HISTONE MODIFICATIONS ACROSS THE CELL CYCLE IN UNDIFFERENTIATED AND DIFFERENTIATING MOUSE EMBRYONIC STEM CELLS

By

Hannah Myren Goss

A thesis submitted to

The University of Birmingham

For the degree of

Doctor of Philosophy

School of Immunity and Infection

Institute of Biomedical Research

University of Birmingham

June 2014

For everyone I consider to be family.

My parents who gave me everything I needed to become who I am today.

For Ander who I'm so proud to call my brother.

For my grandparents for their unwavering love and faith.

For my truest friends for sticking by me and giving me strength.

For Adam, you are my home.

ABSTRACT

Epigenetic mechanisms are credited with a central role in the regulation of gene expression which is as much a part of cell identity as the DNA itself. The role of post translational histone modifications in stem cells has been of increasing interest in recent years, particularly since the generation of induced pluripotent stem cells (iPSCs) and their potential for clinical applications. However, the heritability of histone modifications has not yet been determined, and as such their status as epigenetic remains in question. Here we have taken the novel approach of comparing the enrichment of histone modifications, across specific genes and how they are modulated through the various phases of the cell cycle and in doing so we address this question of heritability from a new perspective. Of the histone modifications analysed, fluctuations of the enrichment was observed on all cell cycle, pluripotent and housekeeping genes across the cell cycle in undifferentiated embryonic stem (ES) cells. In cell cycle regulated genes the patterns of modification enrichment revealed an increase in active marks either pre-emptive or at the point of expression, mostly accompanied by repressive marks that were enriched where the active marks became depleted. Histone modifications in ES cells appear to be highly dynamically regulated, not heritable, perhaps reflective of the plasticity of these cells. Following on from this ES cells were differentiated for seven days, allowing the enforcement of canonical cell cycle regulation and change in the transcription profile towards a less pluripotent and plastic phenotype, at which point changes in some histone modification patterns were observed including H3K4me3 and H3K27me3 which stabilised across the cell cycle. In undifferentiated and differentiated cells we appear to see multiple modes of action for histone modifications which include the heritability of H3K4me3 and H3K27me3 in differentiating cells but not undifferentiated ES cells.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Dr Laura O'Neill. Thank you for all the nudges, pushing, coaxing and patience required to get me here. I really don't know how you find the time to do everything you manage, it is truly spectacular. All of your help has meant a great deal to me, thank you.

Thank you to all the people in the chromatin group past and present for your help, advice, sympathy and general great company. Thank you Karl Nightingale, John Halsall, Maaike Weirsma, Charlotte Rutledge, Richard Bowker, Edith Terrenoire, Nuthana Prathivadi Bhayankaram, Milan Fernando, Rebecca Lauren Smith, Bryan Turner and anyone else I've forgotten.

I would also like to thank all members in the Anderson lab for their generosity with their help and time. You made venturing up a floor a pleasure.

I would finally like to gratefully acknowledge the Medical Research Council for funding this project.

TABLE OF CONTENTS

List of Figures

List of Tables

List of abbreviations

H3, H4 – histone 3, histone 4
K4, K8, K9, K27, S10 – lysine 4, lysine 8, lysine 9, lysine 27, serine 10
me1, me2, me3 – monomethyl, dimethyl, trimethyl histone modification
ac, phos – acetylation, phosphorylation histone modification

AS cells – adult stem cells
BD – bound fraction
C-ChIP – carrier chromatin immunoprecipitation
CDK – cyclin dependent kinases
DNMT – DNA methyl transferase
ES cells – embryonic stem cells
FACs – fluorescence activated cell sorting
HATs – histone acetyl transferases
HDACs – histone deacetylases
HDMs – histone demethylases
hES cells – human embryonic stem cells
HKMTs – histone lysine methyl transferases
HMTs – histone methyl transferases
HP1 – heterochromatin protein 1
HSCs – hematopoietic stem cells
iPSCs – induced pluripotent stem cells
JAK – janus kinase
KDMs – lysine demethylases
LIF – leukaemia inhibitory factor
LIFR – leukaemia inhibitory factor receptor
MEF - mouse embryonic fibroblasts
mES cells – mouse embryonic stem cells
N-ChIP – native chromatin immunoprecipitation
ncRNAs – non coding RNAs
OS25s – male mouse ES cell line
PcGs – polycomb group proteins
PRCs – polycomb repressive complexes
PRMTs – protein arginine methyl transferases
PTM – post translational modifications
Rb protein – retinoblastoma protein
RNAi – RNA interference
SCNT – somatic cell nuclear transfer
STAT – signal transducers and activators of transcription
TSS – transcriptional start sites
Trx G – trithorax group proteins
UB – unbound fraction

1 INTRODUCTION

Complex multicellular organisms such as mammals are made up of hundreds of distinct specialized cell types including a myriad of epithelial, muscle, bone, blood and nerve cells before even touching upon any of the cell types that compose the internal organs. Each cell type is able to carry out a distinct function as a result of its carefully controlled specialized gene expression pattern. The identity of a cell is determined not only by the inheritance of DNA but by its careful control and regulation. All cells within one organism have ultimately developed from a single totipotent cell, the zygote. How a single cell is able to proliferate and form an organism with complex tissue systems and organs has been the subject of intense investigation for a great many years. As the zygote divides cells begin to differentiate and in doing so lose plasticity; totipotency is lost as early as the 8-16 cell stage where inner cells become pluripotent and outer cells go on to form the trophectoderm (Pedersen et al. 1986; Fleming 1987; Torres-Padilla 2013). As they proliferate the genome within each cell remains the same, it is the organisation and select expression of genes within DNA which directs and controls proliferation and differentiation to the point where a complex organism emerges. Cells with the ability to self-renew and differentiate into more than one other type of cell have been termed stem cells and have many unique traits and properties. They have long been used in research, not only as models for early development but also for their ability to rapidly proliferate and the incredible therapeutic potential they hold.

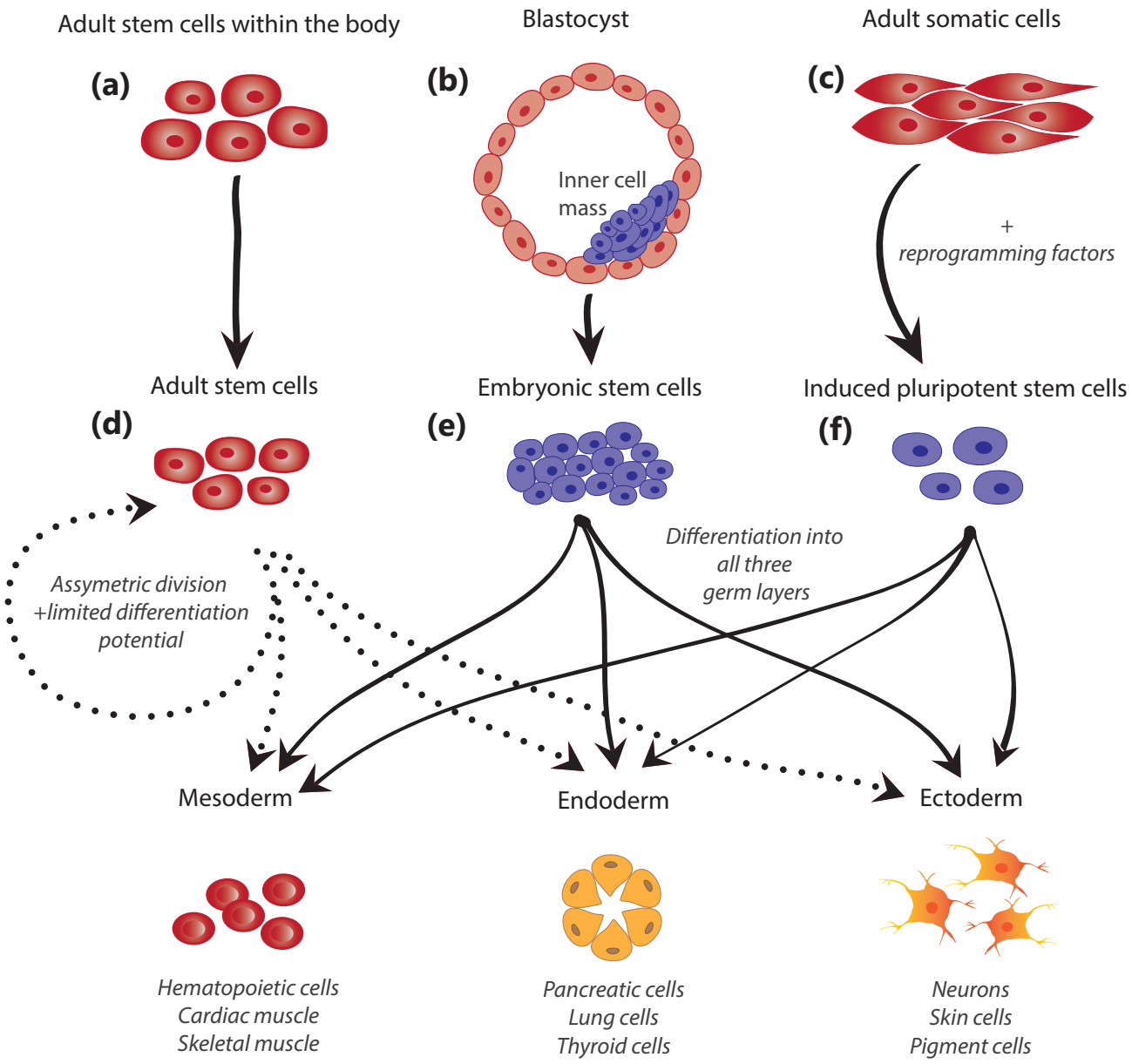
1.1 STEM CELLS

The flexibility of stem cells and consequently their vast potential took the world by storm when they first emerged into the public eye. They have since been portrayed as

the miracle cure for so many medical problems. Stem cells are defined by their plasticity and self-renewal. Many types of stem cell exist, from adult stem cells -which can differentiate into a limited pool of cell types, to cells taken from the inner cell mass of the blastocyst at the 100 cell stage which are known as embryonic stem cells (ES cells), and more recently induced pluripotent stem cells (iPSCs) reprogrammed from somatic cells manipulated *in vitro*. Both ES cells and iPSCs are termed pluripotent because of their ability to differentiate into any cell type within the body excluding extra embryonic tissue; they are able to generate any cell type derivative of the three germ layers; ectoderm, endoderm and mesoderm (figure 1.1)(Bradley et al. 1984). The unique regenerative capabilities and potential of stem cells are being investigated in association with a plethora of diseases including diabetes, heart disease, spinal cord injury and multiple sclerosis (Lodi et al. 2011; Prasongchean & Ferretti 2012; Van Zant & Liang 2012; Deb & Ubil 2014). New tissue could be created from stem cells to replace that which has been injured such as extensive burn injuries or damage to the cornea (Rama et al. 2001; Pellegrini et al. 2007; Blais et al. 2013; Lewis 2013). In addition to therapeutic uses, stem cells or cells differentiated from them, could have a vital role in advancing research, such as the screening of new drugs and the generation of model systems which could be used to study normal development and also identify causes of birth defects (Stice et al. 2006; Lou & Liang 2011).

The potential uses of stem cells are inestimable, however despite the optimism found in media reports, a great deal more research and understanding at the molecular level is needed before most of these goals will become feasible (Daley 2012). Basic stem cell research is advancing at a rapid pace; however, this has not been reflected in the clinical applications. The potential of stem cells lies in their plasticity yet at the same time this very plasticity slows the progress into the use of these cells primarily because plasticity

Figure 1.1 The plasticity of stem cells



The fate of stem cells. (a) Adult stem cells within the body can be isolated and purified to culture (d) Adult stem cells. Adult stem cells divide assymetrically resulting in a stem cell and differentiating cell as daughter cells. The differentiation potential of Adult stem cells is limited, usually resistricted to cells within one germ layer. Cells isolated from (b) the inner cell mass of the blastocyst are cultured as (e) Embryonic stem cells, these cells are able to differentiate into cells from all three germ layers; mesoderm, endoderm and ectoderm. (c) Adult somatic cells can be converted into (f) induced pluripotent stem cells upon reprogramming with pluripotency related transcription factors. (f) Induced pluripotent stem cells share the ability of embryonic stem cells to differentiate into cells from all three germ layers.

and rapid proliferation are linked to genomic instability, properties also associated with cancer cells. Stem cells, particularly those obtained from embryos (ES cells) rather than adults (adult stem cells), are considered to have a high degree of genetic instability possibly linked to their rapid proliferation and unique cell cycle (Martin 1981; Kawai et al. 2010; Lim & Gong 2013; Marzi et al. 2013). *In vitro* culture conditions may also contribute to aberrant growth. Other concerns with the clinical use of stem cells include immune rejection, low cell numbers and ethical considerations when obtaining cells. The particular source of the stem cells can reduce some of these risks but not all of them. For example adult stem cells are less likely to be associated with tumour formation and depending on the source would not be rejected, but are only present in low numbers, whereas iPSCs could be expanded in culture to obtain large cell numbers and would not be rejected if derived from the patient, however they do have a high degree of instability particularly when generated through lenti-virus or retro-virus based delivery systems (Lodi et al. 2011; Daley 2012). In 2009, Amariglio and colleagues reported a critical case of stem cell therapy in humans which resulted in tumourogenesis. A 13 year old boy presented with a multifocal brain tumour 4 years after injection of foetal neural stem cells, upon analysis the tumours were determined to originate from the injected stem cells and grew from at least two separate donors (Amariglio et al. 2009). Such instances highlight the need for a greater understanding and therefore control of the elements regulating pluripotency and differentiation which will be essential in producing safe tissues for clinical application.

1.1.1 Adult stem cells (AS cells)

As the name suggests AS cells are found within the adult body and serve as an internal maintenance and repair system. They retain a limited degree of plasticity compared to

ES cells and are found throughout the adult organism (including but not limited to the brain, bone marrow, peripheral blood, blood vessels, skeletal muscle, skin, teeth, heart, gut, liver, ovarian epithelium and testis). Adult stem cells in some organs, for example intestinal stem cells in the gut and hematopoietic stem cells in bone marrow, are constantly dividing to replenish or replace cells (Van Zant & Liang 2012). Other examples, such as pancreatic and cardiac stem cells, cells remain quiescent for long periods of time, only dividing under particular circumstances (Ye et al. 2013; Ziv et al. 2013). It is important for these cells to adapt to physiological needs. Adult stem cells can undergo both asymmetric division, where one daughter cell remains a stem cell with all the potential of the parent cell whereas the other differentiates (and takes on a specialised function) and symmetric division in which both daughter cells are identical (either stem cells or differentiated cells). It has been proposed that symmetrical division of AS cells within tissue is a mechanism by which cancer may be delayed (Shahriyari & Komarova 2013). The fine balance between self-renewal and differentiation is thought to be facilitated within the stem cell niche. The majority of adult stem cells are believed to reside within special niches in the body where the microenvironment allows them to be maintained in an undifferentiated state (Moore & Lemischka 2006; Rezza et al. 2014). The AS cell niche must provide the ideal conditions for a balance between quiescence and proliferation, and self-renewal and differentiation.

Adult stem cells have a limited capacity to divide *in vitro*, therefore large numbers are hard to obtain from one source which limits their therapeutic use and makes studying them challenging. The best characterised adult stem cells are hematopoietic stem cells (HSCs) and their niche within the bone marrow. In actual fact the first use of stem cells in medicine was hematopoietic stem cells in bone marrow transplants which were first carried out in the 1960s, over 40 years ago (Gatti et al. 1968; Bortin 1970). HSCs are

highly proliferative and give rise to all other blood cells; they are therefore used in the treatment of a long list of blood diseases both congenital and acquired and remain the most successful use of stem cells in human treatments (Hatzimichael & Tuthill 2010).

1.1.2 EMBRYONIC STEM CELLS (ES cells)

ES cells in particular have attracted a great deal of media attention, primarily because of their therapeutic potential, but also because of the ethical considerations associated with obtaining the cells. The first mouse ES (mES) cell line was isolated from the inner cell mass of a blastocyst in 1981, but it was not until 1998 that the first human ES cell line was generated (Martin 1981; Thomson et al. 1998). The use of human embryos to create ES cell lines remains a hotly contested topic and is extremely tightly regulated. In the UK only embryos created *in vitro* are eligible for licensed research, the majority of which are "surplus" from *in vitro* fertilisation (IVF) treatments used with the donor's consent. Generation of human ES cells is currently inefficient, however once established they can theoretically be propagated indefinitely and allow continuous study without the need to harvest new cells.

To be considered a mouse ES cell line, cells must be able to form chimeras when injected into a carrier mouse embryo and also be maintained for a prolonged period of time without differentiating or acquiring genetic abnormalities (Evans 2011). The maintenance of pluripotency initially required the presence of feeder cells (irradiated mouse embryonic fibroblasts) however, it was later discovered that mES cells could remain undifferentiated without a feeder layer if grown in the presence of leukaemia inhibitory factor (LIF). LIF promotes self-renewal in mES cells through various signalling pathways, primarily the JAK/STAT3 pathway (Williams et al. 1988; Smith et al. 1988).

Many genes are thought to be involved in the maintenance of pluripotency, however very few affect ES phenotype significantly when knocked down. Amongst those that do are *Pou5f1* (encodes the Oct4 protein), *Sox2* and *Nanog* pointing to a critical role for these genes (Nichols et al. 1998; Niwa et al. 2000; Chambers et al. 2003; Mitsui et al. 2003; Nishiyama et al. 2013). Pluripotency is maintained by a carefully regulated network with *Pou5f1*, *Sox2* and *Nanog* at the core. Levels of these key regulators are instrumental for the direction of differentiation in ES cells. *Pou5f1* and *Nanog* expression drop off when cells are differentiated down any embryonic line however *Sox2* expression is necessary for neural lineages (Schöler et al. 1989; Chambers et al. 2003; Zhou et al. 2007; Abranches et al. 2013).

1.1.3 INDUCED PLURIPOTENT STEM CELLS (iPSCs)

For a long time stem cell differentiation was believed to be unidirectional. This idea was first challenged by somatic cell nuclear transfer (SCNT); the process of inserting the nucleus from a somatic cell into an egg which had had its nucleus removed (Ball et al. 2009)(Gurdon 1962; Gurdon & Wilmut 2011). The implanted nucleus and chromatin is reprogrammed by cytoplasmic factors from the egg (returning complete plasticity) and allowing the newly created cell to be a totipotent zygote. This newly created zygote, if allowed, can be grown to blastocyst stage or even planted within a surrogate uterus and carried to term. SCNT allows cloning from adult cells, the classic example being Dolly the sheep (Wilmut et al. 1997).

SCNT allows a terminally differentiated nucleus to revert completely to that of a totipotent cell within the environment of an egg. However, not much was known about the exact reprogramming mechanisms. This changed in 2006 when Yamanaka and colleagues were able to induce pluripotency in mouse fibroblast cells by introducing

only four transcription factors; Oct4, Sox2, C-myc and Klf4 and culturing them in ES cell conditions (Takahashi & Yamanaka 2006). Groups lead by S. Yamanaka and J. A. Thompson generated iPSCs from human cells in 2007 using the same four transcription factors, (Takahashi et al. 2007; Yu et al. 2007). In generating iPSCs the specialized transcription patterns that define the role and function of a cell (cellular identity) are completely rewritten including chromatin structure which undergoes extensive remodelling to restore the plasticity of pluripotency. All this is initiated through the introduction of a few key transcription factors.

iPSCs are a new branch in stem cell research; cells generated in this manner have the same differential potential as ES cells but can be created from adult cells. This opens up the possibility of using a patient's own cells for any stem cell based treatment thereby circumventing any autoimmune rejection or ethical considerations. Stem cell lines with a known genetic condition could also be created for research purposes. For their efforts in this field Gurdon and Yamanaka were awarded the 2012 Nobel Prize in physiology or medicine and in recent years media attention has eagerly followed the progress of iPSCs, the newest phenomenon in stem cell research.

Use of viruses in the generation of iPSCs for use in humans has raised concerns, principally that they may affect the normal function of cells. The use of viral vectors to generate iPSCs later placed into mice sometimes leads to the mice developing cancer (Liu et al. 2013). Several non-viral methods for the generation of iPSCs have been proposed including non-integrating viral vectors (Stadtfeld et al. 2008; Fusaki et al. 2009), chemicals/small molecules (O'Doherty et al. 2013), recombinant proteins (Kim et al. 2009; Zhou et al. 2009) and RNA based methods (Warren et al. 2010), one

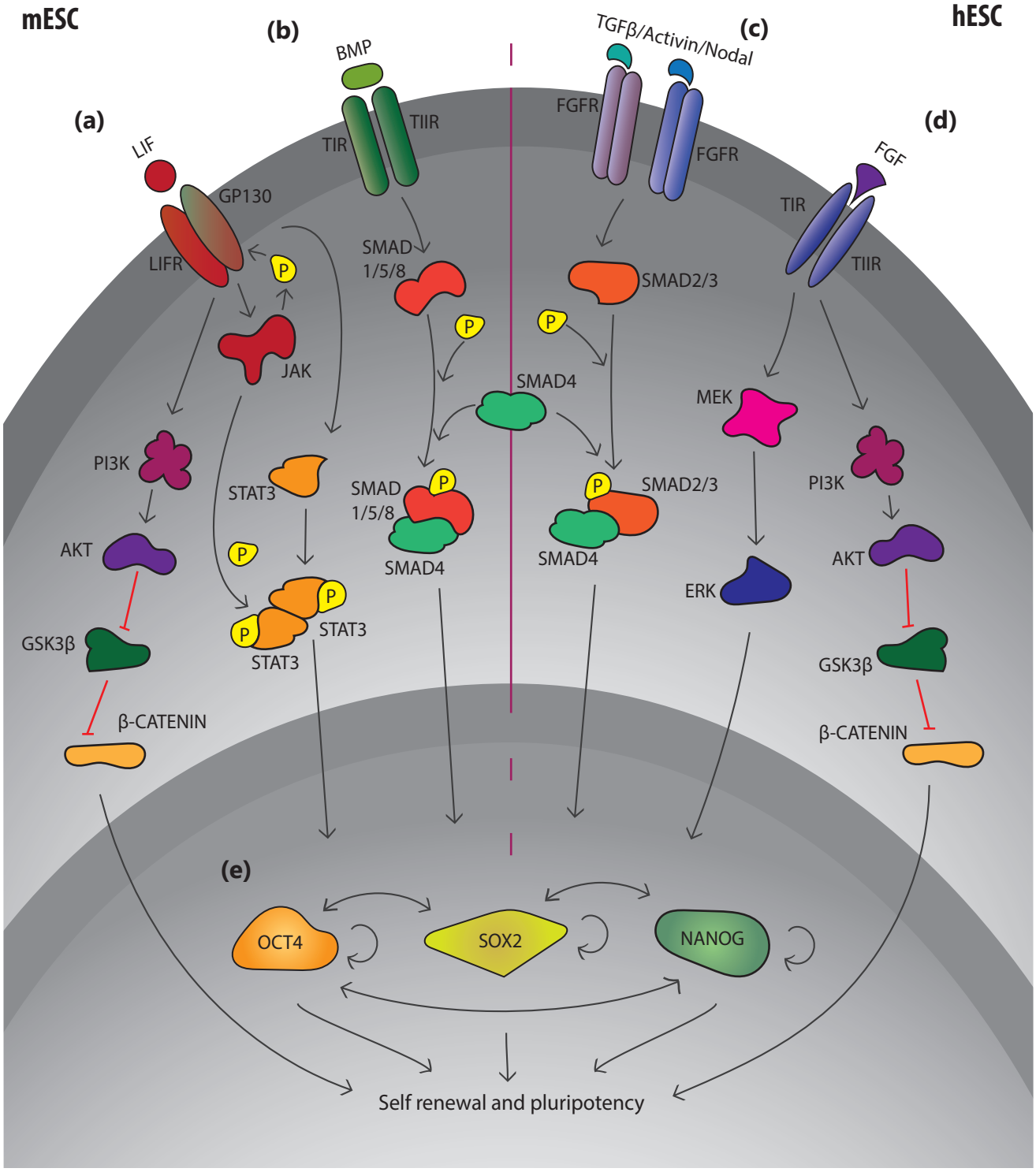
particular example involves the use of a stem cell niche to create the right environmental cues to induce reprogramming (Lim & Gong 2013).

1.2 CELL SIGNALLING PATHWAYS CONTROLLING PLURIPOTENCY

Understanding the molecular mechanisms through which ES and iPS cells are maintained in a pluripotent state is likely to be central in future applications, as a consequence a great deal of progress has been made on this front in recent years. A pluripotent cell's decision to maintain their stem cell identity or differentiate is based on levels the transcription factors Oct4, Sox2 and Nanog, otherwise known as the master regulators (Boyer et al. 2005; Loh et al. 2006; Ding et al. 2009). Levels of the master regulators are controlled via complex extracellular signalling pathways, which cascade down to transcriptional regulatory networks and chromatin remodelling complexes (figure 1.2). Transcriptional control is intricate and multi-layered allowing both flexibility and rapid response to external cues, thereby enabling cells to differentiate into any cell type excluding the extra embryonic tissue. Growth of ES and iPSCs in culture requires specialised culture medium with growth factors to stimulate the extracellular signalling pathways and maintain pluripotency.

Different signalling pathways are required to support pluripotency in mice and humans. LIF, a member of the interleukin6 cytokines has been identified as the key factor in maintaining pluripotency in mES cells by inhibiting their differentiation (figure 1.2)(Smith et al. 1988; Williams et al. 1988). LIF regulates self-renewal in mES cells by binding to the LIF receptor (LIFR) on the cell surface which subsequently forms a heterodimer with GP130 and activates JAK (Janus Kinase). JAK phosphorylates GP130 allowing it to bind STAT3 (signal transducers and activators of transcription 3)(H. Niwa et al. 1998). JAK then phosphorylates STAT3 which homodimerises and passes through

Figure 1.2 Signalling pathways involved in pluripotency



A depiction of the cell signalling pathways involved in regulating pluripotency. The importance of each pathway varies between mouse embryonic stem cells and human embryonic stem cells. For mouse embryonic stem cells (a) LIF via the JAK/STAT pathway, and (b) the BMP/SMAD pathway are essential in maintaining stem cell pluripotency and self renewal. Whereas (c) the TGFβ/Activin/Nodal pathway, and (d) the FGF/MEK pathway are central for human embryonic stem cells. Within the nucleus the master regulators of pluripotency are the same in mice and humans; OCT4, SOX2 and NANOG. These three transcription factors regulate a large number of genes including themselves and each other to maintain a cells pluripotency and self renewal.

to the nucleus. Once in the nucleus phosphorylated STAT3 binds to gene enhancers, often at sites already occupied by the master regulators Oct4, Sox2 and Nanog at the sites of differentially expressed genes presumably supressing expression (Reich & Liu 2006; Chen et al. 2008; Kidder et al. 2008). Knocking down STAT3 was able to activate endodermal genes and mesodermal genes which, along with other early differentiation markers, have been identified as STAT3 targets, demonstrating that pluripotency is maintained through the suppression of lineage specific genes (Bourillot et al. 2009).

LIF has also been found to influence other downstream signalling pathways including the PI3K and the YES pathway (Tamm et al. 2011). Although LIFR and GP130 are also present in human cells and the use of human LIF in culture also produces phosphorylated STAT3 in the nucleus, this is not sufficient to maintain pluripotency in hES cells (Dahéron et al. 2004).

Other cell signalling pathways associated with pluripotency and stem cell fate include the TGFβ superfamily consisting of more than 40 members such as TGFβ, activin, nodal and BMPs (bone morphogenic proteins), Fibroblast growth factor (FGF) and the WNT signalling pathways. BMPs are able to replace serum, if used alongside LIF in the generation of mES cells; BMPs lead to the activation of differentiation gene inhibitors known as the Id (inhibitor of cell differentiation) genes which prevents neutral differentiation and promotes self-renewal (Ying et al. 2003). In contrast BMP signalling in humans has been linked to differentiation towards extra embryonic cells, using BMP inhibitors in culture can aid long term self-renewal and pluripotency in hES cells (Gonzalez et al. 2011). The TGFβ/activin/nodal pathway plays a greater role in hESC maintenance than BMPs, which is evident by the ability of Activin A (a member of the TGFβ super family) to replace a MEF feeder layer in hESC culture (Beattie et al. 2005).

Human ESCs are grown in the presence of FGFs, the inhibition or withdrawal of which rapidly induces differentiation. The up-regulation of all four FGF receptors is observed in undifferentiated hESCs when compared to differentiated hESCs. The exact mechanism through which FGFs operate has not yet been determined but several pathways are thought to be involved including FGF/MEK, phosphatidylinositol 3-kinase (P13K) and protein kinases (MAPKs) (figure 1.2) (Ding et al. 2010).

Interestingly although FGFs are important in maintaining pluripotency in hESCs mainly through the FGF/MEK pathway, the inhibition of this pathway in mESCs promotes pluripotency (Burdon et al. 1999; Bendall et al. 2007; Greber et al. 2007). It appears that the most important pathways for the maintenance of pluripotency in mESCs are STAT3 and BMP whereas TGFβ and FGF/MEK are central for hESCs (figure 1.2).

During differentiation other signalling cascades are triggered resulting in the downstream activation of lineage specific genes that lead to the loss of pluripotency and establish a normal cell cycle. Triggers for this will originate from the cells microenvironment and will either be chemical or physical contact based.

Signalling pathways ultimately allow external factors to influence pluripotency via a cells transcriptional network (figure 1.2); central to this are Oct4, Sox2 and Nanog. The *Pou5f1* gene which encodes the Oct4 protein is highly expressed in both mouse and human ES cells. Upon differentiation *Pou5f1* expression drops and pluripotency is lost. Oct4 is known to regulate a large number of genes including many others associated with self-renewal and proliferation such as *Fgf4, Utf1, Rex1, Sox2, Nanog* and itself (Boyer et al. 2005; Loh et al. 2006; Liang et al. 2008). The stem cell state is believed to exist in the middle region of a bell curve of *Pou5f1* expression. Repression of *Pou5f1* has been associated with differentiation into extra-embryonic tissue, namely

trophectoderm, whereas overexpression results in differentiation towards mesoderm and endoderm (Chen et al. 2008). Oct4 forms complexes with many proteins including Sox2, another of the master regulators, many genes specific to pluripotent stem cells possess an enhancer region which recognises this particular combination (Oct-Sox) including the genes for all three master regulators (*Pou5f1, Sox2, Nanog*), *Utf1, Lefty* and *Fgf4* . Oct4, Sox2 and Nanog have been found to bind sites in close proximity to one and other suggesting interaction between them. Half of all Oct4 binding sites also recognise and bind Sox2 and of these a further 90% bind Nanog (Boyer et al. 2005; Ng & Lufkin 2011). This is further supported by the greater conservation of these linked binding sites between mice and human cells, when compared to binding sites of each component individually (Boyer et al. 2005; Loh et al. 2006; Mathur et al. 2008). However, Sox2 is not only found in pluripotent stem cells but is also important for the development of the nervous system; Sox2 and the other Sox proteins are vital for the correct activation of neural genes (Wegner 2011).

Despite being a core pluripotency factor specifically expressed in ES cells and intimately linked in transcriptional networks alongside Oct4 and Sox2 understanding of the role of Nanog remains remarkably limited. Nanog is able to maintain self-renewal in ES cells independent of the LIF pathway. Nanog depletion leads to differentiation down an extra embryonic endoderm lineage whereas when overexpressed pluripotency is maintained (Chambers et al. 2003; Mitsui et al. 2003). Nanog, like Oct4, is involved in a positive feedback loop through which it activates its own expression (Boyer et al. 2005). Interestingly it was not found to be necessary for the generation of iPSCs, although it has since been used as one of the introduced transcription factors (Takahashi & Yamanaka 2006; Yu et al. 2007).

Although mice and human ES cells appear to have distinct signalling pathways involved in the maintenance of pluripotency they result in very similar transcription factor networks and expression profiles including the same three core transcription factors- Oct4, Sox2 and Nanog- suggesting very similar molecular mechanisms at the gene level. Given that a high degree of conservation in a system is directly linked to the systems biological importance principle finding from mice are generally expected to be applicable to other mammals including humans.

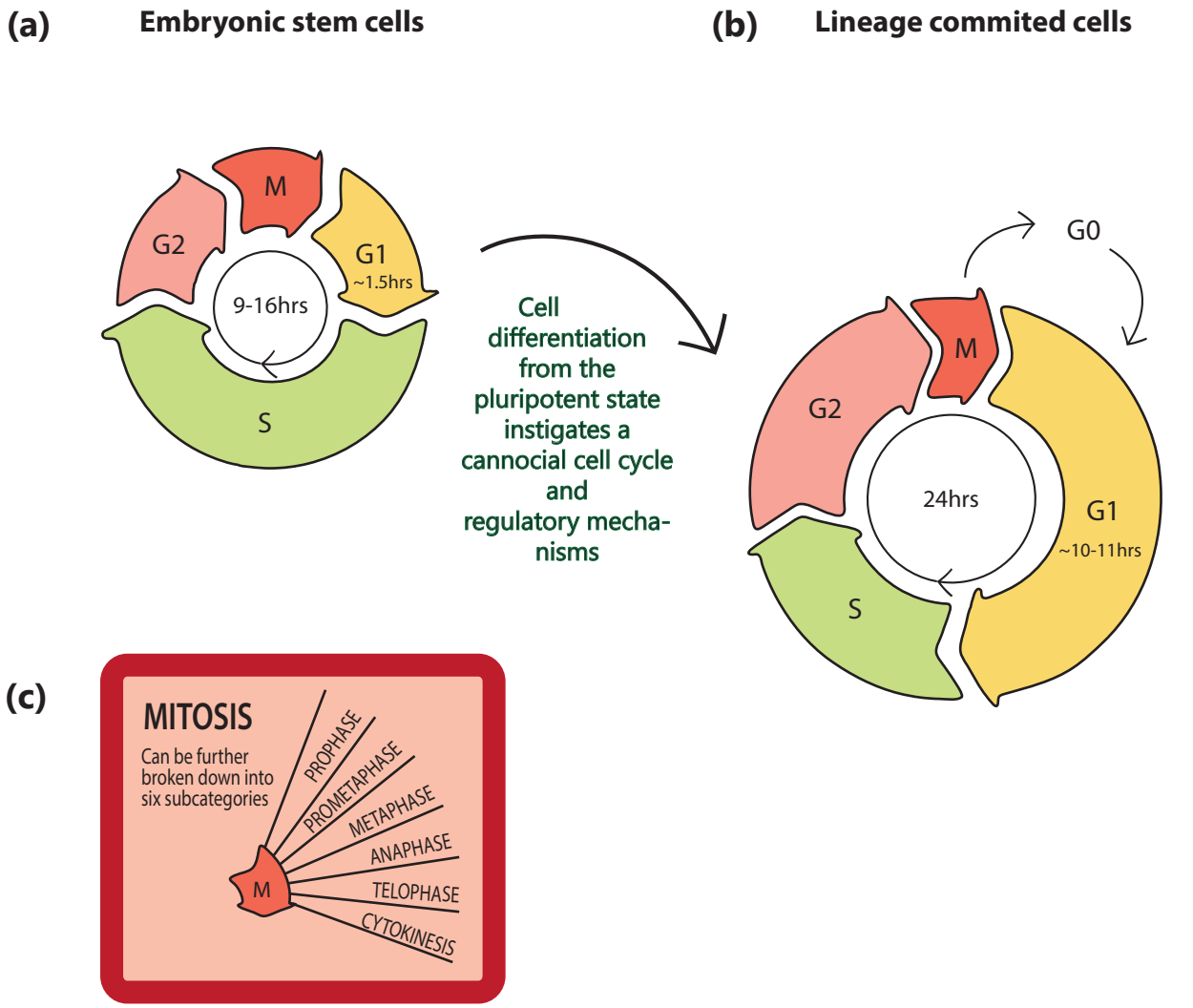
1.3 THE CELL CYCLE

Cells within a developing embryo must be able to divide rapidly and so one of the key defining features of ES cells is their uniquely fast proliferation, facilitated by an unusually short cell cycle. The normal somatic cell cycle is a robust biochemical timer in which cell cycle events occur at the correct time and in the correct order controlled through a complex network of cell signalling pathways and although the vast majority of these pathways are the same in ES cell there are differences.

In ES cells pluripotency and the cell cycle are regulated by signalling pathways with a degree of overlap (Neganova & Lako 2008). The delicate balance between rapid proliferation and differentiation is facilitated by a distinctive cell cycle structure. Several pluripotency associated transcription factors including Nanog, Oct4, Sox2 and Myc directly and indirectly affect the transcriptional regulation of cell cycle genes and even cell cycle related microRNAs (Neganova & Lako 2008; Wang & Blelloch 2009; Zhang et al. 2009; Stein et al. 2012; Nakai-Futatsugi & Niwa 2013).

The cell cycle, otherwise known as the cell division cycle is the process cells go through in order to divide and has been a topic of research for over 40 years (Nurse 2002). It is

Figure 1.3 The cell cycle in embryonic stem cells and lineage commited cells



The cell cycle in embryonic stem cells compared to that of a lineage commited cell (a) The embryonic stem cell cycle can be divided into gap phase 1 (G1), synthesis phase (S), gap phase 2 (G2) and mitosis (M) it is usually between 9 and 16 hours in length (b) Upon differentiation cells will commit to a normal cell cycle which consists of the same phase but a more rigid regulation that involves the enforcement of checkpoints at the G1/S boundary and the G2/M boundary. This results in a much longer cell cycle of approxi-mately 24 hours. (c) In both cell cycles mitosis can be further divided into the following subcategories; prophase, prometaphase, metaphase, anaphase, telophase and cytokine-sis

comprised of four phases (figure 1.3); G1 phase, or growth/gap phase 1, in which cells grow and prepare for DNA synthesis; S phase, or synthesis phase, during which all the DNA within the cell is duplicated; G2 phase, the second growth or gap phase, is when the cell prepares to divide; followed by M phase, or mitosis, the final stage of the cell cycle where duplicated chromosomes are distributed into two daughter nuclei, followed by the division of the cytoplasm (separately known as cytokinesis). The result of cell division is the successful transmission of genetic information in the form of two genetically identical daughter cells. In addition to the successful transmission of DNA, all cells also synthesise the remainder of the cells contents including all macromolecular complexes during the gap phases which are then divided equally into both daughter cells ensuring that the size of the resulting cells are approximately equal to the original. Upon exiting M-phase a cell must commit to re-entering the cell cycle at G1 phase, or alternatively can enter G0 (or resting) phase where it may remain in a quiescent state unless signalled to reinitiate the cell cycle.

Within the cell cycle are also two main checkpoints, the G1/S checkpoint, and the G2/M checkpoint. These checkpoints serve to maintain genomic stability by pausing cell cycle progression until the extent of DNA damage within the cell has been assessed and repaired, should it reach a critical threshold the cell will undergo apoptosis, if not the cell is released to continue to the subsequent phase.

1.3.1 Regulation of the normal cell cycle

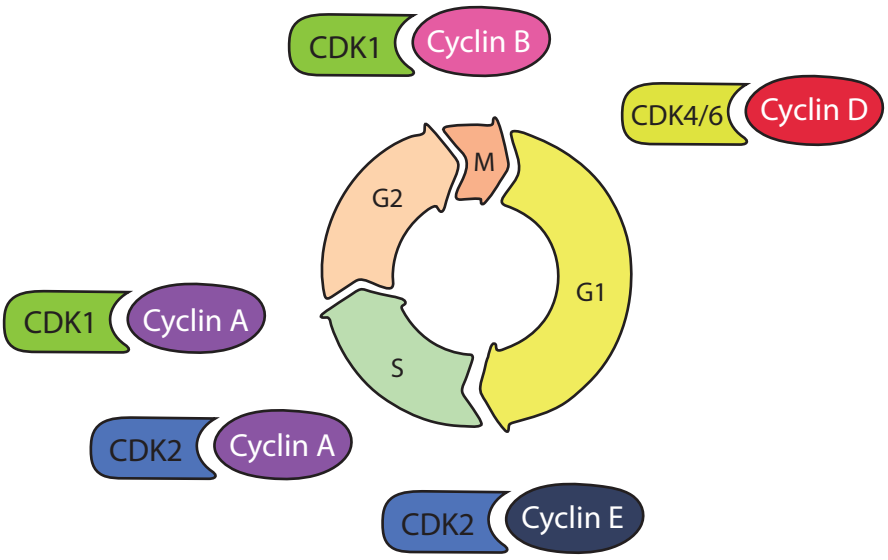
The cell cycle is a highly co-ordinated process. Central to the progression of the cell cycle are the oscillating activities of the cyclin dependent kinases (Cdks). Interestingly the actual levels of the various Cdks remain constant across the cell cycle, it is the expression of the cyclins that fluctuate. Each Cdk has a catalytic subunit that remains

Figure 1.4 Cyclin dependent kinases (CDKs) through the cell cycle

(a)

	Phase	Abbreviation	Cdks	Cyclins
Yeast	First Gap phase	G1	Cdk1	1-3
	Synthesis	S	Cdk1	5 + 6
	Second Gap phase	G2	Cdk1	
	Mitosis	M	Cdk1	1-4
Mammal	First Gap phase	G1	Cdk4, Cdk6	D1, D2, D3
	Synthesis	S	Cdk2	E + A
	Second Gap phase	G2	Cdk1	A
	Mitosis	M	Cdk1	B

(b)



Categorisation of the cyclin dependent kinases throughout the cell cycle. (a) A table of Cdks and corresponding cyclins found at each cell cycle phase in yeast and mammals. (b) A visual representation of the Cdk-cyclin complexes and the point at which they are present across the cell cycle.

inactive until bound to a corresponding regulatory partner, a cyclin (figure 1.4). Formation of Cdk-cyclin complexes activates the catalytic subunit of the Cdk which is then able to phosphorylate a large number of substrates ultimately driving cell cycle progression.

Cyclins D1, D2 and D3 (otherwise known as CCND1, CCND2 and CCND3) are expressed in early G1 phase of the cell cycle and bind to Cdk4 and/or Ckd6 (Cdk4/6) forming Cdk/cyclin complexes that subsequently drive G1 phase events (Sherr 1993; Draetta 1994; Sherr 1994). Amongst the many substrates of this complex is the G1/S checkpoint protein retinoblastoma protein (Rb) (Ewen et al. 1993; Kato & Sherr 1993), which is phosphorylated causing its release from the E2F/Rb complex (Helin et al. 1993; La Thangue 1994a). This step is crucial for the progression from G1 to S-phase, free E2F is a transcription factor which activates the expression of a plethora of genes required in S phase including Cyclins A and E and DNA polymerase (Helin et al. 1993; Johnson et al. 1994; La Thangue 1994b; DeGregori et al. 1995; Duronio & O'Farrell 1995). Cyclin E binds to the already present Cdk2 and the resulting complex also phosphorylates Rb increasing the amount of free E2F within the cell pushing the cell into S-phase. DNA synthesis is triggered and a complete copy of the genome is generated.

During S-phase the release of E2F promotes low levels of cyclin A (Ccna) expression which also binds to Cdk2. The Cdk2/Cyclin A complex has a different binding specificity to Cdk2/Cyclin E and therefore phosphorylates a different range of substrates. As G2 phase commences Cyclin A associates with Cdk1 (otherwise known as Cdc2) which stimulates progression through G2 phase. During G2 phase, expression of Cyclin B is triggered which also forms a complex with Cdk1 (Cdk1/Cyclin B), this begins nuclear envelope breakdown and the initiation of mitosis. (figure 1.4). Cyclin B is expressed as

the cell leaves G2 phase, entering mitosis, and complexes with Cdk1. Cdk1/cyclin B is essential in driving mitosis and its inactivation promotes mitotic exit (Nurse 1990; Coleman & Dunphy 1994; King et al. 1994). Cells in which cyclins cannot be broken down are unable to exit mitosis.

Mitosis can be further broken down into six subcategories (figure 1.3). Prophase; during which the mitotic spindle is formed and chromatin condenses into its most compact form, chromosomes. Prometaphase; as the nuclear membrane is broken down and chromosomes form kinetochores which attach to microtubules from the opposite poles. This is followed by metaphase when chromosomes align at the centre of the cell. Metaphase ends with the destruction of cyclin B via the ubiquitin/proteasome pathway. At anaphase the release of separase is triggered which breaks down cohesin, the protein responsible for keeping sister chromatids together, subsequently allowing them to migrate towards opposite poles. Once separated telophase commences and an individual nuclear envelope is formed around each set of chromatids which then unwind to a less condensed form of chromatin once again. Cytokinesis, the division of the cytoplasm to form two daughter cells, takes place in parallel with nuclear envelope reformation.

As a testament to the importance of cell cycle control Cdk activity is regulated by at least three separate mechanisms; the presence and binding capacity of cyclins as just described; repression via Cdk inhibitory proteins (Dunphy 1994; Elledge & Harper 1994; Sherr & Roberts 1995); and finally phosphorylation of various Cdk subunits (Norbury & Nurse 1991; Reed 1992; Morgan 1995; Nigg 1995; Nasmyth 1996). Cdk inhibitors can be divided into two families; the Cip/Kip family which target cyclin D, E

and A dependent kinases and the INK4 family known only to block the association between Cdk4/6 and cyclin D (Serrano et al. 1993; Xiong et al. 1993).

Phosphorylation of Cdks can occur at a number of locations thereby rendering them active or inactive. Phosphorylation of a threonine residue (the exact location of which varies depending on the specific Cdk) is actually required for the full activation of the kinase by bringing about a conformational change, enhancing stability of the cdk/cyclin complex and allowing ATP access (Ducommun et al. 1991). Negative regulation by phosphorylation is also observed in all Cdks (Solomon 1994; McGowan & Russell 1995; Mueller et al. 1995). This form of negative phosphorylation can be removed by the activity of the cdc25 phosphatases (cdc25A, cdc25B, cdc25C), which are also connected to cell cycle checkpoint pathways and DNA damage response (Millar & Russell 1992; Demetrick & Beach 1993; Jinno et al. 1994).

1.3.2 CELL CYCLE CHECKPOINTS

During the course of the cell cycle DNA can be damaged or altered in many ways, most notably at either DNA replication (S-phase) or chromosome segregation (M-Phase). To ensure that DNA is inherited faithfully from one generation of cells to the next the cell cycle contains a number of checkpoints. The main checkpoints are the G1/S and G2/M checkpoints, a checkpoint also exists within S phase but differs from the other two as it regulates DNA synthesis not just DNA damage, it will not be covered here (Falck et al. 2002). These pathways are present during the entire cell cycle however only result in the inhibition of cell cycle progression at select points, these are the checkpoints.

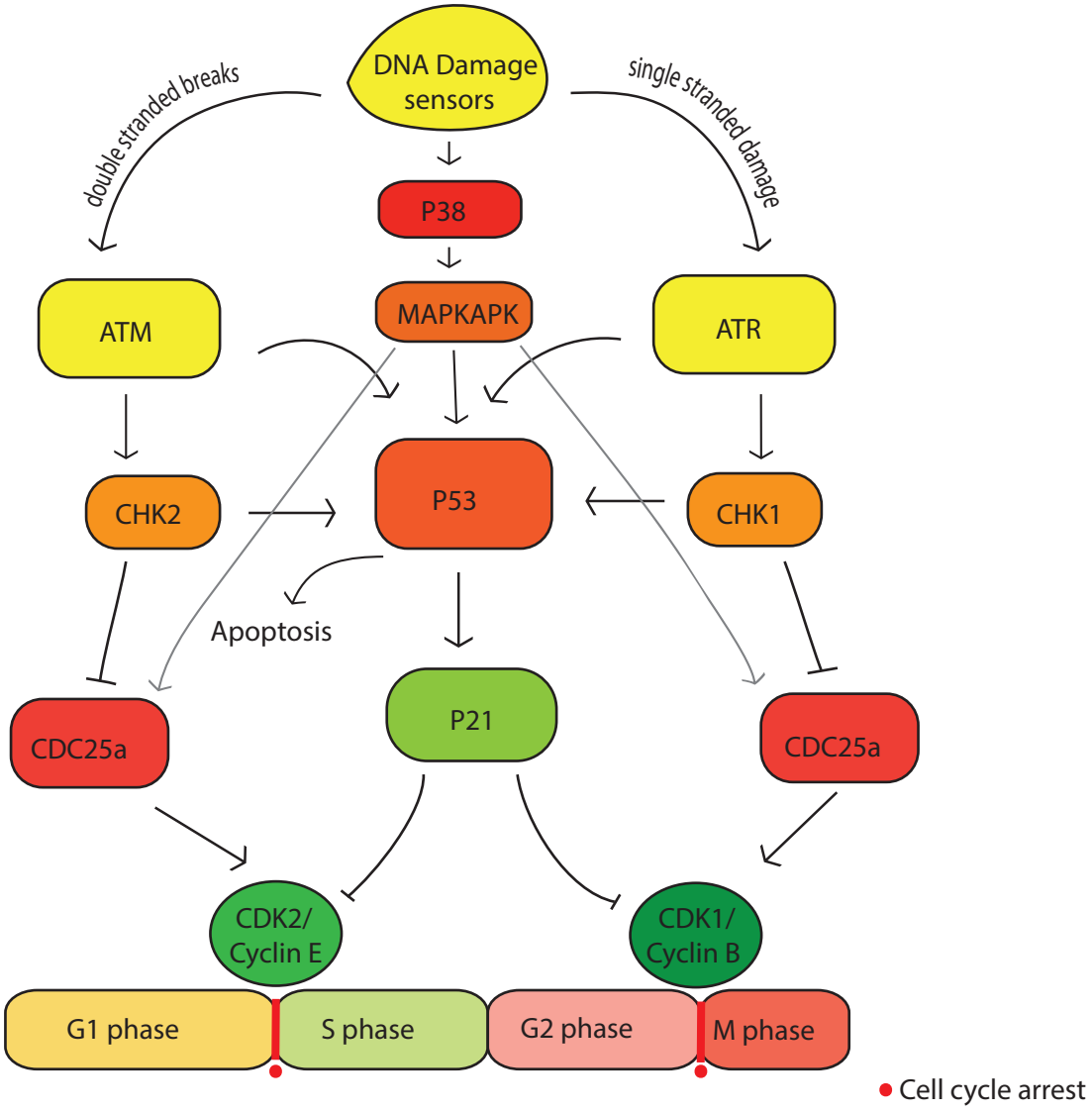
In eukaryotes all of the checkpoints share key elements; however the exact pathway depends on the kind of DNA damage sustained. Any damage is detected by sensor proteins, the signal is passed on to transducer proteins, followed by effector proteins

that are able to initiate cellular events including pausing progression of the cell cycle and triggering DNA repair mechanisms. If the damage is too extensive the cell will either remain blocked (essentially in a senescent state), undergo apoptosis or potentially undergo oncogenesis (Houtgraaf et al. 2006; Bartek & Lukas 2007).

The G1/S checkpoint is primarily enforced by the inhibition of Cdk2 action (Bartek & Lukas 2001; Nojima 2004). This can be brought about through a number of different pathways such as the ATM (ataxia telangiectasia mutated) or ATR (ATM and Rad3 related) repair pathways (figure 1.5). ATM and ATR are sensors that respond to different types of DNA damage, double stranded breaks and UV damage respectively, once triggered they are able to activate by phosphorylation a plethora of downstream proteins including p53 tumour suppressor protein; Chk1 and Chk2. Chk1 and Chk2 phosphorylate Cdc25A thereby labelling it for nuclear exclusion and proteosomal degradation. Cdc25A is responsible for the dephosphorylation of Cdk2 enabling it to bind to cyclin E. Lack of cdc25a therefore halts cell cycle progression prior to DNA synthesis. An alternative pathway is the stabilisation of p53 although a lot of overlap is observed. ATM/ATR → Chk1/2, or p38 → MAPKAPK, phosphorylate p53. This stabilises p53 and promotes p53 transcription factor activity, of which p21 is a target. P21 in turn inhibits Cdk2 activity and subsequently stalls cell cycle progression.

The G2/M checkpoint ensures that mitosis is not entered with any loss of genome integrity. It is also comprised of the same two key pathways, p53 independent pathway acting via removal of cdc25 proteins and a p53/p21 dependent pathway (figure 1.5) (Houtgraaf et al. 2006). There are differences between the two checkpoints, the foremost of which being the target Cdk; the actions of Cdk1 rather than Cdk2 are inhibited by phosphorylation in order to pause cell cycle progression at the end of G2

Figure 1.5 Cell cycle checkpoint signalling pathways in response to DNA damage



Categorisation of the cyclin dependent kinases throughout the cell cycle. (a) A table of Cdks and corresponding cyclins found at each cell cycle phase in yeast and mammals. (b) A visual representation of the Cdk-cyclin complexes and the point at which they are present across the cell cycle.

phase (O'Connor et al. 1994; Agarwal et al. 1995). Ordinarily at the end of G2 phase Cdc25b and Cdc25c, in addition to Cdc25a, are responsible for Cdk1 dephosphorylation. When the G2/M checkpoint pathways are being enforced all Cdc25 family proteins are targeted for removal from the nucleus.

1.3.3 THE EMBRYONIC STEM CELL CYCLE

ES cells do not follow this normal cell cycle strictly, instead they are able to rapidly proliferate as they spend very little time in the gap phases compared to somatic cells and have a mostly absent G1/S checkpoint which is believed to be rate limiting in a canonical cell cycle (White & Dalton 2005; Neganova & Lako 2008; Ballabeni et al. 2011). Only an approximate 1.5 hours is spent in G1 phase with over half of the cell cycle spent in S-phase (Fujii-Yamamoto et al. 2005). The average time spent completing the cell cycle is only 9 hours for mES cells compared to around 24 hours for somatic cells (Nagoshi et al. 2004; Stein et al. 2012).

Many factors come together to bring about this distinctive cell cycle including high expression levels of the cyclins, low expression levels of Cdk inhibitors such as p21 and p27, and no expression of Rb protein (Savatier et al. 1994; Becker et al. 2006; Ballabeni et al. 2011; Li et al. 2012). Rb family proteins are hyperphosphorylated in ES cells rendering them inactive. Inhibition of Rb protein contributes to the lack of a G1/S checkpoint by releasing E2F thus permanently promoting the G1/S transition, this renders cells lacking the Rb protein as deficient for this DNA damage response (Bosco & Knudsen 2005).

It used to be believed that ES cells lacked the cell cycle dependent oscillations in cyclin and Cdk activity of somatic cell. A recent investigation however, has found that these two elements do oscillate in the same manner as observed in somatic cells but the

degree of the oscillations is more muted in ES cells by comparison (Ballabeni et al. 2011). In either case this noticeable disparity undoubtedly plays a role in the differing cell cycle structures.

The inability to fully implement the G1/S checkpoint in ES cells is in part brought about by the sequestering of Chk2 to centrosomes, whereas in somatic cells it is free within the nucleus, this prevents the phosphorylation of downstream proteins including p53 and cdc25A (Hong et al. 2007). Also key to the G1/S transition in ES cells are micro RNAs targeted to key regulators of the G1/S checkpoint, these include the p21 targeting miR-302 cluster in hES cells and the miR-290 cluster in mES cells (Wang & Blelloch 2009; Dolezalova et al. 2012; Krishnakumar & Blelloch 2013; Wang et al. 2013). As ES cells differentiate a proper G1/S checkpoint is activated, similarly expression of these microRNAs drops.

The truncated G1 phase found in ES cells has been linked to maintaining cells in an undifferentiated state as they are believed to be more prone to differentiation signals in G1 phase. Various experiments have been conducted using CDK2 inhibitors to stall cell cycle progression to determine whether G1 phase itself promotes differentiation with conflicting results. A loss of pluripotency markers was observed in some studies, however this was not always the case, further research will be required to determine whether G1 phase in itself promotes differentiation (Miller et al. 2007; Neganova & Lako 2008; Li et al. 2012).

Elements of ES cell pluripotency overlaps with and ties into their exceptional cell cycle. Nanog has been found to bind directly to key cell cycle elements cdk6 and Cdc25A aiding hESCs in the G1/S transition (Zhang et al. 2009). Both pluripotency and cell cycle are reliant upon the cells ability to rapidly and precisely transcribe sections of the

genetic code as required and as such connections have been identified between them and chromatin modifying elements. The Rb protein not only has HDAC (Histone deacetylase) activity but it is known to interact with a histone methyl transferase (SUV39H1) and HP1 (methyl lysine binding protein) targeting it to methylate histone H3 at lysine 9 (a repressive modification) at the cyclin E promoter (Nielsen et al. 2001).

