

Epigenetic events underlying somatic cell reprogramming

by

Carlos Filipe Ribeiro Lemos Pereira

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

MRC Clinical Sciences Centre
Imperial College School of Medicine

September of 2008

“My soul is a hidden orchestra... all I know is the symphony”
(Fernando Pessoa, *in* Livro do Desassossego)

ABSTRACT

Although differentiated cells normally retain cell-type-specific gene expression patterns throughout their lifetime, cell identity can sometimes be modified or reversed *in vivo* by transdifferentiation, or experimentally through cell fusion or by nuclear transfer. Several studies have illustrated the importance of chromatin remodelling, DNA demethylation and dominant transcriptional factor expression for changes in lineage identity. Here the epigenetic mechanisms required to “reset” genome function were investigated using experimental heterokaryons.

To examine the epigenetic changes that are required for the dominant conversion of lymphocytes to muscle, I generated stable heterokaryons between human B-lymphocytes and mouse C2C12 myotubes. I show that lymphocyte nuclei adopt an architecture resembling that of muscle and initiate the expression of muscle-specific genes in the same temporal order as developing muscle. The establishment of this novel gene expression program is coordinated with the shutdown of several lymphocyte-associated genes. Interestingly, inhibition of histone deacetylase (HDAC) activity during reprogramming selectively blocks the silencing of lymphocyte-specific genes but does not prevent the establishment of muscle-specific gene expression.

In order to reprogram somatic cells to pluripotency, I fused human B-lymphocytes and mouse embryonic stem (ES) cells. The conversion of human cells is initiated rapidly, occurring in heterokaryons before nuclear fusion. Reprogramming of human lymphocytes by mouse ES cells elicits the expression of a human ES-specific gene expression profile in which endogenous hSSEA4, hFgf receptors and ligands are expressed while factors that are characteristic of mouse ES cells, such as Bmp4 and Lif receptor are not. Using genetically engineered mouse ES cells I demonstrate that successful reprogramming requires the expression of Oct4, but importantly, does not require Sox2, a factor implicated as critical for the induction of pluripotency. Following reprogramming, mOct4 becomes dispensable for maintaining the multi-potent state of hybrid cells. Finally, I have examined the reprogramming potential of embryonic germ (EG), embryonic carcinoma (EC) and ES cells deficient for the Polycomb repressive complex 2 (PRC2) proteins Eed, Suz12 and Ezh2. While EC and EG cells share the ability to reprogram human lymphocytes with ES cells, the lack of Polycomb proteins abolishes reprogramming. Thus, the repressive chromatin mark (H3K27 methylation) catalysed by PRC2 play a crucial role in keeping ES cells with full reprogramming capacity. Collectively my results underscore the importance of chromatin events during cell fate reprogramming.

ACKNOWLEDGEMENTS

I would first like to give my sincere thank you to Mandy for the stimulating and creative project and for her patience to teach me so much about science and writing. I have been privileged with you as a mentor. I want to thank Matthias for the constant amazing criticisms and for being a model for the unpretentious and committed attitude to science. Once again thanks to Mandy and Matthias for creating such a great scientific and friendly environment; this has made working in the Lymphocyte Development group a great experience in the past years.

I would like to thank Remi for his advice and support during the first year of my PhD, and for introducing and teaching the cell fusion experiments; Natalie, without her efficient help these experiments would not have been possible, for being a perfect “desk-neighbour” and for taking care of the lab social activity; Joana and Kelly for a really nice co-working experience; Tomomi for kindly reading my thesis; and all the past and present members of the lab for help, collaboration, advice or simply because they made work so much fun. I also want to thank the Clinical Sciences Centre for gathering so many enthusiastic people, particularly in the epigenetics, development and cancer section. In this context I want to acknowledge the precious contribution of people from the Developmental Epigenetics group, namely Miguel, Tatyana and Amy and from the Molecular Embryology group, Margarida (for the alkaline phosphatase kit). I would like to thank external collaborators: W. Cui (IRDB in London), A. Surani (Gurdon institute in Cambridge) and H. Niwa (RIKEN in Japan). Thanks also to the MRC football guys for the energetic games.

A special acknowledgement goes to the GABBA PhD Program for scientific training and above all for the freedom of choice. I am indebted to all my colleagues and teachers from the 7th edition for a fantastic year of learning and scientific discussion. Thanks to Prof. Maria de Sousa for being so inspirational. This work would not be possible without the generous financial support from the funding bodies FCT, Portugal, and the Medical Research Council, UK. I also want to thank all my friends in London that made this city a fantastic place to live and so difficult to leave. My very special thanks go to Sandra who has always been there; words are just not enough to acknowledge her contribution.

Finalmente ao meu pai António, minha irmã Liliana e especialmente à minha mãe Margarida agradeço o apoio incondicional, amor e a motivação para enfrentar qualquer desafio.

TABLE OF CONTENTS

ABSTRACT	3
ACKNOWLEDGEMENTS	4
TABLE OF CONTENTS	5
FIGURES AND TABLES	8
ABBREVIATIONS	11
CHAPTER 1. INTRODUCTION	15
1.1. Epigenetic levels of gene regulation	17
1.1.1. Chromatin structure and epigenetic inheritance	17
1.1.2. DNA methylation	20
1.1.3. Covalent histone modifications and chromatin dynamics	22
1.1.4. Transcriptional silencing by Polycomb group proteins	30
1.1.5. Spatial constraints – nuclear organisation	33
1.1.6. Temporal correlates – replication timing	36
1.2. Somatic cell reprogramming	38
1.2.1. Conversion between differentiated cell types	38
1.2.2. Reprogramming somatic cells to pluripotency	43
1.3. Aims of this study	60
CHAPTER 2. MATERIALS AND METHODS	61
2.1. Materials	61
2.1.1. Antibodies	61
2.1.2. Cell lines	62
2.2. Methods	64
2.2.1. Cell culture	64
2.2.2. Cell fusion and experimental heterokaryons	65
2.2.3. Reverse transcription (RT) polymerase chain reaction (PCR)	67
2.2.4. Genomic PCR	69
2.2.5. Cloning and DNA delivery into ES cells	69
2.2.6. Fluorescence activated cell sorting (FACS) analysis	72
2.2.7. Immunofluorescence (IF) analysis	72
2.2.8. Western blot analysis	74
2.2.9. Fluorescence <i>in situ</i> hybridisation (FISH)	75
2.2.10. Replication timing analysis	77

2.2.11. Chromatin immunoprecipitation analysis	78
2.2.12. DNA methylation analysis	79
2.2.13. Genomic context analysis	80

CHAPTER 3. LYMPHOCYTE REPROGRAMMING TO MUSCLE IN STABLE

HETEROKARYONS **81**

3.1. Reprogramming human lymphocytes in muscle heterokaryons	81
3.2. Increased nuclear volume and re-distribution of heterochromatin are early events in the reprogramming of lymphocytes to muscle	83
3.3. Reprogramming initiates a temporally ordered activation of genes encoding human myogenic regulatory factors by lymphocyte-derived nuclei	86
3.4. HDAC activity is required for the silencing of lineage inappropriate genes during reprogramming	87
3.5. Inhibition of HDAC activity results in the co-expression of <i>hCD20</i> and <i>hNCAM</i> in individual reprogrammed nuclei	90
3.6. Discussion and future perspectives	93

CHAPTER 4. HETEROKARYON-BASED REPROGRAMMING OF LYMPHOCYTES

FOR PLURIPOTENCY **98**

4.1. Generation of inter-species heterokaryons between human B-lymphocytes and mouse embryonic stem cells	98
4.2. Reprogramming of gene expression is initiated in ES cell heterokaryons prior to nuclear fusion	99
4.3. Induction of a human ES-specific gene expression profile and multi-lineage potential	103
4.4. Genes not easily re-activated in heterokaryons display heterochromatic features	108
4.5. Inter-species reprogramming of human B cells requires mOct4 but not mSox2	110
4.6. ES-derived factors are dispensable for maintaining the reprogrammed status of somatic cells	115
4.6. Discussion and future perspectives	119

CHAPTER 5. ROLE OF PCG PROTEINS IN THE MAINTENANCE OF ES CELL

REPROGRAMMING ACTIVITY **126**

5.1. Pluripotent cells reprogram human lymphocytes in heterokaryons	126
5.2. <i>Eed</i> is required for embryonic stem cell reprogramming activity	127
5.3. PRC2 activity and H3K27 methylation are required for the reprogramming mediated by ES cells	131
5.4. PRC1 cooperates with PRC2 to maintain the reprogramming ability of ES cells	133

5.5. The inability of PcG-deficient ES cells to reprogram is dominant over wild-type ES cells	134
5.6. Discussion and future perspectives	139
CHAPTER 6. GENERAL DISCUSSION	145
6.1. Inter-species heterokaryons as an experimental system for reprogramming studies	145
6.2. Acquisition and extinction of gene expression programs are separable events during reprogramming	146
6.3. Implications of reprogramming to pluripotency in heterokaryons	148
6.4. Mechanistic aspects of reprogramming towards pluripotency	148
6.5. Epigenetic stability and flexibility of cellular states in development	151
REFERENCES	152
APPENDIX I – PRIMER SEQUENCES	181
APPENDIX II – COMPLETE DATA SETS	186
PUBLICATIONS	194

FIGURES AND TABLES

Figure 1.1. Epigenetic levels of gene regulation.	18
Figure 1.2. Human histone tail regions with some of the post-translational modifications that are known to regulate gene activity.	24
Figure 1.3. The histone deacetylase family and inhibitors.	27
Figure 1.4. Polycomb repressive complexes.	33
Figure 1.5. Regulation of pluripotency.	49
Figure 1.6. Four strategies to induce reprogramming of somatic cells to pluripotency.	53
Figure 2.1. Impact of varying cell ratio in fusion efficiency.	66
Figure 2.2. Expression vectors used in this study.	70
Figure 2.3. Distinction of human and mouse nuclei within heterokaryons.	74
Figure 3.1. Heterokaryon formation between human B-lymphocytes and C2C12 myotubes.	82
Figure 3.2. Reprogramming human B-lymphocytes in C2C12 heterokaryons.	83
Figure 3.3. Nuclear reorganisation is an early event in lymphocyte reprogramming.	84
Figure 3.4. Early redistribution of heterochromatin in hBxC2 heterokaryons.	85
Figure 3.5. Rapid translocation of mouse Myogenin to human nuclei.	85
Figure 3.6. Ordered expression of human myogenic genes by lymphocyte-derived nuclei.	86
Figure 3.7. HDAC inhibition does not impair the activation of muscle gene expression program.	87
Figure 3.8. HDAC inhibition prevents the shutdown of lymphocyte-specific genes during reprogramming.	88
Figure 3.9. TSA treatment does not result in the mis-activation of other lineage related genes.	89
Figure 3.10. <i>hPax5</i> extinction in heterokaryons is not associated with a decline in the expression of <i>hE2A</i> and <i>hEbf</i> .	90
Figure 3.11. Co-expression of lymphocyte- and muscle-specific genes in individual nuclei treated with the HDAC inhibitors TSA and VPA.	91
Table 3.1. RNA FISH detection of <i>hCD20</i> and <i>hNCAM</i> transcripts in day-5 heterokaryons.	92
Figure 3.12. Kinetics of events involved in the multi-step reprogramming of lymphocytes to muscle in heterokaryons.	95
Figure 4.1. Generation of embryonic inter-species heterokaryons through cell fusion.	99
Figure 4.2. Heterokaryons can be maintained up to two days before nuclear fusion.	100
Figure 4.3. Reprogramming of human B-lymphocytes to embryonic state by mouse ES cells does not require nuclear fusion.	101
Figure 4.4. The human <i>Oct4</i> promoter is demethylated in heterokaryons.	103

Figure 4.5. Reprogrammed lymphocytes are comparable to human ES cells rather than mouse ES cells.	104
Figure 4.6. Reprogrammed cells express the human ES marker SSEA4.	105
Figure 4.7. Expression of human-specific embryonic markers in hybrid cells.	106
Figure 4.8. Multi-lineage potential is reset in reprogrammed human lymphocytes.	107
Figure 4.9. Some genes not easily re-activated in heterokaryons show heterochromatic features.	109
Figure 4.10. Oct4 migration to human nuclei precedes reprogramming.	111
Figure 4.11. Oct4 is required for successful reprogramming.	113
Figure 4.12. Sox2 is dispensable for reprogramming.	114
Figure 4.13. Derivation and characterisation of embryonic hybrids.	116
Figure 4.14. Reprogramming is self-sustaining and can be maintained in the absence of ES-derived Oct4.	117
Figure 4.15. Sox2 is not required to maintain the reprogrammed status of hybrid cells.	118
Figure 4.16. Events underlying reprogramming of lymphocytes in embryonic heterokaryons and hybrid cells.	120
Figure 5.1. Reprogramming potential of pluripotent cell lines.	127
Figure 5.2. <i>Eed</i> is required for the ES cell-mediated reprogramming of human lymphocytes.	129
Figure 5.3. Restoring <i>Eed</i> expression in <i>Eed</i> -deficient ES cells rescue their reprogramming defect.	130
Figure 5.4. Methylation of H3K27 in ES cells is required for successful reprogramming.	132
Figure 5.5. Conditional removal of <i>Ring1B</i> results in the decreased reprogramming activity of ES cells.	134
Figure 5.6. PRC2-deficient ES cells do not lack factors implicated in reprogramming.	135
Figure 5.7. De-repression of Polycomb target genes does not induce reprogramming to extra-embryonic lineages.	137
Figure 5.8. <i>Eed</i> -deficient ES cells dominate when fused to wild-type ES cells and human lymphocytes.	138
Figure 6.1. Molecular mechanisms underlying acquisition and silencing of gene expression programs during reprogramming.	147
Figure 6.2. A model for the role of Oct4, Sox2 and PRC2 in the reprogramming to pluripotency.	150
Table A1. Primers used for the selective amplification of human transcripts by semi-quantitative RT-PCR.	181
Table A2. Primers for the selective amplification of human or mouse transcripts by quantitative RT-PCR.	181
Table A3. Primers for the selective amplification of human genomic regions by quantitative PCR.	184
Table A4. Primers for the selective amplification of human bisulfite converted genomic regions by PCR.	185

Figure A1. Replication timing analysis of pluripotency-associated genes in human B-lymphocytes.	186
Figure A2. Kinetic of human lymphocyte reprogramming by mES cells after Oct4 ablation.	186
Figure A3. Kinetic of human lymphocyte reprogramming by mES cells after Sox2 ablation.	187
Figure A4. Reprogramming is self-sustaining and can be maintained in the absence of ES-derived Oct4.	188
Figure A5. Inhibition of <i>de novo</i> protein synthesis in heterokaryons did not alter the overall human pattern of gene expression.	189
Figure A6. Kinetic of human lymphocyte reprogramming by embryonic carcinoma cells.	189
Figure A7. Kinetic of human lymphocyte reprogramming by embryonic germ cells.	190
Figure A8. Kinetic of human lymphocyte reprogramming by mES cells after <i>Eed</i> ablation.	190
Figure A9. Kinetic of human lymphocyte reprogramming by mES cells after <i>Suz12</i> ablation.	191
Figure A10. Kinetic of human lymphocyte reprogramming by mES cells after conditional <i>Ezh2</i> removal.	191
Figure A11. Kinetic of human lymphocyte reprogramming by mES cells after conditional <i>Ring1B</i> removal.	192
Figure A12. m <i>Hand1</i> and m <i>Cdx2</i> are derepressed in PRC2-deficient ES cells.	192
Figure A13. siRNA-mediated knock-down of m <i>Oct4</i> abolishes reprogramming.	193

ABBREVIATIONS

μ	micro
AP	alkaline phosphatase
AT	adenine-thymine
ATP	adenosine tri-phosphate
Ara-C	cytosine β-D arabino furanoside
BAC	bacterial artificial chromosome
bp	base pair
BrdU	5-bromo-2-deoxyuridine
BSA	bovine serum albumin
cDNA	complementary DNA
C2	C2C12 cell line
ChIP	chromatin immunoprecipitation
Chromo	chromatin organisation and modifier
C(T)	threshold cycle
CpG	cytosine-guanine dinucleotide
CREST	calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia
CTCF	CCCTC-binding factor
DAPI	4,6-diamidino-2-phenylindole
DiD	1,1'-dioctadecyl-3, 3, 3', 3' tetramethylindodicarbocyanine
Dil	1,1'-dioctadecyl-3, 3, 3', 3'-tetramethylindocarbocyanine perchlorate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Dnmt	DNA methyltransferase
Dox	doxycycline
dpc	days post-coitum
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylene-dinitrilo tetraacetic acid
EC	embryonic carcinoma
EG	embryonic germ

EGFP	enhanced green fluorescent protein
ES	embryonic stem
FACS	fluorescence-activated cell sorter
FCS	foetal calf serum
FISH	fluorescent <i>in situ</i> hybridisation
FITC	fluorescein isothiocyanate
FGF	fibroblast growth factor
G	gram
GC	guanine-cytosine
GFP	green fluorescent protein
GMEM	Glasgow's modified Eagle's medium
h	hour
H	histone
HAT	histone acetyltransferase
hB	Epstein-Barr virus-transformed human B lymphocytes
HDAC	histone deacetylase
HDACi	histone deacetylase inhibitor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hES	Human embryonic stem
HMT	histone methyltransferase
HSC	haematopoietic stem cell
HP1	heterochromatin protein 1
ICM	inner cell mass
ICR	imprinting control region
IF	immunofluorescence
IL	interleukin
iPS	induced pluripotent stem
IRES	internal ribosome entry site
ISWI	imitation switch family of chromatin remodelling ATPases
K	lysine
Kb	kilobase pair
kDa	kiloDalton
KSR	knockout serum replacement
L	litre
LCR	<i>locus</i> control region
LIF	leukaemia inhibitory factor

LINE	long interspersed element
m	milli
M	molar
Mb	mega base pair
MBD	methyl-CpG binding domain
MEF	mouse embryonic fibroblast
mES	mouse embryonic stem
min	minute
MRF	muscle regulatory factor
mRNA	messenger RNA
miRNA	microRNA
NP-40	nonidet P-40
NuRD	nucleosome remodelling histone deacetylase
O/N	over-night
ORF	open reading frame
PBS	phosphate buffered saline
PcG	Polycomb group
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEV	position effect variegation
PFA	paraformaldehyde
PGC	primordial germ cell
PI	propidium iodide
pol	polymerase
PRC	Polycomb repressive complex
PRE	Polycomb response element
PrE	primitive endoderm
qPCR	real-time quantitative PCR
RA	retinoic acid
RITS	RNA induced transcriptional gene silencing
RNA	ribonucleic acid
RNAi	RNA interference
rpm	rotations per minute
RT	room temperature
RT-PCR	reverse transcription PCR
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i> , budding yeast

<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i> , fission yeast
SAHA	suberoylanilide hydroxamic acid
SDS	sodium dodecyl sulphate
siRNA	small interfering RNA
SET	Su(var)-Enhancer of zeste-Trithorax
Su(var)	suppressor of position effect variegation
SWI/SNF	(mating-type) switch / sucrose non-fermenting family of chromatin remodelling ATPases
TE	trophectoderm
TRE	trithorax response element
Tris	(hydroxymethyl)methylamine
TrxG	trithorax group
TS	trophoblast stem
TSA	trichostatin A
TSS	transcriptional start site
U	units
UTR	untranslated region
VPA	valproic acid
XEN	extra-embryonic endoderm stem
Xist	X inactivation-specific transcript

CHAPTER 1. INTRODUCTION

In higher eukaryotic organisms development is a remarkably orderly process. Starting from the totipotent zygotic cell an impressive array of differentiated cell types is generated in the adult. In humans at least 200 different cell types can be histologically defined with distinctive functions, such as the defence of the organism, muscle contraction or the establishment of neural networks. A central question in developmental biology asks how cells acquire different cell identities during development and how those are stably maintained.

In 1892 Weismann (Weismann, 1892) proposed that heritable elements (genes) are lost as cells became specialised. So according to this argument, a liver cell would lose genetic information except what was needed for liver cell function. Spemann (Spemann, 1938) made an early attempt to test this idea. However, it was following studies of Briggs & King (Briggs and King, 1952) that Gurdon and colleagues showed that, in *Xenopus laevis*, nuclear transfer from differentiated somatic cells to enucleated oocytes supported embryonic development (Gurdon, 1962). They reasoned that if the genetic material were lost, the differentiated state would be irreversible following nuclear transfer. Instead, they showed that the specialised intestinal cell nucleus retained the capacity to direct development of the resulting embryos into tadpoles and adult fertile frogs. The first mammal to be cloned from an adult cell, Dolly (Wilmut et al., 1997), and the cloning of mice from terminally differentiated lymphocytes (Hochedlinger and Jaenisch, 2002), demonstrate the theory of nuclear equivalence; that specialised cells of metazoans possess a gene pool identical to that of the zygote nucleus. Importantly, this established the principle that mechanisms underlying lineage restriction and cell identity are ultimately reversible.

We now know that differences in terminally differentiated cells arise from the selective expression of only a fraction of genes encoded by the genome. But how are differential transcriptional programs established and maintained during the process of cell specialisation? It has been long held, in contrast to lower vertebrate and invertebrate embryos, that early mammalian development is highly regulative and flexible. Experimental manipulation of the early pre-implantation embryo revealed that early blastomeres are remarkably plastic in their lineage potency (Rossant and Tam, 2004; Stern, 2006). The current view, however, is that the mammalian embryo has some order or bias from the very beginning (fertilisation), with inflexibility and cell determination increasing over time (Plusa et al., 2005; Zernicka-Goetz, 2005). Lineage

specification is often associated with signalling events, which in the presence of the appropriate gene regulatory network determine the specialised cell fate. These networks are modulated by the availability of lineage-specific transcriptional factors that leads to the activation and silencing of specific and distinct sets of genes. An example of such regulation can be observed very early in the mouse embryonic development, where the first differentiation event segregates embryonic and extra-embryonic lineages (Rossant, 2001). In the early blastocyst, cells form the inner cell mass (ICM) and trophectoderm (TE) can be distinguished by their requirements for extrinsic and intrinsic regulators. Among these are the transcription factors *Oct4* and *Cdx2* which are essential for development of the ICM and TE, respectively (Niwa et al., 2005). In addition, extrinsic factors such as the leukaemia inhibitory factor (LIF) and BMP4 maintain ES cell self-renewal (derived from ICM) and FGF4 is required for the maintenance of trophoblast stem (TS) cells (Chambers and Smith, 2004; Rossant, 2001). Later in ontogeny, key signalling ligands with well-known roles in embryo patterning are fundamental during myogenesis in muscle progenitor specification. These signals include Shh, Wnts, BMPs, Notch and FGFs, which are known to control the transcription factors *Myf5* and *MyoD* in muscle progenitors (Pownall et al., 2002). *MyoD* has been implicated as a dominant factor for myogenesis as its forced expression is sufficient to convert fibroblasts in skeletal muscle cells (Davis et al., 1987). The importance of single transcription factors for dictating lineage fate is also clear from studies of the transcriptional regulator *Pax5* in B-lymphocytes (Nutt et al., 1999). In this case, *Pax5* functions not only by the activation of the B-cell program, but also by repression of other lineage fates, since its removal favours the development of alternative haematopoietic cell fates from haematopoietic precursors (Busslinger et al., 2000).

There are formally only about 7 distinct signalling pathways through which cells modulate gene expression and the number of examples where a single transcription factor dictate a specific cell fate are limited. How is it then possible to generate such a huge diversity of cell types? Cells seem to accumulate a "history" of what signals they have received over time, and therefore the same signal will have different outcomes depending on the cellular context. For example, LIF signalling contributes to the self-renewal of ES cells but later on in development induces macrophage maturation and enhances the survival of sensory and motor neurons (Metcalf, 2003). The same rationale also applies to intrinsic regulators: when overexpressed in mature B-lymphocytes, the transcription factors *C/EBP α* and *C/EBP β* were shown to promote cell fate conversion to macrophages (Xie et al., 2004). However, the same transcription factors when overexpressed in pancreatic cells promote their conversion to

hepatocytes (Shen et al., 2000). Similarly, while *MyoD* was shown to be sufficient to generate muscle cells from a range of cell types (Choi et al., 1990; Davis et al., 1987), ectopic expression of *MyoD* in more distant related cell lineages to muscle, like ectoderm (Faerman et al., 1993), endoderm (Schafer et al., 1990) or ES cells (Dekel et al., 1992) does not result in complete conversion, indicating that developmental history do indeed impose constraints in tissue-specific transcription factor function. Allelic differences in gene expression (imprinting and allelic exclusion) and mammalian X-chromosome inactivation provided direct evidence that gene expression is not solely regulated in *trans* by diffusible signalling molecules and transcriptional regulators but also regulated in *cis* along the chromosome. Collectively these experiments provided evidence for cellular memory and epigenetic regulation of gene expression.

1.1. Epigenetic levels of gene regulation

The modern usage of the term “epigenetic”, originally coined by Conrad Waddington (Waddington, 1942), refer to heritable traits (over rounds of cell division and sometimes transgenerationally) that do not involve changes to the underlying DNA sequence. The ability to acquire and inherit gene expression programs efficiently is crucial for the process of cell differentiation. The molecular mechanisms that allow differentiated cells to perpetuate the “memory” and the developmental decisions that created it are now beginning to be revealed. These mechanisms involve heritable but potentially reversible modifications in chromatin structure, methylation of DNA, regulation by small interfering RNAs and higher-order genome organisation (Figure 1.1). DNA methyltransferases, histone-modifying enzymes, chromatin-binding proteins, non-coding RNAs and many other regulators may act together with sequence-specific transcriptional factors to establish and convey gene expression patterns during development, ensuring the stability of the “committed” state.

1.1.1. Chromatin structure and epigenetic inheritance

In eukaryotic cells, DNA is complexed with core histones and other chromosomal proteins in the form of chromatin (Figure 1.1). The basic repeating unit of chromatin, the nucleosome, includes the compact octameric histone complex consisting of two copies of each of the four core histones H2A, H2B, H3 and H4 wrapped by 146 basepairs (bp) of DNA (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. The positioning of nucleosomes with respect to the DNA template plays an important role in transcription regulation (Henikoff, 2008; Li et al., 2007a; Schones et al., 2008).

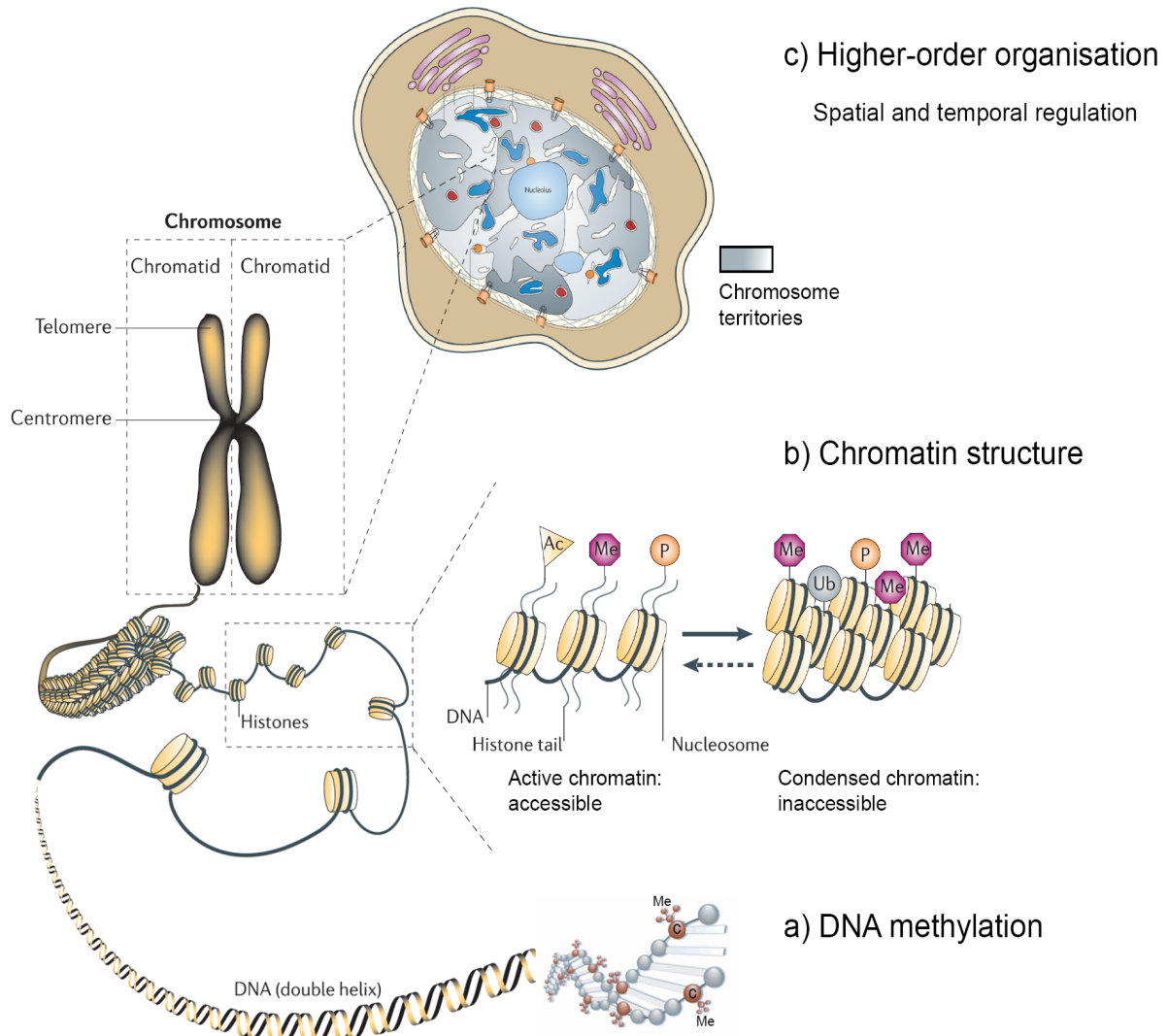


Figure 1.1. Epigenetic levels of gene regulation. Schematic representation of different levels of genome organisation that have an impact on gene activity. **(a)** The DNA double helix carries the genetic information used for development and functioning of all known living organisms. Beyond the sequence, DNA methylation (Me) at the carbon 5 of cytosines (C) provides the most direct epigenetic mechanism of gene regulation (see section 1.1.2). **(b)** DNA is wrapped around histones to form a “beads on a string” structure that is folded into higher order chromatin domains. Specific amino acids within histone tails (black wavy lines) are targets for a number of covalent modifications, including acetylation (Ac), methylation (Me), phosphorylation (P) and ubiquitylation (Ub). Dynamic changes in chromatin structure are influenced by post-transcriptional modifications of the N-terminal tails of histones, chromatin remodelling complexes, histone variants and small interfering RNAs (see sections 1.1.3 and 1.1.4). The nucleosome particles that comprise chromatin are represented as yellow cylinders and DNA as black lines. **(c)** Nucleosomal arrays are packed into several levels of condensed, higher-order chromatin domains and organised in functional nuclear compartments. These three-dimensional structures along with temporal constraints provide additional levels of regulation (see sections 1.1.5 and 1.1.6). Together these mechanisms provide some molecular clues on how a single genome can produce such a diversity of cell types. Adapted from (Lancot et al., 2007; Sparmann and van Lohuizen, 2006).

Specific chromatin remodelling proteins that dynamically regulate nucleosome positioning and/or nucleosome structure were identified. Many of these proteins are ATPases that often function as subunits of large complexes, including the SWI/SNF, ISWI and CDH complexes (Saha et al., 2006; Tsukiyama, 2002). With the aid of additional proteins, including histone H1, nucleosomal arrays are packed into several levels of condensed, higher-order chromatin domains, which are organised in specific nuclear compartments (Figure 1.1) (Maresca and Heald, 2006; Tremethick, 2007). Together, this folding can compact DNA by a factor of 30-40 fold (Widom, 1989), but the topology of higher-order chromatin structure, especially that of the most condensed interphase chromosomes, remain controversial (Goetze et al., 2007; van Holde and Zlatanova, 2007).

Early cytological studies observed that interphase chromatin is present in two distinguishable states: a less condensed form, euchromatin, and a more condensed form, heterochromatin, which stained intensely with specific dyes (Craig, 2005; Heitz, 1928). Euchromatin is a form of chromatin that is decondensed during interphase, highly accessible to nucleases (suggesting that it may also be more accessible to transcription factors) and contains actively transcribed genes (Kornberg and Lorch, 1999). In contrast, the higher-order structures of heterochromatin are thought to favour gene silencing, as it was initially demonstrated by the phenomenon of position effect variegation (PEV). The positioning of genes in close proximity to heterochromatin, often results in gene silencing, giving rise to a mosaic pattern of expression within a given tissue. This heritable silencing phenomenon was first identified in *Drosophila* (Muller and Altenburg, 1930), and has been reported in other organisms from yeast to humans (Dillon and Festenstein, 2002). At least two types of heterochromatin have been distinguished: 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 and facultative heterochromatin which is developmentally regulated and can be formed at *loci* that in other cell-types are found in euchromatic form. One obvious example of facultative heterochromatin formation is the inactivation of one of the two X-chromosomes in early embryonic development (Arney and Fisher, 2004; Heard, 2005).

In the course of cell division the genome is faithfully duplicated and divided between two daughter cells. For this purpose the cell must replicate its DNA sequence but also the underlying information in chromatin structure as well as its higher-order packaging in the nucleus. Given that each of the levels of chromatin organisation can be propagated throughout cell generations, this could represent the molecular basis of heritable traits (Holliday, 1994; Patterton and Wolffe, 1996). Such maintenance of

clonal gene expression patterns achieved by chromatin modifications is referred to as epigenetic inheritance, which potentially provides means to explain how a single genome can produce such a diversity of cell types. However, chromosome structure must also be dynamic and capable of regulated unfolding–folding transitions to allow the accessibility of proteins involved in various cellular processes such as transcription, replication and DNA repair. Indeed, recent studies have revealed the highly dynamic nature of chromatin and chromosome structure at all levels (Misteli, 2007). Uncoupling the chromatin changes that are “locked” and stably propagated throughout cell division from those that are transiently regulated constitutes a big challenge to define the “epigenetic” value of chromatin.

1.1.2. DNA methylation

There is a wide interest in understanding how chromatin structure impacts on gene regulation, and the mechanistic relationships between epigenetic phenomena (Lande-Diner and Cedar, 2005). DNA methylation provides an example of a direct epigenetic mechanism for the maintenance of the repressed state (Figure 1.1). Methylation of DNA within the mammalian genome is catalysed by a family of DNA (cytosine-5)-methyltransferases (Bestor, 2000) and occurs primarily at CpG dinucleotides. CpG methylation, especially within the promoter region of genes, is associated with stable transcriptional repression and plays important roles for X-chromosome inactivation, genomic imprinting and cell differentiation (Bird, 2002; Reik and Dean, 2001). Aberrant DNA methylation contributes to tumorigenesis and other diseases (Jones and Baylin, 2002).

Methylation patterns are propagated in mitotic cells by maintenance methylation activity that takes place in concert with DNA replication. The methyltransferase Dnmt1 performs this maintenance function (Bestor et al., 1988; Li et al., 1992). Consistent with its role, Dnmt1 has a substrate preference for hemimethylated DNA *in vitro* and is usually associated with replication foci in the S-phase of dividing cells (Chuang et al., 1997; Leonhardt et al., 1992). As DNA methylation patterns are stably maintained in differentiated mitotic cells, new patterns arise during embryonic cell differentiation and germ line specification in development. Dnmt3a and Dnmt3b, the two closely related *de novo* methyltransferases, are required for this process (Bestor, 2000). Inactivation of both genes cause a complete failure to establish genome-wide methylation (Okano et al., 1999). Consistent with this, the expression of Dnmt3a and Dnmt3b is developmentally regulated. Both genes are expressed at a high levels in

undifferentiated cells, including embryonic stem cells, and undergo down-regulation upon cell differentiation (Chen et al., 2002; Watanabe et al., 2002). Accordingly, the overall patterns of DNA methylation change dramatically during development. In early embryonic development parental methylation is lost, but methylation levels are restored at the blastocyst stage. Methylation patterns that are established at imprinted *loci* at this stage, such as the *Igf2/H19 locus*, are preserved in all cells of the adult organism, except for germ lineage cells (Edwards et al., 2007). A well characterised example of a *locus* regulated by DNA methylation is the ES-specific factor *Oct4*, which undergoes stable repression with cell differentiation. In differentiating human and mouse EC or ES cells and mouse post-implantation embryos, the inactivation of *Oct4* involves recruitment of transcriptional repressors, alterations in chromatin structure and DNA methylation at the promoter region (Deb-Rinker et al., 2005; Feldman et al., 2006; Gu et al., 2006; Hattori et al., 2004). Dnmt3a and Dnmt3b were shown to function synergistically to methylate the promoter region of the *Oct4* gene (Li et al., 2007b), which along with further recruitment of the methyl-CpG binding proteins MBD2 and MBD3 (Gu et al., 2006) may play an important role in the stable inactivation of *Oct4* in differentiated tissues (Tsuji-Takayama et al., 2004). Functionally, DNA methylation is thought to interfere with gene transcription by at least two independent mechanisms (Bird and Wolffe, 1999). Firstly, methyl-CpGs can directly prevent transcription factors from binding to their recognition sites on the DNA. For example, the parental allele-specific binding of CCCTC-binding factor (CTCF) at the *Igf2/H19* imprinting control region (ICR). The binding of CTCF to the *Igf2/H19* ICR is prevented by CpG methylation (Holmgren et al., 2001). Secondly, methyl-CpGs can recruit a protein family sharing the methyl CpG-binding domain (MBD) which indirectly was shown to induce silencing by recruiting co-repressor complexes to methylated DNA (Boyes and Bird, 1991), including HDACs or histone methyltransferases (HMTs) (Bird, 2002). Likewise, DNA methylation is impaired across pericentric major satellite repeats in ES cells deficient for the HMT Suv39h, due to abrogation of pericentric Dnmt3b localisation (Lehnertz et al., 2003). The links between DNA and histone methylation are thought to be important for propagating (hetero-) chromatin states through mitosis.

It was suggested that DNA methylation is not a primary silencing mechanism, but rather one that makes silencing more stable and potentially irreversible (Jones and Takai, 2001; Li, 2002; Pannell et al., 2000). For example, DNA methylation was shown to follow the establishment of gene silencing and onset of histone 3 lysine 27 methylation during X-chromosome inactivation (Heard, 2005). Likewise, in absence of Dnmt1 X inactivation is unstable, as suggested by random reactivation of an X-linked transgene in mouse embryonic cells and in cultured somatic cells when Dnmt1 was

absent or inhibited (Sado et al., 2000). Upon ES cell differentiation, *Oct4* promoter is methylated after transcriptional repression (Feldman et al., 2006). This view has been recently challenged with evidence that DNA methylation can be dynamically regulated. 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). These results, together with the exceptional occasions when DNA methylation is broadly erased (after fertilisation, during germ cell development and somatic cell reprogramming), implies a mechanism for CpG demethylation. DNA demethylation can be passive, presumably brought about during DNA replication by exclusion of Dnmt1, or active. Despite the intensive research, the mechanism for active erasure of the methylation mark remains largely unknown. 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), so demethylation may proceed by pathways that involve base-excision or mismatch repair (Barreto et al., 2007; Morales-Ruiz et al., 2006; Morgan et al., 2004; Reik, 2007). The DNA deaminases Aid and Apobec1 have been shown *in vitro* to deaminate 5-methylcytosine in DNA to thymine (Morgan et al., 2004); this results in T-G mismatches, which can be repaired by the base-excision repair pathway. Indeed, Aid and Apobec1 are highly expressed by oocytes, stem cells and germ cells (Morgan et al., 2004). Another DNA-damage-responsive gene, *Gadd45* (growth arrest and DNA-damage-inducible 45), might also have a role in demethylation but it remains controversial (Barreto et al., 2007; Jin et al., 2008).

1.1.3. Covalent histone modifications and chromatin dynamics

Chromatin modification also occurs at the core protein components of nucleosomes, the histones (Figure 1.1), which are among the most highly conserved eukaryotic proteins. These are small basic proteins consisting of a globular domain and a more flexible, charged NH₂-terminus (histone tail) that protrudes from the nucleosome. Indeed, a number of enzymes modify histone tails amino acids at specific positions. These post-translational modifications include acetylation (of lysine residues), methylation (lysine and arginine), phosphorylation (serine and threonine), ubiquitylation (lysine), sumoylation (lysine) and ADP-ribosylation, which are thought to modulate the biological role of underlying DNA (Figure 1.2) (Bhaumik et al., 2007; Peterson and Laniel, 2004; Santos-Rosa and Caldas, 2005). Although histone

modifications were identified decades ago, only very recently their functional significance has been elucidated and the enzymes responsible for adding and removing each of these marks, characterised (Kouzarides, 2007). Distinct histone amino-terminal modifications can recruit additional chromatin-associated proteins, which may be important for regulating transition between transcriptionally active and silent chromatin states (Jenuwein and Allis, 2001). These covalent histone modifications provide an attractive “annotation” system for mitotically inheritable information that can be “read” by various effector proteins. By regulating access to underlying DNA, histone modifications and effector proteins can potentially dictate correct cell type-specific gene expression patterns (Jenuwein and Allis, 2001). The total number of possible different combinations of tail modifications is in excess of a thousand. Notably, several studies have indicated that specific combinations of histone modifications correlate well with a particular biological function. For instance, the combination of H4K8 acetylation, H3K14 acetylation, and H3S10 phosphorylation is often associated with transcription. Conversely, tri-methylation of H3K9 and the lack of H3 and H4 acetylation correlates with transcriptional repression in higher eukaryotes (Peterson and Laniel, 2004). The histone H3 tail appears to exist in two distinct states that are likely to be regulated by a “switch” between modifications of neighbouring residues. H3S10 phosphorylation inhibits H3K9 methylation but is synergistically coupled with H3K9 or H3K14 acetylation (Rea et al., 2000). This lead to the hypothesis that distinct histone modifications act sequentially or in combination to form an “histone code” that is read by other proteins to bring about distinct downstream biological events (Jenuwein and Allis, 2001; Spotswood and Turner, 2002).

1.1.3.1. Histone acetylation

One of the most consistent associations between histone modifications and chromatin structure is the underacetylation of all four histones, present at heterochromatin (Jeppesen and Turner, 1993; Maison et al., 2002). Conversely, histone hyperacetylation correlates with chromatin accessibility (Bode et al., 1980; Lee et al., 1993) and transcriptional activity (Allfrey et al., 1964; Pogo et al., 1966). These associations have been recently confirmed by several large-scale studies mapping histone modifications across the genome (Birney et al., 2007; Koch et al., 2007). Moreover, the role of histone acetylation on chromatin dynamics was functionally addressed *in vivo* by the incorporation of H4 homogeneously acetylated at lysine 16 into nucleosome arrays. This led to chromatin compaction defects equivalent to deletion of the entire H4 tail (Shogren-Knaak et al., 2006). Hypoacetylation also seems

containing bromodomains, which include members of the SWI/SNF class of ATP-dependent chromatin remodellers (Agalioti et al., 2002; Loyola and Almouzni, 2004).

Histone acetylation equilibrium is maintained by the balance of histone acetyltransferase (HAT) and HDAC activities (Wade et al., 1997). Turnover cycles of histone acetyl-groups have been shown to be very fast, with half-lives ranging from only a few minutes to several hours. This suggests that histone acetylation is highly dynamic and can rapidly change in response to external cues (Spotswood and Turner, 2002). The identification of histone modifying enzymes themselves remained elusive until the first nuclear HAT, a *Tetrahymena* homolog of yeast Gcn5 (Skvirsky et al., 1986), was identified in 1996 (Brownell et al., 1996). *In vivo* studies in yeast had previously characterised Gcn5 as a transcriptional co-activator protein, and thus its identification as a HAT solidified the view that histone modifications directly regulate transcription. Subsequently, a variety of other transcriptional co-activators, such as CBP/p300, TAF130/250 and P/CAF (a mammalian factor similar to Gcn5) were found to have intrinsic HAT activity (Lee and Workman, 2007; Peterson and Laniel, 2004), and many co-repressors, such as Rpd3, were found to have HDAC activity (Taunton et al., 1996). It is clear that targeting of histone acetylation is achieved by direct interactions between histone modifying enzymes and sequence-specific transcriptional regulators. For instance, the yeast HAT complex SAGA interacts with the transcriptional activation domains of a variety of gene-specific activator proteins, and these interactions target histone acetylation to specific promoter regions *in vivo* (Grant et al., 1998). Likewise, HDACs associate with several co-repressor complexes: HDAC 1 and 2 are found in the ubiquitously expressed mSin3A, NuRD/Mi2/NRD and CoREST co-repressor complexes. HDAC3 associates to and is activated by SMRT and NCoR co-repressors that play an important role in the regulation of gene expression by nuclear hormone receptors (Pazin and Kadonaga, 1997; Yang and Seto, 2003). Moreover, in cycling myoblasts, interactions between HDAC1 and MyoD (Mal et al., 2001), and HDAC4 and 5 and Mef2 (Lu et al., 2000) have been implicated in suppressing differentiation until specific cues (such as serum deprivation) initiate terminal differentiation. Under these conditions, pRb becomes dephosphorylated, sequesters HDAC1 from the MyoD-HDAC1 complex and releases MyoD to promote terminal muscle differentiation (Puri et al., 2001).

HDACs are not redundant in function (Bolden et al., 2006; Marks and Dokmanovic, 2005). Eighteen HDACs have been identified in humans, and they are subdivided into four classes based on their homology to yeast HDACs, their subcellular localisation and their enzymatic activities (Figure 1.3) (Bolden et al., 2006; Thiagalingam et al., 2003). The class I HDACs (1, 2, 3 and 8) are related to the yeast

Rpd3 deacetylase and can generally be detected in the nucleus and show ubiquitous expression in various human cell lines and tissues. Class II HDACs (4, 5, 6, 7, 9 and 10) are related to the yeast Hda1 protein and can shuttle between the nucleus and the cytoplasm and their expression is tissue-restricted. HDAC6 and 10 form a subclass (class IIb) that contain two deacetylase domains and are found predominantly in the cytoplasm. Members of a third class, sirtuins, require NAD⁺ for their enzymatic activity (Blander and Guarente, 2004). The enzymatic activity of class III HDACs is not inhibited by compounds such as trichostatin A (TSA) that broadly inhibit the other HDAC classes, which are zinc-dependent enzymes. HDAC11 is the sole member of class IV, it shares sequence similarity with the catalytic core regions of both class I and II enzymes but does not have strong enough identity to be placed in either class (Gao et al., 2002).

HDACi such as TSA or valproic acid (VPA), disrupt the acetylation equilibrium, leading to a general increase in bulk histone acetylation. HDACi therefore provide good tools to analyse the cellular effects of histone acetylation. Importantly, HDACi selectively alter gene transcription. Studies using cDNA arrays showed only 7–10% of genes were altered in their expression in several different cell lines upon HDAC inhibition (Chambers et al., 2003a; Glaser et al., 2003; Iezzi et al., 2004; Mitsiades et al., 2004), suggesting that the effect of hyper-acetylation in gene expression is cell-type dependent. The structural details of the HDAC–HDACi interaction have been elucidated in studies of a histone deacetylase-like protein from an anaerobic bacterium with TSA and suberoylanilide hydroxamic acid (SAHA) (Finnin et al., 1999). More recently, the crystal structure of HDAC8–hydroxamic inhibitors interaction has been solved (Somoza et al., 2004; Vannini et al., 2004). These studies provided insight into the deacetylation mechanism of acetylated substrates. The hydroxamic acid moiety of the inhibitor directly interacts with the zinc ion at the base of the catalytic pocket. However, the diverse structural characteristics of HDACi translates in differential specificity for HDAC classes or proteins (Xu et al., 2007). While some HDACi such as TSA and SAHA broadly inhibit deacetylase activity, others selectively inhibit different HDACs (Figure 1.3). For example, depsipeptide reportedly possesses stronger activity against HDAC1 and HDAC2 than against the HDAC4 and HDAC6 (Furumai et al., 2002). MS-275 preferentially inhibits HDAC1 with IC₅₀, at 0.3 mM, compared to HDAC3 with an IC₅₀ of about 8 mM, and has no inhibitory effect against HDAC6 and HDAC8 (Hu et al., 2003). Finally, two novel synthetic compounds, SK7041 and SK7068, preferentially target HDAC1 and 2 (Kim et al., 2003).

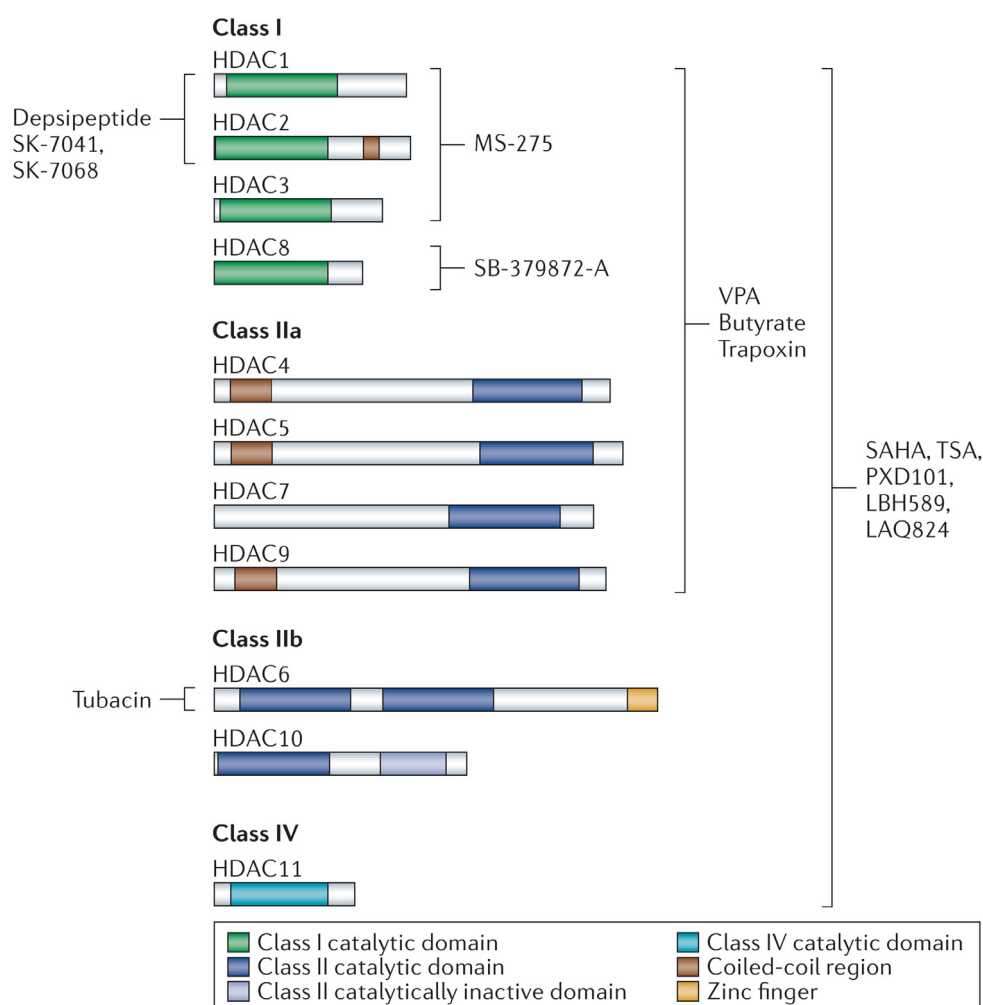


Figure 1.3. The histone deacetylase family and inhibitors. The figure shows the class I (HDAC1, 2, 3 and 8), class IIa (HDAC4, 5, 7 and 9), class IIb (HDAC6 and 10) and class IV (HDAC11) HDACs with the various structural/functional domains listed (lower panel). The capacity of structurally diverse HDACi to inhibit the activity of different HDAC classes or specific HDACs is also shown. SAHA, suberoylanilide hydroxamic acid; TSA, trichostatin A; VPA, valproic acid (adapted from Bolden et al, 2006).

1.1.3.2. Histone methylation

Histone methylation targets selected arginine and lysine residues in the N-terminal tails of histones H3 and H4 (Figure 1.2). The enzymes that catalyse these modifications are known as histone methyltransferases and are either lysine or arginine specific. In contrast to acetylation (which broadly correlates with transcriptional activation), histone methylation is associated with different outcomes depending on the site and also the extent of methylation (one, two or three methyl groups attached to the lysine ϵ -amino group, while arginine can be either mono- or dimethylated) (Kouzarides, 2007). Although several characterised arginine methyltransferases and their identified role in transcriptional activation suggest that this modification is relatively abundant (Boisvert et al., 2005), it has been challenging to detect *in vivo* by mass-spectrometry

(Bannister and Kouzarides, 2005; Strahl et al., 2001). Recently, arginine methylation by the methyltransferase CARM1 has been implicated in the choice between embryonic and extraembryonic commitment of blastomeres in early mouse development (Torres-Padilla et al., 2007). Methylation of lysine residues is known to occur on histone H3 at lysines 4, 9, 27, 36 and 79 and on histone H4 at lysine 20 (Figure 1.2). Histone H3 lysine 4, 36 and 79 methylation has been correlated with transcriptional activity being mostly found in euchromatic regions, whereas methylation of H3K9 and H4K20 are found predominantly but not exclusively at constitutive heterochromatin (Martin and Zhang, 2005; Peters et al., 2003). H3K27, in turn, is enriched at facultative heterochromatin, including the inactive X-chromosome (Peters et al., 2003; Rougeulle et al., 2004). It is unlikely that histone methylation by itself has a pronounced effect on nucleosome structure, as the methyl group is neutrally charged. Instead, histone methyl marks create binding sites for a variety of regulatory proteins that induce changes in chromatin compaction and transcription (Daniel et al., 2005). Several proteins have been shown to specifically recognise and interact with methylated lysines (including chromo domain-, tudor domain- and WD40-repeat domain proteins) (Martin and Zhang, 2005). Those that associate with methylated H3K4 are involved in transcriptional activity (Pray-Grant et al., 2005; Wysocka et al., 2005), whereas those that bind methylated H3K9 or H3K27 have been implicated in transcriptional repression (Fischle et al., 2003; Thiagalingam et al., 2003).

Histone lysine methylation is catalysed by enzymes containing a conserved 130-aminoacid SET-domain (Jenuwein and Allis, 2001). The first identified histone methyltransferases were Suv39h1 and Suv39h2 in mammals and its fission yeast homolog Clr4 (Ivanova et al., 1998; Rea et al., 2000). These HMT specifically catalyse the trimethylation of H3K9 at constitutive heterochromatin. More than 70 other SET-domain proteins have been found since in mammals, including G9a that catalyses the dimethylation of H3K9 at euchromatic areas, the Polycomb group protein Ezh2 that catalyses the trimethylation of H3K27 and Trithorax/MLL that catalyses the trimethylation of H3K4 (Lachner and Jenuwein, 2002). H3K9 methylation creates a high-affinity binding site for the chromodomain of heterochromatin protein 1 (HP1) (Lachner et al., 2001) that recruits additional proteins involved in heterochromatin maintenance. These include Suv39h itself, suggesting a possible mechanism of heterochromatin spreading, whereby HP1 binding to H3K9m3 recruits Suv39h, which methylate adjacent histones (Hiragami and Festenstein, 2005). In addition, this modification was shown to create binding sites for the RITS (RNA induced transcriptional gene silencing) complex, a component of RNAi machinery. It is suggested that RITS and small double-stranded RNAs transcribed from repetitive DNA

elements are able to mediate the targeted recruitment of the fission yeast (*S. pombe*) HMT Clr4 to heterochromatic regions. Deletion of crucial components of the RNAi pathway results in the aberrant accumulation of complementary transcripts from centromeric heterochromatic repeats. This is accompanied by transcriptional de-repression of transgenes integrated at the centromere, loss of histone H3K9 methylation, and impairment of centromere function (Verdel and Moazed, 2005; Volpe et al., 2002). There is also evidence that RNAi is involved in the maintenance of constitutive heterochromatin in higher organisms, including *Drosophila* and humans (Kim et al., 2006; Pal-Bhadra et al., 2004).

Histone methylation was initially thought to be a relatively stable and irreversible modification, as initial studies presented evidence that the turnover rates of histone methyl groups were slower than those of histones themselves (Spotswood and Turner, 2002). However, several families of histone demethylases have been recently identified. The first histone demethylase to be identified was the lysine-specific demethylase 1 (LSD1), a protein previously found to be part of several HDAC complexes, being shown to catalyse mono- or dimethylated (but not trimethylated) H3K4 demethylation by an oxidative demethylation reaction (Shi et al., 2004). LSD1 has been implicated in transcriptional regulation and chromatin boundary formation (Lan et al., 2007b; Metzger et al., 2005). Unlike LSD1, a structurally unrelated family of enzymes has been found capable of removing both mono-, di- and trimethyl marks (Tsukada et al., 2006). These proteins share a JmjC motif that is conserved from *S. cerevisiae* to human and are thought to demethylate histones through an oxidative reaction that requires iron and α -ketoglutarate as co-factors. The methylation marks removed by these family of demethylases include H3K9, H3K27 and H3K36 (Klose and Zhang, 2007). Of particular interest is that the reported H3K27 demethylases JMJD3 and UTX were found to be recruited with the H3K4 methylation machinery (Trithorax complex), resulting in the concerted demethylation of H3K27 (repression mark) and methylation of H3K4 (activation mark) at homeotic (*Hox*) loci during developmental activation (Agger et al., 2007; Lan et al., 2007a). Conversely, the functional interplay between the complex containing the Polycomb methyltransferase Ezh2 (that methylates H3K27) and the H3K4 demethylase Rbp2 was recently shown in mouse ES cells (Pasini et al., 2008).

1.1.3.3. Histone variants

As well as specific modifications at histone tails, certain histone variants that share different degrees of homology to conventional histones are also implicated in

chromatin dynamics. In particular, regions of active transcription have been shown to incorporate the histone variant H3.3, which differs in only four amino acids from H3, but has different biological properties. While the “canonical” H3 is synthesised exclusively in dividing cells and integrated into chromatin during DNA replication, H3.3 is constitutively expressed and deposited into chromatin throughout the cell cycle (Ahmad and Henikoff, 2002a, b). H3.3-specific assembly complex lacks a p150 subunit that is present in the CAF-1 complex (involved in the replication-coupled nucleosome assembly), but instead contains a distinct chaperone termed histone regulator A (HirA) (Ray-Gallet et al., 2002; Tagami et al., 2004). It has been proposed that the replication-independent histone replacement can be involved in the propagation of active chromatin states. Using a model where transcription was induced in a heterochromatic multicopy transgenic array, it was shown that loss of heterochromatic markers upon transgene activation is tightly linked with H3.3 deposition on chromatin (Janicki et al., 2004). In this context, the “memory” of active transcription is retained at some *loci* when endoderm cell nuclei is transplanted into enucleated amphibian eggs. This epigenetic memory was shown to correlate with the association of histone H3.3 at *loci* that display memory but not in those where memory has been lost (Ng and Gurdon, 2008).

Several variants have also been described for histone H2A. For example, H2A.Z in yeast is a euchromatic variant associated with suppression of silencing, but it is localised to heterochromatin and interacts with HP1 in mammals (Fan et al., 2004). Mammalian H2A.Z was shown to be associated with Polycomb-mediated repression (Sarcinella et al., 2007), but was also found at actively transcribed genes (Bruce et al., 2005; Farris et al., 2005). H2A.Z incorporation requires specific nucleosome assembly machinery that in budding yeast includes a member of SWI2/SNF2 chromatin remodelling family, Swr1, that is thought to destabilise the nucleosome and facilitate the exchange of H2A–H2B for H2A.Z–H2B (Mizuguchi et al., 2004). This process was also shown to depend on the transcription-coupled complex FACT, supporting the idea that H2A–H2B exchange occurs during transcription (Belotserkovskaya et al., 2003).

1.1.4. Transcriptional silencing by Polycomb group proteins

During embryogenesis and organogenesis, subsets of genes are activated in a temporally and spatially restricted manner. Shaping an embryo depends critically on the precise temporal and spatial expression of *Hox* genes. Trithorax (TrxG) and

Polycomb (PcG) group proteins were originally identified in *Drosophila* as products of genes that when mutated would affect the spatially restricted expression pattern of *Hox* genes and developmental patterning (Lewis, 1978). Transcription of *Hox* genes is initially achieved by transcription factors and subsequently maintained by PcG or TrxG proteins (Francis and Kingston, 2001). These proteins are part of a widely conserved “cell memory” system that stabilises cell identity by restricting or maintaining transcription patterns set in the first stage of embryonic life, throughout development and in adulthood (Ringrose and Paro, 2004). It is now recognised that PcG proteins regulate many genes in addition to the *Hox* genes and these include cell cycle and developmental-regulated genes such as transcription factors and signalling proteins (Boyer et al., 2006; Schwartz et al., 2006). Consistent with this, several mammalian PcG and TrxG members have been implicated in cell cycle regulation and tumorigenesis (Sparmann and van Lohuizen, 2006). PcG proteins are involved in gene repression and silencing, whereas TrxG proteins counteract repressive PcG functions and maintain gene activity (Orlando and Paro, 1995; Pirrotta, 1998). Notably, both groups encode components of multiprotein complexes that regulate chromatin structure in opposite ways (Mahmoudi and Verrijzer, 2001).

TrxG protein complexes were found to methylate histone H3 lysine 4 (Milne et al., 2002), a modification characteristically associated with active chromatin (Ruthenburg et al., 2007). H3K4 trimethylation has been found at the 5' regions of virtually all actively transcribed genes from budding yeast to humans, whereas H3K4 monomethylation was shown to selectively localise to putative enhancer regions in human cells, but not in yeast (Heintzman et al., 2007). In *Drosophila*, TrxG (and PcG) bind to specialised DNA elements known as Polycomb/Trithorax response elements (PRE/TRE), which include many conserved short motifs, some of which are also recognised by other DNA-binding proteins. These sequences are necessary and sufficient for TrxG/PcG function and often reside within the promoters of target genes (Schwartz and Pirrotta, 2007). In mammals, however, PRE/TRE sequences were not identified suggesting that targeting of these complexes might be more complex. The recruitment of the mammalian TrxG protein MLL is facilitated by basal transcriptional machinery as well as sequence-specific transcription factors (Ruthenburg et al., 2007). The activity of another H3K4 methyltransferase, Ash1, was shown to be mediated by non-coding RNAs (Sanchez-Elsner et al., 2006). H3K4 methylation is known to recruit transcriptional activators and nucleosome remodelling factors, such as the HAT-containing complexes SAGA and the ISWI-containing complex NURF (Pray-Grant et al., 2005; Wysocka et al., 2006).

Methylation of histone H3 lysine 27 is associated with PcG proteins, a family of structurally unrelated proteins involved in the epigenetic maintenance of the transcriptionally silent state (Figure 1.4). PcG proteins were found to be involved in the regulation of other cellular processes including proliferation, X-inactivation and the maintenance of pluripotency (the latter will be further discussed in Section 1.2.2.1) (Heard, 2005; Martinez and Cavalli, 2006; Spivakov and Fisher, 2007). Several Polycomb repressor complexes (PRCs) have been identified. The mammalian PRC2 complex contains the histone methyltransferase Ezh2 and also two other proteins, Eed (embryonic ectoderm development) and Suz12 (suppressor of zeste-12) that are essential for the catalytic activity of the complex (Cao and Zhang, 2004). This complex catalyses the trimethylation of H3K27 (Figure 1.4), creating a binding site for another Polycomb repressor complex, PRC1, that comprises Ring1A/Ring1B, Bmi-1, Polycomb and Polyhomeotic proteins (Wang et al., 2004a). In addition, Ring1A/B in cooperation with Bmi-1 were recently shown to act as ubiquitin ligases catalysing the ubiquitylation of lysine 119 of H2A (Figure 1.4) (de Napoles et al., 2004), an event that was implicated in preventing active transcription of a number of developmental genes in ES cells (Jorgensen et al., 2006; Stock et al., 2007). This suggests that PRC1 and PRC2 collaborate in maintaining silent chromatin. It is possible that PcG proteins can repress transcription by other means. In *Drosophila*, PcG proteins were also shown to interfere with transcriptional initiation and nucleosome remodelling (Dellino et al., 2004; Wang et al., 2004b). PRC1 is thought to maintain *locus* silencing by preventing chromatin remodelling at sites already rendered inactive by conventional DNA-binding factors (Francis and Kingston, 2001; Shao et al., 1999). Binding of PcG proteins might also alter the topology of the DNA itself, leading to the formation of negative superhelical turns (Mohd-Sarip et al., 2006). In addition, there is also evidence that link PcG proteins with RNAi-mediated transcriptional repression: PcG proteins co-localise in human cells with a core component of the RNAi machinery, Argonaute 1 (Kim et al., 2006), and transcription of non-coding RNAs was found from a well-studied PRE-containing transgene, *Fab-7*. Deletion of components of the RNAi pathway such as PIWI, Dicer-2 and Argonaute severely disrupted Polycomb-induced long distance interactions between different *Fab-7* transgenic copies and between known *Hox* target *loci* (Grimaud et al., 2006).

Notably, PcG and TrxG proteins maintain the status of their target genes through mitosis in a manner that is stable but not static. Thus, the interaction of PcG and TrxG proteins is proposed to ensure not only transcriptional memory, but also the capacity to switch if a new transcriptional stimulus is received (Ringrose et al., 2003).

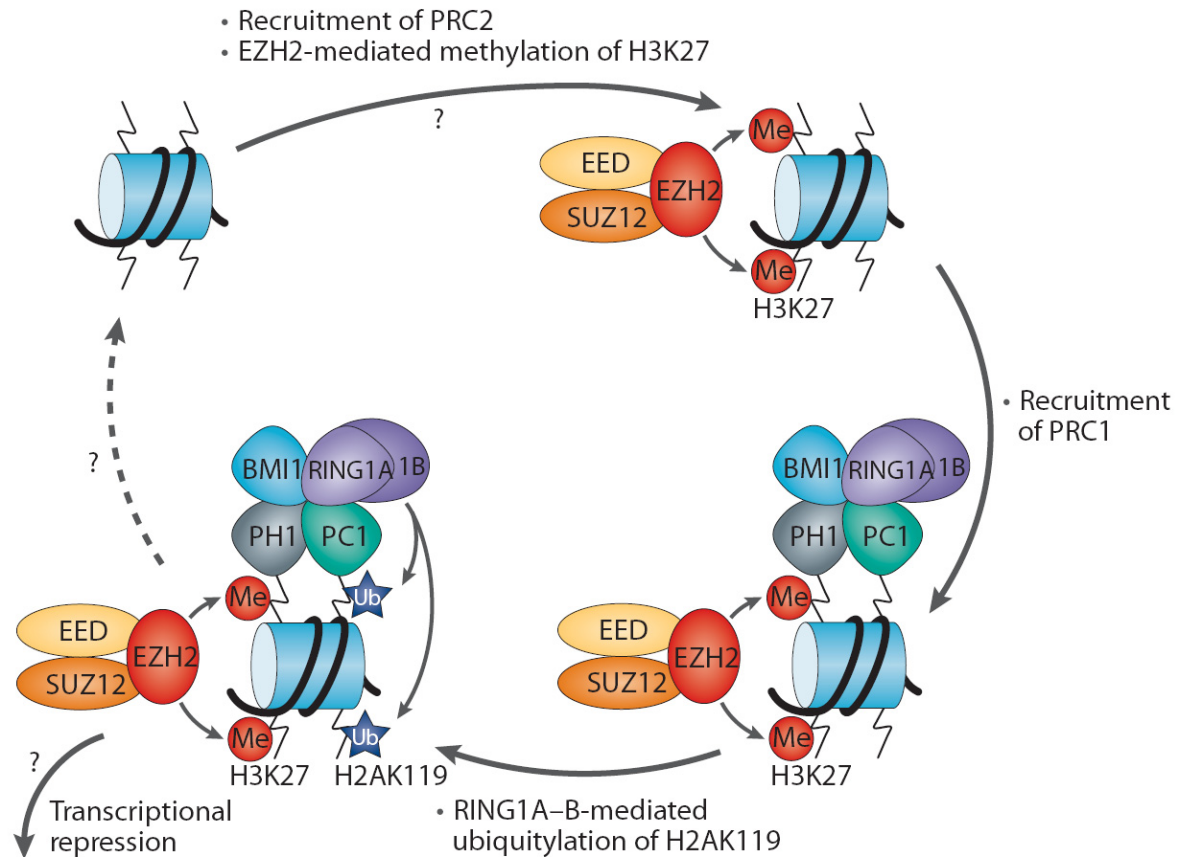


Figure 1.4. Polycomb repressive complexes. Ezh2 methyltransferase, a component of Polycomb Repressor Complex 2 (PRC2) that also contains the Eed and Suz12 proteins, methylates lysine 27 of histone H3 (H3K27me). This modification recruits the Polycomb repressive complex 1 (PRC1) complex containing Polycomb (PC1), Polyhomeotic (PH1) and Bmi1 proteins as well as RING1A and RING1B that function as ubiquitin (Ub) ligases for lysine 119 of histone H2A (H2AK119). It is thought that Polycomb complexes or the histone modifications that they induce can interfere with nucleosome dynamics and prevent RNA polymerase from active transcription. The H3K27me mark and PRC2 binding at promoters can be lost (dotted line) by poorly understood mechanisms (adapted from Spivakov and Fisher, 2007).

