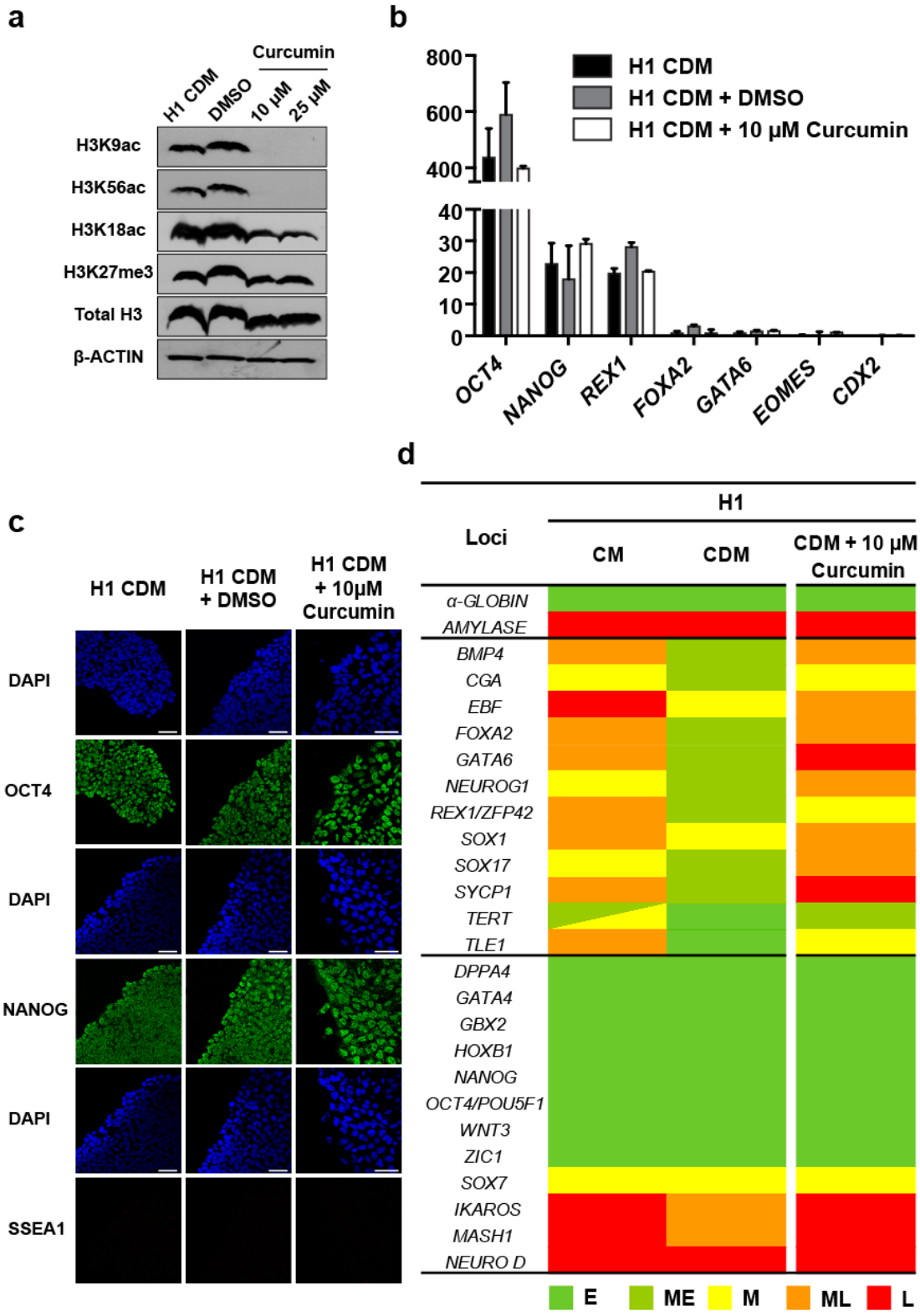


Figure 4.16 Inhibition of P300/CBP histone acetyltransferase activity by Curcumin leads to a reduction of histone acetylation levels and induces early to late replication timing changes in H1 CDM^{activin} cells.
(**a**) Western blot analysis of H1 CDM^{activin} hES cells treated for 24 hours with DMSO (solvent control), 10 and 25 µM of P300/CBP 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/CBP inhibition. Equivalent protein loading is shown by the total H3 and β-actin detection. (**b**) Quantitative RT-PCR analysis of the expression of pluripotency-associated (*OCT4, NANOG* and *REX1*) and differentiated-associated (*FOXA2, GATA6, EOMES* and *CDX2*) genes in untreated H1 CDM^{activin} hES cells as well as cells treated for 4 days with DMSO or 10 µM Curcumin.

Mean and standard error of at least two independent experiments are presented. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Primers used for this analysis are listed in Appendix I, Table A1. (**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. (**d**) 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 μM 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 μM curcumin.

Interestingly, inhibiting P300/CBP activity with 10 μM curcumin treatment for four days was sufficient to significantly alter the replication timing profile of H1 CDM^{activin} cells. In particular, 11 out of 13 loci previously identified as shifting their replication timing patterns, from late to early, when transferred from CM to CDM^{activin} conditions (*BMP4, CGA, EBF, FOXA2, GATA6, NEUROG1, REX1, SOX1, SOX17, SYCP1* and *TLE1),* were consistently found to replicate in the second half of S-phase upon P300/CBP HAT activity inhibition in H1 CDM^{activin} cell cultures (Fig. 4.16d). Most of these loci started to display a replication timing similar to that previously observed in H1 CM cells. 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/CBP activity (Fig. 4.16d), highlighting the target specificity of P300/CBP action. Altogether, these results suggest a key role for P300/CBP-mediated histone acetylation in defining the distinct epigenetic states in hES cell cultures as reflected at the DNA replication timing level.

4.5 Conclusions

Phenotypic and functional heterogeneity has been previously reported in mES cell cultures (Chambers et al. 2007; Hayashi et al. 2008; Toyooka et al. 2008) and it has been indicated that mES cell populations harbour ES cells in naïve primed states (Nichols et al. 2009b). Heterogeneity has also been observed in hES cell cultures with regards to variable expression patterns (Fischer et al. 2010) and very recently for DNA methylation (Tompkins et al. 2012). Here replication timing analysis was used to demonstrate that pluripotent hES cells reversibly adopt distinct epigenetic states when cultured under different cell culture conditions and/or in different laboratories. In particular, dynamic replication timing alterations, from late – to – early were observed when transferring the cells from CM-medium, conditioned on irradiated MEFs and supplemented with Fgf2, to CDM-medium, a chemically defined medium supplemented with Activin A and Fgf2. These alterations were

accompanied by changes in gene expression patterns, with genes linked specifically with human ICM and pre-implantation epiblast being upregulated.

Importantly, a role for the HAT P300 in stably maintaining an early replicating chromatin conformation in hES cells cultured in $CDM^{activin}$ conditions was uncovered. The late to early replication timing alterations in hES cells in response to changing culture conditions were accompanied by enrichment of P300 occupancy at shifting loci as assessed by ChIP. Inhibition of P300/CBP HAT activity by the chemical inhibitor curcumin in $CDM^{activin}$ hES cells was sufficient to revert their replication timing profiles back to a CM-like profile, strongly suggesting that P300 might play a key role in determining the epigenetic status of hES. Furthermore, *P300* expression was found to be directly modulated by TGF-β/Activin signalling in H1 $CDM^{activin}$ cells. Taken together, these findings support the view that distinct epigenetic states exist within hES cell cultures, in response to different environmental cues.

Chapter 5 Characterisation of *REX1/ZFP42* overexpressing human ES cells

5.1 Introduction

Upon changing growth culture conditions from CM to CDM^{activin} media extensive epigenetic alterations were observed in H1 hES cell's replication timing profile. This was accompanied by dynamic fluctuations in gene expression patterns. Though only a limited panel of genes were analysed in these experiments (see Fig. 4.8 and 4.10), among these *NANOG* was found to be transiently upregulated when transferring cells into CDM^{activin} conditions. Interestingly, a significant and steady increase in the level of *REX1* expression was also observed in CDM^{activin} hES cell cultures.

Rex1 gene was first discovered in F9 embryonal carcinoma cells due to its rapid downregulation upon differentiation (Hosler et al. 1989). The *Rex1* gene has arisen through a retrotransposition of *Yy1* and is only present in placental mammals, furthermore DNA-binding zinc-finger domains share homology with those of *Yy1*, but have been shown to bind slightly different target sequences suggesting somewhat divergent roles for the two proteins (Kim et al. 2007). Recent evidence has demonstrated roles for *Rex1* in imprinting and X-inactivation in mouse development (Kim et al. 2011; Gontan et al. 2012). *Rex1* expression has been confirmed in many pluripotent stem cell types including undifferentiated mES and hES cells (Rogers et al. 1991; Henderson et al. 2002), mouse mesenchymal stem cells derived from adult bone marrow (Jiang et al. 2002) as well as *in vivo* in the ICM of early mouse (Rogers et al. 1991) and human (Vassena et al. 2011) embryos. *Rex1* is widely used as a marker of pluripotency both in mouse and human cells (Guallar et al. 2012) and high levels of *Rex1* expression have been linked to increased self-renewal ability and pluripotency in mES cells (Toyooka et al. 2008; Wray et al. 2010; Wray et al. 2011; Marks et al. 2012) as well as complete reprogramming of human iPS cells (Chan et al. 2009). Further insights into the function of *Rex1* in mES cells by means of knockdown and overexpression experiments have, however, produced conflicting results. *Rex1* knockdown has been reported to induce differentiation in mES cells (Zhang et al. 2006; Scotland et al. 2009) and overexpression to destabilise the ES cell self-renewing state (Zhang et al. 2006). Conversely other studies have indicated that *Rex1* is dispensable for the pluripotent state *Rex1*-null mES cells only showing mild defects in the induction of some endoderm markers upon differentiation (Masui et al. 2008). Recently *Rex1* has also been reported to regulate the expression of endogenous

retroviruses in both mES cells and *in vivo* mouse embryos (Guallar et al. 2012). Hence, the exact role played by *Rex1* in both mES and especially hES cells remains largely unclear. The observed upregulation of *REX1* upon growth culture change in H1 hES cells therefore prompted further investigations to examine the possible role of *REX1* in regulating hES cell self-renewal and differentiation in response to cell signalling cues.

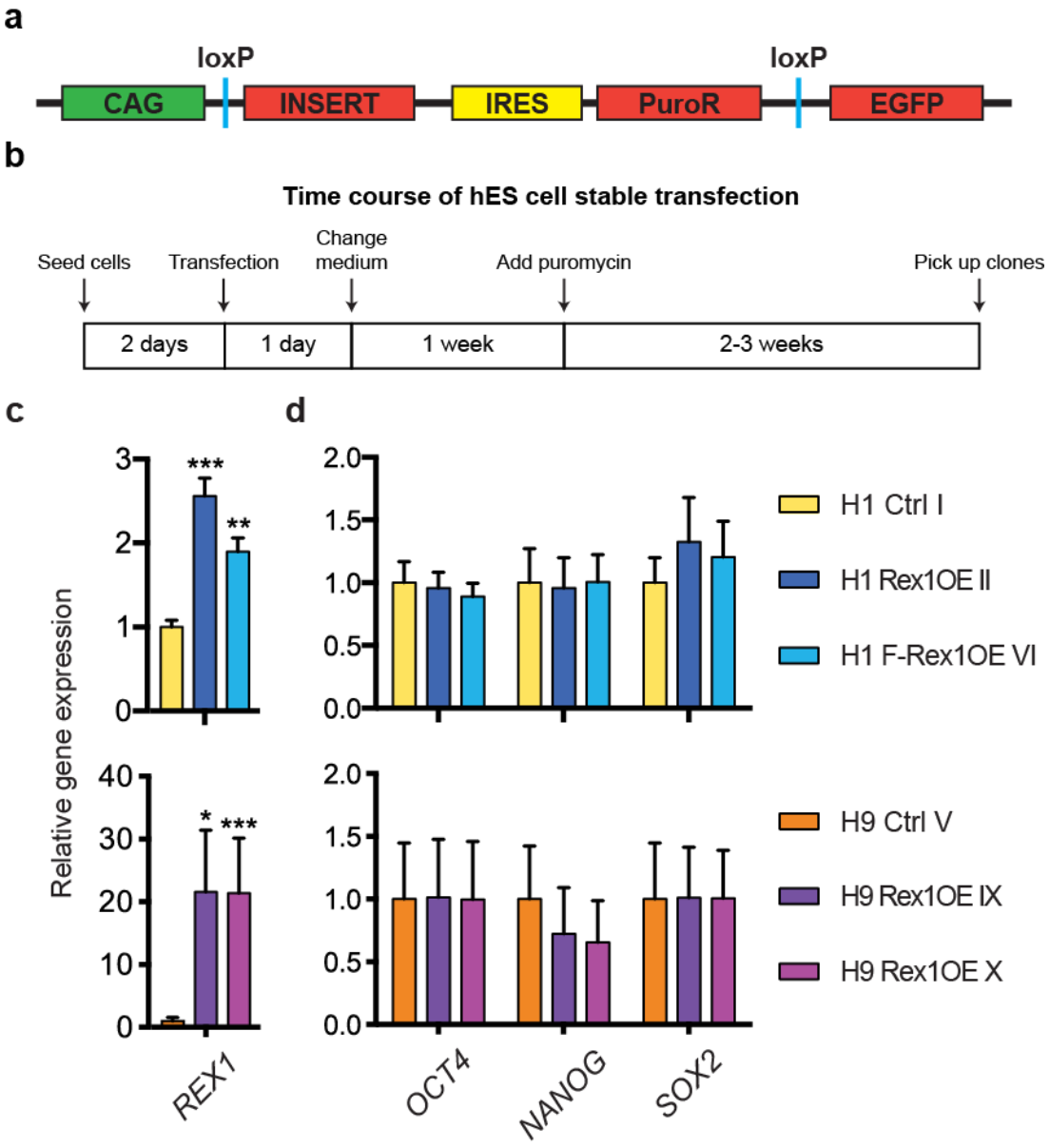


Figure 5.1 Establishment of *REX1* overexpressing (Rex1OE) hES cell clones.
(**a**) Schematic representation of linearised transgene with CAG-promoter, insert, internal ribosomal entry site (IRES), Puromycin resistance gene, enhanced GFP (EGFP) and loxP-sites specified. (**b**) Schematic representation of hES cell stable transfection time-course. (**c-d**) Real-time PCR analysis of Rex1OE and control hES cells for (**b**) *REX1* and (**c**) *OCT4*, *NANOG* and *SOX2*. The two H1 and two H9 Rex1OE clones examined showed significant upregulation of total *REX1* mRNA levels with no significant change in the levels of *OCT4*, *NANOG* and *SOX2* as compared to control cells. Rex1OE clones were established in H1p47ic8+78 and H9p94 grown in $CDM^{activin}$ conditions. Mean and standard error of three independent experiments are shown. Data was normalised to two housekeeping genes (*GAPDH* and *HPRT*) and then to the expression level of the control sample in each cell type. Un-paired two tailed t-test (* = $P < 0.05$, ** = $P < 0.01$ and *** = $P < 0.001$). Primers used for this analysis are listed in Appendix I, Table 1.

5.2 Establishment of *REX1* overexpressing human ES cell clones

To gain further insight into the role of *REX1* in hES cells, *REX1* was constitutively expressed in H1 and H9 hES cell lines. For this analysis, cells were transfected with pPyfloxed expression vector containing either *REX1* or Flag-*REX1* cDNA, or with empty vector as control. Briefly, H1 and H9 hES cell cultures were transfected with the overexpression constructs using Lipofectamine 2000 (see section 2.2.2.4). Approximately one week after transfection puromycin selection was applied to the cells. Cells were then cultured for approximately 2-3 weeks to allow for establishment of clones. Two and six individual H1 hES cell clones overexpressing *REX1* (Rex1OE) and FLAG-*REX1* (F-Rex1OE) were established respectively. Five clones expressing the empty vector, were also picked for further expansion. Eight individual H9 hES cell clones overexpressing *REX1* and seven clones carrying an empty vector (Ctrl), were similarly expanded (see Fig. 5.1a for scheme of transfection). One clone of each H1 Rex1OE II and F-Rex1OE VI as well as two H9 Rex1OE IX and X were selected for subsequent characterisation and analysis. All cells transfected and cultured for these experiments were grown in CDM medium with Activin A and FGF2 supplementation, unless otherwise specified. Furthermore, cells were propagated under puromycin selection to ensure maintained expression of the transgene.

To assess the overexpression level of *REX1* in different clones, qRT-PCR analysis was performed confirming that the total level of *REX1* mRNA in all established clones was significantly higher than in control cells. In particular, *REX1* was found expressed at approximately 2-fold and 20-fold higher levels in H1 and H9 Rex1OE clones, respectively (Fig. 5.1b). In contrast, the core pluripotency genes *OCT4, SOX2* and *NANOG* (Boyer et al. 2005; Loh et al. 2006) retained similar expression levels as control cells (Fig. 5.1c). Protein levels of REX1 were not analysed in absence of specific anti-REX1 antibodies, despite extensive efforts to identify one. Uniform protein expression patterns for OCT4, SOX2 and NANOG were also observed by immunofluorescence in both Rex1OE and control cells (Fig. 5.2a). Phenotypic analysis of the clones was extended to TRA-1-60 and TRA-1-81, two known markers of pluripotent cells (Henderson et al. 2002), as well as OCT4 using flow cytometry (Fig. 5.2b). H1 control and F-Rex1OE cells expressed these pluripotency markers at similar levels with 90-95% of cells staining positive for TRA-1-60 and TRA-1-81 and approximately 62% staining positive for OCT4. Consistently, H9 control and Rex1OE cells expressed OCT4 at similar levels with approximately 73% staining positive. Slightly lower numbers of H9 Rex1OE cells stained positive for TRA-1-60 and TRA-1-81 than control cells with only approximately 79% and 76% staining positive for TRA-1-60 and TRA-1-81 compared to approximately 85%

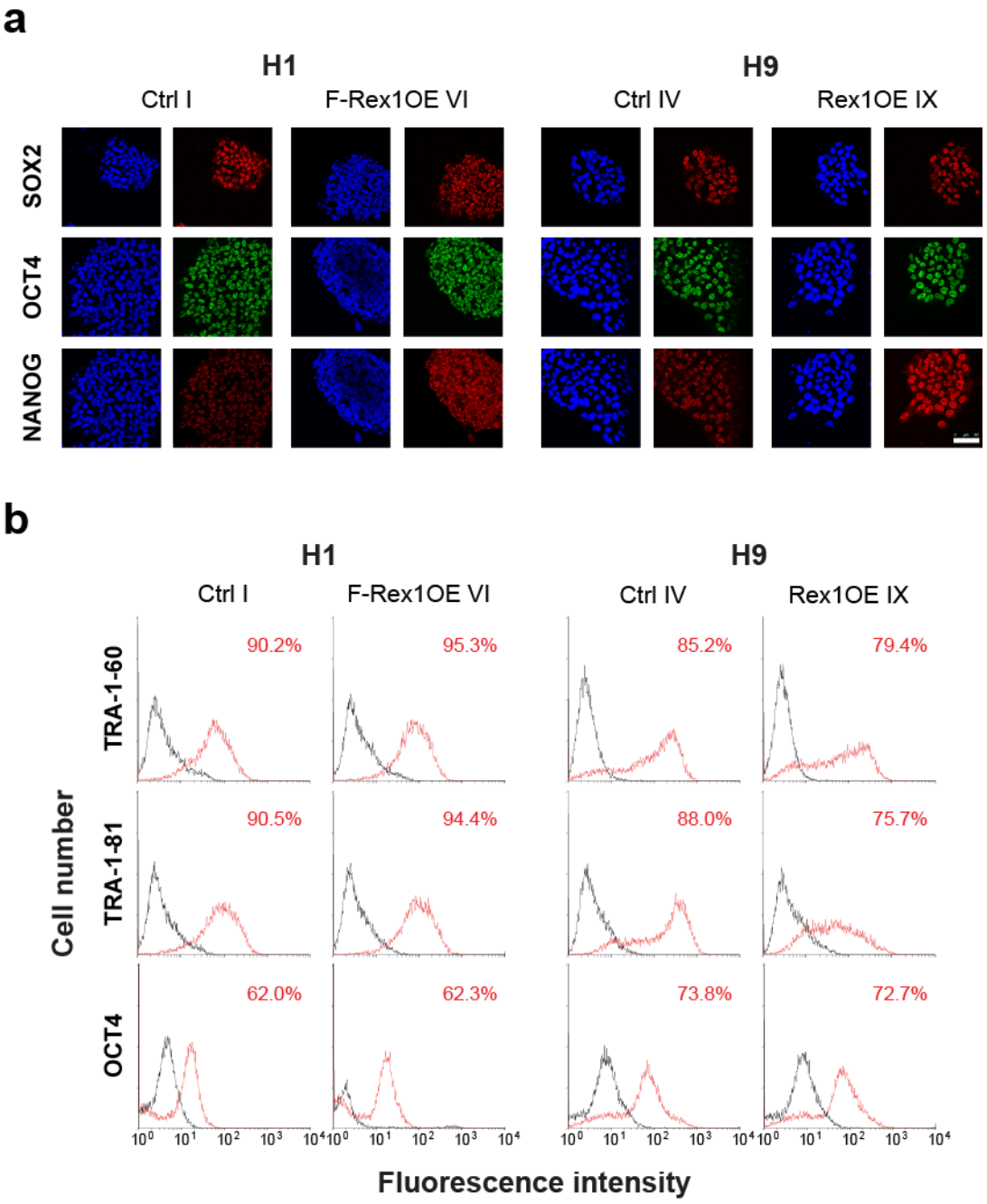


Figure 5.2 *REX1* overexpressing cells retain expression of pluripotency markers in culture.
(**a**) Control and Rex1OE cells established from H1 and H9 cell lines were analysed for the expression of pluripotency markers SOX2, OCT4 and NANOG by immunofluorescence. Nuclei were stained with DAPI. Positive staining was observed for all markers. These images are representative of a single experiment. Scale bar, 50 μm. (**b**) Flow cytometry analysis for H1 and H9 control and Rex1OE cells with TRA-1-60, TRA-1-81 and OCT4. H1 control (Ctrl IV) and Rex1OE (F-Rex1OE VI) cells express these pluripotency markers at similar levels with 90-95% of cells being stained positive for TRA-1-60 and TRA-1-81 and approximately 62% for OCT4, as indicated in histograms. H9 control (Ctrl IV) and Rex1OE (Rex1OE IX) cells stained for OCT4 at similar levels with approximately 73% staining positive. Slightly lower numbers of H9 Rex1OE cells stained positive for TRA-1-60 and TRA-1-81 than control cells (79% & 76% versus 85% & 88%). The profile of positively stained cells (red) was compared to (black) cells stained with an isotype-matched control. Histograms representative of at least three independent experiments for H9 and of a single experiment for H1. Antibodies and dilutions used for this analysis are listed in sections 2.1.1 and 2.1.2

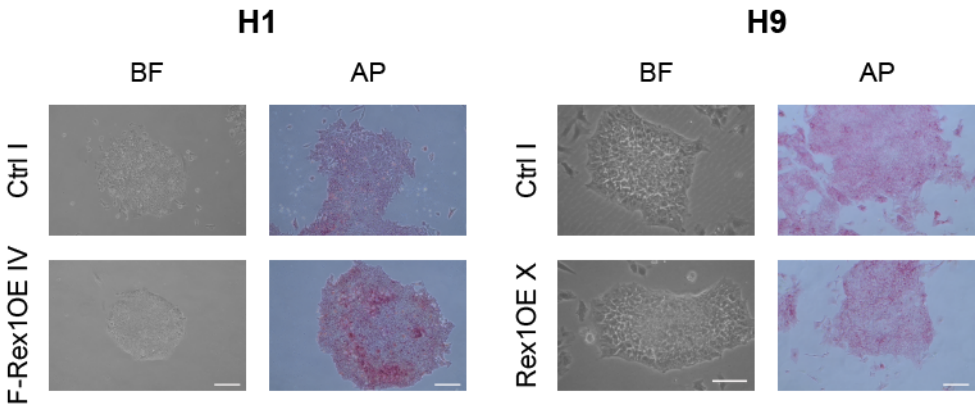


Figure 5.3 Rex1OE hES cells retain normal morphology and stain positive for alkaline phosphatase activity.
Representative bright field (BF) images depicting normal hES cell morphology with flat and well defined colonies and positive staining (pink) following fixation and alkaline phosphatase (AP) treatment for both control and Rex1OE cells established from H1 and H9 cells. AP images are representative of at least three independent experiments. Magnification 20x for H9 BF images and 10x for all others.

and 88% for control cells respectively. Both Rex1OE and control cells similarly formed typical hES cell colonies with a flattened compact morphology (Thomson et al. 1998) and stained positive when assayed for alkaline phosphatase activity, a commonly used indicator for undifferentiated cells (Thomson et al. 1998) (Fig. 5.3). Taken together, these results demonstrate that *REX1* overexpressing hES cells can be established and stably maintained, and that these cells retain an undifferentiated character upon propagation.

5.3 Examination of the growth characteristics of *REX1* overexpressing hES cell clones

Interestingly, a trend showing that Rex1OE clones might grow faster than control clones was observed during routine cell culture. For example, Rex1OE wells would typically reach confluency a day or two days earlier than control wells. This suggested that REX1 might enhance cell proliferation. To assess whether this is the case, a colony formation assay was used to test the self-renewal ability of the Rex1OE and control cells. Rex1OE and control hES clones were trypsinised to a single cell level and seeded into 12-well plates at 5000 cells per cm^2 in normal $CDM^{activin}$ growth conditions and allowed to grow for eight days. At this point cells were stained with alkaline phosphatase and then the number of undifferentiated, mixed and differentiated colonies counted (see Fig. 5.4 for examples of colony morphology). In proliferating conditions both control and Rex1OE clones formed a high proportion of undifferentiated and mixed colonies (75-90%) as expected. However, the percentage of undifferentiated colonies was increased at the expense of differentiated and mixed colonies in Rex1OE wells, most strikingly in H1 cells with approximately 40% versus 55% in H1 and

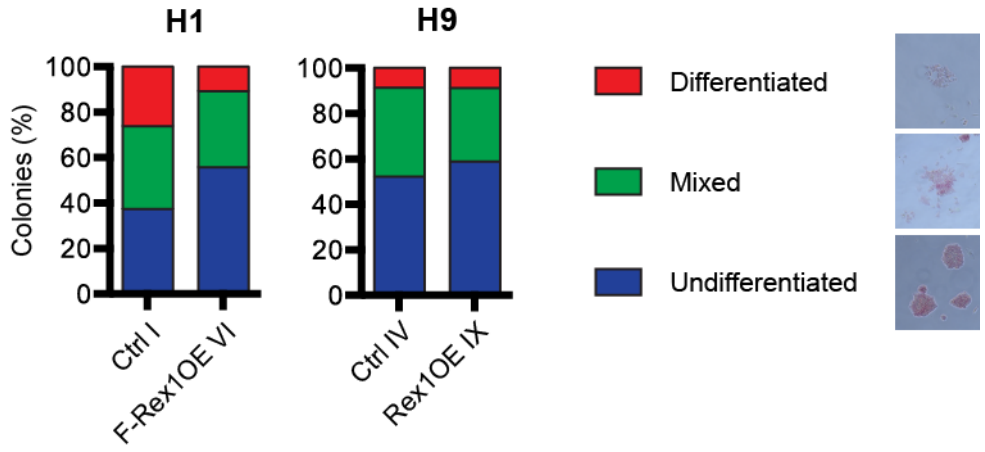


Figure 5.4 *REX1* overexpression does not inhibit hES cell self-renewal ability
Self-renewal capability of both control and Rex1OE cells was tested by colony formation assay. Cells were trypsinised and seeded in single cell suspension at 5000 cells per cm^2 onto 12-well plates under normal $CDM^{activin}$ conditions. After eight days cells were stained with alkaline phosphatase, colonies counted and assessed as being differentiated (red), mixed (green) or undifferentiated (blue). H1 Rex1OE (F-Rex1OE VI) wells had a higher proportion of undifferentiated colonies as compared to control (Ctrl I) wells (55% versus 37%). Similar trend was also observed for H9 when comparing Rex1OE (Rex1OE IX) and control (Ctrl IV) wells (60% versus 50%). All cell populations assessed gave rise to differentiated, mixed and undifferentiated colonies. Data representative of two independent experiments for both H9 and H1 cells. Colony example images magnification 10x

50% versus 60% in H9 clones of undifferentiated colonies counted in control and Rex1OE wells, respectively.