1.4 THE STRUCTURE OF CHROMATIN

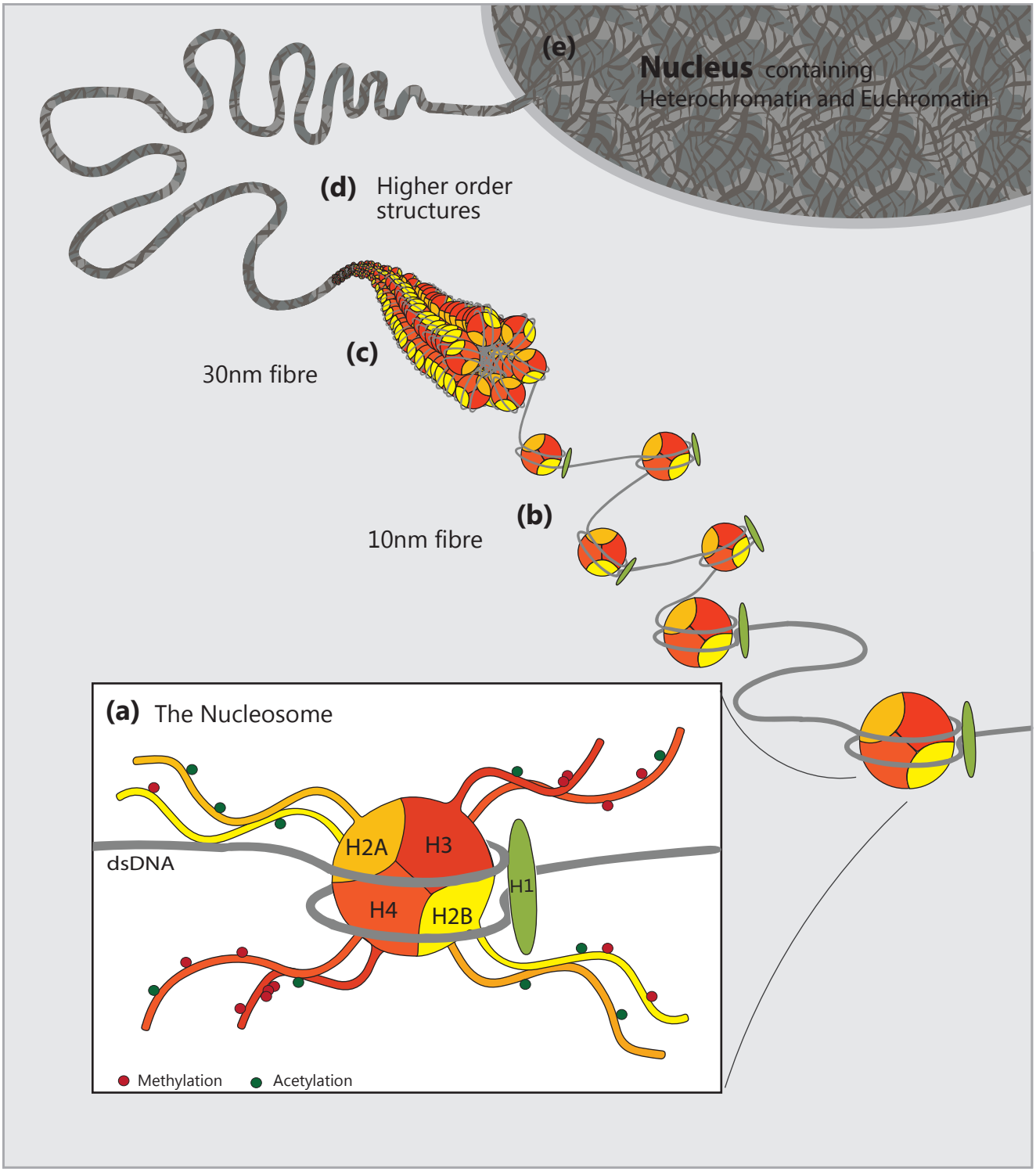
Cell identity is fundamentally the combination of a cells biochemical characteristics and phenotype as brought about by a very specific gene expression profile. A multitude of diverse factors must come together for a cell to achieve and maintain its identity starting with DNA. A vast amount of DNA is stored within the nucleus of all eukaryotes. Metres of DNA is found in most mammalian cells which must be packaged very carefully into the nucleus; organisation of this structure is of vital importance in order to achieve the control over transcription that is required for a functional cell.

DNA is found within the nucleus in the form of heterogeneous chromatin fibres; a complex consisting not only of DNA, but also proteins (primarily histones) and RNAs which condenses into distinctive chromosomes during mitosis (Bassett et al. 2009). Understanding the organisation of chromatin within the nucleus has proven to be extremely difficult and remains an issue of much debate to this day (Woodcock & Ghosh 2010). In humans only around 1.2% of DNA is made up of protein coding exons however, it has recently been found through the ENCODE project that over 80% of genomic sequences do have a functional role (Bernstein et al. 2012). Many more long range regulatory elements such as enhancers, repressors/silencers or insulators exist than was previously anticipated highlighting the importance of the 3D structure of chromatin and its location within the nucleus. Packaging DNA must be done in a way

that not only allows three dimensional interactions of DNA but also enables a variety of essential cell processes to be carried out including the transcription of genes and replication of DNA. These processes require that many elements such as transcription factors, RNA polymerase II complexes and a plethora of histone recognising molecules access relevant parts of DNA as required.

Nucleosomes are the first step in chromatin packaging. The structure of these round disc like protein-DNA complexes was determined to a near atomical level (2.8 Å in 1997 and 1.9 Å in 2002) by x-ray crystallography (Luger et al. 1997; Davey et al. 2002; Richmond & Davey 2003). The nucleosome core is comprised of an octamer of four histones; H2A, H2B, H3 and H4, which are amongst the most highly conserved proteins found in eukaryotes (Felsenfeld & Groudine 2003). Two H3-H4 dimers interact to form a tetramer, $(H3\text{-}H4)_2$, which is sandwiched between two H2A-H2B dimers (Luger et al. 1997)(figure 1.6). This core histone octamer is wrapped in approximately 1.67 superhelical turns, 147 base pairs of DNA from which the N-terminal tails of histone protrude potentially interacting with DNA and DNA affecting molecules (Luger et al. 1997; Davey et al. 2002; Richmond & Davey 2003; Turner 2007; Probst et al. 2009). In most cases the nucleosome is associated to linker histone H1 which serves to stabilise the nucleosome in further compaction. Nucleosomes are positioned on the DNA at intervals of varying length dependent on the organism and compaction state (Bassett et al. 2009). This 10nm fibre represents the primary structure of chromatin. Higher orders of packaging of increasing complexity can form from here such as the 30nm fibre (30nm solenoid) and eventually the chromosome. Chromosomes are formed during M-phase and are the most dense form of DNA observed, compacted by 10,000-20,000 fold (Woodcock & Ghosh 2010).

Figure 1.6 Structure of the nucleosome and higher order packaging of chromatin



(a) The nucleosome is the primary packaging scaffold for DNA within the nucleus. 147 base pairs of double stranded DNA (dsDNA) is wrapped around each nucleosome which is comprised of a core histone octamer of two H2A-H2B dimers and a (H3-H4)2 tetramer. Flexible tails protrude from the core octamer at both the amino and carboxyl end of all histones which are subject to multiple post translational modifications the most prevelent of which are methylation and acetylation on the N-terminal tails. (b) This structure is a 10nm fibre often refered to as 'beads on a string'. Upon further packaging a (c) 30nm fibre is formed followed by further (d) looping and compaction. (e) Within the nucleus both euchromatin or the more compact heterochromatin are found. The most compact structure of DNA is the fully condensed mitotic chromosome.

Different compaction states of chromatin are found preferentially at certain sites within the nucleus which correlate to gene activity (Marshall 2002; Woodcock & Ghosh 2010).

Spatial localisation within the nucleus is believed to play a large role in the regulation of gene expression for example proximity to the nuclear envelope is considered to be a transcriptionally repressive environment (Somech et al. 2005). There are two major types of chromatin found within cells heterochromatin and euchromatin. Heterochromatin is found adjacent to the nucleus edge and euchromatin is dispersed more evenly. Euchromatic DNA is compacted by a degree of ~1,000 fold and is generally more flexible and open than heterochromatin, therefore accessible to transcriptional regulators, and DNA repair and replication complexes. For any DNA process to occur, access to the sections of DNA around nucleosomes must be possible. Although euchromatin is considered open and active not all genes within euchromatin are actively transcribed. There is a high degree of variability between different cell types, however, human tissues generally only express between 60-85% of all protein coding genes and it is thought that less than 10% of these are being expressed at any given time (Ramsköld et al. 2009). Heterochromatin is compacted approximately ~10,000 and has very few genes by comparison. It is rich in A-T nucleotide bases, transcriptionally silent and highly condensed. Heterochromatin offers protection to the centromeres and telomeres and is important for their separation during mitosis (Müller & Almouzni 2013).

1.5 EPIGENETICS

In mammalian cells, DNA, unlike RNAs and proteins, does not vary in sequence or quantity between cell types (with a few key exceptions such as genetic variation as a result of copying errors and also abnormal cell types such as erythrocytes and gametes).

DNA contains the code for everything within a cell, but it is the carefully controlled expression of select parts of the genome that really give a cell its identity and as such epigenetics plays a central role. Epigenetics is defined as the study of heritable changes in gene function that occur without a change in the sequence of nuclear DNA (Berger et al. 2009). Therefore epigenetic factors are essential in the regulation of gene expression patterns. Elements currently considered to be epigenetic are DNA methylation, RNAi's and post translational histone modifications (PTMs).

DNA methylation refers to the methylation of cytosine by DNA methyltransferase (DNMT) enzymes primarily at CpG dinucleotides resulting in two 5-methylcytosine residues sitting diagonally on opposing DNA strands. In mammals 5-methylcytosine is found in 2-7% of the genome particularly at CpG Islands which have a high CpG content. DNA methylation at promoter regions is associated with the repression of expression. The absence of DNA methylation generally associated with active chromatin as mechanistically it prevents the binding of transcription factors and also allows the binding of proteins such as MeCP2 which recruit chromatin silencers (Jones et al. 1998). DNA methylation is essential in silencing DNA during embryonic development, genomic imprinting and X-chromosome inactivation (Cedar & Bergman 2009; Jones 2013). Abnormal DNA methylation has been linked to many diseases including cancer, many of which show hyper methylation of tumour suppressor genes, as well as cell cycle regulatory genes and those involved in DNA repair (Barton & Levine 2008)(McGarvey et al. 2006; Ball et al. 2009; Cedar & Bergman 2009; Jones & Liang 2009; Blomen & Boonstra 2011; Ficz et al. 2011).

RNA interference (RNAi) is emerging as a potential epigenetic element. Growing evidence suggests that RNAi regulates transcription by interacting both directly and

indirectly with the transcriptional machinery. In addition to this, recent studies into *Arabidopsis thaliana,* have led to the hypothesis that some RNAi's move into germ cells and are therefore heritable (Castel & Martienssen 2013). RNAi promise to be an essential component in the regulation of a multitude of cellular processes including transcription, however the exact roles remain to be determined.

As a primary element in the organisation of DNA within the nucleus histones are ideally situated to be able to influence DNA related processes. The four core histones in the nucleosome can all be post-translationally modified at a variety of positions and in a number of ways. The location of nucleosomes on DNA is far from random, the majority of genes have a similar chromatin architecture with favoured positions within the genome. Nucleosome depleted region are often observed within enhancers, promoter and terminator regions. In yeast the +1 and -1 nucleosomes are positioned at preferred sites relative to the transcriptional start site, the precision of nucleosome positioning gradually decreases across the coding region to the 3' end (Struhl & Segal 2013).

1.6 POST TRANSLATIONAL HISTONE MODIFICATIONS

Post translational acetylation of histones was first described around 50 years ago by Allfrey et al (1964) who noted that when histones were acetylated the rate of RNA synthesis increased. Further to this they proposed histones as regulators of chromosomal activity and suggested histone modifications, namely acetylation and methylation, as a mechanism for activation and repression of RNA synthesis (Allfrey et al. 1964). Histones have highly basic flexible amino and carboxyl terminal tails which protrude from the core nucleosome structure (figure 1.6) (Luger et al. 1997). These tails contribute approximately 28% of the histones mass and associate with both DNA around the nucleosome and with adjacent nucleosomes. In particular amino acids 16-25

of H4 interact with the H2A-H2B dimer of the neighbouring nucleosome where it is exposed (Luger et al. 1997). Although modifications have been mapped to the C-terminal domains and within the globular domain of histones, in some instances affecting histone interactions with DNA, they are predominantly located at the N-terminal tails (Kouzarides 2007; Tropberger & Schneider 2010; Tan et al. 2011; Molina-Serrano & Kirmizis 2013).

Since the discovery of acetylation as a post translational histone modification a continually expanding list of post translational histone modifications have been identified including acetylation (of lysine), methylation (of arginine and lysine), phosphorylation, proline isomerization, ubiquitination, ADP ribosylation, citrullination (of arginine), lysine crotonylation, SUMOylation, carbonylation and biotinylation (table 1.1) (Tan et al. 2011; Arnaudo & Garcia 2013). Mass spectrometry is an essential tool in the continuing identification of new post translational histone modifications and has been used in combination with other techniques to identify over twelve new types of modification as well as novel locations for known modifications [such as tyrosine hydroxylation, serine and threonine acetylation, lysine crotonylation, lysine N-formylation, lysine succinylation, lysine malonylation, lysine propionylation, lysine butyrylation, O-glcNAcylation, lysine 5-hydroxylation and cysteine glutathionylation] (Tan et al. 2011; Zhang et al. 2011; Arnaudo & Garcia 2013). Additional histone marks continue to be discovered.

Following on from Allfrey's work, post translational histone modifications are now believed to be involved in many biological processes including but not limited to transcription, DNA repair and recombination, developmental decision making,

Table 1.1: Classes of post translational histone modification

Histone Modification	Residues modified	Putative function
Acetylation	Lysine, serine, threonine	Transcription, repair, replication, condensation
Methylation	Lysine, arginine	Transcription, repair
Phosphorylation	Serine, threonine	Transcription, repair, condensation
Ubiquitination	Lysine	Transcription, repair
Sumoylation	Lysine	Transcription
ADP ribosylation	Glutamic acid	Transcription
Deimination	Arginine	Transcription
Proline isomerisation	Proline	Transcription
Crotonylation	Lysine	Transcription
Carbonylation	Cysteine	Transcription, repair, replication
Biotinylation	Lysine	Gene silencing, proliferation, DNA damage response
Hydroxylation	Lysine, arginine, tyrosine	Nd
Formylation	Lysine	Loss of protein function
Succinylation	Lysine	Structural changes to nucleosome
Malonylation	Lysine	Structural changes to nucleosome
Propionylation	Lysine	Nd
Butyrylation	Lysine	Nd
Glutathionylation	Cysteine	Loss of protein function

A selection of the common and new varieties of post translational histone modification including the residues modified and putative roles for the modifications (nd = not determined). This list covers well known modifications and novel modifications but is not a comprehensive list. Data retrieved from (Kouzarides 2007; Tan et al. 2011; Arnaudo & Garcia 2013).

heterochromatin formation and dosage compensation. Many attempts have been made to categorise these post translational histone modifications, for example several types are associated with either euchromatin or heterochromatin; histone acetylation often correlates with open DNA whereas phosphorylation of histone 3 at serine 10 (H3S10phos) is associated with condensed chromatin, particularly during metaphase. Many histone modifications are highly conserved from yeast to humans, suggesting that such marks are of fundamental importance and although global patterns are very variable across cell types the histone tail domains on H3 and H4 are required for the normal progression of the cell cycle (Morgan et al. 1991; Ernst et al. 2011). Histone modifications have been linked to a wide variety of biological functions and are often distributed in specific patterns (table 1.2) presumably related to function. There are two mechanisms by which histone modifications exert an effect, direct structural change, which is thought to exert a relatively subtle effect compared to the control and regulation of chromatin factor binding. Regardless of the mode of action it is generally accepted that histone modifications exert a functional effect despite controversy over whether or not their role is causative (Henikoff & Shilatifard 2011).

Post translational histone modifications have also been linked to a variety of diseases such as cancer providing yet another reason for their study. Histone modifications are often aberrant in cancer, hypermethylation of the DNA at tumour suppressor gene promoters is often accompanied by a global loss of histone acetylation which has led to the use of drugs that target histone modifying enzymes in treatments (Di Cerbo & Schneider 2013; Bojang & Ramos 2014).

Table 1.2 Common histone modifications, their distribution and function

Histone modification	Signal characteristics	Putative function
H3K4me1	Peak/region	Associated with enhancers and downstream of the TSS
H3K4me2	Peak	Associated with enhancers and promoters
H3K4me3	Peak	Associated with promoters and TSS
H3K9ac	Peak	Associated with promoters
H3K9me1	Region	Located at 5' end of genes
H3K9me2	Peak/region	Repressive mark
H3K9me3	Peak	Repressive mark associated with constitutive heterochromatin
H3K27me3	Region	Repressive mark established by polycomb complex
H3K36me3	Region	Located at 3' regions after 1st intron, associated with elongation
H3K79me2	Region	Transcription associated mark located at 5' end of genes
H4K20me1	Region	Located at 5' end of genes
H4K8ac	Peak	Active mark located at TSS and along gene bodies
H3S10phos	Region	Associated with chromatin condensation at mitosis

A selection of several common histone modifications the functions attributed to these marks. As can be seen within this table different methylation states of the same residue can result in changes to the role of the modification. Also included in the table are the distribution characteristics of each mark. Table based on (Bernstein et al. 2012).

1.6.1 HISTONE ACETYLATION

Post translational acetylation is observed on conserved lysines of the N-terminal tails of all four core histones. Acetylation was the first post translational histone modification to be associated with gene activation and has since been identified as the most frequent class of post translation histone modification (Allfrey et al. 1964). As such acetylation is one of the best characterised histone modifications alongside methylation, an overall enrichment of acetyl marks is generally observed at transcriptionally active promoters. Abundance of histone acetylation is controlled by the levels of two families of enzyme with opposing actions: histone acetyl transferases (HATs) and histone deactylases (HDACs). All of these enzymes have highly conserved domains strongly supporting the case for their importance in cell survival. Histone acetylation has a rapid turnover rate with a half-life of minutes due to the abundance of histone deacetylase enzymes (HDACs), this has been linked to rapid changes in gene expression (de Ruijter et al. 2003; Bannister & Kouzarides 2011).

HATs transfer an acetyl group (utilizing acetyl CoA as a co-factor) to the ε-amino group of lysine chains creating ε-N-acetyllysine (Fuchs et al. 2009). Addition of the acetyl group brings about neutralization of lysine's positive charge potentially weakening interactions between the nucleosome and negatively charged DNA and thus influencing and opening up the structure of chromatin fibres by preventing fibres interacting thereby allowing chromatin remodelling factors greater access to DNA (Li & Shogren-Knaak 2008). As acetylation is the most prevalent type of modification, with a large number of potential site including H3K9, H3K14, H3K18, H4K5, H4K8 and H4K12 (Kouzarides 2007), it has the potential to exert large effects on chromatin structure by reducing or neutralizing the charge on histone tails. Structural perturbation through

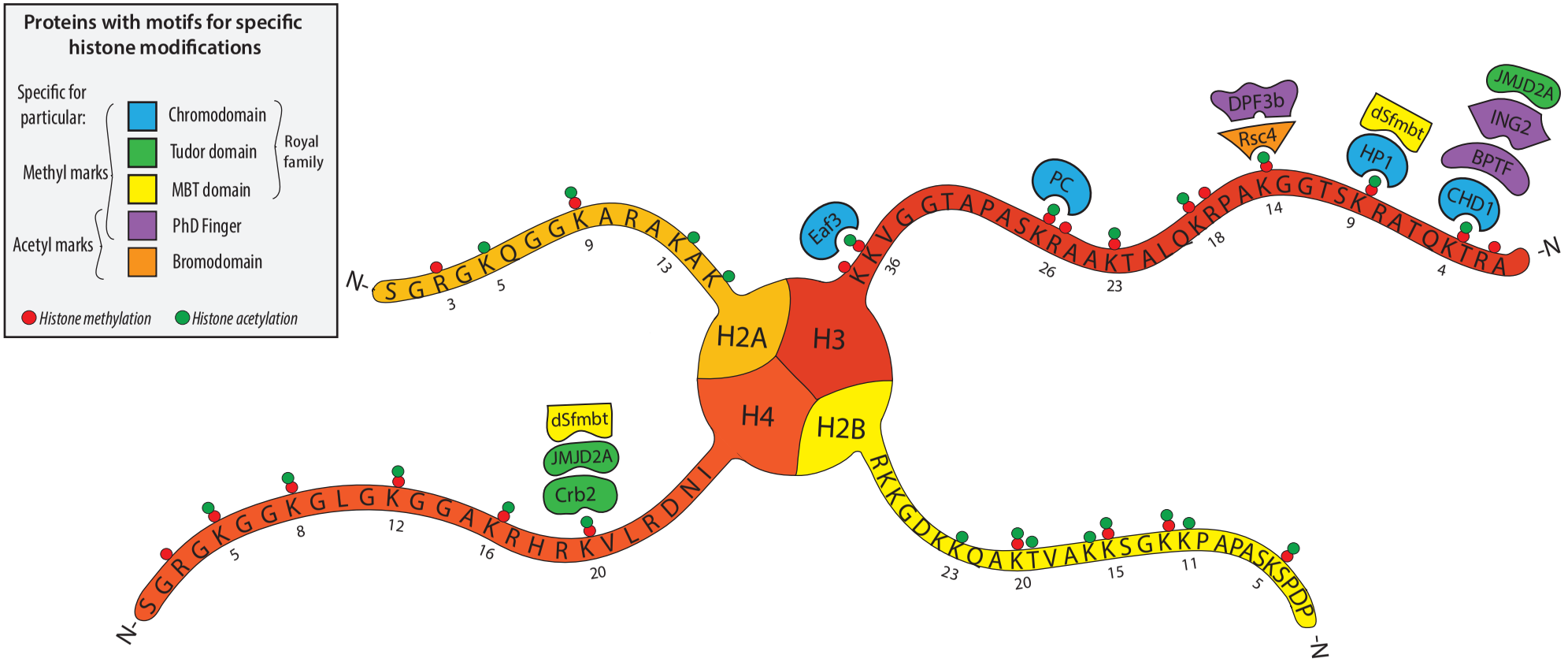
change in charge was initially believed to be the sole mechanism by which acetylation brought about gene activation, by making the DNA more accessible to transcription machinery (Cosgrove 2007). This mechanism is still believed to have an essential role in opening up chromatin structure, however, it is not thought to be the only effect brought about by histone acetylation. Many chromatin remodelling factors have been found to possess bromodomains which recognise and bind to acetylation marks. Binding selectively to particular histone modifications results in a more targeted recruitment of chromatin modifying proteins that can link to functional outcomes (Dhalluin et al. 1999) (Figure 1.7).

The deposition of acetylation is far from random, high levels are generally observed at the promoters regions and 5' ends of actively transcribed genes (Kouzarides 2007; Bannister & Kouzarides 2011). Hyperacetylation is often observed at the promoter region of transcriptionally active genes, an example of which is the GLUT4 promoter which becomes hyperacetylated upon exercise; leading to increased GLUT4 transcription (Richter & Hargreaves 2013). On the H4 tail the preferential acetylation of specific lysine residues is observed, as they are deposited in the following order H4K16 followed by H4K12, H4K8 and H4K5 (Turner et al. 1989; Turner & Fellows 1989).

HATs can be divided into two major classes: type A and type B. Only one type B HAT has been discovered, yeast HAT1/scHAT1, and its homologues in other species. These HATs are found within the cytoplasm and acetylate only free histones. Newly synthesised H4 histones are acetylated at lysine's 5 and 12, which is important for their deposition, after which the marks are removed (Parthun 2007).

Many more type A HATs have been identified and are divided into a further three subgroups GNATs, MYSTs and CBP/p300 HATs as represented in table 1.3

Figure 1.7 Locations of key histone methylations and acetylations and examples of proteins with binding specific to modified histones



The locations of key methylations and acetylations of the N terminal tails of the four core histones, H2A, H2B, H3 and H4. The majority of well studies marks are located on the flexible N-terminal tail regions. Many chromatin binding proteins contain domains for particular modified histone examples of which are included here. The domains include chromodomains, tudor domains and MBT domains that belong to the Royal family and recognise particular methyl marks. PhD domains which have been found in proteins specific to methyl and acetyl marks and bromodomains associate with acetyl marks. Figure adapted from Bannister and Kouzarides 2011. Data from Tan *et al*. 2011 and Arnaudo *et al*. 2013.

Table 1.3: Summary of histone acetyltransferase enzymes and histone deacetylase enzymes

HAT type A		**Organism**	**Histone specificity**
GNAT superfamily			
	KAT2/GCN5	Human/yeast	H3 (K9, K14, K18)/ H2B
	KAT2B/PCAF	Human/Mouse	H3 (K9, K14, K18)/ H2B
	KAT9/ELP3	Human/yeast	H3
	KAT10/Hap2	Yeast	H3 (K14)/H4
MYST related HATs			
	KAT5/TIP60	Human/yeast	H2A/H4
	KAT6/Sas3	Yeast	H3 (K14, K23)
	KAT6A/MOZ	Human/yeast	H3 (K14)
	KAT6B/MORF	Humans	H3 (K14)
	KAT7/HBO1	Human/yeast	H4 (K5, K8, K12) >H3
	KAT8/HMOF	Human/yeast	H4 (K16)
	Esa1	Yeast	H4/H3/H2A
	Sas2	Yeast	H4K16
	KAT4/TAF1	Human/yeast	H2AX (K5)/H4 (K16)
p300/CBP HATs			
	KAT3A/CBP	Various	H2A/H2B/H3/H4
	KAT3B/p300	Various	H2A/H2B/H3/H4
HAT type B			
	KAT1/HAT1	Human/yeast	H4 (K5, K12)
HDAC			
Class I			
	HDAC1	Mammals	H3 (K56)/ H4 (K16)
	HDAC3	Mammals	H3 (K9, K14)/ H4 (K5, K12)
	HDAC8	Mammals	
Class II			
	HDAC 4	Mammals	nd
	HDAC 5	Mammals	nd
	HDAC 6	Mammals	nd
	HDAC 7	Mammals	nd
	HDAC 9	Mammals	nd
	HDAC 10	Mammals	nd
Class III			
	SIRT1	Mammals	H1 (K26)/ H2AX (K5)/ H3 (K56)/ H4K16
	SIRT2	Mammals	H4 (K56) /H4 (K16)
	SIRT3	Mammals	nd
	SIRT5	Mammals	nd
	SIRT6	Mammals	H3 (K9)
Class IV			
	HDAC2	Mammals	H3 (K56)/ H4 (K16)

A selection of common histone acetyltransferases (HATs) and histone deacetylase enzymes (HDACs) divided by classification. Adapted from J. Heath 2010. Data from (Kouzarides 2007; Peserico & Simone 2011; Gong & Miller 2013).

(Hodawadekar & Marmorstein 2007). These enzymes are generally incorporated into large multi-protein complexes which participate in enzyme recruitment along with substrate specificity and enzyme activity (Yang & Seto 2007). An example of such is scGCNS which cannot acetylate histones within nucleosomes when free, however once incorporated into the SAGA complex it is then able to acetylate nucleosomal histones -a human counterpart of this complex has also been identified containing PCAF (the human scGCNS homologue) (Grant et al. 1997; Kornberg & Lorch 1999). As the removal of acetyl marks can cause gene silencing HDACs are often thought of as transcriptional repressors (Bannister & Kouzarides 2011). As HDACs are over expressed in a variety of cancers and are closely correlated to oncogenic factors, HDAC inhibitors such as trichostatin A and valporic acid are used in conjunction with chemotherapy (Kouraklis & Theocharis 2002).

There are four classes of HDACs as represented in table 1.3. Classes I and II are related to the yeast scRpd3 and scHdal. Class III (referred to as the sirtuins) are homologous to yeast scSir2 and requires the cofactor NAD+ for its activity. Finally class IV, of which the only one to be identified thus far is HDAC2.

HDACs are located within multiple distinct complexes and often with other HDACs. This often has functional consequences for the enzyme, for example HDAC1 and HDAC2 are found within NuRD, Sin3a and CoREST complexes, making it difficult to determine which element is responsible for a specific effect. These large complexes often contain DNA binding proteins facilitating interactions between the HDACs and nucleosomal histones. The highly conserved nature of histones, HATs and HDACs is highly indicative of a central role within the cell, further supported by correlations to diverse biological processes including transcription and cell differentiation. Interestingly post

translational acetylation has been found to occur in proteins other than histones some of which are also regulated by HATs and HDACs suggesting that their role is far wider spread than previously thought (Olsen 2013). Many HATs and HDACs are not only able to acetylate multiple different lysine residues but act on proteins other than histones (Nagy & Tora 2007; Haberland et al. 2009). How this will alter the understanding of histone acetylation as epigenetic remains to be seen.

1.6.2 HISTONE METHYLATION

Histone methylations and their associated functions present a far more complex picture than histone acetylation, starting with a larger target group. Both lysine and arginine residues in all four core histones can be methylated, however neither of these results in a net change in charge. Each residue can be methylated in a variety of ways often associated with different functions depending on its specific location and the degree of methylation (Fuchs et al. 2009). A primary examples of this are the opposing actions of the tri-methylation of lysine 4 on histone 3 (H3K4me3) and of lysine 27 also on histone 3 (H3K27me3) which are associated with transcriptional activation and silencing respectively. Additionally the di and tri methylation of H3K9 appear to function independently of each other; they are deposited by different histone methyltransferase and display differing patterns of localisation (Bessler et al. 2010).

Lysine residues can be mono-, di- or tri-methylated at the ε-amino group whereas arginines can be mono- or di-methylated; di-methylation occurring in either a symmetrical or asymmetrical manner (Bannister & Kouzarides 2011). Histone methylation has been linked to a variety of roles in many DNA regulatory processes. As no change of charge is involved, and the specific role of the modification can vary simply based on the location of the residue, the mode of action is thought to be exclusively

through the recruitment of chromatin regulatory factors rather than directly influencing DNA-nucleosomal interactions. In this way (through the recruitment of other factors) histone methylations do still have a role in higher order chromatin structures. Many modifications are essential in the formation and maintenance of euchromatin and heterochromatin including H3K9me2/3, H3K27me3, H3K4me1, H2K20me1/me2 and H4K20me3 which are associated with transcriptional silencing or heterochromatin, and H3K4me2/3 and H3K36me3 which are associated with active chromatin (Noma K et al. 2001; Barski et al. 2007; Kouzarides 2007; Sims & Reinberg 2009).

Methyl groups are deposited by histone methyl transferases (HMTs) which can be further subdivided into protein arginine methyl transferases (PRMTs) and lysine methyl transferases (HKMTs) as in table 1.4. There are 11 members of PRMT in humans, two varieties; type I and type II. Both are able to mono and dimethylate arginine but dimethylation by a type I results in asymmetric distribution whereas dimethylation by a type II results in an exclusively symmetric distribution (Bannister & Kouzarides 2011). All PRMTs transfer a methyl group from S-adenosylmethionine to the ω-guanidino group of arginine. Of the two classes of histone mark, arginine methylation remains largely uncharacterised, most research has focused on lysine methylations.

HKMTs have been extensively investigated and are subsequently fairly well characterised, they are divided into two classes; those with a SET domain and those with a DOT domain. HKMTs catalyse the removal of a methyl group from S-adenosylmethionine onto the ε-amino group of a lysine, the first of which to be identified was Suv39H1 which methylates H3K9 (Rea et al. 2000). The majority of the identified HKMTs to date methylate lysines within the N-terminal tail of which most contain a SET domain. HKMTs are generally highly specific, targeting limited lysine

Table 1.4: Summary of common histone methyl transferase enzymes

HMT		Organism	Histone specificity
PRMT type I			
	PRMT 1	Human	H4 (R3)
	PRMT 4	Human	H3 (R2, R17, R26)/H4 (R3)
	PRMT 6	Human	H3
	PRMT 8	Human	H4
PRMT type II			
	PRMT 5	Human	H2A/H3 (R8) /H4 (R3)
	PRMT 7	Human	H2A/H4
	PRMT 9	Human	H2A/H4
SET domain HKMT			
	KMT1A+B/SUV39H1	Human/Mouse	H3 (K9me3)
	KMT1C/G9a	Human	H3 (K9me1, me2)
	KMT1E/ESET/SETDB1	Human	H3 (K9me3)
	KMT2A/MLL1/Trx	Human/Fly	H3 (K4me1, me2, me3)
	KMT2B/MLL2/Trx	Human/Fly	H3 (K4me1, me2, me3)
	KMT2C/MLL3/Trr	Human/Fly	H3 (K4me1, me2, me3)
	KMT2D/MLL4/Trr	Human/Fly	H3 (K4me1, me2, me3)
	KMT2E/MLL5	Human	H3 (K4me1, me2, me3)
	KMT2F/hSET1A	Human	H3 (K4me1, me2, me3)
	KMT2G/hSET1B	Human	H3 (K4me1, me2, me3)
	KMT3A/SET2	Human	H3 (K36)
	KMT/SET9	Yeast	H4 (K20)
	KMT6/Ezh2/E(Z)	Human/Fly	H3 (K27)
DOT domain KMT			
	KMT4/DOT1L/Dot1	Humans/Yeast	H3 (K79)

A representation of common histone methyltransferase enzymes (HMTs) as divided into protein arginine methyltransferases (PRMTs) and lysine methyltransferases (HKMTs). These are further separated by classification. Adapted from J. Heath 2010. Data from (Allis et al. 2007; Kouzarides 2007; Shilatifard 2008; Wolf 2009).

residues and consistently methylating them to the same extent (i.e. mono, di or tri methylated). In mammals a separate enzyme is responsible for each degree of methylation for example DIM5 can only trimethylate H3K9 and SET7/9 is only able to monomethylate H3K4 (Tamaru et al. 2003; Xiao et al. 2003). Further to this many HKMTs target narrowly defined genomic locations (Thomas et al. 2008; Smith & Denu 2009). Histone lysine methylation was initially considered to be a more stable covalent mark than acetylation as the global turnover is much lower. As a consequence the methylation of histones at lysines has been considered the PTM most likely to play a role in cellular memory (Ng et al. 2009). The identification of histone demethylases (HDMs) has since demonstrated this mark can also be dynamic. There are currently two known classes of HDMs; KDM1, which is limited to the removal of mono or di methyl marks, and HDMs containing a Jumonji-C domain which are able to demethylate mono, di, and tri methylated residues (Cloos et al. 2008).

Many domains have been identified within proteins or protein complexes which bind to methylated residues -more than any other type of mark. These include chromodomains; tudor domains, PWWP and MBT domains (royal-super family modules); as well as WD40 repeats and PhD finger domains (Taverna et al. 2006). Methylated histones provide a binding platform for chromatin remodelling factors and can also block DNA/histone binding molecules such as H3K4me3 which prevents the transcriptionally repressive complex NuRD from binding to the H3 tail. The binding of proteins to histone marks brings proteins into close proximity at the site of the mark allowing complex formation and hence the relevant action can be brought about (Nishioka et al. 2002; Zegerman et al. 2002) (Figure 1.7).

1.6.3 PHOSPHORYLATION OF HISTONE TAILS

Histone phosphorylation is a highly dynamic mark that can take place on serines, threonines or tyrosines and involves the addition of a phosphate group by protein kinases and removal by phosphatases. In the vast majority of cases serine/threonine protein kinases catalyse the placement of a phosphate group from an ATP molecule to the hydroxyl group of an N-terminal serine or threonine residue. How kinases are recruited is largely unknown, however some kinases such as mammalian MAPK1 have been found to incorporate a DNA binding domain potentially facilitating anchoring to DNA (Bannister & Kouzarides 2011).

Placement of a phosphate group results in a significant negative charge and in doing so alters histone-DNA and histone-histone binding ultimately influencing the chromatin structure. However in addition to affecting chromatin structure through charge, phosphorylation much like acetylation and methylation exerts a more targeted effect through the recruitment and/or inhibition of chromatin modifiers. Histone phosphorylation has been found at fewer locations than acetylation or methylation and appears to be carefully targeted.

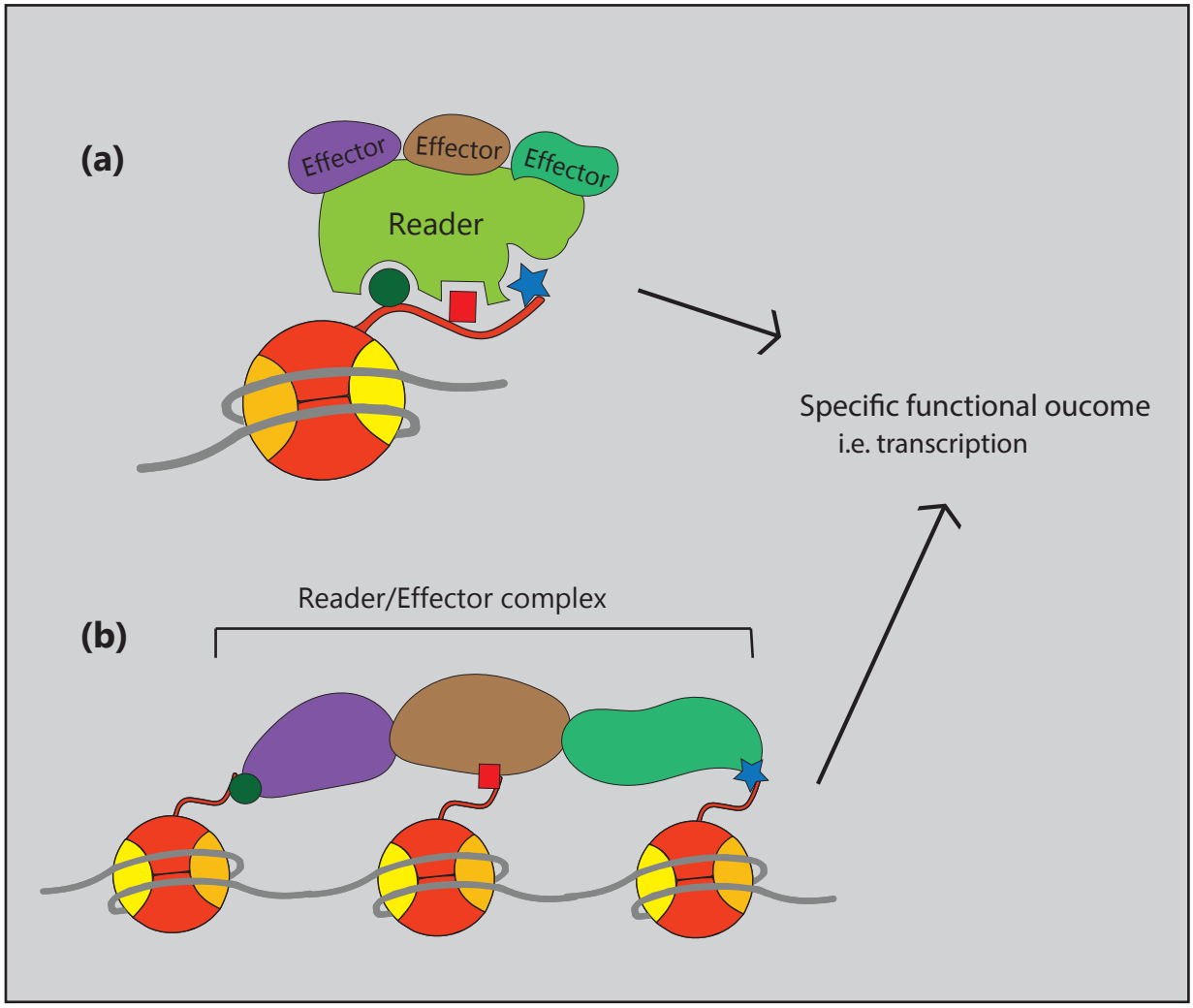
Phosphorylation at H3 Serine 10 is one of the most extensively studied histone modifications. Enrichment of this mark has been found to be cell cycle regulated as it is associated with the condensation of heterochromatin in mitosis (Gurley et al. 1974; Hendzel et al. 1997; Wei et al. 1998). Chromosome condensation is accompanied by phosphorylation of H3S10, which is deposited and removed genome wide by Aurora B kinase (carries out multiple roles in mitosis and chromosome segregation) and Protein phosphatase 1 (PP1) respectively (Hsu et al. 2000; Hirota et al. 2005). However its exact role remains to be determined. H3S10phos has been linked to the loss of

heterochromatin protein 1 (HP1) in mitotic heterochromatin (Fischle et al. 2005). HP1 binds to H3K9me2 or H3K9me3 and is required for the formation of normal heterochromatin, this interaction is weakened when a H3S10phos mark is present. In yeast the inhibition of Aurora B was found to prevent the dissociation of HP1 during mitosis (Fischle et al. 2005; Hirota et al. 2005). This change with the addition of H3S10phos is referred to as a "phospho/methyl switch". Other such "phospho/methyl switches" have subsequently been identified including the dissociation of other H3 binding proteins (SRp20 and ASF/SF2) from chromatin in mitosis upon addition of H3S10phos (Loomis et al. 2009; Wang & Higgins 2013).

1.7 THE EPIGENETIC CODE

Existence of an epigenetic code was first proposed in 1993 by Bryan Turner (Turner 1993) with particular reference to histone acetylation and proteins which recognise acetylation (Turner 1993; Turner 1998; Turner 2000). This idea of a code was developed into the Histone code hypothesis as proposed by Strahl and Allis (2000) and has garnered much support over the years (Strahl & Allis 2000; Jenuwein & Allis 2001; Nightingale et al. 2006; Turner 2007; Lee et al. 2010). The histone code proposes that post translational modifications act sequentially or in combination to facilitate other proteins and complexes ultimately changing or regulating gene expression. Non-histone proteins "readers" bind the histone modifications, which subsequently bind "effectors" that are able to bring about a localised functional effect such as transcription (figure 1.8, Turner 2000; Strahl & Allis 2000; Nightingale et al. 2006). Many proteins have been identified which specifically bind to particular histone modifications (e.g. bromodomains specifically bind acetylated lysines and PHD, Tudor and MBT domains

Figure 1. 8 The histone code



The histone code proposes that a particular set of histone modifications will result in a specific functional outcome and that the outcome will vary depending on the combination histone modifications. The code could be (a) confined to a single nucleosome. A single "reader" protein would recognise particular histone combinations and attract "effector" proteins and produce a functional outcome. Or alternatively (b) proteins with binding sites for particular histone modifications could unite to form a "reader"/"effector" complex bringing about a particular functional outcome. Figure adapted from Nightingale *et al*. 2006.

selectively bind to particular methylated lysines)(Dhalluin et al. 1999; Taverna et al. 2007; Musselman & Kutateladze 2011). Additional modifications on adjacent residues can affect the binding of "readers" and "effectors", either positively or negatively, and hence suppress or change the role of the initial modification. This is referred to as histone crosstalk. Given the number of possible modifications crosstalk between them would vastly increase the complexity of histone signalling. Histone modifications are also thought to work competitively as well as in combination, as in the case of the H4K16ac mark whose presence impedes *in vitro* ribosylation (Messner et al. 2010).

Frequent use of the term epigenetic code is also made in reference to the cross talk between epigenetic pathways; associations have been identified between ncRNAs, DNA methylation and histone modifications. This is clearest in X-chromosome inactivation where all three elements work in concert to repress chromatin but is being increasingly observed in areas where histone modifications are important.

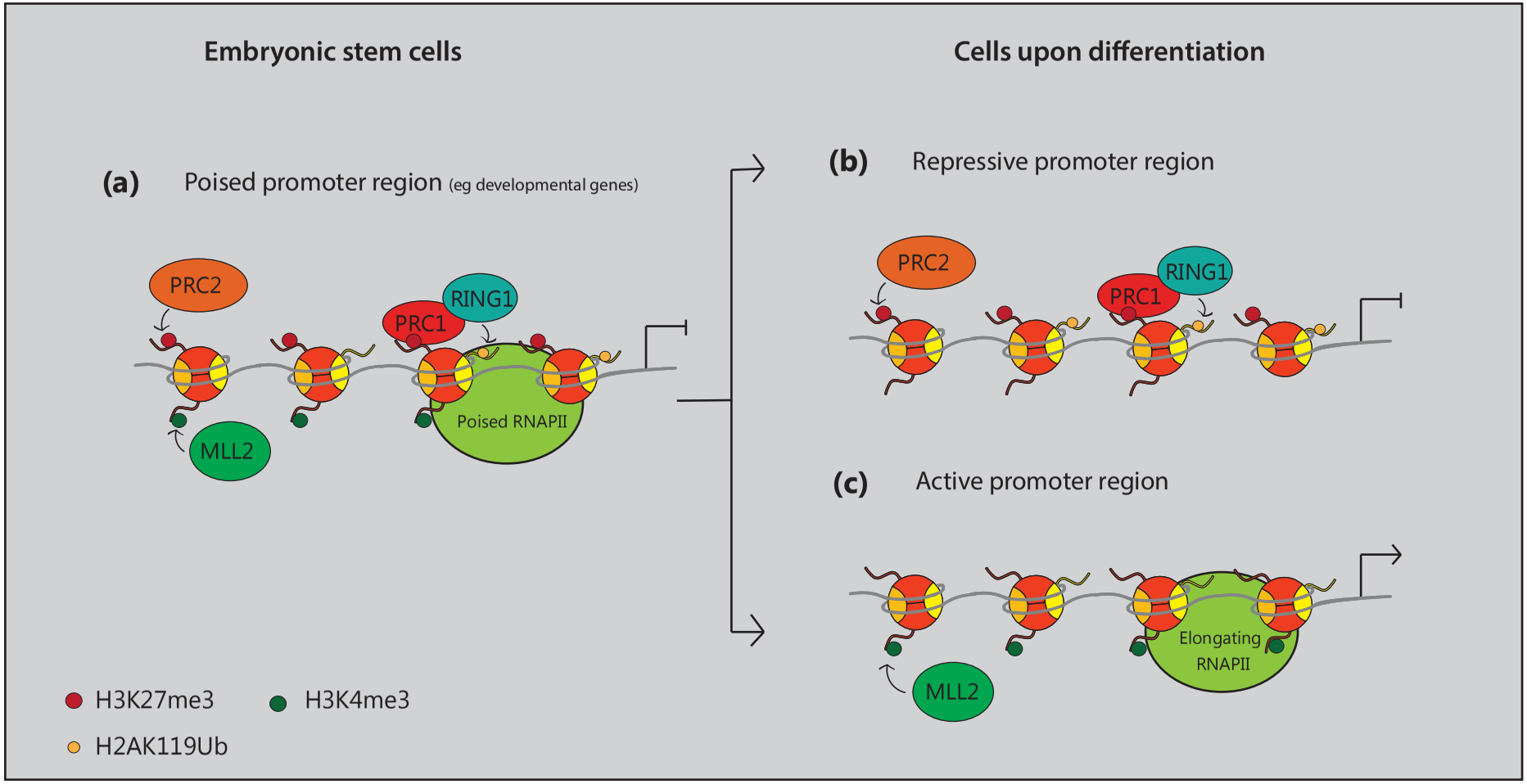
The epigenetic code has evolved into a much broader concept that covers the detection of patterns in histone modifications and other epigenetic factors, not only in the activation or repression of transcription but in a range of DNA regulatory processes. Different histone modifications are seen to act in concert to shield DNA or recruit key effector complexes (Nightingale et al. 2006; Kouzarides 2007) but can this truly be defined as a code? Turner (2007) defines a code in terms of a semiotic system comprised of a sign, its meaning, and a code through which the meaning of the sign is made known. This strict definition requires that without the code there be no direct causal relationship between the sign and its outcome (Turner 2007). Although it remains a viable theory for the role and function of histone modifications it is unlikely that such a rigid definition will apply to all the roles of histone modifications. The terms

histone code and epigenetic code are likely to be too strong a term to describe interacting histone modifications, Henikoff and Shilatiford propose the term "nucleosome signalling" to better describe the phenomenon (Henikoff & Shilatifard 2011). However this does not include any combinationatorial effect observed with other epigenetic factors such as DNA methylation and ncRNAs (non coding RNAs).

1.8 BIVALENT MODIFICATIONS IN EMBRYONIC STEM CELLS

The best understood instance of a combinatorial effect of histone modifications within ES cells is that of bivalency. The term bivalent was first coined by Bernstein et al 2006 to describe domains that contain both active and repressive modifications generally found at the promoter region of genes, the most abundant of which is the dual presence of H3K4me3 (active mark) and H3K27me3 (repressive mark) (Bernstein et al. 2006). The active mark is thought to maintain chromatin in an open configuration ready for transcription whereas the repressive mark prevents the majority of this transcription from taking place. The overall effect of which is to keep genes in a poised or primed state, allowing their rapid activation upon differentiation cues or stimuli. An overlap is also found between bivalency in ES cells and pluripotency factors including Oct4, as such factors often bind promoters with bivalent marks. This is thought to further support the maintenance of genes in an open configuration (Voigt et al. 2013a). In ES cells the vast majority of bivalent modifications are found at the promoters of transcription factors important for development, awaiting differentiation signals (Voigt et al. 2013a). The chromatin state of ES cells is very open, however upon differentiation the nucleus shrinks and a higher proportion of DNA becomes packaged as heterochromatin (Golob et al. 2008). As cells differentiate bivalent domains are believed

Figure 1.9 The bivalent profile of H3K4me3 and H3K27me3 in embryonic stem cells



Many gene promoters in ES cells bear a combinination of conflicting histone modifications known as bivalent marks. The most abundant and best studied case of bivalency is the combination of H3K4me3 (active) and H3K27me3 (repressive). (a) MLL2 and PRC2 deposit H3K4me3 and H3K27me3 respectively. PRC1 then binds to the H3K27me3 mark facilitating the ubiquitination of H2A by RING and therefore restraining the poised RNA polymerase II. Upon differentiation the majority of the poised genes are either (b) lose the active modification becoming repressed, or (c) lose the repressive mark and are consequently actively expressed. Figure adapted from Kraushaar *et al*. 2013.

to lose one of the marks either switching on or off and hence commit one way or the other (figure 1.9).

Although bivalent domains are most commonly associated with ES cells they have also been identified in other types of cell. Essentially all promoters in mESCs are found to possess the H3K4me3 mark, of these approximately 22% also have H3K27me3 and were minimally expressed (Mikkelsen et al. 2007). Upon differentiation bivalency is only found at 8% in neural progenitor cells and 4% in MEFs, and it is likely that some of the bivalent domains in these cells will be located at different points to those found in ES cells (Voigt et al. 2013a).

H3K4me3 and H3K27me3 are regulated by the trithorax group (TrxG) proteins and the polycomb group (PcG) proteins respectively. These two established chromatin modifying systems were first identified in drosophila as activators and repressors of the Hox genes (which control the sequential development of an animal's basic body plan) and work antagonistically (Schuettengruber et al. 2007; Simon & Kingston 2009). In mammals the H3K4me3 mark is deposited by one of 6 non-redundant HKMTs; SET1A, SET1B and MLL proteins 1-4 which all function in conjunction with other proteins, sometimes over lapping, as part of distinct complexes (Shilatifard 2012). PcG proteins in vertebrates and flies are incorporated into the polycomb repressive complexes (PRCs) 1 and 2, of these it is PRC2 deposits H3K27me3 and is associated with repressive chromatin establishment (Margueron & Reinberg 2011). Approximately 10-15% of all H3 histones in ES cells are marked with H3K27me3 when assessed by quantitative mass spectrometry (Peters et al. 2003; Voigt et al. 2012). This figure increases to around 50% upon the inclusion of H3K27me2 (Voigt et al. 2012). H3K27me3 is distributed in broad flat regions, rather than the defined peaks observed at promoters with H3K4me3, and

as such is much more prevalent, found over intergenic regions and inactive genes (Pauler et al. 2009; Young et al. 2011). The presence of H3K27me3 around the transcription start site (TSS) is believed to be almost entirely bivalent (Mikkelsen et al. 2007; Ku et al. 2008; Voigt et al. 2013b). Nucleosomes marked with only one copy of H3H4me3 allow PRC2 to deposit H3K27me3, whereas two copies of the active mark prevent this (Voigt et al. 2012). This along with other evidence strongly point to an asymmetric distribution of bivalent marks on different H3s of the same nucleosome (Voigt et al. 2013a). Recently MLL2 was identified as the Trx group protein responsible for H3K4 trimethylation at bivalent promoters in mouse ES cells (Denissov et al. 2014).

Bivalent domains strongly correlate with CpG Islands in ES cells (Bernstein et al. 2006). Around 70% of all promoters in ES cells are CpG rich and lack DNA methylation. Most H3K4 HKMTs (SET1a/B, MLL1, MLL2) contain domains that specifically bind to unmethylated DNA (Lee et al. 2001; Birke et al. 2002; Bach et al. 2009). All H3K4me3 modifications are located at CpG islands and consequently all bivalent marks that include this histone mark. Further to this both H3K4me3 and H3K27me3 are deposited at sites of artificially introduced CpG islands (Mendenhall et al. 2010; Thomson et al. 2010; Lynch et al. 2012). H3K4me3 prevents de novo DNA methylation, conversely promoters that loose both marks have a high probability of becoming (DNA) hypermethylated. It has still not been definitively determined whether marks are a consequence of the transcriptional status of genes or instructive.

Although the most prevent, H3K4me3 and H3K27me3 are not the only marks believed to be bivalent. Azuara *et. al.* (2006) determined that early replicating but inactive genes in ES cells have H3K4me3 and H3K9ac alongside the H3K27me3 repressive mark (Azuara et al. 2006). Those genes not expressed upon differentiation are subsequently

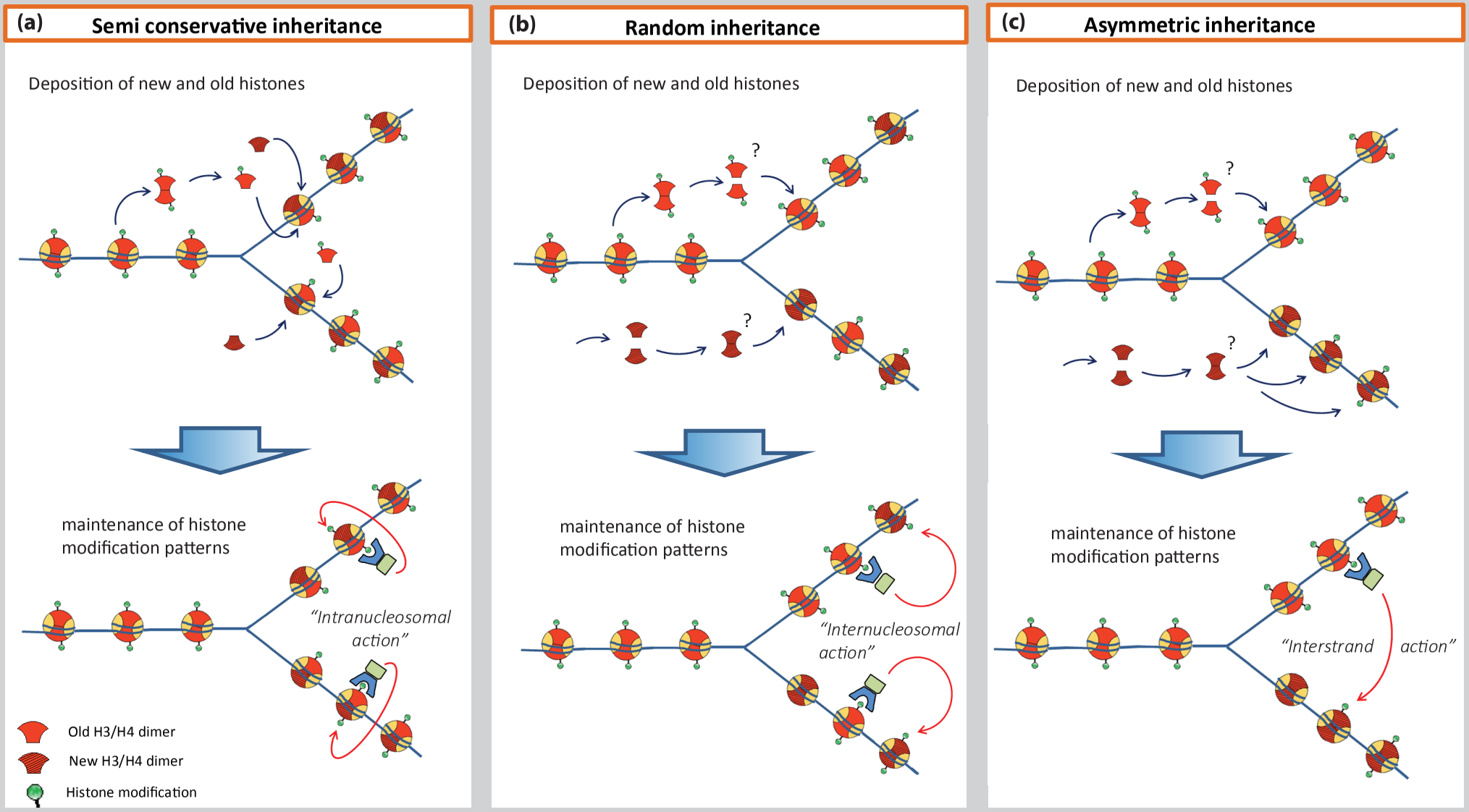
replicated later during S phase in the differentiated cells. H3K9me3 has been observed alongside H3K4me3 and H3K27me3 *in vivo* (Voigt et al. 2013a).