1.1.5. Spatial constraints – nuclear organisation

Beyond histone and DNA modifications there is evidence that suggests that active and inactive chromatin domains occupy different nuclear compartments in some cell types (Brown et al., 1997; Mahy et al., 2002), and that this organisation may be important for maintaining the heritable state of some genes (Brown et al., 1999; Kosak et al., 2002; Su et al., 2004).

A striking feature of nuclear architecture is the existence of distinct structural and functional compartments. Also, components of the machinery that is required for gene transcription or repression are known to have a non-homogeneous distribution in the nucleoplasm (Iborra et al., 1996; Jackson, 1995; Negre et al., 2006). At the level of the genome itself, it is known that individual chromosomes occupy distinct positions in the nucleus, referred to as chromosome territories (Figure 1.1) (Cremer et al., 2006).

As a result of different compaction levels, different chromosome segments adopt a complex organisation and topography within their chromosome territory (Croft et al., 1999; Edelmann et al., 2001). A polarised intranuclear distribution of gene-rich and gene-poor chromosomal segments has been shown to be an evolutionarily conserved principle of nuclear organisation (Cremer et al., 2003; Tanabe et al., 2002). Gene-rich regions tend to be oriented towards the nuclear interior, whereas gene-poor regions tend to be oriented towards the periphery (Foster and Bridger, 2005). Although chromosomes are organised as distinct territories in the interphase nucleus, chromosomes are dynamic structures and can be repositioned during development and cell differentiation, implying that nuclear organisation varies between cell types (Parada et al., 2004). One global example of such is the clustering of heterochromatin in distinct foci in the nucleus of mammalian cells, which can undergo morphological changes during differentiation. For example, when ES cells differentiate towards neural progenitor cells, the number of heterochromatic foci per nucleus decreased while the average cluster size increased, suggesting a global reorganisation of chromatin associated with differentiation (Williams et al., 2006). The same phenomenon is observed during terminal muscle differentiation. The spatial clustering of constitutive heterochromatin is accompanied by enhanced H3K9 and H4K20 tri-methylation across major satellite regions and increased non-coding transcription (Terranova et al., 2005). This clustering was shown to be dependent on histone deacetylation (Terranova et al., 2005) and on the methyl-CpG binding proteins MeCP2 and MBD2 (Brero et al., 2005) that modulate the association of HP1 with heterochromatin (Agarwal et al., 2007).

There is increasing evidence that gene repositioning to specific nuclear compartments is important for regulating expression (Gasser, 2002). Positioning of genes at the nuclear periphery is often correlated with silencing and repositioning towards the nuclear interior allows efficient transcription. Examples of this phenomenon include the *IgH locus* (immunoglobulin heavy chain) in committed B lymphocytes (Kosak et al., 2002), *c-maf* in T cells (Hewitt et al., 2004) and *Mash1* in neuronal cells (Williams et al., 2006) that are repositioned to the nuclear interior upon activation. In addition, the inactive X-chromosome is often observed at the nuclear periphery of female cells (Burke and Stewart, 2002). However, there is not always a strict correlation between movement away from the periphery and gene activation. The *interferon- γ locus*, for instance, is detected at the nuclear periphery irrespective of its transcriptional activity (Hewitt et al., 2004). Furthermore, analysis of a 2-Mb segment encompassing the *Mash1* gene showed that the entire region relocated to the nuclear interior in *Mash1*-expressing neuronal cells, even though some of the genes in this region are not expressed in this cell type (Williams et al., 2006). Findings in yeast have

further complicated our understanding of the relationship between transcriptional activation and gene positioning relative to the nuclear periphery. In *S. cerevisiae*, this region contains silencing compartments (Andrulis et al., 1998), but also activating ones. Indeed, the nuclear pores are important sites of transcription in this species and genome-wide analysis has shown that nuclear pore-associated proteins bind preferentially to active genes (Casolari et al., 2004). Transcriptional activation is frequently accompanied by an apparent movement of genes in the opposite direction to that observed in mammalian cells — that is, from the nuclear interior to the periphery (Brickner and Walter, 2004).

The presence of large amounts of heterochromatin is also thought to constitute another form of structural constraint that influences gene expression in mammalian cells. Several studies have reported a correlation between gene silencing and intranuclear positioning close to constitutive heterochromatin in mammalian cells. It has been suggested that this type of transcriptional repression could be explained by competition between heterochromatic proteins and transcriptional activators at the target *locus* (Brown et al., 1997), or by chromatin compaction that is mediated by heterochromatin-associated remodelling complexes (Grigoryev et al., 2006). Analysis of differentiating cells of the haematopoietic lineage have shown that gene activation is often associated with movement of the *locus* away from the heterochromatin compartment (Delaire et al., 2004; Hewitt et al., 2004) and, conversely, that gene silencing is correlated with repositioning of the *locus* close to heterochromatin (Brown et al., 1999), histone deacetylation and increased H3K9 methylation (Su et al., 2004). Chromosomal context seems to be an important determinant of association between specific *loci* and heterochromatin. This was demonstrated in a study of the human α -globin and β -globin genes (Brown et al., 2001). In lymphocytes, in which both *loci* are silenced, the β -globin but not the α -globin *locus* localises close to centromeric heterochromatin. A possible explanation for this observation is that the two *loci* reside in regions of the genome that differ in several aspects: the α -globin *locus* lies in an early-replicating subtelomeric region that contain genes that are highly and widely expressed, whereas the β -globin cluster is embedded within a late-replicating AT-rich region that contains genes that show tissue-specific expression patterns. It is plausible that the lack of association between the inactive α -globin *locus* and heterochromatin reflects the need for neighbouring genes to be expressed; these might exert a dominant effect on the positioning of the entire genomic region away from heterochromatin.

Genomic interactions between distal regulatory elements and target genes that are located on the same chromosome can also occur. The *β -globin locus* control region (LCR) and the active *β -globin* itself have been shown to physically interact in erythroid cells, causing the intervening 50 kb of DNA sequence to loop (Carter et al., 2002; Tolhuis et al., 2002). This suggests that bending of chromatin facilitates tissue-specific functional interactions between these sites and this spatial proximity might be required for high transcriptional activity. The formation of loops is not simply a consequence of transcriptional activity since its formation precedes gene activation in mouse erythroid progenitors that are committed to this lineage but do not yet express *β -globin* (Palstra et al., 2003). Conversely, chromatin loops can spatially separate genomic regions and contribute to gene silencing (Horiike et al., 2005). Physical interactions in *trans* between different chromosomes can also contribute to gene silencing. Of interest is the reported long-distance pairing of genomic regions that contain PREs (such as *Fab-7* and *Mcp*) in a tissue-specific manner in *D. melanogaster* nuclei. These interactions enhance transcriptional repression that is mediated by PcG proteins (Bantignies et al., 2003; Vazquez et al., 2006) and are not detected in Polycomb-*null* background. Because PcG proteins are known to be compartmentalised in the nucleus (Buchenau et al., 1998; Saurin et al., 1998), these data led to the suggestion that endogenous PcG target genes may undergo physical associations at nuclear PcG bodies dedicated to their regulation.

1.1.6. Temporal correlates – replication timing

In contrast to bacteria, which replicate their genome using a single origin, DNA replication in eukaryotes initiates bi-directionally at multiple chromosomal positions named origins of replication. Activation of all replication origins does not occur synchronously but progresses in a defined temporal order during S-phase of the cell cycle - some origins fire early and some fire late in S-phase (Camargo and Cervenka, 1982). This pattern is clonally inherited through subsequent cell divisions, but can change at genes undergoing developmental activation or silencing (Cimbora and Groudine, 2001; Perry et al., 2004). The temporal order of gene replication during S-phase has been proposed to be important for maintaining gene expression patterns (Gilbert, 2002). Early replication is generally associated with accessible chromatin carrying acetylated histones (Lin et al., 2003; Perry et al., 2004; Vogelauer et al., 2002), whereas late replication is a characteristic of constitutive heterochromatin and some facultative heterochromatin, formed through gene repression (Gilbert, 2002). In

addition to its association with the epigenetic status, replication timing correlates with the genomic context of the *loci*. Genome-wide analyses in *Drosophila* and human cells have shown that GC-rich, gene-dense regions tend to replicate early in S-phase, while AT-rich, gene-poor regions with high LINE (linear interspersed nuclear element) content tend to replicate later (Schubeler et al., 2002; Woodfine et al., 2004). Consistent with this, a study of 54 developmental genes in a model of neural differentiation *in vitro* demonstrated that only genes that map to AT-rich, LINE-poor regions changed replication timing upon transcriptional activation or repression during differentiation, while those mapping to GC-rich backgrounds remained early-replicating irrespective of transcriptional state (Hiratani et al., 2004).

The nature of the correlation between replication timing and genomic and epigenetic features remains unclear. It was hypothesised that histone acetylation renders replication origins more efficient or specifically orchestrates earlier firing leading to earlier completion of *locus* replication (Aparicio et al., 2004; Vogelauer et al., 2002). Histone acetylation is, however, almost completely restricted to early-replicating areas (Azuara et al., 2006; Jorgensen et al., 2007) and therefore other factors are likely to determine whether a *locus* replicates in the middle or end of the S-phase. The cause-and-effect relationship between replication timing and the chromatin status is also not fully established. Single-cell injection experiments in mammalian cells suggest that the timing of replication plays a role in the maintenance of heritable epigenetic states through mitosis. Reporter plasmids injected in early S-phase cells were more transcriptionally active and packaged into chromatin enriched in acetylated histones compared to when injected in late S-phase cells (Zhang et al., 2002). Consistent with this, replication forks are thought to include different subsets of chromatin modifiers in early and late S-phase (Chakalova et al., 2005; Gilbert, 2002). However, recent unpublished findings from our laboratory showing that only a fraction of early-replicating sequences in ES cells is enriched for active histone modifications argue against this as a generalised rule.

During other phases of the cell cycle, other temporal constraints might influence chromatin regulation, such as the length of G1-phase. It is during G1 that interactions between gene *loci* and heterochromatin become established. Therefore the length of G1 (which varies among cell types) might influence gene expression patterns. In *Drosophila* the occurrence of association of the *bw(D)* allele with a heterochromatic compartment was proportional to the time spent in G1/G0 of the cell cycle. According to this model, the longer the time that a *locus* spend in G1 exploring nuclear space, the bigger the probability that it will encounter an heterochromatic compartment and make

stable contacts that leads to stable silencing (Harmon and Sedat, 2005; Thakar and Csink, 2005).

In summary, maintenance of a stable epigenetic memory is a complex task that involves multiple layers of regulation, acting at all levels from DNA and histone modifications to the folding of chromosomal domains in the cell nucleus. The sum of all these epigenetic modifications is thought to confer identity to a cell dependent on their developmental history. Despite apparent stability over multiple cell cycles or long phases of quiescence, these epigenetic modifications are ultimately reversible and their functional significance can be addressed experimentally during somatic cell reprogramming, for example in nuclear transfer or heterokaryon assays.

1.2. Somatic cell reprogramming

Once committed, the differentiated state of a cell is normally stable and can be inherited through cell division. Under certain conditions cell fate can, however, be modified or reversed (DiBerardino, 1988; Simonsson and Gurdon, 2005; Surani, 2001). This exceptional change in cellular identity leading to the establishment of a novel gene expression program is here referred to as reprogramming. Such reprogramming can be demonstrated *in vivo* by *trans*-differentiation and regeneration or experimentally through nuclear transfer and cell fusion (Baron and Maniatis, 1986; Gurdon, 1962; Gurdon and Byrne, 2003; Hochedlinger and Jaenisch, 2002; Wilmut et al., 1997).

1.2.1. Conversion between differentiated cell types

Under exceptional circumstances cells can change their developmental program and give rise to cell-types that differ from their normal fate. Transdifferentiation, the conversion of cells from one differentiated cell type to another can occur particularly during the process of regeneration (Tanaka, 2003; Tsonis and Del Rio-Tsonis, 2004). One well known example is the regeneration of the lens of various species of urodele amphibia (newts and salamanders). Here, after removal of the lens of the eye, a new lens regenerates from the dorsal iris (Eguchi and Kodama, 1993; Tsonis et al., 2004). Iris cells are pigmented epithelial cells which upon regeneration give rise to modified keratinocytes containing high concentrations of crystallin proteins characteristic of the lens. Changes in cell fate between undifferentiated cells committed to form specific tissue types, to another cell lineage

were also observed *in vivo*. This phenomenon, known as transdetermination, has been extensively characterised in *Drosophila melanogaster*. At the end of *Drosophila* larval development, the cells required to produce the adult cuticular structures are found in determined structures called imaginal discs. If these discs are manually fragmented and kept in culture, cells at the wound site undergo proliferation and regenerate the missing part. In rare cases, however, during the regeneration of *Drosophila* imaginal discs, cellular identities can switch fate (McClure and Schubiger, 2007). For example, tissue from a leg disc may produce wing structures (Hadorn, 1968). At the molecular level it was shown that ectopic expression of the transcription factor gene *vestigial* can drive transdetermination of other discs to wing phenotype (Klebes et al., 2005; Maves and Schubiger, 1995). Recent studies indicate that the plasticity of disc cells during regeneration is associated with cell cycle changes (extended S-phase), and the reorganisation of chromatin structure (Sustar and Schubiger, 2005). Downregulation of PcG proteins by JNK signalling in some leg cells may reactivate *vestigial* expression leading to transdetermination (Lee et al., 2005). The fact that regeneration is a widespread phenomenon in metazoans (Brookes and Kumar, 2008) suggests that, despite being determined or differentiated, cells remain remarkably plastic.

In mammals however, conversion of cell fate is more controversial. Work done between the years of 1997 and 2002 in mice suggested that lineage committed progenitors may also undergo transdetermination *in vivo* when cells are placed in environments appropriate for a different developmental program. Differentiated somatic tissues could be populated by cells from bone marrow grafts (Raff, 2003; Tosh and Slack, 2002; Wagers and Weissman, 2004). For example, injection of haematopoietic progenitors into the cardiomyocyte pool could facilitate its regeneration following myocardial infarction (Orlic et al., 2001) as well as give rise to glial cells in the brain (Eglitis and Mezey, 1997; Mezey et al., 2000). Recent research however, has directly challenged the possibility of conversion of genetically unmodified cells *in vivo*, suggesting that earlier studies were simply inaccurate in defining the transdetermined phenotype. In some cases the donor cells were simply lodged in the host tissues but the cells had not actually changed cell fate. In addition, it was shown that the donor cells can spontaneously fuse *in vitro* (Terada et al., 2002; Ying et al., 2002) and *in vivo* (Palermo et al., 2005; Weimann et al., 2003) with host cells, providing an alternative explanation of cell plasticity. Given all these problems it is now generally accepted that haematopoietic progenitors remain committed to their fate even within microenvironments appropriate for other lineages (Balsam et al., 2004; Massengale et al., 2005; Murry et al., 2004).

In humans, it is not infrequent to find *foci* of a particular tissue in the wrong place (metaplasia), but it is difficult to rule out the possibility that the ectopic tissue has arisen by cell migration rather than by transdifferentiation (Slack and Tosh, 2001). Examples include the occurrence of bone in the soft connective tissue, or squamous patches in an epithelium that is normally glandular in histology (Slack, 1985; Slack and Tosh, 2001). Metaplasias usually arise in tissues subject to chronic trauma or infection, hence undergoing continuous regeneration. Some examples of metaplasia are now understood in molecular terms, for example patches of intestinal epithelium in the stomach can be produced by ectopic expression of the transcription factor *Cdx2* (Silberg et al., 2002). Some metaplasias have clinical significance because they predispose to development of cancer, for example squamous metaplasia of the bronchus or intestinal metaplasia of the stomach. However, all these *in vivo* examples of cell identity changes occur in the course of regeneration, which somehow may destabilise the differentiated or committed cell. Whether transdifferentiation and transdetermination represent a rare event in normal development or if it does occur at all in mammals remain controversial.

Experimentally, the constitutive expression of specific transcription factors by certain cell types can sometimes modify or override lineage outcome. For example, expression of individual muscle regulatory factors (MRF) of the *MyoD* family has been shown to convert a range of non-muscle cell types into muscle. However, ectopic expression of *MyoD* in ectoderm lineages *in vivo* (Faerman et al., 1993) or in hepatocytes *in vitro* (Schafer et al., 1990) does not initiate a complete conversion of these lineages into differentiated muscle. Similarly, *MyoD* expression in mouse ES cells and in transgenic embryos promotes activation of endogenous *MyoD* and expression of a few muscle-specific protein genes but does not activate a complete skeletal muscle differentiation program (Dekel et al., 1992). Herein, *MyoD* over-expression can convert ES cells to differentiated muscle only if differentiation is induced by embryoid body formation, indicating that *MyoD* myogenic regulatory functions require changes in cell context in order to initiate a complete differentiation program. Similarly, mis-expression of the erythroid factor *Gata1* in myeloid haematopoietic progenitors resulted in their conversion to erythroid lineages (Heyworth et al., 2002) and the transcription factors *C/EBP α* and *C/EBP β* can reprogram committed mature B-lymphocytes to become macrophages (Xie et al., 2004) or in a different cell context promote the conversion of pancreatic cells to hepatocytes (Shen et al., 2000). More recently, using a strategy of re-expressing key developmental regulators *in vivo*, Zhou et al identified a specific combination of three transcription

factors (*Ngn3*, *Pdx1* and *Mafa*) that reprograms differentiated pancreatic exocrine cells in adult mice into cells that closely resemble beta-cells (Zhou et al., 2008).

In most cases reported here the original and final tissue type are developmentally related in some way, for example being formed as adjacent regions from one sheet of cells in the embryo. Reprogramming between different germ layers has been more challenging, probably due to the developmental history of the cells. To directly analyse more extensive changes in cell fate, another approach has been used: fusion of two distinct somatic cells to form a single cell that contains different nuclei sharing the same cytoplasm.

1.2.1.1. Cell fate conversion in heterokaryons and hybrid cells

Somatic cell fusion, pioneered by H. Harris and N. Ringertz, has been used to illuminate many interesting cellular processes involving cytoplasmic-nuclear interactions (Harris, 1965; Ringertz and Savage, 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, or homokaryons if the parental cells belong to the same cell-type. If heterokaryons are maintained in culture for several days, hybrids will form, resulting from fusion between nuclei and subsequent cell division. For example, studies with hybrids provided early evidence that the malignant state is, in some cases, recessive to the differentiated state (Harris et al., 1969; Peehl and Stanbridge, 1982), a prediction now confirmed at the molecular level for many cancers. In addition, the combination of two cell types through cell fusion to form somatic cell hybrids yielded early evidence that mammalian gene expression can be altered by diffusible *trans*-acting factors. Following fusion with another cell type, the expression of differentiated functions frequently ceased (Mevel-Ninio and Weiss, 1981). For example, in hybrids between hepatocytes or lymphocytes and less specialised cells such as fibroblasts, the expression of cell-type-specific genes derived from the more differentiated parental cell are selectively turned off (Massa et al., 2000). The reason why more undifferentiated cell-types usually dominate over more differentiated fusion partners remains unknown. Somatic-cell hybrids often lose chromosomes in culture which led to the attribution of individual repressive effects to particular chromosomes (Ephrussi, 1972; Massa et al., 2000). Interestingly, the principle of phenotypic exclusion seen during normal development also applies to hybrid cells. When two differentiated cell types are fused, hybrid cells have the potential to express either one or the other phenotype, but not both (Fougere and Weiss, 1978). The molecular mechanisms that make gene expression programs

mutually exclusive remain elusive. In some cases, *trans*-activation of genes was also reported (Peterson and Weiss, 1972; Rankin and Darlington, 1979). However, it is difficult to draw conclusions about mechanisms of gene activation in stable hybrid cell lines because their genetic composition is variable and difficult to control (Ephrussi, 1972). Tetraploid hybrid cells, in particular formed between cells of different species lose chromosomes with cell division. It has been suggested that this happens due to non-synchronous control of DNA replication (Lin and Davidson, 1975). To circumvent this problem, Blau and colleagues performed experiments in heterokaryons which have clearly shown that fusion of one differentiated cell (a muscle cell) with a cell in which muscle genes are not normally expressed (a human amniocyte) leads to the activation of muscle genes in the amniocyte (Blau et al., 1983). Remarkably, cells derived from mesoderm, endoderm and ectoderm could also be dominantly reprogrammed in mouse myotubes to express muscle specific genes (Blau et al., 1985). Cells that were more closely related to muscle – mesodermal derivatives – consistently expressed muscle genes sooner and at a greater extent than the cells of ectodermal or endodermal origin (Blau et al., 1985). These studies have established that differentiated cells do retain flexible lineage potential and that lineage conversion and gene activation can occur in absence of DNA replication (Chiu and Blau, 1984). Gene activation in muscle heterokaryons is rapid (usually within two days), and is dependent on gene dosage, i.e. activation levels are dependent of a threshold concentration of positive regulators (Pavlati and Blau, 1986). These results suggest that factors capable of activating genes can either exchange freely between nuclei or exist in excess within the cytoplasm. Some experiments have shown this to be true for regulatory transcription factors such as the glucocorticoid receptor (Hache et al., 1999). The maintenance of specialised cellular phenotypes through dynamic interplay between positive and negative regulatory molecules could involve either direct interaction by complementing a particular deficiency in one of the cell types in a heterokaryon (Baron, 1993; Blau, 1992), or it could involve indirect effects. Such indirect effects might occur when a positive regulatory factor induces other cell type-specific transcription factors that in turn might activate a diverse group of downstream genes (Hardeman et al., 1986). This latter mechanism appears to operate when erythroid cells are fused with non-erythroid cells (Baron and Farrington, 1994; Baron and Maniatis, 1986).

Experiments on heterokaryons have been interpreted as evidence for continuous regulation of a plastic differentiated state (Blau and Baltimore, 1991), and providing an argument for the continuous need to impose (actively) or maintain the differentiated state (DiBerardino, 1988). Implicit in this model is the idea that all genes are continuously and actively regulated by *trans*-acting factor networks that either

activate or repress genes (Blau and Baltimore, 1991; Blau et al., 1985; Chiu and Blau, 1984). Accordingly, considerable remodelling of chromosomal structure occurs in heterokaryons. For example, the nuclei of chicken erythrocytes consist predominantly of heterochromatin containing the specialised linker histone H5. In heterokaryons formed by fusion of chicken erythrocytes with proliferating mammalian cells, the chicken erythrocyte nuclei once again become transcriptionally active. This process is accompanied by decondensation of chromatin, enlargement of the nucleus and the appearance of nucleoli. The enlargement of the chicken erythrocyte nucleus is due to a massive, but selective, uptake of mammalian nuclear proteins, including RNA polymerases. Conversely histone H5 is partially lost from the chicken erythrocyte nucleus (Ringertz et al., 1985). Histones H2A and H2B also exchange under these circumstances, but histones H3 and H4 do not. These results are predicted based upon the relative affinity of the histones for DNA and their organisation in the nucleosome (Pruss et al., 1995). The stability of DNA methylation states has also been recently explored in stable heterokaryons. Upon cell fusion of human keratinocytes with mouse myotubes, unknown factors alter regulatory regions of genes in keratinocyte nuclei, demethylating and activating a muscle-specific gene (*MyoD*) and methylating and silencing a keratinocyte-specific gene (*Keratin14*) (Zhang et al., 2007). Because these changes occur in the absence of chromosome replication or cell division, they should be mediated by an active mechanism (Zhang et al., 2007). DNA methylation has also been shown to restrict the reprogramming potential of some cell-types (Hela cells) in heterokaryons. When treated with the DNA demethylating agent 5-aza-cytidine prior to cell fusion, the previously non-inducible Hela cell nuclei were capable to activate early and late myogenic proteins (Chiu and Blau, 1985).

1.2.2. Reprogramming somatic cells to pluripotency

By “reversing” development, the reprogramming of somatic cells to embryonic state provides an exciting approach to address the functional significance of epigenetic mechanisms implicated in cellular differentiation and establishment of pluripotency. In addition, reprogramming somatic cells to become pluripotent is an important goal in cell replacement therapy since it affords the opportunity to generate and use patient-specific ES-derived cells as grafts. Using this strategy it would be possible to circumvent the problems of immune rejection that are likely to occur, unless the recipient and donor stem cells are very closely matched. Much of our knowledge of the biology of the uncommitted state came from the derivation of self-renewing cell lines

from human and animal stem cell populations. In particular, the study of the *in vitro* counterparts of undifferentiated cells in the embryo shed light on some of the mechanisms that control the pluripotent state.

1.2.2.1. Genetic and epigenetic regulators of pluripotency

Stem cells are characterised by the ability to self-renew and generate differentiated progeny and are often functionally classified according to their developmental potential. In mammals only the zygote and early blastomeres are totipotent and can generate the entire organism. Pluripotent stem cells – i.e., those that have the potential to generate all cell types of the adult organism excluding the extra-embryonic tissue (Rossant, 2007) – can be obtained from pre- and early post-implantation stages of embryonic development. In particular, ES cells are derived from the ICM of blastocysts (Evans and Kaufman, 1981; Thomson et al., 1998) (Figure 1.5a) that later gives rise to the three germ layers. In contrast, EC cells isolated from naturally occurring or experimentally induced teratocarcinomas have a more restricted lineage potential that varies between different EC cell lines (Andrews et al., 2005; Chambers and Smith, 2004). Later in development, pluripotency is retained only within cells of the primordial germ layer that can give rise to EG cells *in vitro* (Lovell-Badge, 2001). The main difference between EG and ES cells observed so far is that in EG cells, a number of genes have lost allele-specific parental imprints. This reflects the genome-wide reprogramming that occurs during primordial germ cell (PGC) development (Reik et al., 2001) with the purpose to modulate egg and sperm chromatin to a state that allows successful formation of a zygote, which then has the capacity to form an entire organism. PGCs begin to migrate at embryonic day 8.5 to the genital ridge where they arrive at day 11.5 dpc. EG cells are derived from PGC cells from the genital ridges at 11.5 to 13.5 embryonic days, which coincide with the period of genome-wide reprogramming (Morgan et al., 2005). Nonetheless, germline contribution has been demonstrated for at least some EG (Stewart et al., 1994; Tada et al., 1998) and EC (Stewart and Mintz, 1982) cells, thus demonstrating their pluripotency.

Pluripotent ES cells undergo symmetrical self-renewal and can be propagated clonally, producing two identical cells upon cell division. For these cells, self-renewal entails the suppression of differentiation (Ying et al., 2008). When introduced into a blastocyst, ES cells can contribute to formation of all tissues in the organism (including the germ line), whereas their transplantation to adult immunologically compromised

animals gives rise to teratomas – tumours containing a highly diverse array of cell types originated from all three embryonic layers.

Human and mouse embryonic stem cells and their regulation

ES cell pluripotency is thought to be maintained by several key transcription factors. These include the POU-family transcription factor Oct4, the SOX-family transcription factor Sox2 and a homeodomain DNA-binding protein called Nanog (Chambers and Smith, 2004). Embryos lacking *Oct4* do not form ICM, consisting exclusively of trophectoderm tissue (Nichols et al., 1998). *In vitro*, deletion of *Oct4* leads to the conversion of ES cells to trophectoderm (Niwa et al., 2000). Mouse ES cells deficient for *Nanog* are still able to self-renew, but show an increased propensity to differentiate (towards extra-embryonic endoderm) and cannot give rise to mature germ cells (Chambers et al., 2003b; Chambers et al., 2007). The role of Sox2 in early embryonic development is more difficult to study due to the presence of maternally-derived mRNA. It is however known that following implantation, as the maternal component is largely lost, *Sox2-null* embryos generate exclusively trophoblast and extraembryonic ectoderm (Avilion et al., 2003). Similarly to *Oct4*, *Sox2* deletion in ES cells also leads to their conversion in trophectoderm (Masui et al., 2007). Sox2 was shown to heterodimerise with Oct4, acting synergistically to activate Oct-Sox enhancers, and thereby regulating the expression of stem cell-associated genes (Chew et al., 2005; Okumura-Nakanishi et al., 2005; Tomioka et al., 2002). Recent data from H. Niwa's laboratory implicated Sox2 in the regulation of multiple transcription factors that affect *Oct4* expression (Masui et al., 2007). Forced expression of *Oct4* rescues the self-renewal and pluripotency defect of *Sox2-null* ES cells. These results indicate that the essential function of Sox2 is to stabilise ES cells in a pluripotent state by maintaining the requisite level of *Oct4* expression, rather than in the establishment of pluripotency (Masui et al., 2007). In addition to its expression in early embryonic cells, Sox2 is also expressed in neural stem cells and is essential for primitive ectoderm development (Miyagi et al., 2008).

To gain more insight into the transcriptional regulatory network operating in ES cells, different studies have mapped genome-wide *Oct4*, *Nanog* and *Sox2* binding sites in human and mouse ES cells (Boyer et al., 2005; Loh et al., 2006; Mathur et al., 2008). Despite the low overlap between human and mouse datasets, these authors found that *Oct4* and *Nanog* (and in human ES cells, also *Sox2*) co-occupy a large number of developmental genes, including both transcriptionally active and inactive genes in ES cells (Figure 1.5b). Among the active genes, targets included *Oct4*, *Nanog* and *Sox2*

themselves as well as other transcription factors, chromatin remodelling and modifying factors and components of signalling pathways known to be important for ES cell biology. In contrast, a large portion of the inactive targets identified encode developmentally important transcription factors, such as genes involved in the specification of ectoderm, endoderm, mesoderm and extra-embryonic lineages. These findings have been collectively taken to suggest that *Oct4*, *Sox2* and *Nanog* serve as so-called “master regulators” of the pluripotent stem cell identity (Spivakov and Fisher, 2007). However, the network that control pluripotency is more complex and many other players are now being involved (Dejosez et al., 2008; Kim et al., 2008a).

Although human and mouse ES cells seem to share intrinsic regulators, the extrinsic signals required for pluripotency have undergone significant species-specific modification (Figure 1.5a) (Pera and Trounson, 2004). The undifferentiated state of mouse ES cells can be maintained in culture by foetal calf serum in combination with LIF, a cytokine known to act mainly via the JAK-STAT3 signalling pathway (Niwa et al., 1998). The activation of the STAT3 pathway involves heterodimerisation of LIF receptor with the gp130 membrane protein that leads to the activation of JAK tyrosine kinase. Withdrawal of LIF triggers differentiation to a mixture of lineages (Nichols et al., 2001) and the provision of activated STAT3 renders mouse ES cells independent of LIF (Matsuda et al., 1999). In addition, signalling through bone morphogenic proteins (BMPs) maintain ES cell self-renewal. Usually these factors are supplemented by the serum, however, the addition of BMP and LIF to the media allows culturing undifferentiated ES cells in serum-free conditions (Ying et al., 2003). Recently however, it was shown that mouse ES cells can be efficiently derived and maintained in minimal conditions without extrinsic requirements for growth factors or cytokines (Ying et al., 2008). This work showed that ES cells can be maintained in their self-renewing state by blocking differentiation-inducing signalling using a cocktail of inhibitors (for FGF receptor tyrosine kinases, ERK and GSK3 pathway). In contrast to mouse ES cells, LIF signalling is not active in undifferentiated human ES (hES) cells (Daheron et al., 2004; Humphrey et al., 2004) and BMPs induce trophectoderm differentiation, either in the presence (Gerami-Naini et al., 2004; Pera and Trounson, 2004) or absence (Xu et al., 2002) of serum. Conversely, the MEK kinase pathway promotes differentiation in mouse ES cells (Nichols et al., 2001), but in hES cells, fibroblast growth factor 2 (FGF2) (Amit et al., 2000), activates this pathway through FGF receptors, and is required to maintain the undifferentiated state. Several studies suggest that FGF and TGF β signalling interactions are important for maintenance of hES cell pluripotency (Amit et al., 2004; Noggle et al., 2005; Wang et al., 2005; Xu et al., 2005). Thus, long-term maintenance of hES cells can be achieved with a

combination of Activin or Nodal (signalling branch TGF β /Activin/Nodal through Smad2/3 activation) plus FGF2 in the absence of feeder-cell layers, conditioned medium or serum replacer (Beattie et al., 2005; Vallier et al., 2005).

Several phenotypic differences between mouse and hES cells have been reported (Giadrossi et al., 2007; Pera et al., 2000). Mouse cells grow in rounded colonies, and have indistinct borders, whereas hES cell colonies are flatter and often display more distinct borders. The embryonic markers defined by the antibodies SSEA-1, -3 and -4 against cell surface glycolipids are expressed differently in mouse (SSEA1+) and hES (SSEA3/4+) cells. Also the cell-surface proteoglycan recognised by several monoclonal antibodies reactive with hES cells, including TRA-1-60 and TRA-1-81, are not detected in mouse ES cells (Thomson et al., 1998). The question of whether differences between mouse and hES cells are related to species differences or relate to differences in the stages of embryonic development when they were derived remains open (Lovell-Badge, 2007). Some data suggest that the same pattern of antigen expression is seen in the ICM of the human blastocyst and in hES cells. SSEA3, SSEA4 and TRA-1-60 were all found to be expressed in the ICM in a limited series of human pre-implantation blastocysts cultured *in vitro* (Dimos et al., 2008). As these markers are not expressed in the mouse ICM, this finding suggests that there are intrinsic species-specific differences in the expression of these markers in pluripotent cells. Moreover, a general argument in favour of the existence of species-specific differences in pluripotent stem cell phenotype is that pluripotent cells in mice and humans show differences irrespective of whether they originate directly from embryos or from primordial germ cells or through process of teratocarcinogenesis (Pera et al., 2000). Thus, human EC cells resemble hES cells, mouse EC stem cells resemble mouse ES cells, and the two cell-types show the same inter-species differences. However, the recent derivation of pluripotent stem cells from post-implantation mouse embryos, the so-called epiblast stem cells, challenged this view. These novel pluripotent cell lines are distinct from mouse ES cells in their epigenetic state as well as the signals driving their differentiation. Furthermore, they share some features with human ES cells such as patterns of gene expression and signalling responses, suggesting that the differences between human and mouse ES could be potentially explained by their temporal derivation in the embryo (Brons et al., 2007; Tesar et al., 2007).

It has been suggested that chromatin features of pluripotent ES cells reflect a generalised “open” and transcriptionally “permissive” state (Meshorer and Misteli, 2006). For example, Meshorer and colleagues have demonstrated that chromatin proteins such as H2B, H3 and HP1 are more dynamic in ES and EC cells than in differentiated cells (Meshorer et al., 2006). ES cell differentiation resulted in an increase in global levels of the heterochromatic mark H3K9me3 and a decrease of acetylated H3 and H4 (Lee et al., 2004; Meshorer et al., 2006). Also in *Drosophila melanogaster*, living cell measurements showed that, despite similar diffusion coefficients, chromatin mobility is more constrained in the nuclei of differentiated than in the nuclei of undifferentiated counterparts (Thakar and Csink, 2005). ES cells have also been shown to have a characteristic replication timing profile in which many non-expressed tissue-specific regulator genes replicate early in S-phase (Azuara et al., 2006; Hiratani et al., 2004). Consistent with this, several ATP-dependent remodelling factors are highly expressed in ES cells (Kurisaki et al., 2005) and mutations in components of the SWI/SNF chromatin-remodelling complex such as *Brg1*, results in developmental arrest at the 2-cell stage (Bultman et al., 2006).

A series of studies have shown that tissue-specific genes that were thought to be in a repressed state in pluripotent cells might in fact be poised for transcription in ES cells. Szutorisz and co-workers showed that in ES cells, the B-cell specific *λ5-VpreB1* locus is already marked by H3K4me2 and H3 acetylation (active chromatin modifications) prior to activation during B-cell commitment, and that these marks disappear from the locus upon ES cell differentiation into a non-lymphoid lineage (Szutorisz et al., 2005). In addition, PcG proteins were identified to have a crucial role in the regulation of ES cell identity. Chromatin immunoprecipitation studies in mouse and human ES cells revealed that the members of the PRC2 complex Suz12 and Eed bind to the promoters of many developmental genes that are not actively transcribed in ES cells, but are important for the regulation of lineage specification of extra-embryonic lineages or in later stages of development (Boyer et al., 2006; Jorgensen et al., 2006; Lee et al., 2006). Remarkably, Suz12 occupied the majority of non-expressed genes co-bound by transcription factors Oct4, Sox2 and Nanog, suggesting that these regulators may specifically recruit PcG proteins to their repressed targets (Figure 1.5b) (Lee et al., 2006). Supporting this role for PcG proteins, many developmental PcG target genes become derepressed upon genetic disruption of the PRC2 components *Eed* (Figure 1.5c) or *Suz12* (Azuara et al., 2006; Boyer et al., 2006; Lee et al., 2006) or the PRC1 component *Ring1B* (Jorgensen et al., 2006; Stock et al., 2007).

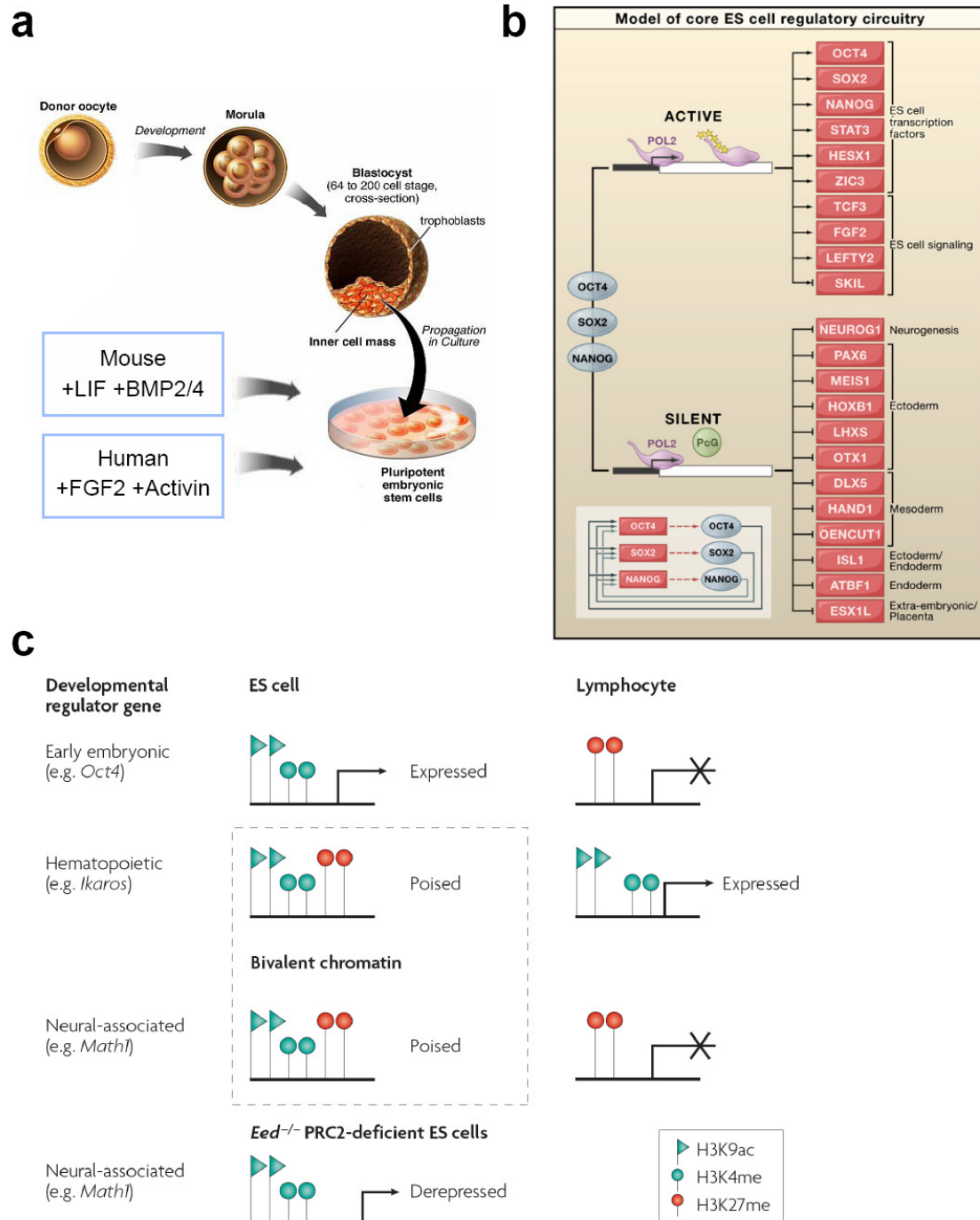


Figure 1.5. Regulation of pluripotency. (a) Pluripotent embryonic stem cells are derived from the inner cell mass of blastocysts. Human and mouse differential extrinsic requirements for the maintenance of the undifferentiated state in culture are indicated. (b) The Oct4, Sox2, and Nanog transcription factors (blue) occupy actively transcribed and silent genes. At this latter set of genes, RNA polymerase II (POL2) is present but not active due to the repressive action of PcG proteins. The interconnected autoregulatory loop, where Oct4, Nanog, and Sox2 bind together at each of their own promoters, is shown (bottom left). (c) The promoters encoding a range of non-transcribed developmental genes bear a combination of “conflicting” histone modifications that are normally associated with either active chromatin states (H3K9ac and H3K4me) or inactive chromatin states (H3K27me). During differentiation, “bivalent” chromatin profiles are generally resolved, leading to transcriptional activation of tissue-specific genes and silencing of *loci* associated with alternative developmental pathways. ES cells that are deficient for Polycomb repressor complex 2 (such as *Eed*^{-/-} cells) do not have the capacity for H3K27me and, consequently, many tissue-specific genes that are bivalent in wild-type cells are derepressed (adapted from Jaenisch and Young, 2008 & Spivakov and Fisher, 2007).

Notably, *Eed*-deficient cells are more prone to differentiate, suggesting a direct functional link between PRC2, gene silencing and the maintenance of the pluripotent state (Montgomery et al., 2005). Interestingly, the promoters of many inactive genes are marked in ES cells by both repressive and active modifications – a so-called “bivalent domain” (consisting of H3K27me3 associated with H3K4me3) (Azuara et al., 2006; Bernstein et al., 2006) (Figure 1.5c). It has been suggested that ‘bivalent’ chromatin may allow tissue-specific regulator *loci* to remain untranscribed (or non-productively transcribed) in pluripotent cells, but ‘primed’ for the rapid activation in response to developmental cues (Giadrossi et al., 2007; Spivakov and Fisher, 2007). Interestingly, this “priming” has been shown to correlate at the molecular level with the presence of poised RNA polymerases (not elongating) at bivalent gene promoters in ES cells (Stock et al., 2007). In ES-derived neural progenitors, as well as haematopoietic cells, bivalent chromatin profiles at many developmental genes were found to be resolved as lineage-appropriate genes became activated and depleted for H3K27me3, while lineage-inappropriate genes lost histone acetylation and H3K4m3 (Bernstein et al., 2006; Mikkelsen et al., 2007). For example, in myoblasts Suz12 is no longer found at the promoters of muscle-specific genes including the key regulator of muscle differentiation *MyoD* (Boyer et al., 2006). Polycomb protein binding was, however, still observed in differentiated cells, including neural progenitors, myoblasts and T cells, indicating that in addition to pluripotency PcG proteins are involved in gene repression at different stages of development (Barski et al., 2007; Lee et al., 2006; Mikkelsen et al., 2007).

The increasing knowledge of how the undifferentiated state is maintained in embryo-derived stem cell lines provided valuable clues for the understanding how re-orchestration of pluripotency occurs during reprogramming (Jaenisch and Young, 2008). Several experimental reprogramming strategies to convert somatic cells to pluripotency (both directly and indirectly) have been outlined including nuclear transfer, cell fusion and the forced expression of factors (Figure 1.6). Apart from the potential value for regenerative medicine, these approaches are instrumental for dissecting the re-establishment of pluripotency in a somatic cell at the cellular and molecular level (Hochedlinger and Jaenisch, 2006; Yamanaka, 2008).

1.2.2.2. Somatic cell nuclear transfer

Evidence of successful nuclear reprogramming of somatic cells was provided by the live birth of animals cloned by injection of differentiated somatic nuclei into eggs

(Figure 1.6). This demonstrates that terminally differentiated cells retain cell plasticity and can be reprogrammed to produce an adult cloned animal (Gurdon, 1962; Hochedlinger and Jaenisch, 2002; Wilmut et al., 1997). Successful nuclear transfer was first reported in 1952 by Briggs and King, who showed that nuclei from blastula stage embryos injected into enucleated *Rana pipiens* eggs resulted in normal hatched tadpoles (Briggs and King, 1952). In 1962 Gurdon successfully generated adult, fertile *Xenopus laevis* by nuclear transplantation of differentiated tadpole intestinal cells, but partial embryo cleavage of cloned embryos was far more frequent than complete cleavage (Gurdon, 1962). However, it has been found that the cells of the partial blastulae can be used as donors for a second, serial set of nuclear transfers to enucleated eggs. When this is done with partial blastulae derived from intestinal epithelium cells of *Xenopus*, it is found that many of the serial nuclear transplant embryos develop remarkably well (Gurdon et al., 2003). More than 30 years later, “Dolly” the sheep became widely recognised for being the first successfully cloned mammal (Wilmut et al., 1997). About a dozen other mammalian species have been cloned since (Gurdon and Byrne, 2003; Vajta and Gjerris, 2006). It remained a formal possibility, however, that donor nuclei which gave rise to the rarely observed clones were actually derived from tissue stem cells, which represent a small proportion of adult tissues. Herein, the successful generation of cloned mice from lymphocytes (Hochedlinger and Jaenisch, 2002; Inoue et al., 2005) unambiguously demonstrated (as lymphocytes undergo irreversible recombination of immunoglobulin or T-cell receptor) that terminal differentiation does not restrict the potential of the nucleus to support development. Cloning from terminally differentiated donor cells is, however, hugely inefficient and in many instances was only successful when a “two-step” procedure, that involved the generation of cloned ES cells was used as an intermediate (Hochedlinger and Jaenisch, 2003). These observations suggested that differentiation state may affect the efficiency of producing cloned animals. For example, blastomeres and ES cells show the highest success rates in direct cloning experiments (10-30%) in sharp contrast to more differentiated cells such as fibroblasts and cumulus cells (1-3%). Terminally differentiated cell-types such as lymphocytes and neurons were cloned using the “two-step” procedure previously described (Hochedlinger and Jaenisch, 2006). This inverse correlation between differentiation and reprogramming is consistent with the observation that a common problem of the cloning procedure is the faulty reactivation of early embryonic genes such as *Oct4* (Bortvin et al., 2003; Jaenisch et al., 2005) and conversely the incomplete silencing of tissue-specific genes (Ng and Gurdon, 2005, 2008). However, because the cloning process is affected by many other parameters, such as cell cycle and the physical characteristics of the donor nucleus, it

has remained unresolved whether cloning efficiency decreases with progressive cell differentiation in all cases. For example, the generation of cloned ES cells from neurons is less efficient than that from neural stem cells (Blueloch et al., 2006; Inoue et al., 2007). In contrast, it has been argued that nuclei from granulocytes are more efficient donors than nuclei from haematopoietic stem cells (Sung et al., 2006).

Nuclear cloning is an inefficient process resulting in the death of most clones soon after implantation or birth of clones with serious abnormalities (Hochedlinger and Jaenisch, 2003; Yang et al., 2007). Faulty reprogramming and epigenetic errors are most likely to be the cause of these developmental defects, since the offspring of cloned animals is not affected (Tamashiro et al., 2002). The dynamics of murine *Oct4* mRNA expression by mouse thymocyte nuclei after transfer to *Xenopus* eggs supports the requirement for complete epigenetic reprogramming (Simonsson and Gurdon, 2004). Chemical removal of thymocyte chromatin proteins substantially accelerated *Oct4* transcript accumulation from 96 hours to 40 hours. Consistent with the observation, injection of unmethylated plasmid DNA containing the genomic *Oct4* region resulted in detection of transcripts within few hours; the authors demonstrated that thymocyte derived *Oct4* transcription was dependent on active chromatin remodelling and DNA demethylation. Following transplantation into an oocyte, somatic nuclei undergo global changes, including nuclear swelling (Barry and Merriam, 1972; Byrne et al., 2003), chromatin re-distribution, induction of DNA synthesis and movement of proteins from the egg cytoplasm into the somatic nucleus (Gurdon, 2006; Merriam, 1969; Tamada et al., 2006). It therefore became important to determine whether faulty reprogramming would affect the therapeutic utility of patient-specific ES cells derived by nuclear transfer. Experiments in mice showed no molecular or biological differences when ES cells derived from fertilised embryos or by nuclear transfer were compared (Brambrink et al., 2006; Wakayama et al., 2007), indicating that nuclear transfer ES cells are as useful for therapeutic application as ES cells derived from fertilised embryos.

Regardless of the recent advances in primate embryo cloning (Byrne et al., 2007), the generation of cloned human embryos remains limited by the availability of oocytes and ethical issues. One approach to overcoming this limitation is to use oocytes from other species such as cow or rabbit. One report (Chen et al., 2003) has documented the establishment of ES cells from rabbit/human inter-species blastocysts. Although these results were not independently reproduced yet, those ES cells phenotypically resembled human ES cells except from carrying human and rabbit mitochondria. In summary, nuclear transfer holds promise for the creation of tailored stem cells for regenerative medicine however, alternative, more amenable approaches

are needed to uncover the molecular mechanisms underlying reprogramming for pluripotency.

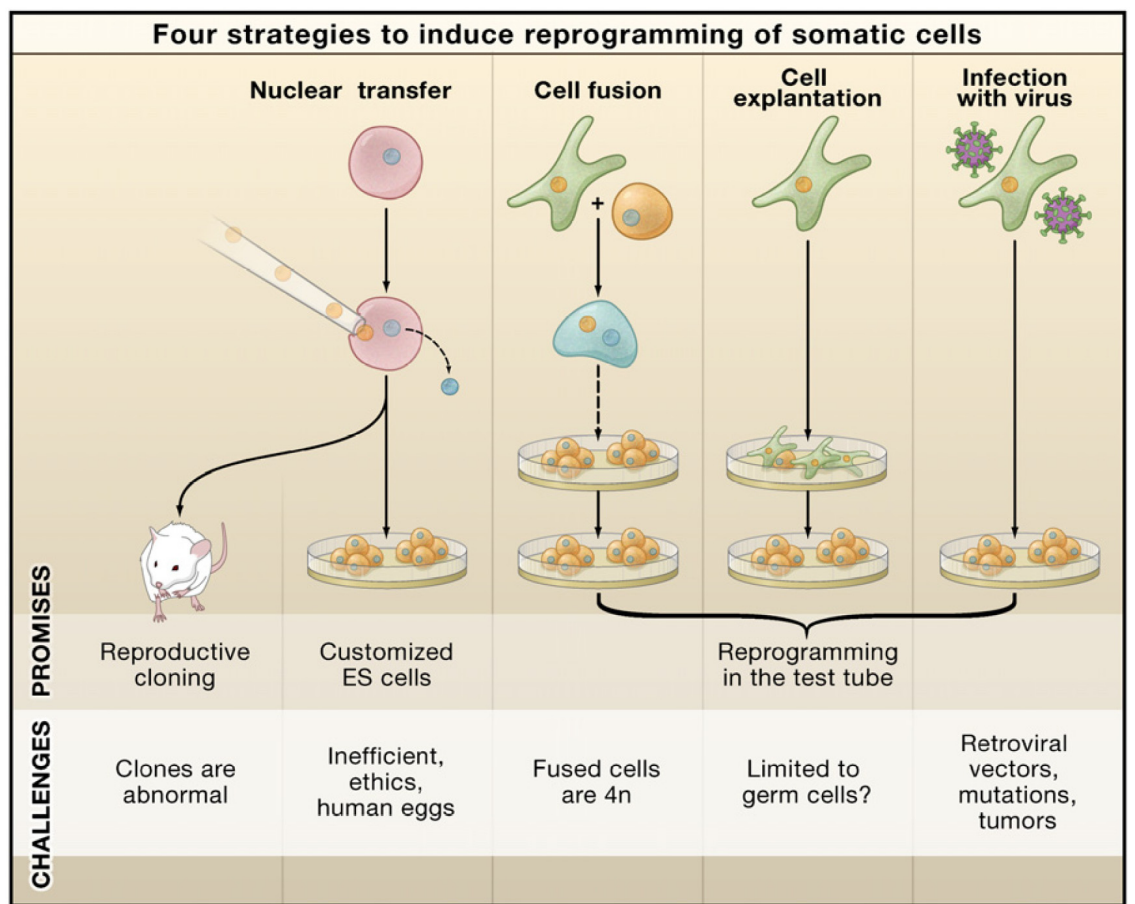


Figure 1.6. Four strategies to induce reprogramming of somatic cells to pluripotency. Nuclear transfer involves the injection of a somatic nucleus into an enucleated oocyte, which, upon transfer into a surrogate mother, can give rise to a clone (“Reproductive cloning”), or upon derivation and expansion in culture can give rise to genetically matched ES cells (“Customized ES cells”). Cell fusion of somatic cells with ES cells results in the generation of tetraploid hybrid cells that show all features of pluripotent ES cells. Explantation of somatic cells in culture selects for immortal cell lines that may be pluripotent or multipotent. At present, spermatogonial stem cells are the only source of pluripotent cells that can be derived from postnatal animals. Transduction of somatic cells with defined factors can initiate reprogramming to a pluripotent state. The promises and challenges associated with each of the strategies are indicated (adapted from Jaenisch and Young, 2008).

1.2.2.3. Cell fusion

Cell fusion has been used to demonstrate the dominant conversion of somatic cells to “early” pluripotent stages (Figure 1.6). Miller and Ruddle showed, in 1976, that the fusion of pluripotent teratocarcinoma cells with primary thymocytes resulted in the formation of pluripotent hybrids that shared all their features with the parental EC cells, including the potential to induce tumors (Miller and Ruddle, 1976). Importantly, the dominance of pluripotent cells over differentiated cells has also been shown in spontaneous (Terada et al., 2002; Ying et al., 2002) and experimental hybrids made

between somatic cells and murine EG (Tada et al., 1997) and ES (Tada et al., 2001) cells. Hybrid cells contribute to the three primary germ layers of chimeric embryos and transplantation into immunodeficient mice results in formation of teratomas. The ability to reprogram somatic cells seems to be conserved in human ES cells (Cowan et al., 2005; Yu et al., 2006) and analysis of mouse and human hybrid cell transcriptome revealed global patterns of gene expression reminiscent of ES cells (Ambrosi et al., 2007; Cowan et al., 2005). A central question raised by these experiments is whether the chromosomes of the somatic cell had been reprogrammed to pluripotency, or whether they were silenced and carried along with the pluripotent genome. To study reprogramming in heterokaryons, hybrid (tetraploid) cells and in nuclear transfer experiments it is important to be able to distinguish and verify events that occur in the donor nucleus from those of the recipient. This can be achieved using at least 3 different approaches (a) using a transgene marker (b) intra-specific polymorphic differences and (c) using differences between species. Activation of silent pluripotency-associated transgenes such as *Oct4-GFP* (Tada et al., 2001) and *Rex1-GFP* (Cowan et al., 2005), or reactivation of the inactive somatic X-chromosome (Kimura et al., 2002) provided some molecular evidence for reprogramming of the somatic genome in hybrid cells. In addition, endogenous genes such as *Cripto* (Cowan et al., 2005), *Oct4* and *Sox2* (Flasza et al., 2003; Han et al., 2008) were also shown to be initiated in the somatic cell nucleus upon cell fusion. In another paper from the T. Tada laboratory, the authors studied the chromatin events that accompany the reprogramming process by fusing ES cells (*M. musculus domesticus*) with thymocytes (*M. musculus molossinus*) (Kimura et al., 2004). By monitoring changes in chromatin structure of thymocyte nuclei upon ES hybridisation the authors observed global and *locus*-specific (at the *Oct4* promoter region) hyperacetylation of histones H3 and H4 and di- and tri- methylation of H3K4, consistent with the chromatin configuration in ES cells. The stability of DNA methylation states has also been explored in embryonic hybrids. Interestingly, there are striking changes in methylation in the somatic nucleus, in which several imprinted, non-imprinted genes (such as *Oct4*) and repetitive DNA were shown to be demethylated (Cowan et al., 2005; Tada et al., 1997; Tada et al., 2001). The somatic DNA methylation pattern at the imprinted *H19/Igf2r locus* is maintained in somatic-ES hybrids, unlike hybrids formed between somatic and EG cells (in which the differential methylation is erased). Thus, ES cells have the capacity to selectively reset certain epigenetic modifications of somatic cells.

Little is known about the molecular mechanisms underlying reprogramming by fusion with ES cells. The factors implicated may reside in the nucleus or in the cytoplasm. This question was initially addressed by separating the nuclear

compartment (karyoplast) from the cytoplasmic compartment (cytoplast) of an ES cell; these elements were then individually fused with neuronal cells isolated from neurospheres (Do and Scholer, 2004). In hybrids produced with the ES-cell karyoplasts, reactivation of an *Oct4*–*GFP* transgene was detected. In contrast, the GFP signal was not detected when neurosphere cells were fused with ES-cell cytoplasts, suggesting that nuclear factors are essential for molecular reprogramming. This conclusion is consistent with cloning experiments in amphibians (Byrne et al., 2003) and mice (Wakayama et al., 1998), which indicate that successful reprogramming depends on direct injection of nuclei into the germinal vesicle or into a metaphase oocyte, where nuclear factors are available in the cytoplasm. However, a recent report suggested that the cytoplasm could be enough for reprogramming (Strelchenko et al., 2006). Smith and colleagues observed a marked increase in reprogrammed cell colonies when they fused neural stem cells with ES cells that overexpress the transcription factor *Nanog* (Silva et al., 2006). The same phenomenon was observed for the transcription factor *Sal14* (Wong et al., 2008), previously identified as a pluripotency-associated factor that regulates *Oct4* in ES cells (Zhang et al., 2006). Also chromatin-associated factors seem to affect reprogramming efficiency since knocking-down G9a, a H3K9 HMT, or over-expressing the H3K9 demethylase Jhdm2a in somatic cells was shown to facilitate reprogramming (Ma et al., 2008). Other questions that have emerged are whether DNA replication and cell division are critical for reprogramming. The time required for *Oct4* activation upon cell fusion (approximately 24 hours) can be correlated with the timing of cell cycle completion in ES cells (approximately 22 hours), suggesting that DNA replication may be essential for reprogramming (Han et al., 2008; Wong et al., 2008). However, nuclear transfer experiments indicated the presence of replication-independent mechanisms. Whether DNA replication, genome inter-mixing and cell division play an important role in reprogramming is unknown.

In addition to providing an *in vitro* system to test the requirements of reprogramming, cell fusion is a potentially attractive approach to generate tailored cells for therapy without relying on nuclear transfer (Cowan et al., 2005). However, for this approach to be viable, the ES-cell nucleus needs to be removed from the heterokaryon in order to generate diploid customised cells for transplantation therapy (Pralong et al., 2005). Although selective elimination of some ES cell-derived chromosomes after hybrid formation is possible (Matsumura et al., 2007), it may be difficult to generate diploid reprogrammed cells in this way due to the risk of generating large-scale genomic instability. If DNA replication and cell division are required for complete reprogramming the applications of this method are limited. In addition, there has been

no convincing functional evidence showing that the somatic donor nucleus has regained the potential to sustain pluripotency in the absence of the ES-cell genome.

1.2.2.4. Reprogramming by cell extracts, culturing and overexpression of defined factors

Using cell-free systems to study reprogramming is an alternative to nuclear transfer or cell fusion. In addition, extracts might allow the purification of protein complexes involved in reprogramming. For example, Allan Wolffe's laboratory subjected permeabilised *Xenopus* tadpole-derived epithelial cells to S-phase *Xenopus* egg extracts (Kikyo et al., 2000). Within two hours donor nuclei lost over 85% of radioactively labelled proteins, concomitant with morphologically visible de-condensation of chromatin and uptake of proteins from the egg cytoplasm. Chromatin de-condensation was shown to be specifically dependent on the ATP-dependent nucleosome remodelling enzyme ISWI, which reversed TATA-binding protein mediated attachment of chromatin domains to the nuclear matrix. In a different set of experiments, human cells were exposed to *Xenopus* cell extracts and *Oct4* transcription was initiated following treatment (Hansis et al., 2004). However, no stable reprogramming was seen in reversibly permeabilised somatic cells that were subsequently passaged in culture, suggesting that a continuous supply of factors might be required for functional reprogramming to a pluripotent state. *Xenopus* egg extracts have proved useful in solving an old enigma in frog cloning: why cloning efficiency was markedly increased when somatic nuclei underwent serial transplantations (Gurdon and Byrne, 2003). Lemaitre and colleagues showed that efficient chromatin remodelling of differentiated nuclei depends on their exposure to mitotic egg extract, thereby facilitating embryonic DNA replication. Indeed pre-incubation of erythrocyte nuclei with *Xenopus* mitotic extract caused DNA replication, origin periodicity to decrease from about 120 Kb to 25 Kb (in early embryos these intervals are around 23 Kb). The authors also provided evidence that the enzyme topoisomerase II is required for chromatin re-organisation and origin recognition complex recruitment (Lemaitre et al., 2005). Mammalian cell extracts were also shown to induce reprogramming. Permeabilised human 293T cells were exposed for 1 hour to extracts from human EC cells, and *Oct4* expression and chromatin remodelling was detected (Freberg et al., 2007; Taranger et al., 2005). However, since human EC cells express *Oct4*, these data cannot exclude the possibility that gene products were detected from the cells that were used to prepare the extracts. To circumvent this problem human 293T cells were permeabilised and incubated in extracts of mouse ES cells. The specific detection of

human pluripotency-associated transcripts occurred 1-8 hours after extract exposure (Bru et al., 2008).

Donor cells from the germ cell lineage such as PGCs or spermatogonial stem cells are known to be unipotent *in vivo*, but it has been shown that pluripotent ES-like cells, can be isolated after prolonged *in vitro* culture from post-natal spermatogonial stem cells (Kanatsu-Shinohara et al., 2004). Spermatogonial stem cell-derived cells are pluripotent, as they were able to form postnatal chimeras and contribute to the germline. More recently, multipotent adult spermatogonial stem cells were derived from testicular spermatogonial stem cells of adult mice (Guan et al., 2006), and these cells were able to differentiate *in vitro* and to generate teratomas *in vivo*, but unable to form chimeras. A concern for any therapeutic application of cells derived from the germ line is their impact on genomic imprinting. It remains an open question whether somatic stem cells derived from adults can be converted to pluripotency by expansion in culture.

Takahashi and Yamanaka recently achieved a significant breakthrough in directly reprogramming somatic cells back to an ES-like state (Takahashi and Yamanaka, 2006). They successfully reprogrammed mouse embryonic and adult fibroblasts to pluripotent ES-like cells after viral-mediated transduction (Figure 1.6) of the four transcription factors *Oct4*, *Sox2*, *c-Myc*, and *Klf4*. These were coined “iPS” (for induced pluripotent stem) cells and were shown to be pluripotent by their ability to form teratomas, although they were unable to generate live chimeras. The original isolation of iPS cells was based on drug-dependent selection for *Fbx15*, a gene targeted by *Oct4* and *Sox2*, and expressed in pluripotent ES cells. When activation of *Oct4* or *Nanog* genes was used as a more stringent selection criterion for pluripotency, the resulting *Oct4*-iPS or *Nanog*-iPS cells, in contrast to *Fbx15*-iPS cells, were fully reprogrammed to pluripotency. Importantly, *Oct4*- and *Nanog*-iPS cells generated postnatal chimeras, contributed to the germ line (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007), and generated late-gestation embryos through tetraploid complementation (Wernig et al., 2007), the most stringent test for developmental potency. In addition, iPS cells were shown to be transcriptionally and epigenetically indistinguishable from ES cells. iPS cells showed reactivation of somatically silenced X-chromosome, DNA demethylation of pluripotency-associated genes and acquisition of chromatin structure that closely resemble ES cells (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). The kinetics of events of reprogramming to iPS was recently addressed. Pluripotency markers such as alkaline phosphatase (3 days), SSEA1 (9 days), and *Oct4* or *Nanog* (20 days) appear sequentially after viral transduction (Brambrink et al., 2008; Stadtfeld et al., 2008a; Wernig et al., 2007). Also,

evidence obtained by using inducible lentivirus-based vectors indicated that the four factors must be expressed for more than 12 days in order to generate iPS cells (Brambrink et al., 2008). Expression of the reprogramming factors in fibroblasts appears to initiate a sequence of events that eventually leads to the conversion of a small percentage of cells. In addition, similarly to what observed in nuclear transfer experiments, iPS reprogramming seems to be influenced by donor cell-type. The same cocktail of 4 transcription factors has been shown to drive mouse hepatocytes, stomach cells (Aoi et al., 2008), pancreatic beta-cells (Stadtfield et al., 2008a) and immature B-lymphocytes to iPS, whereas adult mouse B-lymphocytes (Hanna et al., 2008) required the additional expression of *C/EBP α* or silencing of the B-cell specific factor *Pax5*. Conversely, just 2 factors (*Oct4* and *Klf4*) were sufficient to reprogram adult neural stem cells, which endogenously express *Sox2* and *c-Myc* (Di Stefano et al., 2008; Eminli et al., 2008; Kim et al., 2008b).

Reprogramming of human fibroblasts has also been achieved in a parallel approach using *Oct4*, *Sox2* and either *Nanog* plus *Lin28* (Yu et al., 2007) or *Klf4* plus *c-Myc* (Lowry et al., 2008; Park et al., 2008c; Takahashi et al., 2007). Human iPS cells have been obtained by exposing fibroblasts only to *Oct4*, *Sox2*, and *Lin28* (Yu et al., 2007), although less efficiently. To date, human cell types converted to iPS include foetal fibroblasts, adult dermal fibroblasts, bone-marrow-derived mesenchymal cells and keratinocytes. In order to reprogram keratinocytes a combination of 5 factors (*Oct4*, *Sox2*, *Klf4*, *c-Myc* and *Nanog*) was required (Maherali et al., 2008). Thus, it remains unknown whether the same factors are sufficient to reprogram more specialised cell types such as mature human lymphocytes. Interestingly, generation of human iPS cells was dependent on FGF2 (Park et al., 2008b) rather than LIF (required to generate mouse iPS). Derivation of human iPS cells in mouse ES cell media with LIF was fruitless (Takahashi et al., 2007).

Isolation of iPS cells has been based on retrovirus-mediated transduction of oncogenes and drug-dependent selection for reprogrammed cells. These experimental requirements hinder the eventual medical application of this *in vitro* reprogramming approach. Mice derived from iPS cells frequently developed cancer due to transgene re-activation (Okita et al., 2007) and the isolation of human iPS cells cannot be based on genetically modified donor cells. In an effort to reduce the risk of tumors in iPS cell-derived chimeras, more recent experiments showed that *c-Myc* is dispensable for reprogramming (Nakagawa et al., 2008; Wernig et al., 2008; Yu et al., 2007), though the reprogramming process was significantly delayed and less efficient in the absence of this oncogene. While mice derived from these iPS cells will not develop *c-Myc*-induced tumors (Nakagawa et al., 2008; Wernig et al., 2008), it is not clear whether the

transduction of other retrovirus-transduced transcription factors such as *Oct4* will also cause tumors at later stages (Hochedlinger et al., 2005). In any case, new technologies must be developed to avoid the use of retrovirus. To address the second problem, iPS cells were isolated from genetically unmodified mouse fibroblasts based only on morphological criteria (Blelloch et al., 2006; Meissner et al., 2007).

These pioneering studies have illustrated the importance of several factors for iPS, but also suggested by their low efficiency that additional ones may be needed for efficient conversion to pluripotency. Less than 1% of the cells that have incorporated the four retroviruses can become iPS cells. The identification of the missing factor(s) may enable more efficient and retrovirus-free generation of iPS cells. The big challenge is to adapt what is known for the direct conversion of adult somatic cells to pluripotency to allow the production of genetically matched cells for the study and treatment of human disease (Nishikawa et al., 2008). Recently, a patient with amyotrophic lateral sclerosis was used to successfully generate iPS cells that possess properties of ES cells which could be directed to differentiate into motor neurons, the cell-type destroyed in this disease (Dimos et al., 2008). In addition, a variety of other disease-specific iPS cells were recently generated for research (Park et al., 2008a). Proof-of-principle studies have also shown that disease-specific ES cells produced by somatic cell nuclear transfer or iPS, in combination with gene correction, could be used to cure immunological disorders or sickle cell anemia in mice (Hanna et al., 2007; Rideout et al., 2002). Nonetheless, major drawbacks of reprogramming strategies lay in the low efficiency of the process, the use of virus and ethical concerns, which limit medical applications. It is likely that complex transcriptional and epigenetic events are required for efficient conversion to pluripotency. Therefore one of the most pressing objectives is to increase our knowledge about the reprogramming process, i.e. systematically define the constellation of “reprogramming factors” and interactions that convert the epigenome of a somatic cell into that of an embryonic stem cell.

1.3. Aims of this study

Conversion of one cell-type to another involves the establishment of the novel program of gene expression and the coordinated silencing of lineage inappropriate genes. The molecular mechanisms that control this processes and prevent the expression of lineage inappropriate gene programs (phenotypic exclusion) remain poorly understood. The overall aim of this study is to dissect the epigenetic events and factors required for lineage conversion. For this I will use experimental inter-species heterokaryons as a tractable approach to monitor the progression of reprogramming of human somatic cells.