Other key pluripotency factors, such as *NANOG*, have previously been shown to directly impact on the cell cycle regulation of hES cells (Zhang et al. 2009). To assess whether *REX1* overexpression may similarly impact on hES cell cycle regulation, cells were pulse-labelled with BrdU and co-stained for PI. This allowed the quantification, by flow cytometry, of the relative number of cells in G1, S and G2 phases of the cell cycle and thus the comparison of the cycling state of Rex1OE and control hES cells. As illustrated by figure 5.5, both control and

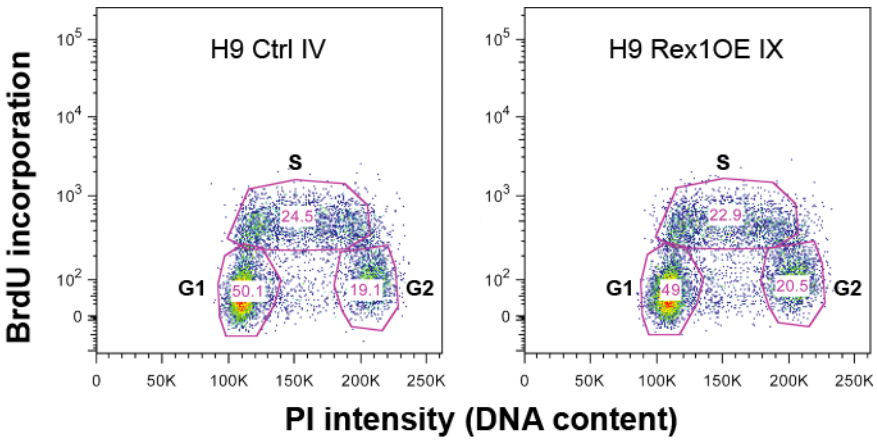


Figure 5.5 H9 control and Rex1OE cells have similar cell cycle profiles.
Bivariate distributions of H9 control and Rex1OE cells grown under $CDM^{activin}$ conditions, showing incorporation of 5-bromo-2-deoxyuridine (BrdU) and propidium iodide (PI) staining reflecting the DNA content. Percentage of cells in G1, S and G2/M phases in each sample is indicated. Data representative of four independent experiments.

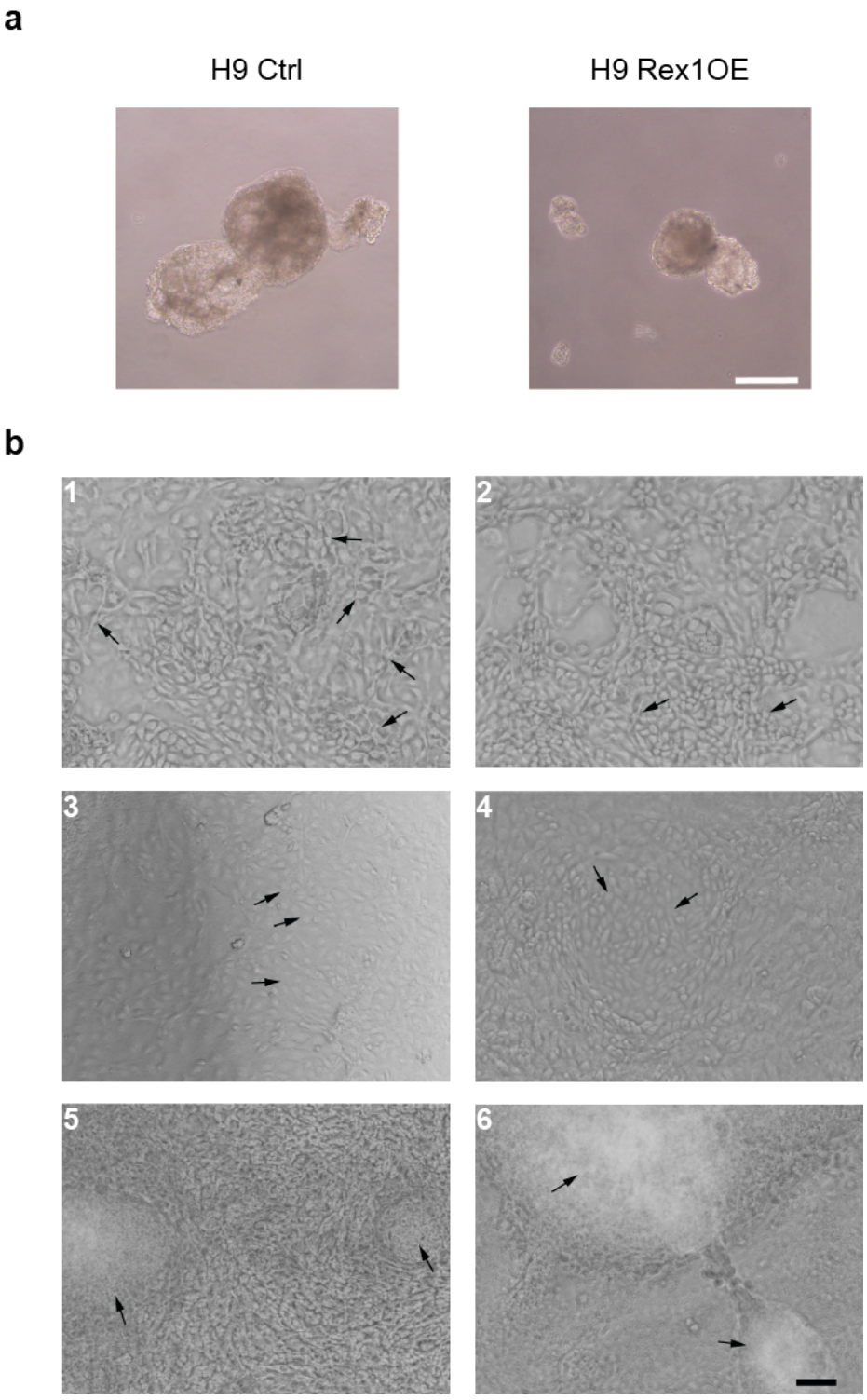
Rex1OE clones were actively cycling and displayed similar percentages of cells in each cell cycle phase (G_1: 49-50%, S: 22.9-24.5% and $G_{2/m}$: 19.1-20.5%). Taken together these data suggest that *REX1* overexpression might confer a self-renewal advantage in hES cells, however no direct effect can be seen on hES cell cycle properties. Note that these data are preliminary will need to be repeated with additional Rex1OE clones as well as with revertant clones to validate a direct role for REX1 as a self-renewal enhancer in hES cells.

5.4 H9 Rex1OE cells form embryoid bodies containing cells derived from all three germ-layers

To next test the ability of Rex1OE cells to give rise to derivatives of all three germ layers, an *in vitro* EB-based assay was used, as described in section 4.2.2. Here, *in vitro* differentiation was induced by growing H9 Rex1OE and control cells as EBs in non-adherent conditions and using differentiation media (Knockout™ DMEM supplemented with 20% FBS, 0.1 mM NEAA, 2 mM L-glutamine and 0.55 mM β-mercaptoethanol) for 10 days (Fig. 5.6a). EBs were then plated onto gelatin-coated plates and allowed to further differentiate for up to an additional 20 days (Fig. 5.6b). At this stage, a wide variety of morphologically different cell types were observed. The most obvious differentiated cell types were beating cardiac muscle-like clusters observed in both control and Rex1OE clones (Fig. 5.6b 5 & 6). Liver-like cells with well-defined cytoplasms and cell bodies were also observed in all populations (Fig. 5.6b 3 & 4). Neuronal-like cells with extended filaments were not observed, however evidence of neuronal differentiation was consistently seen in control wells with a number of cells having short dendritic-like extensions (Fig. 5.6b 1), note however that only a few of these cells were observed in Rex1OE plates (Fig. 5.6b 2). These observations will require further validation using lineage specific markers and immunofluorescence.

To gain more insight into the differentiation ability of Rex1OE cells, qRT-PCR analysis of differentiation markers was conducted on EB samples. Gene expression levels of markers for ectoderm (*PAX6, P63* and *OLIG3*), endoderm (*AFP, PDX1* and *GATA6*) and mesoderm (*BMP4, CD45* and *PECAM1*) were all upregulated in both Rex1OE and control samples after 30 days of differentiation (Fig. 5.6c). Furthermore, expression levels of pluripotency markers (*OCT4, NANOG* and *CRIPTO*) were all reduced to minimal levels in both Rex1OE and control cells as compared to the undifferentiated samples, confirming a loss of their pluripotent status in culture conditions favouring differentiation (Fig. 5.6c). Altogether, these results indicate that H9 Rex1OE hES cells are able to differentiate into derivatives of the three germ layers. It

would be interesting to further assess whether Rex1OE hES cells might be biased in their differentiation ability using directed differentiation protocols towards specific cell lineages.



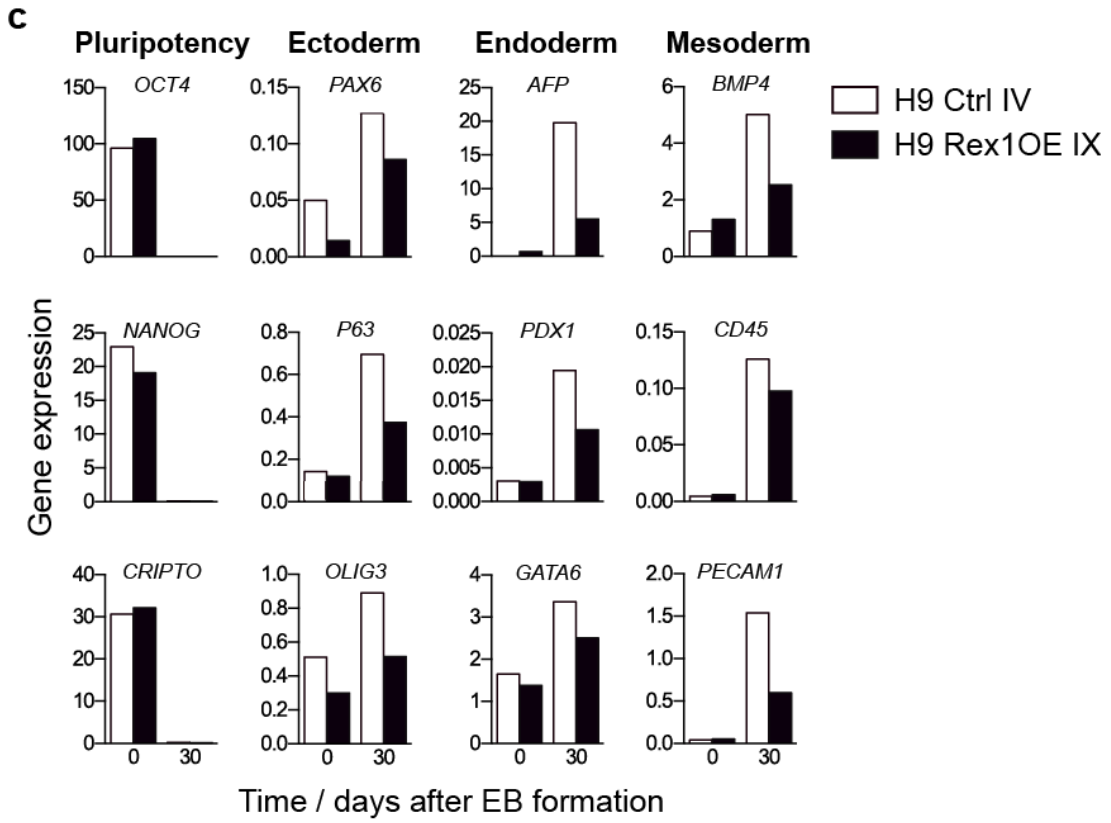


Figure 5.6 H9 Rex1OE hES cells are able to differentiate *in vitro* into derivatives of the three germ layers. (**a**) Morphology EBs formed in non-adherent conditions and cultured for 10 days. Scale bar 10 μm. (**b**) Morphology of EBs plated onto adherent culture dishes and allowed to differentiate for up to 30 days. Neuronal like (1-2) liver like (3-4) and cardiomyocyte like (5-6) cells were observed among control and Rex1OE cultures. Scale bar 10 μm. (**c**) Both control and Rex1OE cells efficiently downregulated the three pluripotency markers *OCT4, NANOG* and *CRIPTO*, and upregulated markers of all three germ lineages *PAX6; P63* and *OLIG3* (ectoderm), *AFP; PDX1* and *GATA6* (endoderm) and *BMP4; CD45* and *PECAM1* (mesoderm), as assessed by qRT-PCR at day 0 and day 30 after of EB assay. Results shown are representative of at least two independent experiments. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Primers used for this analysis are listed in Appendix I, Table 1.

5.5 H9 Rex1OE cells are able to maintain hall-marks of pluripotent cells in culture without Activin A supplementation

It has been reported previously that overexpression of *NANOG* in hES cells confers independence from feeder cells or conditioned media (Darr et al. 2006). This is reminiscent of the impact of overexpression of *Nanog* or another important pluripotency factor *Esrrb* on mES cells self-renewal, negating LIF-dependence (Chambers et al. 2003; Festuccia et al. 2012; Martello et al. 2012). To test whether Rex1OE may confer increased resistance to differentiation in hES cells, colony forming assays were repeated in the absence of Activin A supplementation. Interestingly, the proportion of undifferentiated colonies was increased with H1 F-Rex1OE, from 35% to 60% in control versus Rex1OE cells; a trend also observed to a lesser extent in H9 Rex1OE cells (Fig. 5.7a). To investigate whether Rex1OE hES cells may

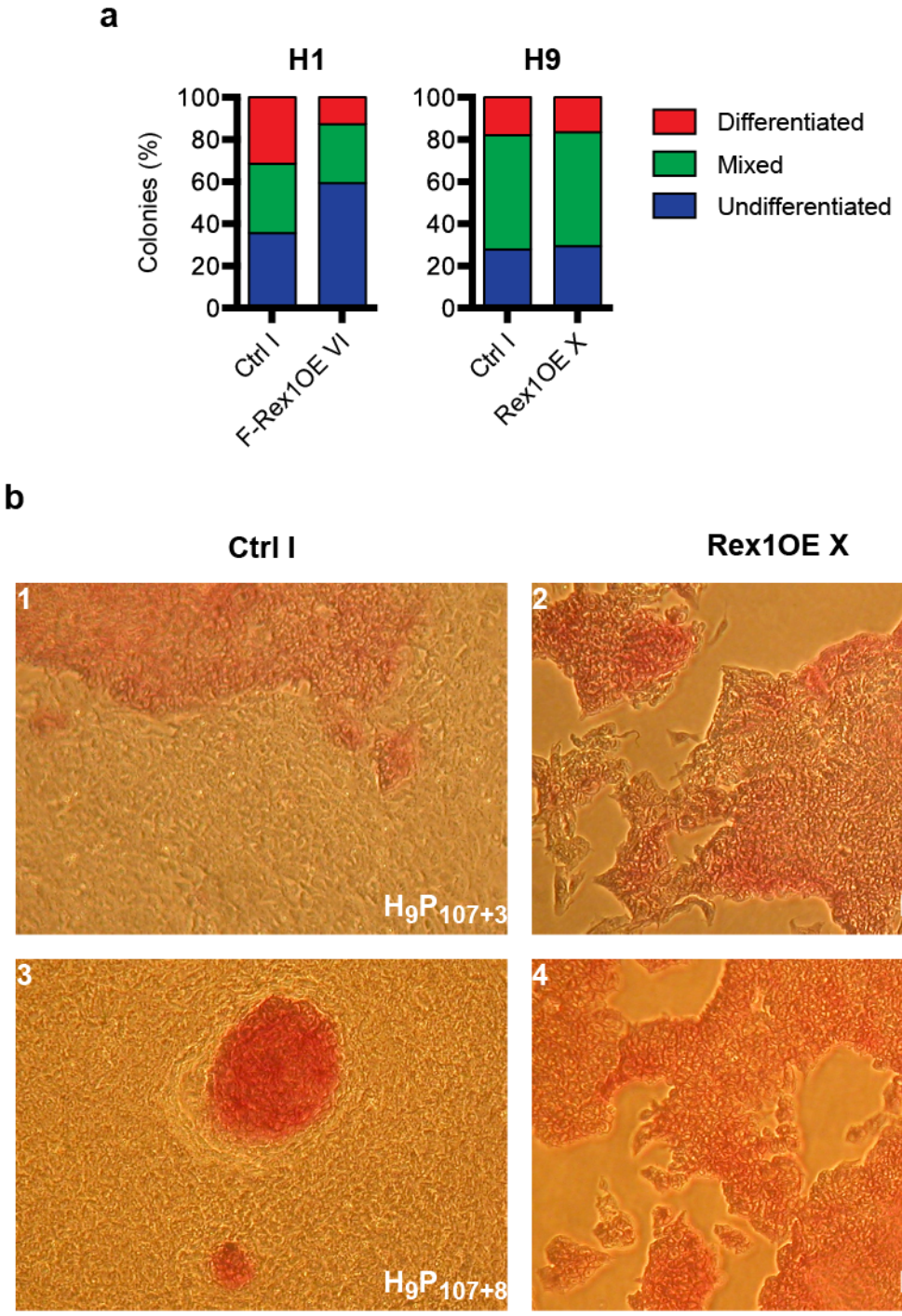
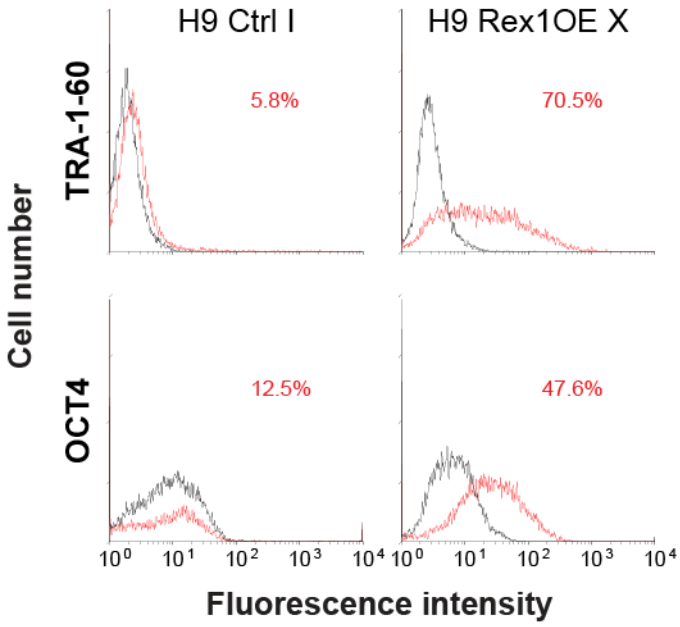


Figure 5.7 H9 Rex1OE hES cells can be serially passaged in absence of Activin A supplementation.
(**a**) H9 and H1 Rex1OE and control cells were assayed for colony formation as described in figure 5.5 in absence of Activin A. In H1 cells a significant increase of undifferentiated cells was observed as compared to control cells (approximately 60% compared to 35% respectively), however for H9 cells a very slightly increased proportion of colonies remained undifferentiated in Rex1OE wells as compared to control cells. Data representative of two independent experiments. (**b**) H9 control and Rex1OE cells were cultured in CDM conditions in absence of Activin A for up to eight passages and tested for AP activity. At both passage three and eight, the majority of control cells were negative for AP activity and only a few undifferentiated, AP-positive (pink), colonies were observed (**1** & **3**). In contrast, Rex1OE cells stained positive for AP with little or no evidence of differentiation at both passage three and eight (**2** & **4**). Images representative for three independent experiments for passage three and a single experiment for passage eight. Magnification 10x

indeed be capable of maintaining self-renewal without Activin A, H9 Rex1OE and control cells were grown for multiple passages without Activin A. High proportion of H9 control cells cultured without Activin A readily differentiated after three passages in these conditions as judged by morphological changes and lack of positive staining for AP activity (Fig. 5.7b 1 & 3). By comparison, H9 Rex1OE cells cultured without Activin A, interestingly, retained typical hES colony morphology and continued to stain positive for AP activity with little evidence of differentiation for more than eight passages (Fig. 5.7b 2 & 4).

Further evidence that *REX1* overexpression may at least partly, inhibit differentiation of hES cells in absence of Activin A came from FACS analysis with TRA-1-60 and OCT4 antibodies. While TRA-1-60 expression was almost completely lost in control cells without Activin A (6% positive cells), approximately 71% of H9 Rex1OE X cells retained positive staining for TRA-1-60, as assessed here after 18 passages in absence of Activin A (Fig. 5.8a). For reference, H9 Rex1OE cells maintained in routine $CDM^{activin}$ conditions stained 79% positive for TRA-1-60 using identical flow cytometry settings (see Fig. 5.2b). Similar trend was observed with regards to OCT4 staining. While only 13% of control cells stained positive for OCT4 after 18 passages in the absence of Activin A signalling, 48% of Rex1OE cells retained OCT4 staining in similar conditions (Fig. 5.8a). Note however that, the percentage of OCT4 positive in these cultures without Activin A was significantly lower than the 73% of OCT4-positive Rex1OE cells detected in routine $CDM^{activin}$ cultures (see Fig. 5.2b), suggesting that Rex1OE cells ultimately differentiate and/or adopt a transitional phenotype. Interestingly, intermediate *OCT4* mRNA expression levels were also observed by qRT-PCR analysis in H9 Rex1OE cells cultured without Activin A for up to 11 passages (Fig. 5.8b). In contrast, expression of *NANOG* mRNA levels gradually decreased in Rex1OE cells when cultured without Activin A, as seen in control cells. Altogether, these results suggest that *REX1* overexpression enhances hES cell's resistance to differentiation upon Activin A withdrawal. Following serial passages in absence of Activin A, Rex1OE cells would, however, adopt an intermediate/transitional phenotype with lower yet detectable *OCT4* expression levels, but no expression of *NANOG*.

a


H9 Ctrl I
H9 Rex1OE X
TRA-1-60
OCT4
5.8%
70.5%
12.5%
47.6%
Cell number
10^0 10^1 10^2 10^3 10^4
Fluorescence intensity

b

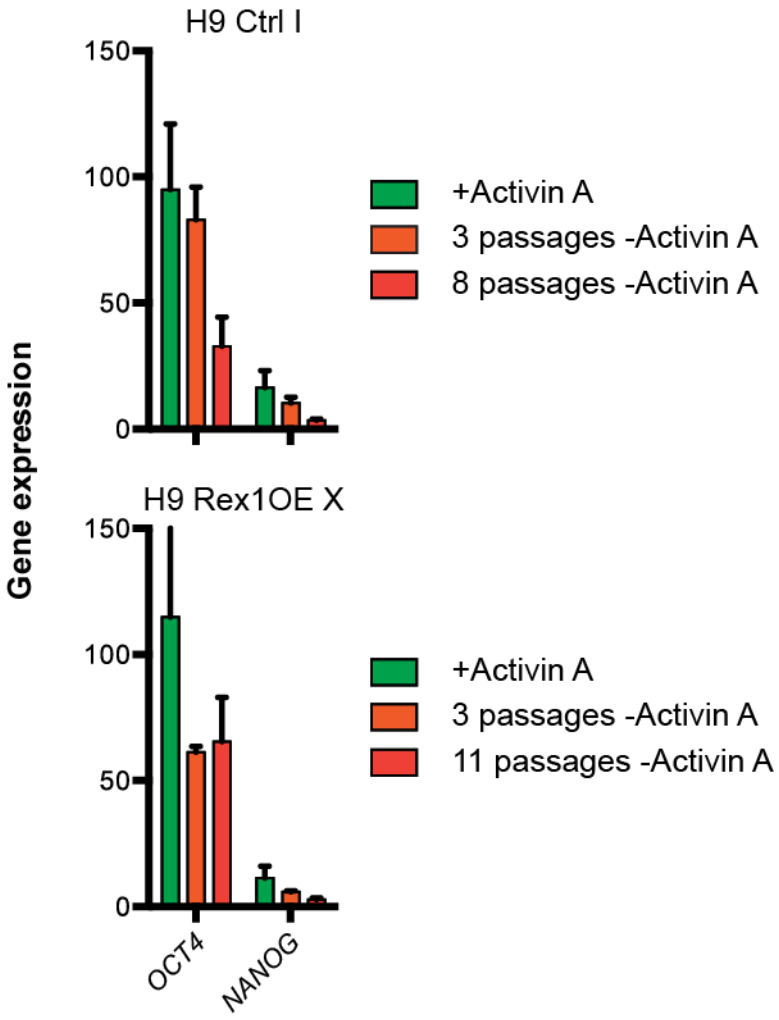

H9 Ctrl I
150
100
50
0
+Activin A
3 passages -Activin A
8 passages -Activin A
H9 Rex1OE X
150
100
50
0
+Activin A
3 passages -Activin A
11 passages -Activin A
Gene expression
OCT4
NANOG

Figure 5.8 Rex1OE hES cells retain intermediate OCT4 expression levels in prolonged culture without Activin A.
(**a**) Expression of pluripotency markers in H9 control and Rex1OE cells cultured without Activin A for up to 18 passages was analysed by flow cytometry. Histograms representative of staining for TRA-1-60 and OCT4 for both H9 control and Rex1OE cells. Up to 70.5% of Rex1OE cells stained positive for TRA-1-60, whereas only 5.8% of control cells were positive. For OCT4 47.6% of Rex1OE cells stained positive, compared to only 12.5% of control cells. The profile of positively stained cells (red) was compared to (black) cells stained with an isotype-matched control. Percentage of positively gated cells from each individual example shown in graphs. Histograms representative of two independent experiments. (**b**) Relative expression levels of *OCT4* and *NANOG* were assayed by qRT-PCR from control and Rex1OE cells cultured in absence of Activin A for up to eleven passages and compared to samples from cells cultured in normal CDM conditions. Both Rex1OE and control H9 cells down regulate *NANOG* in absence of Activin A. Interestingly, while both control and Rex1OE cells downregulate *OCT4* following withdrawal of Activin A, Rex1OE cells retain an intermediate level of expression, even after eleven passages. Mean and standard error of three independent experiments are shown. Data was normalised to two housekeeping genes (GAPDH and HPRT) and then to the expression level of the sample cultured with Activin A supplementation in each cell type. Primers used for this analysis are listed in Appendix I, Table 1.

5.6 Conclusions

REX1 is a widely used as a marker of pluripotency in both mouse and human cells, however its exact role in pluripotent ES cells remains largely unclear. Conflicting results have been reported with *Rex1* being found both necessary and dispensable for mES cell pluripotency (Zhang et al. 2006; Masui et al. 2008; Scotland et al. 2009). Overexpression of *Rex1* in mES cells was similarly suggested to interfere with self-renewal (Zhang et al. 2006). However high endogenous expression of *Rex1* have been linked to high self-renewal ability and to the 'ground-state' of pluripotency in mES cells in other studies (Toyooka et al. 2008; Wray et al. 2010; Wray et al. 2011; Marks et al. 2012). In addition to this, high *REX1* levels have been associated with complete reprogramming of human iPS cells (Chan et al. 2009).

Here the role of *REX1* was investigated in hES cells upon generation and characterisation of *REX1* overexpressing clones. These preliminary results indicate that *REX1* overexpression is not detrimental to the self-renewal ability of hES cells, but might instead confer a self-renewal advantage, as indicated by increased proportion of undifferentiated colonies in a colony forming assay. *REX1* overexpressing hES cells were found to express similar levels of the pluripotency factors *OCT4, SOX2* and *NANOG* and were able to give rise to cells from three different germ lineages, indicating that the cells remain pluripotent. However, a possible bias of differentiation was observed in *REX1* overexpressing cells with only a few or no cells of neuronal morphology being detected following EB differentiation as compared to control EBs (see Fig. 5.6b 2). Further insight into the possible differentiation bias of *REX1* overexpressing cells could be obtained from experiments using lineage-specific differentiation protocols.

Interestingly *REX1* overexpressing cells could be maintained in culture in the absence of Activin A supplementation for numerous passages, where control cells underwent rapid differentiation. Gene expression analysis indicated that *REX1* overexpressing cells retained an intermediate level of *OCT4* expression as compared to control cells. Studies in mES cells have indicated that *Rex1* binds the *Oct4* locus and therefore could modulate *Oct4* levels (Kim et al. 2008). It is possible that this regulatory link is conserved in humans and hence could explain the observed intermediate *OCT4* levels in *REX1* overexpressing as compared to control hES cells, despite the loss of *NANOG* expression in both populations. Further experiments would be required to better understand the role of *REX1* in hES cells. It would be interesting to see the effect of knock down of *REX1* in hES cells as well as mapping of the genome-wide *REX1* binding site distribution, for which Flag-*REX1* overexpressing cells would be instrumental in the absence of commercially available human REX1 specific antibodies.

Chapter 6 Discussion

Pluripotent stem cells offer a tremendous possibility for research into mammalian development *in vitro*. Through the development of reprogramming technologies (Takahashi et al. 2006; Takahashi et al. 2007; Yu et al. 2007), they hold the promise of revolutionising regenerative medicine and novel insights into previously incurable diseases. A growing panel of both mouse and human pluripotent cell lines have been derived from a various embryonic stages (Evans et al. 1981; Martin 1981; Brook et al. 1997; Thomson et al. 1998; Brons et al. 2007; Tesar et al. 2007) as well as postnatal mouse testis (Kanatsu-Shinohara et al. 2004) and by reprogramming somatic cells to pluripotency (Takahashi et al. 2006; Takahashi et al. 2007; Yu et al. 2007). These cells share the expression of key pluripotency factors (such as *Oct4*, *Sox2* and *Nanog*), are able to self-renew indefinitely in culture and give rise to differentiated progeny when induced (Jaenisch et al. 2008; Pera et al. 2010). However, their relationship to each other and whether they correspond to distinct pluripotent states *in vivo* remains unclear. Especially the true lineage affiliation of hES cells in relation to that of mES and mEpiS cells remains a point of debate. Mouse and human ES cells are derived from the ICM of pre-implantation embryos, however hES cells are able to give rise to extra-embryonic tissues (Itskovitz-Eldor et al. 2000; Reubinoff et al. 2000; Xu et al. 2002) whereas mES cells cannot. One possible explanation could be that hES cells originate from a less committed (possibly pre-epiblast) stage while mES cells represent the lineage-restricted epiblast progenitors in the ICM. However such an explanation contradicts with recent reports suggesting that hES cells closely resemble mEpiS cells - a pluripotent stem cell population derived from post-implantation mouse embryos (Brons et al. 2007; Tesar et al. 2007). Furthermore recent evidence indicates that the extra-embryonic tissues which hES cells were thought to give rise to are actually of mesodermal lineage (Bernardo et al. 2011). This study sought to investigate whether distinct pluripotent states could be reliably delineated at the chromatin level by assaying and comparing the DNA replication timing profiles of mES, hES and mEpiS cells. This notably uncovered that hES cells can adopt distinct yet reversible epigenetic states that could be modulated in response to different cell signalling environments. Moreover, these findings identified a role for TGF-β/Activin signalling through SMAD2/3 and the histone acetyltransferase P300 in defining these epigenetic states. Finally, preliminary evidence from this study also hints at a possible role for REX1 in sustaining hES cell self-renewal, possibly via modulating levels of *OCT4*.