1.9 POSSIBLE MECHANISMS FOR THE INHERITANCE OF HISTONE MODIFICATIONS

In order to fulfil the requirements of the term "epigenetic"- in addition to having a functional effect on gene regulation- histone modifications should be heritable. The classic example of an epigenetic mark is DNA methylation specifically at CpG islands, which is mostly maintained in a semi-conservative manner by the actions of DNA methyltransferase 1 (Jones & Liang 2009). DNA methylation is found symmetrically on both parent strands, each of the daughter strands therefore contains one methylated strand which functions as a template allowing the methylation pattern to be faithfully reproduced onto the new strands. As the level of interest into histone modifications continues to grow it is not surprising that attention has turned back to their epigenetic status and this question of heritability which, although generally assumed, has yet to be demonstrated.

In order to be heritable histone modifications must persist through cell division. Most groups currently investigating the heritability of histone modifications are doing so by homing in on S-phase and DNA replication. During DNA replication chromatin is disassembled prior to the replication fork and reassembled on the two daughter strands. In order for information to be faithfully passed from one generation to the next histone patterns must be faithfully restored on both strands. How this occurs and even whether it does is not yet known and the subject of some debate. Three models have been proposed for the transmission of histone modifications from the parent strand to

Figure 1.10 Models for the inheritence of histone modifications



Three possible models for the inheritence of histone modifications. (a) Semi conservative inheritence - following the replication fork the two old H3/H4 histone dimers are divided equally into new nucleosomes on both daughter strands. (b) Random inheritence - incorportation of old and new histones is entirely random. (c) Asymmetric inheritence - old H3/H4 histone dimers are preferentially incorporated into nucleosomes on one of the daughter strands, nucleosomes onthe other strand are composed entirely of new histones.

the two daughter strands of DNA during replication; 1. Semi-conservative; 2. Random; 3. Asymmetric histone inheritance. (figure 1.10)(Probst et al. 2009).

1.9.1 SEMI-CONSERVATIVE INHERITANCE

This is where information is copied from a template in the same manner as CpG methylation or DNA replication, it is an ideally simple model. This model was alluded to as early as 1976, at the time the very structure of the nucleosome was being determined, by Weintraub et al. (1976) who referred to "symmetrically paired half nucleosomes" and the possibility of inheritance of information through histone structure (Weintraub et al. 1976). The model requires histone modifications to exist in a symmetrical manner within individual nucleosomes and also for nucleosomes to separate into two identical halves. In actual fact histones H3 and H4 are believed to be the most stable element of the nucleosome and also where the majority of histone modifications are found. The tetramer formed from two H3-H4 dimers is often referred to as the nucleosome core. It is this core H3-H4 tetramer that would be required to separate alongside parent DNA in a semi conservative manner with one parent H3-H4 dimer being placed at each daughter strand and in the correct position.

It remains extremely difficult to discriminate between symmetrically and asymmetrically modified nucleosomes, be that as it may the existence of multiple marks on the same histone tail in the same chromatin region would suggest that histone modifications are not always symmetrical.

The idea of semi-conservative division of H3-H4 tetramers did not initially seem to be supported by experimental data. Once displaced from DNA H3-H4 dimers were found to be mainly maintained as tetramers (Leffak et al. 1977; Prior et al. 1980; Yamasu & Senshu 1990). However, semi-conservative inheritance was revived as a theory when it

was discovered that histones H3 and H4 are actually deposited onto chromatin as dimers (Tagami et al. 2004; English et al. 2005).

Histone chaperones at the replication fork are responsible for the disruption and consequent redeposition of histones. Up to five distinct histone chaperones have been found in eukaryotes (Alabert & Groth 2012). H3-H4 dimers/tetramers are deposited first, then sandwiched between two H2A-H2B dimers (Polo & Almouzni 2006). Histone chaperones such as FACT and Asf1 are thought to bind the H3-H4 before passing onto the PCNA-associated CAF1 (Chromatin assembly factor 1) for nucleosome reassembly and deposition (Stillman 1986; Smith & Stillman 1989). Although several key chaperones are found to bind only the H3-H4 tetramer; which again makes semi-conservative inheritance seem unlikely; chaperones have been identified which bind to H3-H4 dimers (Bowman et al. 2011; Almouzni et al. 2013). Anti-silencing function 1 (Asf-1), the most conserved histone chaperone in eukaryotes, not only blocks tetramer formation but brakes down tetramers to form H3-H4/Asf-1 heterotrimers (Natsume et al. 2007; Zhu & Reinberg 2011). Asf1 is believed to be a major storage chaperone and has been associated with DNA replication, repair and transcription in yeast (Winkler et al. 2012). PCNA associated Caf1 interacts with both dimers and tetramers and when interacting with dimers it acts to bring them into the best conformation for tetramer formation on the daughter DNA (figure 1.10)(Winkler et al. 2012).

The semi-conservative inheritance theory has again been shown to be unlikely as, regardless of the exact, as yet unknown, histone chaperone pathways the majority of nucleosome H3-H4 tetramer cores are maintained separate from newly formed H3-H4 dimers (Xu et al. 2010; Katan-Khaykovich & Struhl 2011). Some tetramers do split and reform as mixtures of old and new H3-H4 dimers; this has been linked to particular

histone variants. H3.3-H4 tetramers have been found to split in-vivo whereas H3.1-H4 (H3.1 being the most abundant variant) mostly remain intact (Xu et al. 2010). The association of H3.3 with actively transcribed regions, in addition to the low numbers of these events, makes them unlikely to be representative of histone inheritance in general. As experimental evidence suggests that H3-H4 histone cores are mainly conservatively inherited it does not appear that the semi-conservative model of histone modification inheritance is likely for most histone modifications, however it cannot be ruled out as a possibility for all modifications in all scenarios or cell types.

1.9.2 RANDOM INHERITANCE

Random inheritance of histones necessitates that distribution of nucleosomes onto the daughter DNA strands be random and that restoration of the pattern of histone modification still be faithfully maintained. There is much evidence in support of this model of inheritance. The deposition of parent nucleosome cores verses newly synthesised H3-H4 tetramers onto the daughter strands does appear to be random (Annunziato 2005). Studies in *Saccharomyces Cerevisiae* support the idea of heritability as they strongly suggest that the location of histone redisposition on the daughter strand should closely correspond to their location on the parent strand. Parental histones in daughter strands are mostly found within 400 bps of their pre-replication locus on the parent strand (Radman-Livaja et al. 2011), in addition to which synthesis and length of the lagging strand has been linked to immediate histone redisposition (Smith & Whitehouse 2012).

Mechanistically the faithful reproduction of histone patterns via this method would be the most complicated of the proposed models. *Trans*-acting histone modification complexes would be required to bind to the newly deposited parental modifications and

subsequently alter the new histones with the same modification patterns both from the leading strand to the lagging strand and also in the other direction possibly by bringing strands together via a folding mechanism (Almouzni et al. 2013). Evidence in support of this can be found for H3K27 and H3K9 methylation which are both catalysed by methyltransferases within complexes that also bind to the methylated product, otherwise known as reader writer complexes (Lachner et al. 2001; Hansen et al. 2008; Hansen & Helin 2009; Margueron et al. 2009). PRC2-Ezh2 binds to H3K27me3 via the PRC2 subunit Eed and catalyses H3K27 methylation through Ezh2 (Cao et al. 2002; Czermin et al. 2002; Müller et al. 2002; Kuzmichev et al. 2002). Similarly H3K9 methylation inheritance could be mediated via a HP1/Swi6 (which recognises H3K9me3) and Suvar/Clr4 (H3K9 methyltransferase) complex (Rea et al. 2000; Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001). The ATP dependent chromatin remodeller Smarcad1 has also been found to interact directly with Pcna and also histone deacetylases HDAC1/2 and the methyltransferase G9a/GLP, which would not only facilitate histone deacetylation and H3K9methylation but it would target these at the DNA replication fork (Mermoud et al. 2011; Rowbotham et al. 2011). A similar complex has been observed for H4K16 deacetylation, Sir3 recognises the modification and is coupled with Sir2 a H4K16 deacetylase (Hecht et al. 1995; Imai et al. 2000). Such a deacetylation complex is unlikely to be involved in replication coupled inheritance as de novo H3-H4 dimers do not have the H4K16ac, it is an active mark believed to have a role in transcriptional regulation in mouse ES cells. Existence of replication independent complexes throws further questions on complexes that include both chromatin modifying enzymes and also domains to bind and recognise existing modifications. They could be capable of bring about inheritance if located around the replication fork, however it has not been demonstrated that these complexes function *in trans*, in fact it

has been proposed by many that *in cis* is the more likely mode of action, forming chromatin domains rather than faithful replication.

1.9.3 ASYMMETRICAL INHERITANCE

Asymmetrical inheritance (as seen in figure 1.10) has been suggested as an alternative to the random model above (Probst et al. 2009) and is another form in which templated copying could be possible. Asymmetrical DNA synthesis could result in strand bias for parent histone deposition. In order for the strand containing all the original histone modifications to function as a template, histone chaperones and modifying complexes would be required to act in *trans*, but only in one direction, a simpler option than proposed for the random model.

Asymmetrical inheritance is not likely to be responsible for histone inheritance in most cells, however, it remains possible that a variety of mechanisms are involved in establishing the complex chromatin patterns seen, with a greater or lesser degree of fidelity and flexibility. The idea of asymmetrical inheritance is particularly appealing for stem cells and development as after division cells can differ in terms of differentiation and cell fate. Adult stem cells in particular divide asymmetrically resulting in one multipotent stem cell and one specialised cell that will go on to terminally differentiate.

1.10 HISTONE MODIFICATIONS AND THEIR ROLE IN EPIGENETICS

It may be that not all of the many known histone modifications are “epigenetically” heritable; it is even possible that none are. Histones may be heritable in a much looser sense; existing histone modifications may serve as a guide to chromatin modifiers promoting certain modifications in particular areas or even chromatin domains. The association of histone chaperones and chromatin modifying enzymes to the replication

fork suggests a degree of heritability and in addition highlights the importance histone marks.

An alternative to the random inheritance model described would be where histone modification complexes acted not *in trans* but *in cis* spreading along the same strand creating general areas of histone modifications rather than specific locations. This theory relies on the ability of complexes to act *in cis* to nucleosomes nearby, a full mechanism for such would still need to be worked out. This may be the mode of action for the methylation of H3K9 and H3K27 in heterochromatin which are found to be distributed in broader peaks (table 1.2, Huang et al. 2013). The complexes described for the random inheritance model could as easily spread along the chromatin to maintain a regional heterochromatin environment.

Euchromatin and heterochromatin are known to be located within different domains in the nucleus, these domains also fall into the category of early or late replicating domains, with euchromatin being replicated earlier in S phase. Genes that replicate early in S phase have high levels of H3K9ac whereas late replicating genes showed low levels of H3K9ac. Recent experiments which consist of microinjecting DNA into the nucleus during S-phase revealed that deacetylated DNA injected during early S-phase was repackaged as acetylated chromatin and expressed at higher levels (Lande-Diner et al. 2009). In *Drosophila* pericentromeric heterochromatin domains containing H3K9me3 are sensitive to the levels of heterochromatin promoting or opposing factors (Almouzni et al. 2013). Also found in *Drosophila*, large domains of H3K27me3 are found in the bithorax complex region, which is an area of comparatively low histone turnover (Almouzni et al. 2013). Such domains are stable and carried through replication

however it may be difficult to allow rapid changes in gene expression and are too large to accommodate the finer aspects of gene regulation.

These experiments suggest that domains are able to establish different chromatin types however it does not reveal what is happening at the genetic level. Many marks are distributed in much sharper peaks, such as H3K4me3 mostly found at transcriptional start sites, not spread across a larger region (table 1.2).

Different mechanisms may be required for more dynamic changes at actively transcribed chromatin particularly at the promoter regions of active genes. If histone modifications are to play a role in transcriptional regulation they must be changeable, dynamic and correspond to the current cell environment. Histone turnover is fastest in active genes, epigenetic regulatory elements and replication origins (Deal et al. 2010) which makes sense as these are the genes mostly likely to change rapidly. This also highlights the importance of bivalency in ES cells as cells are required to alter gene expression patterns very rapidly upon external stimuli and suggests a functional role for histone modifications in actively transcribed chromatin at least.

Another theory is that the levels of modifications may be simply the result of opposing influences such as histone methylation or acetylation and cell cycle turnover in a replication independent manner (Hathaway et al. 2012; Hodges & Crabtree 2012). Stable isotope labelling followed by quantitative mass spectrometry has enabled the tracking and comparison of modifications on old and new histones. Using this method H4K20 was found to be increasingly methylated after replication from G2-G1 phase of the cell cycle (Pesavento et al. 2008). H4K20 methylation levels are at their highest in M-phase, and lowest at S-phase (which shows negative correlation to H4K16ac) (Rice et al. 2002). The most common H4K20 mark in mammals, H4K20me2, is not observed

after replication in S-phase, the mark is placed by the Suv4-20 enzymes which convert H4K20me1 into H4K20me2. Interestingly deposition of H4K20me1 is only carried out by the cell cycle regulated PR-Set7 enzyme which is present only between G2-phase and early G1 (Scharf et al. 2009; Tardat et al. 2010; Zhu & Reinberg 2011). This suggests that the reestablishment of histone modifications may not be heritable as such but much more flexible, it does however introduce the question of how cellular identity is maintained in long lived quiescent cells that do not divide.

A controversial study by Petruk et. al. (2012) also found that in *Drosophila* embryos, although trithorax protein (Trx) and Polycomb group protein E(z) remain associated with DNA throughout replication, H3K4 and H3K27 methylation were not found at the replication fork or on replicated DNA (Petruk et al. 2012). This study was performed on global chromatin, as with the majority of studies into histone domains. It is very likely that histone inheritance at the gene level follows a different pathway. Data from within our group suggests that actively transcribed genes are protected from global influences on acetylation (Halsall et al. 2012).

Progress is being made regarding the roles of histone modifications and all their associated proteins, as we probe further it appears unlikely that we can fit histone heritability into one simple model. Genetic information by its nature must be very stable- in contrast epigenetic information requires a high degree of plasticity, particularly in the case of histone modifications which must be inherently reversible. Investigations into histone inheritance allude to a much more complicated scenario than exists for DNA replication. While this is to be expected in a system that requires a high degree of plasticity, it does mean that a full understanding of both the role and inheritance of histone modifications will need a great deal more investigation. Even in

the case of CpG methylation, which was believed to be inherited in a very specific manner. Recent investigations have found that unmethylated sites in methylated CpG islands are likely to undergo de novo methylation alluding to a more general pattern to inheritance than previously thought (Jones & Liang 2009).

The question is really do histone modifications contribute solely to short term processes such as DNA transcription, replication and repair, or do they also have a role in long term heritable changes such as the maintenance of cell morphology and cell identity. Histone lysine methylations have a much longer half-life than acetylation or phosphorylation (Pesavento et al. 2008; Barth & Imhof 2010; Zee et al. 2010) and therefore are the more likely group of modifications to be "epigenetically heritable". By examining the enrichment of a panel of post translational histone modifications across the cell cycle it should be possible to gain insight into the heritability of the modifications and also aid in elucidating their mode of action. Further investigations will hopefully help to distinguish between short term instructions and cellular memory. Whether or not histone modifications are found to be classically heritable (i.e. consistent across the cell cycle from one generation to the next) they are essential for all biological processes that require access to DNA and therefore an improved understanding of their mechanisms and regulation will further our knowledge of how we work at a fundamental level.

1.11 AIMS:

The general aim of this thesis is to improve understanding of the nature of histone modifications with particular focus on their heritability.

- To isolate the separate phases of the cell cycle using the mouse embryonic stem cell line OS25s, either by chemically arresting cells at particular points in the cell cycle or by staining cells with propidium iodide and FACs sorting them into the individual phases.
- To determine the relative levels of a panel of post translational histone modifications in these cell cycle fractions via chromatin immunoprecipitation at the promoter of cell cycle regulated genes in order to ascertain the heritability of modifications. Comparing the results for both methods of cell cycle break down.
- To further investigate the genetic landscape of a subset of the cell cycle regulated genes. Looking at histone modification enrichment not only at the promoter region of genes but upstream and downstream of the transcriptional start site.
- To differentiate the mouse embryonic stem cell line OS25s to a point where a canonical cell cycle structure is established in cells as determined by their cell cycle profile upon propidium iodide staining and flow cytometric analysis.
- To isolate separate phases of the cell cycle in the same manner as with the undifferentitated embryonic stem cells and again determine the relative levels of enrichment of histone modifications across the cell cycle in cell cycle regulated genes. Comparing the results from the undifferentiated ES cells to the differentiated ES cells, highlighting the importance of histone modifications in these two cell types.

2 MATERIALS AND METHODS

2.1 CELL CULTURE AND CELL CYCLE ARREST

2.1.1 MUS MUSCULUS OS25 EMBRYONIC STEM CELL CULTURE

OS25s are a murine embryonic stem cell line derived from E14TG2a by sequential gene targeting (Billon et al. 2002). OS25 cells were grown as a monolayer in vented flasks (Sarstedt) pre-gelatinised (0.1% gelatin – Sigma-Aldrich) in 5% CO_2, at 37^{o}c in a humidified atmosphere. Cells were maintained in M3^{+} medium. M3^{+} medium consists of Dulbecco's Modified Eagle Medium (DMEM-Gibco) supplemented with 10^4units/10ml Leukemia Inhibitory Factor (LIF-millipore), 20% heat inactivated fetal calf serum (FCS-Gibco), 10% distilled sterile water, 1% 200mM L-glutamine (Invitrogen), 1% penicillin and streptocillin (Invitrogen), 1% non-essential amino acids (Sigma-Aldrich) and 0.25% 2-mercaptoethanol (Gibco). Cells were supplemented with 50% new medium daily and passaged when 75-80% confluent, detached with 0.05% trypsin-EDTA (Invitrogen) at 37^{o}c , centrifuged at 200 x g for 5min and transferred to new gelatinised vented flasks (Sarstedt) in fresh M3^{+} medium. When harvested OS25 cells were trypsinised at 37^{o}c, centrifuged at 200 x g for 5min, washed twice in ice cold Phosphate buffered saline solution (PBS – 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4 with 5mM sodium butyrate, pH 7.4 (all Sigma-Aldrich)) before resuspending in PBS.

2.1.2 DROSOPHILA MELANOGASTER SL2 CELL CULTURE

SL2 cells are drosophila melanogaster cells derived from late stage embryos. SL2s were grown anaerobically in suspension at 26^{o}c. Cells were maintained in Insect-xpress with L-Glutamine (Lonza) supplemented with 5% FBS (Invitrogen) and 1%

penicillin/streptomycin (Invitrogen). Medium was supplemented every 2-3 days, cultures were passaged once a week.

2.1.3 ACETONE FIXING CELLS

In order to fix cells before FACS or chromatin analysis, OS25 cells were harvested and washed 3 times in PBS/5mM NaButyrate and resuspended in 250µl PBS/5mM NaButyrate per 5 x 10^6 cells. Three times the volume of ice-cold acetone (Fisher) were added drop wise whilst constantly mixing gently. Cells were then left at -20^{o}c for at least 30 minutes, before centrifugation at 35 x g, 4^{o}c, for 10minutes, washed once more in PBS/5mM NaButyrate and centrifuged again at 130 x g at 4^{o}c for 10 minutes.

2.2 ISOLATION OF CELL CYCLE PHASES

2.2.1 ARRESTING OS25S AT CELL CYCLE PHASES

In order to arrest the cell cycle at the G1/S phase, OS25 cells were treated with thymidine (Sigma-Aldrich) at a concentration of 1mM overnight. For cell cycle arrest at M phase OS25s were treated with colcemid (Gibco) at a concentration of 0.13µM overnight. When harvesting OS25 cells were trypsinised (0.05% trypsin-EDTA (Invitrogen)) at 37^{o}c, centrifuged at 200 x g for 5min, washed twice in ice cold Phosphate buffered saline solution (PBS – 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4 with 5mM sodium butyrate, pH 7.4 (all Sigma-Aldrich)) before resuspending in PBS.

2.2.2 FLOW CYTOMETRIC ANALYSIS OF ARRESTED CELL POPULATIONS

In order to determine the efficacy of cell cycle arrest flow cytometric analysis of cells was conducted. Cells were harvested and washed twice in PBS/5mM sodium butyrate before resuspending in 2ml PBS/5mM sodium butyrate and adding 20µl propidium

iodide (1mg/ml -Sigma-Aldrich) and 20 µl RNAse A (20mg/ml). Cells were analyzed on a FACS Calibur flow cytometer using Cell Quest software (BD). A minimum of 100,000 events were recorded per sample with forward and side scatter gates to exclude non-viable cells. Cell counts were plotted against fluorescence to give a cell cycle profile based on DNA content (figure 3.1).

2.2.3 Isolation of cell cycle fractions by Florescence activated cell sorting (FACS) of OS25 cells

Approximately 10^8 acetone fixed OS25 cells were used per cell sort. Cells were washed in PBS/5mM sodium butyrate, centrifuging very slowly at around 35 x g for 10 minutes, before resuspending in PBS/5mM sodium butyrate at a concentration of 5 x 10^6 cells per 100µl. 30-45 minutes before the cell sort 10µl of propidium iodide (1mg/ml -Sigma-Aldrich) was added per ml along with and 20 µl RNAse A (20mg/ml). Cells were isolated based on DNA content as stained by propidium iodide, gating shown in figure 3.8. Sorting was performed on a MoFlo high speed sorter (Dako Cytomations). Forward and side scatter gates were included to exclude nonviable cells. Sorting was performed by Dr R. Bird as part of the University of Birmingham's technology hub facilities. Once sorted into cell cycle phase's cells were counted and reanalysed to determine sample purity. Undifferentiated OS25 cells were washed to remove propidium iodide, snap frozen and stored at -80^{o}c until use. As cell numbers of day 7 embryoid bodies were so low C-ChIP was performed on these samples immediately to minimise cell loss. Carrier cells were added prior to washing to remove propidium iodide.

2.2.4 SNAP FREEZING CELLS AFTER FACS

Cell samples were transferred into 1.5ml microcentrifuge tubes and centrifuged at 200 x g for 10 minutes. The supernatant was removed entirely and the remaining pellet was immediately placed in dry ice and stored at -80^{o}c.

2.3 CHROMATIN IMMUNOPRECIPITATION

2.3.1 NATIVE CHROMATIN IMMUNOPRECIPITATION (N-CHIP)- CHROMATIN ISOLATION

Cultures of OS25 cells were harvested as previously described and washed three times in ice-cold PBS/5mM Nabutyrate. Cell counts were performed and samples were resuspended in TBS (15mM NaCl, 10mM Tris-HCL pH 7.5, 3mM $CaCl_2$, 2mM $MgCl_2$, 5mM Na butyrate (all Sigma-Aldrich)) at approximately 2 x 10^7 cells/ml. An equal volume of 1% Tween40 (Sigma-Aldrich) in TBS and 1/200th volume of 0.1M PMSF (Sigma-Aldrich) was added before stirring on ice for 20 minutes to allow the breakdown of cell membranes and release of the nuclei. Cell lysates were transferred into a Dounce all-glass homogeniser ("tight" pestle) and homogenised with approximately 7 even strokes before checking nuclei emergence under a microscope. This resulted in an approximate yield of 75-80% intact nuclei which were centrifuged (800 x g, 4^{o}c, 20mins) and washed in 5mls of 5% sucrose (Sigma-Aldrich)/TBS before centrifugation at 1800 x g, 4^{o}c, 10mins. Nuclei were then resuspended in 1ml digestion buffer (0.32M Sucrose, 50mM Tris-HCL pH 7.5, 4mM $MgCl_2$, 1mM $CaCl_2$, 0.1mM PMSF, 5mM NaButyrate (all Sigma-Aldrich)). $A_{260/280}$ absorbance values were measured of an aliquot of each sample diluted in 0.1% SDS (Sigma-Aldrich) on a light spectrometer (Amersham/Implen) and used to provide an approximate concentration in order to resuspended samples in digestion buffer at a concentration of 0.5mg/ml.

In order to analyse the chromatin it must be enzymatically cleaved to give a ladder rich in mono-, di-, tri-, tetra-, and greater than penta-nucleosomes. Samples were divided into 1ml aliquots (if there was more than 1ml) to which micrococcal nuclease (Sigma-Aldrich) was added; 50U per 0.5mg/ml and incubated 37^{o}c for 5 mins, digestion was stopped by addition of 5mM EDTA and placing on ice for a further 5 minutes before centrifuging at 16000 x g. The first supernatant was removed (S1) and stored at 4^{o}c. Pellets were resuspended in 1ml (maximum) aliquots in lysis buffer (1mM Tris-HCl pH7.4, 0.2mM Na_2EDTA, 0.2mM PMSF, 5mM Na Butyrate (all Sigma-Aldrich)), and dialyzed overnight against 2l of lysis buffer at 4^{o}c. Samples were centrifuged (800 x g, 4^{o}c, 10min), the supernatant was collected (S2) and the remaining pellet was resuspended in 200µl lysis buffer (P). The concentration of each fraction was again determined by $A_{260/280}$ on a light spectrometer (Amersham/Implen), an aliquot of each diluted in 0.1% SDS. The degree of micrococcal nuclease digestion was determined by analysis on a 1.2% agarose gel (2ng of each sample in the presence of 0.3% SDS). Agarose gel electrophoresis (AGE) was performed before staining the gel with ethidium bromide (5µl of 10 mg/ml). Providing a good nucleosomal ladder was observed in both S1 and S2, these fractions were combined and used as the input for ChIP.

2.3.2 Carrier Chromatin Immunoprecipitation (C-ChIP)- Chromatin Isolation

Carrier chromatin immunoprecipitation is an adapted form of NChIP in which cells from a different organism to the target one (*Mus musculus*), in this case *Drosophila melanogaster* SL2 cells, are used to bulk up low cell numbers (O'Neill et al. 2006). This method relies on the ability to isolate chromatin from SL2 cells, which is more delicate than mammalian cells in addition to efficient immunoprecipitation and species specific primers for PCR.

SL2 cells are harvested, washed 3 times in ice cold PBS and divided into aliquots of 3 x 10^7 cells/pellet which are snap frozen and stored at -80^{o}c until required. Around 10^4 target OS25 cells (cells sorted via FACS) were added to one SL2 pellet and made upto 1ml with ice cold NB buffer (15mM Tris-HCL pH 7.4, 60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 0.5mM 2 mercaptoethanol, 0.1mM PMSF, 5mM NaButyrate). Samples were centrifuged at 200 x g for 7 minutes and washed twice in 0.5ml of ice cold NB buffer before resuspending in a final volume of 750µl NB buffer. Cell lysates were transferred into a Dounce all-glass homogeniser ("tight" pestle) and homogenised, only 2 strokes are required when using frozen SL2 cells and fragile acetone fixed cells. This resulted in an approximate yield of 75-80% intact nuclei. samples were centrifuged at 800 x g for 15 minutes at 4^{o}c then washed twice in 1ml 5% sucrose/NB buffer. Samples were resuspended in 1ml digestion buffer and the concentration was again determined by $A_{260/280}$ on a light spectrometer. Samples were centrifuged at 800 x g, 4^{o}c for 10 minutes then resuspended at 0.5mg/ml in digestion buffer. 50U per 250µg micrococcal nuclease was added to digest chromatin for 5 minutes at 28^{o}c. Digestion was stopped by addition of 5mM EDTA and placing on ice for a further 5 minutes before centrifuging at 16000 x g. The first supernatant was removed (S1) and stored at 4^{o}c. Pellets were resuspended in 1ml (maximum) aliquots in lysis buffer and dialyzed overnight against 2l of lysis buffer at 4^{o}c. Samples were centrifuged (1800 x g, 4^{o}c, 10min), the supernatant was collected (S2) and the remaining pellet was resuspended in 200µl lysis buffer (P). The concentration of each fraction was again determined by $A_{260/280}$ on a light spectrometer, an aliquot of each diluted in 0.1% SDS. The degree of micrococcal nuclease digestion was determined by analysis on a 1.2% agarose gel (2ng of each sample in the presence of 0.3% SDS). Agarose gel electrophoresis (AGE) was performed before staining the gel with ethidium bromide (5µl of 10 mg/ml). Providing a good

nucleosomal ladder was observed in both S1 and S2, these fractions were combined and used as the input for ChIP.

2.3.3 Chromatin Immunoprecipitation (N-ChIP and C-ChIP)

The immunoprecipitation of micrococcal nuclease digested chromatin from both N-ChIP and C-ChIP was conducted in the same manner from this point onwards. The use of siliconised microcentrifuge tubes, 15ml centrifuge tubes and Pasteur pipettes (Appleton Woods) ensured maximum recovery of material at this stage (siliconisation was performed in house). A small amount was removed from samples (~100µl) as a pre-immune (PI) sample and the remaining was divided in two (~50-100µg per antibody). Samples were placed in siliconised 1.5ml centrifuge tubes and an antibody was added to each (antibodies listed in table 2.1). The volume of each sample was made up to 800µl (400µl for PI samples) with Incubation buffer (50mM NaCl, 20mM Tris-HCl pH 7.5, 20mM Na butyrate, 5mM Na_2EDTA, 0.1mM PMSF). Samples were incubated overnight on a very slowly rotating platform at 4^oc before adding 200µl of pre-swollen protein A-sepharose (50% w/v slurry, GE Healthcare) and continuing the rotating incubation for a further 3 hours at room temperature.

Table 2.1: Affinity purified Antibodies used for Native and Carrier Chromatin Immunoprecipitation

Antibody specificity	Origin	Volume per sample
Rabbit anti-H3K4me3	In-house R612 (50% SAS cut)	25ul
Rabbit anti-H3K9ac	In-house R609 (50% SAS cut)	50ul
Rabbit anti-H4K8ac	In-house R403 (50% SAS cut)	25ul
Rabbit H3K27me3	Upstate	10ul
Rabbit H3K9me2	Upstate	7ul
Rabbit H3S10phos	Upstate	10ul
Rabbit Pre-immune	In-house	7ul

The affinity purified antibodies listed were used for both Native and Carrier Chromatin Immunoprecipitation, the origin and volume per sample are included.

Samples were centrifuged (16000 x g, 10 minutes), the supernatant was then removed and kept on ice, this is the unbound (UB) fraction. The pellet of protein A-sepharose was then washed in 10ml Wash Buffer A (50mM Tris-HCl pH 7.5, 10mM EDTA, 5mM sodium butyrate, 50mM NaCl), centrifuged at 800 x g for 5 min before washing in 10ml Wash Buffer B (50mM Tris-HCl pH 7.5, 10mM EDTA, 5mM sodium butyrate, 100mM NaCl) and again centrifuging (800 x g, 5 min). Finally the sample was washed in 10ml Wash Buffer C (50mM Tris-HCl pH 7.5, 10mM EDTA, 5mM sodium butyrate, 150mM NaCl) and centrifuged (800 x g, 5min). To elute the bound fraction (BD) pellets were resuspended in 250µl 1.0% SDS/incubation buffer and incubated for 15mins at room temperature with repeated inversions. Samples were centrifuged (16000 x g, 10 min), the supernatant (BD) was removed and stored on ice. This was repeated with a further 250µl of 1.0% SDS/incubation buffer, the supernatant was again removed and combined with the previous BD extract and 500µl of incubation buffer to reduce the concentration of SDS to 0.5%.

2.3.4 DNA ISOLATION

DNA was isolated from the BD and UB fractions by Qiaquick PCR purification kit (Qiagen). 5 volumes of PB buffer were added to samples. These were transferred to quick columns in collection tubes and centrifuged (16000 x g, 1min) in 800µl increments until all sample has passed through the quick column. Columns were washed once with PE buffer before eluting DNA with 100µl EB buffer for N-ChIP samples and 50µl EB buffer for C-ChIP samples.

2.3.5 PICOGREEN ASSAY OF CHIP DNA

In order to calculate the DNA yield of UB and BD fractions from N-ChIP and C-ChIP a picogreen assay was used. 2µl of DNA was diluted in 198µl PicoGreen (Invitrogen)

(diluted 1:200 in TE Buffer). Samples were analysed using spectrophotometer (Victor3 – Perkin Elmer). Percentage pull downs were calculated using the readings (BD/UBx100). For NChIP samples, the UBs were diluted down to the concentration of the BD so that equal amounts of DNA were analysed. C-ChIP UB samples were diluted twofold in ultrapure d.H_2O to prevent the UB signal over bleaching plates in the ^{32}P-PCR analysis.

2.3.6 Quantitative Real-Time Polymerase Chain Reaction (PCR) analysis of N-ChIP DNA

Quantitative real-time polymerase chain reaction (qPCR) was used for the analysis of the BD and UB DNA from N-ChIP in cell cycle regulated genes and pluripotency associated genes using the Qiagen Quantitect SYBR Green PCR Kit. QPCR was carried out on the ABIPrism 7900HT (Applied Biosystems). This allowed us to look at the relative levels of histone modifications in genes of our choice (see tables 2.2-2.7). QPCR was run on 384 well plates, 2µl of each sample DNA was placed into the plate in triplicate. To this 8µl of master mix (5µl 1xQuantiTect Syber green PCR master mix (Qiagen), 2µl ultra pure d.H_2O, 500nM Forward primer (0.5µl), 500nM reverse primer (0.5µl)) was added per well to give a total volume of 10µl per well. Primers were designed using Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/**primer-blast**) to have a single target sequence, reaction efficiencies of 90-110% and R^2 values greater than or equal to 0.98. Primers were made by Invitrogen. No template control and positive control (mouse genomic DNA) included in each plate.

Table 2.2: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP and C-ChIP DNA – cell cycle regulated genes

	Primer name		NCBI reference	Amplicon length	TM
G1/S-phase associated genes	S-phase kinase-associated protein 2 (Skp2)		NM 013787.2	123bp	60˚c
	Promoter region forward:	5' -GACCCTGACGCACCTCACGC- 3'			
	Promoter region reverse:	5' -AAGGATCCGTGACCCCGGCA- 3'			
	Cyclin D1 (Ccnd1)		NM 007631.2	135bp	60˚c
	Promoter region forward:	5' -CCTCCCTCCTAGCTGTCCTC- 3'			
	Promoter region reverse:	5' -GCTCGGACTGCTTCTCC- 3'			
	Cyclin-dependent kinase 2 (cdk2)		NM 016756.4	101bp	60˚c
	Promoter region forward:	5' -CGCCCTCGTGACGTGAACCA- 3'			
	Promoter region reverse:	5' -TCCCCCACTGCTCCAGACGG- 3'			
S-phase associated genes	Proliferating cell nuclear antigen (Pcna)		NM 011045.2	202bp	55˚c
	Promoter region forward:	5' -AGCTCGATTTGCCTGTGACT- 3'			
	Promoter region reverse:	5' -AATCGTACCCGTGGTTTGAG- 3'			
	DnaJ (Hsp40) homolog, subfamily C, member 2 (Dnajc2)		NM 009584.4	130bp	60˚c
	Promoter region forward:	5' -CGACAAAGGCTGTGCGCCAC- 3'			
	Promoter region reverse:	5' -GCAAGCACTGCAGTATTTTACACCA- 3'			
	Cell division cycle 6 (Cdc6)		NM 001025779.1	168bp	60˚c
	Promoter region forward:	5' -AGACCTGGGGCTGTCCTATT- 3'			
	Promoter region reverse:	5' -TCTCCCGCCACAAATTCTAC- 3'			
G2-phase associated genes	Dlgap5 discs, large (Drosophila) homolog-associated protein 5 (Hurp)		NM 001145949.1	128bp	60˚c
	Promoter region forward:	5' -GCCACGCCTAGCGGGTTTACC- 3'			
	Promoter region reverse:	5' -GCCCCTTCCCCTCGAACGAG- 3'			
	Cyclin B1 (Ccnb1)		NM 172301.3	105bp	60˚c
	Promoter region forward:	5' -CAACAAAGCTTTCGGGAACT- 3'			
	Promoter region reverse:	5' -TGTGGGGAGTACGCATATCA- 3'			
	Cyclin A2 (Ccna2)		NM 009828.2	126bp	60˚c
	Promoter region forward:	5' -CCGGCAGGATCCTATGAG- 3'			
	Promoter region reverse:	5' -GCCCGACGGGAATCTTAC- 3'			
G2/M-phase associated genes	Cyclin A1 (Ccna1)		NM 007628.3	135bp	60˚c
	Promoter region forward:	5' -TCCCGAGCCAGGGTTCTCGG- 3'			
	Promoter region reverse:	5' -GAGGGCTGACGGCGGCTTAC- 3'			
	Cell division cycle 25B (Cdc25b)		NM 001111075.2	119bp	60˚c
	Promoter region forward:	5' -AGAGCGTCGCACCGTCCCTT- 3'			
	Promoter region reverse:	5' -GGAGGTGGGAGCAGCCCTCT- 3'			
	Cell division cycle 25A (Cdc25a)		NM 007658.3	141bp	60˚c
	Promoter region forward:	5' -AGGCCCACGAGACGCTTCCT- 3'			
	Promoter region reverse:	5' -ACTCAGGATAGCCGCCGCCC- 3'			

Primer sets for the promoter region were used for the analysis of N-ChIP and C-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

Table 2.3: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP and C-ChIP DNA – Pluripotency associated genes

	Primer name	NCBI reference	Amplicon length	TM
house keeping gene	beta actin	NM 007393.3	142bp	60˚c
	Promoter region forward:	5' -GCACAGCTTCTTTGCAGCTC- 3'		
	Promoter region reverse:	5' -CATGGTGTCCGTTCTGAGTG- 3'		
Proto-oncogene/master regulator	C-myc	NM 001177352.1	160bp	60˚c
	Promoter region forward:	5' -AGAGGGCGGGGAAGCGAGAG- 3'		
	Promoter region reverse:	5' -TCGGACCCGGCAGCTGCGAG- 3'		
Pluripotency associated gene	Kruppel-like factor 4 (Klf4)	NM 010637.3	111bp	60˚c
	Promoter region forward:	5' -AGCTACCATGGCAACGCGCA- 3'		
	Promoter region reverse:	5' -CTCCCCGCCTCCTCCCTGTG- 3'		
	Nanog homeobox (Nanog)	NM 028016.2	158bp	60˚c
	Promoter region forward:	5' -AGGGAGGCAAAGCTTAGGG- 3'		
	Promoter region reverse:	5' -GCAGCCGTGGTTAAAAGATG- 3'		
	POU domain, class 5, transcription factor 1 (Pou5f1/Oct4)	NM 013633.2	139bp	55˚c
	Promoter region forward:	5' -CTGTAAGGACAGGCCGAGAG- 3'		
	Promoter region reverse:	5' -CAGGAGGCCTTCATTTTCAA- 3'		

Primer sets for the promoter region were used for the analysis of N-ChIP and C-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

Table 2.4: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP DNA

Gene	Primer name	NCBI reference	Amplicon length	TM
Ccnd1	Position 1	NM 007631.2	135bp	60˚c
	Forward:	5' -CCTCCCTCCTAGCTGTCCTC- 3'		
	Reverse:	5' -GCTCGGACTGCTTCTCC- 3'		
	Position2		121bp	60˚c
	Forward:	5' -GATTTTCCGCATGGATGGCAC- 3'		
	Reverse:	5' -CACCAATCTCCTCAACGACC- 3'		
	Position 3		170bp	60˚c
	Forward:	5' -GCCCCCAGTTTGACCTAGAC- 3'		
	Reverse:	5' -TTGGGGTGGGGGATTCTTTG- 3'		
	Position 4		111bp	55˚c
	Forward:	5' -AGATGCACAACTTCTCGGCA- 3'		
	Reverse:	5' -CCCTGGAGCCCTTGAAGAAG- 3'		

Primer sets for the promoter region were used for the analysis of N-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

Table 2.5: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP DNA

Gene	Primer name		NCBI reference	Amplicon length	TM
Pcna	Position 1		NM 011045.2	153bp	60˚c
	Forward:	5' -CTGCCTTGAAGCTGGAGAGT- 3'			
	Reverse:	5' -GAGCGACGCACACAGAAAAG- 3'			
	Position 2			202bp	55˚c
	Forward:	5' -AGCTCGATTTGCCTGTGACT- 3'			
	Reverse:	5' -AATCGTACCCGTGGTTTGAG- 3'			
	Position 3			150bp	60˚c
	Forward:	5' -AATGAGCGTGCAGAGTCACA- 3'			
	Reverse:	5' -CTTACTCTGCGCTCCGAAGG- 3'			
	Position 4			162bp	55˚c
	Forward:	5' -ACTCTCTCAGGGGTGAAGCA- 3'			
	Reverse:	5' -ACGAAAAGGCCAGACGTGAT- 3'			
	Position 5			102bp	55˚c
	Forward:	5' -TAACTTACTTGGTGCTTCGAATACT- 3'			
	Reverse:	5' -TCCAAAATTCTAAAATGTGCTGGTA- 3'			

Primer sets for the promoter region were used for the analysis of N-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

Table 2.6: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP DNA

Gene	Primer name	NCBI reference	Amplicon length	TM
Hurp	Position 1	NM 001145949.1	134bp	60˚c
	Forward:	5' -CGCTGGCCTATGCTCTCATT- 3'		
	Reverse:	5' -CCACTGGGGAATCTGCTCTG- 3'		
	Position 2		170bp	55˚c
	Forward:	5' -AGATTGGCTCATGTGTGGCT- 3'		
	Reverse:	5' -CCGTGTCAGCCATGAAGTCT- 3'		
	Position 3		161bp	60˚c
	Forward:	5' -AGTGCCTTTAGCATCTGGGC- 3'		
	Reverse:	5' -TGTGACCCCATGATGTACGC- 3'		
	Position 4		128bp	60˚c
	Forward:	5' -GCCACGCCTAGCGGGTTTACC- 3'		
	Reverse:	5' -GCCCCTTCCCCTCGAACGAG- 3'		
	Position 5		115bp	60˚c
	Forward:	5' -CGACCCTTCACGGAACTGAA- 3'		
	Reverse:	5' -ACCAACCTATCGCAAGCCTC- 3'		
	Position 6		147bp	60˚c
	Forward:	5' -CCTGCCTGTGCACTCATCTT- 3'		
	Reverse:	5' -TGAGACTCTACCCGACCTGT- 3'		
	Position 7		133bp	60˚c
	Forward:	5' -ACCGAAGTGTCTGTTTCGCT- 3'		
	Reverse:	5' -GGTGTCACGTTTTGCCAGTC- 3'		

Primer sets for the promoter region were used for the analysis of N-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

Table 2.7: Sequences of Mus Musculus genomic DNA primer sets used for the quantitative real-time PCR analysis of N-ChIP DNA

Gene	Primer name	NCBI reference	Amplicon length	TM
Cdc25b	Position 1	NM 001111075.2	111bp	60˚c
	Forward:	5' -ATGAGGCCACCTTAGGGGAT- 3'		
	Reverse:	5' -ATTGACGCGAGAATCCTGGG- 3'		
	Position 2		139bp	55˚c
	Forward:	5' -AGAGCGTCGCACCGTCCCTT- 3'		
	Reverse:	5' -GGAGGTGGGAGCAGCCCTCT- 3'		
	Position 3		106bp	60˚c
	Forward:	5' -GCCACCTCTCGGTCTTTGAG- 3'		
	Reverse:	5' -AGGTTGTGCATGGTCTGTGT- 3'		
	Position 4		130bp	60˚c
	Forward:	5' -AAGGTTGTGCCCTGCTCTTT- 3'		
	Reverse:	5' -ACCTGGCTTTCTAGTGCAGC- 3'		
	Position 5		135bp	60˚c
	Forward:	5' -GCTATGGCACAGTAGAGCCC- 3'		
	Reverse:	5' -GGAGGAAACAGTAACGGCCA- 3'		
	Position 6		143bp	60˚c
	Forward:	5' -CTAGGTGAGGCTGCTGGAAC- 3'		
	Reverse:	5' -TCTGGAAGCGCACATTCTCT- 3'		

Primer sets for the promoter region were used for the analysis of N-ChIP DNA. Amplicon length, NCBI reference code and the annealing temperature are given for each primer set.

The following PCR cycling parameters were followed:

Temperature	Time	Step	Cycles
95°c	15 minutes	Activation	x1
94°c	15 seconds	Denaturation	x40
55-60°c *	30 seconds	Annealing	
72°c	30 seconds	Extension	

[*depending on the primer pair]

A dissociation curve (95°c melt gradient) was used to verify specific amplification of the desired gene. The efficiency of each reaction and the Ct values were calculated by the ABI Prism 7900HT Applied Biosystems software.

The average of each triplicate was used to calculate the relative enrichment of the BD sample compared to the unbound sample ($\Delta\Delta CT=2^{-_{BD\ CT-UB\ CT)}}$. The ratio of BD to UB were then displayed graphically.

2.3.7 P32 Radioactive Polymerase Chain Reaction (PCR) analysis of C-ChIP DNA

In order to analyse DNA from CChIP radioactive PCR (^{32}P PCR) was used as qPCR has previously been found ineffective due to the high degree of interference from *Drosophila Melanogaster* DNA. ^{32}P PCR was run on 94 well plates, 2µl of each sample of DNA was placed into the plates in duplicate. To this 48µl of master mix (45µl 1.1x Reddy Mix PCR master mix (AB gene), 1µl ^{32}P (dCTP radiolabelled with α-^{32}P diluted 1:10 (Perkin Elmer), 1µl Forward primer, 1µl reverse primer) was added per well to give a total volume of 50µl per well. Primers were designed using Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/**primer-blast**) and made by Invitrogen (table 2.2). Controls with *Mus Musculus* genomic DNA and *Drosophila Melanogaster* genomic DNA were included in each plate. It is very important that primers are mouse specific to

ensure no amplification of *Drosophila* DNA occurs (see primer table for primers). The following PCR cycling parameters were followed:

Temperature	Time	Step	Cycles
94°c	60 seconds	Denaturation	
55-60°c*	60 seconds	Annealing	x42
72°c	90 seconds	Extension	

[*depending on the primer pair]

After 39 cycles, 15µl aliquots were removed from each well, the remaining samples were allowed to continue to cycle 42. This is to later confirm that PCR remains within the exponential growth phase.

2.3.8 5% POLYACRYLAMIDE GEL ELECTROPHORESIS

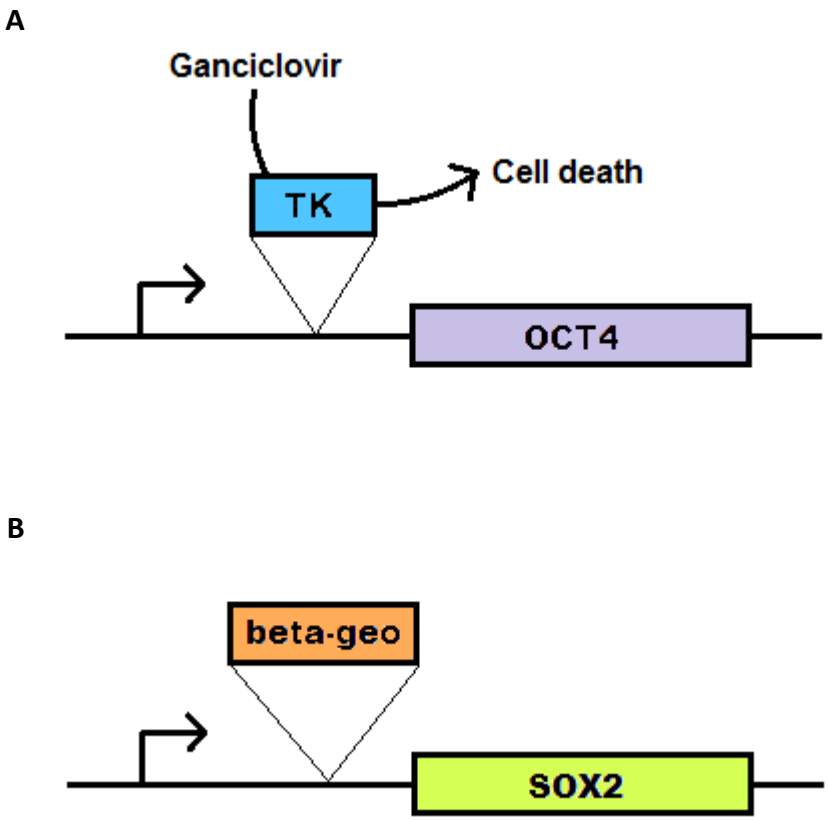
10µl of BD PCR product and 5µl UB, both from 39 cycles and 42 cycles, were run on a 5% polyacrylamide gel. Gels were prepared a minimum of 5 hours in advance to allow full polymerisation; 7ml 30% Acrylamide/1% NN'bisacrylamide, 4ml 10xTBE buffer, 200µl 20% w/v ammonium persulphate, 50µl TEMED, topped upto 40ml with dH_2O per gel. Gels were electrophoresed at 250 volts for approximately 40 minutes. Gels were carefully transferred onto Whatman No1 filter paper and dried on a gel drier for 1 hour minimum (SpeedGel System, Thermo Savant). These were then placed in phosphor storage screens and left to expose overnight before scanning with a Phosphor Imager (Molecular Imager PharosFX -BioRad). The intensity values for each PCR product were determined using Image Quant TL software (Molecular Dynamics), at which point the loading differences and dilution factors were factored in, the UB results were diluted down to the BD and the average of each duplicate was used to calculate the ratio of BD to UB and these results were then displayed graphically.

2.4 DIFFERENTIATION OF OS25S DOWN A NEURAL PATHWAY

OS25s allow for the positive selection of *Sox2* expressing (neuro epithelial) cells with geneticin in neurally differentiated cells and negative selection of any remaining *Oct4* expressing (undifferentiated) cells with ganciclovir (figure 2.1). OS25s were created by integrating a β*geo* construct into the *Sox2* locus conferring resistance to geneticin. And a hydromycin-thymidine kinase construct into the *Oct4* locus which by phosphorylating it, converts ganciclovir into a toxic drug (Billon et al. 2002).

In order to differentiate, OS25s were harvested and seeded at a density of $7x10^6$ cells per dish in 10cm bacterial dishes (Sterillin) (Day 0). The medium used, $M3^-$ consists of Dulbecco's Modified Eagle Medium supplemented with 20% heat inactivated fetal calf serum, 10% distilled sterile water, 1% 200mM L-glutamine, 1% penicillin and streptocillin, 1% non-essential amino acids and 0.25% 2-mercaptoethanol. Cells were supplemented with 50% new medium daily. On day 4 retinoic acid (Sigma-Aldrich) was added to a concentration of 10μM. Embryoid bodies were monitored daily. On day 6 the medium was replaced with a 50:50 mixture of Dulbecco's Modified Eagle Medium/nutrient mixture-F12 (Gibco) with N2 supplement (Gibco) and neural basal medium (Gibco) and B27 supplement (Gibco). To this 2.5μM Ganciclovir and 100μg/ml of geneticin (Gibco) were added. Embryoid bodies were harvested on day 7 by carefully transferring to a 50ml falcon tube and washing gently in cold PBS, they were then gently dissociated by passing through a 200μl pipette tip and washed again in PBS.

Figure 2.1: Representation of the genetic modifications in OS25 murine embryonic stem cell line



OS25 murine embryonic stem cells contain an inserted (**A**) hydromycin thymidine kinase construct in the *Pou5f1* locus which when expressed converts ganciclovir into a toxic form of the drug resulting in cell death and (**B**) a beta-geo contruct in the *Sox2* locus conferring resistance to geneticin. Combined these modifications allow the possitive selection of *Sox2* expressing cells and the negative selection of *Pou5f1* expressing cells (Billon *et. al.* 2002).

3 RESULTS

Histone modifications in stem cells and their differentiation are of great interest, particularly in light of the recent generation of iPS cells and the development of stem cell therapies. The abundance of histone modifications in stem cells has been extensively studied. Several histone modifications have been found to be enriched in stem cells, in particular H3K4me3 and H3K27me3 which are considered to be bivalent in many developmental genes in ES cells (Bernstein et al. 2006; Voigt et al. 2013a). The presence of bivalent marks in ES cells is believed to poise the expression of crucial genes, maintaining them at state of readiness, repressing gene expression whilst holding chromatin in an open configuration (Voigt et al. 2013a). Current understanding of the enzymes responsible for depositing histone marks and removing them is improving rapidly as interest in this area grows. For example there is a complex balance between H3K4me3 and the enzyme Set1A/B, and H3K27me3 and the enzyme PRC1/2, which can change in either direction (either off or on) depending on developmental cues.

As our understanding of the roles played by histone modifications grows, an increasingly complex picture of their networks continues to unfold before us. Post translational histone modifications have been closely linked to a variety of cellular functions including the activation and silencing of gene transcription, DNA replication and DNA repair (Bannister & Kouzarides 2011). However the heritability of such marks, and as such their status as “epigenetic” remains largely unresolved.

By narrowing in on separate phases of the cell cycle and comparing the enrichment of a variety of histone modifications at these different phases it will be possible to determine whether they are maintained consistently throughout the cell cycle and hence are

heritable or function in a more dynamic way. If histone marks are classically heritable, histone modifications would be expected to remain stable across all phases of the cell cycle, however, should marks play a more dynamic role in transcriptional regulation they would be expected to change based on the expression of a particular gene. If histone modifications prove to be more dynamic than heritable then levels of the marks may vary across the cell cycle. By focusing primarily on genes whose expression is cell cycle regulated it could be possible to make inferences regarding their relationship to gene expression, namely are they predictive or consequential.

3.1 THE CELL CYCLE IN EMBRYONIC STEM CELLS

A defining feature of embryonic stem cells is their ability to rapidly proliferate, which is brought about largely by their very characteristic short cell cycle. The ES cell cycle has been studied intensively and the mechanisms differentiating it from a normal cell cycle are now well understood, principally the significantly diminished gap phases in ES cells - 50-60% of the cell cycle is spent in S-phase (Figure3.1A & Savatier et al. 1994) and also the absence of a normally regulated G1/S transition. A major regulator in cell cycle control is the levels of CDK activity. In somatic cells these levels are cell cycle regulated via the MAPK pathway, however in ES cells levels of CDKs are cell cycle independent. These factors combine to allow ES cells to turnover in less than 10 hours (White et al. 2005; White & Dalton 2005).