Initially, I will focus on the characterisation of the epigenetic events of reprogramming of human somatic cells to muscle identity. Experimental cell fusion will be used to reprogram human B-lymphocytes through the formation of stable heterokaryons with mouse myotubes. With this approach I will address large scale events of reprogramming such as changes in nuclear organisation and the importance of chromatin mechanisms for lineage inappropriate gene extinction. Further, I will expand the heterokaryon approach for the reprogramming of somatic cells to pluripotency. Mouse ES cells will be fused to human B-lymphocytes and the requirements/stability of the reprogramming process will be addressed. This approach will also be used to uncover the individual role of factors previously implicated in reprogramming and finally, address the contribution of Polycomb repression in ES cells for somatic cell reprogramming.

This study might contribute to both an increased understanding of the molecular basis of cell identity conversion as well as helping to develop potential applications of nuclear reprogramming in cell replacement therapy.

CHAPTER 2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Antibodies

hLamin A/C	Mouse monoclonal anti-human Lamin A/C (VP-L550; Vector Laboratories Inc., Burlingame, CA). Does not cross-react with mouse Lamin A/C. Used for immunofluorescence (IF) at 1:100 dilution.
GFP	Rabbit polyclonal anti-green fluorescent protein (GFP) (A11122; Molecular Probes, Eugene, OR). Used for IF at 1:200 dilution.
SSEA4	Mouse monoclonal anti-human SSEA4 (MC-813-70; Developmental Hybridoma Studies Bank, Iowa City, IA). Used for IF and fluorescence activated cell sorting (FACS) at 1:3 dilution.
TRA-1-60	Mouse monoclonal anti-human TRA-1-60 (MAB4360; Chemicon/Milipore, Temecula, CA). Used for IF at 1:12 dilution.
TRA-1-81	Mouse monoclonal anti-human TRA-1-81 (MAB4381; Chemicon). Used for IF at 1:20 dilution.
hNestin	Rabbit polyclonal anti-human Nestin (ab28944; Abcam Ltd., Cambridge, UK). Does not cross-react with mouse Nestin protein. Used for IF at 1:100 dilution.
hNanog	Rabbit polyclonal anti-human Nanog (ab21624; Abcam). Does not cross-react with mouse Nanog protein. Used for IF at 1:100 dilution.
mOct4	Goat polyclonal anti-mouse Oct4 (sc-8628; Santa Cruz Biotechnology Inc., Santa Cruz, CA). Used for western blot at 1:2000 dilution. Mouse monoclonal anti-Oct4 (611202; BD Biosciences, San Jose, CA). Used for IF at 1:100.
mSox2	Rabbit polyclonal anti-mouse Sox2 (ab15830; Abcam). Used for western blot at 1:1500 dilution.
mLamin B	Goat polyclonal anti-mouse Lamin B antibody (sc-6216; Santa Cruz Biotechnology).
CREST	Human CREST (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia) autoimmune serum. A Gift from G. Brown, University of Birmingham Medical School, UK. Used for IF at 1:200 dilution.
mMyogenin	Mouse monoclonal anti-mouse Myogenin (M3559; DakoCytomation, Denmark). Used for IF at 1:100.
5.1H11	Mouse anti-human NCAM (neural cell adhesion molecule) (5.1H11 supernatant, Developmental Studies Hybridoma Bank). Does not cross-react with mouse NCAM protein. Used for IF at 1:2 dilution.
mEed	Mouse monoclonal anti-mouse Eed. Antibody kindly provided by A. Otte (University of Amsterdam, The Netherlands). Used for western blot at 1:200 dilution.
mSuz12	Rabbit polyclonal anti-mouse Suz12 (CS-029-050; Diagenode SA, Liège, Belgium). Used for western blot at 1:1000 dilution.

mEzh2	Rabbit polyclonal anti-mouse Ezh2 (CS-039-050; Diagenode). Used for western blot at 1:1000 dilution.
FLAG	Mouse monoclonal anti-FLAG epitope (F3165; Sigma-Aldrich Company Ltd., Gillingham, UK). Used for western blot at 1:2000 dilution.
H3	Rabbit polyclonal anti-histone 3 carboxy terminal (ab1791; Abcam).
H3K9Ac	Rabbit polyclonal anti-histone 3 lysine 9 acetylation (07-352; Upstate Biotechnology /Milipore, Lake Placid, NY).
H3K9me3	Rabbit polyclonal anti-histone 3 lysine 9 trimethylation (07-442; Upstate).
H3K27me3	Rabbit polyclonal anti-histone 3 lysine 27 trimethylation (07-449; Upstate).
H3K27m1	Rabbit polyclonal anti-histone 3 lysine 27 monomethylation (07-448; Upstate).
H3K27m2	Rabbit polyclonal anti-histone 3 lysine 27 dimethylation (07-452; Upstate).
H4K20m3	Rabbit polyclonal anti-histone 4 lysine 20 trimethylation (07-463; Upstate).
H3K4m3	Rabbit polyclonal anti-histone 3 lysine 4 trimethylation (ab8580; Abcam).
H4Ac	Rabbit polyclonal anti-histone 4 acetylation (06-866; Upstate).
BrdU	Mouse monoclonal anti-BrdU antibody (347580; BD Biosciences).
Methyl-C	Mouse monoclonal anti-MethylCytidine (33D3; Eurogentec, Belgium).

2.1.2. Cell lines

C2C12	C2C12 is a subclone from a myoblast cell line (C2) established from normal adult mouse leg muscle; genetic background C3H (Yaffe and Saxel, 1977).
hB	Human Epstein-Barr Virus transformed adult B cell clones obtained from E. Eden (MRC Clinical Sciences Centre, London, UK) (Eden et al., 2002).
hB Puro^r	Human Epstein-Barr Virus transformed adult B cells where the puromycin resistance gene was inserted by electroporation.
HFM 16806	Human foetal muscle derived myoblast cell line.
E14tg2A	Hprt deficient mouse ES cell line derived from Lesch-Nyhan embryos (Hooper et al., 1987); feeder-independent; 129 genetic background.
OS25	The E14tg2A-derivative was obtained from Austin Smith (Wellcome Trust Centre for Stem Cell Research, Cambridge, UK). This cell line was previously subjected to sequential gene targeting to integrate <i>βgeo</i> into the <i>Sox2 locus</i> and hygromycin-thymidine resistance gene into the <i>Oct4 locus</i> (Billon et al., 2002); feeder-independent; 129 genetic background.
ZHBTc4	Genetically engineered mouse ES cell line carrying a tetracycline-regulatable <i>Oct4</i> transgene replacing endogenous <i>Oct4</i> alleles (Niwa et al., 2000); feeder-independent; 129 genetic background.
2TS22c	Genetically engineered mouse ES cell line carrying a tetracycline-regulatable <i>Sox2</i> transgene replacing endogenous <i>Sox2</i> alleles (Masui et al., 2007); feeder-independent; 129 genetic background; a kind gift from Dr. H. Niwa (RIKEN Centre for Developmental Biology, Japan).
2O1	<i>Sox2</i> deficient ES cells constitutively overexpressing <i>Oct4</i> (Masui et al., 2007); feeder-independent; 129 genetic background; a kind gift from Dr. H. Niwa (RIKEN Centre for Developmental Biology, Japan).

CCE	Mouse ES cell line; feeder-dependent; genetic background 129; a gift from Dr. Vasso Episkopou (MRC Clinical Sciences Centre, London, UK).
19G / 23G	Mouse embryonic germ cell lines derived from 11.5 dpc genital ridges; 129/Sv-Rosa26 genetic background; feeder-dependent; a gift from Prof A. Surani (The Gurdon institute, Cambridge, UK).
F9	Mouse embryonic carcinoma cell line; feeder-independent; genetic background 129.
P19	Mouse embryonic carcinoma cell line; feeder-independent; genetic background C3H.
B1.3 / G8.1	<i>Eed</i> ^{-/-} mouse ES cell lines (Azuara et al., 2006), derived from <i>eed</i> ^{3354SB} mutant mice obtained from the Oak Ridge National Laboratory; feeder-dependent; C5Bl/6 genetic background.
B1.3BAC	Rescued <i>Eed</i> deficient mouse ES cell line (B1.3) with a bacterial artificial chromosome (BAC) carrying <i>Eed</i> and neomycin resistance gene (see section 2.2.4); feeder-dependent; C5Bl/6 genetic background.
B1.3Neo	<i>Eed</i> ^{-/-} mouse ES cell line (B1.3) carrying the neomycin resistance gene (see section 2.2.4); feeder-dependent; C5Bl/6 genetic background.
B1.3-Oct4	<i>Eed</i> ^{-/-} mouse ES cell lines overexpressing Oct4 protein (see section 2.2.4); feeder-dependent; C5Bl/6 genetic background.
ESL32	<i>Eed</i> ^{-/-} mouse ES cell line derived from <i>Eed</i> deficient embryos (Wang et al., 2001); feeder-dependent; 129/C5Bl/6 mixed genetic background.
ESL12	<i>Eed</i> ^{+/-} mouse ES cell line; feeder-dependent; 129/C5Bl/6 mixed genetic background.
ESL21	<i>Eed</i> ^{+/+} mouse ES cell line, matching wild-type control for ESL32; feeder-dependent; 129/C5Bl/6 mixed genetic background; ESL cell lines were kindly provided by N. Brockdorff and M. Casanova (MRC clinical Sciences Centre, London, UK).
Suz12KO	<i>Suz12</i> ^{-/-} mouse ES cell line (Fujimura et al., 2006); feeder-dependent; a gift from H. Koseki (RIKEN Yokohama institute, Japan).
Suz12WT	Wild-type mouse ES cell line; matching control for <i>Suz12</i> ^{-/-} ES cells; feeder-dependent; a gift from H. Koseki (RIKEN Yokohama institute, Japan).
Ezh2-1.3	Mouse ES cell line <i>Ezh2</i> ^{flx/flx} expressing Cre-ERT2 recombinase which activation is inducible by tamoxifen. Mice with both <i>Ezh2</i> alleles (SET domain) flanked by loxP sites (Su et al., 2003) were crossed with mice carrying Cre-ERT2 cassette (Seibler et al., 2003). ES cells were then derived and clones isolated, characterised and kindly provided by S. Sauer (MRC clinical Sciences Centre, London, UK); feeder-independent.
ES-ERT2	Mouse ES cell line <i>Ring1A</i> ^{-/-} , <i>Ring1B</i> ^{flx/flx} expressing Cre-ERT2 recombinase which activation is inducible by tamoxifen (Endoh et al., 2008; Stock et al., 2007); feeder-dependent; a gift from H. Koseki (RIKEN Yokohama institute, Japan).
Oct4-GFP B	Abelson transformed mouse pre-B cell lines derived from the transgenic mice GOF18ΔPE (Palmieri et al., 1994) bone marrow; C5Bl/6 genetic background.
MEFs	Mouse embryonic fibroblasts isolated from wild type embryos (E13.5).
SNLs	Clonally derived cell line from a STO cell line (mouse embryonic fibroblasts) that stably expresses LIF and neomycin-resistance gene.

2.2. Methods

2.2.1. Cell culture

All tissue culture reagents were from Invitrogen (Invitrogen Ltd, Paisley, UK), unless stated otherwise. C2C12 cells were maintained as undifferentiated myoblasts in growth medium (GM) [Dulbecco's modified medium (DMEM) supplemented with 20% (v/v) foetal calf serum (FCS; Globepharm, Esher, UK), 4 mM L-Glutamine and antibiotics (10 µg/ml Penicillin and Streptomycin)], and routinely dissociated with trypsin-EDTA. The formation of multinucleated differentiated myotubes was induced in a mitogen poor medium (DM): DMEM, 2% Horse Serum (HS; Globepharm), 4 mM L-Glutamine and antibiotics. After one day in DM, the cells were replated on gelatin coated dishes at $3 \cdot 10^4$ cells per cm^2 . To eliminate proliferating cells 10^{-5} M Ara-C (Cytosine β -D arabino furanoside; Sigma) was added to the differentiating cultures. HFM 16806 cells were maintained as undifferentiated myoblasts in Ham F10 media supplemented with 20% FCS, 4 mM L-glutamine and antibiotics. MEFs and SNLs were cultured in DMEM supplemented with 10% FCS, 2 mM L-Glutamine and antibiotics, grown until confluence, dissociated with trypsin-EDTA, irradiated with γ -irradiation (300 rad) and frozen for future use in 10% dimethyl sulfoxide (DMSO), 90% FCS.

EBV-transformed hB clones were maintained in RPMI-1640 supplemented with 10% FCS, 2 mM L-Glutamine and antibiotics. The Oct4-GFP B-cell line was grown in RPMI-1640 supplemented with 20% FCS, non-essential amino acids, 2 mM L-glutamine, 50 µM β -mercaptoethanol, antibiotics and interleukin (IL)-7 (5 ng/ml; R&D systems, Minneapolis, MN).

ES cells were grown and maintained undifferentiated either on irradiated MEF/SNL feeder layers or directly on 0.1% gelatin-coated surfaces (Sigma), and dissociated with trypsin-EDTA. Feeder dependent ES cells were grown in KO-DMEM medium plus 10% FCS (PAA Laboratories GmbH, Pasching, Austria), non-essential amino acids, L-Glutamine, β -mercaptoethanol, antibiotics and 1000 U/ml of LIF (Esgro, Chemicon/Millipore). Feeder-free ES cells were grown on 0.1% gelatin-coated flasks (Fisher Scientific UK Ltd, Leicestershire, UK) in Glasgow's Minimal Eagle medium (GMEM-BHK 21) supplemented with 10% FCS (PAA Laboratories GmbH), non-essential amino acids, 1 mM sodium pyruvate, 0.075% sodium bicarbonate, antibiotics, 0.1 mM β -mercaptoethanol and LIF (1000 U/ml). Doxycycline (1 µg/ml, Sigma), retinoic acid (10^{-6} M, Sigma), 4-hydroxy-tamoxifen (800 nM, Sigma), cycloheximide (0.1 µg/ml, Sigma) were added to the media when indicated. Human ES cell lines H1, H7 and H9

cells (Thomson et al., 1998) were cultured in medium conditioned by mitotically inactivated MEFs supplemented with 8 ng/ml of bFGF (FGF2; Peprotech, London, UK) on matrigel-coated plates (BD Biosciences), as previously described (Xu et al., 2001). Cells were routinely passaged at a 1:3 dilution by treatment with 200 U/ml collagenase IV and mechanical dissociation. All cells were maintained at 37°C and 5% (v/v) CO₂.

2.2.2. Cell fusion and experimental heterokaryons

2.2.2.1. Muscle-lymphocyte heterokaryons

Heterokaryons were produced by fusing mouse C2C12 myotubes and human B cells with 50% polyethylene glycol, pH 7.4 (PEG 1500, Roche Diagnostics, Mannheim, Germany), according to Chiu and Blau protocol (Chiu and Blau, 1984). C2C12 myoblasts were induced to differentiate (as described in section 2.2.1) and 2 days after the onset of differentiation, non-adherent hB-lymphocytes were plated in serum free DMEM (7.10⁶ cells for 35 mm dish) and centrifuged (400 g for 15 min) to allow them to get in contact with the differentiating muscle culture. The cells were fused with 1 ml of polyethylene glycol (PEG 1500) for 1 min, rinsed three times in serum-free media (1 ml) at 37°C and maintained in differentiation media, in presence of 10⁻⁵ M Ara-C and 10⁻⁵ M Ouabain (g-strophanthin, Sigma) to eliminate proliferating cells and unfused hB cells, respectively. Heterokaryons cultures were maintained for 2-8 days in DM. Where stated, cells were treated immediately after fusion with HDAC inhibitors, 20 nM TSA or 1 mM VPA.

2.2.2.2. ES-lymphocyte heterokaryons

ES cells and hB-lymphocytes were respectively labelled with VibrantTM 1,1'-dioctadecyl-3, 3, 3', 3' tetramethylindodicarbocyanine (DiD) and 1,1'-dioctadecyl-3, 3, 3', 3'-tetramethylindodicarbocyanine perchlorate (DiI) cell labelling solutions (Molecular Probes). Cells were resuspended at 1.10⁶ cells/ml in DMEM and labelled with 5 µl/ml of dye at 37°C, 15 min. ES and hB were then mixed at an appropriate ratio (mES:hB ratio 1:1, Figure 2.1; ES:Oct4-GFP B ratio 1:5), and washed twice in phosphate buffered saline (PBS). The supernatant was completely removed and 1 ml of PEG at 37°C was added to the pellet of cells over 60 sec and incubated at 37°C for 90 sec with constant stirring. Then, 4 ml of serum-free medium (DMEM) were carefully added over a period of 3 min, followed by 10 ml of DMEM and incubation at 37°C for 3 min. After

centrifugation (1350 rpm, 5 min), the pellet was allowed to swell in complete medium for 3 min. Cell mixtures were then resuspended and cultured under conditions promoting the maintenance of undifferentiated mouse ES cells at $0.5 \cdot 10^6$ cells/cm². To eliminate unfused hB cells Ouabain (10^{-5} M) was added to the medium. Proliferating ES cells were eliminated by the addition of 10^{-5} M Ara-C 4-6 hours after fusion and then removed after 16 hours. When *Hprt*^{-/-} ES cells were used, HAT (20 μ M hypoxanthine, 0.08 μ M aminopterin and 3.2 μ M thymidine; Sigma) was added to the medium 24 hours after fusion.

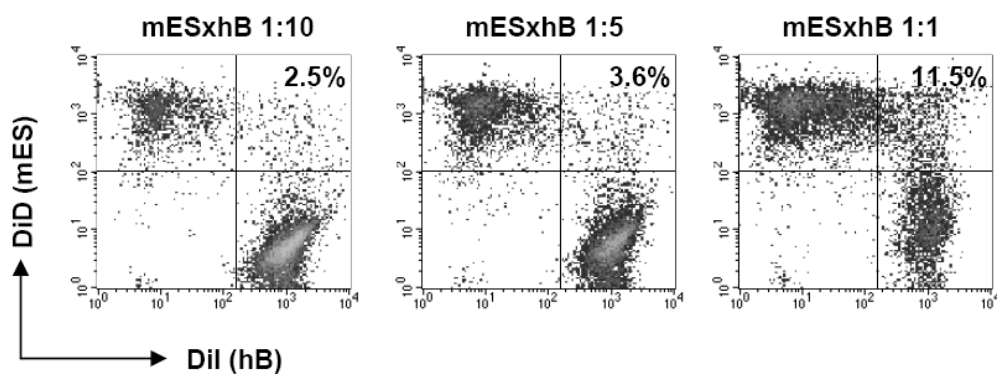


Figure 2.1. Impact of varying cell ratio in fusion efficiency. Inter-species heterokaryons (mESxhB) were generated by cell fusion of mouse embryonic stem cells (mES) with human B-lymphocytes (hB) in the presence of polyethylene glycol. hB and mES cells were respectively labelled with the cell membrane dyes Dil and DiD and analysed by flow cytometry. Fusion efficiency was assessed by FACS analysis as the percentage of double-labelled cells (upper right quadrant of dot plots). Cell fusion with different mES:hB ratios are shown.

2.2.2.3. Embryonic hybrid clones

Embryonic hybrid cells were produced by fusion of Oct4-GFP B lymphocytes with either ZHBTc4 or 2TS22c ES cells. ESxB-lymphocyte hybrids were selected using HAT for 8-12 days. GFP positive colonies were identified under a fluorescence microscope Leica DM IRE2. Colonies were isolated with a pipette, dissociated with trypsin-EDTA, and cultured in mouse ES cell conditions.

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

2.2.3.1. RNA extraction and cDNA synthesis

RNA extraction was performed using RNA-BEE reagent (Tel-Test Inc., Friendswood, TX) and residual DNA was eliminated using the DNA-free kit (Ambion, Austin, TX) according to manufacturer's instructions. RNA was then reverse transcribed using SuperscriptTM First-Strand Synthesis system (Invitrogen). 3 µg of total RNA was diluted in Rnase free water to a final volume of 11 µl and supplemented with 1 µl of 10 mM dNTP mix (Invitrogen) and 1 µl of oligo (dT)₁₂₋₁₈ (Invitrogen). The mixture was incubated at 65°C for 5 min and put on ice for 1 min, when 1 µl of 0.1 M DTT, 4 µl of 5X first strand buffer, 1 µl of RnaseOUT (Invitrogen) and 1 µl of 200 U/µl Superscript III were added. A reaction mixture without the enzyme was also set up as a control (designated "-RT"). The mixture was incubated at 25°C for 5 min, 50°C for 1 h and at 75°C for 15 min. cDNAs of interest were then detected by semi-quantitative PCR or real-time PCR.

2.2.3.2. Semi-quantitative PCR analysis

Semi-quantitative PCR analysis was performed in 50 µl using 0.4 µM primers, 1.25 U of HotStarTaqTM DNA polymerase (reference 203205, Qiagen, Valencia, CA) 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% agarose gels and visualised by ethidium bromide staining. A fifth of the PCR product was loaded and separated on 1.5% agarose gels. Blotted on Hybond-N nylon membrane (Amersham Pharmacia Biotech, Little Chalfont, UK), and hybridised with ³²P-labelled probes. Blots were washed and signals quantified using a phosphoimager. Probes were derived from cDNA preparation by PCR and ³²P-labelled using random priming (Prime-Lt II kit, Stratagene, La Jolla, USA). The primers and conditions used for this analysis can be found in Appendix I, Table A1. For each reaction of PCR amplification, conditions were maintained within the dynamic range (i.e. non-saturating).

2.2.3.3. Real-time quantitative PCR analysis (qPCR)

Real-time PCR analysis was carried out on a Opticon™ DNA engine using Opticon Monitor 3 software (MJ Research Inc., Waltham, MA), under the following cycling conditions: 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. PCR reactions included 2X Sybr-Green PCR Mastermix (Qiagen), 300 nM primers and 2 µl of template in a 35 µl reaction volume. A reaction without DNA was included to control for the formation of primer dimers and each measurement was performed in triplicate. The analysis of the qPCR data was performed with the Opticon Monitor 3 software and the relative abundance of sequences was calculated using the $\Delta\Delta C(T)$ method (Pfaffl et al., 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. *Gapdh* was generally used for gene expression data normalisation. Primer sequences can be found in Appendix I, Table A2.

2.2.3.4. Primer design and testing for quantitative PCR

Primers were designed with the Primer Express software (Applied Biosystems, Warrington, UK), using sequences from the mouse and human database from EnSEMBL. The following design criteria were used: amplicon size of 100-200 bp, GC content of 40-60%, melting temperature 55-65°C. All primer pairs were tested for predicted products within the mouse and human genome using *in silico* PCR tools such as PUNS (<http://okeylabimac.med.utoronto.ca/PUNS>) or the UCSC genome browser (<http://genome.ucsc.edu>). Primers yielding a single predicted PCR product of the correct size were ordered from Sigma Genosys (the sequences for the primers are shown in their 5' to 3' orientation).

The efficiency of amplification for each primer pair was determined by qPCR with sequential 2-fold dilutions of cDNA or genomic DNA (Pfaffl et al., 2001). Primers yielding poor linear fits of the C(T) versus logarithm of concentration ($R^2 < 0.99$) or efficiencies lower than 1.8 or greater than 2.2 were discarded. When species specificity was required (i.e. analysis in inter-species heterokaryons), primers were designed to specifically amplify the human sequences but not the mouse orthologs and then tested in human and mouse positive control cDNA.

2.2.4. Genomic PCR

2.2.4.1. Genomic DNA extraction

Genomic DNA was isolated from cells by overnight incubation in 500 µl of lysis buffer (200 mM NaCl, 100 mM Tris-HCl pH 7.4, 5 mM EDTA, 0.2% SDS, 250 µg Proteinase K) at 55°C and subsequent phenol/chlorophorm/isoamylalcohol and chlorophorm/isoamylalcohol extractions. DNA was then precipitated with isopropanol (50%), followed by 70% ethanol wash, and resuspension in Tris-EDTA (TE; 10 mM Tris-HCl pH 8.0, 1 mM EDTA). DNA concentration was quantified by spectrophotometry.

2.2.4.2. Genomic PCR

PCR analysis was performed in 50 µl using 0.4 µM primers, 1.25 U of HotStarTaq™ DNA polymerase 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% agarose gels and visualised by ethidium bromide staining. Presence of rearranged DNA of the *IgH* locus (D-J region) was checked by PCR using with the following primers: *IgH*-F: 5'-TTCAAAGCACAATGCCTGGCT-3' *IgH*-R: 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3'.

2.2.5. Cloning and DNA delivery into ES cells

2.2.5.1. Engineering of protein expression and RNAi constructs

The mouse *Oct4* open reading frame (ORF) was amplified from ES cells by RT-PCR (see section 2.2.3.2) with the linker primers 5'-CGTATCGGTACCATGGCTGGACACCTG GCTTC-3' and 5'-CCTTACGAAATTCTAACCCCAAAGCTCCAGGT 3', which introduced flanking restriction sites for KpnI and EcoRI (underlined). The PCR product was ligated into the KpnI-EcoRI sites of the pDFLAG-cDNAIII vector (a gift from Dr. B Cobb) using the procedures described on section 2.2.5.2 and sequence verified (MRC Clinical Sciences Centre sequencing facility). To generate a eukaryotic expression construct, the 2xFLAG-Oct4 sequence was subcloned into pCBA (a gift from Dr. Amy Tang) using BglII-EcoRV sites. In the pCBA vector, expression was driven by the chicken β-actin promoter and selected with neomycin resistance gene (Figure 2.2a).

Short hairpin RNA sequences targeting mouse *Oct4* transcript only (and not the human ortholog) were 5'-GATCCCCGAAGGATGTGGTTCGAGTATTCAAGAGATACTCGAACCACA TCCTTCTTTT-3' and 5'-AGCTTAAAAAGAAGGATGTGGTTCGAGTATCTCTTGAATACTCGAACCA CATCCTTCGGG-3' (Chew et al., 2005). A short hairpin with no predictable targets in both human and mouse was included as a control 5'-GATCCCCGCGCGCTTTGTAGGATTCTTCA AGAGACGAATCCTACAAAGCGCGCTTTT-3' and 5'-AGCTTAAAAAGCGCGCTTTGTAGGATTCTT CTCTTGAACGAATCCTACAAAGCGCGCGGG-3', designed with the following on-line software (<http://jura.wi.mit.edu/bioc/siRNAext/home.php>). Together, these primers constituted a double-stranded BglII-HindIII fragment coding for a short hairpin RNA. This fragment was cloned into the pSUPERneo+GFP vector (VEC-PBS-0005; OligoEngine, Inc, Seattle, WA) using the BglII-HindIII restriction sites (Figure 2.2b).

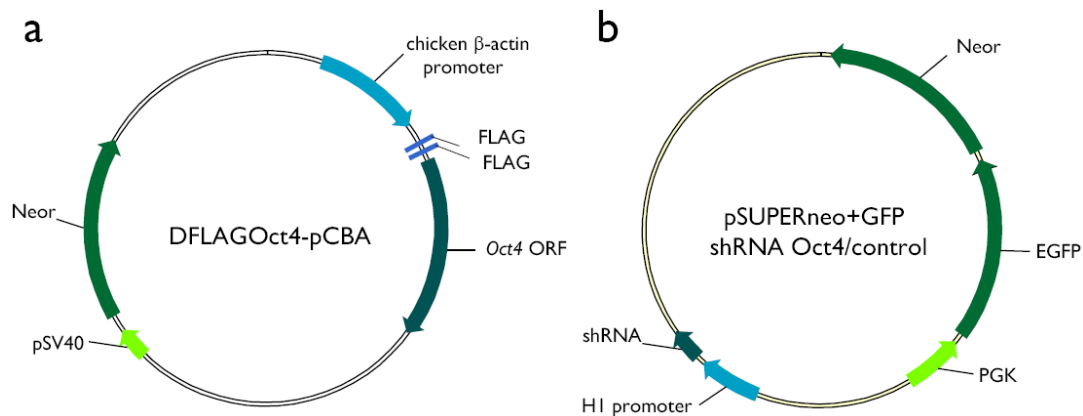


Figure 2.2. Expression vectors used in this study. (a) Vector used for overexpression of tagged *Oct4* protein in ES cells (DFLAGOct4-pCBA). Double FLAG *Oct4* ORF sequence was cloned downstream the chicken β -actin promoter. Neor – Neomycin resistance gene. pSV40 – SV40 promoter. *Oct4* ORF – mouse *Oct4* open reading frame. (b) Vector used for expression of short hairpin RNA (shRNA) in ES cells. Hairpins were cloned downstream the H1 promoter (Polymerase III dependent H1 promoter). EGFP – enhanced green fluorescence protein. PGK – murine phosphoglycerate kinase promoter.

For *Eed* reconstitution, a bacterial artificial chromosome (BAC) was inserted in *Eed* deficient ES cells (B1.3 clone). The *Eed* containing BAC clone was called RP23-370F10 and includes the 97076839-97260535 region of the forward DNA strand of mouse chromosome 7. The BAC clone was ordered in the supplied vector (BACPAC resources, Oakland, CA) which contains a chloramphenicol acetyltransferase (CAT) gene for selection in bacteria. The CAT gene was replaced with the neomycin resistance cassette from the pL452 plasmid (NCI, Frederick, MD) by recombineering. The CAT gene is 660 bp, 330 bp of the 5' and 330 bp of the 3' regions were PCR-amplified and cloned into the pL452 plasmid, 5' and 3' respectively of the neomycin resistance cassette. Recombination was induced by the transformation (as described in section 2.2.5.2) of the recombineering competent bacterial strain SW102 (NCI,

Frederick, MD) (Warming et al., 2005) with the BAC clone and posterior introduction of the pL452 plasmid by electroporation.

2.2.5.2. Plasmid construction and bacterial transformation

Enzymes and buffers were from New England Biolabs (New England Biolabs Ltd, Herts, UK). 1 µg of insert and 1 µg of vector were digested with appropriate restriction enzymes following manufacturer's instruction. Digestion products were run on 1% agarose gels with the 100 bp or 1 Kb ladders and bands of correct size were purified using QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions. 50 ng of vector and a three-molar excess of insert were ligated with 1 µl (400 U) of T4 DNA ligase overnight at 14°C. DNA was added to 20 µl of 5X KCM buffer (0.5 M KCL, 0.15 M CaCl₂, 0.25 M MgCl₂) and the total volume adjusted to 100 µl with H₂O. The mixture was added to 100 µl of competent DHS5α cells (Invitrogen), incubated on ice for 20 min, 35 sec at 42°C and on ice for 2 min. Transformed cells were then plated into Lysogeny Broth (LB) agar supplemented with 50 µg/µl of ampicillin and incubated overnight at 37°C. Individual colonies were picked into LB broth with 50 µg/µl of ampicillin and incubated overnight at 37°C with agitation. DNA was extracted from exponentially growing cultures with the Miniprep or Maxipreps kits (Qiagen) according to manufacturer's instructions and final DNA diluted in double distilled water.

For short hairpin RNA constructions the synthetic primer pairs (HPLC purified by Sigma) were treated separately with T4 polynucleotide kinase for 30 min at 37°C in a mix containing 1 µl of enzyme, 2 µl of 10x ligase buffer and 8.5 µl of 100 mM of each of the two primers. The reaction was then diluted with 180 µl of buffer A (10 mM Tris pH7.4, 50 mM NaCl), incubated at 95°C for 5 min and allowed to cool slowly to room temperature for primer annealing. 2 µl of annealed primers were then ligated to BglII-HindIII digested pSUPERneo+gfp. Bacterial clones were then picked, grown and DNA extracted and sequenced (MRC Clinical Sciences Centre sequencing facility).

2.2.5.3. DNA delivery into ES cells

DNA delivery into ES cells was performed using lipofectamine 2000 (Invitrogen). 1.10⁶ ES cells were plated in a well of a 6-well plate and 24 h later cells were transfected. 4 µg DNA were diluted in 400 µl of OptiMEM (Invitrogen) while

lipofectamine was also diluted 1:40 in OptiMEM. After mixing for 5 min both solutions were combined and left at room temperature for 20 min. The mixture was then added to the cells and incubated at 37°C for 4.5 h. Finally, the transfection mix was replaced for 2 ml of complete ES cell culture medium. The DNA BAC clone was delivered into ES cells by electroporation. ES cells plus DNA (30 µg) were transferred into a cold electroporation cuvette (Gene Pulser Cuvette, 165-2088, Bio Rad Laboratories Inc., Waltham, MA) and left on ice for 5 min before electroporation (200V, 960 µF, BioRad). Cells were incubated on ice for an additional 5 min for recovery and cultured in complete ES cell medium. When required selection was applied 24 h after transfection (G418, 400 µg/ml; Invitrogen) and resistant ES cell colonies manually isolated after 10 days.

2.2.6. Fluorescence activated cell sorting (FACS) analysis

Cell fusion efficiency was assessed using Vibrant™ DiD and DiI cell labelling solutions (Molecular Probes; emission spectra: DiI λ =565nm, DiD λ =665nm). B-lymphocytes and ES cells were labelled with DiI and DiD detected using the FL2 and FL4 detection channels, respectively.

Analysis of the SSEA4 surface antigen was performed as follows: 1.10^6 cells were washed twice in FACS buffer (PBS, 3% FCS), resuspended in anti-SSEA4 antibody (diluted 1:3 in FACS buffer) and incubated 30 min at room temperature. The cells were then washed twice in FACS buffer and incubated for an additional 30 min with a FITC-labelled secondary antibody (1:100 dilution; Jackson Laboratory, Bar Harbor, ME). Finally, cells were washed twice and then resuspended in 100 µl of FACS buffer and analysed in a FACScalibur (BD Biosciences) with CellQuest software. FACS purification was performed using a FACSaria cell sorter (BD Biosciences).

2.2.7. Immunofluorescence (IF) analysis

hB-lymphocytes were attached to glass coverslips pre-coated with poly-L-lysine (Sigma). Heterokaryons were cultured in 0.5% gelatin-coated Thermanox® coverslips (Nalge-Nunc Inc., Rochester, NY). Mouse and human nuclei were distinguished in the resulting heterokaryons by counterstaining with 4,6-diamidino-2-phenylindole (DAPI) and/or human-specific Lamin A/C staining (Figure 2.3). Shared cytoplasm was confirmed by F-actin staining (Phalloidin; A12380, Molecular Probes).

At appropriate time points coverslips were removed, washed in PBS and fixed in 2% paraformaldehyde (PFA) in PBS for 20 min. Fixed samples were washed in PBS and, for intracellular staining, permeabilised with 0.4% Triton X-100 for 5 min. Samples were incubated sequentially in blocking solution [2.5% bovine serum albumin (BSA), 0.05% Tween20, 10% Normal Goat Serum (Vector) in PBS] for 30 min and in primary antibody diluted in blocking solution (2 h at room temperature) in a humid chamber. Coverslips were washed in wash buffer (0.2% BSA, 0.05% Tween20 in PBS; 3x5 min) and incubated with secondary antibodies coupled with appropriated fluorophores (Molecular Probes) diluted in blocking solution for 45 min. Finally, cells were washed twice in wash buffer (5 min), once in PBS (3 min) and mounted in Vectashield (Vector) with DAPI (0.1 µg/ml). Staining with the antibody 5.1H11 was performed on unfixed samples labeled for 1 h at 37°C. 5.1H11-labelling was revealed using biotin coupled with anti-mouse IgG antibody (1 h at 37°C; Vector) followed by Texas-Red-coupled avidin (15 min at 37°C; Vector). All washes were performed at room temperature in DM. 5.1H11-labelled samples were fixed in 2% PFA for 20 min. For alkaline phosphatase assays, hybrid colonies 8 days after cell fusion were stained with alkaline phosphatase assay kit (Sigma) according to manufacturer's instructions.

Samples were visualised using a TCS SP5 Leica laser-scanning confocal microscope. Images were processed using Leica Confocal software and Adobe Photoshop CS2. Microscope settings and laser power were kept constant between the controls and samples. To estimate nuclear volumes, z-stacks spanning individual nuclei were collected on the laser scanning confocal microscope. The volume of each section was obtained by multiplying the xy-surface of each confocal plane by the thickness of sections along the z-axis. To investigate chromocentre organisation, image processing software (Leica Confocal Software and Metamorph 4.0) were used to perform 3D reconstructions and maximal projections of z-series collected through individual nuclei after CREST immunostaining. The number of chromocentres, the number of centromeric signals per chromocentre, and the total number of individual signals per nucleus were analysed in reconstructed projections. Images of live GFP fluorescent colonies and alkaline phosphatase staining were collected using a Leica DM IRE2 microscope running Metamorph software.

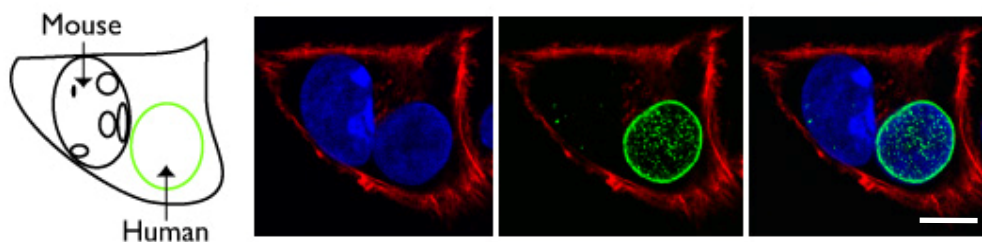


Figure 2.3. Distinction of human and mouse nuclei within heterokaryons. Human and mouse nuclei were distinguished in the resulting heterokaryons by immunofluorescence with human-specific Lamin A/C (green) staining and DAPI (blue) counterstain. Human nucleus shows a diffuse DAPI staining and positive labelling for laminA/C whereas mouse nucleus shows DAPI spotted pattern (with intense foci). Actin staining (red) delineates single cells.

2.2.8. Western blot analysis

Whole cell extracts were prepared by direct lysis of cells (5 min at 95°C and subsequent vortexing) in protein sample buffer (50 mM Tris-HCL pH 6.8, 1% SDS, 10% glycerol, 0.001% Bromophenol Blue, 5% β -mercaptoethanol). For western blot analysis of modified histones, histone proteins were isolated from whole cells by acid extraction. $10 \cdot 10^7$ ES cells were pelleted and resuspended in 1 ml PBS (4°C), centrifuged (500g for 5 min) and the supernatant removed. Cell pellets were resuspended in 180 μ l of ice cold lysis buffer (100 mM HEPES pH 7.9, 1.5 mM $MgCl_2$ 10 mM KCl, 0.5 mM DTT and 1.5 mM PMSF), 20 μ l of 2M HCl and incubated on ice for 30 min. Following acid lysis the solution was centrifuged 11,000g for 10 min at 4°C, the supernatant of acid soluble proteins collected and sequentially dialyzed against 0.1 M acetic acid (twice for 1 hour) and water (1 hour, 3 hours and overnight, respectively). The protein solution was quantified and stored at -70°C.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out with the Bio-Rad minigel system. 20 μ g of protein sample and the benchmark pre-stained protein ladder (Invitrogen) were loaded on an acrylamide (BioRad) stacking gel [5% (w/v) acrylamide, 0.125 M Tris (pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate, and 0.1% (v/v) N,N,N',N'-tetramethylethylenediamine] and separated in a 10% acrylamide resolving gel [10% (w/v) acrylamide, 0.4 M Tris (pH 8.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate, and 0.1% (v/v) N,N,N',N'-tetramethylethylenediamine] using Tris-glycine electrophoresis buffer [1.5% (w/v) Tris, 7.2% glycine, 0.5% (w/v) SDS]. Resolved acrylamide gels were blotted to a Protan nitrocellulose transfer membrane (Schleicher & Schuell Bioscience, Dassel, Germany) in transfer buffer (48 mM Trizma base, 39 mM glycine, 0.037% (w/v) SDS and 20% (v/v) methanol) using the trans-blot semi-dry

electrophoretic transfer apparatus (BioRad). The membranes were incubated for 30 min with blocking buffer [5% (w/v) fat free milk powder (Marvel), 1.2 g/L Tris pH7.4, 8.75 g/L NaCl], followed by primary antibody incubation diluted in blocking buffer for 2 h at room temperature, with agitation. After washing 3 times in wash buffer [1.2 g/L Tris pH7.4, 8.75 g/L NaCl] for 5 min, blots were incubated with horseradish peroxidase-coupled secondary antibodies (anti-rabbit and anti-mouse were from Amersham and used at 1:5000 and 1:2000 dilutions, respectively; anti-goat used at 1:2000 dilution from Santa Cruz) in blocking buffer for 1 h at room temperature. Detection was done with the ECL-Plus western blotting detection kit (Amersham) following manufacturer's instructions and using Kodak X-Omat photographic films.

2.2.9. Fluorescence *in situ* hybridisation (FISH)

2.2.9.1. Probe labelling

DNA probes for DNA and RNA FISH were labelled with either digoxigenin-11-dUTP (DIG) or biotin-16-dUTP nick translation kit (Roche) according to manufacturer's instructions. First, 1 µg of DNA was pre-treated with RNaseA (20 µg/ml) for 30 min at 37°C to remove RNA impurities and subsequently subjected to nick translation for 2 h at 15°C. The reaction was then stopped on ice and fragment size was determined by agarose gel electrophoresis. The labelled probe was then purified via columns (included in the kit) and 50 ng plus 10 µg of salmon sperm DNA (Invitrogen) were used per slide.

2.2.9.2. DNA FISH

Mouse γ -satellite (major-satellite) probe was directly labelled with Fluoro-RED and used in combination with digoxigenin-11-dUTP (DIG)-coupled human α -satellite probe (Q-Biogen, Illkirch, France). α -satellite signal was detected with anti-DIG fluorescein isothiocyanate (AD-FITC) raised in sheep (Boehringer, Mannheim, Germany), followed by anti-sheep FITC (Vector).

Cells were hypotonically treated in 75 mM KCl for 5 min at room temperature before fixation in ice-cold methanol: acetic acid 3:1. Slides were denatured for 2.5 min at 75°C in 70% formamide and 2xSSCP (10xSSC, 0.2M NaH₂PO₄, pH 7.0), immediately passed through an ice-cold ethanol series (70, 90 and 100%) and allowed

to air-dry. The probes were resuspended in 15 µl hybridization solution (2xSSC, 50% formamide, 10% dextran sulphate, 1% Tween-20), denatured at 75°C for 3 min, pre-annealed at 37°C and hybridization occurred overnight at 37°C in a humid chamber after sealing with parafilm and vulcanizing rubber solution (Weldtite, UK). Slides were washed as follows: 50% formamide and 2xSSC (0.3M NaCl, 30 mM sodium citrate pH 7) at 45 °C (3 x 5 min), 0.5xSSC (60°C, 3 x 5 min) and SSCT (4xSSC, 0.05% Tween20; Room temperature, 1 x 5 min). Slides were blocked in 5% (w/v) fat free milk powder (Marvel) in SSCT for 30 min at 37°C in humid chamber. Two layers of DIG and biotin detection were performed with 3 washes (SSCT) in between and afterwards, followed by a final wash in PBS and mounting in DAPI containing Vectashield.

2.2.9.3. RNA FISH

RNA-FISH was performed as described previously by Lawrence *et al.* (Lawrence et al., 1989) with probes synthesised by PCR. hCD20 (a 6.1 kb genomic DNA fragment spanning the last 6 exons of the human CD20 gene) was labelled with DIG, and detected as described in section 2.2.9.2. hNCAM (a 5.6kb genomic DNA fragment spanning exon 2 to exon 7 of the human NCAM gene) probe was labelled with biotin-16-dUTP and detected with Avidin-Texas-Red (Vector), followed by a biotinylated goat anti-avidin antibody (Vector) and Avidin-Texas-Red antibody (Invitrogen).

Cells were permeabilised for 7 min on ice cold CSK (cytoskeletal) buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM PIPES, 0.5% Triton X-100, 1 mM EGTA, 2 mM Vanadyl Ribonucleoside for RNase inhibition) and subsequently fixed in 4% PFA in PBS for 10 min on ice. Coverslips were washed for 3 min in SSCT and blocked with 4 mg/ml BSA in SSCT (blocking buffer) for 30 min at room temperature. After washing, 10 µl of labeled FISH probe in 150 µl of hybridization solution (2XSSC, 50% formamide, 10% dextran sulphate, 2 mM Vanadyl Ribonucleoside) was added and the coverslip was sealed with parafilm and vulcanizing rubber solution and hybridization occurred overnight at 37°C in a humid chamber. In next day, 3 washes (5 min each) in 2xSSC containing 50% formamide were performed followed by 3 washes in 2xSSC, all at 42°C. Slides were transferred into wash buffer and subsequently blocked in blocking buffer for 30 min at 37°C in humid chamber. Next, two layers of DIG and biotin detection were performed with 3 washes (SSCT) in between and afterwards, followed by a final wash in PBS and mounting in DAPI containing Vectashield.

2.2.10. Replication timing analysis

2.2.10.1. BrdU labelling and cell cycle fractionation by flow cytometry

All reagents used were from Sigma unless stated otherwise. Non-synchronized, exponentially growing hB cells were incubated with 50 μ M 5'-bromodeoxyuridine (BrdU) for 30 min at 37°C, washed in cold PBS, fixed in 70% ethanol and stored at 4°C in the dark for at least overnight. Ethanol was removed by washing the cells twice with PBS and collected by centrifugation at 1700 rpm for 7 min, 4°C. BrdU-labelled cells were resuspended in staining buffer (40 mM Tris pH 7.4, 0.8% NaCl and 21 mM MgCl₂) supplemented with 0.05% NP-40, propidium iodide (50 μ g/ml) and RNase A (1 mg/ml), carefully disaggregated by syringing through a 25G needle and allowed to stain in the dark for 30 minutes prior to separation into different phases of the cell cycle on a FACSAria cell sorter (BD Biosciences). 50,000 cells were collected for each of six fractions (G1, S1, S2, S3, S4 and G2/M) into lysis buffer (1 M NaCl, 10 mM EDTA, 50 mM Tris-HCL pH 8.0, 0.5% SDS, 0.2 mg/ml proteinase K, 0.25 mg/ml denatured, sheared salmon sperm DNA) and stored at -20°C following incubation for 2 h at 50°C.

2.2.10.2. Isolation of BrdU-labelled DNA by immunoprecipitation

BrdU-labelled DNA from *Drosophila melanogaster* Schneider cells was added to each fraction (6 ng per 10,000 sorted cells). DNA was purified by phenol, chloroform/isoamyl alcohol extraction and ethanol precipitation. Precipitated DNA was dissolved in 480 μ l of TE with 0.2 mg denatured, sheared salmon sperm DNA (Invitrogen) at 37°C for 1 hour and stored at 4°C in the dark. Samples were then sonicated (VC 130 PB, Sonics & Materials, Inc, Newtown, USA) to generate fragments of an average size of 700 bp (range 250 bp - 2 kb). After denaturation for 3 min at 95°C and cooling down on ice for 2 min, samples were adjusted to 10 mM sodium phosphate (pH 7.0), 0.14 M NaCl, and 0.05% Triton X-100 and then incubated with 80 μ l of mouse anti-BrdU antibody (25 μ g/ml) for 30 min at room temperature with constant rotation. Secondary antibody (rabbit anti-mouse IgG, Sigma) was added in excess (35 μ g) and incubated at room temperature with constant rotation for additional 30 min. DNA-protein complexes were collected by centrifugation at 15000 rpm for 20 min at 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 μ l lysis buffer (50 mM Tris-HCL pH

8, 10 mM EDTA, 0.5% SDS, 0.25 mg/ml proteinase K) and incubated overnight at 37°C. BrdU-labelled DNA was then extracted by phenol/chloroform/isoamyl alcohol and ethanol precipitation before resuspension in TE. The abundance of newly replicated DNA in each cell-cycle fraction was determined by qPCR (see section 2.2.3.3). Primer pairs used for this analysis are listed in Appendix I, Table A3.

2.2.11. Chromatin immunoprecipitation analysis

2.2.11.1. Preparation of extracts from cell lines

All reagents used were from Sigma unless stated otherwise. Approximately 1.10^8 exponentially growing hB cells were harvested, washed twice in PBS and crosslinked with 1% PFA for 10 min at room temperature. After quenching of PFA with glycine at a final concentration of 125 mM for an additional 10 min at room temperature, the cells were washed twice with ice-cold PBS containing protease inhibitor cocktail (Roche) 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. Using a Bioruptor™ 200 (Diagenode), the chromatin was sonicated on ice to an average size of 300 to 500 bp running the settings: output-high, 30 sec on, 30 sec off, for 15 min. Sonication efficiency was tested by running 4 µl of sonicated chromatin (supplemented with 2 µl 5% N-Lauroylsacrosine, 2 µl methyl-orange DNA loading buffer and 2 µl of PBS) on a 1.5% agarose gel. Sonication was repeated until the fragment size was approximately 500 bp. Insoluble proteins were discarded after centrifugation of the lysate at 15,000 rpm for 15 min at 18°C and DNA concentration quantified by spectrophotometry (NanoDrop® ND-1000, Wilmington, USA).

2.2.11.2. Bead preparation

0.2 g of protein A-sepharose beads (P3391; Sigma) per sample were rinsed twice with 2 ml of ice cold dilution buffer (1% Triton X-100, 2 mM EDTA pH 8.0, 150 mM NaCl, 20 mM Tris-HCL pH 8.1 and protease inhibitor cocktail) and blocked overnight, at 4°C with the same solution supplemented with 10 µl of BSA (New England Biolabs, 10 mg/ml) and 4 µl salmon sperm DNA (10 mg/ml), on a spinning wheel. After this blocking step, the beads were washed twice in dilution buffer and stored up to 6 weeks at 4°C in dilution buffer and lysis buffer in a 10:1 ratio.

2.2.11.3. Immunoprecipitation

Fragmented chromatin (140 µg) was diluted 1/10 in dilution buffer with protease inhibitors (Roche). 30 µl of the blocked beads were added per sample and incubated at 4°C for 2h, on a spinning wheel, to pre-clear the solution of any non-specific binding to the beads. The chromatin solution was then filtered to remove the beads. Fragmented chromatin was subjected to immunoprecipitation with 5 µg of the following antibodies: anti-H3K9Ac, -H4Ac, -H3K4m3, -H3K9m3, -H3K27m3, -H4K20m3, rabbit anti-mouse-IgG antiserum (negative-control; DakoCytomation) and anti-H3-carboxy terminal antibody, left incubating overnight at 4°C, on a spinning wheel. Afterwards, 30 µl of beads were once again added to the chromatin solutions and left immunoprecipitating 2 h on a spinning wheel, at 4°C, after which they were washed as follows: 4 times in 1 ml of ice-cold wash buffer (20 mM Tris-HCl pH 8.1, 2 mM EDTA pH 8.0, 150 mM NaCl, 0.1% SDS, 1% Triton X-100) and collected by centrifugation at 4,000 rpm for 1 min at 4°C, and washed once in final wash buffer (20mM Tris-HCl pH 8.1, 2 mM EDTA pH 8.0, 500 mM NaCl, 0.1% SDS, 1% Triton X-100) on ice. After adding elution buffer (1% SDS, 0.1M NaHCO₃) to elute the immunocomplexes from the beads, these were treated with RNase and proteinase K, 2 h at 37°C and 6 h 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 80 µL of TE. Quantification of precipitated DNA was performed using qPCR amplification (section 2.2.3.3) and the primers listed on Appendix I, Table A3. Histone's modification levels were normalised against input DNA and total H3 detected, and the ratio of modified-H3 to total H3 was denoted as relative abundance.

2.2.12. DNA methylation analysis

2.2.12.1. Bisulfite genomic sequencing

Genomic DNA was isolated from cells as described in section 2.2.4.1. Bisulfite modification of genomic DNA was carried out with the EZDNA methylation kit (Zymogenetics Inc., Orange, CA) according to the manufacturer's recommendations. 2 µg of genomic DNA was converted at 50°C (protected from light) with CT conversion reagent and column purified (provided in the kit). DNA was then eluted in 10 µl and 2 µl

of recovered DNA was used for PCR amplification (section 2.2.4.2). PCR primers were designed to recognise bisulfite-converted human DNA only using MethPrimer (<http://www.urogene.org/methprimer/index.html>) and tested *in silico* using MethBLAST (<http://medgen.ugent.be/methBLAST/welcome.html>). Primer pairs used in this analysis can be found in Appendix I, Table A4. Amplified products were cloned (see section 2.2.5.2) into pCR2 (Invitrogen) and ten bacterial clones were randomly picked, cultured overnight with LB at 37°C with agitation. Next day DNA was extracted and sequenced (MRC Clinical Sciences Centre sequencing facility).

2.2.12.2. Methyl DNA immunoprecipitation (mDIP)

mDIP was carried out in hB genomic DNA as described (Weber et al., 2005). Genomic DNA was sonicated (VC 130 PB, Sonics & Materials, Inc) to produce random fragments ranging in size from 300 bp to 1000 bp. 4 µg of fragmented DNA was used for the mDIP assay. After denaturation for 10 min at 95°C and cooling down on ice for 2 min, samples were immunoprecipitated for 2 h at 4°C with 10 µl of anti-methylcytidine antibody in a final volume of 500 µl of IP buffer (10 mM sodium phosphate pH 7.0, 0.14 M NaCl, and 0.05% Triton X-100) with constant rotation. The mixture was then incubated with 30 µl of Dynabeads with M-280 sheep antibody against mouse IgG (DynaL Biotech, Oslo, Norway) for 2 h at 4°C and washed three times with 700 µl of IP buffer. Beads were then treated with proteinase K for 3 h at 50°C and methylated DNA recovered by phenol-chloroform extraction followed by ethanol precipitation. The abundance of methylated DNA was determined by qPCR amplification (see section 2.2.3.3) with 25 ng of input DNA and 1/30 of the immunoprecipitated methylated DNA. Primer pairs used in this analysis are listed in Appendix I, Table A3.

2.2.13. Genomic context analysis

For GC content and LINE density calculations, 200 Kb of surrounding DNA (100 Kb upstream and downstream of the transcription start position) were analysed, based on the May 2004 human genome assembly of the University of California, Santa Cruz, genome browser (<http://genome.ucsc.edu/>) (Hiratani et al., 2004). The number of transcriptional units surrounding the middle position of the gene (500 Kb of surrounding DNA) was counted for gene density calculations.

CHAPTER 3. LYMPHOCYTE REPROGRAMMING TO MUSCLE IN STABLE HETEROKARYONS

3.1. Reprogramming human lymphocytes in muscle heterokaryons

The generation of multinucleated heterokaryons in which different parental cells are fused to generate polyploid cells where nuclei share the same cytoplasm but remain spatially discrete has been demonstrated between myotubes and cells of an unrelated lineage (Blau et al., 1983; Terranova et al., 2006). Upon heterokaryon formation changes in somatic cell identity are induced by muscle cells. For instance, cells derived from the three main germ layers (mesoderm, endoderm and ectoderm) can be dominantly reprogrammed to express muscle-specific genes in heterokaryons (Blau et al., 1985). Here, I will analyse in detail the kinetics of nuclear events that accompany reprogramming of human B lymphocytes (hB) by the mouse myoblast cell line C2C12 (see section 2.1.2 for cell line details). These experiments were performed in collaboration with Remi Terranova. Proliferating C2C12 myoblasts were induced to differentiate by exposure to low mitogen media. After 2 days lymphocytes were added to the myotube cultures and fused in the presence of PEG (Figure 3.1a). In inter-species heterokaryons it is possible to discriminate human nuclei (using selective primers for human genes; α -satellite DNA) from mouse nuclei (γ -satellite DNA). Human and mouse nuclei within the resulting myotubes (analysed 0-8 days after cell fusion) were distinguished by FISH using differentially labelled probes that selectively recognise human γ -satellite or mouse α -satellite DNA. α -satellite DNA (green) was present in the nuclei of human B cells before (Figure 3.1b) and after cell fusion with mouse myotubes (Figure 3.1c) and allowed them to be readily discerned from mouse C2C12 nuclei (labelled red with γ -satellite probe). Human lymphocyte-derived nuclei were also identified by counter-staining with DAPI (Figure 3.1c right panel, arrow), which highlights AT-rich regions that surround mouse centromeres and generates punctuate labelling that is selective for mouse nuclei. Human nuclei, in contrast, show diffuse DAPI labelling (illustrated in Figure 3.2a, arrowed).

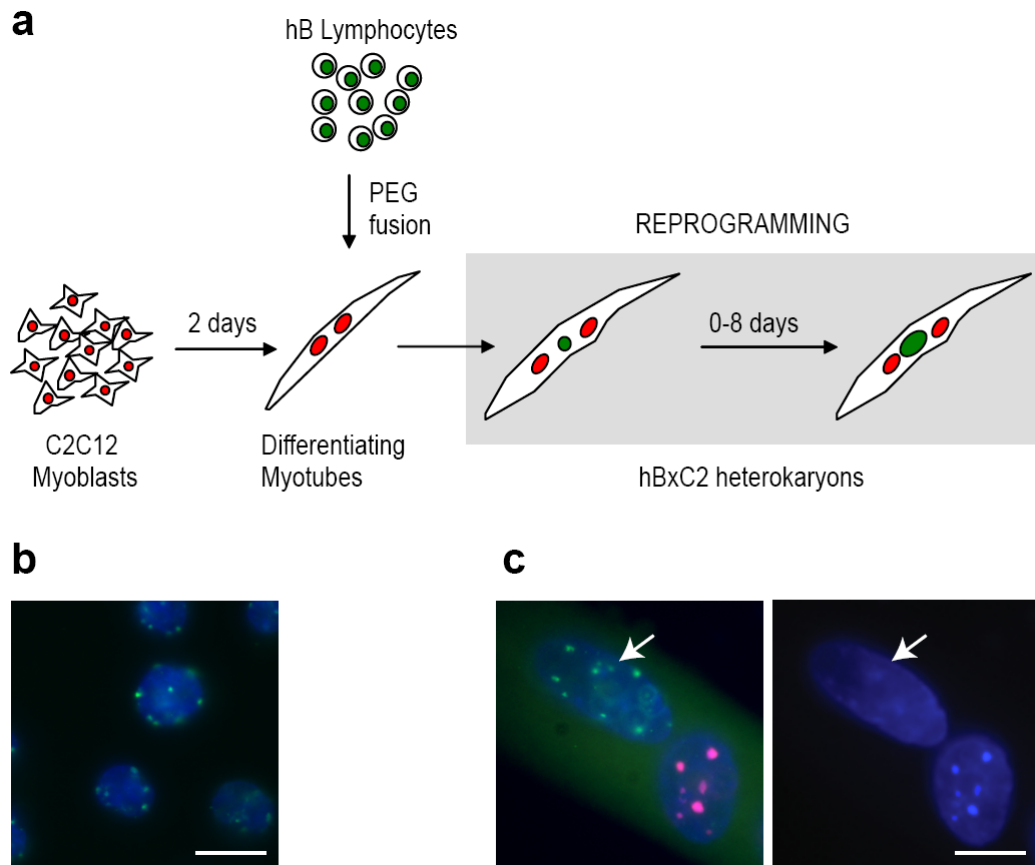


Figure 3.1. Heterokaryon formation between human B-lymphocytes and C2C12 myotubes. (a) Shows the protocol used to generate inter-species (hBxC2) heterokaryons. Mouse C2C12 myoblasts were differentiated into myotubes and fused with human B-lymphocytes (hB). Heterokaryons were then analysed from 0 to 8 days after cell fusion. Mouse and human nuclei were distinguished by FISH using probes that were specific for mouse γ -satellite DNA (red) or human α -satellite DNA (green), or by DAPI staining (blue). (b) Shows a confocal image of hB nuclei prior to fusion and (c) shows confocal images of a myotube containing a human (arrowed) and a mouse nucleus. Scale bars, 10 μ m.

As an indicator of reprogramming, expression of the human muscle marker NCAM (detected using the monoclonal antibody 5.1H11, Figure 3.2a, labelled in red) that recognises human, but not mouse NCAM (Chiu and Blau, 1984) was monitored over time using immunofluorescence labelling. Approximately 37% of heterokaryons expressed human NCAM two days after fusion, 59% after four days and 72% after seven days (Figure 3.2b), consistent with previously reported estimates of the reprogramming of mesoderm-derived cells by C2C12 myotubes (Blau et al., 1985).

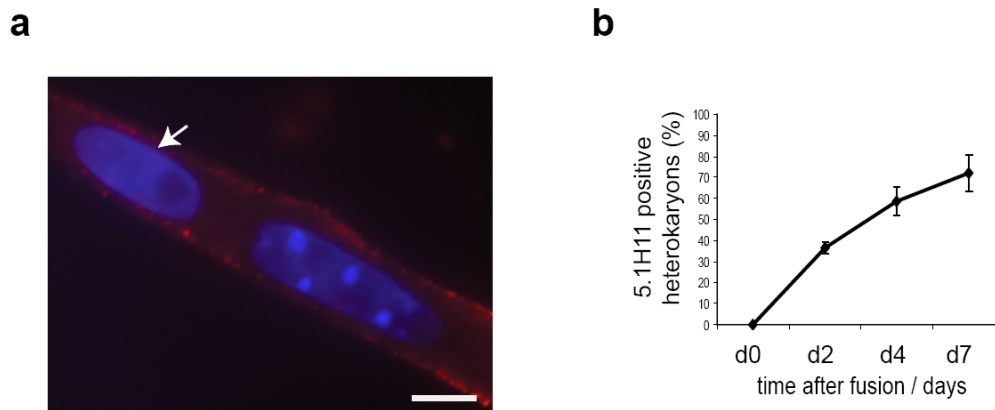


Figure 3.2. Reprogramming human B-lymphocytes in C2C12 heterokaryons. (a) Confocal section of a reprogrammed heterokaryon identified by human NCAM expression (detected with 5.1H11 antibody by immunofluorescence, red). One mouse (DAPI intense foci) and one human nucleus (arrowed) are seen (scale bar, 10 μm). (b) Expression kinetics of human NCAM (5.1H11 antibody) by lymphocyte-derived nuclei (means $\pm$ standard deviations, $n = 3$ with 50 nuclei per experiment).

3.2. Increased nuclear volume and re-distribution of heterochromatin are early events in the reprogramming of lymphocytes to muscle

Two days after fusion the global morphology of the lymphocyte-derived nuclei (in terms of both size and shape) changed, becoming similar to that of neighbouring C2C12 nuclei. This result suggests that the reprogramming of human B-lymphocytes is associated with global changes in nuclear architecture (Figure 3.3a). To investigate this further, I documented the change in the size of human lymphocyte-derived nuclei upon heterokaryon formation. The volume of human B cell nuclei, estimated by immunofluorescence (materials and methods, section 2.2.7), increased from approximately $600 \mu\text{m}^3$ prior to fusion to $1500 \mu\text{m}^3$ and $1800 \mu\text{m}^3$ two and four days after heterokaryon formation, respectively (Figure 3.3b). The distribution of constitutive heterochromatin within these nuclei was analysed by immunofluorescence with CREST (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia) autoimmune serum. Human CREST antiserum was found to recognise three constitutive centromeric proteins CENP-A, -B, and -C (Earnshaw et al., 1986; Moroi et al., 1980).

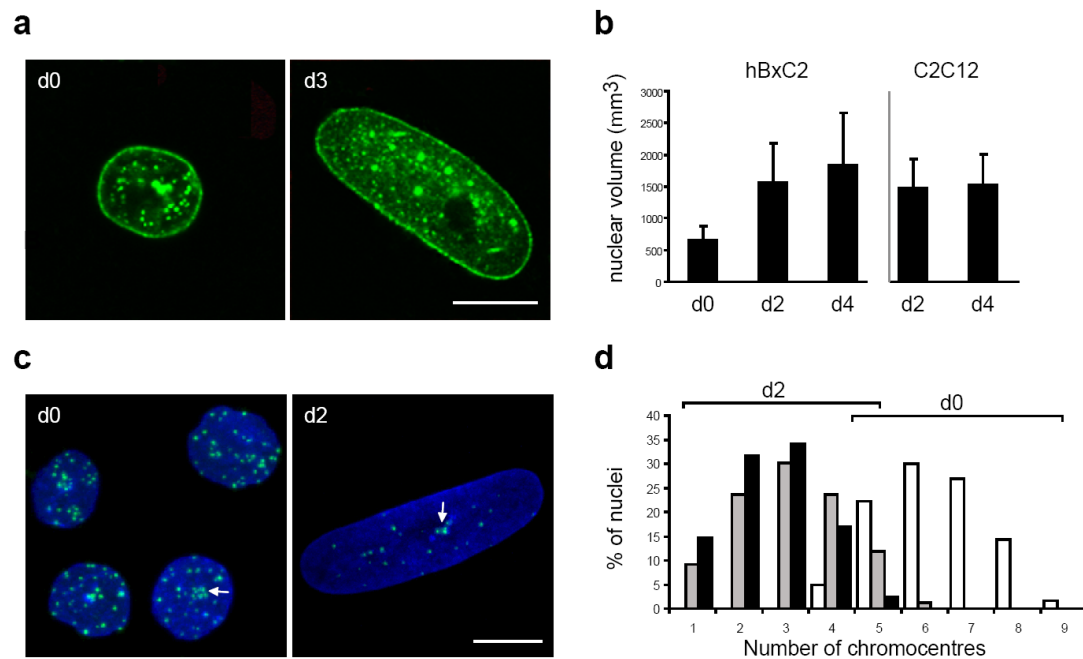


Figure 3.3. Nuclear reorganisation is an early event in lymphocyte reprogramming. (a) Single optical section showing lamin A/C immunofluorescence labelling of human B-lymphocyte nuclei (hB) before (d0) and three days after fusion with C2C12 myotubes (d3). (b) The volume of hB nuclei before (d0), 2 and 4 days after heterokaryon formation (d2, d4) compared to C2C12 myotubes 2 and 4 days after serum withdrawal and differentiation (d2, d4). Nuclear volume was estimated as described in materials and methods (section 2.2.7) (means $\pm$ standard deviations, $n = 3$ with 50 nuclei per experiment). (c) CREST antiserum reveals the centromere distribution (green) in hB nuclei (d0) and two days after heterokaryon formation (d2), where individual chromocentres are indicated (arrowed). Confocal images are maximal projections of multiple optical z-sections. (d) Distribution of constitutive heterochromatin in hB nuclei before (d0, open bars) and 2 days after fusion in hNCAM expressing (black bars) and hNCAM negative (not yet reprogrammed) heterokaryons (grey bars) was compared by assessing the number of discrete CREST signals and number of chromocentres per nucleus ($n = 100$). Scale bars, 10 μ m.

Prior to fusion, human centromeres were clustered to form 5-8 chromocentres per nucleus detected by CREST labelling (one chromocentre was defined as three or more closely juxtaposed CREST signals; Figure 3.3c, green) (Alcobia et al., 2000; Brown et al., 2001; Weierich et al., 2003). Two days after heterokaryon formation, the number of chromocentres per nucleus was markedly reduced (Figure 3.3c, d) as well as CREST spatially discrete signals (Figure 3.4, middle panel). This reorganisation mimicked the spatial distribution of constitutive heterochromatin seen in neighbouring C2C12 nuclei (Figure 3.4, right panel) and occurred rapidly as no further major changes occurred at later time points (between day 2 and day 7, Figure 3.4).

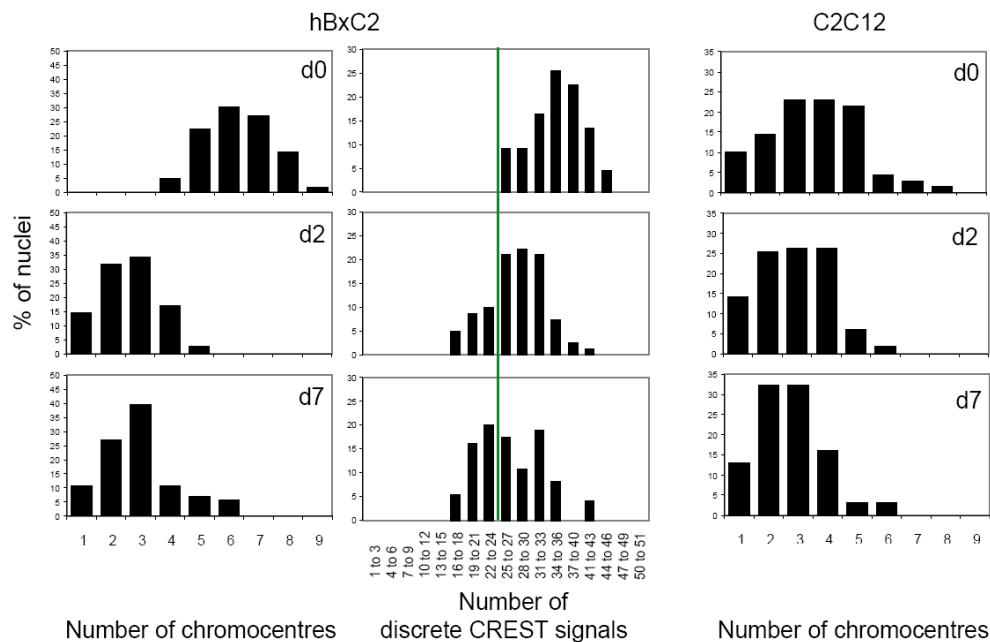


Figure 3.4. Early redistribution of heterochromatin in hBxC2 heterokaryons. CREST immunostaining was used to compare the distribution of constitutive heterochromatin in human B-lymphocyte (hB) -derived nuclei prior to fusion (d0), two (d2), and seven (d7) days after hBxC2 heterokaryon formation and successful reprogramming (5.1H11 positive heterokaryons). The number of chromocentres and the number of discrete CREST signals was assessed using 3D-reconstructed sequential z-series spanning 100 individual nuclei for each time point. Differentiating C2C12 cultures (0, 2 and 7 days) were analysed as controls (right panel).