6.1 Human and mouse ES cells have distinct replication timing profiles

The present study was designed to investigate whether the distinct yet closely related pluripotent developmental states represented by mES, hES and mEpiS cells could be reliably discriminated using candidate locus based replication timing analysis. The candidate loci selected for this analysis mainly encode human and mouse homologs selectively expressed in embryonic and extra-embryonic tissues of the early embryo or involved in early somatic and germ-line lineage specification.

The current study found that replication timing profiles of independently derived mES cells lines from distinct genetic backgrounds were highly conserved across cell lines. Interestingly, this contrasts with the reports describing high transcriptional variability in mES cells, with regards to pluripotency associated genes such as *Nanog*, *Rex1* and *Stella* (Chambers et al. 2007; Hayashi et al. 2008; Toyooka et al. 2008). This could suggest that replication timing profiles may be a more robust identifier of mES cell identity than transcriptional profiles. The replication timing profiles of mES cells were found to be replicating early in S-phase for a majority of the candidate loci studied; a result corroborating previous findings (Azuara et al. 2006; Hiratani et al. 2008) and reflecting a broadly accessible chromatin conformation. By the means of replication timing analysis, mES cells could also be reliably distinguished from closely related pluripotent mEC cells and mES cell derived progenitors, such as neural progenitor cells (Perry et al. 2004; Azuara et al. 2006; Hiratani et al. 2008). This prompted the use of replication timing analysis to characterise hES cells, where lineage affiliation and functional verification of embryonic potential *in vivo* is not possible, due to obvious ethical reasons.

Unexpectedly, the replication timing profiles of hES cell lines were significantly different from those of mES cells. All of the three hES cell lines analysed H1, H7 and H9 displayed a much increased proportion of candidate loci replicating in middle, middle-late or late parts of S-phase as compared to mES cells. Previously such an increase in the proportion of later replicating loci has been associated with differentiation and loss of pluripotent identity (Hiratani et al. 2004; Perry et al. 2004; Azuara et al. 2006). Despite this, hES cultures analysed in this study were confirmed to express high and homogeneous levels of key pluripotency markers such as *OCT4* and *NANOG* and other markers of the undifferentiated hES cell state, thus excluding that increased incidence of spontaneous differentiation could account for the observed differences. Furthermore, H7 hES cells could still be distinguished from neural progenitor cells derived from these hES cells by replication timing (Appendix II,

Fig. A1), as previously reported for mES cells (Perry et al. 2004; Azuara et al. 2006; Hiratani et al. 2008). The observed divergence between mES and hES replication timing profiles correlates with data from genome-wide experiments (Ryba et al. 2010), where mES and hES replication timing profiles were reported highly conserved within regions of conserved synteny, but were found to have significant differences at the genome-wide scale. Interestingly, the coordinated DNA replication in megabase-sized replication domains is conserved between mouse and human cells, however in mES cells early and late replicating domains are of similar sizes and contrastingly in hES cells late replication domains are significantly larger and less numerous than early domains (Ryba et al. 2010).

6.2 Human ES and mouse EpiS cells share a common variable epigenetic state as defined at the replication timing level

Another unanticipated feature of hES cell's replication timing profiles was the significant degree of variability observed between different hES cell lines and cultures. In particular, many differentiation-associated loci were found to replicate at different times during S-phase, regardless of their lineage affiliation and cell lines analysed. This contrasts with mES cells, which instead display stable and largely early replicating profiles across cell lines. Such variability could not be simply explained by differences in derivation and culture histories of the analysed cell lines (H1, H7 and H9) as they were all derived at the same time in the same laboratory (Thomson et al. 1998) and have been maintained in the same culture conditions (Xu et al. 2001). Furthermore, it is worth noting that this replication timing variability was specific only to certain loci (*DPPA4, TERT, TLE1, SYCP1, BMP4, EBF, FOXA2, GATA6, SOX17, SOX3* and *ZIC1*); a number of variable loci that is likely to increase with the analysis of additional candidate genes or hES cell lines. An alternative explanation for the observed variability might be that it reflects the substantial level of genetic variation found in human populations, in contrast to inbred laboratory mouse populations. Recent studies have clearly reported variability between hES cell populations at the transcriptional level (Skottman et al. 2005; Wei et al. 2005; Adewumi et al. 2007; Guenther et al. 2010; Narva et al. 2010; Newman et al. 2010; Amps et al. 2011) and at the epigenetic level (Guenther et al. 2010; Bruck et al. 2011; Nazor et al. 2012). However, analysis of 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 ruled out genetic variation as the primary explanation. The replication timing profiles of human CD34+ haematopoietic stem

cells were shown to be remarkably similar across samples, with no or little evidence of epigenetic variability (Appendix II, Fig. A8).

The loci, which exhibited variable replication timing profiles in hES cells, were also analysed for histone acetylation. Acetylated histones are thought to reflect variations in chromatin accessibility and transcriptional permissiveness (Bickmore et al. 1995; Vogelauer et al. 2002; Donaldson 2005). Analysis with histone acetylation marks (H3K9ac, H3K14ac, H3ac and H4ac), failed to find any correlation with the variable replication timing profiles observed in hES cells (Appendix II, Fig. A4). No correlation was also found with the expression levels of the variable genes (Appendix II, Fig. A4). In addition to this, DNA methylation did not show any correlation with the variable loci (Appendix II, Fig. A5). Analysis of the genomic features (GC content and LINE density) at variable loci, also failed to identify any clear correlation (Appendix II, Table A4), however all variable loci, apart from *TERT*, were found in gene poor regions.

In an attempt to determine the lineage affiliation of hES cells, replication timing analysis was also extended to several mEpiS cell lines and their profiles compared to those of hES and mES cells. Restricting this analysis to a set of known ICM/epiblast-specific markers was not sufficient to clarify the developmental status of hES cells. *Rex1* – an ICM-specific marker that is shutdown in the epiblast – replicates early during S-phase in mES cells as expected, and 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, yet remains late replicating in hES cells, suggesting that hES cells may not strictly resemble mEpiS cells, as previously reported (Brons et al. 2007; Tesar et al. 2007). It should, however, be noted that the selection of these markers was based on gene expression analysis in mouse embryos, in the absence of equivalent information in human embryos at the time. Use of loci determined from human embryonic development could provide a better correlation between hES and mEpiS cells. Replication timing analysis was further extended to other target loci previously analysed in mES and hES cells that are subjected to replication timing variability in hES cells, i.e. differentiation-associated markers such as *Ebf*, *Ikaros*, *Foxa2*, *Sox17*, *Sox3* and *Zic1*. Remarkably, similar variability as in hES cells was also observed across many of these loci. What could be the functional significance of the observed replication timing variability in hES and mEpiS cells? Such a level of plasticity could suggest an unstable or transitional state: primed on the verge of differentiation. Mouse EpiS cells are thought to represent such a transitional state and it has been suggested that hES cells share this state (Rossant 2008; Nichols et al. 2009b). As seen In this study, mEpiS and hES cells

furthermore harbour similar cell cycle properties, most likely reflecting their common growth requirements (Brons et al. 2007; Tesar et al. 2007). Interestingly the replication timing profiles of hES and mEpiS cells were shown to be more significantly aligned to each other than to mES profiles, at the genome-wide scale (Ryba et al. 2010). Hence, the observation of shared variability in hES and mEpiS, cell's profiles further support the view that hES cells might preferentially harbour an epiblast-like phenotype in culture, possibly suggesting a pre-gastrulating developmental stage affiliation *in vivo*.

6.3 Distinct and reversible epigenetic states within hES cell cultures

A key finding from the current study is the demonstration that hES cells can undergo extensive yet reversible epigenetic alterations in response to changing cell culture conditions. Both H9 and H1 cells replication timing profiles were shifted towards early S-phase as seen at many specific loci when transferring cells from CM medium to $CDM^{activin}$ medium. Furthermore, experiments on fully reprogrammed hiPS cells confirmed that both pluripotent hES and hiPS cells share the same level of epigenetic plasticity in response to environmental cues. Importantly, there was a high degree of overlap between loci that advance their replication kinetics toward early S-phase in H9, H1 and hiPS cells, suggesting that specific chromosomal regions may be targeted across the genome and that a conserved underlying mechanism would operate in both natural (hES) and induced (iPS) pluripotent cells. Interestingly, when the *REX1* locus was studied in detail, it was observed that in $CDM^{activin}$ cultured hES cells a small window of early replication centred around *REX1* was embedded within a domain of late replication, where as in CM cultured cells the whole studied region replicated in late S-phase This would suggest the existence of a replication origin within the studied region and that change of the cell culture environment allowed its early firing. It would be interesting in future studies to try to localise the possible origin of replication perhaps by *in situ* hybridisation and immunofluorescence targeted at members of the DNA replication complex, such as PCNA.

Importantly, no such replication timing alterations were observed when comparing the profiles of 'normal' versus late-passage, 'culture adapted' hES cell cultures often associated with karyotypic changes that provide some advantages in terms of clonogenic capacity and self-renewal properties of these cells (Draper et al. 2004; Enver et al. 2005; Baker et al. 2007b; Amps et al. 2011) (Appendix II, Fig. A9). Moreover, the dynamic and reversible nature of the observed replication timing shifts argues against cell adaptation. Nevertheless, the demonstration that different hES cell populations with distinct epigenetic properties can

predominate under varying cell culture conditions is of critical importance and might at least partly explain variation in experimental results observed at different laboratories. Different starting populations of pluripotent cells, according to their epigenetic identity, might respond differently to protocols of differentiation and have unforeseen biases for generation of certain tissues.

Here, it was verified that H1 hES cells cultured under both CM and CDM^{activin} conditions retain the expression of markers of undifferentiated state, and the ability to give rise cells from all three germ lineages. However, phenotypic and functional differences were observed between these cell populations. Interestingly, transcriptional analysis of a limited panel of genes demonstrated that markers associated with the ICM-stage in human embryos became upregulated in the CDM^{activin} cultured cells. Furthermore, when the differentiation kinetics of CM and CDM^{activin} grown H1 hES cells were analysed, CM cultured population was observed to more rapidly upregulate markers of differentiation (Appendix II, Fig. A10). Taken together these results demonstrate that human ES cells can reversibly adopt different epigenetic identities and supports the view that pluripotency is not a fixed state, but that heterogeneous pluripotent sub states can exist with distinct phenotypical and functional properties. Furthermore, this emphasises the importance of determining the true identity of any pluripotent cell population for future applications in regenerative medicine, as responses to differentiation protocols are likely to vary depending on the status of the cell population. It would be of interest to extend this study to other growth culture conditions of hES and iPS cells such as mTeSR (Ludwig et al. 2006) and Essential 8 (Chen et al. 2011) media, which do not contain Activin A, but are dependent on FGF2 and TGFβ1 or FGF2 and Nodal/TGFβ1, respectively.

6.4 Investigating the molecular pathways underlying distinct pluripotent states in hES cells

Previous reports have demonstrated the critical role of TGF-β/Activin signalling in maintaining pluripotent hES cell cultures. Here, upon transfer of H1 hES cells from CM to CDM^{activin} conditions, levels of activated SMAD2/3, the intracellular signalling molecule associated with the TGF-β/Activin pathway, were increased. This increase is likely due to the Activin A supplementation in chemically defined media as well as the lack of inhibitory and/or modulatory signalling factors known to be present in conditioned media (Lim et al. 2002; Chin et al. 2007; Prowse et al. 2007; Villa-Diaz et al. 2009). The increased P-SMAD2/3 levels were also accompanied by an increase in the levels of NANOG, which however were

only transiently upregulated. This transient increase could be due to a negative feedback thought to exist between NANOG expression levels and SMAD2/3 signalling possibly leading to the observed 'self-regulation' (Vallier et al. 2009a). Similar transient increases were also observed in the expression levels of several members of the TGF-β family, such as *NODAL* (Appendix II, Fig. A7), whose expression is already known to be controlled by negative feedback regulation (Besser 2004).

Significantly, profiling of the expression of various histone and chromatin modifiers in H1 hES cells upon changing growth culture conditions demonstrated that P300, a known SMAD2/3 co-activator (Pouponnot et al. 1998; Ross et al. 2006), levels were specifically upregulated in H1 $CDM^{activin}$ cells. Subsequently, it was observed that modulating TFG-β/Activin signalling induced changes in the levels of P300 expression in H1 hES cells; a previously unreported regulation link. Analysis of established hES cell lines with stable *SMAD2* knockdown and either *NANOG* knockdown or overexpression furthermore suggested that the observed effect on P300 expression might be mediated through SMAD2/3 only and not via NANOG. Interestingly, recent studies mapping SMAD2/3 binding sites in hES cells have indicated a SMAD2/3 binding site located 10 kb downstream of the *P300* gene. Further analysis, by means of for example luciferase activity assays and site directed mutation, is required to determine whether this binding site is indeed regulating *P300* levels in hES cells in response to TGF-β/Activin signalling.

The functional significance of P300 upregulation in H1 $CDM^{activin}$ cultures was then further investigated and linked to the acquisition of an early replicating profile across many silenced loci. Notably, a gradual increase in the bulk levels of P300 preferentially acetylated histone targets was observed over serial passages of H1 hES cells under $CDM^{activin}$ conditions. Importantly, it was shown that P300 is selectively recruited to target loci shifting their replication timing from late to early upon change of growth culture conditions, with higher enrichment levels observed in H1 $CDM^{activin}$ cultured cells when compared to H1 CM cell populations. In accordance to this, increased enrichment levels of associated histone H3 acetylation were also observed at these target loci. Histone acetylation is generally linked with increased locus accessibility (Kouzarides 2007), as seen here, in $CDM^{activin}$ cells, at both active (i.e. *REX1*) and inactive genes (i.e. *CGA* and *FOXA2*) and consistently reflected at the level of DNA replication timing; hyperacetylation at these loci positively correlates with an early replication profile., as previously reported (Méndez 2009).

Functionally, inhibiting P300 HAT activity in $CDM^{activin}$ hES cell cultures using the specific P300/CBP inhibitor curcumin, was sufficient to induce early-to-late replication timing

alterations at loci that were previously found to shift their replication timing from late-to-early upon transfer from CM to CDM^{activin} conditions. This suggests a role for P300 and/or CBP in stably maintaining a hyperacetylated and early replicating profile in hES cell cultures. Attempts to establish stable *P300* knockdown hES cell lines and thus validate a direct role for P300 proved unsuccessful, as P300-depleted cultures could only be propagated for a few (three to five) passages before they became either overcome by differentiation or lost due to senescence and cell death. This phenotype might be potentially caused by multitude of factors that relate to the pleiotropic role of P300. Over 400 different proteins have been described to interact with P300/CBP (Bedford et al. 2012), including, for example, the tumour suppressor P53 (Gu et al. 1997), transcriptional activator β-CATENIN (Hecht et al. 2000), cell cycle regulator CDK2 (He et al. 2011) and TF YY1 (Lee et al. 1995). Homozygous P300 or CBP –null mouse embryos die by E10.5, due to heart and neural tube defects (Yao et al. 1998). Interestingly, the compound heterozygous mutant embryos are also embryonic lethal, indicating that gene dosage of P300 and CBP needs to be tightly regulated during normal development (Yao et al. 1998). Perhaps most importantly P300/CBP activity is also required for the G1/S-phase transition of the cell cycle (Ait-Si-Ali et al. 2000). This could explain the observed arrest of cell cultures upon stable P300 knockdown. Furthermore P300 has also been reported to regulate the expression of *Nanog* in mES cells (Zhong et al. 2009). Whether such a link is conserved in hES cells remains to be confirmed, however reduced levels of *NANOG* expression as seen in P300-depleted hES cells might be a plausible reason for the increased incidence of differentiation in these clones.

Transiently transfecting hES cells with siRNA oligos that specifically target *P300*, in contrast, did not lead to cell proliferation arrest and/or differentiation. However, even though levels of *P300* were effectively reduced at both mRNA and protein level four days post transfection, only a very mild reduction in global acetylation levels of H3 histone residues was observed. Subsequently, analysis of the replication timing demonstrated no change between control and siP300-treated hES cell samples. A possible explanation for the lack of reduction in histone acetylation is that *P300* knockdown could be masked by compensation through activity by the closely related CBP. It would be interesting to address this question with a double targeting of CBP and P300 simultaneously. It is equally possible that the effect of P300 depletion on histone acetylation would need several rounds of cell division to manifest. Note that a gradual increase in histone acetylation was evident upon transferring cells from CM to CDM^{activin} conditions after at least five passages under CDM^{activin} conditions.

Inhibition of P300/CBP activity, by curcumin, in H1 CDM^{activin} cells induced early to late shifts of replication timing at ten out of the twelve loci shown to shift upon transfer from CM to CDM^{activin} conditions and had no effect on loci not subject for epigenetic alterations. This suggests a role for P300/CBP activity in specifically determining the epigenetic status of hES cells. Attempts to duplicate the effects of P300/CBP inhibition with a small molecule inhibitor of P300/CBP, C646 (Bowers et al. 2010), failed to induce similar reversion of early replication as seen with curcumin (Appendix II, Fig. A11). An explanation for C646 and curcumin having different efficiencies of inhibiting P300/CBP HAT activity in H1 CDM^{activin} cells could lie with their mechanism of action. C646 has been described to interact with the HAT domain of P300/CBP in a similar manner as another inhibitor Lys-CoA, and thus reduce the binding and HAT activity (Liu et al. 2008; Bowers et al. 2010). Curcumin on the other hand, has been shown to covalently bind P300/CBP and inhibit their activity as well as target them for proteasome-mediated degradation (Marcu et al. 2006), which could explain the higher effects observed with curcumin treatment.

Taken together, these data suggest that P300/CBP HAT activity might play an important role in regulating the epigenetic state of hES pluripotent cells in response to TGF-β/Activin signalling (Fig. 6.1), favouring a stable, hyperacetylated profile and inducing the expression of ICM-related markers. It does however remain to be determined, whether both P300 and CBP are involved in this process, as transiently knocking-down P300 alone did not significantly impact on hES cell epigenetic state. It should also be noted that both chemical inhibitors used in these experiments do not selectively act on P300 but on both HATs and thus cannot shine light on this question. Future experiments refining the dose of C646 used or the use of a different inhibitor such as Lys-CoA (Liu et al. 2008), could provide further

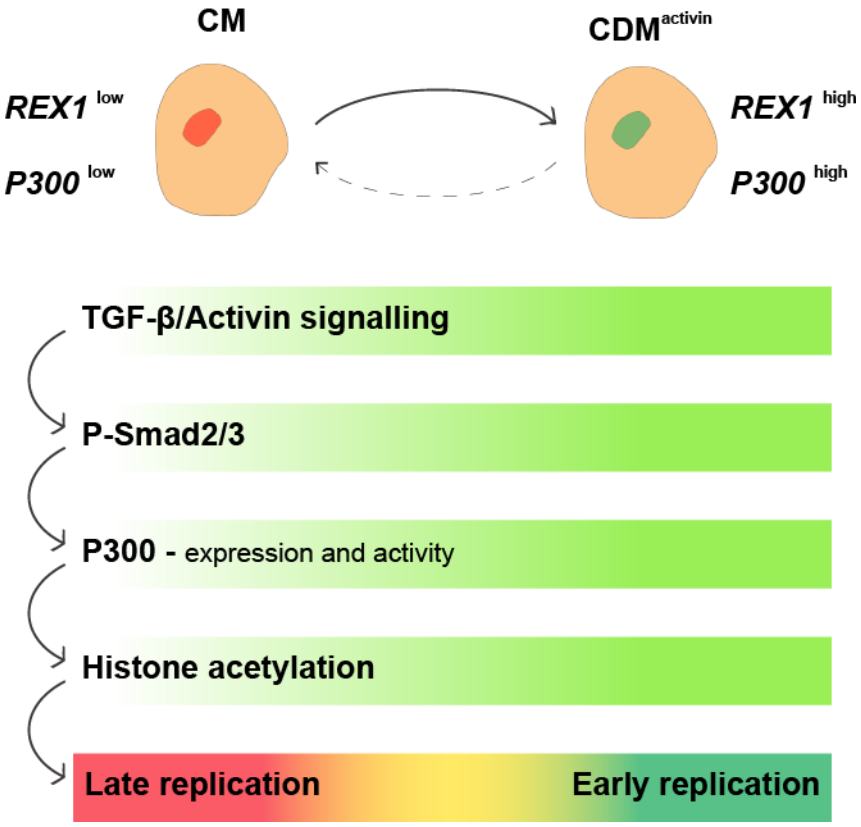


Figure 6.1 Summary for observed interplay between TGF-β signalling and P300 HAT activity in the regulation of distinct epigenetic states in hES cells.
Epigenetically distinct states are determined in hES cell cultures grown under varying culture conditions, as exemplified here by two different cell culture conditions, CM and CDM^{activin}, respectively relying on TGF-β/Activin signalling pathway for maintenance of hES pluripotency. Human ES cells are able to transit from one state to the other when transferred from CM to CDM^{activin} conditions. This was accompanied by increase in TGF-β/Activin signalling activity as indicated by higher levels of active P-SMAD2/3. This leads to the induction of P300 expression and activity upon changing cell signalling environment, favouring and stably maintaining a hyperacetylated state and early replicating profile at many key developmental loci. Interestingly levels of *REX1* were also low in H1 hES CM cells and high in CDM^{activin} cells, suggesting that these reversible epigenetic states may underlie different pluripotent states within hES cell cultures

confirmation on the key role of P300/CBP in regulating the hES epigenetic state.

6.5 Effects of *REX1* overexpression in hES cells

Interestingly, upon change of culture conditions from CM to CDM^{activin} *REX1* was significantly upregulated in H1 hES cells. To gain further insight to the role of *REX1* in hES cells, *REX1* overexpressing cell lines were established in H1 and H9 hES cells. Preliminary characterisation of these cells indicates that they retain key hall-marks of undifferentiated hES cells. They were also able to give rise to a multitude of different cell types via EB differentiation, however *REX1* overexpressing cells may in fact possibly exhibit a differentiation bias as no clear morphological evidence of neuronal-like cells was seen during the EB analysis. Gene expression analysis of markers of differentiation did, however, indicate an upregulation of ectodermal markers, alongside endo- and mesodermal markers. To test whether the Rex1OE cells in fact do have a bias for differentiation they should be tested with lineage-specific protocols of differentiation.

Interestingly, preliminary evidence also indicates that *REX1* overexpression in hES cells might actually confer a slight advantage for self-renewal, as indicated by colony forming assays. Unexpectedly, Rex1OE cells were also found to be able to partially resist differentiation when cultured in the absence of Activin A supplementation. Rex1OE cells cultured without Activin A retained an intermediate level of *OCT4* expression whereas expression levels of *NANOG* were substantially downregulated. A regulatory link between Rex and *Oct4* expression has been reported in mES cells (Kim et al. 2008), whether this link also exists in hES cells yet remains unclear, however it could provide an explanation to the observed phenotype in Rex1OE cells by direct regulation of *OCT4* levels by REX1. Interestingly a binding site for *SMAD2/3* has also been mapped near the *REX1* locus, if this binding site acts to regulate the expression of *REX1*, it could explain the increase of *REX1* expression upon growth culture change in hES. Moreover, Rex1OE cells could be thus less dependent on Activin A signalling. Interestingly evidence from the literature also suggests a possible connection between REX1 and P300. P300 has been shown to interact with YY1 to modulate transcription (Lee et al. 1995). Binding of YY1 to P300 and CBP has been mapped to zinc finger domains (Austen et al. 1997), which are highly conserved between YY1 and REX1 (Kim et al. 2007). Taken together, these preliminary data suggest a putative role for *REX1* in the regulation of hES pluripotent identity. However, the nature of the potential regulation remains undetermined. Further characterisation of the Rex1OE and F-Rex1OE hES cell lines possibly utilising the FLAG-epitope in chromatin binding site mapping and the ability of

establishing revertant cell lines by Cre-Lox excision of the transgene should provide more insight to the role of *REX1* in the regulation of hES cells. Furthermore, it would be highly interesting to elucidate the possible roles REX1 might play during X-inactivation in hES cells as was recently shown in mES cells (Gontan et al. 2012), utilising the established RexOE and Flag-RexOE hES cell lines.

References

Aasen, T., Raya, A., et al. (2008). *Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes*. Nat Biotech **26**(11): 1276-1284. http://www.nature.com/nbt/journal/v26/n11/suppinfo/nbt.1503_S1.html

Abyzov, A., Mariani, J., et al. (2012). *Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells*. Nature **advance online publication**. http://www.nature.com/nature/journal/vaop/ncurrent/abs/nature11629.html - supplementary-information

Adewumi, O., Aflatoonian, B., et al. (2007). *Characterization of human embryonic stem cell lines by the International Stem Cell Initiative*. Nature biotechnology **25**(7): 803.

Ahlgren, U., Jonsson, J., et al. (1996). *The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice*. Development **122**(5): 1409-1416.

Ahmad, K. and Henikoff, S. (2002a). *Histone H3 variants specify modes of chromatin assembly*. Proceedings of the National Academy of Sciences of the United States of America **99**(Suppl 4): 16477-16484. 10.1073/pnas.172403699

Ahmad, K. and Henikoff, S. (2002b). *The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly*. Molecular Cell **9**(6): 1191-1200. 10.1016/s1097-2765(02)00542-7

Ait-Si-Ali, S., Polesskaya, A., et al. (2000). *CBP/p300 histone acetyl-transferase activity is important for the G1/S transition*. Oncogene **19**(20): 2430-2437.

Allfrey, V. G., Faulkner, R., et al. (1964). *ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*. Proceedings of the National Academy of Sciences **51**(5): 786-794.

Amit, M., Carpenter, M. K., et al. (2000). *Clonally Derived Human Embryonic Stem Cell Lines Maintain Pluripotency and Proliferative Potential for Prolonged Periods of Culture*. Developmental Biology **227**(2): 271-278. 10.1006/dbio.2000.9912

Amps, K., Andrews, P. W., et al. (2011). *Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage*. Nat Biotech **29**(12): 1132-1144. http://www.nature.com/nbt/journal/v29/n12/abs/nbt.2051.html - supplementary-information

An, W., Leuba, S. H., et al. (1998). *Linker histone protects linker DNA on only one side of the core particle and in a sequence-dependent manner*. Proceedings of the National Academy of Sciences **95**(7): 3396-3401.

Aparicio, J. G., Viggiani, C. J., et al. (2004). *The Rpd3-Sin3 Histone Deacetylase Regulates Replication Timing and Enables Intra-S Origin Control in Saccharomyces cerevisiae*. Molecular and Cellular Biology **24**(11): 4769-4780. 10.1128/mcb.24.11.4769-4780.2004

Aparicio, O. M., Billington, B. L., et al. (1991). *Modifiers of position effect are shared between telomeric and silent mating-type loci in S. cerevisiae*. Cell **66**(6): 1279-1287. 10.1016/0092-8674(91)90049-5

Arney, K. L. and Fisher, A. G. (2004). *Epigenetic aspects of differentiation*. Journal of Cell Science **117**(19): 4355-4363. 10.1242/jcs.01390

Arnold, S. J. and Robertson, E. J. (2009). *Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo*. Nature Reviews Molecular Cell Biology **10**(2): 91-103. 10.1038/nrm2618

Austen, M., Lüscher, B., et al. (1997). *Characterization of the Transcriptional Regulator YY1*. Journal of Biological Chemistry **272**(3): 1709-1717. 10.1074/jbc.272.3.1709

Avilion, A. A., Nicolis, S. K., et al. (2003). *Multipotent cell lineages in early mouse development depend on SOX2 function*. Genes & Development **17**(1): 126-140. 10.1101/gad.224503

Azuara, V. (2006). *Profiling of DNA replication timing in unsynchronized cell populations*. Nature Protocols **1**(4): 2171-2177.

Azuara, V., Brown, K. E., et al. (2003). *Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication*. Nature Cell Biology **5**(7): 668-674.

Azuara, V., Perry, P., et al. (2006). *Chromatin signatures of pluripotent cell lines*. Nat Cell Biol **8**(5): 532-538.

Babaie, Y., Herwig, R., et al. (2007). *Analysis of Oct4-Dependent Transcriptional Networks Regulating Self-Renewal and Pluripotency in Human Embryonic Stem Cells*. Stem Cells **25**(2): 500-510. 10.1634/stemcells.2006-0426

Baker, D. E., Harrison, N. J., et al. (2007a). *Adaptation to culture of human embryonic stem cells and oncogenesis in vivo*. Nat Biotechnol **25**(2): 207-215.

Baker, D. E. C., Harrison, N. J., et al. (2007b). *Adaptation to culture of human embryonic stem cells and oncogenesis in vivo*. Nat Biotech **25**(2): 207-215. http://www.nature.com/nbt/journal/v25/n2/suppinfo/nbt1285_S1.html

Balasubramanyam, K., Varier, R. A., et al. (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*. Journal of Biological Chemistry **279**(49): 51163-51171. 10.1074/jbc.M409024200

Bannister, A. J. and Kouzarides, T. (2005). *Reversing histone methylation*. Nature **436**(7054): 1103-1106.

Bannister, A. J. and Kouzarides, T. (2011). *Regulation of chromatin by histone modifications*. Cell Res **21**(3): 381-395.

Bannister, A. J., Zegerman, P., et al. (2001). *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain*. Nature **410**(6824): 120-124.