The model system used within these experiments are OS25s, a mouse embryonic stem cell line which has been genetically modified at two loci (Billon 2002). A hydromycin-thymidine construct was inserted into the *Pou5f1* locus. When expressed hydromycin thymidine kinase converts ganciclovir into a toxic form of the drug resulting in cell death. Also a beta-geo construct was inserted into the *Sox2* locus conferring resistance

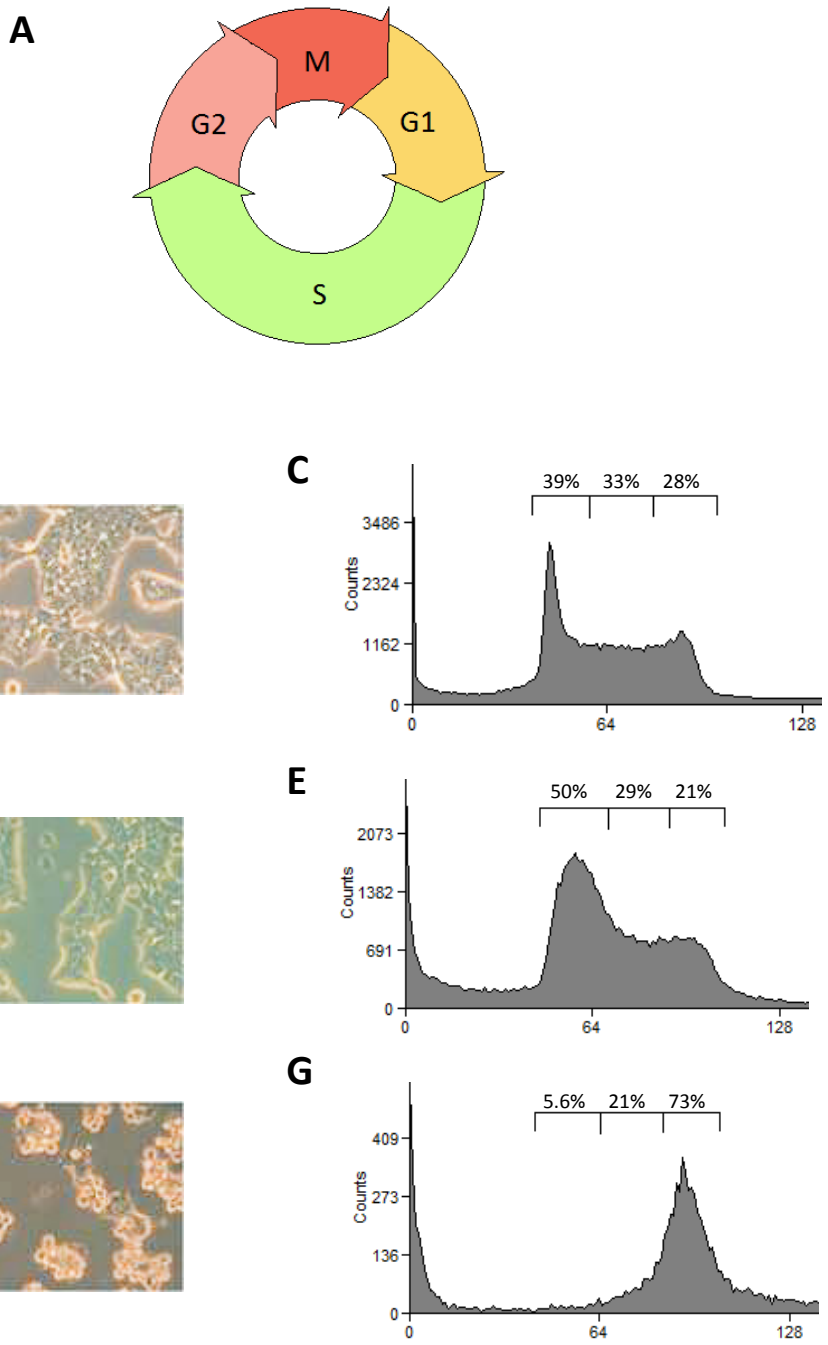
to geneticin upon *Sox2* expression (Figure 2.1). These two modifications allow the positive selection of neurally differentiated cells and the negative selection of remaining undifferentiated cells with the drugs ganciclovir and geneticin. Morphologically OS25 cells are typical of surface adherent mouse ES cells, growing in rounded colonies (Figure 3.1B) and upon staining with propidium iodide and analysing through flow cytometry a typical ES cell cycle profile is observed with a much higher proportion of cells found at S-phase than is seen that seen in somatic cells (Figure 3.1C).

3.1.1 Arresting ES cells at G1/S phase and M phase of the cell cycle

In order to compare histone modifications across the cell cycle it is first necessary to have populations of cells at specific phases of the cell cycle. This was initially achieved through the use of chemicals, thymidine was used to arrest cells at G1/S-phase. High concentrations of thymidine binds to DNA and interferes with nucleotide biosynthesis feedback mechanisms causing a shortage of nucleotides without which DNA replication cannot continue, thus causing cells to arrest at G1/S (Bostock et al. 1971). Thymidine treated, G1/S-phase arrested, cells are morphologically indistinguishable from the asynchronous population (Figure 3.1D). In order to analyse the cell cycle profile of these G1/S arrested populations of OS25s by flow cytometry cells were made permeable by treating with igepal. This allowed propidium iodide to enter cells and stain the DNA. The Asynchronous population as seen in Figure 3.1C has a typical ES cell profile. The propidium iodide cell cycle profile of thymidine treated cells clearly reveals that the majority of cells (79% in Figure 3.1E) are found within G1 or S phase.

Treatment of cells with colcemid caused cell cycle arrest at M-phase. Colcemid is a microtubule depolymerizing drug which prevents the formation of the bipolar spindle

Figure 3.1: The ES cell cycle: photographs and FACs profiles of cells arrested at G1/S (thymidine treated) and M phase (colcemid treated)



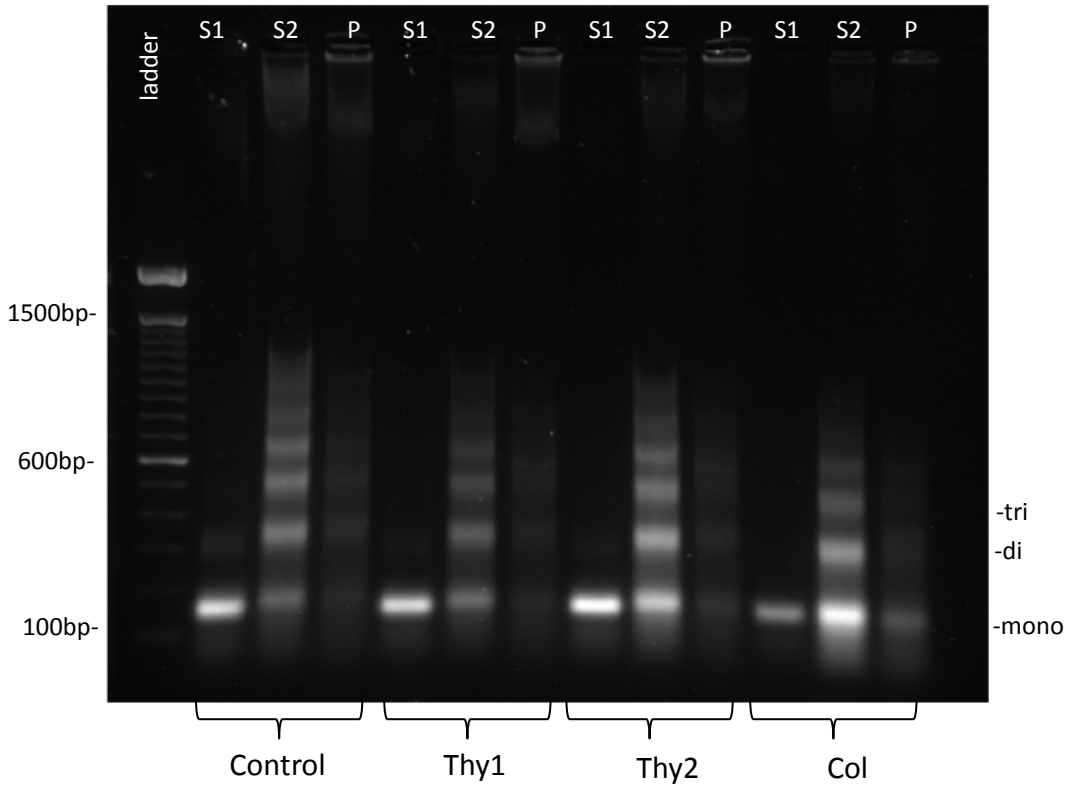
The mouse ES cell cycle. **A**; A representation of the ES cell cycle . **B**; Photograph of an asynchronous population of OS25 mouse ES cells with typical morphology. **C**; Propidium iodide profile of asynchronous OS25 cells. Including the proportion of cells found at G1, S and G2/M-phase (from left to right) **D**; Photograph of OS25 cells treated overnight with thymidine. Morphologically comparable to asynchronous OS25s. **E**; Propidium iodide cell cycle profile of thymidine treated cells. 79% of cells within the profile are found at G1 or S. **F**; Photograph of OS25 cells treated overnight with colcemid. Morphologically cells are more rounded as they detach from the flasks surface upon entering M-phase. **G**; Propidium iodide cell cycle profile of colcemid treated cells. 73% of cells found within G2/M-phase.

and cell cycle arrest at M-phase (Rieder & Palazzo 1992). Cells treated with colcemid appear smaller and rounder, this is caused when cells begin to detach from the flask surface in preparation for division (Figure 3.1F). In order to harvest only those cells arrested at M-phase, colcemid treated samples were harvested by knocking the side of the flask dislodging any cells entering M-phase. The propidium iodide cell cycle profile of colcemid treated cells reveals that the majority are arrested at G2/M-phase (73% in Figure 3.1G).

3.2 Enrichment of post translational histone modifications in ES cells arrested at G1/S phase and M phase

Native chromatin Immunoprecipitation is a well-established and an invaluable tool for looking at the relative levels of histone modifications in specific genes as compared to the whole genome (O'Neill & Turner 2003). Chromatin was isolated from the previously described chemically treated cells along with an untreated control sample (Asynchronous). The chromatin was then digested by micrococcal nuclease and analysed on a 1.2% agarose gel (Figure 3.2). Successful isolation and digestion was confirmed by the presence of nucleosomal ladders in the first and second supernatants (S1 and S2) ranging from mono-, di-, tri-, tetra- to greater than penta-nucleosomes. S1 contains mostly mononucleosomes, whereas S2 contains a larger nucleosomal ladder as seen clearly in each sample in Figure 3.2. The pellet fraction shows a smear of high molecular weight material with negligible amounts of chromatin. S1 and S2 were combined prior to immunoprecipitation, allowing the maximum amount of chromatin to be carried through for the immunoprecipitation. The presence of a nucleosomal ladder is important as some antibodies preferentially bind to mononucleosomes and others to di or tri-nucleosomes. The chromatin samples from the untreated, thymidine and

Figure 3.2 : Chromatin from thymidine treated; G1/S-phase arrested, and colcemid treated; M-phase arrested ES cells



Representative analysis of chromatin from N-ChIP. Chromatin was isolated from four samples of OS25 ES cells: A control population, thymidine treated-G1/S-phase arrested cells (Thy1 and Thy2) and colcemid treated- M-phase arrested cells (Col). Chromatin was digested with micrococcal nuclease generating a ladder of mono-, di-, tri-, tetra- and greater than pentanucleosomes. Equal amounts of each sample were run on a 1.2% agarose gel. First (S1) and second (S2) supernatants were subsequently combined for the immunoprecipitation step. The pellet (P) was discarded.

colcemid treated cells were incubated with antibodies against histone modifications associated with active (H3K4me3, H3K9ac, H4K8ac) and repressed (H3K27me3, H3K9me2) gene expression.

Following immunoprecipitation the unbound and bound DNA was extracted and purified using a Qiagen PCR purification kit and quantified using picogreen. Percentage pull downs per antibody per sample were calculated (table 3.1) and varied based on the abundance of the modification and the efficiency of pull down and recovery. The DNA obtained from N-ChIP was analysed by qPCR, equal concentrations of bound and unbound DNA were used as calculated from the picogreen analysis. Primer sets were for the promoter region of an array of cell cycle regulated genes. The bound to unbound ratio was calculated in each case giving the enrichment of the histone modification in the sample that bound the antibody compared to that which did not (unbound). All results were further normalised to the asynchronous result, therefore asynchronous values are all set at one, those below one represent a decrease and values above one display a fold increase in levels of the histone mark as compared to the asynchronous population.

3.2.1 ENRICHMENT OF HISTONE MODIFICATIONS IN G1 PHASE ASSOCIATED GENES

Relative levels of the active marks H3K4me3, H3K9ac, H4K8ac and the silent marks H3K27me3 and H3K9me2 were compared at the promoter of G1 phase associated genes *Skp2*, *Ccnd1* and *Cdk2* for OS25 cells chemically arrested at G1/S phase and M phase of the cell cycle along with an asynchronous control (Figure 3.3).

The levels of H3K4me3 follow the same pattern across all three genes (Figure 3.3). There is little or no change in its enrichment in G1/S compared to the Asynchronous in

Figure 3.3: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G1/S phase associated genes as determined by native chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.3: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G1/S phase associated genes as determined by native chromatin immunoprecipitation

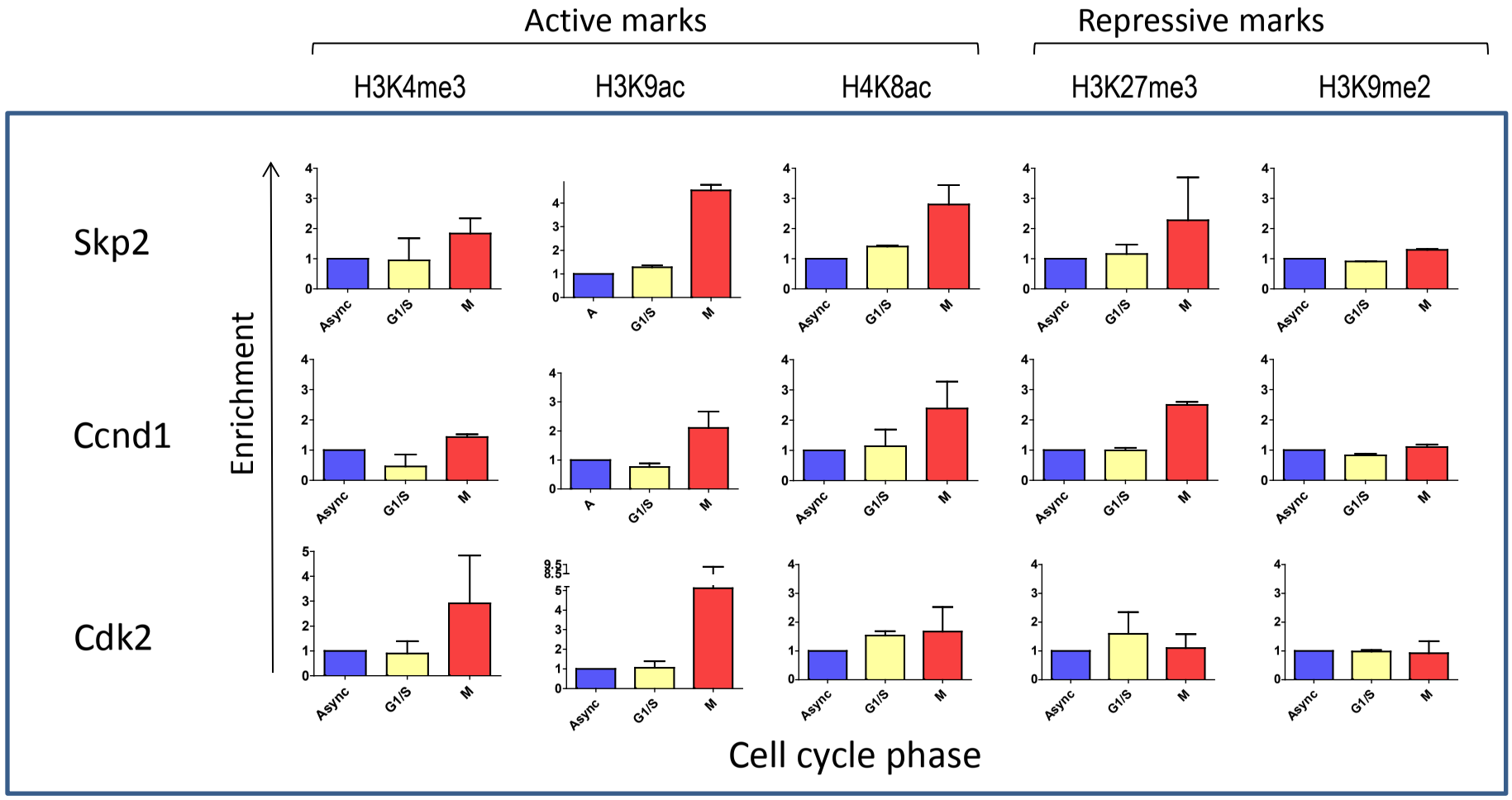


Table 3.1: Percentage pull downs of DNA from N-ChIP of G1/S and M-phase arrested ES cells as calculated by picogreen analysis

Antibody	Sample	N=1	N=2	N=3
H3K4me3	Async	1.111	2.703	3.622
H3K27me3	Async	25.579	3.990	25.700
H3K9ac	Async	0.384	1.120	1.318
H3K9me2	Async	7.922	2.730	10.130
H4K8ac	Async	4.356	7.960	7.834
H3S10phos	Async	16.455	7.604	19.949
H3S10phosK9ac	Async	8.715	2.798	6.844
PreImmune	Async	0.178	0.986	0.990
H3K4me3	G1/S	0.627	1.372	45.647
H3K27me3	G1/S	18.809	2.883	31.284
H3K9ac	G1/S	0.239	0.772	3.204
H3K9me2	G1/S	12.411	1.762	22.963
H4K8ac	G1/S	3.740	4.576	39.111
H3S10phos	G1/S	13.739	7.031	29.102
H3S10phosK9ac	G1/S	20.625	0.886	19.388
PreImmune	G1/S	0.204	0.413	1.972
H3K4me3	M	1.282	0.700	1.595
H3K27me3	M	11.144	2.347	36.030
H3K9ac	M	0.965	0.296	0.466
H3K9me2	M	4.510	2.334	9.783
H4K8ac	M	3.340	4.513	6.771
H3S10phos	M	4.983	4.908	19.049
H3S10phosK9ac	M	11.917	1.015	19.339
PreImmune	M	0.632	0.124	0.504

Relative levels of histone modifications were determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3, H3K9me2, H3S10phos and H3S10phosK9ac as seen above. DNA from antibody bound and unbound fractions were purified and quantified by picogreen. From these values the percentage pull down was calculated for each antibody. A pre-immune no antibody control was also included which had a very low pull down efficiency.

Skp2 and *Cdk2*. In *Ccnd1* H3K4me3 is actually depleted in G1/S when compared to the asynchronous population. Far more noticeable in all three genes however, is the enrichment of the mark at M phase, with all three genes showing a two or more fold enrichment. A similar pattern is observed for H3K9ac and H4K8ac and since these genes are expressed at G1 phase, active histone modifications appear to be pre-emptive of expression.

The repressive mark H3K27me3 shows no real change in enrichment for G1/S when compared to the Asynchronous sample, however there is a twofold increase at M-phase on *Skp2* and *Ccnd1* (Figure 3.3). Elevated levels of H3K27me3 at M-phase mirror those in the active marks. This bivalency has often been observed in ES cells, particularly between H3K4me3 and H3K27me3. The presence of a bivalent repressive mark is thought to keep the active mark in check, silent but poised for expression in this case at the next phase of the cell cycle (Voigt et al. 2013b). H3K9me2, also a repressive mark, does not change across any of the G1 phase associated genes and appears to be stable across the cell cycle (Figure 3.3).

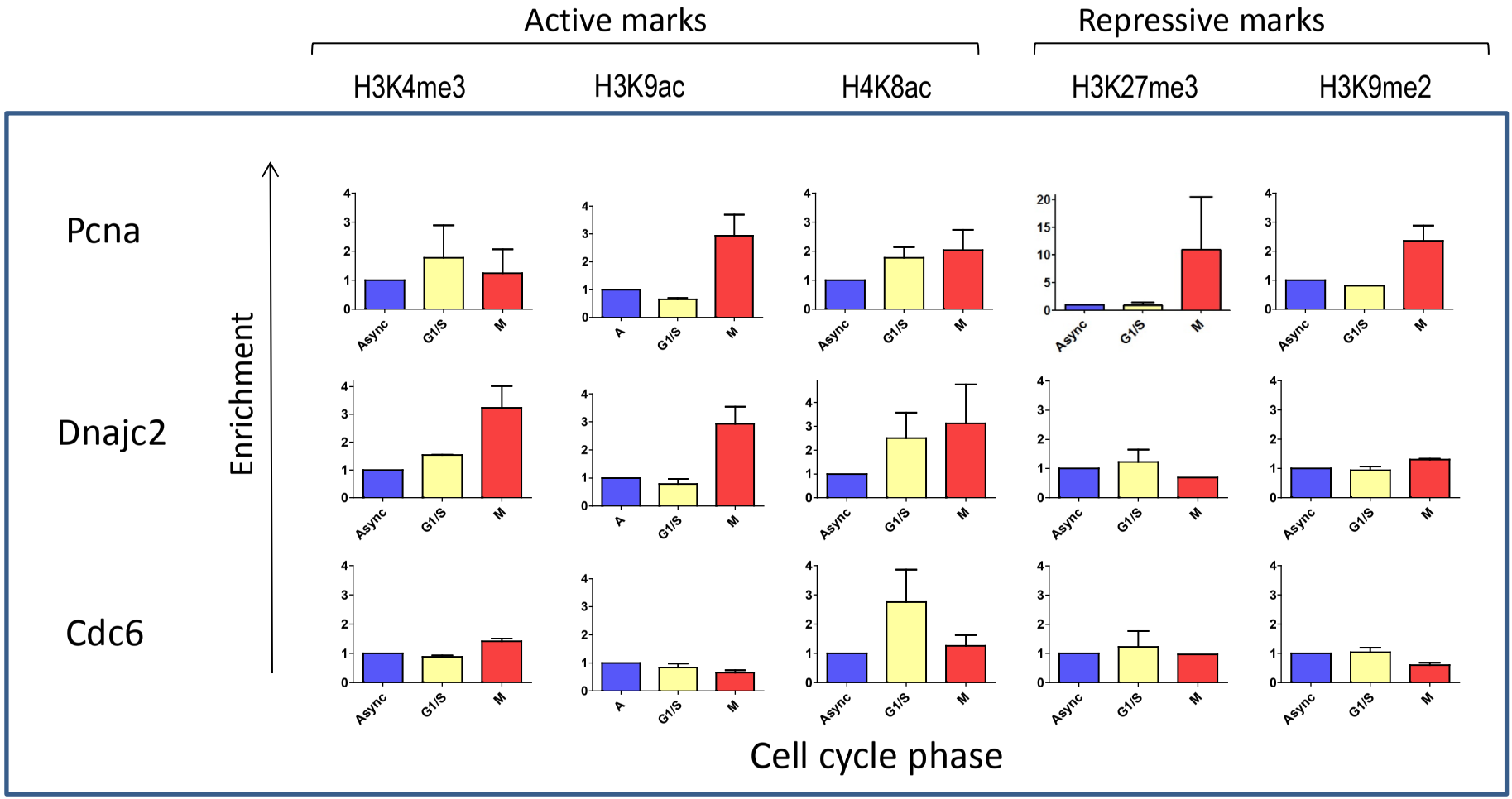
3.2.2 ENRICHMENT OF HISTONE MODIFICATIONS IN S PHASE ASSOCIATED GENES

A comparison was then made between the relative levels of the same active and repressive histone modifications at the promoter of the S phase associated genes *Pcna*, *Dnajc2*, and *Cdc6* (Figure 3.4). In *Pcna* enrichment of the active mark H4K8ac is slightly elevated in both G1/S and M-phase compared to the Asynchronous. Whereas H3K9ac is highly enriched at M-phase when compared to G1/S-phase. However the remaining active mark H3K4me3 peaks at G1/S phase with levels of the mark in M-phase comparable to the Asynchronous population. The repressive mark H3K9me2 has a greater than twofold enrichment in M-phase whereas a slight depletion is observed in

Figure 3.4: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in S phase associated genes as determined by native chromatin immunoprecipitation

Relative levels of histone modifications in S phase associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.4: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in S phase associated genes as determined by native chromatin immunoprecipitation



G1/S phase. Little change in the levels of H3K27me3 is observed in G1/S phase arrested cells compared to the asynchronous sample, however a large enrichment is observed at M-phase. This pattern is repeated for both H3K9ac and H3K9me2, which are also elevated levels in M-phase, potentially bivalently marking the gene at this point, maintaining a poised state pre-emptive of expression. The other active marks investigated are elevated at G1/S, potentially as the gene is beginning to be expressed.

In the next S-phase associated gene, *Dnajc2*, levels of the active modifications are enriched at M-phase when compared to G1/S phase and the Asynchronous population, particularly H3K4me3 and H3K9ac which are both at similar levels in G1/S phase when compared to Asynchronous but approximately threefold enriched at M-phase. H4K8ac is enriched around threefold in M-phase compared to the Asynchronous and in addition has an approximate twofold enrichment of the mark in G1/S when compared to the Asynchronous. In order for both G1/S and M-phase to be enriched compared to the Asynchronous population a depletion of the mark must be present at another part of the cell cycle. This suggests that a depletion in H4K8ac exists at an unobserved point of the cell cycle such as S-phase or G2-phase. The silent marks H3K27me3 and H3K9me2 appear to be stable across the cell cycle. It is possible that the increase in levels of the active marks at M-phase may be pre-emptive of an early S-phase expression. H4K8ac is not only enriched at M-phase but also at G1/S potentially pre-emptive of the point in which the gene is believed to be expressed.

Interestingly when another gene, *Cdc6*, is considered a completely new pattern of histone modifications is observed. The active marks H3K4me3 and H3K9ac, along with the silent marks H3K27me3 and H3K9me2 are all very stable across the cell cycle from G1/S to M-phase when compared to the Asynchronous population. The only mark which

appears to differ across the cell cycle in *Cdc6* is H4K8ac which increases almost threefold at G1/S when compared to M-phase and the Asynchronous population. If *Cdc6* is expressed in early S phase the mark is elevated at the point of expression, however should it be expressed at a later point in S-phase it could be considered pre-emptive.

With the exception of *Pcna*, neither H3K27me3 and H3K9me2 appear to change over the cell cycle. This maybe because any enrichment is very dynamic and is not observed in the cell cycle phases looked at or it may be that these marks are heritable. It is difficult to make any assumptions about the marks observed in S-phase associated genes as each one looked at appears to have a different signature. Without covering more of the cell cycle it is not possible to draw any firm conclusions, particularly as the cells being observed are arrested around the G1/S transition, which in mouse ES cells does not have a checkpoint. From the FACs profile of thymidine treated cells it is clear that cells are arrested at the G1/S transition (Figure 3.1), however the lack of a normal checkpoint is likely to have an effect on genes which are cell cycle regulated particularly those expressed in S-phase.

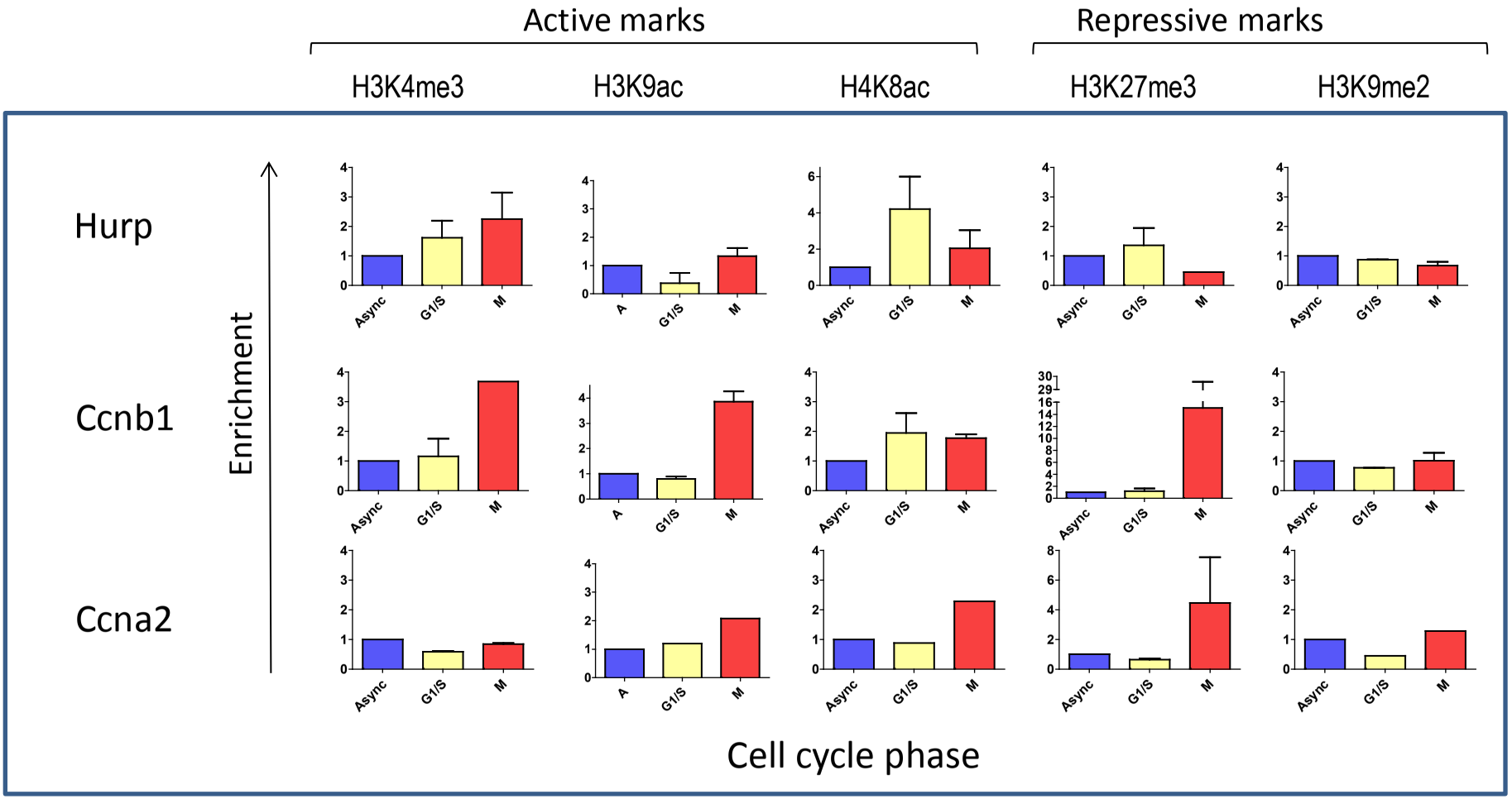
3.2.3 Enrichment of histone modifications in G2 phase expressed genes

The same panel of active and repressive histone modifications were then analysed in G2-phase associated genes *Hurp*, *Ccnb1* and *Ccna2* (Figure 3.5). Interestingly an enrichment of H3K4me3 and H3K9ac is observed at M-phase compared to G1/S for all three G2 phase associated genes. The enrichment of these two marks compared to G1/S phase varies from a very small increase in *Ccna2* upto more than a threefold enrichment in *Ccnb1*. The levels of H3K4me3 and H3K9ac do not vary in G1/S compared to the asynchronous population in either *Ccnb1* or *Ccna2*, however in *Hurp*, H3K4me3 is enriched and H3K9ac depleted when compared to Asynchronous.

Figure 3.5: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G2 phase associated genes as determined by native chromatin immunoprecipitation

Relative levels of histone modifications in G2 phase associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.5: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G2 phase associated genes as determined by native chromatin immunoprecipitation



The final active mark H4K8ac has the same enrichment at M-phase in *Ccna2* but not in *Hurp* or *Ccnb1*. Levels of the mark at G1/S in *Ccnb1* are equal to those in M-phase and greater than the Asynchronous population suggesting a depletion at some point between G1/S and M phase. In *Hurp* H4K8ac enrichment is approximately two fold greater at G1/S than M-phase and 3 fold greater than the Asynchronous.

The enrichment of most active marks at M phase could be reflective of G2 expression where marks remain in place throughout M phase. However this does not explain the enrichment of H4K8ac at G1/S phase in *Hurp* and *Ccnb1*. A very large increase in H3K27me3 is observed in M-phase in *Ccnb1* and *Ccna2*, compared to G1/S and the Asynchronous population. However in *Hurp* H3K27me3 is actually slightly enriched at G1/S and depleted at M-phase. This echoes the enrichment of H4K8ac suggesting that these two marks are acting bivalently in *Hurp*. No change in H3K9me2 levels is observed across the cell cycle in these three genes.

Increased levels of H3K4me3 and H3K9ac at M-phase may be reflective of gene expression. However, as the majority of the genes looked at have this enrichment of active and repressive marks at M-phase it may be that these marks increase (in at least these genes) at M-phase to protect or shield genes during mitosis. It is also important to consider the possibility that the effect is created by the treatment of cells with colcemid. The enrichment of some marks observed at G1/S-phase compared to the Asynchronous, and for H4K8ac and H3K27me3 also compared to M-phase, could be interpreted as pre-emptive of gene expression, however without looking at a further break-down of the cell cycle it is impossible to determine.

3.2.4 ENRICHMENT OF HISTONE MODIFICATIONS IN G2/M-PHASE ASSOCIATED GENES

If histone modifications are classically heritable they would be expected to remain stable across the cell cycle, however this is not being observed. It appears that histone marks have a much more dynamic role in chromatin. By investigating cell cycle regulated genes it may be possible to infer whether histone modifications are predictive (potentially playing an active role in gene expression) or consequential (marking genes as they are being expressed). When observing histone enrichment in G2/M phase associated genes an enrichment of active marks at G1/S (or possibly later such as late S or G2-phase) compared to the Asynchronous would be expected should marks be predictive of expression. However should marks be consequential an enrichment would only be expected at G2/M-phase. G2/M-phase associated genes investigated were *Ccna1*, *Cdc25b* and *Cdc25a* (Figure 3.6).

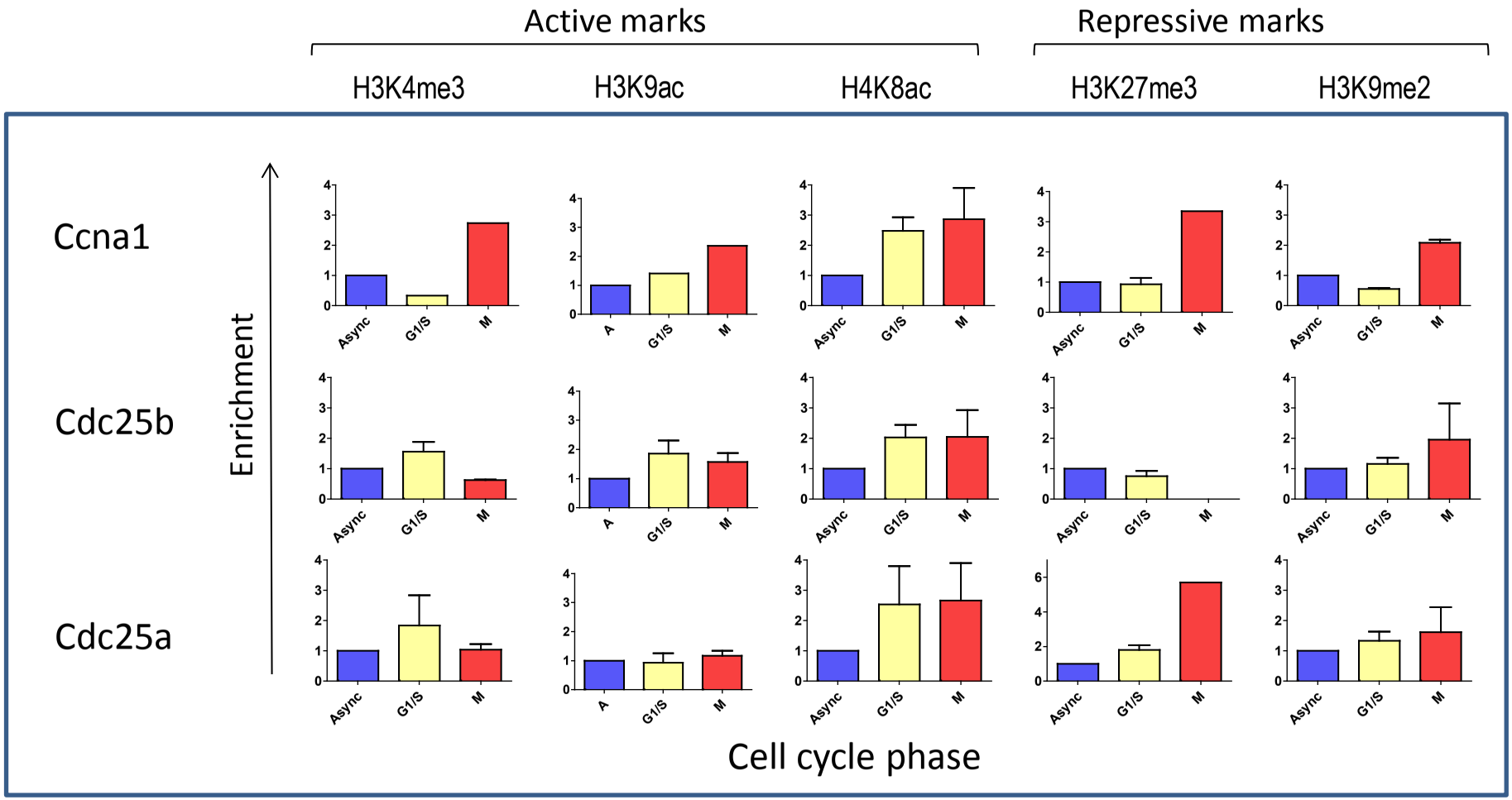
Interestingly both of these described trends are observed for different marks in different genes. In *Ccna1* all of the marks, active and repressive are enriched at M phase when compared to G1/S-phase and the Asynchronous population. However, the levels of H4K8ac in G1/S-phase are almost as high as those in M-phase and over twofold greater than the asynchronous, suggesting depletion at another point in the cell cycle. Levels of the remaining marks are approximately equivalent in G1/S phase compared to the Asynchronous apart from H3K4me3 which is depleted in G1/S.

This pattern is not observed in either *Cdc25a* or *Cdc25b*, which both have a similar pattern of histone modifications. In *Cdc25b* all three active marks are enriched at G1/S-phase compared to Asynchronous, H3K9ac and H4K8ac were also enriched to similar levels at M-phase. Unlike *Ccna1* and *Cdc25b*, expression of *Cdc25a* peaks at G1/S as well

Figure 3.6: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G2/M phase associated genes as determined by native chromatin immunoprecipitation

Relative levels of histone modifications in G2/M phase associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.6: Relative enrichment of a panel of histone marks in G1/S and M-phase arrested OS25 ES cells in G2/M phase associated genes as determined by native chromatin immunoprecipitation



as G2/M. In *Cdc25a* levels of H3K4me3 and H4K8ac are approximately twofold enriched in G1/S when compared to the Asynchronous population, this increase is not observed at M-phase for H3K4me3, but remained twofold enriched for H4K8ac. The repressive marks again appear to have the same pattern in all genes. No enrichment at G1/S-phase but a large enrichment at M-phase.

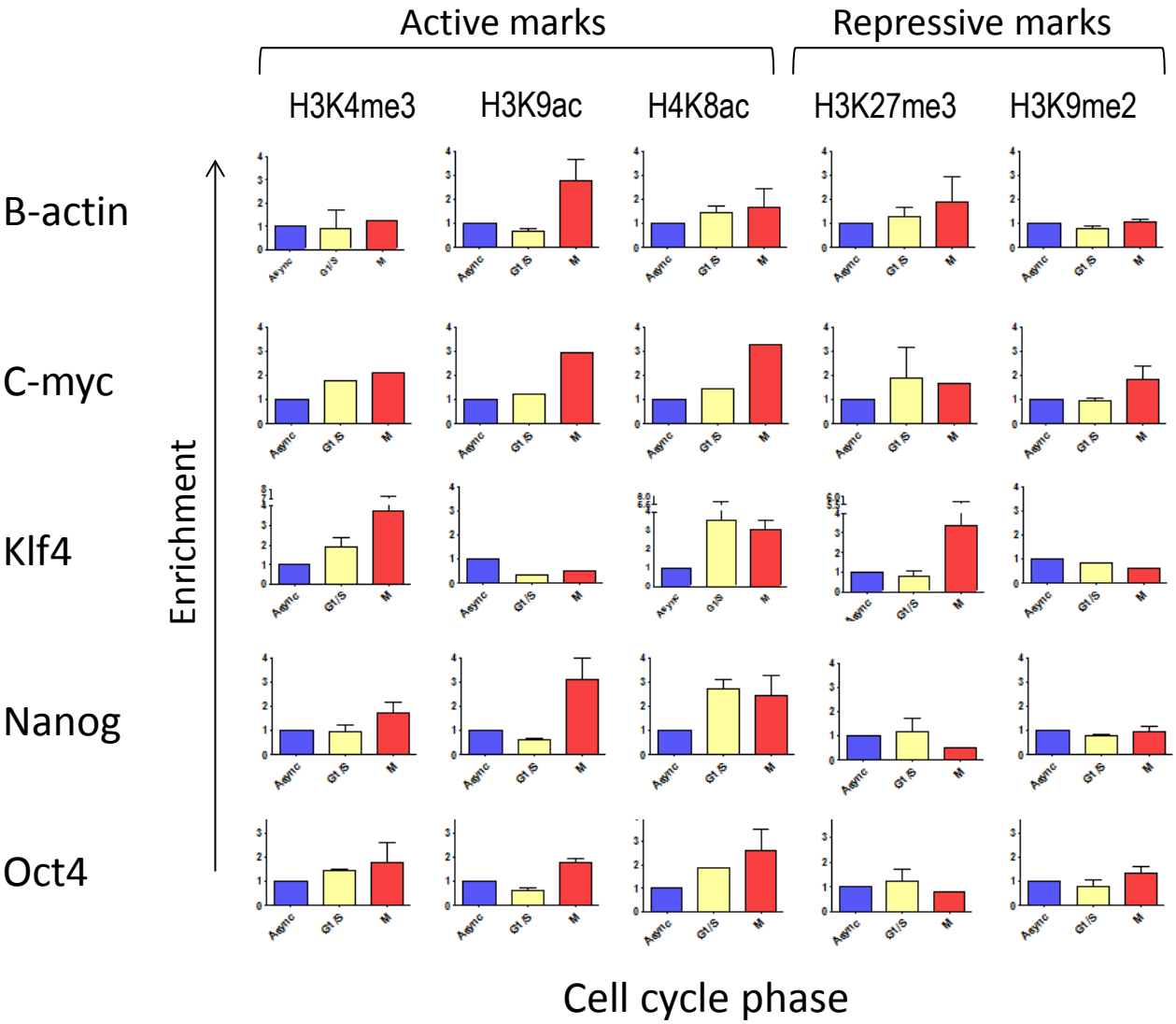
Based on the above description active marks appear pre-emptive in *Cdc25b*, H3K9ac and H4K8ac remaining elevated in M-phase. However a consequential pattern is observed for *Ccna1* and *Cdc25a*.

3.2.5 Enrichment of histone modifications in ES cells arrested at G1/S phase and M phase as observed in pluripotency associated genes

Pluripotency related genes are vital for the maintenance and proliferation of stem cells which includes their truncated G1 phase and distinctive cell cycle profile. The vital role played by these pluripotency transcription factors is reflected in their careful regulation. *Pou5f1*, *Nanog* and *Sox2* all regulate themselves and each other to maintain pluripotency. Between them *Pou5f1, Sox2* and *Nanog* are found bound to over 352 genes (Adjaye et al. 2005; Perera et al. 2006). The enrichment of histone modifications was also looked at across the cell cycle in the pluripotency related genes *Pou5f1, Nanog, Klf4, C-Myc* and a house keeping gene *β-actin*.

Given the careful regulation of these genes it would be expected that any histone marks located on the promoters of these genes would also be subject to the same tight control. Figure 3.7 represents the changes in H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2 at the promoters. It is clear to see that none of the histone modifications looked at are consistently stable across the cell cycle in these genes (figure 3.7). Different marks appear to be enriched at different cell cycle stages for each gene

Figure 3.7 : Histone modifications in pluripotency genes across the ES cell cycle



Relative levels of histone modifications in pluripotency associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

including *β-actin*. For *β-actin* the majority of marks appear stable across the cell cycle with the exception of H3K9ac, which is enriched approximately three fold at M phase compared to G1/S phase, and H3K27me3 which is also slightly more abundant at M-phase. There appears to be at least one active mark which is highly enriched at M-phase in each gene. In *C-myc*, H3K9ac and H4K8ac are both more abundant at M-phase, whereas in *Nanog* and *Pou5f1* the only significantly enriched active mark is H3K9ac. Contrary to these genes, in *Klf4*, H3K4me3 is the active mark enriched at M-phase. Other marks appear constant from G1/S to M-phase. However H4K8ac is consistently enriched at both G1/S and M when compared to the Asynchronous sample. Across these genes the silent marks are mostly stable throughout the cell cycle, with the exception of an increase of H3K27me3 at M-phase in *Klf4*. The expression of each of these genes is essential in maintaining pluripotency, therefore an absence of repressive marks is not unexpected. Even the enrichment of H3K27me3 at M-phase is accompanied by a similar enrichment of H3K4me3 at the same phase suggesting that the marks are bivalent.

In these pluripotency related genes, which are not thought to be cell cycle regulated, we observe variability as in the cell cycle regulated genes. Active marks are enriched at M-phase sometimes with a corresponding bivalent repressive mark. However the particular mark enriched at M-phase varies between the genes indicating that it is not the result of the chemical treatment of cells with colcemid but a genuine increase in marks at this phase of the cell cycle. These genes are known to be highly expressed in stem cells therefore the strong presence of at least one active mark in each gene reflects expectations, however why the marks are predominately found at M-phase is open to speculation. It may be that important genes are “protected” when entering M-phase as chromatin is condensed. The addition of active marks to these genes may enable their transcription to be recommenced as rapidly as possible upon cell division, allowing the

careful controls on such genes to reassert themselves as early as possible. Any repressive marks present are potentially bivalent marks, as the active marks may highlight genes for early transcription, so the repressive marks maintain the gene's silence, poised, allowing rapid transcription should it be required.

3.3 Enrichment of post translational histone modifications in FACs sorted ES cells

By chemically inducing cell cycle arrest at G1/S phase and M-phase and measuring the enrichment of a variety of different histone modifications it has become clear that levels of these marks often vary across the cell cycle. Each of the marks has changed across the cell cycle in at least one of the genes looked at, suggesting that marks are not stable across the cell cycle and calls into question their heritability. However, the pattern exhibited by marks is difficult to interpret and requires additional data. Current models addressing the inheritance of histone modifications have focused on their transmission through DNA replication which occurs during S-phase as this is the point at which histones are removed and re-deposited onto both the new and old strands of DNA including the incorporation of new histones which are believed to already possess post translational modifications on the histone tails (Probst et al. 2009). Many groups have focused on the exact mechanisms for this and there are three suggested models; semi-conservative inheritance, random inheritance and asymmetrical inheritance as described in figure 1.10 (Probst et al. 2009). This approach will lead to a greater understanding of how histone modifications are dealt with during DNA replication however this is not the only point of the cell cycle at which they are deposited or removed. By looking at more distinct phases of the cell cycle it would be possible to

glean a clear picture of the changes in levels of histone modifications in relation to the cell cycle and expression of particular genes.

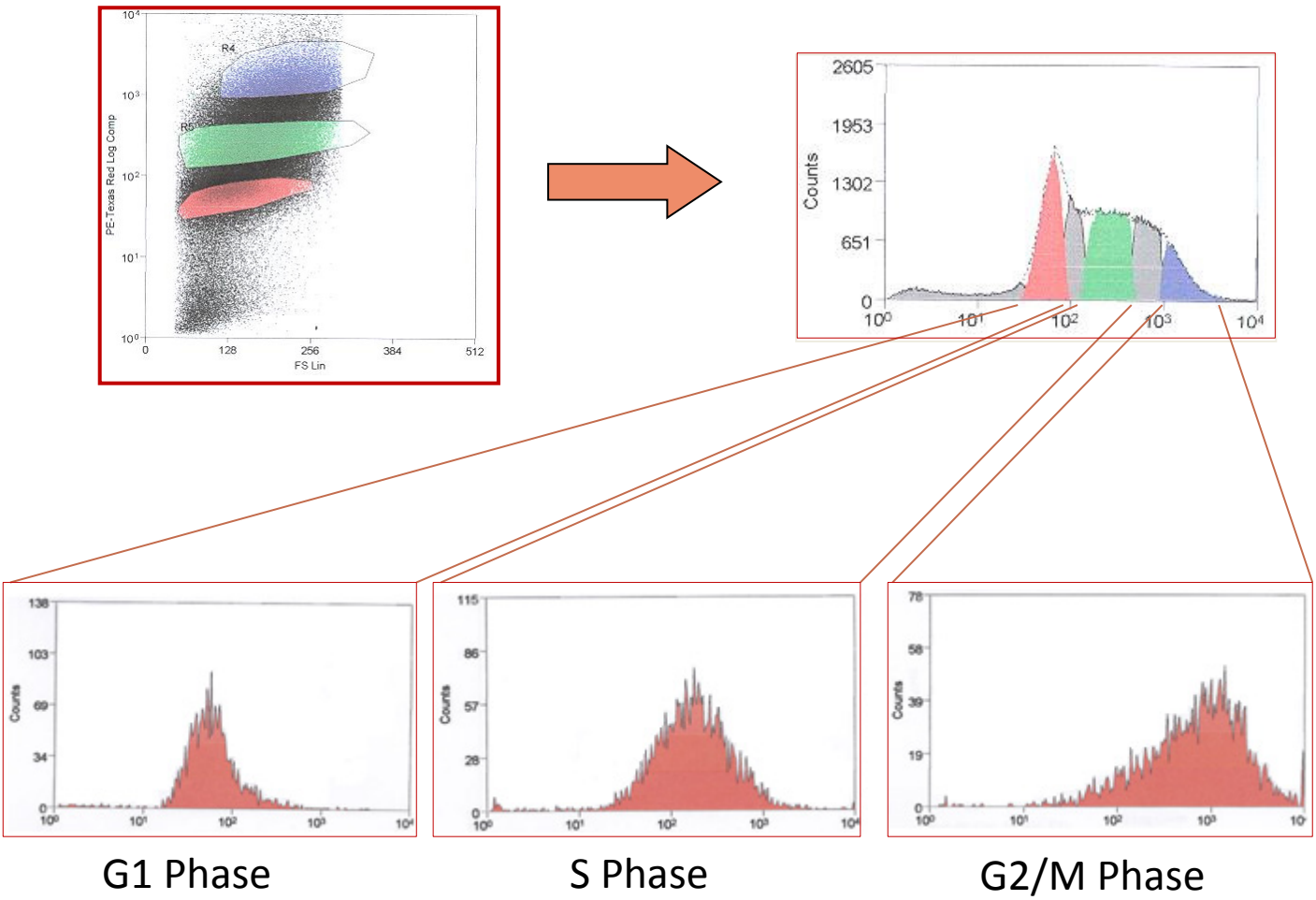
3.3.1 FLORESCENCE ACTIVATED CELL SORTING OF ES CELLS

In order to truly evaluate the enrichment of histones across the cell cycle it is important to include S-phase separately. This is possible through the use of Fluorescence activated cell sorting (FACs). As seen in Figure 3.8 propidium iodide can be used to view the cell cycle based on DNA content, this was one of the earliest applications of flow cytometry. The different phases of the cell cycle can be gated and cells isolated and purified based on this profile (Figure 3.8). As previously described cells in the initial peak of the propidium iodide profile are at G1, cells in the small peak at the end of the profile are at G2/M phase and those between these two peaks are at S phase. These populations were gated for as seen in the representative sort in Figure 3.8. Sorting cells allows the isolation of very pure populations of cells at G1, S and G2/M phase. Once cell cycle populations have been sorted samples of each are analysed again to confirm the purity of each sample. The high purity can also be seen in Figure 3.8. In this representative sort the purified samples are 88.38% at G1 phase, 88.13% at S phase and 89.87% at G2/M phase respectively.

3.3.2 COMPARISON OF FIXED AND UNFIXED CHROMATIN

In order to sort cells via FACS it is first necessary to fix cells as live stem cells actively pump out the propidium iodide dye (Goodell et al. 1996). Cell cycle profiles for N-ChIP are obtained by permeabilizing the cell membrane but this technique is not appropriate for obtaining intact cells via sorting. In order to retain whole cells and permeabilize the membrane, allowing the entry of propidium iodide, cells were fixed in an organic solvent. Initially ethanol was tried as it is routinely used as a fixative. When fixed in

Figure 3.8: Florescence Activated Cell Sorting (FACS) of OS25 ES cells into G1 phase, S phase and G2/M phase of the cell cycle



OS25 mouse embryonic stem cells were acetone fixed before labelling with propidium iodide and cell sorting into G1 phase, S phase and G2/M phase of the cell cycle. The above is a representative cell sort, including typical gating. Purified populations Re-analyzed by flow cyctometry again to confirm the purity of the samples (>86%). The numbers of cells obtained are in the region of 10^5 -10^6 per purified sample.

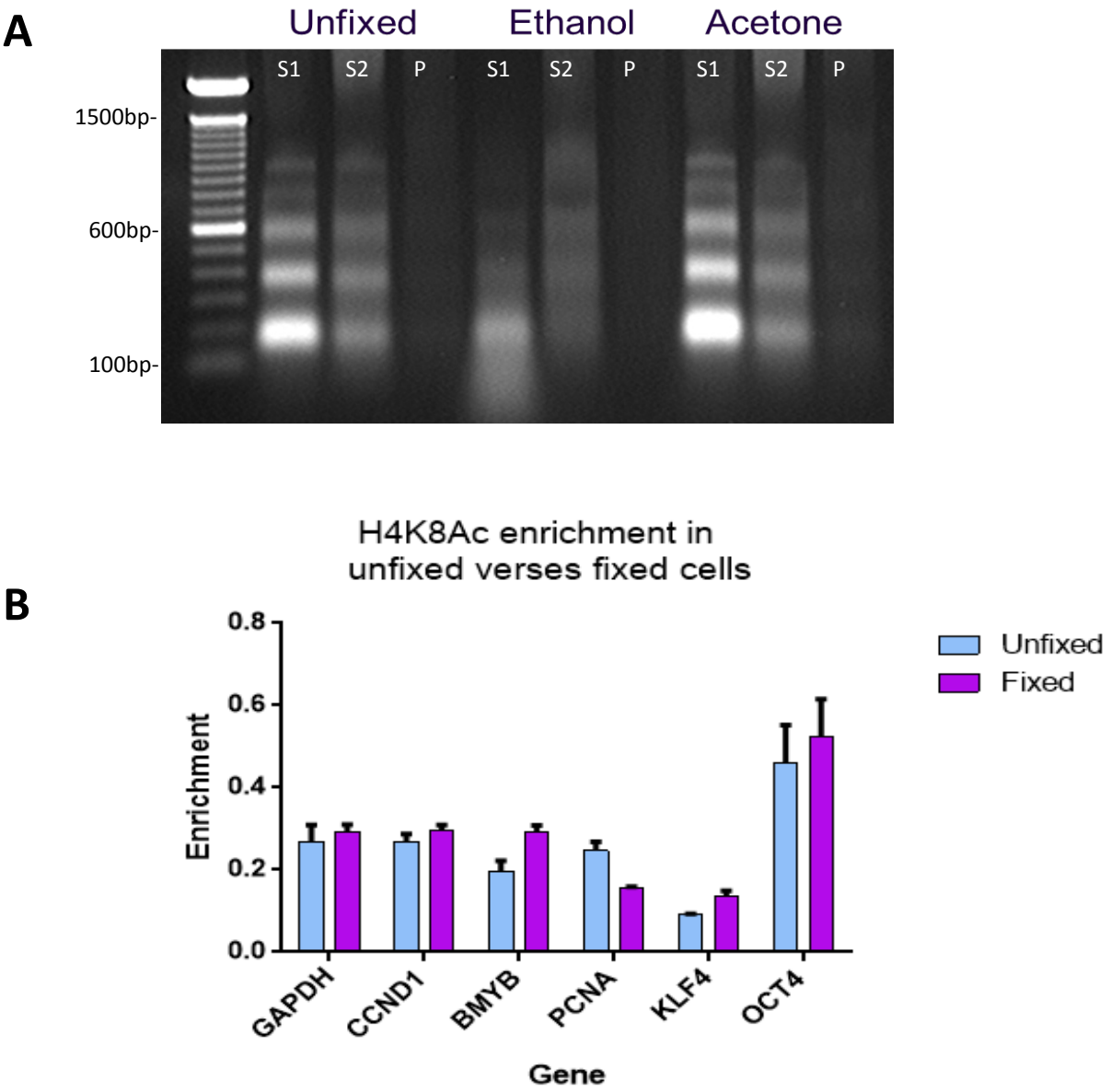
ethanol, cells become much more fragile and are over digested, as can be seen by the smearing of DNA in the gel (Figure 3.9 A, lanes 4-6). Subsequently acetone was used as the fixative. Cells fixed in acetone produce clear nucleosomal ladders when digested with micrococcal nuclease (Figure 3.9 A, lanes 7-9) and isolated chromatin is comparable to those from unfixed cells (Figure 3.9 A, lanes 1-3).

Having established that chromatin made from acetone fixed cells produces clear nucleosomal ladders, it was important to ensure that fixation did not interfere with the histone modifications in any way. Therefore N-ChIP was performed on both acetone fixed and unfixed cells using antisera against H4K8ac. This was analysed by qPCR using primers for Gapdh, *Ccnd1*, *Pcna*, *B-myb*, *Klf4* and *Pou5f1*. The results were displayed graphically as before, with H4K8ac enrichment as bound over unbound (Figure 3.9 B). Levels of the histone mark were extremely comparable in all of the genes tested. This therefore confirmed that the process of acetone fixation did not cause the loss of or alter the histone modifications within the cell, or the ability to immunoprecipitate chromatin from these cells using established protocol. Fixing cells in acetone in this way entirely avoids the use of chemicals to arrest cells at particular points in the cell cycle thereby eliminating the possibility that the chemicals are influencing results.