Changes in nuclear size and architecture were evident prior to the expression of hNCAM (Figure 3.3d, comparison between hNCAM positive and negative heterokaryons; black and grey bars, respectively), suggesting that nuclear reorganisation may precede overt human muscle reprogramming. This rapid timing of response implicates transacting factors derived from mouse myotubes in the nuclear remodelling of lymphocytes. Consistent with this possibility, mouse Myogenin protein was detected in many human nuclei as early as 24 hours after heterokaryon formation (Figure 3.5), several days before endogenous human myogenin transcripts were expressed (Figure 3.6).

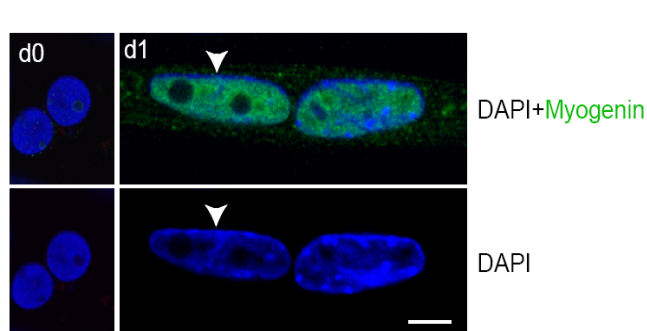


Figure 3.5. Rapid translocation of mouse Myogenin to human nuclei. Immunofluorescence detection of Myogenin protein (green) in a human nucleus (arrowed) 24 hours after cell fusion. Human and mouse nuclei were distinguished on the basis of diffuse and spotted DAPI staining (blue), respectively. Images show single confocal optical sections. Scale bar, 10 µm.

3.3. Reprogramming initiates a temporally ordered activation of genes encoding human myogenic regulatory factors by lymphocyte-derived nuclei

To determine the kinetics of gene activation and silencing induced by heterokaryon formation and reprogramming of human lymphocytes, we analysed the expression of a panel of human muscle- and B cell-associated genes using RT-PCR and primer combinations that were designed to amplify specifically human transcripts only (listed in Appendix I, Table A1). As a control, mouse C2C12 cDNA was included at all times to confirm the specificity of each primer combinations to human transcripts (C2, Figure 3.6). This analysis showed that muscle gene expression was efficiently initiated in human B cell nuclei two days after heterokaryon formation (Figure 3.6). Interestingly, human MRFs were expressed in a sequence that accurately reflects their normal temporal order during myogenic differentiation (Pownall et al., 2002), so that *hMyf5* was transiently expressed at days 2-3, followed by *hMyogenin* (day 3 onwards) and *hMrf4* later (day 4 onwards).

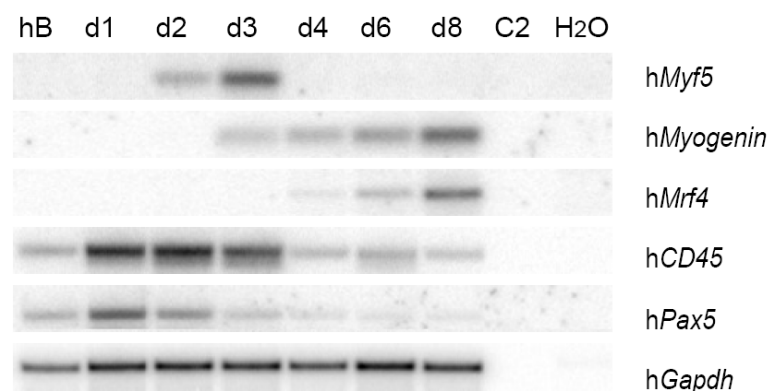


Figure 3.6. Ordered expression of human myogenic genes by lymphocyte-derived nuclei. Expression of human muscle-specific (*hMyf5*, *hMyogenin*, *hMrf4*) and human lymphocyte-specific (*hCD45*, *hPax5*) genes detected by RT-PCR before (hB) and after cell fusion (day 1 to 8). Prior to fusion, hB expressed *hGapdh*, *hCD45* and *hPax5*, but not muscle-specific genes. Following heterokaryon formation, human *Myf5*, *Myogenin* and *Mrf4* expression was initiated and the expression of human lymphocyte-specific genes *CD45* and *Pax5* declined. Mouse C2C12 samples (C2) were used as negative controls to confirm the specificity of primers to human transcripts and *hGapdh* was used to standardise input.

The establishment of a novel gene program by human B cell-derived nuclei was accompanied by declining expression of the B cell-specific regulator *hPax5* and the leukocyte-associated gene *hCD45*. Expression of human *Gapdh*, a ubiquitously expressed gene, remained constant throughout these analyses. This observation suggests that silencing and activation of gene expression programs are *coordinated* during reprogramming.

3.4. HDAC activity is required for the silencing of lineage inappropriate genes during reprogramming

To assess whether HDAC activity is required to dominant reprogramming of lymphocytes to muscle, we treated heterokaryons with the HDAC inhibitor TSA. Application of low doses of TSA to differentiating C2C12 or to primary cultures of mouse muscle has been shown to increase histone H3K9 and H4K20 trimethylation at pericentric DNA and prevent centromeric clustering (Terranova et al., 2005), suggesting a role for HDACs in the formation of repressive heterochromatin environments during terminal differentiation. To assess their importance during lineage conversion, heterokaryons established between human B-lymphocytes and C2C12 cells were treated with low doses of TSA (20 nM). This did not affect the success rate of heterokaryon formation or the proportion of hNCAM positive reprogrammed nuclei seven days after fusion ($79 \pm 6.7\%$, compare with $72 \pm 8.7\%$ in absence of TSA), although a slightly enhanced induction of hNCAM expression was seen in the presence of TSA (Figure 3.7a).

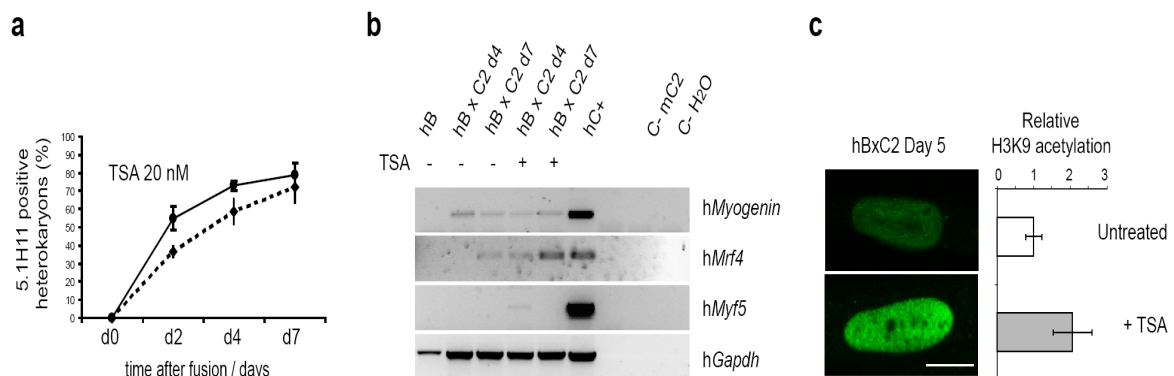


Figure 3.7. HDAC inhibition does not impair the activation of muscle gene expression program. (a) Expression kinetics of human NCAM (assessed by 5.1H11 immunofluorescence) in hBxC2 heterokaryons in the presence (solid line) or absence (broken line) of TSA (20 nM). For each time point, 100 heterokaryons were analysed in two separate experiments, values shown are the mean and standard deviation. (b) RT-PCR analysis of *hMyogenin*, *hMrf4*, *hMyf5* and *hGapdh* gene expression in human B-lymphocytes (hB) and in day 4 (d4) and day 7 (d7) hBxC2 heterokaryons cultured in the presence (+) or absence (-) of 20 nM TSA. Differentiating human muscle was used as positive control (hC+) and mouse C2C12 cells (C- mC2) as negative control for RT-PCR analysis. *hGapdh* was used to standardise input. (c) Immunofluorescence analysis of day 5 heterokaryons (hBxC2 day5) showed increased histone H3 lysine 9 (H3K9) acetylation in human nuclei in response to treatment of heterokaryons with 20 nM TSA (+ TSA). Confocal microscope settings and laser power were kept constant so that the relative abundance of H3K9 acetylation could be directly compared (and quantified, right hand panel) to untreated control (means $\pm$ standard deviations, 50 nuclei per experiment). Scale bar, 10 μ m.

A comparison of *Myf5*, *Myogenin* and *Mrf4* expression in heterokaryons that were treated with TSA or untreated (Figure 3.7b) confirmed that acquisition of muscle gene expression by human lymphocyte-derived nuclei was not compromised. *Myogenin*

expression was detected at days 4 and 7, *Mrf4* was most abundant at later stages while *Myf5* was only detected early after heterokaryon formation (days 2-3). TSA treatment resulted in a slight advance in the timing of *hMrf4* expression, a result that is consistent with a global increase in histone acetylation (Figure 3.7c) and chromatin accessibility induced by TSA and a previous report that showed precocious expression of *mMrf4* in primary mouse muscle treated with HDAC inhibitors (Terranova et al., 2005).

TSA did however have a dramatic effect on the temporal extinction of human lymphocyte gene expression in inter-species heterokaryons (Figure 3.8). Whereas expression of *hCD45*, *hPax5*, *hCD20* and *hCD37* was extinguished (or significantly diminished in the case of *hCD45*) in heterokaryons by day 7, sustained expression of these genes was evident in TSA treated heterokaryons (highlighted in green, compare left and right panels, Figure 3.8). Human *Gapdh*, a gene that is expressed in all human tissues, was detected throughout in both TSA-treated and untreated cultures and is shown for comparison.

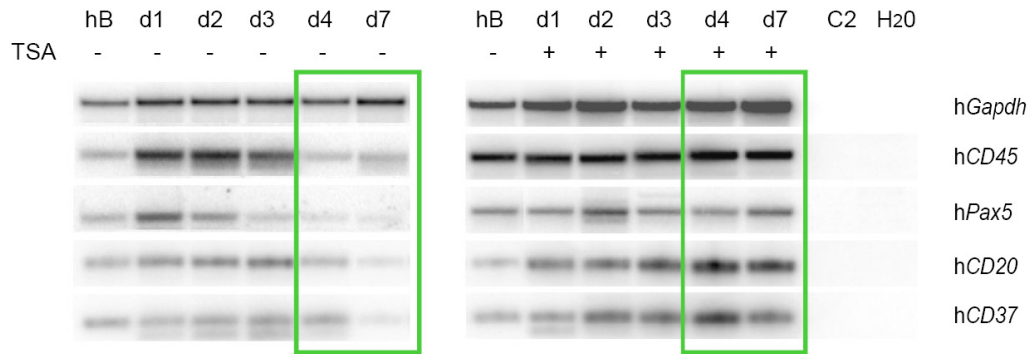


Figure 3.8. HDAC inhibition prevents the shutdown of lymphocyte-specific genes during reprogramming. RT-PCR analysis showed the impact of 20 nM TSA treatment on the expression of *hCD45*, *hPax5*, *hCD20* and *hCD37* in inter-species hBxC2 heterokaryons 0 (hB) to 7 days after cell fusion (d1-d7). *hGapdh* was used to standardise input. Green boxes highlight that lymphocyte-associated genes were not extinguished following TSA treatment (right hand panel).

To evaluate the selectivity of this TSA response in heterokaryons, we analysed the expression of other human genes including *hPax6*, *hRex1*, *hTert*, *hNanog* and *hOct4*. Most genes were not significantly upregulated in heterokaryons upon TSA treatment (Figure 3.9a), with the exception of *hOct4* which was transiently activated (Figure 3.9a, hB x C2 d4) upon cell fusion. Similarly, TSA treatment of mouse myoblasts and mouse myotubes did not induce the inappropriate expression of lymphocyte-associated and neural-associated genes (Figure 3.9b) although 2 of the 11 tested (*mEngrailed* and *mCD4*) showed slight up-regulation. This result is consistent with estimates that approximately 10% of genes may be sensitive to HDAC inhibition in

myoblasts (Iezzi et al., 2004), and suggests that HDAC inhibition in experimental heterokaryons selectively targets the silencing of lymphocyte associated gene expression program. This claim is also supported by the finding that *hPax5* extinction in heterokaryons is not associated with reduced expression of *Pax5*-upstream regulators (such as *hE2A* and *hEbf*) and that *hPax5* (but not *hE2A* or *hEbf*) was sensitive to TSA treatment (Figure 3.10), as confirmed by quantitative RT-PCR (primers listed in Appendix I, Table A2). The expression of *hE2A* and *hEbf* in heterokaryons showed a similar pattern to the one seen in differentiating human muscle cultures (right panel), confirming that these transcriptional regulators are not silenced in the muscle lineage.

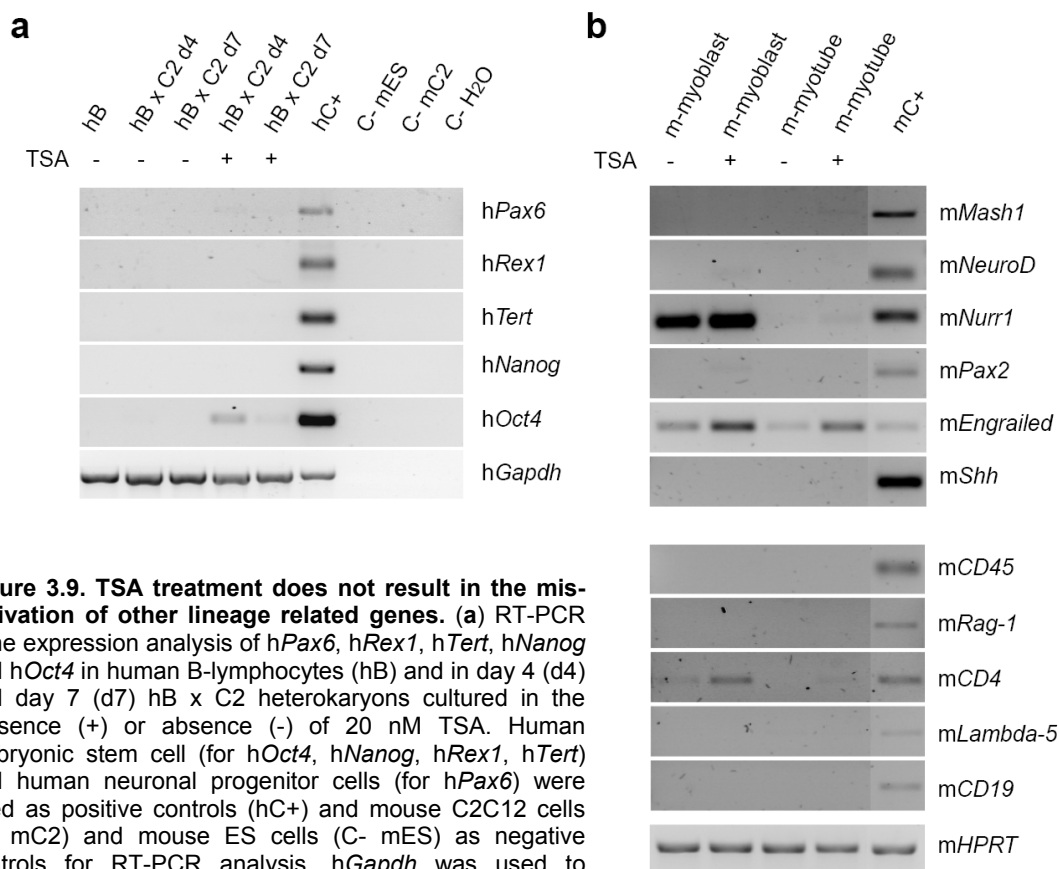


Figure 3.9. TSA treatment does not result in the mis-activation of other lineage related genes. (a) RT-PCR gene expression analysis of *hPax6*, *hRex1*, *hTert*, *hNanog* and *hOct4* in human B-lymphocytes (hB) and in day 4 (d4) and day 7 (d7) hB x C2 heterokaryons cultured in the presence (+) or absence (-) of 20 nM TSA. Human embryonic stem cell (for *hOct4*, *hNanog*, *hRex1*, *hTert*) and human neuronal progenitor cells (for *hPax6*) were used as positive controls (hC+) and mouse C2C12 cells (C- mC2) and mouse ES cells (C- mES) as negative controls for RT-PCR analysis. *hGapdh* was used to standardise input. (b) RT-PCR analysis of mouse transcripts for *mMash1*, *mNeuroD*, *mNurr1*, *mPax2*, *mEngrailed1*, *mShh* (neural-associated) and *mCD45*, *mRag1*, *mCD4*, *mLambda-5*, *mCD19* (lymphocyte-associated) in mouse myoblasts and myotubes cultured in the presence or absence of 20 nM TSA. Mouse brain and foetal liver (mC+) were used as positive controls for this analysis. *mHPRT* was used to standardise input.

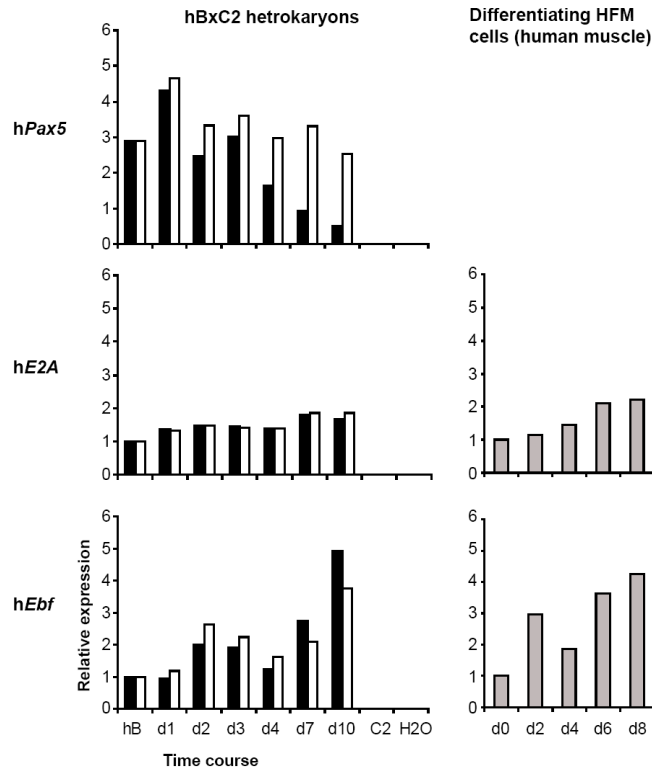


Figure 3.10. *hPax5* extinction in heterokaryons is not associated with a decline in the expression of *hE2A* and *hEbf*. Quantitative RT-PCR analysis of human *hPax5*, *hE2A* and *hEbf* expression in hBxC2 heterokaryons treated with 20 nM TSA (white bars) or untreated (black bars), 0 (hB) to 10 days (d1-10) after cell fusion. Expression of *hE2A* and *hEbf* was compared with control cultures of differentiating human HFM (Human Foetal Muscle) cells (grey bars). Mouse C2C12 samples (C2) were used as negative controls to confirm the specificity of primers for human transcripts and the data were normalised to *hGapdh* expression.

3.5. Inhibition of HDAC activity results in the co-expression of *hCD20* and *hNCAM* in individual reprogrammed nuclei

The observation that TSA prevents the silencing of lymphocyte-associated genes but allows the acquisition of muscle gene expression predicts that TSA treated heterokaryons would contain nuclei that simultaneously express lymphocyte and muscle associated genes. To verify this at the level of individual nuclei, RNA FISH was performed to simultaneously detect human muscle- (*hNCAM*) and lymphocyte- (*hCD20*) specific transcripts in TSA treated experimental heterokaryons. To verify the specificity of probes used to detect human *CD20* (*hCD20*-DIG) RNA transcripts, the detection *hCD20* signals (green) in human B-lymphocyte nuclei were sensitive to RNase but not to DNAase or TSA treatment (Figure 3.11a). Human *NCAM* (*hNCAM*-biotin) was not detected in any of these conditions in human B-lymphocytes (Figure

3.11a) but two hNCAM signals (red) were seen in human myoblasts (Figure 3.11b), consistent with gene expression patterns.

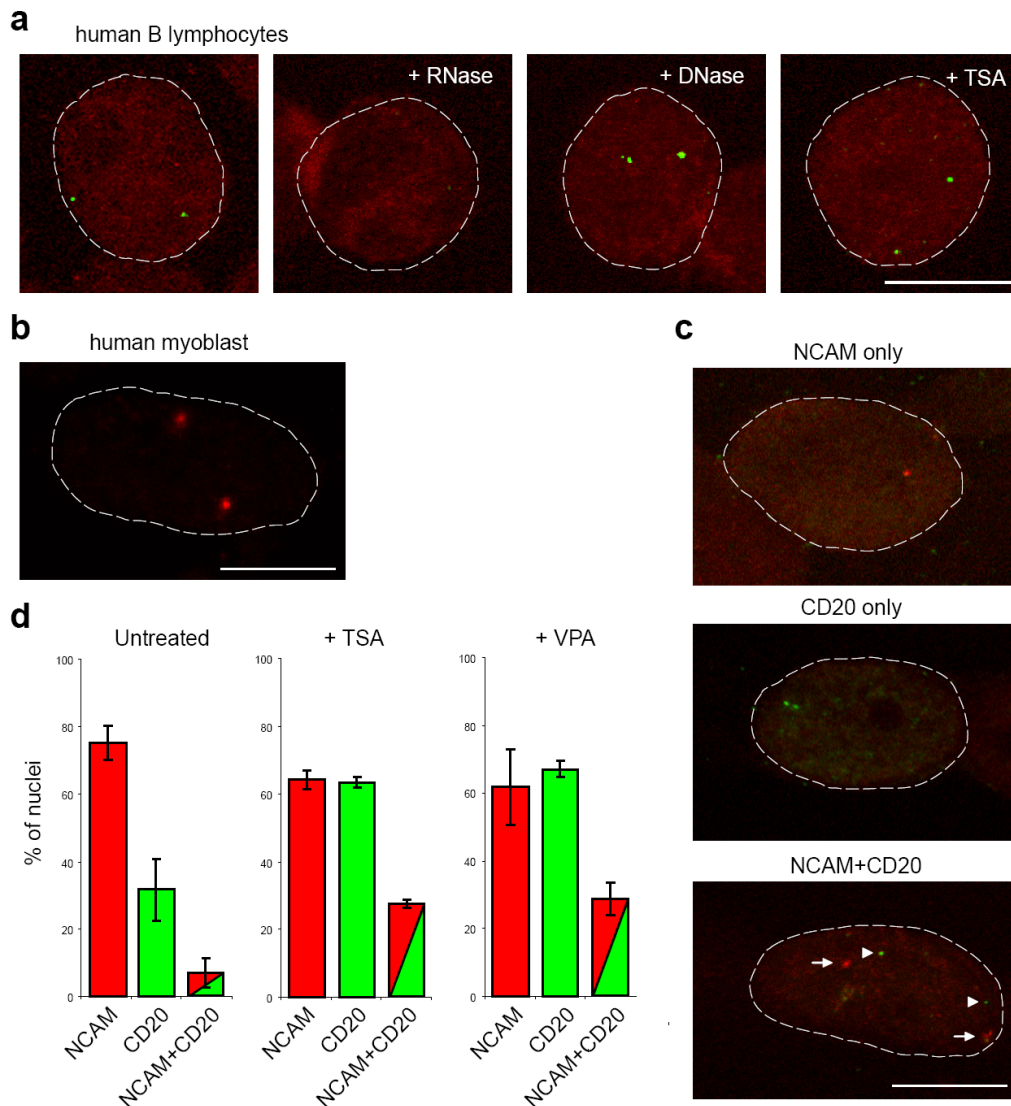


Figure 3.11. Co-expression of lymphocyte- and muscle-specific genes in individual nuclei treated with the HDAC inhibitors TSA and VPA. (a) Human B-lymphocytes were analysed for hCD20 (green) and hNCAM (red) transcripts using RNA FISH. Two foci of hCD20 signal were evident in hB cell nuclei that were removed by pre-treatment of samples with RNase A (+RNase), but not DNase (+DNase). Application of 20 nM TSA did not impair hCD20 signal detection (+TSA). (b) Human NCAM transcripts (red) were detected in human myoblast nuclei but not in lymphocytes. (c) RNA FISH showing examples of nuclei that express hCD20 transcripts alone (green) and hNCAM transcripts alone (red) or both (red/green, arrowed) in lymphocyte-derived nuclei within day5 hBxC2 heterokaryons. Confocal images are maximal projection of multiple optical z-sections (scale bars, 10 μ m). (d) The proportion of nuclei expressing hNCAM (red), hCD20 (green), or both hNCAM and hCD20 (red/green) signals in heterokaryons cultured in the absence (untreated) or presence of 20 nM TSA (+TSA) or 1 mM VPA (+VPA) is expressed as a proportion of human nuclei in which RNA transcripts were detected.

Using these probes, RNA-FISH analysis was used to score nuclei expressing the muscle gene hNCAM alone, the lymphoid gene hCD20 alone, or co-expressing both hNCAM and hCD20 within individual nuclei in heterokaryons at day 5 (illustrated in Figure 3.11c; quantification is shown in Figure 3.11d and Table 3.1). In untreated

heterokaryons in which human RNA signals were found, most nuclei (75%) expressed *hNCAM* and only a minority (32%) expressed *hCD20*. Nuclei in which both transcripts were detected simultaneously were rare (7%), and may reflect nuclei “in transit” from the lymphoid to muscle program. Addition of TSA or the structurally unrelated HDAC inhibitor VPA did not affect muscle-specific *hNCAM* signals (red histograms) but had a dramatic effect on the proportion of lymphocyte-derived nuclei that retained *hCD20* expression (green histograms, Figure 3.11d). Consequently, and in contrast to untreated cultures, many lymphocyte-derived nuclei within TSA- or VPA-treated heterokaryons co-expressed both *hCD20* and *hNCAM* (28 and 29%, respectively; red/green histograms, Figure 3.11d). These results confirm that HDAC activity is important for the silencing of lymphocyte-specific genes during reprogramming.

Table 3.1. RNA FISH detection of *hCD20* and *hNCAM* transcripts in day-5 heterokaryons.

Sample	Nuclei with <i>hNCAM</i> signals only	Nuclei with <i>hCD20</i> signals only	Nuclei with both <i>hCD20/hNCAM</i> signals	Sample size	No signal
Untreated exp. 1	20	5	1	56	30
exp. 2	29	10	2	72	31
exp. 3	20	10	4	60	26
+ TSA (20 nM) exp. 1	20	18	15	75	22
exp. 2	16	17	12	60	15
+ VPA (1 mM) exp. 1	15	16	17	77	29
exp. 2	17	21	11	69	20

3.6. Discussion and future perspectives

Demonstrations of somatic cell reprogramming by nuclear transfer and in experimental heterokaryons have shown that many types of differentiated cells retain a flexible lineage potential (Baron and Maniatis, 1986; Blau et al., 1983; Gurdon, 1962; Hochedlinger and Jaenisch, 2002; Wilmut et al., 1997) [reviewed by DiBerardino (DiBerardino, 1988) and Surani (Surani, 2001)]. Here the epigenetic plasticity of human B-lymphocytes was assessed in stable C2C12-derived muscle heterokaryons. I show that dominant conversion of human lymphocytes to muscle is a multi-step process that includes changes in nuclear size, architecture and gene activity. In muscle heterokaryons, where chromosome replication is precluded (Chiu and Blau, 1984), these steps can be distinguished on the basis of their timing and the requirement for HDAC activity. Within the first 48 hours of heterokaryon formation, lymphocyte nuclei increase in size and heterochromatin domains are redistributed to mimic the spatial arrangement of neighbouring mouse myocytes. Importantly, this reorganisation appears to precede the activation of endogenous human muscle-specific genes. It is possible that these changes in nuclear architecture are the result of physical constraints, or that dominant factors within the myotubes impose a muscle-specific organisation on the introduced nucleus. As cycling and non-cycling lymphocytes display different constitutive heterochromatin organisation (Brown et al., 1999; Solovei et al., 2004) and cell cycle withdrawal is a characteristic feature of myogenic differentiation (Ait-Si-Ali et al., 2004; Shen et al., 2003; Walsh and Perlman, 1997), it is also possible that nuclear reorganisation is a consequence of imposing cell cycle arrest. Previous studies have shown that fusion of nucleated chicken erythrocytes with rat myoblasts results in pronounced nuclear enlargement and chromatin redistribution prior to gene reprogramming (Dupuy-Coin et al., 1976). Similar effects have been reported for mouse lymphocytes injected into *Xenopus* oocytes (Byrne et al., 2007) and bone marrow cells that form stable heterokaryons with Purkinje neurons (Weimann et al., 2003). Chromatin redistribution in developing muscle has also shown to be dependent on the methyl-CpG binding proteins MeCP2 and MBD2 (Brero et al., 2005). MeCP2 interacts with HP1 to modulate heterochromatin association (Agarwal et al., 2007). Taken together, these results suggest that myogenic regulators and chromatin remodelling machinery are both likely to have important roles in conveying dominant cell-type-specific nuclear organisation during reprogramming.

Trans-acting factors derived from mouse myotubes, that are implicated in nuclear remodelling during the myogenic conversion of lymphocytes, have yet to be

characterised. The MRFs: MyoD, Myf5, Myogenin, and Mrf4, are conserved bHLH transcription factors present in vertebrates (Weintraub et al., 1991), that have the remarkable ability to convey muscle-associated properties to many non-muscle cell types (Aurade et al., 1994; Choi et al., 1990; Davis et al., 1987; Ohkawa et al., 2006). Ectopic expression of Myogenin in fibroblasts, induces muscle differentiation in a manner dependent on the chromatin-remodelling enzyme Brg1 (Ohkawa et al., 2006). These results suggest that MRFs convey the muscle identity by cooperating with the SWI/SNF chromatin remodelling machinery to alter chromatin structure (de la Serna et al., 2001). Consistent with this possibility, mouse Myogenin protein (myotube-derived) was detected in human nuclei (lymphocyte-derived) as early as 24 hours after heterokaryon formation. However, to carefully elaborate which factors are required and sufficient to establish muscle-specific identity, will require additional experiments. For example, using inter-species heterokaryons treated with drugs that inhibit the activity of specific enzymes, or using RNA interference to silence mouse myotube-specific genes and alter *trans* reprogramming.

My results show that *de novo* expression of an endogenous muscle gene program by lymphocyte-derived nuclei begins two days (*hMyf5*) after heterokaryon formation. The activation of *hMyf5* is then followed by *hMyogenin* (day 3 onwards) and *hMrf4* later (day 4 onwards) (summarised in Figure 3.12). The pattern of expression of the MRFs (*MyoD*, *Myf5*, *Myogenin* and *Mrf4*) genes in vertebrates provided evidence that these regulators are involved in both the specification and differentiation of myogenic progenitor lineages (Pownall et al., 2002). *Myf5* and *MyoD* are expressed in proliferative myoblasts (Emerson, 1990), and in tissue culture cells this expression is stably maintained through autoregulation (Thayer et al., 1989), which implicates *Myf5* and *MyoD* in the establishment and maintenance of muscle progenitor lineages. However, *Myogenin* and *Mrf4* are activated during myoblast differentiation (Brownell et al., 1996; Rhodes and Konieczny, 1989; Wright et al., 1989) and likely have functions later for the activation of genes associated with muscle contractile proteins (Lassar et al., 1991). Interestingly, human MRFs genes were activated in a lymphoid cell type upon cell fusion in a sequence that accurately reflects their temporal order during normal muscle differentiation. In agreement with my findings, the activation of *hMyoD*, one of the earlier markers of myogenic commitment, was also shown to occur rapidly (transcriptional activation between day 1 and 3) in heterokaryons generated between human keratinocytes and mouse muscle cells (Zhang et al., 2007).

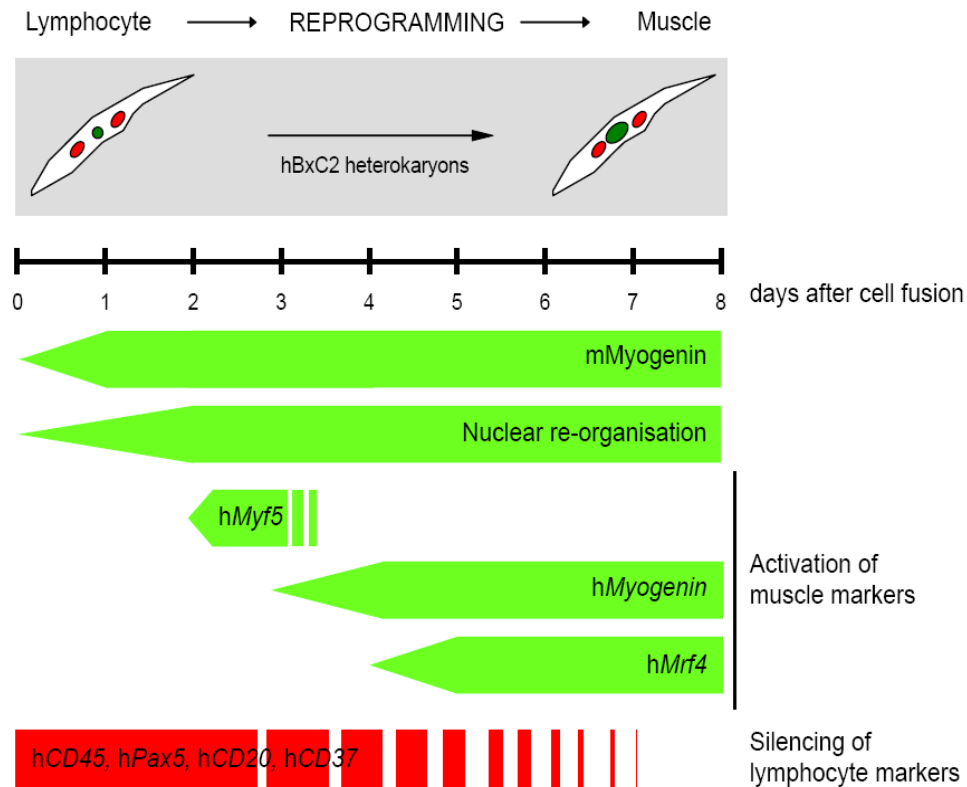


Figure 3.12. Kinetics of events involved in the multi-step reprogramming of lymphocytes to muscle in heterokaryons. *Trans* acting factor (mMyogenin) migration to human nuclei and nuclear re-organisation (nuclear size and re-distribution of heterochromatin domains) are early events of reprogramming (day 1 and 2, respectively). Activation of endogenous human muscle markers begins at day 2 (hMyf5) and is followed by hMyogenin (day 3 onwards) and hMrf4 later (day 4 onwards). These events are indicated by the green arrows. Reduction of lymphocyte-specific gene expression (hCD45, hPax5, hCD20 and hCD37) begins after 2-3 days and continues over a 7-day period (red boxes).

Activation of muscle-specific genes in lymphocyte-derived nuclei was accompanied by the extinction of several lymphoid-associated genes (hCD45, hPax5, hCD20, hCD37). Reduced expression of leukocyte-specific genes began 2-3 days after cell fusion and continued over a 7-day period (Figure 3.12). These kinetics were similar to the reported shut-down of human albumin expression during hepatocyte reprogramming to muscle (Miller et al., 1988). Thus reprogramming to muscle identity appears to be coordinated with the silencing of lineage-inappropriate genes in heterokaryons.

Throughout development, cellular differentiation is characterised by the coordinated activation and extinction of genes at specific times. In developing lymphocytes, the regulated expression and silencing of the terminal deoxynucleotidyl transferase locus, *Dnrt*, provides a particularly well studied example (Wernig et al., 2007). *Dnrt* is expressed by immature lymphocytes but is transcriptionally silenced in response to signalling through the T- or B-cell receptor. A temporal analysis of the chromatin modifications that accompany stable silencing showed that silencing is

nucleated at the promoter by the ordered deacetylation of histone H3 Lysine 9, loss of methylation at H3 Lysine 4 and methylation at H3 Lysine 9. This was followed by repositioning of the *Dnmt* locus to pericentric heterochromatin (Brown et al., 1999) and the bidirectional spreading of repressive histone modifications (Kimura et al., 2004). The importance of histone deacetylation for initiating the assembly of silent chromatin is underscored by a requirement for Ikaros, a DNA binding factor that is essential for lymphocyte development, that interacts with HDAC-containing nucleosome remodelling NuRD complex (Koipally et al., 1999) and is required for the appropriate stage-specific silencing of many lymphocyte-associated genes (Sabbattini et al., 2001; Su et al., 2004). Remarkably, my results show that in experimental heterokaryons, inhibition of HDAC activity by low doses of TSA or VPA blocked the silencing of lineage inappropriate genes without affecting ongoing conversion to muscle. Accordingly, upon HDAC inhibition, approximately 30% of individual human nuclei in heterokaryons were seen to co-express two different lineage-associated gene programs. The RNA-FISH approach used to address gene expression in individual nuclei provides “snapshots” of local transcription at a given time point. Since gene activity has been shown to occur not continuously but in bursts of transcription (Cai et al., 2006; Golding et al., 2005; Raj et al., 2006) and there are massive variations in the number of local mRNA molecules when a gene “transits” from inactive to active state (Raj et al., 2006), it is even possible that the number of co-expressing nuclei is underestimated.

My results underscore the importance of HDACs for correctly extinguishing the gene expression program of nuclei undergoing reprogramming and conversion to muscle. To date, 18 HDACs have been identified in human and categorised in several classes based on their homology with yeast HDACs (Blander and Guarente, 2004; de Ruijter et al., 2003). Moreover, HDACs are not redundant in function, as different classes of HDACs associate with different co-repressors and activators (Bolden et al., 2006; Marks and Dokmanovic, 2005; Xu et al., 2007). Whether a specific HDAC protein(s) or family is involved in gene extinction during reprogramming of lymphocytes to muscle remain an open question. In my experiments I have used broad HDAC inhibitors such as TSA and VPA which inhibit both class I and class II HDACs by binding to their zinc containing catalytic domain (HDACs class III, or sirtuins are NAD⁺ dependent and therefore not inhibited by these compounds). However, it has been reported that class II HDACs are five times less susceptible to inhibition by VPA than class I HDACs (Gottlicher et al., 2001). Since the failure of lymphocyte-specific gene extinction was similar in TSA and VPA treated heterokaryons, it is tempting to speculate that class I HDACs may be involved in gene extinction. In the future, this could be further investigated using more selective inhibitors such as depsipeptide

(Furumai et al., 2002), SK7041 and SK7068 (Kim et al., 2003), which specifically inhibits some Class I HDACs or using RNAi-based approaches.

Collectively, these results underscore the importance of gene silencing in reprogramming but also show that the establishment of muscle gene expression and the extinction of lymphocyte identity are distinct and mechanistically separable components of lineage conversion.

CHAPTER 4. HETEROKARYON-BASED REPROGRAMMING OF LYMPHOCYTES FOR PLURIPOTENCY

4.1. Generation of inter-species heterokaryons between human B-lymphocytes and mouse embryonic stem cells

Spontaneous and experimental cell fusion of differentiated cells with pluripotent ES cells induces the expression of pluripotency-associated markers in the hybrid cells (Cowan et al., 2005; Tada et al., 2001; Ying et al., 2002), re-activation of the silent X-chromosome (Tada et al., 2001) and DNA de-methylation and chromatin remodelling at specific sites in the somatic cell genome (Cowan et al., 2005; Kimura et al., 2004). While these data show that reprogramming occurs through the epigenetic resetting of gene expression programs in the differentiated cell (Simonsson and Gurdon, 2004; Surani, 2001), it has been unclear whether genome duplication is required for conversion to pluripotency (Han et al., 2008). Here I asked whether reprogramming to pluripotency occurs in heterokaryons and investigated the requirements for the dominant reprogramming of human B cells by fusion with mouse ES cells.

The protocol used for “transient” heterokaryon formation is illustrated in Figure 4.1a. Unlike the muscle syncytia detailed in Chapter 3, embryonic stem cell lines do not generate multinucleated cells, and the resulting heterokaryons are transient rather than stable. Human B cells were fused with mouse ES cells using PEG and the nuclear events in fused cells were monitored by fluorescence microscopy and quantitative RT-PCR. To facilitate the identification of fused cells, mouse E14tg2A ES cells were pre-labelled with DiD and human B cells with DiI. The two labelled cell populations were mixed in a 1:1 ratio and inter-species heterokaryons were generated in suspension by PEG-mediated cell fusion. Dual-stained cells were purified by FACS (typically 10-15% of cells, Figure 4.1b) and cultured under conditions promoting the maintenance of undifferentiated mES cells.

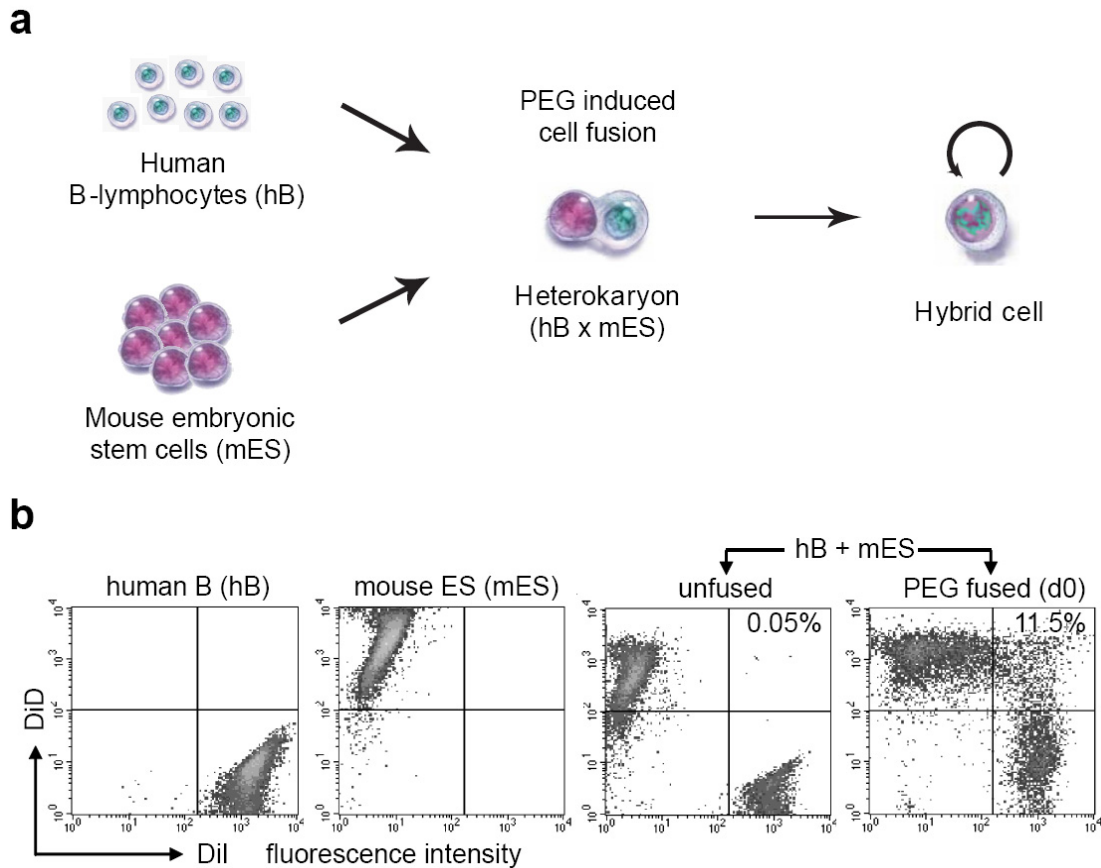


Figure 4.1. Generation of embryonic inter-species heterokaryons through cell fusion. (a) Shows the experimental strategy used to generate inter-species heterokaryons (hB x mES). Human B-lymphocytes (hB) and mouse embryonic stem cells (mES) were fused in the presence of polyethylene glycol (PEG) and monitored over time. (b) Human B-lymphocytes (hB) and mouse embryonic stem cells (mES) were respectively labelled with the cell membrane dyes DiI and DiD and fused in the presence of PEG. Fused cells (double-labelled, upper right quadrant) were FACS sorted and cultured under conditions promoting the maintenance of undifferentiated mES cells.

4.2. Reprogramming of gene expression is initiated in ES cell heterokaryons prior to nuclear fusion

Human (B cell-derived) and mouse (ES cell-derived) nuclei were distinguished on the basis of DAPI and human-specific lamin A/C labelling, and the proportion of cells containing two discrete (heterokaryons) or conjoined nuclei (hybrids) was assessed over time (Figure 4.2a). Up to 2 days following cell fusion 98 – 99% of dual labelled cells were identified as heterokaryons in which a single human and a single mouse nuclei were evident (illustrated in Figure 4.2a, central image). The kinetics of nuclear fusion were also confirmed by FISH analysis in which probes specific for mouse chromosomes (γ -satellite, red) or human chromosomes (α -satellite, green) were

used to detect nuclei containing solely human or solely mouse chromosomes or both (indicative of hybrid formation, Figure 4.2b).

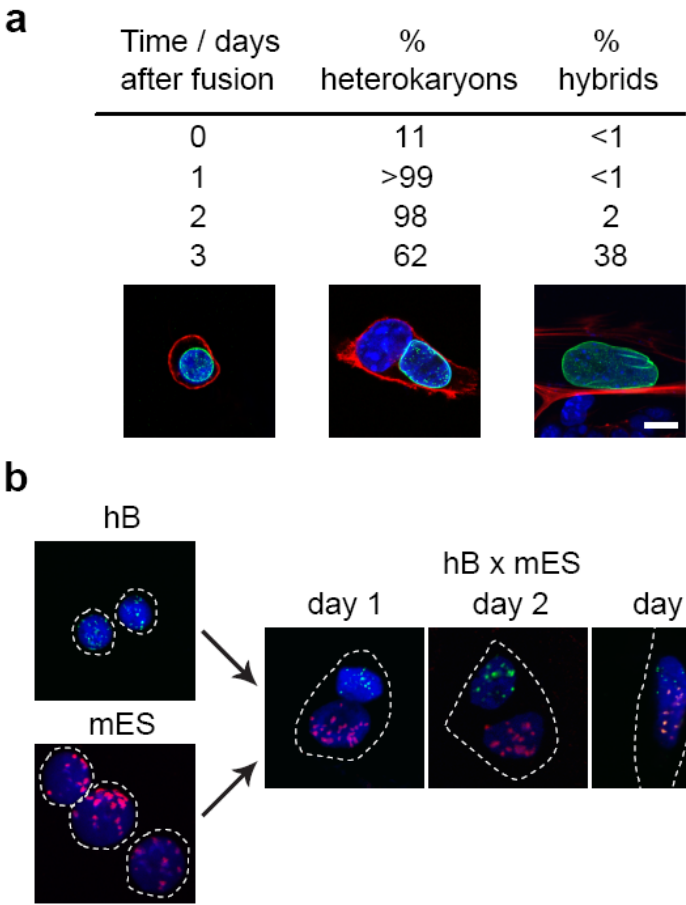


Figure 4.2. Heterokaryons can be maintained up to two days before nuclear fusion. (a) Immunofluorescence analysis of the kinetics of heterokaryon (cells in which parental nuclei share the same cytoplasm but remain discrete) and hybrid formation (where both parental genomes occupy the same nucleus) following PEG-induced fusion. hB-derived nuclei were distinguished from mouse nuclei on basis of DAPI (blue) and human Lamin A/C staining (green), and actin staining (red) delineates individual cells. In the lower panels confocal sections showing a hB cell prior to fusion (left, day 0), a heterokaryon [one mouse (with DAPI intense foci) and one human nucleus (hLamin A/C positive)] (middle, day 2) and a hybrid cell (right, day 3) are shown. $n = 100$ (b) Mouse and human nuclei were distinguished by FISH using probes specific for mouse γ -satellite DNA (red) or human α -satellite DNA (green) along with DAPI (blue). Confocal sections of human B cells (hB) and mouse ES cells (mES) before and after cell fusion (hB x mES) are shown. Heterokaryons were identified up to 2 days after fusion, but by day 3 hybrid formation was detected. Scale bars, 10 μm .

The expression of pluripotency-associated and lymphocyte-associated genes by human B cell-derived nuclei was assessed by RT-PCR, using primers that were designed to selectively amplify each of the human transcripts. This analysis showed that *hOct4* and *hNanog* transcripts arise from hB-derived nuclei, two days after heterokaryon formation (Figure 4.3a, left panel). The establishment of a new gene expression program by hB nuclei was accompanied by a reduction in expression of the lymphocyte-associated genes *hCD20* and *hCD37* (Figure 4.3a, right panel).

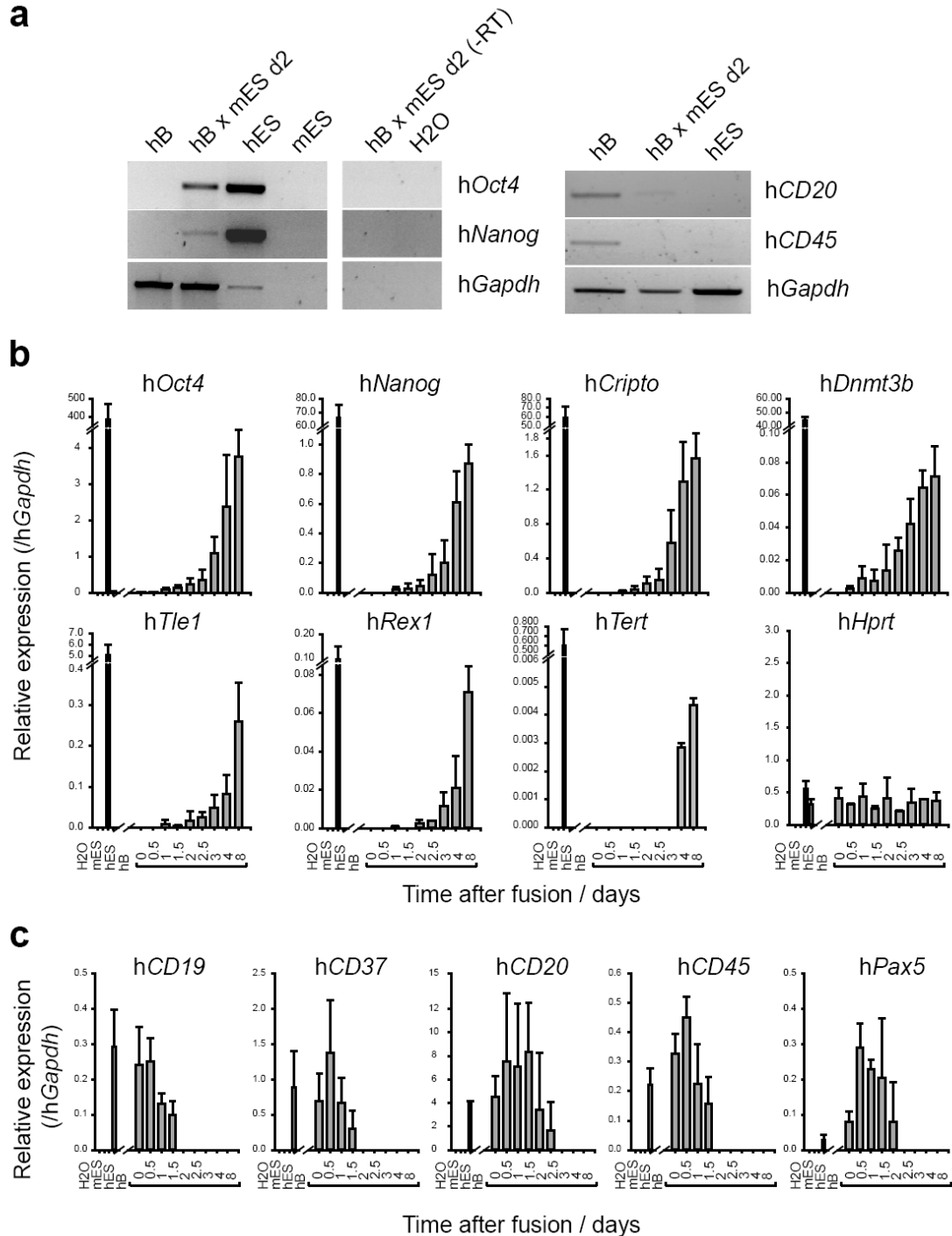


Figure 4.3. Reprogramming of human B-lymphocytes to embryonic state by mouse ES cells does not require nuclear fusion. (a) Expression of human ES-specific (*hOct4*, *hNanog*) and human lymphocyte-specific (*hCD20*, *hCD45*) transcripts detected by RT-PCR using human-specific primers. Prior to fusion, hB cells expressed *hGapdh*, *hCD20* and *hCD45* but not embryonic stem cell-specific genes. Following heterokaryon formation (hB x mES d2), human pluripotency-associated genes *hOct4* and *hNanog* were expressed (left panel) and *hCD20* and *hCD45* were extinguished (right panel). mES, -RT and H₂O were used as negative controls and human embryonic stem cells (hES) as a positive control. *hGapdh* was used to standardise input. (b) The expression of hES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hTle1*, *hRex1* and *hTert*) was assessed by quantitative RT-PCR analysis 0 to 8 days after cell fusion. Positive (hES-NCL1, black bars) and negative (hB, mES) controls for this analysis were included. The constitutively expressed gene *hHprt* remains detectable at identical levels at all time points. (c) Activation of embryonic genes is accompanied by silencing of lymphocyte-specific genes (*hCD19*, *hCD37*, *hCD20*, *hCD45* and *hPax5*). Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

The kinetics of gene activation was then assessed by qRT-PCR with primers recognising a panel of human ES-associated gene products, including transcription factors, DNA modifiers and signalling molecules. The expression of human *Oct4*, *Nanog*, *Cripto*, *Dnmt3b* and *Tle1* was detected in cells as early as 1 day after fusion and human *Rex1* after 2 days (Figure 4.3b). Expression of *hTert* was detected from day 4 onwards (Figure 4.3b). Although the expression levels were low in heterokaryons (<1% of that detected in human ES cells, cell line NCL1), these increased over time and were undetectable in non-fused (or self-fused, not shown) human B cells or control mouse ES cells. Mouse lymphocyte-specific gene transcripts (*mCD19*, *mCD37* and *mCD45*) were not detected throughout the analysis (not shown), confirming ES cell dominance in conversion (Cowan et al., 2005; Tada et al., 2001). The increased expression of human pluripotency-associated genes over this 8-day period was mirrored by a reduction in expression and extinction of several human lymphocyte-associated genes within the second (*hCD45*, *hCD37* and *hCD19*) or third day (*hCD20* and *hPax5*) of heterokaryon formation (Figure 4.3c). *hHprt* expression was equivalent at all stages, as anticipated (Figure 4.3b). Collectively these data show that upon dominant reprogramming, activation and silencing of tissue-specific gene programs begins ahead of, and therefore does not require, nuclear fusion and cell division. In addition, since these results examine gene expression at the population level, it is possible that gene expression varied between individual heterokaryons and hybrid cells.

As the reprogramming of somatic cell has been previously shown to result in altered DNA methylation at specific *loci* (Cowan et al., 2005; Freberg et al., 2007; Simonsson and Gurdon, 2004; Tada et al., 2001), we examined changes in the methylation status of the human *Oct4* gene promoter (Deb-Rinker et al., 2005) and as a control, the *Igf2/H19* imprinting control region (ICR) (Ulaner et al., 2003). As illustrated in Figure 4.4, human B cells prior to fusion showed high levels of DNA methylation throughout the *hOct4* promoter and across a single *Igf2/H19* allele. Following cell fusion, DNA methylation of *hOct4* in reprogrammed B cells declined, consistent with a trend towards a hypomethylated state as seen in the human ES cell line H1. Demethylation of the *hOct4* promoter was rapid, detected prior to nuclear fusion and cell division (heterokaryons at day2) and maintained throughout the experiment (d2-d8), a result that is consistent with active chromatin remodelling of the *locus* prior to expression. In addition, a mosaic pattern of demethylation was observed (some DNA clones completely demethylated while others remain largely methylated) suggesting that only a proportion of heterokaryons are being successfully reprogrammed. This pattern of demethylation was also observed in human 293T cells

upon exposure to ES-cell extract (Freberg et al., 2007) at the human *Oct4* and *Nanog* promoter regions. No changes in DNA methylation at *Igf2/H19* ICR were detected over this period, consistent with its imprinted status and the lack of ability of ES cells to erase imprints (Tada et al., 2001; Ulaner et al., 2003).

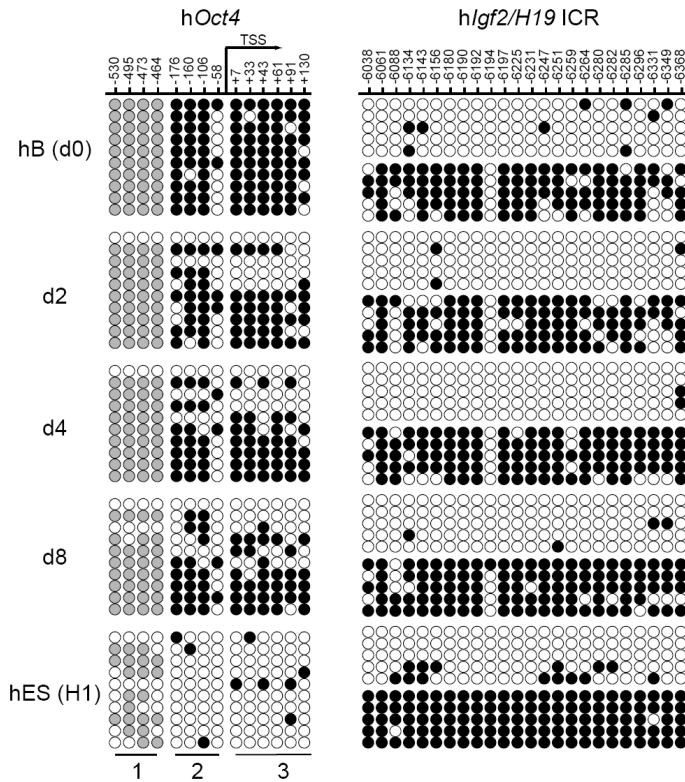


Figure 4.4. The human *Oct4* promoter is demethylated in heterokaryons. Bisulfite genomic sequencing analysis of DNA methylation at the human *Oct4* promoter 0, 2, 4 and 8 days after cell fusion with mES cells. Human ES cells (hES, cell line H1; lower panel) are shown as controls. The methylation pattern of *Igf2/H19* imprinting control region (ICR) was analysed as a control. The position of CpG sites relative to the transcriptional start site (TSS) is indicated. Open circles represent unmethylated cytosines, black closed circles represent methylated cytosines and gray closed circles represent constitutively methylated cytosines. Region 1, 2 and 3 indicate CpG sites that are part of the same PCR product.

4.3. Induction of a human ES-specific gene expression profile and multi-lineage potential

A comparison of the relative abundance of gene-specific transcripts in reprogrammed human B cells (Figure 4.5a, right-hand column), showed a strong similarity with the gene expression profiles of several human ES cell lines [NCL1 (Stojkovic et al., 2004), H1, H7, H9 (Thomson et al., 1998); Figure 4.5a, left-hand column]. For example, while *Oct4* was abundantly expressed in all human and mouse ES cell lines, *Nanog* and *Cripto* expression was consistently much lower than *Oct4* (100-1000 fold) for each of the mouse ES cell lines analysed (OS25, CCE, E14, ZHBTc4; Figure 4.5a, middle panel). In human ES cell lines however, *Oct4*, *Nanog* and *Cripto* transcripts were similarly abundant, a result that was mirrored in reprogrammed

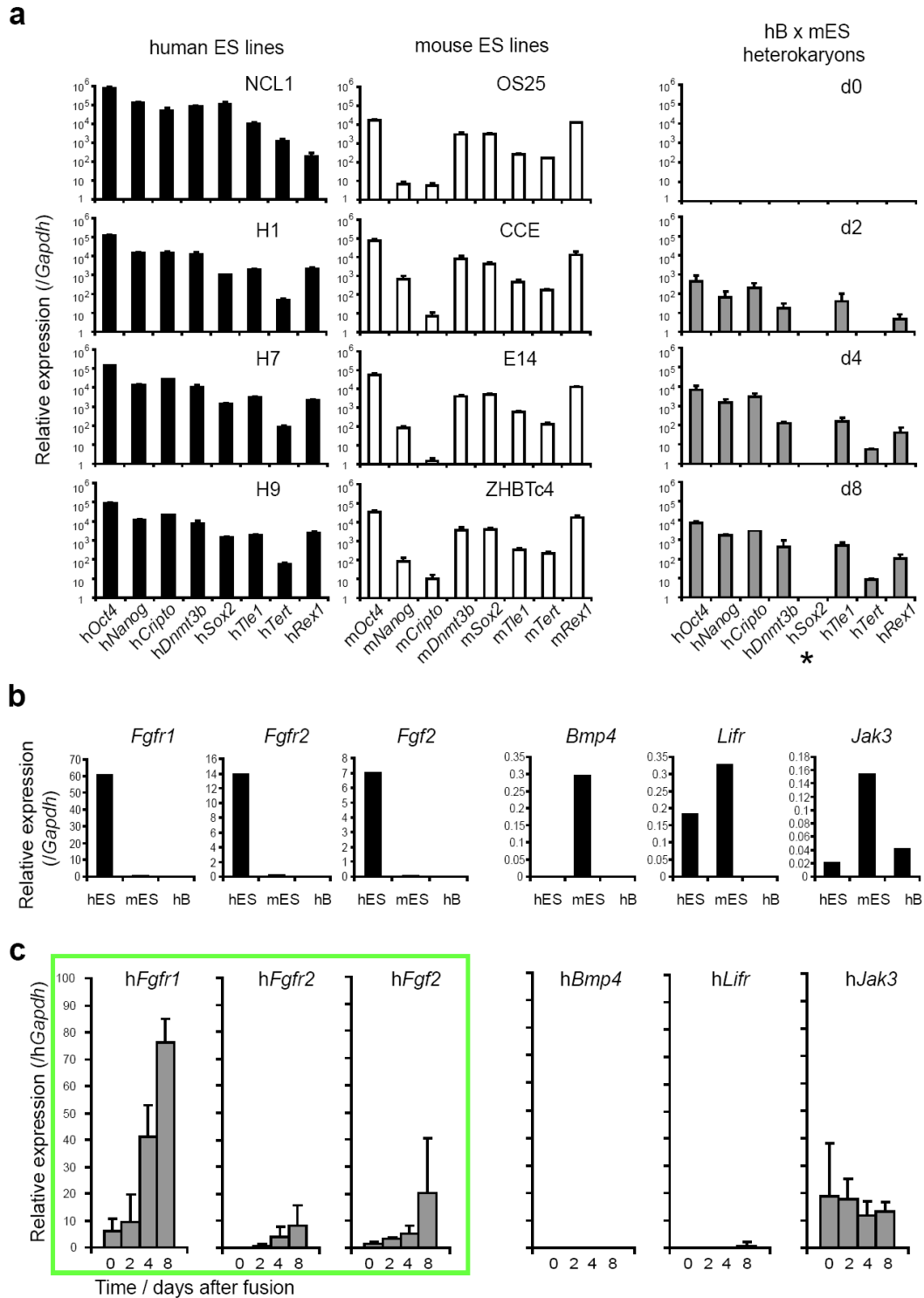


Figure 4.5. Reprogrammed lymphocytes are comparable to human ES cells rather than mouse ES cells. (a) Quantitative RT-PCR analysis of the relative levels of gene expression in several human (NCL1, H1, H7 and H9), mouse (OS25, CCE, E14 and ZHBTc4) ES cell lines and in reprogrammed lymphocytes (hB x mES) at 0, 2, 4 and 8 days after cell fusion. Human ES lines (left panel) and hB x mES (right panel) gene expression data was normalised to h*Gapdh*. Mouse ES lines (middle panel) gene expression data was normalised to m*Gapdh*. Error bars indicate the s.d. of 3-4 independent experiments. (b) Expression of *Fgfr1*, *Fgfr2*, *Fgf2*, *Bmp4*, *Lifr* and *Jak3* was assessed by qRT-PCR in human ES cells (hES, NCL1), mouse ES cells (mES) and human B-lymphocytes (hB). *Fgfr1*, *Fgfr2* and *Fgf2* were uniquely expressed by human ES cells. (c) After cell fusion, the expression of genes involved in the maintenance of undifferentiated human ES cells (h*Fgfr1*, h*Fgfr2* and h*Fgf2*; highlighted with green box) and genes selectively expressed by mouse ES cells (h*Bmp4*, h*Lifr* and h*Jak3*) was analysed by qRT-PCR. Data were normalised to h*Gapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

human B cells. Expression of some pluripotency-associated genes, for example *Sox2*, was variable and often required more time (>8 days) to be detected in hB x mES cultures (not shown).

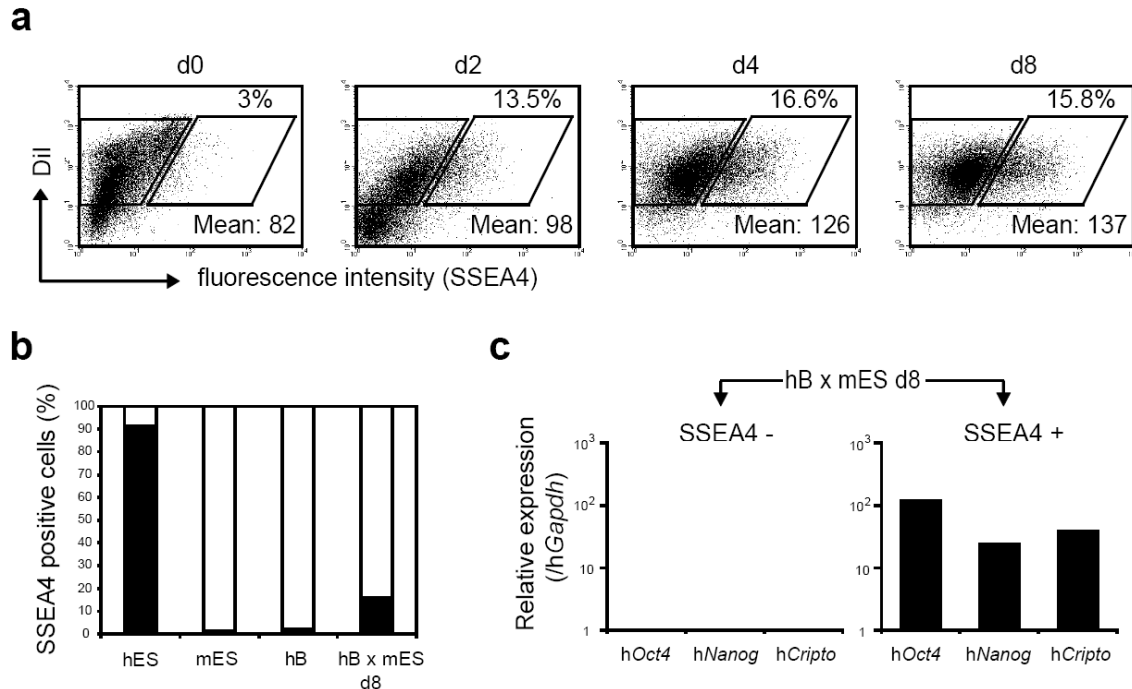


Figure 4.6. Reprogrammed cells express the human ES marker SSEA4. (a) Heterokaryons resulting from human B cell and mouse ES cell fusions (hB x mES) were stained for SSEA4 at 0, 2, 4 and 8 days and expression was analysed by flow cytometry. The results showed that 13.5% (day 2), 16.6% (day 4) and 15.8% (day 8) of total heterokaryons expressed SSEA4, as delineated by the rhomboid gates. Mean intensity fluorescence of positive cells is indicated. (b) FACS analysis showed that >90% of hES cells (H1 cell line) expressed SSEA4, while hB and mES did not (2.1% and 1.5% respectively). A proportion of heterokaryons showed SSEA4 expression (15.8%) 8 days after cell fusion (hB x mES d8). (c) FACS sorting of SSEA4 positive cells co-purifies reprogrammed cells that express *hOct4*, *hNanog* and *hCripto*, as assessed by qRT-PCR. Data were normalised to *hGapdh* expression.