Bao, S., Tang, F., et al. (2009). *Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells*. Nature **461**(7268): 1292-1295. 10.1038/nature08534

Bar-Nur, O., Russ, Holger A., et al. (2011). *Epigenetic Memory and Preferential Lineage-Specific Differentiation in Induced Pluripotent Stem Cells Derived from Human Pancreatic Islet Beta Cells*. Cell Stem Cell **9**(1): 17-23. 10.1016/j.stem.2011.06.007

Barski, A., Cuddapah, S., et al. (2007). *High-Resolution Profiling of Histone Methylations in the Human Genome*. Cell **129**(4): 823-837. 10.1016/j.cell.2007.05.009

Bates, D. L. and Thomas, J. O. (1981). *Histories H1 and H5: one or two molecules per nucleosome?* Nucleic Acids Research **9**(22): 5883-5894. 10.1093/nar/9.22.5883

Bauer, U.-M., Daujat, S., et al. (2002). *Methylation at arginine 17 of histone H3 is linked to gene activation*. EMBO Rep **3**(1): 39-44.

Beattie, G. M., Lopez, A. D., et al. (2005). *Activin A Maintains Pluripotency of Human Embryonic Stem Cells in the Absence of Feeder Layers*. Stem Cells **23**(4): 489-495. 10.1634/stemcells.2004-0279

Beck, F., Erler, T., et al. (1995). *Expression of Cdx-2 in the mouse embryo and placenta: Possible role in patterning of the extra-embryonic membranes*. Developmental Dynamics **204**(3): 219-227. 10.1002/aja.1002040302

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**(4): 733-737.

Bedford, D. C. and Brindle, P. K. (2012). *Is histone acetylation the most important physiological function for CBP and p300?* Aging **4**(4): 1-9.

Bell, O., Tiwari, V. K., et al. (2011). *Determinants and dynamics of genome accessibility*. Nat Rev Genet **12**(8): 554-564.

Bendall, S. C., Stewart, M. H., et al. (2007). *IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro*. Nature **448**(7157): 1015-1021. http://www.nature.com/nature/journal/v448/n7157/suppinfo/nature06027_S1.html

Bernardo, Andreia S., Faial, T., et al. (2011). *BRACHYURY and CDX2 Mediate BMP-Induced Differentiation of Human and Mouse Pluripotent Stem Cells into Embryonic and Extraembryonic Lineages*. Cell Stem Cell **9**(2): 144-155.

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

Bernstein, B. E., Meissner, A., et al. (2007). *The Mammalian Epigenome*. Cell **128**(4): 669-681. 10.1016/j.cell.2007.01.033

Besser, D. (2004). *Expression of Nodal, Lefty-A, and Lefty-B in Undifferentiated Human Embryonic Stem Cells Requires Activation of Smad2/3*. Journal of Biological Chemistry **279**(43): 45076-45084. 10.1074/jbc.M404979200

Bhaumik, S. R., Smith, E., et al. (2007). *Covalent modifications of histones during development and disease pathogenesis*. Nat Struct Mol Biol **14**(11): 1008-1016.

Bhutani, N., Brady, J. J., et al. (2010). *Reprogramming towards pluripotency requires AID-dependent DNA demethylation*. Nature **463**(7284): 1042-1047. http://www.nature.com/nature/journal/v463/n7284/suppinfo/nature08752_S1.html

Bhutani, N., Burns, David M., et al. (2011). *DNA Demethylation Dynamics*. Cell **146**(6): 866-872. 10.1016/j.cell.2011.08.042

Bickmore, W. A. and Carothers, A. D. (1995). *Factors affecting the timing and imprinting of replication on a mammalian chromosome*. Journal of Cell Science **108**(8): 2801-2809.

Bielinska, M., Narita, N., et al. (1999). *Distinct roles for visceral endoderm during embryonic mouse development*. The international journal of developmental biology **43**(3): 183-205.

Bird, A. (2002). *DNA methylation patterns and epigenetic memory*. Genes & Development **16**(1): 6-21. 10.1101/gad.947102

Birger, Y., Shemer, R., et al. (1999). *The imprinting box of the mouse Igf2r gene*. Nature **397**(6714): 84-88.

Boland, M. J., Hazen, J. L., et al. (2009). *Adult mice generated from induced pluripotent stem cells*. Nature **461**(7260): 91-94. 10.1038/nature08310

Bongso, A., Fong, C.-Y., et al. (1994). *Fertilization and early embryology: Isolation and culture of inner cell mass cells from human blastocysts*. Hum. Reprod. **9**(11): 2110-2117.

Booth, C. and Potten, C. S. (2000). *Gut instincts: thoughts on intestinal epithelial stem cells*. The Journal of Clinical Investigation **105**(11): 1493-1499.

Bortvin, A., Eggan, K., et al. (2003). *Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei*. Development **130**(8): 1673-1680.

Bowers, E. M., Yan, G., et al. (2010). *Virtual Ligand Screening of the p300/CBP Histone Acetyltransferase: Identification of a Selective Small Molecule Inhibitor*. Chemistry & Biology **17**(5): 471-482. DOI: 10.1016/j.chembiol.2010.03.006

Boyer, L. A., Lee, T. I., et al. (2005). *Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells*. Cell **122**(6): 947-956. DOI: 10.1016/j.cell.2005.08.020

Boyer, L. A., Plath, K., et al. (2006). *Polycomb complexes repress developmental regulators in murine embryonic stem cells*. Nature **441**(7091): 349-353. http://www.nature.com/nature/journal/v441/n7091/suppinfo/nature04733_S1.html

Boyes, J. and Bird, A. (1991). *DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein*. Cell **64**(6): 1123-1134. 10.1016/0092-8674(91)90267-3

Bradley, A., Evans, M., et al. (1984). *Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines*. Nature **309**(5965): 255-256. 10.1038/309255a0

Branco, M. R. and Pombo, A. (2006). *Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations*. PLoS Biol **4**(5): e138. 10.1371/journal.pbio.0040138

Braun, T., Buschhausen-Denker, G., et al. (1989). *A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts*. EMBO journal **8**(3): 701.

Brons, I. G. M., Smithers, L. E., et al. (2007). *Derivation of pluripotent epiblast stem cells from mammalian embryos*. Nature **448**(7150): 191-195. 10.1038/nature05950

Brook, F. A. and Gardner, R. L. (1997). *The origin and efficient derivation of embryonic stem cells in the mouse*. Proceedings of the National Academy of Sciences **94**(11): 5709-5712. 10.1073/pnas.94.11.5709

Brookes, E., de Santiago, I., et al. (2012). *Polycomb Associates Genome-wide with a Specific RNA Polymerase II Variant, and Regulates Metabolic Genes in ESCs*. Cell Stem Cell **10**(2): 157-170.

Brown, S., Teo, A., et al. (2011). *Activin/Nodal Signaling Controls Divergent Transcriptional Networks in Human Embryonic Stem Cells and in Endoderm Progenitors*. Stem Cells **29**(8): 1176-1185. 10.1002/stem.666

Brownell, J. E., Zhou, J., et al. (1996). *Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation*. Cell **84**(6): 843-851. 10.1016/s0092-8674(00)81063-6

Bruck, T. and Benvenisty, N. (2011). *Meta-analysis of the heterogeneity of X chromosome inactivation in human pluripotent stem cells*. Stem Cell Research **6**(2): 187-193. 10.1016/j.scr.2010.12.001

Burke, B. and Stewart, C. L. (2002). *Life at the edge: the nuclear envelope and human disease*. Nat Rev Mol Cell Biol **3**(8): 575-585.

Byrne, J. A., Pedersen, D. A., et al. (2007). *Producing primate embryonic stem cells by somatic cell nuclear transfer*. Nature **450**(7169): 497-502. 10.1038/nature06357

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**(5): 757.

Cao, R. and Zhang, Y. (2004). *The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3*. Current Opinion in Genetics & Development **14**(2): 155-164. 10.1016/j.gde.2004.02.001

Carlson, L. L., Page, A. W., et al. (1992). *Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting*. Genes & Development **6**(12b): 2536-2541. 10.1101/gad.6.12b.2536

Cedar, H. and Bergman, Y. (2012). *Programming of DNA Methylation Patterns*. Annual Review of Biochemistry **81**(1): 97-117. doi:10.1146/annurev-biochem-052610-091920

Chambers, I., Colby, D., et al. (2003). *Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells*. Cell **113**(5): 643-655. Doi: 10.1016/s0092-8674(03)00392-1

Chambers, I., Silva, J., et al. (2007). *Nanog safeguards pluripotency and mediates germline development*. Nature **450**(7173): 1230-1234. 10.1038/nature06403

Chan, E. M., Ratanasirintrawoot, S., et al. (2009). *Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells*. Nature Biotechnology **27**(11): 1033-1037. 10.1038/nbt.1580

Chang, B., Chen, Y., et al. (2007). *JMJD6 Is a Histone Arginine Demethylase*. Science **318**(5849): 444-447. 10.1126/science.1145801

Chapman, G., Remiszewski, J. L., et al. (1997). *The mouse homeobox gene, Gbx2: genomic organization and expression in pluripotent cells in vitro and in vivo*. Genomics **46**(2): 223-233.

Chen, G., Gulbranson, D. R., et al. (2011). *Chemically defined conditions for human iPSC derivation and culture*. Nat Meth **8**(5): 424-429. http://www.nature.com/nmeth/journal/v8/n5/abs/nmeth.1593.html - supplementary-information

Chen, Y.-G., Li, Z., et al. (2012). *Where PI3K/Akt Meets Smads: The Crosstalk Determines Human Embryonic Stem Cell Fate*. Cell Stem Cell **10**(3): 231-232.

Chin, A. C. P., Fong, W. J., et al. (2007). *Identification of proteins from feeder conditioned medium that support human embryonic stem cells*. Journal of Biotechnology **130**(3): 320-328. 10.1016/j.jbiotec.2007.04.013

Chou, Y.-F., Chen, H.-H., et al. (2008). *The Growth Factor Environment Defines Distinct Pluripotent Ground States in Novel Blastocyst-Derived Stem Cells*. Cell **135**(3): 449-461. 10.1016/j.cell.2008.08.035

Chuang, C.-H., Carpenter, A. E., et al. (2006). *Long-Range Directional Movement of an Interphase Chromosome Site*. Current Biology **16**(8): 825-831. 10.1016/j.cub.2006.03.059

Chuang, L. S.-H., Ian, H.-I., et al. (1997). *Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for p21WAF1*. Science **277**(5334): 1996-2000. 10.1126/science.277.5334.1996

Chung, Y., Klimanskaya, I., et al. (2006). *Embryonic and extraembryonic stem cell lines derived from single mouse blastomeres*. Nature **439**(7073): 216-219. http://www.nature.com/nature/journal/v439/n7073/suppinfo/nature04277_S1.html

Cimbora, D. M. and Groudine, M. (2001). *The Control of Mammalian DNA Replication: A Brief History of Space and Timing*. Cell **104**(5): 643-646. 10.1016/s0092-8674(01)00260-4

Ciruna, B. G. and Rossant, J. (1999). *Expression of the T-box gene Eomesodermin during early mouse development*. Mechanisms of Development **81**(1–2): 199-203. 10.1016/s0925-4773(98)00243-3

Clapier, C. R. and Cairns, B. R. (2009). *The Biology of Chromatin Remodeling Complexes*. Annual Review of Biochemistry **78**(1): 273-304. doi:10.1146/annurev.biochem.77.062706.153223

Clark, A. T., Rodriguez, R. T., et al. (2004). *Human STELLAR, NANOG, and GDF3 Genes Are Expressed in Pluripotent Cells and Map to Chromosome 12p13, a Hotspot for Teratocarcinoma*. Stem Cells **22**(2): 169-179. 10.1634/stemcells.22-2-169

Collignon, J., Sockanathan, S., et al. (1996). *A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2*. Development **122**(2): 509-520.

Colman, A. and Dreesen, O. (2009). *Induced pluripotent stem cells and the stability of the differentiated state*. EMBO reports **10**(7): 714-721. 10.1038/embor.2009.142

Copeland, R. A., Solomon, M. E., et al. (2009). *Protein methyltransferases as a target class for drug discovery*. Nat Rev Drug Discov **8**(9): 724-732.

Costantini, M. and Bernardi, G. (2008). *Replication timing, chromosomal bands, and isochores*. Proceedings of the National Academy of Sciences **105**(9): 3433-3437. 10.1073/pnas.0710587105

Covault, J. and Chalkley, R. (1980). *The identification of distinct populations of acetylated histone*. Journal of Biological Chemistry **255**(19): 9110-9116.

Cowan, C. A., Atienza, J., et al. (2005). *Nuclear Reprogramming of Somatic Cells After Fusion with Human Embryonic Stem Cells*. Science **309**(5739): 1369-1373. 10.1126/science.1116447

Cremer, T., Cremer, M., et al. (2006). *Chromosome territories – a functional nuclear landscape*. Current Opinion in Cell Biology **18**(3): 307-316. 10.1016/j.ceb.2006.04.007

Creyghton, M. P., Markoulaki, S., et al. (2008). *H2AZ Is Enriched at Polycomb Complex Target Genes in ES Cells and Is Necessary for Lineage Commitment*. Cell **135**(4): 649-661. 10.1016/j.cell.2008.09.056

Daheron, L., Opitz, S. L., et al. (2004). *LIF/STAT3 Signaling Fails to Maintain Self-Renewal of Human Embryonic Stem Cells*. Stem Cells **22**(5): 770-778. 10.1634/stemcells.22-5-770

Darr, H., Mayshar, Y., et al. (2006). *Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features*. Development **133**(6): 1193-1201. 10.1242/dev.02286

Das, C., Lucia, M. S., et al. (2009). *CBP/p300-mediated acetylation of histone H3 on lysine 56*. Nature **459**(7243): 113-117. 10.1038/nature07861

Davidson, K. C., Adams, A. M., et al. (2012). *Wnt/β-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4*. Proceedings of the National Academy of Sciences **109**(12): 4485-4490. 10.1073/pnas.1118777109

de Napoles, M., Mermoud, J. E., et al. (2004). *Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation*. Developmental Cell **7**(5): 663-676. 10.1016/j.devcel.2004.10.005

De, S. and Babu, M. M. (2010). *Genomic neighbourhood and the regulation of gene expression*. Current Opinion in Cell Biology **22**(3): 326-333. 10.1016/j.ceb.2010.04.004

Dechat, T., Pfleghaar, K., et al. (2008). *Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin*. Genes & Development **22**(7): 832-853. 10.1101/gad.1652708

Deniaud, E. and Bickmore, W. A. (2009). *Transcription and the nuclear periphery: edge of darkness?* Current Opinion in Genetics & Development **19**(2): 187-191. 10.1016/j.gde.2009.01.005

Dillon, N. and Festenstein, R. (2002). *Unravelling heterochromatin: competition between positive and negative factors regulates accessibility*. Trends in Genetics **18**(5): 252-258. 10.1016/s0168-9525(02)02648-3

Dimitrova, D. S. and Gilbert, D. M. (1999). *The Spatial Position and Replication Timing of Chromosomal Domains Are Both Established in Early G1 Phase*. Molecular Cell **4**(6): 983-993.

Dimos, J. T., Rodolfa, K. T., et al. (2008). *Induced Pluripotent Stem Cells Generated from Patients with ALS Can Be Differentiated into Motor Neurons*. Science **321**(5893): 1218-1221. 10.1126/science.1158799

Djebali, S., Davis, C. A., et al. (2012). *Landscape of transcription in human cells*. Nature **489**(7414): 101-108. http://www.nature.com/nature/journal/v489/n7414/abs/nature11233.html - supplementary-information

Doetschman, T. C., Eistetter, H., et al. (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**(1): 27-45.

Donaldson, A. (2005). *Shaping time: chromatin structure and the DNA replication programme*. Trends Genet **21**: 444 - 449.

Dostie, J. and Bickmore, W. A. (2012). *Chromosome organization in the nucleus – charting new territory across the Hi-Cs*. Current Opinion in Genetics & Development **22**(2): 125-131. 10.1016/j.gde.2011.12.006

Draper, J. S., Moore, H. D., et al. (2004). *Culture and Characterization of Human Embryonic Stem Cells*. Stem Cells and Development **13**(4): 325-336. doi:10.1089/scd.2004.13.325

Dunham, I., Kundaje, A., et al. (2012). *An integrated encyclopedia of DNA elements in the human genome*. Nature **489**(7414): 57-74. http://www.nature.com/nature/journal/v489/n7414/abs/nature11247.html - supplementary-information

Dutrillaux, B., Couturier, J., et al. (1976). *Sequence of DNA replication in 277 R- and Q-bands of human chromosomes using a BrdU treatment*. Chromosoma **58**(1): 51-61. 10.1007/bf00293440

Efroni, S., Duttagupta, R., et al. (2008). *Global Transcription in Pluripotent Embryonic Stem Cells*. Cell Stem Cell **2**(5): 437-447.

Eistetter, H. R. (1989). *Pluripotent Embryonal Stem Cell Lines Can Be Established from Disaggregated Mouse Morulae*. Development, growth & differentiation **31**(3): 275-282. 10.1111/j.1440-169X.1989.00275.x

Enver, T., Pera, M., et al. (2009). *Stem Cell States, Fates, and the Rules of Attraction*. Cell Stem Cell **4**(5): 387-397. 10.1016/j.stem.2009.04.011

Enver, T., Soneji, S., et al. (2005). *Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells*. Hum. Mol. Genet. **14**(21): 3129-3140. 10.1093/hmg/ddi345

Evans, M. J. and Kaufman, M. H. (1981). *Establishment in culture of pluripotential cells from mouse embryos*. Nature **292**(5819): 154-156.

Fabbrizio, E., Messaoudi, S. E., et al. (2002). *Negative regulation of transcription by the type II arginine methyltransferase PRMT5*. EMBO Rep **3**(7): 641-645. http://www.nature.com/embor/journal/v3/n7/suppinfo/embor122_S1.html

Farkash-Amar, S., Lipson, D., et al. (2008). *Global organization of replication time zones of the mouse genome*. Genome Research **18**(10): 1562-1570. 10.1101/gr.079566.108

Feldman, N., Gerson, A., et al. (2006). *G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis*. Nat Cell Biol **8**(2): 188-194. http://www.nature.com/ncb/journal/v8/n2/suppinfo/ncb1353_S1.html

Felsenfeld, G. and Groudine, M. (2003). *Controlling the double helix*. Nature **421**(6921): 448-453.

Ferrai, C., de Castro, I. J., et al. (2010). *Gene Positioning*. Cold Spring Harbor Perspectives in Biology **2**(6). 10.1101/cshperspect.a000588

Festuccia, N., Osorno, R., et al. (2012). *Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells*. Cell Stem Cell **11**(4): 477-490.

Ficz, G., Branco, M. R., et al. (2011). *Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation*. Nature **473**(7347): 398-402. http://www.nature.com/nature/journal/v473/n7347/abs/10.1038-nature10008-unlocked.html - supplementary-information

Fischer, Y., Ganic, E., et al. (2010). *NANOG Reporter Cell Lines Generated by Gene Targeting in Human Embryonic Stem Cells*. PLoS ONE **5**(9): e12533.

Fisher, A. G. and Brockdorff, N. (2012). *Epigenetic memory and parliamentary privilege combine to evoke discussions on inheritance*. Development **139**(21): 3891-3896. 10.1242/dev.084434

Fong, H., Hohenstein, K. A., et al. (2008). *Regulation of Self-Renewal and Pluripotency by Sox2 in Human Embryonic Stem Cells*. Stem Cells **26**(8): 1931-1938.

Fraser, P. and Bickmore, W. (2007). *Nuclear organization of the genome and the potential for gene regulation*. Nature **447**(7143): 413-417.

Gardner, D. K., Vella, P., et al. (1998). *Culture and transfer of human blastocysts increases implantation rates and reduces the need for multiple embryo transfers*. Fertility and Sterility **69**(1): 84-88.

Gardner, R. L. (1998). *Contributions of blastocyst micromanipulation to the study of mammalian development*. BioEssays **20**(2): 168-180. 10.1002/(sici)1521-1878(199802)20:2<168::aid-bies9>3.0.co;2-p

Gaspar-Maia, A., Alajem, A., et al. (2011). *Open chromatin in pluripotency and reprogramming*. Nat Rev Mol Cell Biol **12**(1): 36-47.

Gaspar-Maia, A., Alajem, A., et al. (2009). *Chd1 regulates open chromatin and pluripotency of embryonic stem cells*. Nature **460**(7257): 863-868. http://www.nature.com/nature/journal/v460/n7257/suppinfo/nature08212_S1.html

Gasser, S. M. (2002). *Visualizing Chromatin Dynamics in Interphase Nuclei*. Science **296**(5572): 1412-1416. 10.1126/science.1067703

Georgopoulos, K., Moore, D., et al. (1992). *Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment*. Science **258**(5083): 808-812. 10.1126/science.1439790

Gerami-Naini, B., Dovzhenko, O. V., et al. (2004). *Trophoblast Differentiation in Embryoid Bodies Derived from Human Embryonic Stem Cells*. Endocrinology **145**(4): 1517-1524. 10.1210/en.2003-1241

Gerrard, L., Rodgers, L., et al. (2005a). *Differentiation of Human Embryonic Stem Cells to Neural Lineages in Adherent Culture by Blocking Bone Morphogenetic Protein Signaling*. Stem Cells **23**(9): 1234-1241. 10.1634/stemcells.2005-0110

Gerrard, L., Rodgers, L., et al. (2005b). *Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling*. Stem Cells **23**(9): 1234-1241.

Giadrossi, S., Dvorkina, M., et al. (2007). *Chromatin organization and differentiation in embryonic stem cell models*. Current Opinion in Genetics & Development **17**(2): 132-138. 10.1016/j.gde.2007.02.012

Gilbert, D. M. (2002). *Replication timing and transcriptional control: beyond cause and effect*. Current Opinion in Cell Biology **14**(3): 377-383.

Gilbert, D. M. (2010). *Evaluating genome-scale approaches to eukaryotic DNA replication*. Nat Rev Genet **11**(10): 673-684.

Gilbert, N., Boyle, S., et al. (2004). *Chromatin Architecture of the Human Genome: Gene-Rich Domains Are Enriched in Open Chromatin Fibers*. Cell **118**(5): 555-566. 10.1016/j.cell.2004.08.011

Goll, M. G. and Bestor, T. H. (2005). *EUKARYOTIC CYTOSINE METHYLTRANSFERASES*. Annual Review of Biochemistry **74**(1): 481-514. doi:10.1146/annurev.biochem.74.010904.153721

Gontan, C., Achame, E. M., et al. (2012). *RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation*. Nature **advance online publication**. http://www.nature.com/nature/journal/vaop/ncurrent/abs/nature11070.html - supplementary-information

González, F., Boué, S., et al. (2011). *Methods for making induced pluripotent stem cells: reprogramming à la carte*. Nat Rev Genet **12**(4): 231-242. http://www.nature.com/nrg/journal/v12/n4/suppinfo/nrg2937_S1.html

Gore, A., Li, Z., et al. (2011). *Somatic coding mutations in human induced pluripotent stem cells*. Nature **471**(7336): 63-67.

Goren, A. and Cedar, H. (2003). *Replicating by the Clock*. Nat Rev Mol Cell Biol **4**(1): 25-32. 10.1038/nrm1008

Goren, A., Tabib, A., et al. (2008). *DNA replication timing of the human β-globin domain is controlled by histone modification at the origin*. Genes & Development **22**(10): 1319-1324. 10.1101/gad.468308

Gottschling, D. E., Aparicio, O. M., et al. (1990). *Position effect at S. cerevisiae telomeres: Reversible repression of Pol II transcription*. Cell **63**(4): 751-762. 10.1016/0092-8674(90)90141-z

Greber, B., Lehrach, H., et al. (2007). *Fibroblast Growth Factor 2 Modulates Transforming Growth Factor β Signaling in Mouse Embryonic Fibroblasts and Human ESCs (hESCs) to Support hESC Self-Renewal*. Stem Cells **25**(2): 455-464. 10.1634/stemcells.2006-0476

Greber, B., Wu, G., et al. (2010). *Conserved and Divergent Roles of FGF Signaling in Mouse Epiblast Stem Cells and Human Embryonic Stem Cells*. Cell Stem Cell **6**(3): 215-226. DOI: 10.1016/j.stem.2010.01.003

Gruenbaum, Y., Cedar, H., et al. (1982). *Substrate and sequence specificity of a eukaryotic DNA methylase*. Nature **295**(5850): 620-622.

Gu, W. and Roeder, R. G. (1997). *Activation of p53 Sequence-Specific DNA Binding by Acetylation of the p53 C-Terminal Domain*. Cell **90**(4): 595-606. 10.1016/s0092-8674(00)80521-8

Guallar, D., Pérez-Palacios, R., et al. (2012). *Expression of endogenous retroviruses is negatively regulated by the pluripotency marker Rex1/Zfp42*. Nucleic Acids Research: 1-15. 10.1093/nar/gks686

Guelen, L., Pagie, L., et al. (2008). *Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions*. Nature **453**(7197): 948-951. http://www.nature.com/nature/journal/v453/n7197/suppinfo/nature06947_S1.html

Guenther, M. G., Frampton, G. M., et al. (2010). *Chromatin Structure and Gene Expression Programs of Human Embryonic and Induced Pluripotent Stem Cells*. Cell Stem Cell **7**(2): 249-257. 10.1016/j.stem.2010.06.015

Guo, G., Yang, J., et al. (2009). *Klf4 reverts developmentally programmed restriction of ground state pluripotency*. Development **136**(7): 1063-1069. 10.1242/dev.030957

Gurdon, J. B. (1962). *The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles*. Journal of Embryology and Experimental Morphology **10**(4): 622-640.

Haigis, M. C. and Guarente, L. P. (2006). *Mammalian sirtuins—emerging roles in physiology, aging, and calorie restriction*. Genes & Development **20**(21): 2913-2921. 10.1101/gad.1467506

Han, D. W., Greber, B., et al. (2011). *Direct reprogramming of fibroblasts into epiblast stem cells*. Nat Cell Biol **13**(1): 66-71. http://www.nature.com/ncb/journal/v13/n1/abs/ncb2136.html - supplementary-information

Han, D. W., Tapia, N., et al. (2010). *Epiblast Stem Cell Subpopulations Represent Mouse Embryos of Distinct Pregastrulation Stages*. Cell.

Hanna, J., Cheng, A. W., et al. (2010a). *Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs*. Proceedings of the National Academy of Sciences **107**(20): 9222-9227. 10.1073/pnas.1004584107

Hanna, J., Markoulaki, S., et al. (2009). *Metastable Pluripotent States in NOD-Mouse-Derived ESCs*. Cell Stem Cell **4**(6): 513-524.

Hanna, J. H., Saha, K., et al. (2010b). *Pluripotency and Cellular Reprogramming: Facts, Hypotheses, Unresolved Issues*. Cell **143**(4): 508-525.

Hansen, R. S., Canfield, T. K., et al. (1993). *Association of fragile X syndrome with delayed replication of the FMR1 gene*. Cell **73**(7): 1403-1409.

Hashimshony, T., Zhang, J., et al. (2003). *The role of DNA methylation in setting up chromatin structure during development*. Nat Genet **34**(2): 187-192.

Hassan, A. H., Prochasson, P., et al. (2002). *Function and Selectivity of Bromodomains in Anchoring Chromatin-Modifying Complexes to Promoter Nucleosomes*. Cell **111**(3): 369-379. 10.1016/s0092-8674(02)01005-x

Hatton, K. S., Dhar, V., et al. (1988). *Replication program of active and inactive multigene families in mammalian cells*. Mol. Cell. Biol. **8**(5): 2149-2158.

Haub, O. and Goldfarb, M. (1991). *Expression of the fibroblast growth factor-5 gene in the mouse embryo*. Development **112**(2): 397-406.

Hay, D. C., Sutherland, L., et al. (2004). *Oct-4 Knockdown Induces Similar Patterns of Endoderm and Trophoblast Differentiation Markers in Human and Mouse Embryonic Stem Cells*. Stem Cells **22**(2): 225-235. 10.1634/stemcells.22-2-225

Hayashi, K., Lopes, S. M. C. d. S., et al. (2008). *Dynamic Equilibrium and Heterogeneity of Mouse Pluripotent Stem Cells with Distinct Functional and Epigenetic States*. Cell **3**(4): 391-401.

He, H., Yu, F.-X., et al. (2011). *CBP/p300 and SIRT1 Are Involved in Transcriptional Regulation of S-Phase Specific Histone Genes*. PLoS ONE **6**(7): e22088. 10.1371/journal.pone.0022088

Heard, E. (2005). *Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome*. Current Opinion in Genetics & Development **15**(5): 482-489. 10.1016/j.gde.2005.08.009

Hebbes, T. R., Thorne, A. W., et al. (1988). *A direct link between core histone acetylation and transcriptionally active chromatin.* The EMBO journal **7**(5): 1395-1402.

Hebert, J. M., Boyle, M., et al. (1991). *mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation*. Development **112**(2): 407-415.

Hecht, A., Vleminckx, K., et al. (2000). *The p300/CBP acetyltransferases function as transcriptional coactivators of [beta]-catenin in vertebrates*. EMBO J **19**(8): 1839-1850. http://www.nature.com/emboj/journal/v19/n8/suppinfo/7593010a_S1.html

Hemberger, M., Dean, W., et al. (2009). *Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal*. Nature Reviews Molecular Cell Biology **10**(8): 526-537. 10.1038/nrm2727

Henderson, J. K., Draper, J. S., et al. (2002). *Preimplantation Human Embryos and Embryonic Stem Cells Show Comparable Expression of Stage-Specific Embryonic Antigens*. Stem Cells **20**(4): 329-337. 10.1634/stemcells.20-4-329

Hendrich, B. and Bird, A. (1998). *Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins*. Molecular and Cellular Biology **18**(11): 6538-6547.