3.3.3 CARRIER CHROMATIN IMMUNOPRECIPITATION

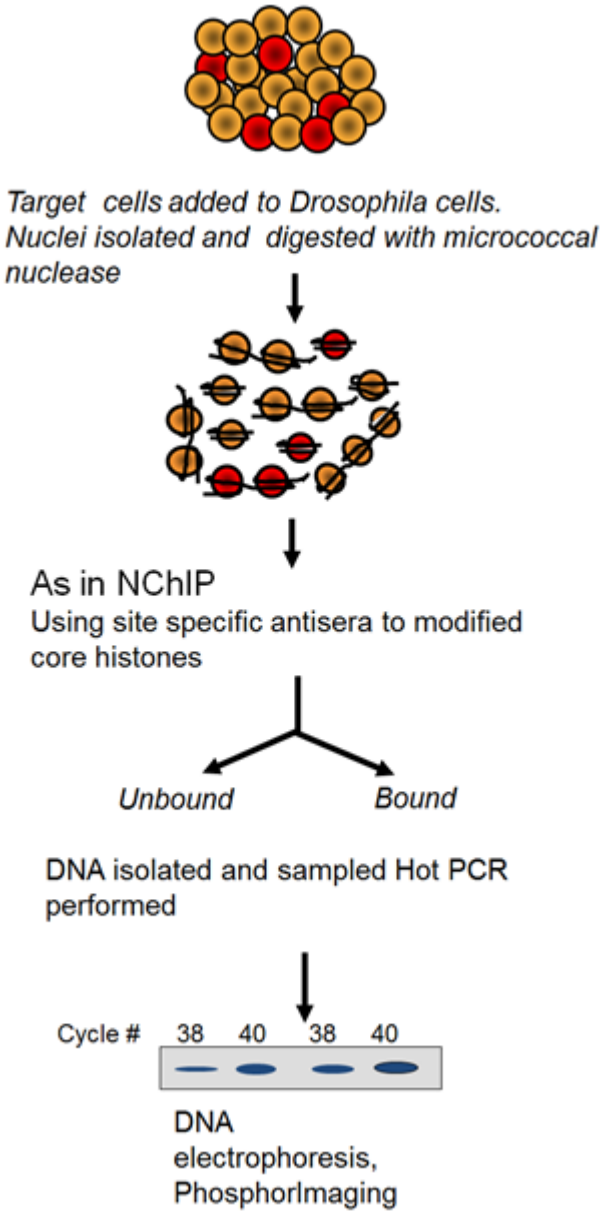
Determining levels of histone modifications in the isolated populations of cells at G1, S and G2/M phase cells was not possible with N-ChIP as cells numbers are so low, particularly as cells were fixed -reducing their size appreciably. It was therefore necessary to use an alternative chromatin immunoprecipitation technique; Carrier Chromatin Immunoprecipitation (C-ChIP) which was pioneered within the lab (O'Neill et al. 2006).

Figure 3.9: Comparison of unfixed and fixed chromatin: Micrococcal nuclease digestion and Enrichment of the H4K8ac mark in a panel of genes



A comparison of unfixed and fixed cells when used in Native Chromatin Immunoprecipitation. **A**, Chromatin from unfixed, ethanol fixed and acetone fixed ES cells was isolated and digested with micrococcal nuclease. This was analysed 1.2% agarose gel electrophoresis as per figure 3.2. Work was carried out by Milan Fernando. **B**, Enrichment of H4K8Ac calculated from unfixed and acetone fixed ES cells. Native Chromatin Immmunoprecipiation was performed using H4K8ac antibodies. The resultant DNA was analysed by quantitiative real-time PCR and relative levels of H4K8ac were calculated from equal concentrations of bound:unbound samples. Data are the mean ±SEM from technical triplicate reactions.

Figure 3.10: Representation of Carrier Chromatin Immunoprecipitation

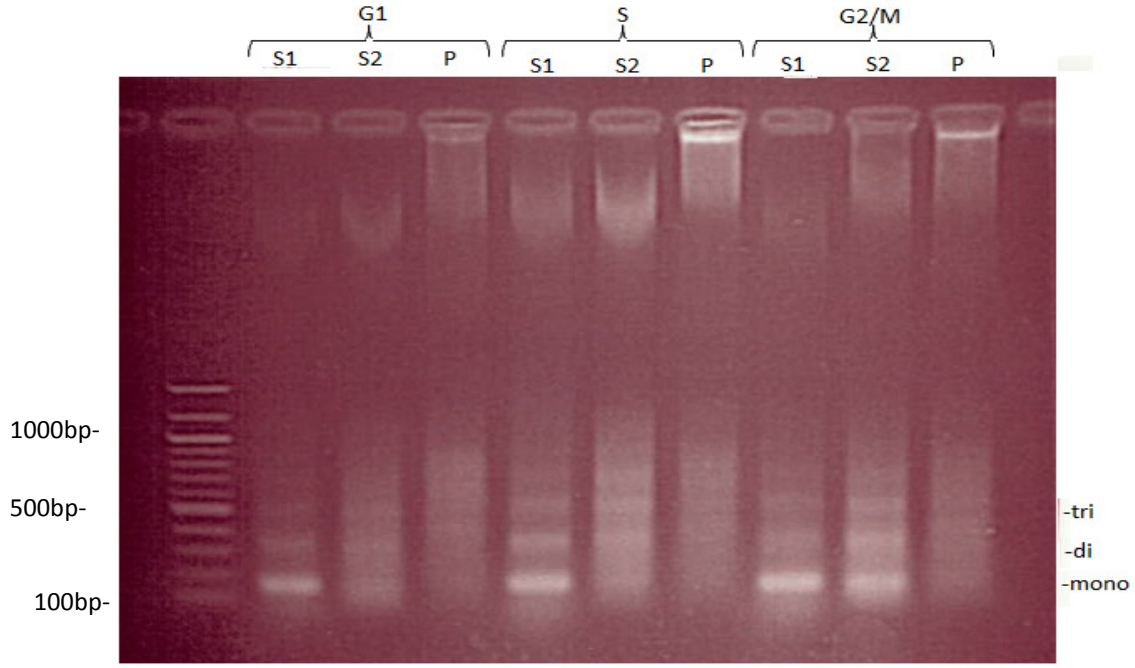


Diagramatic representation of Carrier-chromoatin immunoprecipitation (C-ChIP). Drosophila SL2 are used to bulk up low cell numbers of target cells (OS25 mouse ES cells), chromatin is then isolated and digested with miccrococcal nuclease. Digested chromatin is immunoprecipitated using antibodies for a variety of histone modifications and DNA is isolated and purified from both the antibody bound and unbound fractions. The resulting bound and unbound DNA are then analysed by radioactive PCR using mouse specific primers to look at the target cells only.

C-ChIP utilizes *Drosophila melanogaster* SL2 cells to augment low cell numbers before chromatin extraction. Figure 3.10 is a representation of the technique. In the experiments described here approximately 10^4 FACs OS25 cells were added to 3x 10^7 carrier SL2 cells per sample. The acetone fixed cells are very fragile therefore frozen SL2 cells are used as freezing makes the cells more fragile. The nuclei were isolated from cells. Chromatin was extracted before being digested by micrococcal nuclease and analysed on a 1.2% agarose gel (Figure 3.11). Chromatin isolated from SL2 cells is much more delicate which is reflected in the nucleosomal ladders observed in the agarose gel. Ladders are less defined, however clear ladders can still be observed in the S1 and S2 fractions at the appropriate sizes for mono-, di-, tri-, tetra- and penta-nucleosomes. As in N-chip mono-, di-, tri- and tetra-nucleosomes are seen clearly in each sample in the agarose gel of the chromatin. Both fractions S1 and S2 were combined prior to immunoprecipitation using antisera against H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. After immunoprecipitation the unbound and bound DNA was extracted and purified using a Qiagen PCR purification kit and quantified using picogreen in the same manner as N-ChIP. Percentage pull downs per antibody per sample were calculated as seen in table (table 3.2).

The large volume of *Drosophila melanogaster* DNA present within samples often impedes amplification of the *Mus musculus* DNA through qPCR. Various attempts had been made previously by the lab to use this form of analysis and during the course of this project Taqman probes were tested unsuccessfully as a possible tool, the same problems of inconsistency were found. The analysis of C-ChIP samples was therefore carried out using the highly sensitive radioactive P^{32} PCR. Unbound and bound DNA were analysed using species specific primer sets for the promoter region of selected cell cycle regulated genes. Histone modifications were looked at in a selection of the genes

Figure 3.11: Chromatin from SL2 cells plus 'target' cells to be used in C-ChIP



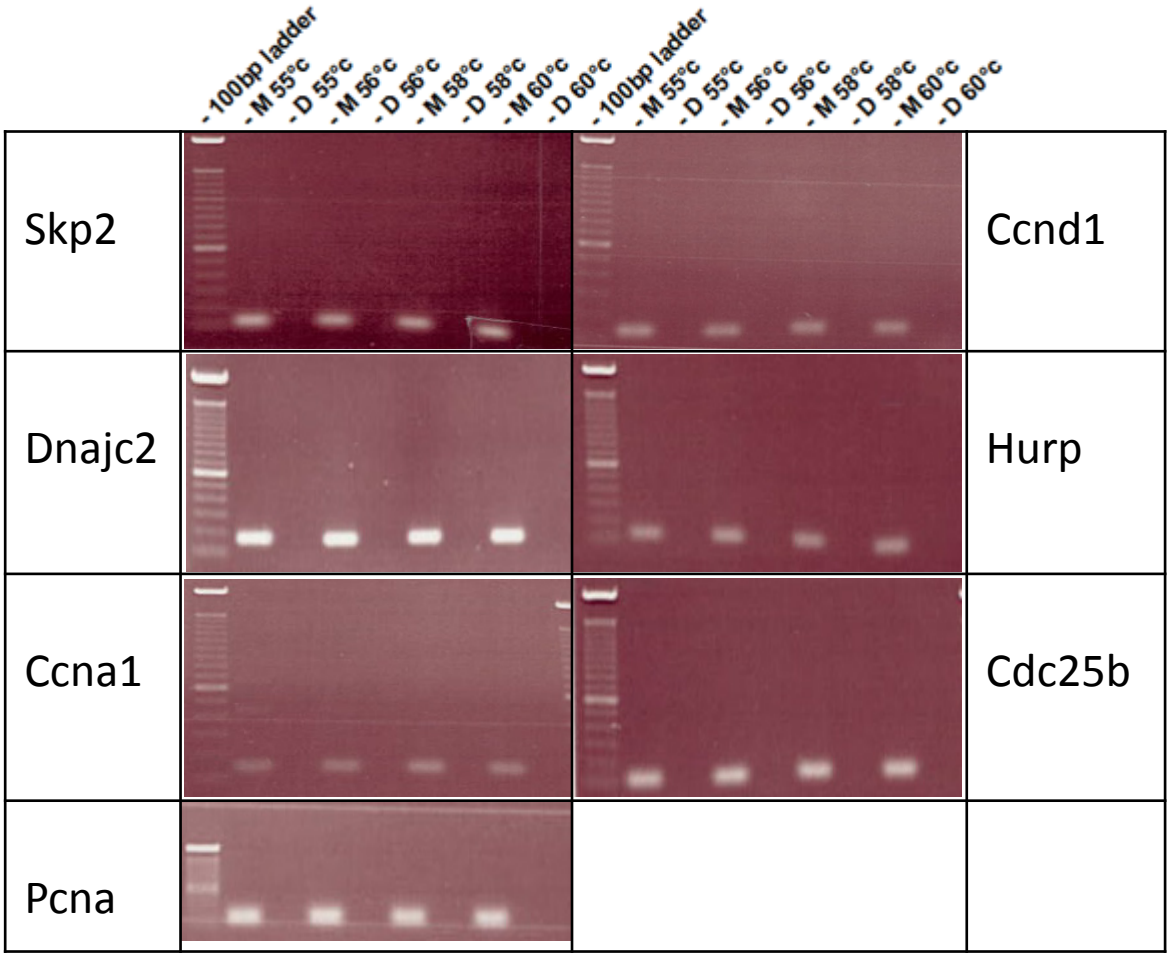
Representative analysis of chromatin isolated for C-ChIP. Chromatin was isolated from three samples of of approximately 3 x 10^7 drosophila SL2 cells, spiked with around 10^4 G1 phase OS25s (G1), S phase OS25 cells (S) and G2/M phase OS25 cells (G2/M) per sample. Chromatin was digested with micrococcal nuclease generating a ladder of mono-, di-, tri-, tetra- and greater than pentanucleosomes. Equal amounts of each sample were run on a 1.2% agarose gel. First (S1) and second (S2) supernatants were subsequently combined for the immunoprecipitation step. The pellet (P) was discarded.

investigated via chemical arrest; these genes were *Skp2* and *Ccnd1* (G1-phase), *Pcna* and *Dnajc2* (S-phase), *Hurp* (G2/M-phase), and *Ccna1* and *Cdc25b* (M-phase).

Species specific primers were needed to ensure that the information gained is strictly from the target cells with no cross reaction with the carrier DNA. Primer specificity was ensured by testing primers using *Mus musculus* genomic DNA and *Drosophila melanogaster* genomic DNA as can be seen in Figure 3.12. In all primer sets an amplicon was observed when the template was *Mus musculus* genomic DNA and not observed for the *Drosophila melanogaster* genomic DNA template.

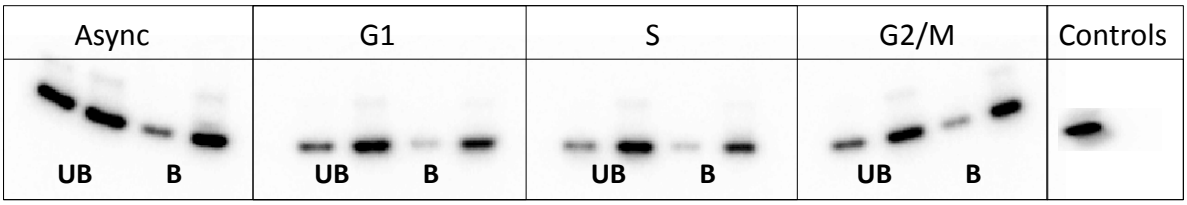
Unbound and bound samples of asynchronous cells, cells at G1, S and G2/M phase were amplified by PCR incorporating P^{32} dCTP into the product. Aliquots of each sample were taken at 39 cycles and then again at 42 cycles. These aliquots were run on 5% PAGE gels alongside *Mus musculus* and *Drosophila melanogaster* genomic DNA controls. The gels were dried onto filter paper and exposed to a phosphor screen overnight before scanning on a phosphoimager. A representative PAGE gel of radioactive PCR products obtained from C-ChIP samples is shown in Figure 3.13. In this gel bands are observed in for all the samples with the exception of the *Drosophila* negative control. Samples are taken at 39 cycles and 42 cycles to ensure that they are still within the exponential phase of PCR. It is quite clear in Figure 3.13 that bands taken at 42 cycles are stronger than those of aliquots taken at 39- cycles and that amplification has not plateaued. The strength of each samples signal was quantified using Biorad phosphoimaging software and enrichment was again calculated as the bound to unbound ratio, which has represented graphically. All values were normalised to an Asynchronous population. A value of 1 suggest that the mark is as present in that sample as in the asynchronous

Figure 3.12: Primer testing for species specificity prior to radioactive PCR analysis of C-ChIP samples



Skp2, Ccnd1, Dnajc2, Hurp, Ccna1, Cdc25b and Pcna primers were tested for species specificity across a temperature gradient by PCR using mouse genomic DNA (M) and Drosophila DNA (D). Amplification products were run on a 1.2% agarose gel with a 100base pair molecular marker. All primers show strong amplification of gene from mouse DNA and no amplification from drosophila DNA.

Figure 3.13: Radioactive PCR analysis: Representative PAGE gel from P^{32}-PCR of C-ChIP samples



PAGE gel of Radioactive PCR products obtained from C-ChIP samples. Unbound (UB) and Bound (B) samples of asynchronous cells, cells at G1, S and G2/M were amplified by PCR incorporating P^{32} into the product. Aliquots of samples were taken at 39 cycles and then again at 42 cycles and run on 5% PAGE gels. This representative gel was from radioactive PCR run with primers for Ccnd1. Gels were dried onto filter paper and exposed on to a phosphor screen overnight before scanning with a phosphoimager (Biorad). Controls were run on a separate gel.

population, less than one indicates a depletion in the mark and more than one indicates an enrichment of the mark.

3.3.4 Enrichment of histone modifications in G1 phase associated genes

C-ChIP and Radioactive P^{32} PCR were used to compare the relative levels of the active marks H3K4me3, H3K9ac and H4K8ac, and the silent marks H3K27me3 and H3K9me2 in FACs sorted cells at G1, S and G2/M phase of the cell cycle. Histone modification enrichment was compared at the promoters of the G1/S associated genes *Skp2* and *Ccnd1* (Figure 3.14).

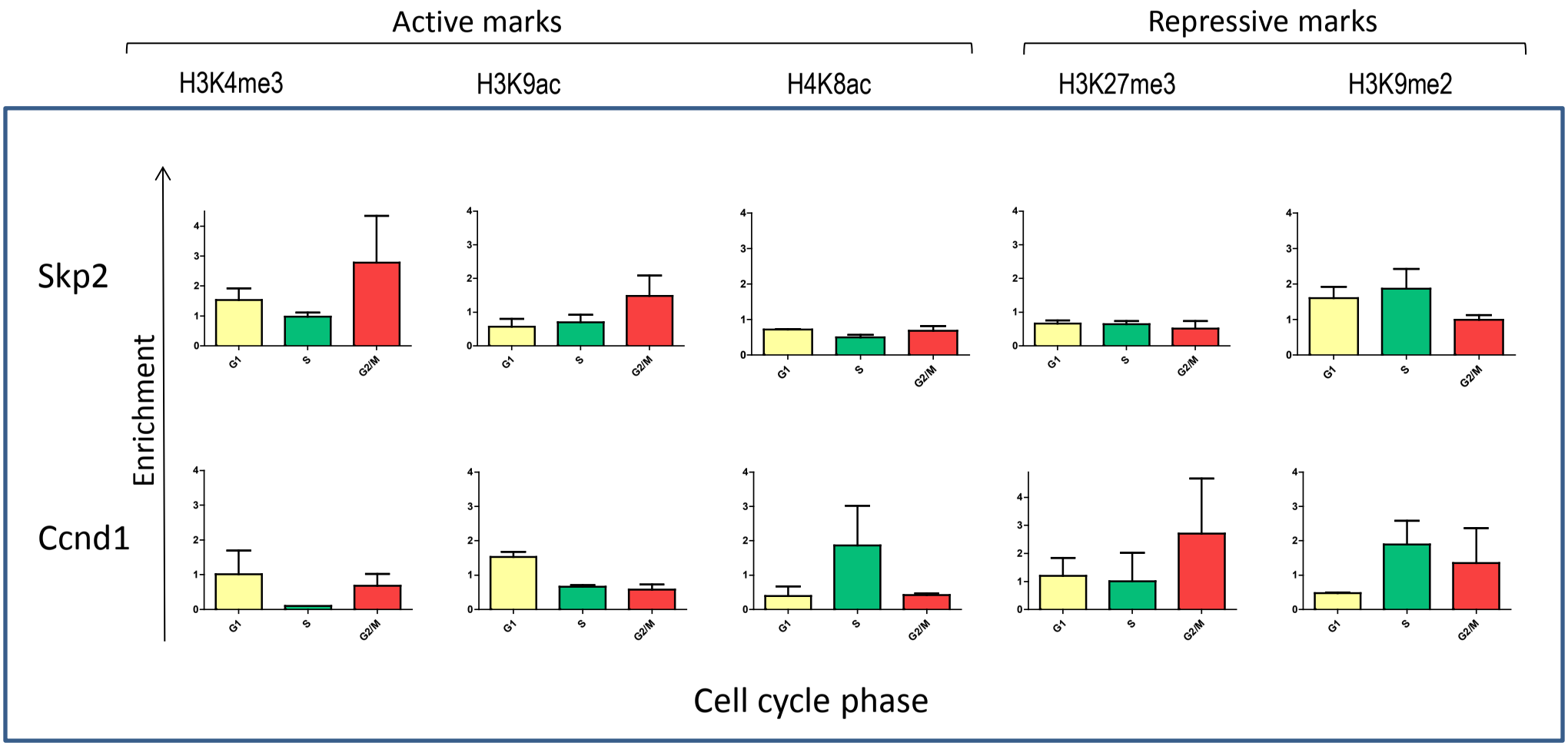
It is important to remember that when using chemicals to obtain samples at specific parts of the cell cycle; cells were arrested at the G1/S phase checkpoint and M phase, whereas populations of cells achieved by FACs sorting were at G1, S-phase and a combination of G2 and M phases.

Skp2 is required for the G1/S transition. It is expressed during G1 phase of the cell cycle (Lin & Diehl 2004; Zhou et al. 2013). At the *Skp2* promoter H3K4me3 is only slightly enriched at G1 compared to S phase, but is followed by a nearly fourfold increase at G2/M. H3K9ac varies very little between G1 and S phase but also has a marked increase at G2/M. This is the same pattern observed in the chemically arrested cells. However, in the initial experiments the H3K9ac mark appears to be the more prominent mark. The other active mark H4K8ac appears stable across the cell cycle - which differs from that seen in chemically arrested cells where levels of H4K8ac were elevated at M phase. However, should histone modifications prove to be dynamic the enrichment at M-phase may not be observed as cells at M phase will only make up a small proportion in a G2/M phase population. The same is true for the repressive modification H3K27me3 which in colcemid treated cells showed enrichment at M phase whereas in the FACs sorted cells

Figure 3.14: Relative enrichment of a panel of histone marks in FACs sorted ES cells in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in OS25 mouse ES cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.14: Relative enrichment of a panel of histone marks in FACs sorted OS25 mouse ES cells in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation



the mark remained stable over all phases of the cell cycle. H3K9me2 on the other hand showed enrichment at G1 and S and decreased going into G2/M phase. Therefore this drop in H3K27me3 at the cell cycle stage before expression is the reverse of pre-emptive active marks H3K4me3 and H3K9ac which are elevated at G2/M phase.

Ccnd1 is also expressed at G1-phase and is essential for the G1/S transition (Dalvai & Bystricky 2010; Lacomme et al. 2012). Levels of H3K4me3 are very low at the promoter of this gene at S phase when compared to G1 and G2/M which do not vary a great deal from each other. H3K9ac is slightly enriched at G1 phase when compared to S and G2/M. H3K8ac on the other hand noticeably increases at S-phase of the cell cycle compared to G1 and G2/M which are comparable. The silent mark H3K27me3 is steady across G1 and S phase but enriched at G2/M. For H3K9me2 this enrichment begins earlier at S phase and carries through to G2/M. The pattern observed for *Ccnd1* are those that would be expected if active marks appeared at the moment of expression, in particular H3K9ac, and repressive marks are depleted at the point of expression – i.e. in G1 phase.

3.3.5 Enrichment of histone modifications in S phase associated genes

S-phase has been the focus of several investigations into the heritability of histone modifications. DNA replication occurs during S-phase and nucleosomes and any modifications need to be faithfully replaced and recreated on the old and newly synthesised strands of DNA. From the data so far it does not appear that modifications remain stable, it seems that marks vary based on the expression of genes in one way or another. This suggests that modifications not only require a mechanism of transmitting information from one strand to another but that they readily alter and adjust in a much more dynamic manner than previously believed. Looking at the pattern of histone

modifications in S-phase associated genes could present useful insights into such a mechanism.

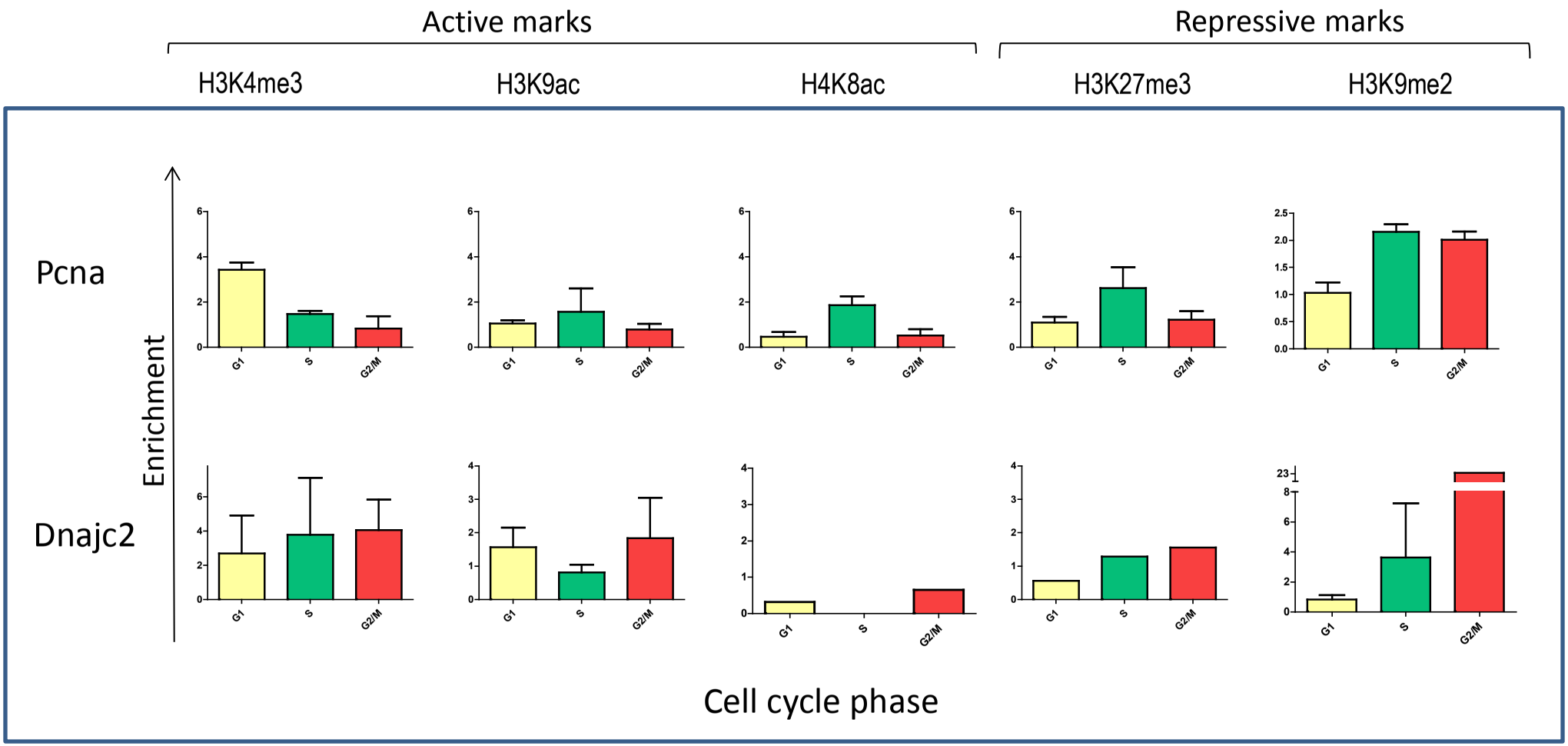
In Figure 3.15 the enrichment of active and repressive marks in *Pcna* (S-phase) can be seen. There is more than a twofold enrichment of H3K4me3 in G1 phase when compared to the rest of the cell cycle. This mark is very strong and is pre-emptive of gene expression. H3K9ac and H4K8ac both increase at S-phase; when the gene is expressed; however the degree of enrichment over the rest of the cell cycle is less than that observed for H3K4me3. Repressive marks for this gene are H3K27me3, which also peaks at S-phase, and H3K9me2 which has a marked decrease a G1 but remains enriched from S through to G2/M phase. *Pcna* appears to have a strong pre-emptive active mark, H3K4me3, and lower but appreciable active marks, H3K9ac and H4K8ac, at S-phase when the gene is expressed. Echoing this H3K27me3 is present to pre-emptively repress expression at S-phase, as is H3K9me2, with the later remaining through G2/M phase and finally falling away at G1- phase. The pattern observed is consistent with observations from the chemically arrested cells with the exception of H3K9ac, H3K8ac and H3K27me3. The difference seen in H3K8ac and H3K27me3 may be attributed to the fact that different phase of the cell cycle are being compared, which due to the nature of the overlaps, may obscure increases in histone marks at times. However H3K9ac appears to be elevated at M-phase in the chemically treated cells and at S-phase in the FACs sorted samples. This also may be the consequence of comparing G2/M to M-phase alone, it may be experimental error or could be the result of the chemical treatment with colcemid.

A different pattern is observed in *Dnajc2*. H3K4me3 appears mostly stable across the cell cycle. H3K9ac appears to be enriched at G1, depleted at S-phase where the gene is

Figure 3.15: Relative enrichment of a panel of histone marks in FACs sorted ES cells in S phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in OS25 mouse ES cells FACs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.15: Relative enrichment of a panel of histone marks in FACs sorted OS25 mouse ES cells in S phase associated genes as determined by carrier chromatin immunoprecipitation



expressed, but enriched again at G2/M. This mark appears enriched at all parts of the cell cycle that *Dnajc2* is not expressed in. H4K8ac appears to be depleted at G1 and G2/M-phases when compared to the Asynchronous control, however despite at least three technical repeats of three biological repeats no signal was observed for this mark at S-phase, suggesting that levels of the modification are below the detectable threshold even by this extremely sensitive technique. The active marks do not follow the pattern observed in *Pcna*, if anything H3K9ac and H4K8ac appear depleted at S-phase. The repressive marks on the other hand are both over two fold enriched from G1 to S-phase, H3K27me3 is maintained through to G2/M-phase whereas levels of H3K9me2 drastically increase again at G2/M-phase. The repressive modifications on *Dnajc2* are consistent with a pre-emptive silencing. Both marks increase from G1 to S-phase where the gene is expressed and the marks are maintained or increased in G2/M. This is consistent with repressive marks in the other S-phase expressed gene *Pcna*.

Little change in the repressive marks is observed in the chemically arrested cells, which could be attributed to the exact phases of the cell cycle looked at as previously postulated. Observations from the chemically arrested cells appears to correspond to these FACs sorted cells for active marks in the cells arrested at G1/S phase, however, the enrichment of marks at M-phase is again absent in the G2/M population. Colcemid functions by inhibiting the formation of spindle microtubules. During this artificially maintained mitosis, chromatin condensation continues which can lead to a 1-1.5x reduction in their length (Rieder & Palazzo 1992). The result of exposure to colcemid may be creating an artificial environment and altering chromatin including histone modifications. However only a small proportion of cells at G2/M are thought to be in M-phase. The combination of G2 and M may be masking any enrichment exclusive to M-phase, particularly if the mark is only enriched briefly. In order to investigate this

further it would be crucial to obtain a population of cells exclusively at M-phase without using colcemid to cause cell cycle arrest.

3.3.6 Enrichment of histone modifications in G2 phase associated gene Hurp

Hurp is expressed at G2/M-phase of the cell cycle (Tsou et al. 2003). At the *Hurp* promoter levels of H3K4me3 are enriched at S-phase compared to G1 and G2/M-phase (Figure 3.16). H3K9ac is also augmented at S-phase compared to G1, and is maintained in G2/M. These both follow the previously observed pattern of marks being enriched pre-emptive of their expression, with marks sometimes being maintained during the point of their expression. The final active mark H4K8ac also has a very slight enrichment at S-phase but is fairly constant across the cell cycle. Interestingly both of the repressive marks looked at are also enriched at S-phase and depleted at G2/M where expression of the gene occurs (potentially a bivalent mark). No band was ever obtained for H3K9me2 at G2/M despite a minimum of three technical repeats of three biological repeats strongly suggesting that the mark is below detectable levels. Active marks are enriched at S-phase as are repressive marks, however where the repressive marks fall away at G2/M-phase the active marks are maintained at the point where expression takes place.

In *Hurp* both active and repressive marks are pre-emptive of gene expression, enriched to a greater or lesser extent. This would support the idea of histone modifications playing an active role in the initiation of transcription, where marks are placed and subsequently genes are activated or silenced.

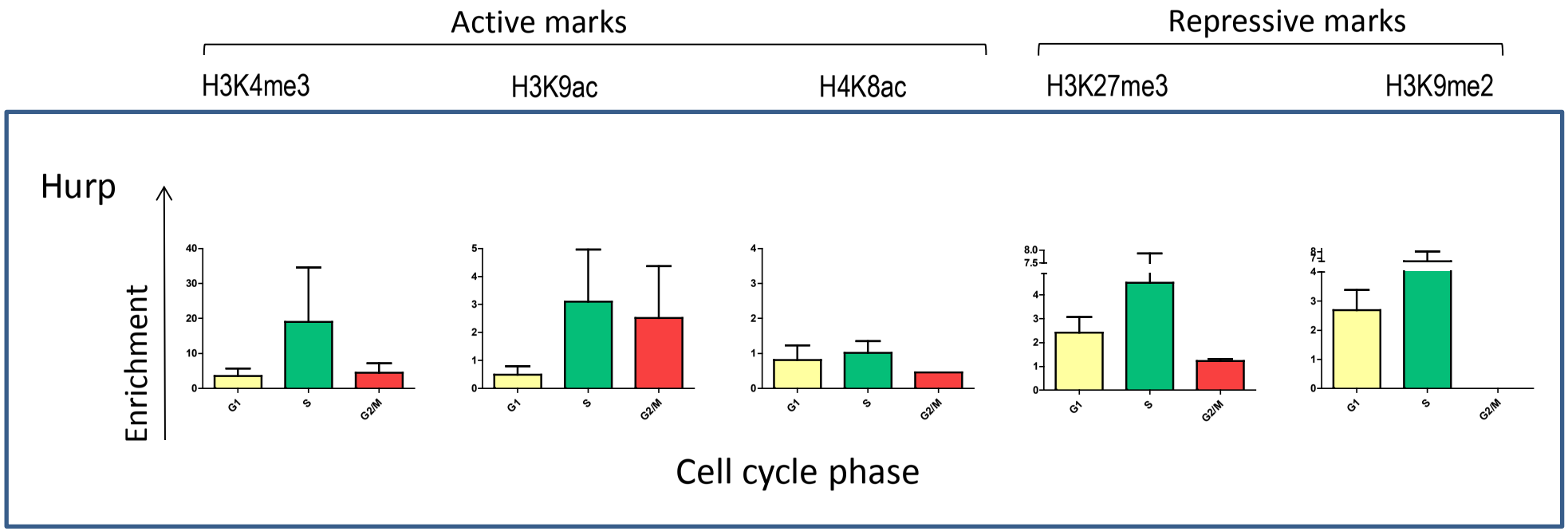
3.3.7 Enrichment of histone modifications in G2/M-phase associated genes

The M-phase associated genes looked at were *Ccna1* and *Cdc25b*. *Ccna1* is a member of

Figure 3.16: Relative enrichment of a panel of histone marks in FACs sorted ES cells in the G2 phase associated gene Hurp as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G2 phase associated gene Hurp as determined by carrier chromatin immunoprecipitation (C-ChIP) in OS25 mouse ES cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.16: Relative enrichment of a panel of histone marks in FACs sorted OS25 mouse ES cells in the G2 phase associated gene Hurp as determined by carrier chromatin immunoprecipitation



the highly conserved cyclin family which controls the progression of the cell cycle by activating the cyclin-dependent kinases (CDKs). It has a role in the G2/M transition. *Ccna1* is expressed throughout the cell cycle however it reaches its highest levels at late S and G2/M-phase (Rivera et al. 2006; Müller et al. 1999).

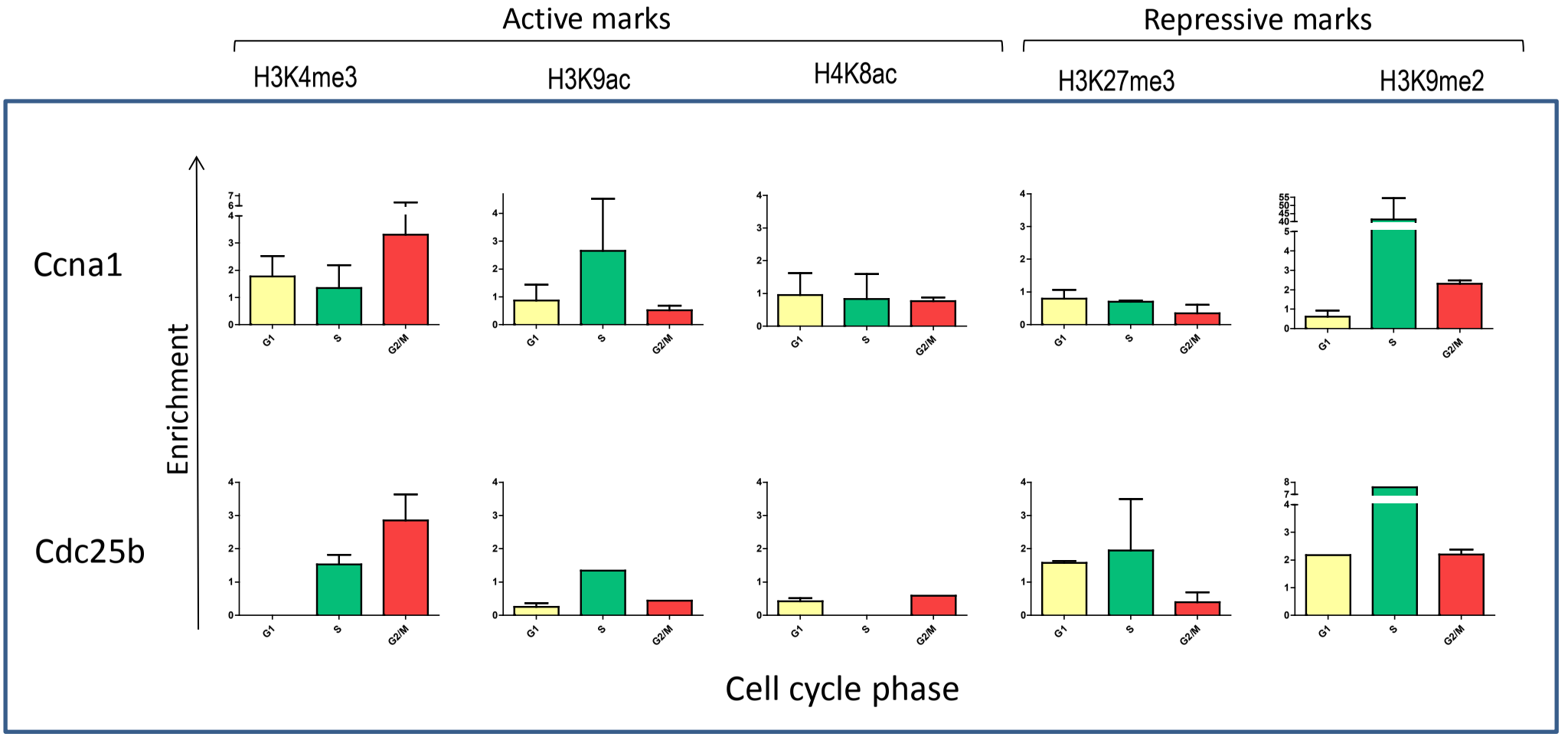
As can be seen in Figure 3.17 *Ccna1* appears to have a pre-emptive active mark, H3K9ac is enriched over twofold at S-phase compared to G1 and G2/M. In addition to this active mark, H3K4me3 appears to have cell cycle specific enrichment of an over threefold increase, this time at the point of expression G2/M-phase. The remaining active mark H4K8ac is consistent across the cell cycle. Of the silent marks H3K27me3 is around 50 percent depleted in G2/M, at the point of expression, whereas H3K9me2 is highly enriched at S-phase, which is mostly lost by M-phase and depleted at G1-phase. This would be consistent with a pre-emptive bivalent silencing mark that falls away as expression is increases. These marks are therefore pre-emptive and also, in the case of H3K4me3, reflective of expression.

Cdc25b is a member of the Cdc25 family of phosphatases and also controls entry into mitosis by activating a cyclin dependent kinase – Cdc2. Expression of *Cdc25b* is believed to peak at G2/M-phase of the cell cycle (Bugler et al. 2010). A dominant mark at the promoter of this gene is H3K4me3 which is strongly enriched at G2/M-phase. H3K9ac is slightly enriched at S phase of the cell cycle but levels of H4K8ac remain consistently low throughout and therefore presumably not involved in the regulation of this genes expression. Whereas the dominant repressive gene is H3K9me2 which is over a sevenfold enriched at S-phase compared to the G1 and G2/M-phases. H3K27me3 is enriched at G1 and S-phase followed by a depletion at G2/M, where expression takes place. This follows the previously observed pattern of an increase in both silent and

Figure 3.17: Relative enrichment of a panel of histone marks in FACs sorted ES cells in G2/M phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G2/M phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in OS25 mouse ES cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.17: Relative enrichment of a panel of histone marks in FACs sorted OS25 mouse ES cells in G2/M phase associated genes as determined by carrier chromatin immunoprecipitation



active marks pre-emptive of gene expression, with an active mark remaining enriched at the point of expression, in this gene H3K4me3 but this varies on gene to gene basis.

The pattern that emerges when looking at histone modification enrichment across the cell cycle in FACs sorted cells is that of an increase in both active and silent marks pre-emptive of gene expression. The repressive marks falls away at the point of expression but enriched active marks often remain enriched.

3.4 EPIGENETIC LANDSCAPE OF REPRESENTATIVE CELL CYCLE REGULATED GENES

The focus to this point has been on the levels of histone modifications at the promoter region of genes. Not only is transcription initiated at the promoter, but many previous studies have found that the H3K4me3, H3K27me3 and histone acetylations are predominantly located around the promoter region. Other marks such as H3K9me2 and H3S10phos are thought to be more prevalent within genic regions. This section aims to look at this particular panel of modifications, not only at the promoter, but within the context of the gene, in doing so it will be possible to observe whether changes are promoter specific as expected (associated with transcriptional regulation) or otherwise. Six histone modifications have been looked at in four representative cell cycle regulated genes. The vast majority of studies into histone modifications, including those determining the predominant location of marks, have been conducted on asynchronous populations of cells, looking at modifications across the cell cycle and across the genes has not been done before and may reveal previously unknown cell cycle specificity. OS25s were chemically arrested at G1/S phase and M-phase through the use of thymidine and colcemid as previously described. These were again analysed by N-ChIP and qPCR in order to determine the relative levels of histone modification in the bound

samples compared to the unbound samples and normalised to the asynchronous control.

3.4.1 Enrichment of histone modifications across the Ccnd1 gene

Ccnd1 is a G1/S specific cyclin, in mice it is found on the antisense strand of chromosome 7 at a length of 9995bps. No known genes are found to overlap *Ccnd1* at any point. Relative levels of H3K4me3, H3K27me3, H3K9ac, H3K9me2, H4K8ac and H3S10phos were determined in G1/S and M-phase arrested ES cells by N-ChIP across *Ccnd1* (Figure 3.18).

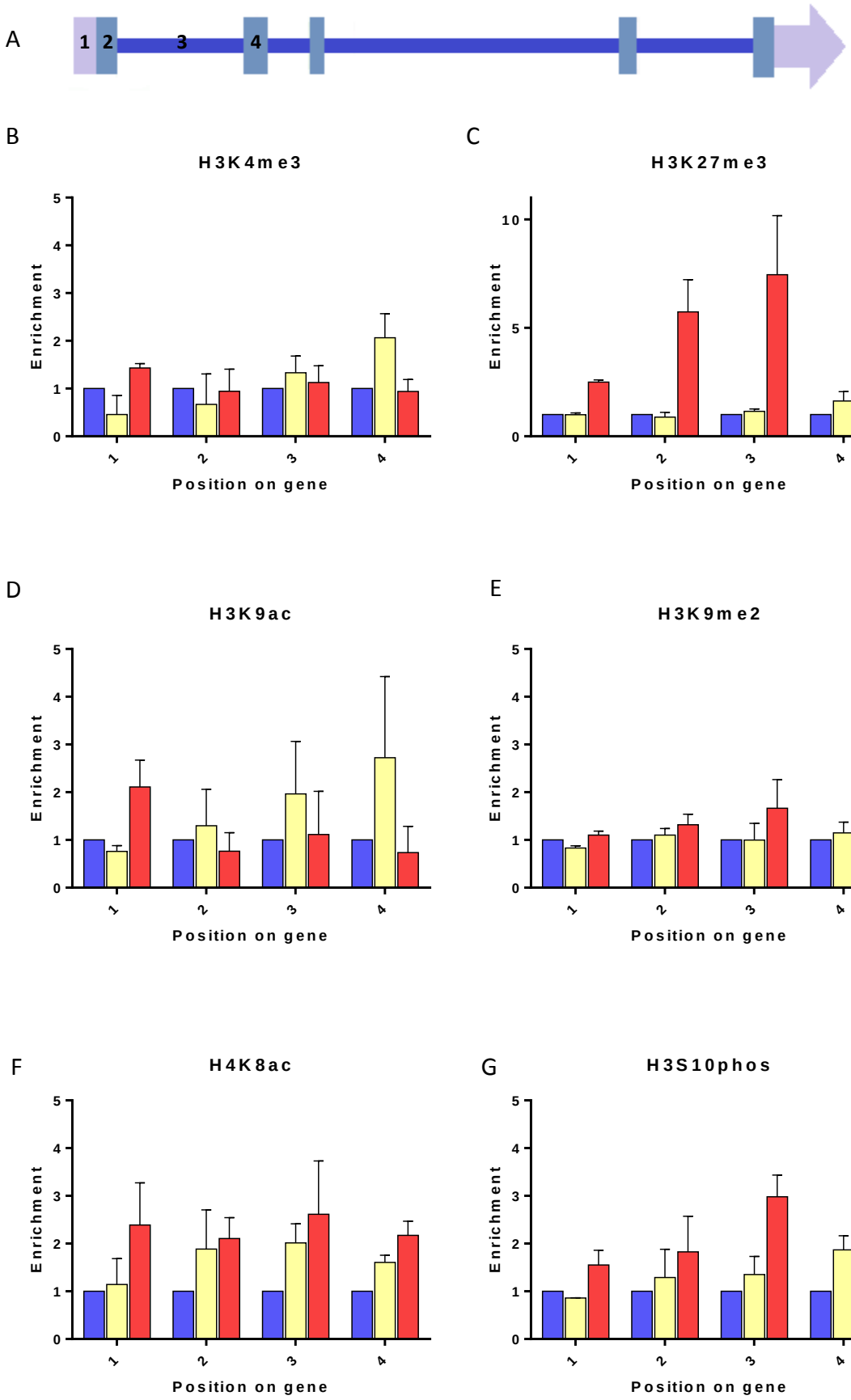
At the *Ccnd1* promoter (position one) levels of H3K4me3 are enriched at M-phase, this drops to around one at the first exon (position two) and remains there in the first intron and second exon (positions three and four respectively). Interestingly levels of the mark are depleted at G1/S phase at the promoter however they increase slightly by the first exon and again at the first intron and second exon, with levels of the marks at G1/S approximately two fold enriched by the second exon. The same is true for H3K9ac. At M-phase H3K9ac is enriched at the promoter, this drops to a relative level of approximately one by the first exon and remains there through the first intron and second exon. On the other hand the mark is depleted at G1/S at the promoter but steadily increases in the G1/S samples through the first exon, first intron and second exon in which it is enriched over two fold.

At the promoter of the *Ccnd1* gene H4K8ac appears in much the same way as it does for the previous two active marks; enriched at M-phase. However, differences appear when looking further across the gene, H4K8ac remains enriched at approximately the same

Figure 3.18: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G1/S expressed Ccnd1 gene

Relative levels of histone modifications across the G1/S expressed mouse gene Ccnd1 (cyclin D1) as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells. Cells were arrested at G1/S phase (yellow bars) by treating with thymidine and arrested at M phase (red bars) by treating with colcemid along side an asynchronous population (blue bars). **A**: Representation of the Ccnd1 gene including 5'UTR (purple), introns (narrow blue), exons (wide pale blue) and the 3'UTR. Immunoprecipitation performed with antibodies to **B**:H3K4me3, **C**:H3K9ac, **D**:H4K8ac, **E**:H3K27me3, **F**:H3K9me2 and **G**:H3S10phos. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.18: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G1/S expressed Ccnd1 gene



levels at M-phase and levels of the marks at G1/S become enriched to nearly the same degree.

The repressive mark H3K27me3 is extremely interesting, levels of the mark remain at the relative level of one at all positions considered for G1/S-phase samples, however in M-phase samples the mark which is enriched at the promoter becomes increasing enriched peaking in the first intron and beginning to drop again by the second exon, although it is still enriched at M-phase compared to G1/S. This suggests that H3K27me3 is most abundant in M-phase (comparatively) at the first exon and intron. The distribution of H3K27me3 in M-phase is reflective of a broader region of the mark as would be expected (Table 1.2) however the fact that this appears to be exclusive to M-phase is an important revelation. Levels of H3K9me2 which are stable across the cell cycle at the promoter remain that way through the gene, perhaps increasing slightly at M-phase through the first intron and second exon. Finally H3S10phos is enriched at M-phase at the promoter. This enrichment at M-phase increases, peaking at the first intron and beginning to drop again at the second exon. In G1/S phase samples a slight increase in enrichment is observed at the first exon, through the first intron and into the second exon, but levels remain below those at M-phase.

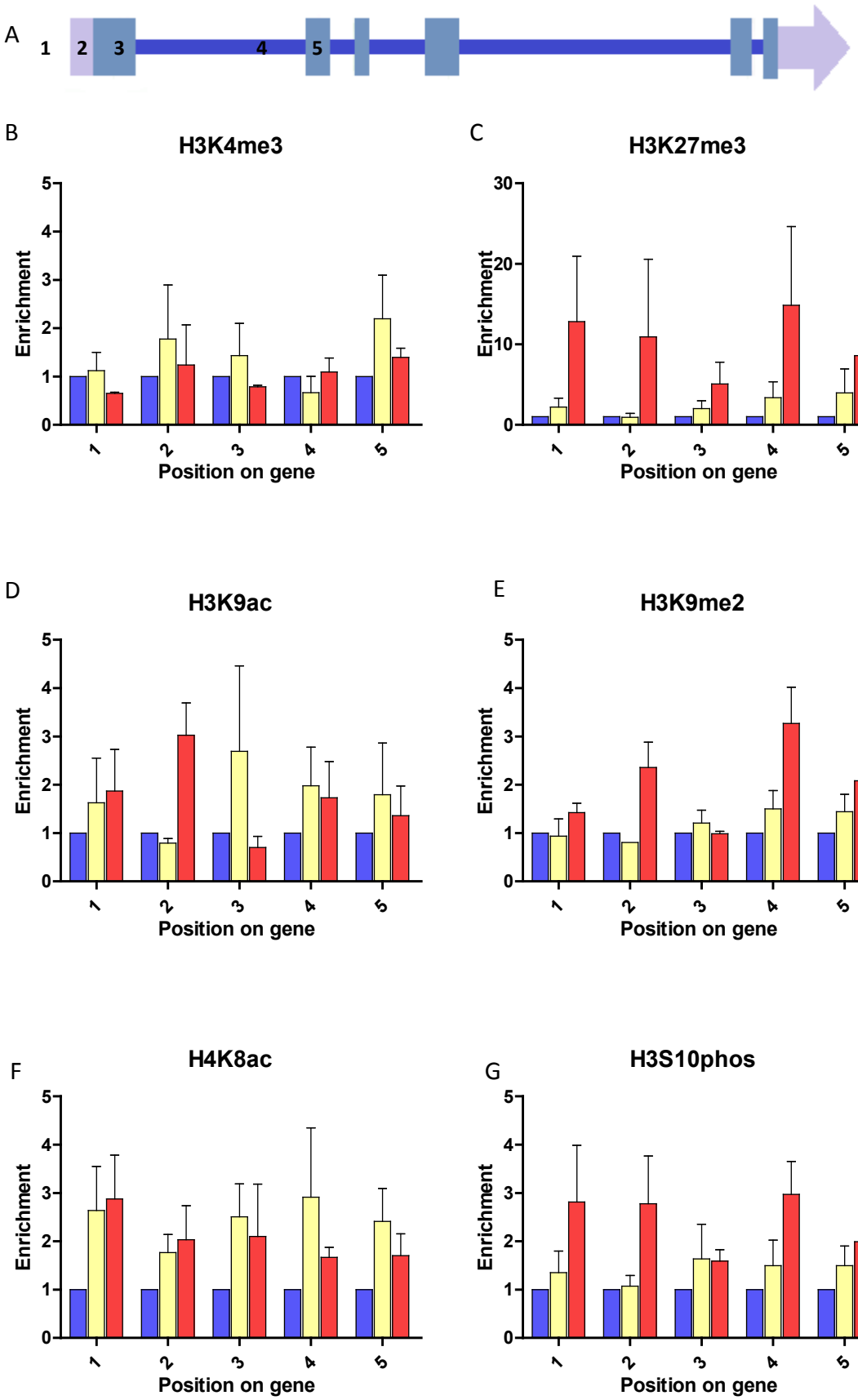
3.4.2 Enrichment of histone modifications across the Pcna gene

Pcna is an S-phase specific co-factor for DNA polymerase delta, in mice it is found on the antisense strand of chromosome 2 and is a short gene at only 3895bps. It does not overlap any known genes. Relative levels of the same panel of histone marks as in the previous section were considered across the *Pcna* gene (Figure 3.19). At the promoter of *Pcna* (position two) levels of H3K4me3 are slightly enriched at G1/S-phase compared to M-phase. This is also observed to a lesser degree just prior to the promoter at

Figure 3.19: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the S expressed Pcna gene

Relative levels of histone modifications across the S expressed mouse gene PCNA (proliferating cell nuclear antigen) as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells. Cells were arrested at G1/S phase (yellow bars) by treating with thymidine and arrested at M phase (red bars) by treating with colcemid along side an asynchronous population (blue bars). **A**: Representation of the Pcna gene including 5'UTR (purple), introns (narrow blue), exons (wide pale blue) and the 3'UTR. Immunoprecipitation performed with antibodies to **B**:H3K4me3, **C**:H3K9ac, **D**:H4K8ac, **E**:H3K27me3, **F**:H3K9me2 and **G**:H3S10phos. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.19: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the S expressed Pcna gene



position one and just afterwards in the first exon (position three). In G1/S-phase arrested cells at the first intron (position 4) levels of the mark drop off and are slightly depleted however at the second exon (position five) they are again enriched. Levels of H3K4me3 at M-phase appear to remain at roughly the same levels as observed in the asynchronous population.

The active mark H3K9ac is enriched at both G1/S and M-phase compared to the asynchronous population just upstream of the promoter, however, levels of the mark at the promoter are not enriched at all in G1/S and around threefold enriched at M-phase. At the first exon this is reversed, the mark is enriched at G1/S but not M-phase. By the first intron both of the marks are enriched but to a lesser extent, this is again seen at the second exon. H3K9ac is only enriched at M-phase at the promoter where it is pre-emptive of gene expression. The H4K8ac mark appears to be enriched at both G1/S and M-phase in all positions (upstream of the promoter through to the second exon) suggesting that it is depleted between G1/S and M-phase, as would be consistent with a pre-emptive enrichment of the mark. However normalising to the asynchronous may highlighting any slight fluctuations on histone modification levels.

As in *Ccnd1* the repressive mark H3K27me3 is interesting in *Pcna*. Levels of the mark at G1/S are at their comparative lowest at the promoter gradually increasing slightly through the first exon to the second exon. At M-phase the mark is highly enriched at all positions, dipping slightly at the first exon compared to the other points, however it is still enriched at M-phase at this point.

H3K9me2 is not enriched at any of the observed points at G1/S. At M-phase it is slightly enriched upstream of the promoter, increasing to an over two fold enrichment at the promoter. At the first exon no enrichment is observed at either G1/S or M-phase.

However enrichment of the mark at M-phase peaks again at the first intron with an over three fold enrichment, this falls by the second exon in which H3K9me2 is only slightly enriched at M-phase compared to G1/S. Exactly the same is observed for the H3S10phos mark in *Pcna* as H3K9me2 with the exception of the point upstream of the promoter where the mark is more enriched at M-phase than was true of H3K9me2. A secondary peak at M-phase of H3K27me3, H3K9me2 and H3S10phos is observed at the first intron. This is consistent within the repressive marks which appear enriched at M-phase at more than just the promoter region. Of the active marks H3K9ac appears enriched at M-phase at only the promoter.