Similarities between gene expression profiles of human ES cell lines and hB x mES fused cells prompted us to examine additional markers that are expressed solely by either human or mouse ES cells (Ginis et al., 2004; Pera and Trounson, 2004; Sato et al., 2003; Xu et al., 2005). These included fibroblast growth factor receptors (*Fgfr1* and *Fgfr2*) and *Fgf2* (expressed by human ES cells, Figure 4.5b), LIF receptor (*Lifr*) and *Bmp4* (expressed by mouse ES cells) and SSEA4, a surface glycoprotein selectively expressed by human ES cells (Thomson et al., 1998). This analysis revealed that reprogrammed cells expressed increasing amounts *hFgfr1*, *hFgfr2* and *hFgf2*, but did not express *hBmp4* or *hLifr* or upregulate the downstream kinase *hJak3* (Figure 4.5c). Thus, these data show that although dominant conversion is driven by mouse ES cells that express *Bmp4* and *Lifr* prior to fusion (Figure 4.5b), reprogrammed heterokaryons and hybrid cells have a remarkably different expression profile that resembles human, rather than mouse ES cell lines. Consistent with this, fusion of

mouse ES cells and human B cells resulted in SSEA4 expression by a proportion of cells (13-16%, days 2-8 as shown in Figure 4.6a and 4.6b) and the isolation of SSEA4-positive cells co-purified successfully reprogrammed cells that express *hOct4*, *hNanog* and *hCripto* (Figure 4.6c), while SSEA4-negative cells were not reprogrammed. The observation that only a proportion of heterokaryons are successfully reprogrammed, as judged by *hOct4* DNA demethylation and SSEA4 expression, might partly explain why the levels of transcripts encoding pluripotency factors are lower in reprogrammed cultures than established hES cell lines.

To ask whether the reprogramming of human B cells by mouse ES cells resets multi-lineage potential, hB x mES cultures were treated with retinoic acid (RA) 6-8 days after cell fusion in order to induce differentiation. Prior to RA treatment, hybrid colonies expressed several pluripotency-associated markers, including *hNanog* protein (detected using a human Nanog-specific antibody) and the human embryonic-specific antigens SSEA4, TRA-1-60 and TRA-1-81 (Thomson et al., 1998) (Figure 4.7).

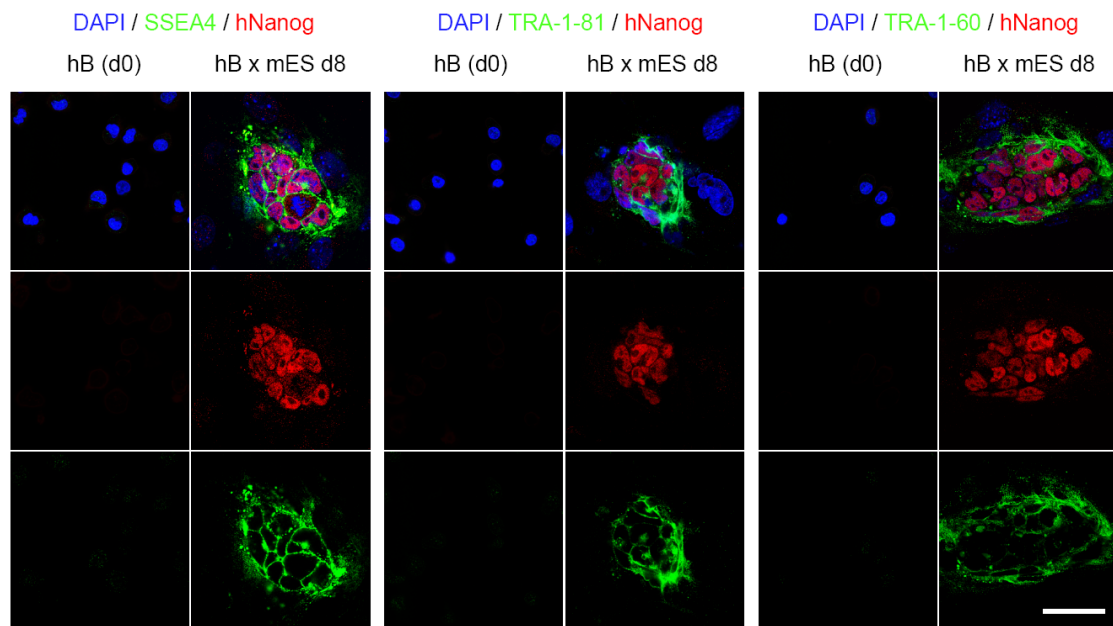


Figure 4.7. Expression of human-specific embryonic markers in hybrid cells. Human B cells (hB) and mouse ES cells (mES) were fused and the resulting colonies (hB x mES, day 8) were analysed by immunofluorescence. Detection of *hNanog* protein (red) and the human ES-specific antigens SSEA4, TRA-1-81 and TRA-1-60 (green) with human-specific antibodies is shown. Control hB cells did not express any of the markers. DAPI staining is shown in blue. Images are single confocal sections. Scale bar, 50 μ m.

In addition, most cells in hB x mES colonies showed alkaline phosphatase (AP) activity (Thomson et al., 1998) (Figure 4.8a), as well as the expression of human AP transcripts (*hALPL*, Figure 4.8b). Following treatment with RA, AP activity and expression of *hOct4*, *hNanog* and *hRex1* was reduced (Figure 4.8a and 4.8c), while morphological heterogeneity within colonies increased.

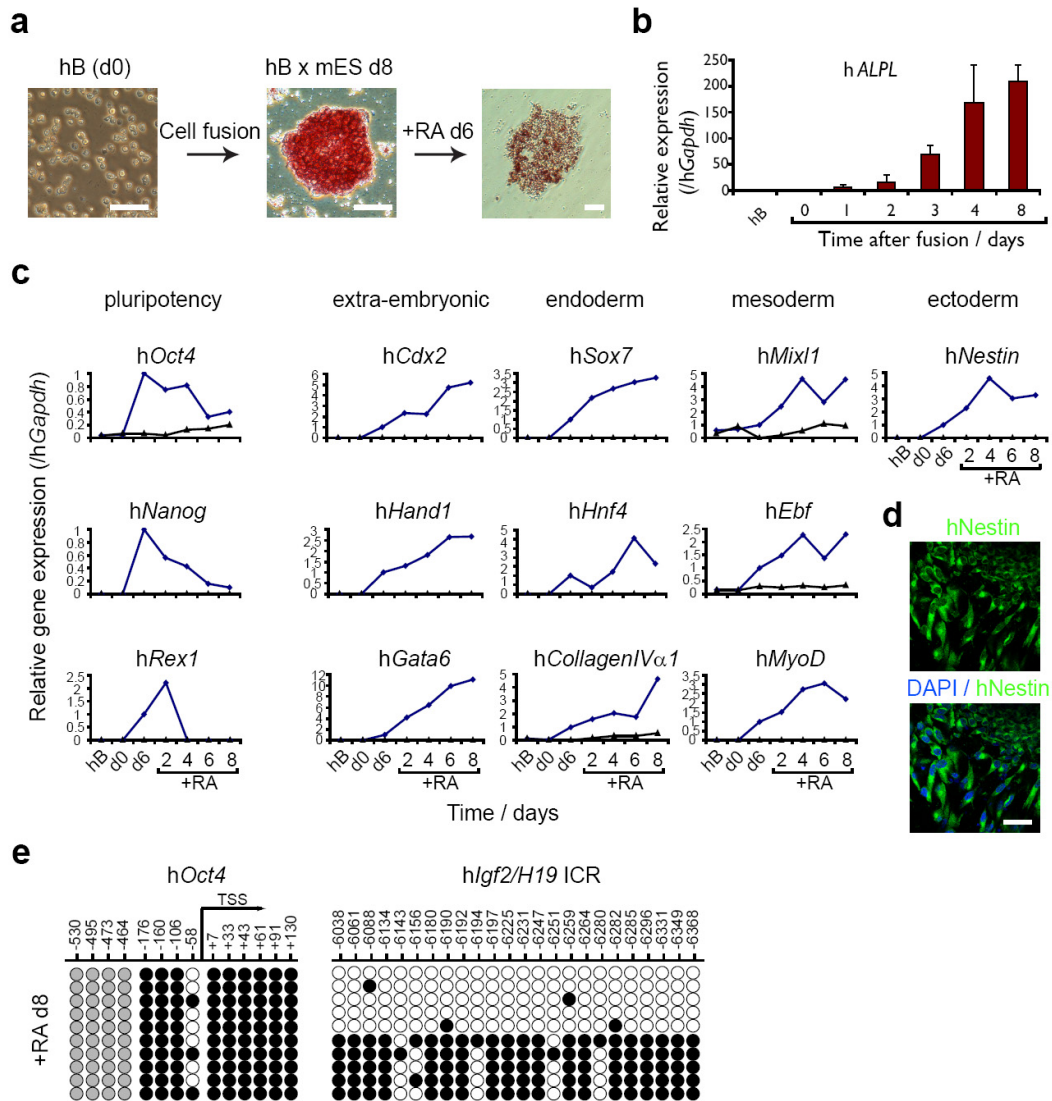


Figure 4.8. Multi-lineage potential is reset in reprogrammed human lymphocytes. (a) Alkaline phosphatase activity assay of hybrid colonies (hB x mES, day 8) resulting from fusion of human B cells (hB) and mouse ES cells. Hybrid colonies showed alkaline phosphatase activity (pink) that was reduced upon retinoic acid treatment (+RA, day 6). Images are bright field pictures. (b) qRT-PCR analysis of human alkaline phosphatase (hALPL) gene expression upon cell fusion. (c) qRT-PCR analysis of gene expression upon RA treatment of hybrid cells (blue line) showed that levels of pluripotency genes (hOct4, hNanog and hRex1) declined while differentiation-associated genes were upregulated [extra-embryonic (hCdx2, hHand1 and hGata6), endoderm (hSox7, hHnf4 and hCollagenIV α 1), mesoderm (hMixl1, hEbf and hMyoD) and ectoderm (hNestin)]. Unfused hB cells were included as controls (black line). Data were normalised to hGapdh expression. (d) Immunostaining of ectodermal differentiating colonies 6 days after RA exposure with anti human-specific Nestin antibody (green). Nuclei were visualised with DAPI staining (blue). Scale bars, 50 μ m. (e) Bisulfite genomic sequencing analysis of DNA methylation at the human Oct4 promoter and the Igf2/H19 imprinted control region (ICR) 8 days after RA treatment. The position of CpG sites relative to the transcriptional start site (TSS) is indicated. Open circles represent unmethylated cytosines, black closed circles represent methylated cytosines and gray closed circles represent constitutively methylated cytosines.

Detection of extra-embryonic (hCdx2, hHand1 and hGata6), endoderm (hSox7, hHnf4 and hCollagenIV α 1), mesoderm (hMixl1, hEbf and hMyoD) and ectoderm (hNestin) associated transcripts in RA treated hB x mES cultures, but not in RA treated hB cells (Figure 4.8c, blue and black lines respectively), confirmed a selective response of reprogrammed cells to multiple differentiation pathways. In addition, cells expressing

hNestin protein were detected (using an anti-human Nestin-specific antibody) at the edge of colonies after RA treatment (Figure 4.8d). Differentiation also resulted in increased DNA methylation of the *hOct4* promoter (Figure 4.8e), similar to levels seen in differentiated human B cells (Figure 4.4). Taken together, these results show that reprogramming of human B-lymphocytes by mouse ES cells resets gene expression and multi-lineage potential.

4.4. Genes not easily re-activated in heterokaryons display heterochromatic features

The observation that genes such as *hSox2* and *hTert*, that were not re-activated in day-2 heterokaryons, could reflect the fact that they require a higher threshold of activators for overt expression, or that these genes are subject to multiple layers of repressive epigenetic modifications in B cells, including DNA and histone methylation (Jenuwein and Allis, 2001; Lande-Diner and Cedar, 2005) and late replication (Gilbert, 2002). This possibility prompted us to examine the epigenetic status of these and additional *loci* in hB-lymphocytes prior to cell fusion (results summarised in Figure 4.9).

My results showed that many silent endogenous genes in hB cells can be re-activated in heterokaryons at day 2 (*hOct4*, *hDppa5*, *hCripto*, *hDnmt3b*, *hNanog*, *hGdf3*, *hGbx2* and *hTle1*; Figure 4.9c, white boxes) albeit at low levels (*hFoxd3* and *hRex1*; Figure 4.9c, grey boxes). However, expression of some genes was not detected (*hTert*, *hSox2* and *hDppa4*; Figure 4.9c, black boxes). I assessed the replication timing of these genes in hB cells, as there is a correlation between replication timing and heritably repressed chromatin (Gilbert, 2002). The timing that each *locus* replicates during S-phase was analysed by a PCR-based method (Azucara et al., 2006) and genes ordered them from the earliest (*hOct4*, green) to the latest (*hRex1*, red) replicating gene (Figure 4.9a; see Appendix II, Figure A1 for detailed analysis). Four out of five ES-associated genes analysed (*hSox2*, *hFoxd3*, *hDppa4* and *hRex1*) that were not easily re-activated in heterokaryons were among the late-replicating cohort of genes (Figure 4.9, highlighted by the red open box). This subset of genes was also preferentially located within isochores with low GC content, high frequency of LINE elements (Figure 4.9d, upper panel) and gene poor regions (Figure 4.9d, lower panel), as previously reported (Hiratani et al., 2004).

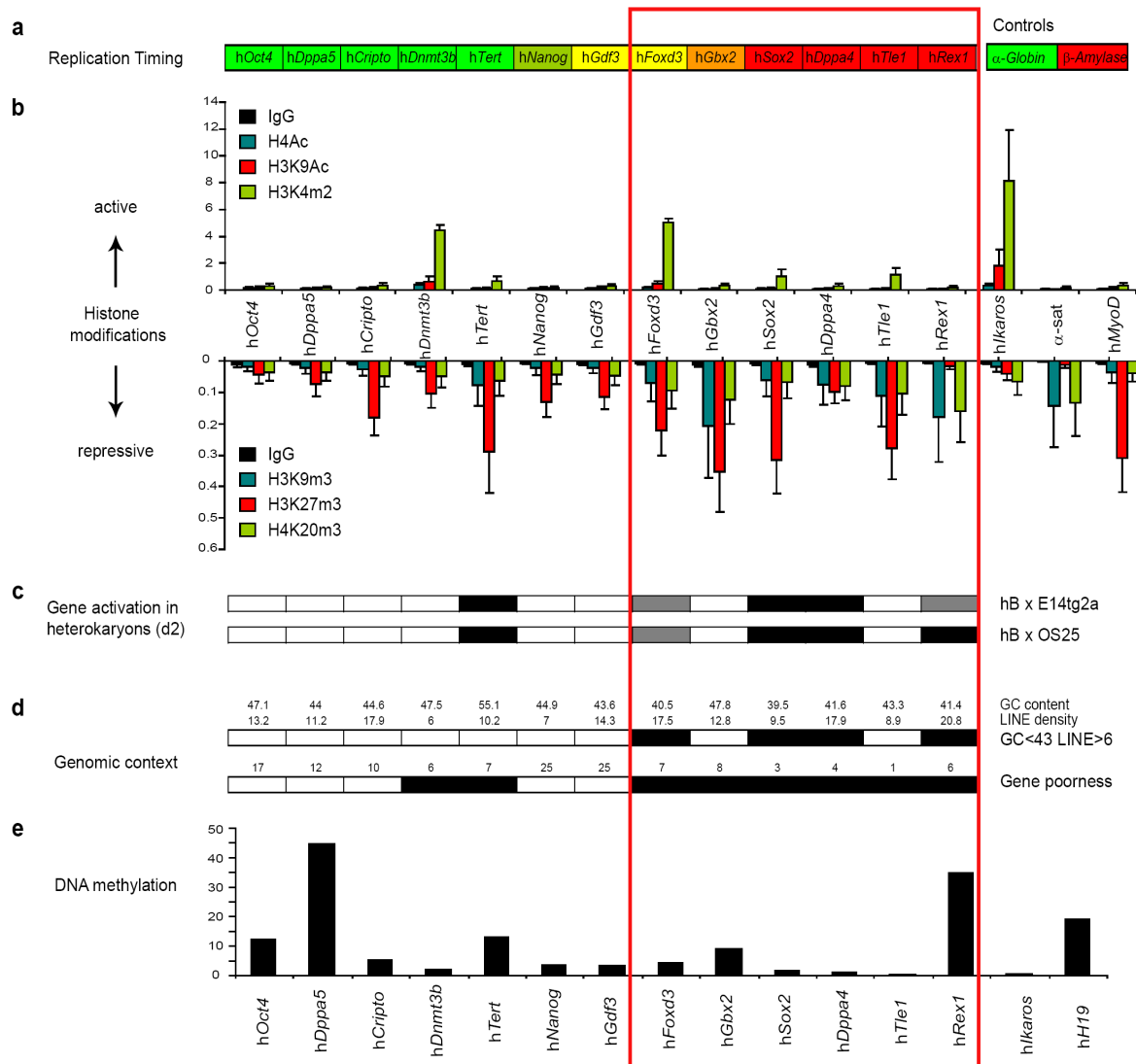


Figure 4.9. Some genes not easily re-activated in heterokaryons show heterochromatic features. (a) The replication timing of each gene (hOct4, hDppa5, hCripto, hDnmt3b, hTert, hNanog, hGdf3, hFoxd3, hGbx2, hSox2, hDppa4, hTle1 and hRex1) was defined accordingly to its peak abundance in G1-S1 phase (early, green), S2 phase (middle-early, dark green), S2-S3 phase (middle, yellow), S3 phase (middle-late, orange) or S4-G2 (late, red), determined in two independent experiments using a PCR-based assay (Azuara et al., 2006) in hB-lymphocytes. Genes are ordered according to their replication timing. α -Globin and β -Amylase were included as control genes. **(b)** The abundance of active (H4Ac, H3K9Ac and H3K4me2) and repressive (H3K9m3, H3K27m3 and H4K20m3) histone modifications was assessed in hB-lymphocytes using chromatin immunoprecipitation and qPCR. The colour code for different histone marks is indicated. Controls for active (hIkaros) and repressive (α -satellite and hMyoD) histone modifications were included. The enrichment indicates the ratio of modified H3 to unmodified H3 detected at each promoter region normalised to input DNA. Error bars indicate the s.d. of 3 independent experiments. **(c)** Heterokaryons were generated by cell fusion of mouse ES cells (E14tg2A or OS25 cell line) with hB-lymphocytes and each of the human ES-associated transcripts was quantified by qRT-PCR two days after fusion. Detectable (white boxes), detectable at low levels (grey boxes) and not detectable (black boxes) genes were categorised based on three independent experiments. **(d)** The genomic context of each gene was analysed and GC content, long interspersed nuclear elements (LINE) and gene density are indicated (see material and methods for details). In the upper panel black boxes highlight region which GC content is less than 43% and LINE density greater than 6% (Hiratani et al., 2004). In the lower panel, black boxes indicate gene poor regions which contain less than 10 genes in the surrounding 500Kb. **(e)** The abundance of CpG DNA methylation at each promoter region was assessed in hB-lymphocytes using methyl-DNA immunoprecipitation (MDIP) (Weber et al., 2005) and qPCR. Controls for unmethylated (hIkaros) and methylated (hH19; hH19/Igf2 imprinting control region) regions were included. The enrichment indicates the ratio of methylated DNA over input DNA. The red open box highlights the late replicating cohort of genes which includes 4 out of 5 not easily re-activated genes in heterokaryons.

Furthermore, active histone modifications (H4Ac, H3K9Ac and H3K4me2; Figure 4.9b, upper panel) as assessed by chromatin immunoprecipitation were absent from the majority of these *loci* (with exception of hDnmt3b and hFoxd3) consistent with transcriptional repression of these genes in hB-lymphocytes. Repressive histone modifications (H3K9m3, H3K27m3 and H4K20m3; Figure 4.9b, lower panel) were also enriched preferentially at these late replicating genes. The *hTert* gene, however, showed high levels of H3K27m3 despite being early-replicating. This may provide an explanation for the lack of *hTert* expression in heterokaryons. Finally, no correlation was observed between the re-activation potential and the abundance of DNA methylation at the promoter regions of the genes (Figure 4.9e) as assessed by methyl-DNA immunoprecipitation (MDIP) (Weber et al., 2005). These results suggest there is a tendency for genes that replicate later in S-phase, with repressive histone modifications, being more difficult to re-activate in heterokaryons. However some genes, such as *hTle1* and *hGbx2*, rapidly reactivated in heterokaryons, are late-replicating and highly enriched for repressive histone marks. This suggests that while the epigenetic status of genes influences their re-expression potential, this is not the only constraint to activate genes in heterokaryons. The right combination and/or the levels of “*trans*” acting factors are also likely to have an important role in conveying ES-specific gene expression.

4.5. Inter-species reprogramming of human B cells requires mOct4 but not mSox2

Oct4 is part of the core regulatory circuitry in ES cells (Boyer et al., 2005) and it is essential for pluripotency and self-renewal (Niwa et al., 2000). To assess the potential role of mouse-derived Oct4 as a dominant “*trans*” acting factor within inter-species heterokaryons, we generated ES cells expressing Flag-tagged mouse Oct4 protein (Figure 4.10a) and fused these with human B lymphocytes (Figure 4.10b). Flag-tagged Oct4 protein (derived from mouse ES cells) was seen to accumulate within human nuclei 3 to 6 hours after cell fusion (Figure 4.10b; complete kinetics shown in Figure 4.10c). In addition, Oct4 protein was present in human heterokaryon nuclei by 3 hours, before transcription of endogenous *hOct4* was initiated (at 24 hours). Thus the translocation of the ES-derived Oct4 protein into human lymphocyte nuclei occurs prior to overt reprogramming. This experiment was done in collaboration with Kelly Morris.

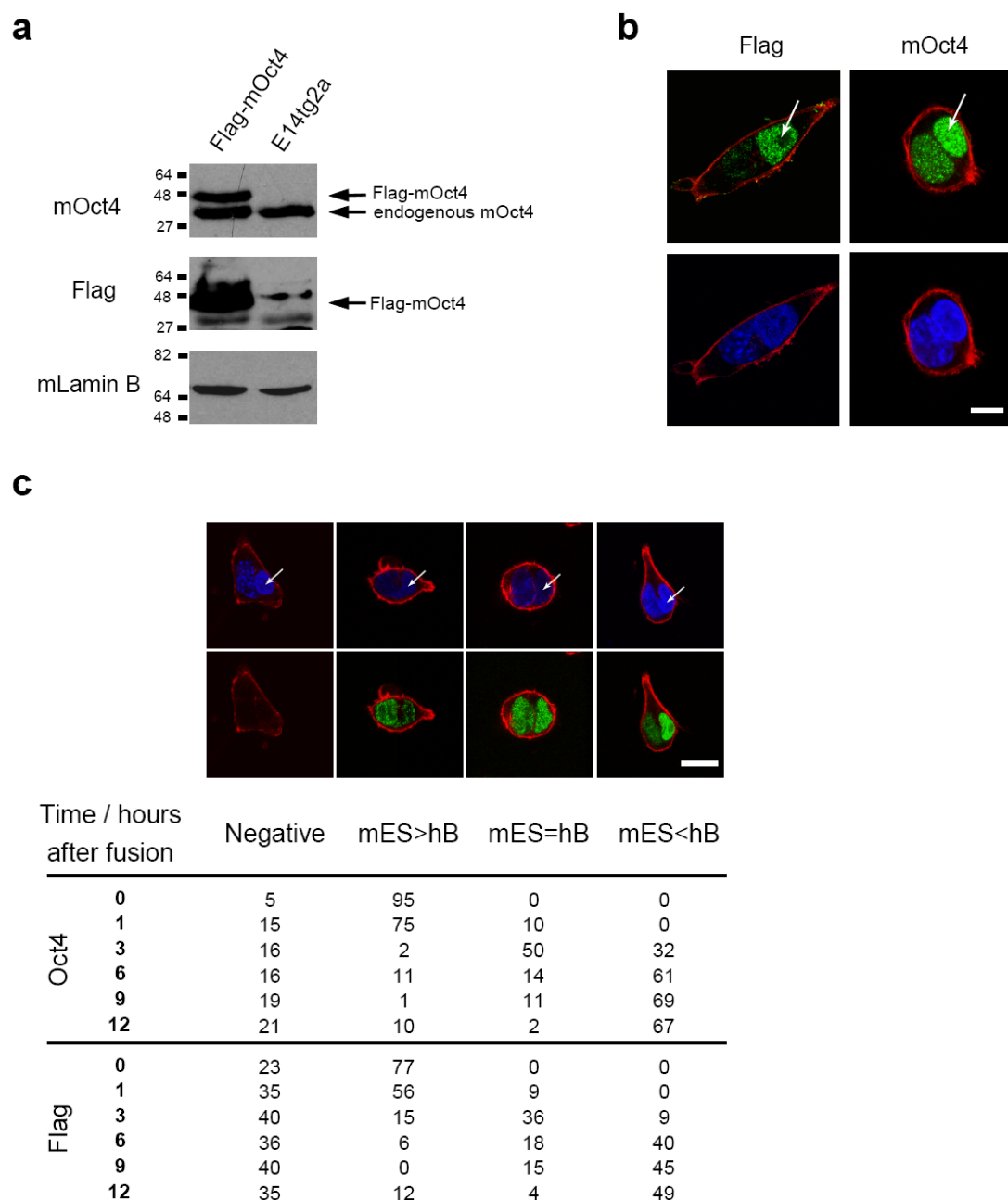


Figure 4.10. Oct4 migration to human nuclei precedes reprogramming. (a) Mouse ES cells expressing a tagged Oct4 protein (Flag-mOct4) were generated by insertion of Flag-tagged mouse *Oct4* cDNA in E14tg2A ES cells (parental cell line). Western blotting with anti-Oct4 and anti-Flag antibodies confirmed the presence of Flag-tagged Oct4 protein by transduced cells. Equivalent protein loading is shown with mouse Lamin B detection. (b) Flag-mOct4 ES cells were fused to hB cells and immunofluorescence analysis of cultured heterokaryons 6 hours after cell fusion showed the presence of ES cell-derived Oct4 (Flag-Oct4, green) in a human nucleus (arrowed). Human nuclei were distinguished from mouse nuclei on basis of diffuse versus punctuate DAPI staining (blue), respectively. Actin labelling (red) delineates the cell membrane. Images are confocal sections of heterokaryons containing a single mouse (with DAPI intense foci) and a single human nucleus. (c) Immunofluorescence kinetic analysis of cultured heterokaryons at 0, 1, 3, 6, 9 and 12 hours with Oct4 or Flag antibodies (green). Heterokaryons were scored according to the following Oct4 distribution: Oct4 protein not detected (Negative), stronger staining in mES-derived nucleus than hB nucleus (mES>hB), nuclei equally labelled (mES = hB), stronger in the human nucleus (mES<hB). Confocal sections of representative heterokaryons from each of the categories are shown (upper panels). Human nuclei were distinguished from mouse nuclei on basis of diffuse versus punctuate DAPI staining (blue), respectively. Actin labelling (red) delineates the cell membrane. Scale bars, 10 μ m. n = 100.

Reprogramming of human fibroblasts to ES-like cells has been shown to require the activation of at least four factors including Oct4, Sox2 and either Nanog plus Lin28 (Yu et al., 2007) or Klf4 plus c-Myc (Masui et al., 2007). Recently it was shown that mouse ES cells that lack Sox2, a factor thought to be vital for preventing extra-embryonic differentiation, can remain pluripotent if provided with additional Oct4 (Masui et al., 2007). To investigate the relative importance of Oct4 and Sox2 in reprogramming, mouse ES cells that are inducible *null* (Tet-off) for mOct4 [ZHBTc4 (Niwa et al., 2000)] or for mSox2 [2TS22C (Masui et al., 2007)] were used as fusion partners with human B cells. These inducible null ES cell lines were kindly provided by H. Niwa and have been characterised previously (Masui et al., 2007; Niwa et al., 2000). Both cell lines display a rapid (within 24 hours) and complete elimination of Oct4 or Sox2 gene/protein expression upon doxycycline (+Dox) treatment. In my hands, pre-treatment of ZHBTc4 cells with Dox for 6 and 12 hours, resulted in a progressive decrease in mOct4 gene expression (Figure 4.11b), without significantly affecting the expression of several other pluripotency-associated genes or the efficiency with which the ES cells fuse with human B cells (Figure 4.11a). Successful reprogramming, as judged by induction of several human genes (*Oct4*, *Nanog*, *Cripto*, *Dnmt3b*, *Sox2*, *Tle1*, *Tert* and *Rex1*) was however reduced (+6 hours) or eliminated (+12 hours) by pre-treatment of ZHBTc4 cells with Dox (Figure 4.11c, a complete kinetic analysis is provided in Appendix II, Figure A2). This result showed that mOct4 expression is critically important for initiating successful reprogramming, in keeping with previous reports (Maherali et al., 2007; Masui et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007; Yu et al., 2007). The extinction of human lymphocyte-specific genes was however not impaired by Oct4 removal (Appendix II, Figure A2), a result that may support the previous findings that the activation and silencing of gene expression programs in heterokaryons are mechanistically distinct processes (Chapter 3).

Eliminating mSox2 expression in mouse ES cell (Figure 4.12a, 2TS22C cell line) had, in contrast, a relatively mild effect on reprogramming efficiency (Figure 4.12b, compare values at 0, 12 and 24 hours of Dox treatment). Furthermore, reprogramming was fully restored in fusions using 2O1 cells, a Sox2-deficient mES cell line in which mOct4 expression is up-regulated (Masui et al., 2007) (Figure 4.12b, values shown in red and complete kinetics shown in Appendix II, Figure A3).

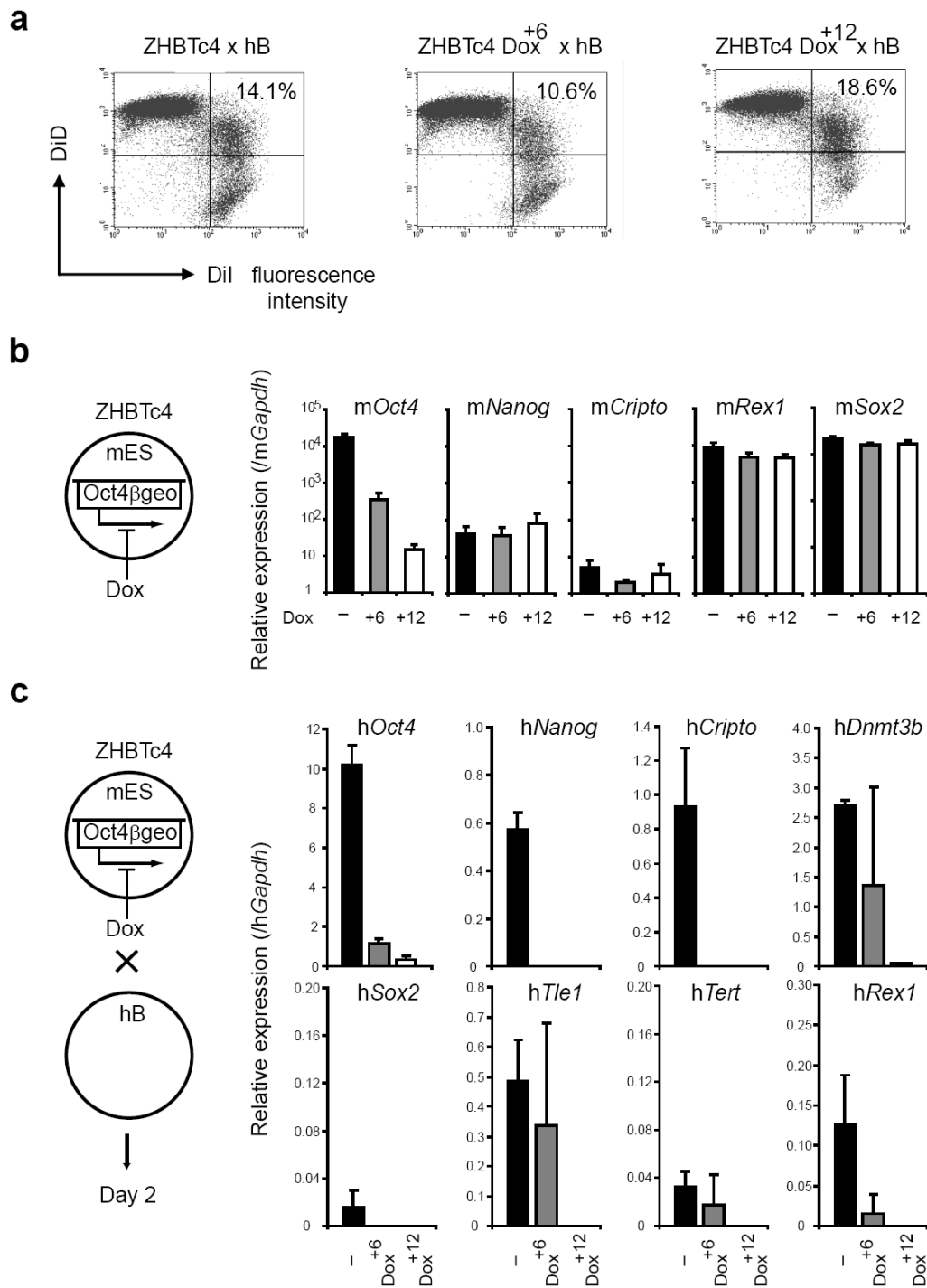


Figure 4.11. Oct4 is required for successful reprogramming. (a) ZHBTc4 [untreated or treated with doxycycline (Dox) for 6 and 12 hours] and hB cells were labelled (with DiD and DiI, respectively) and PEG-fused. Fusion efficiencies (14.1%, 10.6% and 18.6%) were obtained by FACS, as a percentage of double-labelled cells. (b) In ZHBTc4 ES cells endogenous *Oct4* was replaced by an inducible transgene (*Oct4βgeo*) which can be downregulated by addition of Dox (Niwa et al., 2000). Quantitative RT-PCR analysis showed that 6 hours (+6) and 12 hours (+12) after Dox treatment, *mOct4* was progressively downregulated, while expression of other pluripotency-associated genes (*mNanog*, *mCripto*, *mRex1* and *mSox2*) was largely unaffected. (c) ES cells expressing normal levels of *Oct4* (-), partially reduced (*Dox*⁺⁶) or lacking *Oct4* expression (*Dox*⁺¹²) were fused to hB-lymphocytes. Successful reprogramming was assessed by quantifying the abundance of human ES-associated transcripts (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*), two days after fusion by qRT-PCR. Data were normalised to *Gapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

These data show that Oct4, but not Sox2, is critical for the dominant reprogramming activity of mouse ES cells. Interestingly, using 2O1 cells we observed the enhanced induction of hSox2 (Figure 4.12b, red arrow), a result that suggests that mouse-derived Oct4 levels may be important for initiating hSox2 expression from lymphocytes in these transient heterokaryons.

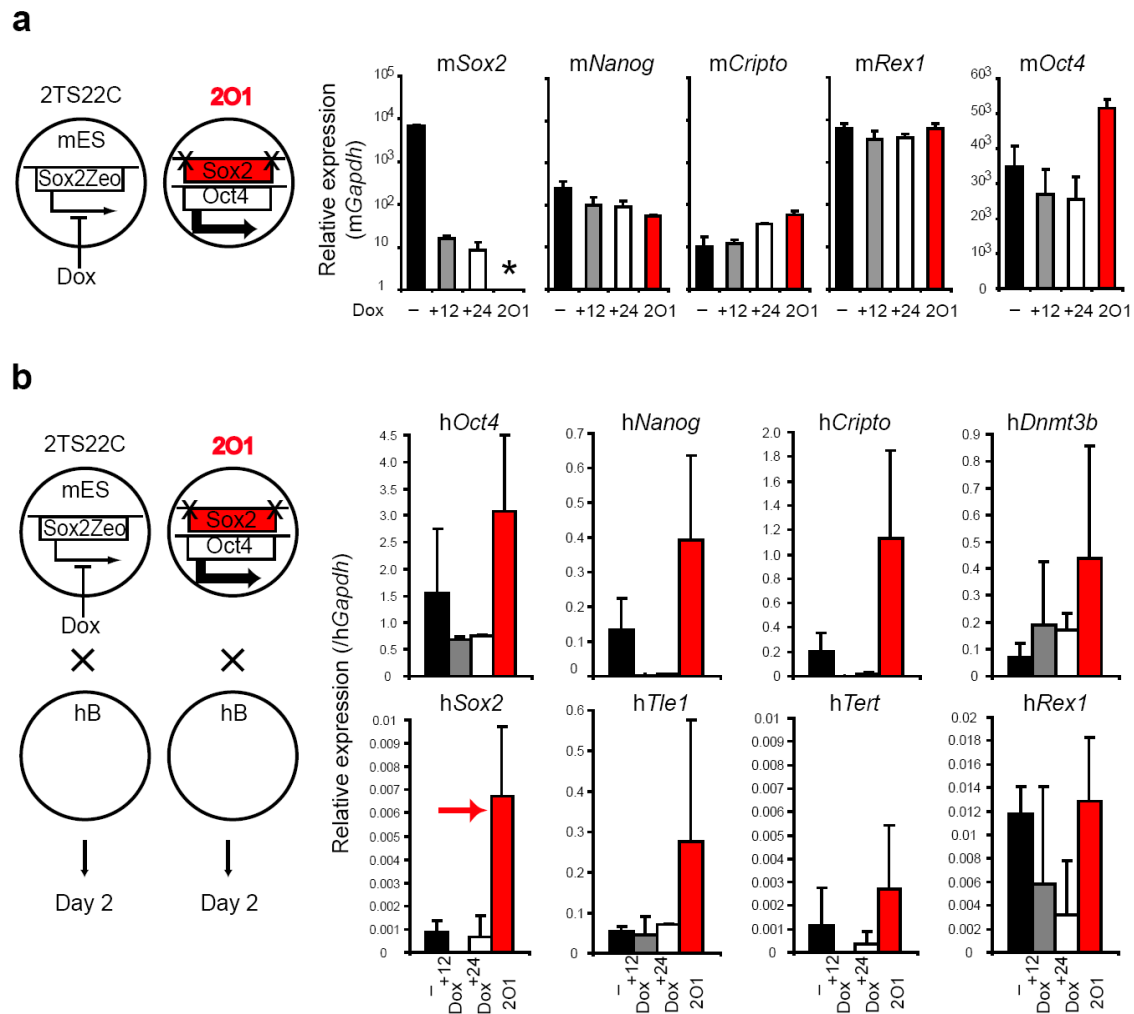


Figure 4.12. Sox2 is dispensable for reprogramming. (a) In 2TS22C ES cells endogenous Sox2 is replaced by an inducible transgene (Sox2Zeo) which can be downregulated by addition of doxycycline (Dox) (Masui et al., 2007). Quantitative RT-PCR analysis showed that 12 hours (+12) and 24 hours (+24) after Dox treatment, mSox2 was downregulated while expression of other pluripotency-associated genes (mNanog, mCripto, mRex1 and mOct4) continued to be expressed. 2O1 ES cells are Sox2-deficient mES cells (asterisk) in which mOct4 expression is up-regulated (red bars). (b) ES cells expressing Sox2 (-), Sox2 depleted cells (Dox⁺¹², Dox⁺²⁴) and 2O1 cells were fused to hB-lymphocytes. Successful reprogramming was assessed by quantifying the abundance of human ES-associated transcripts (hOct4, hNanog, hCripto, hDnmt3b, hSox2, hTle1, hTert and hRex1), two days after fusion by qRT-PCR. An elevated induction of hSox2 using 2O1 cells as a fusion partner is highlighted by an arrow (red). All data was normalised to Gapdh expression and error bars indicate the s.d. of 2-3 independent experiments.

4.6. ES-derived factors are dispensable for maintaining the reprogrammed status of somatic cells

To assess whether gene expression by the reprogrammed cell is stable (self-sustaining) or requires the continuous supply of factors provided by the mouse ES cell, I generated hybrid cells between B-lymphocytes and ES cells in which *Oct4* expression could be conditionally withdrawn following fusion (ZHBTc4, experimental outline shown in Figure 4.13a).

In these experiments, fusions were performed between mouse lymphocytes carrying a silent, *Oct4*-driven GFP transgene (GOF18 Δ PE) and mouse ZHBTc4 ES cells. This allows successfully reprogrammed hybrid cells to be identified on the basis of GFP re-expression (for example by day 10, Figure 4.13a). Hybrid clones contained a rearranged *IgH* locus, confirming their derivation from mouse B cells (Figure 4.13c) and displayed twice the DNA content of diploid cells (4n, Figure 4.13b) consistent with their tetraploid status. As anticipated, hybrid cells expressed ZHBTc4-derived *Oct4* β geo transcripts (Figure 4.14a) and several pluripotency-associated genes, but did not express B cell markers such as *CD19*, *Pax5* and *Ly108* (Figure 4.13d). Two hybrid clones were selected for study, hybrid 4 and 12. These were treated with Dox to selectively ablate expression of ZHBTc4-derived *Oct4* β geo (Figure 4.14; Figure 4.14a shows the strategy for the specific detection of *Oct4* β geo transgene). Withdrawal of ZHBTc4-derived *Oct4* did not alter the expression of *mNanog* and *mSox2* in reprogrammed cells (Figure 4.14b), and did not precipitate differentiation. For example, the trophectoderm associated genes *mCdx2* and *mHand1* were not induced by *mOct4* removal in hybrid cells (Niwa et al., 2000) (Figure 4.14c); while removal of *Oct4* from the parental ZHBTc4 line resulted in *mCdx2* and *mHand1* expression (Figure 4.14c, right hand panel). In addition, characteristic events of ES cell differentiation such as morphological changes and loss of *Oct4* protein (or *Oct4* promoter-driven GFP signal) were not observed in hybrid cells upon Dox treatment (Figure 4.14d and Appendix II, Figure A4). These cells could maintain self-renewal for at least for 15 passages in the presence of Dox without detectable differentiation (not shown).

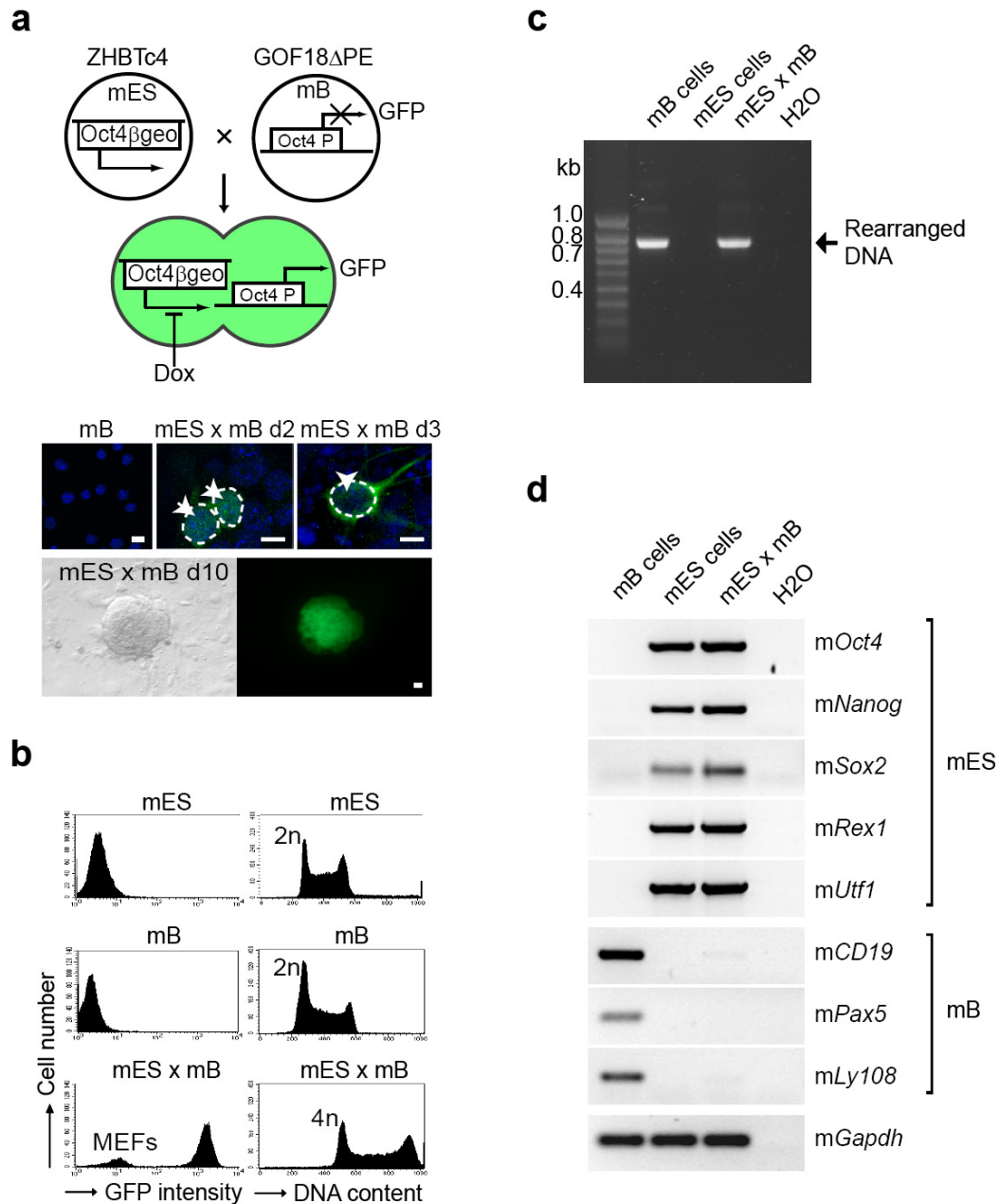


Figure 4.13. Derivation and characterisation of embryonic hybrids. (a) To address whether reprogramming is stable or subject to reversion, we ablated *Oct4* expression after hybrid formation. ZHBTc4 ES cells [mES with endogenous *Oct4* replaced by an inducible transgene (*Oct4βgeo*) which can be downregulated by addition of doxycycline (Dox) (Niwa et al., 2000)] were fused with mouse B-lymphocytes (mB) carrying a GFP transgene under the control of *Oct4* promoter (GOF18ΔPE). Reprogramming of mB results in the re-activation of GFP in hybrid colonies (d10, lower panels). Kinetic analysis of single cells (upper panels) showed that transgene re-activation occurs in heterokaryons (day 2, 2 arrows), and hybrid cells (day3, arrowhead). mB cells are shown as negative controls. Nuclei were visualised with DAPI staining (blue). Scale bars, 10 μm. (b) Hybrid clones (mES x mB, 4n) that re-expressed GFP were isolated and analysed by FACS. mES, mB and mES x mB hybrid cells unstained (left panel) or stained with propidium iodide (right panel) to assess GFP expression and DNA content, respectively. (c) Contribution of the lymphocyte genome within hybrid cells was confirmed by detection of a rearranged *IgH* locus (D-J region). *IgH* rearrangement was seen in B-lymphocytes (mB), hybrid cells (mES x mB) but not in mES cells. The rearranged DNA can be detected by PCR amplification and visualised on the gel as a 750 bp band. (d) Lymphocyte-specific genes (*mCD19*, *mPax5* and *mLy108*) were not detected in hybrid cells although ES-specific genes (*mOct4*, *mNanog*, *mSox2*, *mRex1* and *mUtf1*) remain detectable by RT-PCR.

Similar results were observed following withdrawal of ES-derived Sox2 in hybrid cells (hybrid cells generated between B-lymphocytes and 2TS22C ES cells in which Sox2 expression could be conditionally withdrawn, Figure 4.15a). No change in *mOct4* or *mNanog* expression (Figure 4.15b), or trophectoderm differentiation as *mHand1* expression (Masui et al. 2007) (Figure 4.15c) was not precipitated by the loss of Sox2 protein from hybrid cells (Figure 4.15d). These data show that reprogramming of mouse lymphocytes by mouse ES cells induces an epigenetically stable (and heritable) resetting of gene expression in the lymphocyte nucleus.

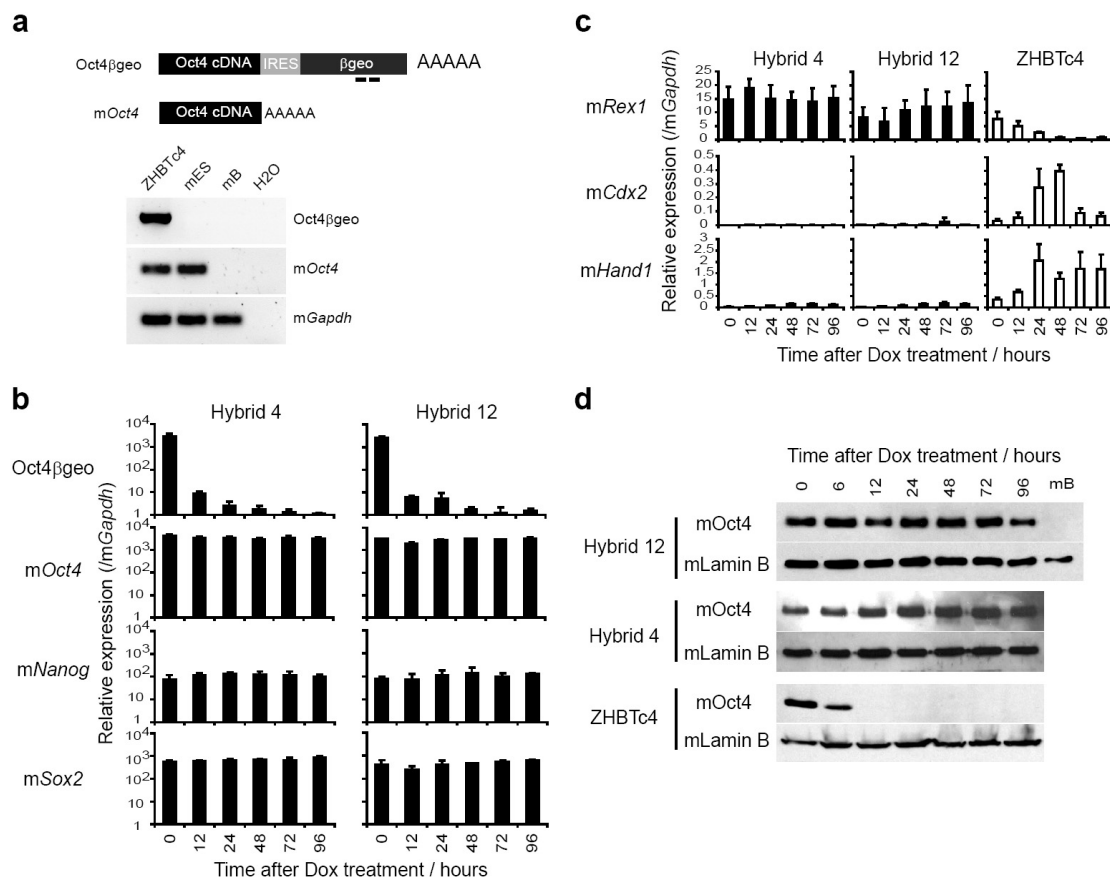


Figure 4.14. Reprogramming is self-sustaining and can be maintained in the absence of ES-derived Oct4. (a) Specific detection of Oct4βgeo transgene by RT-PCR with primers within βgeo cassette, which specifically amplify ZHBTc4-derived *Oct4* but not endogenous *mOct4*. mES and mB cells were included as controls. *mGapdh* was used to standardise input. (b) Hybrid clones (4 and 12) were treated with doxycycline (Dox) to ablate ES-derived *Oct4*, and qRT-PCR confirmed downregulation of Oct4βgeo transcript (upper panels). Removal of mES-derived Oct4βgeo in hybrid clones did not affect gene expression of pluripotency-associated transcripts (lower panels; *mOct4*, *mNanog* and *mSox2*) after 96 hours of Dox treatment. (c) No differentiation was observed after *Oct4* removal in hybrid cells. *mRex1* expression was retained and the extra-embryonic markers *mHand1* and *mCdx2* were not induced. In ZHBTc4 ES cells (open bars) upon Dox treatment, *mHand1* and *mCdx2* were induced (Niwa et al., 2000) and is shown for comparison. Data were normalised to *mGapdh* expression. Error bars indicate the s.d. of 3 independent experiments. (d) Western blotting with anti-Oct4 antibody confirmed that Oct4 protein is rapidly removed after Dox treatment of ZHBTc4 ES cells (lower panel) but remains detectable at all times in hybrid cells (upper panels). mB cells were included as a negative control. Equivalent protein loading is shown with Lamin B detection.

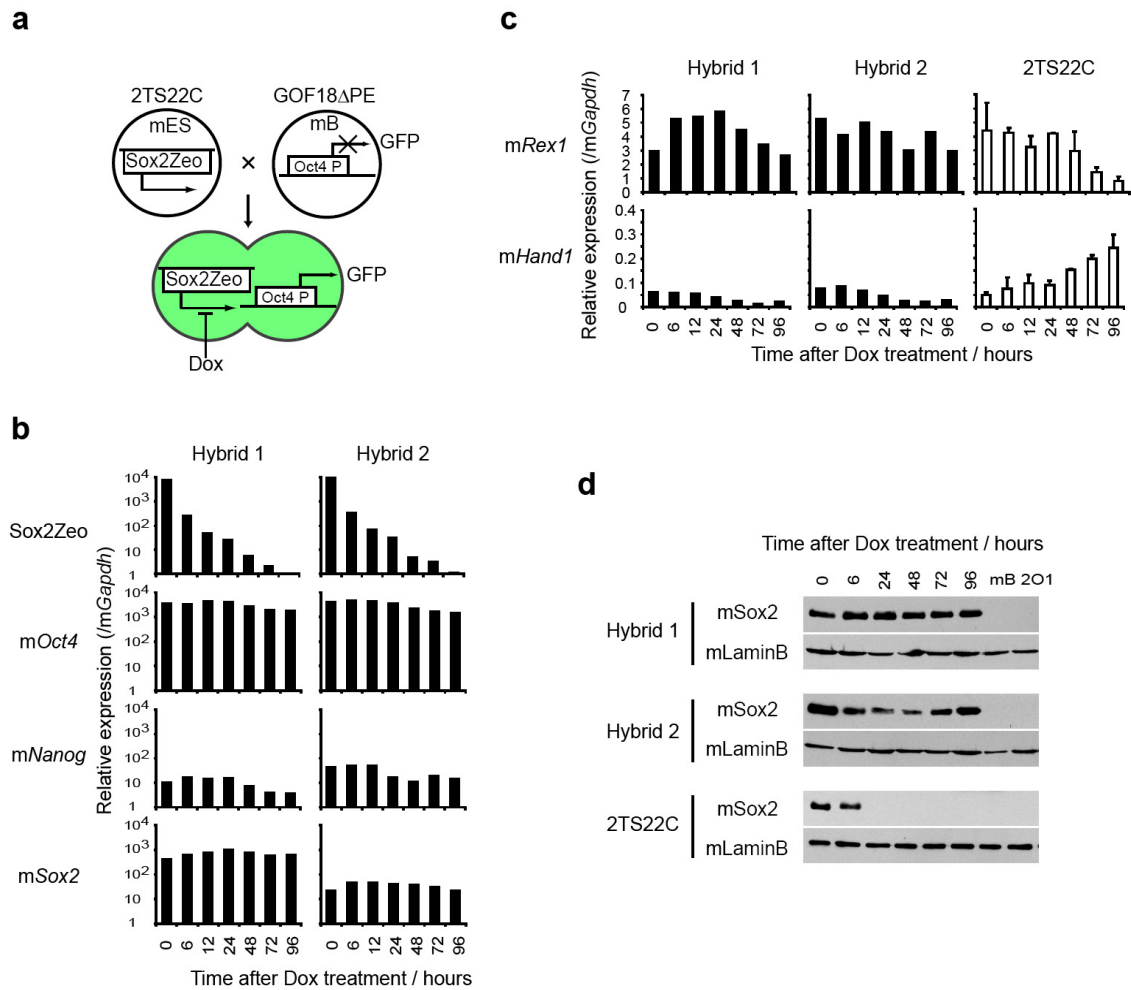


Figure 4.15. Sox2 is not required to maintain the reprogrammed status of hybrid cells. (a) 2TS22C ES cells (mES) were fused with mouse B-lymphocytes (mB) carrying a GFP transgene under the control of Oct4 promoter (GOF18 Δ PE) which undergoes re-activation after cell fusion and successful reprogramming. (b) Hybrid clones (1 and 2) were treated with doxycycline (Dox) to ablate ES-derived Sox2. Quantitative RT-PCR analysis showed the downregulation of Sox2Zeo transcript; in contrast, mSox2 levels were maintained, along with other pluripotency-associated genes (mOct4 and mNanog). (c) Differentiation was not observed after Sox2 removal in hybrid cells. mRex1 expression was sustained and mHand1 was not activated. 2TS22C ES cell line (open bars) is shown for comparison where mHand1 expression is induced upon Dox treatment (Masui et al., 2007). Data is normalised to mGapdh expression. Error bars indicate the s.d. of 2 independent experiments. (d) Western blotting with anti-Sox2 antibody confirmed that Sox2 protein is rapidly removed after Dox treatment of 2TS22C ES cells (lower panel) but remains detectable at all times in hybrid cells (upper panels). mB cells and 2O1 cells (Sox2 deficient cell line) were included as controls. Equivalent protein loading is shown with Lamin B detection.

4.6. Discussion and future perspectives

Previous studies have shown that lymphocyte nuclei can be reprogrammed to a pluripotent state by injection into *Xenopus* (Byrne et al., 2007) or mouse oocytes (Hochedlinger and Jaenisch, 2002) or by forming hybrid cells with pluripotent ES, EC or EG cells (Do et al., 2007; Tada et al., 1997; Tada et al., 2001).

Here I have shown that reprogramming human lymphocytes to a multi-potential state can be achieved by creating heterokaryons with mouse ES cells. This process results in the re-expression of endogenous human genes that are normally associated with human blastocyst development and human ES cell lines. Successful inter-species reprogramming is initiated in heterokaryons prior to nuclear fusion and chromosome intermixing, and generates cells that express human FGF signalling pathway components (Fgfr1, Fgfr2 and Fgf2) and human ES-specific surface molecules such as SSEA4, TRA-1-60 and TRA1-81. Markers that are restricted to mouse ES cells are not expressed by these cells (e.g. Bmp4 and LIF receptor). Conversion of human B cells into ES-like cells results in the re-modelling of the somatic genome with loss of DNA methylation at the *hOct4* locus. However, a subset of ES-associated genes analysed which displayed heterochromatic features such as late replication and repressive histone marks, were not easily re-activated in heterokaryons. Using genetically engineered mouse ES cells I demonstrate that successful reprogramming of human lymphocytes occurs independently of Sox2, a factor thought to be required for iPS cells. In contrast, reprogramming is critically dependent upon Oct4, since Oct4 deletion abolishes the reprogramming capacity of mES cells. Importantly, once reprogramming is initiated by factors produced by the dominant (ES) nucleus, we show that the subsequent withdrawal of mOct4 or mSox2 does not compromise the phenotype of hybrid cells. This result implies that the reprogrammed state, once initiated, is both self-sustaining and heritable (results summarised in Figure 4.16).

My results show that activation of endogenous human ES cell gene expression (e.g. *hOct4*, *hNanog*, *hCripto*, *hTle1*, *hDnmt3b*, *hDppa5*, *hGdf3* and *hGbx2*) by lymphocyte-derived nuclei begins very rapidly (24 hours after fusion), although the majority of these transcripts were not detected during the first 12 hours. Increasing amounts of ES-derived transcripts were detected in heterokaryons maintained in culture over an 8-day period of time. The observation that the percentage of SSEA4 positive cells does not increase over time suggests that successful conversion occurs in a subset of heterokaryons.

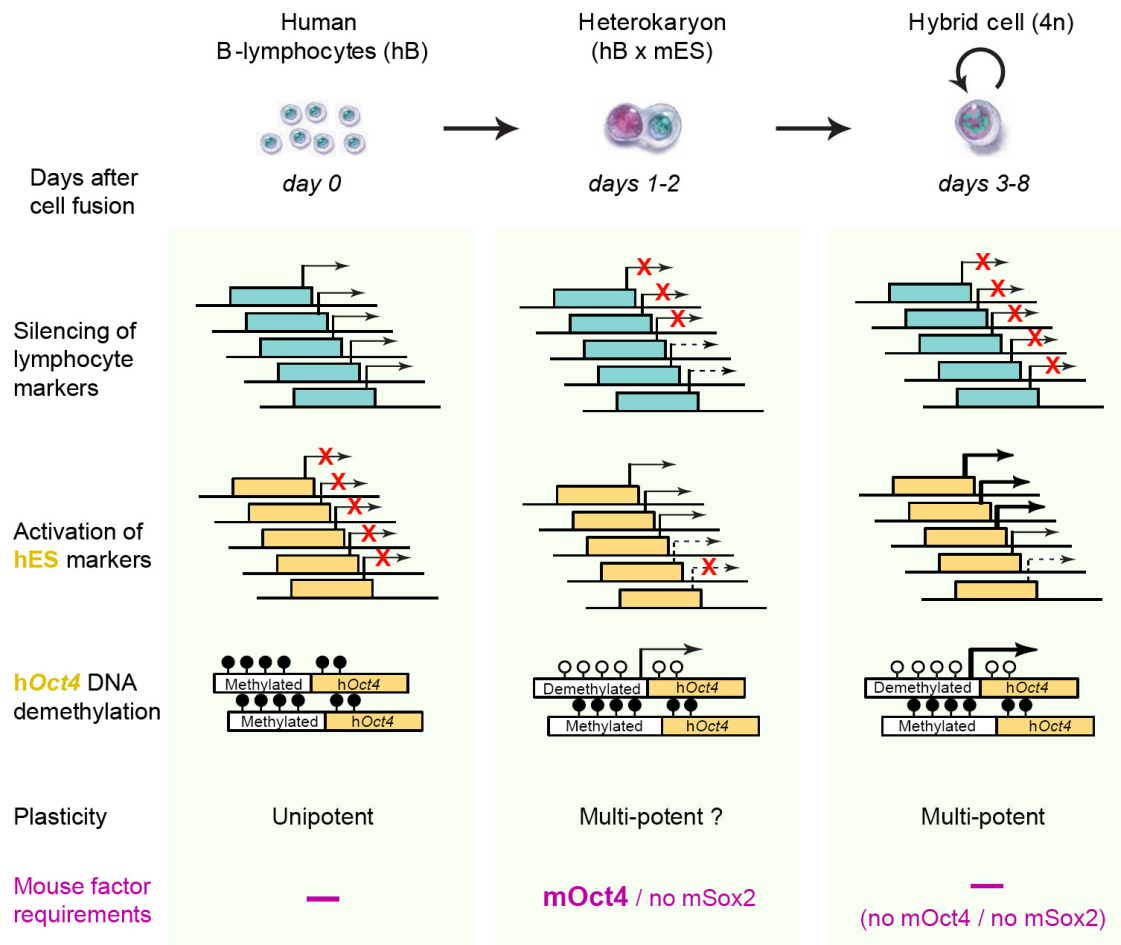


Figure 4.16. Events underlying reprogramming of lymphocytes in embryonic heterokaryons and hybrid cells. Extinction (crossed arrows; *hCD45*, *hCD37* and *hCD19*) or reduction (dashed arrows; *hCD20* and *hPax5*) of lymphocyte-specific gene expression is initiated in heterokaryons and is maintained in hybrid cells. Activation of the majority of endogenous human ES markers begins in heterokaryons (solid arrows) and transcript detection increases (bold arrows) in hybrid cells. However, some genes were not activated (crossed arrow; e.g. *hSox2*;) or detected at low levels (dashed arrows; e.g. *hRex1*) in heterokaryons. Demethylation of the human *Oct4* promoter starts in heterokaryons (in a subset of clones) and is maintained in hybrid cells. mOct4 but not mSox2 is required to initiate the conversion for the multi-potent phenotype in heterokaryons. The maintenance of the reprogrammed status in hybrid cells does not require mouse *trans* acting factors such as mOct4 and mSox2.

Whether the increased transcript detection was a consequence of an accumulation of stable mRNAs in successfully reprogrammed heterokaryons or of an increase in transcription rate with time, cannot be determined in the experiments reported here. When compared to the activation of muscle genes in muscle heterokaryons, where transcripts were firstly detected from B-lymphocytes two days after fusion (Chapter 3), gene re-activation in embryonic heterokaryons is very rapid suggesting that undifferentiated ES cells are potent “reprogrammers” of somatic cells. Although the molecular mechanisms of dominant reprogramming by ES cells is poorly understood, the rapid induction of gene expression could reflect the high expression levels of chromatin remodelling factors found in undifferentiated cells (Ramalho-Santos et al., 2002). To support this idea there is evidence that ES cells possess “hyperdynamic

chromatin" (Meshorer et al., 2006) and are globally transcriptionally active (Efroni et al., 2008) or primed (Stock et al., 2007), as compared with differentiated cells. Extinction of lymphocyte-specific gene expression (hCD45, hCD37, hCD19, hCD20 and hPax5) was achieved after 48-72 hours and was maintained in hybrid cells. This data is consistent with earlier work examining gene extinction in hybrid cells (Massa et al., 2000). Lymphoma cells fused with fibroblasts (Junker and Pedersen, 1985) results in immunoglobulin gene extinction very shortly after fusion. Maintenance of lymphocyte-specific gene silencing in both inter- and intra-species hybrids is also supported by genome-wide analysis of gene expression in human and mouse hybrid cells (Ambrosi et al., 2007; Cowan et al., 2005).

One surprising aspect of the reprogramming data shown here is the rapidity of gene conversion and DNA demethylation that occurs within heterokaryons. Previous studies have shown that reprogramming in experimental heterokaryons using adult cells from different lineages (Blau et al., 1985; Terranova et al., 2006) can be initiated before chromosome replication and cell division. In addition, heterokaryons formed by fusing adult mouse erythroleukemia (MEL) cells and human embryonic/foetal erythroid (K562) cells revealed that transcriptionally inactive mouse globin genes can be re-activated within 24 hours in transient heterokaryons (Baron and Maniatis, 1986). Here I show that conversion of unipotent lymphocytes towards multipotency is achieved in transient heterokaryons before nuclear fusion and cell division. The rapid re-activation of endogenous pluripotency-associated genes seen in inter-species heterokaryons is consistent with transgene re-activation studies that have reported *Oct4gfp* expression (at 24 hours) by MEFs (Wong et al., 2008) or neural stem cells (Han et al., 2008) fused with mouse ES or EC cells. Likewise, re-activation of human *Oct4* and *Nanog* by human somatic nuclei, has been shown to occur rapidly upon DNA de-methylation and Tpt1 activation induced by *Xenopus* oocytes (Koziol et al., 2007; Simonsson and Gurdon, 2004). My results underscore that nuclear fusion and cell division are not required events for reprogramming. Nevertheless, the importance of cell division, and in particular DNA synthesis for both establishing and maintaining programs of gene expression has been widely debated (Egli et al., 2008; Wolffe, 1991). It has been suggested that mitotic conditioning is crucial to reset the chromatin structure of differentiated adult donor cells for embryonic DNA replication in *Xenopus* eggs (Lemaitre et al., 2005). In agreement with this hypothesis, ES cells were shown to be most effective at reprogramming when they contained a high proportion of cells in the S and G2/M phases of the cell cycle (Sullivan et al., 2006). Landmark cell cycle experiments by Rao and colleagues (Rao and Johnson, 1970) established that S-phase cells are able to induce DNA synthesis when fused to a G1 cell. Furthermore,

the same observation has been reported in inter-species heterokaryons (Graves, 1972). These studies suggest that nuclear reprogramming may require an initial round of somatic DNA replication of silent chromatin in the presence of ES-derived factors produced during S and G2/M phases. With the experiments reported here I cannot confirm or refute the importance of DNA synthesis for reprogramming to pluripotency in inter-species heterokaryons. However, as successful reprogramming is only achieved in a proportion of heterokaryons (<13%), and mouse ES cell cultures contain >30% of cells in S-phase, it is possible that genome-wide DNA synthesis or partial DNA replication [or repair (Barreto et al., 2007)] is required for successful lymphocyte conversion. Further experiments will be required to elaborate on this point.

The epigenetic remodelling of the hB-lymphocyte genome for pluripotency includes the loss of DNA methylation at the *hOct4* promoter region as previously reported (Cowan et al., 2005; Freberg et al., 2007; Simonsson and Gurdon, 2004; Tada et al., 2001). Whether the gene-specific de-methylation observed in heterokaryons involves passive or active process [such as DNA repair-based mechanisms (Barreto et al., 2007)], is unknown. In contrast to what is observed for *hOct4* (that is transcriptionally and epigenetically reprogrammed rapidly in heterokaryons), some genes may be subjected to several layers of “repression” such as late replication and repressive histone marks, rendering them more “resistant” to activation (e.g. *hSox2*, *hTert*, *hFoxd3*, *hDppa4*, and *hRex1*). Accordingly, when the expression of 10 *Oct4*-related genes was analysed in individual cumulus cell-derived mouse cloned blastocysts, only 62% correctly expressed all tested genes (Bortvin et al., 2003). Interestingly, *Dppa4* was among the group of genes that showed variable expression. The repressed chromatin configuration of *Dppa4* may offer an explanation as to why some genes are difficult to re-activate.