Henikoff, S. (2008). *Nucleosome destabilization in the epigenetic regulation of gene expression*. Nat Rev Genet **9**(1): 15-26.

Hermann, A., Schmitt, S., et al. (2003). *The Human Dnmt2 Has Residual DNA-(Cytosine-C5) Methyltransferase Activity*. Journal of Biological Chemistry **278**(34): 31717-31721. 10.1074/jbc.M305448200

Hiragami-Hamada, K., Xie, S., et al. (2009). *The molecular basis for stability of heterochromatin-mediated silencing in mammals*. Epigenetics & Chromatin **2**(1): 14.

Hiratani, I. and Gilbert, D. (2009a). *Replication timing as an epigenetic mark*. Epigenetics **4**(2): 93-97.

Hiratani, I., Leskovar, A., et al. (2004). *Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores*. Proceedings of the National Academy of Sciences **101**(48): 16861-16866. 10.1073/pnas.0406687101

Hiratani, I., Ryba, T., et al. (2010). *Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis*. Genome Research **20**(2): 155-169. 10.1101/gr.099796.109

Hiratani, I., Ryba, T., et al. (2008). *Global reorganization of replication domains during embryonic stem cell differentiation*. PLoS Biol **6**(10): e245.

Hiratani, I., Takebayashi, S.-i., et al. (2009b). *Replication timing and transcriptional control: beyond cause and effect—part II*. Current Opinion in Genetics & Development **19**(2): 142-149. 10.1016/j.gde.2009.02.002

Hochedlinger, K. and Jaenisch, R. (2002). *Monoclonal mice generated by nuclear transfer from mature B and T donor cells*. Nature **415**(6875): 1035-1038.

Hochedlinger, K. and Plath, K. (2009). *Epigenetic reprogramming and induced pluripotency*. Development **136**(4): 509-523. 10.1242/dev.020867

Hodawadekar, S. C. and Marmorstein, R. (2007). *Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design*. Oncogene **26**(37): 5528-5540.

Holmgren, C., Kanduri, C., et al. (2001). *CpG methylation regulates the Igf2/H19 insulator*. Current Biology **11**(14): 1128-1130. 10.1016/s0960-9822(01)00314-1

Holmquist, G. P. and Ashley, T. (2006). *Chromosome organization and chromatin modification: influence on genome function and evolution*. Cytogenetic and Genome Research **114**(2): 96-125.

Home, P., Ray, S., et al. (2009). *GATA3 Is Selectively Expressed in the Trophectoderm of Peri-implantation Embryo and Directly Regulates Cdx2 Gene Expression*. Journal of Biological Chemistry **284**(42): 28729-28737. 10.1074/jbc.M109.016840

Hooper, M., Hardy, K., et al. (1987). *HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells*. Nature **326**(6110): 292-295.

Hosler, B. A., LaRosa, G. J., et al. (1989). *Expression of REX-1, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells*. Mol. Cell. Biol. **9**(12): 5623-5629.

Howell, C. Y., Bestor, T. H., et al. (2001). *Genomic Imprinting Disrupted by a Maternal Effect Mutation in the Dnmt1 Gene*. Cell **104**(6): 829-838. 10.1016/s0092-8674(01)00280-x

Hu, M., Krause, D., et al. (1997). *Multilineage gene expression precedes commitment in the hemopoietic system*. Genes & Development **11**(6): 774-785. 10.1101/gad.11.6.774

Huang, Y., Fang, J., et al. (2006). *Recognition of Histone H3 Lysine-4 Methylation by the Double Tudor Domain of JMJD2A*. Science **312**(5774): 748-751. 10.1126/science.1125162

Huberman, J. A. and Riggs, A. D. (1968). *On the mechanism of DNA replication in mammalian chromosomes*. Journal of Molecular Biology **32**(2): 327-341.

Hussein, S. M., Batada, N. N., et al. (2011). *Copy number variation and selection during reprogramming to pluripotency*. Nature **471**(7336): 58-62.

Hyslop, L., Stojkovic, M., et al. (2005). *Downregulation of NANOG Induces Differentiation of Human Embryonic Stem Cells to Extraembryonic Lineages*. Stem Cells **23**(8): 1035-1043. 10.1634/stemcells.2005-0080

Itoh, S., Ericsson, J., et al. (2000). *The transcriptional co-activator P/CAF potentiates TGF-{beta}/Smad signaling*. Nucl. Acids Res. **28**(21): 4291-4298. 10.1093/nar/28.21.4291

Itskovitz-Eldor, J., Schuldiner, M., et al. (2000). *Differentiation of Human Embryonic Stem Cells into Embryoid Bodies Comprising the Three Embryonic Germ Layers*. Molecular Medicine **6**(2).

Jaenisch, R. and Young, R. (2008). *Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming*. Cell **132**(4): 567-582. 10.1016/j.cell.2008.01.015

James, D., Levine, A. J., et al. (2005). *TGF{beta}/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells*. Development **132**(6): 1273-1282. 10.1242/dev.01706

Jenuwein, T. and Allis, C. D. (2001). *Translating the Histone Code*. Science **293**(5532): 1074-1080. 10.1126/science.1063127

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**(2): 281-289. 10.1016/0092-8674(93)90419-q

Jiang, Y., Jahagirdar, B. N., et al. (2002). *Pluripotency of mesenchymal stem cells derived from adult marrow*. Nature **418**(6893): 41-49. http://www.nature.com/nature/journal/v418/n6893/suppinfo/nature00870_S1.html

Jin, C. and Felsenfeld, G. (2007). *Nucleosome stability mediated by histone variants H3.3 and H2A.Z*. Genes & Development **21**(12): 1519-1529. 10.1101/gad.1547707

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**(1): 141-151.

Jones, P. H., Simons, B. D., et al. (2007). *Sic Transit Gloria: Farewell to the Epidermal Transit Amplifying Cell?* Cell Stem Cell **1**(4): 371-381. 10.1016/j.stem.2007.09.014

Jørgensen, H. F., Azuara, V., et al. (2007). *The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells*. Genome Biology **8**(8): R169. 10.1186/gb-2007-8-8-r169

Jørgensen, H. F., Giadrossi, S., et al. (2006). *Polycomb Repressive Complexes Restrain the Expression of Lineage-Specific Regulators in Embryonic Stem Cells*. Cell cycle **5**(13): 1411-1414.

Kahata, K., Hayashi, M., et al. (2004). *Regulation of transforming growth factor-{beta} and bone morphogenetic protein signalling by transcriptional coactivator GCN5*. Genes to Cells **9**(2): 143-151. 10.1111/j.1365-2443.2004.00706.x

Kanatsu-Shinohara, M., Inoue, K., et al. (2004). *Generation of Pluripotent Stem Cells from Neonatal Mouse Testis*. Cell **119**(7): 1001-1012. 10.1016/j.cell.2004.11.011

Kangaspeska, S., Stride, B., et al. (2008). *Transient cyclical methylation of promoter DNA*. Nature **452**(7183): 112-115. http://www.nature.com/nature/journal/v452/n7183/suppinfo/nature06640_S1.html

Karnani, N., Taylor, C., et al. (2007). *Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas*. Genome Research **17**(6): 865-876. 10.1101/gr.5427007

Kikyo, N. and Wolffe, A. P. (2000). *Reprogramming nuclei: insights from cloning, nuclear transfer and heterokaryons*. Journal of Cell Science **113**(1): 11-20.

Kim, J., Chu, J., et al. (2008). *An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells*. Cell **132**(6): 1049-1061. DOI: 10.1016/j.cell.2008.02.039

Kim, J. D., Faulk, C., et al. (2007). *Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1*. Nucleic Acids Research **35**(10): 3442-3452. 10.1093/nar/gkm235

Kim, J. D., Kim, H., et al. (2011). *Rex1/Zfp42 as an epigenetic regulator for genomic imprinting*. Human Molecular Genetics. 10.1093/hmg/ddr017

Klimanskaya, I., Chung, Y., et al. (2006). *Human embryonic stem cell lines derived from single blastomeres*. Nature **444**(7118): 481-485. http://www.nature.com/nature/journal/v444/n7118/suppinfo/nature05142_S1.html

Klose, R. J. and Bird, A. P. (2006). *Genomic DNA methylation: the mark and its mediators*. Trends in Biochemical Sciences **31**(2): 89-97. 10.1016/j.tibs.2005.12.008

Klose, R. J. and Zhang, Y. (2007). *Regulation of histone methylation by demethylimination and demethylation*. Nat Rev Mol Cell Biol **8**(4): 307-318.

Koch, C. M., Andrews, R. M., et al. (2007). *The landscape of histone modifications across 1% of the human genome in five human cell lines*. Genome Research **17**(6): 691-707. 10.1101/gr.5704207

Kontgen, F., Suss, G., et al. (1993). *Targeted disruption of the MHC class II Aa gene in C57BL/6 mice*. Int Immunol **5**(8): 957-964.

Kornberg, R. D. (1977). *Structure of Chromatin*. Annual Review of Biochemistry **46**(1): 931-954. doi:10.1146/annurev.bi.46.070177.004435

Koutsourakis, M., Langeveld, A., et al. (1999). *The transcription factor GATA6 is essential for early extraembryonic development*. Development **126**(4): 723-732.

Kouzarides, T. (2007). *Chromatin Modifications and Their Function*. Cell **128**(4): 693-705. 10.1016/j.cell.2007.02.005

Kriaucionis, S. and Heintz, N. (2009). *The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain*. Science **324**(5929): 929-930. 10.1126/science.1169786

Krysan, P. J., Smith, J. G., et al. (1993). *Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences*. Molecular and Cellular Biology **13**(5): 2688-2696. 10.1128/mcb.13.5.2688

Kunarso, G., Chia, N.-Y., et al. (2010). *Transposable elements have rewired the core regulatory network of human embryonic stem cells*. Nature Genetics **42**(7): 631-634. 10.1038/ng.600

Kunath, T., Arnaud, D., et al. (2005). *Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts*. Development **132**(7): 1649-1661. 10.1242/dev.01715

Labosky, P. A., Barlow, D. P., et al. (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**(11): 3197-3204.

Lachner, M., O'Carroll, D., et al. (2001). *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins*. Nature **410**(6824): 116-120.

Lanctot, C., Cheutin, T., et al. (2007). *Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions*. Nat Rev Genet **8**(2): 104-115.

Lande-Diner, L., Zhang, J., et al. (2009). *Shifts in Replication Timing Actively Affect Histone Acetylation during Nucleosome Reassembly*. Molecular Cell **34**(6): 767-774. 10.1016/j.molcel.2009.05.027

Lander, E. S., Linton, L. M., et al. (2001). *Initial sequencing and analysis of the human genome*. Nature **409**(6822): 860-921.

Laster, K. and Kosak, S. T. (2010). *Genomic Pangea: coordinate gene regulation and cell-specific chromosomal topologies*. Current Opinion in Cell Biology **22**(3): 314-319. 10.1016/j.ceb.2010.04.009

Latham, J. A. and Dent, S. Y. R. (2007). *Cross-regulation of histone modifications*. Nat Struct Mol Biol **14**(11): 1017-1024.

Latt, S. A. (1973). *Microfluorometric Detection of Deoxyribonucleic Acid Replication in Human Metaphase Chromosomes*. Proceedings of the National Academy of Sciences **70**(12): 3395-3399.

Leahy, A., Xiong, J.-W., et al. (1999). *Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation*. Journal of Experimental Zoology **284**(1): 67-81. 10.1002/(sici)1097-010x(19990615)284:1<67::aid-jez10>3.0.co;2-o

Lee, J.-S., Smith, E., et al. (2010). *The Language of Histone Crosstalk*. Cell **142**(5): 682-685. 10.1016/j.cell.2010.08.011

Lee, J. S., Galvin, K. M., et al. (1995). *Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300*. Genes & Development **9**(10): 1188-1198. 10.1101/gad.9.10.1188

Lee, K. K. and Workman, J. L. (2007). *Histone acetyltransferase complexes: one size doesn't fit all*. Nat Rev Mol Cell Biol **8**(4): 284-295.

Lee, T. I., Jenner, R. G., et al. (2006). *Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells*. Cell **125**(2): 301-313. 10.1016/j.cell.2006.02.043

Leonhardt, H., Page, A. W., et al. (1992). *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei*. Cell **71**(5): 865-873. 10.1016/0092-8674(92)90561-p

Levenstein, M. E., Ludwig, T. E., et al. (2006). *Basic Fibroblast Growth Factor Support of Human Embryonic Stem Cell Self-Renewal*. Stem Cells **24**(3): 568-574. 10.1634/stemcells.2005-0247

Lewis, E. B. (1978). *A gene complex controlling segmentation in Drosophila*. Nature **276**(5688): 565-570.

Li, E., Bestor, T. H., et al. (1992). *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality*. Cell **69**(6): 915-926. 10.1016/0092-8674(92)90611-f

Li, G. and Reinberg, D. (2011). *Chromatin higher-order structures and gene regulation*. Current Opinion in Genetics & Development **21**(2): 175-186. 10.1016/j.gde.2011.01.022

Li, H., Ilin, S., et al. (2006). *Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF*. Nature **442**(7098): 91-95. http://www.nature.com/nature/journal/v442/n7098/suppinfo/nature04802_S1.html

Lim, J. W. E. and Bodnar, A. (2002). *Proteome analysis of conditioned medium from mouse embryonic fibroblast feeder layers which support the growth of human embryonic stem cells*. PROTEOMICS **2**(9): 1187-1203. 10.1002/1615-9861(200209)2:9<1187::aid-prot1187>3.0.co;2-t

Lin, C. M., Fu, H., et al. (2003). *Dynamic Alterations of Replication Timing in Mammalian Cells*. Current Biology **13**(12): 1019-1028. 10.1016/s0960-9822(03)00382-8

Lin, G., OuYang, Q., et al. (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**(12): 999-1007. http://www.nature.com/cr/journal/v17/n12/suppinfo/cr200797s1.html

Lister, R., Pelizzola, M., et al. (2011). *Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells*. Nature **471**(7336): 68-73.

Liu, X., Wang, L., et al. (2008). *The structural basis of protein acetylation by the p300/CBP transcriptional coactivator*. Nature **451**(7180): 846-850. http://www.nature.com/nature/journal/v451/n7180/suppinfo/nature06546_S1.html

Livak, K. J. and Schmittgen, T. D. (2001). *Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method*. Methods **25**(4): 402-408. 10.1006/meth.2001.1262

Loh, Y.-H., Agarwal, S., et al. (2009). *Generation of induced pluripotent stem cells from human blood*. Blood **113**(22): 5476-5479. 10.1182/blood-2009-02-204800

Loh, Y.-H., Wu, Q., et al. (2006). *The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells*. Nature Genetics **38**(4): 431-440. 10.1038/ng1760

Ludwig, T. E., Levenstein, M. E., et al. (2006). *Derivation of human embryonic stem cells in defined conditions*. Nat Biotech **24**(2): 185-187. http://www.nature.com/nbt/journal/v24/n2/suppinfo/nbt1177_S1.html

Luger, K., Mader, A. W., et al. (1997). *Crystal structure of the nucleosome core particle at 2.8 Å resolution*. Nature **389**(6648): 251-260.

Maherali, N., Sridharan, R., et al. (2007). *Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution*. Cell Stem Cell **1**(1): 55-70.

Mahy, N. L., Perry, P. E., et al. (2002). *Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH*. The Journal of Cell Biology **159**(5): 753-763. 10.1083/jcb.200207115

Mai, Q., Yu, Y., et al. (2007). *Derivation of human embryonic stem cell lines from parthenogenetic blastocysts*. Cell Res **17**(12): 1008-1019.

Maison, C., Bailly, D., et al. (2002). *Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component*. Nat Genet **30**(3): 329-334.

Malas, S., Duthie, S., et al. (1997). *Cloning and mapping of the human SOX1: a highly conserved gene expressed in the developing brain*. Mammalian Genome **8**(11): 866-868. 10.1007/s003359900597

Maldonado-Saldivia, J., van den Bergen, J., et al. (2007). *Dppa2 and Dppa4 Are Closely Linked SAP Motif Genes Restricted to Pluripotent Cells and the Germ Line*. Stem Cells **25**(1): 19-28. 10.1634/stemcells.2006-0269

Marcu, M. G., Jung, Y.-J., et al. (2006). *Curcumin is an Inhibitor of p300 Histone Acetylatransferase*. Medicinal Chemistry **2**(2): 169-174. 10.2174/157340606776056133

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 **5**(6): 589-591.

Margueron, R., Trojer, P., et al. (2005). *The key to development: interpreting the histone code?* Current Opinion in Genetics & Development **15**(2): 163-176. 10.1016/j.gde.2005.01.005

Marks, H., Kalkan, T., et al. (2012). *The Transcriptional and Epigenomic Foundations of Ground State Pluripotency*. Cell **149**(3): 590-604.

Martello, G., Sugimoto, T., et al. (2012). *Esrrb Is a Pivotal Target of the Gsk3/Tcf3 Axis Regulating Embryonic Stem Cell Self-Renewal*. Cell Stem Cell **11**(4): 491-504.

Martin, C. and Zhang, Y. (2005). *The diverse functions of histone lysine methylation*. Nat Rev Mol Cell Biol **6**(11): 838-849.

Martin, G. R. (1981). *Isolation of a Pluripotent Cell Line from Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells*. Proceedings of the National Academy of Sciences of the United States of America **78**(12): 7634-7638.

Masai, H., Matsumoto, S., et al. (2010). *Eukaryotic Chromosome DNA Replication: Where, When, and How?* Annual Review of Biochemistry **79**(1): 89-130. doi:10.1146/annurev.biochem.052308.103205

Masui, S., Nakatake, Y., et al. (2007). *Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells*. Nat Cell Biol **9**(6): 625-635.

Masui, S., Ohtsuka, S., et al. (2008). *Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells*. BMC Developmental Biology **8**(1): 45.

Masumoto, H., Hawke, D., et al. (2005). *A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response*. Nature **436**(7048): 294-298. http://www.nature.com/nature/journal/v436/n7048/suppinfo/nature03714_S1.html

Mathur, D., Danford, T., et al. (2008). *Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET*. Genome Biology **9**(8): R126.

Matsuda, T., Nakamura, T., et al. (1999). *STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells*. The EMBO journal **18**: 4261–4269. 10.1093/emboj/18.15.4261

Matsui, Y., Zsebo, K., et al. (1992). *Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture*. Cell **70**(5): 841-847. 10.1016/0092-8674(92)90317-6

Mayer, W., Niveleau, A., et al. (2000). *Embryogenesis: Demethylation of the zygotic paternal genome*. Nature **403**(6769): 501-502.

McGhee, J. D. and Felsenfeld, G. (1980). *Nucleosome Structure*. Annual Review of Biochemistry **49**(1): 1115-1156. doi:10.1146/annurev.bi.49.070180.005343

McNairn, A. J. and Gilbert, D. M. (2003). *Epigenomic replication: Linking epigenetics to DNA replication*. BioEssays **25**(7): 647-656. 10.1002/bies.10305

McPherson, J. D., Marra, M., et al. (2001). *A physical map of the human genome*. Nature **409**(6822): 934-941.

Mechali, M. (2001). *DNA replication origins: from sequence specificity to epigenetics*. Nat Rev Genet **2**(8): 640-645.

Méndez, J. (2009). *Temporal regulation of DNA replication in mammalian cells*. Critical Reviews in Biochemistry and Molecular Biology **44**(5): 343-351. 10.1080/10409230903232618

Meshorer, E. and Misteli, T. (2006). *Chromatin in pluripotent embryonic stem cells and differentiation*. Nat Rev Mol Cell Biol **7**(7): 540-546.

Meuwissen, R. L. J., Offenberg, H. H., et al. (1992). *A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. Get the document, find related information or use other SFX services*. EMBO journal **11**: 5091.

Mikkelsen, T. S., Ku, M., et al. (2007). *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells*. Nature **448**(7153): 553-560. http://www.nature.com/nature/journal/v448/n7153/suppinfo/nature06008_S1.html

Mileham, P. and Brown, S. (1996). *YAC clone contigs covering 5 Mb of a repeat sequence island on the mouse X Chromosome*. Mammalian Genome **7**(4): 253-261. 10.1007/s003359900077

Miller, R. A. and Ruddle, F. H. (1976). *Pluripotent teratocarcinoma-thymus somatic cell hybrids*. Cell **9**(1): 45-55. 10.1016/0092-8674(76)90051-9

Mitsui, K., Tokuzawa, Y., et al. (2003). *The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells*. Cell **113**(5): 631-642.

Morgan, H. D., Dean, W., et al. (2004). *Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues*. Journal of Biological Chemistry **279**(50): 52353-52360. 10.1074/jbc.M407695200

Morgan, H. D., Santos, F., et al. (2005). *Epigenetic reprogramming in mammals*. Hum. Mol. Genet. **14**(suppl_1): R47-58. 10.1093/hmg/ddi114

Morris, S. A., Teo, R. T. Y., et al. (2010). *Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo*. Proceedings of the National Academy of Sciences **107**(14): 6364-6369. 10.1073/pnas.0915063107

Morrisey, E. E., Tang, Z., et al. (1998). *GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo*. Genes & Development **12**(22): 3579-3590. 10.1101/gad.12.22.3579

Muller, H. J. and Altenburg, E. (1930). *THE FREQUENCY OF TRANSLOCATIONS PRODUCED BY X-RAYS IN DROSOPHILA*. Genetics **15**(4): 283-311.

Najm, Fadi J., Chenoweth, Josh G., et al. (2011). *Isolation of Epiblast Stem Cells from Preimplantation Mouse Embryos*. Cell Stem Cell **8**(3): 318-325.

Nakashima, K., Hagiwara, T., et al. (2002). *Nuclear Localization of Peptidylarginine Deiminase V and Histone Deimination in Granulocytes*. Journal of Biological Chemistry **277**(51): 49562-49568. 10.1074/jbc.M208795200

Nan, X., Ng, H.-H., et al. (1998). *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex*. Nature **393**(6683): 386-389.

Narva, E., Autio, R., et al. (2010). *High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity*. Nat Biotech **28**(4): 371-377. http://www.nature.com/nbt/journal/v28/n4/abs/nbt.1615.html - supplementary-information

Nazor, Kristopher L., Altun, G., et al. (2012). *Recurrent Variations in DNA Methylation in Human Pluripotent Stem Cells and Their Differentiated Derivatives*. Cell Stem Cell **10**(5): 620-634.

Neph, S., Vierstra, J., et al. (2012). *An expansive human regulatory lexicon encoded in transcription factor footprints*. Nature **489**(7414): 83-90. http://www.nature.com/nature/journal/v489/n7414/abs/nature11212.html - supplementary-information

Neumann, H., Hancock, S. M., et al. (2009). *A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation*. Molecular Cell **36**(1): 153-163. 10.1016/j.molcel.2009.07.027

Newman, A. M. and Cooper, J. B. (2010). *Lab-Specific Gene Expression Signatures in Pluripotent Stem Cells*. Cell Stem Cell **7**(2): 258-262. 10.1016/j.stem.2010.06.016

Ng, S., Yue, W., et al. (2009). *Dynamic protein methylation in chromatin biology*. Cellular and Molecular Life Sciences **66**(3): 407-422. 10.1007/s00018-008-8303-z

Niakan, K. K., Han, J., et al. (2012). *Human pre-implantation embryo development*. Development **139**(5): 829-841. 10.1242/dev.060426

Nichols, J., Chambers, I., et al. (2001). *Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines*. Development **128**(12): 2333-2339.

Nichols, J., Silva, J., et al. (2009a). *Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo*. Development **136**(19): 3215-3222. 10.1242/dev.038893

Nichols, J. and Smith, A. (2009b). *Naive and Primed Pluripotent States*. Cell Stem Cell **4**(6): 487-492. 10.1016/j.stem.2009.05.015

Nichols, J. and Smith, A. (2011). *The origin and identity of embryonic stem cells*. Development **138**(1): 3-8. 10.1242/dev.050831

Nichols, J., Zevnik, B., et al. (1998). *Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4*. Cell **95**(3): 379-391.

Niwa, H., Burdon, T., et al. (1998). *Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3*. Genes & Development **12**(13): 2048-2060. 10.1101/gad.12.13.2048

Niwa, H., Miyazaki, J.-i., et al. (2000). *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells*. Nature genetics **24**(4): 372-376. 10.1038/74199

Noggle, S., Fung, H.-L., et al. (2011). *Human oocytes reprogram somatic cells to a pluripotent state*. Nature **478**(7367): 70-75. http://www.nature.com/nature/journal/v478/n7367/abs/nature10397.html - supplementary-information

O'Keefe, R., Henderson, S., et al. (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**(5): 1095-1110. 10.1083/jcb.116.5.1095

O'Leary, T., Heindryckx, B., et al. (2012). *Tracking the progression of the human inner cell mass during embryonic stem cell derivation*. Nat Biotech **30**(3): 278-282. http://www.nature.com/nbt/journal/v30/n3/abs/nbt.2135.html - supplementary-information

Ohi, Y., Qin, H., et al. (2011). *Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells*. Nat Cell Biol **13**(5): 541-549. http://www.nature.com/ncb/journal/v13/n5/abs/ncb2239.html - supplementary-information

Okano, M., Bell, D. W., et al. (1999). *DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development*. Cell **99**(3): 247-257. 10.1016/s0092-8674(00)81656-6

Okano, M., Xie, S., et al. (1998). *Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells*. Nucleic Acids Research **26**(11): 2536-2540. 10.1093/nar/26.11.2536

Okita, K., Ichisaka, T., et al. (2007). *Generation of germline-competent induced pluripotent stem cells*. Nature **448**(7151): 313-317.

Osafune, K., Caron, L., et al. (2008). *Marked differences in differentiation propensity among human embryonic stem cell lines*. Nat Biotech **26**(3): 313-315. http://www.nature.com/nbt/journal/v26/n3/suppinfo/nbt1383_S1.html

Pal, S., Yun, R., et al. (2003). *mSin3A/Histone Deacetylase 2- and PRMT5-Containing Brg1 Complex Is Involved in Transcriptional Repression of the Myc Target Gene cad*. Molecular and Cellular Biology **23**(21): 7475-7487. 10.1128/mcb.23.21.7475-7487.2003

Palmieri, S. L., Peter, W., et al. (1994). *Oct-4 Transcription Factor Is Differentially Expressed in the Mouse Embryo during Establishment of the First Two Extraembryonic Cell Lineages Involved in Implantation*. Developmental Biology **166**(1): 259-267.

Parada, L., McQueen, P., et al. (2004). *Tissue-specific spatial organization of genomes*. Genome Biology **5**(7): R44.

Park, I.-H., Arora, N., et al. (2008a). *Disease-Specific Induced Pluripotent Stem Cells*. Cell **134**(5): 877-886. 10.1016/j.cell.2008.07.041

Park, I.-H., Lerou, P. H., et al. (2008b). *Generation of human-induced pluripotent stem cells*. Nature Protocols **3**(7): 1180-1186. 10.1038/nprot.2008.92

Park, Y.-J., Dyer, P. N., et al. (2004). *A New Fluorescence Resonance Energy Transfer Approach Demonstrates That the Histone Variant H2AZ Stabilizes the Histone Octamer within the Nucleosome*. Journal of Biological Chemistry **279**(23): 24274-24282. 10.1074/jbc.M313152200

Pastor, W. A., Pape, U. J., et al. (2011). *Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells*. Nature **473**(7347): 394-397. http://www.nature.com/nature/journal/v473/n7347/abs/10.1038-nature10102-unlocked.html - supplementary-information

Pelton, T. A., Sharma, S., et al. (2002). *Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development*. Journal of Cell Science **115**(2): 329-339.

Pera, M. F., Andrade, J., et al. (2004a). *Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin*. Journal of Cell Science **117**(7): 1269-1280. 10.1242/jcs.00970

Pera, M. F. and Tam, P. P. L. (2010). *Extrinsic regulation of pluripotent stem cells*. Nature **465**(7299): 713-720. 10.1038/nature09228

Pera, M. F. and Trounson, A. O. (2004b). *Human embryonic stem cells: prospects for development*. Development **131**(22): 5515-5525. 10.1242/dev.01451

Percharde, M., Lavial, F., et al. (2012). *Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming*. Genes & Development. 10.1101/gad.195545.112

Peric-Hupkes, D., Meuleman, W., et al. (2010). *Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation*. Molecular Cell **38**(4): 603-613. 10.1016/j.molcel.2010.03.016

Perry, P., Sauer, S., et al. (2004). *A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction*. Cell cycle **3**(12): 1645-1650.

Peters, A., Kubicek, S., et al. (2003). *Partitioning and plasticity of repressive histone methylation states in mammalian chromatin*. Mol Cell **12**: 1577 - 1589.

Peters, A. H. F. M. and Schübeler, D. (2005). *Methylation of histones: playing memory with DNA*. Current Opinion in Cell Biology **17**(2): 230-238. 10.1016/j.ceb.2005.02.006

Pfaffl, M. W. (2001). *A new mathematical model for relative quantification in real-time RT–PCR*. Nucleic Acids Research **29**(9): e45. 10.1093/nar/29.9.e45

Piccolo, F. M., Pereira, C. F., et al. (2011). *Using heterokaryons to understand pluripotency and reprogramming*. Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1575): 2260-2265. 10.1098/rstb.2011.0004

Pogo, B. G., Allfrey, V. G., et al. (1966). *RNA synthesis and histone acetylation during the course of gene activation in lymphocytes*. Proceedings of the National Academy of Sciences **55**(4): 805-812.