3.4.3 ENRICHMENT OF HISTONE MODIFICATIONS ACROSS THE HURP GENE

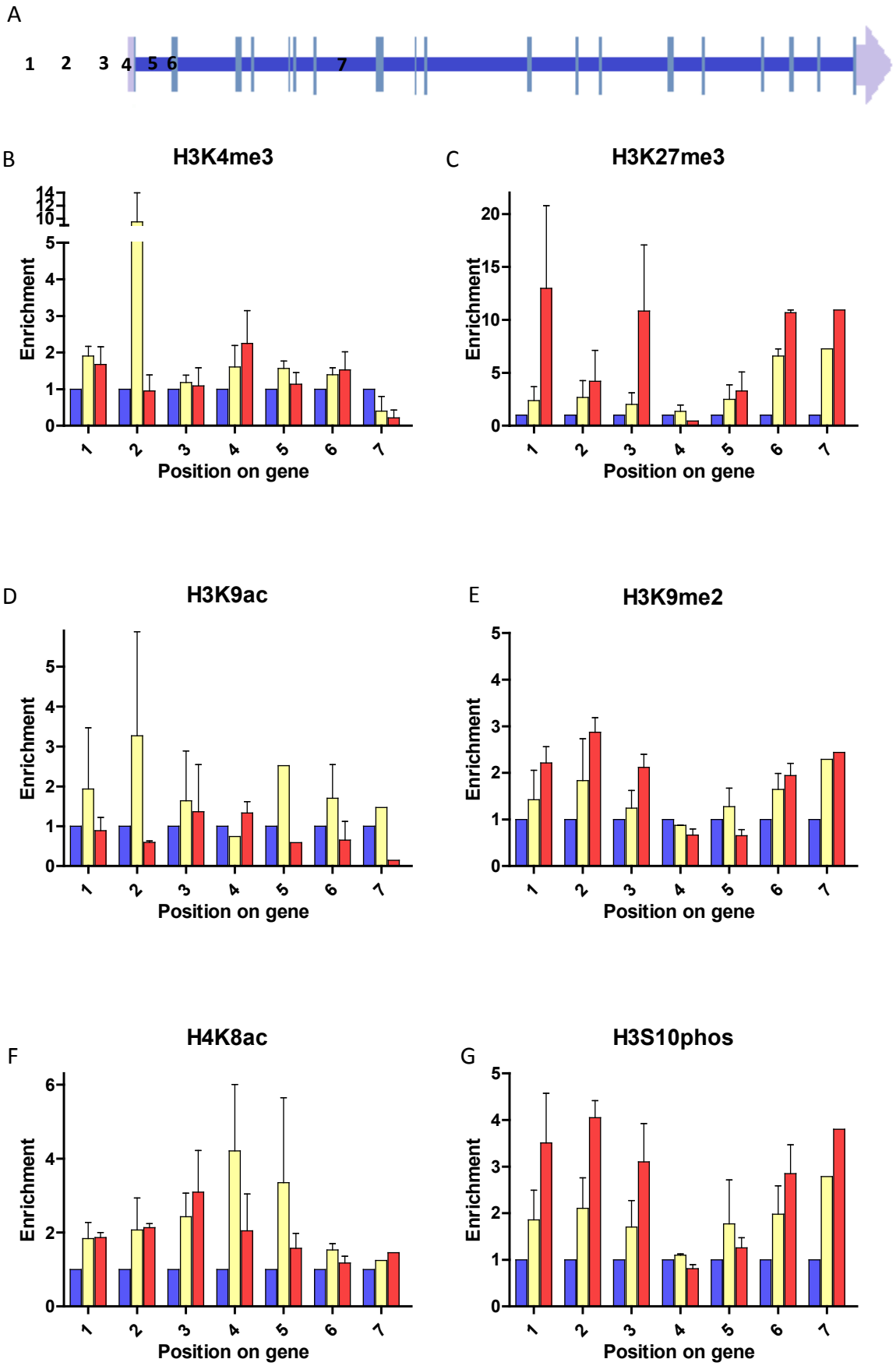
G2 expressed *Hurp* (otherwise known as *Dlgap5*) is located on the negative strand of chromosome 14. It is a large gene at a length of 30,629bps. The *Hurp* gene does not overlap with any other known genes however the 3' end of Lgals3 (lectin, galactose binding, soluble 3) on the sense strand is within 2000bps of the 3' end of *Hurp* and the 3' end of predicted gene 8353 (*Gm8353*) on the antisense strand is within 1500bps of the 5' end of *Hurp*. Relative levels of histone marks were determined in G1/S and M-phase arrested ES cells by N-ChIP. As *Hurp* is such a large gene, and is closely preceded by *Gm8353*, focus was placed on the area around the promoter (located at position four) (Figure 3.20).

Immediately upon looking at levels of H3K4me3 across *Hurp* a large enrichment of the mark is observed at G1/S-phase upstream of the promoter, at position two, this coincides with the promoter region of *Gm8353* and is likely to be related to that gene rather than *Hurp* itself. Apart from this large peak in H3K4me3 a smaller approximately two fold enrichment of the mark is found at the promoter region of *Hurp*. Levels of the

Figure 3.20: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G2 expressed Hurp gene

Relative levels of histone modifications across the G2 expressed mouse gene Hurp (dlgap5) as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells. Cells were arrested at G1/S phase (yellow bars) by treating with thymidine and arrested at M phase (red bars) by treating with colcemid along side an asynchronous population (blue bars). **A**: Representation of the Hurp gene including 5'UTR (purple), introns (narrow blue), exons (wide pale blue) and the 3'UTR. Immunoprecipitation performed with antibodies to **B**:H3K4me3, **C**:H3K9ac, **D**:H4K8ac, **E**:H3K27me3, **F**:H3K9me2 and **G**:H3S10phos. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates with the exception of position seven.

Figure 3.20: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G2 expressed Hurp gene



mark are stable across the cell cycle at the remaining points. A peak of H3K9ac is also observed at G1/S upstream of the promoter in position two. This peak at G1/S disappears entirely by the *Hurp* promoter where at very small peak in H3K9ac is observed at M-phase. Enrichment of the mark at M-phase disappears entirely at the first intron, through the second exon and across to the seventh intron, however enrichment at G1/S returns and persists through these points.

The H3K8ac mark appears to have a role at the *Hurp* promoter but does not appear to have a role in *Gm8353*. Levels of the mark at G1/S phase clearly rise through positions one, two and three upstream of the promoter, peaking at the promoter and dropping again through the first intron, the second exon and across to the seventh intron. Enrichment of the mark at M-phase also follows a gentler curve. Increasing through positions one, two and three upstream of the promoter (at the same levels at seen in G1/S phase) the enrichment begins to drop off at the promoter and the first intron, the second exon and across to the seventh intron.

The repressive mark H3K27me3 is not enriched at G1/S until the first intron where levels rise through the second exon and at the seventh intron. At M-phase, where the mark has been enriched in the previous two genes, we see a high degree of enrichment at position one upstream of the promoter, this mostly fall away at position two (presumable related to *Gm8353*), increases again at position three before becoming completely depleted at the promoter of *Hurp*. Levels of the mark then increase comparatively at M-phase in the first intron, the second exon and across to the seventh intron. It appears that levels of the mark are high at M-phase as in the previous two genes, but in this G2 expressed gene the mark is specifically depleted at the promoter region. This is also observed to a lesser extent for H3K9me2. H3K9me2 like H4K8ac

does not appear to be important for *Gm8353*. Levels of this mark are enriched at M-phase in positions one through to three upstream of the promoter, they are depleted at the promoter before rising again at the second exon and across at the seventh intron. At G1/S the mark is also slightly enriched from positions one, two and three upstream of the promoter, not at promoter but again enriched at the first intron, the second exon and across to the seventh intron. Exactly the same is observed for the H3S10phos mark. It appears that the repressive marks are present before and across the gene but not at the promoter region.

3.4.4 Enrichment of histone modifications across the Cdc25b gene

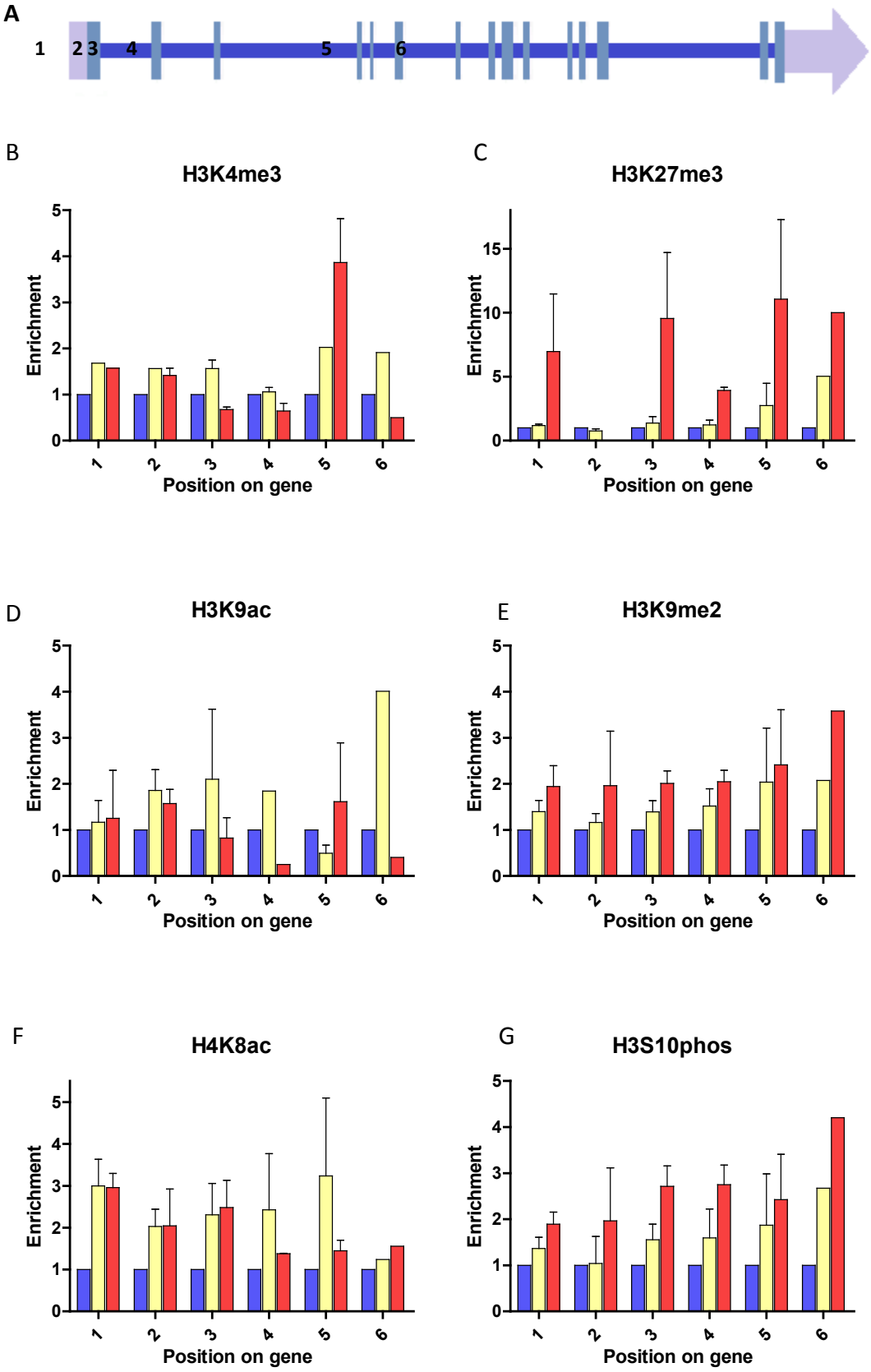
Cdc25b (M-phase inducer phosphatate 2) is required for entry to mitosis. In mice the gene is located on chromosome 2 and is 11,564bps long, it does not overlap with any known genes.

Here it appears that at the promoter (position 2) none of the active marks are elevated more at any particular point of the cell cycle when compared to the surrounding regions (figure 3.21). In all three active marks a slight enrichment is seen in both G1/S and M-phase cell populations when compared to the asynchronous population but this is also observed upstream of the promoter at position 1. This changes at the first exon (position 3) and the first intron (position 4) for H3K4me3 and H3K9ac, where levels of the marks remain enriched to the same degree at G1/S but become slightly depleted at M-phase. Counter to expectations H3K4me3 has a greater than 3 fold enrichment at M-phase within the third intron (position 5), found approximately a third of the way into the gene. Enrichment of H3K9ac at M-phase is also observed at this point. Both H3K4me3 and H3K9ac are slightly more enriched at G1/S-phase at all points except

Figure 3.21: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G2/M expressed Cdc25b gene

Relative levels of histone modifications across the G2/M expressed mouse gene Cdc25b (cell division cycle 25B) as determined by native chromatin immunoprecipitation (N-ChIP) in OS25 mouse ES cells. Cells were arrested at G1/S phase (yellow bars) by treating with thymidine and arrested at M phase (red bars) by treating with colcemid along side an asynchronous population (blue bars). **A**: Representation of the Cdc25B gene including 5'UTR (purple), introns (narrow blue), exons (wide pale blue) and the 3'UTR. Immunoprecipitation performed with antibodies to **B**:H3K4me3, **C**:H3K9ac, **D**:H4K8ac, **E**:H3K27me3, **F**:H3K9me2 and **G**:H3S10phos. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates with the exception of position 6.

Figure 3.21: Relative enrichment of a panel of histone marks in G1/S-phase and M-phase arrested ES cells across the G2/M expressed Cdc25b gene



within the third intron (position 5). Nothing is known of this region of the *Cdc25b* gene but it appears it may have an as yet unknown role.

In *Cdc25b* H4K8ac appears enriched at both G1/S and M immediately upstream of the promoter, at the promoter and the first exon. At the first intron the mark drops in M-phase to approximately one where it remains across the third intron and sixth exon (position 6). At the sixth exon H4K8ac also drops to one at G1/S phase.

Of the repressive marks H3K9me2 and H3S10phos both show very little enrichment at G1/S in prior to the promoter and at the promoter but they steadily increase across the gene to around a twofold enrichment at G1/S by exon 6. These marks are already enriched at M-phase at all observed points. Neither of these marks show any promoter specific enrichment, but are more abundant at M-phase across the entire gene. The same pattern of gradual increase across the gene at G1/S is true of H3K27me3 in keeping with its pattern of distribution (broader peaks – Table 1.2). It is also highly enriched at M-phase in all positions except the promoter, where no signal was observed despite at least three technical repeats being conducted on at least 3 biological repeats, suggesting that levels of H3K27me3 are below the detectable threshold at the promoter in M-phase. This is the only mark in *Cdc25b* which appears to have a promoter specific change and it is also the mark which varies the most across the cell cycle with upto a ten-fold enrichment at M-phase across the gene. Depletion of H3K27me3 at M-phase in the promoter is consistent with the expression of *Cdc25b* at G2/M.

3.5 Differentiation of ES cells into day 7 embryoid bodies

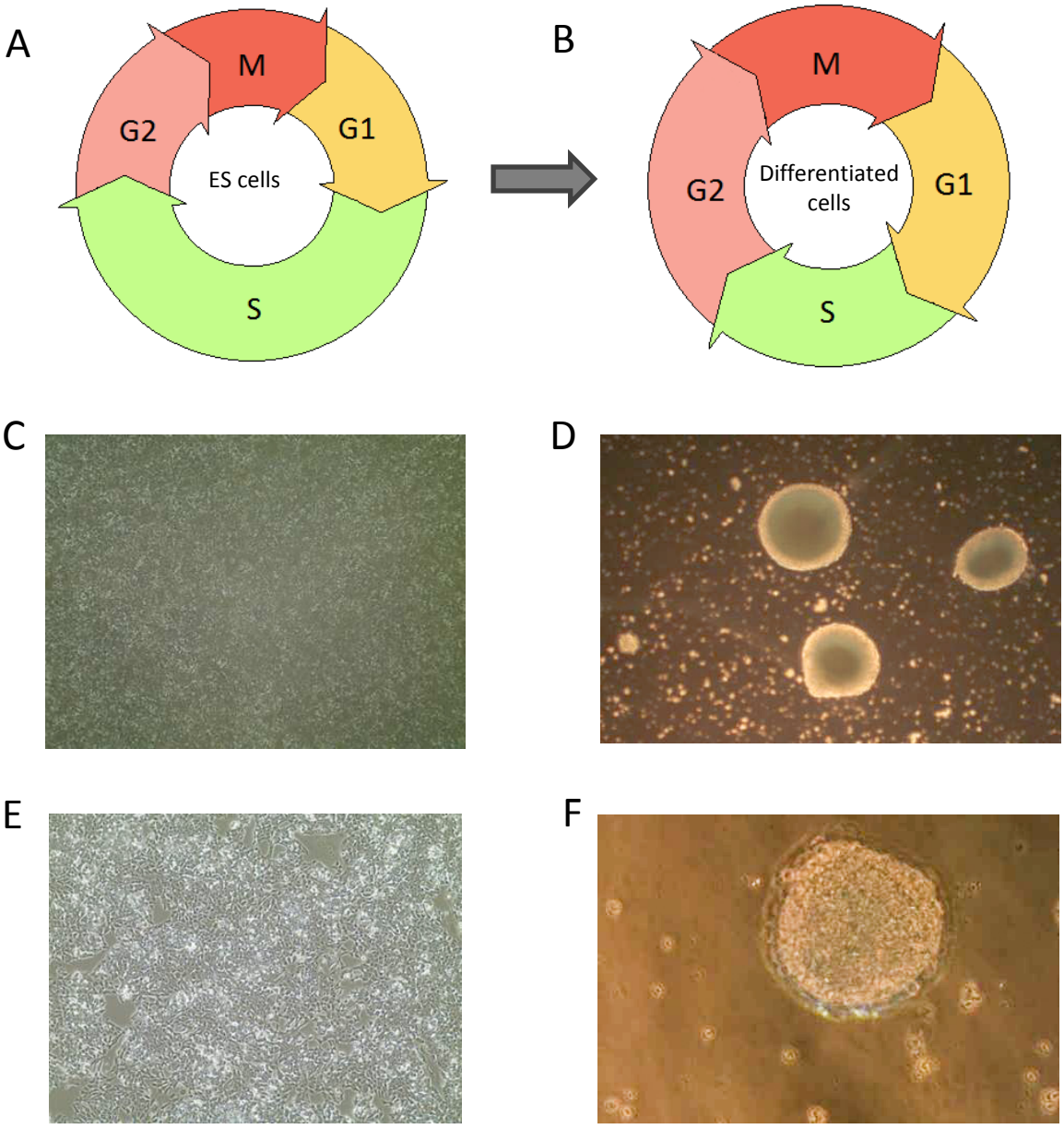
As ES cells differentiate they lose the characteristics that define them as stem cells. Upon differentiation they down regulate the expression of pluripotency related genes and up regulate the expression of more linage specific genes, reducing the plasticity of

the cells the further differentiated they become. Alongside this change is the introduction of normal cell cycle regulation including longer gap phases and more clearly defined cell cycle check points, particularly the enforcement of the G1/S checkpoint which is largely absent in ES cells (Figure 3.22 A & B)(White & Dalton 2005; Savatier et al. 1994; Orford & Scadden 2008). As a consequence differentiated cells are no longer able to proliferate as rapidly.

As can be seen in Figure 3.22, morphological changes are also observed as ES cells are differentiated. OS25s were differentiated by removing LIF from the culture medium, embryoid bodies were formed and directed down a neural pathway through the addition of retinoic acid on day 4. On day 7 cells were harvested by dissociating the embryoid bodies. The differentiating ES cells form free floating large round structures known as embryoid bodies (EBs) (Figure 3.22 D & F) rather than growing as a surface adherent mono layer (Figure 3.22 C & E). Within the embryoid bodies cells become very compact and develop large nuclei surrounded by very little cytoplasm.

Confirmation that a normal cell cycle has been established was determined from the propidium iodide cell cycle profile of differentiated cells. As can be seen in Figure 3.23 A and Figure 3.26, the day 7 differentiated cells have lost the ES like profile with the higher proportion of cells at S-phase and instead the profile retained is typical of somatic cells with a higher proportion of cells in G1.

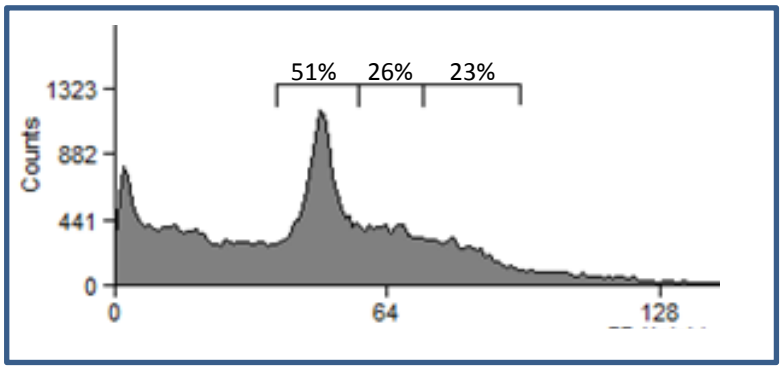
Figure 3.22: Differentiation of ES cells into day 7 embryoid bodies re-asserting normal cell cycle regulation



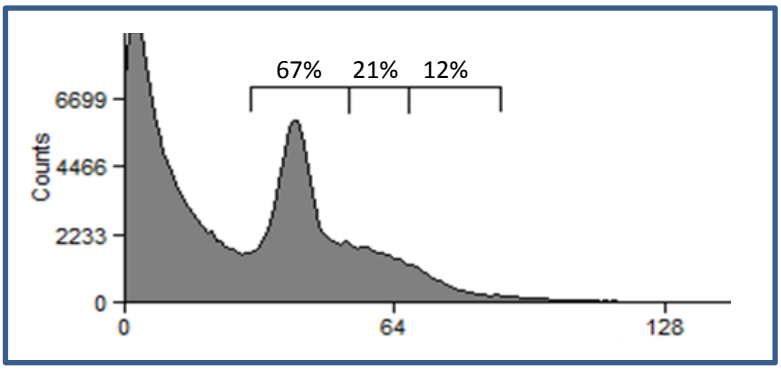
Differentiation of ES cells into day 7 embryoid bodies. **A**: Representation of an ES cell cycle with brief gap phases. **B**: Representation of cell cycle with normal somatic cell cycle regualtion. **C** & **E**: Photographs of OS25 cells. **D** & **F**: Photographs of day 7 embryoid bodies.

Figure 3.23: Propidium iodide cell cycle profiles of day 7 differentiated OS25s treated with thymidine and colcemid

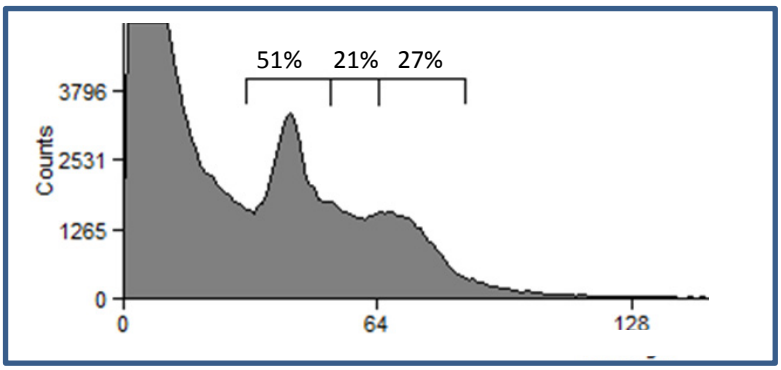
A



B



C



OS25s were differentited down a neural pathway for 7 days producing embryoid bodies which were then treated with thymidine or Colcemid to arrest cells at G1/S or M-phase of the cell cycle respectively before dissociating. Cells were made permeable with igepal and stained with propidium iodide. Flow cyctometry revealed the cell cycle profile of each sample. **A**: An asynchronous control polualtion with a normal cell cycle profile including the percentage of cells within the profile found at G1, S and G2/M-phase. **B**: Thymidine treated cells arrested at G1/S including the percentage of cells within the profile found at G1, S and G2/M-phase. **C**: Colcemid treated cells including the percentage of cells within the profile found at G1, S and G2/M-phase, no cell cycle effect is observed as a result of colcemid treatment but a large amount of cell debrie is observed. These cells were discarded at this point.

3.6 ENRICHMENT OF HISTONE MODIFICATIONS ACROSS THE CELL CYCLE IN CHEMICALLY ARRESTED DIFFERENTIATED ES CELLS

Alongside the change in cell cycle profile as the cells differentiate histone modification patterns change when compared to somatic cells, most notably in the case of bivalency. By looking at the enrichment of histone modifications across the cell cycle in day 7 differentiated cells it will be possible to see whether the regulation of histone modifications also changes upon differentiation and in what way.

3.6.1 ARRESTING DIFFERENTIATED CELLS AT G1/S PHASE AND M PHASE OF THE CELL CYCLE

To compare histone modifications across the cell cycle in differentiated cells thymidine and colcemid were employed to achieve cell cycle arrest at G1/S-phase and M-phase respectively in populations of day 7 differentiated OS25s. Embryoid bodies were treated with thymidine overnight to cause G1/S-phase arrest, cells were again morphologically indistinguishable from the asynchronous population. To analyse the cell cycle profile of the thymidine treated differentiated cells, embryoid bodies were gently dissociated and cells were washed repeatedly in PBS, however a large amount of debris was always present. As seen in Figure 3.23B, thymidine treated cells were found mostly with in G1 or S phase of the cell cycle (88%). Treatment of cells with colcemid however, did not cause the majority of cells to arrest at M-phase. EBs treated with colcemid were much more fragile and began to fray around the edges. As seen in Figure 3.23C a much higher proportion of debris is observed in colcemid treated cells and very little increase is seen in the number of cells found at G2/M-phase. As treatment with colcemid does not successfully cause cells to arrest at G2/M-phase these cells were not taken any further in the experiment.

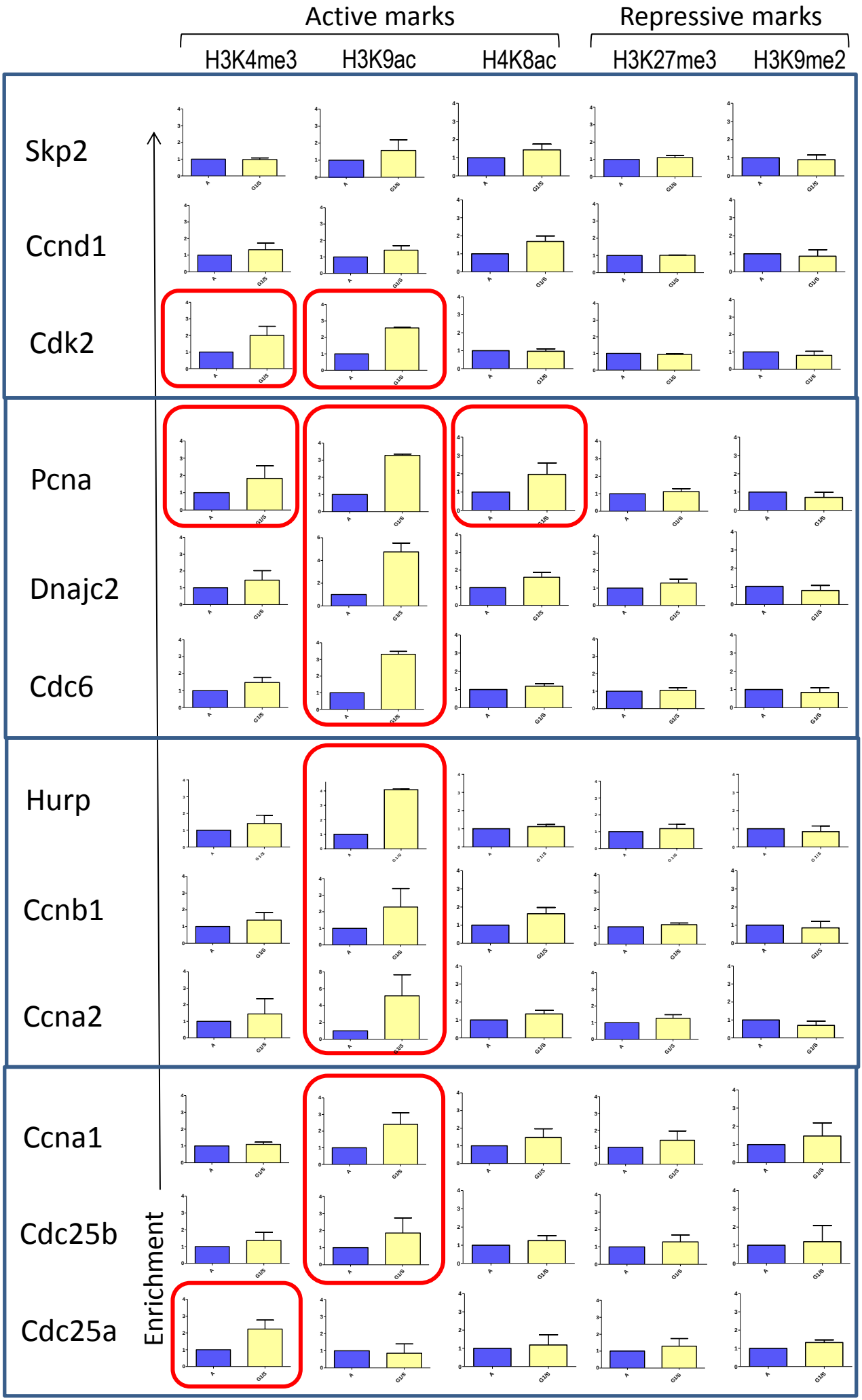
3.6.2 *Enrichment of post translational histone modifications in differentiated ES cells arrested at G1/S phase as observed in cell cycle regulated genes*

By comparing differentiated cells with undifferentiated ES cells it is possible to observe any changes that occurred as normal cell cycle regulation is enforced. Undifferentiated ES cells appear to have at least one active mark in place pre-emptive of gene expression, often accompanied by (a possibly bivalent) repressive mark. Bivalency has been mostly associated with stem cells, by using this approach it will be possible to observe whether the bivalency is also observed at day 7 of differentiation. In addition to this the pre-emptive nature of marks in ES cells may be related to rapid cell cycle turnover and the absence of the normal G1/S and G2/M checkpoints comparing to differentiated cells will assist in determining if this is the case. Native chromatin immunoprecipitation was used exactly as previously described to determine the relative levels of H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2 in the G1/S-phase arrested day 7 differentiated OS25s. Percentage pull downs of each antibody are shown in table 3.3.

The two repressive marks H3K27me3 and H3K9me2 do not change in G1/S when compared to the Asynchronous population across any of the genes looked at (Figure 3.24). This only differs from the undifferentiated cells in a couple of genes. H3K27me3 was slightly enriched in G1/S phase in *Skp2* and *Cdc25a*, which is not seen in the differentiated cells. The removal of the silent marks in these genes at G1/S may be the loss of bivalency, particularly as there is no variation in levels of these marks in G1/S compared to the Asynchronous population, however no conclusions can be made without comparing to the rest of the cell cycle.

Figure 3.24: Relative enrichment of a panel of histone marks in G1/S arrested- day 7 differentiated- OS25 ES cells in cell cycle regulated genes as determined by native chromatin immunoprecipitation

Relative levels of histone modifications in cell cycle regulated genes as determined by native chromatin immunoprecipitation (N-ChIP) in day 7 differentiated OS25 cells arrested at G1/S phase by treating with thymidine along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.


Active marks
Repressive marks
H3K4me3
H3K9ac
H4K8ac
H3K27me3
H3K9me2
Skp2
Ccnd1
Cdk2
Pcna
Dnajc2
Cdc6
Hurp
Ccnb1
Ccna2
Ccna1
Cdc25b
Cdc25a
Enrichment
A
G1/S

The changes observed within the active marks are very interesting, particularly those observed for the two acetyl marks. H3K9ac was not enriched at G1/S in any of the genes except *Cdc25a* in undifferentiated cells, whereas by day 7 this mark becomes enriched at G1/S phase in the G1 associated *Cdk2*, all of the S phase associated genes; *Pcna*, *Dnajc2*, *Cdc6*, the G2 associated genes; *Hurp*, *Ccnb1*, *Ccna2* and G2/M associated *Ccna1*. These marks are circled in red in Figure 3.24. The H4K8ac mark on the other hand is enriched in G1/S phase compared to the Asynchronous in undifferentiated ES cells in *Ccnd1*, *Cdk2*, *Pcna*, *Dnajc2*, *Cdc6*, *Hurp*, *Ccnb1*, *Ccna1*, *Cdc25a* and *Cdc25b*. This enrichment is not present in the majority of genes for differentiated cells, the only genes in which an enrichment of H4K8ac exists at G1/S phase are *Ccnd1*, *Pcna* and *Ccnb1*, and even in these it remains below a twofold increase. For the majority of genes the H4K8ac mark has fallen away at G1/S-phase in differentiated cells.

An enrichment of H3K4me3 in G1/S is only observed in *Cdk2*, *Pcna* and *Cdc25a*, of these the levels observed in *Pcna* and *Cdc25a* are comparable in both ES and day 7 differentiated cells. In *Cdk2* levels of H3K4me3 in G1/S have risen from no enrichment in ES cells to around two-fold enrichment in differentiated cells, this is also observed for H3K9ac.

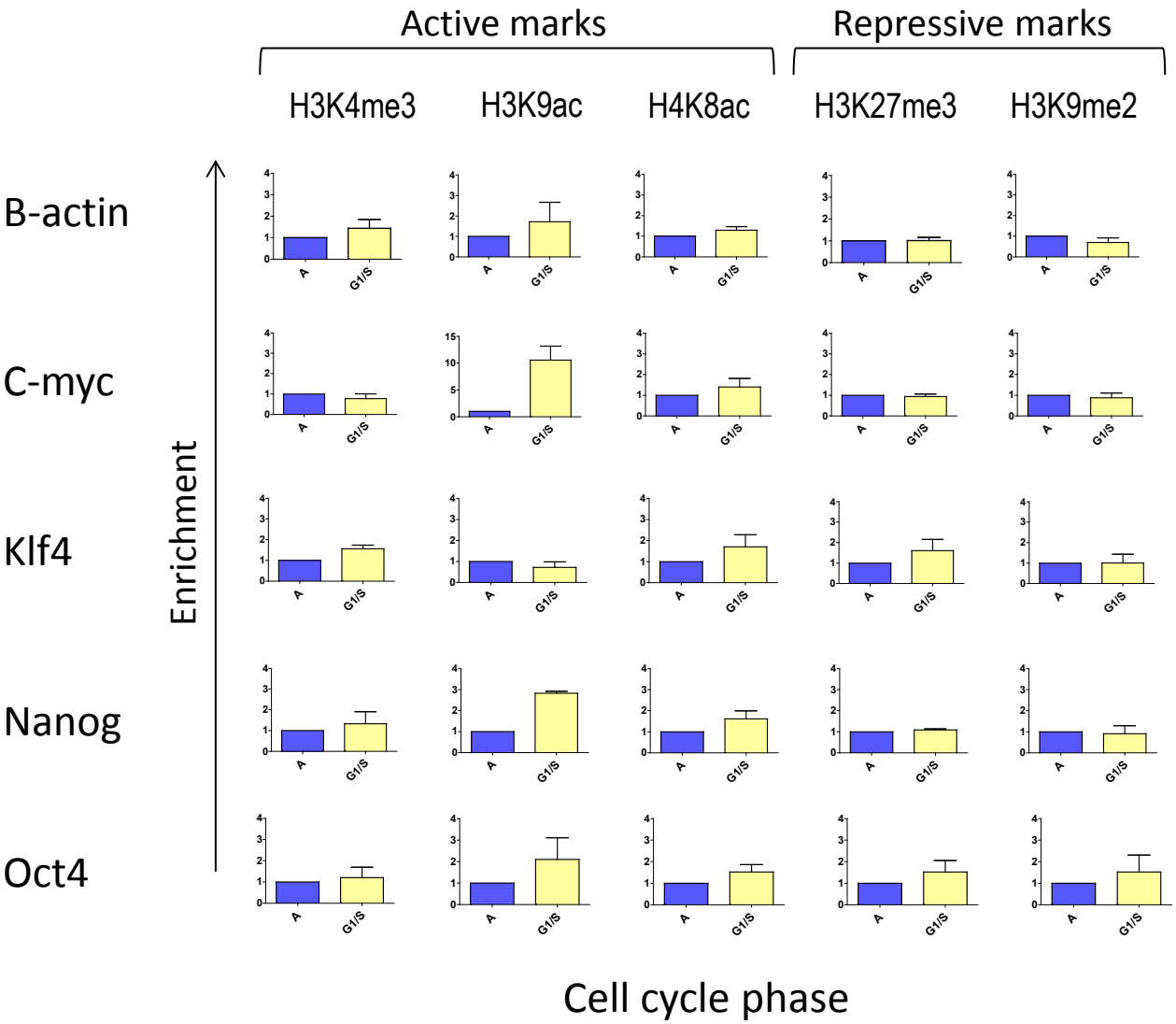
H3K9ac appears to be the predominant active mark at G1/S for differentiated cells in all of the S-phase and G2-phase associated genes, it is between twofold and fivefold enriched at G1/S for all these genes. The particular enrichment of this mark at G1/S phase for S and G2-phase genes appears pre-emptive of expression and seems to replaces the H4K8ac mark which is prevalent at G1/S in ES cells but not differentiated cells. For a clearer picture on these and the other cell cycle regulated genes it will be important to cover more of the cell cycle for firm conclusions to be drawn.

3.6.3 *Enrichment of histone modifications in differentiated ES cells arrested at G1/S phase and M phase as observed in genes associated with pluripotency*

Levels of histone modifications in the pluripotency genes *C-myc*, *Klf4*, *Nanog* and *Pou5f1* were also examined in day 7 differentiated OS25s along with the house keeping gene *β-actin* (Figure 3.25). Similar to the cell cycle regulated genes the repressive marks H3K27me3 and H3K9me2 are stable from the Asynchronous population to cells at G1/S phase in all genes except for *Klf4*. Levels of H3K27me3 in *Klf4* in ES cells are stable from the Asynchronous to G1/S-phase however in the differentiated cells an enrichment of just over one and a half fold is observed at G1/S. The active mark H4K8ac also changes from ES cells to differentiated cells, the large enrichment of the mark at G1/S –phase (over three fold increase) is lost and only a slight increase is evident in the day 7 differentiated cells. The increase in a repressive mark and decrease of an active mark indicate a silencing of the gene which is consistent with current understanding of *Klf4* expression as cells differentiate. *Klf4* prevents differentiation, therefore down regulation is required for it to be successful. Differentiation was induced by withdrawing LIF. As LIF functions via the STAT3 pathway acting on *Klf4*, it's expression is affected first, subsequently acting on other pluripotency factors including *Nanog* and *Pou5f1* (H Niwa et al. 1998; Kidder et al. 2008).

Contrary to this the remaining pluripotency related genes all have a strong enrichment of H3K9ac at G1/S compared to the Asynchronous population. All three genes *C-myc*, *Nanog* and *Pou5f1* had stable levels of H3K9ac in G1/S compared to the Asynchronous population in ES cells which increases over twofold upon differentiation. This is also observed in *β-actin*. The largest change of H3K9ac is at G1/S-phase in *C-myc* which is

Figure 3.25: Relative enrichment of a panel of histone marks in G1/S arrested- day 7 differentiated- OS25 ES cells in pluripotency associcated genes as determined by native chromatin immunoprecipitation



Relative levels of histone modifications in pluripotency associated genes as determined by native chromatin immunoprecipitation (N-ChIP) in day 7 differentiated OS25 cells arrested at G1/S phase by treating with thymidine along side an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA be quantiative real time PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

enriched over tenfold in the differentiated cells. A decrease in H3K27me3 is also observed for *C-myc* as cells are differentiated. Levels of the mark drop at G1/S from an approximate twofold increase to no enrichment at all.

H4K8ac is mostly stable in all genes in differentiated cells. The only differences from ES cells are *Klf4* as described and also *Nanog*, in which an enrichment is present at G1/S in undifferentiated ES cells but this falls away in the differentiated cells. H3K4me3 is not enriched in any gene at G1/S in either the ES cells or the day 7 differentiated cells. As with the cell cycle regulated genes a shift in active marks is observed from H4K8ac to H3K9ac with the exception of *Klf4* which appears to be silenced upon differentiation. It is likely that this is the first step in the down regulation of the pluripotency genes and analysis of cells further down a differentiation pathway would reveal subsequent silencing of other pluripotency genes. However in ES cell many of the marks appear to be most enriched at M-phase - without a greater sample of the cell cycle it is difficult to see the complete picture.

3.7 Enrichment of histone modifications in FACS sorted differentiated ES cells

The levels of histone modifications in the G1/S arrested day 7 differentiated cells are clearly different to those in the undifferentiated ES cells. However it is impossible to draw conclusions from the limited data. By FACs sorting cells into more distinct phases of the cell cycle it should be possible see a clearer picture of histone modifications across the cell cycle in differentiated cells and hence enable a better comparison to the undifferentiated ES cells.

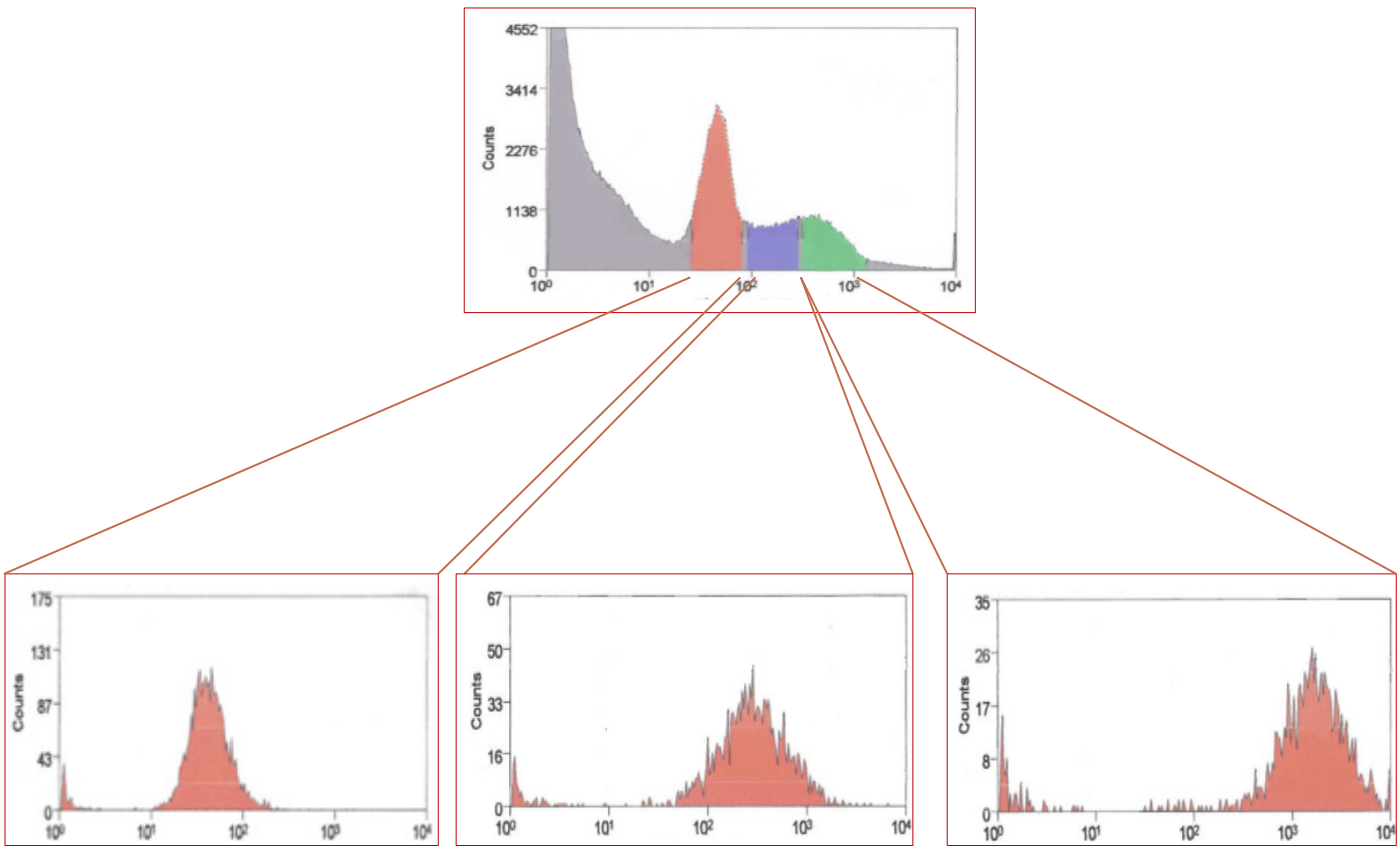
3.7.1 Sorting G1, S and G2/M phase differentiated cell populations by fluorescence-activated cell sorting (FACS)

In order to cover a greater spectrum of the cell cycle FACS was employed again, allowing separate G1, S and G2/M samples to be extracted from an asynchronous day 7 differentiated OS25 population. Dissociating embryoid bodies also results in a large amount of cell debris, which is also eliminated by gating the different section of the cell cycle as can be seen in Figure 3.26. The differentiated state of these cells can be seen in the propidium iodide profile, the G1 peak is more pronounced, fewer cells are found at S phase and the G2/M peak is also more defined. The profile is typical of a non ES/somatic cell cycle. Sorting cells allows the isolation of very pure populations of cells at G1, S and G2/M phase without cell debris. Sorted samples are analysed again to confirm their purity, In this representative sort of differentiated OS25s 92.63% of the G1 phase sample are at G1 phase, 89.71% are at S phase and 89.21% are at G2/M phase. Very low cell numbers in the region of 10^5 cells per purified sample of differentiated OS25s were obtained. As the number of cells obtained from each sort was so low samples were mixed with the carrier cells immediately upon sorting to minimise the loss of cells when washing to remove the propidium iodide. These cells were analysed by C-ChIP and Radioactive PCR as previously described. Percentage pull downs of each antibody are shown in table 3.4.

3.7.2 Enrichment of histone modifications in G1 phase associated genes

C-ChIP and Radioactive P^{32} PCR were used to compare the relative levels of the histone marks H3K4me3, H3K9ac, H3K27me3 and H3K9me2 in FACs sorted day 7 differentiated cells at G1, S and G2/M phases of the cell cycle in *Skp2* and *Ccnd1*(Figure 3.27). Before differentiation active marks in *Skp2* appear enriched at G2/M-phase

Figure 3.26: Florescence Activated Cell Sorting (FACS) of day 7 differentiated OS25 ES cells into G1 phase, S phase and G2/M phase of the cell cycle

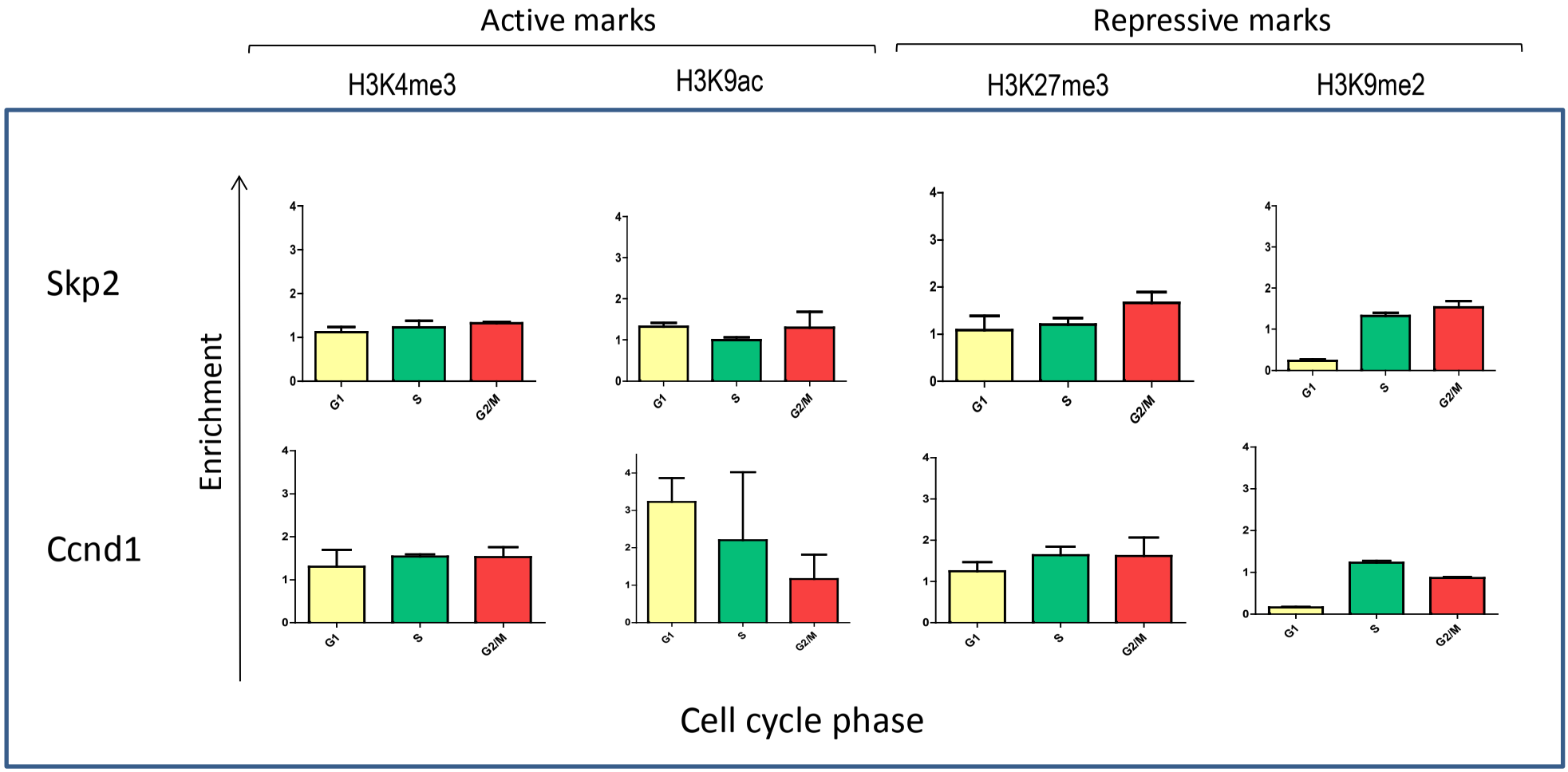


OS25 mouse embryonic stem cells were differentiated for 7 days to produce embryoid bodies before dissociating the embryoid bodies and fixing in acetone. Fixed cells were labelled with propidium iodide and cell sorted into G1 phase, S phase and G2/M phase of the cell cycle. The above is a representative cell sort, including typical gating. Purified populations were re-analyzed by flow cyctometry to confirm the purity of the samples (>88%). The numbers of cells obtained are in the region of 10^5 per purified sample.

Figure 3.27: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in day 7 differentiated OS25 cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.27: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation



particularly H3K4me3. This is not observed in the day 7 differentiated cells, H3K4me3 and H3K9ac are both stable across the cell cycle, G2/M enrichment has fallen away. Changes are also observed with the repressive marks upon differentiation. H3K27me3 is stable across the cell cycle in ES cells, whereas an enrichment of this mark over one and a half fold can be seen at G2/M-phase in differentiated cells. This is a reversal of what is observed in ES cells. Active marks appear not vary across the cell cycle in *Skp2* in differentiated cells whereas H3K27me3 is enriched pre-emptive of expression. These marks are all consistent with the observations from the data obtained by chemically arresting cells at G1/S.

Amongst the cell cycle regulated genes in differentiated cell H3K9me2 stands out in a way not observed in the ES cells. In *Skp2* H3K9me2 had been enriched at G1 and S-phase in ES cells, however in differentiated cells the mark is highly depleted at G1, and enriched approximately one and a half fold at S and G2/M-phase. Regardless of the pattern pre-differentiation and of the gene in question this pattern of depletion at G1 and enrichment of roughly equal proportions at S and G2/M-phase is always observed for H3K9me2.

For *Ccnd1* the bivalent marks H3K4me3 and H3K27me3 are enriched at G2/M-phase in ES cells, with H3K4me3 being retained through to G1 phase. Upon differentiation both of these marks become stable across the cell cycle and bivalency is lost. However H3K9ac, which is enriched at G1 in ES cells increases from a one and a half fold enrichment to an over threefold enrichment upon differentiation which is retained to a degree (approximately twofold) in S-phase. The elevated levels of the H3K9ac mark in G1 and to an extent S-phase are consistent with the expression of *Ccnd1*. H3K9me2 in this gene is consistent with the previously described pattern.

3.7.3 ENRICHMENT OF HISTONE MODIFICATIONS IN S PHASE ASSOCIATED GENES

Should marks be found to be pre-emptive in S-phase expressed genes an enrichment at G1-phase would be expected however this is not observed (Figure 3.28). In *Pcna* a large enrichment of H3K4me3 is observed at G1 in ES cells but once cells are differentiated this mark falls way. H3K4me3 is slightly enriched at S-phase and slightly depleted at G2/M reflecting the expression pattern of the gene. This is also true of the active mark H3K9ac, in both the ES and differentiated cells. H3K27me3 also appears to fall away upon differentiation, in ES cells the mark is highly enriched at S phase but levels remain mostly stable in differentiated cells with levels at S-phase dropping from a nearly threefold enrichment to a one and a third fold enrichment. As with the G1-phase expressed genes H3K4me3 and H3K27me3 bivalency appears to be lost upon differentiation. H3K9ac again appears to be the most significant mark, however the mark appears much stronger at G1/S in chemically arrested cells than S-phase in FACS cells.

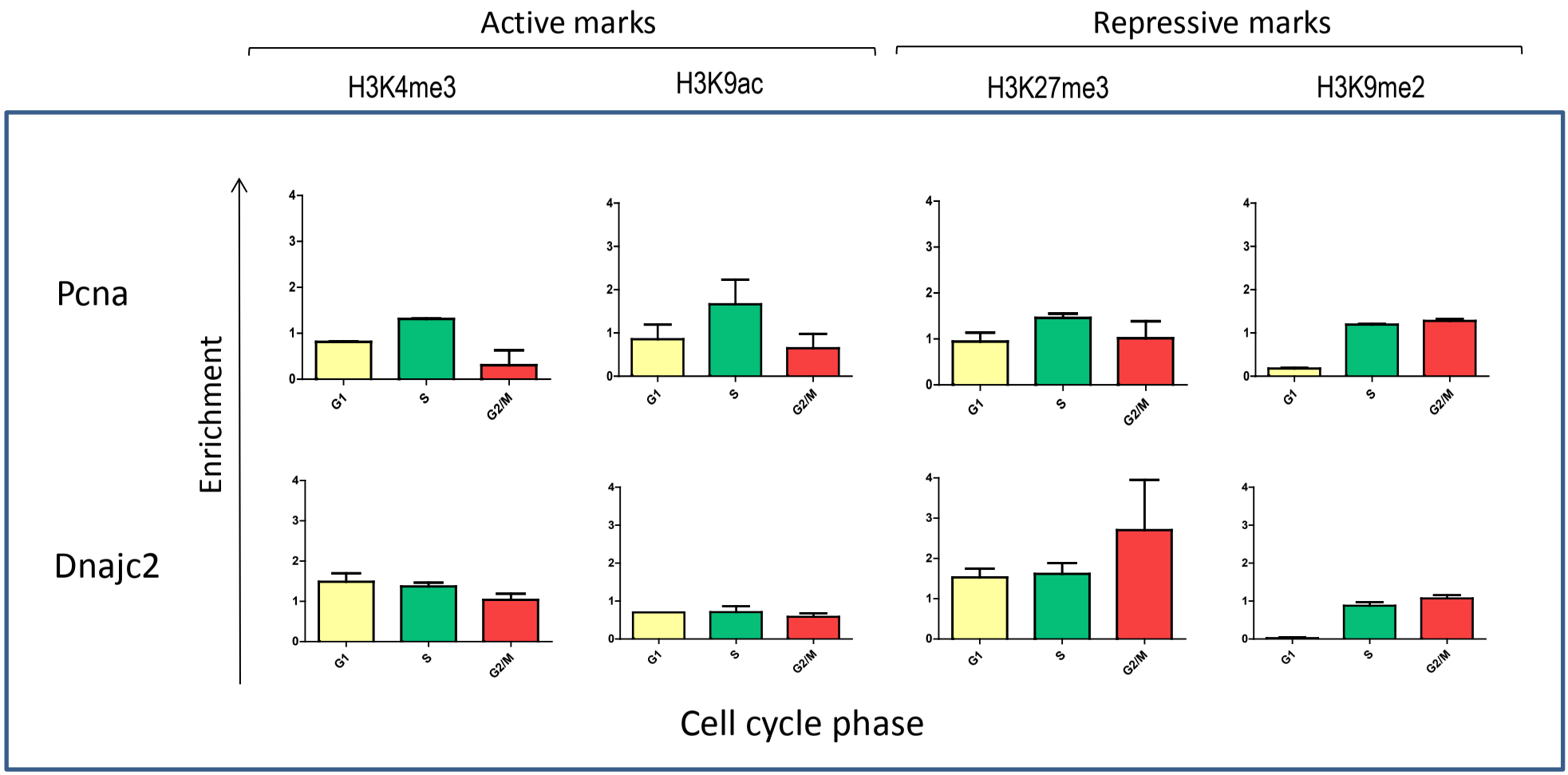
In *Dnajc2* a stabilization of H3K4me3 is also observed across the cell cycle. This is again found for H3K9ac in sorted cells, however it was not the case in the G1/S arrested cells which appeared to have a large enrichment of the mark at G1/S compared to the Asynchronous. Levels of the repressive mark H3K27me3 change from an enrichment at S and G2/M-phase in ES cells to just an enrichment at G2/M-phase in the differentiated cells. This is the phase of the cell cycle after the expression of *Dnajc2*. It appears that bivalency is lost in *Dnajc2* by loss of the H3K4me3 mark whereas the repressive H3K27me3 is maintained.