Importantly, I show that by enhancing *Oct4* levels in *Sox2*-deficient ES cells (ES-2O1), an elevated expression of *hSox2* is induced in reprogrammed human B cells. Recent genome-wide studies have shown that *Sox2* is a direct target of *Oct4* in both human (Boyer et al., 2005) and mouse (Loh et al., 2006) ES cells, a fact that could explain why *hSox2* is efficiently reprogrammed using ES cells that overexpress *mOct4*. In addition to *Sox2*, *Oct4*, *Nanog* and *Cripto* were also shown to be direct *Oct4* targets in both species. Accordingly, these genes were more sensitive to *Oct4* removal (partial *Oct4* removal after 6 hours of Dox treatment) or elevation (using ES-2O1 cells as fusion partner) (see Appendix II, Figure A2 and A3). Together, my results suggest that re-expression of silent *loci* can be overcome by an elevation of the appropriate transcriptional activator (*Oct4*). The exact contribution of chromatin-mediated repression during reprogramming is difficult to address because transcription factors

often recruit co-factors that alter chromatin structure. However, in our hands, overexpression of exogenous *Oct4* in lymphocytes did not induce pluripotent conversion (Pereira & Terranova, unpublished results), a finding that argues that additional DNA-binding and chromatin remodelling factors [including those known to interact with Oct4 (Liang et al., 2008; Wang et al., 2006), or associated with the process of DNA demethylation] may be critical for successful reprogramming. In agreement with this, Oct4 is part of a protein interaction network in ES cells, which is linked to several chromatin co-factors. These include proteins associated with gene repression (such as the histone deacetylation complex NuRD, PcG proteins and the histone demethylase LSD1) and gene activation (the SWI/SNF chromatin remodelling complex proteins SMARCA2, SMARCA4, BAF180 and BAF155) (Liang et al., 2008; Wang et al., 2006). In addition, Oct4 has been shown to be required to maintain the chromatin structure within the extended Nanog *locus* in ES cells (Levasseur et al., 2008).

In the experiments presented here, I have shown that reprogrammed human cells express a profile of transcripts, signalling molecules and surface antigens that are similar to those seen in human ES cells, and different from mouse ES cells. This suggests that an early human embryonic “program” of gene expression is initiated in human nuclei by *trans*-acting (mouse-derived) factors. Differences between the expression profiles of mouse ES cells and human reprogrammed nuclei probably reflect discrepancies in *cis*-acting regions between the mouse and human genomes. However, it is also possible that newly synthesised human regulatory factors (from hB-derived nuclei) could be influencing the pattern of gene expression and directing the hES phenotype (Hardeman et al., 1986). This explanation seems unlikely as (a) the human-specific profile of gene expression was observed very early after cell fusion (24 hours), and (b) the inhibition of *de novo* protein synthesis by treatment with cycloheximide D upon cell fusion did not alter the overall pattern of gene expression (Appendix II, Figure A5). A study in which the entire *hTert* gene was introduced into mice, showed that expression of the transgene was similar to endogenous *hTert* in humans, rather than mouse endogenous *mTert* (Horikawa et al., 2005). It is interesting to speculate that some of the differences between human and mouse ES cells (Ginis et al., 2004; Pera and Trounson, 2004; Sato et al., 2003; Xu et al., 2005) may indeed reflect intrinsic species dissimilarities, rather than temporal differences in the isolation of stem cells (Brons et al., 2007; Tesar et al., 2007). We show that fusion of human lymphocytes to mouse ES cells (that are LIF and Bmp dependent), human ES-like cells are generated that express FGF signalling components (and are not dependent of LIF/Bmp). In agreement with this data, attempts to generate human iPS cells in the

presence of LIF rather than Fgf2 have been unsuccessful (Takahashi et al., 2007). This may suggest that differences between human and mouse ES cells reflect distinct signalling and transcriptional networks, rather than necessarily when or where they were isolated during embryogenesis. It has been documented that only 9.1% of the total genes targeted by Oct4 are shared between human and mouse (Boyer et al., 2005; Loh et al., 2006). Whether the hES gene expression program initiated in reprogrammed lymphocytes is a result of divergent transcriptional networks between human and mouse (Odom et al., 2007), remains to be addressed. The identification of mouse Oct4 genomic targets in human nuclei upon heterokaryon formation would potentially help address this question. For this, the mouse ES cell line constructed during this study (that expresses a Flag-tagged Oct4 protein, Flag-mOct4) would provide a valuable tool. Moreover, the introduction of genetically modified human somatic nuclei in embryonic heterokaryons may help to elucidate the influence of endogenous human factors and regulatory sequences during *cis* reprogramming.

Retroviral-mediated expression of four transcriptional regulators, *Oct4*, *Sox2*, *c-Myc* and *Klf4*, has been shown to drive mouse fibroblasts into an ES-like (iPS) state, albeit at very low frequency (Takahashi and Yamanaka, 2006; Wernig et al., 2007). Mouse hepatocytes, stomach cells (Aoi et al., 2008) and pancreatic beta-cells (Stadtfield et al., 2008a) are also converted using the same strategy, whereas adult mouse B-lymphocytes (Hanna et al., 2008) required the additional expression of *C/EBP α* or silencing of the B-cell specific factor *Pax5*. Reprogramming of human fibroblasts has also been achieved in a parallel approach using *Oct4*, *Sox2* and either *Nanog* plus *Lin28* (Yu et al., 2007) or *Klf4* plus *c-Myc* (Masui et al., 2007; Park et al., 2008c). These pioneering studies have illustrated the importance of these factors for iPS, but also suggested that additional factors may be needed for efficient conversion to pluripotency. I show here that *Sox2*, in contrast to *Oct4*, is not required to convert human B-lymphocytes into a multi-potent state in heterokaryons. This observation contrasts with results obtained using iPS strategies to reprogram mouse and human fully differentiated cells. Whether this is because of differences relating to the overexpression of transcription factor cocktails used in iPS, or that reprogramming occurs over an extended time period [pluripotency-associated genes such as *Oct4*, *Nanog*, *Sox2* and *Tert* are re-activated after 2 weeks of transduction (Brambrink et al., 2008; Stadtfield et al., 2008b)], is not known. However, as *Sox2* was recently shown to be dispensable for the activation of Oct-Sox enhancers in mouse ES cells (Masui et al., 2007), it is also possible that additional Sox family members such as *Sox4*, *Sox11* and *Sox15*, may have redundant functions with *Sox2* in reprogramming.

Finally, I show that withdrawal of mOct4 or mSox2 after reprogramming is initiated does not compromise the phenotype of hybrid cells. This result implies that the reprogrammed state, once initiated, is both self-sustaining and heritable. The targeted elimination of both copies of chromosome 6 (which harbour pluripotency-associated genes including Nanog) from mouse ES-somatic hybrid cells, did also not affect the phenotype of hybrid cells (Matsumura et al., 2007). However, it has been recently shown that ES can self-renew indefinitely in the permanent absence of Nanog, although cells are prone for differentiation (Chambers et al., 2007). Oct4, in turn, is essential for stem cell self-renewal and the reduction of Oct4 levels by 50% leads to differentiation in trophoblast lineage (Niwa et al., 2000). The observation that ES-derived Oct4 is not required for self-renewal of hybrids provides stronger evidence that cell fusion-induced reprogramming is indeed stable. Similarly, withdrawal of exogenous Oct4 after iPS generation does compromise self-renewal (Maherali et al., 2007). Thus reprogramming to pluripotency seems to be a stable heritable condition that does not require the continuous supply of exogenous factors in order to be maintained.

Hence, my results show that human ES-like cells are rapidly and efficiently generated from lymphocytes through fusion with mouse ES cells, that reprogramming is stable, and is achieved through the interplay between transcriptional effectors and unique genomic information provided by human somatic nuclei. Moreover, experimental embryonic heterokaryons offer a powerful experimental approach to trace the contribution of individual factors to the reprogramming of human somatic cells towards a multipotent state.

CHAPTER 5. ROLE OF PCG PROTEINS IN THE MAINTENANCE OF ES CELL REPROGRAMMING ACTIVITY

5.1. Pluripotent cells reprogram human lymphocytes in heterokaryons

The ability of ES cells to induce the expression of a pluripotency-associated markers from differentiated cells, is a property that is shared by other embryonic-derived cell populations (Do et al., 2007; Miller and Ruddle, 1976; Tada et al., 1997). These dominant reprogramming cells include EG cells (Tada et al., 1998), which are derived from PGCs that colonise the genital ridges during early organogenesis (Matsui et al., 1992; Resnick et al., 1992), and EC cells derived from induced or spontaneous teratocarcinomas (Andrews et al., 2005). EG and EC cells share many features with ES cells, such as the ability to self-renewal and follow multiple differentiation pathways. However, EC cells have a more restricted lineage potential (Chambers and Smith, 2004) and EG cells have lost allele-specific parental imprints (Labosky et al., 1994) and are able to reprogram imprinted genes after hybridisation with somatic cells (Tada et al., 1997; Tada et al., 2001). Collectively these reports have encouraged a view that the ability for dominant reprogramming is a shared feature of the undifferentiated state. Here I examine the requirements for reprogramming human lymphocytes towards pluripotency in heterokaryons.

The approach used for heterokaryon formation between pluripotent cell lines and human B-lymphocytes is described in Chapter 4. Human B cells were fused with either mouse ES, EC or EG cells using PEG. Reprogramming was assessed by qRT-PCR of the human embryonic genes *hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*, induced 0 to 3 days after cell fusion (Figure 5.1). This experiment was performed in collaboration with Joana Santos. The reprogramming potential of F9 and P19 embryonic carcinoma cell lines, that express numerous markers in common with ES cells, but have a narrower range of lineage potential (Chambers and Smith, 2004), was examined (Figure 5.1; complete kinetics shown in Appendix II, Figure A6). Reprogramming activity, as measured by the reactivation of human embryonic genes, was similar between the two embryonic carcinoma cell lines (F9 and P19) and ES cells

(E14tg2A and OS25). However, a more detailed comparison between F9 and P19 EC cells revealed slight differences, such as an enhanced ability of P19 cells to initiate the expression of *hSox2* and conversely a reduction in *hRex1* induction kinetics. 19G and 23G embryonic germ cells (Tada et al., 1998) were also able to reprogram human lymphocytes, albeit less efficiently (Figure 5.1; complete kinetics shown in Appendix II, Figure A7). Human pluripotency-associated transcripts were approximately 10-fold less abundant in EGxhB cultures than in ESxhB or ECxhB cultures. In addition, *hRex1*, a marker of the inner cell mass that is downregulated in the epiblast (Pelton et al., 2002; Rogers et al., 1991), was not detected upon cell fusion of EG cells with human lymphocytes. These data suggest that although differences in the reprogramming potential may reflect the developmental origin of cells used, all pluripotent cells share a dominant reprogramming capacity in heterokaryons.

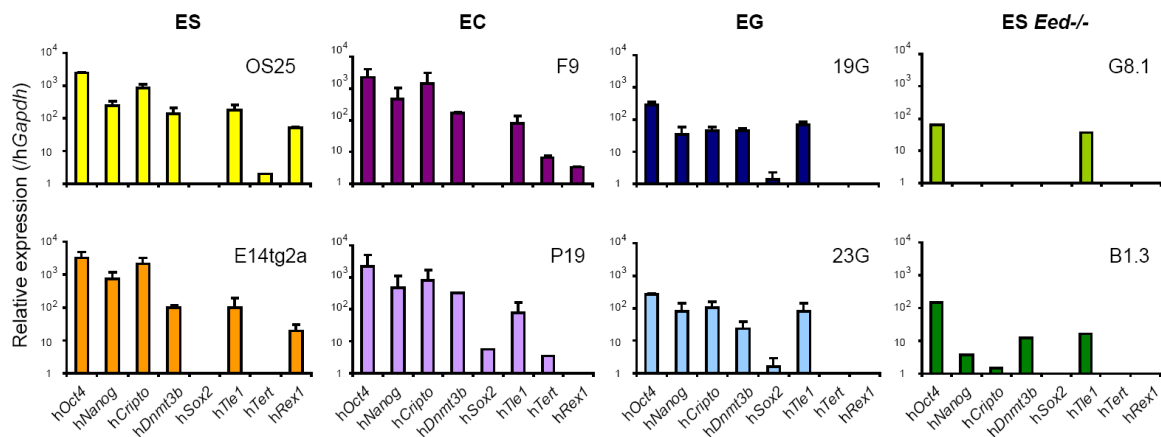


Figure 5.1. Reprogramming potential of pluripotent cell lines. Mouse embryonic stem (ES: OS25 and E14tg2A; yellow and orange bars, respectively), embryonic carcinoma (EC: F9 and P19; dark and light purple bars, respectively), embryonic germ (EG: 19G and 23G; dark and light blue bars, respectively) or *Eed*-deficient ES cells (ES *Eed*^{-/-}: G8.1 and B1.3; dark and light green bars, respectively) were fused to human B-lymphocytes. Successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 3 days after cell fusion. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

5.2. *Eed* is required for embryonic stem cell reprogramming activity

Undifferentiated embryonic cells share many features, including self-renewal, pluripotency and reprogramming activity. These characteristics are translated in the high expression levels of pluripotency-associated transcription factors such as *Oct4*. Withdrawal of *Oct4* from ES cells abolishes self-renewal, pluripotency (Niwa et al., 2000) and reprogramming (see section 4.5). In addition, “stemness” is also reflected at the epigenetic level. For example, Polycomb-mediated repression by the catalysis of

H3K27 methylation at important lineage determinants in undifferentiated ES cells plays an important role keeping these genes in a repressed but primed state (Spivakov and Fisher, 2007). This observation together with the fact that the PcG protein *Eed* is expressed at high levels in undifferentiated ES and *ex vivo* PGCs (Grskovic et al., 2007; Mak et al., 2004), prompted us to examine whether Polycomb repression is required for stem cell-mediated reprogramming. Reprogramming was examined using ES cells that were genetically deficient in *Eed* (see section 2.1.2 for cell line details). *Eed* as part of the PRC2 is essential for the methyltransferase activity of the complex and H3K27 methylation (Azuara et al., 2006; Montgomery et al., 2005; Peters and Schubeler, 2005). Trimethylated and dimethylated H3K27 were not detected in *Eed*-null ES cells, and levels of monomethylated H3K27 were markedly reduced (Figure 5.2a). ES cells lacking *Eed* retain self-renewal and can differentiate *in vitro* (as embryoid bodies) and *in vivo* (within chimeric embryos) into multiple lineages at a reduced efficiency (Chamberlain et al., 2008; Morin-Kensicki et al., 2001). *In vivo* development is compromised in *Eed*-deficient embryos, as these die shortly after gastrulation (Faust et al., 1998). Two independent *Eed*-deficient ES cell lines (G8.1 and B1.3) were fused to hB-lymphocytes and reprogramming assessed by qRT-PCR of the human embryonic genes *hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*. Interestingly, both cell lines showed a severe reduction in the ability to reprogram hB-lymphocytes (Figure 5.1, right panel). The majority of pluripotency-associated transcripts were either reduced or undetectable after 3 days of heterokaryon formation. This failure in reprogramming contrasts with the induction of embryonic-associated transcripts by the other pluripotent cell lines analysed (Figure 5.1, left and central panels). To verify that compromised reprogramming is a consequence of *Eed* loss rather than a reflection of differences in cell line genetic background, *Eed* heterozygous, *Eed*-null and wild-type ES cells were derived from genetically matched blastocysts (Azuara et al., 2006), and each fused to hB-lymphocytes. As shown in Figure 5.2b, successful reprogramming, as judged by induction of several human genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) was abolished in *Eed*-deficient ES cells (ESL32), but not in the heterozygous (ESL12) or wild-type ES cells (ESL21) (a complete kinetic analysis is provided in Appendix II, Figure A8). My results confirm that loss of *Eed* and H3K27 methylation impair the ability of ES cells to dominantly reprogram human lymphocytes.

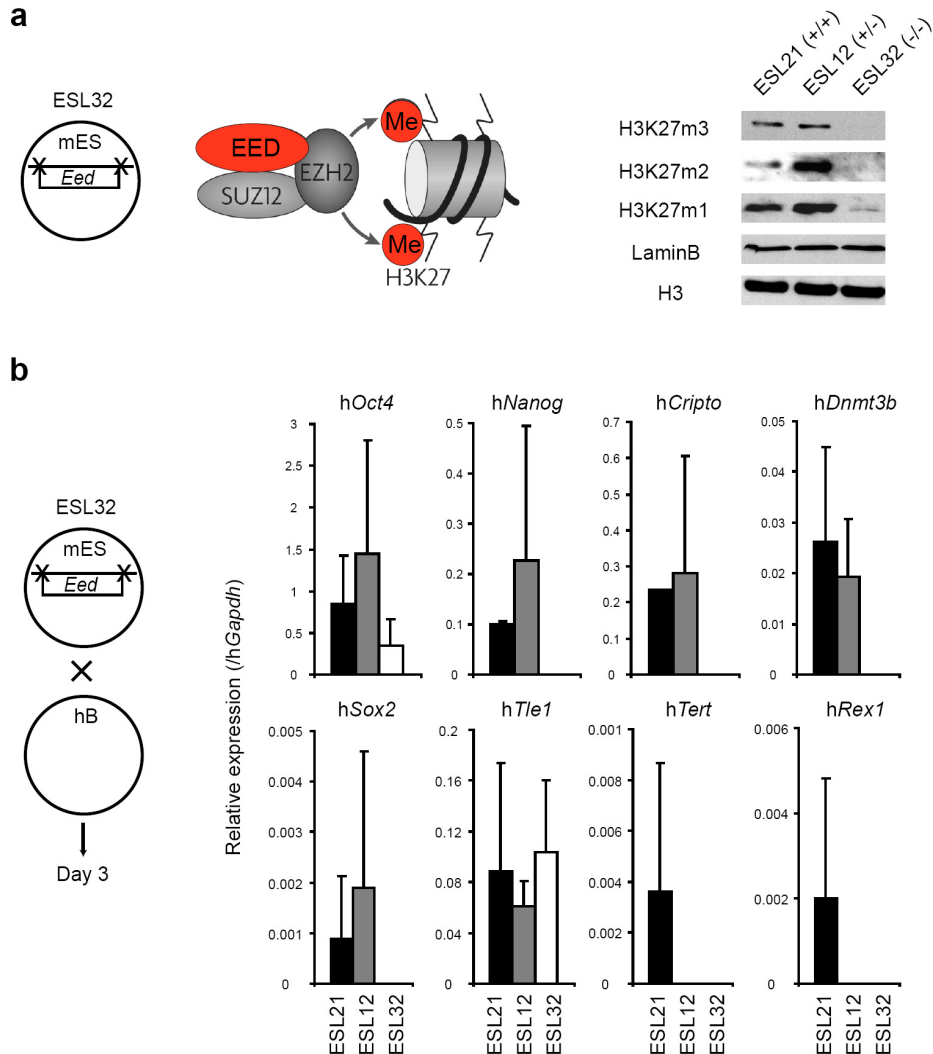


Figure 5.2. *Eed* is required for the ES cell-mediated reprogramming of human lymphocytes. (a) Whole cell lysates from wild-type (ESL21, +/+), heterozygous (ESL12, +/-) and *Eed*-deficient ES cells (ESL32, -/-) were analysed by Western blotting using antibodies to mono- (H2K27m1), di- (H2K27m2), or trimethylated H3K27 (H2K27m3). Equivalent protein loading is shown by Lamin B and total Histone 3 detection. The deletion of the PRC2 member *Eed* in ES cells and the subsequent global reduction of H3K27 methylation levels are highlighted in red. (b) Wild-type (ESL21, black bars), heterozygous (ESL12, grey bars) and *Eed*-deficient ES cells (ESL32, white bars) were fused to human B-lymphocytes. Successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 3 days after cell fusion. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

To assess whether the loss of *Eed* is the cause of the ES cell reprogramming defect, *Eed* expression was restored in *Eed*-null ES cells (B1.3 cell line) using a 180 Kb BAC clone carrying *Eed* gene and neomycin resistance gene (see section 2.2 for cell line derivation details). ES cell colonies resistant to neomycin were screened for the presence of *Eed* protein and H3K27 trimethylation by Western blotting (Figure 5.3a). ES cells in which *Eed* isoforms (Montgomery et al., 2005; Montgomery et al., 2007) were re-expressed and H3K27 trimethylation was re-established (B1.3BAC) were then fused to hB-lymphocytes, and reprogramming assessed by qRT-PCR of human

embryonic genes. Reconstitution of *Eed* expression in *Eed*-deficient ES cells rescued their reprogramming defect (Figure 5.3b, compare solid and open bars on histograms). This data confirm that *Eed* has a specific role in maintaining the reprogramming activity of ES cells. Not all genes were equally affected by *Eed* deletion; for example *hTle1* and *hDnmt3b* activation was variable and their expression was eventually detected (see Figures 5.2 and 5.3).

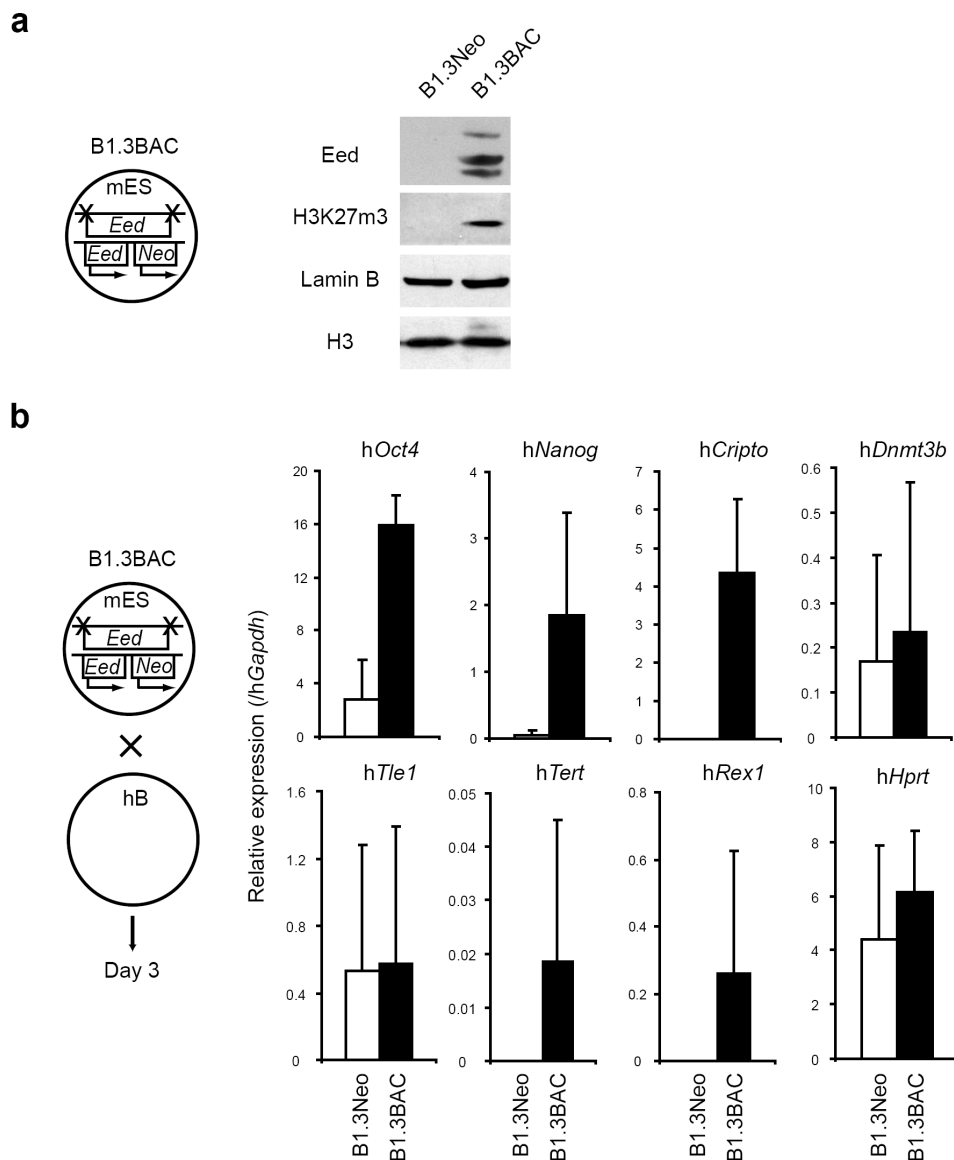


Figure 5.3. Restoring *Eed* expression in *Eed*-deficient ES cells rescue their reprogramming defect. (a) *Eed*-deficient ES cells (B1.3) were rescued by insertion of a BAC clone that carried the *Eed* gene and neomycin resistance gene (B1.3BAC). Whole cell lysates from rescued cell lines (B1.3BAC) and controls (B1.3Neo) were analysed by Western blotting using antibodies to the *Eed* protein and anti-trimethylated H3K27 (H3K27m3). Equivalent protein loading is shown by Lamin B and total H3 detection. (b) B1.3BAC and B1.3Neo ES cells were fused to human B-lymphocytes and successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

5.3. PRC2 activity and H3K27 methylation are required for the reprogramming mediated by ES cells

Methylation of H3K27 is catalysed by PRC2, a complex that depends on three core components: embryonic ectoderm development (*Eed*), suppressor of zeste 12 (*Suz12*) and the HMT enhancer of zeste homologue 2 (*Ezh2*) (Cao et al., 2002; Cao and Zhang, 2004; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). To address whether the lack of reprogramming seen in *Eed*-deficient ES cells is indeed due to the loss of H3K27 methylation, rather than any other putative function of the *Eed* protein, ES cells genetically deficient for *Suz12* (Fujimura et al., 2006) or ES cells that are inducible *null* for *Ezh2* [see section 2.1.2 for cell line details (Su et al., 2003)] were fused to human B cells. Similarly to what was observed in *Eed*-deficient ES cells, trimethylated and dimethylated H3K27 were not detected in *Suz12-null* ES cells and levels of monomethylated H3K27 were markedly reduced (Figure 5.4a, left panel). The *Ezh2*-1.3 ES cell line, that carries tamoxifen-inducible conditional knockout *Ezh2* alleles, was progressively depleted of *Ezh2* protein following addition of tamoxifen (Figure 5.4a, right panel). Although the cells continued to self-renew in the absence of *Ezh2* (*Ezh2*-1.3 Δ/Δ), withdrawal of *Ezh2* resulted in the loss of trimethylated H3K27, a reduction in dimethylated H3K27 and a slight reduction of monomethylated H3K27. *Ezh2* protein levels were reduced following *Suz12* deletion (Figure 5.4a), whereas *Suz12* protein levels were unaffected by *Ezh2* removal, consistent with previous observations (Montgomery et al., 2005; Pasini et al., 2004). *Suz12* and *Ezh2 null* ES cells were fused to lymphocytes and reprogramming was assessed by quantitative RT-PCR of the human ES-associated genes (Figure 5.4b and 5.4c; complete kinetic experiments are provided in Appendix II, Figure A9 and A10, respectively). Deletion of the PRC2 core components *Suz12* and *Ezh2* abolished the reprogramming activity of ES cells. The reprogramming defect seen with *Ezh2-null* ES cells was less pronounced than that seen for *Eed*^{-/-} or *Suz12*^{-/-} cells; *hNanog* and *hCripto* were detected at low levels following fusion with *Ezh2*-1.3 ES cells treated with tamoxifen for 96 hours, or with *Ezh2-null* ES cells. This observation could however, be explained by an incomplete removal of H3K27 methylation following *Ezh2* withdrawal. Collectively, these data illustrate the importance of *Eed*, *Suz12* and *Ezh2* for reprogramming and implicate PRC2 and the methylation of H3K27 as crucial components of the ES cell reprogramming machinery.

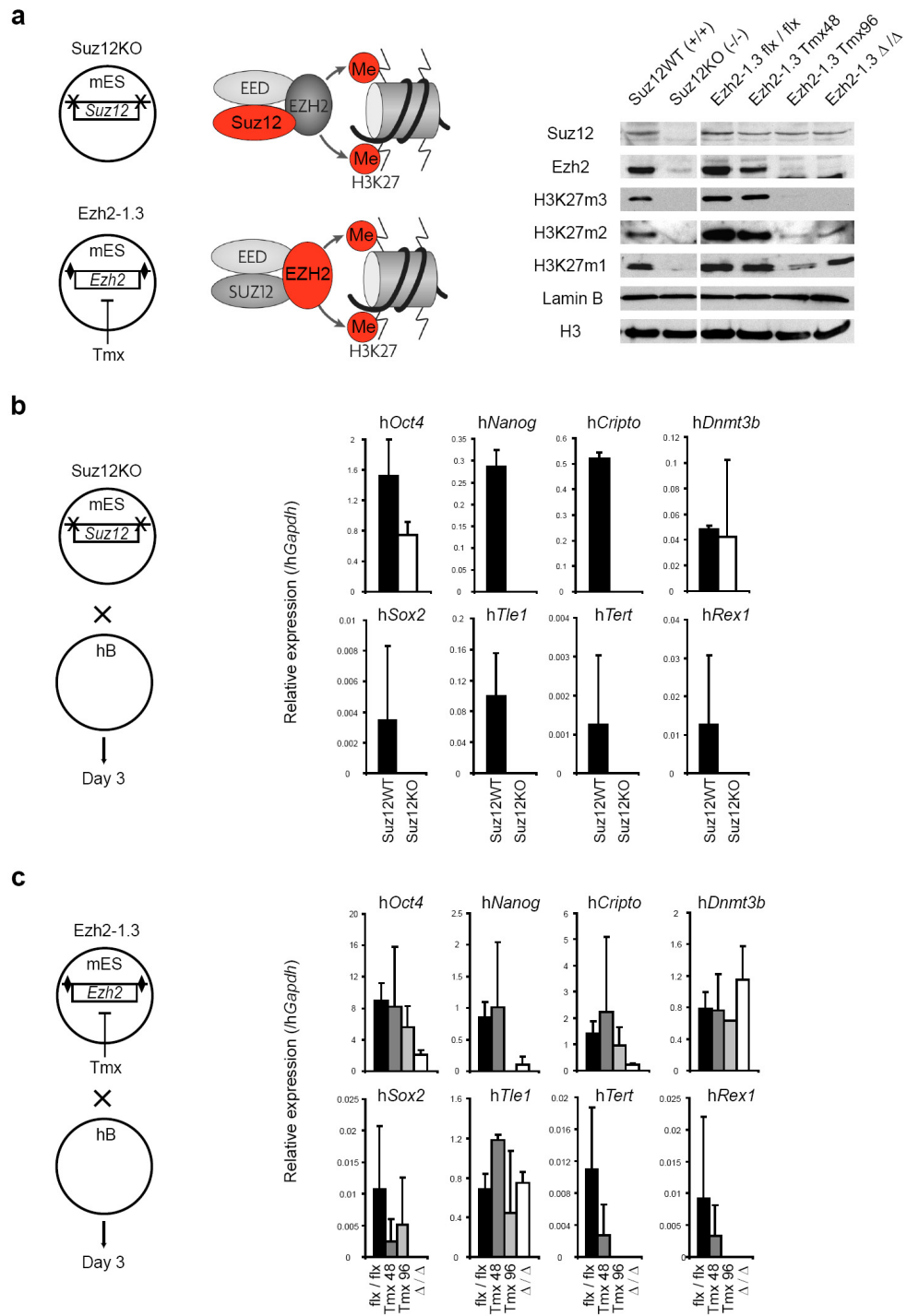


Figure 5.4. Methylation of H3K27 in ES cells is required for successful reprogramming. (a) Whole cell lysates from wild-type (Suz12WT, +/+), *Suz12*-deficient ES cells (Suz12KO, -/-), *Ezh2*-1.3 flx /flx [untreated or treated with 800nM of tamoxifen for 48 (tmx 48) and 96 (tmx 96) hours] and *Ezh2*-deficient ES cells (Δ/Δ) were analysed by Western blotting using antibodies to mono-(H2K27m1), di-(H2K27m2), or trimethylated H3K27 (H2K27m3). The deletion of *Suz12* and *Ezh2* was confirmed using antibodies against the Suz12 and Ezh2 proteins. Equivalent protein loading is shown by Lamin B and total H3 detection. The deletion of the PRC2 members Suz12 and Ezh2 in ES cells and the subsequent global reduction of H3K27 methylation levels are highlighted in red. (b) Wild-type (Suz12WT, black bars) and *Suz12*-deficient ES cells (Suz12KO, white bars) and (c) *Ezh2*-1.3 [untreated (flx / flx, black bars) or treated with 800 nM of tamoxifen for 48 (tmx 48, dark grey bars) and 96 (tmx 96, light grey bars) hours] and *Ezh2*-deficient ES cells (Δ/Δ , white bars) were fused to human B-lymphocytes. Successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 3 days after cell fusion. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

5.4. PRC1 cooperates with PRC2 to maintain the reprogramming ability of ES cells

Although the mechanism through which PRC2 inhibits productive transcription remains unclear, it is known that methylated H3K27 provides a binding site for PRC1 (Cao et al., 2002; Hernandez-Munoz et al., 2005; Kuzmichev et al., 2002; Ringrose and Paro, 2004), a multiprotein complex that includes the Ring1A and Ring1B core proteins. Several studies suggested that Ring1A/B function as ubiquitin ligase for H2AK119 (Cao et al., 2005; de Napoles et al., 2004; Wang et al., 2004a) and that this modification may in turn mediate gene repression. To understand whether the reprogramming activity of stem cells is maintained by PRC1 we used the ES-ERT2 ES cell line that carries tamoxifen-inducible conditional knockout *Ring1B* alleles, and is also homozygous *null* for the functional homologue *Ring1A* (Endoh et al., 2008; Stock et al., 2007). Thus, following addition of tamoxifen, ES-ERT2 cells are progressively depleted of Ring1B protein and global H2A ubiquitylation (within 48 hours), whereas overall levels of PRC2 proteins and associated trimethylated H3K27 are largely unaffected (Endoh et al., 2008; Stock et al., 2007). ES-ERT2 ES cells untreated and treated with tamoxifen for 24, 48 and 96 hours were fused to hB-lymphocytes and reprogramming analysed by the expression of the human embryonic genes *hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1* 0 to 3 days (Figure 5.5; qRT-PCR data in Appendix II, Figure A11). Although *hOct4*, *hDnmt3b* and *hTle1* transcription were initiated normally in *Ring1B-null* ES x hB cultures, *hNanog* and *hCripto* activation was substantially reduced (ES-ERT2 following treatment with tamoxifen for 96 hours) and *hSox2*, *hTert* and *hRex1* expression was not detected. The reprogramming reduction seen in ES-ERT2 cells upon tamoxifen treatment is reminiscent of that of ES cells lacking the core components of PRC2, but less dramatic. Genome-wide studies have shown that PRC1 and PRC2 share many (Boyer et al., 2006; Lee et al., 2006; Negre et al., 2006; Tolhuis et al., 2006) but not all genomic targets (Bracken et al., 2006). This observation suggests that the maintenance of reprogramming capacity of ES cells is mediated, at least partially, by cooperation between PRC2 and PRC1 activity.

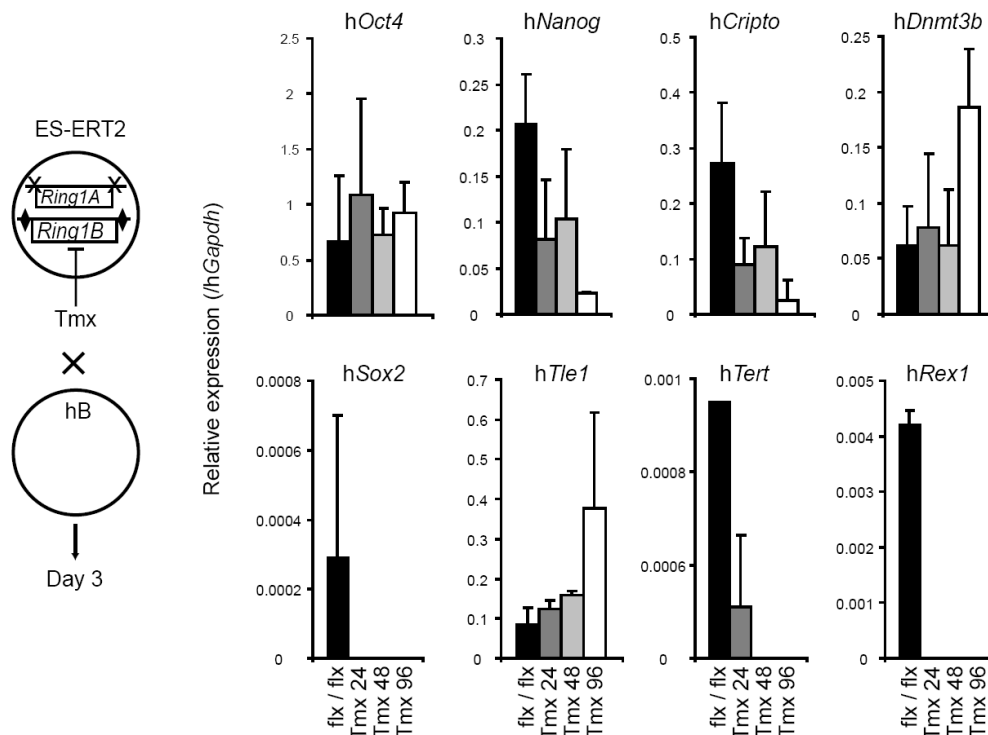


Figure 5.5. Conditional removal of *Ring1B* results in the decreased reprogramming activity of ES cells. ES-Ert2 cells [untreated (flx / flx, black bars) or treated with 800 nM of tamoxifen for 24 (tmx 24, dark grey bars), 48 (tmx 48, light grey bars) and 96 (tmx 96, white bars) hours] were fused to human B-lymphocytes. Successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 3 days after cell fusion. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

5.5. The inability of PcG-deficient ES cells to reprogram is dominant over wild-type ES cells

Polycomb proteins repress many transcriptional regulators which in turn positively and negatively regulate many genes (Azuara et al., 2006; Boyer et al., 2006; Bracken et al., 2006; Jorgensen et al., 2006). To investigate the mechanistic role of Polycomb repressors in reprogramming I asked whether PcG proteins directly or indirectly regulate factors previously implicated in reprogramming (Silva et al., 2006; Takahashi et al., 2007; Yu et al., 2007). Gene expression analysis of *mOct4*, *mNanog*, *mSox2*, *mc-Myc*, *mKlf4* and *mLin28* by quantitative RT-PCR (Figure 5.6a) did not show major differences in expression levels between ES cells lacking the PRC2 components *Eed*, *Suz12* and *Ezh2* (which lack reprogramming activity) and matching wild-type cell lines. Nevertheless, slight variations of *Oct4* levels was shown to have a big impact in reprogramming (see section 4.5), therefore levels of *Oct4* were elevated by the introduction of Flag-tagged mouse *Oct4* in *Eed*-deficient ES cells (B1.3-*Oct4*, Figure 5.6b) and fused these with human B-lymphocytes (Figure 5.6c).

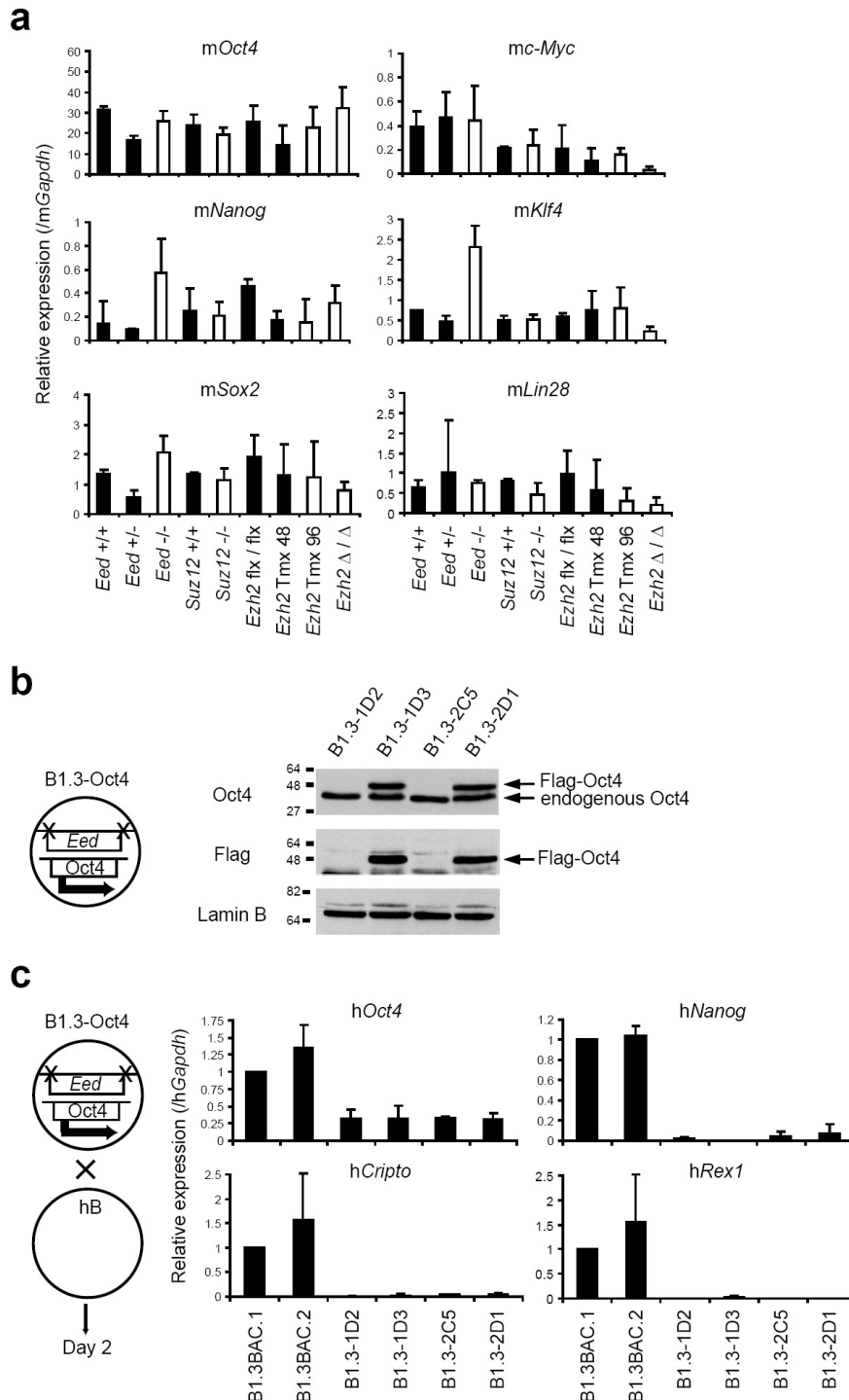


Figure 5.6. PRC2-deficient ES cells do not lack factors implicated in reprogramming. (a) Gene expression analysis by qRT-PCR of *mOct4*, *mNanog*, *mSox2*, *mc-Myc*, *mKlf4* and *mLin28* in ES cells deficient for the PRC2 (white bars) members *Eed* (*Eed* $-/-$), *Suz12* (*Suz12* $-/-$), *Ezh2* (*Ezh2* *tmx* 96; *Ezh2* Δ/Δ) and matching wild-type controls (black bars). (b) *Eed*-deficient ES cells overexpressing Oct4 protein (B1.3-Oct4) were generated by insertion of Flag-tagged mouse *Oct4* cDNA in B1.3 cells (*Eed* $-/-$). Western blotting with anti-Oct4 and anti-Flag antibodies confirmed the presence of Flag-tagged Oct4 protein by transduced cells (B1.3-1D3 and B1.3-2D1) at similar levels as endogenous Oct4 protein. Equivalent protein loading is shown with Lamin B detection. B1.3-Oct4 cells were fused to hB-lymphocytes and the expression of hES-specific genes (*hOct4*, *hNanog*, *hCripto* and *hRex1*) was assessed by qRT-PCR analysis 2 days after cell fusion. Positive (*Eed*-deficient ES cells rescued with a BAC clone containing the *Eed* gene; B1.3BAC.1 and B1.3BAC.2) and negative controls (neomycin resistant *Eed*-deficient ES cells where no exogenous Oct4 protein was detected; B1.3-1D2 and B1.3-2C5) were included for this analysis. Data were normalised to *Gapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

Oct4 overexpression did not rescue the reprogramming defect of *Eed*-deficient ES cells as assessed by the expression analysis of the human embryonic genes *hOct4*, *hCripto*, *hNanog* and *hRex1* by qRT-PCR 2 days after cell fusion. Thus, these experiments suggest that the role of PcG proteins in reprogramming is independent of the pluripotency network of factors in which *Oct4* plays a central role.

In contrast to pluripotency and reprogramming factors, several lineage determinants are de-repressed in ES cells lacking PRC2 activity and H3K27 methylation (Azuara et al., 2006; Boyer et al., 2006; Jorgensen et al., 2006; Lee et al., 2006). These include transcriptional regulators of extra-embryonic lineages such as *Cdx2* and *Hand1* (trophectoderm-specific) or *Gata6* and *Hnf4* (primitive endoderm-specific) [Figure 5.7a and (Jorgensen et al., 2006)]. It is possible that mis-expression of these lineage-specific transcription factors “tip the balance” of reprogramming towards extra-embryonic lineages upon cell fusion with hB-lymphocytes. To directly test this possibility I fused *Eed*-deficient (ESL32), heterozygous (ESL12) and wild-type (ESL21) ES cells with human lymphocytes and analysed the induction of *hCdx2*, *hHand1*, *hGata6* and *hHnf4* (Figure 5.7b). TS cells (Tanaka et al., 1998) and extra-embryonic endoderm stem (XEN) (Tanaka et al., 1998) cells were respectively fused to hB-lymphocytes as positive controls (Santos & Pereira, unpublished data). This analysis showed that lack of *Eed* did not induce reprogramming towards trophectoderm or primitive endoderm, confirming the general inability of *Eed*-deficient ES cells to reprogram lymphocytes.

Failure of *Eed*-deficient ES cells to reprogram has at least 2 explanations: (a) ES cells lack some critical reprogramming factor or (b) mutant ES cells express factors that block or inhibit reprogramming. To distinguish between these possibilities, heterokaryons were generated between human B cells, wild-type mouse ES cells (E14tg2A) and *Eed*-deficient ES cells (B1.3) (Figure 5.8, hB x mES x mES *Eed*^{-/-}). To facilitate the identification of fused cells containing at least 1 lymphocyte-derived, 1 mES-derived and 1 mES *Eed*^{-/-} derived nuclei, mouse E14tg2A ES cells were pre-labelled with DiI and *Eed*^{-/-} ES cells with DiD. The two labelled cell populations and puromycin resistant hB-lymphocytes (see section 2.1.2 for cell line details) were mixed in a 1:1:1 ratio and inter-species heterokaryons were generated in suspension by PEG-mediated cell fusion. After 2 days of culture in the presence of puromycin, dual-stained cells were purified by FACS sorting. Reprogramming was then judged by induction of several human genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hTle1*, *hTert* and *hRex1*) analysed by quantitative RT-PCR.

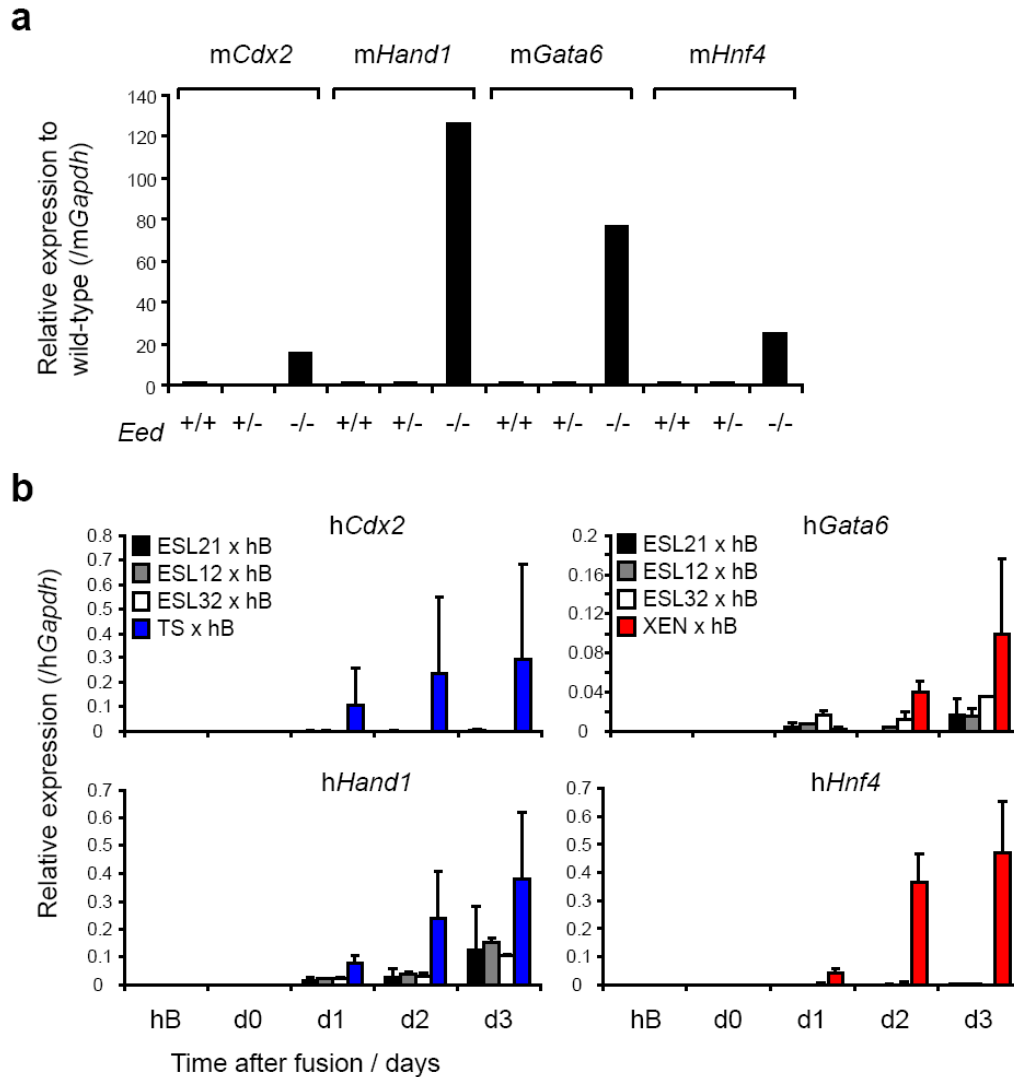


Figure 5.7. De-repression of Polycomb target genes does not induce reprogramming to extra-embryonic lineages. (a) The expression of trophoctoderm-specific (*mCdx2* and *mHand1*) and primitive endoderm-specific (*mGata6* and *mHnf4*) genes was assessed by qRT-PCR in ES cells deficient for *Eed* (ESL32; *Eed* -/-), and matching wild-type (ESL21; *Eed* +/+) and heterozygous (ESL12; *Eed* +/-) controls. *mCdx2*, *mHand1*, *mGata6* and *mHnf4* expression is shown as fold increase relative to the signal obtained in wild-type ES cells. (b) Wild-type (ESL21, black bars), heterozygous (ESL12, grey bars) and *Eed*-deficient ES cells (ESL32, white bars) were fused to human B-lymphocytes. Reprogramming was measured by the activation of human trophoctoderm-specific (*hCdx2* and *hHand1*) and primitive endoderm-specific (*hGata6* and *hHnf4*) human genes by qRT-PCR over 3 days after cell fusion. Trophoctoderm stem cells (TS, blue bars) and extra-embryonic endoderm stem cells (XEN, red bars) were fused to hB-lymphocytes and included as positive controls for this analysis. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

Notably, in heterokaryons where *Eed*-deficient ES cells were present (Figure 5.8; hB x mES x mES *Eed* -/-, highlighted in red) the expression of the majority of reprogrammed human transcripts was reduced approximately 8-fold. This result suggests that *Eed*-deficient ES cells are functionally dominant over their wild-type counterparts. Thus, my data show that PcG proteins maintain ES reprogramming activity by restricting the expression of factors that have a negative impact on this conversion.

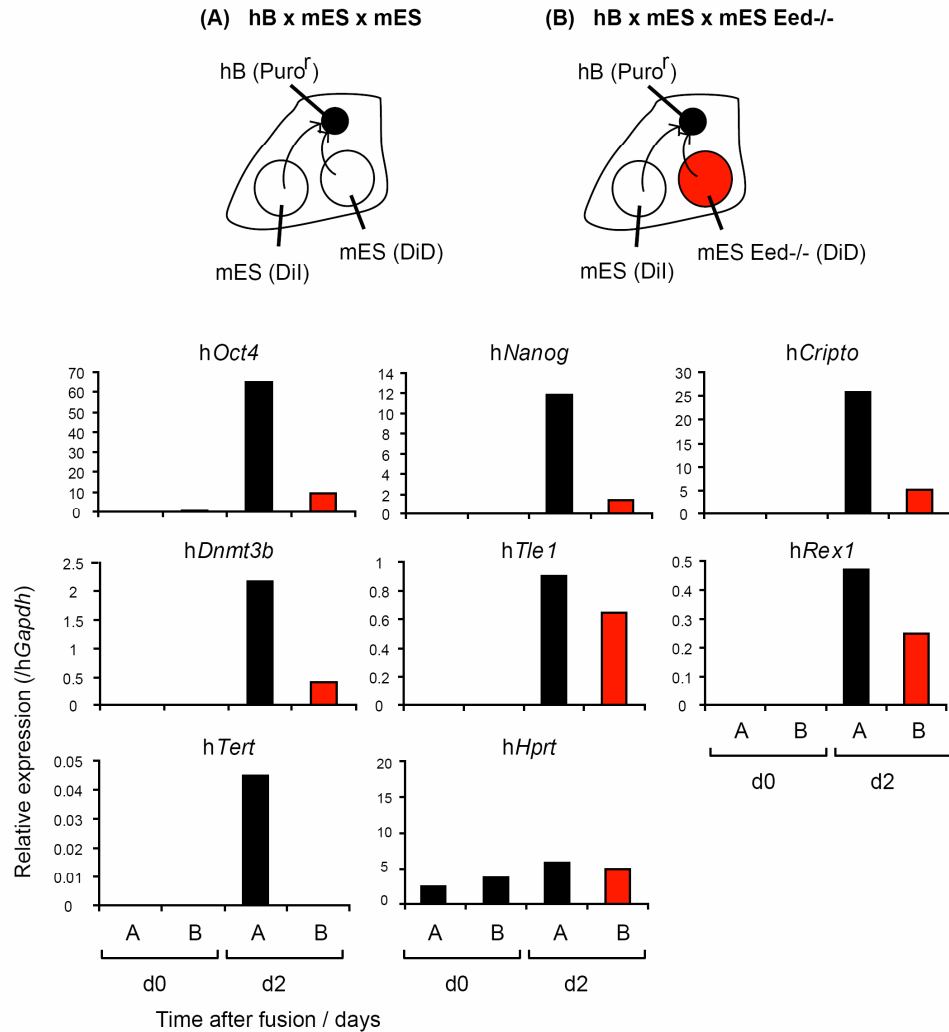


Figure 5.8. *Eed*-deficient ES cells dominate when fused to wild-type ES cells and human lymphocytes. (a) Shows the experimental strategy used to generate inter-species heterokaryons between wild-type ES cells, *Eed*-deficient ES cells and human B-lymphocytes [(B) hB x mES x mES *Eed*^{-/-}; highlighted in red]. E14tg2A mouse embryonic stem cells (mES) and B1.3 *Eed*-deficient ES cells (mES *Eed*^{-/-}, highlighted in red) were respectively labelled with the cell membrane dyes DiI and DiD and fused with puromycin resistant lymphocytes (hB *Puro^R*) in the presence of PEG. Double labelled heterokaryons (resistant to puromycin) were FACS sorted after 2 days of cell fusion. Successful reprogramming was measured by the level of activation of human embryonic genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hTle1*, *hTert* and *hRex1*) by qRT-PCR 0 and 2 days after cell fusion. The control for this analysis included heterokaryons between wild-type ES cells (DiD labelled), wild-type ES cells (DiI labelled) and human B-lymphocytes [(A) hB x mES x mES; black bars]. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression.

5.6. Discussion and future perspectives

Pluripotent stem cells are able to self-renew, generate specialised progeny and reprogram differentiated cells upon hybridoma formation (Do et al., 2007; Tada et al., 1997; Tada et al., 2001). Here, the reprogramming activity of pluripotent ES, EG and EC cells was assessed in embryonic heterokaryons. I show that pluripotent cells share the ability to reprogram human B-lymphocytes, inducing the rapid re-expression of endogenous human genes associated with pluripotency. Using genetically engineered mouse ES cells I demonstrate that the reprogramming activity of stem cells was critically dependent on PcG proteins. Deletion of the PRC2 members Eed, Suz12 and Ezh2 or the PRC1 core components Ring1A/B abolishes ES cell dominant reprogramming of lymphocytes. Remarkably, the reprogramming phenotype of cells with impaired PRC2 function, and therefore lacking H3K27 methylation, is dominant over wild-type ES cells. This result implies that the repressive chromatin mark catalysed by PRC2 play a crucial role in ES cell reprogramming capacity possibly by restricting the expression of factors that inhibit reprogramming.

My results show that activation of endogenous human gene expression program associated with pluripotency by lymphocyte-derived nuclei begins rapidly after fusion with ES, EG or EC cells. These cells share the expression of many pluripotency-associated factors, exogenous requirements to maintain self-renewal and the ability to differentiate into multiple lineages (Andrews et al., 2005; Chambers and Smith, 2004). Previous data showed that when mouse EC cells are fused with somatic cells, notably thymocytes from adult mice, the resulting hybrid cells exhibit the general features of EC cells (Do et al., 2007; Gmur et al., 1980; McBurney, 1977; Miller and Ruddle, 1976; Rousset et al., 1980). My data show that P19 and F9 EC cells were also able to reprogram hB-lymphocytes in heterokaryons similarly to ES cells. EC cells have a more restricted lineage potential than ES cells and the F9 cell line have a propensity to differentiate towards endoderm lineages (Hogan et al., 1981), whereas P19 cells preferentially differentiate towards neuronal lineages (Jones-Villeneuve et al., 1983). The replication timing of a subset of genes was also shown to differ between EC cell lines (Azuara et al., 2006); several neural genes replicated later in F9 cells while *Rex1* replicated later in P19, consistently with the transcriptional inactive status of *Rex1* in P19 cells (Hosler et al., 1989). Do et al (Do et al., 2007) fused F9 or P19 cells with neurosphere cells and the resulting hybrids upregulated endodermal or neural markers, respectively. In line with the previous data, slight but consistent differences were also observed in inter-species heterokaryon analysis: hSox2 and hRex1 were initiated solely

when P19 or F9 cells were respectively used as fusion partners to hB-lymphocytes. Therefore, the reprogramming of precise genes upon heterokaryon formation can reflect the different potential of EC cell lines. EG cells were also able to reprogram hB-lymphocytes in heterokaryons, albeit less efficiently than ES cells. These differences could depend on temporal stem cell origin as EG cells were derived from PGCs at 11.5 dpc and ES derived from ICM of embryos at 3.5 dpc. Contrasting with this observation it was reported that hybrid clones were obtained at a similar rate (2.8×10^{-4}) between EG and mouse thymocytes (Tada et al., 1997) and ES and mouse thymocytes (Tada et al., 2001). It is difficult to know from the experiments presented here whether this discrepancy is due to heterokaryon versus hybrid formation, species differences or even genetic background. This issue will be further addressed in the future. In addition future experiments will examine the potential of EG cells to reset imprinted genes. This may provide further insights into the mechanisms of epigenetic modification in the germ line (Reik et al., 2001). The shared reprogramming potential of different pluripotent cells in heterokaryons implicates dominant reprogramming as a general feature of undifferentiated embryonic cells.

Results show that ES cells lacking PcG proteins lose the ability to induce nuclear reprogramming. PcG proteins were previously characterised in *Drosophila melanogaster* as a series of multiprotein complexes that are crucial for maintaining the inactive state of homeotic, and other important regulators in developing embryos (Lewis, 1978; Pirrotta, 2006; Ringrose and Paro, 2004). Subsequent biochemical analyses have identified distinct Polycomb repressive complexes and revealed the enzymatic activities of some of their components (Levine et al., 2004). In mammals, PRC2 has three central components, Suz12, Eed and Ezh2, a histone methyltransferase that catalyses the methylation of lysine 27 on histone H3 (Cao et al., 2002; Cao and Zhang, 2004; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). The three components of PRC2 are all essential for mouse embryonic development (Erhardt et al., 2003; Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004), but ES cells lacking Eed (Morin-Kensicki et al., 2001), Suz12 (Bracken et al., 2006; Fujimura et al., 2006) and Ezh2 can be derived, despite the lack of global levels of H3K27 methylation. This suggests that PRC2 activity is not essential for ES cell self-renewal. Moreover, ES cells lacking Eed (and PRC2 activity) were recently shown to be functionally pluripotent (Chamberlain et al., 2008). Importantly, my data show that when PRC2 function is compromised (by deletion of Eed, Suz12 or Ezh2), mouse ES cells lose its ability to reprogram human B cells upon cell fusion, arguing the essential function of PRC2 and H3K27 methylation in the reprogramming process. Interestingly, loss of reprogramming correlates with phenotypic severity. The incomplete loss of

H3K27 methylation in *Ezh2 null* ES cells (when compared to *Eed* and *Suz12 null* ES cells) correlates with incomplete loss of reprogramming activity. Both *Eed* and *Suz12* mutations affect the stability of the Ezh2 protein (Montgomery et al., 2005; Pasini et al., 2004) whereas *Ezh2* mutation does not affect the protein levels of Suz12 (and probably not Eed). Therefore partial compensatory effects of the homologous protein Ezh1 (Laible et al., 1997; Su et al., 2003) are likely to occur as the other core components of the complex are still present.

Methylation of H3K27 itself is thought to provide a binding site for PRC1 (Cao et al., 2002; Hernandez-Munoz et al., 2005; Kuzmichev et al., 2002; Ringrose and Paro, 2004), a multiprotein complex that includes the Ring1A and Ring1B core proteins. Several studies suggested that Ring1A/B function as ubiquitin ligase of H2AK119 (Cao et al., 2005; de Napoles et al., 2004; Wang et al., 2004a) and that this histone modification may facilitate gene repression. These observations suggest that PRC1 and PRC2 collaborate in maintaining silent chromatin at developmentally sensitive sites within the genome. According to these findings my results indicate that the PRC1 core components Ring1A/B are also required for reprogramming confirming their co-operative function. However, the reprogramming defect in *Ring1A/B* deficient ES cells was shown to be less pronounced than in PRC2-deficient ES cells. As *Ring1A/B* deletion is incompatible with self-renewal and ES cells in the absence of *Ring1A/B* cannot be stably obtained it is still possible that more time may be required to abolish residual reprogramming activity. In addition, *Ring1A/B*-knockout ES cells show derepression of a number of transcriptional regulators, which are also upregulated in cells that lack PRC2 activity (Endoh et al., 2008; Jorgensen et al., 2006; Stock et al., 2007), but at some target genes (for example, *Math1*), derepression is selectively seen in PRC2 mutant ES cells (Jorgensen et al., 2006). Consistent with this idea, PRC1 binding and H2AK119 ubiquitylation is retained at the inactive X-chromosome even in the absence of PRC2 (Schoeftner et al., 2006).

To the best of my knowledge the involvement of PcG proteins as essential components of the ES cell nuclear reprogramming machinery is a novel finding. Previous studies in *Drosophila melanogaster* have reported the differential regulation of several PcG and TrxG proteins in regenerating leg disc cells (Klebes et al., 2005; Lee et al., 2005; McClure and Schubiger, 2008). Notably these studies have also shown that leg-to-wing induced transdetermination (Maves and Schubiger, 1995) frequency is enhanced in PcG and TrxG heterozygous mutant flies and that downregulation of Polycomb function was observed in transdetermined cells (Lee et al., 2005). These results support the role of PcG and TrxG proteins sustaining lineage-specific transcriptional programs (Ringrose and Paro, 2004), and regulating transdetermination

events (Klebes et al., 2005; Lee et al., 2005; McClure and Schubiger, 2008). My results in turn showed that disruption of Polycomb function in undifferentiated ES cells impair their ability to reprogram differentiated lymphocytes. Taken together these findings highlight the potential binary function of Polycomb in undifferentiated versus somatic cells. It is well established that some Polycomb proteins are over-represented in cancer and stem cells. The distinction however in which Polycomb confer “stemness” to cells, as opposed to the maintenance of differentiated cell fates remain poorly understood. One proposed possibility is that several variations of PcG complexes dictate their different functions (Valk-Lingbeek et al., 2004).

My data suggest that PcG-deficient ES cells are dominant over wild-type ES cells but the underlying molecular mechanism that implicate Polycomb in reprogramming remain unclear and several possibilities deserve consideration. The study of Sustar and Schubiger (Sustar and Schubiger, 2005) reported that transdetermination in *Drosophila* correlates with an extension of the S-phase of cell cycle. It was proposed that this change in cell-cycle regulation is a prerequisite for the change in cell fate. Interestingly, PcG targets include genes involved in cell-cycle regulation (Bracken et al., 2006; Ringrose et al., 2003; Valk-Lingbeek et al., 2004). Moreover, BMI-containing Pc complexes and Brahma-containing trx complexes have been shown to interact with E2F (Trimarchi et al., 2001) and cyclin E (Brumby et al., 2002), respectively, proteins which are essential for S-phase entry. These observations raise the possibility that Polycomb and the S-phase may coordinate a shift in cell commitment. It is not known whether DNA replication plays a role in cell fate conversion in embryonic heterokaryons (discussed in section 4.6). However, this represents a possible mechanism for the involvement of Polycomb proteins in reprogramming and will be further investigated in the future.

Another explanation is that reprogramming is dependent or sensitive to expression of specific transcriptional regulators targets of PcG proteins. Recent genome-wide studies that have mapped the binding sites of PcG components in humans (Bracken et al., 2006; Lee et al., 2006), mice (Boyer et al., 2006) and *Drosophila melanogaster* (Negre et al., 2006; Schwartz et al., 2006; Tolhuis et al., 2006) have shown that many PcG targets encode transcriptional regulators that are inactive in pluripotent cells but become derepressed upon differentiation. Many of these genes carry bivalent chromatin patterns (that is, both “active” and “inactive” histone modifications) in undifferentiated ES cells (Azura et al., 2006; Bernstein et al., 2006). PcG proteins were demonstrated to be important for restricting the expression of these genes, since ES cells that lack PcG proteins showed an inappropriate upregulation of many tissue-specific targets (Azura et al., 2006; Boyer et al., 2006; Jorgensen et al.,

2006). The majority of these genes are transcriptional regulators involved in developmental processes and consequently activate and repress many genes, therefore lacking a factor that “activate” reprogramming is formally possible. My data show that several factors previously implicated in reprogramming are expressed in PcG-deficient ES cells and overexpression of *Oct4* does not rescue the phenotype of *Eed*-deficient ES cells. Importantly, when fused to wild-type ES cells and hB-lymphocytes, *Eed*-deficient cells were shown to be dominant. These results strongly suggest that PcG-deficient ES cells do not lack factors required for reprogramming, neither Polycomb proteins are themselves acting in *trans* in the human lymphocyte nucleus, but mis-expression of some Polycomb target genes may repress reprogramming. Importantly, understanding the roles of such targets and establishing whether these are direct will be essential to rationalise the defect of *PcG* mutants on reprogramming. Among PRC2 direct targets are *Cdx2* and *Hand1* (Boyer et al., 2006; Bracken et al., 2006; Lee et al., 2006), genes involved in trophectoderm specification. Niwa et al. (Niwa et al., 2005) reported that in pre-implantation embryos, *Cdx2* is initially coexpressed with *Oct4* forming a complex for the reciprocal repression of their target genes in ES cells. This suggests that reciprocal inhibition between lineage-specific transcription factors might be involved in the first differentiation steps of mammalian development. Accordingly, *Cdx2* and *Hand1* were derepressed in *Eed*, *Suz12* and at less extent in *Ezh2* deficient ES cells (Appendix II, Figure A12), a result that can be correlated with the reprogramming defect of these cell lines. The possibility that lineage determinants such as *Cdx2* and *Hand1* expressed by PcG-deficient ES cells are directly repressing reprogramming is currently being addressed using RNAi approaches for Polycomb target genes. Whether reprogramming can be restored by the downregulation of such genes in PcG-deficient ES cells is a question that urge to be answered. Recent evidence in the reprogramming of mouse B-lymphocytes and fibroblasts to pluripotency using the iPS strategy provided valuable evidence that lineage regulators can potentially repress reprogramming. For instance, the successful reprogramming of mouse B-lymphocytes (Hanna et al., 2008) require the inactivation of the B cell-specific factor *Pax5* or the forced expression of *C/EBP α* . Both manipulations interfere with the transcriptional network of hB-lymphocytes and has been shown to induce reprogramming into macrophages (Xie et al., 2004). Finally, Mikkelsen et al. (Mikkelsen et al., 2008) suggested that the re-activation or incomplete repression of lineage-specifying factors during the reprogramming process may block the endogenous pluripotency regulatory network. Transient silencing of one or more of these factors, combined with demethylation drugs, improved reprogramming efficiency.

In summary pluripotent stem cells are able to self-renew, generate differentiated progeny and reprogram differentiated cells upon cell fusion. Here I have shown that ES, EG and EC cells share the ability to reprogram differentiated cells in heterokaryons and that this unique feature can be uncoupled from self-renewal and pluripotency by the deletion of Polycomb repression in undifferentiated ES cells.

CHAPTER 6. GENERAL DISCUSSION

During development, cells respond to signals that commit them to a developmental pathway. Transplantation and culture experiments have shown that cellular memory is stable and can be propagated through cell division. In this study, however, I have examined the requirements to override or reverse the differentiated state using experimental heterokaryons.