Polo, J. M., Liu, S., et al. (2010). *Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells*. Nat Biotech **28**(8): 848-855. http://www.nature.com/nbt/journal/v28/n8/abs/nbt.1667.html - supplementary-information

Pouponnot, C., Jayaraman, L., et al. (1998). *Physical and Functional Interaction of SMADs and p300/CBP*. Journal of Biological Chemistry **273**(36): 22865-22868. 10.1074/jbc.273.36.22865

Prieur, A., Besnard, E., et al. (2011). *p53 and p16INK4A independent induction of senescence by chromatin-dependent alteration of S-phase progression*. Nat Commun **2**: 473. http://www.nature.com/ncomms/journal/v2/n9/suppinfo/ncomms1473_S1.html

Probst, A. V., Dunleavy, E., et al. (2009). *Epigenetic inheritance during the cell cycle*. Nature Reviews Molecular Cell Biology **10**(3): 192-206. 10.1038/nrm2640

Prowse, A. B. J., McQuade, L. R., et al. (2007). *Identification of Potential Pluripotency Determinants for Human Embryonic Stem Cells Following Proteomic Analysis of Human and Mouse Fibroblast Conditioned Media*. Journal of Proteome Research **6**(9): 3796-3807. 10.1021/pr0702262

Ralston, A. and Rossant, J. (2005). *Genetic regulation of stem cell origins in the mouse embryo*. Clinical Genetics **68**(2): 106-112. 10.1111/j.1399-0004.2005.00478.x

Rando, O. J. (2012). *Combinatorial complexity in chromatin structure and function: revisiting the histone code*. Current Opinion in Genetics & Development **22**(2): 148-155. 10.1016/j.gde.2012.02.013

Raz, R., Lee, C.-K., et al. (1999). *Essential role of STAT3 for embryonic stem cell pluripotency*. Proceedings of the National Academy of Sciences **96**(6): 2846-2851. 10.1073/pnas.96.6.2846

Rea, S., Eisenhaber, F., et al. (2000). *Regulation of chromatin structure by site-specific histone H3 methyltransferases*. Nature **406**(6796): 593-599.

Reijo Pera, R. A., DeJonge, C., et al. (2009). *Gene expression profiles of human inner cell mass cells and embryonic stem cells*. Differentiation **78**(1): 18-23.

Reik, W. (2007). *Stability and flexibility of epigenetic gene regulation in mammalian development*. Nature **447**(7143): 425-432.

Resnick, J. L., Bixler, L. S., et al. (1992). *Long-term proliferation of mouse primordial germ cells in culture*. Nature **359**(6395): 550-551.

Reubinoff, B. E., Pera, M. F., et al. (2000). *Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro*. Nature biotechnology **18**(4): 399-404.

Revazova, E., Turovets, N., et al. (2007). *Patient-specific stem cell lines derived from human parthenogenetic blastocysts*. Cloning and Stem Cells **9**(3): 432-449.

Richmond, T. J., Finch, J. T., et al. (1984). *Structure of the nucleosome core particle at 7 Å resolution*. Nature **311**(5986): 532-537.

Ringrose, L. and Paro, R. (2004). *EPIGENETIC REGULATION OF CELLULAR MEMORY BY THE POLYCOMB AND TRITHORAX GROUP PROTEINS*. Annual Review of Genetics **38**(1): 413-443. doi:10.1146/annurev.genet.38.072902.091907

Robinson, P. J. J., An, W., et al. (2008). *30 nm Chromatin Fibre Decompaction Requires both H4-K16 Acetylation and Linker Histone Eviction*. Journal of Molecular Biology **381**(4): 816-825. 10.1016/j.jmb.2008.04.050

Robinson, P. J. J., Fairall, L., et al. (2006). *EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure*. Proceedings of the National Academy of Sciences **103**(17): 6506-6511. 10.1073/pnas.0601212103

Robinton, D. A. and Daley, G. Q. (2012). *The promise of induced pluripotent stem cells in research and therapy*. Nature **481**(7381): 295-305. http://www.nature.com/nature/journal/v481/n7381/abs/nature10761.html - supplementary-information

Rogers, M. B., Hosler, B. A., et al. (1991). *Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes*. Development **113**(3): 815-824.

Ross, S., Cheung, E., et al. (2006). *Smads orchestrate specific histone modifications and chromatin remodeling to activate transcription*. The EMBO journal **25**: 4490.

Ross, S. and Hill, C. S. (2008). *How the Smads regulate transcription*. The International Journal of Biochemistry & Cell Biology **40**(3): 383-408. DOI: 10.1016/j.biocel.2007.09.006

Rossant, J. (2008). *Stem Cells and Early Lineage Development*. Cell **132**(4): 527-531. 10.1016/j.cell.2008.01.039

Rossant, J. (2011). *The Impact of Developmental Biology on Pluripotent Stem Cell Research: Successes and Challenges*. Developmental Cell **21**(1): 20-23.

Rossant, J. and Tam, P. P. L. (2009). *Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse*. Development **136**(5): 701-713. 10.1242/dev.017178

Rougeulle, C., Chaumeil, J., et al. (2004). *Differential Histone H3 Lys-9 and Lys-27 Methylation Profiles on the X Chromosome*. Molecular and Cellular Biology **24**(12): 5475-5484. 10.1128/mcb.24.12.5475-5484.2004

Rudnicki, M. A., Schnegelsberg, P. N. J., et al. (1993). *MyoD or Myf-5 is required for the formation of skeletal muscle*. Cell **75**(7): 1351-1359. 10.1016/0092-8674(93)90621-v

Ruggiu, M., Speed, R., et al. (1997). *The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis*. Nature **389**(6646): 73-77.

Ruiz, S., Diep, D., et al. (2012). *Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells*. Proceedings of the National Academy of Sciences. 10.1073/pnas.1202352109

Russ, A. P., Wattler, S., et al. (2000). *Eomesodermin is required for mouse trophoblast development and mesoderm formation*. Nature **404**(6773): 95-99.

Ruthenburg, A. J., Allis, C. D., et al. (2007). *Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark*. Molecular Cell **25**(1): 15-30. 10.1016/j.molcel.2006.12.014

Ryba, T., Hiratani, I., et al. (2010). *Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types*. Genome Research **20**: 761-770. 10.1101/gr.099655.109

Sanchez-Elsner, T., Gou, D., et al. (2006). *Noncoding RNAs of Trithorax Response Elements Recruit Drosophila Ash1 to Ultrabithorax*. Science **311**(5764): 1118-1123. 10.1126/science.1117705

Santos, F., Hendrich, B., et al. (2002). *Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo*. Developmental Biology **241**(1): 172-182. 10.1006/dbio.2001.0501

Sanyal, A., Lajoie, B. R., et al. (2012). *The long-range interaction landscape of gene promoters*. Nature **489**(7414): 109-113. http://www.nature.com/nature/journal/v489/n7414/abs/nature11279.html - supplementary-information

Sato, N., Meijer, L., et al. (2004). *Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor*. Nature Medicine **10**(1): 55-63.

Schaarschmidt, D., Baltin, J., et al. (2004). *An episomal mammalian replicon: sequence-independent binding of the origin recognition complex*. EMBO J **23**(1): 191-201. http://www.nature.com/emboj/journal/v23/n1/suppinfo/7600029a_S1.html

Schmierer, B. and Hill, C. S. (2007). *TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility*. Nat Rev Mol Cell Biol **8**(12): 970-982.

Schones, D. E., Cui, K., et al. (2008). *Dynamic Regulation of Nucleosome Positioning in the Human Genome*. Cell **132**(5): 887-898. 10.1016/j.cell.2008.02.022

Schuettengruber, B., Chourrout, D., et al. (2007). *Genome Regulation by Polycomb and Trithorax Proteins*. Cell **128**(4): 735-745. 10.1016/j.cell.2007.02.009

Schurter, B. T., Koh, S. S., et al. (2001). *Methylation of Histone H3 by Coactivator-Associated Arginine Methyltransferase 1†*. Biochemistry **40**(19): 5747-5756. 10.1021/bi002631b

Schwartz, Y. B., Kahn, T. G., et al. (2006). *Genome-wide analysis of Polycomb targets in Drosophila melanogaster*. Nat Genet **38**(6): 700-705. http://www.nature.com/ng/journal/v38/n6/suppinfo/ng1817_S1.html

Schübeler, D., Scalzo, D., et al. (2002). *Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing*. Nature Genetics **32**(3): 438-442. 10.1038/ng1005

Scotland, K. B., Chen, S., et al. (2009). *Analysis of Rex1 (zfp42) function in embryonic stem cell differentiation*. Developmental Dynamics **238**(8): 1863-1877. 10.1002/dvdy.22037

Shahbazian, M. D. and Grunstein, M. (2007). *Functions of Site-Specific Histone Acetylation and Deacetylation*. Annual Review of Biochemistry **76**(1): 75-100. 10.1146/annurev.biochem.76.052705.162114

Sharma, S., Kelly, T. K., et al. (2010). *Epigenetics in cancer*. Carcinogenesis **31**(1): 27-36. 10.1093/carcin/bgp220

Shi, X., Hong, T., et al. (2006). *ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression*. Nature **442**(7098): 96-99. http://www.nature.com/nature/journal/v442/n7098/suppinfo/nature04835_S1.html

Shi, Y., Lan, F., et al. (2004). *Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1*. Cell **119**(7): 941-953. 10.1016/j.cell.2004.12.012

Shilatifard, A. (2006). *Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression*. Annual Review of Biochemistry **75**(1): 243-269. doi:10.1146/annurev.biochem.75.103004.142422

Shogren-Knaak, M., Ishii, H., et al. (2006). *Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions*. Science **311**(5762): 844-847. 10.1126/science.1124000

Shumaker, D. K., Dechat, T., et al. (2006). *Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging*. Proceedings of the National Academy of Sciences **103**(23): 8703-8708. 10.1073/pnas.0602569103

Simon, I., Tenzen, T., et al. (1999). *Asynchronous replication of imprinted genes is established in the gametes and maintained during development*. Nature **401**(6756): 929-932.

Sims, R. J. and Reinberg, D. (2006). *Histone H3 Lys 4 methylation: caught in a bind?* Genes & Development **20**(20): 2779-2786. 10.1101/gad.1468206

Singh, Amar M., Reynolds, D., et al. (2012). *Signaling Network Crosstalk in Human Pluripotent Cells: A Smad2/3-Regulated Switch that Controls the Balance between Self-Renewal and Differentiation*. Cell Stem Cell **10**(3): 312-326.

Skottman, H., Mikkola, M., et al. (2005). *Gene Expression Signatures of Seven Individual Human Embryonic Stem Cell Lines*. Stem Cells **23**(9): 1343-1356. 10.1634/stemcells.2004-0341

Smith, A. G., Heath, J. K., et al. (1988). *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides*. Nature **336**(6200): 688-690. 10.1038/336688a0

Smith, J. R., Vallier, L., et al. (2008). *Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm*. Developmental Biology **313**(1): 107-117.

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*. Human Molecular Genetics **8**(8): 1373-1386. 10.1093/hmg/8.8.1373

Sparmann, A. and van Lohuizen, M. (2006). *Polycomb silencers control cell fate, development and cancer*. Nature Reviews Cancer **6**(11): 846-856. 10.1038/nrc1991

Spivakov, M. and Fisher, A. G. (2007). *Epigenetic signatures of stem-cell identity*. Nat Rev Genet **8**(4): 263-271. 10.1038/nrg2046

Stevanovlć, M., Lovell-Badge, R., et al. (1993). *SOX3 is an X-linked gene related to SRY*. Human Molecular Genetics **2**(12): 2013-2018. 10.1093/hmg/2.12.2013

Stewart, M. H., Bosse, M., et al. (2006a). *Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment*. Nat Meth **3**(10): 807-815. http://www.nature.com/nmeth/journal/v3/n10/suppinfo/nmeth939_S1.html

Stewart, R., Stojkovic, M., et al. (2006b). *Mechanisms of self-renewal in human embryonic stem cells*. European Journal of Cancer **42**(9): 1257-1272.

Stinchcomb, D. T., Struhl, K., et al. (1979). *Isolation and characterisation of a yeast chromosomal replicator*. Nature **282**(5734): 39-43.

Stock, J. K., Giadrossi, S., et al. (2007). *Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells*. Nat Cell Biol **9**(12): 1428-1435. http://www.nature.com/ncb/journal/v9/n12/suppinfo/ncb1663_S1.html

Stoffers, D. A., Zinkin, N. T., et al. (1997). *Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence*. Nat Genet **15**(1): 106-110.

Stojkovic, M., Lako, M., et al. (2004). *Derivation of Human Embryonic Stem Cells from Day-8 Blastocysts Recovered after Three-Step In Vitro Culture*. Stem Cells **22**(5): 790-797. 10.1634/stemcells.22-5-790

Strahl, B. D., Briggs, S. D., et al. (2001). *Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1*. Current Biology **11**(12): 996-1000. 10.1016/s0960-9822(01)00294-9

Strelchenko, N., Verlinsky, O., et al. (2004). *Morula-derived human embryonic stem cells*. Reproductive BioMedicine Online **9**(6): 623-629. 10.1016/s1472-6483(10)61772-5

Stubblefield, E. (1975). *Analysis of the replication pattern of Chinese hamster chromosomes using 5-bromodeoxyuridine suppression of 33258 Hoechst fluorescence*. Chromosoma **53**(3): 209-221. 10.1007/bf00329172

Suetake, I., Shinozaki, F., et al. (2004). *DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction*. Journal of Biological Chemistry **279**(26): 27816-27823. 10.1074/jbc.M400181200

Tada, M., Takahama, Y., et al. (2001). *Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells*. Current biology : CB **11**(19): 1553-1558.

Tahiliani, M., Koh, K. P., et al. (2009). *Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1*. Science **324**(5929): 930-935. 10.1126/science.1170116

Takahashi, K., Tanabe, K., et al. (2007). *Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors*. Cell **131**(5): 861-872. 10.1016/j.cell.2007.11.019

Takahashi, K. and Yamanaka, S. (2006). *Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors*. Cell **126**(4): 663-676.

Tanaka, S., Kunath, T., et al. (1998). *Promotion of Trophoblast Stem Cell Proliferation by FGF4*. Science **282**(5396): 2072-2075. 10.1126/science.282.5396.2072

Taverna, S. D., Li, H., et al. (2007). *How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers*. Nat Struct Mol Biol **14**(11): 1025-1040. http://www.nature.com/nsmb/journal/v14/n11/suppinfo/nsmb1338_S1.html

Terranova, R., Sauer, S., et al. (2005). *The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity*. Exp Cell Res **310**: 344 - 356.

Tesar, P. J., Chenoweth, J. G., et al. (2007). *New cell lines from mouse epiblast share defining features with human embryonic stem cells*. Nature **448**(7150): 196-199. 10.1038/nature05972

Thomson, J. A., Itskovitz-Eldor, J., et al. (1998). *Embryonic Stem Cell Lines Derived from Human Blastocysts*. Science **282**(5391): 1145-1147. 10.1126/science.282.5391.1145

Thomson, J. A., Kalishman, J., et al. (1995). *Isolation of a Primate Embryonic Stem Cell Line*. Proceedings of the National Academy of Sciences **92**(17): 7844-7848. 10.1073/pnas.92.17.7844

Thomson, J. A., Kalishman, J., et al. (1996). *Pluripotent cell lines derived from common marmoset (Callithrix jacchus) blastocysts*. Biology of Reproduction **55**(2): 254-259. 10.1095/biolreprod55.2.254

Thurman, R. E., Rynes, E., et al. (2012). *The accessible chromatin landscape of the human genome*. Nature **489**(7414): 75-82. http://www.nature.com/nature/journal/v489/n7414/abs/nature11232.html - supplementary-information

Tompkins, J. D., Hall, C., et al. (2012). *Epigenetic stability, adaptability, and reversibility in human embryonic stem cells*. Proceedings of the National Academy of Sciences. 10.1073/pnas.1209620109

Torres-Padilla, M.-E., Parfitt, D.-E., et al. (2007). *Histone arginine methylation regulates pluripotency in the early mouse embryo*. Nature **445**(7124): 214-218. http://www.nature.com/nature/journal/v445/n7124/suppinfo/nature05458_S1.html

Toyooka, Y., Shimosato, D., et al. (2008). *Identification and characterization of subpopulations in undifferentiated ES cell culture*. Development **135**(5): 909-918. 10.1242/dev.017400

Tse, C., Sera, T., et al. (1998). *Disruption of Higher-Order Folding by Core Histone Acetylation Dramatically Enhances Transcription of Nucleosomal Arrays by RNA Polymerase III*. Molecular and Cellular Biology **18**(8): 4629-4638.

Tsukada, Y.-i., Fang, J., et al. (2006). *Histone demethylation by a family of JmjC domain-containing proteins*. Nature **439**(7078): 811-816. http://www.nature.com/nature/journal/v439/n7078/suppinfo/nature04433_S1.html

Tumbar, T. and Belmont, A. S. (2001). *Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator*. Nat Cell Biol **3**(2): 134-139.

Vallier, L., Alexander, M., et al. (2005a). *Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells*. Journal of Cell Science **118**(19): 4495-4509. 10.1242/jcs.02553

Vallier, L., Mendjan, S., et al. (2009a). *Activin/Nodal signalling maintains pluripotency by controlling Nanog expression*. Development **136**(8): 1339-1349. 10.1242/dev.033951

Vallier, L. and Pedersen, R. (2005b). *Human embryonic stem cells*. Stem Cell Reviews and Reports **1**(2): 119-130. 10.1385/scr:1:2:119

Vallier, L., Touboul, T., et al. (2009b). *Signaling Pathways Controlling Pluripotency and Early Cell Fate Decisions of Human Induced Pluripotent Stem Cells*. Stem Cells **27**(11): 2655-2666.

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**(8): 3917-3925. 10.1128/mcb.10.8.3917

van Oosten, A. L., Costa, Y., et al. (2012). *JAK/STAT3 signalling is sufficient and dominant over antagonistic cues for the establishment of naive pluripotency*. Nat Commun **3**: 817. http://www.nature.com/ncomms/journal/v3/n5/suppinfo/ncomms1822_S1.html

Vandesompele, J., De Preter, K., et al. (2002). *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes*. GenomeBiology.com **3**(7): RESEARCH0034.

Vassena, R., Boué, S., et al. (2011). *Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development*. Development **138**(17): 3699-3709. 10.1242/dev.064741

Villa-Diaz, L. G., Pacut, C., et al. (2009). *Analysis of the factors that limit the ability of feeder cells to maintain the undifferentiated state of human embryonic stem cells*. Stem Cells and Development **18**(4): 641.

Vogelauer, M., Rubbi, L., et al. (2002). *Histone acetylation regulates the time of replication origin firing*. Mol Cell **10**: 1223 - 1233.

Voigt, P., LeRoy, G., et al. (2012). *Asymmetrically Modified Nucleosomes*. Cell **151**(1): 181-193. 10.1016/j.cell.2012.09.002

Waddington, C. H. (1942). *The epigenotype*. Endeavour (London) **1**: 18.

Wakayama, S., Hikichi, T., et al. (2007). *Efficient Establishment of Mouse Embryonic Stem Cell Lines from Single Blastomeres and Polar Bodies*. Stem Cells **25**(4): 986-993. 10.1634/stemcells.2006-0615

Wang, H., Huang, Z.-Q., et al. (2001). *Methylation of Histone H4 at Arginine 3 Facilitating Transcriptional Activation by Nuclear Hormone Receptor*. Science **293**(5531): 853-857. 10.1126/science.1060781

Wang, H., Wang, L., et al. (2004). *Role of histone H2A ubiquitination in Polycomb silencing*. Nature **431**(7010): 873-878. http://www.nature.com/nature/journal/v431/n7010/suppinfo/nature02985_S1.html

Wang, J., Zhuang, J., et al. (2012). *Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors*. Genome Research **22**(9): 1798-1812. 10.1101/gr.139105.112

Wang, L., Li, L., et al. (2005). *Human embryonic stem cells maintained in the absence of mouse embryonic fibroblasts or conditioned media are capable of hematopoietic development*. Blood **105**(12): 4598-4603. 10.1182/blood-2004-10-4065

Wang, X., Lin, G., et al. (2009a). *Inhibition of Caspase-mediated Anoikis Is Critical for Basic Fibroblast Growth Factor-sustained Culture of Human Pluripotent Stem Cells*. Journal of Biological Chemistry **284**(49): 34054-34064. 10.1074/jbc.M109.052290

Wang, Z., Zang, C., et al. (2009b). *Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes*. Cell **138**(5): 1019-1031. 10.1016/j.cell.2009.06.049

Wang, Z., Zang, C., et al. (2008). *Combinatorial patterns of histone acetylations and methylations in the human genome*. Nat Genet **40**(7): 897-903. http://www.nature.com/ng/journal/v40/n7/suppinfo/ng.154_S1.html

Waterborg, J. H. (2002). *Dynamics of histone acetylation in vivo. A function for acetylation turnover?* Biochemistry and Cell Biology **80**(3): 363-378. 10.1139/o02-080

Watt, F. and Molloy, P. L. (1988). *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter*. Genes & Development **2**(9): 1136-1143. 10.1101/gad.2.9.1136

Weber, M., Davies, J. J., et al. (2005). *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells*. Nat Genet **37**(8): 853-862.

Wei, C. L., Miura, T., et al. (2005). *Transcriptome Profiling of Human and Murine ESCs Identifies Divergent Paths Required to Maintain the Stem Cell State*. Stem Cells **23**(2): 166-185. 10.1634/stemcells.2004-0162

Wen, B., Wu, H., et al. (2009). *Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells*. Nat Genet **41**(2): 246-250. http://www.nature.com/ng/journal/v41/n2/suppinfo/ng.297_S1.html

Wernig, M., Meissner, A., et al. (2007). *In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state*. Nature **448**(7151): 318-324. 10.1038/nature05944

Wiles, M. V. and Johansson, B. M. (1999). *Embryonic Stem Cell Development in a Chemically Defined Medium*. Experimental Cell Research **247**(1): 241-248. 10.1006/excr.1998.4353

Williams, R., Azuara, V., et al. (2006). *Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus*. J Cell Sci **119**: 132 - 140.

Williams, R. L., Hilton, D. J., et al. (1988). *Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells*. Nature **336**(6200): 684-687. 10.1038/336684a0

Wilmut, I., Schnieke, A. E., et al. (1997). *Viable offspring derived from fetal and adult mammalian cells*. Nature **385**(6619): 810-813.

Wintersberger, E. (2000). *Why is there late replication?* Chromosoma **109**(5): 300-307. 10.1007/s004120000080

Wolf, S. (2009). *The protein arginine methyltransferase family: an update about function, new perspectives and the physiological role in humans*. Cellular and Molecular Life Sciences **66**(13): 2109-2121. 10.1007/s00018-009-0010-x

Woodfine, K., Fiegler, H., et al. (2004). *Replication timing of the human genome*. Human Molecular Genetics **13**(2): 191-202. 10.1093/hmg/ddh016

Wray, J., Kalkan, T., et al. (2011). *Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation*. Nat Cell Biol **13**(7): 838-845. http://www.nature.com/ncb/journal/v13/n7/abs/ncb2267.html - supplementary-information

Wray, J., Kalkan, T., et al. (2010). *The ground state of pluripotency*. Biochem. Soc. Trans. **38**: 1027-1032.

Wu, H., D'Alessio, A. C., et al. (2011). *Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells*. Nature **473**(7347): 389-393. http://www.nature.com/nature/journal/v473/n7347/abs/10.1038-nature09934-unlocked.html - supplementary-information

Wu, Q., Bruce, A. W., et al. (2009). *CARM1 Is Required in Embryonic Stem Cells to Maintain Pluripotency and Resist Differentiation*. Stem Cells **27**(11): 2637-2645. 10.1002/stem.131

Wysocka, J., Swigut, T., et al. (2006). *A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling*. Nature **442**(7098): 86-90. http://www.nature.com/nature/journal/v442/n7098/suppinfo/nature04815_S1.html

Xiao, L., Yuan, X., et al. (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**(6): 1476-1486. 10.1634/stemcells.2005-0299

Xie, W., Song, C., et al. (2009). *Histone H3 Lysine 56 Acetylation Is Linked to the Core Transcriptional Network in Human Embryonic Stem Cells*. Mol Cell **33**(4): 417-427.

Xu, C., Inokuma, M. S., et al. (2001). *Feeder-free growth of undifferentiated human embryonic stem cells*. Nature biotechnology **19**(10): 971-974.

Xu, R.-H., Chen, X., et al. (2002). *BMP4 initiates human embryonic stem cell differentiation to trophoblast*. Nature biotechnology **20**(12): 1261-1264.

Xu, R.-H., Peck, R. M., et al. (2005). *Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells*. Nature Methods **2**(3): 185-190. 10.1038/nmeth744

Xu, R.-H., Sampsell-Barron, T. L., et al. (2008). *NANOG Is a Direct Target of TGF²/Activin-Mediated SMAD Signaling in Human ESCs*. Cell **3**(2): 196-206.

Yamanaka, S. (2009). *Elite and stochastic models for induced pluripotent stem cell generation*. Nature **460**(7251): 49-52. 10.1038/nature08180

Yamanaka, Y., Ralston, A., et al. (2006). *Cell and molecular regulation of the mouse blastocyst*. Developmental Dynamics **235**(9): 2301-2314. 10.1002/dvdy.20844

Yang, X.-J. and Seto, E. (2008). *The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men*. Nat Rev Mol Cell Biol **9**(3): 206-218. http://www.nature.com/nrm/journal/v9/n3/suppinfo/nrm2346_S1.html

Yao, T.-P., Oh, S. P., et al. (1998). *Gene Dosage-Dependent Embryonic Development and Proliferation Defects in Mice Lacking the Transcriptional Integrator p300*. Cell **93**(3): 361-372. Doi: 10.1016/s0092-8674(00)81165-4

Yeo, S., Jeong, S., et al. (2007). *Characterization of DNA methylation change in stem cell marker genes during differentiation of human embryonic stem cells*. Biochemical and Biophysical Research Communications **359**(3): 536-542. 10.1016/j.bbrc.2007.05.120

Ying, Q.-L., Nichols, J., et al. (2003). *BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3*. Cell **115**(3): 281-292.

Ying, Q.-L., Wray, J., et al. (2008). *The ground state of embryonic stem cell self-renewal*. Nature **453**(7194): 519-523. 10.1038/nature06968

Ying, Q. L., Nichols, J., et al. (2002). *Changing potency by spontaneous fusion*. Nature **416**(6880): 545-548.

Yoshida, K., Chambers, I., et al. (1994). *Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways*. Mechanisms of Development **45**(2): 163-171. 10.1016/0925-4773(94)90030-2

Young, Margaret A., Larson, David E., et al. (2012). *Background Mutations in Parental Cells Account for Most of the Genetic Heterogeneity of Induced Pluripotent Stem Cells*. Cell Stem Cell **10**(5): 570-582.

Young, Richard A. (2011). *Control of the Embryonic Stem Cell State*. Cell **144**(6): 940-954.

Yu, J. and Thomson, J. A. (2008). *Pluripotent stem cell lines*. Genes & Development **22**(15): 1987-1997. 10.1101/gad.1689808

Yu, J., Vodyanik, M. A., et al. (2006). *Human Embryonic Stem Cells Reprogram Myeloid Precursors Following Cell–Cell Fusion*. Stem Cells **24**(1): 168-176. 10.1634/stemcells.2005-0292

Yu, J., Vodyanik, M. A., et al. (2007). *Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells*. Science **318**(5858): 1917-1920. 10.1126/science.1151526

Yusa, K., Rashid, S. T., et al. (2011). *Targeted gene correction of [agr]1-antitrypsin deficiency in induced pluripotent stem cells*. Nature **478**(7369): 391-394. http://www.nature.com/nature/journal/v478/n7369/abs/nature10424.html - supplementary-information

Zaehres, H., Lensch, M. W., et al. (2005). *High-Efficiency RNA Interference in Human Embryonic Stem Cells*. Stem Cells **23**(3): 299-305. 10.1634/stemcells.2004-0252

Zhang, J., Xu, F., et al. (2002). *Establishment of transcriptional competence in early and late S phase*. Nature **420**(6912): 198 - 202. 10.1038/nature01150

Zhang, J.-Z., Gao, W., et al. (2006). *Screening for Genes Essential for Mouse Embryonic Stem Cell Self-Renewal Using a Subtractive RNA Interference Library*. Stem Cells **24**(12): 2661-2668. 10.1634/stemcells.2006-0017

Zhang, X., Neganova, I., et al. (2009). *A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A*. The Journal of Cell Biology **184**(1): 67-82. 10.1083/jcb.200801009

Zhang, X., Zhang, J., et al. (2008). *Esrrb Activates Oct4 Transcription and Sustains Self-renewal and Pluripotency in Embryonic Stem Cells*. Journal of Biological Chemistry **283**(51): 35825-35833. 10.1074/jbc.M803481200

Zhong, X. and Jin, Y. (2009). *Critical Roles of Coactivator p300 in Mouse Embryonic Stem Cell Differentiation and Nanog Expression*. Journal of Biological Chemistry **284**(14): 9168-9175. 10.1074/jbc.M805562200

Zhou, H., Wu, S., et al. (2009). *Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins*. Cell Stem Cell **4**(5): 381-384.