Levels of H3K9me2 in both *Pcna* and *Dnajc2* follow the pattern described previously, depletion of the mark at G1, followed by an equal enrichment at both S and G2/M-phase.

Figure 3.28: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in S phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in day 7 differentiated OS25 cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.28: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in S phase associated genes as determined by carrier chromatin immunoprecipitation



In both of these genes this pattern was present before differentiation, with a greater fold enrichment, levels in differentiated cells are comparable in all genes investigated.

3.7.4 Enrichment of histone modifications in the G2 phase associated gene Hurp

The enrichment of H3K4me3, H3K9ac, H3K27me3 and H3K9me2 across the cell cycle in the G2 phase associated gene *Hurp* are displayed clearly in Figure 3.29. In ES cells levels of the marks H3K4me3 and H3K27me3 in *Hurp* are highly enriched at S phase, pre-emptive of the genes expression. This bivalency is lost upon differentiation as H3K4me3 levels remain stable across the cell cycle in the day 7 differentiated cells. Whereas H3K27me3 is enriched at S and G2/M-phase. This appears to be a pre-emptive silencing of the gene which is maintained into G2/M-phase.

In *Hurp* H3K9ac is depleted at G1, S and G2/M when compared to the Asynchronous population, even more than it is in *Dnajc2*. Levels appear stable across the cell cycle in differentiated FACS cells, however there is a large enrichment of the mark at G1/S in cells chemically arrested at G1/S. This may be the result of experimental error as levels of the mark appear stable across the cell cycle in FACs sorted cells. It could also be the result of chemically inducing arrest at the G1/S checkpoint, however it is also a possibility that the mark is extremely dynamic and present at the G1/S checkpoint.

H3K9me2 follows the previously described pattern in *Hurp*. The mark is depleted at G1-phse followed by equal enrichment at S and G2/M-phase.

3.7.5 Enrichment of histone modifications in G2/M phase associated genes

Levels of histone modifications were monitored in the G2/M-phase expressed *Ccna1* and *Cdc25b* (Figure 3.30). H3K4me3 is found to be stable across the cell cycle in *Ccna1* in differentiated cells. In ES cells H3K9ac is found to be enriched at S-phase pre-emptive

Figure 3.29: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in the G2 phase associate gene Hurp as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in day 7 differentiated OS25 cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.29: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in the G2 phase associate gene Hurp as determined by carrier chromatin immunoprecipitation

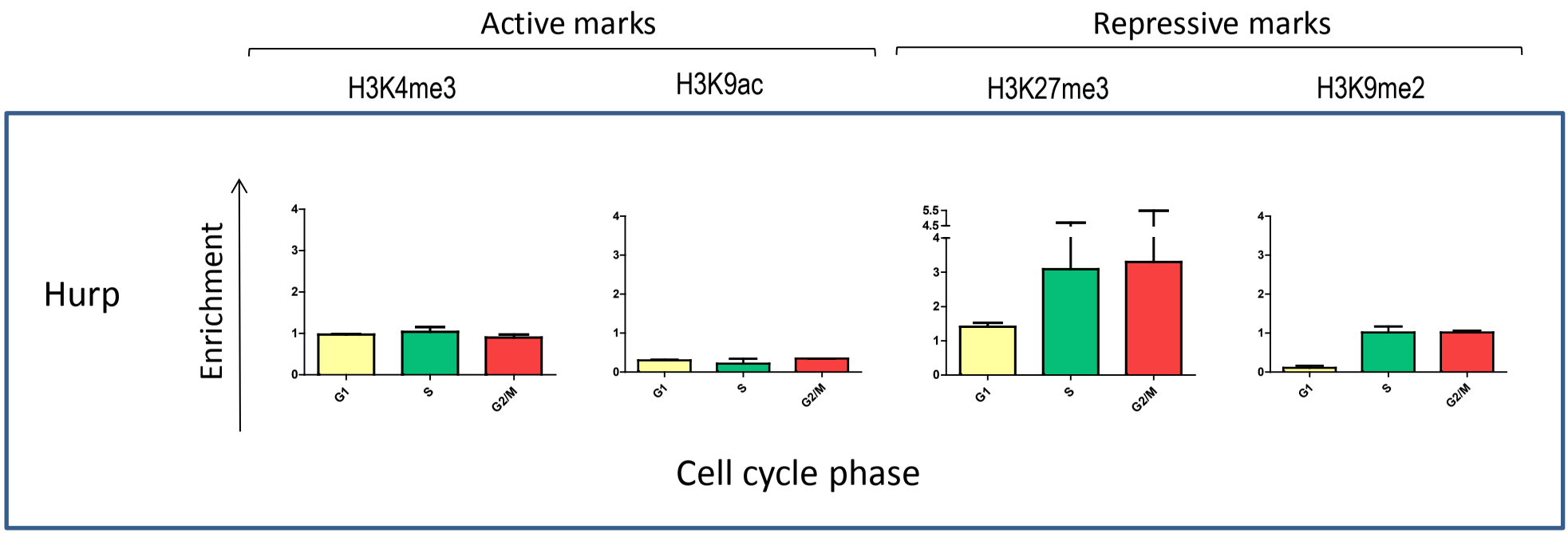
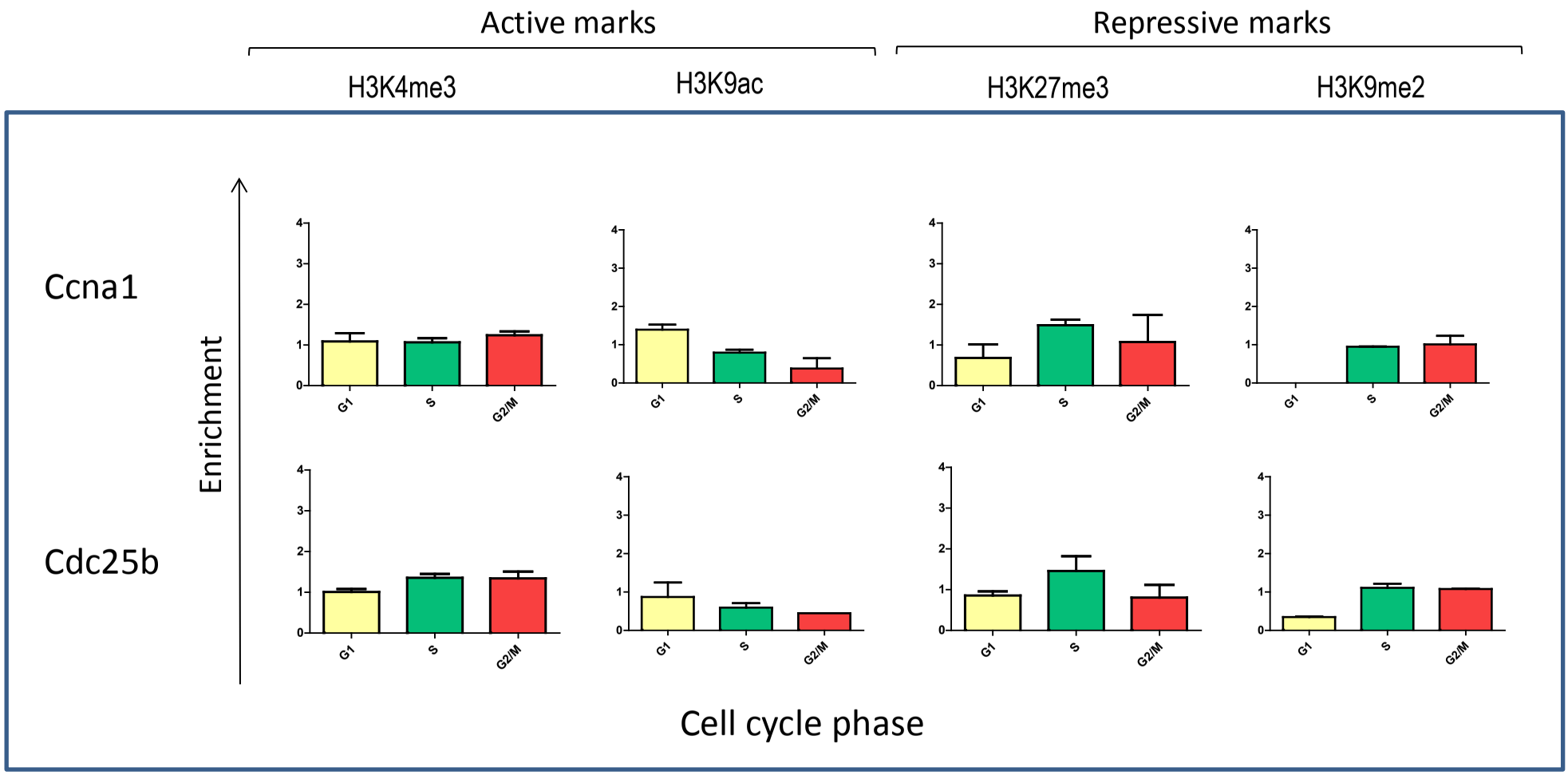


Figure 3.30: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in G2/M phase associated genes as determined by carrier chromatin immunoprecipitation

Relative levels of histone modifications in G1/S phase associated genes as determined by carrier chromatin immunoprecipitation (C-ChIP) in day 7 differentiated OS25 cells Facs sorted into G1, S and G2/M-phase populations. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2. Relative enrichment of modifications were calculated as the bound (immunoprecipitated) to unbound (unprecipitated) DNA by P^{32} radioactive PCR. These values were then normalised to an asynchronous population. Enrichment shows a fold change from the asynchronous population. Data are the mean ±SEM from at least two separate biological replicates.

Figure 3.30: Relative enrichment of a panel of histone marks in FACs sorted day 7 differentiated OS25 cells in G2/M phase associated genes as determined by carrier chromatin immunoprecipitation



of the genes expression, whereas upon differentiation levels of the mark peak at G1 and are depleted at G2/M-phase. An enrichment of the mark is also seen in the G1/S-phase arrested cells.

There is a very slight increase of the repressive mark H3K27me3 at S-phase pre-emptive of expression which is not observed in the undifferentiated ES cells where levels are mostly stable across the cell cycle.

Levels of H3K4me3 have also stabilizes across the cell cycle in *Cdc25b*, the enrichment found at G2/M in undifferentiated ES falling away. This is also observed for H3K27me3 which is enriched at G1 and S-phase, and depleted at G2/M in ES cells, however upon differentiation levels of the mark are comparable to those in *Ccna1*. H3K27me3 has a slight enrichment at S-phase pre-emptive of gene expression. Levels of H3K9ac very slightly decrease from G1 to S to G2/M in the FACs sorted cells, which is not the observation from the chemically arrested differentiated cells. Observations for H3K9me2 in both G2/M-phase expressed genes are the same as in all the other genes, suggesting that in differentiated cells at least this mark has a cell cycle dependent role that is not linked to the regulation of gene expression.

4 DISCUSSION

Cell identity relies on the careful and precise regulation of gene expression. As the primary component of DNA packaging, nucleosomes are ideally positioned to contribute to this regulation; in addition the highly conserved nature of the core histones suggests a vital role for these proteins. Research into the field of epigenetics and histone modifications has grown exponentially in the past decade, a vast volume of work has been carried out covering everything from the different varieties of post translational

modifications (now over 100 types) to the proteins involved in the deposition of histone marks and more recently the proteins that carry out the downstream effects of such modifications (Turner 2014). Many histone modifications have been intricately linked to the activation and repression of gene expression; these marks are believed to have causal roles in such mechanisms however they have not yet been fully elucidated.

This thesis adopts the entirely novel approach of focusing on the changes in histone modification levels across each stage of the cell cycle, narrowing in on the promoter regions of specific genes. In doing so this approach addresses the question of heritability from a new perspective. Should histone modifications remain at consistent levels throughout the cell cycle, particularly both DNA replication in S phase and chromatin condensation in M-phase, then it would be clear that the mark is stable and heritable (figure 4.1 A). It has long been believed that histone modifications are capable of exerting long term effects, maintained on chromatin throughout the cell cycle and onto the next generation. This would be the truly epigenetic role of histone modifications however it has not yet been verified.

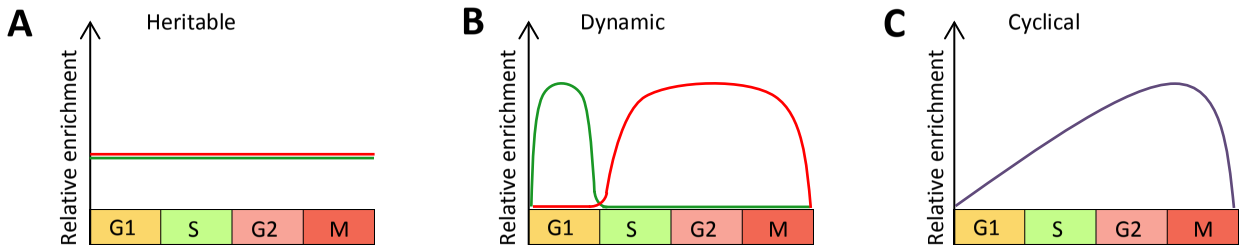
Alternatively variation across the cell cycle could be one of two things; highly dynamic expression dependent changes (for example an enrichment of active marks linked to an increase of gene expression –figure 4.1B), or a cyclical reproducible pattern of histone modifications across the cell cycle in a way that could still be heritable but will depend on the mechanism of histone modification placement rather than the simpler maintenance of marks from one generation to the next as was discussed in section 1.9 (Figure 4.1C). By focusing on cell cycle regulated genes it should be possible to distinguish between the latter two variable patterns of histone modifications.

Figure 4.1: Predicted and observed patterns of histone modifications in undifferentiated ES cells and day 7 embryoid bodies

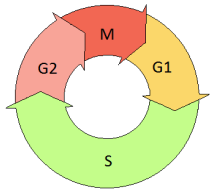
When looking at histone modification enrichment across the cell cycle several patterns may be observed. **A**. Modifications may remain stable across the cell cycle in which case they would be heritable. **B**. Histone modifications may change dynamically, potentially linked to gene expression. **C**. Or a cyclical pattern could exist in which a histone enrichment changes across the cell cycle in a transcription independent manner. In practice all of these were observed. In undifferentiated ES cells a pattern of dynamic regulation linked to gene expression was most prominent, with histone modifications either **D**. pre-emptive or **E**. reflective of gene expression. **F**. A spike in the levels of all histones was observed in cells arrested at M-phase which would be classified as a cyclical pattern. In differentiating ES cells (day 7 embryoid bodies) this changes dramatically. **G**. H3K4me3 and H3K27me3 became stable across the cell cycle and as such are considered heritable. **H**. H3K9ac remains dynamic across the cell cycle in a manner that is reflective of gene expression. **I**. Also observed is a cyclical pattern in the form of H3K9me2 which is depleted at G1-phase and enriched at S and G2/M-phases regardless of the point of gene expression.

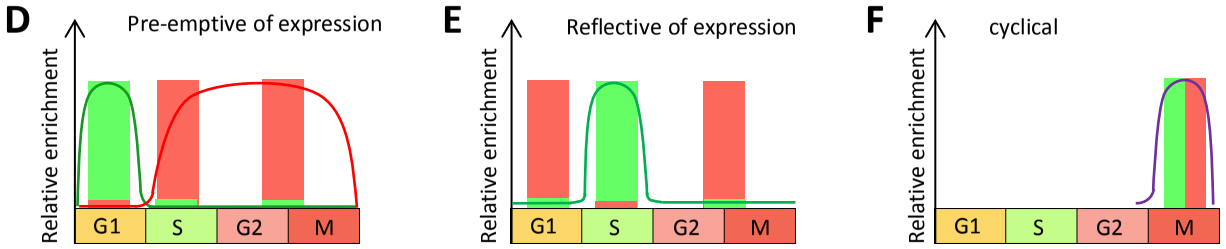
Figure 4.1: Predicted and observed patterns of histone modifications in undifferentiated ES cells and day 7 embryoid bodies

Potential patterns of histone modifications

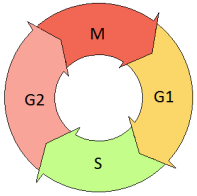


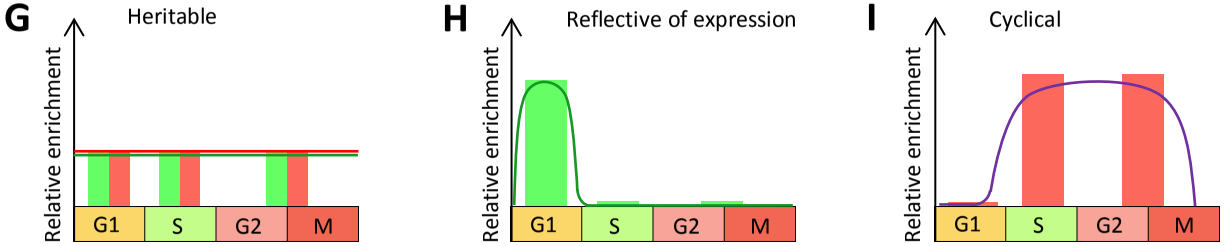
Observed patterns in undifferentiated ES cells





Observed patterns in Day 7 embryoid bodies





To facilitate this approach a selection of active and repressive modifications were chosen alongside a panel of cell cycle regulated genes. In this way it was possible to search for patterns across the cell cycle in cell cycle regulated genes in a way that had not previously been explored. Previous investigations indicate that many histone modifications are highly dynamic in nature, particularly acetylation (Barth & Imhof 2010; Zee et al. 2010; Zentner & Henikoff 2013), however the vast majority of studies have also been conducted on asynchronous populations as modifications have been generally assumed to be more or less stable across the cell cycle. The data within this thesis clearly reveals a much more complex scenario by exploring histone modifications across the cell cycle, patterns or associations between histone modifications that would be missed in an asynchronous population have been subsequently been observed.

The role of histone modifications in stem cells is of particular interest. Stem cells can be used as a model of development in which the precise regulation of gene expression ultimately enables a single cell to develop into an entire organism. Several key differences have been identified between the epigenetic signature of pluripotent stem cells and that of somatic cells including the larger proportion of open/transcription-permissive chromatin and the relative abundance of bivalent marks (Azuara et al. 2006; Spivakov & Fisher 2007). By comparing histone modifications in stem cells and in day 7 embryoid bodies, our understanding of the amazing plasticity these cells possess may improve, thereby ultimately helping to realise their therapeutic potential.

4.1 Histone modification patterns in chemically arrested ES cells

The initial approach adopted was the use of chemicals to induce cell cycle arrest at G1/S and M-phase to determine the relative enrichment of a selection of five histone marks,

three marks associated with activating transcription (H3K4me3, H3K9ac and H4K8ac) and two with repressing transcription (H3K27me3 and H3K9me2).

It is immediately clear by looking at the cell cycle regulated genes that the levels of each of the five marks vary across the cell cycle in at least one of the genes, the majority of the marks varying across most of the genes. The pattern emerging for at least some of these modifications in cell cycle regulated genes is one of highly dynamic modifications that are reflective of regulation of gene expression. The majority of genes appear to have an enrichment of at least one active mark either pre-emptive or at the point of expression (figure 4.1 D+E). However, the large variability of marks between phases and across genes, compounded by the points of observation (G1/S phase rather than separate G1 and S) make interpretation of this data set without further information very difficult.

Interestingly there also appears to be an increase in levels of most of the marks at M-phase regardless of gene expression (figure 4.1 F), this is observed in cells chemically arrested at M-phase and is not seen in the FACs isolated populations. FACs sorted cells are isolated based on DNA content, resulting in a G2/M population that is not comparable to the cells chemically arrested at M-phase. Only a small proportion of G2/M–phase cells will contain cells undergoing mitosis. This could explain how any signal from M-phase cells may be obscured by the greater number of cells at G2–phase, however it is important to also consider the possibility that the use of colcemid to cause cell cycle arrest at M-phase may be resulting in the artificial enrichment of all marks within these cells (Rieder & Palazzo 1992).

4.1.1 CELL CYCLE REGULATED GENES

Relative levels of the histone marks were considered at the promoters of cell cycle regulated genes. The G1-phase associated genes displayed little change in histone modification enrichment in G1/S arrested cells compared to the asynchronous population. However levels of all three active marks increase in M-phase arrested cells. This is also true of H3K27me3 in 2 of the three genes. This suggests that enrichment of histone modifications is potentially pre-emptive of gene expression in this instance, not reflective of expression patterns. The increased presence of both active and repressive marks could be indicative of a temporary bivalency at M-phase.

In S-phase associated genes a greater variation is observed between the asynchronous population and G1/S phase arrested cells. The general pattern observed is one of an increase in the active marks H3K4me3 and H4K8ac in G1/S phase arrested cells when compared to the asynchronous. This is again suggestive of a potentially pre-emptive enrichment of active marks prior to or upon initiation of gene expression. The repressive marks do not vary much across from G1/S arrested cells to the asynchronous. Once again a large proportion of the marks, predominantly active, are again enriched in M-phase arrested cells. This enrichment of marks seen at M-phase both for active and repressive modifications does not appear to relate to the cell cycle dependent expression, it is therefore possible that this pattern is indicative of another mode of action for marks exclusive to M-phase.

The patterns within the G2 phase associated genes are similar to those observed for S-phase associated genes. Two of the three genes (*Hurp* and *Ccnb1*) have an enrichment of the active mark H4K8ac in G1/S cells which could be pre-emptive of an early G2 expression. No real difference is observed between the asynchronous and G1/S for the

repressive marks, however once again the majority of all marks are enriched in the M-phase population. Amongst the G2/M associated genes H4K8ac is enriched at G1/S and M phase, however the overall pattern for the remaining marks is one of enrichment at M-Phase.

4.1.2 PLURIPOTENCY ASSOCIATED GENES

Given that cell cycle regulated genes are likely to require close control across the cell cycle it may be that this is not the type of gene in which long term, heritable marks will be observed. Looking at cell cycle regulated genes facilitates the identification of patterns with respect to gene expression, however those patterns may also be specific to this subset of genes. It was therefore important to look at other genes as well, pluripotency associated genes were selected for their high expression in ES cells and their vital role maintaining cell identity. *C-myc* was included as it was one of the reprogramming factors used in the generation of iPSCs (although they have since been generated in its absence -Nakagawa et al. 2008). It is believed to regulate the expression of 15% of all genes, including many involved in cell division, growth and apoptosis. The same selection of histone modifications were considered in *C-myc*, *Klf4*, *Nanog*, *Pou5f1* and the house keeping gene *β-actin* which are not believed to be in anyway cell cycle regulated.

Again each mark was found to vary across the cell cycle in at least one of the genes, making classic heritability (figure 4.1 A) highly unlikely in this selection of genes as well as the cell cycle regulated genes for at least for this group of modifications. The overall pattern is as it was in previous genes; an enrichment of marks (primarily active marks but also H3K27me3 in *β-actin* and *Klf4*) at M-phase (figure 4.1 F). This is most noticeable of H3K9ac and H4K8ac but is also true of H3K4me3 to a lesser extent. Of the

active marks H4K8ac appears enriched at G1/S-phase in all of the pluripotency genes and H3K4me3 to a lesser extent. Surprisingly H3K9ac was slightly depleted at G1/S in most of this set of genes.

Levels of the repressive marks remain mostly constant across the cell cycle however in these pluripotency related genes this is unlikely to be due to the heritability of the marks. More over this lack of variability across the cell cycle would be expected in genes where the overall presence of marks was negligible. As highly active genes in undifferentiated ES cells, pluripotency genes would not be expected to have repressive marks at the gene promoter (Loh et al. 2006; Liang et al. 2008). If these genes are searched for within the UCSC Mouse Genome Browser (mm9) this is indeed what is observed, H3K27me3 is not enriched at the promoters of *Pou5f1* and *Nanog* the two genes with stable levels of the mark across the cell cycle (in ES cell line Bruce4 - Waterston et al. 2002; Barski et al. 2007).

These genes are believed to be expressed in the same mode as the majority of genes within a cell; ubiquitously across the cell cycle starting at G1 phase until they are condensed during mitosis. The enrichment of active marks at G1/S phase is reflective of an initiation of transcription in G1 and H4K8ac in particular appears to be more abundant in the pluripotency genes which are particularly important for ES cells. It is also clear that the varying patterns of histone enrichment that are observed across the cell cycle for all of these genes further supporting the idea of highly dynamic marks, conflicting with the view of classically heritable histones.

The pattern of enrichment at M-phase is observed in two-thirds of all the histone modifications covered in all of the genes both cell cycle regulated, pluripotency and house-keeping. However one-third of all instances covered thus far do not have this

enrichment at M-phase and its occurrence does not appear entirely random. Most of the repressive marks in pluripotency genes are not enriched at M-phase. As these genes are highly expressed in ES cells uniformity across the cell cycle is more likely to be attributable to a general depletion of repressive marks. This pattern of M-phase enrichment is much more extensive in G1/S associated genes where it is potentially pre-emptive of gene expression and least prevalent (although still present at around 50%) in S-phase associated genes where the enrichment does not appear to be pre-emptive or reflective of expression.

The frequent enrichment of marks at M-phase is ultimately not indicative of long term maintenance of marks but of a much more dynamic control system. A possible explanation for such observations is the deposition of marks prior to, or in conjunction with, chromatin condensation into chromosomes -as a marker of chromatin state. Precise banding patterns observed for many histone modifications in condensed M-phase chromosomes, which could support the idea of mitosis specific chromatin territories (Terrenoire et al. 2010). This would allow the rapid renewal of gene expression patterns immediately after mitosis, particularly important for the maintenance of ES cell identity. However should the enrichment of marks at M-phase prove to be a consequence of the colcemid treatment it would still raise interesting questions about the mechanism by which the enrichments are induced and why they are less prevalent amongst S-phase associated genes or repressive marks in pluripotency genes? Further work would require pure populations of M-phase cells that have not been achieved by a colcemid block which remains challenging. Possible avenues in achieving this include the synchronisation of cells at an earlier point in the cell cycle, such as a G1/S arrest (i.e. with thymidine), release of cells from the block and allowing them to progress simultaneously to M-phase before harvesting. However

during this project the synchronisation of cells by this method was tried and ES cells did not recover from cell cycle arrest simultaneously. Alternatively individual selection of metaphase cells as viewed through an optical microscope could be used followed by C-ChIP which allows analysis of as few as 100 cells; this approach is feasible but presents a significant technical challenge.

The select panel of histone modifications covered within this study are clearly not stable across the cell cycle and therefore are not heritable under the classical interpretation. Patterns of histone modifications appear complex and dynamic; by further breaking down the cell cycle it may be possible to glean further insights into their modes of action.

4.2 HISTONE MODIFICATION PATTERNS IN FACS SORTED ES CELLS

Arresting cells chemically is a commonly employed tool for isolating cells at particular points within the cell cycle. However the cell cycle arrest that is achieved results in differences within the cells and could be criticised for introducing an artificial environment (Rieder & Palazzo 1992). In order to verify that any observations are not an artefact of the chemically induced cell cycle arrest FACs sorting was employed. FACs sorted cells allow a unique opportunity to examine extremely pure fractions of three phases (G1, S and G2/M), separating G1 and S into isolated phases. The development and optimisation of C-ChIP for use on small numbers of acetone fixed cells has enabled a parallel analysis to N-ChIP on the limited numbers of FACs sorted cells and allowed the close examination of histone modifications in these precise pure populations for the first time (O'Neill et al. 2006).

4.2.1 CELL CYCLE REGULATED GENES

The primary difference observed between the FACs sorted samples and the chemically arrested cells is the absence of the frequent enrichment of marks at M-phase. This is not observed in the G2/M-phase cell fraction (primarily composed of cells at G2). Further understanding of this enrichment will require a separate investigation.

The results from the FACs sorted cells support data from the chemically arrested cells; none of the modifications are maintained stably across the cell cycle in all of the genes. This once again presents a picture of dynamically changing histone modifications. In progression from the chemically arrested cells it is soon apparent that the breakdown of the cell cycle by FACs does allow a much clearer picture of G1 and S-phase, in addition to the G2/M-phase fraction. A pattern of histone modifications that is pre-emptive of expression was observed across the majority of the genes looked at (figure 4.1 D). The particular histone modifications to be enriched however, varied from promoter to promoter, emphasizing the value of gene by gene focus.

Skp2 (expressed at early G1-phase) presents an enrichment of the active histone marks (H3K4me3 and H3K9ac) at G2/M-phase, pre-emptive of expression, whereas the repressive mark H3K9me2 is enriched at all other parts of the cell cycle hinting at cooperation between these marks in regulating gene expression, a pattern that is repeated in most genes. At *Ccnd1* (expressed at G1-phase) histone modifications appear reflective of expression rather than pre-emptive, however the same increase of repressive marks (again H3K9me2 but also H3K27me3) is observed at the points where gene expression is not expected.

Amongst S-phase associated genes active marks are observed both pre-emptively and at the point of expression (figure 4.1 D+E). In both cases the repressive marks are also

enriched at S and G2/M-phase potentially cooperating with the regulation of the genes by repressing their expression (beginning at the point of expression and into G2/M-phase). The same general pattern is observed for active marks in the remaining genes; *Hurp* (G2), *Ccna1* (G2/M) and *Cdc25b* (G2/M).

4.2.2 A COMPARISON BETWEEN RESULTS FROM FACS SORTED AND CHEMICALLY ARRESTED CELLS

It is immediately noticeable from this data that none of the modifications observed within these investigations remains constant across the cell cycle in all genes and therefore the idea that a mark is placed and remains unaffected from one generation of cells to the next is incorrect (for these five modifications at least). Should this be the case for all histone modifications then they would not be strictly heritable and therefore epigenetic under the classic understanding of the term. This being said, regardless of categorisation, it has been clearly established that histone modifications play a central role in not only gene expression but DNA replication, elongation and repair (Kouzarides 2007; Yu et al. 2008; Izzo & Schneider 2010; Bannister & Kouzarides 2011; Rothbart & Strahl 2014).

The exact patterns observed within the modifications investigated often varied from gene to gene, highlighting the importance of looking at genes individually. While global approaches such as ChIP-seq are providing enormous amounts of data very rapidly, and although clearly invaluable, it is very likely that in the use of such approaches finer details are slipping through. It is very important that more focused approaches such as this one are still carried out.

With the exception of M-phase data which is not comparable between the two cell cycle isolation methods, the two sets of data reveal the same overall patterns in the

enrichment of histone modifications. However, even within these two sets the importance of breaking down the cell cycle can be observed, for example levels of H3K9me2 in *Ccnd1* appear consistent in the chemically arrested cells whereas when G1/S is separated into isolated populations of G1 and S-phase cells depletion is observed in G1 and an enrichment in S-phase. The same is true of several other marks in other genes (H3K4me3, H3K27me3 and H3K9me2 in *Pcna*, H3K9me2 in *Dnajc2*, H3K4me3, H3K9ac and H3K9me2 in *Hurp*, H3K9me2 in *Ccna1*, and H3K9me2 in *Cdc25b*).

The overall pattern observed in cell cycle regulated genes across the cell cycle is of an enrichment of at least one active mark either pre-emptive of gene expression, or coinciding with the point of expression (figure 4.1 D+E). This suggests a highly dynamic role for histone modifications in nucleosome signalling. Acetylation marks are known to be transient, with modification turnover happening in as few as 3 minutes, however methyl marks are believed to be slower and as such the most likely candidate for heritable histones (Barth & Imhof 2010). A study by Zee and colleagues (2002) catalogued global turnover rates of a variety of histone modifications and determined that the overall half-life for modifications on H3 and H4 are around 1.3 days, a period which would allow for heritability of marks yet this is not observed within our ES cell population. Following on from this histones H2B and H3, in addition to heterochromatin protein 1 (HP1), (which bind to the repressive marks H3K9me2 and H3K9me3) have an increased exchange rate in ES cells when compared to differentiated cells (Meshorer et al. 2006). Interestingly the exchange rate of H3.3 (associated with actively transcribed regions) remains rapid in both ES and differentiated cells. Taking this into consideration the proposed model of a highly dynamic exchange of histone modifications could be applicable to ES cells but may require alteration upon differentiation of cells. This is

further supported by the actions of ATP dependent chromatin remodelling factors, it is estimated that there may be as many as 1 chromatin remodeller per 14 nucleosomes and that they are continuously shuffling nucleosomes back and forth allowing access to the underlying DNA (Varga-Weisz 2010). Hyper-dynamic binding of chromatin modifying proteins would allow cells to be more responsive when in the pluripotent state.

There also appears to be a corresponding pattern amongst repressive marks of enrichment at the points where active marks drop off. This is suggestive of regulation through dual mechanisms both active and repressive on the same gene as transcription is switched on and off. The nucleosome signalling hypothesis that has grown in popularity more recently fits well with the idea of "on" and "off" modifications dynamically regulated as marks respond to changes in the environment around the gene (potentially including feedback mechanisms from the gene itself) (Turner 2012). The pre-emptive enrichment of marks supports the growing evidence of a causal role for at least some histone modifications in gene activation and silencing. With further time and funding it would be helpful to fully characterise the point of transcription in the cell cycle for all the genes to match them to the patterns of histone modifications as precisely as possible.

4.3 THE EPIGENETIC LANDSCAPE OF REPRESENTATIVE CELL CYCLE REGULATED GENES

Looking across the broader gene allowed further insights into histone modifications; many marks showed cell cycle variation in histone enrichment into the body of the gene. Within the literature the active marks H3K4me3, H3K9ac, H4K8ac are linked to promoter regions and in some cases enhancer regions, whereas the repressive marks

H3K27me3, H3K9me2 and H3S10phos are found over broader regions including the gene body, therefore these repressive marks would be expected to stretch into the gene body, but the active marks would not (Barth & Imhof 2010; Zhou et al. 2010). Having said this it is also the case that ES cells possess a greater abundance of active marks and transcriptionally open chromatin than terminally differentiated cell (Meshorer et al. 2006; Lee & Wu 2011).

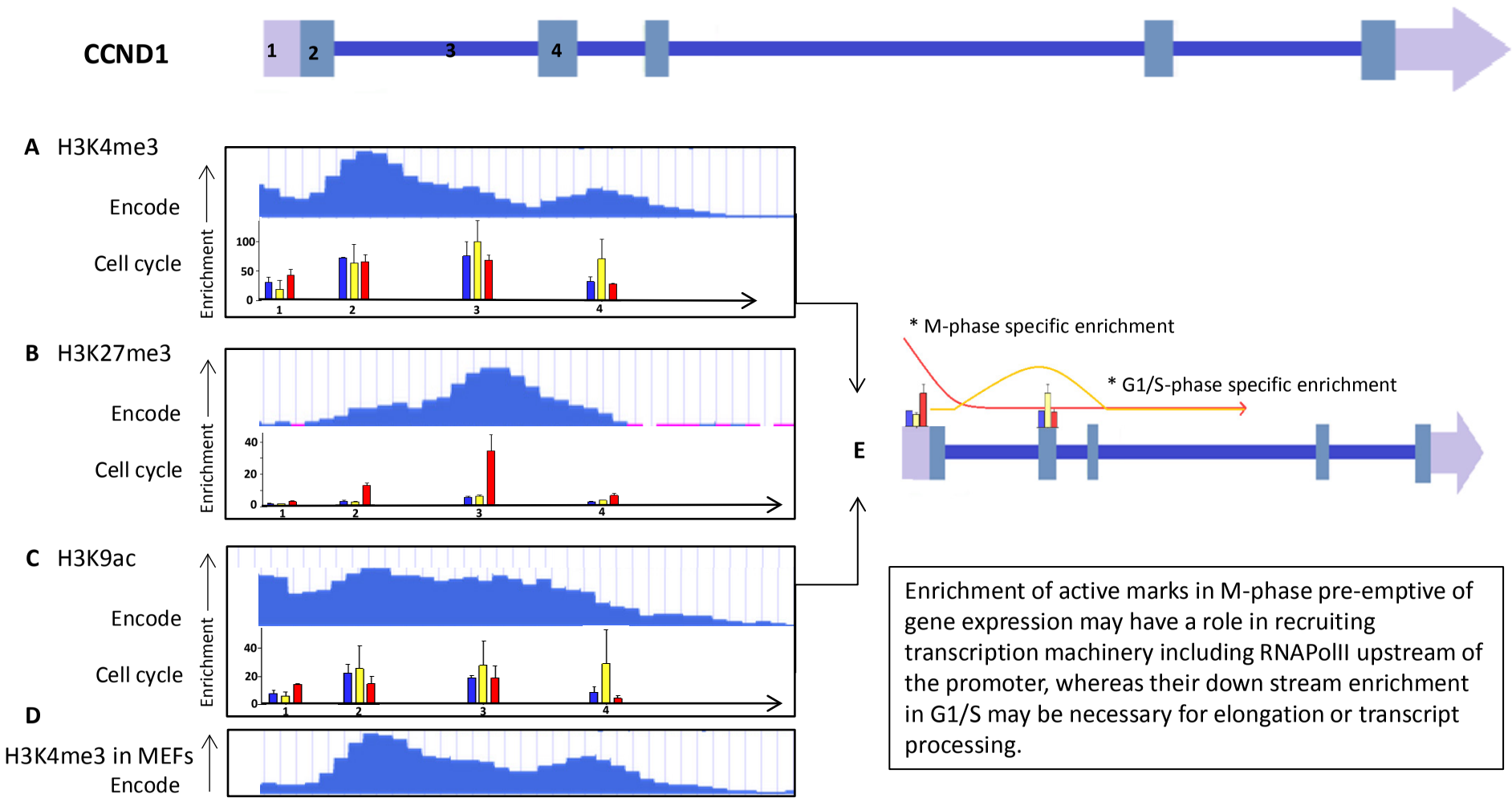
The information gathered was compared to data from the Encode project (Waterston et al. 2002). All the global chromatin profiling that has been, and is being, carried out at present is performed on populations of asynchronous cells, a direct comparison was not made rather a comparison of the relative levels of modifications across the cell cycle to the asynchronous Encode data viewed using the UCSC genome browser (Waterston et al. 2002; Barski et al. 2007).

Ccnd1 has been used to represent this in figure 4.2. The slightly confusing observation that the H3K4me3 (figure 4.2 A) enriched phase at the promoter of this gene is M-phase, whereas this slowly changes to G1/S-phase across the gene becomes much clearer upon comparison with the encode patterns. The enrichment at M-phase appears to correspond perfectly with the peak prior to the TSS, whereas the enrichment at G1/S-phase corresponds to the downstream peak at the second exon. These two peaks contained within the asynchronous Encode data appear to be present at differing time points when the cell cycle is broken down. This is repeated in the other active mark H3K9ac (figure 4.2 C). Only *Pcna* of the remaining genes analysed in this manner appears to have the same pattern in H3K9ac also. The other genes had only single peaks

Figure 4.2: The comparison of the relative enrichment of histone modifications in G1/S-phase and M-phase arrested ES cells to Encode data across the Ccnd1 gene

Relative levels of **A**. H3K4me3, **B**. H3K27me3 and **C**. H3K9ac were compared to the patterns observed across the gene in Encode data (ES cell line Bruce4). Looking at several points at the start of the gene allows a difference in the timing of enrichment peaks to be identified in active marks. Data shown is the ratio of enrichment in the bound sample compared to the unbound sample. H3K4me3 and H3K9ac appear to be enriched at M-phase prior to the transcription start site, whereas enrichment of the mark is at G1/S into the body of the gene peaking at the second exon. Yellow bars = G1/S data, red bars = M-phase data and blue bars = asynchronous. **E**. A potential explanation of this may be found in the timing of the recruitment of various components of transcription machinery, however a great deal more experimentation would be required to investigate this further. **D**. Enrichment of H3K4me3 in MEFs provided for comparison. Encode data obtained from the UCSC genome browser (Waterston *et. al.* 2002; Barski *et. al*. 2007).

Figure 4.2: The comparison of the relative enrichment of histone modifications in G1/S-phase and M-phase arrested ES cells to Encode data across the Ccnd1 gene



of enrichment at around the TSS for the modifications in question in the Encode data (Waterston et al. 2002; Barski et al. 2007).

The H3K27me3 mark in *Ccnd1* (figure 4.2 B) is enriched at M-phase in the same pattern as the Encode data suggesting that the modification is M-phase specific. The same is true of H3K27me3 for *Pcna*, *Hurp* and *Cdc25b*. H3K4me3 enrichment in MEF cells was also included in the comparison as an example of a more differentiated cell type, in all cases the profile was very similar to that of the ES cell lines.

By comparing data from this thesis to that from the Encode project we have uncovered some potentially important insights into these select histone modifications. High throughput methods are providing us with an astonishing amount of data; however this thesis clearly highlights the importance of timing with respect to histone modifications. An aspect that is not incorporated into the experiments that feed into these huge data sets. Not only may enrichment peaks be missed entirely but in looking across select genes it appears that the same histone modifications may be employed at different points of the same gene in a time sensitive manner (figure 4.2 E). This opens an avenue for an entirely new line of investigation.

H3S10phos in this case serves as a control. It has long been documented that this mark is more abundant at M-phase in somatic cells, in ES cells this period of enrichment extends from the end of S-phase to the end of G1 and this is exactly what is observed in all four genes. A slight enrichment at G1/S, as the mark is lost followed by an approximate 2-3 fold increase in M-phase compared to the asynchronous population (Hendzel et al. 1997; Teperek-Tkacz et al. 2010).

4.4 HISTONE MODIFICATION PATTERNS IN DAY 7 EMBRYOID BODIES

ES cells are known to have a particularly permissive and open chromatin environment, this includes a lot of developmental genes which possess bivalent marks to maintain them in a poised state prior to differentiation (Voigt et al. 2013a). Once cells differentiate major chromatin remodelling takes place including the large scale loss of acetylation coupled with the silencing and compression of large parts of the genome. Histone modifications have been linked to short term effects in a variety of ways including the activation or repression of gene expression, DNA replication, repair, recombination, and the higher order structure of chromatin including condensation into chromosomes during metaphase. (Nightingale et al. 2006; Kouzarides 2007; Shilatifard 2008; Yu et al. 2008; Bannister & Kouzarides 2011; Henikoff & Shilatifard 2011; Rothbart & Strahl 2014).

Thus far the histone modifications looked at have changed rapidly in cell cycle regulated genes, either pre-emptive or at the point of expression, and hence a dynamic model was proposed for the regulation of these modifications in ES cells. A different story may be true for differentiated or differentiating cells. It may be possible that after the vast chromatin remodelling that takes place upon the instigation of differentiation, histone modification patterns could constitute part of this change and become classically heritable. To determine whether any changes in histone modification patterns and consequently modes of action take place a repeat of the studies conducted on ES cells was performed on cells that had been differentiated into day 7 embryoid bodies. In doing so it was possible to compare the relative enrichment of histone marks in ES cells to cells that have commenced down a differentiation pathway with all the changes that

this involves including the enforcement of a more canonical cell cycle profile (figure 3.23) and a rapid alteration of the expression profile (Glover et al. 2006).

It is important to note that these cells have only been differentiated for seven days. This time is sufficient for the enforcement of a canonical cell cycle, however reprogramming of the transcriptome is likely to continue until cells reach unipotency or enter a state of cell senescence. The generation of iPSCs has reignited a debate on the plasticity of cells. Once believed to be unidirectional we are now able to reprogram terminally differentiated cells into pluripotent cells. The efficiency of this is currently very low in human cells (0.01% despite 50% transduction rate -Hochedlinger & Plath 2009), however many groups are aiming to improve this rate, including investigations into the heterogeneity of starting populations or the "state" cells are in upon initiation, less focus is placed on the point at which pluripotency is lost (O'Doherty et al. 2013; Tsubouchi & Fisher 2013). With further time and funding it would be extremely interesting and useful to determine at which point along the differentiation timeline cells are unable to return to the pluripotent state without external reprogramming factors. Preliminary work within our lab has suggested that cells still maintain the ability to do this to day 7 of differentiation, the study has not yet been taken further. It has been suggested that terminally differentiated somatic cells are able to regain a large degree of plasticity when place into a specific microenvironment or "niche" (Lim & Gong 2013). There is no doubt that a greater understanding of stem cell/somatic cell plasticity could be infinitely useful in devising novel clinical treatments.

4.4.1 DAY 7 EMBRYOID BODIES CHEMICALLY ARRESTED AT G1/S-PHASE

The cell cycle regulated genes that are considered in this thesis are primarily associated with cell cycle related functions and as such are not silenced upon differentiation. It is

differentiation induced changes to the cell cycle that may affect expression of this particular set of genes. Upon the instigation of differentiation a canonical cell cycle is enforced, G1/S checkpoint is fully regulated and the length of the gap phases increases returning cells to a normal cell cycle profile (figure 3.26). Levels of the cyclins and cdk activity oscillate more gently in ES cells, once cells differentiate these oscillations increase (White & Dalton 2005; Ballabeni et al. 2011). With this in mind it is therefore not surprising to observe that the patterns in levels of histone modification enrichment are often clearer and better defined, with less margin of error between biological replicates in differentiated cells.

The most notable change after differentiation in the chemically arrested cells seems to be the switch from an abundance in enrichment of H4K8ac at G1/S in ES cells (which disappears in the day 7 embryoid bodies) to be almost replaced by a high degree of H3K9ac enrichment at G1/S-phase in many genes. This is also true for the pluripotency and house-keeping genes. Further isolated fractions of the cell cycle are required for a greater analysis.

4.4.2 FACS SORTED DAY 7 EMBRYOID BODIES

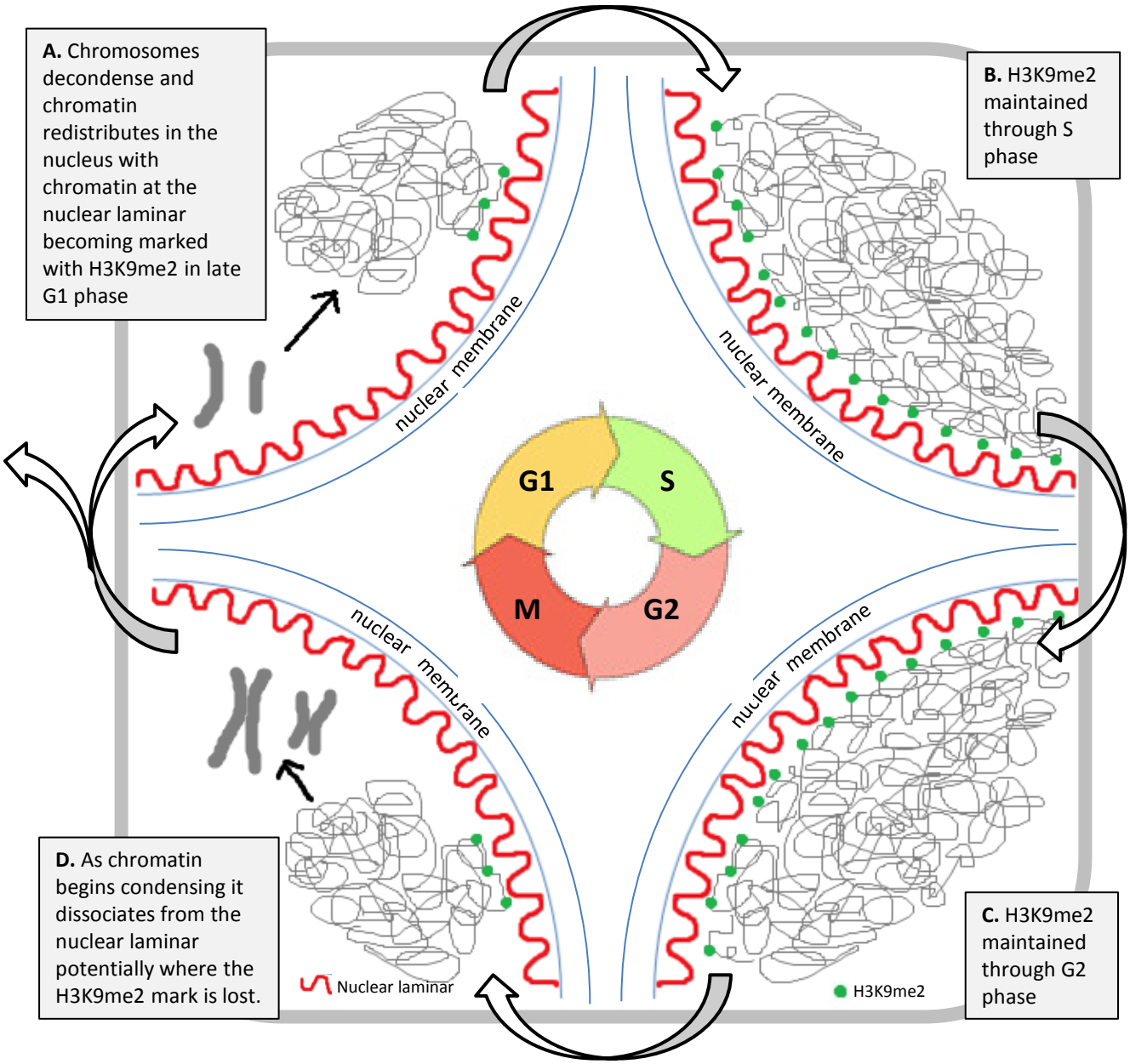
The most obvious difference between the undifferentiated ES cells and the day 7 embryoid bodies is the very clear pattern of enrichment in H3K9me2 at S and G2/M-phases and its depletion in G1-phase which is observed across all genes in the differentiated cells as per figure 4.1 I. Of the repressive marks H3K9me2 appears to be the least cell cycle influenced in ES cells directly contrasting with what is observed in the day 7 embryoid bodies. By comparing the percentage pull down for the antibody to H3K9me2 (Appendix table 6.1 and 6.4) we can clearly see that global levels of the marks do not vary enormously across the cell cycle for ES cells but in the day 7 embryoid

bodies the marks is negligible at G1 but very present at S and G2/M-phase. All of this strongly supports the idea of different roles for the mark in the two cell states, even on a global scale.

However the vast variation between levels of H3K9me2 in G1 and then S and G2/M-phase in embryoid bodies is directly contradicted by Duan *et. al.* (2008) in which a depletion in the mark at prophase, prometaphase and metaphase was found in a number of carcinoma derived cell lines (Duan et al. 2008). However it must be taken into consideration that cells derived from genetically unstable carcinoma cells are likely to present abnormal features.

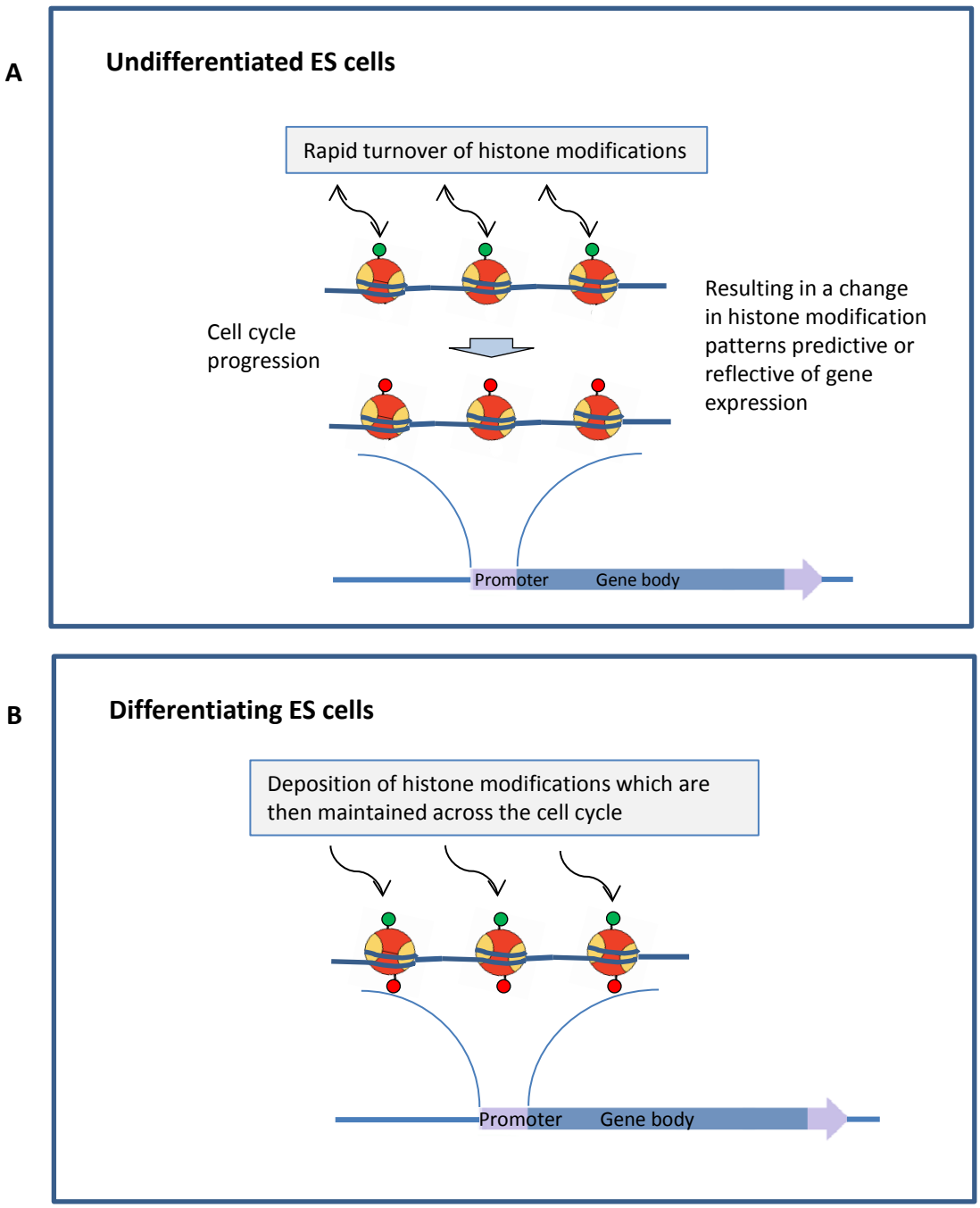
In support of a difference between H3K9me2 in the two cell states Wen *et al*. found that the mark was distributed in large chromatin blacks in differentiated cells but not in ES cells (Wen et al. 2009). H3K9me2 is associated with heterochromatin which is positioned at the nuclear envelope after exiting mitosis (Zhou et al. 2010). The observed cyclical change in H3K9me2 might be explained by recent data to suggest that H3K9me2 is located at the nuclear laminar and chromatin that remains at the nuclear laminar retained the H3K9me2 mark whereas that which moves away does not (Kind et al. 2013). Another study that found that G9a methyltransferase interacts with the nuclear laminar interacting protein BAF (Barrier to auto integration factor -Montes de Oca et al. 2011). This along with the knowledge that association between laminar associated domains (LAD) in chromatin and the nuclear laminar only commence approximately 40 minutes into G1 phase and clear results from this thesis present a compelling argument for the deposition of H3K9me2 at the nuclear laminar late in G1-phase (figure 4.3).

Figure 4.3: Model for the transmission of H3K9me2 in day 7 embryoid bodies



A representation of how H3K9me2 may be deposited across the cell cycle in differentiating or differentiating ES cells. **A**. In G1 phase chromosomes decondense into chromatin and chromatin is reorganised into specific chromatin domains, chromatin associated with the nuclear laminar has high H3K9me2 enrichment, the interaction may be involved in the deposition of the mark. **B.&C**. H3K9me2 is maintained through S and G2 phase. **D.** As chromatin condenses into chromosomes it dissociates from the nuclear laminar potentially the point at which the mark is lost prior to G1-phase

Figure 4.4: Model of the change in histone modification patterns observed for H3K4me3 and H3K27me3 in undifferentiated ES cells compared to differentiating ES cells



H3K4me3 and H3K27me3 present very different patterns of histone modification enrichment in undifferentiated ES cells and differentiating ES cells. **A**. Modification patterns change with cell cycle progression in a manner that is either predictive or reflective of gene expression in undifferentiated ES cells. **B**. Once differentiation is initiated enrichment of these modifications is heritable across the cell cycle.