6.1. Inter-species heterokaryons as an experimental system for reprogramming studies

Although under normal circumstances the pattern of gene expression of differentiated cells is stable and heritable, specialisation can be reversed or changed if regulatory circuits are altered by introduction of exogenous factors or by fusion with different cell partners. In this context, my study demonstrates the usefulness of experimental heterokaryons not only for previous discoveries, for example in cancer biology (Harris et al., 1969), monoclonal antibody production (Kohler and Milstein, 1975), gene activation (Baron and Maniatis, 1986; Blau et al., 1983) and cell cycle (Rao and Johnson, 1970), but also for understanding the plasticity of the differentiated and pluripotent state. In this study, I exploit heterokaryons formed between cell-types of different species to allow the distinction of the events that occur in the donor nucleus from those of the recipient. The use of inter-species heterokaryons provides a useful approach to uncover the mechanisms underlying reprogramming. It is possible to analyse the epigenetic changes at the reprogrammed nucleus over time, including nuclear organisation and DNA/chromatin modifications, in addition to examining the requirement for the type or concentration of *trans*-acting factors. This approach coupled with pharmacological inhibition, gene knock-out or knock-down allows factors that are required to directly induce reprogramming to be defined individually, as well as in combination. My study shows that inter-species heterokaryons offer a powerful approach to trace the contribution of individual factors for direct reprogramming.

6.2. Acquisition and extinction of gene expression programs are separable events during reprogramming

In this study I show that dominant conversion of human lymphocytes to muscle is a multi-step process that includes changes in nuclear architecture and gene activity. Upon heterokaryon formation, nuclei of lymphocytes adopt an architecture that resembles that of muscle and initiates muscle-specific gene expression, in an order that recapitulates ordered gene expression in developing muscle. The establishment of this muscle-specific program was coordinated with the shutdown of several lymphocyte-associated genes. Inhibition of HDACs during reprogramming selectively blocks the silencing of lymphocyte-specific genes, but does not prevent the establishment of muscle-specific gene expression. My results reinforce the established role of HDACs in gene silencing, but also show that the establishment of muscle gene expression and the extinction of lymphocyte identity are distinct and separable components of lineage conversion (Figure 6.1). It is likely that a distinct set of factors are implicated in either the silencing or activation of gene expression programs during reprogramming. In addition to HDACs, transcriptional repressors, HMTs and DNA methyltransferases are also expected to be involved in silencing of lineage inappropriate genes (Figure 6.1, top panel). Conversely, the activation of the novel program of gene expression should be mediated by HATs and histone demethylase activity in concert with sequence-specific transcriptional activators (Figure 6.1, bottom panel). The individual role of each of those factors during lineage conversion of lymphocytes to muscle may be addressed in the future. In 2004, Xie and colleagues reported that enforced expression of *C/EBP α* and *C/EBP β* in differentiated B-lymphocytes led to their reprogramming into macrophages (Xie et al., 2004). C/EBPs induced these changes by inhibiting the B cell commitment transcription factor *Pax5*, leading to the downregulation of its target CD19, and synergising with endogenous *PU.1*, leading to the upregulation of its target Mac-1 and other myeloid markers. The two processes were also shown to be mechanistically separable, since in *PU.1*-deficient pre-B cells, C/EBPs induced CD19 downregulation but not Mac-1 activation. The activation of macrophage gene expression is impaired by limiting the availability of a key transcriptional activator while extinction of lymphocyte identity is unaffected. Similarly, my results show that *Oct4* withdrawal in embryonic heterokaryons results in the impaired activation of pluripotency-associated genes but not the silencing of lymphocyte-specific genes. These observations provide complementary proof that

activation and silencing are indeed mechanistically distinct components of nuclear reprogramming (Figure 6.1).

The molecular mechanisms that prevent the expression of lineage inappropriate genes by differentiated cells (phenotypic exclusion) remain poorly understood (Fougere and Weiss, 1978). Here, I show that two different programs became co-expressed (usually mutually exclusive) upon HDAC inhibition, providing some molecular evidence for the processes that control phenotypic exclusion. Furthermore, my results underscore the importance of gene silencing for successful reprogramming and suggest that, in future, it may be possible to preserve certain characteristics of donor cells through lineage conversion, by restricting the availability of HDACs.

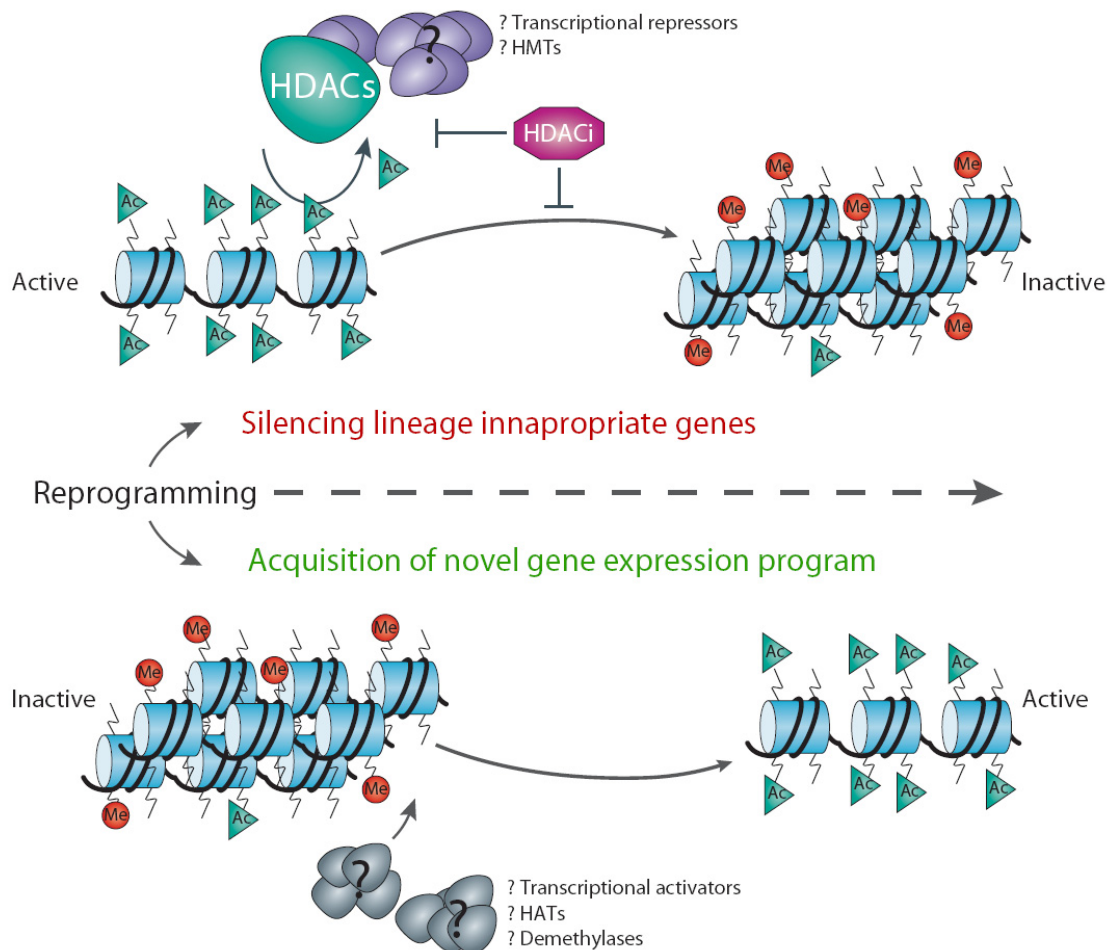


Figure 6.1. Molecular mechanisms underlying acquisition and silencing of gene expression programs during reprogramming. Reprogramming of one cell-type to another involves the establishment of the novel program of gene expression (bottom panel) and the coordinated silencing of lineage inappropriate genes (top panel). Pharmacological inhibition (HDACi) of histone deacetylase (HDAC) activity during the reprogramming of B-lymphocytes to muscle selectively blocks the silencing of lymphocyte-specific genes, but does not prevent the establishment of muscle-specific gene expression. In addition to HDACs, other regulators might be implicated in the extinction of gene expression programs (top panel, question marks) which are probably different from those required for *de novo* gene activation (bottom panel, question marks). Ac – histone acetylation; Me – histone methylation; HMTs – histone methyltransferases; HATs – histone acetyltransferases.

6.3. Implications of reprogramming to pluripotency in heterokaryons

One of the most pressing objectives of medical research today is the development of approaches to restore the function of damaged tissues. An important goal for this work is the isolation of stem cell populations to replace missing or non-functioning cells. Because problems of immune rejection are likely to occur unless the recipient and donor stem cells are very closely matched, a desirable strategy is to convert differentiated cells from patients into immature tailored stem cell populations. Here, I provide evidence that transcriptional and epigenetic conversion of human B-lymphocytes towards a multipotent state is initiated much more rapidly than previously thought, occurring in transient heterokaryons before nuclear fusion and cell division. Reprogramming of human lymphocytes by mouse ES cells elicits the expression of a human ES-specific gene profile, in which markers of human ES cells are expressed but markers that are specific to mouse ES cells are not. Once reprogramming is initiated by factors produced by the dominant mouse ES cell nucleus, is both self-sustaining and heritable. Collectively these results may have an impact for generating human ES-like cells. Proof that mouse ES cells can dominantly reset the multi-lineage potential in human somatic cells, together with evidence that this process begins prior to nuclear fusion and cell division, suggests that improved methods for removing the mouse nucleus from heterokaryons (Pralong et al., 2005) may be applicable for generating human stem cell lines without genetic alterations.

6.4. Mechanistic aspects of reprogramming towards pluripotency

In this study I show that mouse pluripotent cells share the ability to reprogram human B-lymphocytes in heterokaryons, inducing the rapid re-expression of endogenous human genes associated with pluripotency. Using genetically engineered mouse ES cells, I demonstrate that the reprogramming activity of ES cells was critically dependent on the transcription factor Oct4 and PcG proteins (Figure 6.2). However, successful reprogramming of human lymphocytes occurs independently of Sox2, a factor thought to be required for the generation of iPS cells from somatic cells. More recent studies reported the derivation of iPS cells from neural stem cells independently of Sox2 (Di Stefano et al., 2008; Eminli et al., 2008; Kim et al., 2008b). However it is

not clear whether this factor is not required for reprogramming since neural stem cells endogenously express Sox2. Masui et al. have suggested that the essential function of Sox2 is to stabilise ES cells by maintaining the adequate level of *Oct4* expression (Masui et al., 2007). Therefore we propose that Sox2 is not an essential factor for the induction of pluripotency and can be potentially substituted in the future with other factors (Figure 6.2) (Pereira et al., 2008). The reported combination of factors is sufficient to initiate the reprogramming cascade but which of those are essential for the reprogramming process? It is possible that *Oct4*, which in normal development is already expressed in the oocyte, may be the most upstream gene in the molecular circuitry of pluripotency, and the only obligatory factor to initiate reprogramming. It appears that *c-Myc* significantly enhances and accelerates the process but is dispensable in both human and mouse iPS cells. How *c-Myc* function at the molecular level to enhance reprogramming remains unclear. Interestingly, a recent study suggests that, in addition to its transcription factor activity, *c-Myc* is also implicated in the induction of DNA replication (Dominguez-Sola et al., 2007). Since the majority of factors implicated in reprogramming (*Sox2*, *c-Myc*, *Lin28* and *Nanog*) can be dispensable or replaced, they could serve to increase the efficiency of re-wiring the pluripotency network. However, addressing the role of each of these factors individually in heterokaryons would increase our understanding of which factors are essential versus which are sufficient for reprogramming. The iPS studies have illustrated the importance of combinations of factors for reprogramming, but also suggested that additional ones may be needed for efficient conversion to pluripotency. In this study, I show that ES cell-mediated reprogramming is dependent on PRC2 function. The catalysis of the methylation mark on histone H3 lysine 27 by PRC2 represses a range of developmentally-associated genes in ES cells, which when inappropriately expressed may inhibit reprogramming. This can change our view of the reprogramming of somatic cells to the pluripotent state. In addition to the activation of pluripotency-associated gene transcripts in somatic cells, the efficient repression of genes involved in the determination of unrelated lineages might be required for successful reprogramming (Figure 6.2). This provides a rationale for using distinct protein cocktails to directly re-set lineage potential. Thus the PRC2 members, *Eed*, *Suz12* and *Ezh2*, are strong candidates to be directly introduced into somatic cells along with *Oct4* to increase the efficiency of the reprogramming process.

The network of factors implicated in the reprogramming of a somatic cell to pluripotency remains obscure. Experimental heterokaryons coupled with conditionally targeted mouse ES cells offer a powerful approach to trace the contribution of individual factors. Furthermore, knocking down *Oct4* expression in a wild-type ES cell

line with siRNAs, similarly to conditional knock-out ES cells, abolished reprogramming (Appendix II, Figure A13). This provides proof-of-principle that RNAi approaches can be used for screening reprogramming factors. RNAi provide a more feasible approach to unbiasedly identify the network of factors essential for reprogramming to pluripotency. When the optimal combination is found, new methods for the introduction of proteins in cells without any underlying genetic modification must be developed. The use of fusion proteins with the HIV TAT-sequence can render cell-permeant reprogramming factors (Bosnali and Edenhofer, 2008; Glover et al., 2005) and represent an attractive alternative for the orchestration of pluripotency in somatic cells.

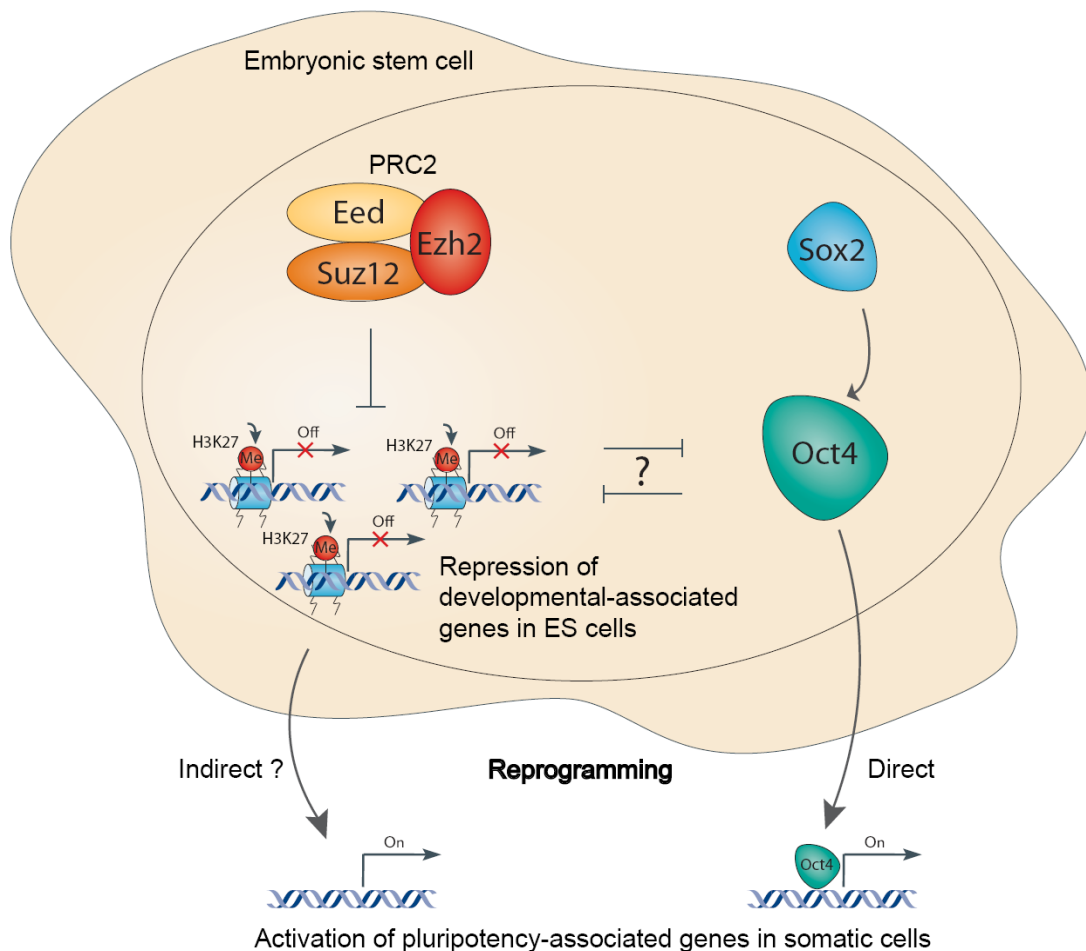


Figure 6.2. A model for the role of Oct4, Sox2 and PRC2 in the reprogramming to pluripotency. Upon cell fusion, embryonic stem cells have the ability to reprogram somatic cells, inducing the rapid re-expression of endogenous genes associated with pluripotency (lower panel). Genetically engineered mouse ES cells were used to address the requirements for conversion of somatic cells to pluripotency. Reprogramming is critically dependent on Oct4, since *Oct4* deletion abolishes the reprogramming capacity of ES cells. This effect is thought to be mediated by the direct activation of Oct4 target genes, essential for the pluripotent-state. Sox2, in contrast to *Oct4*, is not required to convert human B-lymphocytes into a multi-potent state but instead is thought to maintain the adequate *Oct4* levels in ES cells (right panel). Reprogramming activity of stem cells is reliant on Polycomb proteins. Deletion of the Polycomb repressive complex 2 (PRC2) components *Eed*, *Suz12* and *Ezh2* and the histone mark that this complex catalyses (H3K27 – histone H3 lysine 27) impairs reprogramming (left panel). Although the mechanism of action is not clear, it is possible that PRC2 restricts the expression of developmental-associated genes that inhibit ES cell reprogramming network (question mark).

6.5. Epigenetic stability and flexibility of cellular states in development

A common notion in developmental biology is that cell differentiation proceeds through a series of successive binary decisions by which cells adopt alternative phenotypes. Linked to this concept, although by no means essential, is that differentiation is irreversible. However, evidence for *in vivo* and *in vitro* reprogramming continues to challenge these concepts of unidirectional, irreversible differentiation along distinct cell lineages. At the molecular level, the stability of the differentiated state must be ensured by a complex, dynamic network of factors. What are these mechanisms and how are they actively maintained and yet reversed with relative ease during reprogramming? Although autoregulation of regulators seems to play a key role in stability, recent advances in chromatin biology have considerably broadened our understanding of the mechanisms of cell commitment and specification. We now know that there can be different “levels” of regulation defined by different chromatin configurations that can be inherited through rounds of cell division. Recent studies, however, have revealed that the balance between the active and inactive chromatin states in the cell is surprisingly flexible, raising questions as to how chromatin changes in development should be interpreted. Thus, a full understanding of the role of chromatin in developmental plasticity will require knowledge of times, reaction rates, concentrations, half-lives, dissociation constants, and other variables that explain how the reversal of individual epigenetic modifications is achieved at the right place for the right length of time. Perhaps a substantial challenge from the recent observations of nuclear reprogramming will be to develop a more detailed understanding of the chemistry underlying the complex network of factors, to provide a more solid basis for the concept of epigenetic reprogramming.

Taken together my findings highlight the importance of chromatin events during cell fate reprogramming. The proposed future experiments aim to increase our understanding of molecular processes underlying epigenetic reprogramming, and may provide promising perspectives for the development of direct reprogramming strategies.

REFERENCES

- Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. *Cell* 111, 381-392.
- Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., and Cardoso, M.C. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res* 35, 5402-5408.
- Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 449, 731-734.
- Ahmad, K., and Henikoff, S. (2002a). Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16477-16484.
- Ahmad, K., and Henikoff, S. (2002b). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* 9, 1191-1200.
- Ait-Si-Ali, S., Guasconi, V., Fritsch, L., Yahi, H., Sekhri, R., Naguibneva, I., Robin, P., Cabon, F., Polesskaya, A., and Harel-Bellan, A. (2004). A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells. *Embo J* 23, 605-615.
- Alcobia, I., Dilao, R., and Parreira, L. (2000). Spatial associations of centromeres in the nuclei of hematopoietic cells: evidence for cell-type-specific organizational patterns. *Blood* 95, 1608-1615.
- Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* 51, 786-794.
- Ambrosi, D.J., Tanasijevic, B., Kaur, A., Obergfell, C., O'Neill, R.J., Krueger, W., and Rasmussen, T.P. (2007). Genome-wide reprogramming in hybrids of somatic cells and embryonic stem cells. *Stem Cells* 25, 1104-1113.
- Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Developmental biology* 227, 271-278.
- Amit, M., Shariki, C., Margulets, V., and Itskovitz-Eldor, J. (2004). Feeder layer- and serum-free culture of human embryonic stem cells. *Biol Reprod* 70, 837-845.
- Andrews, P.W., Matin, M.M., Bahrami, A.R., Damjanov, I., Gokhale, P., and Draper, J.S. (2005). Embryonic stem (ES) cells and embryonal carcinoma (EC) cells: opposite sides of the same coin. *Biochem Soc Trans* 33, 1526-1530.
- Andrulis, E.D., Neiman, A.M., Zappulla, D.C., and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* 394, 592-595.
- Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of Pluripotent Stem Cells from Adult Mouse Liver and Stomach Cells. *Science*, DOI: 10.1126.
- Aparicio, J.G., Viggiani, C.J., Gibson, D.G., and Aparicio, O.M. (2004). The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in *Saccharomyces cerevisiae*. *Mol Cell Biol* 24, 4769-4780.
- Arney, K.L., and Fisher, A.G. (2004). Epigenetic aspects of differentiation. *J Cell Sci* 117, 4355-4363.
- Aurade, F., Pinset, C., Chafey, P., Gros, F., and Montarras, D. (1994). Myf5, MyoD, myogenin and MRF4 myogenic derivatives of the embryonic mesenchymal cell line C3H10T1/2 exhibit the same adult muscle phenotype. *Differentiation; research in biological diversity* 55, 185-192.

- Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 17, 126-140.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M., *et al.* (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-538.
- Balsam, L.B., Wagers, A.J., Christensen, J.L., Kofidis, T., Weissman, I.L., and Robbins, R.C. (2004). Haematopoietic stem cells adopt mature haematopoietic fates in ischaemic myocardium. *Nature* 428, 668-673.
- Bannister, A.J., and Kouzarides, T. (2005). Reversing histone methylation. *Nature* 436, 1103-1106.
- Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M., and Cavalli, G. (2003). Inheritance of Polycomb-dependent chromosomal interactions in *Drosophila*. *Genes Dev* 17, 2406-2420.
- Baron, M.H. (1993). Reversibility of the differentiated state in somatic cells. *Curr Opin Cell Biol* 5, 1050-1056.
- Baron, M.H., and Farrington, S.M. (1994). Positive regulators of the lineage-specific transcription factor GATA-1 in differentiating erythroid cells. *Mol Cell Biol* 14, 3108-3114.
- Baron, M.H., and Maniatis, T. (1986). Rapid reprogramming of globin gene expression in transient heterokaryons. *Cell* 46, 591-602.
- Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., *et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675.
- Barry, J.M., and Merriam, R.W. (1972). Swelling of hen erythrocyte nuclei in cytoplasm from *Xenopus* eggs. *Exp Cell Res* 71, 90-96.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.
- Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells* 23, 489-495.
- Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. *Science* 301, 1090-1093.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of molecular biology* 203, 971-983.
- Bestor, T.H. (2000). The DNA methyltransferases of mammals. *Human molecular genetics* 9, 2395-2402.
- Bhaumik, S.R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14, 1008-1016.
- Billon, N., Jolicoeur, C., Ying, Q.L., Smith, A., and Raff, M. (2002). Normal timing of oligodendrocyte development from genetically engineered, lineage-selectable mouse ES cells. *J Cell Sci* 115, 3657-3665.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.
- Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. *Cell* 99, 451-454.

- Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816.
- Blander, G., and Guarente, L. (2004). The Sir2 family of protein deacetylases. *Annual review of biochemistry* **73**, 417-435.
- Blau, H.M. (1992). Regulating the myogenic regulators. *Symp Soc Exp Biol* **46**, 9-18.
- Blau, H.M., and Baltimore, D. (1991). Differentiation requires continuous regulation. *The Journal of cell biology* **112**, 781-783.
- Blau, H.M., Chiu, C.P., and Webster, C. (1983). Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* **32**, 1171-1180.
- Blau, H.M., Pavlath, G.K., Hardeman, E.C., Chiu, C.P., Silberstein, L., Webster, S.G., Miller, S.C., and Webster, C. (1985). Plasticity of the differentiated state. *Science* **230**, 758-766.
- Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells* **24**, 2007-2013.
- Bode, J., Henco, K., and Wingender, E. (1980). Modulation of the nucleosome structure by histone acetylation. *European journal of biochemistry / FEBS* **110**, 143-152.
- Boisvert, F.M., Chenard, C.A., and Richard, S. (2005). Protein interfaces in signaling regulated by arginine methylation. *Sci STKE* **2005**, re2.
- Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. *Nature reviews* **5**, 769-784.
- Bortvin, A., Eggan, K., Skaletsky, H., Akutsu, H., Berry, D.L., Yanagimachi, R., Page, D.C., and Jaenisch, R. (2003). Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**, 1673-1680.
- Bosnali, M., and Edenhofer, F. (2008). Generation of transducible versions of transcription factors Oct4 and Sox2. *Biol Chem* **389**, 851-861.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956.
- Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353.
- Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**, 1123-1134.
- Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* **20**, 1123-1136.
- Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151-159.
- Brambrink, T., Hochedlinger, K., Bell, G., and Jaenisch, R. (2006). ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. *Proc Natl Acad Sci U S A* **103**, 933-938.
- Brero, A., Easwaran, H.P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., and Cardoso, M.C. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *The Journal of cell biology* **169**, 733-743.
- Brickner, J.H., and Walter, P. (2004). Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS biology* **2**, e342.

- Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. *Proc Natl Acad Sci U S A* 38, 455-463.
- Brockes, J.P., and Kumar, A. (2008). Comparative Aspects of Animal Regeneration. *Annu Rev Cell Dev Biol*.
- Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., *et al.* (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191-195.
- Brown, K.E., Amoils, S., Horn, J.M., Buckle, V.J., Higgs, D.R., Merckenschlager, M., and Fisher, A.G. (2001). Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. *Nat Cell Biol* 3, 602-606.
- Brown, K.E., Baxter, J., Graf, D., Merckenschlager, M., and Fisher, A.G. (1999). Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. *Mol Cell* 3, 207-217.
- Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merckenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* 91, 845-854.
- Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843-851.
- Bru, T., Clarke, C., McGrew, M.J., Sang, H.M., Wilmut, I., and Blow, J.J. (2008). Rapid induction of pluripotency genes after exposure of human somatic cells to mouse ES cell extracts. *Exp Cell Res* 314, 2634-2642.
- Bruce, K., Myers, F.A., Mantouvalou, E., Lefevre, P., Greaves, I., Bonifer, C., Tremethick, D.J., Thorne, A.W., and Crane-Robinson, C. (2005). The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. *Nucleic Acids Res* 33, 5633-5639.
- Brumby, A.M., Zrally, C.B., Horsfield, J.A., Secombe, J., Saint, R., Dingwall, A.K., and Richardson, H. (2002). Drosophila cyclin E interacts with components of the Brahma complex. *Embo J* 21, 3377-3389.
- Buchenau, P., Hodgson, J., Strutt, H., and Arndt-Jovin, D.J. (1998). The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing. *The Journal of cell biology* 141, 469-481.
- Bultman, S.J., Gebuhr, T.C., Pan, H., Svoboda, P., Schultz, R.M., and Magnuson, T. (2006). Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes Dev* 20, 1744-1754.
- Burke, B., and Stewart, C.L. (2002). Life at the edge: the nuclear envelope and human disease. *Nat Rev Mol Cell Biol* 3, 575-585.
- Busslinger, M., Nutt, S.L., and Rolink, A.G. (2000). Lineage commitment in lymphopoiesis. *Current opinion in immunology* 12, 151-158.
- Byrne, J.A., Pedersen, D.A., Clepper, L.L., Nelson, M., Sanger, W.G., Gokhale, S., Wolf, D.P., and Mitalipov, S.M. (2007). Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* 450, 497-502.
- Byrne, J.A., Simonsson, S., Western, P.S., and Gurdon, J.B. (2003). Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. *Curr Biol* 13, 1206-1213.
- Cai, L., Friedman, N., and Xie, X.S. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature* 440, 358-362.
- Camargo, M., and Cervenka, J. (1982). Patterns of DNA replication of human chromosomes. II. Replication map and replication model. *American journal of human genetics* 34, 757-780.
- Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* 20, 845-854.

- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043.
- Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* 15, 57-67.
- Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nat Genet* 32, 623-626.
- Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117, 427-439.
- Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: shaping the landscape of the genome. *Nat Rev Genet* 6, 669-677.
- Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* 26, 1496-1505.
- Chambers, A.E., Banerjee, S., Chaplin, T., Dunne, J., Debernardi, S., Joel, S.P., and Young, B.D. (2003a). Histone acetylation-mediated regulation of genes in leukaemic cells. *Eur J Cancer* 39, 1165-1175.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003b). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113, 643-655.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450, 1230-1234.
- Chambers, I., and Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* 23, 7150-7160.
- Chen, T., Ueda, Y., Xie, S., and Li, E. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *The Journal of biological chemistry* 277, 38746-38754.
- Chen, Y., He, Z.X., Liu, A., Wang, K., Mao, W.W., Chu, J.X., Lu, Y., Fang, Z.F., Shi, Y.T., Yang, Q.Z., *et al.* (2003). Embryonic stem cells generated by nuclear transfer of human somatic nuclei into rabbit oocytes. *Cell Res* 13, 251-263.
- Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25, 6031-6046.
- Chiu, C.P., and Blau, H.M. (1984). Reprogramming cell differentiation in the absence of DNA synthesis. *Cell* 37, 879-887.
- Chiu, C.P., and Blau, H.M. (1985). 5-Azacytidine permits gene activation in a previously noninducible cell type. *Cell* 40, 417-424.
- Choi, J., Costa, M.L., Mermelstein, C.S., Chagas, C., Holtzer, S., and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. *Proc Natl Acad Sci U S A* 87, 7988-7992.
- Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277, 1996-2000.
- Cimbora, D.M., and Groudine, M. (2001). The control of mammalian DNA replication: a brief history of space and timing. *Cell* 104, 643-646.

- Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309, 1369-1373.
- Craig, J.M. (2005). Heterochromatin--many flavours, common themes. *Bioessays* 27, 17-28.
- Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *The Journal of cell biology* 162, 809-820.
- Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. *Curr Opin Cell Biol* 18, 307-316.
- Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of cell biology* 145, 1119-1131.
- Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185-196.
- Daheron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Andrews, P.W., Itskovitz-Eldor, J., and Daley, G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem Cells* 22, 770-778.
- Daniel, J.A., Pray-Grant, M.G., and Grant, P.A. (2005). Effector proteins for methylated histones: an expanding family. *Cell cycle (Georgetown, Tex)* 4, 919-926.
- Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987-1000.
- de la Serna, I.L., Carlson, K.A., and Imbalzano, A.N. (2001). Mammalian SWI/SNF complexes promote MyoD-mediated muscle differentiation. *Nat Genet* 27, 187-190.
- de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., *et al.* (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* 7, 663-676.
- de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. *The Biochemical journal* 370, 737-749.
- Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M., and Walker, P.R. (2005). Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. *The Journal of biological chemistry* 280, 6257-6260.
- Dejosez, M., Krumenacker, J.S., Zitur, L.J., Passeri, M., Chu, L.F., Songyang, Z., Thomson, J.A., and Zwaka, T.P. (2008). Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell* 133, 1162-1174.
- Dekel, I., Magal, Y., Pearson-White, S., Emerson, C.P., and Shani, M. (1992). Conditional conversion of ES cells to skeletal muscle by an exogenous MyoD1 gene. *The New biologist* 4, 217-224.
- Delaire, S., Huang, Y.H., Chan, S.W., and Robey, E.A. (2004). Dynamic repositioning of CD4 and CD8 genes during T cell development. *The Journal of experimental medicine* 200, 1427-1435.
- Dellino, G.I., Schwartz, Y.B., Farkas, G., McCabe, D., Elgin, S.C., and Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. *Mol Cell* 13, 887-893.
- Di Stefano, B., Prigione, A., and Broccoli, V. (2008). Efficient genetic reprogramming of unmodified somatic neural progenitors uncovers the essential requirement of Oct4 and Klf4. *Stem Cells Dev.*

- DiBerardino, M.A. (1988). Genomic multipotentiality of differentiated somatic cells. *Cell Differ Dev* 25 *Suppl*, 129-136.
- Dillon, N., and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet* 18, 252-258.
- Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., *et al.* (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321, 1218-1221.
- Do, J.T., Han, D.W., Gentile, L., Sobek-Klocke, I., Stehling, M., Lee, H.T., and Scholer, H.R. (2007). Erasure of cellular memory by fusion with pluripotent cells. *Stem Cells* 25, 1013-1020.
- Do, J.T., and Scholer, H.R. (2004). Nuclei of embryonic stem cells reprogram somatic cells. *Stem Cells* 22, 941-949.
- Dominguez-Sola, D., Ying, C.Y., Grandori, C., Ruggiero, L., Chen, B., Li, M., Galloway, D.A., Gu, W., Gautier, J., and Dalla-Favera, R. (2007). Non-transcriptional control of DNA replication by c-Myc. *Nature* 448, 445-451.
- Dupuy-Coin, A.M., Ege, T., Bouteille, M., and Ringertz, N.R. (1976). Ultrastructure of chick erythrocyte nuclei undergoing reactivation in heterokaryons and enucleated cells. *Exp Cell Res* 101, 355-369.
- Earnshaw, W., Bordwell, B., Marino, C., and Rothfield, N. (1986). Three human chromosomal autoantigens are recognized by sera from patients with anti-centromere antibodies. *J Clin Invest* 77, 426-430.
- Edelmann, P., Bornfleth, H., Zink, D., Cremer, T., and Cremer, C. (2001). Morphology and dynamics of chromosome territories in living cells. *Biochimica et biophysica acta* 1551, M29-39.
- Eden, E.R., Patel, D.D., Sun, X.M., Burden, J.J., Themis, M., Edwards, M., Lee, P., Neuwirth, C., Naoumova, R.P., and Soutar, A.K. (2002). Restoration of LDL receptor function in cells from patients with autosomal recessive hypercholesterolemia by retroviral expression of ARH1. *J Clin Invest* 110, 1695-1702.
- Edwards, C.A., Rens, W., Clarke, O., Mungall, A.J., Hore, T., Graves, J.A., Dunham, I., Ferguson-Smith, A.C., and Ferguson-Smith, M.A. (2007). The evolution of imprinting: chromosomal mapping of orthologues of mammalian imprinted domains in monotreme and marsupial mammals. *BMC evolutionary biology* 7, 157.
- Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoepfner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., *et al.* (2008). Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* 2, 437-447.
- Egli, D., Birkhoff, G., and Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat Rev Mol Cell Biol* 9, 505-516.
- Eglitis, M.A., and Mezey, E. (1997). Hematopoietic cells differentiate into both microglia and macroglia in the brains of adult mice. *Proc Natl Acad Sci U S A* 94, 4080-4085.
- Eguchi, G., and Kodama, R. (1993). Transdifferentiation. *Curr Opin Cell Biol* 5, 1023-1028.
- Ekwall, K., Olsson, T., Turner, B.M., Cranston, G., and Allshire, R.C. (1997). Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* 91, 1021-1032.
- Emerson, C.P. (1990). Myogenesis and developmental control genes. *Curr Opin Cell Biol* 2, 1065-1075.
- Eminli, S., Utikal, J.S., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of Neural Progenitor Cells into iPS Cells in the Absence of Exogenous Sox2 Expression. *Stem Cells*.

- Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M., *et al.* (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513-1524.
- Ephrussi, B. (1972). *Hibridization of somatic cells* (Princeton, NJ, Princeton University Press).
- Erhardt, S., Su, I.H., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M.A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* **130**, 4235-4248.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
- Faerman, A., Pearson-White, S., Emerson, C., and Shani, M. (1993). Ectopic expression of MyoD1 in mice causes prenatal lethalties. *Dev Dyn* **196**, 165-173.
- Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Mol Cell* **16**, 655-661.
- Farris, S.D., Rubio, E.D., Moon, J.J., Gombert, W.M., Nelson, B.H., and Krumm, A. (2005). Transcription-induced chromatin remodeling at the c-myc gene involves the local exchange of histone H2A.Z. *The Journal of biological chemistry* **280**, 25298-25303.
- Faust, C., Lawson, K.A., Schork, N.J., Thiel, B., and Magnuson, T. (1998). The Polycomb-group gene *ee* is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development* **125**, 4495-4506.
- Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**, 188-194.
- Finnin, M.S., Donigian, J.R., Cohen, A., Richon, V.M., Rifkind, R.A., Marks, P.A., Breslow, R., and Pavletich, N.P. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* **401**, 188-193.
- Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* **17**, 1870-1881.
- Flasza, M., Shering, A.F., Smith, K., Andrews, P.W., Talley, P., and Johnson, P.A. (2003). Reprogramming in inter-species embryonal carcinoma-somatic cell hybrids induces expression of pluripotency and differentiation markers. *Cloning Stem Cells* **5**, 339-354.
- Foster, H.A., and Bridger, J.M. (2005). The genome and the nucleus: a marriage made by evolution. *Genome organisation and nuclear architecture. Chromosoma* **114**, 212-229.
- Fougere, C., and Weiss, M.C. (1978). Phenotypic exclusion in mouse melanoma-rat hepatoma hybrid cells: pigment and albumin production are not reexpressed simultaneously. *Cell* **15**, 843-854.
- Francis, N.J., and Kingston, R.E. (2001). Mechanisms of transcriptional memory. *Nat Rev Mol Cell Biol* **2**, 409-421.
- Freberg, C.T., Dahl, J.A., Timoskainen, S., and Collas, P. (2007). Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Molecular biology of the cell* **18**, 1543-1553.
- Fujimura, Y., Isono, K., Vidal, M., Endoh, M., Kajita, H., Mizutani-Koseki, Y., Takihara, Y., van Lohuizen, M., Otte, A., Jenuwein, T., *et al.* (2006). Distinct roles of Polycomb group gene products in transcriptionally repressed and active domains of Hoxb8. *Development* **133**, 2371-2381.

- Furumai, R., Matsuyama, A., Kobashi, N., Lee, K.H., Nishiyama, M., Nakajima, H., Tanaka, A., Komatsu, Y., Nishino, N., Yoshida, M., *et al.* (2002). FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. *Cancer research* 62, 4916-4921.
- Gao, L., Cueto, M.A., Asselbergs, F., and Atadja, P. (2002). Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *The Journal of biological chemistry* 277, 25748-25755.
- Gasser, S.M. (2002). Visualizing chromatin dynamics in interphase nuclei. *Science* 296, 1412-1416.
- Gerami-Naini, B., Dovzhenko, O.V., Durning, M., Wegner, F.H., Thomson, J.A., and Golos, T.G. (2004). Trophoblast differentiation in embryoid bodies derived from human embryonic stem cells. *Endocrinology* 145, 1517-1524.
- Giadrossi, S., Dvorkina, M., and Fisher, A.G. (2007). Chromatin organization and differentiation in embryonic stem cell models. *Curr Opin Genet Dev* 17, 132-138.
- Gilbert, D.M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* 14, 377-383.
- Ginis, I., Luo, Y., Miura, T., Thies, S., Brandenberger, R., Gerecht-Nir, S., Amit, M., Hoke, A., Carpenter, M.K., Itskovitz-Eldor, J., *et al.* (2004). Differences between human and mouse embryonic stem cells. *Developmental biology* 269, 360-380.
- Glaser, K.B., Staver, M.J., Waring, J.F., Stender, J., Ulrich, R.G., and Davidsen, S.K. (2003). Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. *Molecular cancer therapeutics* 2, 151-163.
- Glover, D.J., Lipps, H.J., and Jans, D.A. (2005). Towards safe, non-viral therapeutic gene expression in humans. *Nat Rev Genet* 6, 299-310.
- Gmur, R., Solter, D., and Knowles, B.B. (1980). Independent regulation of H-2K and H-2D gene expression in murine teratocarcinoma somatic cell hybrids. *The Journal of experimental medicine* 151, 1349-1359.
- Goetze, S., Mateos-Langerak, J., and van Driel, R. (2007). Three-dimensional genome organization in interphase and its relation to genome function. *Seminars in cell & developmental biology* 18, 707-714.
- Golding, I., Paulsson, J., Zawilski, S.M., and Cox, E.C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell* 123, 1025-1036.
- Gottlicher, M., Minucci, S., Zhu, P., Kramer, O.H., Schimpf, A., Giavara, S., Sleeman, J.P., Lo Coco, F., Nervi, C., Pelicci, P.G., *et al.* (2001). Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *Embo J* 20, 6969-6978.
- Grant, P.A., Sterner, D.E., Duggan, L.J., Workman, J.L., and Berger, S.L. (1998). The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes. *Trends in cell biology* 8, 193-197.
- Graves, J.A. (1972). DNA synthesis in heterokaryons formed by fusion of mammalian cells from different species. *Exp Cell Res* 72, 393-403.
- Grigoryev, S.A., Bulynko, Y.A., and Popova, E.Y. (2006). The end adjusts the means: heterochromatin remodelling during terminal cell differentiation. *Chromosome Res* 14, 53-69.
- Grimaud, C., Bantignies, F., Pal-Bhadra, M., Ghana, P., Bhadra, U., and Cavalli, G. (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* 124, 957-971.
- Grskovic, M., Chaivorapol, C., Gaspar-Maia, A., Li, H., and Ramalho-Santos, M. (2007). Systematic identification of cis-regulatory sequences active in mouse and human embryonic stem cells. *PLoS genetics* 3, e145.
- Gu, P., Le Menuet, D., Chung, A.C., and Cooney, A.J. (2006). Differential recruitment of methylated CpG binding domains by the orphan receptor GCNF initiates the repression and silencing of Oct4 expression. *Mol Cell Biol* 26, 9471-9483.

- Guan, K., Nayernia, K., Maier, L.S., Wagner, S., Dressel, R., Lee, J.H., Nolte, J., Wolf, F., Li, M., Engel, W., *et al.* (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* **440**, 1199-1203.
- Gurdon, J.B. (1962). Adult frogs derived from the nuclei of single somatic cells. *Developmental biology* **4**, 256-273.
- Gurdon, J.B. (2006). From nuclear transfer to nuclear reprogramming: the reversal of cell differentiation. *Annu Rev Cell Dev Biol* **22**, 1-22.
- Gurdon, J.B., and Byrne, J.A. (2003). The first half-century of nuclear transplantation. *Proc Natl Acad Sci U S A* **100**, 8048-8052.
- Gurdon, J.B., Byrne, J.A., and Simonsson, S. (2003). Nuclear reprogramming and stem cell creation. *Proc Natl Acad Sci U S A* **100 Suppl 1**, 11819-11822.
- Hache, R.J., Tse, R., Reich, T., Savory, J.G., and Lefebvre, Y.A. (1999). Nucleocytoplasmic trafficking of steroid-free glucocorticoid receptor. *The Journal of biological chemistry* **274**, 1432-1439.
- Hadorn, E. (1968). Transdetermination in cells. *Sci Am* **219**, 110-114 *passim*.
- Han, D.W., Do, J.T., Gentile, L., Stehling, M., Lee, H.T., and Scholer, H.R. (2008). Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. *Stem Cells* **26**, 445-454.
- Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creighton, M.P., Steine, E.J., Cassady, J.P., Foreman, R., *et al.* (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**, 250-264.
- Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M., *et al.* (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **318**, 1920-1923.
- Hansen, J.C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annual review of biophysics and biomolecular structure* **31**, 361-392.
- Hansis, C., Barreto, G., Maltry, N., and Niehrs, C. (2004). Nuclear reprogramming of human somatic cells by xenopus egg extract requires BRG1. *Curr Biol* **14**, 1475-1480.
- Hardeman, E.C., Chiu, C.P., Minty, A., and Blau, H.M. (1986). The pattern of actin expression in human fibroblast x mouse muscle heterokaryons suggests that human muscle regulatory factors are produced. *Cell* **47**, 123-130.
- Harmon, B., and Sedat, J. (2005). Cell-by-cell dissection of gene expression and chromosomal interactions reveals consequences of nuclear reorganization. *PLoS biology* **3**, e67.
- Harris, H. (1965). Behaviour of differentiated nuclei in heterokaryons of animal cells from different species. *Nature* **206**, 583-588.
- Harris, H., Miller, O.J., Klein, G., Worst, P., and Tachibana, T. (1969). Suppression of malignancy by cell fusion. *Nature* **223**, 363-368.
- Hattori, N., Nishino, K., Ko, Y.G., Hattori, N., Ohgane, J., Tanaka, S., and Shiota, K. (2004). Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *The Journal of biological chemistry* **279**, 17063-17069.
- Heard, E. (2005). Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr Opin Genet Dev* **15**, 482-489.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311-318.
- Heitz, E. (1928). "Das heterochromatin der moose." *Jahrb Wiss Bot* **69**, 762-818.
- Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* **9**, 15-26.

- Hernandez-Munoz, I., Taghavi, P., Kuijl, C., Neeffjes, J., and van Lohuizen, M. (2005). Association of BMI1 with polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. *Mol Cell Biol* 25, 11047-11058.
- Hewitt, S.L., High, F.A., Reiner, S.L., Fisher, A.G., and Merckenschlager, M. (2004). Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *European journal of immunology* 34, 3604-3613.
- Heyworth, C., Pearson, S., May, G., and Enver, T. (2002). Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. *Embo J* 21, 3770-3781.
- Hiragami, K., and Festenstein, R. (2005). Heterochromatin protein 1: a pervasive controlling influence. *Cell Mol Life Sci* 62, 2711-2726.
- Hiratani, I., Leskovar, A., and Gilbert, D.M. (2004). Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores. *Proc Natl Acad Sci U S A* 101, 16861-16866.
- Hochedlinger, K., and Jaenisch, R. (2002). Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* 415, 1035-1038.
- Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. *N Engl J Med* 349, 275-286.
- Hochedlinger, K., and Jaenisch, R. (2006). Nuclear reprogramming and pluripotency. *Nature* 441, 1061-1067.
- Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* 121, 465-477.
- Hogan, B.L., Taylor, A., and Adamson, E. (1981). Cell interactions modulate embryonal carcinoma cell differentiation into parietal or visceral endoderm. *Nature* 291, 235-237.
- Holliday, R. (1994). Epigenetics: an overview. *Developmental genetics* 15, 453-457.
- Holmgren, C., Kanduri, C., Dell, G., Ward, A., Mukhopadhyay, R., Kanduri, M., Lobanenko, V., and Ohlsson, R. (2001). CpG methylation regulates the Igf2/H19 insulator. *Curr Biol* 11, 1128-1130.
- Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* 326, 292-295.
- Horikawa, I., Chiang, Y.J., Patterson, T., Feigenbaum, L., Leem, S.H., Michishita, E., Larionov, V., Hodes, R.J., and Barrett, J.C. (2005). Differential cis-regulation of human versus mouse TERT gene expression in vivo: identification of a human-specific repressive element. *Proc Natl Acad Sci U S A* 102, 18437-18442.
- Horike, S., Cai, S., Miyano, M., Cheng, J.F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet* 37, 31-40.
- Hosler, B.A., LaRosa, G.J., Grippo, J.F., and Gudas, L.J. (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, 5623-5629.
- Hu, E., Dul, E., Sung, C.M., Chen, Z., Kirkpatrick, R., Zhang, G.F., Johanson, K., Liu, R., Lago, A., Hofmann, G., et al. (2003). Identification of novel isoform-selective inhibitors within class I histone deacetylases. *The Journal of pharmacology and experimental therapeutics* 307, 720-728.
- Humphrey, R.K., Beattie, G.M., Lopez, A.D., Bucay, N., King, C.C., Firpo, M.T., Rose-John, S., and Hayek, A. (2004). Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. *Stem Cells* 22, 522-530.
- Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci* 109 (Pt 6), 1427-1436.

- Iezzi, S., Di Padova, M., Serra, C., Caretti, G., Simone, C., Maklan, E., Minetti, G., Zhao, P., Hoffman, E.P., Puri, P.L., *et al.* (2004). Deacetylase inhibitors increase muscle cell size by promoting myoblast recruitment and fusion through induction of follistatin. *Dev Cell* 6, 673-684.
- Inoue, K., Noda, S., Ogonuki, N., Miki, H., Inoue, S., Katayama, K., Mekada, K., Miyoshi, H., and Ogura, A. (2007). Differential developmental ability of embryos cloned from tissue-specific stem cells. *Stem Cells* 25, 1279-1285.
- Inoue, K., Wakao, H., Ogonuki, N., Miki, H., Seino, K., Nambu-Wakao, R., Noda, S., Miyoshi, H., Koseki, H., Taniguchi, M., *et al.* (2005). Generation of cloned mice by direct nuclear transfer from natural killer T cells. *Curr Biol* 15, 1114-1118.
- Ivanova, A.V., Bonaduce, M.J., Ivanov, S.V., and Klar, A.J. (1998). The chromo and SET domains of the Ctr4 protein are essential for silencing in fission yeast. *Nat Genet* 19, 192-195.
- Jackson, D.A. (1995). Nuclear organization: uniting replication foci, chromatin domains and chromosome structure. *Bioessays* 17, 587-591.
- Jaenisch, R., Hochedlinger, K., and Eggan, K. (2005). Nuclear cloning, epigenetic reprogramming and cellular differentiation. *Novartis Found Symp* 265, 107-118; discussion 118-128.
- Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567-582.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074-1080.
- Jeppesen, P., and Turner, B.M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* 74, 281-289.
- Jin, S.G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS genetics* 4, e1000013.
- Jones-Villeneuve, E.M., Rudnicki, M.A., Harris, J.F., and McBurney, M.W. (1983). Retinoic acid-induced neural differentiation of embryonal carcinoma cells. *Mol Cell Biol* 3, 2271-2279.
- Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3, 415-428.
- Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068-1070.
- Jorgensen, H.F., Azuara, V., Amoils, S., Spivakov, M., Terry, A., Nesterova, T., Cobb, B.S., Ramsahoye, B., Merckenschlager, M., and Fisher, A.G. (2007). The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells. *Genome biology* 8, R169.
- Jorgensen, H.F., Giadrossi, S., Casanova, M., Endoh, M., Koseki, H., Brockdorff, N., and Fisher, A.G. (2006). Stem cells primed for action: polycomb repressive complexes restrain the expression of lineage-specific regulators in embryonic stem cells. *Cell cycle (Georgetown, Tex)* 5, 1411-1414.
- Junker, S., and Pedersen, S. (1985). Time course of arrest of immunoglobulin expression in heterokaryons and early hybrids of human lymphoma cells and mouse fibroblasts. A study of transcriptional and translational events. *Exp Cell Res* 158, 349-359.
- Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S., *et al.* (2004). Generation of pluripotent stem cells from neonatal mouse testis. *Cell* 119, 1001-1012.
- Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). Transient cyclical methylation of promoter DNA. *Nature* 452, 112-115.
- Kikyo, N., Wade, P.A., Guschin, D., Ge, H., and Wolffe, A.P. (2000). Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. *Science* 289, 2360-2362.

- Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13, 793-797.
- Kim, D.K., Lee, J.Y., Kim, J.S., Ryu, J.H., Choi, J.Y., Lee, J.W., Im, G.J., Kim, T.K., Seo, J.W., Park, H.J., *et al.* (2003). Synthesis and biological evaluation of 3-(4-substituted-phenyl)-N-hydroxy-2-propenamides, a new class of histone deacetylase inhibitors. *Journal of medicinal chemistry* 46, 5745-5751.
- Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008a). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132, 1049-1061.
- Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., *et al.* (2008b). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 454, 646-650.
- Kimura, H., Tada, M., Hatano, S., Yamazaki, M., Nakatsuji, N., and Tada, T. (2002). Chromatin reprogramming of male somatic cell-derived XIST and TSIX in ES hybrid cells. *Cytogenet Genome Res* 99, 106-114.
- Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2004). Histone code modifications on pluripotential nuclei of reprogrammed somatic cells. *Mol Cell Biol* 24, 5710-5720.
- Klebes, A., Sustar, A., Kechris, K., Li, H., Schubiger, G., and Kornberg, T.B. (2005). Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the Polycomb group, trithorax group and lama genes. *Development* 132, 3753-3765.
- Klose, R.J., and Zhang, Y. (2007). Regulation of histone methylation by demethylination and demethylation. *Nat Rev Mol Cell Biol* 8, 307-318.
- Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., *et al.* (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome research* 17, 691-707.
- Kohler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256, 495-497.
- Koipally, J., Renold, A., Kim, J., and Georgopoulos, K. (1999). Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *Embo J* 18, 3090-3100.
- Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98, 285-294.
- Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296, 158-162.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Koziol, M.J., Garrett, N., and Gurdon, J.B. (2007). Tpt1 activates transcription of oct4 and nanog in transplanted somatic nuclei. *Curr Biol* 17, 801-807.
- Kurisaki, A., Hamazaki, T.S., Okabayashi, K., Iida, T., Nishine, T., Chonan, R., Kido, H., Tsunasawa, S., Nishimura, O., Asashima, M., *et al.* (2005). Chromatin-related proteins in pluripotent mouse embryonic stem cells are downregulated after removal of leukemia inhibitory factor. *Biochemical and biophysical research communications* 335, 667-675.
- Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* 16, 2893-2905.
- Labosky, P.A., Barlow, D.P., and Hogan, B.L. (1994). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with

- embryonic stem (ES) cell lines. *Development (Cambridge, England)* **120**, 3197-3204.
- Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. *Curr Opin Cell Biol* **14**, 286-298.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120.
- Laible, G., Wolf, A., Dorn, R., Reuter, G., Nislow, C., Lebersorger, A., Popkin, D., Pillus, L., and Jenuwein, T. (1997). Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in *Drosophila* heterochromatin and at *S. cerevisiae* telomeres. *Embo J* **16**, 3219-3232.
- Lan, F., Bayliss, P.E., Rinn, J.L., Whetstone, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., *et al.* (2007a). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689-694.
- Lan, F., Zaratiegui, M., Villen, J., Vaughn, M.W., Verdel, A., Huarte, M., Shi, Y., Gygi, S.P., Moazed, D., Martienssen, R.A., *et al.* (2007b). *S. pombe* LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Mol Cell* **26**, 89-101.
- Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**, 104-115.
- Lande-Diner, L., and Cedar, H. (2005). Silence of the genes--mechanisms of long-term repression. *Nat Rev Genet* **6**, 648-654.
- Lassar, A.B., Davis, R.L., Wright, W.E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell* **66**, 305-315.
- Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73-84.
- Lee, J.H., Hart, S.R., and Skalnik, D.G. (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* **38**, 32-38.
- Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284-295.
- Lee, N., Maurange, C., Ringrose, L., and Paro, R. (2005). Suppression of Polycomb group proteins by JNK signalling induces transdetermination in *Drosophila* imaginal discs. *Nature* **438**, 234-237.
- Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301-313.
- Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-1200.
- Lemaitre, J.M., Danis, E., Pasero, P., Vassetzky, Y., and Mechali, M. (2005). Mitotic remodeling of the replicon and chromosome structure. *Cell* **123**, 787-801.
- Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**, 865-873.
- Levasseur, D.N., Wang, J., Dorschner, M.O., Stamatoyannopoulos, J.A., and Orkin, S.H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. *Genes Dev* **22**, 575-580.
- Levine, S.S., King, I.F., and Kingston, R.E. (2004). Division of labor in polycomb group repression. *Trends in biochemical sciences* **29**, 478-485.

- Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570.
- Li, B., Carey, M., and Workman, J.L. (2007a). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 3, 662-673.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-926.
- Li, J.Y., Pu, M.T., Hirasawa, R., Li, B.Z., Huang, Y.N., Zeng, R., Jing, N.H., Chen, T., Li, E., Sasaki, H., *et al.* (2007b). Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Mol Cell Biol* 27, 8748-8759.
- Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., *et al.* (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol*.
- Lin, C.M., Fu, H., Martinovsky, M., Bouhassira, E., and Aladjem, M.I. (2003). Dynamic alterations of replication timing in mammalian cells. *Curr Biol* 13, 1019-1028.
- Lin, M.S., and Davidson, R.L. (1975). Replication of human chromosomes in human-mouse hybrids: evidence that the timing of DNA synthesis is determined independently in each human chromosome. *Somatic cell genetics* 1, 111-122.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., *et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38, 431-440.
- Lovell-Badge, R. (2001). The future for stem cell research. *Nature* 414, 88-91.
- Lovell-Badge, R. (2007). Many ways to pluripotency. *Nature biotechnology* 25, 1114-1116.
- Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci U S A* 105, 2883-2888.
- Loyola, A., and Almouzni, G. (2004). Bromodomains in living cells participate in deciphering the histone code. *Trends in cell biology* 14, 279-281.
- Lu, J., McKinsey, T.A., Zhang, C.L., and Olson, E.N. (2000). Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Mol Cell* 6, 233-244.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Ma, D.K., Chiang, C.H., Ponnusamy, K., Ming, G.L., and Song, H. (2008). G9a and Jhdmd2a regulate embryonic stem cell fusion-induced reprogramming of adult neural stem cells. *Stem Cells* 26, 2131-2141.
- Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C., and Hochedlinger, K. (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 3, 340-345.
- Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., *et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* 1, 55-70.
- Mahmoudi, T., and Verrijzer, C.P. (2001). Chromatin silencing and activation by Polycomb and trithorax group proteins. *Oncogene* 20, 3055-3066.
- Mahy, N.L., Perry, P.E., and Bickmore, W.A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol* 159, 753-763.
- Maison, C., Bailly, D., Peters, A.H., Quivy, J.P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric

- heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 30, 329-334.
- Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. *Science* 303, 666-669.
- Mal, A., Sturniolo, M., Schiltz, R.L., Ghosh, M.K., and Harter, M.L. (2001). A role for histone deacetylase HDAC1 in modulating the transcriptional activity of MyoD: inhibition of the myogenic program. *Embo J* 20, 1739-1753.
- Maresca, T.J., and Heald, R. (2006). The long and the short of it: linker histone H1 is required for metaphase chromosome compaction. *Cell cycle (Georgetown, Tex)* 5, 589-591.
- Marks, P.A., and Dokmanovic, M. (2005). Histone deacetylase inhibitors: discovery and development as anticancer agents. *Expert opinion on investigational drugs* 14, 1497-1511.
- Marmorstein, R. (2001). Protein modules that manipulate histone tails for chromatin regulation. *Nat Rev Mol Cell Biol* 2, 422-432.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6, 838-849.
- Martinez, A.M., and Cavalli, G. (2006). The role of polycomb group proteins in cell cycle regulation during development. *Cell cycle (Georgetown, Tex)* 5, 1189-1197.
- Massa, S., Junker, S., and Matthias, P. (2000). Molecular mechanisms of extinction: old findings and new ideas. *Int J Biochem Cell Biol* 32, 23-40.
- Massengale, M., Wagers, A.J., Vogel, H., and Weissman, I.L. (2005). Hematopoietic cells maintain hematopoietic fates upon entering the brain. *The Journal of experimental medicine* 201, 1579-1589.
- Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., *et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* 9, 625-635.
- Mathur, D., Danford, T.W., Boyer, L.A., Young, R.A., Gifford, D.K., and Jaenisch, R. (2008). Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET. *Genome biology* 9, R126.
- Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *Embo J* 18, 4261-4269.
- Matsui, Y., Zsebo, K., and Hogan, B.L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* 70, 841-847.
- Matsumura, H., Tada, M., Otsuji, T., Yasuchika, K., Nakatsuji, N., Surani, A., and Tada, T. (2007). Targeted chromosome elimination from ES-somatic hybrid cells. *Nature methods* 4, 23-25.
- Maves, L., and Schubiger, G. (1995). Wingless induces transdetermination in developing *Drosophila* imaginal discs. *Development* 121, 1263-1272.
- McBurney, M.W. (1977). Hemoglobin synthesis in cell hybrids formed between teratocarcinoma and Friend erythroleukemia cells. *Cell* 12, 653-662.
- McClure, K.D., and Schubiger, G. (2007). Transdetermination: *Drosophila* imaginal disc cells exhibit stem cell-like potency. *The international journal of biochemistry & cell biology* 39, 1105-1118.
- McClure, K.D., and Schubiger, G. (2008). A screen for genes that function in leg disc regeneration in *Drosophila melanogaster*. *Mechanisms of development* 125, 67-80.
- Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nature biotechnology* 25, 1177-1181.

- Merriam, R.W. (1969). Movement of cytoplasmic proteins into nuclei induced to enlarge and initiate DNA or RNA synthesis. *J Cell Sci* 5, 333-349.
- Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 7, 540-546.
- Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 10, 105-116.
- Metcalf, D. (2003). The unsolved enigmas of leukemia inhibitory factor. *Stem Cells* 21, 5-14.
- Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.
- Mevel-Ninio, M., and Weiss, M.C. (1981). Immunofluorescence analysis of the time-course of extinction, reexpression, and activation of albumin production in rat hepatoma-mouse fibroblast heterokaryons and hybrids. *The Journal of cell biology* 90, 339-350.
- Mezey, E., Chandross, K.J., Harta, G., Maki, R.A., and McKercher, S.R. (2000). Turning blood into brain: cells bearing neuronal antigens generated in vivo from bone marrow. *Science* 290, 1779-1782.
- Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454, 49-55.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Miller, R.A., and Ruddle, F.H. (1976). Pluripotent teratocarcinoma-thymus somatic cell hybrids. *Cell* 9, 45-55.
- Miller, S.C., Pavlath, G.K., Blakely, B.T., and Blau, H.M. (1988). Muscle cell components dictate hepatocyte gene expression and the distribution of the Golgi apparatus in heterokaryons. *Genes Dev* 2, 330-340.
- Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* 10, 1107-1117.
- Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787-800.
- Mitsiades, C.S., Mitsiades, N.S., McMullan, C.J., Poulaki, V., Shringarpure, R., Hideshima, T., Akiyama, M., Chauhan, D., Munshi, N., Gu, X., *et al.* (2004). Transcriptional signature of histone deacetylase inhibition in multiple myeloma: biological and clinical implications. *Proc Natl Acad Sci U S A* 101, 540-545.
- Miyagi, S., Masui, S., Niwa, H., Saito, T., Shimazaki, T., Okano, H., Nishimoto, M., Muramatsu, M., Iwama, A., and Okuda, A. (2008). Consequence of the loss of Sox2 in the developing brain of the mouse. *FEBS letters* 582, 2811-2815.
- Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303, 343-348.
- Mohd-Sarip, A., van der Knaap, J.A., Wyman, C., Kanaar, R., Schedl, P., and Verrijzer, C.P. (2006). Architecture of a polycomb nucleoprotein complex. *Mol Cell* 24, 91-100.
- Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* 15, 942-947.

- Montgomery, N.D., Yee, D., Montgomery, S.A., and Magnuson, T. (2007). Molecular and functional mapping of EED motifs required for PRC2-dependent histone methylation. *Journal of molecular biology* 374, 1145-1157.
- Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias, M.I., Ariza, R.R., and Roldan-Arjona, T. (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci U S A* 103, 6853-6858.
- Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *The Journal of biological chemistry* 279, 52353-52360.
- Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. *Human molecular genetics* 14 Spec No 1, R47-58.
- Morin-Kensicki, E.M., Faust, C., LaMantia, C., and Magnuson, T. (2001). Cell and tissue requirements for the gene *eed* during mouse gastrulation and organogenesis. *Genesis* 31, 142-146.
- Moroi, Y., Peebles, C., Fritzler, M.J., Steigerwald, J., and Tan, E.M. (1980). Autoantibody to centromere (kinetochore) in scleroderma sera. *Proc Natl Acad Sci U S A* 77, 1627-1631.
- Muller, H.J., and Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in *Drosophila*. *Genetics* 15, 283-311.
- Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* 111, 197-208.
- Murry, C.E., Soonpaa, M.H., Reinecke, H., Nakajima, H., Nakajima, H.O., Rubart, M., Pasumarthi, K.B., Virag, J.I., Bartelmez, S.H., Poppa, V., *et al.* (2004). Haematopoietic stem cells do not transdifferentiate into cardiac myocytes in myocardial infarcts. *Nature* 428, 664-668.
- Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochizuki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature biotechnology* 26, 101-106.
- Negre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during *Drosophila* development. *PLoS biology* 4, e170.
- Ng, R.K., and Gurdon, J.B. (2005). Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc Natl Acad Sci U S A* 102, 1957-1962.
- Ng, R.K., and Gurdon, J.B. (2008). Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol* 10, 102-109.
- Nichols, J., Chambers, I., Taga, T., and Smith, A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development* 128, 2333-2339.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379-391.
- Nishikawa, S., Goldstein, R.A., and Nierras, C.R. (2008). The promise of human induced pluripotent stem cells for research and therapy. *Nat Rev Mol Cell Biol* 9, 725-729.
- Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 12, 2048-2060.

- Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24, 372-376.
- Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophoblast differentiation. *Cell* 123, 917-929.
- Noggle, S.A., James, D., and Brivanlou, A.H. (2005). A molecular basis for human embryonic stem cell pluripotency. *Stem Cell Rev* 1, 111-118.
- Nutt, S.L., Heavey, B., Rolink, A.G., and Busslinger, M. (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 401, 556-562.
- O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* 21, 4330-4336.
- Odom, D.T., Dowell, R.D., Jacobsen, E.S., Gordon, W., Danford, T.W., MacIsaac, K.D., Rolfe, P.A., Conboy, C.M., Gifford, D.K., and Fraenkel, E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39, 730-732.
- Ohkawa, Y., Marfella, C.G., and Imbalzano, A.N. (2006). Skeletal muscle specification by myogenin and Mef2D via the SWI/SNF ATPase Brg1. *Embo J* 25, 490-501.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.
- Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313-317.
- Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *The Journal of biological chemistry* 280, 5307-5317.
- Orlando, V., and Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr Opin Genet Dev* 5, 174-179.
- Orlic, D., Kajstura, J., Chimenti, S., Bodine, D.M., Leri, A., and Anversa, P. (2001). Transplanted adult bone marrow cells repair myocardial infarcts in mice. *Annals of the New York Academy of Sciences* 938, 221-229; discussion 229-230.
- Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. (2004). Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* 303, 669-672.
- Palermo, A.T., Labarge, M.A., Doyonnas, R., Pomerantz, J., and Blau, H.M. (2005). Bone marrow contribution to skeletal muscle: a physiological response to stress. *Developmental biology* 279, 336-344.
- Palmieri, S.L., Peter, W., Hess, H., and Scholer, H.R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Developmental biology* 166, 259-267.
- Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet* 35, 190-194.
- Pannell, D., Osborne, C.S., Yao, S., Sukonnik, T., Pasceri, P., Karauskakis, A., Okano, M., Li, E., Lipshitz, H.D., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. *Embo J* 19, 5884-5894.
- Parada, L.A., Sotiropoulos, S., and Misteli, T. (2004). Spatial genome organization. *Exp Cell Res* 296, 64-70.
- Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008a). Disease-Specific Induced Pluripotent Stem Cells. *Cell*.