Appendix I Primer sequences used

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

Human primers		
Species/Gene		***Sequence 5'-3'***
β-TUBULIN	s	CAGATGTTCGATGCCAAGAA
	as	TGCTGTTCTTGCTCTGGATG
hAFP	s	AAATGCGTTTCTCGTTGCTT
	as	GCCACAGGCCAATAGTTTGT
hBMP4	s	TGGTCTTGAGTATCCTGAGCG
	as	GCTGAGGTTAAAGAGGAAACGA
hBRG1	s	TGGACTACAGCGACTCACTGA
	as	TCGATGGCCTTTGACGCAC
hCBP	s	GACAAGCGAAACCAACAAACC
	as	TGGGGTCTATGGGATTTGGGT
hCD19	s	GCTCAAGACGCTGGAAAGTATTATT
	as	GATAAGCCAAAGTCACAGCTGAGA
hCD45	s	CCCCATGAACGTTACCATTTG
	as	GATAGTCTCCATTGTGAAAATAGGCC
hCDX2	s	CAGGACGAAAGACAAATATCGAGTG
	as	CCAGATTTTAACCTGCCTCTCAGA
hC-MYC	s	TTGTACCTGCAGGATCTGAG
	as	CTTGTTCCTCCTCAGAGTCG
hCRIPTO	s	AGAAGTGTTCCCTGTGTAAATGCTG
	as	CACGAGGTGCTCATCCATCA
hDNMT3B	s	GTCAAGCTACACACAGGACTTGACAG
	as	AGTTCGGACAGCTGGGCTTT
hEBF	s	GGAGATCGAGAGGACAGCGT
	as	GTCAATGAGGCGCACGTAGA
hEOMES	s	CAAAGGCGCAAATAACAACAAC
	as	CCACGCCATCCTCTGTAACTTC
hESRRB	s	TGAGGCCAGAGGTGATCCAG
	as	TCTTGATGAAGGAGCCGCAG
hFGF5	s	CAGCACCAAAGGCTCAGCTT
	as	CCTTGCTTCTAACCCATCATATCC
hFOXA2	s	GGGAGCGGTGAAGATGGA
	as	TCATGTTGCTCACGGAGGAGTA
hGAPDH	s	TCTGCTCCTCCTGTTCGACA
	as	AAAAGCAGCCCTGGTGACC
hGATA6	s	ACCACCTTATGGCGCAGAAAC
	as	TTTTTCATAGCAAGTGGTCTGGG
hGCN5	s	GCTACATCAAGGACTACGAG
	as	GCTTCTTGATGATCTCTTTCTG
hGDF3	s	GTACTTCGCTTTCTCCCAGAC

	as	GCCAATGTCAACTGTTCCCTT
hHPRT	s	TCCTTGGTCAGGCAGTATAATCC
	as	GTCAAGGGCATATCCTACAACAAA
hKLF4	s	CACATTAATGAGGCAGCCACC
	as	AAGTCGCTTCATGTGGGAGAG
hLIN28	s	TGCGGGCATCTGTAAGTGG
	as	GGAACCCTTCCATGTGCAG
hNANOG	s	CCAACATCCTGAACCTCAGCTAC
	as	GCCTTCTGCGTCACACCATT
hNODAL	s	AGAAGCAGATGTCCAGGGTAGCT
	as	GAGAGGTTGGAGTAGAGCATAAGGA
hOCT4	s	TCGAGAACCGAGTGAGAGGC
	as	CACACTCGGACCACATCCTTC
hP300	s	TCAGCCAAGCGGCCTAAAC
	as	TCACCACCATTGGTTAGTCCC
hP63	s	TTTCAGAGGCAATCCACACA
	as	ATGCATGCAAATGAGCTCTG
hOLIG3	s	TTTCTTCTCACGTCTTGCCGCG
	as	TGGGGAGGCTTTAGGCGGGAAATT
hPAX6	s	TCCGTTGGAACTGATGGAGT
	as	GTTGGTATCCGGGGACTTC
hPCAF	s	CGAATCGCCGTGAAGAAAGC
	as	GAGGGGTTAGGGTTTTTCCAG
hPDX1	s	CCTTTCCCATGGATGAAGTC
	as	GGAACTCCTTCTCCAGCTCTA
hPECAM1	s	AACAGTGTTGACATGAAGAGCC
	as	TGTAAAACAGCACGTCATCCTT
hPPAR-γ	s	GCTGGCCTCCTTGATGAATA
	as	TTGGGCTCCATAAAGTCACC
hREX1/ZFP42	s	GCGTACGCAAATTAAAGTCCAGA
	as	CAGCATCCTAAACAGCTCGCAGAAT
hSMAD2	s	GCCATCACCACTCAAAACTGT
	as	GCCTGTTGTATCCCACTGATCTA
hSIRT1	s	TAGGCGGCTTGATGGTAATC
	as	TCTGGCATGTCCCACTATCA
hSOX1	s	AACACTTGAAGCCCAGATGGA
	as	GCAGGCTGAATTCGGTTCTC
hSOX17	s	GGCGCAGCAGAATCCAGA
	as	CCACGACTTGCCCAGCAT
hSOX2	s	CACACTGCCCCTCTCACACAT
	as	CATTTCCCTCGTTTTTCTTTGAA
hSOX3	s	TTTGATAAGCCTACCCTTCC
	as	GTTCTTGAGTTCAGTCTCCA
hTERT	s	GCCAGCATCATCAAACCCC
	as	CTGTCAAGGTAGAGACGTGGCTC
hTLE1	s	TGTCTCCCAGCTCGACTGTCT

	as	AAGTACTGGCTTCCCCTCCC
hZIC1	s	CTGGCTGTGGCAAGGTCTTC
	as	CAGCCCTCAAACTCGCACTT

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

Human primers				
Species/Gene	***Tm***	***n***		***Sequence 5'-3'***
hβ-ACTIN	60ºC	25	s	TCACCACCACGGCCGAGCG
			as	TCTCCTTCTGCATCCTGTCG
hGATA6	60ºC	25	s	CCATGACTCCAACTTCCACC
			as	ACGGAGGACGTGACTTCGGC
hNANOG	60ºC	25	s	AGCCTCTACTCTTCCTACCACC
			as	TCCAAAGCAGCCTCCAAGTC
hNESTIN	65ºC	25	s	CAGCTGGCGCACCTCAAGATG
			as	AGGGAAGTTGGGCTCAGGACTG
hOCT4	60ºC	25	s	CTTGCTGCAGAAGTGGGTGGAGGAA
			as	CTGCAGTGTGGGTTTCGGGCA
hPAX6	60ºC	25	s	AACAGACACAGCCCTCACAAACA
			as	CGGGAACTTGAACTGGAACTGAC
hSOX1	60ºC	25	s	CAATGCGGGGAGGAGAAGTC
			as	CTCTGGACCAAACTGTGGCG
hSOX2	60 ºC	25	s	CCCCCGGCGGCAATAGCA
			as	TCGGCGCCGGGGAGATACAT
hT	63ºC	30	s	GTGACCAAGAACGGCAGGAGG
			as	TGTTCCGATGAGCATAGGGGC

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		
Species/Gene		***Sequence 5'-3'***
h*α-GLOBIN* promoter	s	TTTGTCCCCGGACCTGCTGCC
	as	CCCCACGGTCGGCTGACAC
h*AMYLASE* promoter	s	GGTGAGTCTGTGTGGTCAGCAGTCTCT
	as	CCACGGTGCTCTGGTAGATAACGTAAG
h*BMP4* promoter	s	CAGCTCCTTTCTCCATAACCTGA
	as	GGTACTAGAAAGCATGCACCGAC
h*CDX2* promoter	s	AACAACCACTGCTCCTGTCTCC
	as	CGACCTGACACAGCTAAATATTCAA
h*CGA* promoter	s	ATGGCTCCAAACAAAAATGACC
	as	TCCTCACCTGCTTTTATACCAGC
h*CRIPTO* promoter	s	ACTTCAAGTCTGGAGCCCCC
	as	GCCTGAGGAGCTAGGTGTGTGT
h*CXC4R4* promoter	s	GAAGTCACTATGGGAAAAGATGGG

h*DAZL* promoter	as	GCTTGCTTTCTTCAGGAAATTCTG
	s	TCCTGAATGTCAAACTGGCATT
h*DNMT3B* promoter	as	CGCCATCACTTTTAGCTTCACA
	s	TGTGTGTCTCCGTTCGGGTT
h*DPPA4* promoter	as	ATCAGAAGCCCTAAGCGGGA
	s	AAATAGCTTTCTTCAAGCCGCC
h*EBF* promoter	as	CGGAGACATTGGGAGATTGAG
	s	ACTTGAAGTAGCAAAGCCGGCA
h*EOMES* promoter	as	TGTAATGATCACAGGCCGGTGG
	s	AGGAAAAGAAAGTCACAGGCGA
h*ESXL1* promoter	as	CATCGGTCAAGTTGACCACTTG
	s	AGCATGTTCACTACCTGCACCA
h*FGF5* promoter	as	TATAGCGGTACGGTGCCTTCC
	s	AAAAACAGCCTTCTCCCCGAAC
h*FOXA2* promoter	as	CACTGCTTCTACGTCTTTCGGC
	s	CCCTTACTCAAGCTTCAATGCG
h*GATA4* promoter	as	TTCTTGGTTCTGAAAACCTGGC
	s	CGGTAGCACTTGGGCATTTT
h*GATA6* promoter	as	AGCAGGCAAAGTCCAGGCT
	s	ACGACCTGAGCCGTAGCAT
h*GBX2* promoter	as	CCTAGCCCCTTCACAGCAAA
	s	GGCAGGCAAAATGTGAATGAG
h*GDF3* promoter	as	GAAGCCGGCGTACTTATCTCC
	s	AAAGCCATCGTGAGCCTCAT
h*HOXB1* promoter	as	GACAGAGTGAAAAGGAGACATGCA
	s	CTAGCCCACCACCCTAAGACAA
h*IGF2/hH19* ICR	as	GGCTCCTGGGTTATTAGATTTGG
	s	ACATTCACACGAGCATCCAGG
h*IKAROS* promoter	as	GCTCTTTAGGTTTGGCGCAAT
	s	GAACAATGCGAGTGAGCAACTTCAGG
h*MASH1* promoter	as	GGCCGAGCGTCCCGCCCAGGCTG
	s	CAATTTCTAGGGTCACCGAGGA
h*MIXL1* promoter	as	GAACGGTGTGGAGGTGAGGA
	s	GTTTGACCAACCTGGTACAGGG
hMYF5 promoter	as	GCAGGCGACAGAACTCTCTTCT
	s	TGTCTGTAGAGGAAAGGCGGA
h*NANOG* promoter	as	ATATTCTTTCGGGAGGGAGGC
	s	ACGGCCTCCCAATTTACTGG
h*NESTIN* promoter	as	GGTTCAACAGGAATGGGATAAAAC
	s	TTTCCAGGCAGCGTCTCTCTAG
h*NEUROD* promoter	as	GGATGTGGCCAGGTTTTTGTT
	s	CTTCACTGCGTGCCTCAGTCT
h*NEUROG1* promoter	as	GAGAGATTAACCCTTTCAGGCG
	s	TTGTTGCGCCGGGTACTTAAG
h*OCT4/POU5F1* promoter	as	GCCGGTCTCCTGAGTGATGT
	s	TTGGGAGTTGAAAGTTGGGTGT

h*PAX6* promoter	as	AGGCTGGTCTTGAATTCCTGTC
	s	AAAGGACAGCACAGAAACTTGC
h*PDX1* promoter	as	TCAAGTGAAGGTTTCTGGGTTG
	s	GAATGCCAGAGTTTCGTGTGTTT
h*PSC1/RBM27* promoter	as	CCCAGCTCTCTAGGTCAGAAGC
	s	GTGAACGTATTGCACGTTGTGTTGC
h*PRCE/ESPL1* promoter	as	TCAGCCTGGTTCCTTCCTCTCGTTT
	s	AGGAAGGAAGAGCGGAATCGT
h*REX1/ZFP42* promoter	as	GGAAGGAGGCAAAGAATCACAT
	s	CTCATGTGATCCCCCCGTCT
hREX1 promoter -1Mb	as	TACGCGTGGGTGTAATCACATT
	s	GCACACGTTTTTCCAGTGTGCAA
hREX1 promoter -0.5Mb	as	AGGGGAAAAACTGCCTTGTGTGGA
	s	AGGGAGTGTGGTTTTCAGGCGT
hREX1 promoter +0.5Mb	as	ACATGCAGGAGCCGTATGGGTA
	s	TCAAGTCGTGGGCTGCTACACA
hREX1 promoter +1Mb	as	TGTGCTGACAAGCACGCAGA
	s	AAAGCCTTCATGGCCCGACTGT
hSatellite repeats	as	TTGCTTCCCTCCTTTGTCGCGT
	s	ATCGAATGGAAATGAAAGGAGTCA
h*SOX1* promoter	as	GACCATTGGATGATTGCAGTCA
	s	TCTTTGGCAAGTGGTTTGTGC
h*SOX17* promoter	as	GGAGGCAACGACAACAAAAAAA
	s	TATCAACGGTGTCTTTCGCATT
h*SOX2* promoter	as	AAGCCGAGGGTGTTTCAGTG
	s	CGTGAGAGAGTGTTGGCACCT
h*SOX3* promoter	as	TTGTTCTCCCGCTCATCCAC
	s	AATGGCCACAGAGCAGAGACA
h*SOX7* promoter	as	TGTGGAAAGTATGTGTGGGACG
	s	TGGCAGTCGGAGTATTAAGAACG
h*SYCP1* promoter	as	TCAGAAAATCAACCCAACCAGG
	s	GAGCCTGACAGTTTCCCTAGCA
h*TEKT1* promoter	as	CTGGCAGTTCAACGGCTAAATA
	s	GAGTGACTTCTGGAGTGTTTGCC
h*TERT* promoter	as	GAGGGTTAAAAATCAGGAAGCCA
	s	TCACGTCCGGCATTCGTGGT
h*TLE1* promoter	as	TCGAATCGGCCTAGGCTGTG
	s	CTGTTTTCCCCACCATCTCGT
h*WNT3* promoter	as	TCCTTCACCTTCGTGTGCTTCT
	s	CGCGTGATTGACAGGCTG
h*ZIC1* promoter	as	CAGCCGGGTTTGAGGATGT
	s	CTGTGTTTACTAGCCCCATCCC
D. melanogaster EMC	as	ATTCCTCTTCCTCCTTTGCCC
	s	GGTTTCAAGTTGTGGTCCAAGTTCC
	as	GTTCGTTCGGCTGGAGAGAGAT

Mouse specific primers for replication timing, ChIP and MeDIP analysis

Species/Gene		***Sequence 5'-3'***
m*a-globin* promoter	s	CCACAAGCTGCGTGTGGAT
	as	ATGCCGCCTGCCAGGT
m*Amylase* promoter	s	CAAAGCAAAATGAAGTTCGTTCTGCTGC
	as	GAAGTTATCTTACCTGCACCCCTCCAAATC
m*Bmp4* promoter	s	TTGAAGAGGTATCTCCTTCGGG
	as	GATTGGTTTAAAGGGTGGCCT
m*Cdx2* promoter	s	ACCACCTTCTGCCTGAGAATGTAC
	as	CCTCCAATCACAGGTTCAAAGACT
m*Cga* promoter	s	GCCATCCAATCACTGGGTAGA
	as	TCCCACTGTCTTTTATACCAGC
m*Cripto* promoter	s	CTGTATAGGGTCAGCACTTCCAGCC
	as	CCCCGGAGGCACCCTTTACT
m*Cxcr4* promoter	s	CCTTCCTCCTTTGGCTAATCCT
	as	TCACTCTCAAAACTGTGTGCGG
m*Dazl* promoter	s	TGCCGCCCGTTATTAATTTC
	as	CAAAAATGGCACCTCGAAGG
m*Dnmt3b* promoter	s	CGAAAAGCCAATGCAGGGTTGCTGT
	as	GCACACACGCACATACAAGCACATTCAC
m*Dppa4* promoter	s	TCCCAGATCTACAGTTGCCAGG
	as	GGCCAGAAATGCGCTTTGT
m*Ebf* promoter	s	ATTGGCTGTGATAAGGCGGAAGGATGACG
	as	CCCCAGAGCTAAACACAGGCACACTAA
m*Eomes* promoter	s	CCTCCTCTCACCCCAACAGAGCGAA
	as	AACCGATGTCTAGCTTGTTGGTCACAGG
m*Esx1* promoter	s	GCGAACTCGGAGCTGCTAAA
	as	TCCACGTGCTGTTAACTGACTTTT
m*Fgf5* promoter	s	CATGTGTCCCAGCCAACTCACT
	as	TTCCCAGGCTCCAGATGGC
m*Foxa2* promoter	s	CAGCCTAGACTTCTCTGAGATCCTCC
	as	TCAATAATGGGCCTTGTGGGA
m*Gata4* promoter	s	GATCTGAGGCTAGCAAGGCAT
	as	TTTGCTGCTCTGCCAAGGA
m*Gata6* promoter	s	AGCTCTTATTGTCCGCTAGGGC
	as	TGAAGAGAAGGAGGAGGGATCC
m*Gbx2* promoter	s	CGAACAGCTGCTGGAGCTGGAGAAAG
	as	GCACCTCGCTGAGTTTGAGGGCG
m*Gdf3* promoter	s	GCTCCCAGAAGTTATCTTATTGCC
	as	TCAGCAGTGTACAGCGTGAGTG
m*Hoxb1* promoter	s	CTTCCTAGTCATCCTTTTGTCCCA
	as	CTTATCCCAGAACCCCCATTC
m*Ikaros* promoter	s	CCAGTTTCAGGGACTCGGCT
	as	TCGGGGAACACGGGACAC
m*Mash1* promoter	s	CCAGGCTGGAGCAAGGGA
	as	CGGTTGGCTTCGGGAGC

m*Mixl1* promoter	s	CACAGATACTCTAGGCAGGAGCAGC
	as	CCTGTCACCATCAAAGGCGC
hMyf5 promoter	s	GGAGATCCGTGCGTTAAGAATCC
	as	CGGTAGCAAGACATTAAAGTTCCGTA
m*Nanog* promoter	s	CCCTCTGAGTTTGACCGGTGA
	as	CAAGCTAGGATGTTAGGTCTCCCTG
m*Nestin* promoter	s	GGAGCACTGGCAGACTTCCCT
	as	GGTGTTGGAAACCTCGTGAACTG
m*NeuroD* promoter	s	TTCAGGCTAGGACATAGACTCAGTGA
	as	CCCTTTGTGGCAGCAGAAGA
m*Neurog1* promoter	s	TTGTTGCGCGCCGTACTTAAGG
	as	TGGTCTCCTGAGTGATGTCGCC
m*Oct4/Pou5f1* promoter	s	GGCTCTCCAGAGGATGGCTGAG
	as	TCGGATGCCCCATCGCA
m*Pax6* promoter	s	GGTCAGAGGTAATTATGTCACCGCG
	as	TTATTCCAGAGTTGAGCCGTGAGCA
m*Pdx1* promoter	s	GGACTACATCTTGAGTTGCAGGC
	as	CCCCCTCGCTATGTTTTTTGT
m*Psc1/Rbm27* promoter	s	ACATAAACGCACCGCACGTTGT
	as	CCTGGTTACTTCCGCTGGTTTTT
m*Prce/Espl1* promoter	s	GACAGCGCGGCAGATAAACTAC
	as	GCACCTCCTTTCAAATATCCCTAG
m*Rex1/Zfp42* promoter	s	TTTGCGGGAATCCAGCAGT
	as	CGTCCCATCGCCACTCTAGAC
m*Sox1* promoter	s	ACAAGAGGAGGCAGCGAACC
	as	TCGCAGGTGGAAAGTTTCTCC
Sox17 promoter	s	TGAGGCTCGCTGTAGAAGAGTG
	as	AGCTGTGCCCGTAACCATTT
m*Sox2* promoter	s	CCATCCACCCTTATGTATCCAAG
	as	CGAAGGAAGTGGGTAAACAGCAC
m*Sox3* promoter	s	AGGAGGAAGCGGGAATCCAG
	as	AAACCTCCGCTGCCCG
m*Sox7* promoter	s	TCCACAGTGCCAGTTTAGGGAA
	as	CAGTAAACGTGGGCAGCCAC
m*Sycp1* promoter	s	AAGCCATCTTTAGCACCCAGC
	as	TGGTTGTTAGTTGCCTGGTGG
m*Tekt1* promoter	s	GATAAGCAGATACACCCAGCTCATC
	as	CCTTCTGGTGTGCATGAAAACA
m*Tert* promoter	s	GTCGCACCACAATAAAGCCTTAAC
	as	GCTGTGGTTGACACAGACAGTGA
m*Tle1* promoter	s	ATCCCTCCCCCAACAAGAAA
	as	TTGAAGTTTGCAGAAGCCTGTG
m*Wnt3* promoter	s	CATTTTCCTAGCCCCTGATCC
	as	TTTCTGTGATCCCTTGGCACTT
m*Zic1* promoter	s	AGGGAAAAGAAATCGTGGCATT
	as	CGCTTGACATCCCCTTTTGA

Appendix II Additional data

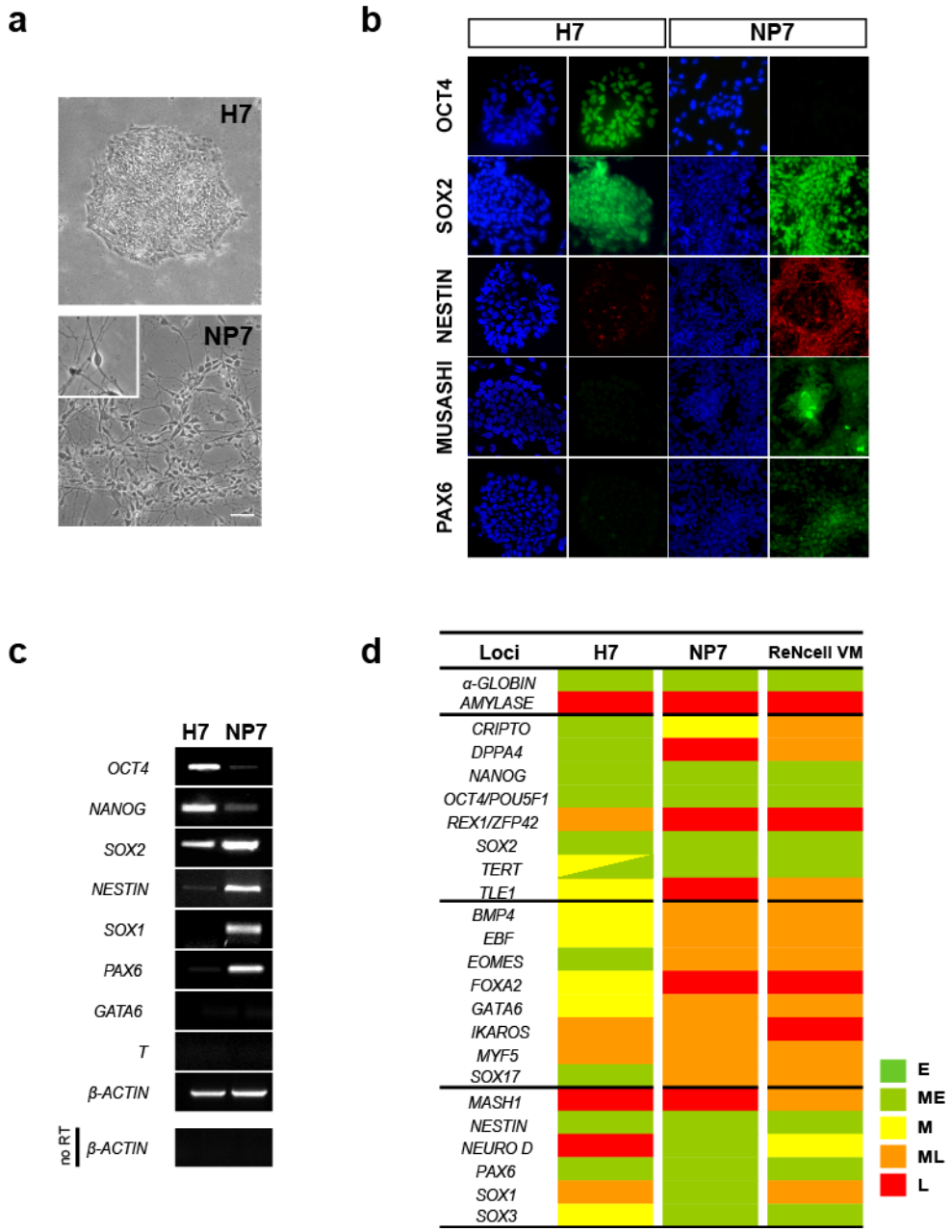


Figure A1 H7 undifferentiated hES cells can be clearly distinguished by replication timing analysis from neural progenitors (NP) H7-derived.

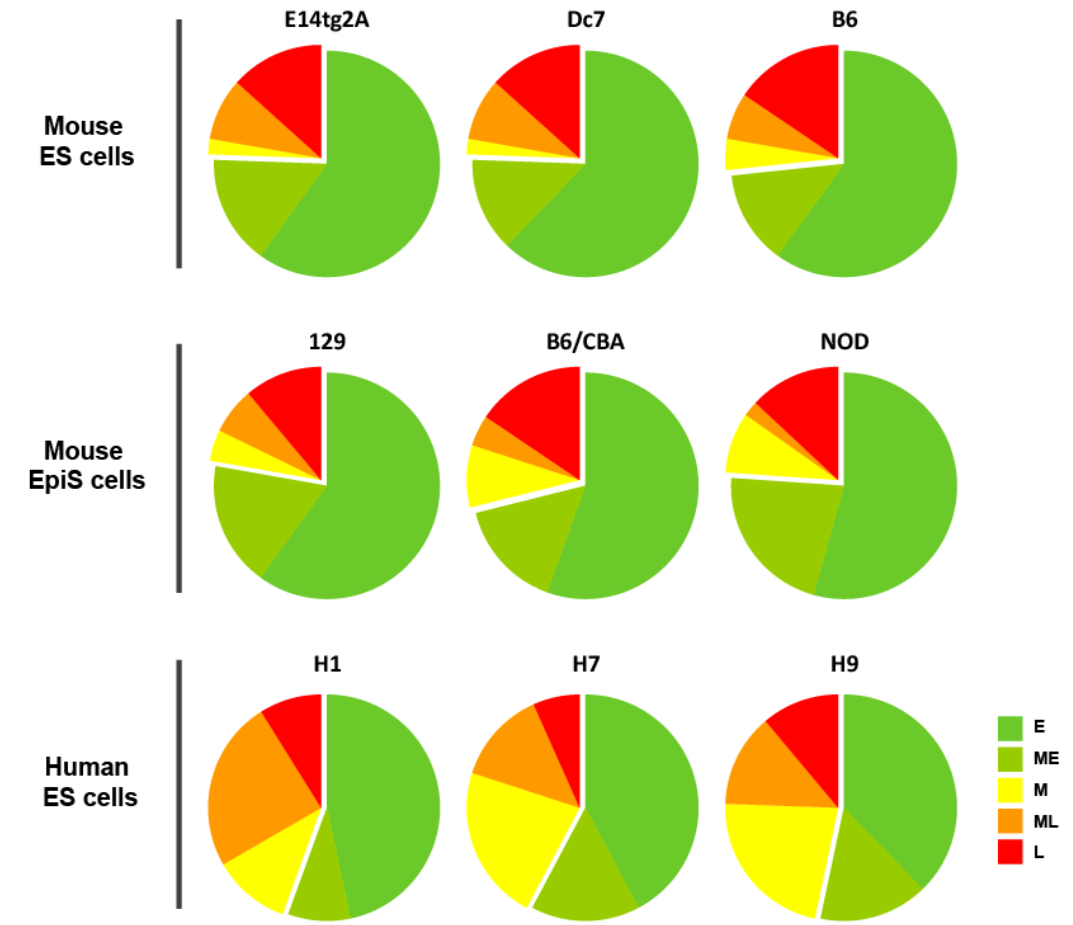
(a) Phase-contrast images of undifferentiated H7 hES cells (top panel) grown on matrigel-coated plates in conditioned media and NP7 cells with typical bipolar morphology (bottom panel) after >60 days of differentiation, induced as previously described (Gerrard et al. 2005b). Scale bar, 75 µm. (**b**) 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**) Semi quantitative 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 foetal 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.

Table A4 - Genomic features of replication timing analysed genes

	Loci	GC content (%)[a]	LINE density (%)[b]	Gene density [c]
hES cell variable loci	*BMP4*	39,7	23,2	2,0
	DPPA4	41,6	17,9	5,0
	EBF	39,2	14,2	5,0
	FOXA2	42,9	15,3	4,0
	GATA6	43,9	9,4	2,0
	SOX17	40,8	24,7	4,0
	SOX3	40,3	25,2	3,0
	SYCP1	36,6	31,9	10,0
	TERT	55,1	10,2	15,0
	TLE1	43,3	8,9	3,0
	ZIC1	38,8	10,0	1,0
hES cell steady loci	*AMYLASE*	35,4	19,5	4,0
	CDX2	44,1	15,1	9,0
	CGA	38,7	17,8	9,0
	CRIPTO	44,6	17,9	19,0
	CXCR4	42,6	13,9	5,0
	DAZL	39,8	22,2	5,0
	DNMT3B	47,5	6,0	17,0
	EOMES	40,5	19,1	2,0
	ESXL1	41,1	15,1	7,0
	FGF5	38,5	22,9	3,0
	GATA4	47,2	15,1	22,0
	GBX2	47,8	12,8	7,0
	GDF3	43,6	14,3	18,0
	HOXB1	48,5	4,1	23,0
	IKAROS	40,9	17,9	6,0
	MASH1	39,9	26,5	3,0
	MIXL1	42,5	15,5	14,0
	MYF5	35,5	15,9	6,0
	NANOG	44,9	7,0	19,0
	NESTIN	51,1	7,6	32,0
	NEUROD	38,3	24,2	6,0
	NEUROG1	47,0	15,3	9,0
	OCT4/POU5F1	47,1	13,2	40,0
	PAX6	41,6	14,8	6,0
	PDX1	44,3	18,8	11,0
	PRCE/ESPL1	50,8	10,5	31,0
	PSC1/RBM27	39,0	4,6	9,0
	REX1/ZFP42	41,4	20,8	2,0
	SOX1	47,4	13,1	1,0
	SOX2	39,5	9,5	0,0
	SOX7	46,6	15,7	7,0
	TEKT1	42,1	39,5	34,0
	WNT3	49,0	7,3	12,0
	α-GLOBIN	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 surronding 500 kb.