ES cells on the other hand have a different nuclear organisation of chromatin which could potentially explain the differences observed from ES cells to day 7 embryoid bodies (Yoshioka et al. 2009). Alternatively H3K9me2 deposition could be coupled to DNA replication, however were this truly the case we would expect the relative enrichment of the mark to gradually increase across S-phase therefore overall it would be lower in G2/M which we are not observing. The more likely hypothesis is that the mark is lost upon chromatin condensation in M-phase (as chromosomes condense they dissociate from the nuclear laminar) and is re-deposited at the nuclear envelope once the spatial re-organisation of chromatin in the nucleus has been achieved (figure 4.3). This pattern follows the principles of figure 4.1 C, in that enrichment of the mark is repeated cyclically in a manner that is not linked to gene expression.

Of the remaining marks both H3K4me3 and H3K27me3 appear to have stabilized across the cell cycle in most genes with the notable exception of *Pcna*. To confirm an enrichment of these marks around the promoter region in differentiated we consulted Encode data and determined that peaks of H3K4me3, H3K27me3 and H3K9ac are observed in all of the genes in both liver cells and brain cells, with the exception of H3K27me3 in *Hurp*, where the mark is negligible around the TSS (Waterston et al. 2002; Barski et al. 2007). This is also the most anomalous gene when observing the stabilisation of H3K27me3 across the cell cycle. It is therefore likely that at low levels of the mark small differences are amplified. Both H3K4me3 and H3K27me3 appear to be heritable across the cell cycle.

When considering histone turnover, methyl marks were considered to be the most likely candidates, this indeed appears to be the case. The half-life of H3K27me3 was determined to be approximately 3 days, much longer than the average of 1.3 marking it

as a likely candidate for heritability (Zee et al. 2010). Other methyl marks with long half-lives include H3K79me2 (≃3.6 days) and H4K20me3 (≃4.8 days), given that H3K27me3 appears to be heritable these two marks are worth investigating, although it must be considered that the initial study was conducted on HeLa cells and will likely vary from other cell types (Waterston et al. 2002; Barski et al. 2007). The half-life of H3K4me3 was not determined in this set of experiments however it was suggested that the presence of both an active mark and repressive mark appeared to slow histone turnover. These results are consistent with the model proposed in figure 4.1 A, and as the two marks appear to be present and yet are stable across the cell cycle they can be considered heritable in the day 7 embryoid bodies. This is a vast change from the undifferentiated ES cells (figure 4.4).

The remaining mark (the only acetyl mark) H3K9ac is more difficult to characterise. In some of the genes (*Ccnd1* and *Pcna*) this active marks is enriched at the point of gene expression. However in others (*Skp2*, *Dnajc2*, *Hurp*, and *Cdc25b*) the marks appear stable across the cell cycle. That being said several of this later class of genes appear to be stable across the cell cycle but at a value of less than one (asynchronous levels). This suggests that an enrichment in the mark is taking place but is not observed in the isolated populations. Further break down and coverage of the cell cycle might help clarify this. The variability observed for H3K9ac along with the fact that acetylation marks have a much faster turnover mean than it is not likely that this mark is heritable like H3K4me3 and H3K27me3.

The results within this thesis raise many more questions such as: Do specific modifications peak at different genetic loci at separate time points and to what end? Which other histone modifications are heritable and by which means is the information

subsequently interpreted? Can these heritable marks be passed from one generation of mice to the next -the dynamic nature of these marks in ES cells would suggest not. Each of these points are a potential starting point for further investigations, all of which would help to further increase our understanding within epigenetics that would over time translate into the clinical setting.

4.5 CONCLUSIONS

The field of epigenetics into which histone modifications are incorporated has been receiving ever more interest particularly in the last couple of decades. As our understanding of the varying roles of histone modifications grows an increasingly complex picture unfolds in which histones modifications cross talk not only with each other but also with other epigenetic factors. It seemed likely given the variety of roles attributed to histone modifications that multi modes of action would be uncovered and this has indeed been the case. This thesis provides evidence towards several different modes of action for a range of modifications, potentially both short term and long term.

The presence of histone marks in metaphase cells as seen by spreads was originally viewed as evidence of their persistence throughout the cell cycle. It was thought that if marks were maintained during this crucial stage of chromatin condensation then they were heritable. This has since been revealed as a much more complicated scenario as the work conducted in this thesis attests. The histone modifications covered within this thesis, H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2, are not heritable in undifferentiated ES cells in the classic understanding of the term. They do not remain consistent across the cell cycle in any of the genes looked into within this set of investigations in ES cells. Rather the modifications appear to be regulated very dynamically in these cells, mostly pre-emptive of, or reflective of, gene expression

patterns. In addition to a transcription associated dynamic change in histone enrichment, a cell cycle specific enrichment was also observed in cells chemically arrested at M-phase. All of the marks investigated appear to be enriched at M-phase in most but not all genes. This would be an interesting avenue for further research.

Upon differentiation H3K4me3 and H3K27me3 both appear to stabilise across the cell cycle despite clear peaks of the modifications at the promoter region. Both of these marks are therefore classically heritable in the majority of genes looked at in the day 7 embryoid bodies. A different pattern was observed for H3K9me2, it appears to be cyclical, absent at G1 and enriched at S and G2/M regardless of the gene. This could also be classified as heritable but not under the classical understanding of the term. The pattern observed for H3K9me2 in differentiated ES cells is consistent with a role in the spatial organisation of chromatin supported by the literature.

The heritability of these marks in differentiated embryoid bodies but not in the pluripotent ES cells themselves suggests that the ES cells require highly dynamic regulation of gene expression but that this becomes fixed and heritable upon entering a less plastic state. H3K9ac on the other hand appears to remain dynamic in accord with the idea that the high turnover of acetyl marks makes them unlikely candidates for long term roles. Of the four histone modifications there appear to be three differing mechanisms of action in differentiated ES cells; heritable marks (H3K4me3 and H3K27me3); a dynamic mark linked to expression (H3K9ac); and a cyclical mark not related to expression (H3K9me2).

Another important point made clear by the results of this thesis, is that work carried out on asynchronous populations may not always give a complete picture, for instance if a mark were highly enriched in an active promoter but only transiently in order to trigger

a downstream event, this could easily be missed. Clearly amongst cell cycle regulated genes in ES cells at least, simply looking at an asynchronous population of cells can give a distorted view. In light of the increasing interest into epigenetic regulation of cell cycle events this should be taken into consideration (Wang & Blelloch 2009; Ballabeni et al. 2011; Momčilović et al. 2011).

5 REFERENCES

Abranches, E., Bekman, E. & Henrique, D., 2013. Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. *PloS one*, 8(3), p.e59928.

Adjaye, J. et al., 2005. Primary differentiation in the human blastocyst: comparative molecular portraits of inner cell mass and trophectoderm cells. *Stem cells (Dayton, Ohio)*, 23(10), pp.1514–25.

Agarwal, M.L. et al., 1995. p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), pp.8493–7.

Alabert, C. & Groth, A., 2012. Chromatin replication and epigenome maintenance. *Nature reviews. Molecular cell biology*, 13(3), pp.153–67.

Allfrey, V.G., Faulkner, R. & Mirsky, A.E., 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 51, pp.786–94.

Almouzni, G. et al., 2013. Chromatin dynamics at the replication fork: there's more to life than histones. *Current Opinion in Genetics & Development*, 23(2), pp.140–146.

Amariglio, N. et al., 2009. Donor-derived brain tumor following neural stem cell transplantation in an ataxia telangiectasia patient. *PLoS medicine*, 6(2), p.e1000029.

Annunziato, A.T., 2005. Split decision: what happens to nucleosomes during DNA replication? *The Journal of biological chemistry*, 280(13), pp.12065–8.

Arnaudo, A.M. & Garcia, B.A., 2013. Proteomic characterization of novel histone post-translational modifications. *Epigenetics & chromatin*, 6(1), p.24.

Azuara, V. et al., 2006. Chromatin signatures of pluripotent cell lines. *Nature cell biology*, 8(5), pp.532–8.

Bach, C. et al., 2009. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. *Oncogene*, 28(6), pp.815–23.

Ball, M.P. et al., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature biotechnology*, 27(4), pp.361–8.

Ballabeni, A. et al., 2011. Cell cycle adaptations of embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48), pp.19252–7.

Bannister, A.J. et al., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, 410(6824), pp.120–4.

Bannister, A.J. & Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Cell research*, 21(3), pp.381–95.

Barski, A. et al., 2007. High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4), pp.823–37.

Bartek, J. & Lukas, J., 2007. DNA damage checkpoints: from initiation to recovery or adaptation. *Current opinion in cell biology*, 19(2), pp.238–45.

Bartek, J. & Lukas, J., 2001. Pathways governing G1/S transition and their response to DNA damage. *FEBS letters*, 490(3), pp.117–22.

Barth, T.K. & Imhof, A., 2010. Fast signals and slow marks: the dynamics of histone modifications. *Trends in biochemical sciences*, 35(11), pp.618–26.

Barton, K.M. & Levine, E.M., 2008. Expression patterns and cell cycle profiles of PCNA, MCM6, cyclin D1, cyclin A2, cyclin B1, and phosphorylated histone H3 in the developing mouse retina. *Developmental dynamics : an official publication of the American Association of Anatomists*, 237(3), pp.672–82.

Bassett, A. et al., 2009. The folding and unfolding of eukaryotic chromatin. *Current opinion in genetics & development*, 19(2), pp.159–65.

Beattie, G.M. et al., 2005. Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem cells (Dayton, Ohio)*, 23(4), pp.489–95.

Becker, K.A. et al., 2006. Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. *Journal of cellular physiology*, 209(3), pp.883–93.

Bendall, S.C. et al., 2007. IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature*, 448(7157), pp.1015–21.

Berger, S.L. et al., 2009. An operational definition of epigenetics. *Genes & development*, 23(7), pp.781–3.

Bernstein, B.E. et al., 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125(2), pp.315–26.

Bernstein, B.E. et al., 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), pp.57–74.

Bessler, J.B., Andersen, E.C. & Villeneuve, A.M., 2010. Differential localization and independent acquisition of the H3K9me2 and H3K9me3 chromatin modifications in the Caenorhabditis elegans adult germ line. *PLoS genetics*, 6(1), p.e1000830.

Billon, N., 2002. Normal timing of oligodendrocyte development from genetically engineered, lineage-selectable mouse ES cells. *Journal of Cell Science*, 115(18), pp.3657–3665.

Birke, M. et al., 2002. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic acids research*, 30(4), pp.958–65.

Blais, M. et al., 2013. Concise review: tissue-engineered skin and nerve regeneration in burn treatment. *Stem cells translational medicine*, 2(7), pp.545–51.

Blomen, V.A. & Boonstra, J., 2011. Stable transmission of reversible modifications: maintenance of epigenetic information through the cell cycle. *Cellular and molecular life sciences : CMLS*, 68(1), pp.27–44.

Bojang, P. & Ramos, K.S., 2014. The promise and failures of epigenetic therapies for cancer treatment. *Cancer treatment reviews*, 40(1), pp.153–69.

Bortin, M.M., 1970. A compendium of reported human bone marrow transplants. *Transplantation*, 9(6), pp.571–87.

Bosco, E.E. & Knudsen, E.S., 2005. Differential role of RB in response to UV and IR damage. *Nucleic acids research*, 33(5), pp.1581–92.

Bostock, C.J., Prescott, D.M. & Kirkpatrick, J.B., 1971. An evaluation of the double thymidine block for synchronizing mammalian cells at the G1-S border. *Experimental cell research*, 68(1), pp.163–8.

Bourillot, P.-Y. et al., 2009. Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. *Stem cells (Dayton, Ohio)*, 27(8), pp.1760–71.

Bowman, A. et al., 2011. The histone chaperones Nap1 and Vps75 bind histones H3 and H4 in a tetrameric conformation. *Molecular cell*, 41(4), pp.398–408.

Boyer, L.A. et al., 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122(6), pp.947–56.

Bradley, A. et al., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, 309(5965), pp.255–6.

Bugler, B. et al., 2010. Unscheduled expression of CDC25B in S-phase leads to replicative stress and DNA damage. *Molecular cancer*, 9(1), p.29.

Burdon, T. et al., 1999. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Developmental biology*, 210(1), pp.30–43.

Cao, R. et al., 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science (New York, N.Y.)*, 298(5595), pp.1039–43.

Castel, S.E. & Martienssen, R.A., 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature reviews. Genetics*, 14(2), pp.100–12.

Cedar, H. & Bergman, Y., 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nature reviews. Genetics*, 10(5), pp.295–304.

Chambers, I. et al., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113(5), pp.643–55.

Chen, X. et al., 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, 133(6), pp.1106–17.

Cloos, P.A.C. et al., 2008. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes & development*, 22(9), pp.1115–40.

Coleman, T.R. & Dunphy, W.G., 1994. Cdc2 regulatory factors. *Current opinion in cell biology*, 6(6), pp.877–82.

Cosgrove, M.S., 2007. Histone proteomics and the epigenetic regulation of nucleosome mobility. *Expert review of proteomics*, 4(4), pp.465–78.

Czermin, B. et al., 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, 111(2), pp.185–96.

Dahéron, L. et al., 2004. LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem cells (Dayton, Ohio)*, 22(5), pp.770–8.

Daley, G.Q., 2012. The promise and perils of stem cell therapeutics. *Cell stem cell*, 10(6), pp.740–9.

Dalvai, M. & Bystricky, K., 2010. Cell cycle and anti-estrogen effects synergize to regulate cell proliferation and ER target gene expression. *PloS one*, 5(6), p.e11011.

Davey, C.A. et al., 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *Journal of molecular biology*, 319(5), pp.1097–113.

Deal, R.B., Henikoff, J.G. & Henikoff, S., 2010. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science (New York, N.Y.)*, 328(5982), pp.1161–4.

Deb, A. & Ubil, E., 2014. Cardiac fibroblast in development and wound healing. *Journal of molecular and cellular cardiology*.

DeGregori, J., Kowalik, T. & Nevins, J.R., 1995. Cellular targets for activation by the E2F1 transcription factor include DNA synthesis- and G1/S-regulatory genes. *Molecular and cellular biology*, 15(8), pp.4215–24.

Demetrick, D.J. & Beach, D.H., 1993. Chromosome mapping of human CDC25A and CDC25B phosphatases. *Genomics*, 18(1), pp.144–7.

Denissov, S. et al., 2014. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development (Cambridge, England)*, 141(3), pp.526–37.

Dhalluin, C. et al., 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, 399(6735), pp.491–6.

Di Cerbo, V. & Schneider, R., 2013. Cancers with wrong HATs: the impact of acetylation. *Briefings in functional genomics*, 12(3), pp.231–43.

Ding, L. et al., 2009. A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell stem cell*, 4(5), pp.403–15.

Ding, V.M.Y. et al., 2010. FGF-2 modulates Wnt signaling in undifferentiated hESC and iPS cells through activated PI3-K/GSK3beta signaling. *Journal of cellular physiology*, 225(2), pp.417–28.

Dolezalova, D. et al., 2012. MicroRNAs regulate p21(Waf1/Cip1) protein expression and the DNA damage response in human embryonic stem cells. *Stem cells (Dayton, Ohio)*, 30(7), pp.1362–72.

Draetta, G.F., 1994. Mammalian G1 cyclins. *Current opinion in cell biology*, 6(6), pp.842–6.

Duan, Q. et al., 2008. Phosphorylation of H3S10 blocks the access of H3K9 by specific antibodies and histone methyltransferase. Implication in regulating chromatin dynamics and epigenetic inheritance during mitosis. *The Journal of biological chemistry*, 283(48), pp.33585–90.

Ducommun, B. et al., 1991. cdc2 phosphorylation is required for its interaction with cyclin. *The EMBO journal*, 10(11), pp.3311–9.

Dunphy, W.G., 1994. The decision to enter mitosis. *Trends in cell biology*, 4(6), pp.202–7.

Duronio, R.J. & O'Farrell, P.H., 1995. Developmental control of the G1 to S transition in Drosophila: cyclin Eis a limiting downstream target of E2F. *Genes & development*, 9(12), pp.1456–68.

Elledge, S.J. & Harper, J.W., 1994. Cdk inhibitors: on the threshold of checkpoints and development. *Current opinion in cell biology*, 6(6), pp.847–52.

English, C.M. et al., 2005. ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3-H4 heterotetramer on DNA. *Biochemistry*, 44(42), pp.13673–82.

Ernst, J. et al., 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, 473(7345), pp.43–9.

Evans, M., 2011. Discovering pluripotency: 30 years of mouse embryonic stem cells. *Nature reviews. Molecular cell biology*, 12(10), pp.680–6.

Ewen, M.E. et al., 1993. Functional interactions of the retinoblastoma protein with mammalian D-type cyclins. *Cell*, 73(3), pp.487–97.

Falck, J. et al., 2002. The DNA damage-dependent intra-S phase checkpoint is regulated by parallel pathways. *Nature genetics*, 30(3), pp.290–4.

Felsenfeld, G. & Groudine, M., 2003. Controlling the double helix. *Nature*, 421(6921), pp.448–53.

Ficz, G. et al., 2011. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, 473(7347), pp.398–402.

Fischle, W. et al., 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature*, 438(7071), pp.1116–22.

Fleming, T.P., 1987. A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. *Developmental biology*, 119(2), pp.520–31.

Fuchs, S.M., Laribee, R.N. & Strahl, B.D., 2009. Protein modifications in transcription elongation. *Biochimica et biophysica acta*, 1789(1), pp.26–36.

Fujii-Yamamoto, H. et al., 2005. Cell cycle and developmental regulations of replication factors in mouse embryonic stem cells. *The Journal of biological chemistry*, 280(13), pp.12976–87.

Fusaki, N. et al., 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, 85(8), pp.348–62.

Gatti, R.A. et al., 1968. Immunological reconstitution of sex-linked lymphopenic immunological deficiency. *Lancet*, 2(7583), pp.1366–9.

Glover, C.H. et al., 2006. Meta-analysis of differentiating mouse embryonic stem cell gene expression kinetics reveals early change of a small gene set. *PLoS computational biology*, 2(11), p.e158.

Golob, J.L. et al., 2008. Chromatin remodeling during mouse and human embryonic stem cell differentiation. *Developmental dynamics : an official publication of the American Association of Anatomists*, 237(5), pp.1389–98.

Gonzalez, R. et al., 2011. Dorsomorphin promotes human embryonic stem cell self-renewal. *Angewandte Chemie (International ed. in English)*, 50(15), pp.3439–41.

Goodell, M.A. et al., 1996. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *The Journal of experimental medicine*, 183(4), pp.1797–806.

Grant, P.A. et al., 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & development*, 11(13), pp.1640–50.

Greber, B., Lehrach, H. & Adjaye, J., 2007. Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem cells (Dayton, Ohio)*, 25(2), pp.455–64.

Gurdon, J.B., 1962. Adult frogs derived from the nuclei of single somatic cells. *Developmental biology*, 4, pp.256–73.

Gurdon, J.B. & Wilmut, I., 2011. Nuclear transfer to eggs and oocytes. *Cold Spring Harbor perspectives in biology*, 3(6).

Gurley, L.R., Walters, R.A. & Tobey, R.A., 1974. Cell cycle-specific changes in histone phosphorylation associated with cell proliferation and chromosome condensation. *The Journal of cell biology*, 60(2), pp.356–64.

Haberland, M., Montgomery, R.L. & Olson, E.N., 2009. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature reviews. Genetics*, 10(1), pp.32–42.

Halsall, J. et al., 2012. Genes are often sheltered from the global histone hyperacetylation induced by HDAC inhibitors. *PloS one*, 7(3), p.e33453.

Hansen, K.H. et al., 2008. A model for transmission of the H3K27me3 epigenetic mark. *Nature cell biology*, 10(11), pp.1291–300.

Hansen, K.H. & Helin, K., 2009. Epigenetic inheritance through self-recruitment of the polycomb repressive complex 2. *Epigenetics : official journal of the DNA Methylation Society*, 4(3), pp.133–8.

Hathaway, N.A. et al., 2012. Dynamics and memory of heterochromatin in living cells. *Cell*, 149(7), pp.1447–60.

Hatzimichael, E. & Tuthill, M., 2010. Hematopoietic stem cell transplantation. *Stem cells and cloning : advances and applications*, 3, pp.105–117.

Hecht, A. et al., 1995. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell*, 80(4), pp.583–92.

Helin, K., Harlow, E. & Fattaey, A., 1993. Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein. *Molecular and cellular biology*, 13(10), pp.6501–8.

Hendzel, M.J. et al., 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an

ordered fashion coincident with mitotic chromosome condensation. *Chromosoma*, 106(6), pp.348–60.

Henikoff, S. & Shilatifard, A., 2011. Histone modification: cause or cog? *Trends in genetics : TIG*, 27(10), pp.389–96.

Hirota, T. et al., 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature*, 438(7071), pp.1176–80.

Hochedlinger, K. & Plath, K., 2009. Epigenetic reprogramming and induced pluripotency. *Development (Cambridge, England)*, 136(4), pp.509–23.

Hodawadekar, S.C. & Marmorstein, R., 2007. Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene*, 26(37), pp.5528–40.

Hodges, C. & Crabtree, G.R., 2012. Dynamics of inherently bounded histone modification domains. *Proceedings of the National Academy of Sciences of the United States of America*, 109(33), pp.13296–301.

Hong, Y. et al., 2007. Protecting genomic integrity in somatic cells and embryonic stem cells. *Mutation research*, 614(1-2), pp.48–55.

Houtgraaf, J.H., Versmissen, J. & van der Giessen, W.J., 2006. A concise review of DNA damage checkpoints and repair in mammalian cells. *Cardiovascular revascularization medicine : including molecular interventions*, 7(3), pp.165–72.

Hsu, J.Y. et al., 2000. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell*, 102(3), pp.279–91.

Huang, C., Xu, M. & Zhu, B., 2013. Epigenetic inheritance mediated by histone lysine methylation: maintaining transcriptional states without the precise restoration of marks? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1609), p.20110332.

Imai, S. et al., 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature*, 403(6771), pp.795–800.

Izzo, A. & Schneider, R., 2010. Chatting histone modifications in mammals. *Briefings in functional genomics*, 9(5-6), pp.429–43.

Jenuwein, T. & Allis, C.D., 2001. Translating the histone code. *Science (New York, N.Y.)*, 293(5532), pp.1074–80.

Jinno, S. et al., 1994. Cdc25A is a novel phosphatase functioning early in the cell cycle. *The EMBO journal*, 13(7), pp.1549–56.

Johnson, D.G., Ohtani, K. & Nevins, J.R., 1994. Autoregulatory control of E2F1 expression in response to positive and negative regulators of cell cycle progression. *Genes & development*, 8(13), pp.1514–25.

Jones, P.A., 2013. Genome wide analysis of DNA methylation and nucleosome positioning. *Epigenetics & Chromatin*, 6(Suppl 1), p.O37.

Jones, P.A. & Liang, G., 2009. Rethinking how DNA methylation patterns are maintained. *Nature reviews. Genetics*, 10(11), pp.805–11.

Jones, P.L. et al., 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature genetics*, 19(2), pp.187–91.

Katan-Khaykovich, Y. & Struhl, K., 2011. Splitting of H3-H4 tetramers at transcriptionally active genes undergoing dynamic histone exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 108(4), pp.1296–301.

Kato, J.Y. & Sherr, C.J., 1993. Inhibition of granulocyte differentiation by G1 cyclins D2 and D3 but not D1. *Proceedings of the National Academy of Sciences of the United States of America*, 90(24), pp.11513–7.

Kawai, H. et al., 2010. Tridermal tumorigenesis of induced pluripotent stem cells transplanted in ischemic brain. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 30(8), pp.1487–93.

Kidder, B.L., Yang, J. & Palmer, S., 2008. Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. *PloS one*, 3(12), p.e3932.

Kim, D. et al., 2009. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell stem cell*, 4(6), pp.472–6.

Kind, J. et al., 2013. Single-cell dynamics of genome-nuclear lamina interactions. *Cell*, 153(1), pp.178–92.

King, R.W., Jackson, P.K. & Kirschner, M.W., 1994. Mitosis in transition. *Cell*, 79(4), pp.563–71.

Kornberg, R.D. & Lorch, Y., 1999. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, 98(3), pp.285–94.

Kouraklis, G. & Theocharis, S., 2002. Histone deacetylase inhibitors and anticancer therapy. *Current medicinal chemistry. Anti-cancer agents*, 2(4), pp.477–84.

Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, 128(4), pp.693–705.

Krishnakumar, R. & Blelloch, R.H., 2013. Epigenetics of cellular reprogramming. *Current opinion in genetics & development*, 23(5), pp.548–55.

Ku, M. et al., 2008. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS genetics*, 4(10), p.e1000242.

Kuzmichev, A. et al., 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes & development*, 16(22), pp.2893–905.

Lachner, M. et al., 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, 410(6824), pp.116–20.

Lacomme, M. et al., 2012. NEUROG2 drives cell cycle exit of neuronal precursors by specifically repressing a subset of cyclins acting at the G1 and S phases of the cell cycle. *Molecular and cellular biology*, 32(13), pp.2596–607.

Lande-Diner, L., Zhang, J. & Cedar, H., 2009. Shifts in replication timing actively affect histone acetylation during nucleosome reassembly. *Molecular cell*, 34(6), pp.767–74.

Lee, J.H., Voo, K.S. & Skalnik, D.G., 2001. Identification and characterization of the DNA binding domain of CpG-binding protein. *The Journal of biological chemistry*, 276(48), pp.44669–76.

Lee, J.-S., Smith, E. & Shilatifard, A., 2010. The language of histone crosstalk. *Cell*, 142(5), pp.682–5.

Lee, Y.H. & Wu, Q., 2011. Chromatin regulation landscape of embryonic stem cell identity. *Bioscience reports*, 31(2), pp.77–86.

Leffak, I.M., Grainger, R. & Weintraub, H., 1977. Conservative assembly and segregation of nucleosomal histones. *Cell*, 12(3), pp.837–45.

Lewis, C.J., 2013. Stem cell application in acute burn care and reconstruction. *Journal of wound care*, 22(1), pp.7–8, 10, 12–6.

Li, S. & Shogren-Knaak, M.A., 2008. Cross-talk between histone H3 tails produces cooperative nucleosome acetylation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47), pp.18243–8.

Li, V.C., Ballabeni, A. & Kirschner, M.W., 2012. Gap 1 phase length and mouse embryonic stem cell self-renewal. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31), pp.12550–5.

Liang, J. et al., 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature cell biology*, 10(6), pp.731–9.

Lim, J.M. & Gong, S.P., 2013. Somatic cell transformation into stem cell-like cells induced by different microenvironments. *Organogenesis*, 9(4).

Lin, D.I. & Diehl, J.A., 2004. Mechanism of cell-cycle control: ligating the ligase. *Trends in biochemical sciences*, 29(9), pp.453–5.

Liu, Z. et al., 2013. The tumourigenicity of iPS cells and their differentiated derivates. *Journal of cellular and molecular medicine*, 17(6), pp.782–91.

Lodi, D., Iannitti, T. & Palmieri, B., 2011. Stem cells in clinical practice: applications and warnings. *Journal of experimental & clinical cancer research : CR*, 30, p.9.

Loh, Y.-H. et al., 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics*, 38(4), pp.431–40.

Loomis, R.J. et al., 2009. Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. *Molecular cell*, 33(4), pp.450–61.

Lou, Y. & Liang, X., 2011. Embryonic stem cell application in drug discovery. *Acta pharmacologica Sinica*, 32(2), pp.152–9.

Luger, K. et al., 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, 389(6648), pp.251–60.

Lynch, M.D. et al., 2012. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *The EMBO journal*, 31(2), pp.317–29.

Margueron, R. et al., 2009. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, 461(7265), pp.762–7.

Margueron, R. & Reinberg, D., 2011. The Polycomb complex PRC2 and its mark in life. *Nature*, 469(7330), pp.343–9.

Marshall, W.F., 2002. Order and disorder in the nucleus. *Current biology : CB*, 12(5), pp.R185–92.

Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 78(12), pp.7634–8.

Marzi, I. et al., 2013. The involvement of a Nanog, Klf4 and c-Myc transcriptional circuitry in the intertwining between neoplastic progression and reprogramming. *Cell cycle (Georgetown, Tex.)*, 12(2), pp.353–64.

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

McGarvey, K.M. et al., 2006. Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. *Cancer research*, 66(7), pp.3541–9.

McGowan, C.H. & Russell, P., 1995. Cell cycle regulation of human WEE1. *The EMBO journal*, 14(10), pp.2166–75.

Mendenhall, E.M. et al., 2010. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS genetics*, 6(12), p.e1001244.

Mermoud, J.E., Rowbotham, S.P. & Varga-Weisz, P.D., 2011. Keeping chromatin quiet: how nucleosome remodeling restores heterochromatin after replication. *Cell cycle (Georgetown, Tex.)*, 10(23), pp.4017–25.

Meshorer, E. et al., 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental cell*, 10(1), pp.105–16.

Messner, S. et al., 2010. PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucleic acids research*, 38(19), pp.6350–62.

Mikkelsen, T.S. et al., 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153), pp.553–60.

Millar, J.B. & Russell, P., 1992. The cdc25 M-phase inducer: an unconventional protein phosphatase. *Cell*, 68(3), pp.407–10.

Miller, J.P. et al., 2007. Interweaving the cell cycle machinery with cell differentiation. *Cell cycle (Georgetown, Tex.)*, 6(23), pp.2932–8.

Mitsui, K. et al., 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113(5), pp.631–42.

Molina-Serrano, D. & Kirmizis, A., 2013. Beyond the histone tail: Acetylation at the nucleosome dyad commands transcription. *Nucleus (Austin, Tex.)*, 4(5), pp.343–348.

Momčilović, O., Navara, C. & Schatten, G., 2011. Cell cycle adaptations and maintenance of genomic integrity in embryonic stem cells and induced pluripotent stem cells. *Results and problems in cell differentiation*, 53, pp.415–58.

Montes de Oca, R., Andreassen, P.R. & Wilson, K.L., 2011. Barrier-to-Autointegration Factor influences specific histone modifications. *Nucleus (Austin, Tex.)*, 2(6), pp.580–90.

Moore, K.A. & Lemischka, I.R., 2006. Stem cells and their niches. *Science (New York, N.Y.)*, 311(5769), pp.1880–5.

Morgan, B.A., Mittman, B.A. & Smith, M.M., 1991. The highly conserved N-terminal domains of histones H3 and H4 are required for normal cell cycle progression. *Molecular and cellular biology*, 11(8), pp.4111–20.

Morgan, D.O., 1995. Principles of CDK regulation. *Nature*, 374(6518), pp.131–4.

Mueller, P.R. et al., 1995. Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15. *Science (New York, N.Y.)*, 270(5233), pp.86–90.

Müller, C. et al., 1999. Cloning of the cyclin A1 genomic structure and characterization of the promoter region. GC boxes are essential for cell cycle-regulated transcription of the cyclin A1 gene. *The Journal of biological chemistry*, 274(16), pp.11220–8.

Müller, J. et al., 2002. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, 111(2), pp.197–208.

Müller, S. & Almouzni, G., 2013. A network of players in H3 histone variant deposition and maintenance at centromeres. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*.

Musselman, C.A. & Kutateladze, T.G., 2011. Methyl fingerprinting of the nucleosome reveals the molecular mechanism of high-mobility group nucleosomal-2 (HMGN2) association. *Proceedings of the National Academy of Sciences of the United States of America*, 108(30), pp.12189–90.

Nagoshi, E. et al., 2004. Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell*, 119(5), pp.693–705.

Nagy, Z. & Tora, L., 2007. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene*, 26(37), pp.5341–57.

Nakagawa, M. et al., 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature biotechnology*, 26(1), pp.101–6.

Nakai-Futatsugi, Y. & Niwa, H., 2013. Transcription factor network in embryonic stem cells: heterogeneity under the stringency. *Biological & pharmaceutical bulletin*, 36(2), pp.166–70.

Nakayama, J. et al., 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science (New York, N.Y.)*, 292(5514), pp.110–3.

Nasmyth, K., 1996. At the heart of the budding yeast cell cycle. *Trends in genetics : TIG*, 12(10), pp.405–12.

Natsume, R. et al., 2007. Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature*, 446(7133), pp.338–41.

Neganova, I. & Lako, M., 2008. G1 to S phase cell cycle transition in somatic and embryonic stem cells. *Journal of anatomy*, 213(1), pp.30–44.

Ng, P.M.-L. & Lufkin, T., 2011. Embryonic stem cells: protein interaction networks. *Biomolecular concepts*, 2(1-2), pp.13–25.

Ng, S.S. et al., 2009. Dynamic protein methylation in chromatin biology. *Cellular and molecular life sciences : CMLS*, 66(3), pp.407–22.

Nichols, J. et al., 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, 95(3), pp.379–91.

Nielsen, S.J. et al., 2001. Rb targets histone H3 methylation and HP1 to promoters. *Nature*, 412(6846), pp.561–5.

Nigg, E.A., 1995. Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 17(6), pp.471–80.

Nightingale, K.P., O'Neill, L.P. & Turner, B.M., 2006. Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. *Current opinion in genetics & development*, 16(2), pp.125–36.

Nishioka, K. et al., 2002. Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes & development*, 16(4), pp.479–89.

Nishiyama, A. et al., 2013. Systematic repression of transcription factors reveals limited patterns of gene expression changes in ES cells. *Scientific reports*, 3, p.1390.

Niwa, H. et al., 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & development*, 12(13), pp.2048–60.

Niwa, H., Miyazaki, J. & Smith, A.G., 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics*, 24(4), pp.372–6.

Nojima, H., 2004. G1 and S-phase checkpoints, chromosome instability, and cancer. *Methods in molecular biology (Clifton, N.J.)*, 280, pp.3–49.

Noma K, Allis, C.D. & Grewal, S.I., 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science (New York, N.Y.)*, 293(5532), pp.1150–5.

Norbury, C. & Nurse, P., 1991. Cyclins and cell cycle control. *Current biology : CB*, 1(1), pp.23–4.

Nurse, P., 2002. Cyclin dependent kinases and cell cycle control (nobel lecture). *Chembiochem : a European journal of chemical biology*, 3(7), pp.596–603.

Nurse, P., 1990. Universal control mechanism regulating onset of M-phase. *Nature*, 344(6266), pp.503–8.

O'Connor, P.M. et al., 1994. Role of the cdc25C phosphatase in G2 arrest induced by nitrogen mustard. *Proceedings of the National Academy of Sciences of the United States of America*, 91(20), pp.9480–4.

O'Doherty, R., Greiser, U. & Wang, W., 2013. Nonviral methods for inducing pluripotency to cells. *BioMed research international*, 2013, p.705902.

O'Neill, L.P. & Turner, B.M., 2003. Immunoprecipitation of native chromatin: NChIP. *Methods (San Diego, Calif.)*, 31(1), pp.76–82.

O'Neill, L.P., VerMilyea, M.D. & Turner, B.M., 2006. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nature genetics*, 38(7), pp.835–41.

Olsen, C.A., 2013. An Update on Lysine Deacylases Targeting the Expanding "Acylome." *ChemMedChem*.

Orford, K.W. & Scadden, D.T., 2008. Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. *Nature reviews. Genetics*, 9(2), pp.115–28.

Parthun, M.R., 2007. Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene*, 26(37), pp.5319–28.

Pauler, F.M. et al., 2009. H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome research*, 19(2), pp.221–33.

Pedersen, R.A., Wu, K. & Bałakier, H., 1986. Origin of the inner cell mass in mouse embryos: cell lineage analysis by microinjection. *Developmental biology*, 117(2), pp.581–95.

Pellegrini, G., De Luca, M. & Arsenijevic, Y., 2007. Towards therapeutic application of ocular stem cells. *Seminars in cell & developmental biology*, 18(6), pp.805–18.

Perera, H.K. et al., 2006. Expression and shifting subcellular localization of the transcription factor, Foxd3, in embryonic and adult pancreas. *Gene expression patterns : GEP*, 6(8), pp.971–7.

Pesavento, J.J. et al., 2008. Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. *Molecular and cellular biology*, 28(1), pp.468–86.

Peters, A.H.F.M. et al., 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Molecular cell*, 12(6), pp.1577–89.

Petruk, S. et al., 2012. TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. *Cell*, 150(5), pp.922–33.

Polo, S.E. & Almouzni, G., 2006. Chromatin assembly: a basic recipe with various flavours. *Current opinion in genetics & development*, 16(2), pp.104–11.

Prasongchean, W. & Ferretti, P., 2012. Autologous stem cells for personalised medicine. *New biotechnology*, 29(6), pp.641–50.

Prior, C.P. et al., 1980. Incorporation of exogenous pyrene-labeled histone into Physarum chromatin: a system for studying changes in nucleosomes assembled in vivo. *Cell*, 20(3), pp.597–608.

Probst, A. V, Dunleavy, E. & Almouzni, G., 2009. Epigenetic inheritance during the cell cycle. *Nature reviews. Molecular cell biology*, 10(3), pp.192–206.

Radman-Livaja, M. et al., 2011. Patterns and mechanisms of ancestral histone protein inheritance in budding yeast. *PLoS biology*, 9(6), p.e1001075.

Rama, P. et al., 2001. Autologous fibrin-cultured limbal stem cells permanently restore the corneal surface of patients with total limbal stem cell deficiency. *Transplantation*, 72(9), pp.1478–85.

Ramsköld, D. et al., 2009. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. L. J. Jensen, ed. *PLoS computational biology*, 5(12), p.e1000598.

Rea, S. et al., 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796), pp.593–9.

Reed, S.I., 1992. The role of p34 kinases in the G1 to S-phase transition. *Annual review of cell biology*, 8, pp.529–61.

Reich, N.C. & Liu, L., 2006. Tracking STAT nuclear traffic. *Nature reviews. Immunology*, 6(8), pp.602–12.

Rezza, A., Sennett, R. & Rendl, M., 2014. Adult Stem Cell Niches: Cellular and Molecular Components. *Current topics in developmental biology*, 107C, pp.333–372.

Rice, J.C. et al., 2002. Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes & development*, 16(17), pp.2225–30.

Richmond, T.J. & Davey, C.A., 2003. The structure of DNA in the nucleosome core. *Nature*, 423(6936), pp.145–50.

Richter, E.A. & Hargreaves, M., 2013. Exercise, GLUT4, and skeletal muscle glucose uptake. *Physiological reviews*, 93(3), pp.993–1017.

Rieder, C.L. & Palazzo, R.E., 1992. Colcemid and the mitotic cycle. *Journal of cell science*, 102 (Pt 3, pp.387–92.

Rivera, A. et al., 2006. Cyclin A1 is a p53-induced gene that mediates apoptosis, G2/M arrest, and mitotic catastrophe in renal, ovarian, and lung carcinoma cells. *Cellular and molecular life sciences : CMLS*, 63(12), pp.1425–39.

Rothbart, S.B. & Strahl, B.D., 2014. Interpreting the language of histone and DNA modifications. *Biochimica et biophysica acta*.

Rowbotham, S.P. et al., 2011. Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. *Molecular cell*, 42(3), pp.285–96.

De Ruijter, A.J.M. et al., 2003. Histone deacetylases (HDACs): characterization of the classical HDAC family. *The Biochemical journal*, 370(Pt 3), pp.737–49.

Savatier, P. et al., 1994. Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. *Oncogene*, 9(3), pp.809–18.

Scharf, A.N.D., Barth, T.K. & Imhof, A., 2009. Establishment of histone modifications after chromatin assembly. *Nucleic acids research*, 37(15), pp.5032–40.

Schöler, H.R. et al., 1989. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *The EMBO journal*, 8(9), pp.2551–7.

Schuettengruber, B. et al., 2007. Genome regulation by polycomb and trithorax proteins. *Cell*, 128(4), pp.735–45.

Serrano, M., Hannon, G.J. & Beach, D., 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature*, 366(6456), pp.704–7.

Shahriyari, L. & Komarova, N.L., 2013. Symmetric vs. Asymmetric Stem Cell Divisions: An Adaptation against Cancer? *PloS one*, 8(10), p.e76195.

Sherr, C.J., 1994. G1 phase progression: cycling on cue. *Cell*, 79(4), pp.551–5.

Sherr, C.J., 1993. Mammalian G1 cyclins. *Cell*, 73(6), pp.1059–65.

Sherr, C.J. & Roberts, J.M., 1995. Inhibitors of mammalian G1 cyclin-dependent kinases. *Genes & development*, 9(10), pp.1149–63.

Shilatifard, A., 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current opinion in cell biology*, 20(3), pp.341–8.

Shilatifard, A., 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry*, 81, pp.65–95.

Simon, J.A. & Kingston, R.E., 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature reviews. Molecular cell biology*, 10(10), pp.697–708.

Sims, R.J. & Reinberg, D., 2009. Processing the H3K36me3 signature. *Nature genetics*, 41(3), pp.270–1.

Smith, A.G. et al., 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, 336(6200), pp.688–90.

Smith, B.C. & Denu, J.M., 2009. Chemical mechanisms of histone lysine and arginine modifications. *Biochimica et biophysica acta*, 1789(1), pp.45–57.

Smith, D.J. & Whitehouse, I., 2012. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature*, 483(7390), pp.434–8.

Smith, S. & Stillman, B., 1989. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell*, 58(1), pp.15–25.

Solomon, M.J., 1994. The function(s) of CAK, the p34cdc2-activating kinase. *Trends in biochemical sciences*, 19(11), pp.496–500.

Somech, R. et al., 2005. The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *Journal of cell science*, 118(Pt 17), pp.4017–25.

Spivakov, M. & Fisher, A.G., 2007. Epigenetic signatures of stem-cell identity. *Nature reviews. Genetics*, 8(4), pp.263–71.

Stadtfeld, M. et al., 2008. Induced pluripotent stem cells generated without viral integration. *Science (New York, N.Y.)*, 322(5903), pp.945–9.

Stein, G.S. et al., 2012. The architectural organization of human stem cell cycle regulatory machinery. *Current pharmaceutical design*, 18(13), pp.1679–85.

Stice, S.L. et al., 2006. Human embryonic stem cells: challenges and opportunities. *Reproduction, fertility, and development*, 18(8), pp.839–46.

Stillman, B., 1986. Chromatin assembly during SV40 DNA replication in vitro. *Cell*, 45(4), pp.555–565.

Strahl, B.D. & Allis, C.D., 2000. The language of covalent histone modifications. *Nature*, 403(6765), pp.41–5.

Struhl, K. & Segal, E., 2013. Determinants of nucleosome positioning. *Nature structural & molecular biology*, 20(3), pp.267–73.

Tagami, H. et al., 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, 116(1), pp.51–61.

Takahashi, K. et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), pp.861–72.

Takahashi, K. & Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), pp.663–76.

Tamaru, H. et al., 2003. Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in Neurospora crassa. *Nature genetics*, 34(1), pp.75–9.

Tamm, C., Böwer, N. & Annerén, C., 2011. Regulation of mouse embryonic stem cell self-renewal by a Yes-YAP-TEAD2 signaling pathway downstream of LIF. *Journal of cell science*, 124(Pt 7), pp.1136–44.

Tan, M. et al., 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, 146(6), pp.1016–28.

Tardat, M. et al., 2010. The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. *Nature cell biology*, 12(11), pp.1086–93.

Taverna, S.D. et al., 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology*, 14(11), pp.1025–40.

Taverna, S.D. et al., 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular cell*, 24(5), pp.785–96.

Teperek-Tkacz, M. et al., 2010. Phosphorylation of histone H3 serine 10 in early mouse embryos: Active phosphorylation at late S phase and differential effects of ZM447439 on first two embryonic mitoses. *Cell Cycle*, 9(23), pp.4674–4687.

Terrenoire, E. et al., 2010. Immunostaining of modified histones defines high-level features of the human metaphase epigenome. *Genome biology*, 11(11), p.R110.

La Thangue, N.B., 1994a. DP and E2F proteins: components of a heterodimeric transcription factor implicated in cell cycle control. *Current opinion in cell biology*, 6(3), pp.443–50.

La Thangue, N.B., 1994b. DRTF1/E2F: an expanding family of heterodimeric transcription factors implicated in cell-cycle control. *Trends in biochemical sciences*, 19(3), pp.108–14.

Thomas, L.R. et al., 2008. Functional analysis of histone methyltransferase g9a in B and T lymphocytes. *Journal of immunology (Baltimore, Md. : 1950)*, 181(1), pp.485–93.

Thomson, J.A. et al., 1998. Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)*, 282(5391), pp.1145–7.

Thomson, J.P. et al., 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature*, 464(7291), pp.1082–6.

Torres-Padilla, M.-E., 2013. Generating different epigenotypes. *Reproductive BioMedicine Online*.

Tropberger, P. & Schneider, R., 2010. Going global: novel histone modifications in the globular domain of H3. *Epigenetics : official journal of the DNA Methylation Society*, 5(2), pp.112–7.

Tsou, A.-P. et al., 2003. Identification of a novel cell cycle regulated gene, HURP, overexpressed in human hepatocellular carcinoma. *Oncogene*, 22(2), pp.298–307.

Tsubouchi, T. & Fisher, A.G., 2013. Reprogramming and the pluripotent stem cell cycle. *Current topics in developmental biology*, 104, pp.223–41.

Turner, B.M., 1993. Decoding the nucleosome. *Cell*, 75(1), pp.5–8.

Turner, B.M., 2007. Defining an epigenetic code. *Nature cell biology*, 9(1), pp.2–6.

Turner, B.M., 2000. Histone acetylation and an epigenetic code. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 22(9), pp.836–45.

Turner, B.M., 1998. Histone acetylation as an epigenetic determinant of long-term transcriptional competence. *Cellular and molecular life sciences : CMLS*, 54(1), pp.21–31.

Turner, B.M., 2014. Nucleosome signalling; An evolving concept. *Biochimica et biophysica acta*.

Turner, B.M., 2012. The adjustable nucleosome: an epigenetic signaling module. *Trends in genetics : TIG*, 28(9), pp.436–44.

Turner, B.M. & Fellows, G., 1989. Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells. *European journal of biochemistry / FEBS*, 179(1), pp.131–9.

Turner, B.M., O'Neill, L.P. & Allan, I.M., 1989. Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS letters*, 253(1-2), pp.141–5.

Varga-Weisz, P.D., 2010. Insights into how chromatin remodeling factors find their target in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46), pp.19611–2.

Voigt, P. et al., 2012. Asymmetrically modified nucleosomes. *Cell*, 151(1), pp.181–93.

Voigt, P., Tee, W.-W. & Reinberg, D., 2013a. A double take on bivalent promoters. *Genes & development*, 27(12), pp.1318–38.

Wang, F. & Higgins, J.M.G., 2013. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends in cell biology*, 23(4), pp.175–84.

Wang, Y. et al., 2013. miR-294/miR-302 promotes proliferation, suppresses G1-S restriction point, and inhibits ESC differentiation through separable mechanisms. *Cell reports*, 4(1), pp.99–109.

Wang, Y. & Blelloch, R., 2009. Cell cycle regulation by MicroRNAs in embryonic stem cells. *Cancer research*, 69(10), pp.4093–6.

Warren, L. et al., 2010. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell stem cell*, 7(5), pp.618–30.

Waterston, R.H. et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–62.

Wegner, M., 2011. SOX after SOX: SOXession regulates neurogenesis. *Genes & development*, 25(23), pp.2423–8.

Wei, Y. et al., 1998. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, 95(13), pp.7480–4.

Weintraub, H., Worcel, A. & Alberts, B., 1976. A model for chromatin based upon two symmetrically paired half-nucleosomes. *Cell*, 9(3), pp.409–17.

Wen, B. et al., 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature genetics*, 41(2), pp.246–50.

White, J. et al., 2005. Developmental activation of the Rb-E2F pathway and establishment of cell cycle-regulated cyclin-dependent kinase activity during embryonic stem cell differentiation. *Molecular biology of the cell*, 16(4), pp.2018–27.

White, J. & Dalton, S., 2005. Cell cycle control of embryonic stem cells. *Stem cell reviews*, 1(2), pp.131–8.

Williams, R.L. et al., 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, 336(6200), pp.684–7.

Wilmut, I. et al., 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385(6619), pp.810–3.

Winkler, D.D. et al., 2012. Yeast CAF-1 assembles histone (H3-H4)2 tetramers prior to DNA deposition. *Nucleic acids research*, 40(20), pp.10139–49.

Woodcock, C.L. & Ghosh, R.P., 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harbor perspectives in biology*, 2(5), p.a000596.

Xiao, B. et al., 2003. Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature*, 421(6923), pp.652–6.

Xiong, Y. et al., 1993. p21 is a universal inhibitor of cyclin kinases. *Nature*, 366(6456), pp.701–4.

Xu, M. et al., 2010. Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science (New York, N.Y.)*, 328(5974), pp.94–8.

Yamasu, K. & Senshu, T., 1990. Conservative segregation of tetrameric units of H3 and H4 histones during nucleosome replication. *Journal of biochemistry*, 107(1), pp.15–20.

Yang, X.-J. & Seto, E., 2007. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene*, 26(37), pp.5310–8.

Ye, L. et al., 2013. Patching the heart: cardiac repair from within and outside. *Circulation research*, 113(7), pp.922–32.

Ying, Q.L. et al., 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell*, 115(3), pp.281–92.

Yoshioka, H., McCarrey, J.R. & Yamazaki, Y., 2009. Dynamic nuclear organization of constitutive heterochromatin during fetal male germ cell development in mice. *Biology of reproduction*, 80(4), pp.804–12.

Young, M.D. et al., 2011. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic acids research*, 39(17), pp.7415–27.

Yu, H. et al., 2008. Inferring causal relationships among different histone modifications and gene expression. *Genome research*, 18(8), pp.1314–24.

Yu, J. et al., 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)*, 318(5858), pp.1917–20.

Van Zant, G. & Liang, Y., 2012. Concise review: hematopoietic stem cell aging, life span, and transplantation. *Stem cells translational medicine*, 1(9), pp.651–7.

Zee, B.M. et al., 2010. In vivo residue-specific histone methylation dynamics. *The Journal of biological chemistry*, 285(5), pp.3341–50.

Zegerman, P. et al., 2002. Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *The Journal of biological chemistry*, 277(14), pp.11621–4.

Zentner, G.E. & Henikoff, S., 2013. Regulation of nucleosome dynamics by histone modifications. *Nature structural & molecular biology*, 20(3), pp.259–66.

Zhang, S. et al., 2011. Modification of histones by sugar β-N-acetylglucosamine (GlcNAc) occurs on multiple residues, including histone H3 serine 10, and is cell cycle-regulated. *The Journal of biological chemistry*, 286(43), pp.37483–95.

Zhang, X. et al., 2009. A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. *The Journal of cell biology*, 184(1), pp.67–82.

Zhou, H. et al., 2009. Generation of induced pluripotent stem cells using recombinant proteins. *Cell stem cell*, 4(5), pp.381–4.

Zhou, J. et al., 2010. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. *Plant molecular biology*, 72(6), pp.585–95.

Zhou, Q. et al., 2007. A gene regulatory network in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 104(42), pp.16438–43.

Zhou, W. et al., 2013. ERα, SKP2 and E2F-1 form a feed forward loop driving late ERα targets and G1 cell cycle progression. *Oncogene*.

Zhu, B. & Reinberg, D., 2011. Epigenetic inheritance: uncontested? *Cell research*, 21(3), pp.435–41.

Ziv, O., Glaser, B. & Dor, Y., 2013. The plastic pancreas. *Developmental cell*, 26(1), pp.3–7.

6 APPENDIX

Table 6.1: Percentage pull downs of DNA from C-ChIP of G1, S and G2/M-phase FACs sorted ES cells as calculated by picogreen analysis

Antibody	Sample	N=1	N=2	N=3
H3K4me3	Async	1.658	1.941	1.271
H3K27me3	Async	6.943	3.351	5.896
H3K9ac	Async	1.745	5.387	0.932
H3K9me2	Async	1.402	4.049	0.320
H4K8ac	Async		0.527	0.478
PreImmune	Async	0.479	0.075	0.0543
H3K4me3	G1	2.639	2.757	2.453
H3K27me3	G1	6.105	7.028	8.077
H3K9ac	G1	1.686	2.786	0.965
H3K9me2	G1	1.25	1.842	0.403
H4K8ac	G1		1.001	1.102
PreImmune	G1	0.165	0.330	0.072
H3K4me3	S	2.471	2.325	2.035
H3K27me3	S	3.423	6.209	8.333
H3K9ac	S	2.793	13.418	1.259
H3K9me2	S	1.678	1.678	0.225
H4K8ac	S		0.419	1.504
PreImmune	S	0.792	0.297	0.670
H3K4me3	G2/M	1.849	2.932	2.105
H3K27me3	G2/M	10.467	6.258	9.969
H3K9ac	G2/M	2.208	5.141	1.621
H3K9me2	G2/M	1.593	15.068	0.483
H4K8ac	G2/M		1.320	2.355
PreImmune	G2/M	0.331	0.386	0.686

Relative levels of histone modifications were determined by carrier chromatin immunoprecipitation (C-ChIP) in OS25 mouse ES cells FACs sorted into G1, S and G2/M phase populations alongside an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3 and H3K9me2 as seen above. DNA from antibody bound and unbound fractions were purified and quantified by picogreen. From these values the percentage pull down for each antibody. A pre-immune no antibody control was also included which had a very low pull down efficiency.

Table 6.2: Percentage pull downs of DNA from N-ChIP of G1/S-phase arrested day 7 differentiated ES cells as calculated by picogreen analysis

Antibody	Sample	N=1	N=2	N=3
H3K4me3	Async	1.483	1.784	1.415
H3K27me3	Async	17.372	18.362	19.427
H3K9ac	Async	0.146	0.536	1.112
H3K9me2	Async	19.007	11.291	1.344
H4K8ac	Async	6.056	7.54	8.704
H3S10phos	Async	3.686	26.919	27.946
PreImmune	Async	0.106	0.028	0.812
H3K4me3	G1/S	0.678	0.924	0.924
H3K27me3	G1/S	16.773	18.465	35.291
H3K9ac	G1/S	0.580	0.241	0.254
H3K9me2	G1/S	16.782	22.867	4.503
H4K8ac	G1/S	5.947	10.142	7.572
H3S10phos	G1/S	20.848	28.906	35.819
PreImmune	G1/S	0.567	0.025	0.275

Relative levels of histone modifications were determined by native chromatin immunoprecipitation (N-ChIP) in day 7 differentiated OS25 mouse ES cells arrested at G1/S phase by treating with thymidine and arrested at M phase by treating with colcemid alongside an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H4K8ac, H3K27me3, H3K9me2 and H3S10phos as seen above. DNA from antibody bound and unbound fractions were purified and quantified by picogreen. From these values the percentage pull down for each antibody. A pre-immune no antibody control was also included which had a very low pull down efficiency.

Table 6.3: Percentage pull downs of DNA from C-ChIP of G1, S and G2/M-phase FACs sorted day 7 differentiated ES cells as calculated by picogreen analysis

Antibody	Sample	N=1	N=2
H3K4me3	Async	0.508	2.118
H3K27me3	Async	2.520	8.341
H3K9ac	Async	0.974	0.692
H3K9me2	Async	11.546	11.783
PreImmune	Async	0.252	0.422
H3K4me3	G1	0.246	0.538
H3K27me3	G1	0.773	1.236
H3K9ac	G1	0.453	0.540
H3K9me2	G1	0.035	0.257
PreImmune	G1	-0.140	0.400
H3K4me3	S	0.200	1.254
H3K27me3	S	0.790	25.636
H3K9ac	S	0.714	0.500
H3K9me2	S	9.253	21.817
PreImmune	S	0.090	0.504
H3K4me3	G2/M	0.684	3.500
H3K27me3	G2/M	0.764	52.977
H3K9ac	G2/M	0.623	0.560
H3K9me2	G2/M	20.030	39.772
PreImmune	G2/M	0.259	0.417

Relative levels of histone modifications were determined by carrier chromatin immunoprecipitation (C-ChIP) in day 7 differentiated OS25 mouse ES cells FACs sorted into G1, S and G2/M phase populations alongside an asynchronous population. Immunoprecipitation performed with antibodies to H3K4me3, H3K9ac, H3K27me3 and H3K9me2 as seen above. DNA from antibody bound and unbound fractions were purified and quantified by picogreen. From these values the percentage pull down for each antibody. A pre-immune no antibody control was also included which had a very low pull down efficiency.