- Park, I.H., Lerou, P.H., Zhao, R., Huo, H., and Daley, G.Q. (2008b). Generation of human-induced pluripotent stem cells. *Nat Protoc* 3, 1180-1186.
- Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008c). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451, 141-146.
- Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *Embo J* 23, 4061-4071.
- Pasini, D., Hansen, K.H., Christensen, J., Agger, K., Cloos, P.A., and Helin, K. (2008). Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev* 22, 1345-1355.
- Patterton, D., and Wolffe, A.P. (1996). Developmental roles for chromatin and chromosomal structure. *Developmental biology* 173, 2-13.
- Pavlat, G.K., and Blau, H.M. (1986). Expression of muscle genes in heterokaryons depends on gene dosage. *The Journal of cell biology* 102, 124-130.
- Pazin, M.J., and Kadonaga, J.T. (1997). What's up and down with histone deacetylation and transcription? *Cell* 89, 325-328.
- Peehl, D.M., and Stanbridge, E.J. (1982). The role of differentiation in the suppression of tumorigenicity in human cell hybrids. *Int J Cancer* 30, 113-120.
- Pelton, T.A., Sharma, S., Schulz, T.C., Rathjen, J., and Rathjen, P.D. (2002). Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. *J Cell Sci* 115, 329-339.
- Pera, M.F., Reubinoff, B., and Trounson, A. (2000). Human embryonic stem cells. *J Cell Sci* 113 (Pt 1), 5-10.
- Pera, M.F., and Trounson, A.O. (2004). Human embryonic stem cells: prospects for development. *Development* 131, 5515-5525.
- Pereira, C.F., Terranova, R., Ryan, N.K., Santos, J., Morris, K.J., Cui, W., Merkenschlager, M., and Fisher, A.G. (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS genetics* 4, e1000170.
- Perry, P., Sauer, S., Billon, N., Richardson, W.D., Spivakov, M., Warnes, G., Livesey, F.J., Merkenschlager, M., Fisher, A.G., and Azuara, V. (2004). A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction. *Cell cycle (Georgetown, Tex)* 3, 1645-1650.
- Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12, 1577-1589.
- Peters, A.H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. *Curr Opin Cell Biol* 17, 230-238.
- Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. *Curr Biol* 14, R546-551.
- Peterson, J.A., and Weiss, M.C. (1972). Expression of differentiated functions in hepatoma cell hybrids: induction of mouse albumin production in rat hepatoma-mouse fibroblast hybrids. *Proc Natl Acad Sci U S A* 69, 571-575.
- Pfaffl, M.W., Lange, I.G., Daxenberger, A., and Meyer, H.H. (2001). Tissue-specific expression pattern of estrogen receptors (ER): quantification of ER alpha and ER beta mRNA with real-time RT-PCR. *Apmis* 109, 345-355.
- Pirrotta, V. (1998). Polycomb the genome: PcG, trxG, and chromatin silencing. *Cell* 93, 333-336.
- Pirrotta, V. (2006). Polycomb silencing mechanisms and genomic programming. *Ernst Schering Research Foundation workshop*, 97-113.
- Plusa, B., Hadjantonakis, A.K., Gray, D., Piotrowska-Nitsche, K., Jedrusik, A., Papaioannou, V.E., Glover, D.M., and Zernicka-Goetz, M. (2005). The first

- cleavage of the mouse zygote predicts the blastocyst axis. *Nature* 434, 391-395.
- Pogo, B.G., Allfrey, V.G., and Mirsky, A.E. (1966). RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc Natl Acad Sci U S A* 55, 805-812.
- Pownall, M.E., Gustafsson, M.K., and Emerson, C.P., Jr. (2002). Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu Rev Cell Dev Biol* 18, 747-783.
- Pralong, D., Mrozik, K., Occhiodoro, F., Wijesundara, N., Sumer, H., Van Bostel, A.L., Trounson, A., and Verma, P.J. (2005). A novel method for somatic cell nuclear transfer to mouse embryonic stem cells. *Cloning Stem Cells* 7, 265-271.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* 433, 434-438.
- Pruss, D., Hayes, J.J., and Wolffe, A.P. (1995). Nucleosomal anatomy--where are the histones? *Bioessays* 17, 161-170.
- Puri, P.L., Iezzi, S., Stiegler, P., Chen, T.T., Schiltz, R.L., Muscat, G.E., Giordano, A., Kedes, L., Wang, J.Y., and Sartorelli, V. (2001). Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis. *Mol Cell* 8, 885-897.
- Raff, M. (2003). Adult stem cell plasticity: fact or artifact? *Annu Rev Cell Dev Biol* 19, 1-22.
- Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS biology* 4, e309.
- Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C., and Melton, D.A. (2002). "Stemness": transcriptional profiling of embryonic and adult stem cells. *Science* 298, 597-600.
- Rankin, J.K., and Darlington, G.J. (1979). Expression of human hepatic genes in mouse hepatoma--human amniocyte hybrids. *Somatic cell genetics* 5, 1-10.
- Rao, P.N., and Johnson, R.T. (1970). Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature* 225, 159-164.
- Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 9, 1091-1100.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-432.
- Reik, W., and Dean, W. (2001). DNA methylation and mammalian epigenetics. *Electrophoresis* 22, 2838-2843.
- Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* 293, 1089-1093.
- Resnick, J.L., Bixler, L.S., Cheng, L., and Donovan, P.J. (1992). Long-term proliferation of mouse primordial germ cells in culture. *Nature* 359, 550-551.
- Rhodes, S.J., and Konieczny, S.F. (1989). Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev* 3, 2050-2061.
- Rideout, W.M., 3rd, Hochedlinger, K., Kyba, M., Daley, G.Q., and Jaenisch, R. (2002). Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell* 109, 17-27.
- Ringertz, N.R., Nyman, U., and Bergman, M. (1985). DNA replication and H5 histone exchange during reactivation of chick erythrocyte nuclei in heterokaryons. *Chromosoma* 91, 391-396.
- Ringertz, N.R., and Savage, R.E. (1976). *Cell hybrids* (New York, Academic Press).

- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* 38, 413-443.
- Ringrose, L., Rehmsmeier, M., Dura, J.M., and Paro, R. (2003). Genome-wide prediction of Polycomb/Trithorax response elements in *Drosophila melanogaster*. *Dev Cell* 5, 759-771.
- Rogers, M.B., Hosler, B.A., and Gudas, L.J. (1991). Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development* 113, 815-824.
- Rossant, J. (2001). Stem cells from the Mammalian blastocyst. *Stem cells* (Dayton, Ohio) 19, 477-482.
- Rossant, J. (2007). Stem cells and lineage development in the mammalian blastocyst. *Reproduction, fertility, and development* 19, 111-118.
- Rossant, J., and Tam, P.P. (2004). Emerging asymmetry and embryonic patterning in early mouse development. *Dev Cell* 7, 155-164.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C.D., Reinberg, D., Avner, P., and Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* 24, 5475-5484.
- Rousset, J.P., Jami, J., Dubois, P., Aviles, D., and Ritz, E. (1980). Developmental potentialities and surface antigens of mouse teratocarcinoma x lymphoid cell hybrids. *Somatic cell genetics* 6, 419-433.
- Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* 25, 15-30.
- Sabbattini, P., Lundgren, M., Georgiou, A., Chow, C., Warnes, G., and Dillon, N. (2001). Binding of Ikaros to the lambda5 promoter silences transcription through a mechanism that does not require heterochromatin formation. *Embo J* 20, 2812-2822.
- Sado, T., Fenner, M.H., Tan, S.S., Tam, P., Shioda, T., and Li, E. (2000). X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Developmental biology* 225, 294-303.
- Saha, A., Wittmeyer, J., and Cairns, B.R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* 7, 437-447.
- Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. (2006). Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to Ultrabithorax. *Science* 311, 1118-1123.
- Santos-Rosa, H., and Caldas, C. (2005). Chromatin modifier enzymes, the histone code and cancer. *Eur J Cancer* 41, 2381-2402.
- Sarcinella, E., Zuzarte, P.C., Lau, P.N., Draker, R., and Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol* 27, 6457-6468.
- Sato, N., Sanjuan, I.M., Heke, M., Uchida, M., Naef, F., and Brivanlou, A.H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. *Developmental biology* 260, 404-413.
- Saurin, A.J., Shiels, C., Williamson, J., Satijn, D.P., Otte, A.P., Sheer, D., and Freemont, P.S. (1998). The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *The Journal of cell biology* 142, 887-898.
- Schafer, B.W., Blakely, B.T., Darlington, G.J., and Blau, H.M. (1990). Effect of cell history on response to helix-loop-helix family of myogenic regulators. *Nature* 344, 454-458.
- Schoeffner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *Embo J* 25, 3110-3122.

- Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887-898.
- Schubeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J., and Groudine, M. (2002). Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. *Nat Genet* 32, 438-442.
- Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat Genet* 38, 700-705.
- Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* 8, 9-22.
- Seibler, J., Zevnik, B., Kuter-Luks, B., Andreas, S., Kern, H., Hennek, T., Rode, A., Heimann, C., Faust, N., Kauselmann, G., *et al.* (2003). Rapid generation of inducible mouse mutants. *Nucleic Acids Res* 31, e12.
- Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* 98, 37-46.
- Shen, C.N., Slack, J.M., and Tosh, D. (2000). Molecular basis of transdifferentiation of pancreas to liver. *Nat Cell Biol* 2, 879-887.
- Shen, X., Collier, J.M., Hlaing, M., Zhang, L., Delshad, E.H., Bristow, J., and Bernstein, H.S. (2003). Genome-wide examination of myoblast cell cycle withdrawal during differentiation. *Dev Dyn* 226, 128-138.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941-953.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-847.
- Silberg, D.G., Sullivan, J., Kang, E., Swain, G.P., Moffett, J., Sund, N.J., Sackett, S.D., and Kaestner, K.H. (2002). Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology* 122, 689-696.
- Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. *Nature* 441, 997-1001.
- Simonsson, S., and Gurdon, J. (2004). DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nat Cell Biol* 6, 984-990.
- Simonsson, S., and Gurdon, J.B. (2005). Changing cell fate by nuclear reprogramming. *Cell cycle (Georgetown, Tex)* 4, 513-515.
- Skvirsky, R.C., Greenberg, M.L., Myers, P.L., and Greer, H. (1986). A new negative control gene for amino acid biosynthesis in *Saccharomyces cerevisiae*. *Current genetics* 10, 495-501.
- Slack, J.M. (1985). Homoeotic transformations in man: implications for the mechanism of embryonic development and for the organization of epithelia. *J Theor Biol* 114, 463-490.
- Slack, J.M., and Tosh, D. (2001). Transdifferentiation and metaplasia--switching cell types. *Curr Opin Genet Dev* 11, 581-586.
- Solovei, I., Schermelleh, L., During, K., Engelhardt, A., Stein, S., Cremer, C., and Cremer, T. (2004). Differences in centromere positioning of cycling and postmitotic human cell types. *Chromosoma* 112, 410-423.
- Somoza, J.R., Skene, R.J., Katz, B.A., Mol, C., Ho, J.D., Jennings, A.J., Luong, C., Arvai, A., Buggy, J.J., Chi, E., *et al.* (2004). Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure* 12, 1325-1334.
- Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6, 846-856.

- Spemann, H. (1938). Embryonic development and induction (New York, Hafner Publishing Company).
- Spivakov, M., and Fisher, A.G. (2007). Epigenetic signatures of stem-cell identity. *Nat Rev Genet* 8, 263-271.
- Spotswood, H.T., and Turner, B.M. (2002). An increasingly complex code. *J Clin Invest* 110, 577-582.
- Stadtfeld, M., Brennand, K., and Hochedlinger, K. (2008a). Reprogramming of pancreatic Beta cells into induced pluripotent stem cells. *Curr Biol* 18, 890-894.
- Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008b). Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem Cell* 2, 230-240.
- Stern, C.D. (2006). Evolution of the mechanisms that establish the embryonic axes. *Curr Opin Genet Dev* 16, 413-418.
- Stewart, C.L., Gadi, I., and Bhatt, H. (1994). Stem cells from primordial germ cells can reenter the germ line. *Developmental biology* 161, 626-628.
- Stewart, T.A., and Mintz, B. (1982). Recurrent germ-line transmission of the teratocarcinoma genome from the METT-1 culture line to progeny in vivo. *J Exp Zool* 224, 465-469.
- Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* 9, 1428-1435.
- Stojkovic, M., Lako, M., Stojkovic, P., Stewart, R., Przyborski, S., Armstrong, L., Evans, J., Herbert, M., Hyslop, L., Ahmad, S., *et al.* (2004). Derivation of human embryonic stem cells from day-8 blastocysts recovered after three-step in vitro culture. *Stem Cells* 22, 790-797.
- Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R., *et al.* (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol* 11, 996-1000.
- Strelchenko, N., Kukhareno, V., Shkumatov, A., Verlinsky, O., Kuliev, A., and Verlinsky, Y. (2006). Reprogramming of human somatic cells by embryonic stem cell cytoplasm. *Reprod Biomed Online* 12, 107-111.
- Su, I.H., Basavaraj, A., Krutchinsky, A.N., Hobert, O., Ullrich, A., Chait, B.T., and Tarakhovskiy, A. (2003). Ezh2 controls B cell development through histone H3 methylation and IgH rearrangement. *Nat Immunol* 4, 124-131.
- Su, R.C., Brown, K.E., Saaber, S., Fisher, A.G., Merckenschlager, M., and Smale, S.T. (2004). Dynamic assembly of silent chromatin during thymocyte maturation. *Nat Genet* 36, 502-506.
- Sullivan, S., Pells, S., Hooper, M., Gallagher, E., and McWhir, J. (2006). Nuclear reprogramming of somatic cells by embryonic stem cells is affected by cell cycle stage. *Cloning Stem Cells* 8, 174-188.
- Sung, L.Y., Gao, S., Shen, H., Yu, H., Song, Y., Smith, S.L., Chang, C.C., Inoue, K., Kuo, L., Lian, J., *et al.* (2006). Differentiated cells are more efficient than adult stem cells for cloning by somatic cell nuclear transfer. *Nat Genet* 38, 1323-1328.
- Surani, M.A. (2001). Reprogramming of genome function through epigenetic inheritance. *Nature* 414, 122-128.
- Sustar, A., and Schubiger, G. (2005). A transient cell cycle shift in Drosophila imaginal disc cells precedes multipotency. *Cell* 120, 383-393.
- Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C.M., Tora, L., and Dillon, N. (2005). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol Cell Biol* 25, 1804-1820.

- Tada, M., Tada, T., Lefebvre, L., Barton, S.C., and Surani, M.A. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *Embo J* 16, 6510-6520.
- Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr Biol* 11, 1553-1558.
- Tada, T., Tada, M., Hilton, K., Barton, S.C., Sado, T., Takagi, N., and Surani, M.A. (1998). Epigenotype switching of imprintable loci in embryonic germ cells. *Dev Genes Evol* 207, 551-561.
- Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51-61.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.
- Tamada, H., Van Thuan, N., Reed, P., Nelson, D., Katoku-Kikyo, N., Wudel, J., Wakayama, T., and Kikyo, N. (2006). Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol Cell Biol* 26, 1259-1271.
- Tamashiro, K.L., Wakayama, T., Akutsu, H., Yamazaki, Y., Lachey, J.L., Wortman, M.D., Seeley, R.J., D'Alessio, D.A., Woods, S.C., Yanagimachi, R., *et al.* (2002). Cloned mice have an obese phenotype not transmitted to their offspring. *Nat Med* 8, 262-267.
- Tanabe, H., Muller, S., Neusser, M., von Hase, J., Calcagno, E., Cremer, M., Solovei, I., Cremer, C., and Cremer, T. (2002). Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci U S A* 99, 4424-4429.
- Tanaka, E.M. (2003). Regeneration: if they can do it, why can't we? *Cell* 113, 559-562.
- Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science* 282, 2072-2075.
- Taranger, C.K., Noer, A., Sorensen, A.L., Hakelien, A.M., Boquest, A.C., and Collas, P. (2005). Induction of dedifferentiation, genomewide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells. *Molecular biology of the cell* 16, 5719-5735.
- Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408-411.
- Terada, N., Hamazaki, T., Oka, M., Hoki, M., Mastalerz, D.M., Nakano, Y., Meyer, E.M., Morel, L., Petersen, B.E., and Scott, E.W. (2002). Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature* 416, 542-545.
- Terranova, R., Pereira, C.F., Du Roure, C., Merckenschlager, M., and Fisher, A.G. (2006). Acquisition and extinction of gene expression programs are separable events in heterokaryon reprogramming. *J Cell Sci* 119, 2065-2072.
- Terranova, R., Sauer, S., Merckenschlager, M., and Fisher, A.G. (2005). The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Exp Cell Res* 310, 344-356.
- Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196-199.

- Thakar, R., and Csink, A.K. (2005). Changing chromatin dynamics and nuclear organization during differentiation in *Drosophila* larval tissue. *J Cell Sci* 118, 951-960.
- Thayer, M.J., Tapscott, S.J., Davis, R.L., Wright, W.E., Lassar, A.B., and Weintraub, H. (1989). Positive autoregulation of the myogenic determination gene MyoD1. *Cell* 58, 241-248.
- Thiagalingam, S., Cheng, K.H., Lee, H.J., Mineva, N., Thiagalingam, A., and Ponte, J.F. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. *Annals of the New York Academy of Sciences* 983, 84-100.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145-1147.
- Tolhuis, B., de Wit, E., Muijters, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. (2006). Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat Genet* 38, 694-699.
- Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10, 1453-1465.
- Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res* 30, 3202-3213.
- Torres-Padilla, M.E., Parfitt, D.E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* 445, 214-218.
- Tosh, D., and Slack, J.M. (2002). How cells change their phenotype. *Nat Rev Mol Cell Biol* 3, 187-194.
- Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* 128, 651-654.
- Trimarchi, J.M., Fairchild, B., Wen, J., and Lees, J.A. (2001). The E2F6 transcription factor is a component of the mammalian Bmi1-containing polycomb complex. *Proc Natl Acad Sci U S A* 98, 1519-1524.
- Tsonis, P.A., and Del Rio-Tsonis, K. (2004). Lens and retina regeneration: transdifferentiation, stem cells and clinical applications. *Experimental eye research* 78, 161-172.
- Tsonis, P.A., Madhavan, M., Tancous, E.E., and Del Rio-Tsonis, K. (2004). A newt's eye view of lens regeneration. *Int J Dev Biol* 48, 975-980.
- Tsuji-Takayama, K., Inoue, T., Ijiri, Y., Otani, T., Motoda, R., Nakamura, S., and Orita, K. (2004). Demethylating agent, 5-azacytidine, reverses differentiation of embryonic stem cells. *Biochemical and biophysical research communications* 323, 86-90.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811-816.
- Tsukiyama, T. (2002). The in vivo functions of ATP-dependent chromatin-remodelling factors. *Nat Rev Mol Cell Biol* 3, 422-429.
- Ulaner, G.A., Yang, Y., Hu, J.F., Li, T., Vu, T.H., and Hoffman, A.R. (2003). CTCF binding at the insulin-like growth factor-II (IGF2)/H19 imprinting control region is insufficient to regulate IGF2/H19 expression in human tissues. *Endocrinology* 144, 4420-4426.
- Vajta, G., and Gjerris, M. (2006). Science and technology of farm animal cloning: state of the art. *Anim Reprod Sci* 92, 211-230.
- Valk-Lingbeek, M.E., Bruggeman, S.W., and van Lohuizen, M. (2004). Stem cells and cancer; the polycomb connection. *Cell* 118, 409-418.

- Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* **118**, 4495-4509.
- van Holde, K., and Zlatanova, J. (2007). Chromatin fiber structure: Where is the problem now? *Seminars in cell & developmental biology* **18**, 651-658.
- Vannini, A., Volpari, C., Filocamo, G., Casavola, E.C., Brunetti, M., Renzoni, D., Chakravarty, P., Paolini, C., De Francesco, R., Gallinari, P., *et al.* (2004). Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. *Proc Natl Acad Sci U S A* **101**, 15064-15069.
- Vazquez, J., Muller, M., Pirrotta, V., and Sedat, J.W. (2006). The Mcp element mediates stable long-range chromosome-chromosome interactions in *Drosophila*. *Molecular biology of the cell* **17**, 2158-2165.
- Verdel, A., and Moazed, D. (2005). RNAi-directed assembly of heterochromatin in fission yeast. *FEBS letters* **579**, 5872-5878.
- Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. *Mol Cell* **10**, 1223-1233.
- Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833-1837.
- Waddington, C.H. (1942). "The epigenotype". *Endeavour* **1**, 18-20.
- Wade, P.A., Pruss, D., and Wolffe, A.P. (1997). Histone acetylation: chromatin in action. *Trends in biochemical sciences* **22**, 128-132.
- Wagers, A.J., and Weissman, I.L. (2004). Plasticity of adult stem cells. *Cell* **116**, 639-648.
- Wakayama, S., Hikichi, T., Suetsugu, R., Sakaide, Y., Bui, H.T., Mizutani, E., and Wakayama, T. (2007). Efficient establishment of mouse embryonic stem cell lines from single blastomeres and polar bodies. *Stem Cells* **25**, 986-993.
- Wakayama, T., Perry, A.C., Zuccotti, M., Johnson, K.R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394**, 369-374.
- Walsh, K., and Perlman, H. (1997). Cell cycle exit upon myogenic differentiation. *Curr Opin Genet Dev* **7**, 597-602.
- Wang, G., Zhang, H., Zhao, Y., Li, J., Cai, J., Wang, P., Meng, S., Feng, J., Miao, C., Ding, M., *et al.* (2005). Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. *Biochemical and biophysical research communications* **330**, 934-942.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004a). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878.
- Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* **28**, 371-375.
- Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368.
- Wang, L., Brown, J.L., Cao, R., Zhang, Y., Kassis, J.A., and Jones, R.S. (2004b). Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell* **14**, 637-646.
- Warming, S., Costantino, N., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2005). Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* **33**, e36.
- Watanabe, D., Suetake, I., Tada, T., and Tajima, S. (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mechanisms of development* **118**, 187-190.

- Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37, 853-862.
- Weierich, C., Brero, A., Stein, S., von Hase, J., Cremer, C., Cremer, T., and Solovei, I. (2003). Three-dimensional arrangements of centromeres and telomeres in nuclei of human and murine lymphocytes. *Chromosome Res* 11, 485-502.
- Weimann, J.M., Johansson, C.B., Trejo, A., and Blau, H.M. (2003). Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. *Nat Cell Biol* 5, 959-966.
- Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benzeira, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., *et al.* (1991). The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* 251, 761-766.
- Weismann, A. (1892). *Das Keimplasma. Eine Theorie der Vererbung*. G. Fischer, Jena. English translation by W.N. Parker and H. Rönnefeldt. 1915. (New York, Scribner's Sons).
- Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2, 10-12.
- Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318-324.
- Widom, J. (1989). Toward a unified model of chromatin folding. *Annual review of biophysics and biophysical chemistry* 18, 365-395.
- Williams, R.R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jorgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., *et al.* (2006). Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J Cell Sci* 119, 132-140.
- Wilmot, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810-813.
- Wolffe, A.P. (1991). Implications of DNA replication for eukaryotic gene expression. *J Cell Sci* 99 (Pt 2), 201-206.
- Wong, C.C., Gaspar-Maia, A., Ramalho-Santos, M., and Reijo Pera, R.A. (2008). High-Efficiency Stem Cell Fusion-Mediated Assay Reveals Sall4 as an Enhancer of Reprogramming. *PLoS ONE* 3, e1955.
- Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I., and Carter, N.P. (2004). Replication timing of the human genome. *Human molecular genetics* 13, 191-202.
- Wright, W.E., Sassoon, D.A., and Lin, V.K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* 56, 607-617.
- Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121, 859-872.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86-90.
- Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise reprogramming of B cells into macrophages. *Cell* 117, 663-676.
- Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. *Nature biotechnology* 19, 971-974.

- Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nature biotechnology* 20, 1261-1264.
- Xu, R.H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nature methods* 2, 185-190.
- Xu, W.S., Parmigiani, R.B., and Marks, P.A. (2007). Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene* 26, 5541-5552.
- Yaffe, D., and Saxel, O. (1977). Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* 270, 725-727.
- Yamanaka, S. (2008). Pluripotency and nuclear reprogramming. *Philos Trans R Soc Lond B Biol Sci* 363, 2079-2087.
- Yang, X., Smith, S.L., Tian, X.C., Lewin, H.A., Renard, J.P., and Wakayama, T. (2007). Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning. *Nat Genet* 39, 295-302.
- Yang, X.J., and Seto, E. (2003). Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Curr Opin Genet Dev* 13, 143-153.
- Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115, 281-292.
- Ying, Q.L., Nichols, J., Evans, E.P., and Smith, A.G. (2002). Changing potency by spontaneous fusion. *Nature* 416, 545-548.
- Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519-523.
- Yu, J., Vodyanik, M.A., He, P., Slukvin, I., and Thomson, J.A. (2006). Human embryonic stem cells reprogram myeloid precursors following cell-cell fusion. *Stem Cells* 24, 168-176.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917-1920.
- Zernicka-Goetz, M. (2005). Cleavage pattern and emerging asymmetry of the mouse embryo. *Nat Rev Mol Cell Biol* 6, 919-928.
- Zhang, F., Pomerantz, J.H., Sen, G., Palermo, A.T., and Blau, H.M. (2007). Active tissue-specific DNA demethylation conferred by somatic cell nuclei in stable heterokaryons. *Proc Natl Acad Sci U S A* 104, 4395-4400.
- Zhang, J., Tam, W.L., Tong, G.Q., Wu, Q., Chan, H.Y., Soh, B.S., Lou, Y., Yang, J., Ma, Y., Chai, L., *et al.* (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* 8, 1114-1123.
- Zhang, J., Xu, F., Hashimshony, T., Keshet, I., and Cedar, H. (2002). Establishment of transcriptional competence in early and late S phase. *Nature* 420, 198-202.
- Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D.A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*.

APPENDIX I – PRIMER SEQUENCES

Table A1. Primers used for the selective amplification of human transcripts by semi-quantitative RT-PCR.

<i>Species/Gene</i>	<i>T_m</i>	<i>n</i>		<i>Sequence 5'-3'</i>
hCD20	60°C	25	s	GCAGCAACGGAGAAAACTC
			as	CTGATCTTGGGGAGGTTCTG
hCD37	60°C	25	s	CGAGGAGAGCTGGGACTATG
			as	AGACGTGGTCCAGGTTTCTG
hCD45	60°C	25	s	TTGCGATTTCCGTGTAAAAG
			as	GTTCCCAAATCATCCTCCAG
hPax5	60°C	25	s	GCTCGTCGTA CTCCATCAGC
			as	CGTTGTACGAGGAATACTGAGG
hGapdh	60°C	19	s	AGCCACATCGCTCAGACAC
			as	AGGGGTCTACATGGCAACTG
hMyogenin	60°C	25	s	CAGCGAATGCAGCTCTCAC
			as	CAGAAGTAGTGGCATCTGTGG
hMrf4	60°C	25	s	TGGGGAAAATGTTACTCTGC
			as	GAAGGCTACTCGAGGCTGAC
hMyf5	60°C	25	s	GCAGGAGTTGCTGAGAGAGC
			as	TTAGGCCCTCCTGGAAGAAG

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

Table A2. Primers for the selective amplification of human or mouse transcripts by quantitative RT-PCR.

Human specific primers for qRT-PCR

<i>Species/Gene</i>		<i>Sequence 5'-3'</i>
hGapdh	s	TCTGCTCCTCCTGTTCGACA
	as	AAAAGCAGCCCTGGTGACC
hHprt	s	TCCTTGGTCAGGCAGTATAATCC
	as	GTCAAGGGCATATCCTACAACAAA
hOct4	s	TCGAGAACCGAGTGAGAGGC
	as	CACACTCGGACCACATCCTTC
hNanog	s	CCAACATCCTGAACCTCAGCTAC
	as	GCCTTCTGCGTCACACCATT
hCripto	s	AGAAGTGTTCCCTGTGTAAATGCTG
	as	CACGAGGTGCTCATCCATCA
hDnmt3b	s	GTCAAGCTACACACAGGACTTGACAG
	as	AGTTCGGACAGCTGGGCTTT
hTert	s	GCCAGCATCATCAAACCCC
	as	CTGTCAAGGTAGAGACGTGGCTC

hTle1	s	TGTCTCCCAGCTCGACTGTCT
	as	AAGTACTGGCTTCCCCTCCC
hSox2	s	CACACTGCCCCTCTCACACAT
	as	CATTTCCCTCGTTTTTCTTTGAA
hRex1	s	GCGTACGCAAATTAAAGTCCAGA
	as	CAGCATCCTAAACAGCTCGCAGAAT
hCD37	s	GTGGCTGCACAACAACCTTATTT
	as	GCCTAACGGTATCGAGCGAG
hCD19	s	GCTCAAGACGCTGGAAAGTATTATT
	as	GATAAGCCAAAGTCACAGCTGAGA
hCD20	s	CCCCATCTACCCAATACTGTTACAG
	as	TGGGTCTGGAGCACGTTCTT
hCD45	s	CCCCATGAACGTTACCATTTG
	as	GATAGTCTCCATTGTGAAAATAGGCC
hPax5	s	AGCTTCCAGTCACAGCATAGTGTC
	as	ACCTTCGTCTCTCTTGCGCTT
hFgfr2	s	GCAAGTGAGAGCTTCCTGAGCT
	as	ACCTTCAATCGAGGCAGGAA
hFgfr2	s	GCCGCCGTGATCAGTTG
	as	GCGCCCTTTATCTGCAAGTACT
hFgf2	s	TCACATCAAGCTACAACCTCAAGCA
	as	AAGCCAGTAATCTTCCATCTTCCTT
hBmp4	s	TGGTCTTGAGTATCCTGAGCG
	as	GCTGAGGTAAAGAGGAAACGA
hLifr	s	CCAACATGACTTGCGACTACGT
	as	CCTGGTCGAACTCATCAGATTCTAT
hJak3	s	AAGCAGCGAGCTTGATGAGC
	as	GTCTATGGCCCCCAGGTGTA
hALPL	s	CCTGGCAGGGCTCACACT
	as	AAACAGGAGAGTCGCTTCAGAGA
hColagenIV α 1	s	TCAGCAGGGCATCGCAT
	as	AAATGTCAATTCAGGCCTAGTGG
hNestin	s	TGTGGCCCAGAGGCTTCTC
	as	CAGGGCTGGTGAGCTTGG
hHnf4	s	ACCAGCTACATCGCCTACCTGA
	as	CAGGAGGAAAACCTTCGTGCT
hMixl1	s	CCGAGTCCAGGATCCAGGTA
	as	CTCTGACGCCGAGACTTGG
hSox7	s	ACGCCGAGCTCAGCAAGAT
	as	TCCACGTACGGCCTCTTCTG
hGata6	s	ACCACCTTATGGCGCAGAAAC
	as	TTTTTCATAGCAAGTGGTCTGGG
hCdx2	s	CAGGACGAAAGACAAATATCGAGTG
	as	CCAGATTTTAACTGCCTCTCAGA
hHand1	s	ACCAGCTACATCGCCTACCTGA

	as	CAGGAGGAAAACCTTCGTGCT
	s	GGAGATCGAGAGGACAGCGT
<i>hEbf</i>	as	GTCAATGAGGCGCACGTAGA
	s	AACTTAAATGCCCCCTCCCC
<i>hMyoD</i>	as	TTCAGTTCTCCGCCTCTCCTAC
	s	GGTCCGTGACATCAACGAGG
<i>hE2A</i>	as	CTTTGGGATTCAGGTTCCGCT

Mouse primers for qRT-PCR

Species/Gene		Sequence 5'-3'
	s	TGCACCACCAACTGCTTAGC
<i>mGapdh</i>	as	GGCATGGACTGTGGTCATGAG
	s	CGTGGAGACTTTGCAGCCTG
<i>mOct4</i>	as	GCTTGGCAAACCTGTTCTAGCTCCT
	s	GAAGTATTCTTGCTTACAAGGGTCTGC
<i>mNanog</i>	as	GCATCTTCTGCTTCCTGGCAA
	s	CACCAACCCAGGGTATCAGTT
<i>mCripto</i>	as	AGAGTTCTGTCCAGTGTCGTC
	s	ACTGCCTGGAGTTCAGTAGGA
<i>mDnmt3b</i>	as	CCCTGTCTGATGGAGTTCGAC
	s	GAGTGGAACTTTTGTCCGAGA
<i>mSox2</i>	as	GAAGCGTGTACTTATCCTTCTTCAT
	s	TCTGAGTGGGCAGTCTCACTT
<i>mTle1</i>	as	GAAGAAGGGTCCTCGTTAGACA
	s	TCAAGAGCAGTAGTCGCCAG
<i>mTert</i>	as	TCTCGGGACAGGATAGCATCT
	s	CTCCTAGCCGCCTAGATTTCCA
<i>mRex1</i>	as	CGTGTCCCAGCTCTTAGTCCATT
	s	GCAGAGCATCAACTGGCTG
<i>mFgfr1</i>	as	GGTCACGCAAGCGTAGAGG
	s	GAGGAATACTTGATCTCACCCAGC
<i>mFgfr2</i>	as	CTGGTGCTGTCCTGTTTGGG
	s	GGCTCTACTGCAAGAACGGC
<i>mFgf2</i>	as	TGGAGTTGTAGTTTGACGTGTGG
	s	GAGGGATCTTTACCGGCTCC
<i>mBmp4</i>	as	GTTGAAGAGGAAACGAAAAGCAG
	s	GCAGAGCATCAACTGGCTG
<i>mLifr</i>	as	GGTCACGCAAGCGTAGAGG
	s	CTTCCGCCTGATCTGCGAC
<i>mJak3</i>	as	TGGCAAAGTCTAACGTGATGG
	s	TCAACCTCGCCACAACCTTCCC
<i>mCdx2</i>	as	TGGCTCAGCCTGGGATTGCT
	s	ACGTGCTGGCCAAGGATGCA
<i>mHand1</i>	as	TGGTTTAGCTCCAGCGCCCA
<i>mCD19</i>	s	GGAGAGCACCCGGTCAGA

	as	CCACACTGCTGACCTTGCAA
mCD37	s	AAGTACTTCCTCTTCGTTTTCAACCT
	as	ACCTGAGACAGCCAGGACCTT
mCD45	s	TGTACCACCAGGGACTGACAAG
	as	TCTGGCTCACAGTGGAGTACATATG
mKlf4	s	TCCTTTCCAACCTCGCTAACCC
	as	CGGATCGGATAGCTGAAGCTG
mc-Myc	s	TTCATCTGCGATCCTGACGAC
	as	AGGGGTCAATGCACTCGGA
mLin28	s	TCACTGGCCCTGGTGGTGTGTT
	as	TGGACACGAGGCCACCATATGG
mGata6	s	GACTCCTACTTCCTCTTCTTAATTCAGA
	as	ACCTGAATACTTGAGGTCACTGTTCTC
mHnf4	s	AATGGACAGATGTGTGAGTGGCC
	as	CCAGCAGCTTGCTAGATGGC
mMsx1	s	GCCTCTCGGCCATTTCTCAG
	as	CGGTTGGTCTTGTGCTTGCG
mPax3	s	AATTACCCACGCAGCGGC
	as	GCACAGGATCTTAGAGACGCAACC
mMath1	s	GGAGAAGCTTCGTTGCACGC
	as	GGGACATCGCACTGCAATGG
mNKX2.2	s	TGTGCAGAGCCTGCCCTTAA
	as	GCCCTGGGTCTCCTTGTCAT
Oct4βgeo transgene (within neo ^r gene)	s	CGGCAGGAGCAAGGTGAGAT
	as	CAAGATGGATTGCACGCAGG
Sox2Zeo transgene (within Zeo ^r gene)	s	AAGTTGACCAGTGCCGTTCC
	as	TGAACAGGGTCACGTCGTCC

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

Table A3. Primers for the selective amplification of human genomic regions by quantitative PCR.

Replication Timing, ChIP and MDIP analysis		
Species/Gene		Sequence 5'-3'
hβ-Amylase promoter	s	GGTGAGTCTGTGTGGTCAGCAGTCTCT
	as	CCACGGTGCTCTGGTAGATAACGTAAG
hα-globin promoter	s	TTTGTCCCCGGACCTGCTGCC
	as	CCCCACGGTCGGCTGACAC
hOct4 promoter	s	TTGGGAGTTGAAAGTTGGGTGT
	as	AGGCTGGTCTTGAATTCCTGTC
hDppa5 promoter	s	CAGGTCAATGCAGGTGGAAC
	as	CCAGCAGGTTGCCAACTCTT
hCripto promoter	s	ACTTCAAGTCTGGAGCCCCC
	as	GCCTGAGGAGCTAGGTGTGTGT
hDnmt3b promoter	s	TGTGTGTCTCCGTTCCGGTT

	as	ATCAGAAGCCCTAAGCGGGA
	s	TCACGTCCGGCATTCTGTGGT
hTert promoter	as	TCGAATCGGCCTAGGCTGTG
	s	ACGGCCTCCCAATTTACTGG
hNanog promoter	as	GGTTCAACAGGAATGGGATAAAAC
	s	AAAGCCATCGTGAGCCTCAT
hGdf3 promoter	as	GACAGAGTGAAAAGGAGACATGCA
	s	TCCCGCTGGATCTTTAAGTAGAA
hFoxd3 promoter	as	CGGTAGTCGAAAGTCCCGTTT
	s	GGCAGGCAAAATGTGAATGAG
hGbx2 promoter	as	GAAGCCGGCGTACTTATCTCC
	s	CGTGAGAGAGTGTTGGCACCT
hSox2 promoter	as	TTGTTCTCCCGCTCATCCAC
	s	AAATAGCTTTCTTCAAGCCGCC
hDppa4 promoter	as	CGGAGACATTGGGAGATTGAG
	s	CTGTTTTCCCAACCATCTCGT
hTle1 promoter	as	TCCTTCACCTTCGTGTGCTTCT
	s	CTCATGTGATCCCCCGTCT
hRex1 promoter	as	TACGCGTGGGTGTAATCACATT
	s	GAACAATGCGAGTGAGCAACTTCAGG
hIkaros promoter	as	GGCCGAGCGTCCCGCCAGGCTG
	s	ACATTACACGAGCATCCAGG
hIgf2/hH19 ICR	as	GCTCTTTAGGTTTGGCGCAAT
	s	AATTCGTTGGAAACGGGAT
h α -Satellite	as	TTCAAAGAGGTCTACATGTCCC
	s	AAACGTGACTGAGAACTTGATCGC
hMyoD promoter	as	TTTCCTTTGCCTGTCCAAACTCTG
	s	GGTGCAGATCATCCCCTTGA
D. melanogaster Gbe locus	as	TTACCCGACGGCGAAAG

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

Table A4. Primers for the selective amplification of human bisulfite converted genomic regions by PCR.

Bisulfite genomic sequencing analysis		
Species/Gene region		Sequence 5'-3'
hOct4 region 1	s	TAGTTGGGATGTGTAGAGTTTGAGA
	as	TAAACCAAAACAATCCTTCTACTCC
hOct4 region 2	s	AAGTTTTTGTGGGGGATTTGTAT
	as	CCACCCACTAACCTTAACCTCTA
hOct4 region 3	s	GTTAGAGGTTAAGGTTAGTGGGTG
	as	AAACCTTAAAACTTAACCAATCC
hIgf2/hH19 ICR	s	TGTTGAAGGTTGGGGAGATGGGA
	as	CCCAAACCATAACACTAAAACCCTC

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

APPENDIX II – COMPLETE DATA SETS

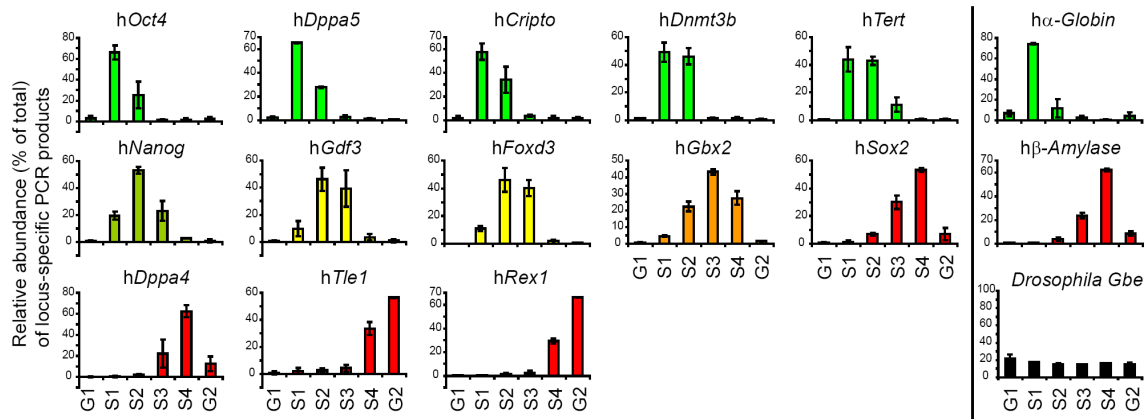


Figure A1. Replication timing analysis of pluripotency-associated genes in human B-lymphocytes. The replication timing of each gene (*hOct4*, *hDppa5*, *hCripto*, *hDnmt3b*, *hTert*, *hNanog*, *hGdf3*, *hFoxd3*, *hGbx2*, *hSox2*, *hDppa4*, *hTle1* and *hRex1*) was defined accordingly to its peak abundance in G1-S1 phase (early, green), S2 phase (middle-early, dark green), S2-S3 phase (middle, yellow), S3 phase (middle-late, orange) or S4-G2 (late, red), using a quantitative PCR-based assay (Azuara et al., 2006). Genes are ordered according to their replication timing. *hα-Globin* and *hβ-Amylase* were included as controls for an early and a late replicating gene, respectively. *Drosophila melanogaster Gbe* is an internal control for uniform recovery of BrdU-labelled DNA. Error bars indicate the s.d. of 2 independent experiments.

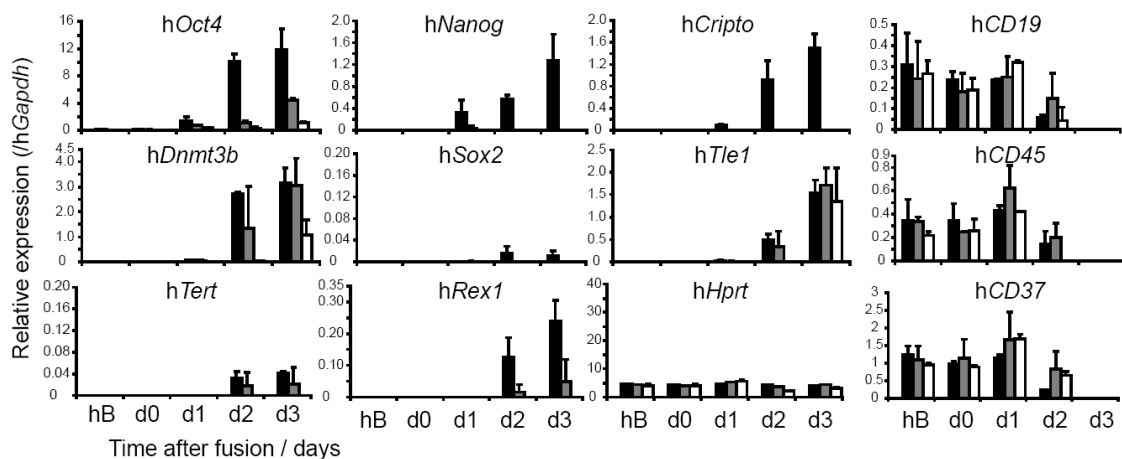


Figure A2. Kinetic of human lymphocyte reprogramming by mES cells after Oct4 ablation. ZHBTc4 ES cells expressing *Oct4* (black bars), or in which *Oct4* expression has been partially or completely ablated by the addition of doxycycline for 6 and 12 hours (grey and white bars, respectively) were fused to hB-lymphocytes. The activation of human ES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) and silencing of lymphocyte-specific genes (*hCD19*, *hCD45* and *hCD37*) were quantified by qRT-PCR over the period of 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

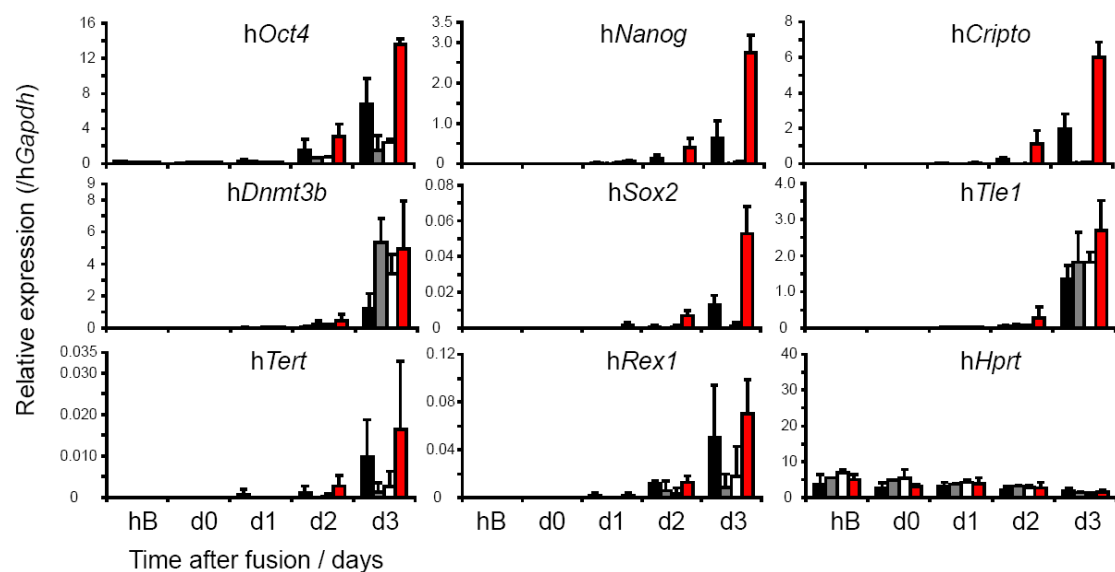


Figure A3. Kinetic of human lymphocyte reprogramming by mES cells after Sox2 ablation. 2TS22C (black bars), Sox2 depleted cells (grey and white bars; doxycycline 12 and 24 hours, respectively) and 2O1 cells (red bars; Sox2-deficient mES cells in which mOct4 expression is constitutively up-regulated) were used as fusion partners with hB cells and reprogramming was assessed by quantification of human-ES transcripts (hOct4, hNanog, hCripto, hDnmt3b, hSox2, hTle1, hTert and hRex1) using qRT-PCR over 3 days after cell fusion. hHprt was added as a control gene. Data were normalised to hGapdh expression. Error bars indicate the s.d. of 2-3 independent experiments.

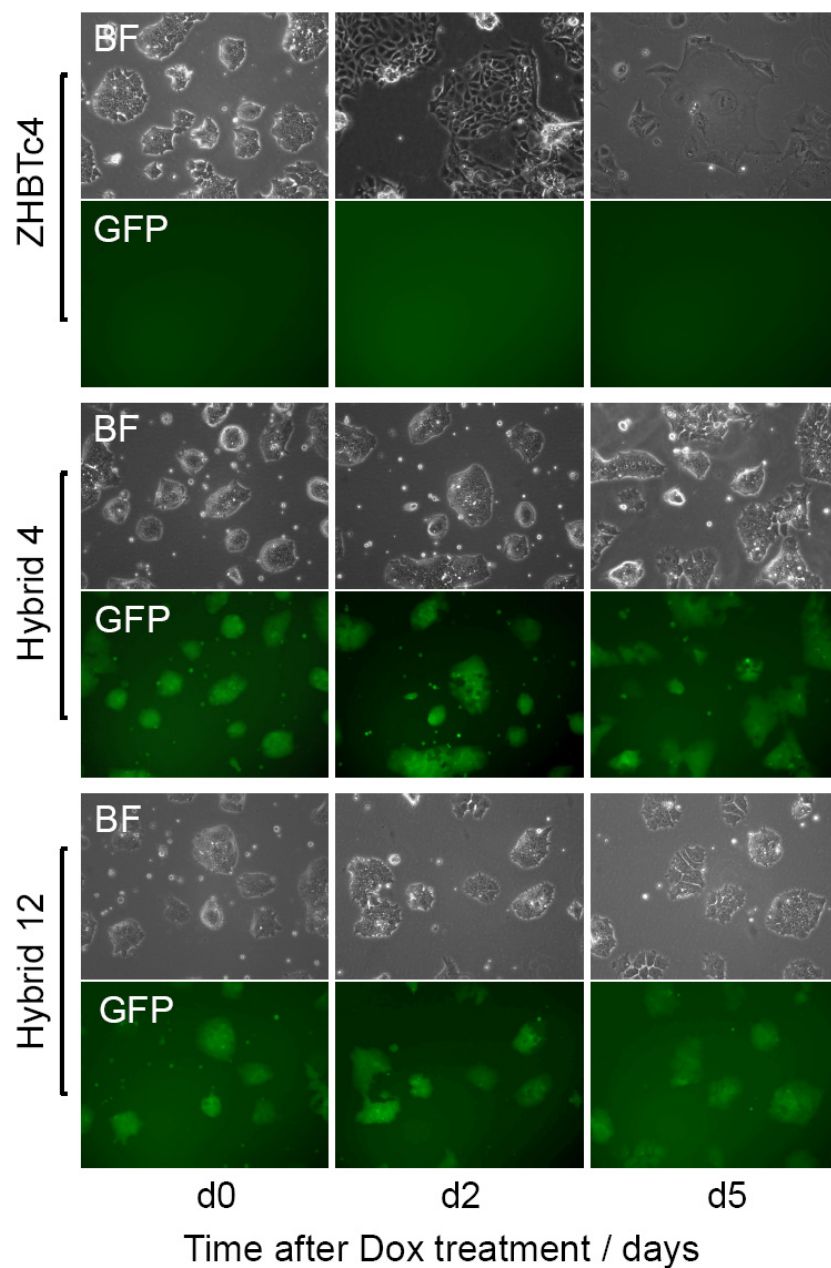


Figure A4. Reprogramming is self-sustaining and can be maintained in the absence of ES-derived *Oct4*. Dox treatment of ZHBTc4 ES cells results in morphological changes characteristic of trophectoderm differentiation (upper panel) which were not observed in hybrid clone 4 and 12 under the same conditions after 2 (d2) or 5 (d5) days of Dox treatment. GFP protein (*Oct4* promoter-driven) remains detectable in hybrid cells throughout the experiment.

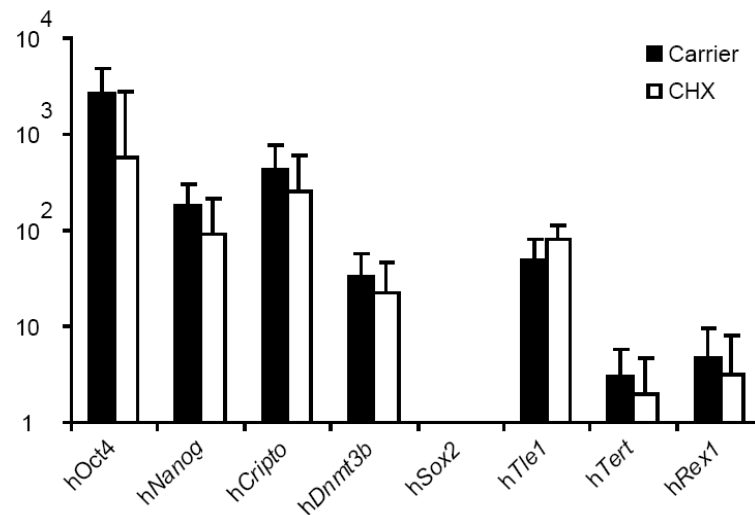


Figure A5. Inhibition of *de novo* protein synthesis in heterokaryons did not alter the overall human pattern of gene expression. OS25 ES cells were fused to hB cells and reprogramming was assessed by quantification of human-ES transcripts (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) using qRT-PCR 2 days after cell fusion. After PEG treatment heterokaryons were treated continuously with 0.1 μ g/ml of cyclohexime D (white bars) or with carrier control (black bars) for two days. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

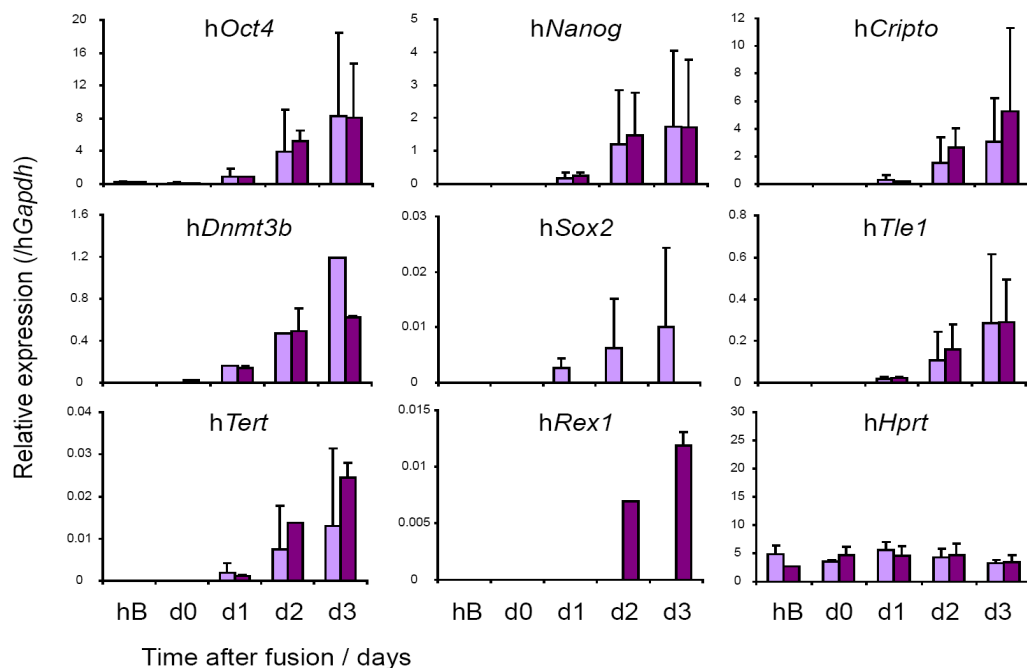


Figure A6. Kinetic of human lymphocyte reprogramming by embryonic carcinoma cells. P19 (light purple bars) and F9 (dark purple bars) embryonic carcinoma cell lines were fused to hB-lymphocytes. The activation of human ES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) was quantified by qRT-PCR over the period of 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

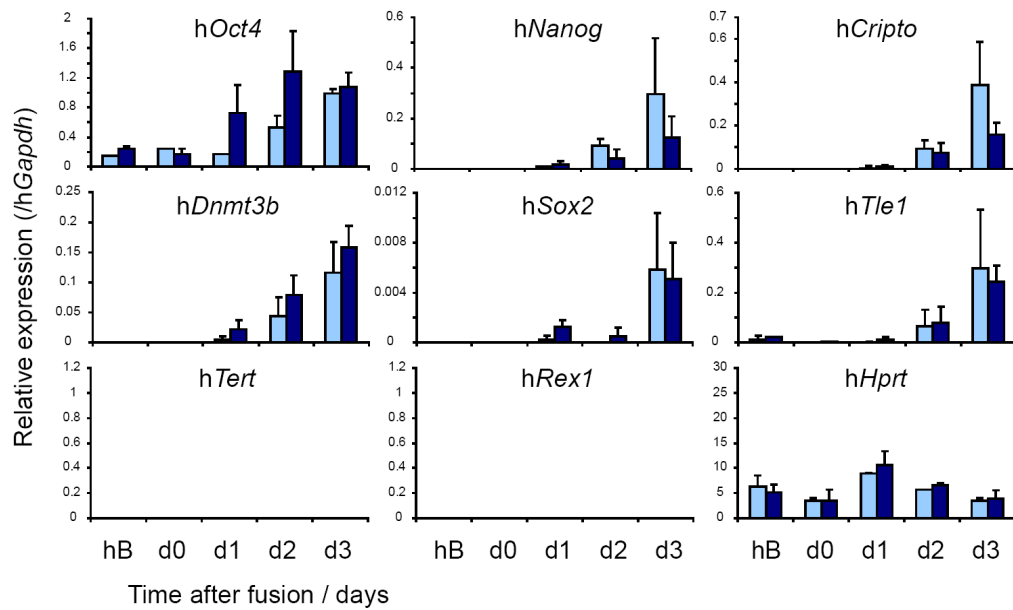


Figure A7. Kinetic of human lymphocyte reprogramming by embryonic germ cells. 23G (light blue bars) and 19G (dark blue bars) embryonic germ cell lines were fused to hB-lymphocytes. The activation of human ES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) was quantified by qRT-PCR over the period of 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2-3 independent experiments.

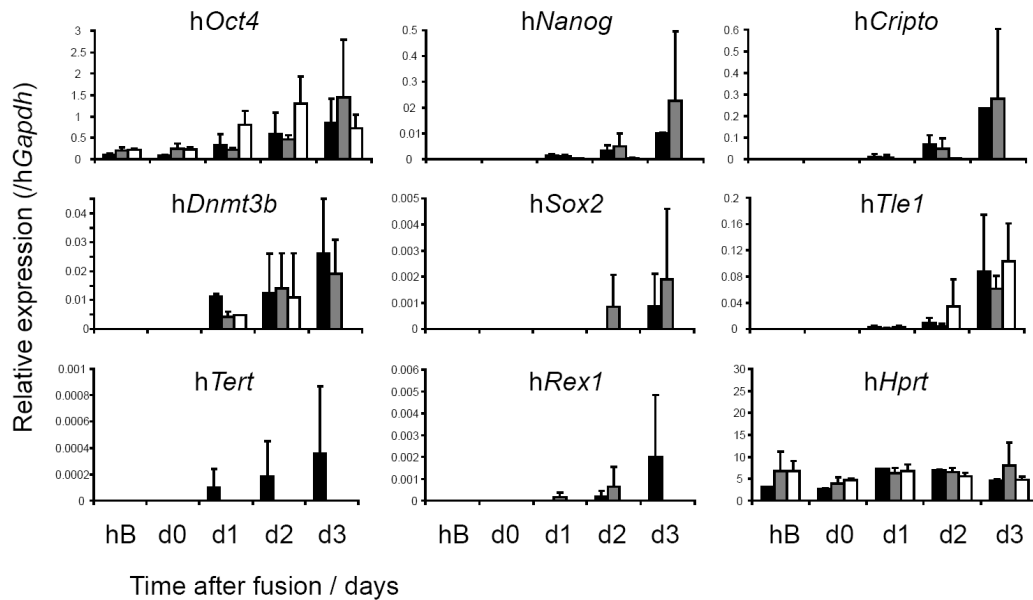


Figure A8. Kinetic of human lymphocyte reprogramming by mES cells after *Eed* ablation. Wild-type (black bars), heterozygous (grey bars) and *Eed*-deficient (white bars) embryonic stem cells were fused to hB-lymphocytes. The activation of human ES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) was quantified by qRT-PCR over the period of 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

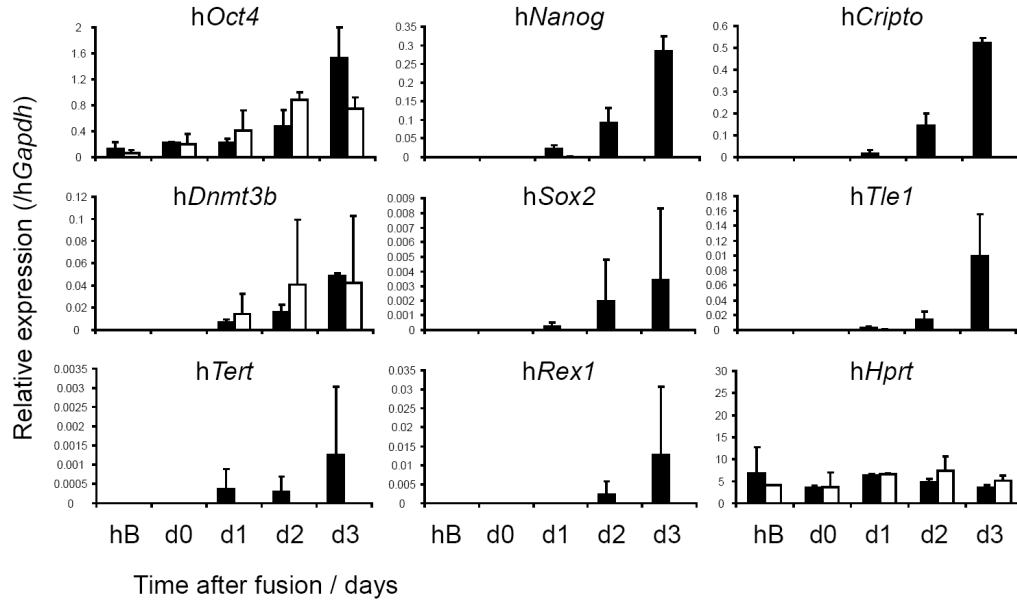


Figure A9. Kinetic of human lymphocyte reprogramming by mES cells after *Suz12* ablation. Wild-type (black bars) and *Suz12*-deficient (white bars) embryonic stem cells were fused to hB-lymphocytes. The activation of human ES-specific genes (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) was quantified by qRT-PCR over the period of 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

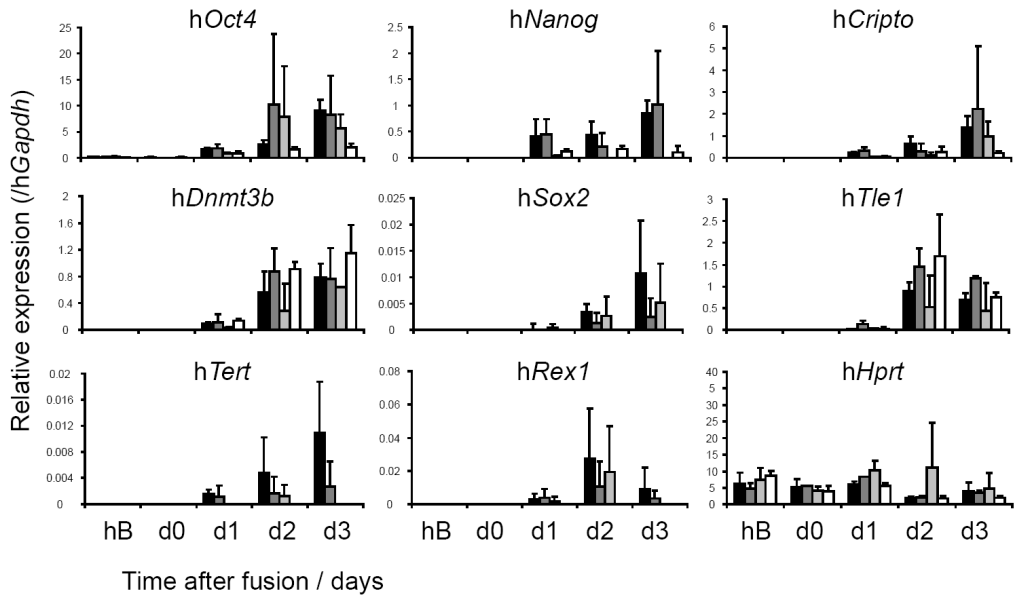


Figure A10. Kinetic of human lymphocyte reprogramming by mES cells after conditional *Ezh2* removal. *Ezh2*-1.3 ES cells expressing *Ezh2* (black bars), or in which *Ezh2* expression has been partially or completely ablated by the addition of 800 nM tamoxifen for 48 and 96 hours (dark and light grey bars, respectively) and *Ezh2*-deficient ES cells (white bars) were fused to hB-lymphocytes. Reprogramming was assessed by quantification of human-ES transcripts (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) using qRT-PCR over 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 3 independent experiments.

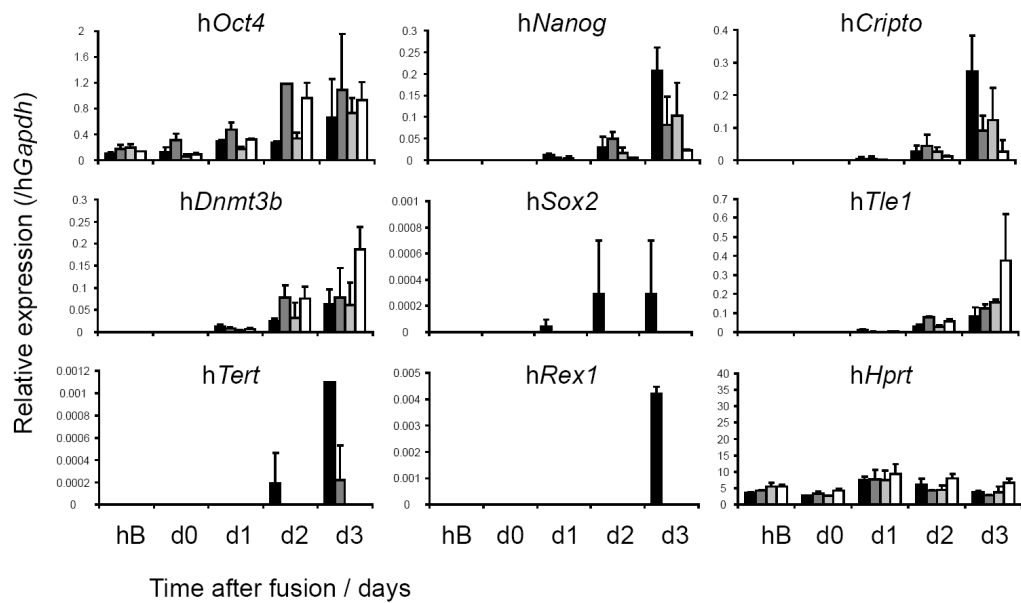


Figure A11. Kinetic of human lymphocyte reprogramming by mES cells after conditional *Ring1B* removal. ES-ERT2 ES cells expressing *Ring1B* (black bars), or in which *Ring1B* expression has been partially or completely ablated by the addition of 800 nM tamoxifen for 24, 48 and 96 hours (dark grey, light grey and white bars, respectively) were fused to hB-lymphocytes. Reprogramming was assessed by quantification of human-ES transcripts (*hOct4*, *hNanog*, *hCripto*, *hDnmt3b*, *hSox2*, *hTle1*, *hTert* and *hRex1*) using qRT-PCR over 3 days after cell fusion. *hHprt* was added as a control gene. Data were normalised to *hGapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

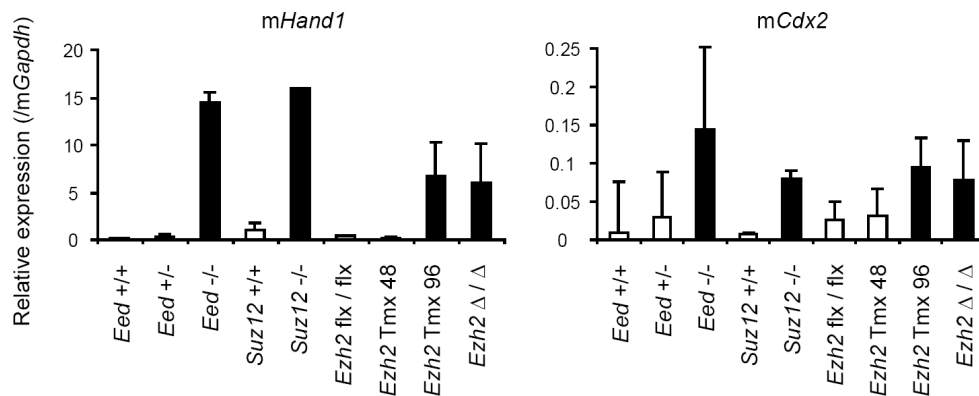


Figure A12. *mHand1* and *mCdx2* are derepressed in PRC2-deficient ES cells. (a) Gene expression analysis by qRT-PCR of *mHand1* and *mCdx2* in ES cells deficient for the PRC2 (black bars) members Eed (*Eed* $-/-$), Suz12 (*Suz12* $-/-$), Ezh2 (*Ezh2* tamoxifen 96 hours; *Ezh2* Δ/Δ) and matching wild-type controls (white bars). Data were normalised to *mGapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

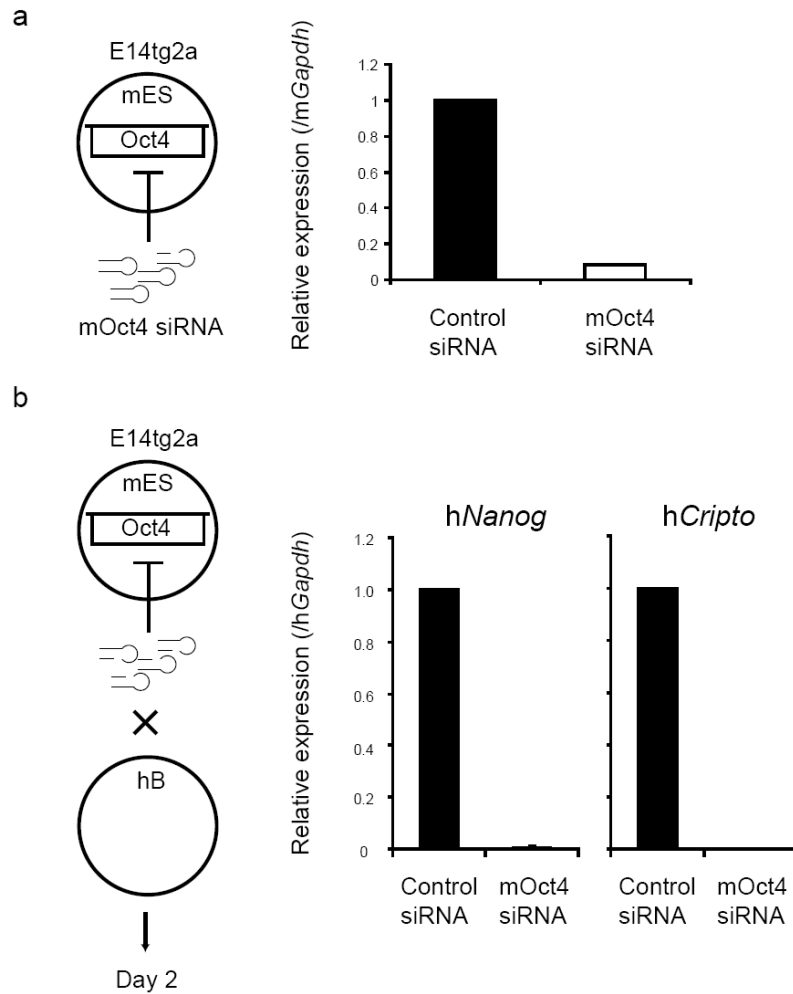


Figure A13. siRNA-mediated knock-down of mOct4 abolishes reprogramming. (a) E14tg2A ES cells were transfected with either mOct4-siRNA or target-less-siRNA (a negative control siRNA designed to have no expected targets in human and mouse cells) vectors. 48 hours later, transfected cells (GFP+) were FACS sorted and analysed by quantitative RT-PCR analysis. mOct4-siRNA targeted cells showed a >90% reduction in Oct4 transcript levels as compared to cells transfected with target-less-siRNA (control). (b) E14tg2A ES cells expressing mOct4-siRNA or control-siRNA were fused to hB-lymphocytes, and successful reprogramming was assessed by quantifying the abundance of human ES-associated transcripts (*hNanog* and *hCripto*) two days after fusion by qRT-PCR. Successful reprogramming judge by the activation of human pluripotency-associated transcripts was abolished by pre-treatment of mES cells with Oct4-siRNAs. Data were normalised to *Gapdh* expression. Error bars indicate the s.d. of 2 independent experiments.

PUBLICATIONS

Parts of the work presented here have been published as follow:

Terranova, R., Pereira, C.F., Du Roure, C., Merkschlager, M., and Fisher, A.G. (2006). Acquisition and extinction of gene expression programs are separable events in heterokaryon reprogramming. *J Cell Sci* 119, 2065-2072.

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