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 A2 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.

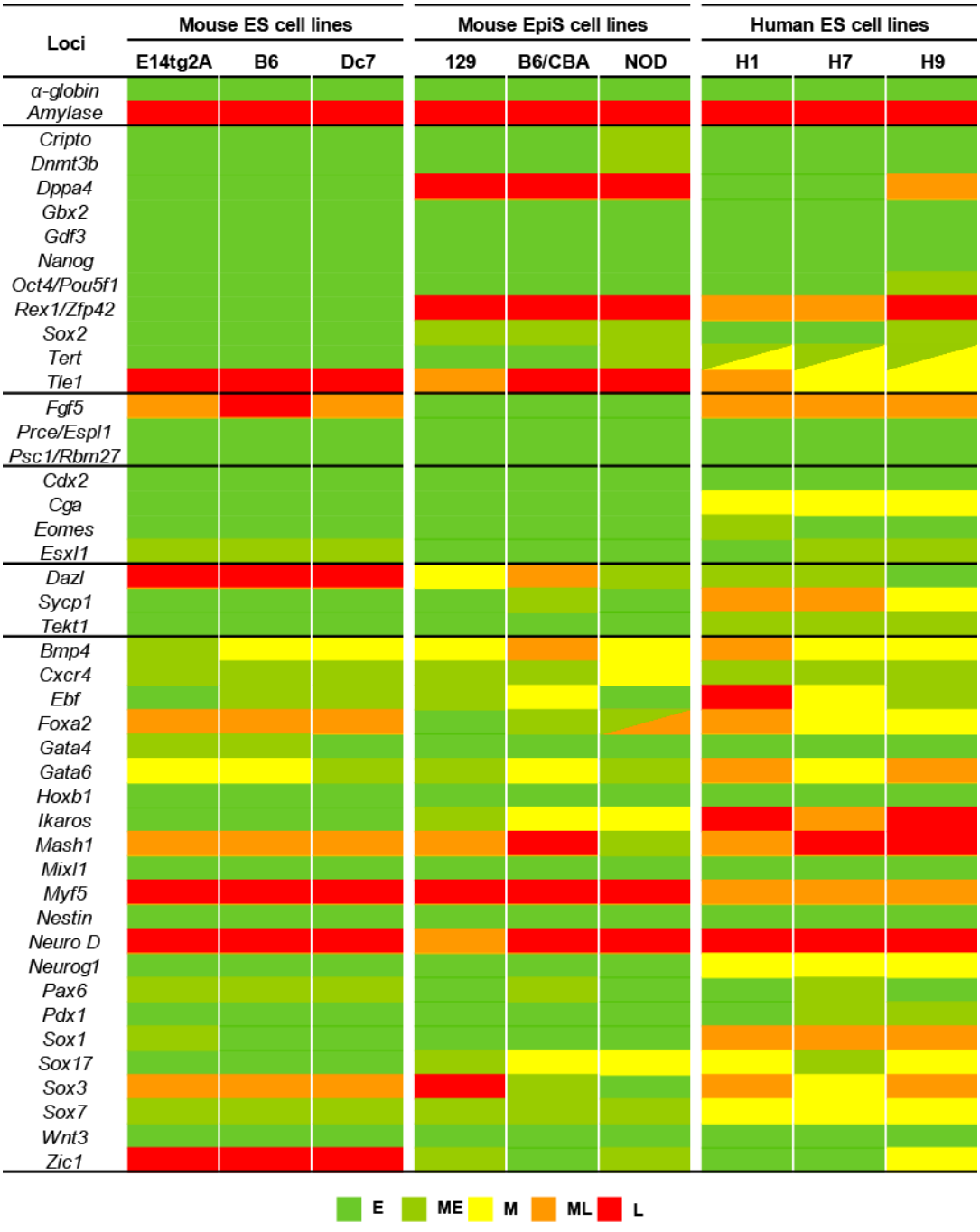


Figure A3 Comparative replication timing analysis of mES (E14tg2A, B6 and Dc7), mEpiS (129, B6/CBA and NOD) and hES (H1, H7 and H9) cell lines.
The replication timing of each gene was defined as early (E), middle-early (ME), middle (M), middle-late (ML) or late (L) and shown as a gradation of colours from green (earliest) to red (latest).

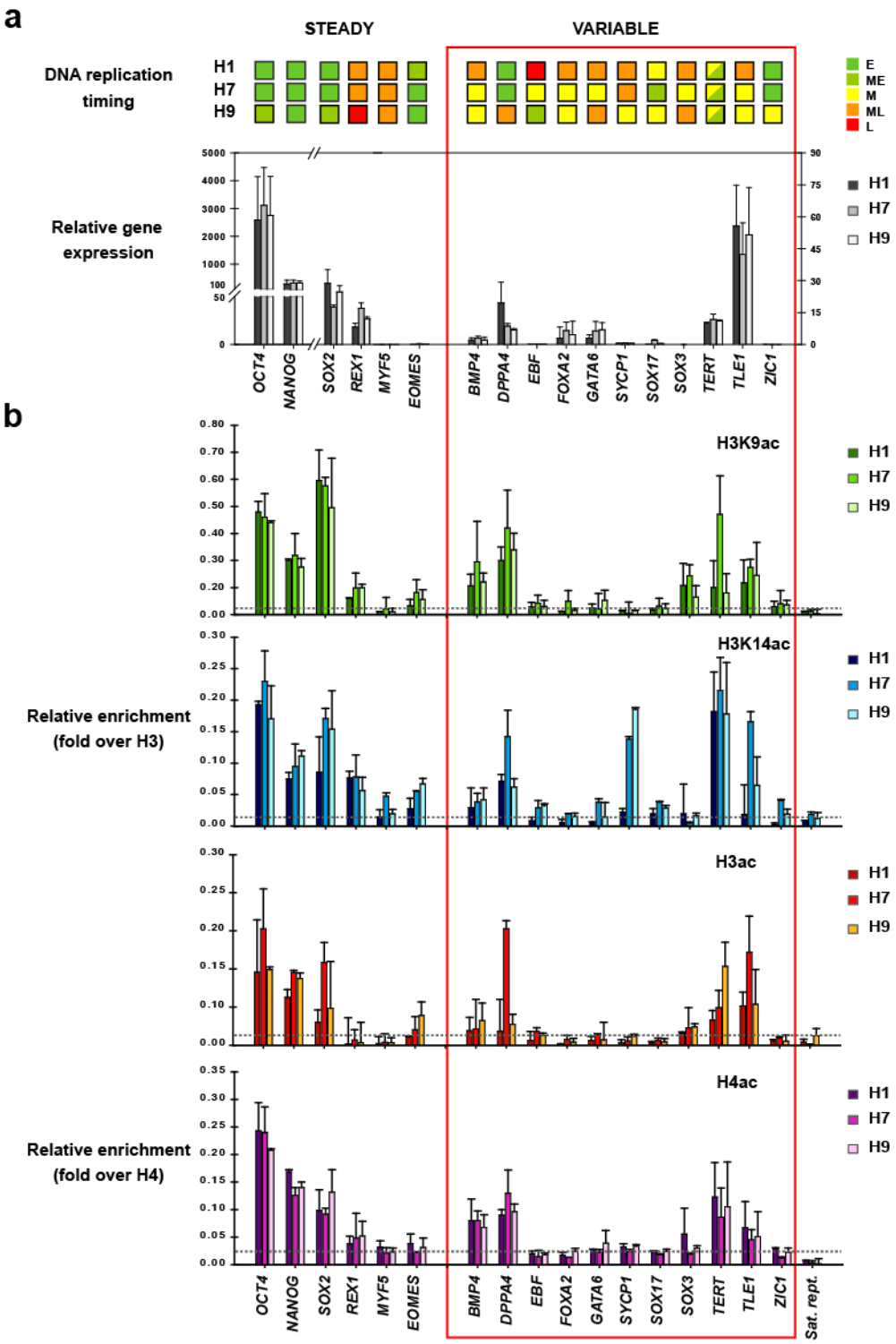


Figure A4 Human ES cells replication timing variability is not associated with gene expression or histone acetylation specific patterns.

(**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 qRT-PCR 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 qRT-PCR. 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 immunoprecipitation. 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.

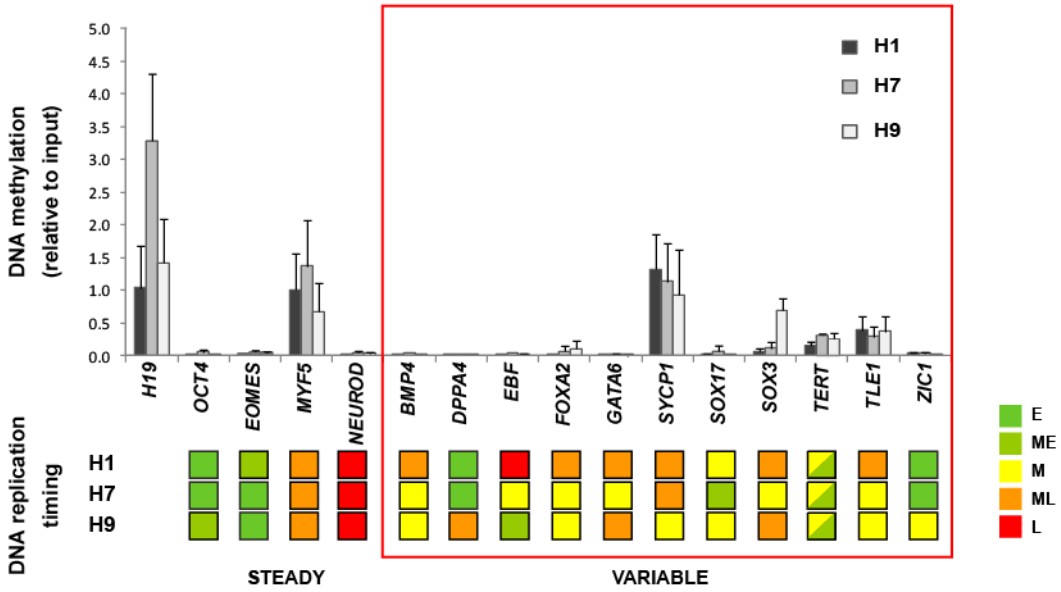


Figure A5 Human ES cells replication timing variability is not associated with DNA methylation specific patters.

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 qRT-PCR 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 analysed. 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.

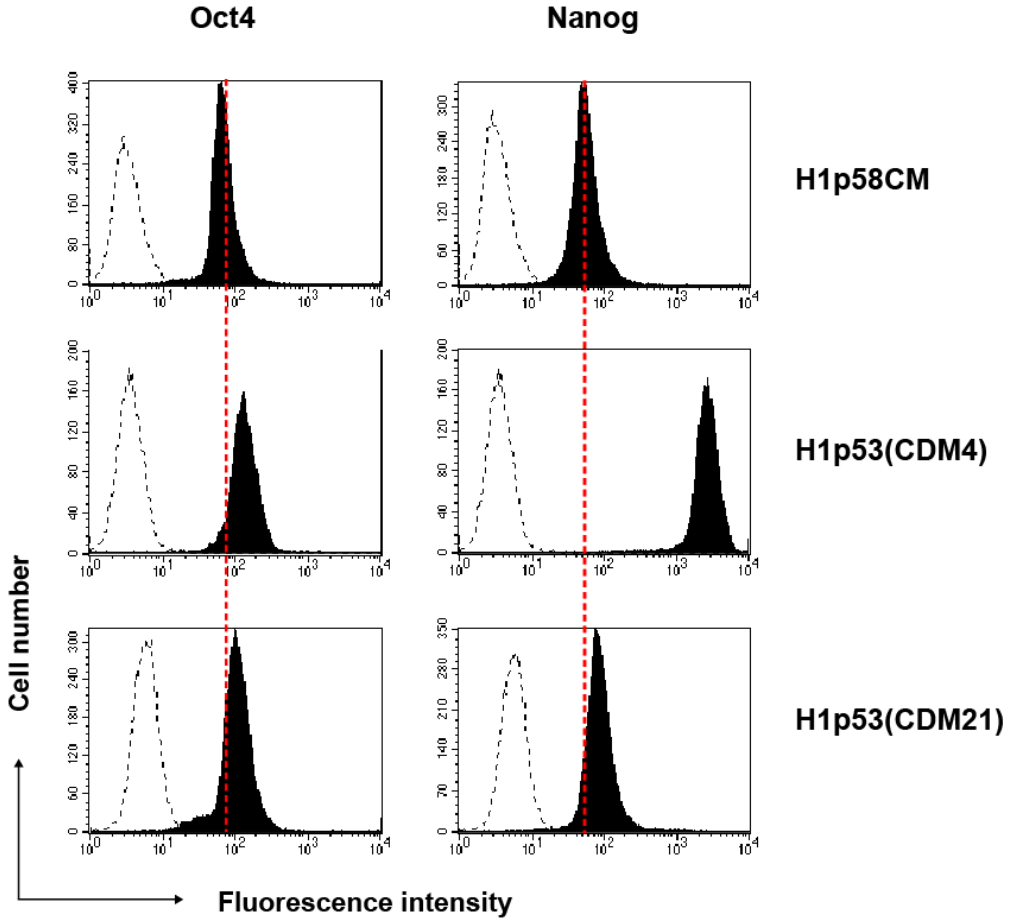


Figure A6 Transient increase in the expression levels of Nanog in H1 CDM^{activin} hES cells
Flow cytometry analysis of OCT4 (left panel) and NANOG (right panel) levels in H1 CM (H1p58CM) and CDM^{activin} early [H1p53(CDM4)] and long term cultures of CDM^{activin} [H1p53(CDM21)]. The red dotted lines 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.

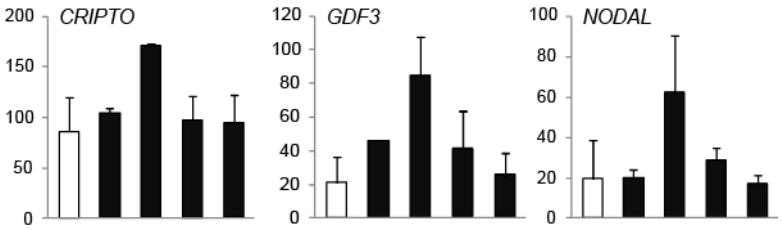


Figure A7 Differential gene expression profiles of TGF-β members in H1 CM and CDM^{activin} hES cell populations.
Real-time PCR analysis of selected genes in H1 CM (H1p49CM, white bars) and cells transferred and cultured for different passage numbers in CDM^{activin} conditions [H1p47(CDM3, 6, 14 and 25), black bars]. TGF-β superfamily members transiently increased (*CRIPTO*, *GDF3* and *NODAL*) in CDM^{activin} cells. Mean and standard error of at least two experiments are shown. Data was normalised to the expression of two housekeeping genes (*GAPDH* and *HPRT*).

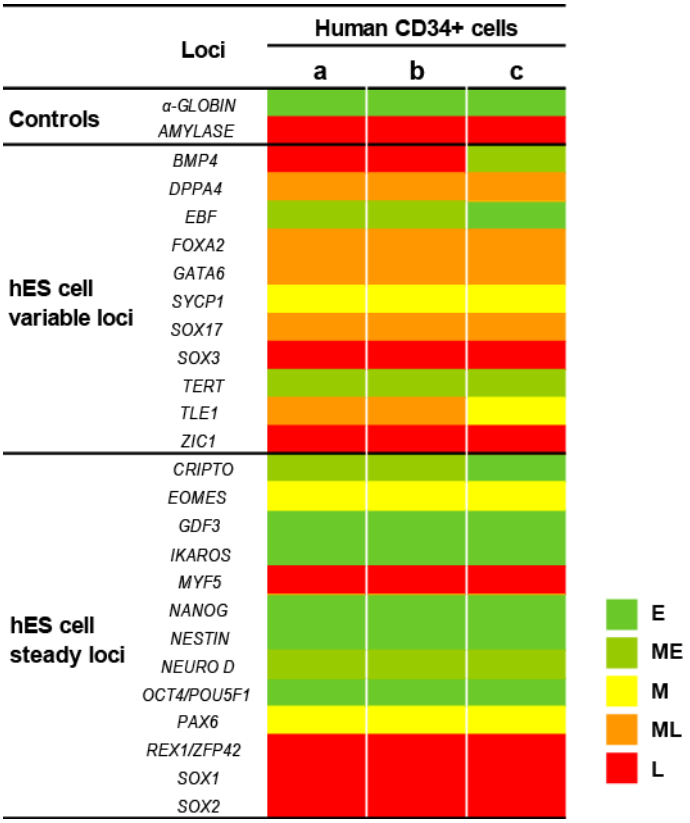


Figure A8 Replication timing variability is not observed across human CD34+ haematopoietic stem cells.
The replication timing of variable loci across different hES cell lines was analysed in CD34+ cells isolated from peripheral blood of three different healthy volunteers (here represented as **a**, **b** 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.**

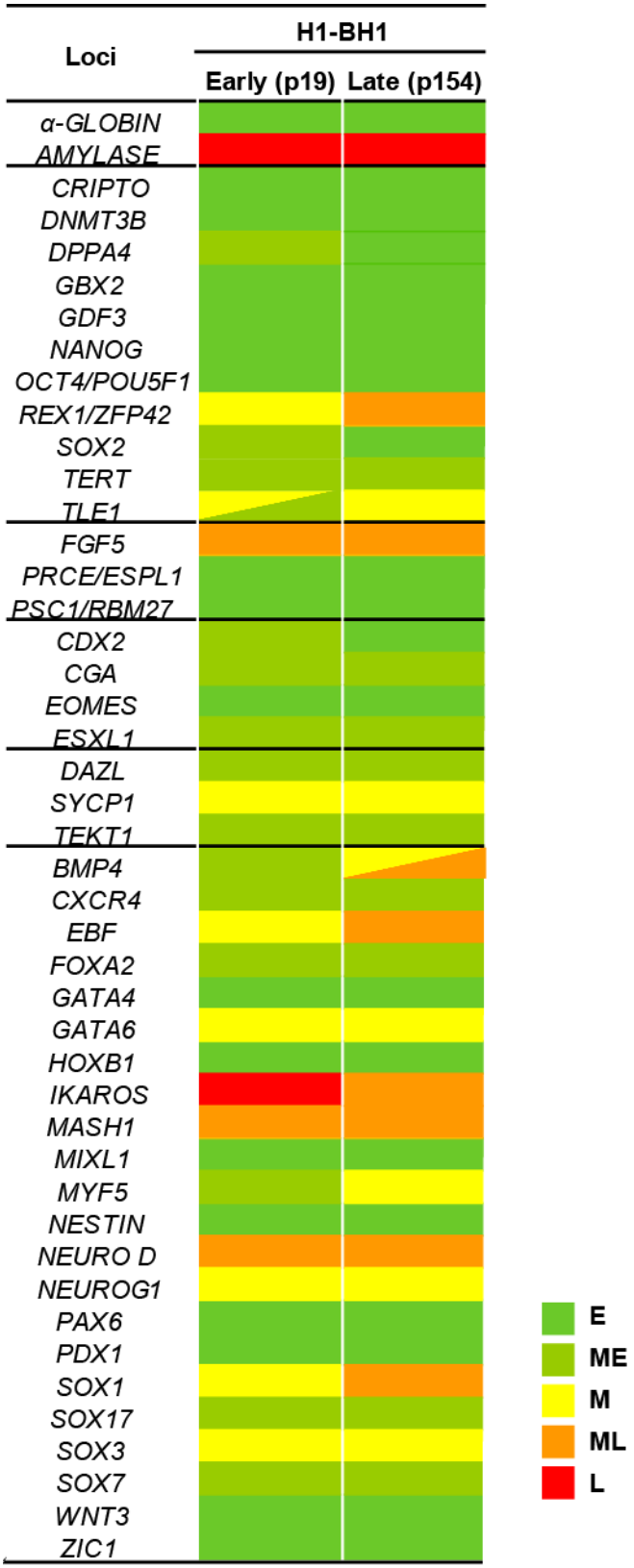


Figure A9 Cell culture adaptation and karyotype abnormalities are not associated with earlier replication timing profiles in hES cells.

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. 2007a)

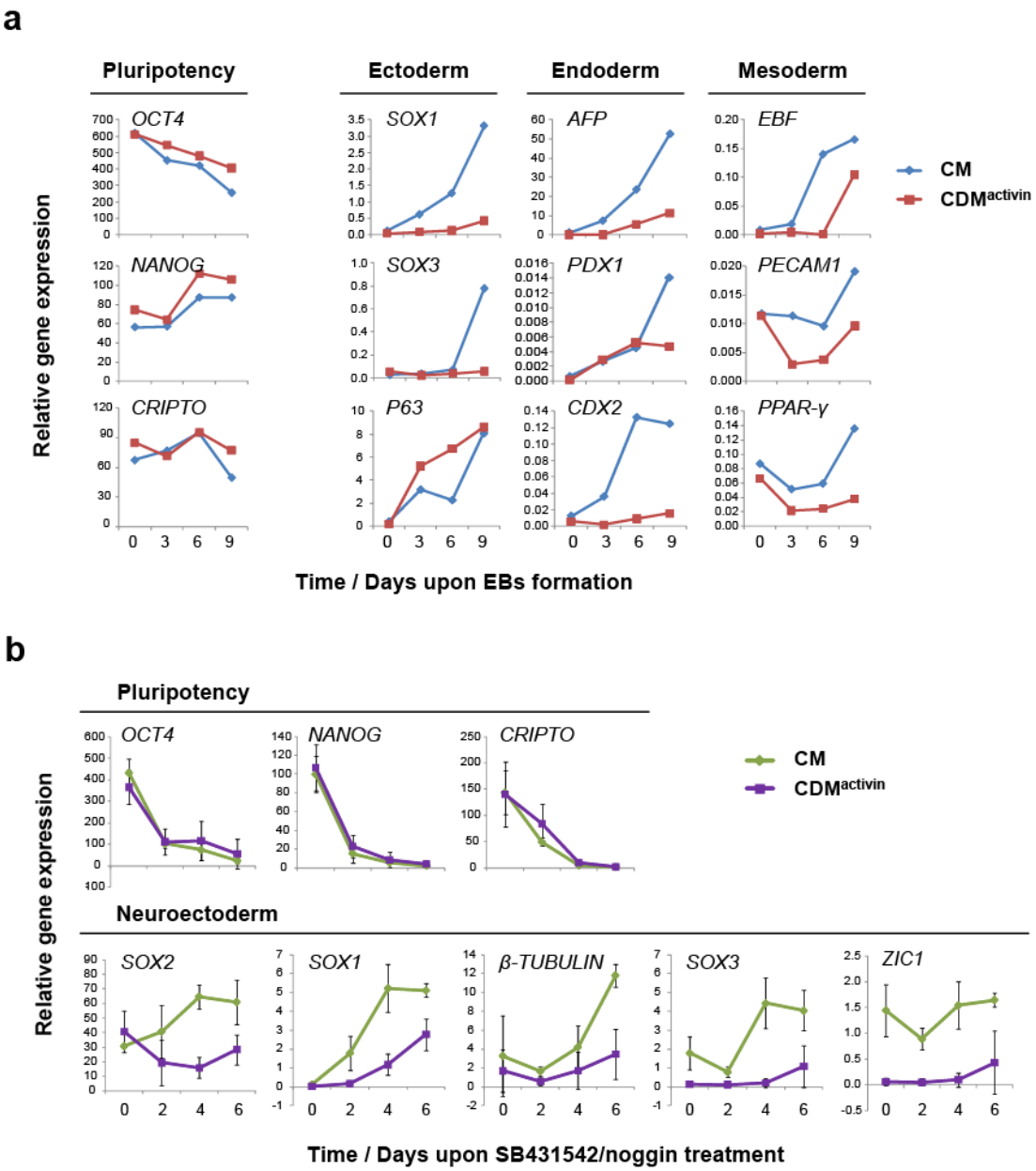


Figure A10 Human ES cells cultured in CDM^{activin} differentiate more slowly by EBs differentiation than hES cells cultured in CM

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 qRT-PCR analysis of three independent experiments. Data was normalized to the expression of two housekeeping genes (*GAPDH* and *HPRT*). Primers used for this analysis are listed in Appendix I, Table A1.

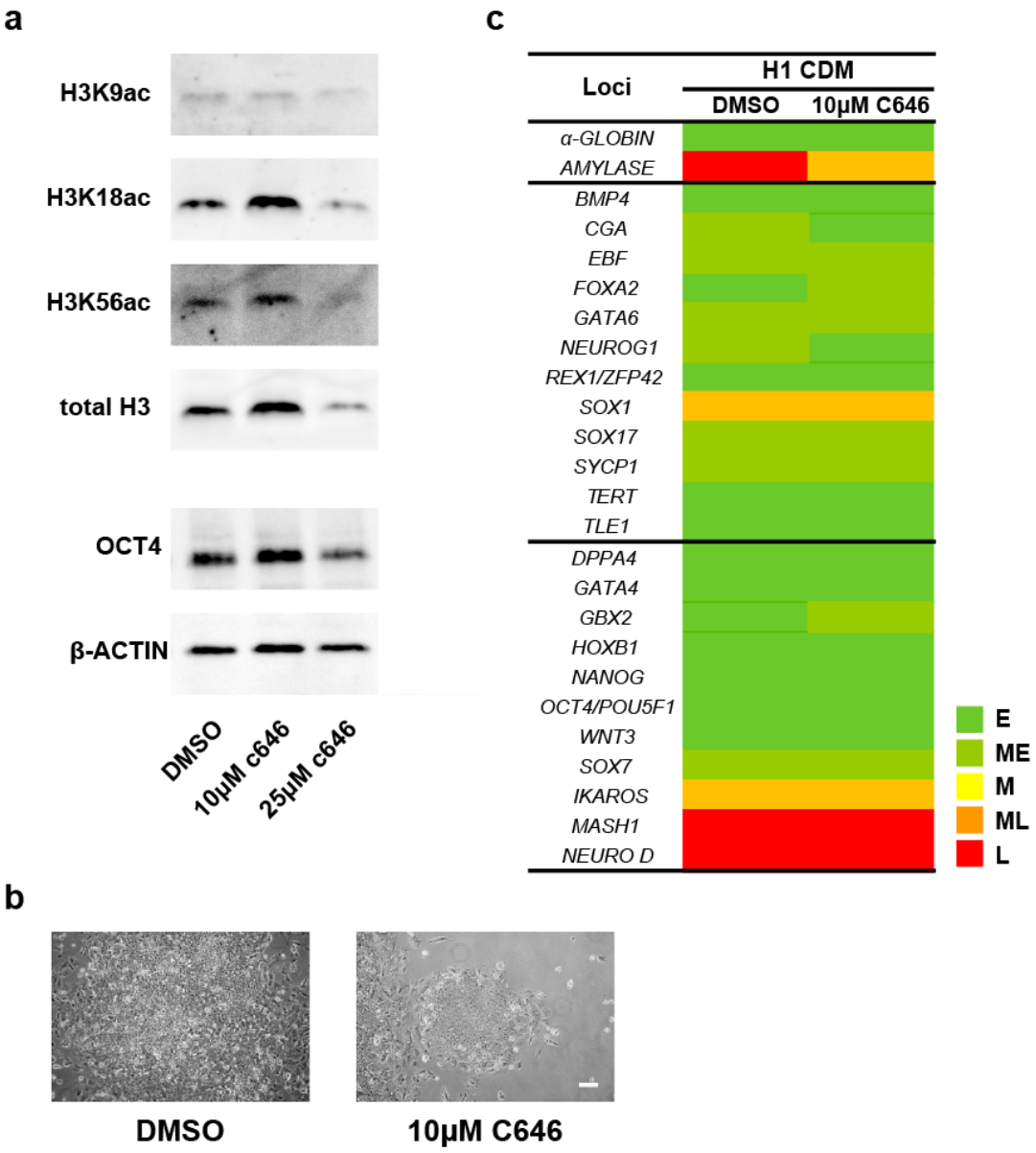


Figure A11 Inhibition of P300/CBP histone acetyltransferase activity by small molecule inhibitor C646 does not affect the replication timing profiles of H1 CDM^{activin} cultured cells.
(**a**) Western blot analysis of H1 CDM^{activin} hES cell populations treated for 24 hours with DMSO (solvent control), 10 μM or 25 μM C646, P300/CBP histone acetyltransferase activity inhibitor, and analysed after four days of culture. Membranes were probed with antibodies specific for p300/CBP preferentially acetylated histone H3 targets (H3K9ac, H3K18ac and H3K56ac) and total H3 and OCT4 and β-ACTIN (as a control for differentiation and protein loading). Equivalent protein loading is shown by the β-ACTIN detection. (**b**) Representative bright field images of H1 CDM^{activin} hES cell colonies treated with either DMSO or 10 μM C636 for four days. Scale bar 10 μm. (**c**) The replication timing of individual loci compared between H1 CDM^{activin} cell populations treated with DMSO or 10 μM C646 and analysed after four days of culture. No effect on replication timing was observed in 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*) or other loci that 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}. Data representative of two independent experiments.